Towards selection on F4 enterotoxigenic *Escherichia coli* resistance in pigs for prevention of neonatal and post-weaning diarrhea

Tiphanie Goetstouwers

Promoter: Prof. Dr. L.J. Peelman

Co-promoter: Prof. Dr. E. Cox

Department of Nutrition, Genetics and Ethology

Faculty of Veterinary Medicine

Ghent University

Thesis submitted in order to obtain the degree of Doctor in Veterinary Science (PhD),
Faculty of Veterinary Medicine, Ghent University, 2014

Table of contents

LIST OF ABBREVIATIONS

5-HT	5-hydroxytryptamine
ACTB	actin, beta
ADP	adenosine diphosphate
AGP	antibiotic growth promoter
ANPEP	aminopeptidase N
ARSA	arylsulfatase A
B3GNT5	lactosylceramide 1,3-N-acetyl-beta-D-glucosaminyl transferase
B3GALT3	beta-1,3-galactosyltransferase 3
B4GALT4	beta-1,4-galactosyltransferase 4
B4GALT6	beta-1,4-galactosyltransferase 6
BAC	bacterial artificial chromosome
BBMV	brush border membrane vesicle
bp	base pair(s)
cAMP	cyclic adenosine monophosphate
cDNA	complementary DNA
Cer	ceramide
CFTR	cystic fibrosis transmembrane conductance regulator
cGMP	cyclic guanylate monophosphate
CMH	Cochran-Mantel-Haenszel
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
E. coli	*Escherichia coli*
ELISA	enzyme-linked immunosorbent assay
ETEC	enterotoxigenic *Escherichia coli*
F4R	F4 receptor
$F4R^-$	F4 receptor-negative
$F4R^+$	F4 receptor-positive
GAL3ST1	galactose-3-O-sulfotransferase 1
Gal	galactose
GALC	galactosylceramidase
GalNAc	*N*-acetylgalactosamine
GBA	glucosidase, beta
GCC	guanylyl cyclase C
GLA	galactosidase, alpha
GLB1	galactosidase, beta 1
GLB1L	galactosidase, beta 1-like
Glc	glucose
GlcNAc	*N*-acetylglucosamine
GM1	monosialotetrahexosylganglioside
GP74	74 kDa glycoprotein
GSL	glycosphingolipid
GTP	guanosine triphosphate
GWAS	genome-wide association study
HBD-1	β-defensin 1
HEG1	heart of glass 1
HexNAc	*N*-acetylhexosamine
IBS	identity-by-state

IMTGP	intestinal mucin-type sialoglycoprotein
KPAN1	karyopherin alpha 1
kb	kilo base pairs
kDa	kilo Dalton
KIAA0226	beclin-1 associated RUN domain containing protein
KLH	keyhole limpet hemocyanin
KPAN1	karyopherin alpha 1
LL-37	peptides cathelicidin
LMLN	leishmanolysin-like gene
LT	heat-labile enterotoxin
MAF	minor allele frequency
Mb	mega base pairs
MDS	multidimensional scaling
MUC4	mucin 4
MUC13	mucin 13
MUC20	mucin 20
MW	molecular weight
MYLK	myosin light chain kinase
NEU1	sialidase 1
NEU2	sialidase 2
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PGE2	prostaglandin E2
PTS	proline-threonine-serine
RNA	ribonucleic acid
RT	reverse transcription
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SLC12A8	solute carrier family 12 member 8
SNP	single nucleotide polymorphism
SO3	sulfur trioxide
SSC	Sus scrofa chromosome
ST	heat-stable enterotoxin
TRFC	transferrin receptor
UGCG	UDP-glucose ceramide glucosyltransferase
UGT8	UDP glycosyltransferase 8
UTR	untranslated region
ZDHHC19	zinc finger DHHC type containing 19 protein gene

CHAPTER 1

REVIEW OF THE LITERATURE

1. REVIEW OF THE LITERATURE

1.1. Introduction

Global pork production in carcass weight is 110.7 million tons and accounts for about 43.3 % of the meat production (beef, pig, and poultry) worldwide. After China, the European Union (EU) is the second biggest producer of pig meat with a yearly production of 22.3 million tons (USDA, 2014). Therefore, maintaining of animal health in the pig industry is very important in pig production. Neonatal diarrhea and post-weaning diarrhea caused by *Escherichia coli* (*E. coli*), in particular enterotoxigenic *E. coli* (ETEC), are among the most important economical diseases in pigs (Francis, 2002; USDA, 2002; Fairbrother et al., 2005; USDA, 2007). It has been estimated that in the United States (US) alone post-weaning diarrhea caused by ETEC cost $90 million annually due to the death of up to 5 % of young pigs (Zhang, 2007). In addition, other factors causing significant economic losses to the pig industry due to diarrhea caused by ETEC are reduced growth rates, morbidity, and cost of medication (Fairbrother et al., 2005).

Antimicrobial drugs have been widely used in human and veterinary medicine to treat and prevent bacterial infections and to improve production efficiency in food-producing animals (WHO, 1997; US FDA, 2012). The use of antimicrobial drugs as antibiotic growth promoters (AGPs) in animal diets has been documented to improve growth performance and to reduce morbidity and mortality by approximately 50 % in pigs. Although the exact mechanisms by which AGPs influence the performance of the pig are still unclear, the beneficial effects of AGPs are probably due to the suppression of pathogenic bacteria in the digestive tract, increased feed utilization and stimulation of metabolic processes (Kil and Stein, 2010). The use of AGPs increases the prevalence of antimicrobial resistance in farm animals and the risk that resistant bacteria or resistance determinants might be transferred from animals to humans (WHO, 1997). Antimicrobial resistance within a wide range of infectious agents is a growing public health threat of broad concern to public health (WHO, 2014). Based on the current available data, more than two million people are every year infected with antimicrobial-resistant pathogens in the US, with at least 23,000 deaths as a result (CDC, 2013).

The use of antibiotics as growth promoters in animal production has been totally banned within the EU in 2006 (EC Regulation 1831/2003[1]). Some countries outside of Europe also banned the use of antimicrobials for growth promotion, but in many countries, such as the US and Canada, antimicrobials are still approved for production purposes like growth promotion (Maron et al., 2013). However, the FDA has recommended voluntary guidelines limiting the use of medically important antimicrobial drugs in food-producing animals to those uses that are considered necessary for assuring animal health and that include veterinary oversight or consultation (US FDA, 2012). The discontinuation of AGP use in the EU has led to a reduced pig performance and an increased morbidity and mortality rate among pigs, mostly associated with enteric infections. This has substantially increased the use of therapeutic antibiotics in pigs in Europe (Casewell et al., 2003).

Colistin, a polymyxin antibiotic, has been used for decades in veterinary medicine, especially in swine to treat diarrhea caused by *E. coli* and is the most frequent oral administered antimicrobial drug (30.7 %) used in the Belgian pig industry (Callens et al., 2012; EMA, 2013). In human, colistin is used as a last resort drug for nosocomial infections caused by the following multiresistant Gram-negative bacteria: carbapenemase-producing *Enterobacteriaceae* (*E. coli*, *Klebsiella*), and multidrug-resistant *Pseudomonas* and *Acinetobacter* species (non-fermenters) (EMA, 2013). Although antimicrobial resistance prevalence to ampicilin, sulphomethoxazole, tetracycline, and trimethoprim in porcine *E. coli* strains can exceed 50%, the antimicrobial resistance prevalence to colistin in porcine *E. coli* strains is still low (9.6 %) (Boyen et al., 2010; Habrun et al. 2010; Morales et al., 2012; Chantziaras et al., 2014). Therefore, the resistance to colistin should be closely monitored. Also, the occurrence of multi-drug resistant porcine *E. coli* strains has rapidly increased in recent years (Van Driessche et al., 1995; Maynard et al., 2003; Boerlin et al., 2005; Fairbrother et al., 2005; Costa et al., 2010; Li et al., 2012). Recent studies reported that approximately 87% of the porcine *E. coli* isolates were resistant to 4 or more antimicrobials (Lee et al., 2009; Habrun et al., 2010). Therefore, alternative strategies effectively reducing diarrhea caused by *E. coli* in pigs are urgently needed.

[1] EC Regulation 1831/2003. European Union legislation on feed additives: http://ec.europa.eu/food/food/animalnutrition/feedadditives/legisl_en.htm

1.2. Neonatal and post-weaning diarrhea

Neonatal diarrhea usually occurs during the first three to five days of life, while post-weaning diarrhea (PWD) is commonly observed at three to ten days after weaning (Alexander, 1994; Hampson, 1994).

A survey conducted by the National Animal Health Monitoring System (NAHMS) in the US found diarrhea to be the most common cause of pre-weaning morbidity, 42 % of which occurred during the first three days after birth, and an important cause of pre-weaning mortality (10.8 %) (National swine survey, 1992; Tubbs et al., 1993). The most common agent associated with neonatal diarrhea as well as PWD is enterotoxigenic *Escherichia coli* (ETEC) (Fairbrother et al., 2005; Fairbrother, 2006; Jackson and Cockcroft, 2007). For PWD caused by ETEC, the mortality rate is often 1.5 to 2 % and can be up to 25 % when no treatment is given whereas the morbidity rate may be over 50 % among weaned piglets (Hampson, 1994; Fairbrother and Gyles, 2006; Jackson and Cockcroft, 2007). The clinical signs of ETEC are watery diarrhea, dehydration, emaciation and death (Alexander, 1994; Hampson, 1994).

In addition to ETEC, neonatal diarrhea and PWD can also be caused by other microorganisms. Transmissible gastroenteritis virus, *Clostridium perfringens* type C, porcine epidemic diarrhea virus, rotavirus A and *Cystoisospora suis* (previously known as *Isospora suis*) are known to be associated with diarrhea during the neonatal period (Edfors-Lilja and Wallgren, 2000; Kongsted, 2014). At weaning age, *Salmonella* spp., *Lawsonia intracellularis*, *Trichuris suis*, *Brachyspira* spp., rotaviruses and coronaviruses are other common causes of diarrhea in pigs (Edfors-Lilja and Wallgren, 2000; Jackson and Cockcroft, 2007; Hopwood and Hampson, 2013).

1.3. *Escherichia coli*

ETEC are the leading cause of diarrhea in newborn and recently weaned pigs causing great economic loss due to decreased growth rate and considerable morbidity and mortality (Alexander, 1994; USDA, 2002; Flores, 2004; Fairbrother et al., 2005; USDA, 2007). The group of pathogenic *E. coli* causing diarrheal diseases is referred to as diarrheagenic *E.coli* and based on their specific virulence factors and phenotypic traits are divided into six pathotypes: (i) enteropathogenic *E. coli* (EPEC), (ii) enterotoxigenic *E. coli* (ETEC), (iii) verocytotoxigenic *E. coli* (VTEC), which is synonymous with the term verocytotoxic or Shiga

toxin-producing *E. coli* (STEC), and also includes the enterohaemorrhagic *E. coli* (EHEC) category; (iv) enteroinvasive *E. coli* (EIEC); (v) enteroaggregative *E. coli* (EAEC); and (vi) diffusely adherent *E. coli* (DAEC) (Nataro and Kaper, 1998; Clements et al., 2012). The ETEC, VTEC and EPEC strains represent the three major classes of enteric pathogens leading to diarrhea in pigs (Chin and Chapman, 2009). ETEC represent about 65 to 70 % of the *E. coli* strains causing diarrhea (Martins et al., 2000; Cheng et al. 2006; Vu-Khac et al., 2007; Alustiza et al., 2012).

ETEC are characterized by their ability to proliferate in the small intestine and by the production of enterotoxins. The adhesion of ETEC to the small intestinal epithelial cells is mediated by fimbrial adhesins and allows the bacteria to resist the flushing action of intestinal peristalsis (Gaastra and de Graaf, 1982; Nagy and Fekete, 1999). Five fimbrial types, F4 (formerly K88), F5 (K99), F6 (987P), F18 (F107) and F41, have been characterized in porcine ETEC strains (Nagy and Fekete, 1999; Vu-Khac et al., 2007). Two types of enterotoxins can be distinguished on the basis of their thermostability, namely the heat-labile (LT) and the heat-stable (ST) enterotoxins. These enterotoxins stimulate the epithelial cells of the small intestine to secrete fluid into the lumen of the gut causing diarrhea (Gaastra and de Graaf, 1982; Nagy and Fekete, 1999). Recent studies have indicated that enterotoxins, such as LT, also enhance enteric bacterial colonization (Berberov et al., 2004; Zhang et al., 2006; Glenn et al., 2009). Some F18 ETEC strains inducing post-weaning diarrhea produce both enterotoxins and verotoxin-2e, also named Shiga toxin-2e, and are referred to as ETEC/VTEC (Bertschinger and Gyles, 1994; DebRoy et al., 1999; Nagy and Feteke, 1999). This verotoxin-2e (edema disease toxin) damages the vascular endothelium of the digestive tract, subcutis and brain, which leads to subcutaneous edema (MacLeod et al., 1991; Wadell et al., 1997). Neurological signs occur shortly after subcutaneous edema is apparent and consist of mental confusion, ataxia and a staggering gait. Finally, pigs are unable to rise and display a range of signs including tremors, paddling of the limbs, extensor rigidity, convulsions and coma (Gannon et al., 1989).

1.4. F4 enterotoxigenic *Escherichia coli*

ETEC strains expressing F4 fimbriae are currently worldwide associated with neonatal and post-weaning diarrhea in pigs (Francis, 2002; Frydendahl, 2002; Do et al., 2005; Do et al., 2006; Vu-Khac et al., 2007; Zhang et al., 2007; Amezcua et al., 2008; Madoroba et al., 2009; Vidotto et al., 2009; De la Fe Rodriguez et al., 2011; Byun et al., 2013). In *E. coli* isolates from pigs with neonatal diarrhea, the prevalence of F4 fimbriae was 53.1 % in North Vietnam (Do et al., 2006), 31.9 % in Korea (Byun et al., 2013), and 28.4 % in Zimbabwe (Madoroba et al., 2009). In Belgium, no recent data are available on the prevalence of F4 fimbrial *E. coli* isolates from pigs with neonatal diarrhea. The prevalence of F4 fimbriae in *E. coli* isolates from pigs with PWD was 60 % in Hungary (Nagy et al., 1990), 50.5 % in Switzerland (Sarrazin et al., 2000; Rampoldi, 2013), 64.4 % in US (Zhang et al., 2007), 44 % in Brazil (Vidotto et al., 2009), 53 % in Belgium (data unpublished) and 19 % in Slovakia (Vu-Khac et al., 2007).

F4 ETEC do not adhere to the brush border from all piglets. Pigs that showed F4 ETEC adhesion to their brush borders were designated as F4 ETEC susceptible and pigs with no F4 ETEC adhesion were designated as F4 ETEC resistant (Figure 1.1) (Rutter et al., 1975; Sellwood et al., 1975). The two phenotypes, susceptible and resistant, seem to be inherited in a simple Mendelian manner (Sellwood et al., 1975; Gibbons et al., 1977).

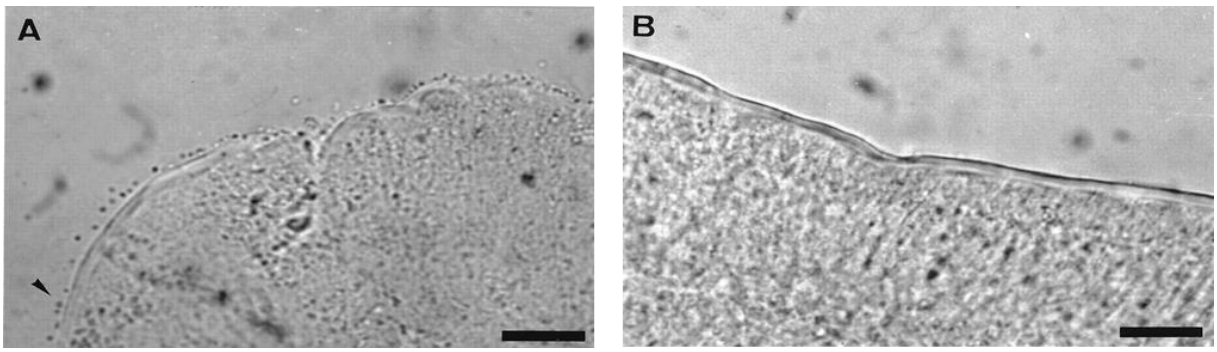

Figure 1.1: Small intestinal villous brush borders after the *in vitro* adhesion assay with F4ac ETEC. (A) F4 receptor-positive ($F4R^+$) brush border with strong adhesion; (B) F4 receptor-negative ($F4R^-$) brush border without adhesion. Bars, 10 µm. *From Van den Broeck et al. (1999a).*

Even though infection with ETEC can occur at any age, the susceptibility towards F4 ETEC appears to be age-dependent resulting in a high susceptibility in the neonatal stage and immediately after weaning (day 21-28) (Jones and Rutter, 1972; Wilson and Hohman, 1974; Willemsen and de Graaf, 1992; Alexander, 1994; Hampson, 1994). This age-dependent susceptibility is due to a number of factors. First, the porcine mucosal immune system is functionally immature at birth and develops gradually during the first six weeks of life (Stokes et al., 2004; Bailey and Haverson, 2006). Therefore, the period from birth to weaning is a critical time. Also, the low levels of digestive enzymes and the high pH levels in the stomach and small intestine in the newborn piglets provides a favorable environment for bacteria (Alexander, 1994). Changes in the small intestine structure and enterocyte brush border enzyme activities immediately after weaning and the withdrawal of the passive lactogenic protection are important predisposing factors related to PWD (van Beers-Schreurs et al., 1992; Hampson, 1994). Furthermore, an age-dependent expression of F4 receptors in the mucus can play a role in an increased resistance towards F4 ETEC in weaned pigs (Conway et al., 1990).

Neonatal piglets can be passively protected against infection with F4 ETEC by pathogen-specific antibodies in colostrum and milk of the sow (Gaastra and de Graaf, 1982; Sellwood, 1982). However, if either intake or quality is too low then F4 ETEC can still adhere to the small intestine and cause diarrheal illness (Alexander, 1994). Because resistant sows are not susceptible towards F4 ETEC infection, no F4-specific antibodies are produced and secreted in their colostrum and milk (Sellwood, 1982). Therefore, heterozygous susceptible offspring of F4 ETEC resistant sows are highly susceptible to F4 ETEC (Alexander, 1994). However, the passive maternal protection can be obtained and increased by vaccinating the sows during pregnancy, protecting the piglets against F4 ETEC infection during the suckling period (Rutter et al., 1976; Fürer et al., 1982). At weaning, lactogenic immunity disappears and an active mucosal immunity would be required in order to protect the piglets against F4 ETEC. By vaccinating the piglets via parenteral route, no mucosal protection can be obtained, because parenteral vaccines stimulate in general the systemic rather than the mucosal immune system. Oral vaccination has shown to induce protective intestinal mucosal immunity against F4 ETEC (Bianchi et al., 1996; Van den Broeck et al., 1999b; Melkebeek et al., 2013). However, this vaccination should occur during the suckling period to obtain protection at weaning when piglets become infected, because PWD generally occurs three to ten days after weaning (Hampson, 1994; Melkebeek et al., 2013). Unfortunately, no commercial oral

vaccines are currently available that can overcome neutralization by maternal milk antibodies and/or other milk factors that can interfere with vaccination (Melkebeek et al., 2013).

Besides vaccination, other prevention strategies for F4 ETEC diarrhea are improving the management in intensive piggeries and improving the intestinal microenvironment and resistance by using probiotics, prebiotics, acidifiers, neutraceuticals, plant extracts and antimicrobial peptides in the diet (Alexander, 1994; Fairbrother et al., 2005; Chin et al., 2009; Vondruskova et al., 2010; Thacker, 2013). The supplementation of piglet diets with high levels of zinc oxide (3000 mg/kg) also reduces the incidence and severity of post-weaning diarrhea caused by ETEC and improves the growth performance of the pig (Hill et al., 2001; Roselli et al., 2003; Fairbrother et al., 2005). However, the application of pig manure containing high levels of zinc to soil can negatively impact the environment. Therefore, alternatives with low levels of zinc oxide are currently being explored (Kim et al., 2010; Wang et al., 2012; Kim et al., 2015).

Another strategy is selecting and breeding pigs for hereditary resistance towards F4 ETEC using genetic markers associated with the F4 receptor locus (Sellwood et al., 1975; Jørgensen et al., 2003). It should be noted that selection towards F4 ETEC resistance might negatively affect genetic improvement in traits of economic importance, thus caution should be exercised. Contrary to the report of Baker et al. (1997), Yan et al. (2009a) indicated that pigs expressing the F4ab receptor, the F4ac receptor or both, exhibit desirable production traits by showing a higher performance during the fattening period than pigs non-adhesive for F4ab/ac ETEC. This observation is in agreement with previous research (Edfors-Lilja et al., 1986). The influence of the F4ad receptor on growth was also evaluated. Although pigs expressing the F4ad receptor showed poorer performance during the fatting period than pigs without the F4ad receptor, other unknown beneficial traits of economic importance might be correlated with the presence of the F4ad receptor (Yan et al., 2009a).

Treatment of F4 ETEC diarrhea is usually based on the administration of water, electrolytes and antimicrobials. Electrolyte therapy is given to restore the ionic imbalance, to prevent dehydration and to maintain the body condition (Alexander, 1994). Antibiotics have been commonly used in the treatment of F4 ETEC, but their intensive use has resulted in rapid increase in the percentage of multi-drug resistant porcine ETEC strains (Van Driessche et al., 1995; Maynard et al., 2003; Boerlin et al., 2005; Fairbrother et al., 2005; Lee et al., 2009; Costa et al., 2010; Li et al., 2012).

1.4.1. Virulence factors

The pathogenicity of F4 ETEC is determined by 2 virulence factors, namely F4 fimbriae and enterotoxins (Gaastra and de Graaf, 1982). F4 fimbriae are proteinaceous surface appendages mediating the attachment of F4 ETEC to the small intestine. By adhering to the small intestine, F4 ETEC can overcome the flushing effect of small intestinal peristalsis, facilitating colonization of the small intestine (Gaastra and de Graaf, 1982; Smyth et al., 1994). The enterotoxins, heat-labile (LT) and heat-stable (ST) enterotoxins produced by F4 ETEC are responsible for the diarrhea in pigs by decreasing absorption and increasing secretion of fluids and electrolytes in the small intestine (Nagy and Fekete, 1999).

1.4.1.1. F4 fimbriae

Fimbriae are long, nonflagellar, filamentous, threadlike protein polymers found on the surface of many strains of *Escherichia coli* (Figure 1.2.A). The ability of ETEC to adhere to the small intestine by fimbriae is an important initial factor in the establishment of diarrheal disease. Adhesion to the host tissue allows the bacteria to resist the flushing action of intestinal peristalsis and to colonize the intestinal epithelium *in vivo* (Jones and Rutter, 1972; Gaastra and de Graaf, 1982; Klemm, 1985). The F4 fimbriae, composed of major (FaeG) and minor (FaeF, FaeH, FaeC, and FaeI) subunit structures, exist in three serological variants, namely F4ab, F4ac and F4ad (Figure 1.2.B) (Orskov et al., 1964; Guinee and Jansen, 1979; Van den Broeck et al., 2000). These variants differ in the amino acid composition of the major fimbrial subunit FaeG, which contains conserved regions (a) and variable regions (b, c, or d) (Gaastra et al., 1981; Gaastra et al., 1983; Josephsen et al., 1984; Van den Broeck et al., 2000). The FaeG subunit has adhesive properties and recognizes glycoconjugates on the surface of enterocytes (Sellwood, 1980; Bakker et al., 1992; Van den Broeck et al., 1999c). Both conserved and variable regions of the FaeG subunit contribute to the receptor binding site of F4 ETEC, but everything indicates that the variable regions (b, c, or d) of the FaeG subunit involved in the receptor binding site explains the different F4 receptor profiles observed in pigs (Bakker et al., 1992; Sun et al., 2000; Van den Broeck et al., 2000; Zhang et al., 2009).

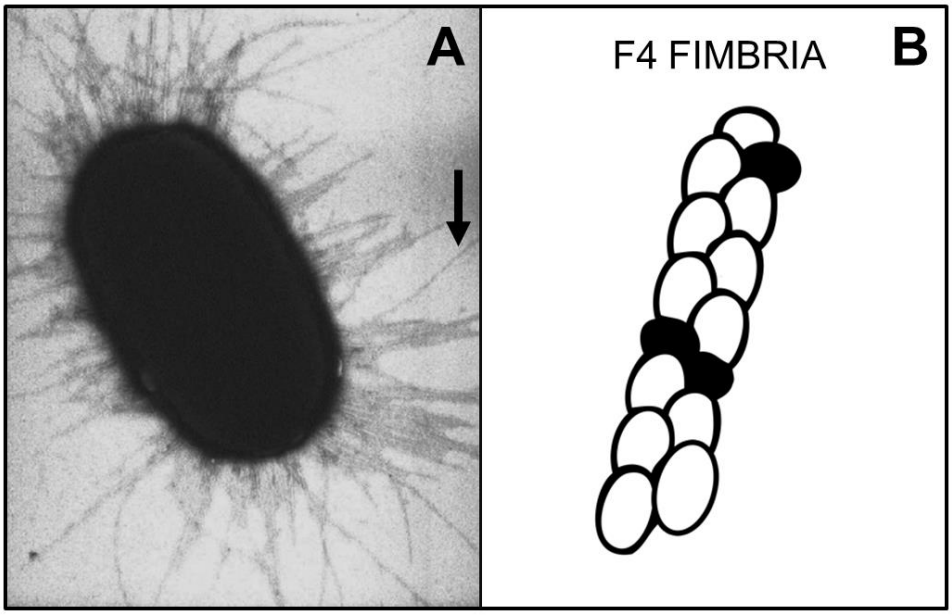

Figure 1.2: (A) *Escherichia coli* with fimbriae. The arrow indicates a fimbria. (B) Schematic presentation of the F4 fimbria. The dark bulbs indicate the minor subunits and the white bulbs indicate the major subunits. *Adapted from Gross (2006) and Van den Broeck et al. (2000).*

1.4.1.2. F4 enterotoxins

Two types of enterotoxins can be distinguished on the basis of their thermostability, namely the heat-stable (ST) and the heat-labile (LT) enterotoxins.

Two types of heat-stable (ST) enterotoxins have been described, namely STa and STb (Burgess et al., 1978). STa toxins produced by human and animal strains of ETEC seem to differ in some amino acids (Moseley et al., 1983) and are labelled in literature as STaH and STaP. Human ETEC may produce both STaH and STaP, while ETEC from calves and pigs produce STaP. STaP will be named STa in this thesis. F4 ETEC strains can produce STa, STb, LT, or a combination of enterotoxins (Gaastra and de Graaf, 1982; Cox et al., unpublished results). Most commonly F4 ETEC strains produce both LT and STb (Gyles, 1971; Moon et al., 1986; Nagy et al., 1990; Cox et al., unpublished results) where LT is a more significant contributor to diarrhea in very young piglets than STb (Casey et al., 1998; Berberov et al., 2004; Zhang et al., 2006; Erume et al., 2008). However, STb is the major contributor to fluid loss and inflammation in the first hours after infection of newly-weaned piglets with an F4 ETEC STa, STb and LT producing strain (Loos et al., 2012). STa is less frequently identified in ETEC F4 isolates from piglets with PWD.

STa and STb differ in methanol solubility and in mechanisms of action (Burgess et al., 1978; Kennedy et al., 1984). Through binding of STa to the extracellular domain of the membrane-associated guanylyl cyclase C (GCC) in the pig intestine, cytoplasmic cyclic guanylate monophosphate (cGMP) is produced (Katwa et al., 1991; Wada et al., 1996). Increased levels of cGMP lead to a stimulation of chloride secretion and an inhibition of sodium re-absorption, subsequently causing water release into the intestinal lumen (Weiglmeier et al., 2010). STa also stimulates the chloride and bicarbonate secretion via an alternative non-GCC receptor in a cystic fibrosis transmembrane conductance regulator (CFTR)-independent manner (Sellers et al., 2008). Unlike LT and STa toxins, there is no evidence of an association between STb and an intracellular cyclic nucleotide elevation (Kennedy et al., 1984; Hitotsubashi et al., 1992; Peterson and Whipp, 1995). STb, a bacterial pore-forming toxin with membrane-permeabilization properties, binds to a functional receptor, glycosphingolipid sulfatide, present on the epithelial cell surface of the pig jejunum (Rousset et al., 1998; Gonçalves et al., 2007). After internalization, this toxin leads to opening the guanosine triphosphate (GTP)-binding protein-dependent calcium channels in the plasma membrane, resulting in an influx of extracellular calcium in the cell and activating calmodulin-dependent protein kinase II (Dreyfus et al., 1993; Fujii et al., 1997; Labrie et al., 2002). The activation of calmodulin-dependent protein kinase II may stimulate and open plasma membrane chloride channels, leading to the secretion of water and electrolytes (Fujii et al., 1997). In addition, the calcium influx in the cell seems to activate production of prostaglandin E2 (PGE2) and serotonin (5-hydroxytryptamine [5-HT]), leading to intestinal secretion and increasing the intestinal motility (Hitotsubashi et al., 1992; Harville and Dreyfus, 1995; Peterson and Whipp, 1995).

The heat-labile enterotoxins (LT) can be divided in two antigenic groups, LTI and LTII, based on the reactivity to antisera against cholera toxin (Gilligan et al., 1983; Green et al., 1983; Pickett et al., 1986). Although LTII-producing *E. coli* have been isolated in pigs, no F4 adhesins were until now present in these isolates (Celemin et al., 1994; Celemin et al., 1995; Casey et al., 2012). The porcine heat-labile enterotoxin (LTI) is composed of one A subunit and five identical B subunits (Nataro and Kaper, 1998). The B subunits are arranged in a ring and preferentially bind to ganglioside GM1 (monosialotetrahexosylganglioside), also a known receptor for the cholera toxin (Holmgren et al., 1975; Nataro and Kaper, 1998; Grange et al., 2006). The A subunit is responsible for activating adenylate cyclase through catalyzation of the adenosine diphosphate (ADP) ribosylation of a regulatory component of the GTP-binding protein. This constant activation of adenylate cyclase increases the intracellular cyclic

adenosine monophosphate (cAMP) levels (Chang et al., 1987; Nataro and Kaper, 1998). Subsequent cAMP-dependent protein kinase A is activated which phosphorylates the R domain of the cystic fibrosis transmembrane conductance regulator (CFTR), resulting a Cl- and water efflux into the intestinal lumen (Nataro and Kaper, 1998). In addition, recent studies have indicated that enterotoxins, such as LT, promote the colonization of the intestinal epithelium by ETEC strains (Berberov et al., 2004; Zhang et al., 2006; Glenn et al., 2009). This colonization enhancement is primarily mediated by the increased intracellular cAMP levels caused by ADP ribosylation activation (Johnson et al., 2009). This suggests that the cAMP activation of the enterocyte contributes to an important survival advantage on ETEC (Glenn et al., 2009). Glenn et al. (2009) proposed that the toxin-mediated fluid secretion caused by LT disrupts the protective mucosal barrier by hydrating and washing away the mucus, leaving the intestinal surface vulnerable to colonization (Figure 1.3). In parallel, the antimicrobial peptides cathelicidin (LL-37) and β-defensin 1 (HBD-1) are downregulated by LT. These antimicrobial peptides form the first line of host defense against infections over the epithelial surfaces and play a critical role in the immune responses of the mucosal surface (Chakraborty et al., 2008).

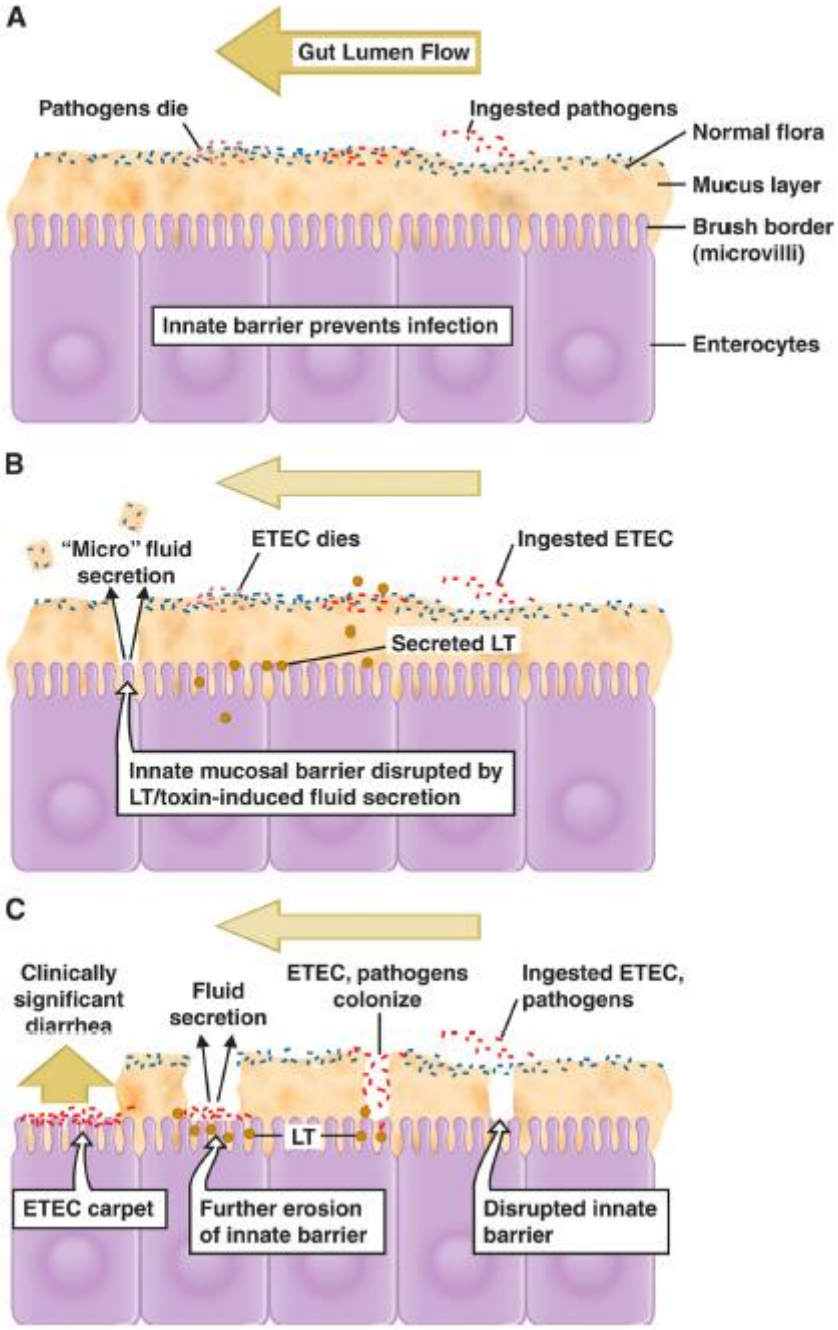

Figure 1.3: Proposed model of toxin-mediated effects on the innate mucosal barrier. (A) The healthy enterocyte brush border is coated with gel mucus secreted by goblet cells containing sIgA, defensins, cell surface mucins, and normal flora. Ingested pathogens generally do not survive in this environment. (B) ETEC is ingested and secretes heat-labile toxin (LT). Toxin penetrates the innate mucosal barrier to the microvillar space, and subsequent enterocyte activation and fluid secretion result in areas of barrier disruption. (C) The disrupted barrier renders the intestine susceptible to pathogen attachment, leading to colonization, pathogen proliferation, and clinically apparent infection. *From Glenn et al. (2009).*

1.4.2. F4 receptor in pigs

In addition to adhesive and enterotoxic virulence factors of F4 ETEC, adhesion of F4 ETEC to specific receptors on the brush borders of epithelial cells in the porcine small intestine is a key step in the pathogenesis (Sellwood, 1982; Nagy and Fekete, 1999). Presence of these receptors in the brush borders is age-independent. They are equally present at the age of one week, five weeks and six months (Willemsen and de Graaf, 1992). This observation is supported by previous results from Bijlsma et al. (1985) showing that the F4 receptors on brush borders are present in slaughter pigs (sows and boars). Also, the results of a recent study (Rampoldi et al., 2014) confirmed the presence of F4 receptors on isolated enterocytes in pigs between one and five years of age. However, not all piglets are susceptible to infection with F4 ETEC. Certain pigs do not express the F4-specific receptors on the intestinal epithelial cells and are therefore resistant to F4 ETEC infection (Gibbons et al., 1977; Sellwood, 1982).

The distribution of F4-specific receptors differs along the small intestinal tract. All 3 variants of F4 ETEC adhered significantly more to the jejunal villi than to the duodenal and ileal villi in F4 ETEC susceptible pigs (Cox and Houvenaghel, 1993). Subsequently, Chandler et al. (1994) reported that the porcine mid-small intestine contained the highest amounts of F4 receptor. In the proximal and distal part of the small intestine, the F4 receptors were found in lower amounts and were entirely absent in the caecum and colon of susceptible pigs.

Enteric pathogens must overcome the innate mucosal barrier to cause disease (Glenn et al., 2009). Therefore, F4 ETEC must pass through the mucus layer functioning as a dynamic defensive barrier in order to bind to the F4 receptors present on brush borders of villous enterocytes (Wilson and Hohmann, 1974; Sellwood et al., 1975; Sellwood, 1984a; Conway et al., 1990; Deplancke and Gaskins, 2001). This protective mucus layer covers the gastrointestinal epithelium and is composed predominantly of mucin glycoproteins that are synthesized and secreted by goblet cells (Deplancke and Gaskins, 2001). Due to enterocyte turnover, F4 receptors are released into the mucus. The presence of F4 receptors in mucus is age-dependent. The amount of F4 receptors in the mucus of 35-day-old unweaned susceptible pigs is 16 times higher than that in newborn susceptible piglets. One-week-old and 35-days-old post-weaning piglets were shown to contain F4 receptors in their mucus while these receptors were hardly detectable in the mucus of six-month-old pigs (Willemsen and de

Graaf, 1992). It appears that high concentrations of F4 receptors in mucus are protective by preventing F4 ETEC adhesion to the underlying epithelium, whereas the concentrations of F4 receptor in newborn ileal mucus are insufficient to prevent F4 ETEC adhesion (Conway et al., 1990). As described above, it has also been suggested that the innate mucosal barrier can be disrupted by the toxin effects of LT, resulting in an enhanced susceptibility of the underlying epithelium to F4 ETEC adhesion (Figure 1.3) (Glenn et al., 2009).

1.4.2.1. F4R inheritance

The F4ab/ac ETEC adhesion phenotype is inherited as an autosomal dominant trait in a simple Mendelian manner, involving a single locus with two alleles (Sellwood et al., 1975; Gibbons et al., 1977). Some studies suggest that one locus is controlling both F4ab and F4ac ETEC susceptibility (Bijlsma and Bouw, 1987; Python et al., 2002; Jørgensen et al., 2003), while others suggest two linked but distinct loci located on sus scrofa chromosome 13 (SSC13) (Guérin et al., 1993; Edfors-Lilja et al., 1995; Li et al., 2007a; Wang et al., 2007; Zhang et al., 2008; Yan et al., 2009b). For F4ad ETEC, a separate locus independent from the F4ab/acR locus is responsible for the F4ad adhesion phenotype (Bijlsma and Bouw, 1987). Unlike the F4ab/acR locus, the F4adR locus is not localized on SSC13 (Peelman, 1999; Python, 2003).

Besides the strong adhesive phenotype and the non-adhesive phenotype, another type of adhesion, the weak adhesive phenotype was observed for Fab, F4ac and F4ad ETEC (Sellwood, 1980; Sellwood, 1984b; Hu et al., 1993; Baker et al., 1997; Li et al., 2007a; Yan et al., 2009b). The inheritance of this weak adhesion phenotype remains unclear. However, it was suggested that both strong and weak adhesion to F4ab, F4ac and F4ad was controlled by one receptor, but that this weak adhesion arises from the influence of epistatic genes on the receptor expression or from post-expression inhibition or modification on the receptor sites (Bijlsma and Bouw, 1987; Python et al., 2005; Li et al., 2007a). For F4ad ETEC, a recent study confirmed the existence of more than one gene controlling the F4ad receptor based on segregation analysis (Rampoldi et al., 2014). In addition, Hu et al. (1993) suggested that, unlike the high affinity receptor co-segregating with F4ab/acR, the low affinity receptor associated with the adhesion phenotype D ($F4adR^+$) is only expressed in pigs under 16 weeks of age. This hypothesis was rejected by Rampoldi et al. (2014) who showed that the low affinity receptor present on enterocytes is expressed independent of age.

1.4.2.2. Determination of the F4 phenotype

Different F4 ETEC adhesion phenotypes have been identified in pigs based on the observed adhesion to the intestinal brush borders (Table 1.1). Bijlsma et al. (1982) identified five F4 ETEC adhesion phenotypes, namely phenotype A ($F4abR^{+}/F4acR^{+}/F4adR^{+}$), phenotype B ($F4abR^{+}/F4acR^{+}$), phenotype C ($F4abR^{+}/adR^{+}$), phenotype D ($F4adR^{+}$), and phenotype E ($F4abR^{-}/F4acR^{-}/F4adR^{-}$). Then, a sixth phenotype, namely phenotype F ($F4abR^{+}$), was described by Baker et al. (1997). The additional and infrequent phenotypes, phenotype G ($F4acR^{+}$) and H ($F4acR^{+}/F4adR^{+}$), were mainly observed in eastern breeds (Bonneau et al., 1990, Li et al., 2007a; Yan et al., 2009b).

Table 1.1: Summary of F4 ETEC adhesion phenotypes.

PHENOTYPE[1]	ADHESION[2]		
	F4ab	**F4ac**	**F4ad**
A	●	●	●
A1	FA	FA	FA[3]
A2	FA	FA	PA
B	●	●	-
B1	FA	FA	-
C	●	-	●
C1	PA	-	FA[3]
C2	PA	-	PA
D	-	-	●
D1	-	-	FA[3]
D2	-	-	PA
E	-	-	-
F	●	-	-
G	-	●	-
H	-	●	●

[1]Phenotype designations in grey are according to Bijlsma et al. (1982), Baker et al. (1997), Bonneau et al. (1990), Li et al. (2007a) and Yan et al. (2009b). Phenotype designations in white are according to Rampoldi et al. (2014).
[2]Adhesion in grey is defined as follows: '●' denotes adhesion, '-' denotes no adhesion. Adhesion in white is defined as follows: 'FA' denotes fully adhesive (>85 % adhesive enterocytes), 'PA' denotes partially adhesive (>0 % to 85 %), '-' denotes no adhesion.
[3]$F4adR^{FA}$ phenotype masks the expression of the $F4adR^{PA}$ phenotype (Rampoldi et al., 2014).

Based on the weak adhesion to F4ab ETEC and F4ad ETEC, Rampoldi et al. (2014) further divided the phenotypes A to D into phenotype A1 ($F4abR^{FA}/F4acR^{FA}/F4adR^{FA}$), phenotype A2 ($F4abR^{FA}/F4acR^{FA}/F4adR^{PA}$), phenotype B1 ($F4abR^{FA}/F4acR^{FA}/F4adR^{-}$), phenotype C1 ($F4abR^{PA}/F4acR^{-}/F4adR^{FA}$), phenotype C2 ($F4abR^{PA}/F4acR^{-}/F4adR^{PA}$),

phenotype D1 (F4abR$^-$/F4acR$^-$/F4adR^{FA}), and D2 (F4abR$^-$/F4acR$^-$/F4adR^{PA}) with 'FA' as the fully adhesive receptor and 'PA' as the partially adhesive receptor.

The distribution of F4 ETEC adhesion phenotypes seems to be breed specific. In general, phenotypes A, B and E are common adhesion phenotypes in European pig breeds, phenotype A and E in American breeds and phenotype A, D and E in Chinese pig breeds (Table 1.2) (Baker et al., 1997; Li et al., 2007a; Joller, 2009; Yan et al., 2009b; Rampoldi et al., 2014). Within the European breeds, Large White and Landrace breeds showed a similar distribution of the F4 ETEC adhesion phenotypes with phenotype A, B and E as the most prevalent phenotypes (Li et al., 2007; Yan et al., 2009b).

It can be concluded that Chinese breeds show a lower susceptibility rate toward F4ac ETEC (27.78 - 33.78 %) than European (37.86 - 79.59 %) or American breeds (36.69 - 47.91 %). It has also been reported that the Chinese breeds, such as Meishan and Fengjing, are resistant to F4ac ETEC (Michaels et al., 1994).

Table 1.2: Distribution of F4 ETEC adhesion phenotypes among western and Chinese breeds.

REF.[1]	BREEDS[2]		PHENOTYPE								TOTAL
			A	B	C	D	E	F	G	H	
[1]	American breeds (CW, D, H, Y)	No. pigs	41	5	6	11	27	6	0	0	96
		%	42.7	5.21	6.25	11.46	28.13	6.25	0	0	100
[2]	European breeds (LW, LWxLR)	No. pigs	299	31	26	37	96	0	0	0	489
		%	61.20	6.30	5.30	7.60	19.6	0	0	0	100
[3]	European breeds (LW, LR)	No. pigs	27	50	0	3	17	0	1	0	98
		%	27.55	51.02	0	3.06	17.35	0	1.02	0	100
	American breed (D)	No. pigs	14	2	1	0	25	3	1	0	46
		%	30.44	4.35	2.17	0	54.35	6.52	2.17	0	100
	12 Chinese indigenous breeds	No. pigs	21	20	5	21	70	2	3	6	148
		%	14.19	13.51	3.38	14.19	47.3	1.35	2.03	4.05	100
[4]	European breeds (LW, LR, LWxLR)	No. pigs	367	227	197	235	493	50	0	0	1569
		%	23.39	14.47	12.55	14.98	31.42	3.19	0	0	100
[5]	European breeds (LW, LR)	No. pigs	125	39	11	32	48	9	9	3	276
		%	45.29	14.13	3.99	11.59	17.39	3.26	3.26	1.09	100
	Chinese breed (Songliao Black)	No. pigs	12	8	5	26	26	8	0	5	90
		%	13.33	8.89	5.56	28.89	28.89	8.89	0	5.56	100

[1][1] Baker et al. (1997), [2] Rampoldi et al. (2014), [3] Yan et al. (2009b), [4] Joller (2009) and [5] Li et al. (2007a).

[2]Breeds are defined as follows: 'CW' denotes Chester White, 'D' denotes Duroc, 'H' denotes Hampshire, 'Y' denotes Yorkshire, 'LW' denotes Large White, 'LR' denotes Landrace, 'LWxLR' denotes Large White x Landrace crossbreds.

1.4.2.3. Intestinal host receptors

1.4.2.3.1. Glycoproteins and glycolipids

It was observed that F4ab and F4ac fimbriae prefer to bind to glycoproteins and that F4ad fimbriae prefer to bind to glycolipids, suggesting that the lipid and protein moieties of the receptors could play a role in mediating the binding of the fimbria to the receptor (Erickson et al., 1992; Erickson et al., 1994; Grange and Mouricout, 1996; Billey et al., 1998; Grange et al., 1999). It is possible that the protein and lipid moieties act as an integral part of the F4 binding epitope, serve as a site for a secondary interaction between the F4 variant and receptor that stabilizes the fimbriae-carbohydrate interaction, or present the carbohydrates in the correct conformation for recognition by the F4 fimbriae (Grange et al., 1999). For instance, it was observed that the ceramide (lipid) component of glycosphingolipids to which carbohydrate moieties are attached, interacts with the specificity of the binding of *E. coli* to glycosphingolipids (Bäckhed et al., 2002; Teneberg et al., 2004).

1.4.2.3.1.1. ANPEP

Aminopeptidase N, ANPEP, belonging to the M1 family of zinc metallopeptidases, is a 936 amino acid membrane glycoprotein and is widely expressed on the surface of various cell types, including porcine enterocytes (Delmas et al., 1992; Rawlings and Barrett, 1993; Olsen et al., 1997). Previously, it has been shown that ANPEP is a functional cell surface receptor for group I coronaviruses and that this interaction is dependent on species-specific amino acid differences in the receptor protein (Delmas et al., 1992; Yeager et al., 1992; Delmas et al., 1994; Tresnan et al., 1996; Benbacer et al., 1997; Li et al., 2007b; Tusell et al., 2007). Recently, ANPEP has been identified as an F4ac receptor by comparative proteomic analysis of the brush border proteins in F4ac receptor-positive ($F4acR^+$) and F4ac receptor-negative ($F4acR^-$) pigs and by adherence/internalization experiments on ANPEP-transfected cells. F4ac binding to ANPEP is mediated by the carbohydrate moieties of ANPEP in a sialic acid-dependent manner, resulting in clathrin-mediated endocytosis of the fimbriae (Melkebeek et al., 2012). Binding of F4ab/ad ETEC to ANPEP has not yet been investigated.

The gene encoding ANPEP [GenBank: NC_010449.4] is located on chromosome 7 (SSC7) and is composed of 20 exons (Poulsen et al., 1991). Although the locus controlling F4ab/ac ETEC susceptibility has been mapped on SSC13, it is possible that variation in gene

expression of *ANPEP*, localized on SSC7, is regulated by *trans*-acting factors present in the candidate region on SSC13 and therefore explains the observed F4ac ETEC binding profile to ANPEP in F4ac phenotyped pigs (Douglas and Wood, 2011; Melkebeek et al., 2012; Ren et al., 2012; Schroyen et al., 2012a; Nguyen and Seoighe, 2013).

1.4.2.3.1.2. Other F4 receptors

Two glycoproteins with a weight of 210 and 240 kDa isolated from the brush border were identified as F4ab/ac ETEC receptor in pigs exhibiting phenotype A and B (Erickson et al., 1992; Billey et al., 1998; Francis et al., 1998). These glycoproteins are intestinal mucin-type sialoglycoproteins (IMTGP-1 (210 kDa) and IMTGP-2 (240 kDa)), and contain O-linked oligosaccharides composed of galactosyl (β-1,3) *N*-acetylgalactosamine, α-linked fucose, galactosyl (β-1,4) *N*-acetylglucosamine, sialic acid, galactose and *N*-acetylgalactosamine (Erickson et al., 1994; Grange et al 1998).

A 74 kDa glycoprotein (GP74), belonging to the transferrin family, was recognized by F4ab fimbriae, and not by F4ac or F4ad fimbriae. Also, the intestinal GP74 receptor was detected in F4ab ETEC adhesive pigs and not in F4ab ETEC non-adhesive pigs. Compared to the gastric and serum tranferrins, the intestinal GP74 was characterized by high amounts of mannose, galactose and *N*-acetylglucosamine (Grange et al., 1996). Another F4ab ETEC receptor, galactosylceramide, was identified in the mucus from the small intestine (Blomberg et al., 1993). Further, other glycoproteins (26kDa, 40 to 42 kDa, 41 kDa, and 80 kDa) present in the mucus of the small intestine of the pigs were recognized by F4ab fimbriae (40 to 42 kDa) or F4ac fimbriae (26kDa, 41 kDa, and 80 kDa) (Metcalfe et al., 1991; Fang et al., 2000; Jin et al., 2000).

For F4ad ETEC, a neutral glycosphingolipid (GSL), neolactotetraosylceramide, originating from intestinal epithelial cells has been identified as F4ad ETEC receptor. This GSL with a terminal β-linked galactose was found in the adhesive phenotypes A and D, but not in phenotype C (Grange et al., 1999). The terminal β-linked galactose is an essential component for the recognition of the neutral GSL by F4ad fimbriae as well as for the recognition of intestinal mucin-type sialoglycoproteins IMTGP-1 and IMTGP-2 by F4ac fimbriae (Grange et al., 1998; Grange et al., 1999). However, no studies have been performed on genes coding for these F4-binding glycoproteins or glycolipids to elucidate the genetic basis of F4 ETEC resistance.

1.4.2.3.2. Carbohydrate-binding specificity of F4 ETEC

Based on the presence of different F4 ETEC adhesion phenotypes in pigs containing different F4 receptor profiles and based on the different haemagglutination patterns of the F4 variants, each variant possesses a related, but different carbohydrate-binding specificity (Bijlsma and Frik, 1987; Grange et al., 2002). Previous studies determining the carbohydrate-binding specificity of the F4 variants indicated that *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), *N*-acetylmannosamine, D-galactosamine and galactose (Gal) may play a role in the interaction of the fimbria with the receptor (Gibbons et al., 1975; Anderson et al., 1980; Sellwood, 1980; Blomberg et al., 1993; Payne et al., 1993). In an attempt to characterize the carbohydrate receptor structure of the F4 fimbrial variants by using porcine serum transferrin and GSL standards, it was found that the minimal binding epitope for all three variants contains a β-linked *N*-acetylhexosamine (HexNAc), and that the presence of a terminal β-linked galactose enhances the binding of F4 variants (Grange et al., 2002). Also, the binding of F4 fimbriae to mucus proteins was found to be blocked by a lectin of *Euonymus europeaus* which specifically recognizes the Galα(1–3)Gal sequence, indicating that this disaccharide forms a significant part of the receptor structure (Willemsen and de Graaf, 1992).

Recently, the target cell receptors of the F4 fimbriae were further investigated by binding of isolated F4ab, F4ac and F4ad fimbriae, and F4ab, F4ac, and F4ad ETEC, to reference glycosphingolipids, glycosphingolipids from erythrocytes and from porcine small intestinal epithelium. Specific interactions between the F4ab, F4ac and F4ad fimbriae and both acid and non-acid glycosphingolipids were obtained (Table 1.3). All three F4 variants recognized GSLs characterized as galactosylceramide (GalNAcα3GalNAcß3Galß4Glcß1Cer and GalNAcα3GalNAcß3Galß4GlcNAcß3Galß4Glcß1Cer) isolated from chicken erythrocytes, and lactosylceramide (Galß4Glcß1Cer) isolated from the porcine small intestinal epithelium.
Probably due to differences in the adhesive subunit (FaeG) sequence between these F4 variants, two distinct binding patterns for F4ab and F4ac were observed. The F4ac fimbriae and the F4ac ETEC selectively bound to galactosylceramide (Galß1Cer), while the GSL recognized by F4ab fimbriae and F4ab ETEC were more divers. F4ab fimbriae and F4ab ETEC bound to galactosylceramide (Galß1Cer), sulfatide (SO3-3Galß1Cer), sulf-lactosylceramide (SO3-3Galß4Glcß1Cer), globotriaosylceramide (Galα4Galß4Glcß1Cer), and galabiaosylceramide (Galα4Galβ1Cer). The F4ad fimbriae and the F4ad-fimbriated ETEC

bound to reference gangliotriaosylceramide (GalNAcß4Galß4Glcß1Cer), gangliotetraosylceramide (Galß3GalNAcß4Galß4Glcß1Cer), isoglobotriaosylceramide (Galα3-Galß4Glcß1Cer), and neolactotetraosylceramide (Galß4GlcNAcß3Galß4Glcß1Cer). Although distinct carbohydrate binding patterns were defined for each fimbrial subtype, further characterization of these glycans of the F4 binding GSL would be an important next step (Coddens et al., 2011).

Three genes (*B3GNT5*, *B4GALT4* and *B3GALT3*) of the glucosyl/galactosyltransferse family located on SSC13 were examined by comparing their cDNA sequences in F4acR$^+$ pigs and F4acR$^-$ pigs. Unlike lactosylceramide 1,3-N-acetyl-β-D-glucosaminyl transferase (*B3GNT5*) and β-1,3-galactosyltransferase 3 (*B3GALT3*), β-1,4-galactosyltransferase 4 (*B4GALT4*) was only partially sequenced with approximately 100 base pairs missing at the 5' end (Python et al., 2005). *B3GNT5* encodes an enzyme that initiates the formation of the lactoseries GSL by transferring GlcNAc in a β-1,3-linkage to lactosylceramide (Henion et al., 2001). *B4GALT4*, a member of the β-1,4- galactosyltransferase gene family, and *B3GALT3*, a member of the β-1,3-galactosyltransferase gene family, encode enzymes that transfer galactose as a β-anomer through 1,3 or 1,4 linkage to a variety of acceptors (Zhou et al., 1999). No polymorphisms in these genes were found to be associated with F4ac ETEC susceptibility (Python et al., 2005). In a study of Ouyang et al. (2011), the complete coding sequence of *B3GNT5* was also sequenced in F4ab/acR$^+$ and F4ab/acR$^-$ pigs, concluding the same that no polymorphisms were causal for F4ab/ac ETEC susceptibility (Ouyang et al., 2011).
In addition, an expression study of *B3GNT5* and *B4GALT4* was performed on cDNA isolated from the mid-jejunum of piglets with and without diarrhea. Although no expression difference was found, it should be noted that the piglets were not F4 phenotyped and that is possible that the cause of diarrhea may not be due to F4 ETEC (Schroyen et al., 2008).

Table 1.3: Summary of GSL binding specificities of F4ab, F4ac, and F4ad fimbriae. *Adapted from Coddens et al., 2011*.

No.	Trivial name	Structure	F4ab[a]	F4ac[a]	F4ad[a]
I. Chicken erythrocyte Glycosphingolipids					
1.	Galactosylceramide	Galß1Cer	+++	+++	–
2.		GalNAcα3GalNAcß3Galß4Glcß1Cer	+++	+++	+++
3.		GalNAcα3GalNAcß3Galß4GlcNAcß3Galß4Glcß1Cer	+++	+++	+++
II. Porcine intestinal Glycosphingolipids					
1.	Galactosylceramide	Galß1Cer	+++	+++	–
2.	Sulfatide	SO3-3Galß1Cer	+++	+	–
3.	Sulf-lactosylceramide	SO3-3Galß4Glcß1Cer	+++	–	–
4	Globotriaosylceramide (t18:0-h24:0)	Galα4Galß4Glcß1Cer	+++	–	–
III. Reference Glycosphingolipids					
1.	Lactosylceramide (t18:0-h16:0-h24:0)	Galß4Glcß1Cer	+++	+++	+++
2.	Galabiaosylceramide	Galα4Galß1Cer	+++	–	–
3.	Isoglobotriaosylceramide	Galα3Galß4Glcß1Cer	–	–	+
4.	Gangliotriaosylceramide	GalNAcß4Galß4Glcß1Cer	–	–	+++
5.	Gangliotetraosylceramide	Galß3GalNAcß4Galß4Glcß1Cer	–	–	+++
6.	Neolactotetraosylceramide	Galß4GlcNAcß3Galß4Glcß1Cer	–	–	+

[a]F4ab denotes binding obtained with both F4ab fimbriae and F4ab-fimbriated ETEC, F4ac binding obtained with both F4ac fimbriae and F4ac-fimbriated ETEC, and F4ad binding obtained with both F4ad fimbriae and F4ad-fimbriated ETEC.
Binding is defined as follows: +++ denotes an intense and highly reproducible staining when 4 mg of the glycosphingolipid was applied on the thin-layer chromatogram, + denotes an occasional staining, while − denotes no binding even at 4 mg.

1.4.2.4. Genetic mapping of the F4 receptor locus

The first linkage studies reported linkage of the F4ab/ac receptor (F4ab/acR) locus with the transferrin locus on SSC13 (Guérin et al., 1993; Edfors-Lilja et al., 1995). By fine mapping, the localization of the F4ab/acR locus was refined to the interval between the microsatellites *S0068* and *SW1030* (Python et al., 2002) and later to the interval between microsatellites *SW207* and *SW225* (Figure 1.4) (Jørgensen et al., 2003).

Based on the results of the linkage studies, Jørgensen et al. (2004) proposed mucin 4 (*MUC4*; GenBank: DQ848681) as a strong candidate gene.

The linkage study of Joller et al. (2009) refined the candidate region for the F4ab/acR locus to a 10 Mb region (5.7 cM) between microsatellites *SW207* and *S0075*. Through haplotype mapping, the candidate region for the F4ab/acR locus was then refined to a 3.1 Mb region between the zinc finger DHHC type containing 19 protein gene (*ZDHHC19*) and microsatellite *S0075* (Jacobsen et al., 2009). Recently, the candidate region of F4ab/acR locus was located between the leishmanolysin-like gene (*LMLN*) and microsatellite *S0283* based on a recombination event in the *MUC4*-F4ab/acR interval (Rampoldi et al., 2011). In addition, no significant single nucleotide polymorphisms (SNPs) were detected in the *MUC4-LMLN* region in the genome-wide association study of Fu et al. (2012) with the Porcine SNP60 BeadChip. Based on the presence of several haplotypes associated with resistance to F4ab/ac ETEC, Rampoldi (2013) argued that the F4ab/acR locus should be located between *ALGA0072075* (located close to *HEG1*) and *MUC13-813* (located in exon 11 of *MUC13*).

Although different polymorphisms were identified in the candidate region for F4ab/acR locus, none of these polymorphisms were causative. Jørgensen et al. (2004) found that the g.8227C>G polymorphism in *MUC4* was strongly associated with F4ab/ac ETEC susceptibility. Based on this polymorphism, a DNA marker based test has been developed to allow genotyping for F4ab/ac ETEC susceptibility or resistance (Jørgensen et al., 2004). However, the genotypes, especially the CC genotype which was identified as the resistant genotype to F4ab/ac ETEC adhesion, were not completely consistent with the results of the adhesion assay. This result confirms that the g.8227G>C polymorphism is a marker but not the actual causative mutation (Rasschaert et al., 2007; Li et al., 2008).

Also, other polymorphisms in *MUC4* (Peng et al., 2007), mucin 13 (*MUC13*) (Zhang et al., 2008; Huang, 2010; Ren et al., 2012), mucin 20 (*MUC20*) (Ji et al., 2011), transferrin receptor (*TRFC*) (Wang et al., 2007; Jacobsen et al., 2011), solute carrier family 12 member 8 (*SLC12A8*), myosin light chain kinase (*MYLK*), karyopherin alpha 1 (*KPAN1*) (Huang et al.,

2008), beclin-1 associated RUN domain containing protein (*KIAA0226*) (Jacobsen et al., 2011), lactosylceramide 1,3-N-acetyl-β-D-glucosaminyl transferase (*B3GNT5*) (Ouyang et al., 2011) were found to be associated with F4ab/ac ETEC phenotypes, but none of these polymorphisms were causative.

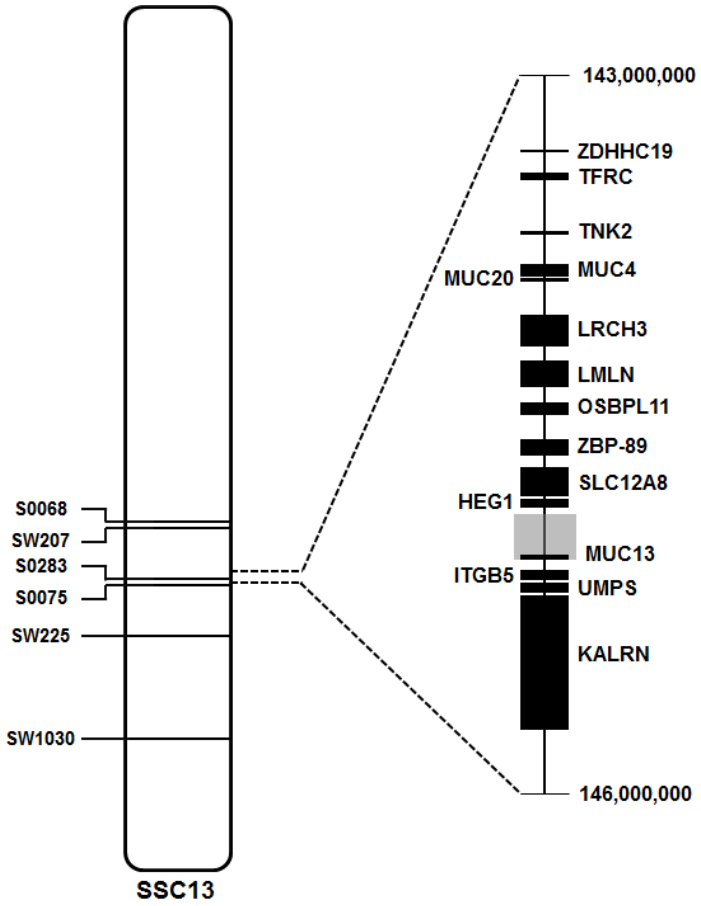

Figure 1.4: Localization of the F4ab/acR candidate region on SSC13. The markers, gene annotations, order and approximate scale are deduced from the porcine map in Ensembl (Sscrofa10.2). The grey box indicates the candidate region for the F4ab/acR locus located between *ALGA0072075* (located close to *HEG1*) and *MUC13-813* interval.

In addition, a genome-wide scan was carried out in 33 pigs to find any linkage of microsatellites spread over the 18 pig autosomes with the receptor locus for F4ad ETEC, but no linked markers were identified during this study. A larger pedigree (136 pigs) was then used for chromosomes 4, 11 and 13 linkage analysis, concluding that no linkage was found between F4adR locus and the microsatellite markers on chromosomes 4, 11 and 13 (Python, 2003).

1.4.2.4.1. Candidate genes based on the genetic mapping of the F4ab/ac R locus

Based on the study of Rampoldi (2013), two annotated porcine genes are located in the proposed candidate region for the F4ab/acR locus, namely heart of glass 1 (*HEG1*) and *MUC13* (Figure 1.4).

1.4.2.4.1.1. HEG1

HEG1 (GenBank: NC_010455.4; chr13: 144,760,509-144,829,493) was recently proposed a candidate gene by Fu et al. (2012). HEG1, a transmembrane protein, is evolutionary related to MUC13 and its extracellular domain has been predicted to be highly glycosylated (Lang et al., 2006; Faurobert and Albiges-Rizo, 2010). Although the exact function remains unknown, HEG1 plays a role in regulating vascular development and integrity (Faurobert and Albiges-Rizo, 2010; Gingras et al., 2012). As in human, *HEG1* showed low expression levels in the porcine small intestine and revealed alternative splicing not related to the F4ab/ac ETEC adhesion phenotype (Fu et al., 2012; Rampoldi, 2013). Because the region between *SLC12A8* and the microsatellite *KVL1293* (located in intron 12 of *HEG1*) was excluded as candidate region for F4ab/acR locus and because there were no associated SNPs identified in the interval between microsatellite *KLV1293* and 3'UTR of *HEG1*, it was suggested that *HEG1* is no longer a candidate gene for F4ab/ac ETEC susceptibility (Rampoldi, 2013).

1.4.2.4.1.2. MUC13

Mucins are glycoproteins characterized by a tandem repeat structure, which comprises most of the protein backbone and is the scaffold for a large number of complex O-linked carbohydrate side chains (Williams et al., 2001). As in human, *MUC13* is highly expressed in the porcine small intestine (Williams et al., 2001; Schroyen et al., 2012b; Rampoldi, 2013). Recently, *MUC13* was suggested as the causal gene for F4ac ETEC susceptibility (Ren et al., 2012). The MUC13 glycoprotein consists of an α and β subunit interconnected by covalent links which can be cleaved by proteases. The ß subunit refers to the C-terminal subunit containing the cytoplasmic tail and forms a homodimer (Williams et al., 2001). The α subunit contains the tandem repeat domain where the O-linked glycosylation takes place via glycosyltransferases and is enriched with serine, threonine and also proline (PTS tandem repeat region) (Figure 1.5). Glycans, including mannose, xylose, *N*-acetylglucosamine

(GlcNAc), *N*-acetylgalactosamine (GalNAc), and others, are mostly attached to the serine and threonine residues (Williams et al., 2001; Bergstrom and Xia, 2013).

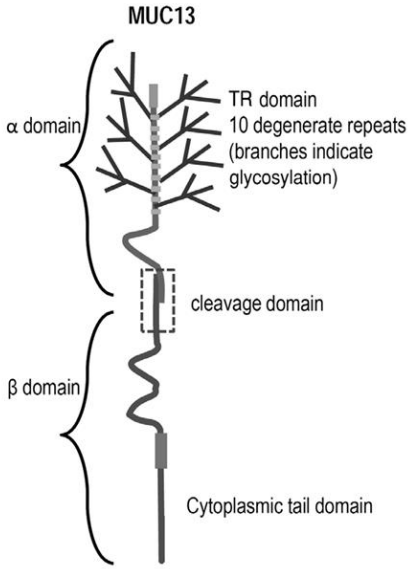

Figure 1.5: Schematic diagram of MUC13 showing the α subunit (extracellular) containing the tandem repeat (TR) domain and the β subunit containing the cytoplasmic tail domain with in between the cleavage site. *Adapted from Maher et al. (2011).*

Two isoforms of MUC13 were recently identified in pig, namely MUC13A and MUC13B. Unlike MUC13A that lacks an O-glycosylation site, MUC13B is predicted to be heavily O-glycosylated and therefore would contain the binding site for F4ac ETEC. These different O-glycosylation patterns between MUC13A and MUC13B would be due to the distinct tandem repeats in exon 2 of *MUC13*. To allow genotyping for F4ac ETEC susceptibility or resistance, a DNA marker based test was then developed based on the indel of 68 bp in intron 2 of *MUC13* with the longer sequence for *MUC13A* and the shortened sequence for *MUC13B* (Ren et al., 2012). Also, in the study of Rampoldi (2013), it was suggested that exon 2 of *MUC13* is the most likely location for the F4ab/acR locus.

1.4.2.5. Methods for genetic mapping of the F4ab/ac receptor locus

Linkage studies and association studies have both been used in the search for the genetic causes underlying F4 ETEC susceptibility. These approaches can either be hypothesis-based (i.e., candidate gene/region studies) or hypothesis-free (i.e., genome-wide studies) (Wang et al., 2011). Both methods (linkage and association analysis) use genetic markers to discover loci, but differ in experimental design. Genetic markers, such as restriction fragment length polymorphisms, simple sequence length polymorphisms (minisatellites and microsatellites), and single nucleotide polymorphisms (SNPs) are genotyped in a pedigree (linkage) or in a population (association) (Brown, 2002).

Linkage is the co-segregation of a genetic marker and the disease phenotype within a pedigree. This was demonstrated by Morgan (1911) in Drosophila *melanogaster*, the species of fruit fly. Morgan (1911) proposed that the further apart two genes are, the higher the probability of recombination between these two genes. By determining the probability of recombination between two genes, the genetic distance between these two genes can be measured allowing the construction of a genetic linkage map (Sturtevant, 1913). By using polymorphic DNA markers for the construction of a genetic linkage map, genetic loci responsible for disease can be mapped with respect to the DNA marker loci by linkage analysis in a multigenerational pedigree (Botstein et al., 1980). The recombination frequency (proportion of recombinant progeny arising from a genetic cross) is calculated which is proportional to the physical distance between the marker and the disease loci (Sturtevant, 1913; Griffiths et al., 1999). Linkage analysis generally uses a low-density map of microsatellite markers that are spaced at intervals of ~10 cM across the genome (Sellick et al., 2004). Microsatellite markers are short (up to several hundred base pairs) segments of DNA consisting of multiple tandem repeats of a two- or three base pair sequence (Chang, 2011). Further fine mapping (by further marker genotyping and/or sequencing of candidate genes) will then be required to identify the exact causative variant (Duncan et al., 2014). However, recent studies have shown that linkage analysis using a dense map of SNPs or microsatellites can result in a substantial gain in the power to detect linkage (Evans and Cardon, 2004; Sellick et al., 2004). SNPs (single nucleotide polymorphisms) are specific positions in a genome where the nucleotide (A, T, C, or G) differs between and within species (Griffiths et al., 1999; Chang, 2011).

Association analysis correlates genetic markers, usually SNPs, with the disease phenotype across a population and determines whether a certain allele occurs at a frequency higher than the one expected by change in animals with a particular phenotype (Yamada et al., 2011). This nonrandom pattern is called linkage disequilibrium (LD) and is dependent on the recombination frequency and the number of generations since the mutation was introduced into the population (Jorde, 1995). Recently, the Porcine SNP60 DNA BeadChip (Illumina, USA; Ramos et al., 2009) containing 62,163 SNPs that uniformly span the whole genome was developed. Due to this uniform genome-wide coverage, this Porcine SNP60 DNA BeadChip enables a broad range of applications, including whole-genome association studies, and can be used to determine genetic variation in porcine breeds such as Duroc, Landrace, Piétrain, and Large White.

A candidate region on SSC13 has long been identified as harboring the F4ab/acR locus by linkage and association studies using genetic markers, but further detailed study of all genetic variants in the locus is required to discover the causal variant(s), to quantify their contribution to F4ab/ac ETEC susceptibility, and to elucidate their roles in functional pathways. Recent advances in DNA sequencing technology, such as next generation sequencing (NGS), have enabled the sequencing of individual genomes and are able to provide a much deeper, more uniform picture of genetic variations (1000 Genomes Project Consortium, 2010). NGS platforms perform massively parallel sequencing, during which millions of fragments of DNA or RNA from a single sample are sequenced in unison and allow an entire genome or exome to be sequenced within a day. Besides whole-genome and whole-exome sequencing, targeted sequencing of specific genes or genomic regions is also possible using NGS which would be of interest to find the causal variant(s) located in the previous identified candidate region on SSC13 (Grada and Weinbrecht, 2013). In addition, transcript expression levels can be assessed using NGS on RNA in $F4R^+$ and $F4R^-$ pigs and can contribute to the further elucidation of the genetic basis of F4 ETEC resistance.

1.5. References

1000 Genomes Project Consortium. 2010. A map of human genome variation from population-scale sequencing. Nature. 467:1061-1073.

Alexander TJL. 1994. Neonatal diarrhoea in pigs. In: Gyles CL (ed), *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, UK, 151-170.

Alustiza FE, Picco NY, Bellingeri RV, Terzolo HR, Vivas AB. 2012. Frequency of virulence genes of *Escherichia coli* among newborn piglets from an intensive pig farm in Argentina. Rev Argent Microbiol. 44:250-254.

Amezcua R, Friendship RM, Dewey CE. 2008. An investigation of the presence of *Escherichia coli* O149:K91:F4 on pig farms in southern Ontario and the use of antimicrobials and risk factors associated with the presence of this serogroup. Can Vet J. 49:39-45.

Anderson MJ, Whitehead JS, Kim YS. 1980. Interaction of *Escherichia coli* K88 antigen with porcine intestinal brush border membranes. Infect Immun. 29:897-901.

Bäckhed F, Ahlsén B, Roche N, Ångström J, von Euler A, Breimer ME, Westerlund-Wikström B, Teneberg S, Richter-Dahlfors A. 2002. Identification of target tissue glycosphingolipid receptors for uropathogenic, F1C-fimbriated *Escherichia coli*, and its role in mucosal inflammation. J Biol Chem. 277:18198-18205.

Bailey M, Haverson K. 2006. The postnatal development of the mucosal immune system and mucosal tolerance in domestic animals. Vet Res. 37:443-453.

Baker DR, Billey LO, Francis DH. 1997. Distribution of K88 *Escherichia coli*-adhesive and nonadhesive phenotypes among pigs of four breeds. Vet Microbiol. 54:123-132.

Bakker D, Willemsen PT, Simons LH, van Zijderveld FG, de Graaf FK. 1992. Characterization of the antigenic and adhesive properties of FaeG, the major subunit of K88 fimbriae. Mol Microbiol. 6:247-255.

Benbacer L, Kut E, Besnardeau L, Laude H, Delmas B. 1997. Interspecies aminopeptidase-N chimeras reveal species-specific receptor recognition by canine coronavirus, feline infectious peritonitis virus, and transmissible gastroentertis virus. J Virol. 71:734-737.

Berberov EM, Zhou Y, Francis DH, Scott MA, Kachman SD, Moxley RA. 2004. Relative importance of heat-labile enterotoxin in the causation of severe diarrheal disease in the gnotobiotic piglet model by a strain of enterotoxigenic *Escherichia coli* that produces multiple enterotoxins. Infect Immun. 72:3914-3924.

Bergstrom KSB, Xia L. 2013. Mucin-type O-glycans and their roles in intestinal homeostasis. Glycobiology. 23:1026-1037.

Bertschiner HU, Gyles CL. 1994. Oedema disease of pigs. In: Gyles CL (ed). *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, UK, 193-220.

Bianchi AT, Scholten JW, van Zijderveld AM, van Zijderveld FG, Bokhout BA. 1996. Parenteral vaccination of mice and piglets with $F4^+$ *Escherichia coli* suppresses the enteric anti-F4 response upon oral infection. Vaccine. 14:199-206.

Bijlsma IGW, deNijs A, van der Meer C, Frik JF. 1982. Different pig phenotypes affect adherence of *Escherichia coli* to jejunal brush borders by K88ab, K88ac, or K88ad antigen. Infect Immun. 37:891-894.

Bijlsma IGW, deNijs A, Frik JF. 1985. The prevalence of different porcine phenotypes in the Netherlands concerning adherence of K88-positive *Escherichia coli* to intestinal epithelium. Vet Quart. 7:246-248.

Bijlsma IG, Bouw J. 1987. Inheritance of K88-mediated adhesion of *Escherichia coli* to jejunal brush borders in pigs: a genetic analysis. Vet Res Commun. 11:509-518.

Bijlsma IG, Frik JF. 1987. Haemagglutination patterns of the different variants of *Escherichia coli* K88 antigen with porcine, bovine, guinea pig, chicken, ovine and equine erythrocytes. Res Vet Sci. 43:122-123.

Billey LO, Erickson AK, Francis DH. 1998. Multiple receptors on porcine intestinal epithelial cells for the three variants of *Escherichia coli* K88 fimbrial adhesin. Vet Microbiol. 59:203-212.

Blomberg L, Krivan HC, Cohen PS, Conway PL. 1993. Piglet ileal mucus contains protein and glycolipid (galactosylceramide) receptors specific for *Escherichia coli* K88 fimbriae. Infect Immun. 6:2526-2531.

Boerlin P, Travis R, Gyles CL, Reid-Smith R, Janecko N, Lim H, Nicholson V, McEwen SA, Friendship R, Archambault M. 2005. Antimicrobial resistance and virulence genes of *Escherichia coli* isolates from swine in Ontario. Appl Environ Microbiol. 71:6753-6761.

Bonneau M, Duval-Iflah Y, Guérin G, Ollivier L, Renard C, Renjifo X. 1990. Aspects génétiques et microbiologiques de la colibacillose K88 chez le porc. Ann Rech Vet. 21:302-303.

Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 32:314-331.

Boyen F, Vangroenweghe F, Butaye P, De Graef E, Castryck F, Heylen P, Vanrobaeys M, Haesebrouck F. 2010. Disk prediffusion is a reliable method for testing colistin susceptibility in porcine *E. coli* strains. Vet Microbiol. 144:359-362.

Brown TA. 2002. Chapter 5: Mapping Genomes. In: Brown TA (ed), Genomes, 2nd edition. Wiley-Liss, Oxford, UK, Available: http://www.ncbi.nlm.nih.gov/books/NBK21116/.

Burgess MN, Bywater RJ, Cowley CM, Mullan NA, Newsome PM. 1978. Biological evaluation of a methanol-soluble, heat stable *Escherichia coli* enterotoxin in infant mice, pigs, rabbits, and calves. Infect Immun. 21:526-531.

Byun JW, Jung BY, Kim HY, Fairbrother JM, Lee MH, Lee WK. 2013. O-serogroups, virulence genes of pathogenic *Escherichia coli* and pulsed-field gel electrophoresis (PFGE) patterns of O149 isolates from diarrhoeic piglets in Korea. Vet Med-Czech. 58:468-476.

Callens B, Persoons D, Maes D, Laanen M, Postma M, Boyen F, Haesebrouck F, Butaye P, Catry B, Dewulf J. 2012. Prophylactic and metaphylactic antimicrobial use in Belgian fattening pig herds. Prev Vet Med. 106:53-62.

Casewell M, Friis C, Marco E, McMullin P, Phillips I. 2003. The European ban on growth-promoting antibiotics and emerging consequences for human and animal health. J Antimicrob Chemother. 52:159-161.

Casey TA, Herring CJ, Schneider RA, Bosworth BT, Whipp SC. 1998. Expression of heat-stable enterotoxin STb by adherent *Escherichia coli* is not sufficient to cause severe diarrhea in neonatal pigs. Infect Immun. 66:1270-1272.

Casey TA, Connell TD, Holmes RK, Whipp SC. 2012. Evaluation of heat-labile enterotoxins type IIa and type IIb in the pathogenicity of enterotoxigenic *Escherichia coli* for neonatal pigs. Vet Microbiol. 159:93-89.

CDC (Centers for Disease Control and Prevention). 2013. Antibiotic resistance threats in the United States, 2013. CDC, Atlanta, US.

Celemin C, Anguita J, Naharro G, Suarez S. 1994. Evidence that *Escherichia coli* isolated from healthy pigs intestine hybridize with LT-II, ST& and SLT-II DNA probes. Microbial Pathogen. 16:77-81.

Celemin C, Rubio P, Echeverria P, Suarez S. 1995. Gene toxin patterns of *Escherichia coli* isolated from diseased and healthy piglets. Vet Microbiol. 45:121-127.

Chakraborty K, Ghosh S, Koley H, Mukhopadhyay AK, T. Ramamurthy, Saha DR, Mukhopadhyay D, S. Roychowdhury, Hamabata T, Takeda Y, Das S. 2008. Bacterial exotoxins downregulate cathelicidin (hCAP-18/LL-37) and human beta-defensin 1 (HBD-1) expression in the intestinal epithelial cells. Cell Microbiol. 10:2520-2537.

Chandler DS, Mynott TL, Luke RKJ, Craven JA. 1994. The distribution and stability of *Escherichia coli* K88 receptor in the gastrointestinal tract of the pig. Vet Microbiol. 38:203-215.

Chang PP, Moss J, Twiddy EM, Holmes RK. 1987.Type II heat-labile enterotoxin of Escherichia coli activates adenylate cyclase in human fibroblasts by ADP ribosylation. Infect Immun. 55:1854-1858.

Chang B. 2011. Determination of genomic locations of target genetic loci. In: Gu W, Wang Y (eds), Gene discovery for disease models. John Wiley & Sons, New Jersey, USA, 111-138.

Chantziaras I, Dewulf J, Boyen F, Callens B, Butaye P. 2014. Antimicrobial resistance prevalence of pathogenic and commensal *Escherichia coli* in food-producing animals in Belgium. VDT. 83:225-233.

Cheng D, Sun H, Xu J, Gao S. 2006. PCR detection of virulence factor genes in *Escherichia coli* isolates from weaned piglets with edema disease and/or diarrhea in China. Vet Microbiol. 115:320-28.

Chin JJC, Chapman TA. 2009. Impact on animal health: Designer probiotics for the management of intestinal health and colibacillosis in weaner pigs. In: Lee YK, Salminen S (eds), Handbook of probiotics and prebiotics. John Wiley & Sons, New Jersey, USA, 406-440.

Clements A, Young JC, Constantinou N, Frankel G. 2012. Infection strategies of enteric pathogenic *Escherichia coli*. Gut microbes. 3:71-87.

Coddens A, Valis E, Benktander J, Ångström J, Breimer ME, Cox E, Teneberg S. 2011. Erythrocyte and porcine intestinal glycosphingolipids recognized by F4 fimbriae of enterotoxigenic *Escherichia coli*. PloS ONE. 6:e23309.

Conway PL, Welin A, Cohen PS. 1990. Presence of K88-specific receptors in porcine ileal mucus is age dependent. Infect Immun. 58:3178-3182.

Costa MM, Drescher G, Maboni F, Weber SS, Schrank A, Vainstein MH, Schrank IS, Vargas AC. 2010. Virulence factors, antimicrobial resistance, and plasmid content of *Escherichia coli* isolated in swine commercial farms. Arq Bras Med Vet Zootec. 62:30-36.

Cox E, Houvenaghel A. 1993. Comparison of the *in vitro* adhesion of K88, K99, F41, and P987 positive *Escherichia coli* to intestinal villi of 4- to 5-week-old pigs. Vet Microbiol. 34:7-18.

Cox E. Van Praet W, Vandersmissen T. Claerebout E, Butaye P, Vangroenweghe F, Coddens A. Pathogens associated with post-weaning diarrrhea in Belgium. Unpublished results.

DebRoy C, Roberts E, Scheuchenzuber W, Kariyawasam S, Jayarao BM. 2009. Comparison of genotypes of *Escherichia coli* strains carrying F18ab and F18ac fimbriae from pigs. J Vet Diagn Invest. 21:359-64.

De la Fe Rodriguez PY, Coddens A, Del Fava E, Abrahantes JC, Shkedy Z, Martin LOM, Muñoz EC, Duchateau L, Cox E, Goddeeris BM. 2011. High prevalence of $F4^+$ and $F18^+$ *Escherichia coli* in Cuban piggeries as determined by serological survey. Trop Anim Health Prod. 43:937-946.

Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjöström H, Noren O, Laude H. 1992. Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. Nature. 357:417-419.

Delmas B, Gelfi J, Sjöström H, Noren O, Laude H. 1994. Further characterization of aminopeptidase N as a receptor for coronaviruses. Adv Exp Med Biol. 342:293-298.

Deplancke B, Gaskins HR. 2001. Microbial modulation of innate defense: goblet cells and the intestinal mucus layer. Am J Clin Nutr. 73:1131S-1141S.

Do T, Stephens C, Townsend K, Wu X, Chapman T, Chin J, McCormick B, Bara M, Trott D. 2005. Rapid identification of virulence genes in enterotoxigenic *Escherichia coli* isolates associated with diarrhoea in Queensland piggeries. Aust Vet J. 83:293-299.

Do TN, Cu PH, Nguyen HX, Au TX, Vu QN, Driesen SJ, Townsend KM, Chin JJ, Trott DJ. 2006. Pathotypes and serogroups of enterotoxigenic *Escherichia coli* isolated from pre-weaning pigs in north Vietnam. J Med Microbiol. 55:93-99.

Douglas AGL, Wood MJA. 2011. RNA splicing: disease and therapy. Brief Funct Genomics. 10:151-164.

Dreyfus LA, Harville B, Howard DE, Shaban R, Beatty DM, Morris SJ. 1993. Calcium influx mediated by the *Escherichia coli* heat-stable enterotoxin B (ST_B). Proc Natl Acad Sci USA. 90:3202-3206.

Duncan E, Brown M, Shore EM. 2014. The revolution in human monogenic disease mapping. Genes. 5:792-803.

Edfors-Lilja I, Petersson H, Gahne B. 1986. Performance of pigs with or without the intestinal receptor for *Escherichia coli* K88. Anim Prod. 42:381-387.

Edfors-Lilja I, Gustafsson U, Duval-Iflah Y, Ellergeen H, Johansson M, Juneja RK, Marklund L, Andersson L. 1995. The porcine intestinal receptor for *Escherichia coli* K88ab, K88ac: regional localization on chromosome 13 and influence of IgG response to the K88 antigen. Anim Genet. 26:237-242.

Edfors-Lilja I, Wallgren, P. 2000. *Escherichia coli* and *Salmonella* Diarrhoea in Pigs. In: Axford RFE, Bishop SC, Nicholas FW, Owen JB (eds), Breeding for disease resistance in farm animals, 2nd edition. CABI publishing, Wallingford, UK, 253-267.

EMA (European Medicines Agency). 2013. Use of colistin products in animals within the European Union: development of resistance and possible impact on human and animal health. EMA/755938/2012.

Erickson AK, Willgohs JA, McFarland SY, Benfield DA, Francis DH. 1992. Identification of two porcine brush border glycoproteins that bind the K88ac adhesin of *Escherichia coli* and correlation of these glycoproteins with the adhesive phenotype. Infect Immun. 60:983-988.

Erickson AK, Baker DR, Bosworth BT, Casey TA, Benfield DA, Francis DH. 1994. Characterization of porcine intestinal receptors for the K88ac fimbrial adhesin of *Escherichia coli* as mucin-type sialoglycoproteins. Infect Immun. 62:5404-5410.

Erume J, Berberov EM, Kachman SD, Scott MA, Zhou Y, Francis DH, Moxley RA. 2008.Comparison of the contributions of heat-labile enterotoxin and heat-stable enterotoxin b to the virulence of enterotoxigenic *Escherichia coli* in F4ac receptor-positive young pigs. Infect Immun. 76:3141-3149.

Evans DM, Cardon LR. 2004. Guidelines for genotyping in genomewide linkage studies: single-nucleotide–polymorphism maps versus microsatellite maps. Am J Hum Genet. 75:687-692.

Fairbrother JM, Nadeau E, Gyles CL. 2005. *Escherichia coli* in postweaning diarrhea in pigs: An update on bacterical types, pathogenesis and prevention strategies. Anim Health Res Rev. 6:17-39.

Fairbrother JM. 2006. Neonatal *Escherichia coli* diarrhea. In: Straw BE, Zimmerman JJ, D'Allaire S, Taylor DJ (eds), Diseases of swine, 9th edition. Blackwell Publishing Professional, Iowa, USA, 641-649.

Fairbrother JM, Gyles CL. 2006. Postweaning *Escherichia coli* diarrhea and edema disease. In: Straw BE, Zimmerman JJ, D'Allaire S, Taylor DJ (eds), Diseases of swine, 9th ed. Blackwell Publishing Professional, Iowa, USA, 649-662.

Fang L, Gan ZB, Marquardt RR. 2000. Isolation, affinity purification, and identification of piglet small intestine mucosa receptor for enterotoxigenic *Escherichia coli* K88ac^{+} fimbriae. Infect Immun. 68:564-569.

Faurobert E, Albiges-Rizo C. 2010. Recent insights into cerebral cavernous malformations: a complex jigsaw puzzle under construction. FEBS J. 277:1084-1096.

Flores A. 2004. New way to control *E. coli* in weaned pigs. Agric Res. 52:9.

Francis DH, Grange PA, Zeman DH, Baker DR, Sun R, Erickson AK. 1998. Expression of mucin-type glycoprotein K88 receptors strongly correlates with piglet susceptibility to K88+ enterotoxigenic *Escherichia coli*, but adhesion of this bacterium to brush borders does not. Infect Immun. 66:4050-4055.

Francis DH. 2002. Enterotoxigenic *Escherichia coli* infection in pigs and its diagnosis. J Swine Health Prod. 10:171-175.

Frydendahl K. 2002. Prevalence of serogroups and virulence genes in *Escherichia coli* associated with postweaning diarrhoea and edema disease in pigs and a comparison of diagnostic approaches. Vet Microbiol. 85:169-182.

Fu WX, Yang L, Xin L, Xiao-Yan N, Xiang-Dong D, Liu JF, Zhang Q. 2012. A genome-wide association study identifies two novel promising candidate genes affecting *Escherichia coli* F4ab/ac susceptibility in swine. PLoS ONE. 7:e32127.

Fujii Y, Nomura T, Yamanaka H, Okamoto K. 1997. Involvement of Ca(2+)-calmodulin-dependent protein kinase II in the intestinal secretory action of *Escherichia coli* heat-stable enterotoxin II. Microbiol Immunol. 41:633-636.

Fürer E, Cryz SJJ, Dorner F, Nicolet J, Wanner M, Germanier R. 1982. Protection against colibacillosis in neonatal piglets by immunization of dams with procholeragenoid. Infect Immun. 35:887-894.

Gaastra W, Mooi FR, Stuitje AR, de Graaf FK. 1981. The nucleotide sequence of the K88ab protein subunit of porcine enterotoxigenic *Escherichia coli*. FEMS Microbiol Lett. 12:41-46.

Gaastra W, de Graaf FK. 1982. Host-specific fimbrial adhesins of noninvasive enterotoxigenic *Escherichia coli* strains. Microbiol Rev. 46:129-161.

Gaastra W, Klemm P, de Graaf FK. 1983. The nucleotide sequence of the K88ad protein subunit of porcine enterotoxigenic *Escherichia coli*. FEMS Microbiol Lett. 18:177-183.

Gannon VP, Gyles CL, Wilcock BP. 1989. Effects of *Escherichia coli* Shiga-like toxins (verotoxins) in pigs. Can J Vet Res. 53:306-312.

Gibbons RA, Jones GW, Sellwood R. 1975. An attempt to identify the intestinal receptor for the K88 adhesin by means of a haemagglutination inhibition test using glycoproteins and fractions of sow colostrum. J Gen Microbiol. 86:228-240.

Gibbons RA, Sellwood R, Burrows M, Hunter PA. 1977. Inheritance of resistance to neonatal *E. coli* diarrhea in the pig: examination of the genetic system. Theor Appl Genet. 51:65-70.

Gilligan PH, Brown JC, Robertson DC. 1983. Immunological relationships between cholera toxin and *Escherichia coli* heat-labile enterotoxin. Infect Immun. 42:683-691.

Gingras AR, Liu JJ, Ginsberg MH. 2012. Structural basis of the junctional anchorage of the cerebral cavernous malformations complex. JCB. 199:39-48.

Glenn GM, Francis DH, Danielsen EM. 2009. Toxin-mediated effects on the innate mucosal defenses: implications for enteric vaccines. Infect Immun. 77:5206.

Gonçalves C, Vachon V, Schwartz JL, Dubreuil JD. 2007. The *Escherichia coli* enterotoxin STb permeabilizes piglet jejunal brush border membrane vesicles. Infect Immun. 75:2208-2213.

Grada A, Weinbrecht K. 2013. Next-generation sequencing: methodology and application. J Invest Dermatol. 133:e11.

Grange PA, Mouricout MA. 1996. Transferrin associated with the porcine intestinal mucosa is a receptor specific for K88ab fimbriae of *Escherichia coli*. Infect Immun. 64:606-610.

Grange PA, Erickson AK, Anderson TJ, Francis DH. 1998. Characterization of the carbohydrate moiety of intestinal mucin-type sialoglycoprotein receptors for the K88ac fimbrial adhesin of *Escherichia coli*. Infect Immun. 66:1613-1621.

Grange PA, Erickson AK, Levery SB, Francis DH. 1999. Identification of an intestinal neutral glycosphingolipid as a phenotype-specific receptor for the K88ad fimbrial adhesin of *Escherichia coli*. Infect Immun. 67:165-172.

Grange PA, Mouricout MA, Levery SB, Francis DH, Erickson AK. 2002. Evaluation of receptor binding specificity of *Escherichia coli* K88 (F4) fimbrial adhesin variants using porcine serum transferrin and glycosphingolipids as model receptors. Infect Immun. 70:2336-2343.

Grange PA, Parrish LA, Erickson AK. 2006. Expression of putative *Escherichia coli* heat-labile enterotoxin (LT) receptors on intestinal brush borders from pigs of different ages. Vet Res Commun. 30:57-71.

Green BA, Neill RJ, Ruyechan WT, Holmes RK. 1983. Evidence that a new enterotoxin of *Escherichia coli* which activates adenylate cyclase in eukaryotic target cells is not plasmid mediated. Infect Immun. 41:383-390.

Griffiths AJF, Gelbart WM, Miller JH, Lewontin RC. Modern genetic analysis. New York: W. H. Freeman. Available: http://www.ncbi.nlm.nih.gov/books/NBK21358/.

Gross L. 2006. Bacterial Fimbriae Designed to Stay with the Flow. PLoS Biol. 4:e314.

Guérin G, Duval-Iflah Y, Bonneau M, Bertaud M, Guillaume P, Ollivier L. 1993. Evidence for linkage between K88ab, K88ac intestinal receptors to *Escherichia coli* and transferrin loci in pigs. Anim Genet. 24:393-396.

Guinee PA, Jansen WH. 1979. Behavior of *Escherichia coli* K antigens K88ab, K88ac, and K88ad in immunoelectrophoresis, double diffusion, and hemagglutination. Infect Immun. 23:700-705.

Gyles CL. 1971. Discussion: Heat-labile and heat-stable forms of the enterotoxin from *E. coli* strains enteropathogenic for pigs. Ann N Y Acad Sci. 176:314-322.

Habrun B, Kompes G, Cvetnić Ž, Špičić S, Benić M, Mitak M. 2010. Antimicrobial susceptibility of *Escherichia coli*, *Salmonella* spp., *Pasteurella multocida*, *Streptococcus suis* and *Actinobacillus pleuropneumoniae* isolated from diagnostic samples from large pig breeding farms in Croatia. Vet Arhiv. 80:571-583..

Hampson DJ. 1994. Post-weaning *E. coli* diarrhoea in pigs. In: Gyles CL (ed), *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, UK, 171-191.

Harville BA, Dreyfus LA. 1995. Involvement of 5-hydroxytryptamine and prostaglandin E2 in the intestinal secretory action of *Escherichia coli* heat-stable enterotoxin B. Infect Immun. 63:745-750.

Henion TR, Zhou D, Wolfer DP, Jungalwala FB, Hennet T. 2001. Cloning of a mouse β1,3N-acetylglucosaminyltransferase GlcNAc(β1,3)Gal(β1,4)Glc-ceramide synthase gene encoding the key regulator of lacto-series glycolipid biosynthesis. J Biol Chem. 276:30261-30269.

Hill GM, Mahan DC, Carter SD, Cromwell GL, Ewan RC, Harrold RL, Lewis AJ, Miller PS, Shurson GC, Veum TL. 2001. Effect of pharmacological concentrations of zinc oxide with or without the inclusion of an antibacterial agent on nursery pig performance. J Anim Sci. 79:934-941.

Hitotsubashi S, Fujii Y, Yamanaka H, Okamoto K. 1992. Some properties of purified *Escherichia coli* heat-stable enterotoxin II. Infect Immun. 60:4468-4474.

Holmgren J, Lonnroth I, Mansson J, Svennerholm L. 1975. Interaction of cholera toxin and membrane GM1 ganglioside of small intestine. Proc Natl Acad Sci USA. 72:2520-2524.

Hopwood DE, Hampson DJ. 2003. Interactions between the intestinal microflora, diet and diarrhoea, and their influences on piglet health in the immediate post-weaning period. In: Pluske JR, Le Dividich,

J, Verstegen MWA (eds), Weaning the pig: concepts and consequences. Wageningen Academic Publishers, Wageningen, Netherlands, 199-218.

Hu ZL, Hasler-Rapacz J, Huang SC, Rapacz J. 1993. Studies in swine on inheritance and variation in expression of small intestinal receptors mediating adhesion of the K88 enteropathogenic *Escherichia coli* variants. J Hered. 84:157-165.

Huang X, Ren J, Yan X, Peng Q, Tang H, Zhang B, Ji H, Yang S, Huang L. 2008. Polymorphisms of three gene-derived STS on pig chromosome 13q41 are associated with susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac in pigs. Sci China Ser C. 51:614-619.

Huang L. 2010. *MUC13* molecular marker for identifying the F4ac adhesin-caused diarrhea resistance of the weanling pig and the use thereof. International application published under the patent cooperation treaty (PCT), PCT/CN2009/070574 or WO 2010/066118.

Jackson PGG, Cockcroft PD. 2007. Population medicine. In: Jackson PGG, Cockcroft PD (eds), Handbook of pig medicine. Elsevier Ltd, Philadelpia, USA, 16-31.

Jacobsen M, Kracht SS, Esteso G, Cirera S, Edfors I, Archibald AL, Bendixen C, Andersson L, Fredholm M, Jørgensen CB. 2009. Refined candidate region specified by haplotype sharing for *Escherichia coli* F4ab/F4ac susceptibility alleles in pigs. Anim Genet. 41:21-25.

Jacobsen M, Cirera S, Joller D, Esteso G, Kracht SS, Edfors I, Bendixen C, Archibald AL, Vögeli P, Neuenschwander S, Bertschinger HU, Rampoldi A, Andersson L, Fredholm M, Jørgensen CB. 2011. Characterisation of five candidate genes within the ETEC F4ab/ac candidate region in pigs. BMC Res Notes. 4:225.

Ji H, Ren J, Yan X, Huang X, Zhang B, Zhang Z, Huang L. 2011. The porcine *MUC20* gene: molecular characterization and its association with susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac. Mol Biol Rep. 38:1593-1601.

Jin LZ, Marquardt RR, Baidoo SK, Frohlich AA. 2000. Characterization and purification of porcine small intestinal mucus receptor for *Escherichia coli* K88ac fimbrial adhesin. FEMS Immunol Med Microbiol. 27:17-22.

Johnson AM, Kaushik RS, Francis DH, Fleckenstein JM, Hardwidge PR. 2009. Heat-labile enterotoxin promotes *Escherichia coli* adherence to intestinal epithelial cells. J Bacteriol. 191:178-186.

Joller D. 2009. Comparative molecular approaches to identify host determinants mediating adhesion of *E. coli* F4 strains in pigs. ETH. Available: http://dx.doi.org/10.3929/ethz-a-005931566.

Joller D, Jørgensen CB, Bertschinger HU, Python P, Edfors I, Cirera S, Archibald AL, Bürgi E, Karlskov-Mortesen P, Andersson L, Fredholm M, Vögeli P. 2009. Refined localization of the *Escherichia coli* F4ab/F4ac receptor locus on pig chromosome 13. Anim Genet. 40:749-752.

Jones GW, Rutter JM. 1972. Role of the K88 antigen in the pathogenesis of neonatal diarrhea caused by *Escherichia coli* in piglets. Infec Immun. 6:918-927.

Jorde LB. 1995. Linkage disequilibrium as a gene-mapping tool. Am J Hum Genet. 56:11-14.

Josephsen J, Hansen F, de Graaf FK, Gaastra W. 1984. The nucleotide sequence of the protein subunit of the K88ac fimbriae of porcine enterotoxigenic *Escherichia coli*. FEMS Microbiol Lett. 25:301-306.

Jørgensen CB, Cirera S, Anderson SI, Archibald AL, Raudsepp T, Chowdhary B, Edfors-Lilja I, Andersson L, Fredholm M. 2003. Linkage and comparative mapping of the locus controlling susceptibility towards *E. coli* F4ab/ac diarrhoea in pigs. Cytogenet Genome Res. 102:157-162.

Jørgensen CB, Cirera S, Archibald AL, Anderson L, Fredholm M, Edfors-Lilja I. 2004. Porcine polymorphisms and methods for detecting them. International application published under the patent cooperation treaty (PCT), PCT/DK2003/000807 or WO2004/048606A2.

Katwa LC, Parker CD, Dybing JK, White AA. 1991. Solubilization and reprecipitation from intestinal brush border membranes of a complex containing guanylate cyclase activatable by the heat stable enterotoxin. Arch Biochem Biophys. 290:397-406.

Kennedy DJ, Greenberg RN, Dunn JA, Abernathy R, Ryerse JS, Guerrant RL. 1984. Effects of *Escherichia coli* heatstable enterotoxin STb on intestines of mice, rats, rabbits, and piglets. Infect Immun. 46:639-643.

Kil DY, Stein HH. 2010. Invited review: Management and feeding strategies to ameliorate the impact of removing antibiotic growth promoters from diets fed to weanling pigs. Can J Anim Sci. 90:447-460.

Kim JC, Hansen CF, Pluske JR, Mullan BP, No B. 2010. Evaluating the replacement of zinc oxide with an encapsulated zinc oxide product as a means of controlling post-weaning diarrhoea in piglets. Pork CRC. Project 2C-114:1-16.

Kim SJ, Kwon CH, Park BC, Lee CY, Han JH. 2015. Effects of a lipid-encapsulated zinc oxide dietary supplement, on growth parameters and intestinal morphology in weanling pigs artificially infected with enterotoxigenic *Escherichia coli*. J Anim Sci Technol. 1:4.

Klemm P. 1985. Fimbrial adhesins of *Escherichia coli*. Rev Infect Dis. 7:321-340.

Kongsted H. 2014. New Neonatal Porcine Diarrhoea Syndrome - A study on its aetiology, epidemiology and clinical manifestations. Frederiksberg, SL Grafik. Available: http://curis.ku.dk/ws/files/122989106/Hanne_Kongsted_B5_PhD_thesis_2014.pdf.

Labrie V, Harel J, Dubreuil JD. 2002. *Escherichia coli* heat-stable enterotoxin b (STb) *in vivo* internalization within rat intestinal epithelial cells. Vet Res. 33:223-228.

Lang T, Hansson GC, Samuelsson T. 2006. An inventory of mucin genes in the chicken genome shows that the mucin domain of Muc13 is encoded by multiple exons and that ovomucin is part of a locus of related gel-forming mucins. BMC Genomics. 7:197.

Lee SI, Rayamahji N, Lee WJ, Cha SB, Shin MK, Roh YM, Sang HY. 2009. Genotypes, antibiogram, and pulsed-field gel electrophoresis profiles of *Escherichia coli* strains from piglets in Korea. J Vet Diagn Invest. 21:510-516.

Li YH, Qiu XT, Li HJ, Zhang Q. 2007a. Adhesive patterns of *Escherichia coli* F4 in piglets of three breeds. J Genet Genomics. 34:591-599.

Li BX, Ge JW, Li YL. 2007b. Porcine aminopeptidase N is a functional receptor for the PEDV coronavirus. Virology. 365:166-172.

Li H, Li Y, Qiu X, Niu X, Liu Y, Zhang Q. 2008. Identification and screening of gene(s) related to susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac in piglets. AJAS. 21:489-493.

Li XT, Wang B, Li JL, Zeng FQ, Gong X. 2012. Prevalence and antimicrobial susceptibility of bacterial pathogens isolates from diseased swine in southwest, China. Afr J Microbiol Res. 6:1579-1583.

Loos M, Geens M, Schauvliege S, Gasthuys F, van der Meulen J, Dubreuile JD, Goddeeris BM, Niewold T, Cox E. 2012. Role of heat-stable enterotoxins in the induction of early immune responses in piglets after infection with enterotoxigenic *Escherichia coli*. PLOS One. 7:e41041.

MacLeod DL, Gyles CL, Wilcock BP. 1991. Reproduction of edema disease of swine with purified Shiga-like toxin-II variant. Vet Pathol. 28:66-73.

Madoroba E, Van Driessche E, De Greve H, Mast J, Ncube I, Read J, Beeckmans S. 2009. Prevalence of enterotoxigenic *Escherichia coli* virulence genes from scouring piglets in Zimbabwe. Trop Anim Health Prod. 9:1539-1547.

Maher DM, Gupta BK, Nagata S, Jaggi M, Chauhan SC. 2011. Mucin 13: Structure, function and potential roles in cancer pathogenesis. Mol Cancer Res. 9:531-537.

Maron DF, Smith TJS, Nachman KE. 2013. Restrictions on antimicrobial use in food animal production: an international regulatory and economic survey. Global Health. 9:48.

Martins M, Martinez Rossi N, Ferreira A, Brocchi M, Yano T, Castro A, Silveira W. 2000. Pathogenic characteristics of *Escherichia coli* strains isolated from newborn piglets with diarrhea in Brazil. Vet Microbiol. 76:51-9.

Maynard C, Fairbrother JM, Bekal S, Sanschagrin F, Levesque RC, Brousseau R, Masson L, Lariviere S, Harel J. 2003. Antimicrobial resistance genes in enterotoxigenic *Escherichia coli* O149:K91 isolates obtained over a 23-year period from pigs. Antimicrob Agents Chemother. 47:3214-3221.

Melkebeek V, Rasschaert K, Bellot P, Tilleman K, Favoreel H, Deforce D, De Geest BG, Goddeeris BM, Cox E. 2012. Targeting aminopeptidase N, a newly identified receptor for F4ac fimbriae, enhances the intestinal mucosal immune response. Mucosal Immunol. 5:635-645.

Melkebeek V, Goddeeris BM, Cox E. 2013. ETEC vaccination in pigs. Vet Immunol Immunopathol. 152:37-42.

Metcalfe JW, Krogfelt KA, Krivan HC, Cohen PS, Laux DC. 1991. Characterization and identification of a porcine small intestine mucus receptor for the K88ab fimbrial adhesin. Infect Immun. 59:91-96.

Michaels RD, Whipp SC, Rothschild MF. 1994. Resistance of Chinese Meishan, Fengjing, and Minzhu pigs to the K88ac+ strain of *Escherichia coli*. Am J Vet Res. 55:333-338.

Moon HW, Schneider RA, Moseley SL. 1986. Comparative prevalence of four enterotoxin genes among *Escherichia coli* isolated from swine. Am J Vet Res. 47:210-212.

Morales AS, Fragoso de Araújo J, de Moura Gomes VT, Reis Costa AT, dos Prazeres Rodrigues D, Porfida Ferreira TS, Nogueira de Lima Filsner PH, Felizardo MR, Moreno AM. 2012. Colistin resistance in *Escherichia coli* and *Salmonella enterica* strains isolated from swine in Brazil. The Scientific World J. 2012:1-4.

Morgan TH. 1911. Random segregation vs. coupling in Mendelian inheritance. Science. 34:384.

Moseley SL, Samadpour-Motalebi M, Falkow S. 1983. Plasmid association and nucleotide sequence relationships of two genes encoding heat-stable enterotoxin production in *Escherichia coli* H-10407. J Bacteriol. 156:441-443.

Nagy B, Casey TA, Moon HW. 1990. Phenotype and genotype of *Escherichia coli* isolated from pigs with postweaning diarrhea in Hungary. J Clin Microbiol. 28:651-653.

Nagy B, Fekete PZ. 1999. Enterotoxigenic *Escherichia coli* (ETEC) in farm animals. Vet Res. 30:259-284.

Nataro JP, Kaper JB. 1998. Diarrheagenic *Escherichia coli*. Clin Microbiol Rev. 11:142-201.

National swine survey, Technical Report. 1992. Morbidity/mortality and health management of swine in the United States. USDA-APHIS-VS-NAHMS, Fort Collins Colorado.

Nguyen TT, Seoighe C. 2013. Integrative analysis of mRNA expression and half-life data reveals *trans*-acting genetic variants associated with increased expression of stable transcripts. PloS ONE. 8:e79627.

Olsen J, Kokholm K, Noren O, Sjöström H. 1997. Structure and expression of aminopeptidase N. Adv Exp Med Biol. 421:47-57.

Orskov I, Orskov F, Sojka WJ, Wittig W. 1964. K antigens K88ab(L) and K88ac(L) in *E. coli*. A new O antigen: 0147 and a new K antigen: K89(B). Acta Pathol Microbiol Scand. 62:439-447.

Ouyang J, Zeng W, Ren J, Yan X, Zhang Z, Yang M, Han P, Huang X, Ai H, Huang L. 2011. Association of *B3GNT5* polymorphisms with susceptibility to ETEC F4ab/ac in the white Duroc × Erhualian intercross and 15 outbred pig breeds. Biochem Genet. 50:19-33.

Payne D, O'Reilly M, Williamson D. 1993. The K88 fimbrial adhesin of enterotoxigenic *Escherichia coli* binds to β1-linked galactosyl residues in glycosphingolipids. Infec Immun. 61:3673-3677.

Peelman LJ. 1999. Genetic investigation of the resistance mechanisms of the pig against diarrhea caused by *E. coli*. Verh K Acad Geneeskd Belg. 61:489-515.

Peng QL, Ren J, Yan XM, Huang X, Tang H, Wang YZ, Zhang B, Huang LS. 2007. The g.243A>G mutation in intron 17 of *MUC4* is significantly associated with susceptibility/resistance to ETEC F4ab/ac infection in pigs. Anim Genet. 38:397-400.

Peterson JW, Whipp SC. 1995. Comparison of the mechanisms of action of cholera toxin and the heat-stable enterotoxins of *Escherichia coli*. Infect Immun. 63:1452-1461.

Pickett CL, Twiddy EM, Belisle BW, Holmes RK. 1986. Cloning of genes that encode a new heat-labile enterotoxin of *Escherichia coli*. J Bacteriol. 165:348-352.

Poulsen PH, Thomsen PD, Olsen J. 1991. Assignment of the porcine aminopeptidase N (*PEPN*) gene to chromosome 7cen----q21. Cytogenet Cell Genet. 57:44-46.

Python P, Jörg H, Neuenschwander S, Hagger C, Stricker C, Bürgi E, Bertschinger HU, Stranzinger G, Vögeli P. 2002. Fine-mapping of the intestinal receptor locus for enterotoxigenic *Escherichia coli* F4ac on porcine chromosome 13. Anim Genet. 33:441-447.

Python P. 2003. Genetic host determinants associated with the adhesion of *E. coli* with fimbriae F4 in swine. ETH. Available: http://dx.doi.org/10.3929/ethz-a-004630253.

Python P, Jorg H, Neuenschwander S, Asai-Coakwell M, Hagger C, Bürgi E, Bertschinger HU, Stranzinger G, Vögeli P. 2005. Inheritance of the F4ab, F4ac and F4ad *E. coli* receptors in swine and examination of four candidate genes for F4acR. J Anim Breed Genet. 122:5-14.

Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu ZL, Kerstens HH, Law AS, Megens HJ, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TP, Schnabel RD, Van Tassell CP, Taylor JF, Wiedmann RT, Schook LB, Groenen MA. 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PloS One. 4:e6524.

Rampoldi A, Jacobsen MJ, Bertschinger HU, Joller D, Bürgi E, Vögeli P, Andersson L, Archibald AL, Fredholm M, Jørgensen CB, Neuenschwander S. 2011. The receptor locus for *Escherichia coli* F4ab/F4ac in the pig maps distal to the *MUC4-LMLN* region. Mamm Genome. 22:122-129.

Rampoldi A. 2013. Approach to the gene controlling the porcine receptor for *Escherichia coli* with fimbriae F4ab/F4ac and inheritance of the receptor for F4ad. ETH. Available: http://dx.doi.org/10.3929/ethz-a-009907149.

Rampoldi A, Bertschinger HU, Bürgi E, Dolf G, Sidler X, A. Bratus A, P. Vögeli P, Neuenschwander S. 2014. Inheritance of porcine receptors for enterotoxigenic *Escherichia coli* with fimbriae F4ad and their relation to other F4 receptors. Animal. 8:859-866.

Rasschaert K, Verdonck F, Goddeeris BM, Duchateau L, Cox E. 2007. Screening of pigs resistant to F4 enterotoxigenic *Escherichia coli* (ETEC) infection. Vet Microbiol. 123:249-253.

Rawlings ND, Barrett AJ. 1993. Evolutionary families of peptidases. Biochem J. 290:205-218.

Ren J, Yan X, Ai H, Zhang Z, Huang X, Ouyang J, Yang M, Yang H, Han P, Zeng W, Chen Y, Guo Y, Xiao S, Ding N, Huang L. 2012. Susceptibility towards enterotoxigenic *Escherichia coli* F4ac diarrhea is governed by the MUC13 gene in pigs. PloS ONE. 7:e445739.

Roselli M, Finamore A, Garaguso I, Britti MS, Mengheri E. 2003. Zinc oxide protects cultured enterocytes from the damage induced by *Escherichia coli*. J Nutr. 133:4077-4082.

Rousset E, Harel J, Dubreuil JD. 1998. Sulfatide from the pig jejunum brush border epithelial cell surface is involved in binding of *Escherichia coli* enterotoxin b. Infect Immun. 66:5650-5658.

Rutter JM, Burrows MR, Sellwood R, Gibbons RA. 1975. A genetic basis for resistance to enteric disease caused by *E. coli*. Nature. 257:135-136.

Rutter JM, Jones GW, Brown GTH, Burrows MR, Luther PD. 1976. Antibacterial activity in colostrum and milk associated with protection of piglets against enteric disease caused by K88-positive *Escherichia coli*. Infect Immun. 13:667-676.

Sarrazin E, Fritzsche C, Bertschinger HU. 2000. Hauptvirulenzfaktoren bei *Escherichia coli*-isolaten von über zwei wochen alten schweinen mit ödemkrankheit und/oder colidiarrhöe. Schweizer Archiv Tierheilk. 142:625-630.

Sellers ZM, Mann E, Smith A, Ko KH, Giannella R, Cohen MB, Barrett KE, Dong H. 2008. Heat-stable enterotoxin of *Escherichia coli* (STa) can stimulate duodenal HCO_3^- secretion via a novel GC-C- and CFTR-independent pathway. FASEB J. 22:1306-1316.

Sellick GS, Longman C, Tolmie J, Newbury-Ecob R, Geenhalgh L, Hughes S, Whiteford M, Garrett C, Houlston RS. 2004. Genomewide linkage searches for Mendelian disease loci can be efficiently conducted using high-density SNP genotyping arrays. Nucl Acids Res. 32:e164.

Sellwood R, Gibbons RA, Jones GW, Rutter JM. 1975. Adhesion of enteropathogenic *Escherichia coli* to pig intestinal brush borders: the existence of two pig phenotypes. J Med Microbiol. 8:405-411.

Sellwood R. 1980. The interaction of the K88 antigen with porcine intestinal epithelial cell brush borders. Biochim Biophys Acta. 632:32-335.

Sellwood R. 1982. *Escherichia coli*-associated porcine neonatal diarrhea: antibacterial activities of colostrum from genetically susceptible and resistant sows. Infect Immun. 35:396-401.

Sellwood R. 1984a. An intestinal receptor for the K88 antigen of porcine enterotoxigenic *Escherichia coli*. In: Boedeker EC (ed), Attachment of organisms to the gut mucosa, vol. 2. CRC Press, Florida, USA, 167-175.

Sellwood R. 1984b. The K88 adherence system in swine. In: Boedeker EC (ed), Attachment of organisms to the gut mucosa, vol 1. CRC Press, Florida, USA, 21-29.

Schroyen M, Verhelst R, Niewold TA, Buys N. 2008. Quantification of *B3GNT5* and *B4GALT4* expression reveals no difference between piglets with or without diarrhea. Commun Agric Appl Biol Sci. 73:205-208.

Schroyen M, Stinckens A, Verhelst R, Geens M, Cox E, Niewold T, Buys N. 2012a. Susceptibility of piglets to enterotoxigenic *Escherichia coli* is not related to the expression of *MUC13* and *MUC20*. Anim Genet. 43:324-327.

Schroyen M, Stinckens A, Verhelst R, Niewold T, Buys N. 2012b. The search for the gene mutations underlying enterotoxigenic *Escherichia coli* F4ab/ac susceptibility in pigs: a review. Vet Res. 43:70.

Smyth CJ, Marron M, Smith SGJ. 1994. Fimbriae of *Escherichia coli*. In: Gyles CL (ed.), *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, UK, 399-435.

Stokes CR, Bailey M, Haverson K, Harris C, Jones P, Inman C, Pié S, Oswald IP, Williams BA, Akkermans ADL, Sowa E, Rothkötter HJ, Miller BG. 2004. Postnatal development of intestinal immune system in piglets: implications for the process of weaning. Anim Res. 53:325-334.

Sturtevant AH. 1913. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. J Exp Zool. 14:43-59.

Sun R, Anderson TJ, Erickson AK, Nelson EA, Francis DH. 2000. Inhibition of adhesion of *Escherichia coli* K88ac fimbria to its receptor, intestinal mucin-type glycoproteins, by a monoclonal antibody directed against a variable domain of the fimbria. Infec Immun. 68:3509-3515.

Teneberg S, Ångström J, Ljungh A. 2004. Carbohydrate recognition by enterohemorrhagic *Escherichia coli*: characterization of a novel glycosphingolipid from cat small intestine. Glycobiology. 14:187-196.

Thacker PA. 2013. Alternatives to antibiotics as growth promoters for use in swine production: a review. J Anim Sci Biotechnol. 4-35.

Tresnan DB, Levis R, Holmes KV. 1996. Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. J Virol. 70:8669-8674.

Tubbs RC, Hurd HS, Dargatz D, Hill G. 1993. Preweaning morbidity and mortality in the United States swine heed. Swine Health Prod. 1:21-28.

Tusell SM, Schittone SA, Holmes KV. 2007. Mutational analysis of aminopeptidase N, a receptor for several group I coronaviruses, identifies key determinants of viral host range. J Virol. 81:1261-1273.

US FDA (United States Food and Drug Administration), Center for Veterinary Medicine. 2012. Guidence for industry #209: The judicious use of medically important antimicrobial drugs in food-producing animals. Docket Number FDA-2010-D-0094.

USDA (United States Department of Agriculture). 2002. Swine 2000, Part II: Reference of swine health and health management in the United States, 2000.

USDA (United States Department of Agriculture). 2007. Swine 2006, Part II: Reference of swine health and health management in the United States, 2006.

USDA (United States Department of Agriculture). 2014. Livestock and Poultry: World Markets and Trade - April 18, 2014.

van Beers-Schreurs HMG, Vellenga L, Wensing T, Breukink HJ. 1992. The pathogenesis of the post-weaning syndrome in weaned piglets; a review. Vet Quart. 14:29-34.

Van den Broeck W, Cox E, Goddeeris BM. 1999a. Receptor-dependent immune responses in pigs after oral immunization with F4 fimbriae. Infect Immun. 67:520-526.

Van den Broeck W, Cox E, Goddeeris BM. 1999b. Induction of immune responses in pigs following oral administration of purified F4 fimbriae. Vaccine. 17:2020-2029.

Van den Broeck W, Cox E, Goddeeris BM. 1999c. Receptor-specific binding of purified F4 to isolated villi. Vet Microbiol. 68:255-263.

Van den Broeck W, Cox E, Oudega B, Goddeeris BM. 2000. The F4 fimbrial antigen of *Escherichia coli* and its receptors. Vet Microbiol. 71:223-244.

Van Driessche E, Sachez R, Dieussaert I, Kanarek L, Lintermans P, Beeckmans S. 1995. Enterotoxigenic fimbrial *Escherichia coli* lectins and their receptors: targets for probiotic treatment of diarrhoea. In: Pusztai A, Bardicz S (eds), Lectins: biomedical perspectives. Taylor & Francis, London, UK, 235-292.

Vidotto Marilda C, de Lima Natália CS, Fritzen Juliana TT, de Freitas Júlio C, Venâncio Emerson J, Ono Mario A. 2009. Frequency of virulence genes in *Escherichia coli* strains isolated from piglets with diarrhea in the North Parana State, Brazil. Braz J Microbiol. 40:199-204.

Vondruskova H, Slamova R, Trckova M, Zraly Z, Pavlik I. 2010. Alternatives to antibiotic growth promoters in prevention of diarrhoea in weaned piglets: a review. Vet Med Czech. 55:199-224.

Vu-Khac H, Holoda E, Pilipcinec E, Blanco M, Blanco JE, Dahbi G, Mora A, Lopez C, Gonzalez EA, Blanco J. 2007. Serotypes, virulence genes, intimin types and pfge profiles of *Escherichia coli* isolated from piglets with diarrhoea in Slovakia. Vet J. 174:176-187.

Wada A, Hirayama T, Kitaura H, Fujusawa J, Hasegawa M, Hidaka Y, Shimonishi Y. 1996. Identification of ligand recognition sites in heat-stable enterotoxin receptor, membrane-associated guanylyl cyclase C by site-directed mutational analysis. Infect Immun. 64:5144-5150.

Waddell TE, Lingwood CA, Gyles CL. 1997. Interaction of verotoxin 2e with pig intestine. Infect Immun. 64:1714–1719.

Wang Y, Ren J, Lan L, Yan X, Huang X, Peng Q, Tang H, Zhang B, Ji H, Huang L. 2007. Characterization of polymorphisms of transferrin receptor and their association with susceptibility to ETEC F4ab/ac in pigs. J Anim Breed Genet. 124:225-229.

Wang X, Prins BP, Sõber S, Laan M, Snieder H. 2011. Beyond genome-wide association studies: new strategies for identifying genetic determinants of hypertension. Curr Hypertens Rep.13:442-451.

Wang C, Xie P, Liu LL, Dong XY, Lu JJ, Zou XT. 2012. Use of lower level of capsulated zinc oxide as an alternative to pharmacological dose of zinc oxide for weaned piglets. Asian J Anim Vet Adv. 7:1290-1300.

Weiglmeier PR, Rosch P, Berkner H. 2010. Cure and curse: *E. coli* heat-stable enterotoxin and its receptor guanylyl cyclase C. Toxins. 2:2213-2229.

WHO (World Health Organisation). 1997. The medical impact of the use of antimicrobials in food animals. WHO, Berlin, Germany.

WHO (World Health Organization). 2014. Antimicrobial resistance: global report on surveillance 2014. WHO, Geneva, Switzerland.

Willemsen PTJ, de Graaf FK. 1992. Age and serotype dependent binding of K88 fimbriae to porcine intestinal receptors. Microb Pathogen. 12:367-375.

Williams SJ, Wreschner DH, Tran M, Eyre HJ, Sutherland GR, McGuckin MA. 2001. MUC13, a novel human cell surface mucin expressed by epithelial and hemopoietic cells. J Biol Chem. 276:18327-18336.

Wilson MR, Hohman AW. 1974. Immunity to *Escherichia coli* in pigs: Adhesion of enteropathogenic *Escherichia coli* to isolated intestinal epithelial cells. Infec Immun. 10:776-782.

Zhang W, Berberov EM, Freeling J, He D, Moxley RA, Francis DH. 2006. Significance of heat-stable and heat-labile enterotoxins in porcine colibacillosis in an additive model for pathogenicity studies. Infect Immun. 74:3107-3114.

Zhang W. 2007. Studies are trying to answers for postweaning diarrhea caused by enterotoxigenic *Escherichia coli* (ETEC). Available: http://nationalhogfarmer.com/health-diseases/diseaseprevention/answers-postweaning-diarrhea/index.html. Accessed October 15, 2014.

Zhang W, Zhao M, Ruesch L, Omot A, Francis D. 2007. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. Vet Microbiol. 123:145-152.

Zhang B, Ren J, Yan X, Huang X, Ji H, Peng Q, Zhang Z, Huang L. 2008. Investigation of the porcine *MUC13* gene: isolation, expression, polymorphisms and strong association with susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac. Anim Genet. 39:258-266.

Zhang W, Fang Y, Francis DH. 2009. Characterization of the binding specificity of K88ac and K88ad fimbriae of enterotoxigenic *Escherichia coli* by constructing K88ac/K88ad chimeric FaeG major subunits. Infect Immun. 77:699-706.

Zhou D, Berger EG, Hennet T. 1999. Molecular cloning of a human UDP-galactose:GlcNAcβ1,3GalNAcβ1,3 galactosyltransferase gene encoding an O-linked core3-elongation enzyme. Eur J Biochem. 263:571-576.

Yamada Y, Ichihara S, Takemura M. 2008. Human functional genomics and proteomics. In: Kumar D (ed), Genomics and clinical medicine. Oxford University Press, New York, USA, 31-43.

Yan X, Ren J, Huang X, Zhang ZY, Ouyang J, Zheng WH, Zou ZZ, Yang SJ, Yang B, Xiao S, Huang LS. 2009a. Comparison of production traits between pigs with and without the *Escherichia coli* F4 receptors in a White Duroc × Erhualian intercross F2 population. J Anim Sci. 87:334-339.

Yan X, Huang X, Ren J, Zou Z, Yang S, Ouyang J, Zeng W, Yang B, Xiao S, Huang L. 2009b. Distribution of *Escherichia coli* F4 adhesion phenotypes in pigs of 15 Chinese and Western breeds and a White Duroc x Erhualian intercross. J Med Microbiol. 58:1112-1117.

Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT, Holmes KV. 1992. Human aminopeptidase N is a receptor for human coronavirus 229E. Nature. 357:420-422.

CHAPTER 2

AIM OF THE STUDY

2. AIM OF THE STUDY

F4 enterotoxigenic *Escherichia coli* (F4 ETEC) are worldwide an important cause of diarrhea in neonatal and newly-weaned pigs. F4 fimbriae mediate the adherence of F4 ETEC to the small intestine by binding to specific receptors. Although previous linkage and binding studies were performed on this important economical disease, no causal mutation has yet been identified suggesting that this disease is much more complex than initially thought.
The overall aim of the study was to identify the causal mutation(s) responsible for F4 ETEC susceptibility as well as its target gene(s) and to develop a new diagnostic genotyping test for the identified causal mutation(s). Two different approaches were used in parallel to achieve this goal.

The first approach was to perform a genome-wide association study (GWAS) using the Porcine SNP60 DNA BeadChip (**Chapter 3**) to refine the candidate region for F4ab/ac ETEC susceptibility using well-phenotyped piglets. An accurate assessment of the F4 phenotypes in the piglets is important to precisely pinpoint the phenotype-causing mutation(s). In addition, a selected region harboring the SNPs with the highest evidence of association was sequenced in F4ab/ac receptor-positive and F4ab/ac receptor-negative pigs in order to identify the causal mutation(s) (**Chapter 6**).

The second approach was to investigate whether genetic variation in the gene encoding ANPEP, identified as an F4ac receptor, (**Chapter 4**) and in genes involved in the assembly of the F4 binding carbohydrate moiety of glycosphingolipids (**Chapter 5**) could account for the F4 ETEC binding patterns previously observed in F4 phenotyped piglets. Although these genes are not located in the candidate region for F4ab/ac ETEC susceptibility on chromosome 13 (SSC13), it is possible that the causal mutation exerts remote regulatory effects on these genes whose coding regions lie outside the boundaries of this candidate region.

CHAPTER 3

REFINED CANDIDATE REGION FOR F4ab/ac ETEC SUSCEPTIBILITY SITUATED PROXIMAL TO *MUC13* IN PIGS

Based on PloS ONE 9:e105013 (2014).

Tiphanie Goetstouwers[1], Mario Van Poucke[1], Wouter Coppieters[2], Van Ut Nguyen[3], Vesna Melkebeek[3], Annelies Coddens[3], Katleen Van Steendam[4], Dieter Deforce[4], Eric Cox[3] and Luc J. Peelman[1]

[1]Laboratory of Animal Genetics, Faculty of Veterinary Medicine, Ghent University,, Belgium

[2]Unit of Animal Genomics, GIGA-R & Faculty of Veterinary Medicine, University of Liège (B34), Liège (Sart Tilman), Belgium

[3]Laboratory of Immunology, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium

[4]Laboratory of Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, Ghent University, Ghent, Belgium

3. REFINED CANDIDATE REGION FOR F4ab/ac ETEC SUSCEPTIBILITY SITUATED PROXIMAL TO *MUC13* IN PIGS

3.1. Abstract

F4 enterotoxigenic *Escherichia coli* (F4 ETEC) are an important cause of diarrhea in neonatal and newly-weaned pigs. Based on the predicted differential O-glycosylation patterns of the two MUC13 variants (MUC13A and MUC13B) in F4ac ETEC susceptible and F4ac ETEC resistant pigs, the *MUC13* gene was recently proposed as the causal gene for F4ac ETEC susceptibility. Because the absence of MUC13 on Western blot from brush border membrane vesicles of F4ab/acR$^+$ pigs and the absence of F4ac attachment to immunoprecipitated MUC13 could not support this hypothesis, a new GWAS study was performed using 52 non-adhesive and 68 strong adhesive pigs for F4ab/ac ETEC originating from 5 Belgian farms. A refined candidate region (chr13: 144,810,100-144,993,222) for F4ab/ac ETEC susceptibility was identified with *MUC13* adjacent to the distal part of the region. This candidate region lacks annotated genes and contains a sequence gap based on the sequence of the porcine GenomeBuild 10.2.

We hypothesize that a porcine orphan gene or *trans*-acting element present in the identified candidate region has an effect on the glycosylation of F4 binding proteins and therefore determines the F4ab/ac ETEC susceptibility in pigs.

3.2. Introduction

F4 enterotoxigenic *Escherichia coli* (F4 ETEC) diarrheal disease in neonatal and newly-weaned pigs is an economically important genetic disease inherited in an autosomal dominant pattern (Gibbons et al., 1997; Fairbrother et al., 2005; Zhang et al., 2007). F4 ETEC possess F4 fimbriae acting as ligands for specific carbohydrate receptors on the epithelial surface of the small intestine. This interaction results in colonization of the small intestine and in production of enterotoxins inducing a secretory diarrhea in young pigs (Fairbrother et al., 2005). F4 fimbriae exist in 3 antigenic variants: F4ab, F4ac and F4ad, of which F4ac is the most prevalent, except in central China where F4ad is the most prevalent fimbrial variant (Fairbrother et al., 2005; Wang et al., 2006). In the search for the causal mutation for F4ab/ac ETEC susceptibility, several linkage studies mapped the causal mutation(s) on swine chromosome 13 (SSC13) (Fu et al., 2012; Ren et al., 2012; Schroyen et al., 2012). Some studies suggest that one locus is controlling both F4ab and F4ac ETEC susceptibility (Bijlsma and Bouw, 1987; Python et al., 2002; Jørgensen et al., 2003), while others suggest 2 linked but distinct loci (Guérin et al., 1993; Edfors-Lilja et al., 1995; Li et al., 2007; Wang et al., 2007; Zhang et al., 2008; Yan et al., 2009). The g.8227G>C polymorphism in *MUC4* [Genbank: DQ848681] was found to be strongly associated with F4ab/ac ETEC susceptibility with the dominant G allele representing F4ab/ac ETEC susceptibility and the recessive C allele representing F4ab/ac ETEC resistance. A genotyping test was proposed for distinguishing F4ab/ac ETEC susceptible and resistant pigs (Jørgensen et al., 2004). However, Rasschaert et al. (2007) could not confirm this when comparing with *in vitro* adhesion to villi. Recently, Ren et al. (2012) suggested that 2 MUC13 variants (MUC13A and MUC13B) are responsible for F4ac ETEC susceptibility of Duroc x Erhualian, Chinese conventional Sutai and Duroc x Landrace x Large White hybrids populations, due to a (predicted) different O-glycosylation pattern between these two proteins. In this study, we investigated if *MUC13* is responsible for F4ab/ac ETEC susceptibility and we performed a GWAS study to identify a refined candidate region.

3.3. Material and methods

3.3.1. Sample collection

Experimental and animal management procedures were approved by the animal care and ethics committee of the Faculty of Veterinary Medicine, University of Ghent (EC2010/042). Pigs originating from five Belgian farms were euthanized at 6-18 weeks of age. The breeds of these pigs were Large White, Belgian Landrace, Large White x Belgian Landrace crossbreds, Large White x Piétrain crossbreds, and crossbreds of multiple breeds. Before euthanasia, blood samples were collected in EDTA blood tubes for DNA analysis and stored at -20 °C. After euthanasia, two-meter mid-jejunum samples were washed two times with Krebs-Henseleit buffer (0.12 M NaCl, 0.014 M KCl, 0.001 M KH2PO4, 0.025 M NaHCO3, pH 7.4) and once with Krebs-Henseleit buffer containing 1 % (v/v) formaldehyde at 4 °C. After washing, the villi were scraped from the mucosa of a 20 cm segment and stored as mentioned in Van den Broeck et al. (1999a). These villi were used for the *in vitro* villous adhesion assay. The rest of the mid-jejunal sample was used to isolate brush border membrane vesicles (BBMVs) as described in Nguyen et al. (2013).

3.3.2. F4ab/acR phenotyping based on the *in vitro* villous adhesion assay

All pigs were phenotyped for the presence of the F4ab/ac receptor (F4ab/acR) using the *in vit*ro villous adhesion assay with 4×10^8 F4ac *E. coli* (strain GIS26, serotype O149:K91, $F4ac^+$) or F4ab *E. coli* (strain G7, serotype O8:K87, $F4ab^+$) and an average of 50 villi in a volume of 0.5 ml PBS with 1 % (W/V) D-mannose (Van den Broeck et al., 1999a). Adhesion of more than 30 bacteria per 250 µm villous brush border length was noted as strong adhesive for F4ab/ac ETEC and less than 5 bacteria per 250 µm brush border length was noted as non-adhesive for F4ab/ac ETEC (Cox and Houvenaghel, 1993). A total of 120 pigs from 43 different litters were included in the GWAS study based on the two F4ab/acR phenotypes: non-adhesive ($F4R^-$; n= 52) and strong adhesive ($F4R^+$; n=68). The two F4ab/acR phenotypes were present in 23 litters, in 10 litters only the non-adhesive F4ab/acR phenotype and in 10 other litters only the strong adhesive F4ab/acR phenotype was present. Nine boars in this study had multiple litters (Table 3.1).
Pigs that showed a weak adhesion towards F4ab/ac ETEC were excluded.

3.3.3. *MUC4* TaqMan assay and the Indel *MUC13* marker test

DNA isolation of the blood samples for the *MUC4* TaqMan assay and the Indel *MUC13* marker was performed as described by Van Poucke et al. (2015).

The *MUC4* TaqMan assay was carried out as described by Nguyen et al. (2013) and is based on the g.8227G>C mutation of *MUC4* [Genbank: DQ848681] associated with F4ab/ac ETEC susceptibility (Jørgensen et al., 2004)..

The Indel *MUC13* marker test is based on an indel of 68 bp in intron 2 of *MUC13* [GenBank: NC_010455.4; chr13: 144,993,222-144,993,289] differentiating *MUC13A* (the longer sequence) and *MUC13B* (the shortened sequence) and was performed as described by Ren et al. (2012).

Table 3.1: Information about the pigs used in the *MUC4* TaqMan assay, the Indel *MUC13* marker test, and the GWAS study.

F4R	Pig	Litter	Sow	Boar	Breed[a]	*MUC4*[b] g.8227G>C	Indel *MUC13*[c]	F4R	Pig	Litter	Sow	Boar	Breed[a]	*MUC4*[b] g.8227G>C	Indel *MUC13*[c]
$F4R^+$	V1	1	S1	B1	Hybrid	SS	BB	$F4R^+$	V61	42	S37	B22	LW	SR	AB
$F4R^+$	V2	1	S1	B1	Hybrid	SS	BB	$F4R^+$	V62	42	S37	B22	LW	SR	AB
$F4R^+$	V3	1	S1	B1	Hybrid	SS	AB	$F4R^+$	V63	45	S39	B22	LW	SR	AB
$F4R^+$	V4	1	S1	B1	Hybrid	SR	BB	$F4R^+$	V64	45	S39	B22	LW	SR	BB
$F4R^+$	V5	1	S1	B1	Hybrid	SS	AB	$F4R^+$	V65	45	S39	B22	LW	SR	BB
$F4R^+$	V6	1	S1	B1	Hybrid	RR	BB	$F4R^+$	V66	47	S41	B23	LW x BL	SR	AB
$F4R^+$	V7	2	S2	B2	Hybrid	RR	BB	$F4R^+$	V67	48	S42	B23	LW x BL	SR	BB
$F4R^+$	V8	2	S2	B2	Hybrid	SS	BB	$F4R^+$	V68	49	S43	B23	LW x BL	SR	BB
$F4R^+$	V9	2	S2	B2	Hybrid	SR	BB	$F4R^-$	V69	3	S3	B1	hybrid	SR	BB
$F4R^+$	V10	2	S2	B2	Hybrid	SS	BB	$F4R^-$	V70	4	S4	B3	hybrid	RR	AB
$F4R^+$	V11	3	S3	B1	Hybrid	SR	BB	$F4R^-$	V71	4	S4	B3	hybrid	RR	AB
$F4R^+$	V12	5	S5	B4	Hybrid	SR	BB	$F4R^-$	V72	4	S4	B3	hybrid	RR	AB
$F4R^+$	V13	5	S5	B4	Hybrid	SR	BB	$F4R^-$	V73	4	S4	B3	hybrid	RR	AB
$F4R^+$	V14	6	S6	B5	Hybrid	SR	BB	$F4R^-$	V74	4	S4	B3	hybrid	RR	AB
$F4R^+$	V15	7	S7	B6	Hybrid	SR	BB	$F4R^-$	V75	4	S4	B3	hybrid	RR	AB
$F4R^+$	V16	8	S8	B7	LW x P	SS	BB	$F4R^-$	V76	5	S5	B4	hybrid	RR	BB
$F4R^+$	V17	8	S8	B7	LW x P	SR	BB	$F4R^-$	V77	5	S5	B4	hybrid	RR	BB
$F4R^+$	V18	8	S8	B7	LW x P	SS	BB	$F4R^-$	V78	6	S6	B5	hybrid	RR	AB
$F4R^+$	V19	8	S8	B7	LW x P	SS	BB	$F4R^-$	V79	6	S6	B5	hybrid	RR	AB
$F4R^+$	V20	11	S9	B8	LW x BL	SR	BB	$F4R^-$	V80	7	S7	B6	hybrid	RR	AB
$F4R^+$	V21	12	S10	B9	LW	SR	AB	$F4R^-$	V81	8	S8	B7	LW x P	RR	AA
$F4R^+$	V22	13	S11	B8	LW x BL	SS	BB	$F4R^-$	V82	11	S9	B8	LW x BL	RR	AB
$F4R^+$	V23	14	S12	B8	LW x BL	SS	BB	$F4R^-$	V83	12	S10	B9	LW	RR	AA
$F4R^+$	V24	17	S14	B8	LW x BL	SR	AB	$F4R^-$	V84	13	S11	B8	LW x BL	RR	BB
$F4R^+$	V25	17	S14	B8	LW x BL	SS	BB	$F4R^-$	V85	14	S12	B8	LW x BL	RR	BB
$F4R^+$	V26	18	S15	B10	LW x P	SR	AB	$F4R^-$	V86	15	S13	B10	LW x P	RR	AA
$F4R^+$	V27	24	S19	B14	Hybrid	RR	BB	$F4R^-$	V87	17	S14	B8	LW x BL	RR	AB
$F4R^+$	V28	25	S20	B15	Hybrid	RR	BB	$F4R^-$	V88	18	S15	B11	LW x P	RR	AB
$F4R^+$	V29	25	S20	B15	Hybrid	SR	BB	$F4R^-$	V89	21	S16	B12	LW x BL	RR	AB
$F4R^+$	V30	25	S20	B15	Hybrid	RR	BB	$F4R^-$	V90	22	S17	B13	hybrid	RR	AB
$F4R^+$	V31	25	S20	B15	Hybrid	RR	BB	$F4R^-$	V91	23	S18	B13	hybrid	RR	BB
$F4R^+$	V32	26	S21	B15	Hybrid	SR	BB	$F4R^-$	V92	24	S19	B14	hybrid	RR	BB
$F4R^+$	V33	26	S21	B15	Hybrid	SR	BB	$F4R^-$	V93	25	S20	B15	hybrid	RR	BB
$F4R^+$	V34	27	S22	B15	Hybrid	SS	BB	$F4R^-$	V94	29	S24	B12	LW	RR	AA

F4R	Pig	Litter	Sow	Boar	Breed[a]	*MUC4*[b] g.8227G>C	Indel *MUC13*[c]	F4R	Pig	Litter	Sow	Boar	Breed[a]	*MUC4*[b] g.8227G>C	Indel *MUC13*[c]
F4R$^+$	V35	27	S22	B15	Hybrid	SR	BB	F4R$^-$	V95	30	S25	B16	LW x P	RR	AA
F4R$^+$	V36	27	S22	B15	Hybrid	SS	BB	F4R$^-$	V96	31	S26	B17	hybrid	RR	AB
F4R$^+$	V37	28	S23	B15	Hybrid	SR	BB	F4R$^-$	V97	34	S29	B17	hybrid	RR	AB
F4R$^+$	V38	28	S23	B15	Hybrid	SR	BB	F4R$^-$	V98	34	S29	B17	hybrid	RR	AB
F4R$^+$	V39	29	S24	B12	LW	SR	AB	F4R$^-$	V99	35	S30	B17	hybrid	RR	AB
F4R$^+$	V40	29	S24	B12	LW	SS	BB	F4R$^-$	V100	37	S32	B19	BL	RR	AB
F4R$^+$	V41	30	S25	B16	LW x P	SR	AB	F4R$^-$	V101	38	S33	B20	LW	RR	AA
F4R$^+$	V42	30	S25	B16	LW x P	SS	BB	F4R$^-$	V102	39	S34	B18	BL	RR	AB
F4R$^+$	V43	32	S27	B17	Hybrid	SR	BB	F4R$^-$	V103	40	S35	B21	LW	RR	AA
F4R$^+$	V44	32	S27	B17	Hybrid	RR	BB	F4R$^-$	V104	40	S35	B21	LW	RR	AA
F4R$^+$	V45	32	S27	B17	Hybrid	SS	BB	F4R$^-$	V105	40	S35	B21	LW	RR	AA
F4R$^+$	V46	33	S28	B17	Hybrid	SR	BB	F4R$^-$	V106	40	S35	B21	LW	RR	AA
F4R$^+$	V47	33	S28	B17	Hybrid	RR	AB	F4R$^-$	V107	41	S36	B22	LW	RR	AA
F4R$^+$	V48	34	S29	B17	Hybrid	SR	BB	F4R$^-$	V108	41	S36	B22	LW	RR	AB
F4R$^+$	V49	35	S30	B17	Hybrid	SR	BB	F4R$^-$	V109	41	S36	B22	LW	SR	AA
F4R$^+$	V50	35	S30	B17	Hybrid	RR	AB	F4R$^-$	V110	41	S36	B22	LW	SR	AA
F4R$^+$	V51	36	S31	B18	LW x BL	SS	BB	F4R$^-$	V111	42	S37	B22	LW	RR	AB
F4R$^+$	V52	36	S31	B18	LW x BL	SR	AB	F4R$^-$	V112	43	S38	B22	LW	RR	AB
F4R$^+$	V53	36	S31	B18	LW x BL	SR	AB	F4R$^-$	V113	43	S38	B22	LW	RR	AB
F4R$^+$	V54	37	S32	B19	BL	SR	AB	F4R$^-$	V114	43	S38	B22	LW	RR	AB
F4R$^+$	V55	37	S32	B19	BL	RR	BB	F4R$^-$	V115	43	S38	B22	LW	RR	AA
F4R$^+$	V56	38	S33	B20	LW	SR	AB	F4R$^-$	V116	45	S39	B22	LW	RR	AA
F4R$^+$	V57	38	S33	B20	LW	SR	AB	F4R$^-$	V117	46	S40	B23	LW	RR	AB
F4R$^+$	V58	39	S34	B18	BL	SR	BB	F4R$^-$	V118	46	S40	B23	LW	RR	AA
F4R$^+$	V59	39	S34	B18	BL	SR	BB	F4R$^-$	V119	46	S40	B23	LW	RR	AA
F4R$^+$	V60	42	S37	B22	LW	SR	BB	F4R$^-$	V120	48	S42	B23	LW x BL	SR	AB

[a] Breeds are defined as follows: 'LW' denotes Large White, 'BL' denotes Belgian Landrace, 'LW x BL' denotes Large White x Belgian Landrace crossbreds, 'LW x P' denotes Large White x Piétrain crossbreds, 'hybrid' denotes crossbreds of multiple breeds; [b]*MUC4* genotypes are defined as follows: 'SS' denotes homozygous susceptible (S allele is corresponding with the G allele), 'SR' denotes heterozygous susceptible (R allele is corresponding with the C allele), 'RR' denotes homozygous resistant; [c]*MUC13* alleles are defined as follows: 'AA' denotes homozygous for the *MUC13A* allele, 'AB' denotes both *MUC13A* and *MUC13B* alleles are present, 'BB' denotes homozygous for the *MUC13B* allele.

3.3.4. Expression of the MUC13 glycoprotein and F4ac binding proteins in BBMVs

In a recent study Nguyen et al. (2013) demonstrated that two high molecular weight (MW) glycoproteins (>250 kD), which bind F4ab and F4ac fimbriae, are consistently present in pigs with the *MUC4* homozygote or heterozygote susceptible genotype that show an immune response upon oral immunization with F4ac fimbriae. Here, these high MW glycoproteins were purified from F4ab/acR$^+$ BBMVs group I (Nguyen et al., 2013) using anion exchange (Pierce strong anion exchange spin column, Thermo Scientific) followed by gel filtration (HiPrep 16/60 Sephacryl S-200, GE Healthcare). Eluates were tested by one-dimensional immunoblotting for the presence of the high MW glycoproteins with biotinylated F4ac fimbriae as described in Nguyen et al. (2013).

A peptide (FPKIKVDISRGGQP) selected from the C-terminus of the porcine MUC13B sequence was synthesized and conjugated to keyhole limpet hemocyanin (KLH) (Genscript, Piscataway, USA) (Ren et al., 2012). The peptide conjugate emulsified in incomplete Freund's adjuvant was used to immunize rabbits. Immunoglobulin G was purified using a Hitrap™ Protein G HP column (GE Healthcare) (Van den Broeck et al., 1999b). These antibodies dissolved in PBS were used to stain immunoblots of BBMVs from F4ab/acR$^+$ pigs separated by SDS-PAGE under reducing and non-reducing conditions (Nguyen et al., 2013). In parallel, the binding pattern of biotinylated F4ac fimbriae to these blots was determined.

Additionally, MUC13 was precipitated from BBMVs by incubating 1 mg of the vesicles with 50 µg/ml anti-MUC13 antibodies for 1 h at 4 $^{\circ}$C followed by 50 µl protein A conjugated sepharose (Protein A Sepharose ™ CL-4B, GE Healthcare) for 1 h at 4 $^{\circ}$C. Precipitated MUC13 was eluted from the sepharose by boiling it for 10 min at 95 $^{\circ}$C in 100 µl SDS-PAGE reducing loading buffer. Subsequently SDS-PAGE and immunoblotting with F4ac fimbriae were performed to determine the F4ac binding pattern (Nguyen et al., 2013).

3.3.5. Genome-wide association study

DNA isolation of the blood samples for the genome-wide association study was performed as described by Dupuis et al. (2011). A total of 120 F4ab/acR phenotyped pigs were genotyped using the Porcine SNP60 BeadChip (Illumina) containing 62,163 SNPs, according to the manufacturer's protocol. The position of the SNPs was based on the current pig genome assembly (Sscrofa10.2). Quality scores were analyzed from allele

cluster definitions for each SNP as determined by the Illumina GenomeStudio Genotyping Module version 1.0 (Illumina). All genotype calls were extracted from the raw data with a minimum gencall score threshold of 0.2. An additional quality control was performed using PLINK (Purcell et al., 2007). SNPs were filtered with a call rate > 95 %, GenTrain Score < 0.7 and minor allele frequency (MAF) ≥0.01. A total of 6874 (11 %) SNPs were excluded from further analysis. Two SNPs, namely the *MUC4* SNP and the *MUC13* SNP, were added in our association analysis based on the g.8227G>C mutation of *MUC4* [Genbank: DQ848681] associated with F4ab/ac ETEC susceptibility and based on the indel of 68 bp in intron 2 of *MUC13* [GenBank: NC_010455.4; chr13: 144,993,222-144,993,289] (Jørgensen et al., 2004; Ren et al., 2012). Pigs were genotyped for the *MUC4* SNP by performing the *MUC4* TaqMan assay as described by Nguyen et al. (2013).

No samples were excluded due to frequency of missing genotypes > 5 %. Population stratification based on pair-wise identity-by-state (IBS) distances was also quantified using the PLINK software (Purcell et al., 2007). Multidimensional scaling (MDS) identified five clusters corresponding to breed with no genotypic outliers and the two F4ab/acR phenotypes were present in each cluster. These five clusters were used as stratification criteria for following association analysis.

Association between the SNPs and the F4ab/ac receptor-positive (F4ab/acR^+) status was assessed using the 2x2xK Cochran-Mantel-Haenszel (CMH) test for 5 clusters.

The Bonferroni correction and the Max(T) permutation procedure (10,000 permutations) within the breeds were used to correct for experiment-wise error rate (Table 3.2) (Zhang et al., 2012). Manhattan plots of the results were generated using Haploview (Figure 3.1) (Barret et al., 2005).

Table 3.2: Significant SNPs for F4ab/ac ETEC susceptibility.

Chr.	Position (bp)[a]	SNP name	Nearest gene[b]	Genomic control corrected *P*-values[c]	Bonferroni corrected *P*-values	*max*(T) empirical P-*values*
13	144,946,742	*ASGA0089965*	*MUC13* (45,916 bp)	1.29E-20	6.19E-21	1.00E-04
13	144,981,309	*ASGA0091537*	*MUC13* (11,349 bp)	1.29E-20	6.19E-21	1.00E-04
13	145,009,805	*ALGA0106330*	*MUC13* (within)	1.14E-16	5.97E-16	1.00E-04
13	144,299,267	*ASGA0095873*	*LMLN* (within)	9.24E-13	5.11E-11	1.00E-04
13	143,820,612	*MUC4*	*MUC4* (within)	2.87E-13	1.17E-11	1.00E-04
13	144,832,256	*ALGA0072075*	*HEG1* (31,668 bp)	3.17E-12	2.42E-10	1.00E-04
13	144,733,031	*ASGA0058925*	*SLC12A8* (within)	8.48E-12	8.37E-10	1.00E-04
13	145,772,058	*MARC0088848*	*HEG1* (within)	1.90E-11	2.31E-09	1.00E-04
13	145,732,401	*ASGA0058958*	*SLC12A8* (within)	2.80E-11	3.78E-09	1.00E-04
13	145,671,763	*ALGA0072105*	*SLC12A8* (within)	3.64E-11	5.26E-09	1.00E-04
13	143,656,188	*ASGA0058885*	*SLC12A8* (within)	5.09E-11	8.03E-09	1.00E-04
13	145,398,474	*MARC0045442*	*UMPS* (243582 bp)	1.51E-10	3.16E-08	1.00E-04
13	144,197,577	*ALGA0072072*	*LMLN* (within)	7.37E-10	2.33E-07	1.00E-04
13	144,167,475	*ALGA0072065*	*IQCG* (within)	7.37E-10	2.33E-07	1.00E-04
13	144,145,817	*ALGA0072067*	*IQCG* (within)	7.37E-10	2.33E-07	1.00E-04
13	144,946,317	*ALGA0122555*	*MUC13*(46,341 bp)	1.72E-09	6.77E-07	1.00E-04
13	144,993,222-144,993,289	Indel *MUC13*	*MUC13* (within)	1.72E-09	6.77E-07	1.00E-04
0	0	*ALGA0122702*	NA	2.49E-09	1.08E-06	1.00E-04
13	144,094,647	*MARC0089106*	*IQCG* (40,326 bp)	3.48E-09	1.65E-06	1.00E-04
13	144,126,389	*MARC0043596*	*IQCG* (8,584 bp)	3.48E-09	1.65E-06	1.00E-04
13	143,866,440	*ALGA0072062*	*MUC20* (within)	8.20E-09	4.85E-06	1.00E-04
13	147,328,480	*ALGA0072162*	*SEMA5B* (16,459 bp)	8.30E-09	4.92E-06	1.00E-04
13	147,911,293	*ISU10000469*	*CASR* (within)	5.42E-08	5.23E-05	0.0004
13	146,433,577	*H3GA0037388*	*PTPLB* (174,712)	5.64E-08	5.50E-05	0.0004
13	142,313,068	*ASGA0058867*	*APOD* (218,113)	6.64E-08	6.75E-05	0.0005999
13	145,414,240	*DIAS0000584*	*UMPS* (259,348 bp)	7.50E-08	7.87E-05	0.0006999

Chr.	Position (bp)[a]	SNP name	Nearest gene[b]	Genomic control corrected *P*-values[c]	Bonferroni corrected *P*-values	*max*(T) empirical P-*values*
13	143,624,457	*M1GA0017682*	*TNK2* (22,741 bp)	1.03E-07	0.00012	0.0011
13	143,618,378	*MARC0012378*	*TNK2* (28,820 bp)	1.03E-07	0.00012	0.0011
13	148,020,127	*MARC0031951*	*CASR* (85,108 bp)	2.38E-07	0.00034	0.0013
13	146,824,849	*H3GA0037402*	*PDIA5* (18,255 bp)	2.49E-07	0.00036	0.0013
13	143,638,483	*MARC0093203*	*TNK2* (8,715 bp)	3.11E-07	0.00047	0.0013
13	143,825,858	*ASGA0058906*	*MUC4* (within)	5.75E-07	0.00102	0.0025
13	147,415,740	*MARC0032449*	*HASPBAP1* (within)	7.35E-07	0.00139	0.0031
13	147,295,491	*DIAS0001133*	*SEMA5B* (within)	7.55E-07	0.00144	0.0034
13	145,473,321	*H3GA0037371*	*UMPS* (318,429)	7.88E-07	0.00152	0.0034
13	146,909,376	*M1GA0017695*	*PDIA5* (within)	9.70E-07	0.00197	0.0043
13	145,096,895	*ALGA0072090*	*ITGB5* (within)	1.50E-06	0.00339	0.005999
13	143,307,737	*H3GA0037321*	*SLC51A* (119 bp)	2.23E-06	0.00559	0.007899
0	0	*M1GA0027009*	NA	3.61E-06	0.01024	0.013
13	144,611,608	*MARC0067282*	*SLC12A8* (19,579 bp)	5.85E-06	0.0188	0.0201
13	144,488,410	*MARC0099692*	*ZBP-89* (28,751 bp)	5.85E1-06	0.0188	0.0201
13	147,927,896	*ASGA0059010*	*CASR* (within)	6.62E-06	0.02197	0.0232
13	144,781,809	*ASGA0058923*	*HEG1* (within)	7.22E-06	0.0245	0.0247
13	147,536,783	*MARC0105487*	*PARP15* (5097 bp)	1.22E-05	0.04727	0.0374

[a] Derived from porcine GenomeBuild 10.2; [b] The nearest annotated porcine gene to the significant SNP based on the porcine GenomeBuild 10.2. Numbers in parentheses indicate distance in base pairs (bp); [c] Genomic control corrected significance value. This is based on a simple estimation of the inflation factor based on median chi-square statistic (genomic inflation factor λ is 1.26636) (Purcell et al., 2007).

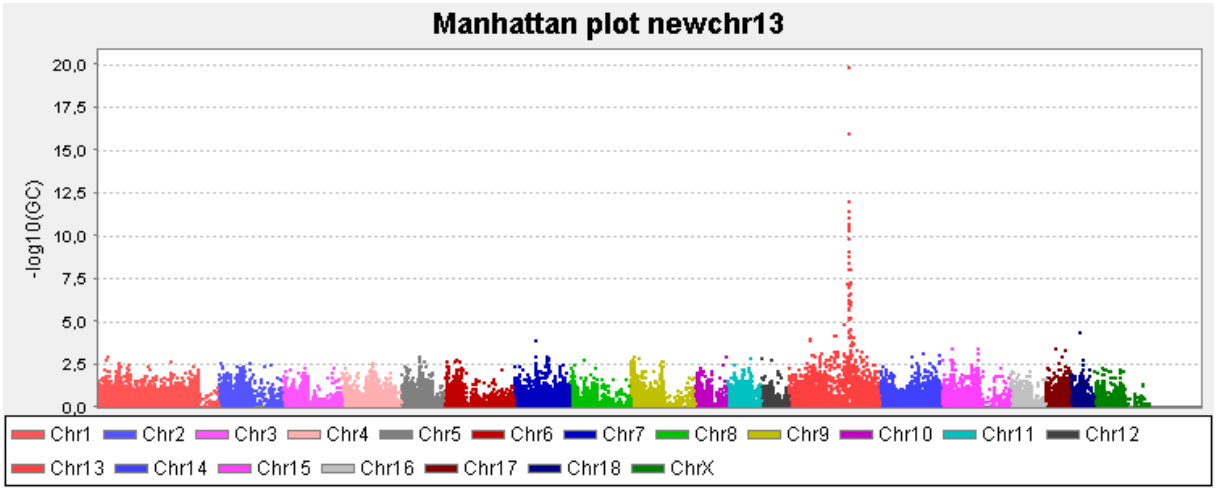

Figure 3.1: Manhattan plot obtained from the *P*-values for F4ab/ac ETEC susceptibility in 120 pigs. SNPs are plotted on the X-axis ordered by chromosomal position. Genome-wide -log(10) *P*-values adjusted to genomic control are plotted on the Y-axis.

3.4. Results

3.4.1. *MUC4* TaqMan assay and the Indel *MUC13* marker test

Using the *MUC4* TaqMan assay 17 pigs (25 %) were genotyped as homozygous susceptible (SS), 41 pigs (60.3 %) as heterozygous susceptible (SR) and 10 pigs (14.7 %) as homozygous resistant (RR) in the F4ab/ac ETEC strong adhesive group (n=68). In the F4ab/ac ETEC non-adhesive group (n=52), 4 SR pigs (7.7 %) and 48 RR pigs (92.3 %) were present (Table 3.1). The Indel *MUC13* marker test showed that 50 pigs (73.5 %) were homozygous for the *MUC13B* allele (BB) and 18 pigs (26.5 %) had a *MUC13A* as well as a *MUC13B* allele (AB) in the F4ab/ac ETEC strong adhesive group (n=68). In the F4ab/ac ETEC non-adhesive group (n=52), 8 pigs (15.4 %) were BB, 27 pigs (51.9 %) were AB and 17 pigs (32.7 %) were homozygous for the *MUC13A* allele (AA) (Table 3.1).

3.4.2. F4ac fimbriae do not bind to MUC13 glycoprotein of BBMVs

Separating BBMVs by SDS-PAGE under reducing (Figure 3.2.A) and non-reducing conditions (Figure 3.2.B) followed by immunoblotting with biotinylated F4ac fimbriae (lanes 2 and 4) or anti-MUC13 antibodies (lanes 3 and 5) revealed bands with a similar molecular weight (47, 34 and <25 KDa). These bands are not F4-specific bands since they are presented

in both $F4R^+$ and $F4R^-$ BBMV blots (Nguyen et al., 2013). However, it is also clearly demonstrated that the anti-MUC13 antibodies did not bind to the high MW glycoproteins which are specifically recognized by F4ac fimbriae being only present in $F4R^+$ and not in $F4R^-$ BBMV blots. Purifying and enriching the high MW glycoproteins by a combination of anion exchange chromatography and gel filtration did not change this (Figure 3.3 lane 3). Immunoprecipitation of MUC13 with the anti-MUC13 antibodies enriched a protein with MW of 110 kDa (Figure 3.4.A lane 2) but this was not recognized by F4ac fimbriae (Figure 3.4.A and B lane 2). Furthermore, not the immunoprecipitated fraction, but the non-immunoprecipitated fraction contained the F4-specific high MW glycoproteins (Figure 3.4.B lane 3). This result excludes MUC13 as one of the F4-specific high MW glycoproteins.

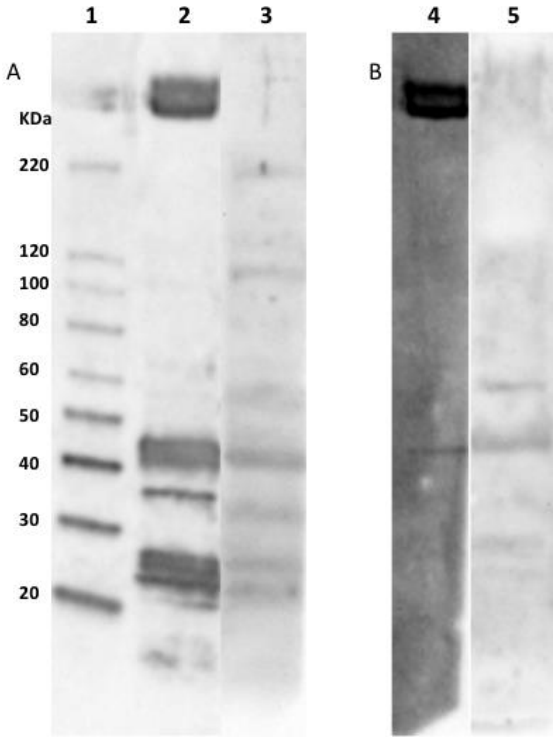

Figure 3.2: Immunoblotting of F4ab/acR^+ BBMVs with F4ac fimbriae (lane 2 and 4) or anti-MUC13 antibodies (lane 3 and 5). Proteins were separated under reducing (A) and non-reducing (B) conditions. F4ac fimbriae bound to the F4-specific high molecular weight glycoproteins (only present in $F4R^+$ pigs) and several non-specific F4-binding bands <50kDa (present in $F4R^+$ and $F4R^-$ pigs) (Nguyen et al., 2013). Anti-MUC13 antibodies recognized BBMV protein bands of 55, 47, 34 and <25 kDa under reducing and non-reducing conditions, but bands of 200, 110 kDa only under reducing conditions. Lane 1 = protein standards.

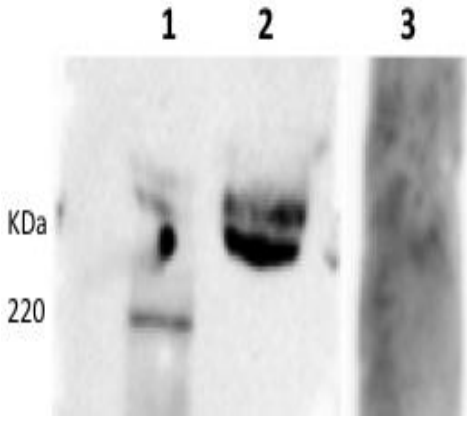

Figure 3.3: Absence of detection of intestinal MUC13 (lane 3) by anti-mucin 13 antibodies in the purified high molecular weight (MW) fraction of F4ab/acR$^+$ BBMVs. Purification occurred by anion exchange chromatography followed by gel filtration chromatography. Strong binding of F4ac fimbriae to the high MW glycoproteins can be seen in lane 2. Lane 1: protein standard.

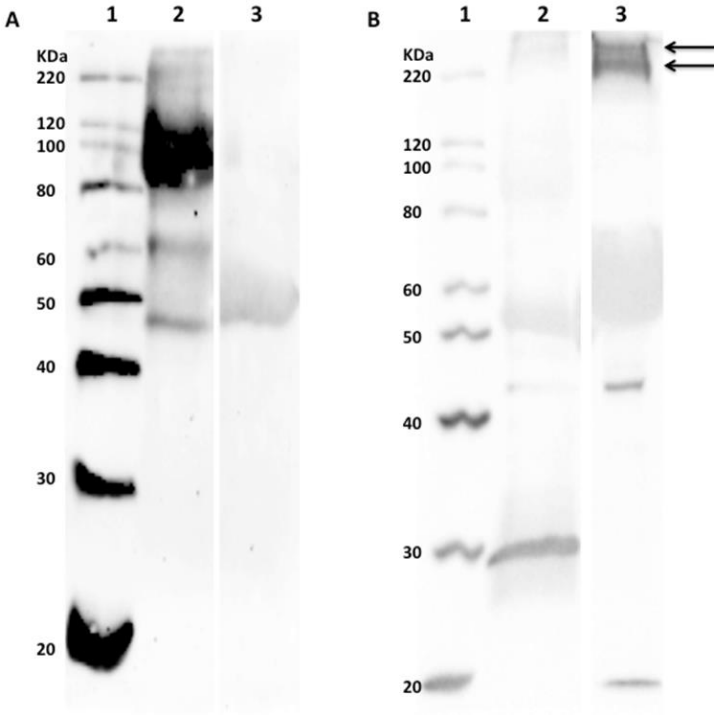

Figure 3.4: Absence of F4ac fimbriae binding to immunoprecipitated intestinal MUC13. Intestinal MUC13 was purified from 1 mg F4ab/acR$^+$ brush border membrane vesicles (BBMVs) by immunoprecipitation with anti-MUC13 antibodies and protein A sepharose. The eluate (lane 2) and the non-precipitated fraction (lane 3) were immunoblotted with anti-MUC13 antibodies (A) or with F4ac fimbriae (B). Immunoprecipitation enriched a protein with a band of 110 kDa but this was not recognized by F4ac fimbriae. The F4-specific high MW glycoproteins were found in the non-precipitated fraction. Lane 1 = protein standards. Arrows: position of the high MW glycoproteins.

3.4.3. Association analysis

The dataset presented in this article has been submitted to the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE57981. The results of the SNP association significance are visualized in a Manhattan plot of genome-wide-log(10) *P*-values adjusted to genomic control (Figure 3.1). Outliers displaying the highest association *P*-values of the SNPs are only visible on SSC13. After controlling for multiple testing, 42 SNPs of the Porcine SNP60 BeadChip reached genome-wide significance (<0.05): 40 SNPs are located within an interval of 5.7 Mb on SSC13 and the position of 2 SNPs (*ALGA0122702* and *M1GA0027009*) are unknown in the GenomeBuild 10.2 (Table 3.2). The 3 most significant SNPs are located in a region of 63,063 bp on SSC13 close to *MUC13*. SNP1 (*ASGA0089965*) and SNP2 (*ASGA0091537*) have the same *P*-value and are completely linked, only 1 strong adhesive F4Rab/ac pig (1/68) and 1 non-adhesive F4ab/ac pig (1/52) have a different genotype for these SNPs than expected. For SNP3 (*ALGA0106330*), 7 strong adhesive F4ab/ac pigs (7/68) and 1 non-adhesive F4ab/ac pig (1/52) have a different genotype than expected (Figure 3.5.B). A region of 213,267 bp (chr13: 144,810,100-145,023,367) was considered as the candidate region for F4ab/ac ETEC susceptibility. The borders of this region are 2 unassociated SNPs, namely *MARC0002946* (SNPa) and *ALGA0106230* (SNPb) (Figure 3.5.A).

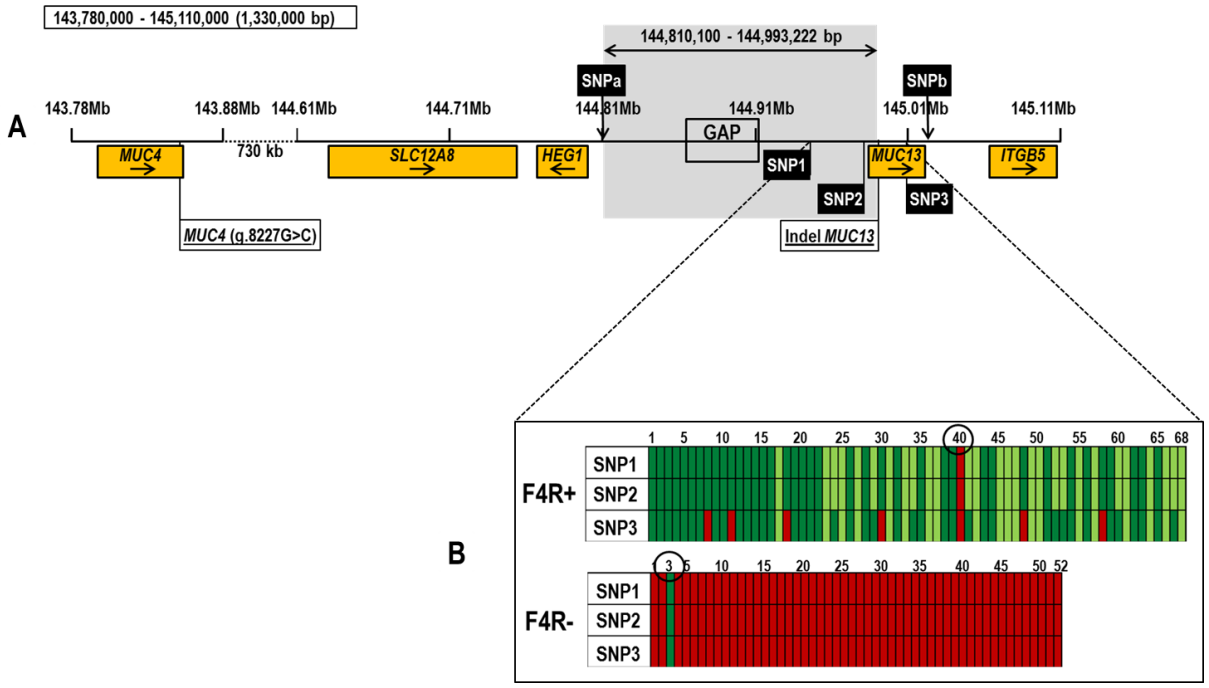

Figure 3.5: Schematic representation showing the identified candidate region (chr13: 144,810,100-144,993,222) of F4ab/ac ETEC susceptibility between *MARC0002946* (SNPa) and Indel *MUC13* marker on SSC13 (chr13: 143,780,000-145,110,000).
(A) SNP1 (*ASGA0089965*), SNP2 (*ASGA0091537*) and SNP3 (*ALGA0106330*) are the most significant SNPs in the association study. *MARC0002946* (SNPa) and *ALGA0106230* (SNPb) are not associated with F4ab/ac ETEC susceptibility. The orange boxes represent all the annotated genes in the 1.33 Mb region of SSC13. The gray box represent the candidate region where no annotated genes were found during the *in silico* comparative mapping.
(B) Schematic representation showing the genotypes of SNP1 (*ASGA0089965*), SNP2 (*ASGA0091537*) and SNP3 (*ALGA0106330*) of 68 strong adhesive ($F4R^+$) and 52 non-adhesive ($F4R^-$) for F4ab/ac ETEC.
For SNP1 and SNP3, the dark green boxes represent CC genotype, light green boxes represent CT genotype and red boxes represent TT genotype. For SNP2, the dark green boxes represent TT genotype, light green boxes represent CT genotype and red boxes represent CC genotype. Pig 40 ($F4aR^+$; See Table 3.1: V27) and pig 3 ($F4R^-$; See Table 3.1: V120) show different genotypes for the markers than expected.

3.5. Discussion

The results of the *MUC4* TaqMan assay confirm that the g.8227G>C mutation of *MUC4* is associated with F4ab/ac ETEC susceptibility (Jørgensen et al., 2004). Nevertheless, the genotypes, especially the RR genotype, were not completely consistent with the results of the *in vitro* villous adhesion assay. Ten of the RR pigs (14.7 %) showed adhesion towards F4ab/ac ETEC and 4 SR pigs (7.7 %) showed no adhesion. This result confirms earlier

findings that the g.8227G>C mutation is a marker but not the actual causative mutation (Rasschaert et al., 2007; Li et al., 2008).

Recently, *MUC13* was suggested as the causal gene for F4ac ETEC susceptibility (Ren et al., 2012). The MUC13 glycoprotein consists of an α and β subunit interconnected by covalent links which can be cleaved by proteases (Williams et al., 2001). The ß subunit refers to the C-terminal subunit containing the cytoplasmic tail and forms as homodimer. The α subunit carries weighty O-glycosylations, which are thought to be responsible for the F4 ETEC binding. Unlike MUC13A that lacks an O-glycosylation site, MUC13B is predicted to be heavily O-glycosylated and therefore would contain the binding site for F4ac ETEC (Ren et al., 2012). This hypothesis suggests that the *MUC13B* allele (present in F4ac ETEC susceptible pigs) is dominant over the *MUC13A* allele (present in the F4ac ETEC resistant pigs). By performing the Indel *MUC13* marker test in 120 F4ab/acR phenotyped pigs, the presence of the *MUC13B* allele was identified in 35 F4ab/ac ETEC non-adhesive pigs (67.3 %), rejecting the hypothesis (Table 3.1).

The predicted mass of the entire MUC13B protein is about 44 kDa based on the MUC13B sequence (Ren et al., 2012), but the expected mass on SDS gel should be much higher due to massive glycosylation and a gap in the domain rich in the amino acids Proline, Threonine and Serine (PTS domain) of the DNA sequence.

The peptide used to immunize rabbits was selected from the MUC13B sequence described by Ren et al. (2012). It is located in the cytoplasmic tail of the ß subunit and is predicted not to contain glycosylation sites based on DictyOGlyc and NetNGlyc (Ren et al., 2012). Consequently, in theory, the produced antibodies should bind to the cytoplasmic parts of MUC13 without interference of glycans at the glycosylated sites. Binding of the MUC13 antibodies to BBMV proteins was tested under non-reducing and reducing conditions so that both intact MUC13 and the individual monomers with digestion products should be observed. As expected, more bands were recognized in the reducing condition. Nguyen et al. (2013) demonstrated that the three F4 variants bound specifically the two high molecular weight (MW) glycoproteins (>250 kD), while the other bands (130, 110, 75, 64, 50, 47, 43, 40, 34, 32, and <25 kDa) were also recognized by F4 fimbriae but not specifically since they were present in both $F4R^+$ and $F4R^-$ BBMV blots (Nguyen et al., 2013). In the present study, anti-MUC13 antibodies and F4ac fimbriae also bound the 47, 34, and <25 KDa bands, which are not F4-specific. However, in none of the conditions MUC13 antibodies recognized the F4-specific high MW glycoproteins. Furthermore, the immunoprecipitated MUC13 was not recognized by F4ac fimbriae. These results exclude MUC13 as one of the F4-binding high

molecular weight glycoproteins and F4 fimbriae do not bind specifically to MUC13. Also by using NetOGlyc 4.0 which is a mucin-type (GalNAc) glycosylation predictor trained on mammalian protein sequences (Steentoft et al., 2003), it was predicted that MUC13A (AEO00194.1) as well as MUC13B (AEO00200.1) are O-glycosylated.

Based on our results, we could conclude that the 2 MUC13 variants are not responsible for susceptibility towards F4ab/ac ETEC.

Our GWAS study demonstrates a strong association between 3 SNPs (*ASGA0089965*, *ASGA0091537*, *ALGA0106330*) and the F4ab/acR locus. In 118 pigs (98.33 %), 2 markers, namely SNP1 (*ASGA0089965*) and SNP2 (*ASGA0091537*), were in complete linkage disequilibrium (LD) with the F4ab/acR locus. For 2 pigs (1 strong adhesive F4ab/ac pig and 1 non-adhesive F4ab/ac pig), the F4ab/acR phenotype showed the opposite genotype for these markers (Figure 3.5.B). It is possible that the causal mutation is located proximal of SNP1 or that the F4ab/acR phenotype is more complex and regulated by multiple factors.

Our results confirm that *MUC4* and *MUC13* are not completely associated with F4ab/ac ETEC susceptibility. In our association analysis, the *MUC4* marker (g.8227G>C mutation) and the Indel *MUC13* marker [GenBank: NC_010455.4; chr13: 144,993,222-144,993,289] were included. The *MUC4* marker showed a lower significant *P*-value than 4 markers located more distal from the g.8227G>C mutation (Table 3.2; Figure 3.5.A). The low significant *P*-value of the Indel *MUC13* marker (Table 3.2; Figure 3.5.A) confirmed our previous results that *MUC13* is not the causal gene for F4ab/ac ETEC susceptibility. Four markers of the Porcine SNP60 BeadChip (*ALGA0072075* [GenBank: NC_010455.4; chr13: 144,832,256], *ALGA0106330* (SNP3), *DIAS0000584* [GenBank: NC_010455.4; chr13: 145,414,240] and *MARC0006918* [unknown position]), and 2 additional markers (*MUC13*-226 [GenBank: NC_010455.4; chr13: 145,010,437] and *MUC13*-813 [GenBank: NC_010455.4; chr13: 145,016,914]) were in complete LD with the F4ab/acR locus in a Swiss experimental herd (Rampoldi et al., 2011). Except for one sow and some of her offspring, markers *ALGA0106330* (SNP 3), *MUC13*-226 and *MUC13*-813 were not in LD with the F4ab/acR locus (Rampoldi, 2013). In our study, the markers *ALGA0072075*, *ALGA0106330* (SNP 3), and *DIAS0000584* all had statistically significant *P*-values (Table 3.2), but showed a weaker F4ab/acR association than SNP1 and SNP2. The unmapped marker *MARC0006918* was not associated with the F4ab/acR locus in our study. Because *MUC13*-226 and *MUC13*-813 are positioned distal to the Indel *MUC13* marker, we refined the candidate region to 183,122 bp (chr13: 144,810,100-144,993,222; Figure 3.5.A).

Also, our 3 most significant SNPs were significant in another F4ab/ac ETEC susceptibility study (Fu et al., 2012), but other more proximal located SNPs (*MARC0012378*, *M1GA0017682*, *ALGA0072075*) in this study showed a higher significant *P*-value.
This result could be due to the fact that they included the F4ab/acR phenotype weak adhesive as non-adhesive pigs.
The refined candidate region of 183,122 bp (chr13: 144,810,100-144,993,222) on SSC13 is almost in complete linkage with the F4ab/acR phenotype. No porcine annotated genes are present in our candidate region based on the available genome sequence of the porcine GenomeBuild 10.2 (Figure 3.5.A). Comparison of the candidate region with orthologous regions of human, mouse and rat genomes (GenBank: NC_000003.11 (human); NC_0000826 (mouse); NC_005110.3 (rat)) using BLAST analysis was performed to identify non-annotated porcine genes as well as identifying the gap sequence (between NW_003611795.1 and NW_003617796.1) present in the candidate region (Benson et al., 2009). Performing the interspecies comparison, no new non-annotated porcine genes were identified and we were unable to identify the gap sequence. The presence of a porcine orphan gene in the candidate region cannot be ruled out. Fang et al. (2012) identified 240 orphan genes with no counterpart in any other organism (human, horse, dog, cat, cattle, rat and mouse) during analysis of the genome sequence of the Wuzhishan miniature pig. Also, the candidate region could contain a *trans*-acting element interacting with a distant gene influencing F4 ETEC adherence in pigs.
Based on previous F4 ETEC binding studies, it was concluded that the carbohydrate moiety of glycoconjugates appears to be necessary for establishing adhesion with the F4 adhesin (Erickson et al., 1992; Grange and Mouricout, 1995; Grange et al., 1999; Coddens et al., 2011; Melkebeek et al., 2012). We propose that a porcine orphan gene or a *trans*-acting element present in the candidate region (chr13: 144,810,100-144,993,222) has an effect on the glycosylation of F4 binding proteins and therefore determines the F4 ETEC susceptibility in pigs.

3.6. Acknowledgements

The authors wish to thank Linda Impe, Dominique Vander Donckt and Ruben Van Gansbeke for excellent technical assistance. This work was supported by an agricultural research grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen).

3.7. References

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 21:263-265.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. Nucleic Acids Res. 37:D26-31.

Bijlsma IG, Bouw J. 1987. Inheritance of K88-mediated adhesion of *Escherichia coli* to jejunal brush borders in pigs: a genetic analysis. Vet Res Commun. 11:509-518.

Coddens A, Valis E, Benktander J, Ångström J, Breimer ME, Cox E, Teneberg S. 2011. Erythrocyte and porcine intestinal glycosphingolipids recognized by F4 fimbriae of enterotoxigenic *Escherichia coli*. PloS ONE. 6:e23309.

Cox E, Houvenaghel A. 1993. Comparison of the *in vitro* adhesion of K88, K99, F41 and P987 positive *Escherichia coli* to intestinal villi of 4-week-old to 5-week-old pigs. Vet Microbiol. 34:7-18.

Dupuis MC, Zhang Z, Druet T, Denoix JM, Charlier C, Lekeux P, George M. 2011. Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. Mamm Genome. 22:613-620.

Edfors-Lilja I, Gustafsson U, Duval-Iflah Y, Ellergeen H, Johansson M, Juneja RK, Marklund L, Andersson L. 1995. The porcine intestinal receptor for *Escherichia coli* K88ab, K88ac: regional localization on chromosome 13 and influence of IgG response to the K88 antigen. Anim Genet. 26:237-242.

Erickson AK, Willgohs JA, McFarland SY, Benfield DA, Francis DH. 1992. Identification of two porcine brush border glycoproteins that bind the K88ac adhesin of *Escherichia coli* and correlation of these glycoproteins with the adhesive phenotype. Infect Immun. 60:983-988.

Fairbrother JM, Nadeau E, Gyles CL. 2005. *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. Anim Health Res Rev. 6:17-39.

Fang X, Mu Y, Huang Z, Li Y, Han L, Zhang Y, Feng Y, Chen Y, Jiang X, Zhao W, Sun X, Xiong Z, Yang L, Liu H, Fan D, Mao L, Ren L, Liu C, Wang J, Li K, Wang G, Yang S, Lai L, Zhang G, Li Y, Wang J, Bolund L, Yang H, Wang Jian, Feng S, Li S, Du Y. 2012. The sequence and analysis of a Chinese pig genome. GigaScience. 1:16.

Fu WX, Yang L, Xin L, Xiao-Yan N, Xiang-Dong D, Liu JF, Zhang Q. 2012. A genome-wide association study identifies two novel promising candidate genes affecting *Escherichia coli* F4ab/ac susceptibility in swine. PLoS ONE. 7:e32127.

Gibbons RA, Sellwood R, Burrows M, Hunter PA. 1977. Inheritance of resistance to neonatal *E. coli* diarrhea in the pig: examination of the genetic system. Theor Appl Genet. 51:65-70.

Grange PA, Mouricout MA. 1995. Transferrin associated with the porcine intestinal mucosa is a receptor specific for K88ab fimbriae of *Escherichia coli*. Infect Immun. 64:606-610.

Grange PA, Erickson AK, Levery SB, Francis DH. 1999. Identification of an intestinal neutral glycosphingolipid as a phenotypic-specific receptor for K88ad fimbrial adhesin of *Escherichia coli*. Infect Immun. 37:165-172.

Guérin G, Duval-Iflah Y, Bonneau M, Bertaud M, Guillaume P, Ollivier L. 1993. Evidence for linkage between K88ab, K88ac intestinal receptors to *Escherichia coli* and transferrin loci in pigs. Anim Genet. 24:393-396.

Jørgensen CB, Cirera S, Anderson SI, Archibald AL, Raudsepp T, Chowdhary B, Edfors-Lilja I, Andersson L, Fredholm M. 2003. Linkage and comparative mapping of the locus controlling susceptibility towards *E. coli* F4ab/ac diarrhoea in pigs. Cytogenet Genome Res. 102:157-162.

Jørgensen CB, Cirera S, Archibald AL, Anderson L, Fredholm M, Edfors-Lilja I. 2004. Porcine polymorphisms and methods for detecting them. International application published under the patent cooperation treaty (PCT), PCT/DK2003/000807 or WO2004/048606A2.

Li Y, Qiu X, Li H, Zhang Q. 2007. Adhesive patterns of *Escherichia coli* F4 in piglets of three breeds. J Genet Genomics. 34:591-599.

Li H, Li Y, Qiu X, Niu X, Liu Y, Zhang Q. 2008. Identification and screening of gene(s) related to susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac in piglets. Asian-Austral J Anim. 21:489-493.

Melkebeek V, Rasschaert K, Bellot P, Tilleman K, Favoreel H, Deforce D, De Geest BG, Goddeeris BM, Cox E. 2012. Targeting aminopeptidase N, a newly identified receptor for F4ac fimbriae, enhances the intestinal mucosal immune response. Mucosal Immunol. 5:635-645.

Nguyen VU, Goetstouwers T, Coddens A, Van Poucke M, Peelman L, Deforce D, Melkebeek V, Cox E. 2013. Differentiation of F4 receptor profiles in pigs based on their mucin 4 polymorphism, responsiveness to oral F4 immunization and *in vitro* binding of F4 to villi. Vet Immunol Immunopathol. 152:93-100.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Gen. 81:559-575.

Python P, Jörg H, Neuenschwander S, Hagger C, Stricker C, Bürgi E, Bertschinger HU, Stranzinger G, Vögeli P. 2002. Fine-mapping of the intestinal receptor locus for enterotoxigenic *Escherichia coli* F4ac on porcine chromosome 13. Anim Genet. 33:441-447.

Steentoft C, Vakhrushev SY, Vester-Christensen MB, Schjoldager KT, Kong Y, Bennett EP, Mandel U, Wandall H, Levery SB, Clausen H. 2003. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 32:1478-1488.

Rampoldi A, Jacobsen MJ, Bertschinger HU, Joller D, Bürgi E, Vögeli P, Andersson L, Archibald AL, Fredholm M, Jørgensen CB, Neuenschwander S. 2011. The receptor locus for *Escherichia coli* F4ab/F4ac in the pig maps distal to the *MUC4-LMLN* region. Mamm Genome. 22:122-129.

Rampoldi A. 2013. Approach to the gene controlling the porcine receptor for *Escherichia coli* with fimbriae F4ab/F4ac and inheritance of the receptor for F4ad. ETH. Available: http://dx.doi.org/10.3929/ethz-a-009907149.

Rasschaert R, Verdonck F, Goddeeris B, Duchateau L, Cox E. 2007. Screening of pigs resistant to F4 enterotoxigenic *Escherichia coli* (ETEC) infection. Vet Microbiol. 123:249-253.

Ren J, Yan X, Ai H, Zhang Z, Huang X, Ouyang J, Yang M, Yang H, Han P, Zeng W, Chen Y, Guo Y, Xiao S, Ding N, Huang L. 2012. Susceptibility towards enterotoxigenic *Escherichia coli* F4ac diarrhea is governed by the MUC13 gene in pigs. PloS ONE. 7:e445739.

Schroyen M, Stinckens A, Verhelst R, Niewold T, Buys N. 2012. The search for the gene mutations underlying enterotoxigenic *Escherichia coli* F4ab/ac susceptibility in pigs: a review. Vet Res. 43:70.

Van den Broeck W, Cox E, Goddeeris BM. 1999a. Receptor-specific binding of purified F4 isolated villi. Vet Microbiol. 68:255-263.

Van den Broeck W, Cox E, Goddeeris BM. 1999b. Induction of immune responses in pigs following oral administration of purified F4 fimbriae. Vaccine. 17:2020-2029.

Van Poucke M, Vandesompele J, Mattheeuws M, Van Zeveren A, Peelman LJ. 2005. A dual fluorescent multiprobe assay for prion protein genotyping in sheep. BMC Infect Dis. 5:13.

Wang J, Jiang SW, Chen XH, Liu ZL, Peng J. 2006. Prevalence of fimbrial antigen (K88 variants, K99 and 987P) of enterotoxigenic *Escherichia coli* from neonatal and postweaning piglets with diarrhea in central China. Asian-Aust J Anim Sci. 19:1342-1346.

Wang Y, Ren J, Lan L, Yan X, Huang X, Peng Q, Tang H, Zhang B, Ji H, Huang L. 2007. Characterization of polymorphisms of transferrin receptor and their association with susceptibility to ETEC F4ab/ac in pigs. J Anim Breed Genet. 124:225-229.

Williams SJ, Wreschner DH, Tran M, Eyre HJ, Sutherland GR, McGuckin MA. 2001. MUC13, a novel human cell surface mucin expressed by epithelial and hemopoietic cells. J Biol Chem. 276:18327-18336.

Zhang W, Zhao M, Ruesch L, Omot A, Francis D. 2007. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. Vet Microbiol. 123:145-152.

Zhang B, Ren J, Yan X, Huang X, Ji H, Peng Q, Zhang Z, Huang L. 2008. Investigation of the porcine *MUC13* gene: isolation, expression, polymorphisms and strong association with susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac. Anim Genet. 39:258-266.

Zhang X, Huang S, Zhang Z, Wang W. 2012. Chapter 10: Mining genome-wide genetic markers. PLoS Comput Biol. 8:e1002828.

Yan X, Huang X, Ren J, Zou Z, Yang S, Ouyang J, Zeng W, Yang B, Xiao S, Huang L. 2009. Distribution of *Escherichia coli* F4 adhesion phenotypes in pigs of 15 Chinese and Western breeds and a White Duroc x Erhualian intercross. J Med Microbiol. 58:1112-1117.

CHAPTER 4

F4ac-RELATED EXPRESSION ANALYSIS OF THE AMINOPEPTIDASE N GENE IN PIGS

Based on Journal of Animal Science 92:1866-1873 (2014).

Tiphanie Goetstouwers[1], Mario Van Poucke[1], Ut Nguyen Van[2], Vesna Melkebeek[2], Annelies Coddens[2], Dieter Deforce[3], Eric Cox[2] and Luc J. Peelman[1]

[1] Laboratory of Animal Genetics, Faculty of Veterinary Medicine, Ghent University, Heidestraat 19, B-9820 Merelbeke, Belgium

[2] Laboratory of Immunology, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, B-9820 Merelbeke, Belgium

[3]Laboratory of Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium

4. F4ac-RELATED EXPRESSION ANALYSIS OF THE AMINOPEPTIDASE N GENE IN PIGS

4.1. Abstract

Intestinal infections with F4 enterotoxigenic *Escherichia coli* (F4 ETEC) are worldwide an important cause of diarrhea in neonatal and recently weaned pigs. F4 fimbriae mediate the adherence of F4 ETEC to the small intestine by binding to specific receptors. Porcine aminopeptidase N (ANPEP) was recently identified as a new F4ac receptor. Because *ANPEP* is located on chromosome 7 (SSC7) and the causal mutation for F4ab/ac susceptibility is shown to be located on chromosome 13 (SSC13) (but not yet identified), we investigated if F4ac susceptibility is determined by differential expression of *ANPEP*, regulated by the causal mutation in a *trans*-acting manner. The expression of *ANPEP* was investigated by RT-PCR and sequencing, using three F4ac receptor-positive ($F4acR^+$) and two F4ac receptor-negative ($F4acR^-$) pigs, which were F4ac phenotyped based on the oral immunization assay with F4ac fimbriae and the *in vitro* villous adhesion assay for F4ac ETEC. No F4ac adhesion phenotype explaining differential gene expression was found in the small intestine of $F4acR^+$ and $F4acR^-$ pigs. In conclusion, we hypothesize that the differences in F4ac binding to ANPEP might be due to modifications in its carbohydrate moieties by a genetic variation involved in the glycosylation of ANPEP.

4.2. Introduction

Enterotoxigenic *Escherichia coli* (ETEC) expressing F4 fimbriae are worldwide a major cause of diarrhea in neonatal and recently weaned pigs (Do et al., 2006; Zhang et al., 2007; Amezcua et al., 2008; Madoroba et al., 2009; De la Fe Rodriguez et al., 2011). The bacteria colonize the small intestine of the pig by binding with their F4 fimbriae to specific receptors and produce enterotoxins, inducing diarrhea (Jones and Rutter, 1972; Alexander, 1994). The F4 fimbriae, composed of major and minor subunit structures, exist in three serological variants, namely F4ab, F4ac and F4ad (Orskov et al., 1964; Guinee and Jansen, 1979; Mooi and de Graaf, 1985). These variants differ in the amino acid composition of the major fimbrial subunit FaeG, which has adhesive properties and recognize glycoconjugates on the surface of enterocytes (Kearns and Gibbons et al., 1979; Sellwood, 1980; Bakker et al., 1992; Van den Broeck et al., 1999a). Although different F4 receptor profiles were observed due to different F4 ETEC adhesion phenotypes, no causal mutation in previously proposed candidate genes has yet been identified (Ren et al., 2012; Schroyen et al., 2012).

Recently, aminopeptidase N (ANPEP) has been found to act as an endocytotic F4ac receptor by comparative proteomic analysis of the brush border proteins in F4ac receptor-positive ($F4acR^+$) and F4ac receptor-negative ($F4acR^-$) pigs (Melkebeek et al., 2012). ANPEP, belonging to the M1 family of zinc metallopeptidase, is a 936 amino acid membrane glycoprotein and is widely expressed on the surface of various cell types, including porcine enterocytes (Delmas et al., 1992; Rawlings and Barrett., 1993; Olsen et al., 1997). The gene encoding ANPEP [GenBank: NC_010449.4] is located on SSC7 and is composed of 20 exons (Poulsen et al., 1991). Although the locus controlling F4ab/ac ETEC susceptibility has been mapped on SSC13, it is possible that variation in gene expression (no expression or expression of an isoform to which binding is impossible) of *ANPEP*, localized on SSC7, is regulated by *trans*-acting factors present in the candidate region on SSC13 and therefore explain the F4ac ETEC binding profile to ANPEP in F4ac phenotyped pigs (Douglas and Wood, 2011; Ren et al., 2012; Schroyen et al., 2012; Nguyen and Seoighe, 2013).

4.3. Material and methods

4.3.1. Animals, sample collection and F4ac phenotyping

Forty-six mixed-breed pigs from seven different litters were phenotyped for the presence or absence for the F4ac receptor (Table 4.1). Blood samples for DNA analysis were collected before euthanasia at 6-18 weeks of age in EDTA blood tubes and stored at -20 °C. After euthanasia, mid-jejunum samples for RNA analysis were collected, washed 3 times with Krebs–Henseleit buffer (0.12 M NaCl, 0.014 M KCl, 0.001 M KH_2PO_4, 0.025 M $NaHCO_3$, pH 7.4). Next, they were frozen in liquid nitrogen and stored at -80 °C. Mid-jejunum samples for the *in vitro* villous adhesion assay were washed twice with ice cold Krebs-Henseleit buffer, followed by one washing step with Krebs–Henseleit buffer containing 1 % (v/v) formaldehyde. Villi were then scraped from the mucosa and stored as mentioned in Van den Broeck et al. (1999b).
For the oral immunization, pigs were orally given 1 mg of F4ac fimbriae (strain Gis26) in 10 ml phosphate buffered saline (PBS) on three consecutive days and once again at 15 days post primary immunization. Blood was collected before the first immunization and after the second immunization at 15 and 21 days from the jugular vein to determine seropositivity via an F4ac-specific ELISA (Van den Broeck et al., 1999b).
The presence of the F4ac receptor was determined by performing the *in vitro* villous adhesion assay for the F4ac variant (Van den Broeck et al., 1999b; see Table 4.1).
Experimental and animal management procedures were approved by the animal care and ethics committee of the Faculty of Veterinary Medicine at Ghent University (EC2010/042).

4.3.2. RNA isolation and cDNA synthesis

Total RNA isolation of mid-jejunum samples of five pigs (pig 5, 8, 15, 33, 36) was performed by using the Aurum™ Total RNA Fatty and Fibrous Tissue kit (Bio-Rad) according to the instructions of the manufacturer. Residual DNA was removed with an on-column DNase treatment. The concentration of total RNA (between 400 and 1600 ng/µl) and the $OD_{260/280}$ ratio for purity (between 2.09 and 2.14) were measured by the Nanodrop ND-1000 Spectrophotometer (Isogen). RNA quality was assessed by evaluation of the 28S and the 18S ribosomal bands from 1 µg of total RNA on a 0.8 % agarose gel stained with ethidium bromide. All RNA samples were confirmed to be DNA-free by a minus RT-PCR with

SscrANPEP-F5/R5 primer pair on 20 ng RNA (Table 4.2). Positive and negative controls were included. The PCR product was examined by gel electrophoresis.

Then cDNA was synthesized from 1 µg high quality RNA with the Improm-II Reverse Transcriptase kit (Promega) using Random and Oligo dT primers (each 0.5 µg per reaction), verified by PCR (similar to the minus RT-PCR but with 10× diluted cDNA as a template and 40 PCR cycles) and examined by gel electrophoresis.

Table 4.1: Detailed list of the pigs used in this study.

Information pig				**F4ac binding profiling**		**Information pig**				**F4ac binding profiling**	
Pig	Litter	Sow	Boar	IR[1]	F4ac adhesion[2]	Pig	Litter	Sow	Boar	IR[1]	F4ac adhesion[2]
1	1	Sow A	Boar A	+	+++	24	2	Sow B	Boar B	+	+++
2	1	Sow A	Boar A	+	+++	25	2	Sow B	Boar B	+	+
3	1	Sow A	Boar A	+	+++	26	4	Sow D	Boar C	+	+++
4	1	Sow A	Boar A	+	+++	27	4	Sow D	Boar C	+	+++
5	1	Sow A	Boar A	+	+++	28	4	Sow D	Boar C	+	+++
6	2	Sow B	Boar B	+	+++	29	4	Sow D	Boar C	+	+
7	2	Sow B	Boar B	+	+++	30	5	Sow E	Boar D	+	+
8	2	Sow B	Boar B	+	+++	31	7	Sow G	Boar F	+	+
9	3	Sow C	Boar A	+	+++	32	4	Sow D	Boar C	-	-
10	3	Sow C	Boar A	+	+++	33	4	Sow D	Boar C	-	-
11	5	Sow E	Boar D	+	+++	34	4	Sow D	Boar C	-	-
12	5	Sow E	Boar D	+	+++	35	4	Sow D	Boar C	-	-
13	5	Sow E	Boar D	+	+++	36	5	Sow E	Boar D	-	-
14	5	Sow E	Boar D	+	+++	37	5	Sow E	Boar D	-	-
15	5	Sow E	Boar D	+	+++	38	4	Sow D	Boar C	+	-
16	6	Sow F	Boar E	+	+++	39	4	Sow D	Boar C	+	-
17	6	Sow F	Boar E	+	+++	40	6	Sow F	Boar E	+	-
18	6	Sow F	Boar E	+	+++	41	6	Sow F	Boar E	+	-
19	7	Sow G	Boar F	+	+++	42	7	Sow G	Boar F	+	-
20	7	Sow G	Boar F	+	+++	43	3	Sow C	Boar A	+	-
21	7	Sow G	Boar F	+	+++	44	4	Sow D	Boar C	-	+++
22	1	Sow A	Boar A	+	+++	45	4	Sow D	Boar C	-	+++
23	2	Sow B	Boar B	+	+	46	5	Sow E	Boar D	-	+++

[1]IR (immune response) is defined as follows: '+' denotes positive immune response upon oral immunization with F4ac fimbriae, '-' denotes negative immune response.

[2]F4ac ETEC adhesion is defined as follows: '+++' denotes strong adhesion, '+' denotes weak adhesion, '-' denotes no adhesion.

4.3.3. Expression analysis and sequencing the complete coding sequence of *ANPEP* by RT-PCR

cDNA was used for amplifying the whole coding sequence of *ANPEP* in four overlapping amplicons with four sets of primer pairs (SscrANPEP-F1/R1, -F2/R2, -F3/R3 and -F4/R4) designed with Primer3Plus (Untergasser et al., 2007; Table 4.2) in three F4acR$^+$ piglets (pig 5, 8, 15) and two F4acR$^-$ piglets (pig 33, 36). The composition of the PCR mix and the PCR conditions with 40 PCR cycles were similar as described above. PCR products were examined by gel electrophoresis for differential gene expression and alternative splicing.
After purification of the PCR products with Exonuclease I (4 U) and Antarctic Phosphatase (2 U) (New England Biolabs) at 37 °C for 30 min and 80 °C for 15 min, 1 µl PCR product was directly sequenced with 2 pmol of the corresponding *ANPEP* primers and an additional sequence primer (Table 4.2) using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The sequence reactions were analyzed on a 16-capillary 3130xl DNA Analyzer (Applied Biosystems) according to the manufacturer's protocol. Sequences were processed using BioEdit software (Hall et al., 1999).

Table 4.2: Primer pairs for amplification and sequencing of *ANPEP* (Genbank: NM_214277.1).

Primer	**F-primer (5'→3')** **R- primer (5'→3')**	**Ta (°C)** **Length cDNA / gDNA (bp)**	**Exon**
SscrANPEP-F1 SscrANPEP-R1	GCTCCCTTCTCACCCTCACCAT AGTCGGGCAAGGCAATCTGG	68 1050 / -	EX1→EX5
SscrANPEP-R1b	GGACGATGCTTTTGCCCTTGA	Seq primer	EX1
SscrANPEP-F2 SscrANPEP-R2	TGGCCTACATCGTGAGCGAGTT AGATGAGATGGGAACAATCCAGAGGT	64 935 / -	EX3→EX11
SscrANPEP-F3 SscrANPEP-R3	AGATGGGCTTCCCCGTCATC CTGTTCAGGAGCCAGACCTCGTT	68 895 / -	EX10→EX17
SscrANPEP-F4 SscrANPEP-R4	GCGGTCCACCATCTACTGCAA TTCTCAGGTTCAGGGTCTTTATGGAA	64 900 / -	EX17→EX20
SscrANPEP-F5 SscrANPEP-R5	GAATGACCTGTGGCTGAATGAGG GAGTCAAACATCTCGCTGATCTGG	66 210 / 307	EX6→EX7

4.4. Results

4.4.1. F4ac phenotyping in pigs

To obtain a reliable F4ac phenotype, two tests were performed in 46 pigs, the oral immunization assay with F4ac fimbriae and the *in vitro* villous adhesion assay for F4ac ETEC (Table 4.1). In this study, 31 seropositive pigs showed adhesion towards F4ac ETEC and six seronegative pigs showed no F4ac ETEC adhesion. Six seropositive pigs showed no F4ac ETEC adhesion and three seronegative pigs showed F4ac adhesion. Five pigs were selected for the RT-PCR experiments: three F4acR$^+$ pigs (seropositive and strong F4ac ETEC adhesion) and two F4acR$^-$ pigs (seronegative and no F4ac ETEC adhesion).

4.4.2. Expression analysis and sequencing the complete coding sequence of *ANPEP* by RT-PCR

Expression of *ANPEP* in the small intestine was analyzed via gel electrophoresis of the RT-PCR products. For all four primer pairs, covering the complete coding sequence, no obvious expression difference (expression or not, or expression of isoforms with a detectable amplicon size difference) was identified between the F4acR$^+$ and the F4acR$^-$ pigs, as illustrated in Figure 4.1 for primer pair SscrANPEP-F3/R3. All amplicons were sequenced to detect isoforms with a similar amplicon size than the wild type amplicon, but none were found. Only eight point mutations were found in the three F4acR$^+$ and two F4acR$^-$ pigs sequenced (Table 4.3): two silent mutations (exon 1), five missense mutations (exon 1, 4 and 12), and one mutation in the beginning of the partially sequenced 3'UTR (exon 20). The mutation in the 3'UTR was found in a region where no microRNA binding sites were detected by miRbase (Griffiths-Jones, 2004).

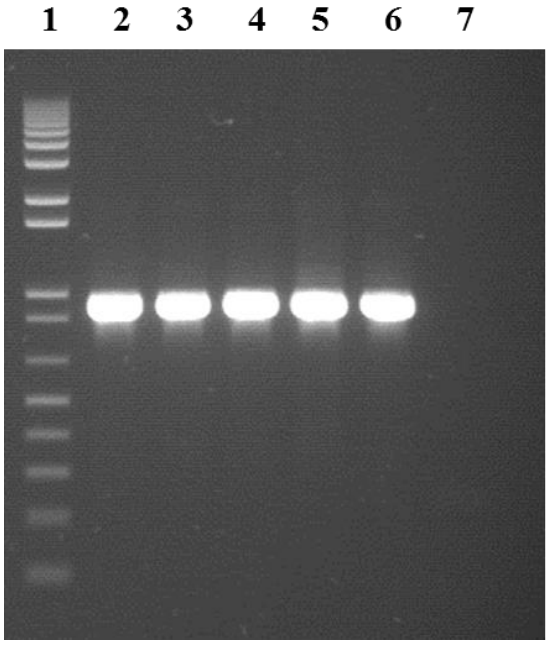

Figure 4.1: Agarose gel of SscrANPEP-F3/R3 PCR products using cDNA as a template representing the absence of differential gene expression and obvious alternative splice variants in the small intestine of F4acR$^+$ and F4acR$^-$ pigs. Lane 1: 1Kb Plus DNA ladder (Invitrogen); lane 2: F4acR$^+$ pig 5; lane 3: F4acR$^+$ pig 8; lane 4: F4acR$^+$ pig 15; lane 5: F4acR$^-$ pig 33; lane 6: F4acR$^-$ pig 36; lane 7: water.

Table 4.3: *ANPEP* mutations detected in 3 F4acR$^+$ and 2 F4acR$^-$ pigs during sequencing.

	EX1 g.2380 C>T (Y33Y)	**EX1 g.2419 G>A (Q46Q)**	**EX1 g.2602 C>A (F107L)**	**EX1 g.2603 A>C (I108L)**	**EX1 g.2615 C>T (P112S)**	**EX4 g.4328 C>T (P330S)**	**EX 12 g.8214 A>G (I621V)**	**EX 20 g.16875 C>G (-)**
NCBI_ss#	**831884247**	**831884248**	**831884249**	**831884250**	**831884251**	**831884252**	**831884254**	**831884255**
PIG 5 (R$^+$)	TAC/TAC (Y33Y)	CAG/CAG (Q46Q)	TTA/TTA (L107L)	CTT/CTT (L108L)	CCC/CCC (P112P)	CCC/TCC (P330S)	ATC/ATC (I621I)	C/C
PIG 8 (R$^+$)	TAC/TAC (Y33Y)	CAG/CAG (Q46Q)	TTA/TTA (L107L)	CTT/CTT (L108L)	CCC/CCC (P112P)	CCC/CCC (P330P)	GTC/GTC (V621V)	C/C
PIG 15 (R$^+$)	TAC/TAT (Y33Y)	CAG/CAA (Q46Q)	TTA/TTA (L107L)	CTT/CTT (L108L)	CCC/CCC (P112P)	CCC/CCC (P330P)	GTC/GTC (V621V)	C/C
PIG 33 (R$^-$)	TAC/TAC (Y33Y)	CAG/CAG (Q46Q)	TTA/TTA (L107L)	CTT/CTT (L108L)	CCC/CCC (P112P)	CCC/CCC (P330P)	GTC/GTC (V621V)	C/C
PIG 36 (R$^-$)	TAC/TAT (Y33Y)	CAG/CAA (Q46Q)	TTA/TTA (L107L)	CTT/CTT (L108L)	CCC/TCC (P112S)	TCC/TCC (S330S)	GTC/GTC (V621V)	C/G

Mutation in codon and subsituted amino acid are underlined.

4.5. Discussion

For thirty-seven pigs (80.43 %) the results of the oral immunization assay and the *in vitro* villous adhesion assay were in agreement, whereas for nine pigs (19.56 %) the results were conflicting.

Six pigs (13.04 %) showed *in vitro* no F4ac ETEC adhesion and became nevertheless seropositive. Blotting of brush border membrane proteins of these pigs with F4ac fimbriae displayed similar F4ac binding patterns as the F4acR$^-$ group described in Nguyen et al. (2013). It has been described that oral administration of F4ac antigens to F4acR$^-$ pigs (based on the *in vitro* villous adhesion test) can prime the immune system resulting in a secondary antibody response following a subsequent intramuscular immunization (Van den Broeck et al., 2001). Looking at the F4ac-specific serum IgA response of these piglets seven days after boosting, a weaker response occurred (from a titer of seven to 30) in comparison with F4acR$^+$ piglets (titer increased till 43). Consequently, we hypothesize that, at least during the second oral immunization, F4ac antigens boosted the immune system via small wounds in the mouth or mucosa of these pigs.

Three seronegative pigs (6.52 %) were F4ac ETEC adhesion-positive. Although the *in vitro* villous adhesion test has been proven to be reliable (Van den Broeck et al., 1999b), it has been described that the correlation between *in vitro* F4 ETEC adhesion to isolated brush border vesicles and F4 ETEC susceptibility is not absolute (Francis et al., 1998).

Having well-characterized case-control groups is necessary in order to find a strong genotype-phenotype correlation in genetic association studies. We propose to use pigs with corresponding results for the oral immunization and the *in vitro* villous adhesion assay for F4ac phenotyping the case-control groups for F4ac ETEC. We decided not to include genotyping data of *MUC4* (Jørgensen et al., 2004) nor *MUC13* (Ren et al., 2012) because their association with the F4ac adhesion phenotype was found to be too low, at least in our Belgian breeds (Goetstouwers et al., 2014).

Because ANPEP was recently identified as an endocytotic receptor for F4ac ETEC (Melkebeek et al., 2012), the expression of *ANPEP* was investigated in the small intestine by RT-PCR and sequencing, because only an obvious difference in expression (expression or no expression) between F4acR$^+$ and F4acR$^-$ pigs could explain binding or no binding of F4ac ETEC in pigs. We could conclude that no apparent expression difference was observed.

Also, alternative splicing could play a role in F4ac ETEC susceptibility by removing or adding specific domains of the F4ac fimbriae binding site, or by altering the steric structure of

ANPEP. Although human ANPEP is subjected to alternative splicing (Dybkaer et al., 2001) we observed no porcine *ANPEP* alternative splice variants in the small intestine during the RT-PCR experiment and sequencing of all the fragments.
ANPEP mutations cannot play a role in the F4ac ETEC susceptibility because previous linkage studies have shown that the locus controlling F4ab/ac ETEC susceptibility is located on SSC13 (Ren et al., 2012; Schroyen et al., 2012). As a result we can conclude that the eight identified mutations do not have an effect on the F4ac binding properties of ANPEP.
It was shown that F4ac ETEC, like porcine transmissible gastroenteritis coronaviruses, is binding on the carbohydrate moieties of ANPEP in a sialic acid-dependent manner (Schultze et al., 1996; Melkebeek et al., 2012). Given the results described above, F4ac susceptibility is not governed by *ANPEP* expression regulated by the as of yet unidentified causal mutation on SSC13. We hypothesize that the susceptibility towards F4ac ETEC mediated by ANPEP is not due to modifications in the protein itself, but due to modifications in the carbohydrate structures attached to it by a genetic variation influencing the glycosylation of ANPEP.

4.6. Acknowledgements

The authors wish to thank Linda Impe, Dominique Vander Donckt and Ruben Van Gansbeke for excellent technical assistance. This research was funded by an agricultural research grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen).

4.7. References

Alexander TJL. 1994. Neonatal Diarrhoea in Pigs. In: Gyles CL (ed), *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, UK, 151-170.

Amezcua R, Friendship RM, Dewey CE. 2008. An investigation of the presence of *Escherichia coli* O149:K91:F4 on pig farms in southern Ontario and the use of antimicrobials and risk factors associated with the presence of this serogroup. Can Vet J. 49:39-45.

Bakker D, Willemsen PT, Simons LH, van Zijderveld FG, de Graaf FK. 1992. Characterization of the antigenic and adhesive properties of FaeG, the major subunit of K88 fimbriae. Mol Microbiol. 6:247-255.

De la Fe Rodriguez PY, Coddens A, Del Fava E, Abrahantes JC, Shkedy Z, Martin LOM, Muñoz EC, Duchateau L, Cox E, Goddeeris BM. 2011. High prevalence of $F4^+$ and $F18^+$ *Escherichia coli* in Cuban piggeries as determined by serological survey. Trop Anim Health Prod. 43:937-946.

Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjöström H, Noren O, Laude H. 1992. Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. Nature. 357:417-419.

Delmas B, Gelfi J, Sjöström H, Noren O, Laude H. 1994. Further characterization of aminopeptidase N as a receptor for coronaviruses. Adv Exp Med Biol. 342:293-298.

Do TN, Cu PH, Nguyen HX, Au TX, Vu QN, Driesen SJ, Townsend KM, Chin JJ, Trott DJ. 2006. Pathotypes and serogroups of enterotoxigenic *Escherichia coli* isolated from pre-weaning pigs in north Vietnam. J Med Microbiol. 55:93-99.

Douglas AGL, Wood MJA. 2011. RNA splicing: disease and therapy. Brief Funct Genomics. 10:151-164.

Dybkaer K, Kristensen JS, Pedersen FS. 2001. Single site polymorphisms and alternative splicing of the human CD13 gene - different splicing frequencies among patients with acute myeloid leukaemia and healthy individuals. Br J Haematol. 112:691-696.

Francis DH, Grange PA, Zeman DH, Baker DR, Sun R, Erickson AK. 1998. Expression of mucin-type glycoprotein K88 receptors strongly correlates with piglet susceptibility to K88+ enterotoxigenic *Escherichia coli*, but adhesion of this bacterium to brush borders does not. Infect Immun. 66:4050-4055.

Goetstouwers T, Van Poucke M, Coppieters W, Nguyen VU, Melkebeek V, Coddens A, Van Steendam K, Deforce D, Cox E, Peelman LJ. 2014b. Refined Candidate Region for F4ab/ac enterotoxigenic *Escherichia coli* Susceptibility situated proximal to *MUC13* in Pigs. PloS ONE 9:e105013.

Griffiths-Jones S. 2004. The microRNA registry. Nucleic Acids Res. 32:D109-D111.

Guinee PA, Jansen WH. 1979. Behavior of *Escherichia coli* K antigens K88ab, K88ac, and K88ad in immunoelectrophoresis, double diffusion, and hemagglutination. Infect Immun. 23:700-705.

Gupta R, Jung E, Brunak S. 2004. Prediction of N-glycosylation sites in human proteins. unpublished.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 41:95-98.

Jones GW, Rutter JM. 1972. Role of the K88 antigen in the pathogenesis of neonatal diarrhea caused by *Escherichia coli* in piglets. Infec Immun. 6:918-927.

Jørgensen CB, Cirera S, Archibald AL, Anderson L, Fredholm M, Edfors-Lilja I. 2004. Porcine polymorphisms and methods for detecting them. International application published under the patent cooperation treaty (PCT), PCT/DK2003/000807 or WO2004/048606A2.

Kearns M, Gibbons RA. 1979. The possible nature of pig intestinal receptor for the K88 antigen of *Escherichia coli.* FEMS Microbiol Lett. 6:165-168.

Madoroba E, Van Driessche E, De Greve H, Mast J, Ncube I, Read J, Beeckmans S. 2009. Prevalence of enterotoxigenic *Escherichia coli* virulence genes from scouring piglets in Zimbabwe. Trop Anim Health Prod. 9:1539-1547.

Melkebeek V, Rasschaert K, Bellot P, Tilleman K, Favoreel H, Deforce D, De Geest BG, Goddeeris BM, Cox E. 2012. Targeting aminopeptidase N, a newly identified receptor for F4ac fimbriae, enhances the intestinal mucosal immune response. Mucosal Immunol. 5:635-645.

Mooi FR, de Graaf FK. 1985. Molecular biology of fimbriae of enterotoxigenic *Escherichia coli*. Curr Top Microbiol Immunol. 188:199-138.

Nguyen TT, Seoighe C. 2013. Integrative analysis of mRNA expression and half-life data reveals *trans*-acting genetic variants associated with increased expression of stable transcripts. PloS ONE. 8:e79627.

Nguyen VU, Goetstouwers T, Coddens A, Van Poucke M, Peelman L, Deforce D, Melkebeek V, Cox E. 2013. Differentiation of F4 receptor profiles in pigs based on their mucin 4 polymorphism, responsiveness to oral F4 immunization and *in vitro* binding of F4 to villi. Vet Immunol Immunopathol. 152:93-100.

Olsen J, Kokholm K, Noren O, Sjöström H. 1997. Structure and expression of aminopeptidase N. Adv Exp Med Biol. 421:47-57.

Orskov I, Orskov F, Sojka WJ, Wittig W. 1964. K antigens K88ab(L) and K88ac(L) in *E. coli*. A new O Antigen: 0147 and a new K antigen: K89(B). Acta Pathol Microbiol Scand. 62:439-447.

Poulsen PH, Thomsen PD, Olsen J. 1991. Assignment of the porcine aminopeptidase N (*PEPN*) gene to chromosome 7cen----q21. Cytogenet Cell Genet. 57:44-46.

Rasschaert K, Verdonck F, Goddeeris BM, Duchateau L, Cox E. 2007. Screening of pigs resistant to F4 enterotoxigenic *Escherichia coli* (ETEC) infection. Vet Microbiol. 123:249-253.

Rawlings ND, Barrett AJ. 1993. Evolutionary families of peptidases. Biochem J. 290:205-218.

Ren J, Yan X, Ai H, Zhang Z, Huang X, Ouyang J, Yang M, Yang H, Han P, Zeng W, Chen Y, Guo Y, Xiao S, Ding N, Huang L. 2012. Susceptibility towards enterotoxigenic *Escherichia coli* F4ac diarrhea is governed by the MUC13 gene in pigs. PloS ONE. 7:e445739.

Schroyen M, Stinckens A, Verhelst R, Geens M, Cox E, Niewold T, Buys N. 2012. Susceptibility of piglets to enterotoxigenic *Escherichia coli* is not related to the expression of *MUC13* and *MUC20*. Anim Genet. 43:324-327.

Schultze B, Krempl C, Ballesteros ML, Shaw L, Schauer R, Enjuanes L, Herrler G. 1996. Transmissible gastroenteritis coronavirus, but not the related porcine respiratory coronavirus, has a sialic acid (*N*-glycolylneuraminic acid) binding activity. J Virol. 70:5634-5637.

Sellwood R. 1980. The interaction of the K88 antigen with porcine intestinal epithelial cell brush borders. Biochim Biophys Acta. 632:32-335.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35:W71-W74.

Van den Broeck W, Cox E, Goddeeris BM. 1999a. Receptor-specific binding of purified F4 to isolated villi. Vet Microbiol. 68:255-263.

Van den Broeck W, Cox E, Goddeeris BM. 1999b. Receptor-dependent immune responses in pigs after oral immunization with F4 fimbriae. Infect Immun. 67:520-526.

Van den Broeck W, Bouchaut H, Cox E, Goddeeris BM. 2001. Receptor-independent priming of the systemic immune system of pigs by low oral doses of F4 fimbriae. Vet Immunol Immunopathol. 85:171-178.

Zhang W, Zhao M, Ruesch L, Omot A, Francis D. 2007. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. Vet Microbiol. 123:145-152.

CHAPTER 5

VARIATION IN 12 PORCINE GENES INVOLVED IN THE CARBOHYDRATE MOIETY ASSEMBLY OF GLYCOSPHINGOLIPIDS DOES NOT ACCOUNT FOR DIFFERENTIAL BINDING OF F4 ETEC AND THEIR FIMBRIAE

Adapted from BMC Genetics 15:103 (2014).

Tiphanie Goetstouwers[1], Mario Van Poucke[1], Annelies Coddens[2], Ut Nguyen Van[2], Vesna Melkebeek[2], Dieter Deforce[3], Eric Cox[2] and Luc J. Peelman[1*]

[1]Laboratory of Animal Genetics, Faculty of Veterinary Medicine, Ghent University, Heidestraat 19, 9820 Merelbeke, Belgium

[2]Laboratory of Immunology, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, 9820 Merelbeke, Belgium

[3]Laboratory of Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium

5. VARIATION IN 12 PORCINE GENES INVOLVED IN THE CARBOHYDRATE MOIETY ASSEMBLY OF GLYCOSPHINGOLIPIDS DOES NOT ACCOUNT FOR DIFFERENTIAL BINDING OF F4 ETEC AND THEIR FIMBRIAE

5.1. Abstract

Glycosphingolipids (GSLs) are important membrane components composed of a carbohydrate structure attached to a hydrophobic ceramide. They can serve as specific membrane receptors for microbes and microbial products, such as F4 enterotoxigenic *Escherichia coli* (F4 ETEC) and isolated F4 fimbriae. The aim of this study was to investigate the hypothesis that variation in genes involved in the assembly of the F4 binding carbohydrate moiety of GSLs (i.e. *ARSA, B4GALT6, GAL3ST1, GALC, GBA, GLA, GLB1, GLB1L, NEU1, NEU2, UGCG* and *UGT8*) could account for differential binding of F4 ETEC and their fimbriae.

RT-PCR could not reveal any differential expression of the 12 genes in the jejunum of F4 receptor-positive ($F4R^+$) and F4 receptor-negative ($F4R^-$) pigs. Sequencing the complete open reading frame of the 11 expressed genes (*NEU2* was not expressed) identified 72 mutations. Although some of them might have a structural effect, none of them could be associated with a F4R phenotype.

We conclude that no regulatory or structural variation in any of the investigated genes is responsible for the genetic susceptibility of pigs towards F4 ETEC.

5.2. Introduction

Glycosphingolipids (GSLs) are membrane components that participate in many intracellular and extracellular biological processes (Lingwood, 2011). They are located in the outer leaflet of the plasma membrane in mammalian cells and are composed of a carbohydrate moiety linked to a lipid (ceramide). Biosynthesis of GSL occurs by the stepwise addition of carbohydrates first to the ceramide component, then to the growing carbohydrate chain (Varki et al., 2009). The genes from the cerebroside-sulfatid region of the sphingolipid metabolism pathway are directly involved in synthesizing the carbohydrate core structure of GSLs (Figure 5.1). The cell surface carbohydrate structure of GSL can serve as specific binding sites for pathogens and their toxins, leading to subsequent adhesion (Schengrund, 2003). Recently, it has been shown that the carbohydrate moiety of GSL interacts with F4 enterotoxigenic *Escherichia coli* (F4 ETEC) and their fimbriae (Coddens et al., 2011). F4 ETEC infections are a major cause of neonatal and post-weaning diarrhea in pigs (Fairbrother et al., 2005). Following attachment with their F4 fimbriae to specific receptors in the small intestine, they colonize the small intestine and produce enterotoxins (heat-labile and heat-stable enterotoxins) which stimulate fluid secretion of epithelial cells, causing diarrhea in young pigs. Three antigenic F4 variants (F4ab, F4ac and F4ad) have been described (Gaastra and de Graaf, 1982). The F4ac variant is worldwide the most common variant, except in central China where the F4ad variant is the most prevalent (Fairbrother et al., 2005, Wang et al., 2006). Susceptibility towards F4 ETEC is inherited as an autosomal dominant Mendelian trait and the locus controlling F4ab/ac ETEC susceptibility has been mapped on chromosome 13 (SSC13). Recently, a new refined candidate region for F4ab/ac ETEC susceptibility has been identified on SSC13 indicating that the causal mutation for F4ab/ac ETEC susceptibility is not located in the previous suggested candidate genes on SSC13 (Goetstouwers et al., 2014a). The locus controlling F4ad ETEC susceptibility has not been mapped yet. So far, no causal mutation explaining the F4ab/ac/ad ETEC susceptibility in pigs has been identified (Schroyen et al., 2012; Rampoldi, 2013; Goetstouwers et al., 2014a). Therefore, the purpose of this study is to determine if the F4 ETEC binding differences observed by Coddens et al. (2011) could be explained by differential expression (for F4ab/ac/ad) or structural variation (for F4ad) of genes involved in the assembly of the carbohydrate moiety of GSLs.

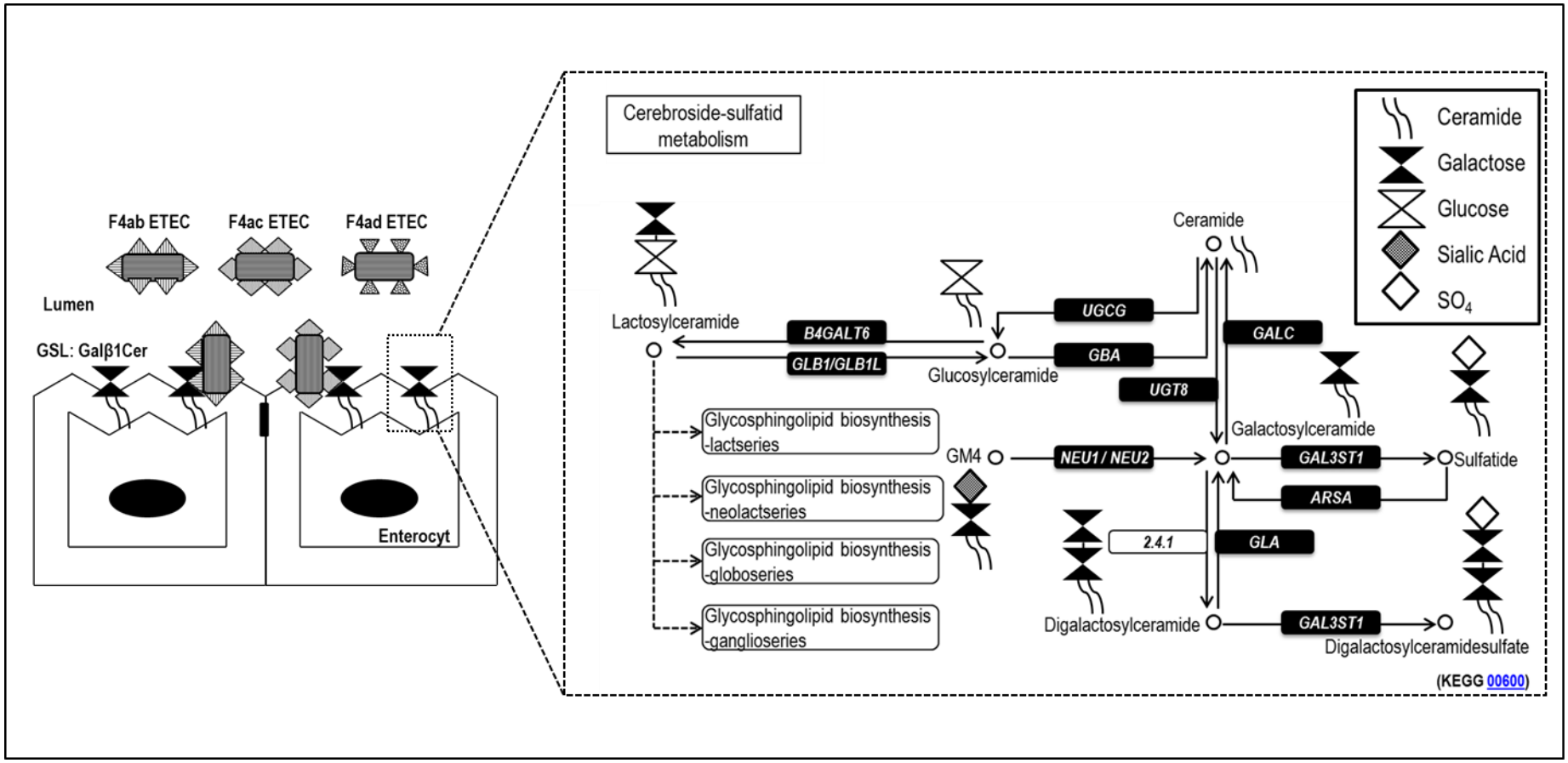

Figure 5.1: F4 ETEC binding on GSLs and the 12 investigated genes of the cerebroside-sulfatid pathway.
The carbohydrate moiety of GSLs has been shown to bind F4 ETEC and their fimbriae. According to Coddens et al. (2011), galactosylceramide Galβ1Cer binds to F4ab/ac ETEC and fimbriae. Twelve genes involved in the carbohydrate moiety assembly of glycosphingolipids were selected from the cerebroside-sulfatid region of the sphingolipid metabolism pathway (adapted from KEGG pathway 00600).
The solid lines represent molecular interaction or relation, the dashed lines represent linked to another map (see http://www.genome.jp/kegg-bin/show_pathway?map00600 for further details).

5.3. Methods

5.3.1. Sample collection

Crossbred pigs from different litters were euthanized at 5-18 weeks of age. Before euthanasia, blood samples were collected in EDTA blood tubes and stored at -20 °C for DNA isolation. After slaughter, samples of mid-jejunum were collected using protocols approved by the animal care and ethics committee of the Faculty of Veterinary Medicine, University of Ghent (EC2010/042). Mid-jejunum samples for RNA isolation were washed three times with Krebs–Henseleit buffer (0.12 M NaCl, 0.014 M KCl, 0.001 M KH2PO4, 0.025 M NaHCO3, pH 7.4), immediately frozen in liquid nitrogen and stored at -80 °C until RNA isolation. Villi from mid-jejunum samples for the *in vitro* villous adhesion assay were isolated and stored as described by Van den Broeck et al. (1999a).

5.3.2. Animal selection based on the *in vitro* villous adhesion assay

The *in vitro* villous adhesion assay for F4ab/ac/ad ETEC was carried out as described by Van den Broeck et al. (1999a). Adhesion of more than 30 bacteria per 250 µm villous brush border length was noted as strong adhesive for F4 ETEC ($F4R^+$) and less than 5 bacteria per 250 µm brush border length was noted as non-adhesive for F4 ETEC ($F4R^-$) (Cox and Houvenaghel, 1993).

Eight pigs, representing six different F4 adhesion phenotypes, were selected for the expression study and mutation detection. These phenotypes were previously described as phenotype A (F4ab/ac/adR^+; pig 1 and 2), B (F4ab/acR^+; pig 3), C (F4ab/adR^+; pig 4), D (F4adR^+; pig 5), E (F4ab/ac/adR^-; pig 6 and 7) and F (F4abR^+; pig 8) (Bijlsma et al., 1982; Baker et al., 1997). The phenotypes G (F4acR^+) and H (F4ac/adR^+), mainly observed in eastern breeds, were absent in our study (Bonneau et al., 1990, Li et al., 2007; Yan et al., 2009). Fourteen additional pigs were selected, only based on the presence of the F4ad receptor (7 F4adR^+ and 7 F4adR^-), for the GALC (c.979T>C) mutation screening.

5.3.3. DNA isolation, RNA isolation and cDNA synthesis

DNA isolation from frozen blood samples was performed as described by Van Poucke et al. (2005). RNA isolation and cDNA synthesis of frozen mid-jejunum samples was performed as described by Goetstouwers et al. (2014b).

Table 5.1: Details of the investigated genes involved in the assembly of the F4 binding carbohydrate moiety of GSLs.

Gene Symbol	Gene Description	NCBI gene ID	RefSeq status	Chr.	AccNo
ARSA	arylsulfatase A	396973	Non-curated	5	NM_213933.1
B4GALT6	beta-1,4-galactosyltransferase 6	100517222	Non-curated	6	XM_003127886.3
GAL3ST1	galactose-3-O-sulfotransferase 1	100155265	Non-curated	14	NM_001244429.1
GALC	Galactosylceramidase	100156917	Curated	7	NM_001243631.1
GBA	glucosidase, beta	449572	Non-curated	4	NM_001005730.1
GLA	galactosidase, alpha	407057	Non-curated	X	NM_001177925.1
GLB1	galactosidase, beta 1	/	/	15[α]	AK230951.1
GLB1L	galactosidase, beta 1-like	100154356	Non-curated	15	XM_001928375.3
NEU1	sialidase 1	100124381	Non-curated	7	NM_001101822.1
NEU2	sialidase 2	100738467	Non-curated	15	XM_003483766.2
UGCG	UDP-glucose ceramide glucosyltransferase	100152737	Non-curated	1	XM_001925267.5
UGT8	UDP glycosyltransferase 8	/	/	8[α]	GU991196.1

[α] chromosome location based on human comparative mapping information

5.3.4. *In silico* gene analysis and experimental validation

Non-curated porcine gene sequences (Table 5.1) from NCBI databases were (re)checked manually using BLAST analysis (genomic and mRNA) for a human-pig and a pig-pig comparison (Benson et al., 2009). Primers were designed with Primer3Plus (Untergasser et al., 2007), generating overlapping amplicons that cover the complete coding sequence. RT-PCR products were generated with porcine mid-jejunum cDNA as a template (Table 5.2),

of which 2 µl was used to check the amplicon length using agarose gel electrophoresis. The rest of the product (8 µl) was cleaned up with 4 U Exonuclease I and 2 U Antarctic Phosphatase (New England Biolabs) at 37 °C for 30 min and 80 °C for 15 min, and sequenced for verification. Forward and reverse sequencing reactions were performed with the PCR primers as described by Goetstouwers et al. (2014b).

Table 5.2: Oligonucleotide sequences of primers with their specifications (*ANPEP*: NM_214277.1, *ARSA*: NM_213933.1, *B4GALT6*: XM_003127886.3, *GAL3ST1*: NM_001244429.1, *GALC*: NM_001243631.1, *GBA*: NM_001005730.1, *GLA*: NM_001177925.1, *GLB1*: AK230951.1, *GLB1L*: XM_001928375.3, *NEU1*: NM_001101822.1, *NEU2*: XM_003483766.2, *UGCG*: XM_001925267.5, *UGT8*: JQ650526).

Gene (Chr.)	**Name Primer pair**	**F-primer (5'→3') R-primer (5'→3')**	**Ta (°C) Length cDNA / gDNA (bp)**	**Exon (EX) / Intron (IN)**
ANPEP (7)	SscrANPEP-F1	GAATGACCTGTGGCTGAATGAGG	66 210 / 307	EX 6 → EX 7
	SscrANPEP-R1	GAGTCAAACATCTCGCTGATCTGG		
ARSA (5)	SscrARSA-F1	GTCAGTGTGCGGGGACTTCT	66 578 / -	EX 1 → EX 4
	SscrARSA-R1	GGAAGCAGGTCAGGTTCTGG		
	SscrARSA-F2	ATGTGCCGGTGTCTCTGTGC	66 960 / -	EX 2 → EX 8
	SscrARSA-R2	TCATCTGGGTAGGCCGGGTA		
	SscrARSA-F3	AGACAACGGACCCGAGACGA	66 783 / -	EX 5/6 → EX 9
	SscrARSA-R3	TCGAGCCACCTTCATCCACTC		
B4GALT6 (6)	SscrB4GALT6-F1	TGCAGTCTCGGGGTTTGTT	62 1009 / -	EX 1 → EX 7
	SscrB4GALT6-R1	CCAAAGGTCATCGTCTTCTCC		
	SscrB4GALT6-F2	GCAGCCTTTTAACCGTGCAAT	62 751 / -	EX 6 → EX 9
	SscrB4GALT6-R2	CGCCGAAGGGAATGAGTCTA		
GAL3ST1 (14)	SscrGAL3ST1-F1	TGCAGGGTACAGGGACATGC	64°C 866 / -	EX 1 → EX 3
	SscrGAL3ST1-R1	GAAGAGCAGGTTGCGGAGGT		
	SscrGALT3ST1-F2	GCTTCCACTACGACGAGGTT	62°C 962 / -	EX 3 → EX 3
	SscrGAL3ST1-R2	AGTCCTGGCTGGTTGCAC		
GALC (7)	SscrGALC-F1	AACTCCTGCCCTTCTCCATCA	56 1009 / -	EX 1 → EX 9
	SscrGALC-R1	CTAGCCACTAAATTCCAAGCAATGG		
	SscrGALC-F2	GATGCTAGGTTGACTGAGAAGAAGC	64 965 / -	EX 8 → EX 15
	SscrGALC-R2	TTACCCTTCCGGCAATGAAC		
	SscrGALC-F3	TGTCTATGAGGACGATTTCAACG	64 967 / -	EX 13 → EX 17
	SscrGALC-R3	CGACGAGGACACAGCTCACT		
	SscrGALC-F4	CATTTGTGACCTCTCTGGTACTGG	62 - / 520 [α]	IN 8 → IN 9
	SscrGALC-R4	AGGTGCTGTTTGCCTTGTCTC		
GBA (4)	SscrGBA-F1	ACCAATAAGAAGTGCGGAAAGG	60 833 / -	EX 1 → EX 6
	SscrGBA-R1	AACTGGAAGTCATCAGGGGTGT		
	SscrGBA-F2	GATTTGGAGGGGCCATGA	64 909 / -	EX 5 → EX 11
	SscrGBA-R2	CAGTCAGTCCAGCCAACCAC		
	SscrGBA-F3	AGGCGGCTAAGTACGTTCA	64 965 / -	EX 10 → EX 13
	SscrGBA-R3	CGGCTTTCCTAGTCTCTTCC		
GLA (X)	SscrGLA-F1	AAGAACCTAGAAGCCCAGGTGACT	66 961 / -	EX 1 → EX 6
	SscrGLA-R1	TCGGAGGTCATTGGACATGAG		
	SscrGLA-F2	TTGGACTGGACATCTTCTAACCA	64 702 / -	EX 5 → EX 7
	SscrGLA-R2	GCTGCGGTTGTGACCTACAC		
GLB1 (15)	SscrGLB1-F1	CTGTCCCGGCGCTGACTG	64 932 / -	EX 1 → EX 8
	SscrGLB1-R1	CGTGGAGGAAAGAAGCCACCA		
	SscrGLB1-F2	TGAACCCAGAGGACCCTTGA	64 1101 / -	EX 7 → EX 16
	SscrGLB1-R2	GTGCATGCTCCAGCTCCAG		
	SscrGLB1-F4	ATCTTCCCGATGGACACTGA	64 594 / -	EX 15 → EX 16
	SscrGLB1-R4	GCAGGAAATCCTTGGGTGA		

Gene (Chr.)	Name Primer pair	F-primer (5'→3') R-primer (5'→3')	Ta (°C) Length cDNA / gDNA (bp)	Exon / Intron
GLB1L (15)	SscrGLB1L-F1	CGGTTCCCTCCCAGGAATCT	70 971 / -	EX 1 → EX 8
	SscrGLB1L-R1	CATGTTCACACTGGCTCCCAAC		
	SscrGLB1L-F2	TATGAACCCCACGGGCCACT	62 816 / -	EX 7 → EX 15
	SscrGLB1L-R2	TGGATGGGAAACCACCACTTTACA		
	SscrGLB1L-F3	CGTGCCTACGTCATGGTGGA	62 883 / -	EX 13 → EX 16
	SscrGLB1L-R3	AACCGGAGGGCATTTGGAAG		
NEU1 (7)	SscrNEU1-F1	TGCTGTGGATTTGAGGGTGA	64 812 / -	EX 1 → EX 4
	SscrNEU1-R1	ATCGCTGAGGAGGCAGAAGA		
	SscrNEU1-F2	TCCTTGGATATAGGCACTGAGATGT	64 786 / -	EX 3 → EX 6
	SscrNEU1-R2	GGAACTCTCTTCCAGGCTCCTC		
NEU2 (15)	SscrNEU2-F3	GCCTACGCTTACCGCAACCT	64 381 / -	EX 2 → EX 2
	SscrNEU2-R3	GGATCTTCGCTTCGGGTCA		
UGCG (1)	SscrUGCG-F1	AGCCACTAGGCTGCGGGAAG	66 955 / -	EX 1→ EX 4
	SscrUGCG-R1	GGACCTTGGATGAGAGGTTCCAA		
	SscrUGCG-F2	GATGCTAGATTGTTCATAGGTGGCAAA	64 813 / -	EX 3/4 → EX 9
	SscrUGCG-R2	CGGCCAGTTCTCCAGCTTATTG		
	SscrUGCG-F3	CCTGGCGTGGTTTATATTTGACT	62 463 / -	EX 8 → EX 9
	SscrUGCG-R3	CCAATTCTCTTGATTCTCTACTTCCAC		
UGT8 (8)	SscrUGT8-F1	CAGCCGAAGGAGCAGGAG	62 1580 / -	EX 1→ EX 5
	SscrUGT8-R1	TATCGTAATGGTCTCCAAAGAGTGG		
	SscrUGT8-R1a	CAGCAGGATACCAAAGGCCAGT	62	EX 1 → EX 1
	SscrUGT8-F2	GGGGAACAATACCAAGCTCA	62 791 / -	EX 4 → EX 6
	SscrUGT8-R2	CAGCCATCTTAATTCCACAGAA		

[α]Positions of GALC±4 and amplicon size based on Acc. Nr. NC_010449.4

5.3.5. Semi-quantitative expression study via RT-PCR

All above mentioned primers, generating overlapping amplicons covering the complete open reading frame of the 12 investigated genes, were used to perform RT-PCR (see above) on cDNA of the mid-jejunum samples of the 8 selected animals. Agarose gel electrophoresis was used to analyze the number and the length of the PCR products to check for phenotype explaining alternative splicing, and to compare the intensity of the bands to check for phenotype explaining differential expression (semi-quantitatively). *ACTB* was used as a validated reference gene (Erkens et al., 2006).

5.3.6. Mutation detection via sequencing of the RT-PCR products

All RT-PCR products from the expression study were sequenced (see above) to check for F4ad phenotype explaining structural mutations.

5.3.7. *GALC* (c.979T>C) mutation screening

The *GALC* (c.979T>C) mutation was screened in 14 additional F4ad phenotyped pigs (7 F4adR$^+$ and 7 F4adR$^-$) via PCR with primer pair SscrGALC±4 and DNA as a template (Table 5.1), and direct sequencing with the reverse sequence primer after PCR amplicon clean-up (see above).

5.4. Results and discussion

For all 12 genes (Table 5.1) there is a curated human reference sequence available in the public databases. In pig however, this is so far only the case for *GALC*. For 9 genes (i.e. *ARSA*, *B4GALT6*, *GAL3ST1*, *GBA*, *GLA*, *GLB1L*, *NEU1*, *NEU2* and *UGCG*) there was a predicted porcine sequence. These sequences were subjected to an *in silico* gene analysis and experimental validation. The coding sequence of all predicted porcine sequences was found to be correct since the exact sequence was found to be expressed in the jejunum, except for *NEU2* that was not expressed. We neither observed *NEU2* expression in porcine lymph node, heart, lung, dorsal muscle, diaphragm, liver, spleen, gall bladder, kidney, adrenal gland, bladder, duodenum, jejunum, ileum, colon and rectum (*NEU2* assay was validated with DNA as a template; data not shown), which resembles the situation in human where extremely low levels of mRNA expression were found in all human tissues, except for testis, placenta and ovary (Koseki et al., 2012). Interspecies sequence comparison revealed the complete porcine *GLB1* coding sequence in a non-annotated mRNA sequence [GenBank: AMP010068C04]. The exact sequence of 1992 bp (encoding a protein of 663 amino acids) was found to be expressed in the jejunum and shows 85 % sequence identity with its human ortholog [GenBank: NM_000404.2]. The complete coding sequence of porcine *UGT8* (1623 bp; encoding a protein of 541 amino acids) was amplified by RT-PCR from jejunum cDNA with primers based on its human ortholog [GenBank: NM_001128174.1]. Interspecies comparison showed only high sequence identities with *UGT8* orthologs (93 % with its human ortholog) and the sequence was submitted to NCBI as the first porcine *UGT8* mRNA sequence [GenBank: JQ65026].

The eight pigs used in this study were solely phenotyped based on the *in vitro* villous adhesion test that has been proven to be reliable (Cox and Houvenaghel, 1993; Van den Broeck et al., 1999a; Van den Broeck et al., 1999b). Phenotyping of the pigs based on the associated markers identified in previous linkage studies or based on the associated mutations

in *MUC4* and *MUC13* would not be precise, because they are not in complete linkage disequilibrium with the F4ab/ac locus (Jørgensen et al., 2004; Rasschaert et al., 2007; Li et al., 2008; Goetstouwers et al., 2014a). Although linkage studies mapped the causal locus for the F4ab/ac susceptibility on SSC13 (Schroyen et al., 2012; Rampoldi, 2013; Goetstouwers et al., 2014a), it is possible that the expression of any of the 12 investigated genes is influenced by a *trans*-acting element present in this candidate region (Goetstouwers et al., 2014a). As no positional information is available for the F4ad ETEC receptor, a regulatory mutation impairing expression of any of the investigated GSL genes could also be responsible for the F4ad ETEC susceptibility in pigs. Because an obvious difference in expression between F4 receptor-positive ($F4R^+$) and F4 receptor-negative ($F4R^-$) pigs was expected, semi-quantitative measurements using 8 pigs with different F4 adhesion phenotypes were performed. For every amplicon a single fragment was generated with the same intensity for all samples. We can conclude that F4 ETEC susceptibility is not caused by any mutation affecting the expression level of any of the investigated genes nor by the expression of splice variants.

All amplicons generated in the expression study were sequenced to investigate if a structural mutation in any of these genes could be responsible for F4ad ETEC susceptibility. In total, 72 mutations were found: 45 silent mutations, 24 missense mutations, 2 mutations in the 3'UTR and 1 nonsense mutation (Table 5.3). Only the silent mutation c.979T>C in *GALC* was differential for the presence of the F4ad receptor in this sample set. The CC homozygotes and CT heterozygotes were present in the $F4adR^+$ pigs and only TT homozygotes were present in the $F4R^-$ pigs. We expected a homozygous genotype in the $F4adR^-$ pigs, because resistance to F4 adhesion ($F4R^-$) is inherited in a recessive Mendelian way (Gibbons et al., 1977). We screened this mutation in 14 additional F4ad phenotyped pigs. Four TT homozygotes and 3 CT heterozygotes were observed in the $F4adR^+$ pigs (n=7) and 7 TT homozygotes in de $F4adR^-$ group (n=7). Because 4 TT homozygotes were present in the $F4adR^+$ pigs, we can conclude that this mutation is not associated with F4ad ETEC susceptibility. For completeness we also looked for association with the F4ab/acR phenotype, but as could be expected from the chromosomal position of the GSL genes none of the 72 mutations were differential in $F4ab/acR^+$ and $F4ab/acR^-$ pigs.

Table 5.3: Prevalence of the differential structural mutations in 11 investigated genes based on F4R type.

	ARSA						*B4GALT6*						*GAL3ST1*																											*GALC*														
	c.1348C>T			c.1546C>T			c.453T>A			c.1210T>C			c.320A>G			c.402C>T			c.484C>A			c.512G>A			c.542C>T			c.831C>T			c.963C>T			c.1007G>A			c.1228C>T			c.52G>A			c.130C>T			c.449T>C			c.664C>T			**c.979T>C**		
	EX 7			EX 9			EX 3			EX 9			EX 2			EX 2			EX 3			EX 3			EX 3			EX 3			EX 3			EX 3			EX 3			EX 1			EX 1			EX 4			EX 7			**EX 9**		
	L361L			L427L			F95Y			I347I			Q5R			A32A			A60T			G69E			S79F			Y175Y			F219F			V234E			R308C			R13R			P39P			F146L			G217G			**C322C**		
	CC	CT	TT	CC	CT	TT	TT	TA	AA	TT	TC	CC	AA	AG	GG	CC	CT	TT	CC	CA	AA	GG	GA	AA	CC	CT	TT	CC	CT	TT	CC	CT	TT	GG	GA	AA	CC	CT	TT	GG	GA	AA	CC	CT	TT	TT	TC	CC	CC	CT	TT	**TT**	**TC**	**CC**
F4abR+	4	1	0	4	1	0	5	0	0	5	0	0	2	3	0	2	2	1	4	1	0	3	1	1	0	0	5	4	0	1	2	2	1	4	1	0	5	0	0	1	0	4	1	1	3	0	0	5	0	0	5	2	1	2
F4abR-	2	1	0	3	0	0	2	1	0	1	2	0	1	2	0	1	2	0	3	0	0	3	0	0	0	0	3	3	0	0	1	2	0	3	0	0	2	1	0	0	0	3	1	2	0	0	0	3	0	0	3	2	1	0
F4acR+	2	1	0	2	1	0	3	0	0	3	0	0	1	2	0	1	2	0	3	0	0	2	1	0	0	0	3	2	0	1	1	2	0	2	1	0	3	0	0	1	0	2	1	0	2	0	0	3	0	0	3	1	0	2
F4acR-	4	1	0	5	0	0	4	1	0	3	2	0	2	3	0	2	2	1	4	1	0	4	0	1	0	0	5	5	0	0	2	2	1	5	0	0	4	1	0	0	0	5	1	3	1	0	0	5	0	0	5	3	2	0
F4adR+	4	0	0	4	0	0	4	0	0	4	0	0	1	3	0	1	2	1	3	1	0	3	1	0	0	0	4	3	0	1	1	2	1	3	1	0	4	0	0	1	0	3	0	1	3	0	0	4	0	0	4	**0**	**2**	**2**
F4adR-	2	2	0	3	1	0	2	1	0	2	2	0	2	2	0	2	2	0	4	0	0	3	0	1	0	0	4	4	0	0	2	2	0	4	0	0	3	1	0	0	0	4	2	2	0	0	0	4	0	0	4	**4**	**0**	**0**
	6	2	0	7	1	0	7	1	0	6	2	0	3	5	0	3	4	1	7	1	0	6	1	1	0	0	8	7	0	1	3	4	1	7	1	0	7	1	0	1	0	7	2	3	3	0	0	8	0	0	8	4	2	2

	GALC																		*GBA*												*GLA*																					*GLB1*		
	c.1192A>G			c.1426T>C			c.1468C>T			c.1549C>T			c.1927C>T			*2328G>A			c.1007G>A			c.1178T>C			c.1445C>T			c.1636G>A			c.207T>C			c.541A>G			c.612T>C			c.850C>G			c.885G>T			c.928G>A			c.1278A>G			c.29G>A		
	EX 10			EX 13			EX 13			EX 13			EX 16			EX 16			EX 7			EX 7			EX 8			EX 10			EX 2			EX 3			EX 4			EX 6			EX 6			EX 6			EX 7			EX 1		
	L383L			T471T			P485P			G512G			G638G			3'UTR			K225K			T282T			F371F			R435H			S69S			I181V			Y204Y			Q284E			A295A			E310K			S426S			V5I		
	AA	AG	GG	CC	CT	TT	CC	CT	TT	CC	CT	TT	CC	CT	TT	GG	GA	AA	GG	GA	AA	TT	TC	CC	CC	CT	TT	GG	GA	AA	TT	TC	CC	AA	AG	GG	TT	TC	CC	CC	CG	GG	GG	GT	TT	GG	GA	AA	AA	AG	GG	GG	GA	AA
F4abR+	0	2	3	5	0	0	2	3	0	0	0	5	2	1	2	0	0	5	1	3	1	5	0	0	5	0	0	4	1	0	0	0	5	0	0	5	0	0	5	0	0	5	0	0	5	0	0	5	0	0	5	4	1	0
F4abR-	0	3	0	1	2	0	1	2	0	0	2	1	0	3	0	0	0	3	0	3	0	2	1	0	2	1	0	3	0	0	0	0	3	0	0	3	0	0	3	0	0	3	0	0	3	0	0	3	0	0	3	3	0	0
F4acR+	0	1	2	3	0	0	2	1	0	0	0	3	2	0	1	0	0	3	0	2	1	3	0	0	3	0	0	2	1	0	0	0	3	0	0	3	0	0	3	0	0	3	0	0	3	0	0	3	0	0	3	2	1	0
F4acR-	0	4	1	3	2	0	1	4	0	0	2	3	0	4	1	0	0	5	1	4	0	4	1	0	4	1	0	5	0	0	0	0	5	0	0	5	0	0	5	0	0	5	0	0	5	0	0	5	0	0	5	5	0	0
F4adR+	0	1	3	4	0	0	3	1	0	0	0	4	2	2	0	0	0	4	1	2	1	4	0	0	4	0	0	3	1	0	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4	3	1	0
F4adR-	0	4	0	2	2	0	0	4	0	0	2	2	0	2	2	0	0	4	0	4	0	3	1	0	3	1	0	4	0	0	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4	4	0	0
	0	5	3	6	2	0	3	5	0	0	2	6	2	4	2	0	0	8	1	6	1	7	1	0	7	1	0	7	1	0	0	0	8	0	0	8	0	0	8	0	0	8	0	0	8	0	0	8	0	0	8	7	1	0

	GLB1																											*GLB1L*																										
	c.139T>C			c.151G>A			c.310G>T			c.982C>T			c.1042T>C			c.1577C>G			c.1580T>C			c.1902A>G			c.1948C>T			c.80G>A			c.167C>T			c.422G>A			c.491C>T			c.731C>A			c.836A>G			c.1662A>C			c.1952T>G			*2088C>T		
	EX 2			EX 2			EX 3			EX 10			EX 10			EX 15			EX 15			EX 16			EX 16			EX 1			EX 2			EX 3			EX 5			EX 6			EX 8			EX 15			EX 16			EX 16		
	L41L			Q45Q			T98T			P322P			L342L			R521G			Y522H			E629G			R656X			P15P			V44V			E129E			F152F			L232L			T267T			T543P			D639E			3'UTR		
	TT	TC	CC	GG	GA	AA	GG	GT	TT	CC	CT	TT	TT	TC	CC	CC	CG	GG	TT	TC	CC	AA	AG	GG	CC	CT	TT	GG	GA	AA	CC	CT	TT	GG	GA	AA	CC	CT	TT	CC	CA	AA	AA	AG	GG	AA	AC	CC	TT	TG	GG	CC	CT	TT
F4abR+	0	0	5	2	3	0	2	3	0	4	1	0	0	2	3	1	4	0	0	3	2	4	1	0	1	4	0	4	1	0	0	1	4	0	1	4	4	1	0	5	0	0	4	1	0	4	1	0	5	0	0	4	1	0
F4abR-	0	0	3	2	1	0	3	0	0	2	1	0	0	2	1	3	0	0	0	3	0	2	1	0	3	0	0	2	1	0	1	1	1	1	1	1	1	0	2	2	0	1	3	0	0	3	0	0	2	1	0	3	0	0
F4acR+	0	0	3	1	2	0	1	2	0	3	0	0	0	2	1	0	3	0	0	2	1	3	0	0	0	3	0	2	1	0	0	1	2	0	1	2	2	1	0	3	0	0	2	1	0	3	0	0	3	0	0	2	1	0
F4acR-	0	0	5	3	2	0	4	1	0	3	2	0	0	2	3	4	1	0	0	4	1	3	2	0	4	1	0	4	1	0	1	1	3	1	1	3	3	0	2	4	0	1	5	0	0	4	1	0	4	1	0	5	0	0
F4adR+	0	0	4	2	2	0	3	1	0	2	2	0	0	2	2	2	2	0	0	4	0	2	2	0	2	2	0	3	1	0	1	1	2	1	1	2	2	1	1	3	0	1	3	1	0	4	0	0	3	1	0	3	1	0
F4adR-	0	0	4	2	2	0	2	2	0	4	0	0	0	2	2	2	2	0	0	2	2	4	0	0	2	2	0	3	1	0	0	1	3	0	1	3	3	0	1	4	0	0	4	0	0	3	1	0	4	0	0	4	0	0
	0	0	8	4	4	0	5	3	0	6	2	0	0	4	4	4	4	0	0	6	2	6	2	0	4	4	0	6	2	0	1	2	5	1	1	5	5	1	2	7	0	1	7	1	0	7	1	0	7	1	0	7	1	0

	NEU1																		*UGCG*																								UGT8											
	c.196C>G			c.296A>G			c.318T>C			c.951T>C			c.1215A>G			c.1337G>A			c.637G>A			c.694C>T			c.874A>G			c.937G>A			C.1204C>T			c.1537A>G			c.1547A>C			c.1586G>A			c.142T>C			c.561A>G			c.755T>C			c.926C>T		
	EX 1			EX 1			EX 2			EX 5			EX 6			EX 6			EX 2			EX 3			EX 5			EX 5			EX 7			EX 9			EX 9			EX 9			EX 1			EX 2			EX 3			EX 3		
	L20G			N53S			T60T			L271L			K359K			R400Q			T72T			D91D			P151P			P172P			S261S			T372T			R376R			E389K			H46H			D186G			P251S			L308L		
	CC	CG	GG	AA	AG	GG	CC	CT	TT	TT	TC	CC	AA	AG	GG	GG	GA	AA	GG	GA	AA	CC	CT	TT	AA	AG	GG	GG	GA	AA	CC	CT	TT	AA	AG	GG	AA	AC	CC	GG	GA	AA	TT	TC	CC	AA	AG	GG	TT	TC	CC	CC	CT	TT
F4abR+	5	0	0	2	2	1	3	2	0	2	2	1	2	2	1	0	2	3	4	1	0	1	0	4	4	0	1	1	1	3	4	0	1	2	2	1	3	1	1	4	1	0	3	2	0	4	1	0	5	0	0	4	1	0
F4abR-	2	1	0	0	2	1	3	0	0	1	1	1	1	1	1	1	1	1	3	0	0	1	0	2	3	0	0	0	1	2	3	0	0	1	2	0	2	1	0	3	0	0	3	0	0	3	0	0	2	1	0	3	0	0
F4acR+	3	0	0	2	0	1	1	2	0	2	0	1	2	0	1	0	1	2	3	0	0	1	0	2	3	0	0	0	1	2	3	0	0	1	2	0	2	1	0	3	0	0	1	2	0	2	1	0	3	0	0	2	1	0
F4acR-	4	1	0	0	4	1	5	0	0	1	3	1	1	3	1	1	2	2	4	1	0	1	0	4	4	0	1	1	1	3	4	0	1	2	2	1	3	1	1	4	1	0	5	0	0	5	0	0	4	1	0	5	0	0
F4adR+	3	1	0	2	2	0	2	2	0	3	1	0	3	1	0	1	2	1	4	0	0	1	0	3	4	0	0	0	1	3	4	0	0	2	2	0	3	1	0	3	1	0	2	2	0	4	0	0	4	0	0	4	0	0
F4adR-	4	0	0	0	2	2	4	0	0	0	2	2	0	2	2	0	1	3	3	1	0	1	0	3	3	0	1	1	1	2	3	0	1	1	2	1	2	1	1	4	0	0	4	0	0	3	1	0	3	1	0	3	1	0
	7	1	0	2	4	2	6	2	0	3	3	2	3	3	2	1	3	4	7	1	0	2	0	6	7	0	1	1	2	5	7	0	1	3	4	1	5	2	1	7	1	0	6	2	0	7	1	0	7	1	0	7	1	0

5.5. Conclusions

Overall, we can conclude that no structural or regulatory variation in any of the 12 investigated genes is associated with F4 ETEC susceptibility. However, some of the mutations found (e.g. a nonsense mutation (c.1577C>G) in exon 5 of *GLB1*, introducing a premature stop codon (R656X) truncating the GLB1 protein with 8 amino acids at the C-terminus) may be of importance for other GSL-related diseases (Hanada, 2005; Taube et al., 2010; Xu et al., 2010).

5.6. Acknowledgements

The authors wish to thank Linda Impe, Dominique Vander Donckt and Ruben Van Gansbeke for excellent technical assistance. This work was supported by an agricultural research grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen).

5.7. References

Baker DR, Billey LO, Francis DH. 1997. Distribution of K88 *Escherichia coli*-adhesive and nonadhesive phenotypes among pigs of four breeds. Vet Microbiol. 54:123-132.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. Nucleic Acids Res. 37:D26-31.

Bijlsma IGW, de Nijs A, van der Meer C, Frik JF. 1982. Different pig phenotypes affect adherence of *Escherichia coli* to jejunal brush borders by K88ab, K88ac, or K88ad antigen. Infect Immun. 37:891-894.

Bonneau M, Duval-Iflah Y, Guérin G, Ollivier L, Renard C, Renjifo X. 1990. Aspects genetiques et microbiologiques de la colibacillose K88 chez le porc. Ann Rech Vet. 21:302-303.

Coddens A, Valis E, Benktander J, Ångström J, Breimer ME, Cox E, Teneberg S. 2011. Erythrocyte and porcine intestinal glycosphingolipids recognized by F4 fimbriae of enterotoxigenic *Escherichia coli*. PloS ONE. 6:e23309.

Cox E, Houvenaghel A. 1993. Comparison of the *in vitro* adhesion of K88, K99, F41 and P987 positive *Escherichia coli* to intestinal villi of 4-week-old to 5-week-old pigs. Vet Microbiol. 34:7-18.

Erkens T, Van Poucke M, Vandesompele J, Goossens K, Van Zeveren A, Peelman LJ. 2006. Development of a new set of reference genes for normalization of real-time RT-PCR data of porcine backfat and longissimus dorsi muscle, and evaluation with *PPARGC1A*. BMC Biotechnol. 6:41.

Fairbrother JM, Nadeau E, Gyles CL. 2005. *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. Anim Health Res Rev. 2005 6:17-39.

Gaastra W, de Graaf FK. 1982. Host-specific fimbrial adhesins of noninvasive enterotoxigenic *Escherichia coli* strains. Microbiol Rev. 46:129-161.

Gibbons RA, Sellwood R, Burrows M, Hunter PA. 1977. Inheritance of resistance to neonatal *E. coli* diarrhea in the pig: examination of the genetic system. Theor Appl Genet. 51:65-70.

Goetstouwers T, Van Poucke M, Coppieters W, Nguyen VU, Melkebeek V, Coddens A, Van Steendam K, Deforce D, Cox E, Peelman LJ. 2014a. Refined Candidate Region for F4ab/ac enterotoxigenic *Escherichia coli* Susceptibility situated proximal to *MUC13* in Pigs. PloS ONE 9:e105013.

Goetstouwers T, Van Poucke M, Nguyen VU, Melkebeek V, Coddens A, Dieter Deforce D, Cox E, Peelman LJ. 2014b. F4-related expression analysis of the aminopeptidase N gene in pigs. J Anim Sci. 92:1866-1873.

Hanada K. 2005. Sphingolipids in infectious diseases. Jpn J Infect Dis. 58:131-148.

Jørgensen CB, Cirera S, Archibald AL, Anderson L, Fredholm M, Edfors-Lilja I. 2004. Porcine polymorphisms and methods for detecting them. International application published under the patent cooperation treaty (PCT), PCT/DK2003/000807 or WO2004/048606A2.

Koseki K, Wada T, Hosono M, Hata K, Yamaguchi K, Nitta K, Miyagi T. 2012. Human cytosolic sialidase *NEU2*-low general tissue expression but involvement in PC-3 prostate cancer cell survival. Biochem Biophys Res Commun. 428:142-149.

Li YH, Qiu XT, Li HJ, Zhang Q. 2007. Adhesive patterns of *Escherichia coli* F4 in piglets of three breeds. J Genet Genomics. 34:591-599.

Li H, Li Y, Qiu X, Niu X, Liu Y, Zhang Q. 2008. Identification and screening of gene(s) related to susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac in piglets. Asian-Austral J Anim. 21:489-493.

Lingwood CA. 2011. Glycosphingolipid functions. Cold Spring Harb Perspect Biol. 3:a004788.

Schengrund CL. 2003. "Multivalent" saccharides: Development of new approaches for inhibiting the effects of glycosphingolipid-binding pathogens. Biochem Pharmacol. 65:699–707.

Schroyen M, Stinckens A, Verhelst R, Niewold T, Buys N. 2012. The search for the gene mutations underlying enterotoxigenic *Escherichia coli* F4ab/ac susceptibility in pigs: a review. Vet Res. 43:70.

Rampoldi A. 2013. Approach to the gene controlling the porcine receptor for *Escherichia coli* with fimbriae F4ab/F4ac and inheritance of the receptor for F4ad. ETH. Available: http://dx.doi.org/10.3929/ethz-a-009907149.

Rasschaert R, Verdonck F, Goddeeris B, Duchateau L, Cox E. 2007. Screening of pigs resistant to F4 enterotoxigenic *Escherichia coli* (ETEC) infection. Vet Microbiol. 123:249-253.

Taube S, Jiang M, Wobus CE. 2010. Glycosphingolipids as receptors for non-enveloped viruses. Viruses. 2:1011-1049.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35:W71-W74.

Van den Broeck W, Cox E, Goddeeris BM. 1999a. Receptor-specific binding of purified F4 isolated villi. Vet Microbiol. 68:255-263.

Van den Broeck W, Cox E, Oudega B, Goddeeris BM. 1999b. Receptor-dependent immune responses in pigs after oral immunization with F4 fimbriae. Infect Immunol. 67:520-526.

Van Poucke M, Vandesompele J, Mattheeuws M, Van Zeveren A, Peelman LJ. 2005. A dual fluorescent multiprobe assay for prion protein genotyping in sheep. BMC Infect Dis. 5:13.

Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. 2009. Essentials of Glycobiology. New York (NY), Cold Spring Harbor Press.

Wang J, Jiang SW, Chen XH, Liu ZL, Peng J. 2006. Prevalence of fimbrial antigen (K88 variants, K99 and 987P) of enterotoxigenic *Escherichia coli* from neonatal and postweaning piglets with diarrhea in central China. Asian-Aust J Anim Sci. 19:1342-1346.

Xu YH, Barnes S, Sun Y, Grabowski GA. 2010. Multi-system disorders of glycosphingolipid and ganglioside metabolism. J Lipid Res. 51:1643-1675.

Yan X, Huang X, Ren J, Zou Z, Yang S, Ouyang J, Zeng W, Yang B, Xiao S, Huang L. 2009. Distribution of *Escherichia coli* F4 adhesion phenotypes in pigs of 15 Chinese and Western breeds and a White Duroc x Erhualian intercross. J Med Microbiol. 58:1112-1117.

CHAPTER 6

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

6. GENERAL DISCUSSION AND FUTURE PERSPECTIVES

6.1. Introduction

Enterotoxigenic *Escherichia coli* (ETEC) expressing F4 fimbriae are worldwide a major cause of diarrhea in neonatal and recently weaned pigs causing great economic loss due to decreased growth rate, cost of medication, and considerable morbidity and mortality (Alexander et al., 1994; Fairbrother et al., 2005; Do et al., 2006; Zhang et al., 2007; Amezcua et al., 2008; Madoroba et al., 2009; Kim et al., 2010; De la Fe Rodriguez et al., 2011). Adherence of F4 ETEC is species specific, because F4 fimbriae are only found on *E. coli* strains that cause diarrhea in pigs (Holland, 1990). The bacteria colonize the small intestine by binding with their F4 fimbriae to specific receptors and produce enterotoxins inducing a secretory diarrhea in young pigs (Fairbrother et al., 2005). The F4 fimbriae, composed of major and minor subunits, exist in three serological variants, namely F4ab, F4ac and F4ad (Orskov et al., 1964; Guinee and Jansen, 1979; Mooi and de Graaf, 1985). These variants differ in the amino acid composition of the major fimbrial subunit FaeG, which has adhesive properties and recognizes glycoconjugates on the surface of enterocytes (Kearns and Gibbons et al., 1979; Sellwood, 1980; Bakker et al., 1992; Van den Broeck et al., 1999a). Although different F4 receptor profiles were observed due to different F4 ETEC adhesion phenotypes, no causal mutation in previously proposed candidate genes has been identified yet (Ren et al., 2012; Schroyen et al., 2012).

Neonatal piglets can be protected against F4 ETEC infections by pathogen-specific antibodies in the sow's colostrum and milk (Gaastra and de Graaf et al., 1982; Sellwood, 1982). This passive maternal protection can be increased by vaccination of the sow (Rutter et al., 1976; Fürer et al., 1982). However at weaning, the lactogenic maternal protection ceases making the recently weaned piglets susceptible to an F4 ETEC infection. Although mucosal protection can be obtained by vaccinating F4 ETEC susceptible piglets via the oral route, no commercial oral vaccines are available yet (Bianchi et al., 1996; Van den Broeck et al., 1999b; Melkebeek et al., 2013).

Therefore, diarrhea caused by F4 ETEC remains a serious problem in weaned piglets. Furthermore, an increase in multi-drug resistant porcine ETEC strains due to intensive use of therapeutic antibiotics in order to suppress infections has been observed (Van Driessche et al., 1995; Maynard et al., 2003; Boerlin et al., 2005, Fairbrother et al., 2005; Lee et al., 2009; Costa et al., 2010; Li et al., 2012). Thus, alternative prevention strategies are urgently needed

to reduce the antibiotic use and the economic losses caused by F4 ETEC infections in swine production.

A strategy for reducing the F4 ETEC diarrhea in pigs is to design a diagnostic test identifying the F4 ETEC resistant pigs to preferentially use these pigs in breeding programs. The *in vitro* villous adhesion test or similar *in vitro* test (Sellwood et al., 1975; Van den Broeck et al., 1999a) are difficult to use as routine test because it either needs major intestinal surgery or slaughter of the pigs to sample intestinal mucosa. Therefore, a genotyping test based on a genetic polymorphism linked to F4 ETEC resistance would be preferable, because it would allow predicting the F4 ETEC adhesion phenotype in live animals (Jørgensen et al., 2004).
Previously, Jørgensen et al. (2004) proposed *MUC4* as a potential candidate gene for the F4ab/ac receptor and found that the *MUC4* g.8227C>G polymorphism in intron 7 was highly associated with F4ab/ac ETEC adhesion phenotypes. Based on the results of previous linkage studies (Rampoldi et al., 2011; Fu et al., 2012) and based on the result of our GWAS including the g.8227G>C mutation of *MUC4* polymorphism (**Chapter 3**), it was concluded that the causal mutation for the F4ab/ac ETEC susceptibility is located distal to *MUC4*. Polymorphisms in the transmembrane mucin 13 (*MUC13*) gene, including the patented *MUC13-813* polymorphism, were found to be in strong linkage disequilibrium with the ETEC F4ab/ac receptor locus. However, none of these *MUC13* polymorphisms are causal mutations (Zhang et al., 2008; Huang, 2010; Ren et al., 2012; Rampoldi, 2013).
Therefore, the main focus of this thesis was to identify the causal mutation(s) responsible for F4 ETEC susceptibility as well as its target gene(s) and to develop a new diagnostic genotyping test for the identified causal mutation(s).

6.2. Determination of the F4 phenotype

Recently, Rampoldi et al. (2014) divided the phenotypes A to E into eight receptor phenotypes based on the weak adhesion to F4ab ETEC and F4ad ETEC (Table 1.1; **Chapter 1**). The F4 adhesion phenotypes were assessed using the microscopic adhesion test according to Vögeli et al. (1996) and Python et al. (2002). For pigs with more than 85 % adhesive enterocytes for an F4 variant, the term fully adhesive (FA) was used. For pigs with more than 0 % and less than 85 % adhesive enterocytes, the term partially adhesive (PA) was used.

In our experiment, the presence of the F4 receptor was determined in 191 pigs (3- to 18-week-old) by performing the *in vitro* villous adhesion assay as described in Van den Broeck et al. (1999b) (data unpublished). Adhesion of less than 5, less than 30 and more than 30 bacteria per 250 μm villous brush border length was noted as non-adhesive, weak adhesive and strong adhesive (Cox and Houvenaghel, 1993). The specificity of the adhesion test used in our studies (**Chapter 3 to 5**) was assessed by blocking experiments with F4-specific monoclonal antibodies and purified F4 fimbriae, and by adhesion studies with bacteria expressing no F4 fimbriae. The data proved that this adhesion test is a reliable tool to demonstrate the presence of the F4 receptor in pigs (Van den Broeck et al., 1999c).

The adhesive phenotype designations (A to F) according to Bijlsma et al. (1982) and Baker et al. (1997) can be further divided into subphenotypes (A-I to A-VIII, B-I to B-IV, C-I to C-II, D-I to D-II, F-I to F-II) based on the F4 ETEC adhesion pattern observed in our pigs of the following breeds: Large White, Belgian Landrace, Large White x Belgian Landrace crossbreds, Large White x Piétrain crossbreds, and crossbreds of multiple breeds (Table 6.1). Phenotype A (60.21 %) and E (23.56 %) were predominantly detected in our population. A similar result was observed in the study of Rampoldi et al. (2014) using Large White purebreds and Large White x Landrace crossbreds (Table 1.2; **Chapter 1**). Phenotype G and H, mainly observed in eastern breeds, were absent in our population (Bonneau et al., 1990, Li et al., 2007; Yan et al., 2009b). In a recent study, the prevalence of F4 fimbriae in *E. coli* isolates from pigs with PWD was 53 % (31 of the 59 isolates) in Belgium (unpublished data). No current data on the prevalence of F4 fimbrial E. coli isolates from pigs with neonatal diarrhea in Belgium are available. Therefore, it can be concluded that significant economic losses will arise when F4 ETEC infection occurs in this population due to the high F4 ETEC susceptibility rate (76.44 %).

Table 6.1: Summary of F4 ETEC adhesion phenotypes observed in our thesis.

PHENOTYPE[1]	ADHESION[2]			No. of pigs	%
	F4ab	F4ac	F4ad		
A	●	●	●	**115**	**60.21**
A-I	+++	+++	+++	55	28.80
A-II	+++	+++	+	25	13.10
A-III	+	+++	+++	8	4.19
A-IV	+	+++	+	8	4.19
A-V	+++	+	+++	10	5.24
A-VI	+	+	+++	3	1.57
A-VII	+++	+	+	3	1.57
A-VIII	+	+	+	3	1.57
B	●	●	-	**12**	**6.28**
B-I	+++	+++	-	9	4.71
B-II	+	+++	-	1	0.52
B-III	+++	+	-	1	0.52
B-IV	+	+	-	1	0.52
C	●	-	●	**5**	**2.62**
C-I	+	-	+++	4	2.10
C-II	+	-	+	1	0.52
D	-	-	●	**12**	**6.28**
D-I	-	-	+++	5	2.62
D-II	-	-	+	7	3.67
E	-	-	-	**45**	**23.56**
F	●	-	-	**2**	**1.05**
F-I	+++	-	-	1	0.52
F-II	+	-	-	1	0.52
G	-	●	-	**0**	**0**
H	-	●	●	**0**	**0**

[1]Phenotype designations in grey are according to Bijlsma et al. (1982), Baker et al. (1997), Bonneau et al. (1990), Li et al. (2007) and Yan et al. (2009b). Phenotype designations in white are according to our thesis.
[2]Adhesion in grey is defined as follows: '●' denotes adhesion, '-' denotes no adhesion. Adhesion in white is defined as follows: '+++' denotes strong adhesion, '+' denotes weak adhesion, '-' denotes no adhesion.

The main observation of determining the F4 ETEC adhesion phenotypes is that 39.8 % of the pigs showed at least weak adhesion towards one of the F4 variants supporting the hypothesis that mutations controlling the expression of the F4 receptor or controlling the post-expression inhibition or modification on the F4 receptor sites play a role in F4 ETEC susceptibility (Bijlsma and Bouw, 1987; Python et al., 2005; Li et al., 2007). When comparing the different F4 ETEC adhesion strengths between the F4 variants in our population, a similar pattern can be observed between the F4ab and F4ac variant (See Table 6.2). For the F4ad variant, a higher percentage of pigs showing weak adhesion was found than for the F4ab and F4ac variants.

Table 6.2: Distribution of the adhesion strength within each F4 variant observed in our thesis.

F4 ETEC	ADHESION STRENGTH			No. of pigs	%
	STRONG ADHESION	WEAK ADHESION	NO ADHESION		
F4ab ETEC	104 (54.45)	30 (15.71)	57 (29.84)	191	100
F4ac ETEC	106 (55.50)	21 (10.99)	64 (33.51)	191	100
F4ad ETEC	85 (44.50)	47 (24.61)	59 (30.89)	191	100

Numbers in parentheses indicate the percentage of pigs showing the adhesion strength.

6.3. Identification of causal mutation(s) located on chromosome 13

A genome-wide association study (GWAS) using the Porcine SNP60 DNA BeadChip was performed (**Chapter 3**) to refine the candidate region for F4ab/ac ETEC susceptibility in well-phenotyped piglets. The reliable assessment of the F4 ETEC phenotype using the adhesion test as described by Van den Broeck et al. (1999b) is important to accurately pinpoint the phenotype-causing mutation(s). In addition, a selected region harboring the SNPs with the highest evidence of association was sequenced in F4ab/ac receptor-positive and F4ab/ac receptor-negative pigs in order to identify the causal mutation(s) (**Chapter 6**; see below).

MUC13 was recently proposed as the causal gene for F4ac ETEC susceptibility based on the predicted differential O-glycosylation patterns of the 2 MUC13 variants (MUC13A and MUC13B) in F4ac ETEC susceptible and F4ac ETEC resistant pigs. Unlike MUC13A that lacks an O-glycosylation site, MUC13B is predicted to be heavily O-glycosylated and therefore would contain the binding site for F4ac ETEC (Ren et al., 2012). This hypothesis suggests that the *MUC13B* allele (present in F4ac ETEC susceptible pigs) is dominant over the *MUC13A* allele (present in the F4ac ETEC resistant pigs). In addition, Rampoldi (2013) also suggested that exon 2 of *MUC13* is the most likely location for the F4ab/acR locus. However, the results of the Indel *MUC13* marker test, the GWAS and the F4ac fimbriae binding assay in our study independently confirmed that MUC13 does not play a role in F4ab/ac ETEC susceptibility, and therefore rejects the hypothesis that *MUC13* is the causal gene for F4ab/ac ETEC susceptibility.

For our GWAS using the Porcine SNP60 BeadChip, the goal was to use the matching control-case design to control confounding and to reduce the potential impact of population

stratification by selecting non-adhesive pigs and strong adhesive pigs originating from the same litter (Schaid, 2002). The pigs were phenotyped for the presence of the F4ab/ac receptor (F4ab/acR) using the *in vitro* villous adhesion assay as described by Van den Broeck et al. (1999b). Only pigs with a clearly defined phenotype were used in this study. Pigs that were phenotyped as weak adhesive towards F4ab/ac ETEC were excluded from our study design, because the inheritance of this phenotype is still unclear. The result of our GWAS slightly differs from the results from previous studies (Rampoldi et al., 2011; Fu et al., 2012; Ren et al., 2012) using the Porcine SNP60 BeadChip. This could be due to the different cut-off values used to differentiate the different phenotypes (adhesive and non-adhesive) measured by the number of bacteria adhering to brush borders prepared from enterocytes (Ren et al., 2012), isolated enterocytes (Rampoldi et al., 2011; Fu et al., 2012) or a length of villous brush border (Goetstouwers et al., 2014) and due to the different composition of buffers used in the adhesion tests (Edfors and Torremorell, 2010). Consequently, the outcome of these tests may be influenced and may complicate interpretations concerning the genetic basis of F4 ETEC resistance (Edfors and Torremorell, 2010).

Based on the result of our GWAS, the candidate region could be refined to an 183,122 bp region between the 2 SNP markers, *MARC0002946* (SNPa) and the Indel *MUC13* marker (chr13: 144,810,100-144,993,222; Figure 3.5. in **Chapter 3**). This candidate region containing the SNPs with the strongest association (SNP1 and SNP2) lacks annotated genes indicating that previously suggested candidate genes on chromosome 13 (SSC13) are not responsible for F4ab/ac ETEC susceptibility. Also, the sequence of the porcine GenomeBuild 10.2 (between NW_003611795.1 and NW_003617796.1) contains a sequence gap (gap 1) in the candidate region. A multispecies comparison of this region to identify the gap sequence (between NW_003611795.1 and NW_003617796.1; chr13: 144,865,325 - 144,915,326) as well as non-annotated porcine genes was performed. No new non-annotated porcine genes were identified and we were unable to identify the gap sequence. Based on these results, it is very likely that the candidate region contains a *trans*-acting element interacting with a distant gene influencing F4ab/ac ETEC adherence in pigs. However, the presence of a porcine orphan gene in the candidate region cannot be ruled out. Fang et al. (2012) identified 240 orphan genes with no counterpart in any other organism (human, horse, dog, cat, cattle, rat and mouse) during analysis of the genome sequence of the Wuzhishan miniature pig.

A region of 85,155 bp (chr13: 144,936,341 - 145,021,495 bp; Figure 6.1) harboring the SNPs with the highest evidence of association (SNP1, SNP2 and SNP3) was sequenced as a starting

point in an attempt to find the causal mutation(s) for F4ab/ac ETEC susceptibility (data unpublished). It should be noted that this region does not contain the previously mentioned sequence gap (gap 1; between NW_003611795.1 and NW_003617796.1). Animals from four different litters of different breeds comprising one F4ab/acR$^+$ pig and one F4ab/acR$^-$ pig were used to reduce the confounding effect of different genetic background. The two pigs exhibiting different genotypes than expected (1 strong adhesive F4Rab/ac pig and 1 non-adhesive F4ab/ac pig) were also included in this study. All mutations identified during sequencing showed the opposite genotype than expected for these two pigs. Although the *in vitro* villous adhesion test has been proven to be reliable (Van den Broeck et al., 1999c), it is possible that the correlation between the *in vitro* F4-mediated adhesion to isolated villi and F4 ETEC susceptibility is not absolute or that an accidental sample swap has occurred. Elucidation might be provided when the candidate region is sequenced in pigs originating from the same litter as the two pigs exhibiting different genotypes. For further discussion, the results of these pigs were disregarded.

In total 725 mutations were found in the selected region, 274 mutations correlated with the F4ab/ac ETEC adhesion phenotype (38 %), 230 mutations correlated with the 2 MUC13 isoforms (MUC13A and MUC13B) (32 %), and 221 mutations that were not correlated with F4ab/ac ETEC adhesion phenotype or the 2 MUC13 isoforms (30 %) (Figure 6.1). In order to identify highly conserved regions and potential regulatory sequences in the associated region, a comparison of orthologous sequences from multiple species was performed using BLAST analysis (Benson et al., 2009). This comparison did not reveal any potential important mutations identified in the associated region.

The selected candidate region for sequencing contains the tandem repeat region in exon 2 of *MUC13* (gap 2 in Figure 6.1). To fill this tandem repeat region, the porcine bacterial artificial chromosome (BAC) bank at the CRB GADIE facility in France (http://crb-gadie.inra.fr/) was screened with *MUC13*-specific primers. Unfortunately, no *MUC13* clones were found during the screening (data not published) and thus the sequence of the tandem repeat region remains unknown. The attempt to close this sequence gap (gap 2) was carried out prior to genotyping the piglets for the Indel *MUC13* marker and the F4ac fimbriae binding assay. Based on the overall result of our study, it can be concluded that this region can be excluded from further analysis.

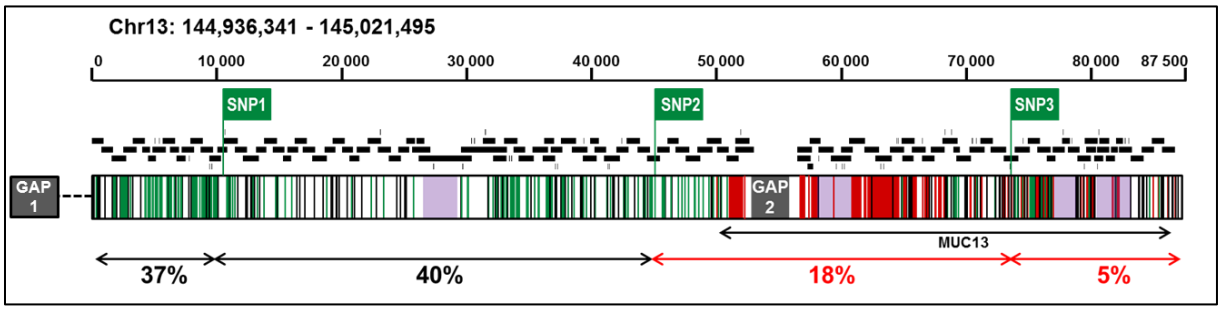

Figure 6.1: Schematic representation of the mutations found in an 85 kb region on SSC13 by sequencing.
The black boxes above the beam represent the primer pairs and sequence primers used for sequencing. Gap 1 before the beam indicates the region between NW_003611795.1 and NW_003617796.1 (chr13: 144,865,325 - 144,915,326). Gap 2 within the beam indicates the tandem repeat region in exon 2 of *MUC13*. The purple regions within the beam represent the regions requiring further sequencing. The green regions within the beam represent the location of the mutations correlated with the F4ab/ac ETEC adhesion phenotype, the red regions represent the location of the mutations correlated with the 2 MUC13 isoforms, and the black regions represent the location of the mutation with no correlation with the F4ab/ac ETEC adhesion phenotype or the 2 MUC13 isoforms.
The percentages under the beam indicate the distribution of the mutations correlated to the F4ab/ac ETEC adhesion phenotype.

Because SNP1 (*ASGA0089965*) and SNP2 (*ASGA0091537*) in our GWAS are almost completely linked with the F4ab/ac ETEC adhesion phenotype, except for 1 strong adhesive F4Rab/ac pig (1/68) and 1 non-adhesive F4ab/ac pig (1/52), it could be useful to genotype these SNPs in live animals used in breeding programs. However, during sequencing of the amplicon harboring SNP1, no variation was found at the location of SNP1 (*ASGA0089965*, chr13: 144,946,742). It was concluded that both strands of DNA were sequenced, because another heterozygous polymorphism was detected in the amplicon. The result might be explained by a sequence duplication of this region within the genome. If SNP1 is located in duplicated regions, the heterozygous genotype (GT) assigned to this SNP may arise from the homozygous genotypes present at the 2 targeted loci (GG at locus 1 and TT at locus 2). To determine whether duplicated regions harboring SNP1 were present in the pig's genome, BLAST analysis was performed to identify differences in the adjacent regions of SNP1 (Benson et al., 2009). No sequences indicating that a duplication event had occurred were found. A next step to examine whether SNP1 is located in duplicated regions would be to

perform Southern blotting analysis using an oligonucleotide probe with the identical sequence as the sequence of the SNP marker *ASGA0089965* (Southern, 1975). After cloning the DNA fragments of interest, the clones can be further sequenced identifying whether duplicated regions containing SNP1 are present in the pig's genome.

Nevertheless, it can be concluded that SNP2 can be used in a genotyping test and that a more reliable result will be obtained than the DNA-based test based on the *MUC4* g.8227C>G polymorphism or the Indel *MUC13* test. The SNP2 genotyping test is currently performed at the Laboratory of Animal Genetics to identify the F4ab/ac ETEC phenotypes (resistant and susceptible) in pigs.

6.4. Identification of target genes responsible for F4 ETEC binding

Regulatory sequences can be divided into 2 broad classes, namely *cis*-acting or *trans*-acting regulatory sequences. While *cis*-acting regulatory elements affect the expression of neighboring genes, *trans*-acting regulatory elements affect the expression of genes distributed across the genome (Strachan et al., 2014). The hypothesis that a genetic variation exerts remote regulatory effects on genes whose coding regions lie outside the boundaries of the region of maximal association was examined by investigating whether a still undiscovered genetic variation in the proposed candidate region of SSC13 could influence the expression of genes involved in previously observed F4 ETEC binding differences.

First, the expression pattern of porcine aminopeptidase N (*ANPEP*) was investigated (**Chapter 4**). ANPEP has been found to act as an endocytotic F4ac receptor by comparative proteomic analysis of the brush border proteins in F4ac receptor-positive ($F4acR^+$) and F4ac receptor-negative ($F4acR^-$) pigs (Melkebeek et al., 2012). The gene encoding ANPEP [GenBank: NC_010449.4] is located on SSC7 and not on SSC13 where the locus controlling F4ab/ac ETEC susceptibility has been mapped on. No expression difference (expression or not, or expression of isoforms with a detectable amplicon size difference) of *ANPEP* was detected in our small group of carefully F4 phenotyped pigs, rejecting the hypothesis that the expression of *ANPEP* is regulated by *trans*-acting factors present in the candidate region on SSC13. Previously, it was concluded that the carbohydrate moiety of glycoconjugates appears to be necessary for establishing adhesion with the F4 adhesin (Erickson et al., 1992; Grange and Mouricout, 1996; Grange et al., 1999; Verdonck et al., 2004; Coddens et al., 2011; Melkebeek et al., 2012). Therefore, we suggest that the susceptibility towards F4ac ETEC

mediated by ANPEP is not due to modifications in the protein itself, but due to the post-translational modifications of ANPEP by a genetic variation influencing the glycosylation of ANPEP.

Second, the interaction between the carbohydrate moiety of glycosphingolypids (GSLs) and F4 ETEC and their fimbriae has been confirmed by Coddens et al. (2011). Twelve genes (i.e. *ARSA*, *B4GALT6*, *GAL3ST1*, *GALC*, *GBA*, *GLA*, *GLB1*, *GLB1L*, *NEU1*, *NEU2*, *UGCG* and *UGT8*) were selected from the cerebroside-sulfatid region of the sphingolipid metabolism pathway (**Chapter 5**), because these genes are involved in the assembly of the F4 binding carbohydrate moiety of GSLs. None of the investigated genes are positioned on SSC13.
In addition, sialic acid-containing glycans are considered to play a role in receptor recognition by F4ac fimbria. Previous studies performing sialidase (also known as neuraminidase) treatment of purified F4ac receptors (Erickson et al., 1994) and BBMVs (Melkebeek et al., 2012) observed a reduction in binding of F4ac fimbria. Therefore, it was highly interesting to investigate whether genetic variation in sialidase 1 (*NEU1*) and sialidase 2 (*NEU2*), exoglycosidases hydrolyzing terminal sialic acid residues, could explain the previous observed findings (Monti et al., 2010).
It can be concluded that F4ab/ac/ad ETEC susceptibility is not caused by any mutation affecting the expression level of any of the investigated genes nor by the expression of splice variants. In addition, no differential structural mutation located within the selected genes explaining the F4ad binding difference was found.

Although previous results (Python et al., 2005; Schroyen et al., 2008; Ouyang et al., 2011) did not indicate that the following three genes of the glucosyl/galactosyltransferse family located on SSC13 play a role in F4 ETEC susceptibility: *B3GNT5*, *B4GALT4* and *B3GALT3*, it would be of interest to analyze the gene expression of these genes as described in the studies above (**Chapter 4 to 5**) to identify whether the gene expression pattern of these genes is consistent with the F4 ETEC phenotype.

6.5. Further perspectives

The economic losses caused by F4 ETEC diarrhea, the increasing prevalence of antimicrobial resistance in porcine *E. coli* strains and the lack of an effective vaccine for prevention of post-weaning diarrhea caused by F4 ETEC would justify the genetic selection against F4 ETEC susceptibility in pigs. When selecting for disease resistance, possible adverse effects on the susceptibility to other diseases and on production traits must be taken into account. It has been reported that the genetic improvement in resistance to ketosis and retained placenta are a correlated response to selection against chronic mastitis in dairy cattle (Heringstad et al., 2007). Therefore, the association between the SNP2 marker, the susceptibility to other diseases, and production traits should be closely monitored.

The mechanism of action by which the associated candidate region influences the F4ab/ac ETEC susceptibility in pig is still unclear. It is unknown which genetic variation present in the candidate region could influence the different F4 ETEC adhesion phenotypes observed in pigs. The result of our genome-wide association study is therefore the starting point for future genetic and functional studies. Also, a GWAS for F4ad ETEC by selecting F4ad phenotyped pigs would be a good starting point to reveal the relevant mutation(s) for F4ad ETEC susceptibility. It is possible that the associated region for F4ad ETEC susceptibility would demand additional exploration as described below for F4ab/ac ETEC.

To facilitate future genetic and functional studies, it is essential to refine the location of causal variants as sharply as possible (McCarthy and Hirschhorn, 2008). Because the gap (gap 1) located within the identified candidate region between NW_003611795.1 and NW_003617796.1 prevents further analysis, filling in the missing sequence by sequencing is required. One approach to obtain the sequence of the gap is to use long-range PCR followed by high throughput sequencing (also known as next-generation sequencing). By modifying the polymerases, long-range PCR can amplify up to 30 kb or longer genomic DNA (Jia et al., 2014). However, in practice it is often difficult to design a robust assay (Taylor and Delaney, 2010). Another approach to fill this gap with unknown size is screening a porcine genomic library with primers adjacent to the gap in order to identify a clone spanning this critical region (Bogden et al., 2011). By completely sequencing the clone using direct Sanger sequencing (Sanger et al., 1977) or high throughput sequencing (Myllykangas et al., 2012), the sequence of the gap can be identified. After identifying the sequence of the gap, BLAST analysis can be performed to identify whether a sequence duplicate of SNP1 is present in the

gap region (Benson et al., 2009). If this is the case, the SNP1 marker can be included in the genotyping test to identify F4ab/ac ETEC resistant pigs.

The candidate region for F4ab/ac ETEC susceptibility can be further refined by sequencing across the entire region of association. Because GWAS uses a dense set of markers in and close to the associated region, it would be difficult to increase the strength of association by fine mapping (McCarthy and Hirschhorn, 2008). Targeted sequencing of the candidate region using next generation sequencing in F4 phenotyped pigs would then be of interest in order to have a much more complete catalogue of potential causal DNA variation(s) (Grada and Weinbrecht, 2013). Further validation of the identified variant(s) in a large data set is required in order to find the true causal variant(s). Because it is likely that the causal variant(s) will be located in a non-coding region influencing gene regulation, whole-exome sequencing identifying the transcript expression levels in $F4R^+$ and $F4R^-$ pigs could help to find the target gene(s) of the regulatory element located on SSC13. In addition, novel orphan genes in the porcine genome can be identified by providing sequence information for protein-coding regions in the candidate region (Grada and Weinbrecht, 2013).

In order to provide further insight into the genetic basis of F4ab/ac ETEC susceptibility and to confirm the target gene(s) of the causal mutation(s), functional studies identifying the exact receptor structures of F4ab/ac ETEC and elucidating the potential mechanisms influencing this interaction between the F4 receptors and F4ab/ac ETEC should be performed. It has been proposed to characterize the glycans of the F4 binding glycoconjugates as an important next step (Coddens et al., 2011). For instance, blocking experiments using lectins or glycan-binding antibodies can profile the glycan structures on the F4 glycoconjugates. Lectins are non-immunological glycan-binding proteins exhibiting specificity towards a defined glycan motif or structural feature whereas antiglycan antibodies are generally highly specific for a single glycan determinant (Stebbins and Sasisekharan, 2014). A screening using a glycan array was already performed via the consortium for functional glycomics (CFG) without success (unpublished results). It should be noted that the protein and lipid moieties of the F4 binding glycoconjugates can also play a role in the F4 fimbriae-carbohydrate recognition and must also be taken into account in further studies (Grange et al., 1999). For instance, the causal variant can influence the expression levels of these proteins and lipids harboring the F4 binding glycan structures and thus explaining the observed F4 adhesion phenotypes.

It can be concluded that identifying the causal variant(s) for F4 ETEC susceptibility and unravelling its mechanism will be challenging, but it will bring us closer to elucidating the

genetic basis of this complex disease leading to new strategies to reduce diarrhea caused by F4 ETEC in swine production.

6.6. References

Alexander TJL. 1994. Neonatal Diarrhoea in Pigs. In: Gyles CL (ed), *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, UK, 151-170.

Amezcua R, Friendship RM, Dewey CE. 2008. An investigation of the presence of *Escherichia coli* O149:K91:F4 on pig farms in southern Ontario and the use of antimicrobials and risk factors associated with the presence of this serogroup. Can Vet J. 49:39-45.

Baker DR, Billey LO, Francis DH. 1997. Distribution of K88 *Escherichia coli*-adhesive and nonadhesive phenotypes among pigs of four breeds. Vet Microbiol. 54:123-132.

Bakker D, Willemsen PT, Simons LH, van Zijderveld FG, de Graaf FK. 1992. Characterization of the antigenic and adhesive properties of FaeG, the major subunit of K88 fimbriae. Mol Microbiol. 6:247-255.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. Nucleic Acids Res. 37:D26-31.

Bianchi AT, Scholten JW, van Zijderveld AM, van Zijderveld FG, Bokhout BA. 1996. Parenteral vaccination of mice and piglets with F4+ *Escherichia coli* suppresses the enteric anti-F4 response upon oral infection. Vaccine. 14:199-206.

Bijlsma IGW, deNijs A, van der Meer C, Frik JF. 1982. Different pig phenotypes affect adherence of *Escherichia coli* to jejunal brush borders by K88ab. K88ac, or K88ad antigen. Infect Immun. 37:891-894.

Bijlsma IG, Bouw J. 1987. Inheritance of K88-mediated adhesion of *Escherichia coli* to jejunal brush borders in pigs: a genetic analysis. Vet Res Commun. 11:509-518.

Boerlin P, Travis R, Gyles CL, Reid-Smith R, Janecko N, Lim H, Nicholson V, McEwen SA, Friendship R, Archambault M. 2005. Antimicrobial resistance and virulence genes of *Escherichia coli* isolates from swine in Ontario. Appl Environ Microbiol. 71:6753-6761.

Bogden R, Stormo K, Dobry J, Mraz A, Tao Q, van Eijk M, van Oeveren J, Prins M, Wittendorp J, van Haaren M. 2011. Next-generation sequencing of bacterial artificial chromosome clones for next-generation physical mapping. In: Harbers M, Kahl H (eds), Tag-based next generation sequencing. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 287-298.

Bonneau M, Duval-Iflah Y, Guérin G, Ollivier L, Renard C, Renjifo X. 1990. Aspects genetiques et microbiologiques de la colibacillose K88 chez le porc. Ann Rech Vet. 21:302-303.

Coddens A, Valis E, Benktander J, Ångström J, Breimer ME, Cox E, Teneberg S. 2011. Erythrocyte and porcine intestinal glycosphingolipids recognized by F4 fimbriae of enterotoxigenic *Escherichia coli*. PloS ONE. 6:e23309.

Costa MM, Drescher G, Maboni F, Weber SS, Schrank A, Vainstein MH, Schrank IS, Vargas AC. 2010. Virulence factors, antimicrobial resistance, and plasmid content of *Escherichia coli* isolated in swine commercial farms. Arq Bras Med Vet Zootec. 62:30-36.

Cox E, Houvenaghel A. 1993. Comparison of the *in vitro* adhesion of K88, K99, F41, and P987 positive *Escherichia coli* to intestinal villi of 4- to 5-week-old pigs. Vet Microbiol. 34:7-18.

De la Fe Rodriguez PY, Coddens A, Del Fava E, Abrahantes JC, Shkedy Z, Martin LOM, Muñoz EC, Duchateau L, Cox E, Goddeeris BM. 2011. High prevalence of $F4^+$ and $F18^+$ *Escherichia coli* in Cuban piggeries as determined by serological survey. Trop Anim Health Prod. 43:937-946.

Do TN, Cu PH, Nguyen HX, Au TX, Vu QN, Driesen SJ, Townsend KM, Chin JJ, Trott DJ. 2006. Pathotypes and serogroups of enterotoxigenic *Escherichia coli* isolated from pre-weaning pigs in north Vietnam. J Med Microbiol. 55:93-99.

Edfors I, Torremorrel M. 2010. *Escherichia coli* and *Salmonella* in pigs. In: Bishop SC, Axford RFE, Nicholas FW, Owen JB (eds), Breeding for disease resistance in farm animals, 3rd edition. CABI publishing, Wallingford, UK, 232-250.

Erickson AK, Willgohs JA, McFarland SY, Benfield DA, Francis DH. 1992. Identification of two porcine brush border glycoproteins that bind the K88ac adhesin of *Escherichia coli* and correlation of these glycoproteins with the adhesive phenotype. Infect Immun. 60:983-988.

Erickson AK, Baker DR, Bosworth BT, Casey TA, Benfield DA, Francis DH. 1994. Characterization of porcine intestinal receptors for the K88ac fimbrial adhesin of *Escherichia coli* as mucin-type sialoglycoproteins. Infect Immun. 62:5404-5410

Fairbrother JM, Nadeau E, Gyles CL. 2005. *Escherichia coli* in postweaning diarrhea in pigs: An update on bacterial types, pathogenesis and prevention strategies. Anim Health Res Rev. 6:17-39.

Fang X, Mu Y, Huang Z, Li Y, Han L, Zhang Y, Feng Y, Chen Y, Jiang X, Zhao W, Sun X, Xiong Z, Yang L, Liu H, Fan D, Mao L, Ren L, Liu C, Wang J, Li K, Wang G, Yang S, Lai L, Zhang G, Li Y, Wang J, Bolund L, Yang H, Wang Jian, Feng S, Li S, Du Y. 2012. The sequence and analysis of a Chinese pig genome. GigaScience. 1:16.

Fu WX, Yang L, Xin L, Xiao-Yan N, Xiang-Dong D, Liu JF, Zhang Q. 2012. A genome-wide association study identifies two novel promising candidate genes affecting *Escherichia coli* F4ab/ac susceptibility in swine. PLoS ONE. 7:e32127.

Fürer E, Cryz SJJ, Dorner F, Nicolet J, Wanner M, Germanier R. 1982. Protection against colibacillosis in neonatal piglets by immunization of dams with procholeragenoid. Infect Immun. 35:887-894.

Gaastra W, de Graaf FK. 1982. Host-specific fimbrial adhesins of noninvasive enterotoxigenic *Escherichia coli* strains. Microbiol Rev. 46:129-161.

Gibbons RA, Sellwood R, Burrows M, Hunter PA. 1977. Inheritance of resistance to neonatal *E. coli* diarrhea in the pig: examination of the genetic system. Theor Appl Genet. 51:65-70.

Goetstouwers T, Van Poucke M, Coppieters W, Nguyen VU, Melkebeek V, Coddens A, Van Steendam K, Deforce D, Cox E, Peelman LJ. 2014. Refined Candidate Region for F4ab/ac enterotoxigenic *Escherichia coli* Susceptibility situated proximal to *MUC13* in Pigs. PloS ONE 9:e105013.

Grada A, Weinbrecht K. 2013. Next-generation sequencing: methodology and application. J Invest Dermatol. 133:e11.

Grange PA, Mouricout MA. 1996. Transferrin associated with the porcine intestinal mucosa is a receptor specific for K88ab fimbriae of *Escherichia coli*. Infect Immun. 64:606-610.

Grange PA, Erickson AK, Levery SB, Francis DH. 1999. Identification of an intestinal neutral glycosphingolipid as a phenotype-specific receptor for the K88ad fimbrial adhesin of *Escherichia coli*. Infect Immun. 67:165-172.

Guinee PA, Jansen WH. 1979. Behavior of *Escherichia coli* K antigens K88ab, K88ac, and K88ad in immunoelectrophoresis, double diffusion, and hemagglutination. Infect Immun. 23:700-705.

Heringstad B, Klemetsdal G, Steine T. 2007. Selection responses for disease resistance in two selection experiments with Norwegian Red cows. J Dairy Sci. 90:2419-2426.

Holland RE. 1990. Some infectious causes of diarrhea in young in farm animals. Clin Microbiol Rev. 3:345-375.

Huang L. 2010. *MUC13* molecular marker for identifying the F4ac adhesin-caused diarrhea resistance of the weanling pig and the use thereof. International application published under the patent cooperation treaty (PCT), PCT/CN2009/070574 or WO 2010/066118.

Jia H, Guo Y, Zhao W, Wang K. 2014. Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. Scientific Reports. 4:5737.

Jørgensen CB, Cirera S, Archibald AL, Anderson L, Fredholm M, Edfors-Lilja I. 2004. Porcine polymorphisms and methods for detecting them. International application published under the patent cooperation treaty (PCT), PCT/DK2003/000807 or WO2004/048606A2.

Kearns M, Gibbons RA. 1979. The possible nature of pig intestinal receptor for the K88 antigen of *Escherichia coli.* FEMS Microbiol Lett. 6:165-168.

Kim YJ, Kim JH, Hur J, Lee JH. 2010. Isolation of *Escherichia coli* from piglets in South Korea with diarrhea and characteristics of the virulence genes. Can J Vet Res. 74:59-64.

Lee SI, Rayamahji N, Lee WJ, Cha SB, Shin MK, Roh YM, Sang HY. 2009. Genotypes, antibiogram, and pulsed-field gel electrophoresis profiles of *Escherichia coli* strains from piglets in Korea. J Vet Diagn Invest. 21:510-516.

Li YH, Qiu XT, Li HJ, Zhang Q. 2007. Adhesive patterns of *Escherichia coli* F4 in piglets of three breeds. J Genet Genomics. 34:591-599.

Li XT, Wang B, Li JL, Zeng FQ, Gong X. 2012. Prevalence and antimicrobial susceptibility of bacterial pathogens isolates from diseased swine in southwest, China. Afr J Microbiol Res. 6:1579-1583.

Madoroba E, Van Driessche E, De Greve H, Mast J, Ncube I, Read J, Beeckmans S. 2009. Prevalence of enterotoxigenic *Escherichia coli* virulence genes from scouring piglets in Zimbabwe. Trop Anim Health Prod. 9:1539-1547.

Maynard C, Fairbrother JM, Bekal S, Sanschagrin F, Levesque RC, Brousseau R, Masson L, Lariviere S, Harel J. 2003. Antimicrobial resistance genes in enterotoxigenic *Escherichia coli* O149:K91 isolates obtained over a 23-year period from pigs. Antimicrob Agents Chemother. 47:3214-3221.

McCarthy MI, Hirschhorn JN. 2008. Genome-wide association studies: potential next steps on a genetic journey. Hum Mol Gen. 17:R156-R165.

Melkebeek V, Rasschaert K, Bellot P, Tilleman K, Favoreel H, Deforce D, De Geest BG, Goddeeris BM, Cox E. 2012. Targeting aminopeptidase N, a newly identified receptor for F4ac fimbriae, enhances the intestinal mucosal immune response. Mucosal Immunol. 5:635-645.

Melkebeek V, Goddeeris BM, Cox E. 2013. ETEC vaccination in pigs. Vet Immunol Immunopathol. 152:37-42.

Monti E, Bonten E, D'Azzo A, Bresciani R, Venerando B, Borsani G, Schauer R, Tettamanti G. 2010. Sialidases in vertebrates: a family of enzymes tailored for several cell functions. Adv Carbohydr Chem Biochem. 64:403-479.

Mooi FR, de Graaf FK. 1985. Molecular biology of fimbriae of enterotoxigenic *Escherichia coli*. Curr Top Microbiol Immunol. 188:199-138.

Myllykangas S , Buenrostro J, Ji HP. 2012. Overview of sequencing technology platforms. In: Rodriguez-Ezpeleta N, Hackenberg M, Aransayet AM (eds), Bioinformatics for high throughput sequencing. Springer-Verlag, New York, USA, 11-26.

Orskov I, Orskov F, Sojka WJ, Wittig W. 1964. K antigens K88ab(L) and K88ac(L) in *E. coli*. A new O antigen: 0147 and a new K antigen: K89(B). Acta Pathol Microbiol Scand. 62:439-447.

Ouyang J, Zeng W, Ren J, Yan X, Zhang Z, Yang M, Han P, Huang X, Ai H, Huang L. 2011. Association of *B3GNT5* polymorphisms with susceptibility to ETEC F4ab/ac in the white Duroc × Erhualian intercross and 15 outbred pig breeds. Biochem Genet. 50:19-33.

Python P, Jörg H, Neuenschwander S, Hagger C, Stricker C, Bürgi E, Bertschinger HU, Stranzinger G, Vögeli P. 2002. Fine-mapping of the intestinal receptor locus for enterotoxigenic *Escherichia coli* F4ac on porcine chromosome 13. Anim Genet. 33:441-447.

Python P, Jorg H, Neuenschwander S, Asai-Coakwell M, Hagger C, Bürgi E, Bertschinger HU, Stranzinger G, Vögeli P. 2005. Inheritance of the F4ab, F4ac and F4ad *E. coli* receptors in swine and examination of four candidate genes for F4acR. J Anim Breed Genet. 122:5-14.

Rampoldi A, Jacobsen MJ, Bertschinger HU, Joller D, Bürgi E, Vögeli P, Andersson L, Archibald AL, Fredholm M, Jørgensen CB, Neuenschwander S. 2011. The receptor locus for *Escherichia coli* F4ab/F4ac in the pig maps distal to the *MUC4-LMLN* region. Mamm Genome. 22:122-129.

Rampoldi A, Bertschinger HU, Bürgi E, Dolf G, Sidler X, A. Bratus A, P. Vögeli P, Neuenschwander S. 2014. Inheritance of porcine receptors for enterotoxigenic *Escherichia coli* with fimbriae F4ad and their relation to other F4 receptors. Animal. 8:859-866.

Rampoldi A. 2013. Approach to the gene controlling the porcine receptor for *Escherichia coli* with fimbriae F4ab/F4ac and inheritance of the receptor for F4ad. ETH. Available: http://dx.doi.org/10.3929/ethz-a-009907149.

Ren J, Yan X, Ai H, Zhang Z, Huang X, Ouyang J, Yang M, Yang H, Han P, Zeng W, Chen Y, Guo Y, Xiao S, Ding N, Huang L. 2012. Susceptibility towards enterotoxigenic *Escherichia coli* F4ac diarrhea is governed by the *MUC13* gene in pigs. PloS ONE. 7:e445739.

Rutter JM, Jones GW, Brown GTH, Burrows MR, Luther PD. 1976. Antibacterial activity in colostrum and milk associated with protection of piglets against enteric disease caused by K88-positive *Escherichia coli*. Infect Immun. 13:667-676.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc Natn Acad Sci USA. 74:5463-5467.

Schaid DJ. 2002. Disease-marker association. In: Elston RC, Olson JM, Palmer L (eds), Biostatistical genetics and genetic epidemiology. John Wiley & Sons, New York, USA, 206-217.

Schroyen M, Verhelst R, Niewold TA, Buys N. 2008. Quantification of *B3GNT5* and *B4GALT4* expression reveals no difference between piglets with or without diarrhea. Commun Agric Appl Biol Sci. 73:205-208.

Schroyen M, Stinckens A, Verhelst R, Geens M, Cox E, Niewold T, Buys N. 2012. Susceptibility of piglets to enterotoxigenic *Escherichia coli* is not related to the expression of *MUC13* and *MUC20*. Anim Genet. 43:324-327.

Sellwood R, Gibbons RA, Jones GW, Rutter JM. 1975. Adhesion of enteropathogenic *Escherichia coli* to pig intestinal brush borders: the existence of two pig phenotypes. J Med Microbiol. 8:405-411.

Sellwood R. 1980. The interaction of the K88 antigen with porcine intestinal epithelial cell brush borders. Biochim Biophys Acta. 632:32-335.

Sellwood R. 1982. *Escherichia coli*-associated porcine neonatal diarrhea: antibacterial activities of colostrum from genetically susceptible and resistant sows. Infect Immun. 35:396-401.

Southern EM. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol. 98:503-517.

Stebbins NW, Sasisekharan. 2014. Functional glycomics analysis: challenges and methodologies. In: Simó C, Cifuentes A, García-Cañas V (eds), Fundamentals of advanced omics technologies: From Genes to Metabolites. Elsevier, Amsterdam, The Netherlands, 255-281.

Strachan T, Goodship J, Chinnery P. 2014. Principles of Gene Regulation and Epigenetics. In: Strachan T, Goodship J, Chinnery P (eds), Genetics and Genomics in Medicine. Garland Science/Taylor & Francis Group, New York, US, 149-188.

Taylor GR, Delaney CA. 2010. Detection of genomic duplications and deletions. In: Patrinos GP, Ansorge WJ (eds), Molecular Diagnostics, 2nd edition. Elsevier, London, UK, 171-182.

Van den Broeck W, Cox E, Goddeeris BM. 1999a. Receptor-specific binding of purified F4 to isolated villi. Vet Microbiol. 68:255-263.

Van den Broeck W, Cox E, Goddeeris BM. 1999b. Induction of immune responses in pigs following oral administration of purified F4 fimbriae. Vaccine. 17:2020-2029.

Van den Broeck W, Cox E, Oudega B, Goddeeris BM. 1999c. Receptor-dependent immune responses in pigs after oral immunization with F4 fimbriae. Infect Immunol. 67:520-526.

Van Driessche E, Sachez R, Dieussaert I, Kanarek L, Lintermans P, Beeckmans S. 1995. Enterotoxigenic fimbrial *Escherichia coli* lectins and their receptors: targets for probiotic treatment of diarrhoea. In: Pusztai A, Bardicz S (Eds.), Lectins: biomedical perspectives. Taylor & Francis, London, UK, 235-292.

Verdonck F, Cox E, Vancaeneghem S, Goddeeris BM. 2004. The interaction of F4 fimbriae with porcine enterocytes as analysed by surface plasmon resonance. FEMS Immunol Med Mic. 41:243-248.

Vögeli P, Bertschinger HU, Stamm M, Stricker C, Hagger C, Fries R, Rapacz J, Stranzinger G. 1996. Genes specifying receptors for F18 fimbriated *Escherichia coli*, causing oedema disease and postweaning diarrhoea in pigs, map to chromosome 6. Anim Genet. 27:321-328.

Zhang W, Zhao M, Ruesch L, Omot A, Francis D. 2007. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. Vet Microbiol. 123:145-152.

Zhang B, Ren J, Yan X, Huang X, Ji H, Peng Q, Zhang Z, Huang L. 2008. Investigation of the porcine *MUC13* gene: isolation, expression, polymorphisms and strong association with susceptibility to enterotoxigenic *Escherichia coli* F4ab/ac. Anim Genet. 39:258-266.

Yan X, Ren J, Huang X, Zhang ZY, Ouyang J, Zheng WH, Zou ZZ, Yang SJ, Yang B, Xiao S, Huang LS. 2009a. Comparison of production traits between pigs with and without the *Escherichia coli* F4 receptors in a White Duroc × Erhualian intercross F2 population. J Anim Sci. 87:334-339.

Yan X, Huang X, Ren J, Zou Z, Yang S, Ouyang J, Zeng W, Yang B, Xiao S, Huang L. 2009b. Distribution of *Escherichia coli* F4 adhesion phenotypes in pigs of 15 Chinese and Western breeds and a White Duroc x Erhualian intercross. J Med Microbiol. 58:1112-1117.

CHAPTER 7

SUMMARY - SAMENVATTING

7. SUMMARY – SAMENVATTING

7.1. Summary

Escherichia coli is the leading cause of diarrhea in newborn and recently weaned piglets causing significant economic losses in the pig husbandry due to decreased growth rate, treatment costs, and considerable mortality and morbidity. The main causative agent of neonatal diarrhea and post-weaning diarrhea is enterotoxigenic *Escherichia coli* expressing F4 fimbriae (F4 ETEC). Because the expression of F4-specific receptor is genetically determined in pig, a diagnostic genotyping test identifying the F4 ETEC resistant pigs can be developed. By using preferentially F4 ETEC resistant pigs in breeding programs, the number of F4 ETEC diarrhea outbreaks in the pig industry can be reduced. However, no causal mutation explaining the F4 ETEC susceptibility in pigs has been identified yet.

Chapter 1 reviews the current literature on F4 ETEC infection in piglets. The first part of **Chapter 1** represents an overview of the virulence factors, F4 fimbriae and enterotoxins. F4 fimbriae (F4ab, F4ac and F4ad) mediate the attachment of F4 ETEC to F4 receptors in the small intestine. By adhering to the small intestine, F4 ETEC can overcome the intestinal peristalsis, facilitating colonization of the small intestine. The enterotoxins, heat-labile (LT) and heat-stable (ST) enterotoxins produced by F4 ETEC are responsible for the diarrhea in piglets by decreasing absorption and increasing secretion of fluids and electrolytes in the small intestine.

The second part of **Chapter 1** represents the current knowledge of F4 receptors and the genetic aspect of the inheritance of these receptors. F4 fimbriae bind to glycoprotein (e.g. ANPEP for F4ac ETEC) or glycolipid receptors where the carbohydrate moiety of these glycoconjugates seems to be involved in the adhesion of F4 ETEC to the small intestine. In addition, the presence of F4ab/ac receptors in pigs is genetically determined. The expression of the F4ab/ac receptors is inherited as an autosomal dominant trait in a simple Mendelian manner. For F4ad receptors, the exact mode of inheritance is still unknown. Genetic mapping of the F4ab/ac receptor (F4ab/acR) locus has revealed that the causal mutation responsible for F4ab/ac ETEC susceptible is located on chromosome 13 (SSC13). Based on recent studies, it was suggested that the most likely location for the F4ab/acR locus is in exon 2 of mucin 13 (*MUC13*).

The overall aim of the study was to identify the causal mutation(s) responsible for F4 ETEC susceptibility as well as its target gene(s) and to develop a new diagnostic genotyping test for the identified causal mutation(s). Two different approaches were used in parallel (**Chapter 2**).

The first approach was to perform a genome-wide association study (GWAS) using the Porcine SNP60 DNA BeadChip (**Chapter 3**) to refine the candidate region for F4ab/ac ETEC susceptibility using well-phenotyped piglets. Recently, *MUC13* was proposed as the causal gene for F4ac ETEC susceptibility. The results of the Indel *MUC13* marker test, the GWAS and the F4ac fimbriae binding assay in our study independently confirmed that *MUC13* does not play a role in F4ab/ac ETEC susceptibility. The refined candidate region lacks annotated genes indicating that previous suggested candidate genes on SSC13 are also not responsible for F4ab/ac ETEC susceptibility. In addition, a selected region within the candidate region harboring the SNPs with the highest evidence of association was sequenced in F4ab/ac receptor-positive and F4ab/ac receptor-negative piglets in order to identify the causal mutation(s) (**Chapter 6**). During sequencing, 725 mutations were found in the selected region of which 274 mutations (38 %) were correlated with the F4ab/ac ETEC phenotype.

The second approach was to investigate whether genetic variation in the gene encoding ANPEP, identified as an F4ac receptor, (**Chapter 4**) and in genes involved in the assembly of the F4 binding carbohydrate moiety of glycosphingolipids (**Chapter 5**) could account for the F4 ETEC binding patterns previously observed in F4 phenotyped piglets. Although these genes are not located in the candidate region for F4ab/ac ETEC susceptibility on SSC13, it is possible that the causal mutation exerts remote regulatory effects on these genes whose coding regions lie outside the boundaries of this candidate region.

In **Chapter 4**, the expression of *ANPEP* was investigated in F4 phenotyped piglets by RT-PCR and sequencing, but no obvious differential expression difference was detected suggesting that the F4ac ETEC binding to ANPEP is not due to modifications in the protein itself, but due to post-translational modifications of ANPEP by a genetic variation influencing the glycosylation of ANPEP.

In addition, the interaction between the carbohydrate moiety of glycosphingolypids (GSLs) and F4 ETEC and their fimbriae has been previously confirmed. In **Chapter 5**, twelve genes (i.e. *ARSA*, *B4GALT6*, *GAL3ST1*, *GALC*, *GBA*, *GLA*, *GLB1*, *GLB1L*, *NEU1*, *NEU2*, *UGCG*, and *UGT8*) involved in the assembly of the F4 binding carbohydrate moiety of GSLs were selected. The expression of these genes in piglets with different F4 ETEC adhesion

phenotypes was also analyzed by RT-PCR and sequencing. The result of this study concluded that F4ab/ac/ad ETEC susceptibility is not caused by a genetic variation in the investigated genes.

In **Chapter 6**, the general discussion and future perspectives are presented.
The F4 ETEC susceptibility rate in our population was 76.44 % which indicates the importance of this study. Although no causal mutation(s) or target gene(s) for F4ab/ac ETEC susceptibility were identified, it is very likely that the identified candidate region located on SSC13 contains a *trans*-acting element interacting with a distant gene influencing F4ab/ac ETEC adherence in pigs. Also, a new marker, SNP2, was suggested to be used in a genotyping test providing a more reliable result than the DNA-based test based on the MUC4 g.8227C>G polymorphism or the Indel *MUC13* marker test. This SNP2 genotyping test is currently performed at the Laboratory of Animal Genetics to identify the F4ab/ac ETEC phenotypes (resistant and susceptible) in pigs.

7.2. Samenvatting

Escherichia coli is de belangrijkste oorzaak van diarree bij pasgeboren en pasgespeende biggen en is verantwoordelijk voor grote economische verliezen in de varkenshouderij door groeiachterstand, behandelingskosten, hoge mortaliteit en morbiditeit. De belangrijkste verwekker van neonatale diarree en speendiarree is enterotoxigene *Escherichia coli* die F4 fimbriae bezitten (F4 ETEC). Omdat de expressie van F4-specifieke receptoren genetisch bepaald is in het varken, kan een diagnostische genotyperingstest voor de selectie van F4 ETEC resistente varkens ontwikkeld worden. Door deze dieren bij voorkeur te gebruiken in fokprogramma's, kan het aantal uitbraken van F4 ETEC diarree in de varkensindustrie gereduceerd worden. Echter werd er tot op heden geen oorzakelijke mutatie voor F4 ETEC gevoeligheid in biggen gevonden.

Hoofdstuk 1 geeft een overzicht van de huidige literatuur over F4 ETEC infectie bij biggen weer. Het eerste deel van **Hoofdstuk 1** handelt over de virulentiefactoren van F4 ETEC, namelijk F4 fimbriae en enterotoxines. De F4 ETEC bacteriën hechten zich met F4 fimbriae (F4ab, F4ac en F4ad) vast aan de F4 receptoren in de dunne darm. Door deze vasthechting kan F4 ETEC de intestinale peristaltiek overwinnen waardoor bacteriële kolonisatie in de

dunne darm kan plaatsvinden. Vervolgens produceren de F4 ETEC bacteriën enterotoxines, hitte-labiele (LT) en hitte-stabiele (ST) enterotoxines, die verantwoordelijk zijn voor de symptomen van diarree. Het tweede deel van **Hoofdstuk 1** geeft de huidige kennis van de F4 receptoren en het genetische aspect van de overerving van deze receptoren weer. F4 fimbriae binden aan glycoproteïne receptoren (vb. ANPEP voor F4ac ETEC) of glycolipide receptoren waarvan de suikergroep van deze glycoconjugaten betrokken lijkt te zijn in de vasthechting van F4 ETEC aan de dunne darm. De aanwezigheid van deze F4 receptoren in varkens is bovendien genetisch bepaald waarvan de expressie van F4ab/ac receptoren als een dominante eigenschap wordt overgeërfd op een Mendeliaanse manier. De exacte wijze van overerving van de F4ad receptoren is echter nog niet gekend. Door het genetisch in kaart te brengen van de F4ab/ac receptor (F4ab/acR) locus werd er aangetoond dat de causale mutatie verantwoordelijk voor F4ab/ac ETEC gevoeligheid op chromosoom 13 gelegen is. Op basis van recente studies werd er gesuggereerd dat de F4ab/acR locus meest waarschijnlijk gelokaliseerd is in exon 2 van mucine 13 (*MUC13*).

Het doel van dit doctoraatsonderzoek was om de oorzakelijke mutatie(s) verantwoordelijk voor F4 ETEC gevoeligheid evenals het doelwitgen of -genen te identificeren en om zo een nieuwe diagnostische genotyperingstest te ontwikkelen op basis van de geïdentificeerde causale mutatie(s). Twee verschillende benaderingen werden in parallel toegepast (**Hoofdstuk 2**).

De eerste benadering was om via een genoomwijde associatiestudie (GWAS) met behulp van de Porcine SNP60 DNA BeadChip (**Hoofdstuk 3**) de kandidaatregio voor F4ab/ac ETEC gevoeligheid verder te verfijnen in goed gefenotypeerde biggen. *MUC13*, gelokaliseerd op chromosoom 13, werd onlangs voorgesteld als het oorzakelijk gen voor F4ac ETEC gevoeligheid. De resultaten van de Indel *MUC13* merker-test, de GWAS en de F4ac fimbriae bindingstest in onze studie bevestigden onafhankelijk dat *MUC13* geen rol speelt in de F4ab/ac ETEC gevoeligheid. De verfijnde kandidaatregio geïdentificeerd in deze thesis bevat geen geannoteerde genen waardoor vastgesteld kon worden dat de eerder voorgestelde kandidaatgenen op chromosoom 13 niet verantwoordelijk zijn voor F4ab/ac ETEC gevoeligheid. Daarnaast werd een gebied met de sterkst geassocieerde SNPs geselecteerd binnen de kandidaatregio en werd deze gesequeneerd in F4ab/ac receptor-positieve en F4ab/ac receptor-negatieve biggen om de causale mutatie(s) (**Hoofdstuk 6**) te identificeren.

Tijdens deze sequenering werden er 725 mutaties in het geselecteerde gebied geïdentificeerd waarvan 274 mutaties (38 %) een correlatie vertoonden met het F4ab/ac ETEC fenotype.

De tweede benadering was om te onderzoeken of een genetische variatie in het gen coderend voor ANPEP, geïdentificeerd als F4ac receptor, en in genen betrokken in de assemblage van het F4-bindend suikerdeel van glycosfingolipiden verantwoordelijk kan zijn voor de F4 ETEC bindingspatronen eerder geobserveerd in F4 gefenotypeerde biggen. Hoewel deze genen niet gelokaliseerd zijn in de kandidaatregio voor F4ab/ac ETEC gevoeligheid, is het mogelijk de oorzakelijke mutatie regulerende effecten uitoefent op genen waarvan de coderende gebieden buiten de grenzen van de geïdentificeerde kandidaatregio voor F4ab/ac ETEC gevoeligheid liggen.
In **Hoofdstuk 4** werd de expressie van *ANPEP* onderzocht in F4 gefenotypeerde biggen door middel van RT-PCR en sequenering. Er werd geen differentieel expressieverschil waargenomen wat erop wijst dat de F4ac ETEC binding aan ANPEP niet te wijten is aan wijzigingen in het eiwit zelf, maar het gevolg is van post-translationele modificaties van dit eiwit door een genetische variatie die de glycosylering van ANPEP beïnvloedt.
Daarnaast werd de interactie tussen de suikergroep van glycosfingolypiden (GSLs) en F4 ETEC en hun fimbriae eerder vastgesteld. In **Hoofdstuk 5**, werden twaalf genen (*ARSA*, *B4GALT6*, *GAL3ST1*, *GALC*, *GBA*, *GLA*, *GLB1*, *GLB1L*, *NEU1*, *NEU2*, *UGCG* en *UGT8*) geselecteerd die betrokken zijn in de assemblage van het F4-bindend suikerdeel van GSLs. De expressie van deze genen werd geanalyseerd in biggen met verschillende F4 ETEC adhesieve fenotypes door RT-PCR en sequenering. Het resultaat van deze studie concludeert dat F4ab/ac/ad ETEC gevoeligheid niet veroorzaakt wordt door een genetische variatie in de onderzochte genen.

In **Hoofdstuk 6** worden de algemene discussie en toekomstperspectieven gepresenteerd.
Hier wordt het belang van deze studie aangetoond door de hoge F4 ETEC gevoeligheidsgraad (76,44 %) in onze varkenspopulatie. Hoewel er geen causale mutatie(s) of doelwitgen(en) voor F4ab/ac ETEC gevoeligheid geïdentificeerd werden in deze thesis, is het zeer waarschijnlijk dat de kandidaatregio op chromosoom 13 een *trans*-werkend element bevat dat genen die een rol spelen in de adhesie van F4ab/ac ETEC aan de dunne darm beïnvloedt. Ook werd er een nieuwe merker, SNP2, voorgesteld om te gebruiken in een genotyperingstest met een betrouwbaarder resultaat tot gevolg dan de DNA-test op basis van het MUC4 g.8227C>G polymorfisme of de Indel *MUC13* merker-test. Deze SNP2-genotyperingstest om de F4ab/ac

ETEC fenotypes (resistent en gevoelig) in varkens te identificeren, wordt momenteel uitgevoerd in het Laboratorium voor Dierlijke Genetica.

CURRICULUM VITAE

Tiphanie Goetstouwers werd geboren op 2 juni 1985 te Wilrijk. In 2003 behaalde zij het diploma hoger secundair onderwijs aan het Sint-Ursula-Instituut te Onze-Lieve-Vrouw-Waver in de richting Moderne Talen-Wetenschappen.

In datzelfde jaar startte zij de studie Diergeneeskunde aan de Campus Groenenborger van de Universiteit Antwerpen waar zij het diploma Bachelor in de Diergeneeskunde behaalde. Drie jaar later studeerde ze af als Master in de Diergeneeskunde (optie Grote Huisdieren) aan de Universiteit Gent.

Geboeid door het wetenschappelijk onderzoek, trad zij op 1 november 2009 in dienst bij de Vakgroep Voeding, Genetica en Ethologie aan de Faculteit Diergeneeskunde. Zij werkte er gedurende vier jaar als doctoraatsstudent in het kader van een IWT-landbouwproject getiteld "Naar selectie op F4 en F18 ETEC resistentie bij biggen voor bestrijding van speendiarree", waarbij het Laboratorium Immunologie, Laboratorium Dierlijke Genetica en het Laboratorium Farmaceutische Biotechnologie de handen in elkaar sloegen. Dit onderzoek heeft geleid tot dit doctoraat met als promotor Prof. Dr. L.J. Peelman en medepromotor Prof. Dr. E. Cox en werd gedurende vier jaar gefinancierd door het agentschap voor Innovatie door Wetenschap en Technologie in Vlaanderen (IWT-Vlaanderen).

Tiphanie Goetstouwers is auteur of medeauteur van meerdere wetenschappelijke publicaties in internationale tijdschriften en woonde het ISAG congres in 2012 bij.

BIBLIOGRAPHY

Publications in international journals

Nguyen VU, **Goetstouwers T**, Coddens A, Van Poucke M, Peelman L, Deforce D, Melkebeek V, Cox E. 2013. Differentiation of F4 receptor profiles in pigs based on their mucin 4 polymorphism, responsiveness to oral F4 immunization and *in vitro* binding of F4 to villi. Vet Immunol Immunop. 152:93-100.

Goetstouwers T, Van Poucke M, Nguyen VU, Melkebeek V, Coddens A, Deforce D, Cox E, Peelman LJ. 2014. F4-related expression analysis of the aminopeptidase N gene in pigs. J Anim Sci. 92:1866-1873.

Goetstouwers T, Van Poucke M, Coppieters W, Nguyen VU, Melkebeek V, Coddens A, Van Steendam K, Deforce D, Cox E, Peelman LJ. 2014. Refined candidate region for F4ab/ac enterotoxigenic *Escherichia coli* susceptibility situated proximal to *MUC13* in pigs. PloS ONE. 9:e105013.

Goetstouwers T, Van Poucke M, Coddens A, Nguyen VU, Melkebeek V, Deforce D, Cox E , Peelman LJ. 2014. Variation in 12 porcine genes involved in the carbohydrate moiety assembly of glycosphingolipids does not account for differential binding of F4 *Escherichia coli* and their fimbriae. BMC Genet. 15:103.

Abstracts and posters at national and international scientific congresses

Nguyen VU, **Goetstouwers T**, Van Poucke M, Coddens A, Melkebeek V, Peelman L, Cox E. 2011. Mucin 4 polymorphism, F4 binding to villi and brush border membrane proteins and F4 immunisation can differentiate pigs as tools towards identifying F4R. ECMIS: *Escherichia coli* and the Mucosal Immune System: interaction, modulation and vaccination. Ghent, Belgium.

Nguyen VU, Melkebeek V, **Goetstouwers T**, Van Poucke M, Peelman L, Goddeeris B, Cox E. 2013. Elispot on enriched IgA+ B-lymphocytes can be used to detect the IgA response following oral immunization with F4 fimbriae of pigs with colostral F4-specific serum antibodies. MUCOVAD: Mucosal Vaccines, Adjuvants and Delivery. Copenhagen, Denmark.

Nguyen VU, Melkebeek V, **Goetstouwers T**, Coddens A, Van Poucke M, Peelman L, Cox E. 2012. Towards identifying porcine F4 ETEC receptors by comparing F4-binding brush border membrane proteins between well characterized groups of pigs. 4th European symposium of Porcine Health Management. Brugge, Belgium.

Goetstouwers T, Van Poucke M, Coddens A, Nguyen VU, Melkebeek V, Cox E, Peelman LJ. 2012. Identification of functional genetic variation in porcine cerebroside-sulfatid metabolism. 33rd International Society for Animal Genetics (ISAG) Conference. Cairns, Australia.

DANKWOORD

Nu wil ik iedereen bedanken die me geholpen heeft om dit doctoraatswerk te voltooien.

Eerst en vooral wil ik Prof. Dr. Luc Peelman en Prof. Dr. Eric Cox bedanken voor hun steun en bijdrage aan dit doctoraatswerk. Luc, bedankt voor de kans om te werken aan dit boeiende project, de enthousiaste begeleiding, uw goede raad en de vele wetenschappelijke adviezen. Eric, bedankt voor de vele waardevolle discussies en de fijne samenwerking.

Daarnaast wil ik Prof. Dr. Dieter Deforce bedanken voor de feedback tijdens de meetings. Prof. Dr. Nadine Buys, Rattlerow-Seghers en andere bedrijven wil ik bedanken voor de hulp met betrekking tot de selectie van de geschikte biggen voor dit onderzoek.
De uitstekende begeleiding bij het analyseren van de associatiestudie door Dr. Wouter Coppieters is zeker het vermelden waard. Ook bedankt aan de leden van de begeleidings- en examencommissie om dit proefschrift kritisch na te lezen.

Een van de leukste aspecten van mijn doctoraat was de werksfeer op het labo Genetica. Ik prijs me gelukkig een paar jaar te hebben mogen meedraaien in de stimulerende sfeer van dit labo, en dank allen die hier hebben toe bijgedragen. Mario, Domi, Linda, Caro, Ruben, Karen, Alice, Tim, Jozefien, Wenwen, Cindy, Christel, Steven en Sarah, jullie waren toffe collega's en vrienden. De koffiepauzes, de vele sportactiviteiten, traktaties en feestjes zorgden voor de nodige ontspanning in deze vier jaar. Mijn bureaugenootjes Mario, Domi en bijna-bureaugenootje Linda wil ik nog speciaal bedanken voor de hilarische momenten (spin op toetsenbord) en goede sfeer in onze bureau. Mario, je praktische en professionele ondersteuning en je kritische kijk op resultaten en publicaties zal ik niet snel vergeten. Ik heb heel veel van je geleerd op wetenschappelijk gebied. Domi, ik koester onze babbelmomenten en naast een goede collega ben je een echte vriendin geworden. Linda, bedankt voor je waardevolle hulp in het labo tijdens mijn doctoraat.

Ut, also special thanks to you. Thank you for being my "partner in crime" during sampling and for helping me during my PhD.

Vesna en Annelies, bedankt om jullie visies en ervaringen over F4 ETEC te delen met mij. Ook wil ik jullie bedanken voor jullie hulp en voor de waardevolle discussies over de resultaten.

Hoewel ze niets essentieels hebben bijgedragen tot deze thesis is ook de inbreng van mijn vrienden niet te onderschatten. Dankjewel allemaal voor de vele opbeurende mails en bemoedigende telefoontjes. De gezellige etentjes en een terrasje van tijd tot tijd gaven mij telkens weer nieuwe energie.

Zonder de liefde en de voortdurende steun en aanmoediging van mijn ouders zou ik nooit geworden zijn wie ik nu ben, daarom draag ik dit doctoraat graag aan jullie op. Jullie zijn de beste ouders die iemand zich kan indenken. Ook mijn zus, Esterelle, en mijn broer, Michaël, zijn er steeds voor mij geweest.

Een laatste woordje van dank is voor Frederik, jouw aanwezigheid en luisterend oor zijn de laatste jaren mijn grootste steun geweest. Dankjewel, gewoon om er te zijn wanneer ik je het meeste nodig had!