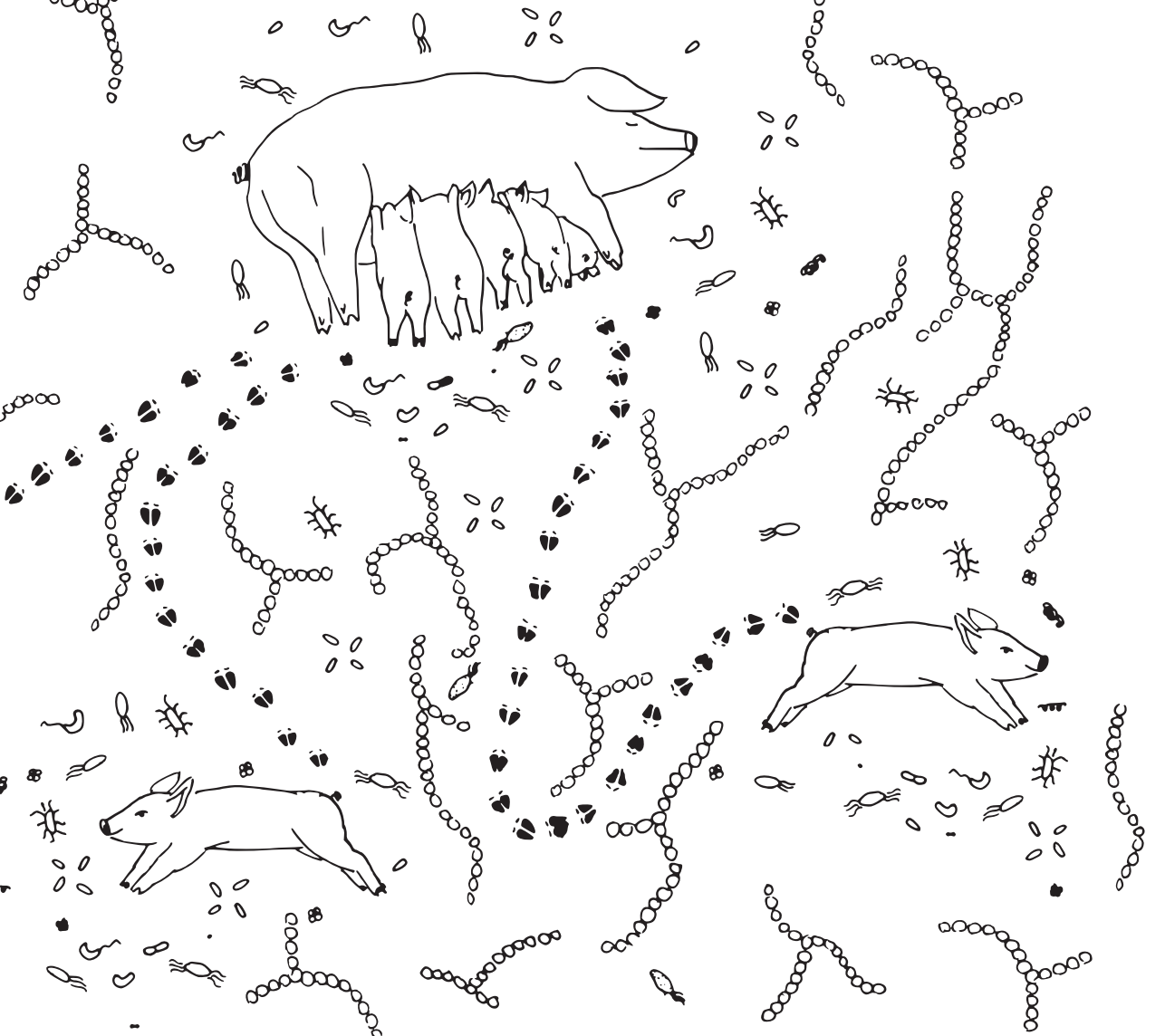
The piglet tonsillar microbiome and *Streptococcus suis* disease

Simen Fredriksen



Propositions

1. Prevention of opportunistic pathogens in industrial pig farming is not economically feasible.
   (this thesis)

2. Confounding factors make most pig microbiome studies uninterpretable.
   (this thesis)

3. Outsourcing DNA sequencing to the cheapest bidder is a recipe for disaster.

4. Public funding is better spent on development of high-quality scientific software than applied research.

5. The impact of livestock research will decrease as alternative protein sources outcompete animal farming.

6. Soap and disinfectants are overused.

Propositions belonging to the thesis, entitled:

The piglet tonsillar microbiome and *Streptococcus suis* disease

Simen Fredriksen
Wageningen, 31 January 2023

The piglet tonsillar microbiome and *Streptococcus suis* disease

Simen Fredriksen

Thesis committee

Promotor
Prof. Dr Jerry M. Wells
Professor of Host-Microbe Interactomics
Wageningen University & Research

Copromotors
Dr Peter van Baarlen
Researcher, Host-Microbe Interactomics Group
Wageningen University & Research

Dr Jos Boekhorst
Researcher, Host-Microbe Interactomics Group
Wageningen University & Research

Other members
Prof. Dr Hauke Smidt, Wageningen University & Research
Prof. Dr Rob J.L. Willems, University Medical Centre Utrecht
Prof. Dr Alexander W. Tucker, University of Cambridge, United Kingdom
Dr Lucy A. Weinert, University of Cambridge, United Kingdom

This research was conducted under the auspices of the Graduate School Wageningen Institute of Animal Sciences (WIAS).

The piglet tonsillar microbiome and *Streptococcus suis* disease

Simen Fredriksen

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 31 January 2023
at 1.30 p.m. in the Omnia Auditorium.

Simen Fredriksen
The piglet tonsillar microbiome and *Streptococcus suis* disease,
146 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2023)
With references, with summary in English

ISBN: 978-94-6447-511-1
DOI: https://doi.org/10.18174/582380

Table of contents

Chapter 1

General introduction

Introduction

Global health challenges caused by new and (re-)emerging infectious diseases of animals and humans are expected to increase in the coming decades [1, 2], with antibiotic-resistant bacteria projected to cause up to 10 million annual human deaths by 2050 [3]. In humans, respiratory infections are a major challenge, and while most cases are viral in nature, bacterial pathogens such as *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Moraxella catarrhalis* are of great clinical relevance. Treatment of these and other bacterial pathogens depend on antibiotics, and thus their impact on human health will increase with rising antimicrobial resistance levels. Bacterial infectious disease in livestock is usually caused by species closely related to those infecting humans. Transferrable resistance mechanisms present a threat to human health, in particular when the same antibiotics are used in human and veterinary medicine.

Pig farming involves heavy use of antibiotics to prevent and treat infectious disease [4], and this has driven increased antibiotic resistance [5]. Following One Health principles, antibiotic resistance may not only be transferred between porcine pathogenic bacteria, but also from porcine to human pathogens by horizontal gene transfer. The oropharyngeal cavity of pigs and humans share high abundance of closely related species within genera such as *Streptococcus*, *Haemophilus*, *Neisseria*, *Pseudomonas*, and *Acinetobacter*. These genera all contain pathogens on the WHO priority list for development of new antibiotics [6]. Some pig colonising species such as *Staphylococcus aureus* and *Streptococcus suis* can also cause zoonotic disease in humans [7].

S. suis is an opportunistic pathogen that colonizes the upper respiratory tract of pigs at high prevalence but also may cause invasive disease dependent on as yet undetermined host- and environmental factors [8]. *S. suis* is also an emerging zoonotic pathogen causing meningitis in humans, with the majority of cases occurring in southeast Asia [9, 10]. Prevention of (porcine) *S. suis* infectious disease is difficult due to lack of cross-protective vaccines, and outbreaks are managed by metaphylactic antimicrobial treatment [8, 11]. This contributes to the emergence of antibiotic resistance, and multi-resistance is common among contemporary *S. suis* isolates [12]. This thesis investigates the biology of *Streptococcus suis* and the bacterial community inhabiting its main habitat, the tonsils of piglets.

The microbiota and infectious disease

Microbiota is a term used to describe to the collection of all microorganisms colonizing an environmental or host-associated habitat. In humans and animals, the gastrointestinal tract (gut) microbiota is the most studied. The intestinal microbiota plays important roles in nutrient uptake, communication via the gut-brain axis, and the pathophysiology of several diseases [13]. The microbiota of the gut and other sites is also key to training and inducing host immunity and immune regulation [14, 15].

Another way the microbiota may impact disease occurrence is by conferring resistance to colonisation and invasion by opportunistic pathogens [16]. Opportunists cause disease in compromised hosts, and a dysbiotic gut microbiota depleted of commensal colonizers is associated with increased infection risk [17]. Commensals may outcompete pathogens by being more specialized in specific non-invasive niches with highly efficient metabolism and competition strategies [18]. In the human gut, commensals produce secondary bile salts and short-chain fatty acids (SCFA) that inhibit opportunistic pathogens [19, 20]. Commensals and their protective effect may be depleted under disruptive conditions. Notably, medical use of antibiotics may not only inhibit targeted pathogens but also the growth of susceptible commensals, thereby disturbing microbiota homeostasis and inducing dysbiosis. With less competition from commensals, opportunistic pathogens such as *Escherichia coli*, *Salmonella Typhimurium*, and *Clostridioides difficile* may rapidly increase in abundance after surviving antibiotic treatment by resistance or sporulation [21–23].

There are fewer studies on the microbiota of the upper respiratory tract (including the nasal cavity and mouth), despite its clinical importance, interaction with the mucosal-associated immune system, and inclusion of opportunistic pathogens [24–26]. Infections by opportunistic pathogens cause inflammation in the upper respiratory tract, manifesting as pharyngitis and tonsillitis. Some bacteria may also cause local or systemic infection by crossing the host epithelium and entering the bloodstream. In humans, common opportunistic colonizers of the upper respiratory tract include *Streptococcus pyogenes*, *S. pneumoniae*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus*. In pigs, species such as *S. suis*, *Glaeserella parasuis*, and *Actinobacillus pleuropneumoniae* share equivalent niches. These and other bacterial and viral pathogens may cause disease both during early life and during the grower-finisher stage in what is termed the porcine respiratory disease complex (PRDC) [27].

Antibiotic resistance and the resistome

Since the discovery of penicillin in 1950, antibiotics have revolutionised treatment of bacterial infections, but introduction of antibiotics was followed by rapid spread of antibiotic-resistant strains. Bacteria continue to gain resistance, raising concerns about a post-antibiotic future where bacterial infections will once again become a leading cause of mortality in humans [28, 29]. Because resistance to antibiotics emerges as an adaptation to antibiotic exposure, reduced antibiotic usage is key to limiting resistance. Although antibiotics are crucial for the treatment of bacterial infectious diseases they are also misused. For instance, antibiotics targeting *Streptococcus pyogenes* have been prescribed in the majority of humans pharyngitis cases in the USA, despite most infections being caused by viruses [30]. In low-income countries such as India and Bangladesh, antibiotics are available and commonly used without prescription or guidance on appropriate use, leading to high resistance levels in these countries [31, 32]. In livestock, antibiotics have been

used prophylactically in the absence of any disease symptoms as "growth promotors" [33, 34]. While now banned in the EU, antibiotic growth promotors are still used elsewhere despite contributing to increasing antibiotic resistance [5].

Acquisition of antibiotic resistance genes (ARGs) is a common mechanism for bacteria to gain antibiotic resistance. ARGs can spread rapidly by horizontal gene transfer (HGT) via phages, plasmids, and transposable elements, further described in textbox 1.1 [35–37]. The collection of ARGs within the metagenome (i.e., all genetic material) of a microbiota is termed the resistome. The epidemiology of ARGs in pathogenic bacteria has been under intense scrutiny due to their clinical relevance, and in recent years the role of the commensal microbiota and resistome in the spread of ARGs has gained attention [38]. While pathogenic bacteria may colonize hosts only transiently, the commensal microbiota is a stable reservoir of ARGs. ARGs carried in harmless commensals are not a problem but for the fact that they can be rapidly acquired by con- and intraspecific pathogens upon infection and antibiotic treatment [39, 40]. The commensal microbiome is continuously shared within populations [41, 42], and with a One Health perspective in mind, it follows that antibiotic use may have consequences beyond the individual human or animal host. Surveying and limiting the expansion of host and environmental resistomes is thus relevant to the health of humans and livestock.

High prevalence of multi-resistant bacteria in pig farms may thus present a danger to human health not only by spread of zoonotic pathogens, but also by transfer of commensals and by HGT of ARGs between pig- and human-specific species. This is great concern because of overlapping antibiotics use to treat human and porcine infections. Contact with livestock and bacterial transfer from animal to human may influence human microbiota composition and the associated resistome [43–46]. Notably, pig-associated methicillin-resistant *Staphylococcus aureus* (MRSA) is prevalent among farm workers and their household members [47]. The resistome of pigs may also transfer to humans indirectly via use of pig manure as fertilizer, because this impacts on the resistome of soil and crop-associated bacteria [48–50].

Textbox 1.1: Horizontal gene transfer:
Horizontal gene transfer (HGT) describes exchange of genetic material between organisms other than from parent to offspring (i.e., vertical transfer). It is usually implied that the recipient integrates the exchanged DNA into their genome. HGT impacts on prokaryote evolution by creating mosaic genomes that contain genes from different strains or taxa. Acquisition of new genetic material enables rapid gain of new phenotypes and may facilitate adaption to new or changing environments. Notably, strains may receive genomic "island" regions containing multiple genes, sometimes related to pathogenicity and involving virulence genes and ARGs.

HGT may occur via several different mechanisms, most notably by conjugation, transduction, and transformation. Transformation is the direct uptake of DNA from the environment, and is thought to be a common mechanism among *Streptococcus* species [51, 52]. Conjugation describes transfer of plasmids by direct bacterial cell-to-cell contact. DNA may also be incorporated into the recipient chromosome by transposons. Transduction involves mis-packaging of bacterial DNA into a bacteriophage which is then transferred into a new strain by infection. DNA may also be transferred by a range of other less well-defined ways [53].

Streptococcus suis

One of the major causes of antibiotic usage in pig farming is *Streptococcus suis*, a Gram-positive opportunistic pathogen. *S. suis* strains harmlessly colonize the oral cavity of piglets in high abundance, but sometimes cause invasive disease. In piglets, *S. suis* infections are most common around weaning age with symptoms including arthritis, endocarditis, sepsis, and meningitis. In older pigs, *S. suis* disease more commonly manifests as respiratory symptoms [10, 54, 55]. *S. suis* is also a zoonotic pathogen that can be transmitted to humans by close contact with pigs and raw pig products [10].

S. suis strains typically possess one of several different variants (serotypes) of a polysaccharide capsule, and this has traditionally been used as a phenotypic marker to classify *Streptococcus* isolates [56, 57]. Variation in the structure and expression level of the streptococcal capsule is key to both colonization and evasion of immune defences during host invasion [58–60]. Both commensal and pathogenic strains possess capsules, but specific serotypes are associated with disease causing isolates. Most zoonotic *S. suis* disease is caused by serotype 2 strains [61, 62].

Recent advances in DNA sequencing have enabled large scale genome sequencing of *S. suis* isolates (see definitions in Textbox 1.2), providing more details on *S. suis* phylogeny and genetics. *S. suis* strains can be grouped into clades, and these have varying associations with clinical and non-clinical isolation sources [63–65]. This may be interpreted as clades being variably adapted to commensal and pathogenic niches. The phylogenetic relationship of *S. suis* strains used in this thesis and relevant reference genomes is shown in Figure 1.

The great diversity of putatively commensal *S. suis* indicates a long evolutionary history within the oral cavity of pigs. Pathogenic *S. suis*, in particular the clade most commonly causing zoonotic disease, are less diverse, have smaller genome sizes [63], and appear to have emerged recently following intensification of pig farming [64]. Pathogenic *S. suis* clades may have reduced ability to persist as colonizers in competition with commensals [66], although they are sometimes isolated from the respiratory tract of healthy pigs.

The oral cavity, in particular the palatine tonsils, have been suggested as a main habitat of *S. suis*, and a putative site of entry into the host bloodstream [66–73]. The oral and respiratory microbiome may be relevant to *S. suis* disease via co-infection and/or by providing colonization resistance. It is thought that co-infection by bacterial and viral pathogens may predispose to *S. suis* invasive disease, especially in post-weaning respiratory disease where *S. suis* is considered part of the porcine respiratory disease complex (PRDC) [74–77]. The host immune system and abiotic stressors such as abrupt maternal separation may also predispose piglets to *S. suis* disease [75]. Reduced maternal immunity and change in diet from sow milk to starch rich dry feed upon weaning has been proposed to lead to microbiota disruption and to facilitate exponential growth and infections by *S. suis* [8, 78]. As *S. suis* can be found at high abundance on the tonsils of healthy pigs, it is clear that at least some strains are able to persist in a commensal niche distinct from host invasion. It is not clear to what degree virulent strains persist as commensal colonisers within sows, rooms, and farms between outbreaks.

Textbox 1.2: Definition of terms used to describe *S. suis* strains:
An isolate is a pure culture started from a single bacterial cell. The term strain is used interchangeably with isolate, but several isolates with high genomic similarity may be considered as the same strain.

Clinical strains are isolated from blood or infected organs of diseased hosts, usually by necropsy.

Non-clinical strains are isolated from non-invasive sites such as on the tonsils and nasal cavity, usually from healthy hosts.

A lineage is an ancestor-descendant line of bacterial cells, while a clade is a monophyletic group of lineages that share a common ancestor. These terms have somewhat limited meaning in bacteria, and in particular streptococci, due to horizontal gene transfer. In practice, the term clade is often used to describe a set of related strains that are highly similar at the level of DNA sequence identity while lineage is used to describe their evolutionary history.

Pathogenic or disease-associated clades are putatively adapted to a niche of host invasion and can be identified by predominantly consisting of clinical strains.

Commensal or non-disease-associated clades are predominantly comprised of non-clinical strains and are assumed to be specialized in colonizing the oral cavity of pigs while having limited potential for host invasion.

Virulent strains have high potential for causing harm to the host e.g., through systemic disease. This may be assessed observationally by repeated isolation from several diseased hosts and experimentally by experimental infection. Virulence genes are considered important to strain virulence.

Systemic disease is used to describe all symptoms where infection takes place within the body, such as when *S. suis* can be found in blood, the heart, brain, or joints. This may manifest with symptoms such as sepsis, meningitis, and arthritis. Respiratory disease indicates pneumonia symptoms and that *S. suis* can be isolated from the lungs but not systemic sites.

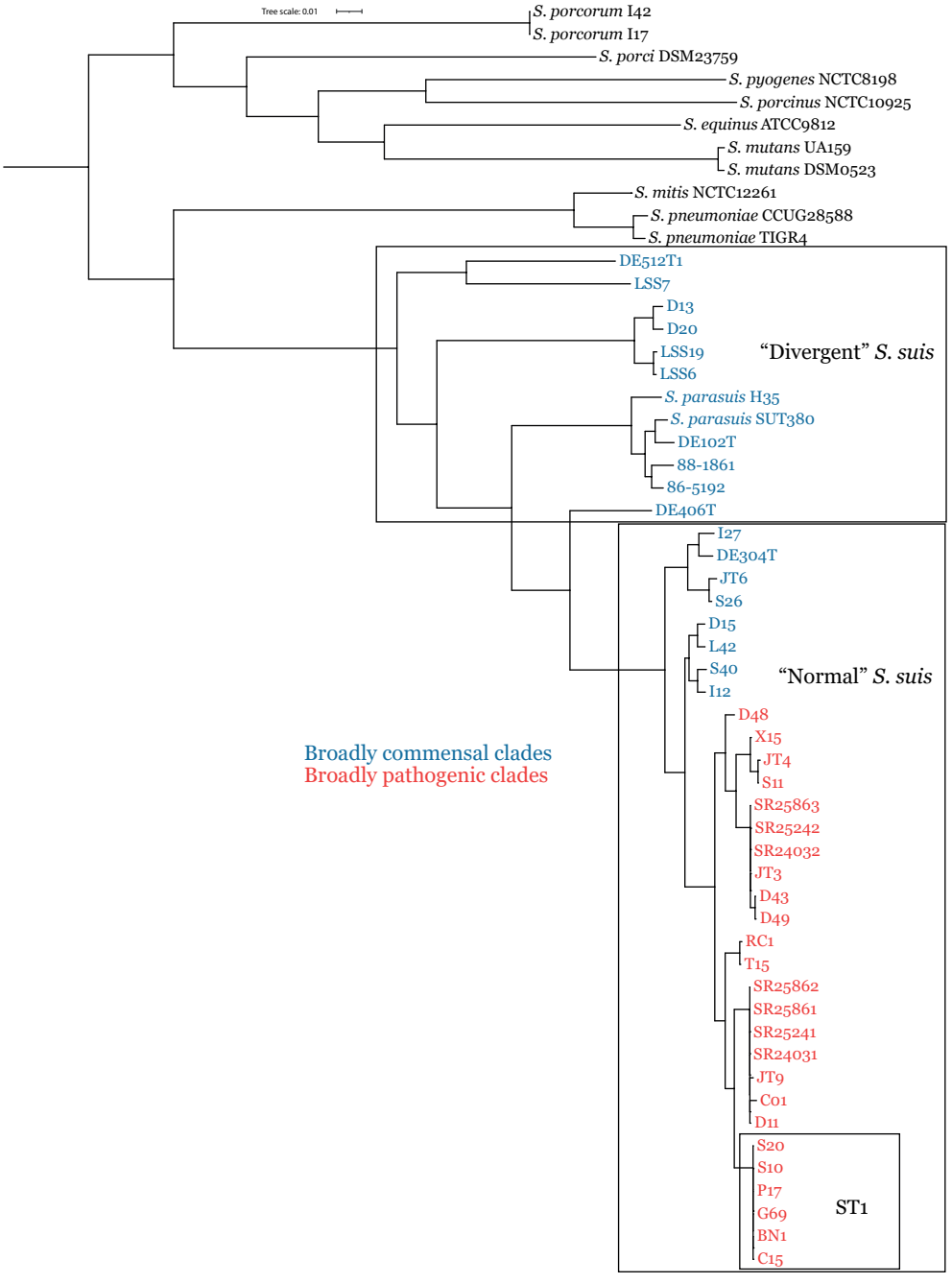


Figure 1: Phylogenetic tree showing the relationship between strains used in this thesis and key *Streptococcus* reference genomes. The tree was constructed with STAG [79] on core genes and mid-point rooted. ST1 = sequence type 1, a highly pathogenic group which includes zoonotic isolates.

Prevention and treatment of *Streptococcus suis* disease

S. suis disease is endemic and costly to farmers, making surveillance, prevention, and treatment a priority. There is no cross-protective vaccine available, owing to the large diversity of *S. suis* strains [8]. In the absence of effective and cross-protective vaccines, autogenous bacterin vaccines are used in some regions [80]. Autogenous bacterins consist of dead bacteria and are usually generated from clinical strains isolated by necropsy on the individual farm where they are applied. Bacterins may be used both directly to vaccinate piglets and indirectly on sows with the aim of providing passive immunity to their offspring. Immunoglobulins can be transferred via sow colostrum and taken up across the piglet intestine in the first two days of life [81]. The effectiveness of bacterins in preventing post-weaning disease in piglets is controversial [8, 80, 82–84].

Considering that the emergence of *S. suis* disease appears linked to modern industrial farming practices, reverting to less intensive farming may reduce the *S. suis* disease burden. Since this comes at an economic cost, it is important to identify the factors most relevant to disease. Farmers and veterinarians currently take measures such as adjusting weaning age, regulating temperature and air quality, improving hygiene, and separating animals, but the effectiveness of these appear situational and are not scientifically proven. Furthermore, existing research on *S. suis* carriage and transmission has largely been culture-based [85–90], which is unlikely to accurately detect *S. suis* prevalence and abundance. Existing work is also limited by use of lab-based determination of *S. suis* diversity. Recent genomic studies have greatly advanced our knowledge of *S. suis*, and other next generation sequencing based approaches may build on this to further increase our understanding of *S. suis*.

Metagenomics based surveillance and diagnosis of disease

Culture-independent microbiome sequencing has enabled discovery of new microbes and genetic material as well as great advances in our understanding of microbial ecology. Amplicon sequencing of the bacterial 16S rRNA gene and fungal internal transcribed spacers (ITS) have allowed cost-effective assessment of taxonomic abundance and diversity in different environments. Metagenomic shotgun sequencing additionally allows assessment of genetic functional potential within microbiomes and assembly of metagenome-assembled genomes (MAGs). MAGs are especially useful for taxa without cultured representatives. Metagenomics has also been put forward as the future of clinical diagnostics and surveillance [24, 91–93]. Metagenomics can detect all microbes present in a sample without culturing and predict their resistance to antibiotic or antiviral treatment. This is useful for rapid detection of pathogens and selection of appropriate antibiotics, which may improve treatment outcomes and reduce antibiotic usage.

In the context of academic research, sequencing the microbiomes of disease cases and comparing them with those of healthy controls can provide less biased information on taxonomic and genetic associations than traditional culture-based approaches

targeting known pathogens. We can now gain information on both known and novel disease-associated microbes and assess how the occurrence and abundance of these correlate with disease symptoms, the environment, and the abundance of other species. Detection of health-associated bacteria is of great interest because they may have probiotic potential. Polymicrobial disease and co-infection can also be assessed, and disease-causing but non-invasive species can be discovered. A large effort has gone towards discovering the role of the microbiome in human diseases such as inflammatory bowel disease [94, 95] and colorectal cancer [96, 97]. Studies have also utilized metagenomics for infectious disease in joint fluid [98], blood [99], and respiratory samples [24]. Research on the respiratory microbiome is, however, lagging behind gut microbiome research despite the relevance to infectious disease [91]. Metagenomic studies on infectious disease in livestock are also lacking.

Challenges and open questions

Although a large effort has been dedicated to understanding the ecology and evolution of *S. suis*, there are significant knowledge gaps. There is a lack of research on how commensal and pathogenic *S. suis* lineages colonize piglets and interact with biotic and abiotic factors to cause disease. The transcriptomic and phenotypic differences between commensal and pathogenic strains are not well described, and the exact (combination of) mechanisms involved in *S. suis* virulence and transmission remain unknown. This hinders disease surveillance and the design of effective preventative measures and treatments. Recent amplicon-sequencing studies have shown high *S. suis* prevalence and abundance in the oral cavity of piglets, and conclusions from culture-based work need to be considered with caution because these likely found many false negatives of *S. suis* presence. It is not known whether *S. suis* colonizes the oral cavity of piglets at high abundances in all farms in different countries, or what proportion of the observed *S. suis* load consists of pathogenic and commensal strains.

Knowledge is also lacking on the polymicrobial nature of *S. suis* disease. It is known that *S. suis* infection often co-occurs with infection by other bacteria and viruses, such as PRRSV, but little quantitative data exists on taxa co-occurring with *S. suis* disease. Associations between the commensal microbiota and *S. suis* are also of great interest. Activity of specific bacterial species may prevent *S. suis* invasive disease by interactions with the host immune system or by providing colonization resistance against pathogenic *S. suis* strains. Identifying taxa positively or negatively associated with *S. suis* disease may aid in development of prevention strategies, vaccines, and probiotics. These measures may contribute to reducing the global *S. suis* disease burden and antibiotic use in pig farming.

Aims and outline of this thesis

The main aims of this thesis were to assess the tonsil microbiota and relative abundance of *S. suis* in piglets across different farms and ages, and to determine the tonsil microbiota associated with *S. suis* invasive disease. **Chapter 1** (this chapter) provides a general background on the topics researched in the following chapters. In **chapter 2** we conducted a field study comparing the microbiota of case piglets displaying symptoms of *S. suis* infectious disease with paired asymptomatic control piglets from the same farms. This case-control study included piglets that differed in environment, breed, age, weaning age, diet, and antibiotic treatment, factors which all may shape tonsillar microbiota composition. In **chapter 3** we aimed to study the factors associated with the tonsil microbiota composition of piglets during early life within a single farm. This study was conducted with a planned study setup and controlled environment, and without the use of antibiotics. Chapters 2 and 3 showed that *S. suis* case piglets are characterised by a significantly different microbiome composition compared to asymptomatic controls, but that factors such as the individual farm, room, and litter explained far larger proportions of the total variation in piglet tonsillar microbiota composition.

Chapter 2 also showed that the microbiome of *S. suis* symptomatic piglets had increased resistomes. In **chapter 4** we aimed to assess whether equivalent disease-associated resistome expansion is found in other studies. As metagenomic studies on livestock and/or the oral microbiome are rare, we opted to conduct a meta-analysis of the human gut resistome. In line with our own results, we found that case participants within studies on human diseases treated with antibiotics had larger resistomes than healthy controls. However, some studies on diseases not treated with antibiotics also had case-control differences in ARG abundance.

Furthermore, chapter 2 found that pathogenic *S. suis* strains appear to be rare in the tonsil microbiota of piglets, while commensal *S. suis* colonize at high abundance. Several previous studies have compared the genomes of clinical and non-clinical *S. suis* strains, and others studied the transcriptome of the most pathogenic clades. Commensal *S. suis* clades have received little attention. **Chapter 5** compares the transcriptome of 21 *S. suis* strains from different clades, including both putatively pathogenic and commensal strains. We found that pathogenic and commensal strains shared high transcriptome similarity despite large genomic sequence variation, but that regulatory mechanisms described in pathogenic *S. suis* model strains may not apply to other *S. suis* clades.

Finally, **chapter 6** summarises and discusses the research chapters in a common context, with special attention for the overlap in chapter content shown in Figure 2. I discuss the impact of antibiotic use in human and veterinary medicine on the microbiome, and the ecological niche of commensal and pathogenic *S. suis*. I also speculate on how possible interactions between *S. suis* and the tonsillar microbiota may facilitate invasive disease. I further discuss challenges and opportunities for future research on *S. suis* and the oral microbiome of piglets.

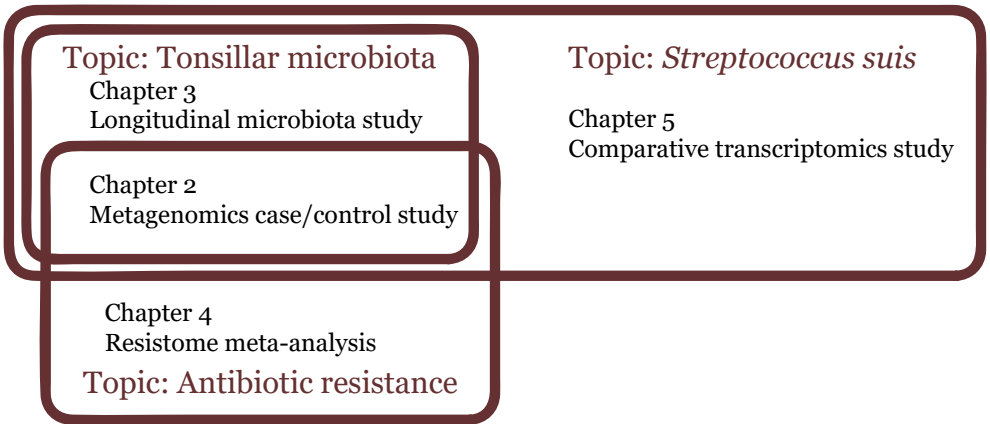


Figure 1.2: Thematic overlap of the thesis research chapters.

Chapter 2

Streptococcus suis infection on European farms is associated with an altered tonsil microbiome and resistome

Simen Fredriksen[1], Carlos Neila-Ibáñez[2,3,4], Isabel Hennig-Pauka[5], Xiaonan Guan[1,6], Jenelle Dunkelberger[7], Isabela Fernandes de Oliveira[1], Maria Laura Ferrando[1], Florencia Correa-Fiz[2,3,4], Virginia Aragon[2,3,4], Jos Boekhorst[1], Peter van Baarlen[1], Jerry M. Wells[1,8]

1. Host-Microbe Interactomics Group, Animal Sciences Department, Wageningen University, Wageningen, The Netherlands
2. Unitat mixta d'Investigació IRTA-UAB en Sanitat Animal, Centre de Recerca en Sanitat Animal (CReSA), Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, 08193, Catalonia, Spain
3. IRTA, Programa de Sanitat Animal. Centre de Recerca en Sanitat Animal (CReSA), Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, 08193, Catalonia, Spain
4. OIE Collaborating Centre for the Research and Control of Emerging and Re-emerging Swine Diseases in Europe (IRTA-CReSA), Bellaterra, 08193 Barcelona, Spain
5. Field Station for Epidemiology, University of Veterinary Medicine Hannover, 49456 Bakum, Germany
6. Schothorst Feed Research B.V., Lelystad, The Netherlands
7. Topigs Norsvin USA, Burnsville, MN, United States
8. Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom

Under review. Available as preprint:
bioRxiv 2022.08.01.500980
https://doi.org/10.1101/2022.08.01.500980

Introduction

Streptoccocus suis is a Gram-positive opportunistic pathogen causing systemic disease in piglets around weaning age. Outbreaks of *S. suis* disease are controlled by metaphylactic use of antibiotics, leading to high levels of antimicrobial resistance in *S. suis* isolates. This is an issue for both animal and human health due to the zoonotic disease potential of *S. suis*. The mechanisms facilitating invasive disease are not known but may involve host and environmental factors. The palatine tonsils are considered a portal of entry for pathogenic strains to cause systemic disease. We hypothesised that tonsil colonization by pathogenic and commensal bacteria may impact on disease risk via colonization resistance and co-infections. We conducted a case-control study on 9 European farms, comparing the tonsil microbiome of piglets with *S. suis* systemic disease with asymptomatic controls. We also compared these to piglets on control farms and piglets reared naturally in a forest.

Results

We found a small but significant difference in the tonsil microbiota composition of case and control piglets. Case-control associations varied between amplicon sequence variants (ASVs) and metagenome assembled genomes (MAGs) within the same species. Variants of putatively commensal taxa including *Rothia nasimurium* were reduced in abundance in case piglets compared to asymptomatic controls. Case piglets had higher relative abundance of *Fusobacterium gastrosuis*, *Bacteroides heparinolyticus*, and uncultured *Prevotella* and *Alloprevotella* species. There was, however, no higher abundance of *S. suis* itself at the species-level or of clinical strain marker genes in case piglets. Piglets sampled prospectively weeks prior to developing clinical signs had reduced microbiota alpha diversity. Despite case-control pairs receiving equal antimicrobial treatment, case piglets had higher abundance of antimicrobial resistance genes (ARGs) conferring resistance to antimicrobial classes used to treat *S. suis*.

Conclusions

The tonsillar microbiota of *S. suis* case piglets had increased abundance of taxa not previously linked to *S. suis* disease. This coincided with increased ARG abundance in case piglets, possibly due to adaptation of disease-associated strains to frequent antimicrobial treatment.

Abstract

Introduction

Streptococcus suis is a Gram-positive bacterium colonizing the upper respiratory tract of pigs. It is one of the major bacterial causes of disease in pigs and a zoonotic pathogen causing sepsis and meningitis in humans [10, 11, 54, 55]. Swine infections are prevented by metaphylactic use of antimicrobials. This has led to increased antimicrobial resistance (AMR) in *S. suis* isolates [100], with macrolide and tetracycline resistance genes *erm(B)* and *tet(O)* being the most common [12]. The spread of antimicrobial resistance genes (ARGs) in zoonotic *S. suis* and to other streptococci is of concern to veterinary and human medicine [101].

The palatine tonsils have been suggested as a main habitat for *S. suis* colonization and a site of entry into the host bloodstream [66–73]. A recent study identified differences in taxonomic composition of the tonsillar microbiota in piglets diagnosed with *S. suis* disease [102]. Microbiota associations with *S. suis* infectious disease are of great interest to understand co-infection dynamics and to identify probiotic candidate species that may provide colonization resistance. It is thought that co-infection by bacterial and viral pathogens of the porcine respiratory disease complex (PRDC) may predispose *S. suis* invasive disease [74, 76, 77].

Tonsillar colonization by different *S. suis* strains may also have an impact on invasive disease risk [66, 103]. While *S. suis* has been described as an opportunistic pathogen, different strains have varying virulence potential and can be grouped into commensal and pathogenic clades [64, 65, 104]. Strains from pathogenic *S. suis* clades have reduced genome sizes [64] and putatively reduced ability to persist as colonizers in competition with commensals. Strains from pathogenic clades can also be isolated from the tonsils of healthy pigs, but the prevalence of asymptomatic carriage and whether *S. suis* invasive disease is preceded by outgrowth of disease-associated strains is not known.

This study was aimed at identifying the tonsillar microbiota associated with *S. suis* systemic disease in weaning age piglets. We utilized a case-control study design and next generation sequencing to characterise the composition of the tonsillar microbiota of piglets with *S. suis* systemic infections on nine European farms and compared these to piglets from four farms without *S. suis* disease. We used 16S rRNA gene amplicon- and metagenomic shotgun sequencing to quantify the tonsillar microbiota and resistome and analysed clinical and non-clinical *S. suis* strains from the farms by whole genome sequencing. This study design allowed us to assess microbiome composition and predict functionality, as well as to assess the abundance of marker genes prevalent in clinical or non-clinical *S. suis* strains.

Results

Sequencing

We sampled 45 case-control pairs of piglets with *S. suis* systemic disease clinical signs and asymptomatic control piglets from the same pen. Case-control pairs were equivalent in age, genetic background, diet, and antimicrobial treatment. Additional control piglets from case farms and from control farms without *S. suis* outbreaks were sampled to assess farm differences. We also sampled one US farm as well as piglets raised organically in a forest for comparison. Amplicon sequencing of the full set of 295 samples yielded an average of 76757 reads per sample after processing with DADA2 [105]. We shotgun sequenced metagenomic DNA from all case-control pairs and a subset of piglets from control farms. A total of 109 metagenomic samples yielded on average 7.5 million reads per sample after removal of host DNA and plant DNA from feed.

The *Streptococcus suis* disease-associated microbiota

Case and control piglets on outbreak farms had significantly different tonsillar microbiota composition, but the effect size was small ($R2 = 0.01$ and $p = 0.02$, PERMANOVA on Bray-Curtis dissimilarity comparing case and control piglets within outbreak farms). Analysis per country found varying case-control results for Spain ($R2 = 0.030$, $p = 0.01$) and Germany ($R2 = 0.017$, $p = 0.93$), which had fewer samples. The largest difference in tonsil microbiota composition was found on the Dutch farm NL1 ($R2 = 0.26$, $p < 0.01$), but this might be due to all 3 case piglets from this farm being sampled shortly after death allowing opportunistic pathogens to bloom. We did, however, not see a general association between the microbiota and clinical sign severity or (future) death within the other farms. There was no consistent difference in alpha diversity between case and control piglets (Figure S1).

We identified amplicon sequence variants (ASVs) significantly associated with piglet case-control status (Figure 1A). SIAMCAT [106] identified three ASVs with significantly ($p < 0.05$) higher abundance in case piglets: ASV 29 (1.8% vs 0.7% mean abundance, *Prevotella* sp.), ASV 3 (3.6% vs 2.2%, *Alloprevotella* sp.), and ASV 21 (3.6% vs 1.5%, *Fusobacterium gastrosuis*). These were all part of the core microbiota in the European farms (minimum 99% prevalence in control piglets). The case-associations were stronger in the Dutch and Spanish farms, where the outbreaks were larger and/or symptoms more severe than in the German farms. These three ASVs also had the strongest association with case piglets using Wilcoxon Rank Sum Test ($p < 0.005$), but no ASVs were significantly different when applying multiple testing correction (FDR > 0.05). No genera were significantly different between case and control piglets (FDR > 0.05).

While the three case-associated ASVs had the strongest case-control associations, few other ASVs trended towards case-associations. A larger number of different ASVs had weaker, non-significant (FDR > 0.05), control-associations. Among core microbiota members, ASV 48 (0.38% vs 0.74%, *Clostridium disporicum*), ASV 36 (0.47% vs 0.91%, *Rothia*

nasimurium), and ASV 6 (1.98% vs 2.62%, *S. suis*) had the strongest control associations. ASV 11 (0.27% vs 0.63%, *Actinobacillus minor*, not present in farms ES4 and NL1) and ASV 81 (0.22% vs 0.40%, *Terrisporobacter mayombei*) were the non-core ASVs with the strongest control-associations. In terms of total relative abundance, ASV 1 (3.5% vs 8.1%, *Moraxella porci/pluranimalium*) showed the largest difference between case and control piglets. ASV 1 had strong control associations in farm ES2 (0.9% vs 5.6%, being completely absent in several case piglets) and ES3 (15% vs 24%) but was absent from farm ES4 and NL1 and equally abundant in German case and control piglets.

Porcine respiratory disease complex (PRDC) and its range of bacteria thought to co-infect with *S. suis* is mainly associated with respiratory disease in older pigs, while this study investigates weaning age systemic disease. Still, PRDC-associated bacteria were of interest in possible co-infections. *S. suis* itself and *Glaesserella parasuis* were among the most abundant tonsil microbiota members. Relative abundance of *S. suis* ASVs trended towards being higher in healthy controls, while *G. parasuis* ASVs had mixed associations. ASV 174 (100% identical to both *G. parasuis* and *Actinobacillus indolicus*) trended towards case-association (0.27% vs 0.16%). *Pasteurella multocida* was relatively prevalent (50%) but at low abundance (mean 0.06%), although two outlier case piglets had over 1% abundance. *Actinobacillus pleuropneumoniae* had 34% prevalence and 0.08% mean abundance but was most abundant in control piglets. *Bordetella bronchiseptica* had lower prevalence (17%) but higher abundance (0.10%) due to some control piglets having up to 12% abundance. *Mycoplasma hyopneumoniae* was virtually absent from the dataset, with only 6 reads from a single control piglet.

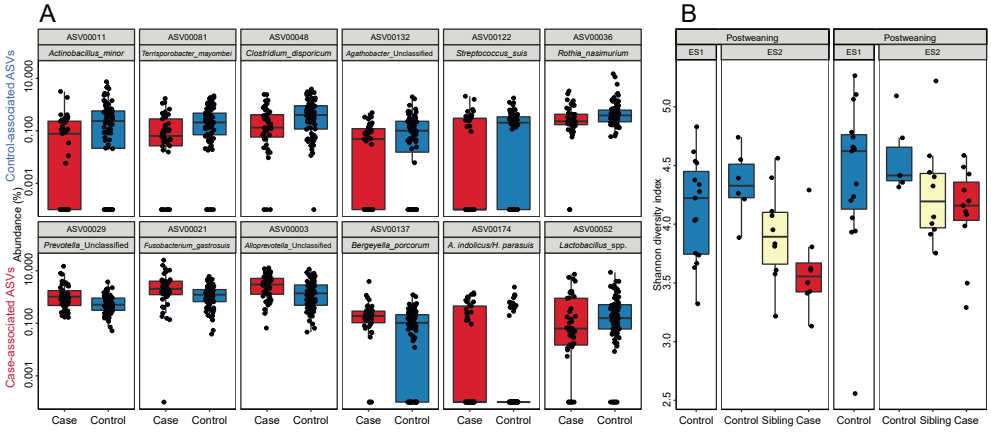


Figure 1: Microbiota case-control associations. **A.** ASVs significantly associated with case-control status on the outbreak farms, sorted by strongest case-control association with Wilcoxon Rank Sum Test. Top row: control-associated ASVs, bottom row: case-associated ASVs. **B.** Shannon diversity (amplicon sequencing) of prospective samples collected pre-weaning from piglets that developed *S. suis* clinical signs 2-4 weeks later. Samples from future case piglets had lower Shannon diversity than control piglets that remained asymptomatic. Piglets from control farm ES1, which was managed by the same company and used sows from ES2, also had high mean diversity.

Early life microbiota diversity predicted clinical sign appearance weeks later

To assess the value of tonsillar microbiota composition to predict future invasive *S. suis* disease, we collected prospective samples on farm ES2, which had a history of post-weaning *S. suis* outbreaks, and farm ES1, a control farm without recorded *S. suis* problems despite using sows produced on farm ES2. A cohort of 40 piglets were sampled on farm ES2 one week before weaning, prior to visible clinical signs. Eight of the 40 piglets developed *S. suis* clinical signs 1-3 weeks post-weaning, and tonsil samples collected from these piglets post-weaning were included in the 45 case-control pairs.

Comparison of prospective samples collected from case and control piglets one week before weaning, 2-4 weeks before disease onset, revealed that the tonsil microbiota of case piglets had significantly lower Shannon diversity compared to control piglets that remained asymptomatic ($p = 0.005$; Figure 1B). Case piglets trended towards having lower diversity also post-weaning, during the outbreak, but the effect was smaller and less significant ($p = 0.12$). Asymptomatic siblings of case piglets also had a lower diversity than control piglets from litters without case piglets and piglets from control farm ES1 (Figure 1B). The difference in composition was, however, smaller pre-weaning ($R2 = 0.06$, $p = 0.15$, PERMANOVA on Bray-Curtis dissimilarity) than during the outbreak ($R2 = 0.13$, $p < 0.01$). The Spearman correlation between ASV ratio of case-control abundance pre- and post-weaning on the farm was low ($R = 0.09$, $p = 0.4$), showing that separate ASVs were differentially abundant and driving case-control compositional differences pre- and post-weaning. The strongest pre-weaning prospective case-associations were ASV 88 (*Pasteurellaceae*, 1.4% vs 0.55%) and ASV 186 (*Actinobacillus indolicus/minor*, 1.1% vs 0.21%), while ASV 77 (Leptotrichia, 0.11% vs 0.33%) and ASV 22 (Actinobacillus minor, 0.06% vs 0.20%) had the strongest control associations.

Metagenome-assembled genomes (MAG) analysis

We shotgun sequenced all case-control pairs and a subset of piglets from control farms. Co-assembly per farm produced 802 metagenome-assembled genomes (MAGs) (>70% completeness and <10% contamination). Read mapping to the MAGs yielded an overall case-control compositional difference similar to that found by amplicon sequencing analysis (Figure 2, $R = 0.03$, $p < 0.01$, Bray-Curtis dissimilarity PERMANOVA). The case-control associated MAGs largely corresponded to the ASVs identified by amplicon-based analysis, but some taxa, including case-associated *Bacteroides heparinolyticus*, were only identified using MAGs. The MAG approach resulted in stronger case-control associations (FDR < 0.0001) than the ASV approach, suggesting that the case-associated ASVs represented several strains with variable case-control associations. The two ASVs with strongest disease-association putatively corresponded to a collection of different *Bacteroidales* MAGs. The MAGs with the strongest disease-associations were classified as *Bacteroidales* UBA1309 and *Alloprevotella* F0040, clades poorly described and with few or no publicly available genomes from related isolates.

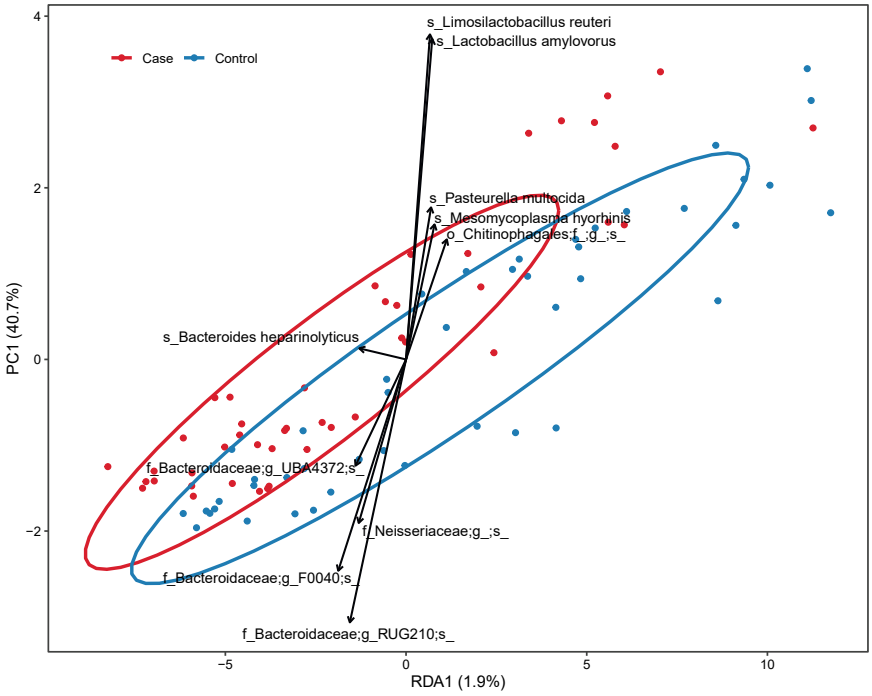


Figure 2: Redundancy analysis (RDA) on the abundance of metagenome-assembled genomes (MAGs) constrained by case-control status. The points represent individual samples; samples in the direction of arrows have more of that taxon. The eclipse represents 75% confidence level.

The tonsillar microbiome of piglets is similar between farms but diverged in free-range forest piglets

The tonsillar microbiota composition was more similar between piglets on the same farm than to that of piglets from other farms (Figure 3A-B). Compared to tonsillar microbiota compositions of farm piglets, Tamworth free-range piglets living outdoors in a forest in the Netherlands had a strikingly different microbiota and resistome (Figure 3C-D). While European farm piglets shared a large core microbiota with 83 ASVs being present in 80% of piglets or more, 44 of these were not found in the free-range piglets. Vice versa, of the 157 ASVs found in all 5 forest piglets, 117 were not found in any farm piglet. The free-range piglets had low abundance of the genera most abundant in farm piglets, in particular *Moraxella* (0.6% vs 13%) and *Streptococcus* (0.9% vs 12%), and higher abundance of Acinetobacter (10% vs 2.3%), "*Rikenellaceae* RC9 gut group" (9.9% vs 0.1%), and *Treponema pedis* (3.6% vs 0.04%). The sample with the lowest *S. suis* abundance in the study, 0.3% abundance of a single ASV (ASV 1050), was from a free-range piglet. This ASV did not have 100% identity to the 16S rRNA gene V3-V4 region of any *S. suis* strain publicly available in SILVA or NCBI assembly databases. The other free-range piglets were colonised by *S. suis* ASVs shared with farm piglets.

Increased tetracycline ARG abundance in case piglets

We quantified the abundance of antibiotic resistance genes (ARGs), collectively called the resistome, by mapping metagenomic reads to the Resfinder database [107] and normalizing abundance by fragments per kilobase reference per million fragments (FPKM). Piglets on most farms received antibiotics via feed or water, and these were included for resistome analysis. Twenty-two piglets received intramuscular injections, and these were excluded from the main resistome analysis. All case-control pairs included in the main analysis had received equal antimicrobial treatment. Farms varied in total ARG abundance (Figure 3D), but shared high abundances of the most abundant ARGs, such as *blaROB-1* conferring resistance to penicillin/amoxicillin/ampicillin (411 FPKM mean abundance), *sul2* conferring sulfamethoxazole resistance (338 FPKM), and streptomycin resistance genes *aph(3")-Ib* (312 FPKM) and *aph(6)-Id* (220 FPKM). The ARGs most common in *S. suis*, *erm(B)* and *tet(O)* [12], were less abundant (66 and 111 FPKM, respectively). When overall ARG composition was compared by PCA, samples clustered by country, except for Spanish pre- and post-weaning samples which clustered separately by age and not by farm (Figure 3C).

ARG abundance may, in addition to by taxonomic composition, be influenced by both historical antimicrobial usage on the farm and by direct antimicrobial treatment of the sampled piglets. This study could not disentangle these two effects since most antimicrobial treatments were given equally to all sampled piglets within each farm. Farm NL1, a high health status research farm where the sampled piglets were not treated, and where piglets are rarely treated with antimicrobials, had the lowest ARG abundance (except for the free-range forest piglets). In Germany, high health status farms DE1 and DE2 had low ARG abundance, but so did DE6 despite a history of severe *S. suis* disease. Farms DE3, ES3, and ES4 were assessed to have the lowest health status and highest historical antimicrobial usage by veterinarians, and piglets on these farms were also administered antimicrobials before sampling. These three farms had the highest ARG abundance (Figure 3D). There was no consistent link between the antimicrobials administered and the abundance of ARGs conferring resistance to these. On farm DE1, piglets that had received tetracycline had lower tetracycline ARG abundance than untreated piglets (810 vs 391 FPKM). On farm DE3, where all piglets had received tetracycline, tetracycline ARG abundance was lower than on DE2 and DE5 where piglets had received no antimicrobial treatment (804 vs 805 and 926 FPKM, respectively).

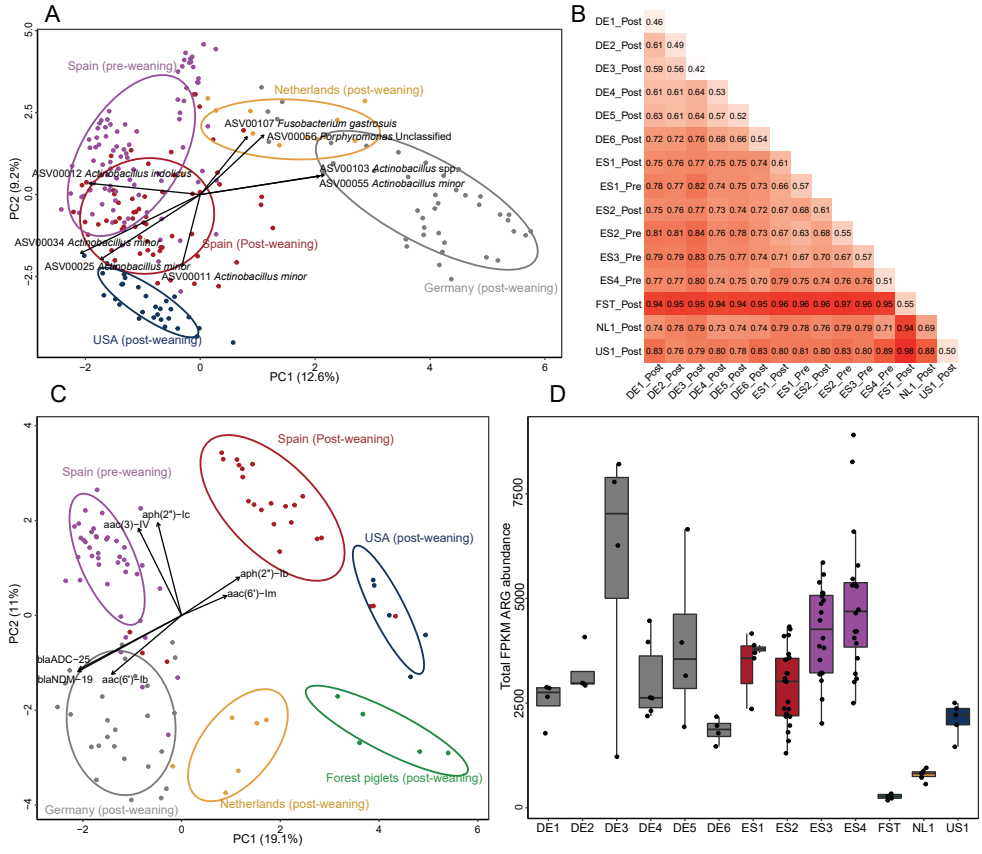


Figure 3: Microbiome differences between farms. **A.** PCA on microbiota composition (ASV abundance) of all farm samples and countries, pre- and post-weaning (transformed with log(1000*abundance+1)). Samples clustered broadly by country, but samples from farm DE6 clustered with NL1, and samples from ES4 clustered away from the other Spanish farms. **B.** Mean pairwise Bray-Curtis dissimilarity between samples from the different farms. **C.** PCA on ARG abundance (transformed with log(1000*abundance+1)). The pre- and post-weaning Spanish samples separated in 2 clusters. **D.** Total abundance of all ARGs per farm.

The total abundance of ARGs was 15% higher in case piglets compared to controls (3473 vs 3026 mean FPKM), but this overall effect was not statistically significant ($p = 0.36$). ARGs conferring resistance to tetracycline (class level) were strongly case-associated (788 vs 620 FPKM, $p = 0.01$), especially within farms ES2 and ES3. Specifically, tetracycline and doxycycline resistance, conferred largely by the same ARGs, were the most abundant Resfinder ARG phenotype categories and had the strongest case-associations (Figure 4). *Tet(Q)* was the individual gene with the strongest case-association (134 vs 75 FPKM, FDR < 0.01). ARGs of classes aminoglycoside, beta-lactam, and macrolide, which like tetracyclines are commonly used to treat *S. suis* disease, also trended towards higher abundance in case piglets.

We investigated the presence of case-associated ARGs in MAGs to determine if the higher ARG abundance was linked to the case-associated taxa. *Tet(Q)* was only found in case-associated *Prevotella* MAGs, and case-associated ARGs *ant(6)-Ib* and *tet(44)* were only found in case-associated *Fusobacteriales* MAGs. ARGs found in control-associated *Rothia* and *Clostridium* MAGs, *lnu(P)*, *erm(Q)*, *tetA(P)*, and *erm(X)*, were not case-associated.

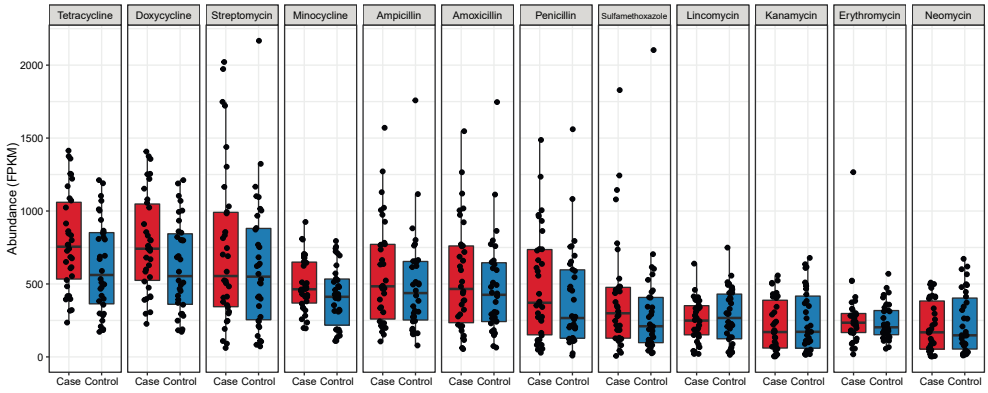


Figure 4: The abundance of the 12 most abundant Resfinder ARG phenotype categories in case and control piglets. Note that most genes confer resistance to several antimicrobials, making the sum of phenotype abundances greater than the total ARG abundance. For example, tetracycline and doxycycline resistance is conferred by the same genes.

Comparison of the tonsil and nasal microbiota

To evaluate whether differences in microbiota composition between case and control piglets were specific to the tonsillar microbiota or a general trend in the upper respiratory tract, we collected 41 nasal swabs from three Spanish farms, ES1-3. Nasal and tonsillar swab microbiota composition was significantly different, but more variation was explained by the farm ($R2 = 0.08$, $p < 0.001$, Bray-Curtis PERMANOVA) than nasal/tonsil sample location ($R2 = 0.06$, $p < 0.001$). The nasal microbiota was characterised by higher abundance of genera *Moraxella* (32% vs 16%) and *Bergeyella* (5% vs 1%) and lower abundance of most other genera. Case and control piglets had a significantly different nasal microbiota ($R2 = 0.07$, $p = 0.04$, Bray-Curtis PERMANOVA). This effect size was comparable to the difference found in the tonsillar microbiota samples for the same subset of piglets ($R2 = 0.07$, $p = 0.09$). However, the ASVs that were differentially abundant in the nasal microbiota of case and control piglets were not the same ASVs that were differentially present in tonsillar microbiota. Some ASVs had inverse case-control associations: ASV 21 (*F. gastrosuis*), ASV 29 (*Prevotella* sp.), and ASV 3 (*Alloprevotella* sp.) were disease-associated in the tonsillar microbiota, both overall and within ES2 and ES3, but health-associated in the nasal microbiota.

S. suis diversity in the tonsillar microbiota

Amplicon sequencing data showed higher *S. suis* abundance in control piglets compared to case piglets within farms with pre-weaning outbreaks (farms ES3 and ES4, 3.9% vs 6.4%, $p = 0.02$). This effect was not significant within post-weaning outbreak farms (DE1, DE3-6, ES2, NL1, 4.4% vs 5.1%, $p = 0.27$).

In total 89 ASVs were classified as *S. suis*, and 52 of these were present in two or more tonsil samples, suggesting that they are unlikely to be sequencing artefacts. Farm piglets were on average colonised by 7 *S. suis* ASVs, with a range between 2 and 16. *Streptococcus suis* relative abundance was similar in outbreak and control farms (Figure 5A). Comparison of the 89 *S. suis* ASVs to the 16S rRNA gene V3-V4 region of 2463 *S. suis* assemblies available on NCBI assembly revealed that many ASVs were poorly or not at all represented by sequenced genomes. This likely relates to the fact that many clinical but few non-clinical strains have been sequenced. ASVs are, however, not good markers for assessing *S. suis* strain diversity as the 16S rRNA gene V3-V4 region correlates poorly with whole genome phylogeny, and as both clinical and non-clinical strains share the same ASVs. 82% of all *S. suis* assemblies had a 16S rRNA gene V3-V4 region amplicon identical to ASV 6, which comprised 30% of *S. suis* in the tonsillar microbiota. ASV 17, the second most abundant *S. suis* ASV, comprised 21% of all *S. suis* in the microbiota but was only found in 0.08% of the assemblies. In addition to isolate DE512T1 collected in the present study, ASV 17 was only found in two isolates recently sampled from diseased pigs in China (GCF_019793915.1 and GCF_019794525.1).

We used metagenomic data to assess the relative proportion of commensal and pathogenic *S. suis* in the tonsillar microbiota of each piglet. We created a *S. suis* pangenome by clustering protein coding genes from a previously published genome collection [104] at 80% identity, and mapped metagenomic reads to the representative sequence of each gene cluster to assess their abundance in each sample. We calculated the ratio of prevalence of each cluster in clinical and non-clinical genomes to assess their putative association with pathogenicity. We found that in the tonsillar microbiota, genes predominantly found in non-clinical strains occurred at higher abundance than genes common in clinical isolates (Figure 5B). There was a small positive correlation between gene clinical/non-clinical genome presence ratio and case/control sample abundance ratio (Figure 5C, Spearman $R = 0.05$, $p < 0.01$), due to higher abundance of commensal *S. suis* in control samples.

One often used *S. suis* marker gene for strain virulence is the gene encoding suilysin (*sly*), and this gene was found in at least 1 clinical strain genome from all farms where we sequenced clinical strains. *Sly* is well suited for use in metagenomics analysis due to high sequence conservation across *S. suis* strains. *Sly* abundance was low in the tonsillar metagenomes, and similarly abundant in case and control piglets (Figure 5D). On some farms we did not detect *sly* in any piglets, but this may be due to small sample size and insufficient sequencing depth.

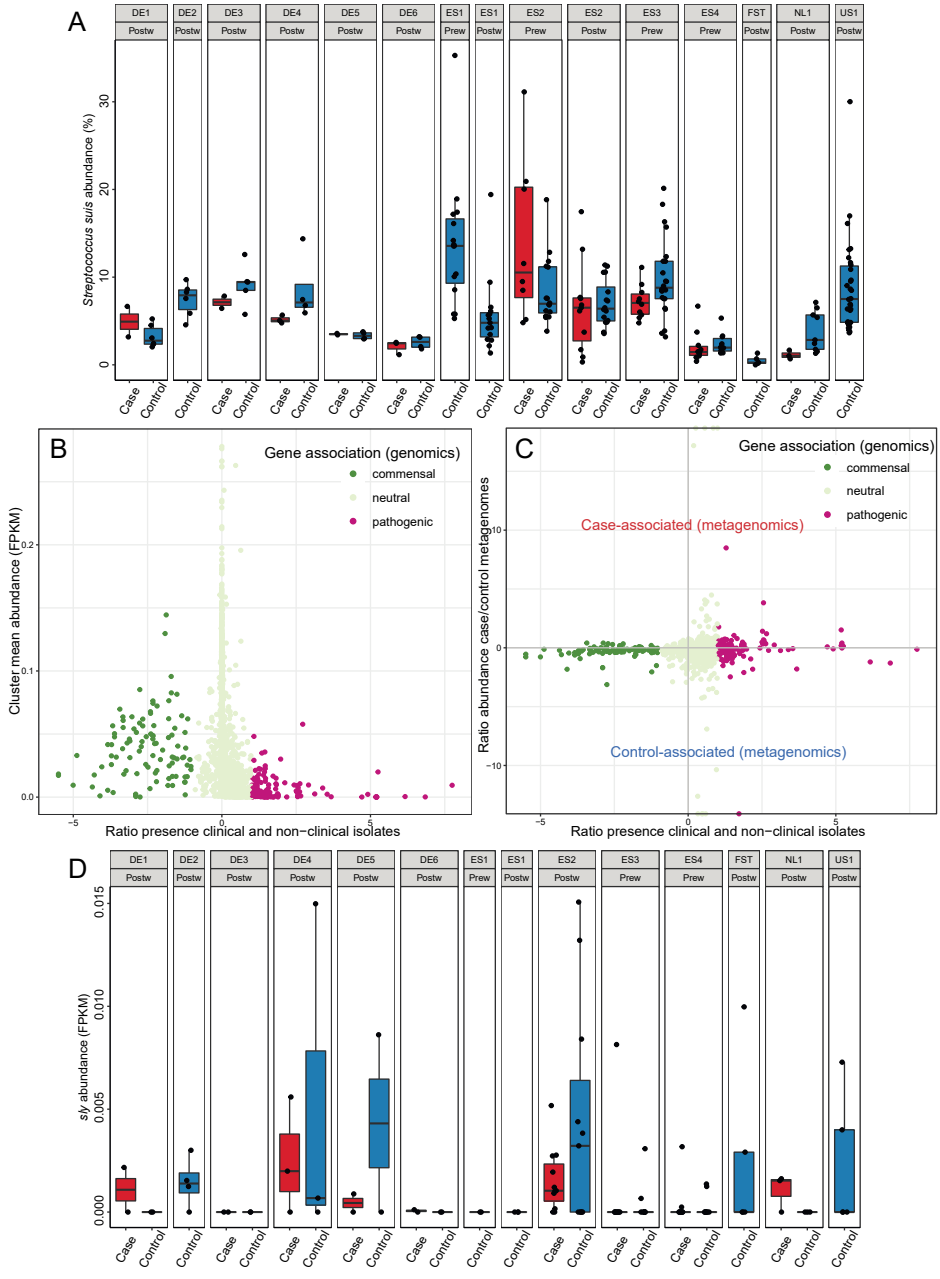


Figure 5: *S. suis* abundance and diversity. **A)** Total *S. suis* relative abundance per farm, amplicon sequencing data. **B)** Within the *S. suis* pangenome, genes associated with non-clinical strains (x-axis) were more abundant (y-axis) in metagenomic samples from the tonsillar microbiota. Genes with ratio 0 are equally common in clinical and non-clinical strains, genes with a ratio of 1 are twice as common in clinical strains. **C)** The correlation between gene association with clinical/non-clinical genomes (x-axis) and association with abundance in case-control metagenomic samples (y-axis). **D)** The abundance of suilysin encoding gene *sly*, a gene highly conserved in clinical strains, in case and control samples per farm.

Discussion

In this study, we found that the tonsil microbiota composition of case piglets with *S. suis* clinical signs is significantly different from the microbiota composition of asymptomatic controls. *Streptococcus suis* disease may occur as a part of polymicrobial infections collectively known as porcine respiratory disease complex (PRDC), which includes porcine reproductive and respiratory syndrome (PRRS) virus, swine influenza A, and potentially bacterial primary and secondary (opportunistic) pathogens [27, 76]. We did not find disease-associations with the tonsil microbiota abundance of any species linked to PRDC, such as *G. parasuis* or *S. suis* itself. This study included only piglets with weaning age systemic *S. suis* disease, and PRDC associated taxa may be more relevant to respiratory disease in older finisher pigs.

We identified novel disease-associations with *Fusobacterium gastrosuis*, *Bacteroides heparinolyticus*, and uncultured *Prevotella* and *Alloprevotella* species. *Fusobacterium gastrosuis* has previously been linked to *Helicobacter suis* infection in the gastric microbiota and shown to have genes involved in adhesion, invasion and induction of cell death as well as in immune evasion in other *Fusobacterium* species [108]. The uncultured case-associated *Alloprevotella*/*Prevotella*/*Bacteroides* species lack isolate genomes and are unknown in relation to *S. suis* disease, possibly due to being unculturable. Case-associated taxa may also interact with the host to facilitate *S. suis* to cross epithelial barriers without entering the bloodstream themselves, thus remaining undetected by necropsy. Alternatively, disease-associated taxa may increase in abundance due to host immune status and dysbiosis, as suggested for oral *Prevotella*/*Alloprevotella* species in humans [109].

Case piglets had lower tonsillar *S. suis* abundance than control piglets. We assessed that this was due to a reduced abundance of strains from commensal clades, most of which are poorly represented among sequenced *S. suis* strains. While *S. suis* genes predominantly found in non-clinical strains were more abundant in control piglets, genes predominantly found in clinical strains, such as *sly*, were low in abundance and more equally distributed between cases and controls. This shows that the majority of *S. suis* colonising tonsils are commensal, lacking genes required for invading the host, but also confirms that strains carrying genes most prevalent in clinical strains are also colonising asymptomatic piglets at low abundance. Based on these results, we conclude that tonsillar colonisation by *S. suis* itself cannot be used to reliably predict or even confirm ongoing *S. suis* invasive disease.

Associations between *S. suis* disease and the tonsillar microbiota have been investigated using amplicon sequencing in a previous study by Niazy et al. [102]. *Bacteroides* and *Lachnospiraceae* were found to be more abundant in controls, while case piglets had higher abundance of *Campylobacter* and *Porphyromonas*. This may be due to European and North American piglets having fundamentally different microbiomes. Differences between studies may largely be due to methodological differences, but in the present study we included one US farm and found the sampled piglets to have low diversity, and

that while microbiota members were shared at the ASV level, composition was dominated by high *Actinobacillus* and *Streptococcus* abundance. The different findings in Niazy et al. might also be due to their sampling of whole tonsillar tissue, while we used swabs. Furthermore, piglets suffering from other pathologies such as rectal prolapse and hernia were sampled as controls, so the case piglets were not compared with healthy controls as in the present study. While their study may include control-associated taxa associated with other disease, our results may not be specific only to *S. suis* disease but include microbiome traits associated with low health status in general. Niazy et al. also sampled uneven numbers of case and control piglets per farm, and from some farms no controls. This confoundment of case-control and farm comparison may have affected the results.

We found case piglets to have higher abundance of antimicrobial resistance genes than controls, despite case-control pairs being treated with the same antimicrobials. ARGs conferring resistance to doxycycline and tetracycline had the strongest case-association, and ARGs conferring resistance to other antimicrobials used against *S. suis* also trended towards being more abundant in case piglets. ARGs found in MAGs of disease-associated taxa *Fusobacterium*, *Prevotella*, and *Alloprevotella* had strong disease-associations compared to those found in health-associated *Rothia* and *Clostridium* MAGs. It is possible that frequent antimicrobial usage has caused selection pressure on disease-associated taxa, leading to accumulation of ARGs.

While antimicrobial treatment may have caused strong selective pressure over time, we found limited increases in abundance of ARGs conferring resistance to the antimicrobials used to treat the sampled piglets. It was unexpected to find that antimicrobials administered by water and feed did not appear to have led to increases in ARG abundance. Antimicrobial usage starting in the first few days after birth has previously been shown to influence piglet nasal microbiota composition and diversity [110], but in the present study most antimicrobial treatment started later in life, and only a few days before sampling. Another plausible explanation for the limited effect of antimicrobial treatment on the tonsillar microbiota in our study is the intrinsic resistance of biofilms to antibiotics [111]. The route of antimicrobial administration is known to influence the impact on the gut microbiota [112, 113], but it is not known to what extent antimicrobials provided in water, feed, or by injection are able to penetrate oral biofilm.

Farms had significant differences in microbiota composition, but clustered by country. This may be due to both sampler bias (a single person collected all samples in each country) and farm environment, practices, and regulations that vary by country. The piglet microbiota may for instance be influenced by factors such as cleaning, feed composition, temperature, and antimicrobial treatment [90, 114–120]. However, the microbiota composition was not more similar on farms with comparable management practices. The genetics and source of the sows, and their vertically transferred microbiota may be a more important determinant of the piglet microbiota than farm conditions. Among the Spanish farms, ES1 and ES2, which were operated by the same company and had frequent exchange of

animals, had the most similar microbiota composition to each other. However, unrelated German farms shared a more similar microbiota composition than any of the Spanish farms. In both Spain and Germany, farms were spread over a large geographic area and managed by different companies.

We sampled the tonsillar microbiota of five piglets living outdoors in a forest in the Netherlands. These piglets have no recorded problems with diseases commonly affecting pigs on intensive farms, including *S. suis* associated disease. We found the tonsillar microbiota of the forest piglets to be fundamentally different from farm piglets, and to have lower ARG abundance. Core ASVs in farm piglets were completely absent in the forest piglets, and core ASVs from the forest piglets were not found in farm piglets. The ecological farm DE4, where piglets had straw bedding and outside access, did not have a more similar microbiota to the forest piglets than other farms. Compared to DE4 and other farms, living outdoors may shape the tonsillar microbiota via exposure to environmental microbes from soil, diverse natural feed sources, high air quality, and lower exposure to bacterial transfer from other piglets. Previous studies have shown differences in the faecal [121–123] and nasal [124] microbiota of wild pigs, but we are not aware of studies on specific factors that drive differences between the natural and domestic microbiota or between farms with varying disease problems. Understanding these factors may be key to preventing disease by opportunistic pathogens in pig farming.

We conducted a longitudinal sampling on farm ES2, collecting prospective samples from pre-weaning, prior to clinical sign development. Future case piglets had reduced alpha diversity compared to asymptomatic controls and piglets from control farm ES1. Asymptomatic siblings of case piglets had alpha diversity intermediate between symptomatic siblings and control piglets from litters without any *S. suis* cases. *S. suis* disease cases were concentrated in a limited number of litters, despite most of the herd remaining unaffected. This suggests that maternal effects involving early life immunity- or microbiota may be important in predisposing *S. suis* disease and potential dysbiosis. This may be due to vertical transmission of a disease-prone microbiota, but also differences in maternal immunity, with antibodies depleting prematurely [125]. *S. suis* disease most commonly occurs around the time colostral antibodies to *S. suis* start to diminsh. Various studies have found other colostral antibodies to have similar half-lives as well as high variation in antibody abundance and half-life between individual piglets and litters of different sows [126–130]. Thus, it is possible that during *S. suis* outbreaks, some piglets have sufficiently high levels of maternal antibodies to opsonise invading *S. suis*, whereas other piglets lack a sufficient level of maternal immunity.

In conclusion, there are small but significant differences between the tonsillar microbiota of *S. suis* case piglets and asymptomatic controls. We discovered novel taxa associated with case piglets, while *S. suis* abundance was higher in controls. The microbiota differences may originate from dysbiosis starting during early life prior to disease outbreak, but further research is needed to assert this. It is also not conclusively

known whether *S. suis* invades through the tonsils or other locations. Case piglets had higher abundance of ARGs conferring resistance against classes commonly used to treat *S. suis* disease. This may be linked to high ARG prevalence in case-associated taxa driven by more frequent exposure antimicrobial treatment than control-associated taxa.

Materials and methods

Animals

We utilized a case-control study design to assess the association between tonsillar microbiota composition and incidence of *S. suis* invasive disease. To search for consistent correlations between microbiota composition and *S. suis* disease incidence we included farms from different countries with different livestock management systems (see supplementary text 1). Tonsil swabs were obtained from 3- to 10-week-old piglets at 13 farms, of which 9 had ongoing *S. suis* disease outbreaks. Three sampled farms had no history of *S. suis* disease. We compared our European farms to samples obtained at one US farm with a history of *S. suis* disease but no cases at the time of sampling. Lastly, we sampled 5 piglets free-living in a forest in the Netherlands at approximately 1 month after separation from the sows.

We selected 45 pairs of case-control piglets for metagenomic sequencing (Table 1). The pairs were selected to be as equal as possible, coming from the same pen, room, and/or sow, and having received equivalent antimicrobial treatment. The cross-sectional case-control study design was extended with longitudinal sampling at two farms, ES1 and ES2, to determine whether microbiota differences in early life may predict future *S. suis* disease occurrence. ES1 is a production farm free of *S. suis* disease despite all sows and their carried microbiota and *S. suis* strains originating from ES2, which had severe *S. suis* disease problems.

Clinical signs consistent with *S. suis* infection were recorded at each farm visit. The cases fell into two categories: arthritis (typically presenting as lameness) and meningitis (including otitis and sepsis, typically presenting as loss of balance, paralysis, paddling, shaking, and convulsing). Confirmation of *S. suis* infection was not carried out on the individual piglets due to welfare reasons. Most sampled piglets recovered after antimicrobial treatment. *S. suis* disease was confirmed by necropsy in all 3 case piglets on farm NL1.

This study and all animal procedures were approved by the appropriate ethical committees. Sampling of the forest piglets was conducted according to the restrictions of the Animal and Human Welfare Codes in The Netherlands (2019.W-0026.001). On other farms, sampling was carried out for diagnostic purposes (covered by EU Directive 2010/63/EU).

Whole genome sequencing of bacterial isolates

We selected 9 clinical (isolated from lesions observed at necropsy on the farms, but not necessarily from the same piglets sampled for tonsillar microbiota) and 7 non-clinical (from tonsillar swabs) *S. suis* isolates for whole-genome sequencing. The isolates were grown in Todd-Hewitt broth with yeast extract overnight, and DNA was isolated with PowerSoil DNA Isolation Kit (Qiagen) with 0.1mm silica bead beating. Isolated DNA was paired-end Illumina sequenced. Reads were trimmed with trimmomatic v0.39 [131], assembled with Spades v3.14.1 [132], and annotated with Prokka v1.14.5 [133].

Sample collection

The palatine tonsil microbiota of piglets was sampled by gently scraping the tonsillar surface with HydraFlock swabs (Puritan, ME, USA) for 10 seconds. Swabs were immediately put in vials containing Powerbead solution (Qiagen, the Netherlands) and transported at -20 ºC before being stored at -80 ºC. DNA was isolated using the PowerSoil DNA Isolation Kit with 0.7mm garnet bead beating according to the manufacturer's recommended protocol.

Table 1: Number of case and control samples collected from each farm.

			Amplicon		Metagenomics	
Farm	**Country**	**Age**	**Case**	**Control**	**Case**	**Control**
DE1	Germany	Postweaning	2	6	2	2
DE2	Germany	Postweaning		6		4
DE3	Germany	Postweaning	2	5	2	2
DE4	Germany	Postweaning	3	4	3	3
DE5	Germany	Postweaning	2	4	2	2
DE6	Germany	Postweaning	3	4	2	2
ES1	Spain	Postweaning		15		3
ES1	Spain	Preweaning		15		2
ES2	Spain	Postweaning	11	16	11	11
ES2	Spain	Preweaning	8 prospective	16		
ES3	Spain	Preweaning	10	22	10	10
ES4	Spain	Preweaning	10	10	10	10
FST	Forest	Postweaning		5		5
NL1	Netherlands	Postweaning	3	9	3	3
US1	USA	Postweaning		30		5

Amplicon sequencing

The V3-V4 region of the 16S rRNA gene was amplified with primers 341F (5'-CCTAYGGGRBGCASCAG-3') and 806R (5'-GGACTACNNGGGTATCTAAT-3') and paired-end 250 bp sequenced using either Illumina HiSeq 2500 or Novaseq 6000. Reads were trimmed with cutadapt 2.3 [134] using default settings before being processed in DADA2 [105] following the v1.4 workflow for paired-end big data. Different sequencing batches were run separately before being merged to account for differences in the learnErrors step. Taxonomy was assigned with SILVA database v138 [135]. Genus level taxonomy was

assigned by the DADA2 pipeline using the RDP Naive Bayesian Classifier algorithm, and we used mmseqs2 easy-search with default settings to detect the species in SILVA with the highest identity alignment to each ASV [136]. The highest identity species alignment above 98.5% was assigned. When several species had equally high identity all were assigned separated by a slash. Amplicon sequence variants (ASVs) with taxonomic assignment as eukaryote, mitochondria, or chloroplast were discarded. Alpha and beta diversity were calculated on rarefied data (24325 reads) using R packages Phyloseq [137] and vegan [138], and the vegan::adonis function was used to perform PERMANOVA to determine the overall compositional differences between groups. Vegan function RDA was used for principal component analysis (PCA) and redundancy analysis (RDA).

Shotgun sequencing

Metagenomic libraries were prepared with NEB Next Ultra DNA Library Prep Kit (New England Biolabs, ME, USA) following the manufacturer's instructions. DNA was fragmented to 350 bp, purified with AMPure XP (Beckman Coulter, CA, USA) and sequenced with 150 bp paired-end sequencing on an Illumina NovaSeq 6000 machine.

Analysis of metagenomic data

Pig and plant (feed) reads were removed with kneaddata (https://github.com/biobakery/kneaddata) using the genomes of pig (GCF_000003025.6), wheat (GCF_002162155.1), and maize (GCF_902167145.1). Since some samples still had a large proportion on host or plant reads left after this, we further normalized read counts by the proportion of plant and pig reads found by kraken analysis [139]. This prevented samples with small proportions of bacterial reads from being outliers when mapping metagenomic reads to marker genes and metagenome assembled genomes (MAGs). We created MAGs using MetaWRAP v1.3.2 [140] with SPAdes v3.14.1 [132], and identified MAG taxonomy with GTDBtk v1.3.0 [141] and ARGs with resfinder software v4.1 [107]. We also used the resfinder database to quantify ARG abundance directly from metagenomics reads. Before mapping we clustered the genes in the database to 90% identity using mmseqs2 [136] easy-cluster with settings "--min-seq-id 0.9 --cov-mode 0". Reads were mapped to the representative sequences of the clusters using mmseqs2 easy-search with default settings, and reads aligning with minimum 100 bp and at 95% identity were accepted.

To assess the *S. suis* population in the tonsillar microbiota we created a *S. suis* pangenome by clustering all protein coding genes and mapping metagenomic reads to representative sequences. We annotated the genomes using Prokka v1.14.5 and clustered all protein coding genes at 80% residue identity using mmseqs easy-cluster (–min-seq-id 0.8 –cluster-mode 2 –cov-mode 1). We used previously published metadata covering 1703 assemblies [104]. We determined the association of each cluster with presence in clinical and non-clinical strains by calculating the ratio of percent presence in each group. The ratio of clusters more present in clinical strains were calculated by (% presence clinical/%

presence non-clinical) - 1, and clusters more present in non-clinical strains by ((% presence non-clinical/% presence clinical) * -1) + 1, so that clusters equally prevalent in clinical and non-clinical strains had a ratio of 0. We mapped metagenomic reads to the representative sequence of each cluster as described above and accepted reads mapping at 80% identity and 80% length.

Availability of data and materials

All microbiome sequencing data generated by this study is deposited in the NCBI BioProject portal under accession number PRJNA854341. Genome assemblies are available under accession numbers PRJNA849547 and PRJNA849577.

Acknowledgements

We thank the farmers and veterinarians that participated in the study. Most farms are anonymized. Farm NL1 is Schothorst Feed Research B.V. (*www.schothorst.nl*) and the forest piglets (FST) were collected from Boeren in het Bos (*www.boereninhetbos.nl*).

Supplementary material

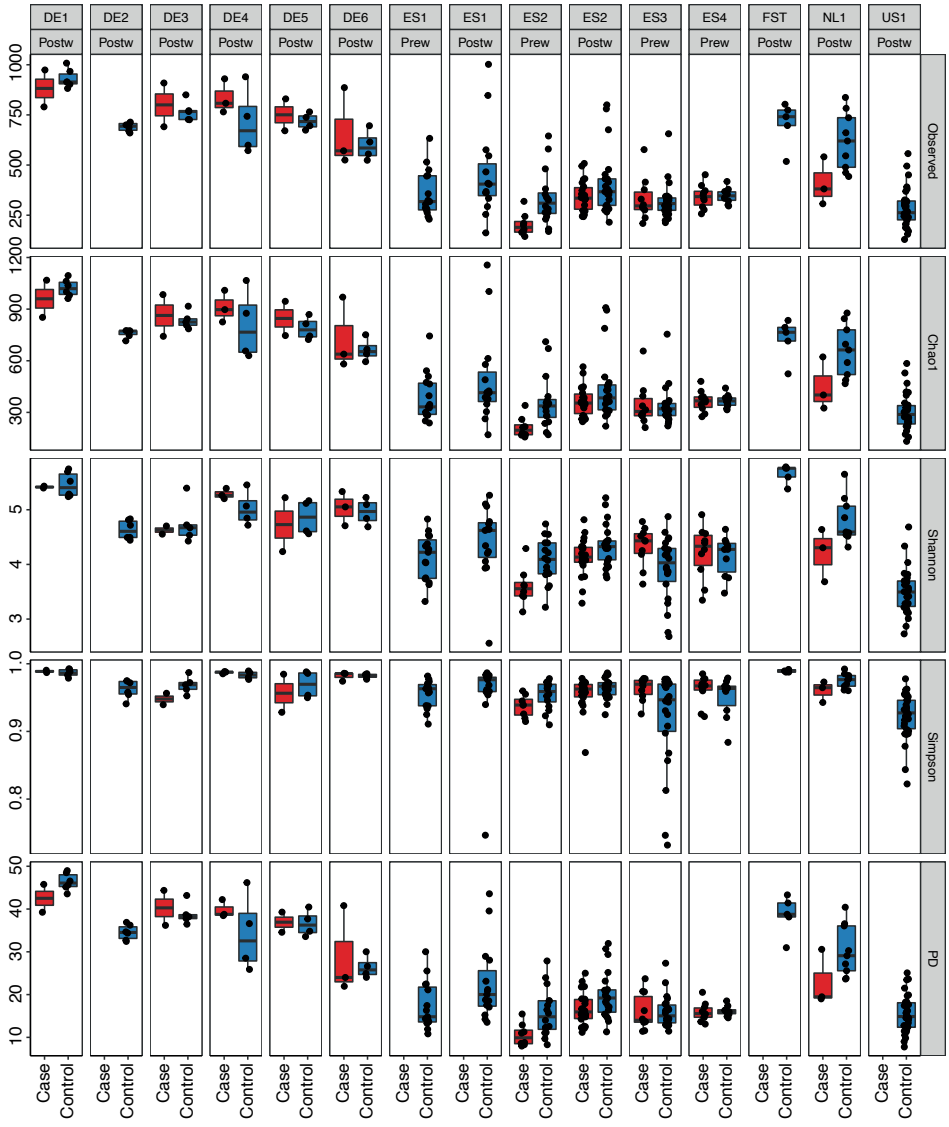


Figure S1: Comparison of case and control piglet alpha diversity within farms. Rows show different diversity metrics (amplicon sequencing data).

Supplementary text 1

Farm description:

DE1 (Germany)

Regular commercial farm. Intermittent *S. suis* problems. Tetracycline treatment since day 1 for piglet 1-4 but not 5-8. Farm size: Sows: 800, Nursery: 2000.

DE2 (Germany)

Regular commercial farm. Very high health status, and piglets had no problems with any disease in the period around sampling. Prestarter feed provided in both dry and liquid form and open water source provided. The piglets did not have their teeth ground or cut. All piglets were treated with amoxicillin the first day after birth as is typical in German farms. Farm size: Sows: 500, Nursery: 1200.

DE3 (Germany)

Regular commercial farm with some PRRSV problems (no vaccination) and intermittent *S. suis* problems. Low health farm with high antimicrobial usage for months/years. Amoxicillin and Tetracycline treatment for 5 days before sampling of piglets 1-5 but not piglet 6-7. Farm size: Sows: 600, Nursery: 2800.

DE4 (Germany)

Ecological farm utilizing vaccines. The piglets have outside access and straw bedding. The farm has reduced biosecurity, not only due to the outside access and lower level of cleanliness but also due to a lack of separation of airflow between age groups, which may facilitate transfer of pathogens. In addition to *S. suis*, the piglets have a high burden of parasites as well as rotavirus problems. No antimicrobial treatment of the sampled piglets. Farm size: Sows: 300, Nursery: 1200.

DE5 (Germany)

Regular commercial farm. Intermittent *S. suis* problems. No antimicrobial treatment of the sampled piglets. Farm size: Sows: 1200, Nursery: 5000.

DE6 (Germany)

Regular commercial farm. This farm has a history of severe *S. suis* problems. PRRSV and PCV2 (outbreak ongoing at sampling, may relate to failure in vaccination), and other pathogens also cause problems at this farm. Only 1 pig treated with antimicrobials (Amoxicillin) in the days prior to sampling. Farm size: Sows: 100, Nursery: 2000.

ES1 (Spain)
Regular commercial farm. Production farm receiving sows from ES2. No history of *S. suis* problems, although this is not a high health status farm in general. No antimicrobials used on sampled piglets. This farm was visited two times, with microbiota sampling of the same piglets at day 14 (1 week before weaning) and day 51 (3 weeks post-weaning). Weaning at day 21.

ES2 (Spain)
Regular commercial breeding farm. This farm delivers dams to farm ES1. This farm has a history of severe *S. suis* problems. This farm was followed from 1 week before weaning to 3 weeks post-weaning, with 1 microbiota sampling per week for 5 weeks. Weaning day 26-29. Sampling at day 18-21 (week -1), day 24-27 (week 0), day 32-35 (week +1), day 39-42 (week +2), and day 47-50 (week +3). All piglets were given amoxicillin in water at day 39, 12 days after weaning and 1 day before sampling point week 2 after weaning. Piglets in some nursery pens were treated with intramuscular marbofloxacin injections, see metadata.

ES3 (Spain)
Regular commercial farm. We sampled pre-weaning *S. suis* cases. This farm has a history of severe *S. suis* problems. All piglets were given prestarter feed with amoxicillin and neomycin.

ES4 (Spain)
Regular commercial farm. This farm has a history of severe *S. suis* problems. All piglets were given amoxicillin + gentamicin injections 2 days before sampling.

NL1 (Netherlands)
Research farm, Schothorst Feed Research B.V., Lelystad, Netherlands. High health status and low antimicrobial usage. Intermittent *S. suis* problems. No antimicrobials were given to the sampled piglets.

FST (Netherlands)
Piglets free roaming in a forest, with no access to indoor housing. Breed: Tamworth. No vaccination or antimicrobial treatment. The piglets forage for food, but this is also supplemented with some dry feed, fruit, and vegetables. Piglets were weaned and living in a different section than the sows at the time of sampling but spent considerably longer time with the sow than in regular farms (until naturally weaned). The outdoor conditions, rooting in soil, varied diet, gradual weaning, and rare breed may all have contributed to the diverged tonsillar microbiota.

US1 (USA)

Regular commercial farm. A farrow-to-finish farm located in the Midwest USA. Overall high health status. Intermittent *S. suis* problems, usually occurring late nursery or early finisher stage. All piglets were treated with gentamicin (day 26) and amoxicillin (day 45). The piglets were weaned at day 26 and sampled at day 55/62.

Chapter 3

Environmental and maternal factors shaping tonsillar microbiota development in piglets

Simen Fredriksen[1], Xiaonan Guan[1,2], Jos Boekhorst[1], Francesc Molist[2], Peter van Baarlen[1], Jerry M. Wells[1,3]

1. Host-Microbe Interactomics Group, Animal Sciences Department, Wageningen University, Wageningen, The Netherlands
2. Schothorst Feed Research B.V., Lelystad, The Netherlands
3. Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom

BMC microbiology 22.1 (2022): 1-12
https://doi.org/10.1186/s12866-022-02625-8

Background

The palatine tonsils are part of the mucosal immune system and stimulate immune responses through M cell uptake sampling of antigens and bacteria in the tonsillar crypts. Little is known about the development of the tonsillar microbiota and the factors determining the establishment and proliferation of disease-associated bacteria such as *Streptococcus suis*. In this study, we assess the tonsillar microbiota development in piglets during the first five weeks of life and identify the relative importance of maternal and environmental farm parameters influencing the tonsillar microbiota at different ages. Additionally, we studied the effect of sow vaccination with a bacterin against *S. suis* on microbiota development and *S. suis* colonisation in their offspring.

Results

Amplicon sequencing of the 16S rRNA gene V3-V4 region revealed that a diverse tonsillar microbiota is established shortly after birth, which then gradually changes during the first five weeks of life without a large impact of weaning on composition or diversity. We found a strong litter effect, with siblings sharing a more similar microbiota compared to non-sibling piglets. Co-housing in rooms, within which litters were housed in separate pens, also had a large impact on microbiota composition. Sow parity and prepartum *S. suis* bacterin vaccination of sows had weaker but significant associations with microbiota composition, impacting the abundance of *Streptococcus* species before and after weaning. Sex and birthweight had limited impact on the tonsillar microbiota, and none of the measured factors had consistent associations with microbiota diversity.

Conclusions

The piglet tonsillar microbiota is established shortly after birth. While microbiota development is associated with both environmental and maternal parameters, weaning has limited impact on microbiota composition. Intramuscular vaccination of sows pre-partum had a significant effect on the tonsillar microbiota composition of their piglets. These findings provide new insights into the mechanisms shaping the tonsillar microbiota.

Introduction

The palatine tonsils contain lymphoid tissue covered by a stratified squamous epithelium that extends into tonsillar crypts. M cells located in the crypt epithelial layer sample environmental antigens, including bacteria, for presentation by resident dendritic cells or macrophages and induction of immune (IgA) responses in saliva and the respiratory tract [142–144]. The tonsillar microbiota is therefore a major stimulus for mucosal immune responses but can also harbour pathogens which may use the tonsils as a portal of entry into the host [145].

A better understanding of the tonsillar microbiota and factors influencing its development is of particular importance in pigs due to colonization by and persistence of pathogenic bacteria such as *Actinobacillus pleuropneumoniae* [146], *Salmonella enterica* [147], and *Streptococcus suis* [148]. *Streptococcus suis* is especially relevant to microbiota research because of its high carriage rates in healthy piglets and ability to cause sepsis, meningitis, and arthritis [8, 70, 102]. Previous studies have found a large increase in the relative abundance of the family *Streptococcaceae* after weaning, coinciding with the abrupt change in piglet diet from milk to starch-rich dry feed at weaning [69, 72, 73]. These studies used antimicrobial growth promoters, and the described microbiota development may vary from Europe where these are banned. Regardless, the increase in *Streptococcaceae* family abundance is of interest because correlations between *S. suis*, carbohydrate availability, and microbiota composition have been linked to *S. suis* virulence and disease risk [8, 78, 102, 149, 150].

There are no cross-protective vaccines against *S. suis*. However, vaccination of sows with autogenous bacterins, vaccines made of dead cells of cultured clinical strains isolated from the individual farm, is used to provide passive immunity against *S. suis* in piglets. Their effectiveness is a matter of debate [8, 80, 82–84]. A recent study showed no lasting protection of piglets after sow vaccination, and vaccination of piglets failed to induce an active immune response [125]. The effect of bacterin vaccination on *S. suis* colonisation and tonsil microbiota development is less well described. In humans, a meta-analysis on infants vaccinated against *S. pneumoniae* showed a reduced carriage of the vaccine capsule types [151], but intramuscular bacterin vaccination of piglets has shown no effect on *S. suis* colonisation [87, 125]. Systemic vaccines rarely impact on mucosal immunity and colonization [152], but vaccination of sows may also impact on *S. suis* colonisation and the tonsil microbiota in their offspring via opsonising antibodies in colostrum and milk.

The aim of this study was to investigate the impact of maternal and environmental effects on the neonate tonsillar microbiota development. We utilized 16S rRNA gene V3-V4 region amplicon sequencing to assess the microbiota composition of 63 piglets from 21 different litters shortly after birth and at week 1, 3 and 5. One group of sows was injected with a multi-strain *S. suis* autogenous bacterin prior to farrowing, allowing us to assess the effects of sow vaccination on the microbiota of their offspring.

Methods

Experimental design

This study was carried out at Schothorst Feed Research BV, a high-health status research farm in the Netherlands. We selected 21 sows of varying parity (including 1 gilt) with normal body condition for the study. Eighteen of the sows were Landrace x Large White, the remaining 3 sows Large White x Large White. All piglets were from the same sire line: Tempo (TOPIGS Norsvin). For analysis, sows were grouped as young parity (parity 0, 1, and 2) or old parity (parity 4, 5, and 6). Nine sows received two 2 mL (between 10^7 and 10^8 cfu/mL) intramuscular injections of a bacterin vaccine approximately six and two weeks before farrowing, while 12 unvaccinated sows served as controls. No piglets received the vaccine. The bacterin was made by Dopharma BV (the Netherlands) and prepared from two *Streptococcus suis* strains (serotypes 7 and 9) isolated from autopsy cases of *S. suis* invasive disease on the farm.

All sows were inseminated on the same day. Sows were housed separately, one sow per pen, in three different farrowing rooms from day 109 of gestation. Each room housed 8 pens, each with a size of 0.60 x 2.50 m for the sow and 2.25 x 2.50 m total surface. Three litters from one of the rooms (room 5) were not included for microbiota sequencing. These litters were, however, also a part of the experimental design and no known factors may have influenced the microbiota in this room differently than the other rooms. Large litters were reduced in size by cross-fostering within 48 hours after birth, but no litters in the study received piglets from another sow. Of the total 362 piglets born, 28 were born dead, 15 cross-fostered, and 49 died before weaning. Creep feed was provided from 7 days after birth until weaning. Piglets were weaned at approximately 26 days of age, and litters were moved to nursery pens (2.00 x 1.13 m) each containing 6 piglets, with the piglets selected for microbiota sampling remaining co-housed to avoid microbiota transfer. This study design did not allow determination of the post-weaning room and pen effect.

Three piglets from each litter were selected for microbiota sequencing based on birth weight. We excluded piglets that died or developed disease symptoms throughout the study period. We selected one low birth weight (LBW), one median birth weight (NBW), and one high birth weight (HBW) piglet per litter. Birth weight varied between litters, meaning that the weight of some LBW piglets was heavier than NBW or HBW piglets of other litters. Piglets were first sampled with tonsil swabs and weighed as soon as possible after birth. Because piglets were born over several days, and only handled during working hours, this occurred separately per litter and up to 48 hours after birth. The study cohort was next sampled all at the same time at week 1 after birth, the timepoint being set as the average age of all piglets. This meant that some piglets were 4 days old and others 9 days old. Subsequently, all piglets were sampled at week 3. At week 4, all piglets were weaned and weighed. One week later, at week 5, all piglets were sampled for the last timepoint for microbiota sampling. At week 7, all piglets were again weighed, but not sampled, before

being mixed and utilized in other studies. Because of confounding treatments in these studies, we calculated post-weaning growth rate only for the first 3 weeks post-weaning.

To compare the piglet microbiota with that from adult sows we collected tonsil swab samples from 12 unrelated sows from the same farm at slaughter. We did not collect tonsil swabs from the original sows used in the study because these sows were not sacrificed, and sampling from living sows requires anaesthesia treatment that would have interfered with the experiment.

Microbiota sequencing

We sampled the tonsillar microbiota of piglets by gently scraping the palatine tonsils with a HydraFlock swab (Puritan, USA) for 10 seconds. Swabs were placed in vials containing Powerbead solution (Qiagen, The Netherlands) and stored at -80 °C. DNA was isolated using the PowerSoil DNA Isolation Kit (Qiagen) with 0.7mm garnet beads. DNA yield was similar between timepoints (week 0: 44 ng/µl, SD=27; week 1: 40 ng/µl, SD=23; week 3: 48 ng/µl, SD=24; week 5: 34 ng/µl, SD=22). Subsequently the V3-V4 region of the 16S rRNA gene was amplified with primers 341F (5'-CCTAYGGGRBGCASCAG-3') and 806R (5'-GGACTACNNGGGTATCTAAT-3'), using 35 PCR cycles with Phusion High-Fidelity PCR Master Mix (New England Biolabs, USA). DNA libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit. DNA quantity and fragment size was checked with Qubit and BioAnalyzer before 250 bp paired-end sequencing on an Illumina NovaSeq 6000 machine. ZYMO mock community #D6300/#D6305 was used to verify the accuracy of our protocol and we processed blank swabs together with the tonsil microbiota samples as controls. Analysis of the abundance of the Amplicon sequence variants (ASVs) identical to the mock community members 16S rRNA gene V3-V4 region showed that DNA from all community members was successfully isolated and amplified. We found a moderate bias towards disproportionally high abundance of Gram-negatives *Escherichia coli* and *Salmonella enterica*, while *Bacillus subtilis* was underrepresented (Figure S1). Three blank control swabs processed together with the tonsil microbiota samples produced no visible bands following PCR amplification.

After sequencing, primers were trimmed with Cutadapt 2.3 [134] and processed in DADA2 [105] with taxonomic assignment to SILVA database v138 [135]. Amplicon sequence variants with taxonomic assignment as eukaryote, chloroplast, or mitochondria were discarded before carrying out further analyses. The ASV abundance per sample was rarefied to the sample with the lowest read count (42537 reads). R packages Phyloseq [137] and vegan [138] were used to calculate alpha and beta diversity. All results shown in the main text are based on Shannon diversity and Bray-Curtis dissimilarity. Vegan function RDA was used for principal component analysis (PCA) and redundancy analysis (RDA), and function adonis was used for permutational analysis of variance (PERMANOVA). We used LEfSe (Linear discriminant analysis Effect Size) [153] to find taxa associated with different timepoints. ASV correlation with numerical variables (such as growth rates and parity) was calculated with Spearman's

rank correlation coefficient, while Wilcoxon Rank Sum Test was used for categorical variables (such as sow vaccination and pairwise comparison of timepoints).

A previous longitudinal study on the piglet tonsillar microbiota [72] reported taxonomy only at the family level. We reanalysed their dataset (BioProject PRJNA391812) with the approach used in the present study as described above. The reanalysed study used different primers, sequencing the 16S rRNA gene V4 region while we used the V3-V4 region. Both primer sets did, however, allow us to distinguish between the two species of interest, *S. suis* and *S. porcorum*, in SILVA database v138.

In vitro growth experiment

Carbon source utilization of *Streptococcus* species from the oral cavity of pigs was tested *in vitro* by measuring growth in a complex medium (CM) with defined carbon sources. The complex medium was made as previously described [154] and supplemented with 0.2 µm filter sterilized D-glucose, lactose, maltotriose, or pullulan at 0.2 g/100 mL as carbon source. Overnight cultures were prepared in Todd-Hewitt yeast broth (THY) inoculated with single colonies from THY agar plates and diluted 1:50 in PBS. 10 µL diluted overnight culture was inoculated in 190 µL fresh THY medium in 96-well plates. In total, six plates, each containing 2 replicates of each treatment, were incubated at 37 °C for 15 hours and optical density measured at 600 nm every 15 minutes using a SpectraMax M5 (Molecular Devices) spectrophotometer.

Results

Overview of the tonsillar microbiota composition

We sequenced the tonsillar microbiota of 63 piglets from 21 different litters with 16S rRNA gene V3-V4 region amplicon sequencing. Tonsil microbiota samples were collected within 48 hours after birth, and when the average age of the cohort was of 1, 3, and 5 weeks old. Nine of the 21 sows were vaccinated with a *S. suis* bacterin vaccine 6 and 2 weeks before expected parturition. All samples were successfully processed and 250 bp paired-end sequenced, resulting in 252 piglet and 12 sow tonsillar microbiota samples with a minimum of 42537 reads and in total 25267 ASVs after processing with DADA2 (Figure S2).

The most abundant genera of the piglet tonsillar microbiota were *Actinobacillus*/*Haemophilus* (two abundant genera indistinguishable with 16S rRNA gene V3-V4 region amplicons), *Moraxella*, *Porphyromonas*, and *Streptococcus*; these genera exceeded 10% mean abundance across the dataset. *Rothia*, *Fusobacterium*, *Neisseria*, *Alloprevotella*, and *Acinetobacter* each contributed more than 3% abundance (Figure 1). Figure S3 shows the abundance of the most abundant genera per sample. Most of the abundant genera were represented by a large number of different ASVs (Figure 1), indicating that the tonsillar microbiota is colonized by multiple closely related strains. Nineteen ASVs contributed 1% abundance or more, and 167 ASVs were present at 0.1% or higher abundance. Thirty four

of these 167 ASVs were classified as *Actinobacillus*/*Haemophilus*, 21 as *Porphyromonas*, and 15 as *Streptococcus*. Three of the six most abundant ASVs were classified as *Moraxella* and differed by only a single SNP.

Tonsillar microbiota development during the first 5 weeks of life

Microbiota composition changed with age (PERMANOVA; R2 = 0.13, p < 0.001), but the age effect was not strong enough to fully separate samples by timepoint with PCA (Figure 2). At the initial timepoint (week 0) we found high abundance of *Porphyromonas* sp. and *Neisseria* sp. ASVs, but these ASVs decreased in abundance before week 1 (*Neisseria* ASV 3 from 7.3% to 3.4%, *Porphyromonas* ASV 7 and ASV 14 from 3.9% to 2.1% and 2.6% to 0.66%, respectively). These 3 ASVs all had weak negative correlations with birthweight at the first timepoints. *Neisseria* ASV 3 had the strongest negative correlation with birthweight at the initial sampling (R = -0.19, p = 0.13), while *Porphyromonas* ASVs 7 and 14 had stronger correlations at week 1 (R = -0.31, p = 0.01 and -0.15, p = -0.24, respectively).

Week 0	Week 1	Week 3	Week 5	Mean	ASVs	
18.1	14.7	16.3	20.3	17.2	165	*Actinobacillus*
11.1	15.8	13.7	14.2	13.3	28	*Moraxella*
16.4	10.1	10.9	8.6	11.4	93	*Porphyromonas*
8.2	10.7	9.8	13.0	10.0	84	*Streptococcus*
6.6	8.3	6.3	3.9	6.1	6	*Rothia*
3.6	3.0	7.0	3.5	4.8	41	*Acinetobacter*
4.7	5.8	4.8	4.0	4.7	38	*Fusobacterium*
7.3	3.6	3.2	2.6	4.1	27	*Neisseria*
3.1	2.7	4.5	4.4	3.6	73	*Alloprevotella*
1.4	2.7	2.6	2.0	2.1	115	*Prevotella*
2.0	2.4	2.0	1.9	2.0	34	*Leptotrichia*
1.5	2.6	1.8	2.3	2.0	58	*Bacteroides*
1.2	0.8	1.2	1.4	1.4	14	*Conchiformibius*
1.4	0.7	1.1	0.9	1.1	31	*Alysiella*
0.7	0.7	1.2	0.7	0.9	15	*Dysgonomonas*
0.8	0.6	0.5	1.6	0.9	38	*Veillonella*
0.4	0.4	0.6	1.4	0.8	31	*Lactobacillus*
0.5	0.7	0.6	0.5	0.6	9	*Gemella*
0.8	0.4	0.6	0.3	0.5	5	*Dielma*
0.5	0.5	0.4	0.5	0.5	2	*Pelistega*

Figure 1: Heatmap showing relative abundance (%) of the most abundant genera at different timepoints and the number of ASVs with more than 100 reads for each genus.

At week 1 the abundance of ASV 1 (*Rothia nasimurium*), *Moraxella pluranimalium/porci* (ASV *2* and ASV 5), and *Streptococcus suis* (ASV 4) increased compared to week 0 (4.1% to 6.4%, 3.4% to 4.6%, 2.5% to 5.7%, and 3.1% to 5.0%, respectively). LEfSe analysis on the abundance of taxa at different timepoints also found that *Porphyromonas* and *Neisseria* were associated with the initial timepoint, *Rothia* and *Moraxella* with week 1, and *Streptococcus* species with week 5 (Figure S4).

Piglets were weaned at week 4, at approximately 26 days old, changing their diet from sow milk with creep feed supplementation to starch-rich dry feed. This might be expected to reshape the microbiota, in part due to the change in available carbohydrates. Between week 3 and 5 the most abundant *Streptococcus suis* ASV (ASV 4) decreased in abundance (3.9% to 2.4%), while ASV 9 (*Streptococcus porcorum)* increased from 0.8% to 5.5% abundance, becoming the most abundant ASV at week 5. Several other less abundant *Streptococcus* ASVs also increased in abundance between week 3 and 5, including ASV 445 (*Streptococcus caballi*, 0.007% to 0.03%), ASV 97 (*Streptococcus hyointestinalis*, 0.1% to 0.6%), and ASV 195 (*Streptococcus porci*, 0.04% to 0.21%).

Overall change in microbiota composition measured by Bray-Curtis dissimilarity was, however, similar between timepoints. The smallest change occurred between week 1 and 3 (mean Bray-Curtis dissimilarity = 0.46), the second smallest change occurred between week 0 and 1 (mean = 0.48), and the largest between week 3 and 5 (mean = 0.49). The number of the most abundant ASVs changing in abundance was also similar between timepoints. Of the 167 ASVs with 0.1% or higher overall abundance, 72, 66, and 72 significantly changed in abundance between timepoints 0 to 1, 1 to 3, and 3 to 5 respectively. The impact of weaning was more evident on rare ASVs. Among the 3724 ASVs with 0.001% or higher overall abundance, 69, 103, and 134 ASVs changed in abundance between timepoints 0 to 1, 1 to 3, and 3 to 5 (Wilcoxon Rank Sum test FDR < 0.05).

A previous study on the piglet tonsillar microbiota in the USA [72] found a similar but stronger increase in overall *Streptococcacae* abundance after weaning but did not report changes at the species level. To evaluate whether *S. porcorum* was responsible for the increase in *Streptococcacae* we re-analysed their dataset (BioProject PRJNA391812) with the approach used in the present study. Our reanalysis found a strong *S. porcorum* increase across weaning, from 0.1% to 29.8%. The increase in *S. porcorum* coincided with an overall decrease in *S. suis* from 5.0% to 2.8%, similar to the decrease from 6.1% to 4.6% in the present study (Figure S5).

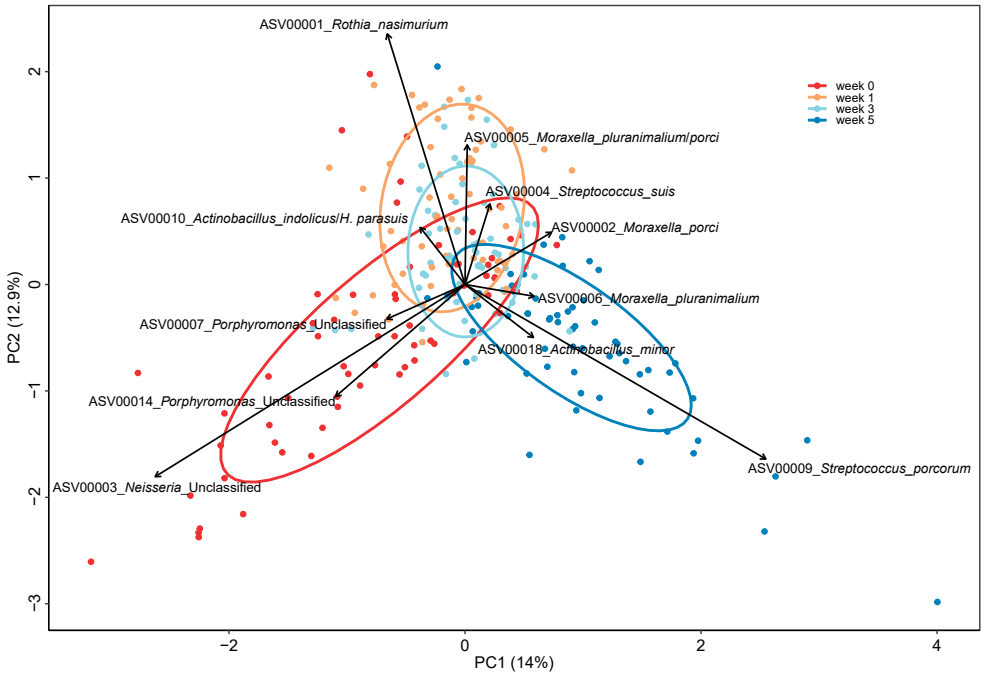


Figure 2: PCA plot showing differences in tonsillar microbiota composition by age. Arrows show ASVs driving the separation; samples in the direction the arrow is pointing have higher abundance of that ASV. Each point represents a sample. At the first timepoint (week 0), the tonsillar microbiota was characterised by a high abundance of *Neisseria* and *Porphyromonas* ASVs, while post-weaning samples (week 5) had a high abundance of *Streptococcus porcorum*.

To determine whether the decrease in *S. suis* and increase of *S. porcorum* might be related to carbohydrate utilisation we conducted an *in vitro* growth experiment in complex medium (CM) containing different carbohydrates present before and/or after weaning. We included 6 different strains of 5 *Streptococcus* species found in the oral cavity of pigs. All strains grew to higher OD_{600} values with all added carbohydrates compared to base complex media (CM) alone, showing that all the tested strains can utilize lactose (present in sow milk), maltotriose (produced by the breakdown of starch by salivary amylase), and the starch dextran pullulan (Figure 3). Some strains did, however, reach higher OD_{600} values on specific carbon sources. *S. parasuis* reached higher OD_{600} values on glucose and maltotriose compared to pullulan and lactose, and *S. porcinus* grew more rapidly to high OD_{600} values on maltotriose and pullulan compared to the other media. The two *S. porcorum* strains differed in the growth on lactose and pullulan. Although we cannot ascertain the phenotype of the strains present *in-vivo*, the tested *S. porcorum* and *S. suis* strains grew on lactose, maltotriose, and pullulan *in vitro,* suggesting that factors other than availability of lactose and starch explain the change in abundance after weaning.

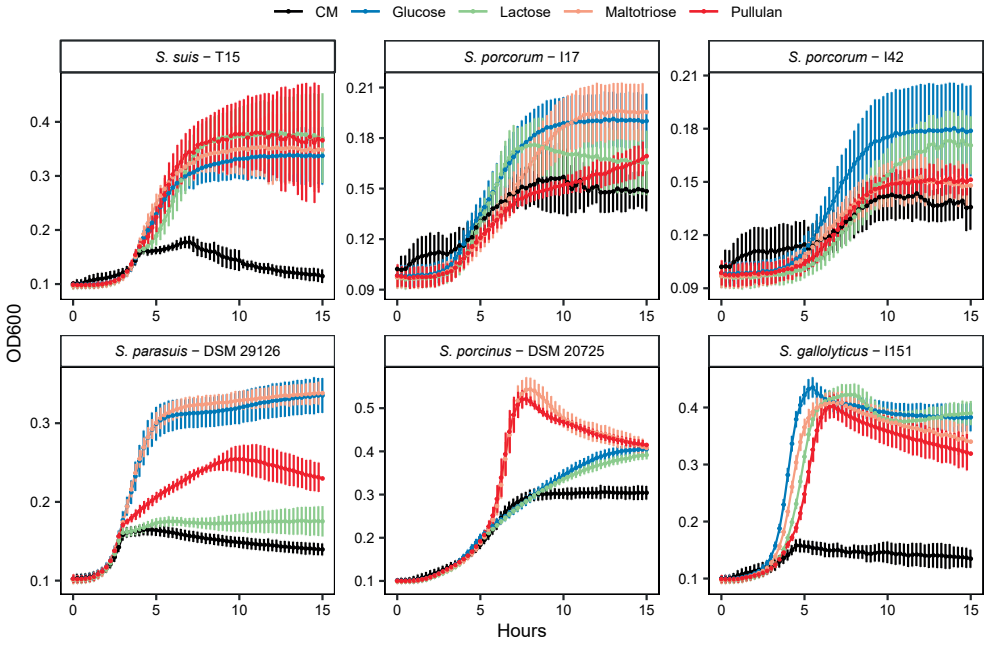


Figure 3 Growth curves (note: independent y-axis per strain) of 6 strains from five *Streptococcus* species found in the oral cavity of pigs on complex media (CM) supplemented with different carbohydrate sources. While growth curves differ between strains, all screened strains reached higher OD_{600} values with all added carbohydrates than CM alone.

Alpha diversity increases during the first 5 weeks of life

Alpha diversity increased with age but did not reach as high Shannon diversity as sows by week 5 (Figure 4a). The diversity increase with age was consistent between alpha diversity metrics, but while sows had high diversity, their ASV richness was lower (Figure S6). To assess the impact of different factors on microbiota diversity, we used a multivariate ANOVA on Shannon diversity with the model BW_kg (numeric BW) + BW_relative (low, median, or high BW within litter) + Sex + Farrowing_room + Parity_group (0, 1 and 2 vs 4, 5, and 6) + Sow_vaccination + Sow_ID + Daily_growth_pre_weaning + Daily_growth_post_weaning for each timepoint. None of the included factors were consistently associated with alpha diversity over time. Birth weight had a small impact on diversity directly after birth (relative within litter: $p < 0.05$, absolute: $p = 0.09$ at week 0), but this association decreased with age. Pre-partum sow vaccination also had the strongest association with alpha diversity at the first timepoint ($p = 0.18$). At the later timepoints, the farrowing room, sow parity, and the individual sow were significant factors ($p < 0.05$) at one or more timepoints. These factors all saw a dip in significance at week 3 before increasing to week 5. Pre- and post-weaning growth rate was not significantly ($p < 0.05$) associated with diversity at any timepoint, although diversity at week 1 had a positive correlation with post-weaning growth rate ($p = 0.051$). Redundancy analysis (RDA) on microbiota

composition with growth rate as constraint did not reveal any significant interaction, and no individual ASV had a significant Spearman correlation with growth rate at any timepoint (FDR < 0.05).

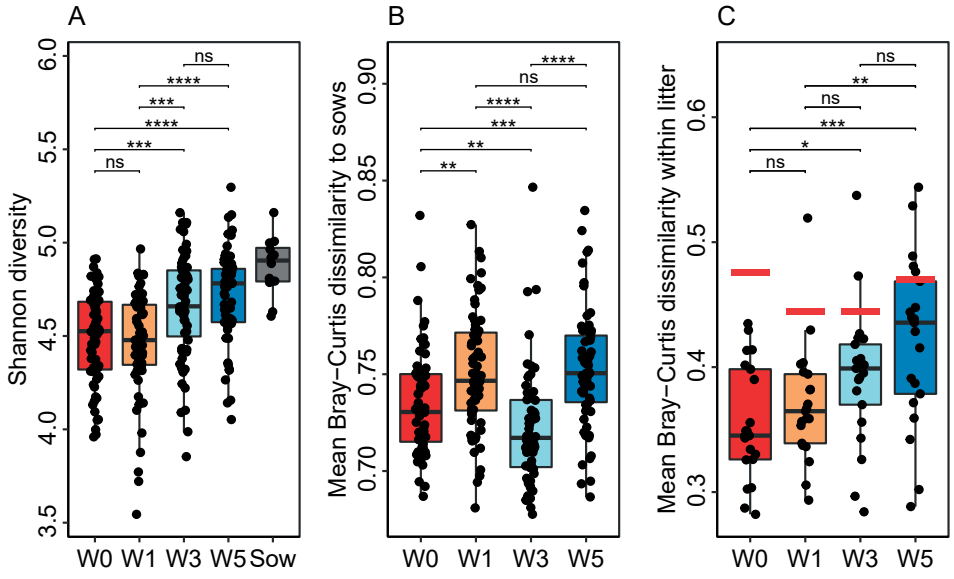


Figure 4: A) Shannon diversity increases with age and approaches adult sow microbiota diversity. **B**) Mean Bray-Curtis dissimilarity between piglets and unrelated sows on the same farm. Significance calculated by Wilcoxon Rank Sum Test is indicated, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. **C**) Total litter effect measured in mean Bray-Curtis dissimilarity between siblings at each timepoint. While each point in A) and B) represents a single piglet, each point in C) represents a litter, and its y-axis value is the mean of all possible pairwise comparisons between the siblings. Red horizontal lines indicate the mean dissimilarity between all piglets at the timepoint. The mean dissimilarity between all piglets was comparable to that between the unrelated sows (mean = 0.48, not shown).

Comparison of the tonsillar microbiota composition in piglets and sows

The compositional changes over the 5 first weeks of piglet life did not show any signs of converging towards the adult sow tonsillar microbiota (Figure 4B) and remained significantly different at week 5 (PERMANOVA; $R2 = 0.30$, $p < 0.001$). The sow tonsillar microbiota differed from that of piglets most notably by lower abundance of *Rothia* (4.1% vs 1.5%), *Streptococcus* (13.6% vs 2.3%), and *Moraxella* (13.8% vs 7.2%) and higher abundance of *Acinetobacter* (3.7% vs 16.7%), *Conchiformibius* (1.5% vs 8.5%), and *Alysiella* (0.87% vs 4.1%) (all FDR < 0.01). Most abundant ASVs in the piglet tonsillar microbiota were present in the sow tonsillar microbiota, although at lower abundance. The most abundant ASVs not present in the sow tonsillar samples were ASV 5 (*Moraxella pluranimalium/porci*), ASV 19 (*Moraxella* sp.), ASV 23 (*Actinobacillus porcinus*), and ASV 22 (*Rothia nasimurium*). *Streptococcus porcorum* abundance was also low in sows (mean: 0.02%).

Statistical associations between the tonsillar microbiota and farm parameters

The tonsillar microbiota composition of piglets may be influenced by a range of factors, and we observed a strong litter effect, with a lower mean Bray-Curtis dissimilarity between siblings than between random piglets (Figure 4C). This overall litter effect may include a range of factors linked to both the sow and shared environment. We determined the associations between microbiota composition and all measured parameters with PERMANOVA on Bray-Curtis dissimilarity at each timepoint with the model described above for alpha diversity (Figure 5A). The farrowing room played a significant role ($p < 0.01$), starting at $R2 = 0.08$ after birth and remaining at $R2 = 0.06$ at week 5 (Figure 5B-D). Sow parity and pre-partum vaccination had smaller but significant associations ($p < 0.01$ at week 1). As for alpha diversity, the effect of parity, sow vaccination, and farrowing room decreased in strength at week 3 before recovering at week 5. Birth weight had a small but significant ($p < 0.05$) association both measured in absolute terms and relative to littermates, but this association decreased with age. The sow ID, in the model representing residual litter effect not explained by the previous factors, was the strongest factor at week 0 ($R2 = 0.45$, $p < 0.001$). The variation explained by the individual sow decreased between each timepoint but remained the strongest factor at week 5 ($R2 = 0.32$, $p < 0.01$).

To further investigate associations between microbiota composition and the model variables we used RDA on each variable for each timepoint. This analysis confirmed the findings from the PERMANOVA, with strong associations between microbiota composition and farrowing room and sow parity (maximum RDA1 = 5.7% and 4.9%, respectively). Absolute and relative birth weight and sex had weaker and non-significant associations (RDA1 < 2% and $p > 0.2$). Sow vaccination remained significant at week 1 and 3 (RDA1 = 3.3% and 3.2%, respectively).

While the overall microbiota had significant associations with several factors with RDA and measured in Bray-Curtis dissimilarity, few specific ASVs had consistent associations between timepoints. The microbiota difference between farrowing rooms were not caused by ASV presence/absence, but different abundance of shared ASVs (Figure 5B-C).

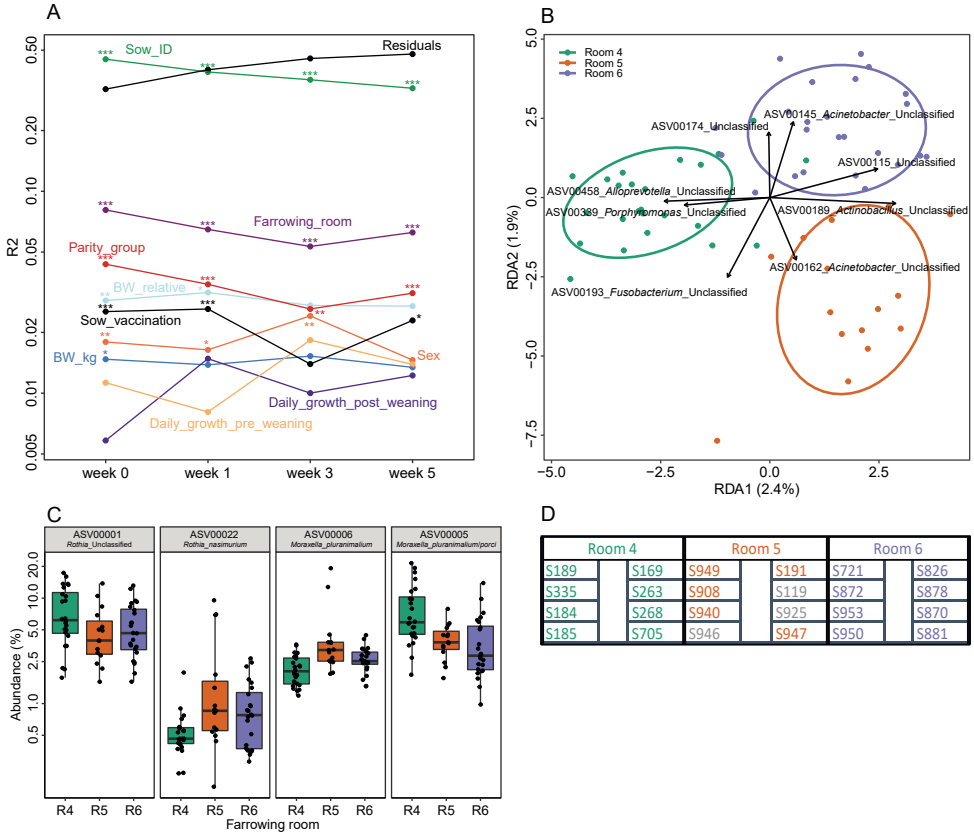


Figure 5: A) The additive effect of factors on the piglet tonsillar microbiota composition calculated by PERMANOVA on Bray-Curtis dissimilarity. Most variation in the dataset remains unexplained by the included factors, as shown in the high values of Sow_ID and residuals (i.e., variation specific to the individual sow (litter) or piglet and not explained by the other factors). Statistical significance is indicated, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **B**) The farrowing room effect at week 1 shown by RDA, abundance transformed by log(1000*abundance+1). **C**) The abundance of the 4 strongest drivers of farrowing room separation in RDA without abundance transformation. **D**) Schematic presentation of the rooms and pens. 3 litters from room 5 were not included for microbiota sequencing (marked in grey).

Bacterin vaccination altered the abundance of *Streptococcus* species

Pre-partum vaccination of sows with a *S. suis* bacterin affected the microbiota composition of their piglets (PERMANOVA; R2 = 0.026, p < 0.05 at week 1). This was in part driven by differences in the abundance of ASV 4 (*S. suis*); this ASV matched the 16S rRNA gene of both bacterin strains at 100% identity. ASV 4 was less abundant in the piglets of vaccinated sows during the first week of life, but by week 3 its abundance was similar in vaccine and control litters (Figure 6). The overall difference in microbiota composition was again significant at week 5, after weaning (PERMANOVA; R2 = 0.026, p < 0.05). This was driven in part by a higher abundance of ASV 9 (*S. porcorum,* 4.31% vs 6.96%) and ASV 97 (*S. hyointestinalis*, 0.39% vs 0.83%). The vaccine effect was additive with a parity effect on the same ASVs; the vaccine effect occurred within both high- and low-parity sows. For instance, ASV 4 (*S. suis*) abundance at week 1 was 3.64% in vaccinated high-parity litters, 4.79% for [no vaccine + high parity], 4.47% for [vaccine + low parity], and 7.10% for [no vaccine + low parity].

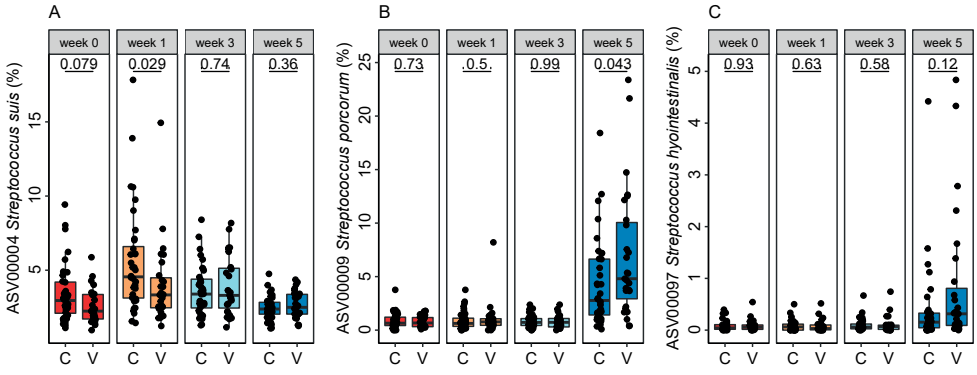


Figure 6: Sow vaccination altered the abundance of *Streptococcus* species. ASV 4 (*S. suis* ASV matching the 16S rRNA gene of the bacterin vaccine strains) abundance was lower in piglets of vaccinated sows (V) than controls (C) at the first two timepoints (**A**), while (**B**) ASV 9 (*S. porcorum*) and (**C**) ASV 97 (*S. hyointestinalis*) were more abundant in the control group post-weaning. Wilcoxon Rank Sum Test p-values are shown.

Discussion

In this study, we followed changes in the tonsillar microbiota of piglets from shortly after birth until one week after weaning. We investigated the association between tonsillar microbiota composition and diversity and environmental, maternal, and individual piglet parameters. We found a strong litter effect, i.e., piglets born by the same sow sharing a similar microbiota. This is consistent with previous studies on the oral [72, 155] and gut [156, 157] microbiota of piglets.

The litter effect on the piglet microbiota may be caused by a variety of factors. For example, environmental factors related to the pen and microbiota transfer from

the environment and via aerosols [158] may be important determinants of microbiota composition. Horizontal microbiota transfer between siblings likely also plays a large role. In this study, strong environmental effects were indicated by the large effect of the farrowing rooms, where several litters were housed in separate pens. The farrowing room differences were not caused by presence/absence of different taxa, but small differences in abundance of shared ASV. These differences may be stochastic and driven by microbiota transmission within rooms. Pens were divided with solid separators preventing direct contact between piglets in neighbouring pens, but airborne transfer is possible [88, 158]. A previous study found inoculated *S. suis* serotype 9 to transmit between separated pens, albeit significantly slower than via direct contact [88]. While the environment on the research farm used for this study is controlled, small differences in temperature or ventilation between the rooms may also have played a role. Further research on this topic is warranted.

Sow factors such as genetics, behaviour, milk and colostrum content, and vertical microbiota transfer from sow to litter may also cause litter effects, as shown by the impact of sow parity and pre-partum vaccination of sows. This is consistent with a previous study showing that the nasal microbiota of piglets is strongly influenced by sow contact [155]. Parity has been reported to influence nasal and faecal microbiota composition of piglets [149, 156], and is positively associated with piglet performance and negatively with piglet mortality [159].

We found a weak association between tonsillar microbiota composition and birthweight. Previous studies on the gut microbiota of piglets reported stronger associations [160–163], and linked this to the effects of intrauterine growth restriction. Intrauterine growth restriction and/or low birth weight has been associated with improper development of intestinal mucosal immunity [164] and gut epithelial barrier function [165], as well as reduced colostrum, milk, and creep feed intake [159, 166–168]. The weaker association between birthweight and the tonsillar microbiota may be due to less developmental impact of intrauterine growth restriction on the tonsils compared to the gut, but also due to the tonsillar microbiota being less influenced by milk and feed intake compared to gut microbiota.

Prepartum vaccination of sows with a *S. suis* bacterin impacted the tonsillar microbiota composition of their piglets, initially by reducing the abundance of the *S. suis* ASV matching the strains used in the vaccine during the first week after birth. A recent study found similar effects of pre-partum sow vaccination on the piglet nasal microbiota and *Glaesserella parasuis* abundance [169]. The vaccination effect may be due to alterations in the vertically transferred sow microbiota, although intramuscular vaccination rarely impacts mucosal immunity [152]. It is also possible that maternal antibodies in colostrum and milk from vaccinated sows influence *S. suis* persistence through immune exclusion. Apart from the effect of the sow vaccination on *S. suis*, the tonsil microbiota composition of piglets was significantly different to the control group also after weaning, at week 5,

in part due to higher abundance of *S. porcorum* and *S. hyointestinalis*. More complex and indirect microbe-microbe interactions may underlie the higher *Streptococcus* abundance post-weaning, as the piglets had stopped suckling. Future studies may want to include analysis of sow microbiota and antibody titres to better assess the mechanisms behind the strong sow influence on the piglet microbiota.

Streptococcus porcorum also stood out as the main species increasing in abundance after weaning, while *S. suis* abundance decreased. *Streptococcus porcorum* is closely related to *S. suis* but has shown differential efficiency in carbohydrate utilisation [170]. We considered it possible that *S. porcorum* might have gained a competitive advantage upon the increased availability of starch after weaning, but an *in vitro* experiment showed that neither *S. porcorum* nor *S. suis* gained a distinct advantage when provided lactose, maltotriose (available in the oral cavity after breakdown of starch by host amylase activity), or the starch dextran pullulan as carbon sources. Thus, the mechanism behind the large increase in *S. porcorum* abundance remains unknown.

Contrary to previous work on both the tonsillar [72] and gut microbiota [171–173], there was no striking change in the microbiota composition after weaning, indicating that the impact of host diet on tonsillar microbiota composition is modest. The larger change in tonsillar microbiota composition previously found [72] was also largely driven by *S. porcorum*, but the increase and change in the overall microbiota composition was far larger. This might be explained by piglets in the present study having access to pre-weaning creep feed, allowing post-weaning-associated bacteria to gain an early foothold, and by the previous study including feed supplementation with the antimicrobial Carbadox. The persistence of a relatively undisturbed microbiota in this study may also have contributed to persistence of the strong litter effect after weaning.

Conclusions

Our results show that the tonsillar microbiota of piglets is established within the first day after birth, and that microbiota composition and diversity changes gradually during the first five weeks of life without converging on the adult sow tonsillar microbiota composition. The impact of weaning and host diet change on the tonsillar microbiota appears to be limited. Tonsillar microbiota composition was associated with a dominant litter effect linked to both environmental and maternal factors. Prepartum *S. suis* bacterin vaccination of sows resulted in reduced abundance of *S. suis* in the tonsillar microbiota in litters of vaccinated sows. We considered that this effect was most likely due to immune exclusion from anti-*S. suis* antibodies in colostrum and sow milk saturating the tonsil surface. Piglets of vaccinated sows also showed significantly altered microbiota composition and increased *S. porcorum* abundance post-weaning, suggesting that the bacterin vaccination had fundamentally altered microbiota development. Our findings show promise for the

potential of microbiome manipulation by probiotics and vaccination strategies to avoid infectious disease in young piglets.

Availability of data and materials

All sequencing data generated in the current study are available in the DANS repository at https://doi.org/10.17026/dans-xxb-56zp.

Acknowledgements

We thank Xandra Benthem de Grave for assistance with sampling of sows and Joyce Bisdom for assistance with DNA isolation.

Supplementary material

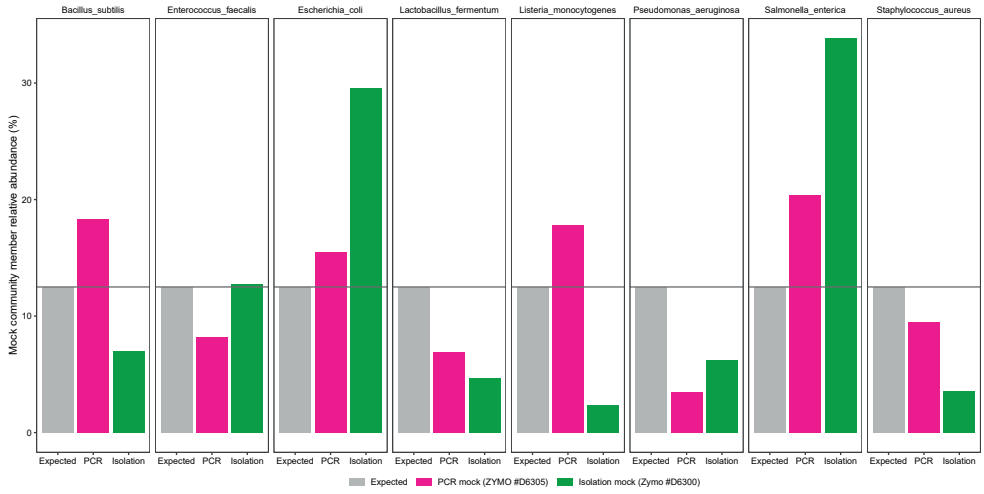


Figure S1. Boxplot showing the relative abundance of ASVs with 100% identity to the 16S rRNA gene V3-V4 region of mock community members. We used both a DNA isolation mock community (ZymoBIOMICS Microbial Community Standard ZYMO #D6300) and PCR mock community ZymoBIOMICS Microbial Community DNA Standard ZYMO #D6305).

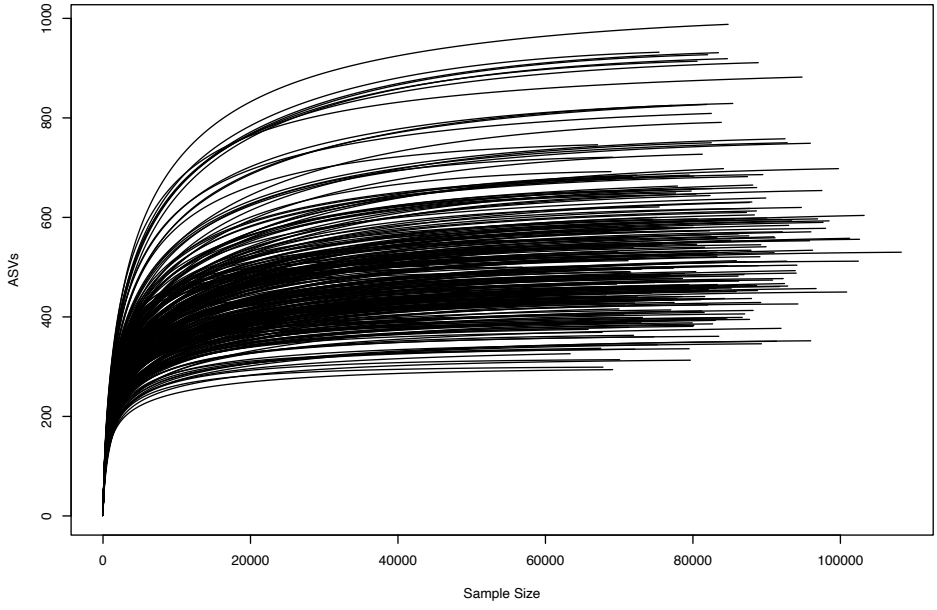


Figure S2. Rarefaction curves for all samples. Constructed with Vegan function rarecurve.

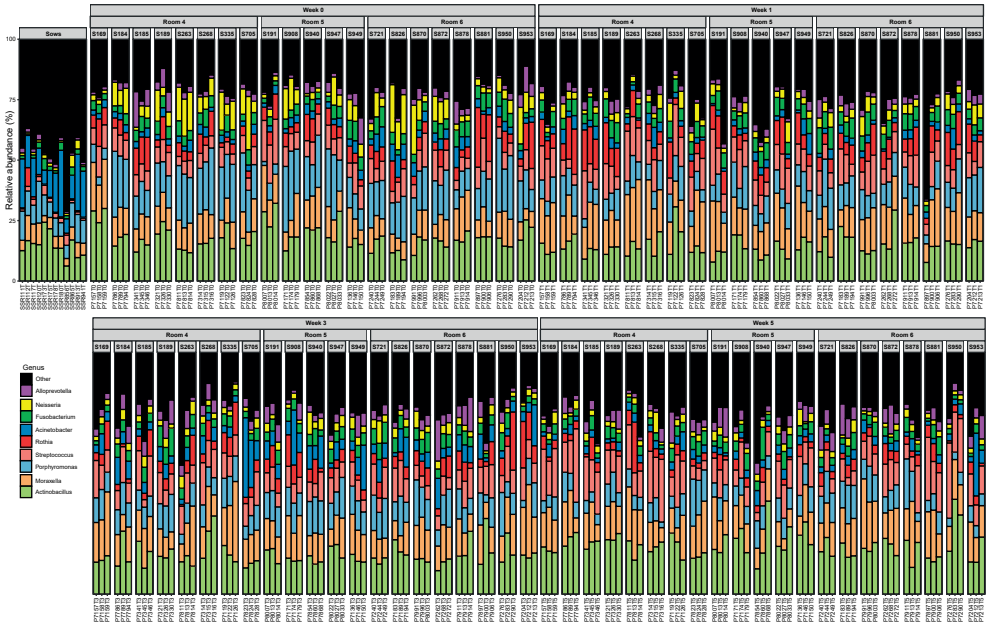


Figure S3. Stacked barplot showing the abundance of the most abundant genera in all samples.

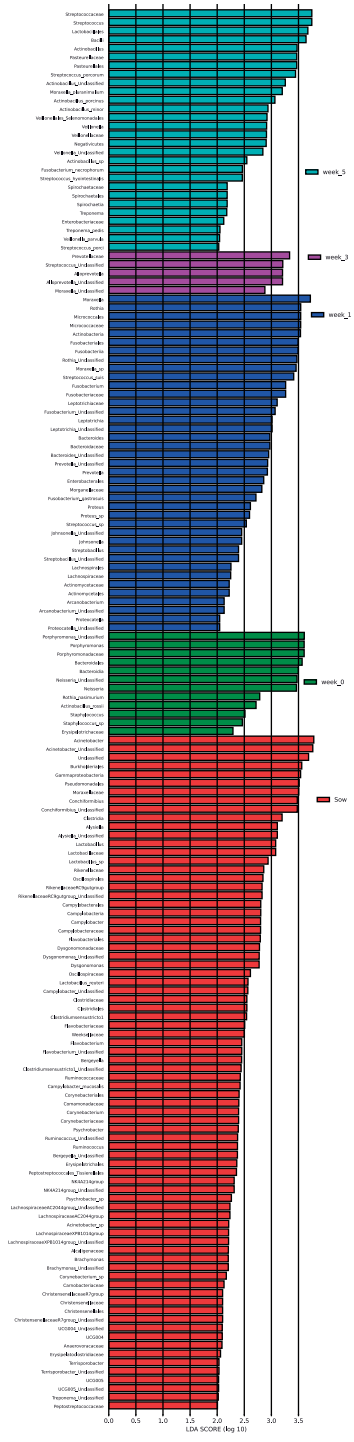


Figure S4. LEfSe analysis on taxa association with timepoints.

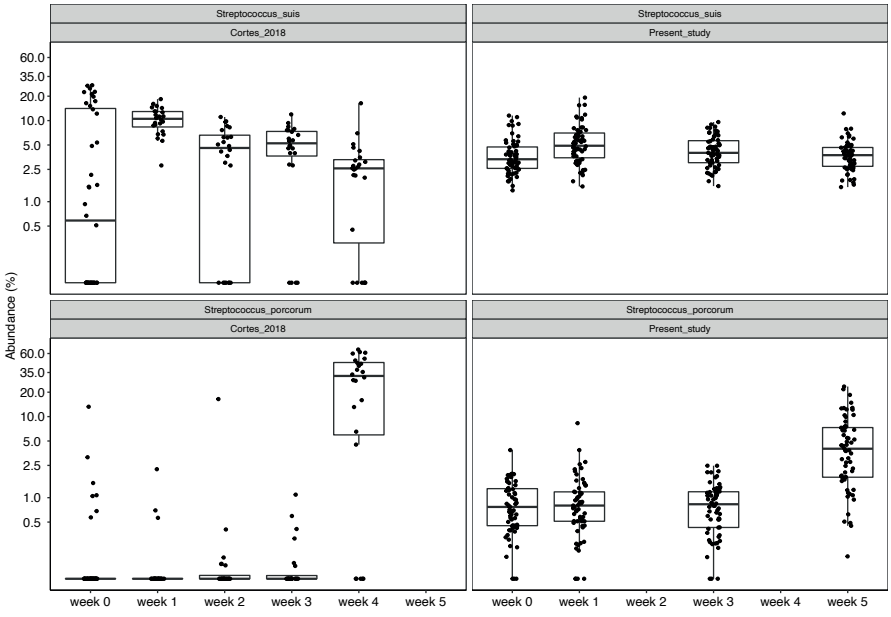


Figure S5. Comparison of *S. suis* and *S. porcorum* abundance at different timepoints in the present study and Pena Cortes et. al. 2018 (NCBI BioProject PRJNA391812). *S. suis* abundance decreases before and across weaning, while *S. porcorum* increases in abundance at weaning. The higher number of zero-counts found in the dataset of Cortes et al. is in part due to larger variation in sequencing depth. 12 samples had less than 1000 reads.

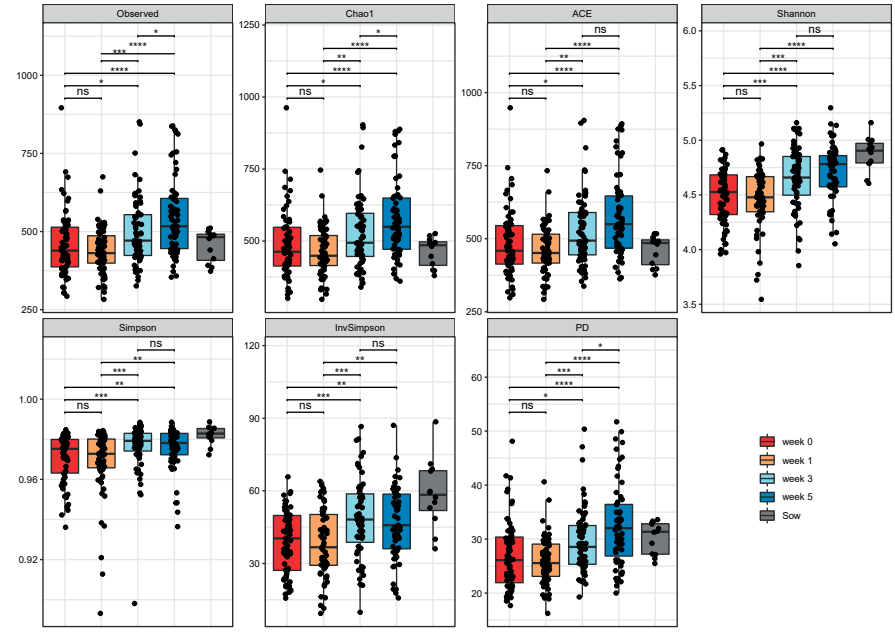


Figure S6. Boxplots of different alpha diversity measures per timepoint.

Chapter 4

Resistome expansion in disease-associated human gut microbiomes

Simen Fredriksen[1], Stef de Warle[1], Peter van Baarlen[1], Jos Boekhorst[1], Jerry M. Wells[1]

1. Host-Microbe Interactomics Group, Animal Sciences Department, Wageningen University, Wageningen, The Netherlands

Under review

Background

The resistome, the collection of antibiotic resistance genes (ARGs) in a microbiome, is increasingly being viewed as relevant to the development of antibiotic resistance (ABR). Many metagenomic studies have reported resistome differences between groups, often in connection with disease and/or antibiotic treatment. However, the consistency of resistome associations with different antibiotic- and non-antibiotically treated diseases has not been established. In this study we re-analysed human gut microbiome data from 26 case-control studies to assess the link between disease and the resistome.

Results

The human gut resistome is highly variable between individuals both within and between studies. We found that for diseases commonly treated with antibiotics, namely cystic fibrosis and diarrhoea, patient microbiomes had significantly elevated ARG abundances compared to controls. Disease-associated resistome expansion was found even when ARG abundance was high in controls, suggesting ongoing and additive ARG acquisition in disease-associated strains. We also found a trend for high ARG abundance in cases from some studies on diseases that are not treated with antibiotics, such as colorectal cancer.

Conclusions

Total ARG abundances can vary between case and control groups, also in the absence of large taxonomic differences in microbiome composition. Resistome expansion in case-control studies on diseases commonly treated with antibiotics suggests that exposure to antibiotics exerts considerable selective pressure for ARG acquisition. It also suggests strong strain-level association with the disease. Weaker but variable case-control resistome associations in studies on non-antibiotic treated diseases suggest that unbalanced cohort selection and confounding factors such as host health and lifestyle may influence resistome studies.

Abstract

Introduction

Antibiotic production and resistance are ancient traits important to competition between bacteria [174]. However, medical antibiotic use has driven an increase in antibiotic resistance (ABR) in human- and livestock-associated bacteria [38], and ABR in pathogenic bacteria has become a major problem for human and veterinary medicine [175]. With a One Health perspective in mind, identifying factors driving the spread of ABR in humans, livestock, and the environment is of great importance [176, 177].

Antibiotic resistance is often based on acquisition of antibiotic resistance genes (ARGs). ARGs can spread rapidly in bacterial populations via horizontal transfer (within and across species boundaries) by bacteriophages, plasmids, and transposable genetic elements [36, 37]. The epidemiology and spread of ARGs has been studied in clinically relevant bacteria, but the role of the commensal microbiome in the spread of ARGs is of increasing interest [38]. The microbiome contains a stable reservoir of ARGs, collectively termed the resistome, that can be acquired by con- and intraspecific bacteria. This reservoir provides genetic material that can enable pathogenic strains to rapidly adapt upon infection and antibiotic treatment [39, 40, 178].

Human, animal, and environmental microbiome studies have revealed differences in abundance and diversity of ARGs (i.e., the resistome) between sites, groups, and populations, suggesting recent or ongoing selective pressure for antibiotic resistance. Antibiotic usage induces positive selection for ARG acquisition, but other forces work to reduce ARG carriage. ARGs can impart a fitness cost in the absence of antibiotic exposure, and this is considered to select for loss of resistance after cessation of antibiotic treatment [5, 179]. Strain level microbiome composition and the resistome may also be equalized within populations by horizontal microbiota transfer [45, 180–182].

Resistome differences are of particular interest when related to disease because of the clinical relevance and impact on choice of antibiotic treatments. If a disease is treated with an antibiotic, then any disease-associated microbiome members that acquire an ARG conferring resistance to that antibiotic have a selective growth advantage. ARGs can co-occur with virulence genes on genomic islands [183], and the two classes of genes may confer synergistic selective advantages to disease-associated strains when occurring together. These processes may lead to increased ARG abundance (resistome expansion) in disease-associated microbiomes. While some studies have theorised this mechanism to explain differences in ARG abundance between groups [150, 184–187], a comprehensive overview of disease-associated resistomes is lacking. It is not known what resistome differences can be expected based on factors other than natural selection from antibiotic exposure. Host inflammation and oxidative stress may promote phage lysogeny and increased horizontal gene transfer and select for disease-associated bacteria rich in ARGs [184, 188–190]. Unbalanced cohort selection may also bias microbiome studies [191, 192].

It is not known whether there is a consistent link between host disease and an expanded resistome. The difference between the number of studies reporting a positive association and studies reporting a negative association could reflect publication bias, as positive associations may be more likely to be published. In this study we aimed to obtain a comprehensive overview of disease-associated resistomes in human gut microbiome studies. We re-analysed 26 studies with publicly available metagenomic data from healthy controls and cases with various morbidities. Some of the included studies investigated diarrhoea and cystic fibrosis, which are commonly treated with antibiotics. Other studies investigated diseases not commonly treated with antibiotics, and thus not expected to be associated with an expanded resistome.

Results and discussion

Cases with diseases mainly treated with antibiotics feature enlarged resistomes

To analyse the relationship between different disease diagnosis (such as colorectal cancer vs benign adenoma) and cohorts within each study, we divided the data from the 26 studies into 35 datasets. We excluded samples from individuals that had been recently treated with antibiotics, except for cystic fibrosis cases as many patients take antibiotics daily to prevent lung infections [193]. We used kraken 2 [139] to assess taxonomic composition and the proportion of bacterial and host DNA, and mapped reads to the Resfinder database [107] to assess ARG abundances. To assess case-control differences in ARG abundance and composition within each dataset, we compared the summed abundance of all ARGs in each sample and compared case and control cohorts with Wilcoxon rank sum test. We also assessed taxonomic and ARG composition with PERMANOVA (Figure 1).

Case-control studies on cystic fibrosis and diarrhoea, diseases where antibiotics are the main treatment, showed greater disease-associated resistome expansion (higher ARG abundance in cases than healthy controls within the dataset) than studies on diseases not treated with antibiotics ($p < 0.0001$, Wilcoxon rank-sum test). The 4 datasets of antibiotic-treated diseases were all among the 5 studies with the greatest resistome expansion in cases. Of the 35 datasets, 22 had higher total ARG abundance in cases and 13 had higher abundance in controls. 9 datasets had significantly ($p < 0.05$, Wilcoxon rank sum test) higher total ARG abundance in cases while only 1 study had significantly lower ARG abundance in cases.

The largest difference in ARG abundance between case and control participants was found in cystic fibrosis (CF) patients. While a range of studies have investigated the CF-associated microbiome [194], only a single, Australian, dataset met the inclusion criteria of being a gut microbiome, metagenomic, case-control study [195, 196]. The Australian CF cases showed high total ARG abundance compared to the study controls but overlapped with both case and control samples collected in other studies from countries with higher antibiotic usage, such as India, Bangladesh, and China. It is possible that horizontal

microbiota transfer between individuals within the Australian population, where antibiotic usage and ABR levels are low, limits ARG abundance in patients by continuously introducing susceptible strains. CF patients in countries with high baseline ARG abundance in the general population likely reach higher ARG abundances.

A

Metadata				Total ARG abundance				Composition PERMANOVA			
Study	Condition	Country	n	Case	Control	Ratio	Wilcox_p	R2_species	R2_ARG	p_species	p_ARG
Wang_2018	Cystic fibrosis	AU	35	1253	348	3.60	0.00	0.20	0.17	0.00	0.00
DavidLA_2015	Diarrhoea	BD	19	1777	583	3.05	0.00	0.51	0.47	0.00	0.00
KieserS_2018	Diarrhoea	BD	27	3043	1302	2.34	0.00	0.35	0.28	0.00	0.00
ThomasAM_2018a	CRC	IT	53	1548	850	1.82	0.05	0.04	0.03	0.04	0.06
SotoGiron_2021	Diarrhoea	EC	77	1458	908	1.60	0.01	0.04	0.04	0.02	0.00
GuptaA_2019	CRC	IN	60	821	596	1.38	0.02	0.08	0.15	0.00	0.00
NagySzakalD_2017	ME/CFS	US	100	746	545	1.37	0.23	0.03	0.01	0.02	0.33
JieZ_2017	Atherosclerosis	CN	385	1086	825	1.32	0.00	0.02	0.04	0.00	0.00
FengQ_2015	CRC	AT	107	390	306	1.28	0.01	0.02	0.03	0.01	0.00
ThomasAM_2018a	Adenoma	IT	51	1068	850	1.26	0.03	0.02	0.03	0.44	0.05
LiJ_2017	Hypertension	CN	140	767	615	1.25	0.10	0.01	0.02	0.22	0.03
QinN_2014	Cirrhosis	CN	237	971	821	1.18	0.00	0.05	0.03	0.00	0.00
WirbelJ_2018	CRC	DE	82	424	376	1.13	0.35	0.03	0.02	0.03	0.22
KarlssonFH_2013	T2D	DE, SE	88	332	300	1.10	0.19	0.02	0.02	0.05	0.11
HMP_2019_ibdmdb	IBD	US	91	485	441	1.10	0.34	0.01	0.01	0.23	0.29
QinJ_2012	T2D	CN	195	948	867	1.09	0.44	0.01	0.01	0.03	0.10
ThomasAM_2018b	CRC	IT	60	606	568	1.07	0.45	0.04	0.03	0.03	0.05
KarlssonFH_2013	IGT	DE, SE	87	314	300	1.05	0.59	0.02	0.02	0.03	0.02
YuJ_2015	CRC	CN	112	916	889	1.03	0.87	0.01	0.01	0.11	0.53
NielsenHB_2014	IBD	ES	141	705	685	1.03	0.99	0.01	0.01	0.13	0.65
HanniganGD_2017	CRC	CA, US	55	299	296	1.01	0.90	0.05	0.02	0.00	0.46
FengQ_2015	Adenoma	AT	108	308	306	1.01	0.35	0.01	0.01	0.77	0.59
LeChatelierE_2013	Obesity	DK	278	346	347	1.00	0.37	0.00	0.01	0.38	0.06
HallAB_2017	IBD	US	32	342	344	0.99	0.80	0.05	0.06	0.13	0.03
SankaranarayananK_2015	T2D	US	37	569	578	0.98	0.07	0.05	0.06	0.05	0.00
YachidaS_2019	Adenoma	JP	313	569	578	0.98	0.85	0.00	0.00	0.30	0.43
VogtmannE_2016	CRC	US	103	374	387	0.97	1.00	0.02	0.01	0.05	0.88
ZellerG_2014	CRC	FR	114	620	643	0.96	0.92	0.04	0.01	0.00	0.08
HanniganGD_2017	Adenoma	CA, US	54	284	296	0.96	0.87	0.03	0.01	0.15	0.72
YachidaS_2019	CRC	JP	504	554	578	0.96	0.14	0.00	0.00	0.04	0.09
ZhuF_2020	Schizofrenia	CN	171	927	996	0.93	0.24	0.01	0.01	0.10	0.17
YuJ_2015	T2D	CN	82	791	889	0.89	0.30	0.02	0.02	0.24	0.09
ZellerG_2014	Adenoma	FR	103	523	643	0.81	0.06	0.01	0.01	0.21	0.15
YeZ_2018	Behcet's disease	CN	65	961	1226	0.78	0.74	0.01	0.02	0.71	0.20
RubelMA_2020	Helminth infection	CM	175	509	780	0.65	0.00	0.06	0.05	0.00	0.00

B

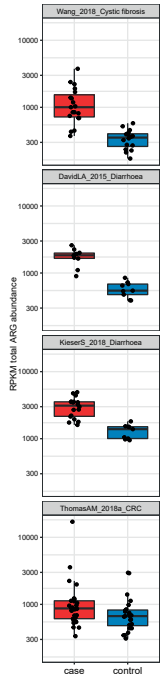


Figure 1: Resistome case-control associations. **A)** Summary statistics per study/disease, sorted from strongest to weakest total ARG abundance case-association. The columns under total ARG abundance show mean reads per kilobase per million reads (RPKM) total ARG abundance in case and control samples, the ratio of these, and Wilcoxon rank sum test p-value for case vs control samples per study. The PERMANOVA columns show Bray-Curtis dissimilarity PERMANOVA R2 and p values for species level taxonomy and ARG composition. CRC = colorectal cancer, adenoma = non-cancerous tumour, ME/CFS = chronic Fatigue Syndrome, T2D = type 2 Diabetes, IGT = impaired glucose tolerance, IBD = inflammatory bowel disease. **B)** Boxplot of the total ARG abundance for the 4 datasets with the strongest case-association in ARG abundance. Boxplots for all datasets are shown in Figure S1.

The diarrhoea-associated resistome

The diarrhoea-associated microbiome has been reported to have a distinct compositional profile influenced by frequent application of antibiotics [185, 186, 197–200]. Our re-analysis of 3 publicly available diarrhoea datasets found that, while the causative agent may vary, diarrhoea cases share increased abundance of *Enterobacteriaceae* species including *Escherichia coli*, *Salmonella enterica*, *Shigella dysenteriae*, and *Klebsiella pneumoniae* (Figure 2A). Occurrence and abundance of *Vibrio cholerae* was limited except for in the dataset of David et. al. 2015, which specifically studied *Vibrio cholerae* associated diarrhoea.

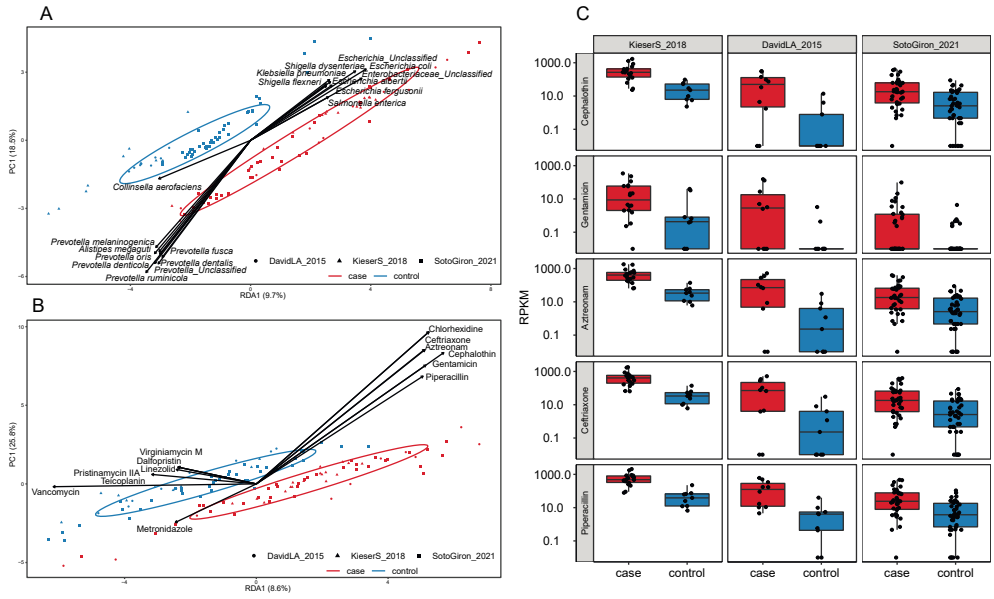


Figure 2: The diarrhoea-associated microbiome and resistome. **A)** Redundancy analysis (RDA) showing the species most strongly separating case and control samples in diarrhoea studies. Each label represents a single sample, eclipses represent 75% confidence level, arrows show genes driving the separation of samples; samples in the direction the arrow is pointing have higher expression of the gene. Species relative abundance input data was transformed by log(1000*abundance+1) and the study was used as RDA covariate. **B)** RDA on the summed abundance of all ARGs per Resfinder conferred resistance phenotype. **C)** Boxplots of the main Resfinder ARG phenotypes separating case samples on the RDA axis. Disinfectant ARGs omitted.

Despite excluding samples collected after antibiotic treatment, we found strong expansion of the diarrhoea-associated resistome. The dataset from Soto-Girón et. al. 2021 [200], which assessed both urban and rural diarrhoea cases in Ecuador, showed a 60% increase in total ARG abundance in cases. The datasets from Kieser et. al. 2018 [186] and David et. al. 2015 [197], both using samples collected in Bangladesh, showed 2-3 fold increase in total ARG abundance in cases. In both studies, case and control participant cohorts were not entirely equivalent. In Kieser et. al. 2018 there was a mismatch in age and social class, and David et. al. 2015 included two cohorts sampled at different times, where only cohort 1 included healthy controls and only cohort 2 included cases sampled prior to antibiotic treatment. In all 3 studies the cases had high abundance of ARGs conferring resistance to cephalothin, gentamicin, spiramycin, and aztreonam (Figure 2B-C). Ampicillin resistance was the most significantly case-associated ARG phenotype in the datasets from Kieser et. al. 2018 and Soto-Girón et. al. 2021 (652 vs 67 RPKM and 240 vs 67 RPKM per study, respectively). In the data from David et. al. 2015, chloramphenicol resistance genes were most significantly associated with cases (120 vs 6 RPKM). Single genes such as bla_{TEM} (conferring resistance to beta-lactams) contributed up to 20% of the total ARG abundance increase in cases in all studies, but case-associated resistome expansion was also, to a

large degree, driven by less abundant genes. In all 3 studies, a larger number of ARGs were (regardless of significance) case-associated than control-associated (174 vs 82, 159 vs 102, and 189 vs 96, respectively). In David et. al. 2015 and Soto-Girón et. al. 2021, ARGs not significantly (FDR > 0.05) associated with either cases or controls had a larger total net contribution to increased ARG abundance in cases than ARGs significantly different (FDR < 0.05). Non-significant and significantly different ARGs summed up to 1025 vs 169 RPKM and 337 vs 212 RPKM in David et. al. 2015 and Soto-Girón et. al. 2021, respectively. Kieser et. al. 2018 showed a more equal contribution of non-significant and significantly case-control associated ARGs to case-associated resistome expansion (839 vs 902 RPKM).

Differences in ARG abundance in diseases not treated with antibiotics

In addition to exposure to antibiotics, the resistome may differ between case and control cohorts due to confounding factors. This may for instance involve biases in participant selection or inflammation facilitating colonization by generalist disease-associated strains enriched in ARGs (i.e., strains associated with the hospital environment or several different diseases).

Inflammatory bowel disease (IBD, including both ulcerative colitis and Crohn's disease) involves bouts of intestinal inflammation. Early-life antibiotic usage has been suggested to predispose IBD [201] and antibiotics may in some cases be used to treat IBD complications [202]. As previously described [203, 204], the abundance of taxa such as *Faecalibacterium prausnitzii*, *Ruminococcus gnavus*, and *Bacteroides fragilis* is associated with IBD status, although there is high inter-individual variation and differing results between studies. All 3 re-analysed IBD datasets [95, 203, 205, 206] showed a neutral resistome case-control association, suggesting that gut inflammation and dysbiosis is by itself not, or only weakly, associated with increased ARG abundance.

Several studies have investigated the microbiome associated with colorectal cancer (CRC) [96, 97, 207–213]. CRC is not treated with antibiotics, so the CRC-associated microbiome should not face increased selection pressure for ARG expansion. However, the CRC case-associated resistome might be influenced by confounding factors such as lifestyle and diet, which predispose CRC development in humans with a CRC risk genotype [214]. Most CRC datasets showed no significant difference in total or individual ARG abundance between cases and controls. However, Gupta et. al. 2019 (India) [208], Feng et. al. 2015 (Austria) [207], and Thomas et. al. 2018 cohort 1 (Italy) [96] showed higher total ARG abundance in cases, in particular for ARGs conferring resistance to spectinomycin, gentamicin, and ticarcillin (Figure 3).

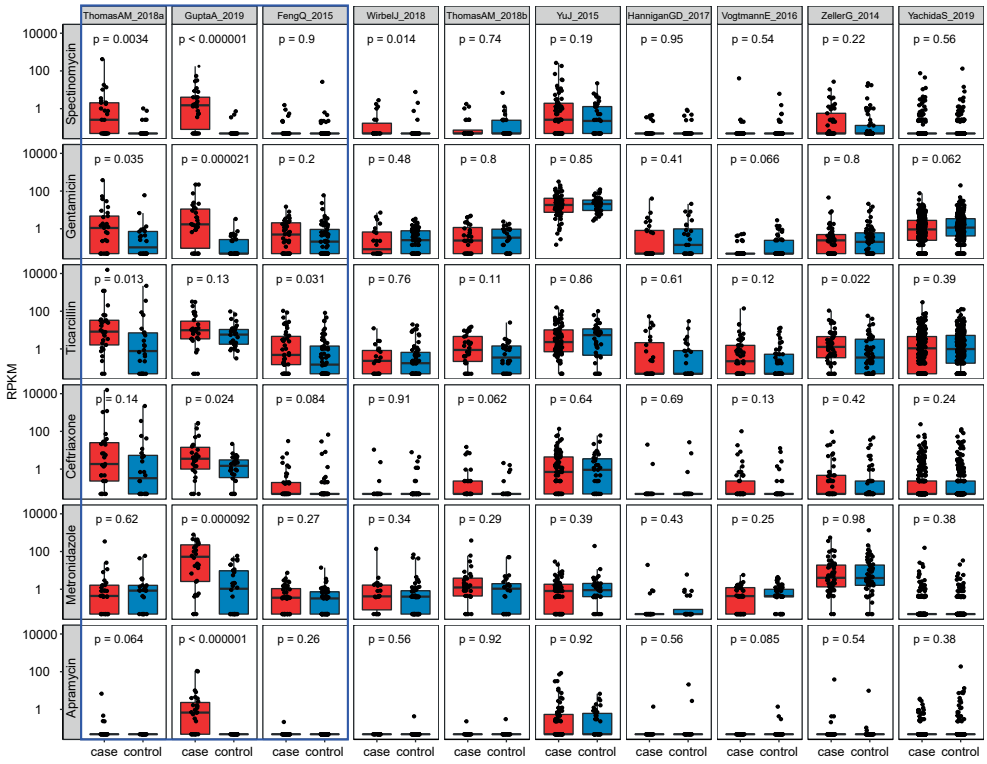


Figure 3: Boxplot of the Resfinder ARG phenotypes that showed strongest case-associations in the 3 CRC studies with overall case resistome expansion (left-most columns, highlighted in blue rectangle). Columns are sorted from strongest to weakest total RPKM ARG abundance case-association. Uncorrected Wilcoxon rank sum test p-values are shown.

Confounding variables and unbalanced study designs

Theoretically, a case-control study design should include case and control populations identical in all aspects except for the investigated disease. Such designs are feasible in lab experiments, but research on humans present large individual variation and lifestyle-associated environmental differences [191, 192]. This poses challenges because the microbiome and risk of disease development may be independently correlated with factors not accounted for, such as dietary habits [215] and socioeconomic status [216, 217]. Case participants may also acquire hospital-associated strains rich in ARGs by horizontal transfer.

We found that different studies on the same disease gave varying resistome case-control associations despite species level taxonomic change being in agreement. Studies investigating different colorectal cancer cohorts showed (trends towards) both higher and lower ARG abundance in cases, despite seemingly equivalent participant selection criteria. The most striking difference was between the two Italian CRC cohorts reported in Thomas et. al 2019 [96]. Cohort 1 (cohort a, collected in Vercelli) was characterised by

high overall high ARG abundance which was expanded in cases compared to controls. ARG abundance in cohort 2 (cohort b, collected in Milan) was lower and equal in cases and controls. Several individual ARGs were significantly differentially abundant in cohort 1, but no ARGs were significantly different within cohort 2. Previous meta-analysis found both cohorts to have a CRC-associated taxonomic compositional profile in agreement with other CRC studies [96, 97]. Gupta et. al. 2019 collected cases and controls from the same locations, but controls were collected as part of a separate study, and this may have contributed to larger taxonomic and resistome differences than the other CRC studies. All microbiome studies apply inclusion criteria to limit the impact of confounding variables on results, but it is possible that interpretation and application of sampling criteria by clinicians vary within and between studies. Stricter participant selection criteria and collection of additional metadata are warranted in human resistome studies.

Linking taxonomic composition and the resistome

For assessment of the immediate clinical relevance of an ARG it is key to determine its host and genomic context. Unfortunately, short-read metagenomic sequencing data is not well suited for assessing the genomic context of ARGs [178, 218–220]. Identical ARG copies occur in different taxa due to horizontal gene transfer, and adjacent genomic regions may also be shared if included in a mobile genetic element or genomic island. Among the re-analysed datasets, some studies showed large differences in both taxonomic composition and ARG abundance. Other studies, such as the Jie et. al. 2017 (on atherosclerosis), had 20-30% higher ARG abundance in cases despite case/control status only explaining 2% of the species level dataset variation (Figure 1). Differences in the resistome and taxonomic composition may occur due to interdependencies (e.g., colonization by different species inherently carrying different ARGs) or independently (abundance shifts among closely related strains with variable ARG content), but these scenarios cannot be differentiated by short-read metagenomic sequencing. This makes it challenging to determine whether increased ARG abundance occurs due to conferring an ecologically relevant trait in disease-associated strains or by correlation with taxonomic change. Taxa may differ in ARG content due to intrinsic resistance and variable selective pressures exerted by antibiotics, as well as different ARGs conferring the same trait. ARGs are often located on plasmids as a part of the accessory genome, and variably present in closely related lineages [12, 221], meaning that resistomes can differ despite limited observed change in taxonomic composition.

Furthermore, while some ARGs may have strong correlations linking them to specific species [222], this is rarely the case for the ARGs driving resistome expansion in cases. Among datasets with resistome expansion in cases, we found strong (auto)correlation in the case/control association of ARGs and the species they had the strongest correlation with, despite weak sample-by-sample correlations (Figure S2). Future studies may address these issues by utilizing long-read, Hi-C, and single-cell sequencing.

Resistome studies may also find differences in ARG abundance as a technical artefact of metagenomic sequencing methodology. Bacteria may carry (multiple) ARGs on (high-copy-number) plasmids, thus contributing greatly to the observed resistome compared to strains with a single chromosomal copy of the same ARG. Variable genome size may also influence strain contribution to the observed resistome. A microbiome with high abundance of taxa with small genome sizes but average numbers of ARGs will appear to have an abundant resistome. We do not expect that this contributed meaningfully to the strongest case-control differences reported in the present study, as commonly case-associated *Gammaproteobacteria* species have larger genomes than commonly control-associated *Lactobacillus* and *Prevotella* species. Plasmid carriage may play a modest role as *Enterobacteriaceae* commonly carry ARGs on plasmids, although these are large and occur in low copy numbers [223]. Data on plasmid copy numbers between (commensal) taxa could be of great relevance to resistome research, but we are not aware of any comprehensive studies on this topic.

Conclusions

The human gut resistome is highly variable between individuals, but strong between-group differences can still be observed in case-control studies. Antibiotic treatment of disease induces positive selection pressure for acquisition and maintenance of ARGs on the disease-associated microbiome, driving expansion of the disease-associated resistome. This implies that the disease-associated microbiomes contain strains with high disease-specificity that while transmitted between individuals and a part of the resident microbiota are more prevalent and abundant during disease and antibiotic treatment. High baseline resistome abundance in controls does not appear to limit further (additive) resistome expansion, underpinning the importance of limiting antibiotic usage in populations with high resistance levels. The resistome of case and control groups may show differences without any convincing biological explanation, and future resistome research should take great care in selecting equivalent study cohorts.

Methods

Study inclusion and data selection

We aimed to include human gut case-control metagenomic shotgun sequencing studies investigating any disease or morbidity with publicly available raw data and metadata. First, we included case-control studies from the curatedMetagenomicData database [224]. Then, we conducted a systematic literature review to identify additional metagenomic studies on cystic fibrosis and diarrhoea, diseases commonly treated with antibiotics by searching PubMed for (((microbiota OR microbiome OR metagenomics) (cystic fibrosis)

AND (shotgun))) NOT (Review[Publication Type]). Studies indexed by PubMed before 2022-07-22 were included. Studies with less than 10 case samples were excluded.

We reviewed the metadata of all studies to select either the full sample set or a subset of samples appropriate for case-control comparison. In longitudinal studies, we selected a single sample from each participant. In studies that collected samples from several cohorts and/or different countries we divided the dataset for separate analysis or excluded samples causing unbalanced study designs. For instance, if the majority of samples were collected from one country and additional cases but no controls were collected from a second country, we excluded these cases. We removed samples or cohorts that had received antibiotic treatment (unless it was not stated which subjects were treated). Some studies did not provide information on antibiotic treatment, and we opted not to exclude these because they did not show significant resistome case-control differences. Cystic fibrosis inherently features antibiotic treatment of case subjects, so these were also all included.

Data processing

We adapted publicly available data from the 26 case-control studies to balance the cases and controls and, where possible, excluding participants that had received antibiotic treatment prior to sampling. We also divided studies into separate datasets when several different diagnoses were investigated, and when samples were collected from separate cohorts and/or differed in geography or methodology. We used the Resfinder database [107] for ARG classification and kraken [139] taxonomic assignment to assess the abundance of bacterial and eukaryotic DNA. Although host-DNA contamination is limited in stool microbiome samples we normalized ARG abundances for the proportion of human DNA.

We used NCBI fastq-dump to download 5 million reads per sample. We limited the reads per sample to reduce computational time required and to avoid bias from uneven sequencing depth. We quantified taxonomic abundance of bacteria, archea, virus, fungi, protozoa, plant, and human DNA using kraken2 and the PlusPFP database 2021-01-27 [139]. To accurately assess the resistome relative abundance we normalised for proportion of host-DNA contamination in each sample (a sample with 50% host reads would otherwise appear to have half the ARG abundance of an equivalent sample with no host DNA contamination).

We used the Resfinder database [107] to identify antibiotic resistance genes (ARGs) and quantify their abundance. To reduce noise from ambiguous mapping to highly similar gene variants we clustered all sequences to 90% identity using MMseqs2 [136] easy-cluster with settings "--min-seq-id 0.9 --cov-mode 0". We mapped metagenomic reads to the clustering representative sequences clusters using MMseqs2 easy-search with setting -s 4.500 and accepted the best hit with minimum 50 bp alignment and 80% identity. ARG abundance was normalized to reads per kilobase per million reads (RPKM).

Statistical analysis

To determine whether the case and control groups of each study differed in overall ARG abundance we summed the RPKM abundance of all ARGs in each sample and calculated the mean per group. We used unpaired Wilcoxon rank-sum test to assess statistical significance. We used R package vegan [138] function RDA for principal component analysis (PCA) and redundancy analysis (RDA), function vegdist to calculate Bray-Curtis dissimilarity, and function adonis was used to perform PERMANOVA to determine the overall compositional difference.

Supplementary material

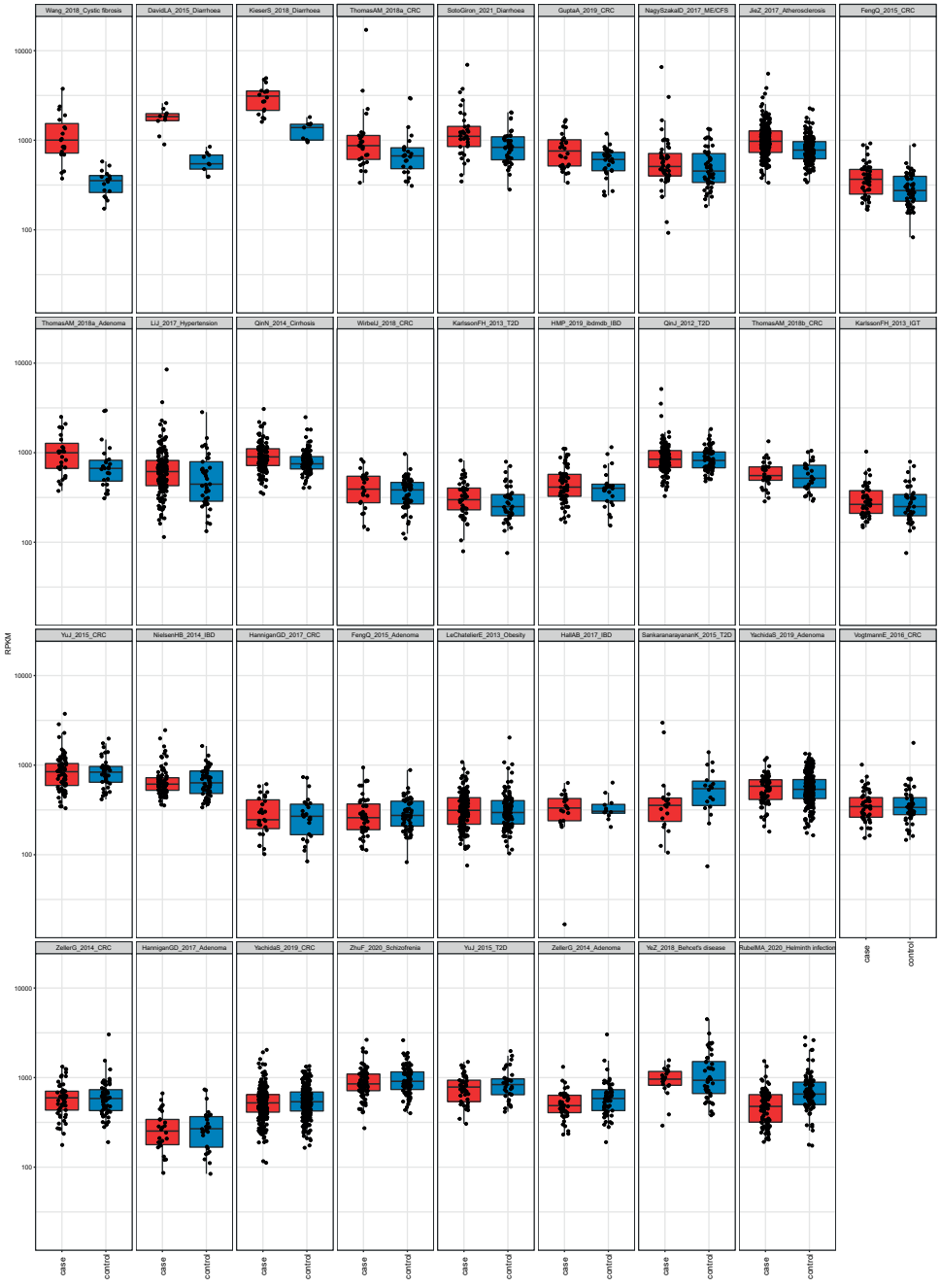


Figure S1: Boxplots comparing the summed abundance of all ARGs in case and control samples from all datasets. Some studies included different diseases; samples with several diseases and the corresponding controls are shown in several facets.

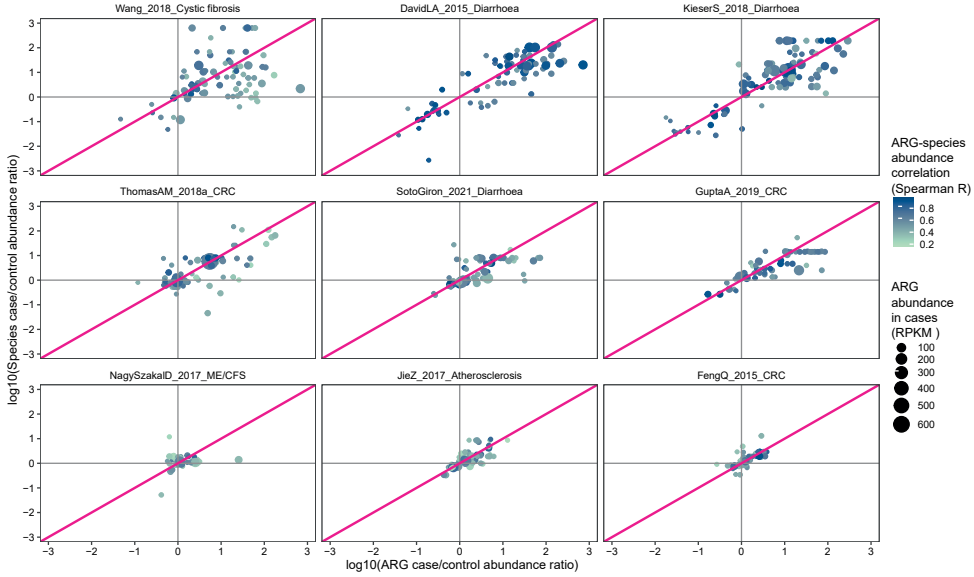


Figure S2: Strong (auto)correlation between the case-association of ARGs and species despite limited sample-by-sample co-occurrence. The figure shows the relationship between the disease-association of ARGs and the species they are most strongly correlated with (regardless of strength and significance of this correlation). Points represent pairs of each ARG (abundance of > 1 RPKM) and the species (mean abundance of > 0.01%) it has the strongest positive Spearman's rank correlation coefficient with. The purple line indicates 1:1 equal case/control association of ARG and species, which could be expected if the ARG is only found on the chromosome of a single species of average genome size. David et. al. 2015 has several, likely genuine, strong correlations due to consistently high abundance of *Vibrio cholerae* in cases.

Chapter 5

Transcriptomics in serum and culture medium reveal shared and differential gene regulation in pathogenic and commensal *Streptococcus suis*

Simen Fredriksen[1], Suzanne D. E. Ruijten[1], Gemma G. R. Murray[2], Maria Juanpere Borràs[1], Peter van Baarlen[1], Jos Boekhorst[1], Jerry M. Wells[1,2]

1. Host-Microbe Interactomics Group, Animal Sciences Department, Wageningen University, Wageningen, The Netherlands
2. Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom

Under review. Available as preprint:
bioRxiv 2022.10.16.512421
https://doi.org/10.1101/2022.10.16.512421

Streptococcus suis colonizes the upper respiratory tract of healthy pigs at high abundance but can also cause opportunistic respiratory and systemic disease. Disease-associated *S. suis* reference strains are well studied, but less is known about commensal lineages. It is not known what mechanisms enable some *S. suis* lineages to cause disease while others persist as commensal colonizers, or to what extent gene expression in disease-associated and commensal lineages diverge. In this study we compared the transcriptomes of 21 *S. suis* strains grown in active porcine serum and Todd-Hewitt yeast broth. These strains included both commensal and pathogenic strains, including several strains of sequence type (ST) 1, which is responsible for most cases of human disease and considered the most pathogenic *S. suis* lineage. We sampled the strains during their exponential growth phase and mapped RNA-sequencing reads to the corresponding strain genomes. We found that the transcriptomes of pathogenic and commensal strains with large genomic divergence were unexpectedly conserved when grown in active porcine serum, but that regulation and expression of key pathways varied. Notably, we observed strong variation of expression across media of genes involved in capsule production in pathogens, and of the agmatine deiminase system in commensals. ST1 strains displayed large differences in gene expression between the two media compared to strains from other clades. Their capacity to regulate gene expression across different environmental conditions may be key to their success as zoonotic pathogens.

Abstract

Introduction

Streptococcus suis is an opportunistic pathogen that can cause septicaemia and meningitis in pigs and humans, but also colonizes the upper respiratory tract of healthy pigs in high abundance [8]. Different *S. suis* lineages appear to be specialized to different niches, being made up of either predominantly clinical strains isolated from necropsy (pathogenic clades), or non-clinical strains isolated from the oral cavity of pigs (commensal clades) [64]. Clinical and non-clinical strains are, however, found in all clades, showing that pathogenic lineages can colonize the oral cavity of asymptomatic pigs at low abundance and that commensal lineages can invade the host in certain circumstances. Recent microbiome studies have confirmed that commensal *S. suis* colonize the oral cavity of piglets at high abundance [70, 150, 225].

Limited research has been conducted on commensal *S. suis* and the mechanisms underlying their great success as piglet oral biofilm colonizers. Commensal *S. suis* clades have high genomic diversity and larger genome sizes compared to pathogenic clades [63], and production of secondary metabolites may aid in antagonising con- and heterospecific competitors [66]. Pathogens have conserved virulence-associated genes but reduced accessory genomes [63]. To successfully infect hosts, pathogens need to rapidly respond to different environments and shifts in nutrient availability and host immune responses. Firstly, they need to colonize the oral cavity to gain access to the host tissue and to persist in herds between outbreaks to ensure transmission. Secondly, they need to cross host epithelium, enter the bloodstream, and survive and proliferate in host tissues and body fluids with varying nutrient availability while avoiding eradication by host defences. To cause meningitis, *S. suis* also crosses the blood-brain barrier. *S. suis* niche differentiation is poorly understood, and the exact combination of virulence-associated genes and gene-host-environment interactions necessary to cause invasive disease have not been established. It is possible that regulation of gene transcription by phase variation and master switches such as carbon catabolite repression are key to rapidly adapting to a changing environment [78, 226, 227].

Despite the great heterogeneity of *S. suis* lineages, most research has been focused on a few closely related and highly pathogenic strains. Transcriptomic studies have used sequence type 1 (serotype 2) strains associated with zoonotic disease, such as P1/7 and S10 [51, 228–232]. This has led to knowledge of how the most common pathogenic *S. suis* clade adapts to the host but left other pathogenic and commensal clades understudied. It is not known how well results from the commonly studied strains translate to other clades especially given the high functional redundancy of many virulence factors described for *S. suis* [233]. A greater understanding of the species-wide *S. suis* transcriptome may yield insight into the differences between commensal and pathogenic *S. suis* clades and increase our understanding of *S. suis* ecology and evolution.

In this study we compare the transcriptomes of 21 *S. suis* strains from a wide phylogenetic background and different isolation sources, including not only clinical strains from pathogenic clades and non-clinical strains from commensal clades, but also clinical strains from commensal clades and non-clinical strains from pathogenic clades. We determined growth curves in Todd-Hewitt Yeast Broth (THY) and active porcine serum (APS), which contains complement and scarce amounts of essential metals [51], and extracted RNA during the exponential growth phase. We mapped RNA-seq data to the individual strain genomes and compared normalized sequencing coverage per gene. The resulting dataset allowed us to gain a better understanding of shared and niche-specific gene expression in commensal and pathogenic *S. suis*.

Materials and methods

Strain selection

We selected 21 *S. suis* strains that are broadly representative of the species, including both clinical and non-clinical strains from different phylogenetic backgrounds (Table 1). Where available we included closely related clinical and non-clinical strains for comparison (i.e., not only clinical strains from pathogenic clades and non-clinical strains from commensal clades, but also clinical strains from commensal clades and non-clinical strains from pathogenic clades). While the isolation source and phylogenetic clade of a strain gives an indication of its virulent potential, this has only been experimentally tested *in vivo* for few strains. Strains P1/7 and S10 are known to be from a pathogenic clade and highly virulent, while strain T15 has been experimentally shown to have low virulence in pigs despite being from a pathogenic clade [234].

This study aimed to assess the overall species transcriptome and compare groups of strains rather than focus on specific gene variants or individual strains. Thus, we prioritized RNA-sequencing of a single sample from many different strains (biological replicates) rather than technical replicates of each strain. To benchmark the repeatability of our methods we performed RNA sequencing on 3 technical replicates of S10 grown in each medium, processed separately on different days.

Table 1: Strain overview. Shortened 3-character name used in figures and tables, full original strain name, serotype, sequence type (ST), number of plasmids in genome assembly, and strain metadata. NT = non-typeable.

Name	Strain	Country	Serotype	ST	Clade type	Isolation source	Plasmids
C01	M101999_C1	Spain	8	NT	Pathogenic	Non-Clinical	1
C15	SS15055_N2_C15	Spain	2	1	Pathogenic	Non-Clinical	0
D11	DNS11	Denmark	9	16	Pathogenic	Systemic/Brain	1
D13	DNC13	Denmark	NT	NT	Commensal	Non-Clinical	2
D15	DNC15	Denmark	16	NT	Commensal	Non-Clinical	0
D20	DNS20	Denmark	NT	NT	Commensal	Systemic/Brain	2
D43	DNR43	Denmark	2	28	Pathogenic	Respiratory	1
D48	DNR48	Denmark	8	NT	Pathogenic	Respiratory	1
D49	DNC49	Denmark	2	28	Pathogenic	Non-Clinical	0
G69	DE609B	Germany	2	1	Pathogenic	Systemic/Brain	0
I12	21435_1	Spain	31	NT	Commensal	Non-Clinical	0
I27	21437_3	Spain	4	NT	Commensal	Non-Clinical	0
L42	LSS42	UK	16	NT	Pathogenic	Non-Clinical	0
P17	P1/7	UK	2	1	Pathogenic	Systemic/Brain	0
S10	S10	Netherlands	2	1	Pathogenic	Systemic/Brain	0
S11	M102942_S11	Spain	9	123	Pathogenic	Systemic/Brain	1
S20	M104300_S20	Spain	2	1	Pathogenic	Systemic/Brain	0
S26	M105052_S26	Spain	19	NT	Commensal	Systemic/Brain	0
S40	M106471_S40 1	Spain	30	NT	Commensal	Systemic/Brain	0
T15	T15	Netherlands	2	19	Pathogenic	Non-Clinical	2
X15	1521251	Canada	7	NT	Pathogenic	Systemic/Brain	0

Hybrid genome assembly

We created new Illumina-Nanopore hybrid genome assemblies for 19 of the 21 strains, as only strains P1/7 and S10 had complete genomes available. T15 had a circular genome assembly that lacked plasmids. We isolated DNA from overnight cultures using the PowerSoil DNA Isolation Kit (Qiagen). Generation of short-read sequencing data for the strains from Germany, Canada, and the UK has been previously reported [12, 64, 150]. Strains without available short-read sequencing data were 250 bp paired-end sequenced using an Illumina HiSeq 2500 instrument by MicrobesNG (Birmingham, UK). Nanopore sequencing was done with the SQK-LSK-109 Ligation Sequencing Kit and Guppy Software v5.0.16 with high accuracy base calling. Genomes were assembled using Unicyler v0.4.9 [235] with default settings. See table S1 for genome statistics.

RNA extraction and sequencing

We sampled the strains in the mid exponential growth phase during growth in Todd-Hewitt Yeast Broth (THY) and active porcine serum (APS). THY is a rich lab medium used to rapidly grow streptococci to high density, and consists of several ingredients including meat infusion, tryptone, glucose, and yeast extract. Serum is extracted from blood by removing cells and clotting factors and although it contains high glucose concentrations essential minerals are scarce. For instance, iron is sequestered by host transferrin to limit bacterial growth in blood. Active serum is not heat-treated and may thus contain

complement factors and antibodies binding to *S. suis* as this pathogen is endemic on farms. *S. suis* has several mechanisms to evade complement activation and formation of the membrane attack complex, including the capsule, factor H binding proteins, and other less well understood mechanisms [233, 236].

Growth curves were started with an overnight culture of *S. suis* in THY and adjusted to OD_{600} 0.5 by dilution in PBS and subsequently inoculated 1:10 in either THY or APS. For determination of growth curves the strains were grown in 96-well plates with 200 µL total volume per well and incubated at 37 °C with 5% CO2. The plates were shaken to prevent sedimentation and measured with a SpectraMax M5 (Molecular Devices) every 30 minutes. Each strain/medium combination was grown in 3 separate plates with 2 technical replicates per plate.

For RNA isolation 10 mL cultures were grown in 15 mL falcon tubes at 37 °C with 5% CO_2. All experiments were carried out using a single batch of THY and APS. Based on the growth curves we identified a timepoint where all strains were in the mid-exponential growth phase. To account for different growth in falcon tubes compared to 96-well plates, we reduced the incubation time and confirmed that the culture was in the correct growth phase by measuring OD_{600} before RNA isolation. Cultures in APS were sampled at 95 minutes and THY cultures at 125 minutes. The APS and THY cultures were started at different times to enable RNA isolation at the same time. Cultures were pelleted by centrifugation and resuspended in QIAzol Lysis Reagent (Qiagen). After bead beating twice for 40 s with 0.1 mm silica beads (MP biomedicals) RNA was isolated using the miRNeasy kit (Qiagen). Trace DNA was digested using the DNase Max Kit (Qiagen). RNA quantity and integrity was confirmed with nanodrop (ThermoFisher) and TapeStation (Agilent). Library preparation was done with the Illumina Stranded Total RNA Prep kit. rRNA was enzymatically depleted and the remaining RNA fragmented and translated to cDNA. The libraries were sequenced with 150 bp paired-end sequencing on the Illumina NovaSeq 6000 system with Illumina NovaSeq 6000 SP reagent kit at iGenSeq (Institut du Cerveau, France).

Bioinformatics

The strain genomes were annotated with Prokka v1.14.5 [133]. We further identified antimicrobial resistance genes with Resfinder v4.1 [107] and biosynthetic gene clusters with antiSMASH v5.1.2 [237] and BiG-SCAPE v1.1.2 [238]. We used clinker [239] for gene cluster comparison. To compare the transcriptome between strains we grouped homologous and paralogous genes using Orthofinder v2.3.12 [79, 240], and the gene expression for paralogous genes was summed. The "orthogroups" found by Orthofinder were given static names from the RefSeq locus tags of *S. suis* reference strains P1/7 (GCF_000091905.1) and D12 (GCF_000231905.1). GO terms for each Orthogroup were found with InterProScan v5.54-87.0 [241], and quantitative analysis of GO term expression was done on the whole transcriptome, including accessory genes. The RNA sequencing data was adapter and

quality trimmed with Trimmomatic v0.39 [131] before being mapped to the genome of the individual strain with bowtie2 [242]. FeatureCounts 2.0.1 [243] with default settings was used to count the number of reads mapping to each gene. Further analysis was carried out with R v4.1.3 [244].

Analysis

Combining the transcriptomes of different strains into a single analysis requires additional normalization compared to single-strain transcriptomic studies. In addition to variable sequencing depth per sample, strains vary in genome size and presence/absence and length of genes. We opted to use TPM read count normalization to facilitate comparison between strains. We identified differentially expressed genes using Wilcoxon Rank Sum Test with FDR correction. Overall transcriptome difference was calculated separately using the whole genome transcriptome (expression of accessory genes absent in strain set to 0) and the core genome transcriptome (using only core genes shared by all strains). To quantify and visualise overall transcriptome difference we used R package vegan [138] Bray-Curtis dissimilarity, principal component analysis (PCA), and redundancy analysis (RDA). To determine if the transcriptome conservation differed significantly between groups, we used estimated marginal means on linear models with R function emmeans [245].

Results and discussion

Dataset description

We sequenced the transcriptome of a genetically diverse set of 21 *S. suis* strains (Figure 1A) in the mid exponential growth phase in Todd Hewitt Yeast Broth (THY) and active porcine serum (APS). Overall, commensal and pathogenic strains grew equally well in both media (Figure S1). All samples were successfully sequenced with a minimum of 20 million 150 bp paired-end reads (1114-1902x coverage). All samples were free from contamination, with >99.8% reads mapping to the corresponding strain genome. One sample, S11 grown in THY, had excessively high expression of the arginine deiminase system (ADS, Figure S2). ADS is previously described as vital for *S. suis* growth in acidic medium [246, 247]. We considered that the high ADS expression was likely to be linked to *in vitro* culturing and to have limited *in vivo* relevance. Considering the excessive impact of this single gene cluster on the total transcriptome we excluded the sample from the main analysis.

Comparison of S10 triplicate samples showed that our methods were highly reproducible, with a maximum pairwise Bray-Curtis dissimilarity of 0.1 between replicates. Distinct but very closely related strains within sequence type 1 (ST1) had similar pairwise dissimilarities, but this was expected as their genomes are virtually identical. It does, however, show that small differences in gene expression between strains should be interpreted with care. For the remaining analysis only a single S10 sample from each medium was included.

Serum opacification

Three commensal strains increased the OD_{600} rapidly and without lag phase when grown in APS, reaching far higher OD_{600} values than the other strains (Figure S1). This was not due to an increase of bacterial biomass, but opacification of the supernatant. Formation of large lipid particles by the protein opacification factor of serum (*ofs*, SSU_RS07445) has previously been described in *S. suis*, and the gene is known to occur in different variants with and without opacifying function [248, 249]. The three strains with strong opacification activity had high expression of *ofs* variants diverged from those previously described [249]. While the shorter *ofs* type-1 [249] variant found in ST1 is expressed at low levels and has limited opacification capacity, it is conserved in pathogenic clades and may have a function related to binding to host cells [248, 250].

Conservation of core genome expression in serum

Commensal clades are diverse, with long (core- and whole genome-based phylogeny) branch lengths between strains, while pathogenic clades, in particular the commonly researched ST1 clade, consist of closely related strains [63]. Our strain selection included both ST1 strains and a wider range of strains from other pathogenic and commensal clades. Strains D13 and D20 form an outgroup to the other strains in our dataset (Figure 1A). The phylogenetic status of this "divergent" outgroup clade is debated (see closely related strains in "clade 2", Baig et. al. 2015) [251]. Despite large genomic differences (less than 86% ANI) to other included strains, their transcriptome was conserved in APS (measured in Bray-Curtis dissimilarity, Figure 1B). Overall, the *S. suis* transcriptome was significantly more conserved in APS compared to THY, and in the expression of the core genome compared to the whole genome transcriptome (estimated marginal means, $p < 0.01$). Despite the large genomic differences between outgroup D13+D20 and the remaining strains, pairwise transcriptome Bray-Curtis dissimilarities to the other strains was overlapping with that between strains from different pathogenic clades. The 13 most expressed GO terms in THY were all more expressed in APS (Figure S3). This suggests that the conserved APS transcriptome reflects upregulation of core physiological functions related to growth and cell division. This may be relevant in facilitating exponential growth during host invasion.

Large transcriptome differences between media in ST1 strains

Transcriptome difference between the two media was largest for ST1 strains (Figure 1C). Redundancy analysis (RDA) on log2fold change of gene expression between THY and APS cultures constrained by clade type (ST1 vs other pathogenic clades vs commensal clades) also showed that regulation of gene expression in ST1 strains was distinct, while other pathogenic clades and commensals were more similar (Figure 1D). The transcriptome of ST1 strains was, however, not strongly divergent from other clades in either medium. PCA

on all samples (Figure 1E) showed that the pairs of samples from each strain separated in similar directions on the PC1 and PC2 axis, except for the two outgroup commensal strains D13 and D20. ST1 strains showed stronger regulation of several genes. This included downregulation of SSU_RS07200 (SprT-like protein) and upregulation of SSU_RS02105 (cysteine synthase) in THY (Figure S4). ST1 and other pathogenic *S. suis* are thought to regulate gene expression related to virulence factors via a phase variable type I DNA methyltransferase system [226, 252]. However, this system is unlikely to explain the differences seen in the present study as the THY and APS cultures grew separately for only a few generations. This is unlikely to be sufficient time for selection to drive change across the experimental population. We did not observe differences in growth rates of the strains, indicating that selection pressure for either phase variant was limited, and that phase variation did not occur. Moreover, non-ST1 strains with the same phase variation system had small transcriptome differences between media. These results suggest that ST1 strains may have additional undescribed mechanisms enabling strong regulation of gene expression in different environmental conditions.

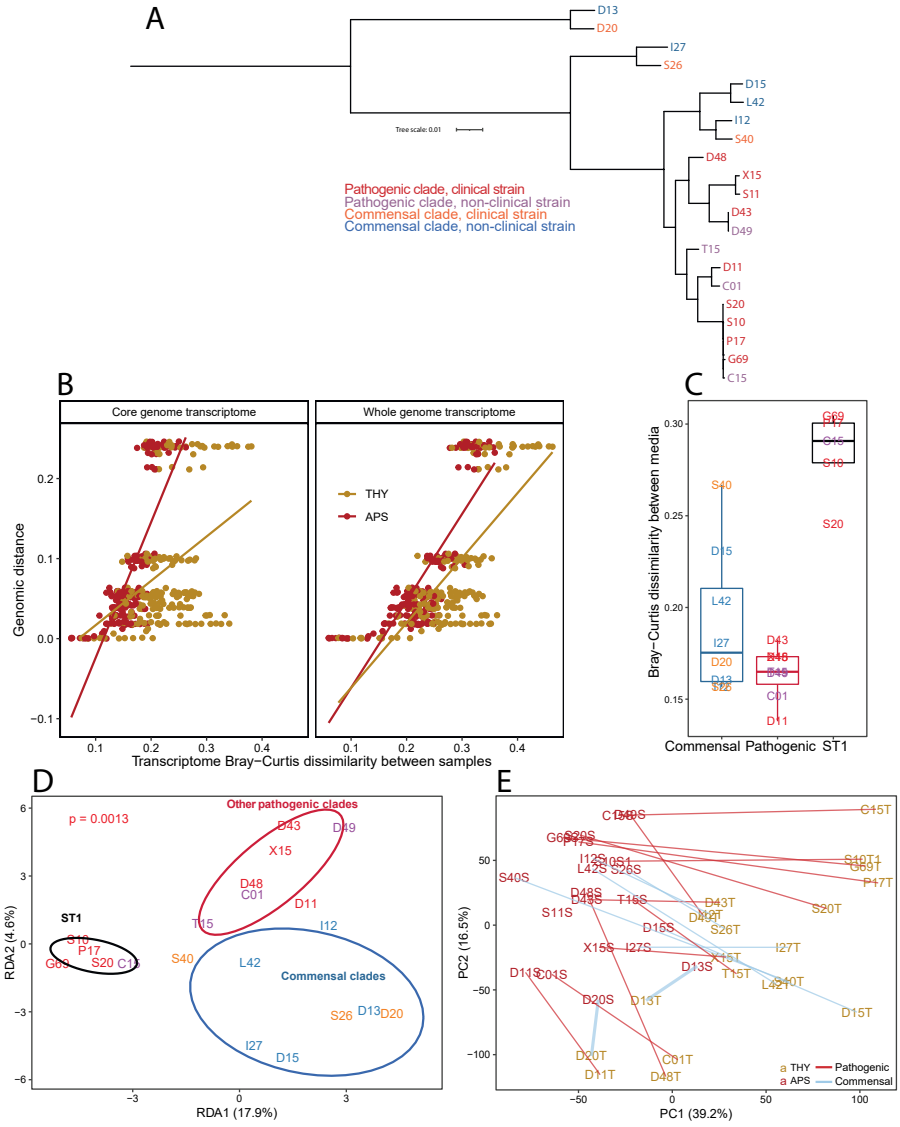


Figure 1. Strain phylogeny and overall transcriptome variation. **A)** Core genome phylogenetic tree of the 21 strains included in the study, constructed by Orthofinder and STAG with default settings and mid-point rooted. While ST1 pathogenic strains such as P1/7 and S10 are closely related, commensal clades have high genomic diversity. **B)** Transcriptome Bray-Curtis dissimilarity between strains correlates with genome phylogenetic distance. The transcriptome is more conserved in APS than in THY. Each point represents a pairwise comparison. Strains were only compared within the same medium. **C)** Boxplot of core genome transcriptome Bray-Curtis dissimilarity between samples from each strain in the two media. **D)** Redundancy analysis (RDA) on log2(fold-change) on the core genome transcriptome constrained by [ST1 vs other pathogenic clades vs commensal clades] showed that ST1 clade strains had a distinct transcriptome change between media. Eclipses represent 75% confidence level. **E)** Principal component analysis (PCA) showing core genome transcriptome separation between samples grown in THY and APS. The pairs of samples were well separated on PC1 in all strains except for strain D13 and D20. Each label (coloured by medium) is one sample, and the two samples of each strain are joined by lines coloured by clade type.

High cps expression in serotype 2 strains

The *S. suis* polysaccharide capsule (CPS) exists in many variants (serotypes). Serotype 2, a capsule type with terminal sialic acid, is the most studied serotype due to its association with high (zoonotic) virulence [103, 253, 254]. In this study all ST1 strains were serotype 2, in addition to strains T15 and D43+D49. Genes involved in capsule production (cps) were upregulated in APS in most strains (Figure 2A). Cps expression was high and strongly upregulated in APS in serotype 2 strains compared to other serotypes, although serotype 2 strains D43 and D49 appeared to have divergent regulation. D49 downregulated cps expression less, and D43 had higher cps expression in THY, opposite of other serotype 2 strains. Non-typeable strains D13 and D20 had the lowest cps expression, but apart from these and serotype 2 strains, cps expression overlapped between commensal (serotype 4, 16, 19, 30, and 31) and pathogenic clades (serotype 7, 8, and 9). The pathogenic clades had higher expression levels than the commensals, but this was only significant in APS ($p = 0.03$) and not in THY ($p = 0.15$).

Streptococcal *cps* expression has been linked to negative regulation by a protein (SSU_RS02755) variably named as *cpsR*/*gntR*/*orf2Y* [255–257]. In *S. pneumoniae*, *cpsR* has been shown to interact with the cps promoter dependent on glucose concentration, negatively controlling *cps* expression and CPS production [256]. We found a negative correlation between *cpsR* and *cps* expression only in serotype 2 strains (Figure 2B). This, and the lack of *cps* regulation in the serotype 2 strains D43 and D49, may be due to variation in *cpsR* (Figure S5).

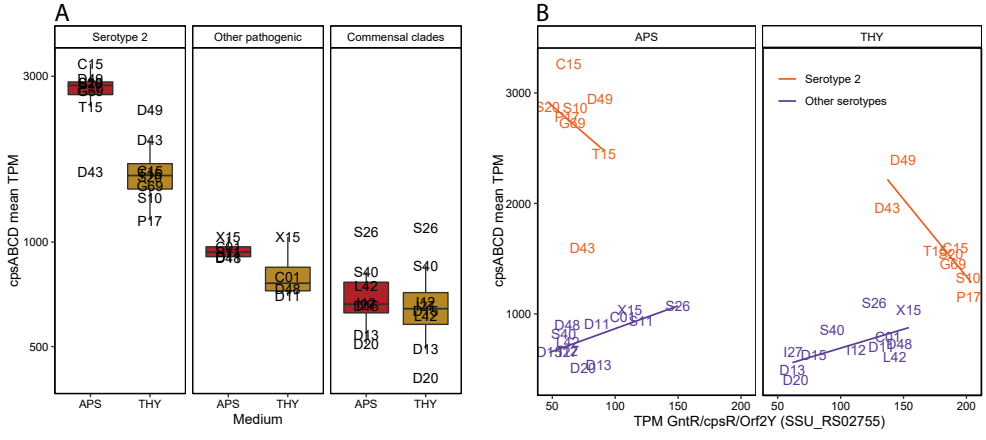


Figure 2. Capsule gene cluster expression. **A)** Mean expression of *cpsABCD* (SSU_RS02765-SSU_RS02780), the initial 4 genes in the cps gene cluster which are shared by all strains. Grouped by serotype 2 (all from pathogenic clades), other pathogenic clade strains, and commensal clade strains. **B)** Scatterplot with regression lines comparing the mean expression of *cpsABCD* with the *cpsR* (SSU_RS02755) regulator. Only the serotype 2 strains had a negative correlation between *cpsR* and *cps* expression, although serotype 2 strains D43 and D49 also appeared to be regulated differently as *cpsABCD* expression was similar in the two media.

High expression of the agmatine deiminase system in serum in commensal strains

As one may expect, many of the genes and pathways most differentially expressed between THY and APS related to acquisition and metabolism of nutrients that are differentially present in the media (Figure 3A). Metal ion acquisition was upregulated in APS while trehalose import and metabolism were highly expressed in some strains in THY. All strains had high expression of the *lac* operon (downstream from SSU_RS04590) in THY, but this was strongly downregulated in APS, except in outgroup strain D13. This strain expresses the *lac* operon constitutively, possibly due to truncations and inversions in the genes encoding repressor *lacR*. Lactose metabolism is likely to be important for *S. suis* during commensal colonisation of pre-weaning piglets, and the fitness cost of constitutive *lac* operon expression appears to be limited as the strain persisted post-weaning (D13 was collected from a piglet approximately 42 days old).

One of the largest overall differences between growth in THY and APS was in amino acid metabolism. The genomes of 6 of 8 commensal strains, but no pathogens, encode an agmatine deiminase system (AgDS, SSUD12_RS06980-SSUD12_RS07005). In the strains in which it was present, AgDS was highly expressed and upregulated in APS compared to in THY. This gene cluster has not previously been described in *S. suis* but is known from *S. pneumoniae* [258] and *S. mutans* [259–261]. The *S. suis* AgDS share high similarity to the AgDS of *S. pneumoniae* reference strain TIGR4 (SP_RS04525-SP_RS04560), *S. porci*, and *S. parasuis* (Figure 3B). The *S. mutans* AgDS is diverged, with several indels and less than 50% amino acid residue identity to the *S. suis* agmatine deiminase gene. In *S. suis* the AgDS appear to be linked to a set of unrelated genes including DNA primase (Figure 3B), and in some strains these make up a 11.5 kb genomic island flanked by transposases.

The arginine deiminase system (ADS/arcABC operon, SSU_RS03045-SSU_RS03060) was found in all strains and in contrast to AgDS it was more expressed in THY than in APS. Agmatine is formed upon decarboxylation of arginine, and agmatine and arginine metabolic pathways show considerable overlap. Both the arginine and agmatine deiminase system produce ammonia, increasing intracellular pH and tolerance to low pH. High ADS and AgDS expression may be linked to acidification of the medium during growth. ADS has been shown to be important for *S. suis* survival in acidic medium [246, 262]. AgDS is thought to be relevant to acidic stress in *S. mutans* [261], but does not appear to increase pH tolerance in *S. pneumoniae* [258]. It is possible that AgDS activity provides a competitive advantage to commensal *S. suis* by increasing intracellular pH. Agmatine released by competing taxa inhibits the growth of *S. mutans* [259], and a similar mechanism may apply to *S. suis*. Considering its variable presence, it may be relevant to both inter- and intra-species competition.

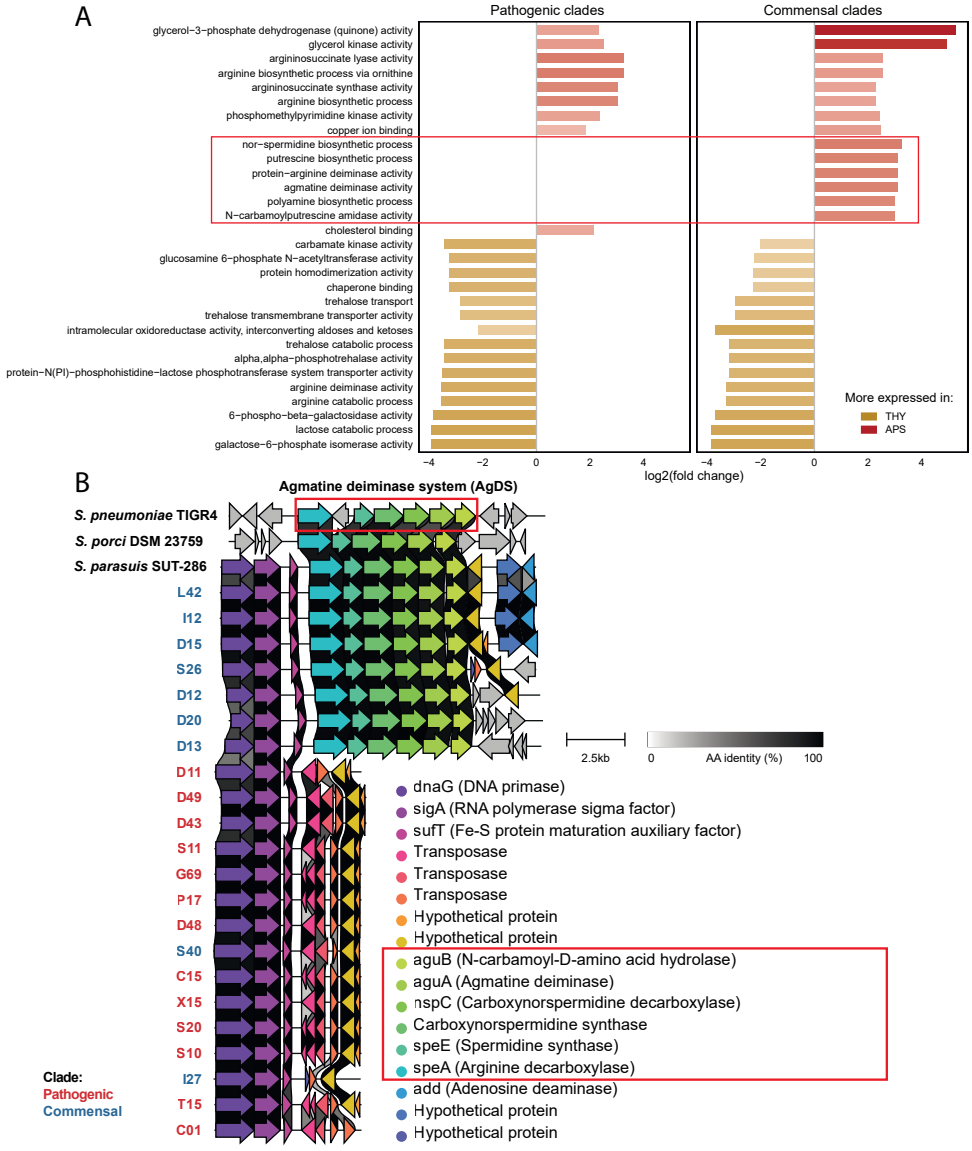


Figure 3. The accessory AgDS system has a large influence on transcriptome differences between THY and APS in commensals. **A)** The 30 GO terms with the largest log2fold changes between THY and APS cultures (average across all strains, GO terms with <10 mean TPM excluded). Several GO terms highly expressed in APS occurred only in commensal clades due to presence/absence of the agmatine deiminase system, suggesting that this gene cluster may be important for commensal *S. suis*. GO terms associated with the agmatine deiminase system highlighted in red. **B)** The AgDS gene cluster is found only in commensal clade strains and has high similarity to the AgDS of *S. pneumoniae*, *S. porci*, and *S. parasuis*.

High expression of Mn^{2+} binding lipoprotein *troA* in pathogens

Transition metal ion homeostasis is vital to bacteria. While commensals sequester metal ions in competition with other microbes, pathogens must overcome host metal ion chelation during infection. Lactic acid bacteria, including *S. suis* and S. *pneumoniae*, can utilize manganese in place of iron, providing them with a competitive advantage [263–266]. We found that genes putatively involved in both iron (SSU_RS03155-SSU_RS03170) and manganese (SSU_RS09395-SSU_RS09415) scavenging were more expressed in APS than THY, although the manganese import gene cluster was 5 times more expressed than the iron import gene cluster. *TroA*, a putative scavenger protein for the *troBCD* ABC transport system [266–268], was also 3-4 times more expressed than the *troBCD* ABC transporter genes in the gene cluster (Figure 4A). This indicates that expression of this gene is regulated separately.

TroA expression varied greatly between strains (232-4165 TPM in THY and 1899-12228 TPM in APS). Some commensals had reduced *troA* expression compared to ST1 strains despite having similar or higher *troBCD* expression (Figure 4A). This may be due to sequence variation of previously described regulators and the gene cluster itself. *S. suis troABCD* expression has been reported to be repressed by *dtxR* family metalloregulator *troR* (SSU_RS09420) and fur family regulator *perR* (SSU_RS01575) depending on metal ion concentrations and oxidative stress [231, 268–270]. Additionally, *mntE* (SSU_RS05010) has been identified as a manganese efflux system [271]. Both *troR* and *perR* may contribute to *troABCD* upregulation in APS compared to THY. Variation in *troA* expression levels between strains grown in the same medium are unlikely to be caused by *troR*, because strains with identical variants varied greatly in expression. *PerR* and *troA* itself also had sequence variation, and the strains with the lowest *troA* expression had a truncated gene cluster structure (Figure 4B, Figure S6). A region between *troA* and *troB* contains a putative oligopeptide transporter pseudogene in most strains, but this was deleted in the strains with the lowest *troA* expression, D13, D20, and I27 (Figure 4B). The initial part of this region has higher expression than *troA*, and its conservation across most of the included *S. suis* clades suggests that it may have a function irrespective of (the length of) the longest predicted open reading frame.

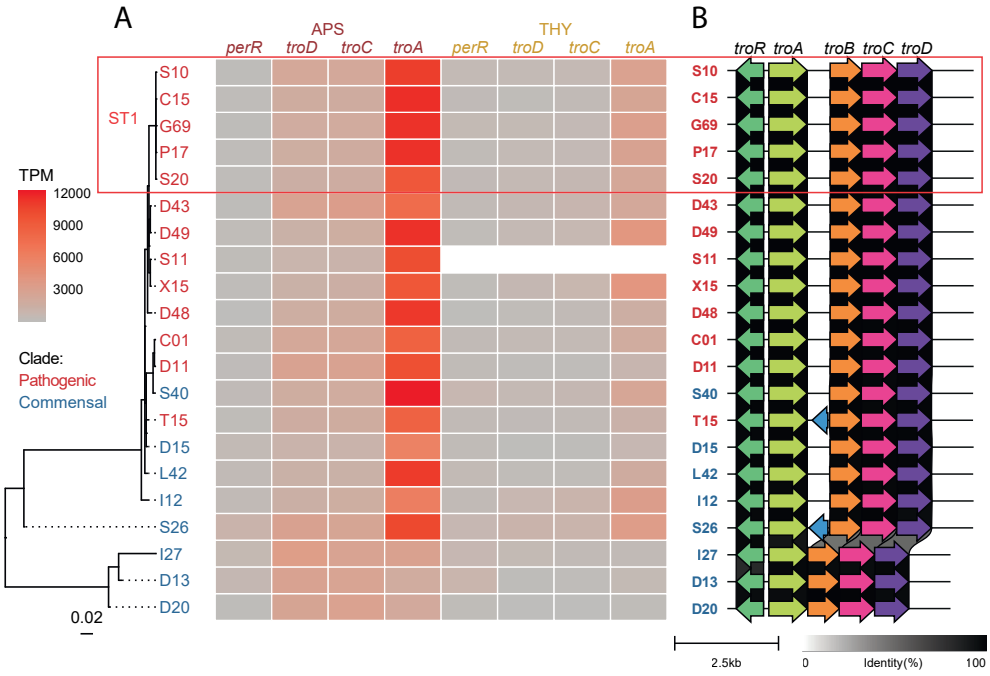


Figure 4. Variation in sequence and expression level of manganese import gene cluster *troABCD* and its regulators. **A)** Regulator *perR* maximum-likelihood tree with heatmap of gene expression TPM values. Mn^{2+} binding lipoprotein *troA* was more expressed than *troBCD* and more expressed in APS compared to THY. Some commensal strains with diverged *perR* and *troRABCD* copies had reduced *troA* expression. **B)** In addition to sequence variation to the other strains, strains I27, D13, and D20 shared a truncated *troABCD* gene cluster which lacked SSU_RS09410, a putative membrane protein variably annotated as a pseudogene (shown in blue) or unannotated by Prokka in different strains, likely due to its short length and truncation.

Conclusions

We found that *S. suis* strains with large genomic divergence have unexpectedly conserved transcriptomes when grown in APS. More variation was observed in THY, and this should be considered when selecting growth medium for *in vitro* assays. Despite overall conservation, the transcriptome of the strains varied in key functions, including well described regulatory mechanisms. In most clades the manganese import and capsule gene clusters may be regulated differently than described for ST1 strains. In general, ST1 strains displayed larger changes in gene expression between THY and APS cultures compared to other clades, and these differences in gene expression may help them rapidly adapt to environmental changes, most notably when changing between upper respiratory tract colonization and host invasion. The gene cluster encoding production of the capsule, which is key in avoiding the host complement attack complex, was among the genes most strongly upregulated and highest expressed in ST1 strains.

Data Summary

The genome assemblies generated in this study are available under BioProject PRJNA855487. The BioSample accession number of each strain is listed in Table S1. RNA-sequencing data is available under BioProject accession number PRJNA863843.

Acknowledgements:

We thank Maria Laura Ferrando and Isabela Fernandes de Oliveira for providing strains and Illumina sequencing data.

Supplementary material

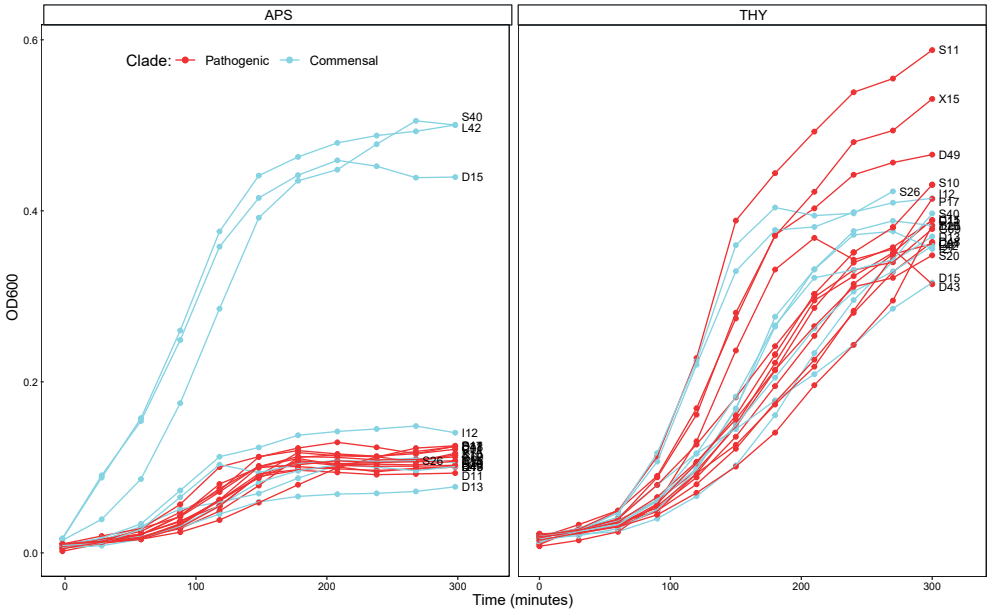


Figure S1: Growth curves. The strains were grown in 96-well plates with 200 µL total volume per well and incubated at 37°C with 5% CO2. Each point represents the mean value of the strain sampled in minimum 3 separate plates with 2 technical replicates per plate.

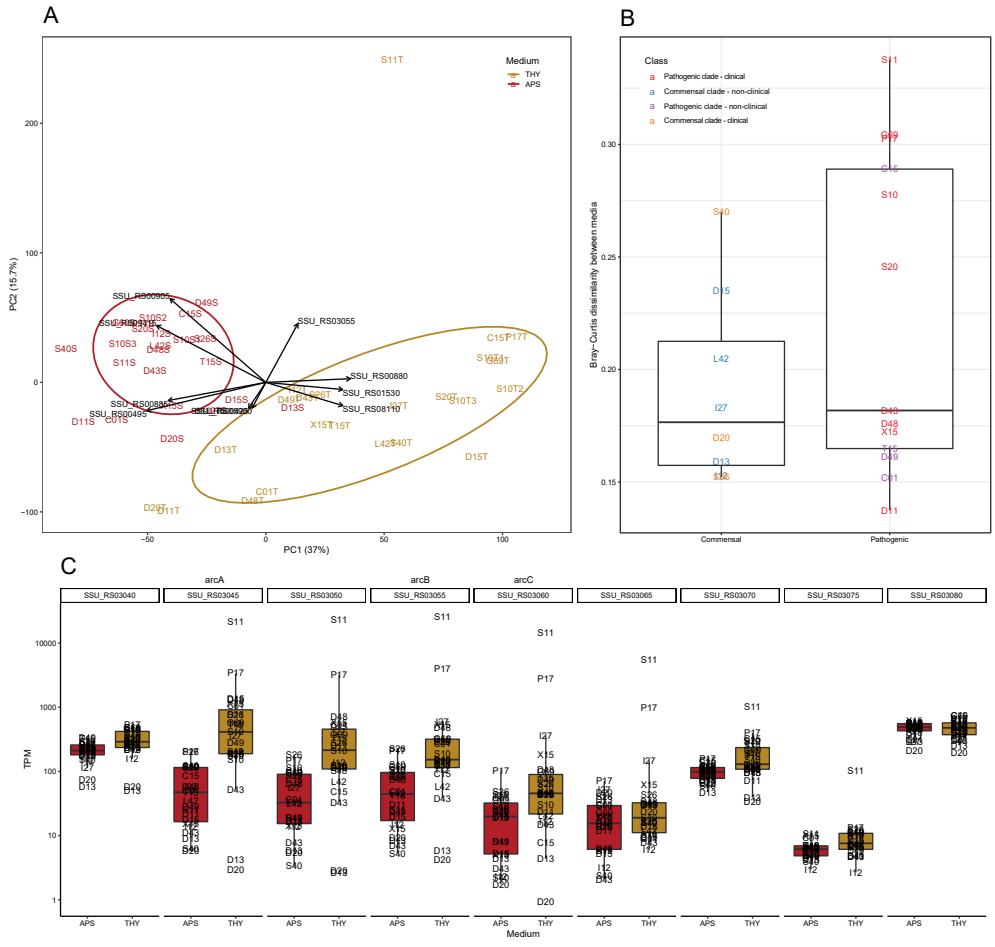


Figure S2: High ADS expression in S11 grown in THY. A) PCA on the core genome transcriptome of all strains shows sample S11T (strain S11 grown in THY) as an outlier. Each label represents a single sample, eclipses represent 75% confidence level, arrows show genes driving the separation of samples; samples in the direction the arrow is pointing have higher expression of the gene. B) Strain S11 also had the largest transcriptome difference between THY and APS cultures of the same strain (core genome transcriptome Bray-Curtis dissimilarity). C) Boxplot of expression shown as log10(TPM) of arcABC and adjacent genes.

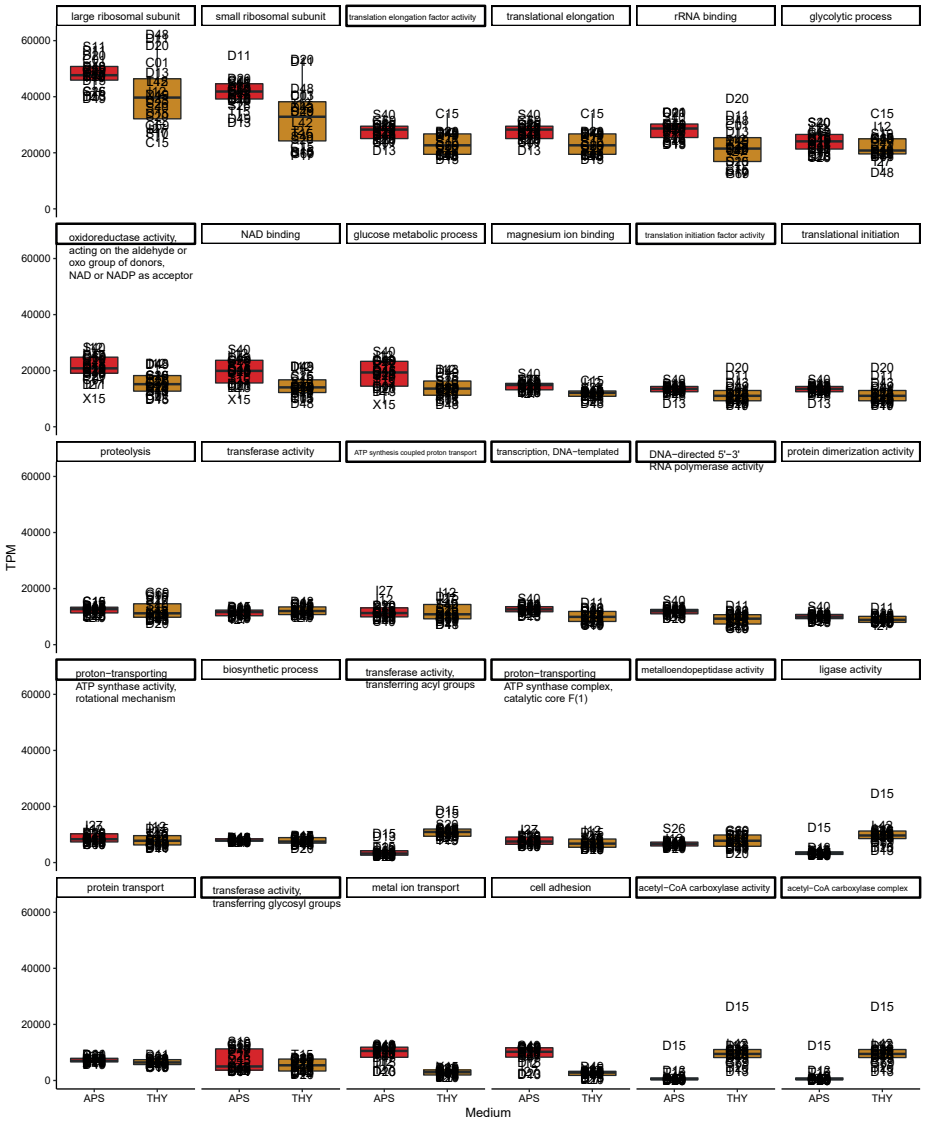


Figure S3: Boxplots showing the expression of the overall most expressed GO terms in APS vs THY.

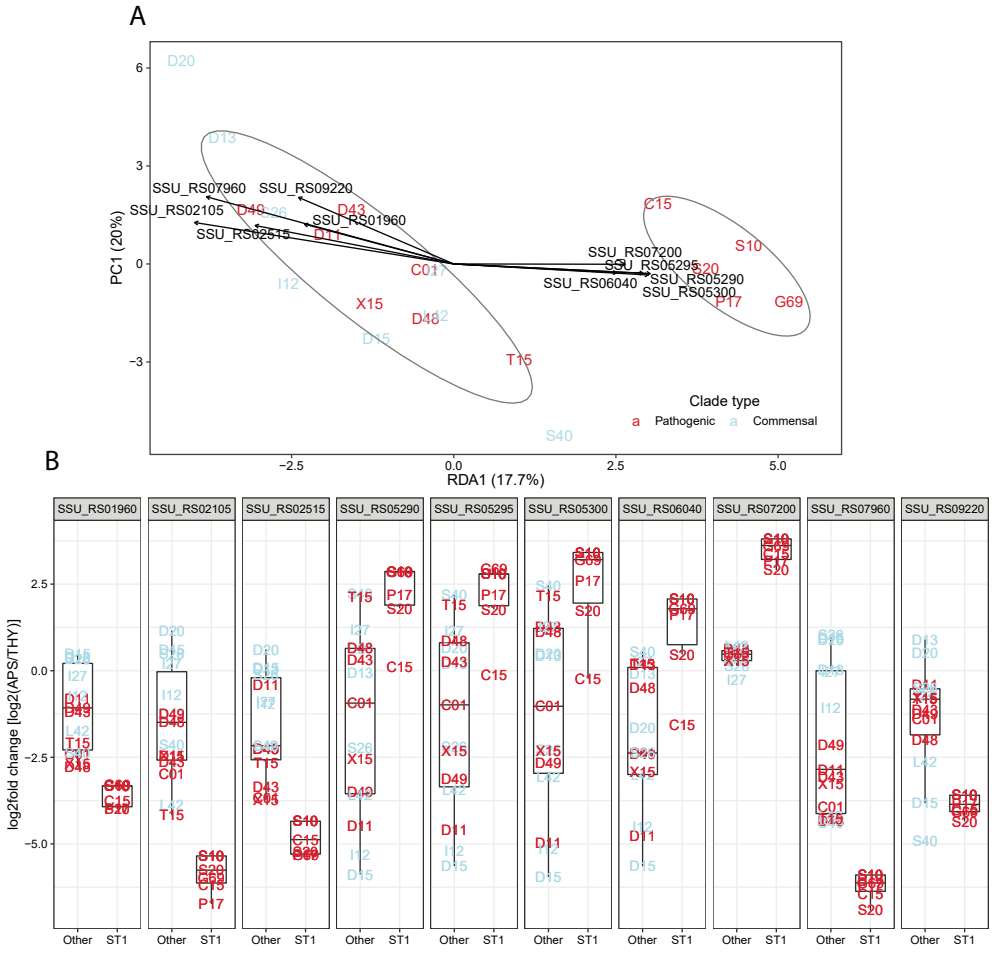


Figure S4: Genes strongly regulated between media in sequence type 1 strains. A) RDA on log2fold gene expression change constrained by ST1 clade strains vs all other strains. Each label represents a single sample, eclipses represent 75% confidence level, arrows show genes driving the separation of samples; samples in the direction the arrow is pointing have higher expression of the gene. B) Boxplot of the log2fold change of the 10 strongest drivers (genes) of the RDA1 axis.

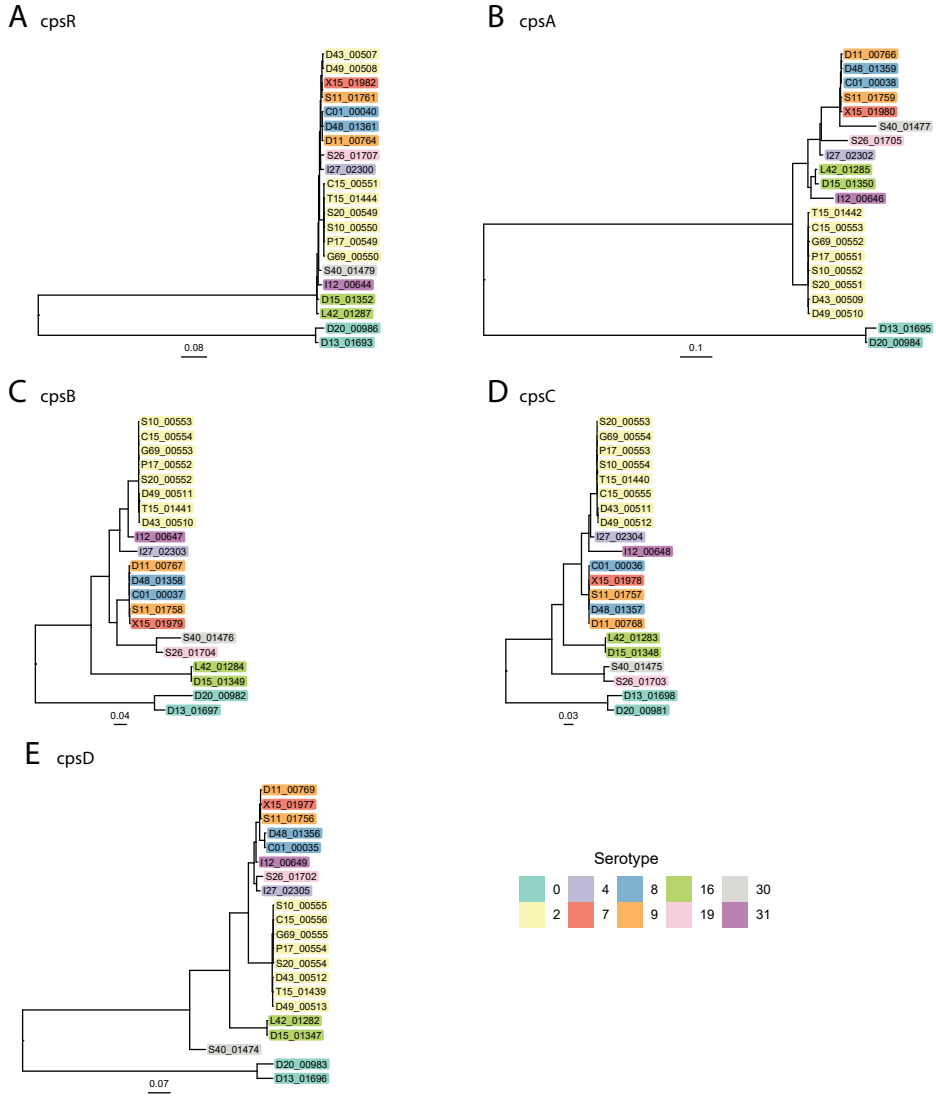


Figure S5: Maximum likelihood trees of cps genes. A) SSU_RS02755, *CpsR/GntR*/Orf2Y B) SSU_RS02765, *cpsA* C) SSU_RS02770, *cpsB* D) SSU_RS02775, *cpsC* E) SSU_RS02780, *cpsD*. Tip labels: [Strain]_[gene number]. Except from outliers D13 and D20, *cpsR* is more conserved than *cpsABCD*. Diverged *cpsR* copies in strains D43 and D49 may cause different *cps* regulation compared to other serotype 2 strains

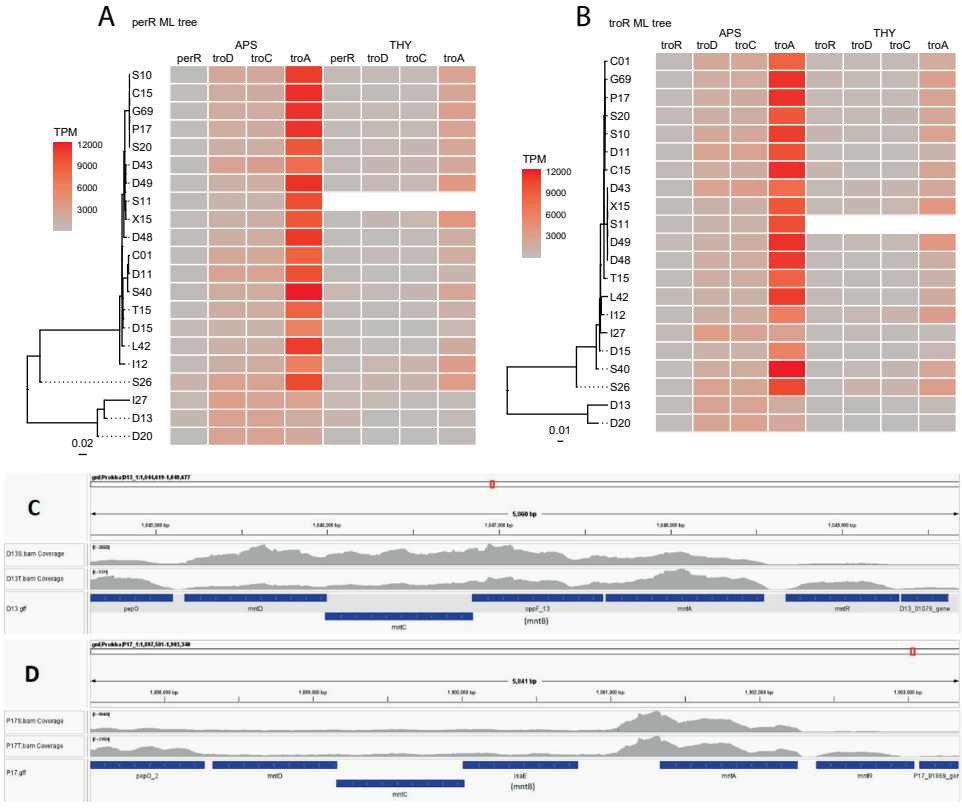


Figure S6: Gene expression of the *tro/mnt* gene cluster. A) ML tree of *perR* with tro expression heatmap. B) ML tree of *troR* with *tro* expression heatmap. C) Raw RNA-seq coverage of the *tro* gene cluster in strain D13 (visualized in IGV). D) Raw RNA-seq coverage of the *tro* gene cluster in strain P1/7.

Table S1: Genome statistics and accession numbers. ST = sequence type, NT = non-typeable.

Name	Strain	Country	Serotype	ST	Clade type	Isolation source	Total length	Largest contig	Contigs	Chromosomal contigs	Plasmids	GC (%)	Biosample accession
C01	M101999_C1	ES	8	NA	Pathogenic	Non-Clinical	2281630	1928593	4	3	1	40.95	SAMN30065789
C15	SS15055_N2_C15	ES	2	1	Pathogenic	Non-Clinical	2007031	2007031	1	1	0	41.3	SAMN30065792
D11	DNS11	DK	9	16	Pathogenic	Systemic/Brain	2169531	2165246	2	1	1	41.22	SAMN30065793
D13	DNC13	DK	NT	NA	Commensal	Non-Clinical	2297247	2264071	3	1	2	43.3	SAMN30065794
D15	DNC15	DK	16	NA	Commensal	Non-Clinical	2209947	2209947	1	1	0	41.28	SAMN30065795
D20	DNS20	DK	NT	NA	Commensal	Systemic/Brain	2188565	2156661	3	1	2	43.68	SAMN30065796
D43	DNR43	DK	2	28	Pathogenic	Respiratory	2231285	2225263	2	1	1	41.16	SAMN30065797
D48	DNR48	DK	8	NA	Pathogenic	Respiratory	2312146	2306687	2	1	1	40.84	SAMN30065798
D49	DNC49	DK	2	28	Pathogenic	Non-Clinical	2137486	2137486	1	1	0	41.26	SAMN30065799
G69	DE609B	DE	2	1	Pathogenic	Systemic/Brain	2055648	2055648	1	1	0	41.27	SAMN29093728
I12	21435_1	ES	31	NA	Commensal	Non-Clinical	2198399	2198399	1	1	0	41.47	SAMN29938211
I27	21437_3	ES	4	NA	Commensal	Non-Clinical	2479314	2163365	9	9	0	41.47	SAMN29938213
L42	LSS42	UK	16	NA	Pathogenic	Non-Clinical	2169516	2169516	1	1	0	41.3	SAMN14932654
P17	P1/7	UK	2	1	Pathogenic	Systemic/Brain	2007491	2007491	1	1	0	41.3	SAMEA3138299
S10	S10	NL	2	1	Pathogenic	Systemic/Brain	2042889	2042889	1	1	0	41.31	SAMEA6227224
S11	M102942_S11	ES	9	123	Pathogenic	Systemic/Brain	2127731	2121233	2	1	1	41.26	SAMN30065800
S20	M104300_S20	ES	2	1	Pathogenic	Systemic/Brain	2074370	2074370	1	1	0	41.22	SAMN30065801
S26	M105052_S26	ES	19	NA	Commensal	Systemic/Brain	2411696	2411696	1	1	0	41.63	SAMN30065802
S40	M106471_S40 1	ES	30	NA	Commensal	Systemic/Brain	2226093	2226093	1	1	0	41.14	SAMN30065803
T15	T15	NL	2	19	Pathogenic	Non-Clinical	2252116	2240209	3	1	2	40.96	SAMN02603633
X15	1521251	CA	7	NA	Pathogenic	Systemic/Brain	2310771	2310771	1	1	0	41.02	SAMN14932458

Chapter 6

General discussion

Biological interpretation

Infectious disease, antibiotics, and the microbiome

The interplay between pathogenic bacteria, disease, antibiotics, and the commensal microbiome is key to the emergence of antibiotic resistance and thus the health of humans and livestock in coming decades. It is of great interest to find alternative solutions to antibiotics and to understand how antibiotic treatment drives emergence of resistance in different contexts. The positive effect of antibiotic growth promoters on piglet growth and mortality rate highlights the potential to improve farm productivity and animal health through the development of strategies to combat opportunist pathogens. In this thesis, I utilized next-generation sequencing and bioinformatic analysis to study the opportunistic pathogen *Streptococcus suis*, the tonsil microbiota associated with *S. suis* disease in piglets, and expansion of antibiotic resistance genes (ARGs) in disease-associated microbiomes.

In **chapter 2** I conducted a case-control study and reported higher abundance of ARGs conferring resistance against antibiotics used to treat *S. suis* disease in *S. suis* diseased piglets compared to asymptomatic controls. This finding is of interest because it shows a link between strain level disease-association, antibiotic-induced selective pressure, and the emergence of antibiotic resistance. Moreover, this (predicted) functional trait underpinned our finding of a disease-associated taxonomic microbiota profile. It also implies that over time, the studied farms had applied antibiotics with relatively high specificity to diseased cohorts. Since antibiotics are usually provided in water or feed to whole cohorts independent of individual disease symptoms, asymptomatic case-associated piglets are also treated, and resistome expansion in the case-associated microbiome may indicate that disease-associated strains have increased prevalence or abundance also in siblings, pen- and roommates of case animals. In **chapter 3** we confirmed that, within a single farm, piglets within litters and rooms share highly similar microbiota composition. Other recent studies have also reported sow or litter effects on the piglet gut microbiome [157, 272], and differences in fecal resistome between farms with different farm management practices and antibiotic usage [119, 120].

In **chapter 4** I further investigated disease-associated resistome expansion by meta-analysis of human gut microbiome studies. I aimed to assess mechanisms of resistome emergence and the impact of these on observational resistome studies. The results indicated that antibiotic treatment of disease induces selective pressure driving acquisition and maintenance of ARGs in disease-associated strains strong enough to be detected in metagenomic studies. These findings were consistent with results from **chapter 2** which suggested that antibiotic treatment targeting *S. suis* disease drives expansion of the resistome in the disease-associated microbiome. Surprisingly, our analysis of human gut microbiome datasets showed up to 300% increase in ARG abundance in cases, far more than the overall 15% increase of ARGs in the tonsillar microbiome of *S. suis* disease piglets.

This was counterintuitive because antibiotic use in piglets is higher than for human diseases. However, antibiotic use in humans is based on more precise diagnosis and only given to individuals exhibiting symptoms, whereas piglets are commonly treated collectively within farms or rooms via feed or water. Moreover, our case-control effect size in microbiome taxonomic composition was small, so the resistome cannot be expected to differ greatly, even if selection pressure is strong. The oral microbiome may also be less susceptible to antibiotics than the gut microbiome due to biofilm formation [273–275].

The ecology of commensal and pathogenic *Streptococcus suis*

Streptococcus species colonizing the human oral cavity are known to be highly niche-specific, such that a strain abundant on the teeth or tongue may be absent on the tonsils [276]. There is consensus that the main habitat of *S. suis* is in the porcine respiratory tract in general and the palatine tonsils in particular [66–73]. In **chapter 2 and 3** we investigated the relative abundance of *S. suis* and other bacteria and found that while commensal *S. suis* is one of the most abundant members of the tonsil microbiota of piglets, pathogenic strains are rare. Previous studies have investigated genomic differences between pathogenic and putatively commensal strains and discussed these in the context of possible niches [63, 64]. Building on this concept, **chapter 5** shows how gene expression differs between commensals and pathogens in pathways that are key to commensal colonization and host invasion. The large genomic divergence and difference in regulation of gene expression within *S. suis* suggests that all clades may not share the same niche. Different clades may have adaptions either to separate niches within the upper respiratory tract or in secondary niches key for transmission within sows or the environment.

S. suis and other streptococci share traits beneficial for commensal niches, but also adaptions facilitating parasitism and host invasion, including a capsule that is key to both niche colonization and host invasion [58–60, 255]. Capsule mutants have demonstrated that even disease-associated serotypes are beneficial for colonization [277], but their distinct structure and increased expression *in vivo* (**chapter 5**) is likely an adaption to host invasion and immune evasion [59, 278]. Production of suilysin (and homologous cytolysins in other streptococci) is also a trait difficult to reconcile with a commensal niche. The gene encoding this pore-forming toxin is highly prevalent and conserved in clinical isolates and considered a key virulence factor in *S. suis* [279, 280]. The capsule is known to shield pathogenic bacteria from antibody binding, complement, and host antimicrobial peptides and polypeptides, as well as limiting desiccation outside the host [103, 281, 282]. Inflammation at mucosal surfaces attracts neutrophils and macrophages to kill invading microbes leading to activation and production of reactive oxidative species. In this context, pathogenic strains may be less sensitive to killing than commensal strains due to the presence of virulence factors that avoid host defenses [283]. (Un)specific host immune responses elicited by pathogenic strains may be to the detriment of solely commensal

strains. This may drive the evolution of commensal strains that antagonize pathogens with bacteriocins [66].

Swab samples give an overview of the bacteria colonizing the tonsil surface, and perhaps the upper part of the tonsillar crypts. However, it does not capture micron-level spatial variation that may play an important role in host-microbe and microbe-microbe interactions. Research on human dental plaque has shown that consortia of different bacterial species form spatial structures where colonizers likely depend on each other for substrate and a favourable extracellular environment [284]. Different *Streptococcus* species in human dental plaque appear to occupy separate niches [285], although in general streptococci seem to occupy distal niches in oral biofilms. Thus, they are exposed to relatively high oxygen content and nutrients in feed and saliva. *Streptococcus* species may be important within bacterial consortia by metabolising host feed and by reducing pH of the extracellular environment, inhibiting competitors and facilitating the growth of collaborative species [284]. *S. suis* is oxygen-tolerant and grows efficiently on lactose and starch derivatives (**chapter 3**), and in **chapter 2** we recovered high proportions of DNA from feed plants on seemingly "clean" tonsillar swabs suggesting that these nutrient sources may be continuously available. *S. suis* is less likely to be competitive in the tonsillar crypts and proximal biofilm layer in direct contact with host epithelial cells. *S. suis* occupying a distal biofilm niche would not have direct access to the host epithelium, and suilysin concentrations at the epithelial layer would be unlikely to reach high enough concentrations to cause cell death. Considering the large genomic diversity of *S. suis*, different clades may have separate niches with different selective pressures and correspondingly different metabolism, mode of transmission, and interaction with the host and microbiota. It is possible that *S. suis* mainly has access to the host epithelium and host invasion during dysbiosis.

It is unclear how pathogenicity is adaptive for *S. suis*. During oral cavity colonization and respiratory disease *S. suis* may spread via direct contact, saliva, and aerosols, but systemic invasion leaves limited opportunity for transmission. Although virulent strains may replicate rapidly to high biomass within systemic sites, they are stuck in an evolutionary one-way street. While they may shed back to the respiratory or intestinal tract, induction of disease in the host is also detrimental given the increased probability of antibiotic treatment. Moreover, the host may be euthanized or die and be removed from the farm environment. With this in mind it is difficult to explain how the most pathogenic clade appears to have evolved in recent decades alongside intensified farming practices [64]. It is possible that invasive disease is not adaptive but a side effect of mechanisms beneficial for survival in other niches. For instance, the mechanisms facilitating host invasion may aid in a parasitic "scavenging" niche of acquiring nutrients from damaged mucosa, and the capsule may aid in protecting from the subsequent inflammation and specific immune responses.

Spread and transmission of *Streptococcus suis*

To remain viable, both commensal and pathogenic *S. suis* need to transmit between piglets, batches, and farms. The exact routes by which this occur are not known. Transmission within farms likely occurs via sows, as piglets of different ages are usually kept separate to prevent transmission of bacteria and viruses. The results from **chapter 3** strongly suggest maternal transfer of the piglet microbiota because of its rapid establishment after birth and because siblings share similar microbiota composition. Sow colonization may also be key for between-farm and global dissemination of all microbes colonizing pigs and pig farms. Modern industrial pig farming is dependent on continuously improved breeds developed by a few global companies, and these are disseminated via a pyramid structured system of nucleus and multiplier farms that deliver gilts to production farms. Breeding farms employ biosecurity measures that may limit transmission and thus the evolutionary success of species ecologically successful in intensively farmed piglets on production farms. The overall high microbiome similarity between farms and recent or ongoing horizontal gene transfer of ARGs does, however, suggest that strains still manage to spread between farms and across the "breeding pyramid". Notably, the most pathogenic *S. suis* lineage has managed to spread globally [64], indicating high ability for transmission between farms and/or continuous re-introduction to breeding farms. Considering the zoonotic ability of *S. suis* it is possible that this occurs via (asymptomatic) colonization of veterinarians or their equipment.

The rate of asymptomatic carriage of pathogenic *S. suis* in piglets, sows, and humans has great implications for *S. suis* transmission within and between outbreaks, and for the development of the design of preventative measures. Future studies should assess the prevalence of pathogenic strains in piglets of different ages, sows, and the farm environment. It should be ascertained whether the putatively pathogenic strains we detect in tonsillar samples are virulent, disease-causing strains that can be isolated from necropsy of disease piglets on the farm, or if they are mostly commensals carrying some virulence-associated genes. This is a challenging research topic because *S. suis* may colonize a range of different body sites and occur in low abundance, making it difficult to detect by culturing, sequencing, and qPCR. Future work should aim to assess in detail whether the tonsils are the main habitat of pathogenic and commensal *S. suis*, or if they colonize other sites in equal or higher abundance. Discovery of sites with high abundance of pathogenic *S. suis* might focus future research on the initial stage of invasive disease. Local variation in the abundance of pathogenic *S. suis* within sites and its direct interaction with the microbiome may also be investigated by microscopic spatial mapping of the tonsillar biofilm and crypts.

The tonsillar microbiota of piglets

The (palatine) tonsils of piglets and its associated microbiota have been a focus of research in veterinary medicine due to being considered the main habitat for *S. suis* and other opportunistic pathogens [69, 72, 73, 90, 286]. The human tonsils and their associated microbiome are also considered important to host health and disease [287]. In this thesis, one key question was whether *S. suis* infectious disease is predisposed by increased abundance of *S. suis* on the tonsils, in particular following stress and change in diet, environment, and/or microbiota dysbiosis at weaning. We aimed to identify health-associated bacteria that may be valuable as candidate probiotics that may antagonise *S. suis*, prevent dysbiosis, or restore the microbiota following disruption. Our results showed that exponential growth (or a "bloom") of *S. suis* can be excluded as a factor leading to invasive systemic disease. Abundance of *S. suis* in the tonsillar microbiota decreases at weaning (**chapter 3**), and *S. suis* diseased piglets are colonized by lower relative abundance of *S. suis* than asymptomatic controls (**chapter 2**). However, it remains possible that the oral microbiota and its interactions with the immune system is critical to disease development. In **Chapter 2** we show that putative microbiota dysbiosis occurs pre-weaning, weeks prior to symptom development. Dysbiosis may occur due to a range of factors including both host-microbe and microbe-microbe interactions, and may also be related to environmental factors, as is the case with overcrowding and *Campylobacter* abundance in chickens [288].

Previous necropsy- and culture-based work has identified co-infections between *S. suis*, viruses, and other bacteria. This complex of disease-associated microbes (mainly associated with grower/finisher stage respiratory disease) is collectively termed porcine respiratory disease complex (PRDC). Co-infections are, however, so common at all ages that *S. suis* is considered a secondary pathogen dependent on primary pathogens increasing host susceptibility to invasion. In particular, porcine reproductive and respiratory syndrome virus (PRRSV) is known to facilitate *S. suis* disease [289]. In **chapter 2** we identified several novel and uncultured disease-associated bacterial species that colonized the tonsils of *S. suis* diseased piglets at higher abundance than asymptomatic control piglets. This finding could be explained in different ways. Firstly, the abundance of the disease-associated species may be causative to *S. suis* disease development, but without themselves invading the host and therefore not being isolated during necropsy. Secondly, they may be indirectly associated with *S. suis* disease by correlation, for instance by expanding in niches that become available following depletion of health-associated species by host inflammation. Thirdly, they may be undescribed because they are unculturable by current microbiological methods (or more specifically the methods routinely used in veterinary diagnostic labs). Future *in vivo* experimental studies on *S. suis* invasive disease may include inoculation of disease-associated taxa to test whether they impact on disease development.

Methodological considerations and approaches

Study design

Study design and participant selection for microbiome studies is more convenient in pigs than humans. Piglets within farms share a similar genetic background, lifelong identical feed intake and antibiotic treatment, and are co-housed in rooms and pens that make up ideal blocked experimental units. One caveat is that the microbiota is strongly associated with the individual litter and room (**chapter 3**). Studies on infectious disease, immunology, and the microbiome need to avoid confoundment between strong maternal effects and small effects of disease and experimental treatment. Experimental studies should consider that samples collected from piglets within the same litter are not independent observations. In case-control studies it is key to be conscious on whether to select control piglets from within litters, pens, or rooms, in particular when using small sample sizes. In **chapter 2** asymptomatic siblings of case piglets had microbiome composition intermediate between their symptomatic siblings and controls from unaffected litters. Disease-associated microbiome traits may be specific to the symptomatic piglet or be shared within the pen (i.e., pen-mates are not true controls but case-associated piglets). However, comparison of piglets from different pens or rooms risks confoundment, in particular because several piglets within single pens often develop symptoms simultaneously. Although bacteria specific to case piglets may be of main interest to understand the mechanisms of host invasion, the room-, sow-, and litter-specific microbiome and environmental factors may be key to development of disease prevention strategies. Disease-associated microbiota traits may also be shared by the whole farm, warranting comparison between farms with and without the disease.

Technical challenges

Taxonomic assessment of microbiomes is dependent on reference data, but there is a lack of cultured, described, and sequenced representatives of the oral microbiome of pigs. This problem is exacerbated by a lack of data from the oral microbiome of humans and other species. The lack of reference data is not a major problem to amplicon sequencing approaches, because each amplicon sequence variant is a distinguishable unit for statistical analysis and itself taxonomically informative. For (short-read) shotgun sequencing approaches, the lack of representative genomes is a bigger problem because reads originate from a random location in the genome and provide no information unless they can be mapped to a known gene or genome. Since most taxa in the porcine oral microbiota lack representative genomes, high proportions of the (shotgun-sequenced) microbiome remain taxonomically unclassified, even at the phylum and domain level. There is also limited data available on the function of individual genes of oral colonizers. This means that microbiome functionality analysis, including pathways and gene ontology

term analysis, yield poor results. Overall, out of the box tools have limited use for analysis of the piglet tonsillar microbiome.

The need for reference genomes in taxonomic analysis may be partly circumvented by construction of metagenome-assembled genomes (MAGs) and mapping of reads to these. In **chapter 2** I recovered 800 MAGs by co-assembly per farm, and these were key to identifying novel disease-associated species. However, the quality of MAGs is not equal to the genomes of isolates, and they should be used with caution. The oral microbiome has high strain-level diversity and thus presents a major challenge to genome assemblers, resulting in mosaic genomes that poorly represent real individual strains. The number of MAGs recovered per taxon is also not proportional to their prevalence or abundance. I assembled several putatively high-quality MAGs of *Mesomycoplasma hyorhinis*, a species with a small genome and no close relatives in the tonsil microbiome, despite its low abundance. Conversely, *Streptococcus*, *Actinobacillus*, and *Moraxella* are abundant genera represented by many species within each sample and are underrepresented among MAGs. Only 14 *Moraxella* MAGs passed the quality threshold despite this taxon being the second most abundant genus. Three separate species or strains of *Moraxella* that differed by only a single SNP in the 16S rRNA gene V3-V4 region were among the most abundant ASVs in tonsillar microbiota samples. I recovered in total 57 *Streptococcus* MAGs from the 12 farms, but this clearly did not represent the full diversity of streptococci in the piglet tonsillar microbiota. Each piglet could be colonized by up to 16 different *S. suis* ASVs, which may each represent several strains, as well as several other *Streptococcus* species including *S. parasuis*, *S. plurextorum*, *S. hyointestinalis*, *S. porcorum, S. porci*, *S. ferus*, and *S. hyovaginalis*. The high rate of horizontal gene transfer among streptococci may present an additional challenge. It is likely that *Streptococus* MAGs are mosaic despite passing quality thresholds, although it is difficult to quantify this without exhaustively sequencing isolates from each piglet. A wide range of different tools exist for MAG assembly, and the choice of assembler may impact on results [290, 291]. In addition to MetaWRAP [140] we tested STRONG [292], which is specialized in strain level genome resolution, but without improvement to the results. This may in part be due to insufficient sequencing coverage.

Variable proportions of bacterial and host DNA impacts on analysis of oral swabs to a much higher degree than faecal samples [293]. In **chapter 2** this was a major issue because host and plant-derived DNA contamination was not only highly variable between samples, but also differed between case and control piglets. This might have been a result of host inflammation causing more epithelial cells to be included in the swabs, or due to case and control piglets behaving differently during sampling. We considered that confounding sampler effects are unlikely to have any significant impact because case-control taxonomic differences were small and to a large degree driven by shifts in abundance of closely related species or strains. Sampling biases may be more likely to find differences in higher level taxa. For instance, applying more pressure while swabbing may release more anaerobic bacteria from the crypts. Regardless, it is important to address host- and feed-

DNA contamination. I first removed reads mapping to genomes of pig and feed plants, but this standard practice was not sufficient to remove all the contaminating DNA. Not all relevant feed plants have available genomes, and low complexity regions (that are characteristic for plant genomes) are masked, preventing removal. Genetic variation to the reference genomes may also result in reads not mapping. We additionally assessed remaining eukaryote DNA using kraken 2 [139] and used this to normalize the abundance of bacterial taxa and ARGs within each sample. Future works may want to utilize host DNA removal and increased sequencing depth to reduce the impact of contaminating DNA.

Closing remarks

Although this thesis has advanced current knowledge on the tonsil microbiota of piglets and the biology of *S. suis*, many questions remain. While there is consensus that *S. suis* is a multifactorial disease predisposed by biotic and abiotic factors, these are poorly characterised and quantified. Ultimately, only large studies with multifactorial approaches are suited for addressing multifactorial research questions. To date, studies on *S. suis* are too small and narrow in scope to properly assess the nature of *S. suis* disease. In the future it may be advantageous to correlate disease occurrence, abundance of pathogenic *S. suis*, host genetic and immunological measurements, the microbiome, and environmental factors such as air quality, temperature, and sanitation. Knowledge of the tonsillar microbiome may be expanded further by including virome and metatranscriptomics sequencing to assess co-infection with viruses and gene expression of *S. suis* and other microbiome members *in vivo*.

Appendices

References
Summary
Acknowledgements
About the author
Overview of completed training activities

References

1. Bloom DE, Cadarette D. Infectious Disease Threats in the Twenty-First Century: Strengthening the Global Response. Frontiers in Immunology. 2019;10.
2. Cohen ML. Changing patterns of infectious disease. Nature. 2000;406:762–7.
3. O'Neill J. Review on antimicrobial resistance. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. 2014;2014.
4. Lekagul A, Tangcharoensathien V, Yeung S. Patterns of antibiotic use in global pig production: A systematic review. Veterinary and Animal Science. 2019;7:100058.
5. Aarestrup FM, Wegener HC, Collignon P. Resistance in bacteria of the food chain: epidemiology and control strategies. Expert Review of Anti-infective Therapy. 2008;6:733–50.
6. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. The Lancet Infectious Diseases. 2018;18:318–27.
7. Christou L. The global burden of bacterial and viral zoonotic infections. Clinical Microbiology and Infection. 2011;17:326–30.
8. Segura M, Aragon V, Brockmeier SL, Gebhart C, Greeff A de, Kerdsin A, et al. Update on Streptococcus suis Research and Prevention in the Era of Antimicrobial Restriction: 4th International Workshop on S. suis. Pathogens. 2020;9:374.
9. Gottschalk M, Xu J, Calzas C, Segura M. Streptococcus suis: a new emerging or an old neglected zoonotic pathogen? Future Microbiology. 2010;5:371–91.
10. Lun Z-R, Wang Q-P, Chen X-G, Li A-X, Zhu X-Q. Streptococcus suis: an emerging zoonotic pathogen. The Lancet Infectious Diseases. 2007;7:201–9.
11. Neila-Ibáñez C, Casal J, Hennig-Pauka I, Stockhofe-Zurwieden N, Gottschalk M, Migura-García L, et al. Stochastic Assessment of the Economic Impact of Streptococcus suis-Associated Disease in German, Dutch and Spanish Swine Farms. Frontiers in Veterinary Science. 2021;8.
12. Hadjirin NF, Miller EL, Murray GGR, Yen PLK, Phuc HD, Wileman TM, et al. Large-scale genomic analysis of antimicrobial resistance in the zoonotic pathogen Streptococcus suis. BMC Biol. 2021;19:191.
13. Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. Curr Opin Gastroenterol. 2015;31:69–75.
14. Belkaid Y, Hand TW. Role of the Microbiota in Immunity and Inflammation. Cell. 2014;157:121–41.
15. van der Hee B, Wells JM. Microbial Regulation of Host Physiology by Short-chain Fatty Acids. Trends in Microbiology. 2021;29:700–12.
16. Sommer F, Anderson JM, Bharti R, Raes J, Rosenstiel P. The resilience of the intestinal microbiota influences health and disease. Nat Rev Microbiol. 2017;15:630–8.
17. Sorbara MT, Pamer EG. Interbacterial mechanisms of colonization resistance and the strategies pathogens use to overcome them. Mucosal Immunol. 2019;12:1–9.

18. Pamer EG. Resurrecting the intestinal microbiota to combat antibiotic-resistant pathogens. Science. 2016;352:535–8.
19. Jacobson A, Lam L, Rajendram M, Tamburini F, Honeycutt J, Pham T, et al. A Gut Commensal-Produced Metabolite Mediates Colonization Resistance to Salmonella Infection. Cell Host & Microbe. 2018;24:296-307.e7.
20. Shelton CD, Yoo W, Shealy NG, Torres TP, Zieba JK, Calcutt MW, et al. Salmonella enterica serovar Typhimurium uses anaerobic respiration to overcome propionate-mediated colonization resistance. Cell Reports. 2022;38.
21. Fuentes S, van Nood E, Tims S, Heikamp-de Jong I, ter Braak CJ, Keller JJ, et al. Reset of a critically disturbed microbial ecosystem: faecal transplant in recurrent Clostridium difficile infection. ISME J. 2014;8:1621–33.
22. Kim S, Covington A, Pamer EG. The intestinal microbiota: Antibiotics, colonization resistance, and enteric pathogens. Immunological Reviews. 2017;279:90–105.
23. Naaber P, Smidt I, Štšepetova J, Brilene T, Annuk H, Mikelsaar M 2004. Inhibition of Clostridium difficile strains by intestinal Lactobacillus species. Journal of Medical Microbiology. 2004;53: 551–4.
24. Charalampous T, Kay GL, Richardson H, Aydin A, Baldan R, Jeanes C, et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. Nat Biotechnol. 2019;37:783–92.
25. Freire M, Nelson KE, Edlund A. The Oral Host–Microbial Interactome: An Ecological Chronometer of Health? Trends in Microbiology. 2021;29:551–61.
26. Lamont RJ, Koo H, Hajishengallis G. The oral microbiota: dynamic communities and host interactions. Nat Rev Microbiol. 2018;16:745–59.
27. Opriessnig T, Giménez-Lirola LG, Halbur PG. Polymicrobial respiratory disease in pigs. Animal Health Research Reviews. 2011;12:133–48.
28. Laxminarayan R, Duse A, Wattal C, Zaidi AKM, Wertheim HFL, Sumpradit N, et al. Antibiotic resistance—the need for global solutions. The Lancet Infectious Diseases. 2013;13:1057–98.
29. Spellberg B, Bartlett JG, Gilbert DN. The Future of Antibiotics and Resistance. New England Journal of Medicine. 2013;368:299–302.
30. Barnett ML, Linder JA. Antibiotic Prescribing to Adults With Sore Throat in the United States, 1997-2010. JAMA Internal Medicine. 2014;174:138–40.
31. Radyowijati A, Haak H. Improving antibiotic use in low-income countries: an overview of evidence on determinants. Social Science & Medicine. 2003;57:733–44.
32. Van Boeckel TP, Gandra S, Ashok A, Caudron Q, Grenfell BT, Levin SA, et al. Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data. The Lancet Infectious Diseases. 2014;14:742–50.
33. Casewell M, Friis C, Marco E, McMullin P, Phillips I. The European ban on growth-promoting antibiotics and emerging consequences for human and animal health. Journal of Antimicrobial Chemotherapy. 2003;52:159–61.

34. McEwen SA, Collignon PJ. Antimicrobial Resistance: a One Health Perspective. Microbiology Spectrum. 2018;6:6.2.10.
35. Hu Y, Yang X, Li J, Lv N, Liu F, Wu J, et al. The Bacterial Mobile Resistome Transfer Network Connecting the Animal and Human Microbiomes. Applied and Environmental Microbiology. 2016;82:6672–81.
36. Redondo-Salvo S, Fernández-López R, Ruiz R, Vielva L, de Toro M, Rocha EPC, et al. Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. Nature Communications. 2020;11:3602.
37. von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, et al. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. Front Microbiol. 2016;7.
38. MacLean RC, San Millan A. The evolution of antibiotic resistance. Science. 2019;365:1082–3.
39. Brinkac L, Voorhies A, Gomez A, Nelson KE. The Threat of Antimicrobial Resistance on the Human Microbiome. Microb Ecol. 2017;74:1001–8.
40. Sommer MOA, Dantas G, Church GM. Functional Characterization of the Antibiotic Resistance Reservoir in the Human Microflora. Science. 2009;325:1128–31.
41. Brito IL, Gurry T, Zhao S, Huang K, Young SK, Shea TP, et al. Transmission of human-associated microbiota along family and social networks. Nat Microbiol. 2019;4:964–71.
42. Sarkar A, Harty S, Johnson KV-A, Moeller AH, Archie EA, Schell LD, et al. Microbial transmission in animal social networks and the social microbiome. Nat Ecol Evol. 2020;4:1020–35.
43. Ding D, Zhu J, Gao Y, Yang F, Ma Y, Cheng X, et al. Effect of cattle farm exposure on oropharyngeal and gut microbial communities and antibiotic resistance genes in workers. Science of The Total Environment. 2022;806:150685.
44. Kraemer JG, Ramette A, Aebi S, Oppliger A, Hilty M. Influence of Pig Farming on the Human Nasal Microbiota: Key Role of Airborne Microbial Communities. Applied and Environmental Microbiology. 2018;84:e02470-17.
45. Sun J, Liao X-P, D'Souza AW, Boolchandani M, Li S-H, Cheng K, et al. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. Nat Commun. 2020;11:1427.
46. Van Gompel L, Luiken REC, Hansen RB, Munk P, Bouwknegt M, Heres L, et al. Description and determinants of the faecal resistome and microbiome of farmers and slaughterhouse workers: A metagenome-wide cross-sectional study. Environment International. 2020;143:105939.
47. George AN, Stewart JR, Evans JC, Gibson JM. Risk of Antibiotic-Resistant Staphylococcus aureus Dispersion from Hog Farms: A Critical Review. Risk Analysis. 2020;40:1645–65.
48. Heuer H, Smalla K. Plasmids foster diversification and adaptation of bacterial populations in soil. FEMS Microbiology Reviews. 2012;36:1083–104.
49. Marti R, Scott A, Tien Y-C, Murray R, Sabourin L, Zhang Y, et al. Impact of Manure Fertilization on the Abundance of Antibiotic-Resistant Bacteria and Frequency of Detection of Antibiotic Resistance Genes in Soil and on Vegetables at Harvest. Applied and Environmental Microbiology. 2013;79:5701–9.

50. Xie W-Y, Yuan S-T, Xu M-G, Yang X-P, Shen Q-R, Zhang W-W, et al. Long-term effects of manure and chemical fertilizers on soil antibiotic resistome. Soil Biology and Biochemistry. 2018;122:111–9.
51. Ferrando ML, Gussak A, Mentink S, Gutierrez MF, van Baarlen P, Wells JM. Active Human and Porcine Serum Induce Competence for Genetic Transformation in the Emerging Zoonotic Pathogen Streptococcus suis. Pathogens. 2021;10:156.
52. Straume D, Stamsås GA, Håvarstein LS. Natural transformation and genome evolution in Streptococcus pneumoniae. Infection, Genetics and Evolution. 2015;33:371–80.
53. Arnold BJ, Huang I-T, Hanage WP. Horizontal gene transfer and adaptive evolution in bacteria. Nat Rev Microbiol. 2022;20:206–18.
54. Gottschalk M, Segura M. The pathogenesis of the meningitis caused by Streptococcus suis: the unresolved questions. Veterinary Microbiology. 2000;76:259–72.
55. Vötsch D, Willenborg M, Weldearegay YB, Valentin-Weigand P. Streptococcus suis – The "Two Faces" of a Pathobiont in the Porcine Respiratory Tract. Front Microbiol. 2018;9.
56. Lancefield RC. A SEROLOGICAL DIFFERENTIATION OF HUMAN AND OTHER GROUPS OF HEMOLYTIC STREPTOCOCCI. Journal of Experimental Medicine. 1933;57:571–95.
57. Okura M, Osaki M, Nomoto R, Arai S, Osawa R, Sekizaki T, et al. Current Taxonomical Situation of Streptococcus suis. Pathogens. 2016;5:45.
58. Feng Y, Cao M, Shi J, Zhang H, Hu D, Zhu J, et al. Attenuation of Streptococcus suis virulence by the alteration of bacterial surface architecture. Sci Rep. 2012;2:710.
59. Hyams C, Camberlein E, Cohen JM, Bax K, Brown JS. The Streptococcus pneumoniae Capsule Inhibits Complement Activity and Neutrophil Phagocytosis by Multiple Mechanisms. Infection and Immunity. 2010;78:704–15.
60. Stollerman GH, Dale JB. The Importance of the Group A Streptococcus Capsule in the Pathogenesis of Human Infections: A Historical Perspective. Clinical Infectious Diseases. 2008;46:1038–45.
61. Goyette-Desjardins G, Auger J-P, Xu J, Segura M, Gottschalk M. Streptococcus suis, an important pig pathogen and emerging zoonotic agent—an update on the worldwide distribution based on serotyping and sequence typing. Emerging Microbes & Infections. 2014;3:1–20.
62. Prüfer TL, Rohde J, Verspohl J, Rohde M, Greeff A de, Willenborg J, et al. Molecular typing of Streptococcus suis strains isolated from diseased and healthy pigs between 1996-2016. PLOS ONE. 2019;14:e0210801.
63. Murray GGR, Charlesworth J, Miller EL, Casey MJ, Lloyd CT, Gottschalk M, et al. Genome Reduction Is Associated with Bacterial Pathogenicity across Different Scales of Temporal and Ecological Divergence. Molecular Biology and Evolution. 2021;38:1570–9.
64. Weinert LA, Chaudhuri RR, Wang J, Peters SE, Corander J, Jombart T, et al. Genomic signatures of human and animal disease in the zoonotic pathogen Streptococcus suis. Nat Commun. 2015;6:1–10.
65. Wileman TM, Weinert LA, Howell KJ, Wang J, Peters SE, Williamson SM, et al. Pathotyping the Zoonotic Pathogen Streptococcus suis: Novel Genetic Markers To Differentiate Invasive

Disease-Associated Isolates from Non-Disease-Associated Isolates from England and Wales. Journal of Clinical Microbiology. 2019;57:e01712-18.

66. Liang Z, Wu H, Bian C, Chen H, Shen Y, Gao X, et al. The antimicrobial systems of Streptococcus suis promote niche competition in pig tonsils. Virulence. 2022;13:781–93.
67. Arends JP, Hartwig N, Rudolphy M, Zanen HC. Carrier rate of Streptococcus suis capsular type 2 in palatine tonsils of slaughtered pigs. Journal of Clinical Microbiology. 1984;20:945–7.
68. Clifton-Hadley F, Alexander T. The carrier site and carrier rate of Streptococcus suis type II in pigs. Veterinary Record. 1980;107:40–1.
69. Lowe BA, Marsh TL, Isaacs-Cosgrove N, Kirkwood RN, Kiupel M, Mulks MH. Defining the "core microbiome" of the microbial communities in the tonsils of healthy pigs. BMC Microbiol. 2012;12:20.
70. Murase K, Watanabe T, Arai S, Kim H, Tohya M, Ishida-Kuroki K, et al. Characterization of pig saliva as the major natural habitat of Streptococcus suis by analyzing oral, fecal, vaginal, and environmental microbiota. PLOS ONE. 2019;14:e0215983.
71. Kernaghan S, Bujold AR, MacInnes JI. The microbiome of the soft palate of swine. Animal Health Research Reviews. 2012;13:110–20.
72. Pena Cortes LC, LeVeque RM, Funk J, Marsh TL, Mulks MH. Development of the tonsillar microbiome in pigs from newborn through weaning. BMC Microbiology. 2018;18.
73. Pena Cortes LC, LeVeque RM, Funk JA, Marsh TL, Mulks MH. Development of the Tonsil Microbiome in Pigs and Effects of Stress on the Microbiome. Front Vet Sci. 2018;5.
74. Mathieu-Denoncourt A, Letendre C, Auger J-P, Segura M, Aragon V, Lacouture S, et al. Limited Interactions between Streptococcus Suis and Haemophilus Parasuis in In Vitro Co-Infection Studies. Pathogens. 2018;7:7.
75. Neila-Ibáñez C, Napp S, Pailler-García L, Franco-Martínez L, Cerón JJ, Aragon V, et al. Risk factors associated to Streptococcus suis cases in swine farms in Spain. Research Square. 2022.
76. Obradovic MR, Segura M, Segalés J, Gottschalk M. Review of the speculative role of co-infections in Streptococcus suis-associated diseases in pigs. Veterinary Research. 2021;52:49.
77. Thanawongnuwech R, Brown GB, Halbur PG, Roth JA, Royer RL, Thacker BJ. Pathogenesis of Porcine Reproductive and Respiratory Syndrome Virus-induced Increase in Susceptibility to Streptococcus suis Infection. Vet Pathol. 2000;37:143–52.
78. Ferrando ML, Baarlen P van, Orrù G, Piga R, Bongers RS, Wels M, et al. Carbohydrate Availability Regulates Virulence Gene Expression in Streptococcus suis. PLOS ONE. 2014;9:e89334.
79. Emms DM, Kelly S. STAG: Species Tree Inference from All Genes. 2018;:267914.
80. Rieckmann K, Pendzialek S-M, Vahlenkamp T, Baums CG. A critical review speculating on the protective efficacies of autogenous Streptococcus suis bacterins as used in Europe. Porcine Health Management. 2020;6:12.
81. Sangild PT. Uptake of Colostral Immunoglobulins by the Compromised Newborn Farm Animal. Acta Veterinaria Scandinavica. 2003;44:S105.

82. Hopkins D, Poljak Z, Farzan A, Friendship R. Field studies evaluating the direct, indirect, total, and overall efficacy of Streptococcus suis autogenous vaccine in nursery pigs. Can Vet J. 2019;60:386–90.
83. Li Y, Gottschalk M, Esgleas M, Lacouture S, Dubreuil JD, Willson P, et al. Immunization with recombinant Sao protein confers protection against Streptococcus suis infection. Clin Vaccine Immunol. 2007;14:937–43.
84. Wisselink HJ, Vecht U, Smith HE, Stockhofe-Zurwieden N. Protection of pigs against challenge with virulent Streptococcus suis serotype 2 strains by a muramidase-released protein and extracellular factor vaccine. Veterinary Record. 2001;148:473–7.
85. Amass SF, Clark LK, Knox K, Wu CC, Hill MA. Streptococcus suis colonization of piglets during parturition. Journal of Swine Health and Production. 1996;4:269–72.
86. Berthelot-Hérault F, Gottschalk M, Labbé A, Cariolet R, Kobisch M. Experimental airborne transmission of Streptococcus suis capsular type 2 in pigs. Veterinary Microbiology. 2001;82:69–80.
87. Dekker CNT, Bouma A, Daemen AJJM, van Leengoed LAMG, Jonker FH, Wagenaar JA, et al. Homologous whole bacterin vaccination is not able to reduce Streptococcus suis serotype 9 strain 7997 transmission among pigs or colonization. Vaccine. 2012;30:1379–87.
88. Dekker N, Bouma A, Daemen I, Klinkenberg D, Leengoed L van, Wagenaar JA, et al. Effect of Spatial Separation of Pigs on Spread of Streptococcus suis Serotype 9. PLOS ONE. 2013;8:e61339.
89. Madsen LW, Nielsen B, Aalbæk B, Jensen HE, Nielsen JP, Riising HJ. Experimental Infection of Conventional Pigs with Streptococcus suis serotype 2 by Aerosolic Exposure. Acta Veterinaria Scandinavica. 2001;42:303.
90. Zou G, Zhou J, Xiao R, Zhang L, Cheng Y, Jin H, et al. Effects of Environmental and Management-Associated Factors on Prevalence and Diversity of Streptococcus suis in Clinically Healthy Pig Herds in China and the United Kingdom. Applied and Environmental Microbiology. 2018;84:e02590-17.
91. Carr VR, Chaguza C. Metagenomics for surveillance of respiratory pathogens. Nat Rev Microbiol. 2021;19:285–285.
92. Gardy JL, Loman NJ. Towards a genomics-informed, real-time, global pathogen surveillance system. Nat Rev Genet. 2018;19:9–20.
93. Schmidt K, Mwaigwisya S, Crossman LC, Doumith M, Munroe D, Pires C, et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. Journal of Antimicrobial Chemotherapy. 2017;72:104–14.
94. Halfvarson J, Brislawn CJ, Lamendella R, Vázquez-Baeza Y, Walters WA, Bramer LM, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. Nature Microbiology. 2017;2:1–7.
95. Hall AB, Yassour M, Sauk J, Garner A, Jiang X, Arthur T, et al. A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. Genome Medicine. 2017;9:103.

96. Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nat Med. 2019;25:667–78.
97. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nature Medicine. 2019;25:679–89.
98. Thoendel MJ, Jeraldo PR, Greenwood-Quaintance KE, Yao JZ, Chia N, Hanssen AD, et al. Identification of Prosthetic Joint Infection Pathogens Using a Shotgun Metagenomics Approach. Clinical Infectious Diseases. 2018;67:1333–8.
99. Decuypere S, Meehan CJ, Puyvelde SV, Block TD, Maltha J, Palpouguini L, et al. Diagnosis of Bacterial Bloodstream Infections: A 16S Metagenomics Approach. PLOS Neglected Tropical Diseases. 2016;10:e0004470.
100. Aarestrup FM, Rasmussen SR, Artursson K, Jensen NE. Trends in the resistance to antimicrobial agents of Streptococcus suis isolates from Denmark and Sweden. Veterinary Microbiology. 1998;63:71–80.
101. Huang J, Ma J, Shang K, Hu X, Liang Y, Li D, et al. Evolution and Diversity of the Antimicrobial Resistance Associated Mobilome in Streptococcus suis: A Probable Mobile Genetic Elements Reservoir for Other Streptococci. Front Cell Infect Microbiol. 2016;6.
102. Niazy M, Hill S, Nadeem K, Ricker N, Farzan A. Compositional analysis of the tonsil microbiota in relationship to Streptococcus suis disease in nursery pigs in Ontario. Animal Microbiome. 2022;4:10.
103. Segura M, Calzas C, Grenier D, Gottschalk M. Initial steps of the pathogenesis of the infection caused by Streptococcus suis: fighting against nonspecific defenses. FEBS Lett. 2016;590:3772–99.
104. Roodsant TJ, Van Der Putten BCL, Tamminga SM, Schultsz C, Van Der Ark KCH. Identification of Streptococcus suis putative zoonotic virulence factors: A systematic review and genomic meta-analysis. Virulence. 2021;12:2787–97.
105. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods. 2016;13:581–3.
106. Wirbel J, Zych K, Essex M, Karcher N, Kartal E, Salazar G, et al. Microbiome meta-analysis and cross-disease comparison enabled by the SIAMCAT machine learning toolbox. Genome Biology. 2021;22:93.
107. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020. https://doi.org/10.1093/jac/dkaa345.
108. De Witte C, Demeyere K, De Bruyckere S, Taminiau B, Daube G, Ducatelle R, et al. Characterization of the non-glandular gastric region microbiota in Helicobacter suis-infected versus non-infected pigs identifies a potential role for Fusobacterium gastrosuis in gastric ulceration. Veterinary Research. 2019;50:39.

109. Wolff B, Boutin S, Lorenz H-M, Ueffing H, Dalpke A, Wolff D. FRI0698 Prevotella and alloprevotella species characterize the oral microbiome of early rheumatoid arthritis. Annals of the Rheumatic Diseases. 2017;76 Suppl 2:754–754.
110. Correa-Fiz F, Santos JMG dos, Illas F, Aragon V. Antimicrobial removal on piglets promotes health and higher bacterial diversity in the nasal microbiota. Sci Rep. 2019;9:1–9.
111. Hall CW, Mah T-F. Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria. FEMS Microbiology Reviews. 2017;41:276–301.
112. Ricker N, Trachsel J, Colgan P, Jones J, Choi J, Lee J, et al. Toward Antibiotic Stewardship: Route of Antibiotic Administration Impacts the Microbiota and Resistance Gene Diversity in Swine Feces. Frontiers in Veterinary Science. 2020;7:255.
113. Zhang L, Huang Y, Zhou Y, Buckley T, Wang HH. Antibiotic Administration Routes Significantly Influence the Levels of Antibiotic Resistance in Gut Microbiota. Antimicrobial Agents and Chemotherapy. 2013. https://doi.org/10.1128/AAC.00670-13.
114. Law K, Lozinski B, Torres I, Davison S, Hilbrands A, Nelson E, et al. Disinfection of Maternal Environments Is Associated with Piglet Microbiome Composition from Birth to Weaning. mSphere. 2021;6:e00663-21.
115. Holman DB, Gzyl KE, Mou KT, Allen HK. Weaning Age and Its Effect on the Development of the Swine Gut Microbiome and Resistome. mSystems. 2021;6:e00682-21.
116. Choudhury R, Middelkoop A, Boekhorst J, Gerrits WJJ, Kemp B, Bolhuis JE, et al. Early life feeding accelerates gut microbiome maturation and suppresses acute post-weaning stress in piglets. Environmental Microbiology. 2021;23:7201–13.
117. Vo N, Tsai TC, Maxwell C, Carbonero F. Early exposure to agricultural soil accelerates the maturation of the early-life pig gut microbiota. Anaerobe. 2017;45:31–9.
118. Megahed A, Zeineldin M, Evans K, Maradiaga N, Blair B, Aldridge B, et al. Impacts of environmental complexity on respiratory and gut microbiome community structure and diversity in growing pigs. Sci Rep. 2019;9:13773.
119. Munk P, Knudsen BE, Lukjancenko O, Duarte ASR, Van Gompel L, Luiken REC, et al. Abundance and diversity of the faecal resistome in slaughter pigs and broilers in nine European countries. Nat Microbiol. 2018;3:898–908.
120. Mencía-Ares O, Cabrera-Rubio R, Cobo-Díaz JF, Álvarez-Ordóñez A, Gómez-García M, Puente H, et al. Antimicrobial use and production system shape the fecal, environmental, and slurry resistomes of pig farms. Microbiome. 2020;8:164.
121. Correa-Fiz F, Blanco-Fuertes M, Navas MJ, Lacasta A, Bishop RP, Githaka N, et al. Comparative analysis of the fecal microbiota from different species of domesticated and wild suids. Sci Rep. 2019;9:13616.
122. Huang J, Zhang W, Fan R, Liu Z, Huang T, Li J, et al. Composition and functional diversity of fecal bacterial community of wild boar, commercial pig and domestic native pig as revealed by 16S rRNA gene sequencing. Arch Microbiol. 2020;202:843–57.

123. Petrelli S, Buglione M, Maselli V, Troiano C, Larson G, Frantz L, et al. Population genomic, olfactory, dietary, and gut microbiota analyses demonstrate the unique evolutionary trajectory of feral pigs. Molecular Ecology. 2022;31:220–37.
124. Correa-Fiz F, Fraile L, Aragon V. Piglet nasal microbiota at weaning may influence the development of Glässer's disease during the rearing period. BMC Genomics. 2016;17.
125. Corsaut L, Misener M, Canning P, Beauchamp G, Gottschalk M, Segura M. Field Study on the Immunological Response and Protective Effect of a Licensed Autogenous Vaccine to Control Streptococcus suis Infections in Post-Weaned Piglets. Vaccines. 2020;8:384.
126. Curtis J, Bourne FJ. Half-lives of immunoglobulins IgG, IgA and IgM in the serum of new-born pigs. Immunology. 1973;24:147–55.
127. Morris CR, Gardner IA, Hietala SK, Carpenter TE, Anderson RJ, Parker KM. Persistence of passively acquired antibodies to Mycoplasma hyopneumoniae in a swine herd. Preventive Veterinary Medicine. 1994;21:29–41.
128. Procházka Z, Fránek M, Krejčí J. Duration of the Persistence of 131I-labelled Colostral and Serum IgG in the Blood of Newborn Piglets. Zentralblatt für Veterinärmedizin Reihe B. 1979;26:366–70.
129. Vandeputte J, Too HL, Ng FK, Chen C, Chai KK, Liao GA. Adsorption of colostral antibodies against classical swine fever, persistence of maternal antibodies, and effect on response to vaccination in baby pigs. Am J Vet Res. 2001;62:1805–11.
130. Vigre H, Ersbøll AK, Sørensen V. Decay of Acquired Colostral Antibodies to Actinobacillus pleuropneumoniae in Pigs. Journal of Veterinary Medicine, Series B. 2003;50:430–5.
131. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
132. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology. 2012;19:455–77.
133. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
134. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17:10–2.
135. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590–6.
136. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol. 2017;35:1026–8.
137. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE. 2013;8:e61217.
138. Oksanen J, Kindt R, Legendre P, O'Hara B, Stevens MHH, Oksanen MJ, et al. The vegan package. Community ecology package. 2007;10:631–7.
139. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20:257.

140. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018;6:158.
141. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2020;36:1925–7.
142. Brandtzaeg PER. Molecular and cellular aspects of the secretory immunoglobulin system. APMIS. 1995;103:1–19.
143. Nave H, Gebert A, Pabst R. Morphology and immunology of the human palatine tonsil. Anat Embryol (Berl). 2001;204:367–73.
144. Scadding GK. Immunology of the tonsil: a review. J R Soc Med. 1990;83:104–7.
145. Fujimura Y. Evidence of M cells as portals of entry for antigens in the nasopharyngeal lymphoid tissue of humans. Virchows Arch. 2000;436:560–6.
146. Gram T, Ahrens P, Andreasen M, Nielsen JP. An Actinobacillus pleuropneumoniae PCR typing system based on the apx and omlA genes — evaluation of isolates from lungs and tonsils of pigs. Veterinary Microbiology. 2000;75:43–57.
147. Volf J, Stepanova H, Matiasovic J, Kyrova K, Sisak F, Havlickova H, et al. Salmonella enterica serovar Typhimurium and Enteritidis infection of pigs and cytokine signalling in palatine tonsils. Veterinary Microbiology. 2012;156:127–35.
148. Gottschalk M, Segura M. The pathogenesis of the meningitis caused by Streptococcus suis: the unresolved questions. Veterinary Microbiology. 2000;76:259–72.
149. Correa-Fiz F, Neila-Ibáñez C, López-Soria S, Napp S, Martinez B, Sobrevia L, et al. Feed additives for the control of post-weaning Streptococcus suis disease and the effect on the faecal and nasal microbiota. Scientific Reports. 2020;10:20354.
150. Fredriksen S, Neila-Ibáñez C, Hennig-Pauka I, Guan X, Dunkelberger J, Fernandes de Oliveira I, et al. Streptococcus suis infection on European farms is associated with an altered tonsil microbiome and resistome. bioRxiv. 2022;:2022.08.01.500980.
151. Fleming-Dutra KE, Conklin L, Loo JD, Knoll MD, Park DE, Kirk J, et al. Systematic review of the effect of pneumococcal conjugate vaccine dosing schedules on vaccine-type nasopharyngeal carriage. Pediatr Infect Dis J. 2014;33 Suppl 2:S152-160.
152. Su F, Patel GB, Hu S, Chen W. Induction of mucosal immunity through systemic immunization: Phantom or reality? Human Vaccines & Immunotherapeutics. 2016;12:1070–9.
153. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biology. 2011;12:R60.
154. Santi I, Pezzicoli A, Bosello M, Berti F, Mariani M, Telford JL, et al. Functional Characterization of a Newly Identified Group B Streptococcus Pullulanase Eliciting Antibodies Able to Prevent Alpha-Glucans Degradation. PLOS ONE. 2008;3:e3787.
155. Obregon-Gutierrez P, Aragon V, Correa-Fiz F. Sow Contact Is a Major Driver in the Development of the Nasal Microbiota of Piglets. Pathogens. 2021;10:697.
156. Berry ASF, Pierdon MK, Misic AM, Sullivan MC, O'Brien K, Chen Y, et al. Remodeling of the maternal gut microbiome during pregnancy is shaped by parity. Microbiome. 2021;9:146.

157. Gaio D, DeMaere MZ, Anantanawat K, Chapman TA, Djordjevic SP, Darling AE. Post-weaning shifts in microbiome composition and metabolism revealed by over 25 000 pig gut metagenome-assembled genomes. Microbial Genomics. 2021;7.
158. Tang Q, Huang K, Liu J, Jin X, Li C. Distribution characteristics of bioaerosols inside pig houses and the respiratory tract of pigs. Ecotoxicology and Environmental Safety. 2021;212:112006.
159. Muns R, Nuntapaitoon M, Tummaruk P. Non-infectious causes of pre-weaning mortality in piglets. Livestock Science. 2016;184:46–57.
160. Gaukroger CH, Stewart CJ, Edwards SA, Walshaw J, Adams IP, Kyriazakis I. Changes in Faecal Microbiota Profiles Associated With Performance and Birthweight of Piglets. Front Microbiol. 2020;11.
161. Huang S-M, Wu Z-H, Li T-T, Liu C, Han D-D, Tao S-Y, et al. Perturbation of the lipid metabolism and intestinal inflammation in growing pigs with low birth weight is associated with the alterations of gut microbiota. Science of The Total Environment. 2020;719:137382.
162. Li N, Huang S, Jiang L, Wang W, Li T, Zuo B, et al. Differences in the Gut Microbiota Establishment and Metabolome Characteristics Between Low- and Normal-Birth-Weight Piglets During Early-Life. Front Microbiol. 2018;9.
163. Zhang W, Ma C, Xie P, Zhu Q, Wang X, Yin Y, et al. Gut microbiota of newborn piglets with intrauterine growth restriction have lower diversity and different taxonomic abundances. Journal of Applied Microbiology. 2019;127:354–69.
164. Dong L, Zhong X, Ahmad H, Li W, Wang Y, Zhang L, et al. Intrauterine Growth Restriction Impairs Small Intestinal Mucosal Immunity in Neonatal Piglets. J Histochem Cytochem. 2014;62:510–8.
165. Tao S, Bai Y, Li T, Li N, Wang J. Original low birth weight deteriorates the hindgut epithelial barrier function in pigs at the growing stage. The FASEB Journal. 2019;33:9897–912.
166. Amdi C, Krogh U, Flummer C, Oksbjerg N, Hansen CF, Theil PK. Intrauterine growth restricted piglets defined by their head shape ingest insufficient amounts of colostrum. J Anim Sci. 2013;91:5605–13.
167. Huting AMS, Sakkas P, Wellock I, Almond K, Kyriazakis I. Once small always small? To what extent morphometric characteristics and post-weaning starter regime affect pig lifetime growth performance. Porcine Health Management. 2018;4:21.
168. Quesnel H, Farmer C, Devillers N. Colostrum intake: Influence on piglet performance and factors of variation. Livestock Science. 2012;146:105–14.
169. Blanco-Fuertes M, Correa-Fiz F, López-Serrano S, Sibila M, Aragon V. Sow vaccination against virulent Glaesserella parasuis shapes the nasal microbiota of their offspring. Sci Rep. 2022;12:3357.
170. Vela AI, Sánchez V, Mentaberre G, Lavín S, Domínguez L, Fernández-Garayzábal JF. Streptococcus porcorum sp. nov., isolated from domestic and wild pigs. Int J Syst Evol Microbiol. 2011;61 Pt 7:1585–9.
171. Bergström A, Skov TH, Bahl MI, Roager HM, Christensen LB, Ejlerskov KT, et al. Establishment of Intestinal Microbiota during Early Life: a Longitudinal, Explorative Study of a Large Cohort of Danish Infants. Appl Environ Microbiol. 2014;80:2889–900.

172. Frese SA, Parker K, Calvert CC, Mills DA. Diet shapes the gut microbiome of pigs during nursing and weaning. Microbiome. 2015;3:28.
173. Pajarillo EAB, Chae J-P, Balolong MP, Kim HB, Kang D-K. Assessment of fecal bacterial diversity among healthy piglets during the weaning transition. The Journal of General and Applied Microbiology. 2014;60:140–6.
174. D'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, et al. Antibiotic resistance is ancient. Nature. 2011;477:457–61.
175. Aarestrup FM. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. Philosophical Transactions of the Royal Society B: Biological Sciences. 2015;370:20140085.
176. Hernando-Amado S, Coque TM, Baquero F, Martínez JL. Defining and combating antibiotic resistance from One Health and Global Health perspectives. Nature Microbiology. 2019;4:1432–42.
177. Baker S, Thomson N, Weill F-X, Holt KE. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. Science. 2018;360:733–8.
178. Crits-Christoph A, Hallowell HA, Koutouvalis K, Suez J. Good microbes, bad genes? The dissemination of antimicrobial resistance in the human microbiome. Gut Microbes. 2022;14:2055944.
179. Andersson DI, Hughes D. Antibiotic resistance and its cost: is it possible to reverse resistance? Nat Rev Microbiol. 2010;8:260–71.
180. Bengtsson-Palme J, Angelin M, Huss M, Kjellqvist S, Kristiansson E, Palmgren H, et al. The Human Gut Microbiome as a Transporter of Antibiotic Resistance Genes between Continents. Antimicrobial Agents and Chemotherapy. 2015;59:6551–60.
181. Davies M, Galazzo G, van Hattem JM, Arcilla MS, Melles DC, de Jong MD, et al. Enterobacteriaceae and Bacteroidaceae provide resistance to travel-associated intestinal colonization by multi-drug resistant Escherichia coli. Gut Microbes. 2022;14:2060676.
182. Frost I, Van Boeckel TP, Pires J, Craig J, Laxminarayan R. Global geographic trends in antimicrobial resistance: the role of international travel. Journal of Travel Medicine. 2019;26:taz036.
183. Pan Y, Zeng J, Li L, Yang J, Tang Z, Xiong W, et al. Coexistence of Antibiotic Resistance Genes and Virulence Factors Deciphered by Large-Scale Complete Genome Analysis. mSystems. 2020;5:e00821-19.
184. Yi X, Li Y, Liu H, Liu X, Yang J, Gao J, et al. Inflammatory Endotype-Associated Airway Resistome in Chronic Obstructive Pulmonary Disease. Microbiology Spectrum. 2022;10:e02593-21.
185. Yang S, Liu Y, Yang N, Lan Y, Lan W, Feng J, et al. The gut microbiome and antibiotic resistome of chronic diarrhea rhesus macaques (Macaca mulatta) and its similarity to the human gut microbiome. Microbiome. 2022;10:29.
186. Kieser S, Sarker SA, Sakwinska O, Foata F, Sultana S, Khan Z, et al. Bangladeshi children with acute diarrhoea show faecal microbiomes with increased Streptococcus abundance, irrespective of diarrhoea aetiology. Environmental Microbiology. 2018;20:2256–69.

187. Qiu Q, Wang J, Yan Y, Roy B, Chen Y, Shang X, et al. Metagenomic Analysis Reveals the Distribution of Antibiotic Resistance Genes in a Large-Scale Population of Healthy Individuals and Patients With Varied Diseases. Frontiers in Molecular Biosciences. 2020;7.
188. Bose D, Chatterjee S, Older E, Seth R, Janulewicz P, Saha P, et al. Host gut resistome in Gulf War chronic multisymptom illness correlates with persistent inflammation. Commun Biol. 2022;5:1–15.
189. Diard M, Bakkeren E, Cornuault JK, Moor K, Hausmann A, Sellin ME, et al. Inflammation boosts bacteriophage transfer between Salmonella spp. Science. 2017;355:1211–5.
190. Stecher B, Denzler R, Maier L, Bernet F, Sanders MJ, Pickard DJ, et al. Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae. Proceedings of the National Academy of Sciences. 2012;109:1269–74.
191. Nearing JT, Comeau AM, Langille MGI. Identifying biases and their potential solutions in human microbiome studies. Microbiome. 2021;9:113.
192. Schloss PD. Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. mBio. 2018;9:e00525-18.
193. Kidd TJ, Canton R, Ekkelenkamp M, Johansen HK, Gilligan P, LiPuma JJ, et al. Defining antimicrobial resistance in cystic fibrosis. Journal of Cystic Fibrosis. 2018;17:696–704.
194. Pust M-M, Wiehlmann L, Davenport C, Rudolf I, Dittrich A-M, Tümmler B. The human respiratory tract microbial community structures in healthy and cystic fibrosis infants. npj Biofilms Microbiomes. 2020;6:1–10.
195. Taylor SL, Leong LEX, Sims SK, Keating RL, Papanicolas LE, Richard A, et al. The cystic fibrosis gut as a potential source of multidrug resistant pathogens. Journal of Cystic Fibrosis. 2021;20:413–20.
196. Wang Y, Leong LEX, Keating RL, Kanno T, Abell GCJ, Mobegi FM, et al. Opportunistic bacteria confer the ability to ferment prebiotic starch in the adult cystic fibrosis gut. Gut Microbes. 2019;10:367–81.
197. David LA, Weil A, Ryan ET, Calderwood SB, Harris JB, Chowdhury F, et al. Gut Microbial Succession Follows Acute Secretory Diarrhea in Humans. mBio. 2015;6:e00381-15.
198. McDonald LC. Effects of short- and long-course antibiotics on the lower intestinal microbiome as they relate to traveller's diarrhea. Journal of Travel Medicine. 2017;24 suppl_1:S35–8.
199. Monaghan TM, Sloan TJ, Stockdale SR, Blanchard AM, Emes RD, Wilcox M, et al. Metagenomics reveals impact of geography and acute diarrheal disease on the Central Indian human gut microbiome. Gut Microbes. 2020;12:1752605.
200. Soto-Girón MJ, Peña-Gonzalez A, Hatt JK, Montero L, Páez M, Ortega E, et al. Gut Microbiome Changes with Acute Diarrheal Disease in Urban Versus Rural Settings in Northern Ecuador. The American Journal of Tropical Medicine and Hygiene. 2021;104:2275–85.
201. Nguyen LH, Örtqvist AK, Cao Y, Simon TG, Roelstraete B, Song M, et al. Antibiotic use and the development of inflammatory bowel disease: a national case-control study in Sweden. The Lancet Gastroenterology & Hepatology. 2020;5:986–95.

202. Sartor RB, Wu GD. Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflammatory Bowel Diseases and Therapeutic Approaches. Gastroenterology. 2017;152:327-339.e4.
203. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature. 2019;569:655–62.
204. Abbas-Egbariya H, Haberman Y, Braun T, Hadar R, Denson L, Gal-Mor O, et al. Meta-analysis defines predominant shared microbial responses in various diseases and a specific inflammatory bowel disease signal. Genome Biology. 2022;23:61.
205. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol. 2014;32:822–8.
206. Schirmer M, Franzosa EA, Lloyd-Price J, McIver LJ, Schwager R, Poon TW, et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. Nat Microbiol. 2018;3:337–46.
207. Feng Q, Liang S, Jia H, Stadlmayr A, Tang L, Lan Z, et al. Gut microbiome development along the colorectal adenoma–carcinoma sequence. Nat Commun. 2015;6:6528.
208. Gupta A, Dhakan DB, Maji A, Saxena R, P.K. VP, Mahajan S, et al. Association of Flavonifractor plautii, a Flavonoid-Degrading Bacterium, with the Gut Microbiome of Colorectal Cancer Patients in India. mSystems. 2019;4:e00438-19.
209. Hannigan GD, Duhaime MB, Ruffin MT, Koumpouras CC, Schloss PD. Diagnostic Potential and Interactive Dynamics of the Colorectal Cancer Virome. mBio. 2018;9:e02248-18.
210. Vogtmann E, Hua X, Zeller G, Sunagawa S, Voigt AY, Hercog R, et al. Colorectal Cancer and the Human Gut Microbiome: Reproducibility with Whole-Genome Shotgun Sequencing. PLOS ONE. 2016;11:e0155362.
211. Yachida S, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. Nat Med. 2019;25:968–76.
212. Yu J, Feng Q, Wong SH, Zhang D, Liang Q yi, Qin Y, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. Gut. 2017;66:70–8.
213. Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. Mol Syst Biol. 2014;10:766.
214. Song M, Chan AT, Sun J. Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. Gastroenterology. 2020;158:322–40.
215. Wilson AS, Koller KR, Ramaboli MC, Nesengani LT, Ocvirk S, Chen C, et al. Diet and the Human Gut Microbiome: An International Review. Dig Dis Sci. 2020;65:723–40.
216. Bowyer RCE, Jackson MA, Le Roy CI, Ni Lochlainn M, Spector TD, Dowd JB, et al. Socioeconomic Status and the Gut Microbiome: A TwinsUK Cohort Study. Microorganisms. 2019;7:17.

217. Renson A, Herd P, Dowd JB. Sick Individuals and Sick (Microbial) Populations: Challenges in Epidemiology and the Microbiome. Annual Review of Public Health. 2020;41:63–80.
218. Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, et al. Interconnected microbiomes and resistomes in low-income human habitats. Nature. 2016;533:212–6.
219. van Schaik W. The human gut resistome. Philosophical Transactions of the Royal Society B: Biological Sciences. 2015;370:20140087.
220. Maguire F, Jia B, Gray KL, Lau WYV, Beiko RG, Brinkman FSL. Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic Islands. Microb Genom. 2020;6:mgen000436.
221. Stracy M, Snitser O, Yelin I, Amer Y, Parizade M, Katz R, et al. Minimizing treatment-induced emergence of antibiotic resistance in bacterial infections. Science. 2022;375:889–94.
222. Mac Aogáin M, Lau KJX, Cai Z, Kumar Narayana J, Purbojati RW, Drautz-Moses DI, et al. Metagenomics Reveals a Core Macrolide Resistome Related to Microbiota in Chronic Respiratory Disease. Am J Respir Crit Care Med. 2020;202:433–47.
223. Rozwandowicz M, Brouwer MSM, Fischer J, Wagenaar JA, Gonzalez-Zorn B, Guerra B, et al. Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. Journal of Antimicrobial Chemotherapy. 2018;73:1121–37.
224. Pasolli E, Schiffer L, Manghi P, Renson A, Obenchain V, Truong DT, et al. Accessible, curated metagenomic data through ExperimentHub. Nat Methods. 2017;14:1023–4.
225. Fredriksen S, Guan X, Boekhorst J, Molist F, van Baarlen P, Wells JM. Environmental and maternal factors shaping tonsillar microbiota development in piglets. BMC Microbiology. 2022;22:224.
226. Atack JM, Weinert LA, Tucker AW, Husna AU, Wileman TM, F. Hadjirin N, et al. Streptococcus suis contains multiple phase-variable methyltransferases that show a discrete lineage distribution. Nucleic Acids Research. 2018;46:11466–76.
227. Zhong X, Ma J, Bai Q, Zhu Y, Zhang Y, Gu Q, et al. Identification of the RNA-binding domain-containing protein RbpA that acts as a global regulator of the pathogenicity of Streptococcus suis serotype 2. Virulence. 2022;13:1304–14.
228. Arenas J, Bossers-de Vries R, Harders-Westerveen J, Buys H, Ruuls-van Stalle LMF, Stockhofe-Zurwieden N, et al. In vivo transcriptomes of Streptococcus suis reveal genes required for niche-specific adaptation and pathogenesis. Virulence. 2019;10:334–51.
229. Koczula A, Jarek M, Visscher C, Valentin-Weigand P, Goethe R, Willenborg J. Transcriptomic Analysis Reveals Selective Metabolic Adaptation of Streptococcus suis to Porcine Blood and Cerebrospinal Fluid. Pathogens. 2017;6:7.
230. Ni H, Li M, Wang Q, Wang J, Liu X, Zheng F, et al. Inactivation of the htpsA gene affects capsule development and pathogenicity of Streptococcus suis. Virulence. 2020;11:927–40.
231. Zheng C, Wei M, Qiu J, Jia M, Zhou X, Jiao X. TroR Negatively Regulates the TroABCD System and Is Required for Resistance to Metal Toxicity and Virulence in Streptococcus suis. Applied and Environmental Microbiology. 2021. https://doi.org/10.1128/AEM.01375-21.

232. Zhong X, Zhang Y, Zhu Y, Dong W, Ma J, Pan Z, et al. Identification of an Autorepressing Two-Component Signaling System That Modulates Virulence in Streptococcus suis Serotype 2. Infection and Immunity. 2019. https://doi.org/10.1128/IAI.00377-19.
233. Segura M, Fittipaldi N, Calzas C, Gottschalk M. Critical Streptococcus suis Virulence Factors: Are They All Really Critical? Trends in Microbiology. 2017;25:585–99.
234. Vecht U, Arends J, van der Molen E, van Leengoed L. Differences in virulence between two strains of Streptococcus suis type II after experimentally induced infection of newborn germ-free pigs. American journal of veterinary research. 1989;50:1037–43.
235. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Computational Biology. 2017;13:e1005595.
236. Deng S, Xu T, Fang Q, Yu L, Zhu J, Chen L, et al. The Surface-Exposed Protein SntA Contributes to Complement Evasion in Zoonotic Streptococcus suis. Frontiers in Immunology. 2018;9.
237. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Research. 2019;47:W81–7.
238. Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, et al. A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol. 2020;16:60–8.
239. Gilchrist CLM, Chooi Y-H. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics. 2021;37:2473–5.
240. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:238.
241. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30:1236–40.
242. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology. 2009;10:R25.
243. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30.
244. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2022.
245. Lenth RV. emmeans: Estimated Marginal Means, aka Least-Squares Means. 2022.
246. Fulde M, Willenborg J, Huber C, Hitzmann A, Willms D, Seitz M, et al. The arginine-ornithine antiporter ArcD contributes to biological fitness of Streptococcus suis. Frontiers in Cellular and Infection Microbiology. 2014;4:107.
247. Fulde M, Willenborg J, de Greeff A, Benga L, Smith HE, Valentin-Weigand P, et al. ArgR is an essential local transcriptional regulator of the arcABC operon in Streptococcus suis and is crucial for biological fitness in an acidic environment. Microbiology. 2011;157:572–82.
248. Baums CG, Kaim U, Fulde M, Ramachandran G, Goethe R, Valentin-Weigand P. Identification of a Novel Virulence Determinant with Serum Opacification Activity in Streptococcus suis. Infection and Immunity. 2006;74:6154–62.

249. Takamatsu D, Osaki M, Tharavichitkul P, Takai S, Sekizaki T 2008. Allelic variation and prevalence of serum opacity factor among the Streptococcus suis population. Journal of Medical Microbiology. 2008;57:488–94.

250. Courtney HS, Pownall HJ. The Structure and Function of Serum Opacity Factor: A Unique Streptococcal Virulence Determinant That Targets High-Density Lipoproteins. Journal of Biomedicine and Biotechnology. 2010;2010:e956071.

251. Baig A, Weinert LA, Peters SE, Howell KJ, Chaudhuri RR, Wang J, et al. Whole genome investigation of a divergent clade of the pathogen Streptococcus suis. Frontiers in Microbiology. 2015;6.

252. Atack JM, Tan A, Bakaletz LO, Jennings MP, Seib KL. Phasevarions of Bacterial Pathogens: Methylomics Sheds New Light on Old Enemies. Trends in Microbiology. 2018;26:715–26.

253. Van Calsteren M-R, Gagnon F, Lacouture S, Fittipaldi N, Gottschalk M. Structure determination of Streptococcus suis serotype 2 capsular polysaccharide. Biochem Cell Biol. 2010;88:513–25.

254. Okura M, Auger J-P, Shibahara T, Goyette-Desjardins G, Van Calsteren M-R, Maruyama F, et al. Capsular polysaccharide switching in Streptococcus suis modulates host cell interactions and virulence. Sci Rep. 2021;11:6513.

255. Smith HE, Damman M, van der Velde J, Wagenaar F, Wisselink HJ, Stockhofe-Zurwieden N, et al. Identification and Characterization of the cps Locus of Streptococcus suis Serotype 2: the Capsule Protects against Phagocytosis and Is an Important Virulence Factor. Infection and Immunity. 1999;67:1750–6.

256. Wu K, Xu H, Zheng Y, Wang L, Zhang X, Yin Y. CpsR, a GntR family regulator, transcriptionally regulates capsular polysaccharide biosynthesis and governs bacterial virulence in Streptococcus pneumoniae. Sci Rep. 2016;6:29255.

257. Y, Nie R, Liu X, Kong J, Wang X, Li J. GntR is involved in the expression of virulence in strain Streptococcus suis P1/7. FEMS Microbiology Letters. 2018;365:fny091.

258. Shah P, Nanduri B, Swiatlo E, Ma Y, Pendarvis K. Polyamine biosynthesis and transport mechanisms are crucial for fitness and pathogenesis of Streptococcus pneumoniae. Microbiology. 2011;157:504–15.

259. Griswold AR, Jameson-Lee M, Burne RA. Regulation and Physiologic Significance of the Agmatine Deiminase System of Streptococcus mutans UA159. Journal of Bacteriology. 2006;188:834–41.

260. Griswold AR, Chen Y-YM, Burne RA. Analysis of an Agmatine Deiminase Gene Cluster in Streptococcus mutans UA159. Journal of Bacteriology. 2004. https://doi.org/10.1128/JB.186.6.1902-1904.2004.

261. Liu Y, Zeng L, Burne RA. AguR Is Required for Induction of the Streptococcus mutans Agmatine Deiminase System by Low pH and Agmatine. Applied and Environmental Microbiology. 2009;75:2629–37.

262. Gruening P, Fulde M, Valentin-Weigand P, Goethe R. Structure, Regulation, and Putative Function of the Arginine Deiminase System of Streptococcus suis. Journal of Bacteriology. 2006. https://doi.org/10.1128/JB.188.2.361-369.2006.

263. Johnston JW, Myers LE, Ochs MM, Benjamin WH, Briles DE, Hollingshead SK. Lipoprotein PsaA in Virulence of Streptococcus pneumoniae: Surface Accessibility and Role in Protection from Superoxide. Infection and Immunity. 2004;72:5858–67.
264. Lisher J, Giedroc D. Manganese acquisition and homeostasis at the host-pathogen interface. Frontiers in Cellular and Infection Microbiology. 2013;3:91.
265. Niven DF, Ekins A, Al-Samaurai AA-W. Effects of iron and manganese availability on growth and production of superoxide dismutase by Streptococcus suis. Can J Microbiol. 1999;45:1027–32.
266. Schreur PJW, Rebel JMJ, Smits MA, van Putten JPM, Smith HE. TroA of Streptococcus suis Is Required for Manganese Acquisition and Full Virulence. Journal of Bacteriology. 2011;193:5073–80.
267. Zheng B, Zhang Q, Gao J, Han H, Li M, Zhang J, et al. Insight into the Interaction of Metal Ions with TroA from Streptococcus suis. PLOS ONE. 2011;6:e19510.
268. Peng W, Yang X, Wang N, Gao T, Liu Z, Liu W, et al. PerR-Regulated Manganese Import Contributes to Oxidative Stress Defense in Streptococcus suis. Applied and Environmental Microbiology. 0:e00086-22.
269. Zhang T, Ding Y, Li T, Wan Y, Li W, Chen H, et al. A Fur-like protein PerR regulates two oxidative stress response related operons dpr and metQIN in Streptococcus suis. BMC Microbiology. 2012;12:85.
270. Peng W, Yang X, Wang Y, Wang N, Li X, Chen H, et al. Mn uptake system affects the virulence of Streptococcus suis by mediating oxidative stress. Veterinary Microbiology. 2022;272:109518.
271. Xu J, Zheng C, Cao M, Zeng T, Zhao X, Shi G, et al. The manganese efflux system MntE contributes to the virulence of Streptococcus suis serotype 2. Microbial Pathogenesis. 2017;110:23–30.
272. de Greeff A, Schokker D, Roubos-van den Hil P, Ramaekers P, Vastenhouw SA, Harders F, et al. The effect of maternal antibiotic use in sows on intestinal development in offspring. J Anim Sci. 2020;98.
273. Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. International Journal of Antimicrobial Agents. 2010;35:322–32.
274. Tuganbaev T, Yoshida K, Honda K. The effects of oral microbiota on health. Science. 2022;376:934–6.
275. Zaura E, Brandt BW, Mattos MJT de, Buijs MJ, Caspers MPM, Rashid M-U, et al. Same Exposure but Two Radically Different Responses to Antibiotics: Resilience of the Salivary Microbiome versus Long-Term Microbial Shifts in Feces. mBio. 2015;6.
276. McLean AR, Torres-Morales J, Dewhirst FE, Borisy GG, Welch JLM. Site-tropism of streptococci in the oral microbiome. Molecular Oral Microbiology. 2022.
277. Magee AD, Yother J. Requirement for Capsule in Colonization by Streptococcus pneumoniae. Infection and Immunity. 2001;69:3755–61.
278. Hathaway LJ, Brugger SD, Morand B, Bangert M, Rotzetter JU, Hauser C, et al. Capsule Type of Streptococcus pneumoniae Determines Growth Phenotype. PLOS Pathogens. 2012;8:e1002574.

279. Barnett TC, Cole JN, Rivera-Hernandez T, Henningham A, Paton JC, Nizet V, et al. Streptococcal toxins: role in pathogenesis and disease. Cellular Microbiology. 2015;17:1721–41.
280. Tenenbaum T, Asmat TM, Seitz M, Schroten H, Schwerk C. Biological activities of suilysin: role in Streptococcus suis pathogenesis. Future Microbiology. 2016;11:941–54.
281. Hamaguchi S, Zafar MA, Cammer M, Weiser JN. Capsule Prolongs Survival of Streptococcus pneumoniae during Starvation. Infection and Immunity. 2018;86:e00802-17.
282. Charland N, Harel J, Kobisch M, Lacasse S, Gottschalk M 1998. Streptococcus suis serotype 2 mutants deficient in capsular expression. Microbiology. 144:325–32.
283. Fittipaldi N, Segura M, Grenier D, Gottschalk M. Virulence factors involved in the pathogenesis of the infection caused by the swine pathogen and zoonotic agent Streptococcus suis. Future Microbiology. 2012;7:259–79.
284. Welch JLM, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. Biogeography of a human oral microbiome at the micron scale. PNAS. 2016;113:E791–800.
285. Morillo-Lopez V, Sjaarda A, Islam I, Borisy GG, Mark Welch JL. Corncob structures in dental plaque reveal microhabitat taxon specificity. Microbiome. 2022;10:145.
286. Horter DC, Yoon K-J, Zimmerman JJ. A review of porcine tonsils in immunity and disease. Animal Health Research Reviews. 2003;4:143–55.
287. Woo JH, Kim ST, Kang IG, Lee JH, Cha HE, Kim DY. Comparison of tonsillar biofilms between patients with recurrent tonsillitis and a control group. Acta Oto-Laryngologica. 2012;132:1115–20.
288. McKenna A, Ijaz UZ, Kelly C, Linton M, Sloan WT, Green BD, et al. Impact of industrial production system parameters on chicken microbiomes: mechanisms to improve performance and reduce Campylobacter. Microbiome. 2020;8:128.
289. Meng F, Wu NH, Nerlich A, Herrler G, Valentin-Weigand P, Seitz M. Dynamic Virus-Bacterium Interactions in a Porcine Precision-Cut Lung Slice Coinfection Model: Swine Influenza Virus Paves the Way for Streptococcus suis Infection in a Two-Step Process. Infection and Immunity. 2015;83:2806–15.
290. Vollmers J, Wiegand S, Kaster A-K. Comparing and Evaluating Metagenome Assembly Tools from a Microbiologist's Perspective - Not Only Size Matters! PLOS ONE. 2017;12:e0169662.
291. Yang C, Chowdhury D, Zhang Z, Cheung WK, Lu A, Bian Z, et al. A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. Computational and Structural Biotechnology Journal. 2021;19:6301–14.
292. Quince C, Nurk S, Raguideau S, James R, Soyer OS, Summers JK, et al. STRONG: metagenomics strain resolution on assembly graphs. Genome Biol. 2021;22:214.
293. Marotz CA, Sanders JG, Zuniga C, Zaramela LS, Knight R, Zengler K. Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome. 2018;6:42.

Summary

Infectious disease and antibiotic resistance are global health problems projected to increase in coming decades. The emergence of antibiotic resistance is intrinsically linked to use of antibiotics, which in turn is linked to bacterial infectious disease. Thus, increased understanding of mechanisms driving emerging resistance is of great importance. This thesis aimed to expand knowledge on the zoonotic opportunistic pathogen *Streptococcus suis* and the tonsillar microbiome of piglets. *S. suis* is of great interest in a One Health perspective due to its high prevalence, zoonotic potential, antimicrobial use being the main treatment, and potential to spread antimicrobial resistance genes to closely related *Streptococcus* species. **Chapter 1** introduced the current knowledge on these topics.

In **Chapter 2** we conducted a field study comparing *S. suis* symptomatic case piglets with asymptomatic controls, utilizing amplicon and shotgun sequencing to assess tonsil microbiome associations. We found differences in species level composition and antibiotic resistance gene (ARG) abundance. These results indicated that specific microbiome members may be involved in preventing or facilitating *S. suis* invasive disease. Furthermore, antibiotic usage targeting *S. suis* disease symptoms appear to drive positive selection for ARGs in disease-associated strains. We further assessed the population biology of *S. suis*. Amplicon sequencing of the 16S rRNA gene V3-V4 region showed that up to 16 different variants likely to be *S. suis* could be found in individual samples. This indicated the presence of a great number of strains as each sequence variant may represent several different clades. We used shotgun sequencing to determine that the majority of *S. suis* in the tonsillar microbiota belong to commensal lineages. These findings were novel because previous works on *S. suis* prevalence were culture based and assessed presence and absence of *S. suis*, while our work shows that commensal *S. suis* colonises the tonsils of all sampled piglets.

Chapter 3 aimed to assess the development of the tonsillar microbiota of piglets during the first weeks of life. Of particular interest was the abundance of *Streptococcus* species, because the interplay between weaning and introduction of dry feed, disruption of microbiota homeostasis, and increased abundance of *S. suis* on the tonsils may predispose *S. suis* virulence and disease. Previous research from the USA observed a large increase in *Streptococcaceae* abundance after weaning, and although they did not show taxonomic shifts at the species level, this fits well with the hypothesis of streptococcal outgrowth. We found that increased *Streptococcaceae* abundance after weaning is driven by *S. porcorum* and other less well described *Streptococcus* species, while *S. suis* decreases in abundance. Our larger sample size and inclusion of more separate litters than previous studies enabled us to assess the relative impact of different within-farm parameters on shaping microbiota composition. We found that microbiota composition differed between rooms, and that siblings shared highly similar microbiota composition. This indicated that sows may have a large impact on the microbiota of their offspring. To experimentally test this hypothesis, we treated one group of sows with a pre-partum bacterin vaccination

against two pathogenic *S. suis* strains from the farm. The sow vaccination resulted in alterations in the abundance of *Streptococcus* species in their offspring. The great impact of sows on their piglets is of interest because it is easier to apply treatments to sows than each individual piglet.

Chapter 4 followed up on resistome results from chapter 2 by conducting a meta-analysis on human gut microbiome datasets. We aimed to better assess resistome dynamics and to assert whether there is a consistent link between antibiotic treatment of disease and the abundance of ARGs in the disease-associated microbiome. We found that within case-control studies, cases with antibiotic treated diseases, namely cystic fibrosis and diarrhoea, had larger resistomes than healthy controls. Other diseases not commonly treated with antibiotics had limited and variable case-control differences. Thus, it appears that our finding of expanded resistomes in *S. suis* diseased piglets is consistent with other studies. We also noted that resistome analysis is prone to a range of biases that may be induced by study design and sequencing methodology. While our study on piglets featured selection of case-control pairs from the same pens and rooms, and thus groups equivalent in environment, diet, and antimicrobial treatment throughout their life, human studies face great challenges in participant selection.

Chapter 5 aimed to assess how pathogenic and commensal *S. suis* strains differ at the transcriptomic level. This was of interest because although commensal *S. suis* are highly successful colonisers of piglet tonsils, limited knowledge exists on what differentiates them from commonly studied pathogens. We compared the transcriptome of 21 *S. suis* strains, including both commonly studied sequence type 1 pathogens, other pathogenic clades, and understudied commensals. We found that all strains shared high transcriptome similarity despite large genomic sequence variation. Even divergent strains, the inclusion of which in the *S. suis* species is debated, shared transcriptome conservation overlapping with that between normal *S. suis* clades. While overall gene expression was similar between *S. suis* strains, regulation and expression of key pathways also varied. When grown in porcine serum, pathogens had strong upregulation of genes involved in capsule production while commensals upregulated expression of the agmatine deiminase system. Sequence type 1 strains appeared to regulate gene expression more strongly than other clades, and regulatory mechanisms well described in sequence type 1 strains may not apply to other *S. suis* clades.

In **chapter 6** I discuss the results from the research chapters in common context. While this thesis made substantial advances in our knowledge of the tonsillar microbiota and the biology of *S. suis*, our whole-tonsil swab samples may miss micron-level spatial variation key to both microbe-microbe and host-microbe interactions. Different *S. suis* clades may have fundamentally different ecological niches but at the same time antagonise each other via bacteriocins and by eliciting host immune responses. Future studies on the piglet microbiome and/or *S. suis* carriage should take great care in study design and selection of appropriate cases and controls.

Acknowledgements

The most important people supporting a successful PhD candidate are the supervisors. All my (co-)promotors made genuine contributions to the content of this thesis. Thank you to Jerry for sharing your knowledge and for being a promotor with a direct involvement and interest in the research projects. You have guided me towards becoming an independent scientist. Thank you, Peter, for your reliably fast feedback on emails and manuscripts and because you were the one that made the decision to hire me for the PhD position in the first place. Thank you, Jos, for guidance in bioinformatics and statistics which has had a large impact on my approach to data analysis. Furthermore, my PhD was a collaborative effort, and the research described in this thesis would not have been possible without contributions and supervision from external collaborators. In particular, thank you to Carlos, Flor, and Virginia at IRTA-CReSA.

A special thanks also to my students, Stef, Suzanne, and Loïs, for their contributions.

The group is also key to developing as a scientist. The most important HMI group member has been Kristiina for providing large quantities of mental support and for designing my thesis cover. Thank you, Nan, for being a close collaborator on many projects and for organizing animal experiments. Thank you, Kim, for saving me from becoming homeless during my initial days in Wageningen. Most of my office time was spent in the zen office with Laura and Alex. No other office can beat the zen office atmosphere. Isabela has been a key collaborator on many projects. Investigadora Blanca has been important in increasing my wet-lab skills. Towards the end of my PhD, Maria has been my lab champion. Thank you to past and present HMI group members, for direct and indirect contributions; Tiantong, Zhan, Nuning, Berdien, Jinya, Bart, Jori, Michiel, Tim, Raka, Aïsha, Jeroen, Evelien, Maaike, Nadya, Linda, and Edo. Thank you to Loes for help with administrative matters, Joyce for taking care of the lab, and Ellen for help with ethical approval applications.

My life as a PhD student in the Netherlands has been good, although not how I imagined it. What I expected to be a mild experience with a related culture turned out to be an international jungle that drives ever increasing reverse culture shocks each time in travel back home. Although not as expected, I would not be without my years in the Wageningen (international) community. In particular during the times of covid restrictions I found a family away from home.

After my first month in Wageningen I moved into "Plantsoen", a run-down office building converted to student housing. The material conditions were sub-par, but we made it a home by the community of people living there. Some of the first people I met there and spent a lot of time with during my early times in Wageningen were Patrick and Stamatis.

Thank you for bike rides and cooking lessons. The following time period I met many people important to my life in Wageningen include Haymanti and Wouter, Jiakun, Andrea (x2), Antonio, Alvaro, Gil, Ana, Bea, Steina, and Dursun.

During early covid lockdown Wageningen emptied out and those of us left behind formed an even more tight knit community. In addition to the old core of Ahmad, Monica, and Stamatis, the community included Lucas, my training partner - after you left I my physical condition rapidly declined, vegan Helena, Ivan, Emmanouil, and Enrico the plant man. Thank you, Mr. Dong, for your Chinese tea parties. They usually featured oolong tea and cookies, but often also an (alleged) covid-protective tea variety. Thank you, Mirella, for waking me up in the morning with coffee countless times. Later on, our community was joined by more members including Caro and Pepe, Caixia, Kazem, Carmelo, and Amit, people that have all also been important to my life in Wageningen.

Threatened by rent increase I moved to exile in the "Firehouse", another Wageningen community of international PhD students sharing the struggle. Firstly, one reason to move there in was to live with office neighbour and kitchen gossip queen Elisa. Thank you, Antonio, for many dinners with sushi, chicken and aggressive voice messages scaring people into cleaning up their mess. Thank you, Fang, for sharing your lunch with me many times, usually starting cooking while I was still eating breakfast. Jiaqi and Melania are inspirations for every PhD student with their extreme work ethics. Berte is Berte, no further description is possible. Hugo provided south African braai culinary experiences. Chandan brought unlimited enthusiasm. The community would also not have been the same without Guichao the Sichuan pepper abuser, big Brecht, Dane the ladies man, Jesus, Jihane, Anne, Leo, Tina, Julia, Deepak, Daniel, Raghavendra, Ymke, Gulio, Elizabeth, Lisanne, Naraya, Claire, Hongrui, Annalisa, Nuran, BBQ Guy, Dimitri, Khalid, Charlotte, Ketan, Elien, Ludo, Weiyi, Tianyu, Canan, Omer, and Aegina.

Thank you to Natassa for your company, use of your heating blanket, and for many cakes and spicy meals.

Lastly, thank you to Elizabeth and Bjørn for proofreading this thesis.

About the author

Simen Fredriksen was born in Oslo, Norway on August 7th, 1992. After finishing his secondary education, he started studying mechanical engineering, but soon discovered that this was not his true path. He took the risky decision to switch to a BSc in Biology at the University of Oslo (2012-2015). Subsequently, he continued with a MSc in Ecology and Evolution at the same institution (2015-2017). His MSc thesis was on the microbiota in hybrid species system of sparrows. During this time, he gained experience in several aspects of molecular ecology research, including the fundamentals of bioinformatic and statistical analysis. He realized that life is more comfortable sitting in a chair than running around the lab or conducting fieldwork starting before sunrise. This motivated him to pursue a career in research focused on data analysis and bioinformatics.

In 2018 Simen moved to Wageningen to start his PhD in the Host-Microbe Interactomics chair group (HMI). In the 1658 days between starting and handing in his PhD thesis he worked on the biology of the opportunistic *Streptococcus suis*. Much of this work is presented in this thesis entitled "The piglet tonsillar microbiome and *Streptococcus suis* disease". During his time at HMI, he also contributed to several other projects in collaboration with partners from academia and industry. Simen has no long-term life plans, but we are sure that he has a bright future with many contributions to science.

By Maria Juanpere Borràs (paranymph)

Overview of completed training activities

The Basic Package (2 ECTS)
Introduction Day, WIAS, Wageningen, 2018
Scientific Integrity course, WGS, Wageningen, 2021
Ethics in Animal Sciences course, WGS, Wageningen, 2021

Disciplinary Competences (13 ECTS)
Literature survey, WIAS, 2018
International Summer School on Networks and Evolution, Sorbonne University, France, 2019
External period, IRTA-CReSA, Barcelona, Spain, 2019
Course: Metagenomics bioinformatics, EMBL-EBI, online, 2020
Course: Bayesian Statistics, PE&RC, Wageningen, 2020
Course: Advanced statistics, PE&RC, Wageningen, 2020
Course: Rmarkdown, WGS, Wageningen, 2021

Professional Competences (5 ECTS)
Course: Supervising BSc and MSc thesis students, ESD, Wageningen, 2021
Course: Efficient Writing Strategies, WGS, Wageningen, 2021
Course: Scientific Writing, WGS, Wageningen, 2021
Course: Scientific Publishing, WGS, Wageningen, 2021
Course: Reviewing a Scientific Manuscript, WGS, Wageningen, 2021
Course: Research Data Management, WGS, Wageningen, 2021
Course: The Final Touch, WIAS, Wageningen, 2021

Presentation Skills (4 ECTS)
Oral presentation, Topigs Norsvin Robustness and Animal Welfare Platform, Ås, Norway, 2018
Oral presentation, WIAS End-of-the-year event-Symposium on collaboration in clusters, 2018
Oral presentation, PIGSs Scientific Meeting 2019, Sitges, Spain
Oral presentation, WIAS Annual Conference 2020, Lunteren
Oral presentation, PIGSs Scientific Meeting 2021, Online
Oral presentation, PIGSs Scientific Meeting 2022, Sitges, Spain

Teaching competences (6 ECTS)
Supervising 2 MSc major thesis
Supervising 1 MSc minor thesis
Teaching assistant

The research described in this thesis was funded by a PhD studentship from the Netherlands Centre for One Health and by EU Horizon 2020 (grant agreement ID 727966).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design: Kristiina Vind
Printed by: ProefschriftMaken.nl