

Institut für Tierwissenschaften  
der Rheinischen Friedrich – Wilhelms – Universität Bonn

---

**Molecular characterization of porcine genes encoding complement  
components of the terminal lytic pathway and their association  
with hemolytic complement activity**

I n a u g u r a l – D i s s e r t a t i o n

zur

Erlangung des Grades  
Doktor der Agrarwissenschaften  
(Dr. agr.)

der

Hohen Landwirtschaftlichen Fakultät

der

Rheinischen Friedrich – Wilhelms – Universität

zu Bonn

vorgelegt im Mai 2008

von

Do Vo Anh Khoa

aus

Cantho, Vietnam

Referent: Prof. Dr. Klaus Wimmers  
Korreferent: Prof. Dr. Karl Schellander and Prof. Dr. Jéns Léon  
Tag der mündlichen Prüfung: 28. August 2008  
Erscheinungsjahr: 2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn unter  
[http://hss.ulb.uni-bonn.de/diss\\_online](http://hss.ulb.uni-bonn.de/diss_online) elektronisch publiziert.

*Dedicated to all beloved members in my large family including great grandmamma,  
parents, parents-in-law, elder brother, younger sister,  
especially, wife and daughter*

## Abstract

Molecular characterization of porcine genes encoding complement components of the terminal lytic pathway and their association with hemolytic complement activity

Activation of the complement system from three different pathways (classical, alternative and lectin pathway) results in the generation of the C3-convertase enzyme, which plays a key role in formation of the membrane attack complex (C5b-C9) causing the death of target cells. The porcine C3 and C5 complement components were characterized and studied for association with hemolytic complement activity (Kumar et al. 2004, Wimmers et al. 2003). In order to gain understanding for the membrane attack complex action in the innate immune mechanism, in this study it was focussed on the terminal complement components C6, C7, C8, and C9 to characterize their molecular structure, to detect single nucleotide polymorphisms (SNPs), to establish their location on chromosome, and to associate their genetic variation with hemolytic complement activity in both classical and alternative pathway in the pig.

The entire length of cDNA sequence of the candidate genes C6, C7, C8A, C8B, C8G and C9 were identified with 3306, 3561, 2146, 2461, 840 and 2536 bp encoding 935, 843, 589, 611, 202 and 543 amino acids, respectively. The porcine deduced protein sequence of the candidate genes showed 67-83% identities with human analogue. Respectively, screening the coding region revealed five, six, seven, nine, and two SNPs in the porcine C6, C7, C8A, C8B, and C9 but non in C8G by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP). Most of the SNPs belong to the functional protein domains such as TSP1, LDLa, MACPF, CCP and FIMAC. Genotyping for several SNP sites in three porcine breeds German Landrace (LR), Pietrain (PIE) and Muong Khuong (MK) showed that European breeds (LR and PIE) had higher allelic variation than the Asian breed (MK). All genotypic frequencies fit to Hardy-Weinberg equilibrium rule.

Using the INRA-Minnesota porcine Radiation Hybrid mapping panel, the porcine C6, C7, and C9 were assigned to the q-arm of chromosome 16 (q1.4) whereas the porcine

C8A, and C8B were mapped to chromosome 6 (q3.1-q3.5). Particularly the porcine C8G was located on chromosome 1 (q2.13).

Genetic association with hemolytic complement activity in both classical (CH50) and alternative pathway (AH50) was carried out in 417 animals of a F<sub>2</sub> DUMI resource population derived from cross between Duroc and Berlin Miniature Pig. Therefore, the F<sub>2</sub> DUMI animals were immunized with *Mycoplasma hyopneumoniae* (Mh), Aujeszky (ADV) and porcine reproductive and respiratory syndrome (PRRSV) vaccine. Sera were isolated from blood samples taken prior and post vaccinations and measurement for CH50 and AH50 was conducted thereafter. For each gene except the porcine C8G, the SNP site with amino acid substitution 862A→G for C6, 881A→G for C7, 1544C→T for C8A, 222C→T for C8B, and 407C→G for C9, segregating in the DUMI, were used for genotyping the F<sub>2</sub> animals using PCR-RFLP with the restriction enzymes TaqI, MboII, Hin6I, FnuDII, and HpyCH4III, respectively. The association results illustrated that significant difference in hemolysis among genotypes was found in CH50 for C7 (p=0.0080), and C9 (p=0.0488). However, this was close to significance for C6 (p=0.0853) and C8A (p=0.0650) in CH50. Therefore between homozygous genotypes CC and TT for C8A hemolytic activity showed significant difference (p=0.0522). There was no association of any of the candidate gene with hemolytic complement activity in the alternative pathway. Analyzing the interaction between genotypes and eight different immunization time points in AH50 revealed significant differences for C8A (p=0.0027), C8B (p=0.0231), and C9 (p=0.0340) whereas in CH50 this interaction was found significant for C8B (p=0.0048). Hemolytic complement activity showed the highest values at the fourth day after immunization with ADV vaccine for CH50 whereas linear increment during the experiment was performed for AH50. Along the vaccination program after each of complement stimulation by different vaccines, a short termed increment of complement activity was found, especially with ADV vaccine. Also male animals always performed higher hemolysis than females in both pathways. These results show that hemolytic complement activity depends on the genetic variation, sex, age, kind of vaccine, and interaction of complement components.

In summary, the obtained results provide the means for further understanding the role of C6, C7, C8, and C9 in natural immune response of the host against pathogens. It also

promotes the porcine C6, C7, C8, and C9 as candidate genes in efforts to genetically improve general animal health, a goal of breeding programmes for food animals.

Contents	Page
Abstract	IV
List of abbreviations	XI
List of tables	XIV
List of figures	XVII
1 Introduction	1
2 Literature review	3
2.1 Innate immunity and adaptive immune response	3
2.2 Complement system	4
2.2.1 The classical pathway	4
2.2.2 The lectin pathway	5
2.2.3 The alternative pathway	6
2.2.4 Membrane attack complex (MAC)	6
2.3 Molecular characterization of candidate genes	8
2.3.1 The sixth complement component (C6)	9
2.3.2 The seventh complement component (C7)	10
2.3.3 The eighth complement component (C8)	11
2.3.4 The ninth complement component (C9)	14
2.4 Functional protein domains in candidate genes	15
2.5 Polymorphisms and deficiency of candidate complement components	17
2.5.1 The sixth complement component (C6)	17
2.5.2 The seventh complement component (C7)	18
2.5.3 The alpha and gamma chain of the eighth complement component (C8A-G)	19
2.5.4 The beta chain of the eighth complement component (C8B)	20
2.5.5 The ninth complement component (C9)	21
2.6 Position of candidate genes on chromosome	22

2.7	Hemolytic complement activity	23
3	Materials and methods	27
3.1	Materials	28
3.1.1	Materials on station	28
3.1.1.1	Experimental animals, tissue and data collection	28
3.1.1.2	Phenotyping of the immune responsiveness: Sampling design	31
3.1.2	Materials for laboratory analysis	32
3.1.2.1	Chemicals, biological materials, kits and others	32
3.1.2.2	Buffers and reagents	33
3.1.2.3	Used softwares	34
3.1.2.4	Equipments	35
3.2	Methods	36
3.2.1	Immunological methods: complement activity measurement	36
3.2.2	Molecular genetic methods	37
3.2.2.1	DNA extraction	37
3.2.2.2	RNA isolation	37
3.2.2.3	Complementary DNA (cDNA) synthesis	38
3.2.2.4	Amplicon and primer design	38
3.2.2.5	Polymerase chain reaction (PCR)	39
3.2.2.6	PCR product purification	40
3.2.2.7	Sequencing	40
3.2.2.8	Screening for single nucleotide polymorphisms (SNPs)	40
3.2.2.9	SNP genotyping	41
3.2.2.10	Assignment of candidate genes to chromosomes	41
3.2.2.11	Phylogenetic analysis	43
3.2.2.12	Statistical analysis	43
4	Results	50
4.1	Molecular characterization and detection of polymorphisms	50
4.1.1	The sixth component of the porcine complement system (C6)	50



4.1.1.1	Complementary DNA and predicted protein sequence	50
4.1.1.2	Single nucleotide polymorphisms	53
4.1.2	The seventh component of the porcine complement system (C7)	53
4.1.2.1	Complementary DNA and predicted protein sequence	53
4.1.2.2	Single nucleotide polymorphisms	55
4.1.3	The alpha chain of the eighth component of the porcine complement system (C8A)	56
4.1.3.1	Complementary DNA and predicted protein sequence	56
4.1.3.2	Single nucleotide polymorphisms	57
4.1.4	The beta chain of the eighth component of the porcine complement system (C8B)	59
4.1.4.1	Complementary DNA and predicted protein sequence	59
4.1.4.2	Single nucleotide polymorphisms	60
4.1.5	The gamma chain of the eighth component of the porcine complement system (C8G)	62
4.1.5.1	Complementary DNA and predicted protein sequence	62
4.1.5.2	Single nucleotide polymorphisms	63
4.1.6	The ninth component of the porcine complement system (C9)	64
4.1.6.1	Complementary DNA and predicted protein sequence	64
4.1.6.2	Single nucleotide polymorphisms	65
4.2	Structural relationship analysis among the porcine homologous complement candidate sequences	67
4.3	Detection of genotypes and assignment of candidate genes to pig chromosome	71
4.3.1	Genotype analysis	71
4.3.2	Location of the candidate genes on porcine chromosome	76
4.4	Association of candidate genes with hemolytic complement activity	78
4.4.1	The classical lytic pathway	78
4.4.2	The alternative pathway	84
5	Discussion	85

5.1	Structure of candidate genes	85
5.1.1	Characteristic of the porcine C6 and C7 complement genes	85
5.1.2	Characteristic of the porcine C8 complement gene	87
5.1.3	Characteristic of the porcine C9 complement gene	90
5.1.4	Common and specific structural and functional features of components of the terminal lytic complement pathway	94
5.2	Polymorphisms, genotyping and assignment of candidate genes to chromosome	97
5.3	Association analysis with hemolytic complement activity	102
5.3.1	Hemolytic complement activity depending on genotypes of the candidate genes	102
5.3.2	Other factors on hemolytic complement activity in both pathways	105
5.3.2.1	Effect of gender	107
5.3.2.2	Effect of age	108
5.3.2.3	Interaction of complement components	108
5.3.2.4	Effect of vaccination	109
6	Summary	114
7	Zusammenfassung	116
8	References	120
9	Annex	153

## List of abbreviations

a.a or A.A	: Amino acid
Ala	: Alanine
Arg	: Arginine
Asn	: Asparagine
Asp	: Asparagine acid
bp	: Base pair
BMP	: Berlin Mininature Pig
<i>Bos taurus (B. taurus)</i>	: Cattle
<i>Branchiostoma belcheri (B. belcheri)</i>	: Japanese lancelet
°C	: Degree Celcius
<i>Canis lupus familiaris (C. familiaris)</i>	: Dog
CCP	: The complement control protein
cds	: Coding region
Cys	: Cysteine
C6	: Complement component C6
C7	: Complement component C7
C8A	: Complement component C8A
C8B	: Complement component C8B
C8G	: Complement component C8G
C9	: Complement component C9
cDNA	: Complementary deoxyribonucleic acid
<i>Ctenopharyngodon idella (C. idella)</i>	: Grass carp
<i>Danio rerio (D. rerio)</i>	: Zebrafish
DNA	: Deoxyribonucleic acid
dNTP	: Deoxyribonucleotide triphosphate
DR	: Duroc porcine breed
DUMI	: Duroc Berlin Miniature pig resource population
EGF	: Epidermal growth factor module
<i>Equus caballus (E. caballus)</i>	: Horse

Fc	: Crystallizable fragment of immunoglobulin G molecule
FIMAC	: Factor I membrane attack complex
<i>Fundulus heteroclitus</i> ( <i>F. heteroclitus</i> )	: Killifish
<i>Gallus gallus</i> ( <i>G. gallus</i> )	: Red jungle fowl
Glu	: Glutamine acid
Gly	: Glycine
h	: Hour
His	: Histidine
<i>Homo sapiens</i> ( <i>H.sapiens</i> )	: Human
HS	: Hampshire porcine breed
Ile	: Isoleucine
IMpRH	: INRA-Minnesota porcine radiation hybrid panel
LDLa	: Cysteine-rich repeat in the low-density lipoprotein
Leu	: Leucine
Lipocalin	: Lipocalin protein domain
LOD	: Log of the odds
LR	: German Landrace porcine breed
LSM	: Least square mean
Lys	: Lysine
MAC	: Membrane attack complex
<i>Macaca mulatta</i> ( <i>M. mulatta</i> )	: Rhesus monkey
MACPF	: Membrane-attack complex / perforin
Met	: Methionine
min	: Minute
MK	: Muong Khuong porcine breed
ml	: Milliliter
<i>Monodelphis domestica</i> ( <i>M. domestica</i> )	: Gray short-tailed opossum
mRNA	: Messenger ribonucleic acid
<i>Mus musculus</i> ( <i>M. musculus</i> )	: House mouse
ng	: Nanogram

nt	: Nucleotide
<i>Oncorhynchus mykiss</i> ( <i>O. mykiss</i> )	: Rainbow trout
ORF	: Open reading frame
<i>Oryctolagus cuniculus</i> ( <i>O. cuniculus</i> )	: rabbit
<i>Pan troglodytes</i> ( <i>P. troglodytes</i> )	: Chimpanzee
<i>Paralichthys olivaceus</i> ( <i>P. olivaceus</i> )	: Bastard halibut
PCR	: Polymerase chain reaction
PCR-RFLP	: Polymerase chain reaction-Restriction fragment length polymorphism
PIE	: Pietrain porcine breed
<i>Pongo pygmaeus</i> ( <i>P. pygmaeus</i> )	: Orangutan
Pro	: Proline
<i>Rattus norvegicus</i> ( <i>R. norvegicus</i> )	: Norway rat
Rec. frags	: Recombination fractions
rpm	: Rotations per minute
s	: Seconds
SAS	: Statistic analysis system software
SE	: Standard error
Seq	: Sequence
Ser	: Serine
SNP	: Single nucleotide polymorphism
<i>Sus scrofa</i> ( <i>S. scrofa</i> )	: Pig
TAE	: Tris-acetate-EDTA
<i>Takifugu rubripes</i> ( <i>T. rubripes</i> )	: The Japanese pufferfish fugu rubripes
Thr	: Threonine
TSP1	: Type 1 repeats in thrombospondin-1
Tyr	: Tyrosine
U	: Unit
UTR	: Untranslated region
Val	: Valine
μl	: Microliter
<i>Xenopus tropicalis</i> ( <i>X. tropicalis</i> )	: Silurana tropicalis

## List of tables

	Page
Table 1 Oligonucleotide primer pairs used for determination of the coding sequence and 5' and 3' untranslated regions (UTR) and for screening SNP of the candidate genes	44
Table 2 Oligonucleotide primers used for radiation hybrid mapping	46
Table 3 List of restriction enzyme used for PCR-RFLP	47
Table 4 Oligonucleotide primers used to discriminate alleles for genotyping	48
Table 5 Percent nucleotide sequence and amino acid sequence identities of the porcine C6 with different species	52
Table 6 Characterization of SNPs in the porcine C6 candidate gene	53
Table 7 Percent nucleotide sequence and amino acid sequence identities of the porcine C7 with different species	55
Table 8 Characterization of SNPs in the porcine C7 candidate gene	56
Table 9 Percent nucleotide sequence and amino acid sequence identities of the porcine C8A with different species	57
Table 10 Characterization of SNPs in the porcine C8A candidate gene	59
Table 11 Percent nucleotide sequence and amino acid sequence identities of the porcine C8B with different species	61
Table 12 Characterization of SNPs in the porcine C8B candidate gene	62
Table 13 Percent nucleotide sequence and amino acid sequence identities of the porcine C8G with different species	63
Table 14 Percent nucleotide sequence and amino acid sequence identities of the porcine C9 with different species	67
Table 15 Number of cysteine residues in each functional protein domain of the candidate genes	70
Table 16 Genotypic frequencies of the candidate genes in different breeds	71
Table 17 Allelic frequencies at several SNPs in different breeds of pig	73
Table 18 Summary of RH mapping results of the candidate genes	78
Table 19 Least squares means of hemolytic activity in both the classical (CH50) and alternative (AH50) complement pathways for different genotypes	79

Table 20	Chromosomal assignments of the genes in difference species	102
Table 21	Position and length of predicted porcine C8A exons	153
Table 22	Position and length of predicted porcine C6 exons	154
Table 23	Position and length of predicted porcine C7 exons	155
Table 24	Position and length of predicted porcine C8B exons	156
Table 25	Position and length of predicted porcine C8G exons	156
Table 26	Position and length of predicted porcine C9 exons	157
Table 27	Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C6 gene	157
Table 28	Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C7 gene	158
Table 29	Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C8A gene	158
Table 30	Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C8B gene	158
Table 31	Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C9 gene	159
Table 32	Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C6 gene	159
Table 33	Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C7 gene	159
Table 34	Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C8A gene	160

Table 35	Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C8B gene	160
Table 36	Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C9 gene	160



## List of figures

	Page
Figure 1 Response to invading microorganisms by innate and adaptive immune systems	3
Figure 2 Assembly of the MAC	6
Figure 3 Structural organization of the MAC protein family	8
Figure 4 Functional sites in C8	12
Figure 5 Locations of published human C6, C7 or C9 mutations leading to deficiencies (to June 2003) in relation to the modules	22
Figure 6 General schema of experimental approach	27
Figure 7 Three different phenotypes in Muong Khuong pig breed	29
Figure 8 Sampling in Muong Khuong animals on farm	30
Figure 9 The structure of F <sub>2</sub> DUMI resource population based on reciprocal crossing of Duroc and Berlin Miniature Pig	31
Figure 10 Vaccination program and time of blood sampling before and after immunization in F <sub>2</sub> DUMI population	32
Figure 11 The phylogenic tree of C6 proteins	51
Figure 12 Location of protein domains in the porcine C6 sequence	52
Figure 13 The phylogenic tree of C7 proteins	54
Figure 14 Location of protein domains in the porcine C7 sequence	55
Figure 15 Location of protein domains in the porcine C8A sequence	58
Figure 16 The phylogenic tree of C8A proteins	58
Figure 17 The phylogenic tree of C8B proteins	60
Figure 18 Location of protein domains in the porcine C8B sequence	61
Figure 19 Location of protein domains in the porcine C8G sequence	63
Figure 20 The phylogenic tree of C8G proteins	64
Figure 21 Location of protein domains in the porcine C9 sequence	65
Figure 22 The phylogenic tree of C9 proteins	66
Figure 23 Similarity among amino acid sequences of the candidate genes	68

Figure 24	The phylogenic tree of homologous proteins of the terminal lytic pathway	68
Figure 25	Multiple alignments of homologous protein domains in both the amino terminus and carboxyl-terminus of the candidate genes C6, C7, C8A, C8B and C9	69
Figure 26	A representative pattern of comparative sequencing	74
Figure 27	Representative patterns of detection of SNP using PCR-RFLP	75
Figure 28	A representative pattern for PCR result of 118 DNA clones of the IMpRH	77
Figure 29	Position of the candidate genes on porcine chromosomes	77
Figure 30	Histogram of least squares means of hemolytic complement activity in the classical pathway depending on genotypes of the candidate genes	80
Figure 31	Histogram of least squares means of hemolytic complement activity in the alternative pathway depending on genotypes of the candidate genes	81
Figure 32	Histogram of least squares means of hemolytic complement activity in the classical pathway along vaccination for the interaction of genotypes of each candidate gene and eight different time points	82
Figure 33	Histogram of least squares means of hemolytic complement activity in the alternative pathway along vaccination for the interaction of genotypes of each candidate gene and eight different time points	83
Figure 34	3D-structure of C6 (a.a 180-515), C7 (a.a 128-411), C8A (136-497), C8B (a.a 168-508) and C9 (122-506) homologous protein fragments	92
Figure 35	Comparison of the porcine C6 (a.a 22-631) and C7 (a.a 23-575) amino-termini with whole mature porcine C8A, C8B and C9 proteins	97

Figure 36	Comparison of the porcine C6 (a.a 632-935) and C7 (a.a 576-843) carboxyl-termini with whole mature porcine C8A, C8B and C9 proteins	97
Figure 37	Correspondences between human and pig chromosomal segments. Location of the porcine C6, C7 and C9 gene on the q-arm of chromosome 16	99
Figure 38	Correspondences between human and pig chromosomal segments. Location of the porcine C8G gene on the q-arm of chromosome 1	100
Figure 39	Correspondences between human and pig chromosomal segments. Location of the porcine C8A and C8B gene on the q-arm of chromosome 6	101
Figure 40	Nucleotide and predicted protein sequence of the porcine C6 gene	161
Figure 41	Nucleotide and predicted protein sequence of the porcine C7 gene	164
Figure 42	Nucleotide and predicted protein sequence of the porcine C8A gene	167
Figure 43	Nucleotide and predicted protein sequence of the porcine C8B gene	169
Figure 44	Nucleotide and predicted protein sequence of the porcine C8G gene	172
Figure 45	Nucleotide and predicted protein sequence of the porcine C9 gene	173
Figure 46	Alignment of the C6 proteins among species	175
Figure 47	Alignment of the C7 proteins among species	179
Figure 48	Alignment of the C8A proteins among species	181
Figure 49	Alignment of the C8B proteins among species	184
Figure 50	Alignment of the C8G proteins among species	187
Figure 51	Alignment of the C9 proteins among species	188

## 1 Introduction

Infectious diseases in farm animals and their control by therapeutic, prophylactic and metaphylactic treatments and veterinary management schemes cause immense costs, animal welfare concerns and increasingly burden man and environment with chemotherapeutics. Disease infections are a main problem in animal production systems causing customer's demands that the use of medicals and vaccines should be reduced. In pig production, producers are using antibiotics to treat illness, to prevent the spread of disease and to improve the nutritional efficiency of animals. The use of antibiotics in farm animals affects the treatment of human diseases due to bacterial resistance. The question is how to reduce antibiotics in pork? Stopping the use of antibiotics for animal growth is one of the best ways done in recent years but not at all. Particularly, *Mycoplasma hyopneumoniae* (Mh) infection causes the disease Porcine Enzootic Pneumonia, a highly contagious and chronic disease affecting pigs and a significant reduction in the growing performance of pigs with estimated economic losses of \$0.2-1 billion per year in the United States (Clark et al. 1991). Mh is spread worldwide and present in almost pig herd (Minion 2002). Antibiotics are less effect in disease treatment because they do not completely remove the infection. Vaccines have been found to reduce the severity of the disease but do not prevent the disease from occurring in infected pigs (Haesebrouck et al. 2004). Sadly, the immune response in presence of Mh in pigs is slow and ineffective (Minion 2002). In the case of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) infection, during outbreak periods the farrowing rate declined by 10.92 points. The number of pigs weaned per litter declined by 1.5 pigs per litter. Mortality level of nursery pigs increased by 10.65% points. Mortality level in grow-finish pigs increased by 6.05% points. Feed efficiency and average daily gain for nursery and grow-finish pigs were also impacted. Average daily gain declined by 0.21 pounds per day (25% decline) for the nursery pigs and 0.20 pounds per day (12% decline) for the grow-finish pigs...The total economic impact on production costs to pig producers in the United States is projected to be \$560.32 million annually (Kliebenstein 2004). Another question is what are better livestock management techniques to improve the health status in both animal and human? Immunologically and genetically improved health and natural disease resistance leads to higher product quality, reduced danger of transmission of infectious agents or traces of antibiotics and has positive impact on pro-

duction efficiency and animal welfare. Disease reduction by genetic means has certain advantages through cumulative and permanent effects. The use of powerful methodologies in molecular biology, biochemistry, and physiology in the last two decades has led to impressive progress in our understanding of the mechanisms of complement activation and its role in immune responses for preventing pathogen in human and some different animal species (Makrides 1998). Direct selection for disease incidence requires, besides a unique pig identification and disease registration system, challenge routines that are inconvenient in intensive pig production. Indirect selection for the expression of immune capacity may be an alternative but requires detailed knowledge of the different components of the immune system. Complement is an important humoral innate immune defence system against invading microorganisms. The activation of the complement cascade allows the direct killing of microbes, the disposal of immune complexes, and the regulation of other immune processes. In pig, the complement components C7 and C8A have been completely characterized at the molecular level (Agah et al. 2000, Nakajima et al. 1998). Agah et al (2000) could show that addition of purified porcine C7 restored the hemolytic activity of C7-depleted human sera in a dose-dependent manner. Genetic polymorphism of the sixth complement component was also detected in several different breeds of pig (Shibata et al. 1993). Furthermore, the porcine C8A gene was physically assigned to chromosome 6q3.3-q3.5 by in situ hybridization using porcine bacterial artificial chromosome (BAC) clone as a hybridization probe (Nakajima et al. 1998) and the porcine C9 gene has been mapped to chromosome 16q1.4 by fluorescence in situ hybridization (Thomsen et al. 1998).

In order to gain more knowledge about regulation of pig complement system as well as to contribute to the improvement of farm animal health status, an investigation on the complete cDNA molecular structure of the porcine terminal complement components C6, C7, C8 and C9 genes was conducted to detect single nucleotide polymorphisms within the candidate genes and to study their association with hemolytic complement activity in the classical and alternative pathway in the F<sub>2</sub> DUMI resource population and to assign them to porcine chromosomes by radiation hybrid and linkage mapping.

## 2 Literature review

### 2.1 Innate immunity and adaptive immune response

The two primary immune responses to microorganisms and their antigens are generated by the innate and acquired or adaptive immune systems. The innate immune system is considered the first line of defense in protecting the host from invading organisms (Thacker 2003). This means that the cells of the innate system recognize, and respond to, pathogens in a generic way in a non-specific manner and not confer long-lasting or protective immunity to the host. Only if the organism escapes and survives the innate immune response an adaptive or acquired immune response occur, resulting in the generation of a specific immune response that will target that particular organism or antigen. The response increases with each exposure to the antigen (Thacker 2003).

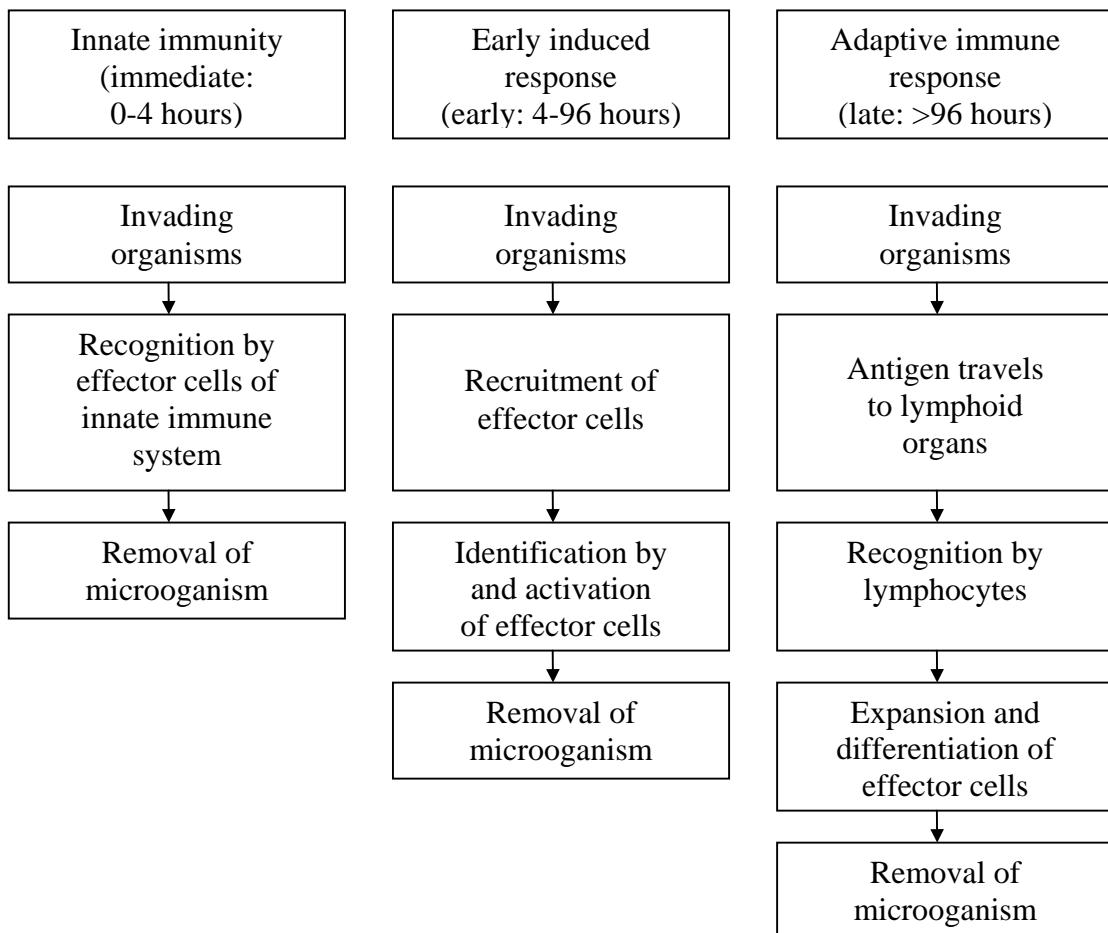


Figure 1 Response to invading microorganisms by innate and adaptive immune systems (Thacker 2003)

The major component of the natural immunity consists of the natural killer (NK) cells, phagocytes, dendritic cells, and the complement system (Heyworth 2001, Medzhitov 2001). Lymphocytes (B and T cells) superimposed on innate immunity to improve host defense against microbes are the primary cell type of the adaptive immune response (Thacker 2003). The adaptive immune response is distinguished from innate immune mechanisms by a higher degree of specific reactivity for the including agent and recall memory (Doenhoff 2000). Its function is mediated by specific antibody or humoral immunity and a specific cellular immune response or cell mediated immunity (Corbeil 1991). This study only refers to the activity of the complement system.

## 2.2 Complement system

The complement system, an effector mechanisms of humoral immunity as well as an important mediator of the innate immune system (Thacker 2003), is a highly regulated and complex set of interacting proteins in blood plasma and on the cell surfaces that can recognize, bind to and kill or remove invading microbes. Normally, these proteins are inactive until they are cleaved by a protease, which, in turn, converts them into another protease. On the other hand, some components of the system serve as the substrate of a prior component and then as an enzyme to activate a subsequent component. There are three pathways of complement activation classical pathway, lectin pathway and alternative pathway leading to the formation of the membrane attack complex (MAC). Assembly of the MAC on target cells results in the formation of transmembrane pores that can lead to cell killing (Müller-Eberhard 1986).

### 2.2.1 The classical pathway

The classical pathway of complement activation is a sequentially acting, multistep cascade in which the complement component C1, C4, C2, and C3 are involved (Loos et al. 1998). The component C1 circulates in blood serum as a molecular complex containing 6 molecules of C1q, 2 molecules of C1r and 2 molecules of C1s. The activation of the complement system, through the classical pathway is initiated by interaction of the C1q component with the Fc regions of antigen-antibody complexes (Wimmers et al. 2003, Reid et al. 1981). Once the globular ‘heads’ of subcomponent C1q are bound by anti-

body, the 'stalks' bind and activate C1r and C1s to generate an enzyme (Haeney, 1998) that cleaves two serum proteins (1) C4 to release C4a and C4b, which binds covalently to sugar residues on cell-surface glycoproteins and (2) C2 to release C2b and the glycoprotein C2a, which binds noncovalently to a site on C4b to give the C4b2a complex, 'the C3 convertase' (Loos et al. 1998). The C3-convertase enzyme catalyzes the cleavage of C3 into C3a and C3b, which binds covalently to glycoproteins scattered across the cell surface. Macrophages and neutrophils have receptors for C3b and can bind the C3b-coated cell or particle preparatory to phagocytosis. This effect qualifies C3b as an opsonin. C3a is a small fragment, which can bind to receptors on basophils and mast cells triggering them to release their vasoactive contents (e.g., histamine). Because of the role of these materials in anaphylaxis, C3a is called an anaphylatoxin. Some of the C3b binds C4b2a to form the C4b2a3b enzymatic molecule (C5 convertase of classical pathway) for initiation of the lytic pathway (Sunyer and Lambris 2001)

### 2.2.2 The lectin pathway

The lectin pathway is mainly driven by manose-binding lectin (MBL) (Thiel et al. 1997), one of the most important constituents of the innate immune system (Fraser et al 1998). Activation of the lectin pathway is similar to the classical pathway, with binding of the MBL serum protein, which is homologous to C1q of classical pathway, to manose-containing carbohydrates on bacteria or viruses that activates the complement cascade (Thacker 2003). The binding results in the association of two serine proteases, MASP1 and MASP2 (MBL-associated serine proteases), which are respectively homologous with C1r and C1s in the classical pathway, to form a complex molecule in the presence of a calcium. The MASP2 enzyme subsequently cleaves C4 into C4a and C4b as well as C2 into C2a and C2b to form the C4b2a complex playing a role as a C3 convertase in classical pathway activation (Fujita 2002). The C3 convertase then cleaves C3 into C3a and C3b, which binds to the membrane in association with C4b and C2a to form C4b2a3b (C5 convertase of lectin pathway) for lytic pathway activation. By contrast, MASP1 is able to cleave C3 directly which results in activation of the alternative pathway (Fujita 2002).



### 2.2.3 The alternative pathway

The alternative pathway represents an important part of the innate immunological system containing plasma proteins C3, factor B (a single-polypeptide chain glycoprotein), factor P (properdin) and factor D (adipsin). The alternative complement pathway is directly activated via binding of spontaneously activated C3b to the surface of a pathogen (Thacker 2003, Reid et al. 1981). It is initiated by the spontaneous hydrolysis of C3, which is abundant in the plasma activated by a H<sub>2</sub>O molecule and reacting with factor B in the presence of Mg<sup>2+</sup> ions to form the C3(H<sub>2</sub>O)B molecule. Factor D then splits the bound Factor B into Bb and Ba, to form the C3(H<sub>2</sub>O)Bb enzymatic molecule (Haeney 1998). The C3(H<sub>2</sub>O)Bb enzyme is at low level in plasma but it can directly cleave the other C3 molecules to release C3a from the C3b molecule which reacts repeatedly with factor B and D to form the C3bBb enzymatic molecule. The C3bBb binds a serum protein called properdin to form the C3bBbP stable enzymatic molecule complex (C3 convertase), which in return can split molecules of C3 into C3a and C3b. Factor P binding to C3bBb extends the half-life of the complex (Fearon et al. 1975). The C3b subsequently binds with C3bBb to form a C3bBb3b molecule (C5 convertase of alternative pathway), which initiates the lytic pathway (Parish 2001, Zipfel 2001).

### 2.2.4 Membrane attack complex (MAC)

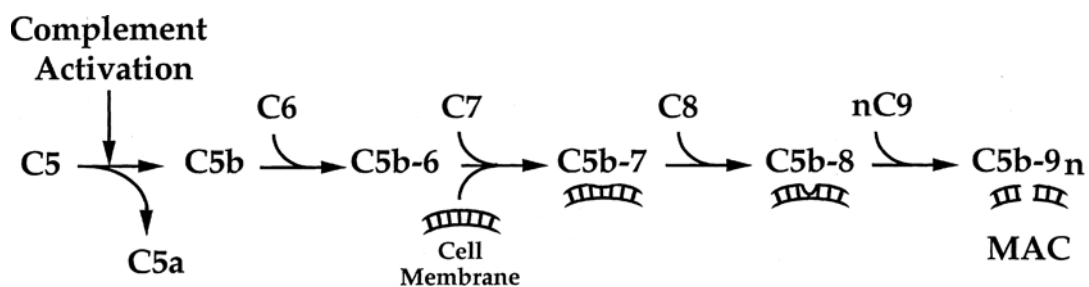


Figure 2 Assembly of the MAC. Activation of complement leads to the formation of C5b-9, the cytolytic membrane attack complex of complement, or MAC (Schreck et al. 2000)

Activation of the three pathways of the complement system generates the C5 convertase as a source of inspiration in formation of the membrane attack complex (MAC) on tar-

get-cell membranes. This complex is a multiprotein set assembled by sequential addition of complement components C5b, C6, C7, C8, and C9 (DiScipio et al. 1984, Podack et al. 1982). In human the formation of MAC begins with specific cleavage of component C5 between residues 74 and 75 into C5a and C5b by C5 convertases: C4b2a3b from classical pathway, C4b2a3b from lectin pathway or C3bBb3b from alternative pathway (Hofsteenge et al. 1999, Makrides 1998, Müller-Eberhard 1988). The C5a is a potent anaphylatoxin and mediates various inflammatory events while the ability of C5b to bind C6 decays rapidly in a C5b-6 metastable bimolecular complex (Wuerzner 2003, DiScipio 1992, 1988, Podack et al. 1979, 1976). Then C7 associates mainly with the C6 subunit of the C5b-6 complex (DiScipio et al. 1988, Müller-Eberhard 1988) forming C5b-7. Upon binding of C7 to C5b-6, the trimolecular C5b-7 complex undergoes a conformational transition and allows insertion of this complex into lipid membranes (DiScipio 1992, 1988, Podack et al. 1979, 1976). Both C5b-6 and C7 are soluble proteins, whereas the C5b-7 complex is amphiphilic and capable of tightly binding phospholipid membranes (DiScipio et al. 1988). Then C8B binds to C5b-7 complex by a specific recognition site on C5b (Stewart et al. 1987). Self-polymerization of C9 is initiated by binding C5b-8 to C9 via C8A to form C5b-C9 complex of the terminal complement components (Stewart et al. 1987, Podack et al. 1982) and constitutes a supramolecular organization (Tschopp et al. 1986). The C5b-9 complex can contain one molecule of each of component C5b, C6, C7, C8 and 6-18 molecules of C9 that polymerize into barrel-like structures and can vary dramatically in size (Esser 1994, Tschopp 1984). Of these, components C6 to C9 are related plasma proteins, which differ in size and complexity (Hofsteenge et al. 1999, Müller-Eberhard 1986). Insertion of further C9 molecules into the target membrane causes local distortion of the phospholipid bilayer permeability resulting in 'leaky patches' (Esser 1991) or forming a hydrophilic channel through the membrane (Bhakdi et al. 1991). It ultimately results in cellular activation, cell lysis or cell death (Morgan 1989, Esser 1987, DiScipio et al. 1984). However, deposition of small amounts of the MAC on cell membranes of nucleated cells may mediate a range of cellular processes without causing cell death (Nicholson-Weller et al. 1993, Morgan 1992). Gram-positive organisms are already protected against complement lysis by the presence of their peptidoglycan cell walls. MAC inhibition may not be the sole function of streptococcal inhibitor of complement (Fernie-King et al. 2001). The cytolytic activities of the membrane attack complex (MAC) are inhibited through the bind-

ing of CD59 antigen to C8 and C9 (Meri et al. 1990). The MAC has an approximate  $M_r$  of  $1.7 \times 10^6$  (Tschopp et al. 1986)

### 2.3 Molecular characterization of candidate genes

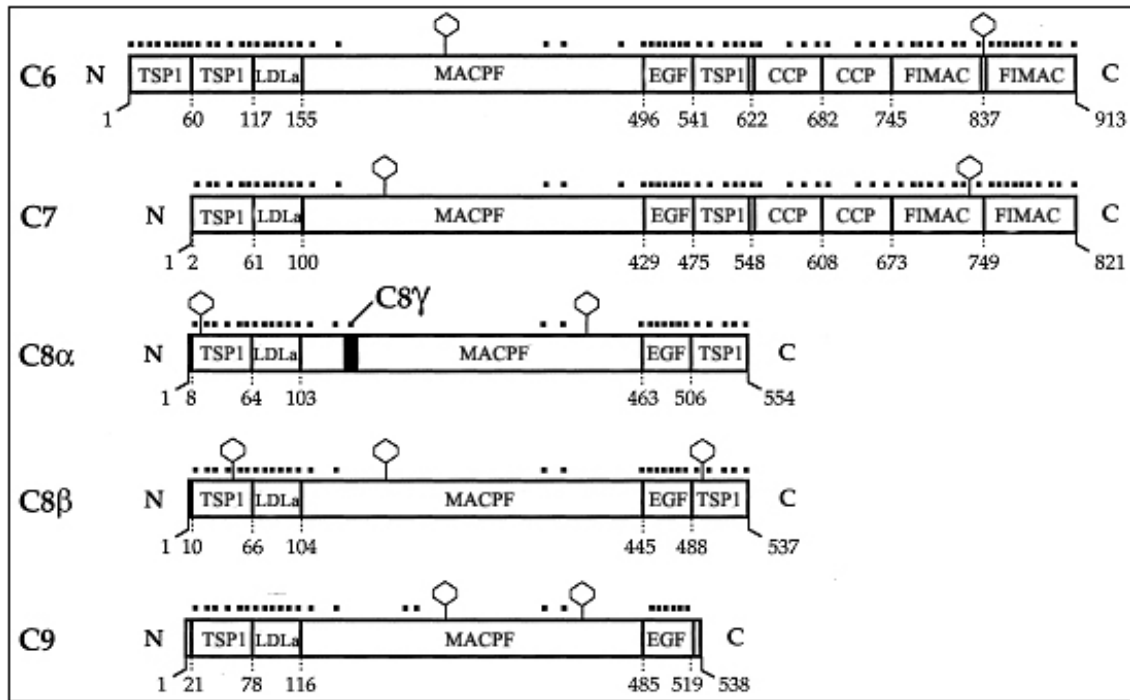


Figure 3 Structural organization of the MAC protein family. Shown are maps based on sequences of the mature human proteins and the module boundaries listed in the SWISS-PROT Protein Sequence Data Bank. Abbreviations correspond to thrombospondin type I (TSP1), low-density lipoprotein receptor class A (LDLa), epidermal growth factor (EGF), complement control protein (CCP) and Factor I membrane attack complex (FIMAC) modules. The membrane attack complex/perforin segment is designated MACPF. Numbers correspond to the first residue in each module. Dots above each map indicate the approximate location of Cys residues, which are highly conserved. All are involved in intrachain disulfide bonds except Cys<sup>164</sup> in C8A, which is linked to C8G. The shaded region fanking Cys<sup>164</sup> corresponds to a 17-residue insertion that is unique to C8A. Hexagonal symbols designate Asn residues that are potential N-glycosylation sites (Schreck et al. 2000)

The four terminal components of the complement system (C6, C7, C8 and C9) are glycoproteins (Witzel-Schloemp et al. 1998, Dewald et al. 1996, Hobart et al. 1995, Chakravarti et al. 1988, Müller-Eberhard et al. 1986) and C-mannosylated on multiple tryptophan residues. From the 113 Trp residues in the complete membrane attack complex, 50 were found to undergo C-mannosylation (Hofsteenge et al. 1999).

### 2.3.1 The sixth complement component (C6)

Complement C6 is one of five plasma proteins that are incorporated into the lytic terminal complement complex on lipid membranes upon activation of the complement cascade (Haefliger et al. 1989). Human C6 is a single-chain glycoprotein with an estimated  $M_r$  of 95-128 kDa and a carbohydrate content of 3.8-11.3% (DiScipio et al. 1982, Kolb et al. 1982, Podack et al. 1979, 1976). C6 is larger than C7, C8A, C8B, and C9 (Chakravarti et al. 1989). The structure of human C6 sequence is encoded with 18 exons (Hobart et al. 1993). The complete polypeptide structure of mature C6 deduced from the cDNA sequence consists of 913 amino acid residues (DiScipio et al. 1989, Haefliger et al. 1989) preceded by a typical 21-residue signal peptide. The protein is homologous with the other terminal components of complement, C7-C9 (DiScipio et al. 1989) and most similar in structure to complement C7, sharing 33.5% identical residues with C7 including all 56 cysteine residues (Haefliger et al. 1989). It is noteworthy that cysteine is clearly the most conserved of all the amino acids because 55 of the 56 cysteines found in C7 match those in C6 (DiScipio et al. 1989). Human C6 contains a large number of cysteine residues, 64 in total in a sequence of two domains, an amino-terminal region, C6a (a.a 22-631) that is related to C8 and C9, and a carboxyl-terminal region, C6b (a.a 632-934) that has partial homology to the complement regulatory proteins factor H and factor I (DiScipio et al. 1989, Haefliger et al. 1989). The absence of any free sulfhydryl groups in the C6 protein implies that all of the 64 cysteines are linked in 32 disulfide bonds. Disulfide linkages may exert a stabilizing influence on C6 structure because the activity of this molecule is partially resistant to treatment with chaotrophs (DiScipio et al. 1989).

The amino-terminal fragment of C6 has 30, 26, 22, and 18% residues identical with those of C7, C8A, C8B, and C9, respectively, while the carboxyl-terminal polypeptide

has 27% of its amino acids identical to those from the corresponding region of C7. The sequence of 490 first amino-terminal half of the C6 protein has 47-52% similarity with C7, C8A, C8B, and C9, as well as 31-38% similarity with thrombospondin (TSP1), epidermal growth factor (EGF), and low density lipoprotein receptor (LDL<sub>r</sub>). It is therefore a member of the channel-forming complement proteins (Chakravarti et al. 1989). Most of the predicted C6 alpha-helical structure resides in a portion of the polypeptide chain that is free of cysteine and which shares homology with C9 and perforin (DiScipio et al. 1989). A site of asparaginyl glycosylation of C6 contained in the sequence Asn-Phe-Thr was located at position 324 (Chakravarti et al. 1989). C6 deficiency can not form the membrane attack complex (MAC) (Mead et al. 2002)

### 2.3.2 The seventh complement component (C7)

Complement C7 is one of the components of membrane attack complex (MAC) generated by the terminal complement cascade. The derived human C7 protein sequence consists of 821 amino acids, and in addition, the precursor protein contains a 22-residue hydrophobic leader peptide. The C7 is highly gene homologous to C6, C8A, C8B, and C9 (Hobart et al. 1995, DiScipio et al. 1988). The human seventh component of complement is a single chain plasma glycoprotein of  $M_r$  97.3 kDa, encoded by 18 exons whose sizes vary from 56 to 244 bp (Hobart et al. 1995, DiScipio et al. 1988). The human C7 single polypeptide chain is glycosylated at two sites and has 56 cysteines presumed to form 28 disulfide bonds. Disulfide bonds are important for stabilizing the structure of this protein. Virtually all the cysteines are found in small units of 35-77 amino acids that exhibit homology with those of various proteins including the low-density lipoprotein receptor, epidermal growth factor precursor and thrombospondin. The carboxyl-terminal segment of C7 is between amino acids 575-843 (DiScipio et al. 1988). Human C7 plays a major role in bringing about the hydrophilic- amphiphilic transition during the formation of the membrane attack complex, and it serves as a membrane anchor for the C5b-7 complex (DiScipio et al. 1988).

The porcine C7 was a single glycoprotein with an approximate molecular mass of 90-100 kDa. Similar to human C7, the predicted amino acid sequence of precursor porcine C7 also consists of 843 amino acid residues. Human (GenBank acc. no. NM 000587)

and porcine C7 share an overall identity of 80% at the protein level. The two potential asparaginyl-linked glycosylation sites (Asn-X-Ser/Thr) at positions 202 and 754 in human C7 are also conserved in porcine C7. Both human and porcine C7 have 56 cysteine residues, the majority of which are located either at the amino or carboxyl terminus of the protein, giving rise to cysteine-rich domains (Agah et al. 2000). Although the function of these cysteine-rich motifs remains unclear, one speculation is that they play a role in the stabilization of  $\beta$ -turns that are presumably involved in protein-protein interactions (DiScipio et al. 1988).

Although liver has been demonstrated to be the major site for the synthesis of C6 and C8, the major site of C7 synthesis has not yet been elucidated. Human cDNA was obtained from liver, but C7 is not detected in the human hepatoma-derived cell line HepG2, thus suggesting that human hepatocytes do not primarily synthesize C7 (DiScipio et al. 1998, Wuerzner et al. 1994). Interestingly, circulating C7 initially switches to 50% of the transplanted allotype following allogeneic liver transplantation, but completely reverts to the original allotype by 6 weeks (Wuerzner et al. 1994). These data suggest that mature mononuclear white blood cells (i.e. Kupffer cells in the liver) may be a major source of C7 production (Agah et al. 2000)

### 2.3.3 The eight complement component (C8)

Human C8 oligomeric protein, is one of five components of the membrane attack complex of complement (MAC) which consists of three different polypeptides C8A (alpha chain,  $M_r = 64$  kDa), C8B (beta chain,  $M_r = 64$  kDa) and C8G (gamma chain,  $M_r = 22$  kDa) (Plumb et al. 1999, Steckel et al. 1980). It is arranged asymmetrically as a disulfide-linked C8alpha-gamma heterodimer that is noncovalently associated with beta chain (Haefliger et al. 1987, Howard et al. 1987, Rao et al. 1987).

The entire human C8A cDNA length was determined to be 2443 base pairs which, when translated, yields an apparent leader sequence of 30 amino acids. This sequence contains an apparent initiation methionine, signal peptide, and propeptide which ends with an arginine-rich sequence that is characteristic of proteolytic processing sites found in the pro form of protein precursor (Rao et al. 1987). Cysteine-rich repeat sequence found

in low-density lipoprotein receptor and to the cysteine-rich epidermal growth factor type sequence found in a number of proteins. There are 1 or 2 asparagine-linked in C8A gene (Rao et al. 1987). Both proteins C8A and C9 contain nearly the same number of amino acids and cysteine-rich amino- and carboxyl-terminal domains (Rao et al. 1987). The complete amino acid sequence of pig C8A was determined by characterizing the full-length cDNA clone isolated from a porcine liver cDNA library. The 2145 bp sequence (GenBank acc. no. AB008156) of this clone have a 558 amino acid open reading frame and a 31 amino acid leader sequence. Porcine C8A was found to be similar to human and rabbit C8A in length, leader sequence, conserved cysteine residues (29 cysteines), cysteine-rich modules (TSP1, LDLa, EGF) (Nakajima et al. 1998). The location of these modules is very similar among pig, rabbit, and human. Differences in the amino acid sequence among the three species were detected in the proposed candidate site for CD59 recognition (amino acids 382-419) (Nakajima et al. 1998). During MAC formation, C8A mediates binding and self-polymerization of C9 to form a pore-like structure on the membrane of target cells (Slade et al. 2006, Scibek et al. 2002).

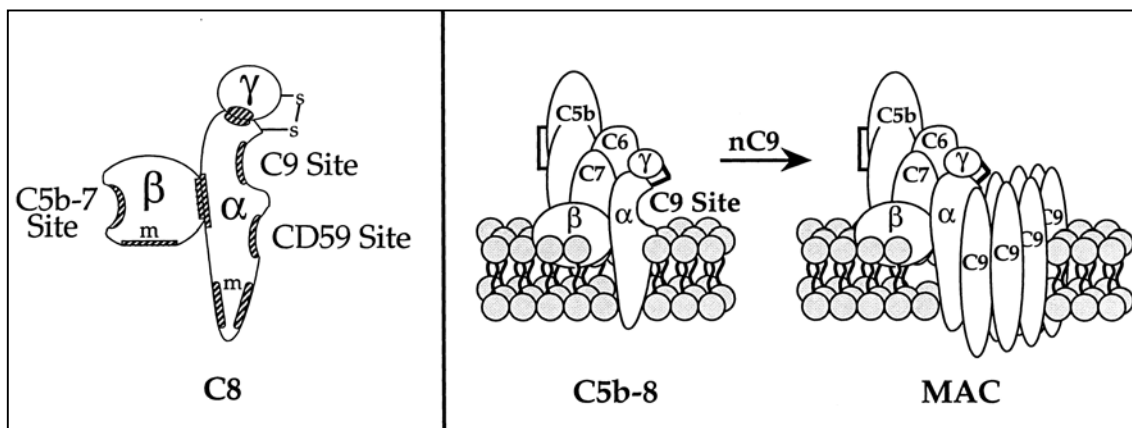


Figure 4 Functional sites in C8. (Left panel) Schematic representation of sites of interaction (hatched) in each subunit of C8. Sites are arbitrarily located within each subunit but are depicted as distinct and nonoverlapping because they function simultaneously within the MAC. Membrane interaction sites are designated 'm'. (Right panel) Relative location of the C8 subunits within membrane-bound C5b-8 and the MAC (Plumb et al. 1998). C9 is depicted as a polymer composed of an arbitrary number of monomeric units (Schreck et al. 2000).

Human C8B-cDNA-positive clones were partially sequenced to characterize the 12 exons of the gene with sizes from 69 to 347 bp (Kaufmann et al. 2004). In serum, the beta-subunit is non-covalently bound to the disulfide-linked alpha-gamma subunit (Kaufmann et al. 1993). Amino-linked carbohydrate on C8 beta is not necessary for association with C8 alpha-gamma or for C8 activity (Letson et al. 1996). Analysis of the derived amino acid sequence revealed several membrane surface seeking-segments that may facilitate C8B interaction with target membranes during complement-mediated cytotoxicity. Determination of the carbohydrate composition indicated 1 or 2 asparagine-linked but no O-linked oligosaccharide chains (Howard et al. 1987). Comparison with the amino acid sequence of C9 shows an overall homology with few deletions and insertions. In particular, the cysteine-rich domains and membrane-inserting regions of C9 are well conserved. These findings are discussed in relation to a possible mechanism of membrane attack complex formation (Haefliger et al. 1987). Analysis of the C8B sequence reveals several segments with the potential to form amphipathic C8B-structures that are capable of interacting with membrane surfaces. These segments may serve two functions. One is that they may assist in stabilizing interactions of C8 with C5b-7 by direct association with the membrane surface. Second, these segments may contribute directly to the lytic activity of C5b-8 and C5b-9 through membrane perturbation (Howard et al. 1987).

For human C8A and C8B, the overall homology is 33% on the basis of identity and 53% when conserved substitutions are allowed. For C8B and C9, the values are 26% and 47%, respectively (Howard et al. 1987). For C8A and C9, they are 24% and 46%, respectively (Rao et al. 1987). The human C8A and C8B are homologous to C6, C7, and C9 and form the MAC family of proteins (Plumb et al. 1998, Hobart et al. 1995). C8A and C8B have correspondingly similar roles in MAC-mediated lysis of erythrocytes and bacterial killing. C8G is not required for complement-mediated killing of Gram-negative bacteria (Parker et al. 2002).

The gamma chain whose functional role remains undefined is not related to any other complement protein but is a member of the lipocalins, a family of proteins that bind small hydrophobic ligands (Dewald et al. 1996). C8G is structural homology to protein



HC (alpha-1-microglobulin/bikunin precursor) (Hunt et al. 1987). Human C8 gamma is a 22 kDa subunit of complement component C8 (Schreck et al. 2000). C8G has one internal disulfide bridge between Cys<sup>76</sup> and Cys<sup>168</sup> within the carboxy-terminal 12 kDa fragment, whereas the remaining cysteine residue 40 forms the disulfide bridge with C8 alpha (Haefliger et al. 1987). C8G may act as a retinol transporting protein in plasma (Haefliger et al. 1991) and plays a role for binding an as yet unidentified ligand (Dewald et al. 1996, Kaufman et al. 1994, Ng et al. 1987).

#### 2.3.4 The ninth complement component (C9)

Human C9 is the most abundant component of the membrane attack complex of the complement system of immune defense. It is the last protein that binds to the assembling membrane attack complex (MAC) of complement, completing the sequence of events that leads to the destruction of target membranes. The single chain serum protein is encoded by a gene composed of 11 exons and located on chromosome 5p13 (Witzel-Schloemp et al. 2001). The nucleotide sequence of the ninth component of human complement coding 537 amino acids in a single polypeptide chain was determined (DiScipio et al. 1984).

There are the presence of surface-binding domains in the amino terminal half and channel-forming domains in the carboxyl terminal portion of native, monomeric C9 (Shiverg et al. 1986). Human monomeric native C9 ( $M_r = 66$  kDa) can be cleaved by  $\alpha$ -thrombin into two single-chain polypeptide fragments: C9a ( $M_r = 28$  kDa) and C9b ( $M_r = 38$  kDa), which are the amino and carboxyl-terminal segments of the protein, respectively (DiScipio et al. 1985). The carboxyl terminal fragment (C9b) enters lipid bilayer membranes and form single channels of defined but heterogeneous sizes and lyses erythrocytes in the absence of a fully assembled C5b-8 complex. The C9a fragment contains several regions of possible amphipathic  $\beta$ -structures. It is conceivable that such areas in C9a serve to bind the peptide sufficiently strongly to the membrane surface to cause leakiness (Shiverg et al. 1986). Human C9 water-soluble glycoprotein will insert quickly into lipid bilayers after binding to the membrane bound C5b-8 complex (Steckel et al. 1983, Hu et al. 1981, Podack et al. 1981). The principal binding site for C9 lies within the MACPF domain of C8alpha (Slade et al. 2006). The molecular weight calcu-

lated from the sequence for the nonglycosylated C9 protein is 60,700 (DiScipio et al. 1984).

#### 2.4 Functional protein domains in candidate genes

The membrane attack complex (MAC) is an assembly of terminal complement component C5b-9 in lytic pathway of complement system (NCBI homepage). The membrane attack is important for mammalian immune defense against invading microorganisms and infected host cells. Proteins of the complement MAC and the protein perforin share a common MACPF domain that is responsible for membrane insertion and pore formation (Hadders et al. 2007). C8A is the first complement protein to insert into the membrane. The size of the MACPF domain of C8A is consistent with known C9 pore sizes (Hadders et al. 2007). Therefore, the principal binding site for C9 lies within the MACPF domain of C8A (Slade et al. 2006)

Perforin is a protein found in cytolytic T-cell and killer cells. In the presence of calcium, perforin polymerises into transmembrane tubules and is capable of lysing, non-specifically, a variety of target cells. There are a number of regions of similarity in the sequences of complement components C6, C7, C8A, C8B, C9 and perforin (NCBI homepage).

The complement control protein (CCP) modules are defined by a consensus sequence within a stretch of about 60 amino acid residues. These modules have been identified more than 140 times in over 20 proteins, including 12 proteins of the complement system (Norman et al. 1991). Typically, two to four modules contribute to a binding site, implying that the orientation of the modules to each other is critical for function (NCBI homepage). The CCP modules mediate specific protein-protein and protein-carbohydrate interactions that are key to the biological function of the regulators of complement activation and, paradoxically, provide binding sites for numerous pathogens (O'Leary et al. 2004). The carboxyl-terminal domains of C6 and C7, which have cysteine-rich modules homologous to those found in factors H and I, have the capacity to link specifically with C5.

The factor I membrane attack complex (FIMAC) domain is found in the terminal complement component proteins (Figure 3), complement component factor I and agrin. Factor I is responsible for cleaving alpha chains of C4b and C3b in the presence of the cofactors, C4-binding protein and factor H, respectively. Agrin is a component of the basal lamina that causes the aggregation of acetylcholine receptors and acetylcholinesterase on the surface of muscle fibres of the neuromuscular junction (<http://www.ebi.ac.uk/interpro>). The FIMACs in C7 mediate reversible binding to C5 (Thai 2004). The C5b-binding domain of C6 was located in the 34 kDa carboxyl terminal fragment consisting of two short consensus repeats and two factor I modules (Haefliger et al. 1989). Although the FIMACs are not required absolutely for C6 activity, these modules promote interaction of C6 with C5 enabling a more efficient bimolecular coupling ultimately leading to the formation of the C5b-6 complex (DiScipio and Berlin 1999).

Both TSP1 and TSP2 are induced in response to injury (Bornstein 2001). Antisense treatment reduced the number of TSP1-positive macrophages by 50% or more (DiPietro et al. 1996).

In human C6, most of the cysteines are located in short (34-77 amino acids) discrete segments that exhibit homology with a wide variety of other proteins such as TSP1, LDLa at the carboxyl terminus (DiScipio et al. 1989, Haefliger et al. 1989), EGF, and CCP and FIMAC (DiScipio et al. 1989). In contrast to C7 and other terminal complement proteins, the TSP1 module at the amino terminus occurs as a tandem repeat in C6. The last tandem repeat at the carboxyl terminus of C6 and C7 has been identified as a new distinct module (FIMAC module), which is closely related to a segment in the heavy chain of complement control factor I (Chakravarti et al. 1989).

Both human C8A and C8B subunits contain a pair of amino-terminal modules (TSP1 and LDLa), a pair of C-terminal modules (EGF+TSP1) and an extended central segment referred to as the membrane attack/perforin (MACPF) domain (Musingarimi et al. 2002, Plumb et al. 1998, Hobart et al. 1995). Binding to C8A-G is dependent on the TSP1 + LDLa + MACPF segment of C8B. Within this segment, the TSP1 module and MACPF domain are principally involved and act cooperatively to mediate binding. Residues

within this segment also mediate binding and incorporation of C8 into the MAC (Musingarimi et al. 2002).

Human C9 is a typical mosaic protein with thrombospondin (TSP1) and low density lipoprotein receptor (LDLa) domains at its amino-terminus and an epidermal growth factor-like (EGF) domain at its carboxyl-terminus (Smith et al. 1992).

## 2.5 Polymorphisms and deficiency of candidate complement components

### 2.5.1 The sixth complement component (C6)

Genetic polymorphisms of the complement component C6 have been reported in human (Ferne et al. 1995, Wuerzner et al. 1995, Dewald et al. 1993, Hobart et al. 1975), particularly, in a Japanese population (Soejuma et al. 2005, Nishimukai et al. 2003, 1986, Nakamura et al. 1984, Tokunaga et al. 1983), in 203 Chinese Han people in Liaoning Province of northeast China (Tsunenari et al. 1992), among Russian rural inhabitants of the Tomsk region (Kucher et al. 1993), in a group of Africans (Ferne et al. 1996), in the Senegal population (Corbo et al. 1994), in Korean living in Seoul (Park et al. 1988), in individuals of African descent living in the United States or Europe (Parham et al. 2007), in 167 unrelated Norwegian Lapps (Olving et al. 1980), in the German population (Schwarzenholz et al. 2000, Kühnl et al. 1980, Kunstmann et al. 1980), in a African ethnic group (Ferne et al. 1996). Genetic variation was not only found in exon but also in intron fragments of human C6 (Parham et al. 2007, Nishimukai et al. 2003, Ferne et al. 1997, 1995, Wuerzner et al. 1995). In other species, polymorphisms were also detected, in the rhesus monkey (Hall 1977), in the chimpanzee (Raum et al. 1980), in the rabbit (Goldman et al. 1982, Kunstmann et al. 1980), in the common marmoset (Whitehouse 1984), in the dog (Shibata et al. 1995, Anderson et al. 1983, Eldridge et al. 1983), in the Manx shearwater (Whitehouse 1982), in the rat (Granados et al. 1884), in the mouse (Hayakawa et al. 1984, Orren et al. 1989), in the pig (Shibata et al. 1993). The porcine C6 locus was found highly polymorphic (Shitaba et al. 1993).

Deficiency of the sixth component of complement is frequently associated with recurrent neisserial infections, especially meningitis caused by *Neisseria meningitidis* (Par-

ham et al. 2007, Nishizaka et al. 1996) or candidiasis and toxoplasmosis (Morgan et al. 1989). Various kinds of polymorphisms were found in different species: *Neisseria* infections in human can be caused by a 'silent' alleles at the C6 loci (Alvarez et al. 1995) or a compound heterozygote for two C6 gene mutations resulting in premature termination codons and C6 null alleles in a 16 year-old African-American male (Zhu et al. 1998) or three frameshift mutations in African-Americans and in two Dutch C6-deficient kindreds (Hobart et al. 1998) or a homozygous single-base deletion in one African American and a heterozygous single base deletion in one Japanese resulting in frame shifts and premature termination of the C6 polypeptide (Nishizaka et al. 1996). C6 deficiency caused by unstable mRNA or a point mutation in the C6 gene resulting in an aberrant transcription of the C6 gene was also studied in the rat (Van Dixhoorn et al. 1997). In addition, point mutations observed in the Peru-Coppock mouse strain C6 gene may result in C6 deficiency (Bhole et al. 2004). C6 deficiency due to a blank allele for C6 was detected in the chimpanzee (Raum et al. 1980) and C6 deficiency involved in atherosclerotic lesion progression in rabbit (Schmiedt et al. 1998). Furthermore, C6 deficiency in two of 241 individuals in six breeds of pig (Landrace, Large White, Duroc, Berkshire, Meishan, Jinhua) and cross breeds tested was found, which suggested the presence of a null allele in pig populations. Marked breed differences among the gene frequencies and heterozygosities at C6 locus were observed (Shitaba et al. 1993).

### 2.5.2 The seventh complement component (C7)

Genetic polymorphisms of the seventh component of complement (C7) were found in human (Ferne et al. 1999, Horiuchi et al. 1999, Fernie et al. 1997, Fernie et a. 1995), especially in Caucasian population (Dewald 1988), in Asian populations (Horiuchi et al. 2002), in a Japanese population (Nishimukai et al. 2003, Wuerzner et al. 1991, Komatsu et al. 1989, Nishimukai et al. 1986, Washio et al. 1986, Nakamura et al. 1984), among Russian rural inhabitants of the Tomsk region (Kucher et al. 1993), in group of Africans (Ferne et al. 1997, 1996), in two Korean families (Kang et al. 2006), in Korean living in Seoul (Park et al. 1988), in 203 Chinese Han population in Liaoning Province of northeast China (Tsunenari et al. 1992), in Senegal population (Corbo et al. 1994), in the Cayapa Indians living in north-western Ecuador (Soelder et al. 1996). Different genetic mutations have also been described in intron segments of human complement

component C7 (Nishimukai et al. 2003, Fernie et al. 1999, 1996, 1995, Nishizaka et al. 1996), in the domestic dog (Eldridge et al. 1983), in the common marmoset (Whitehouse 1984).

The complement C7 deficiency was found in human diseases such as candidiasis and toxoplasmosis (Morgan et al. 1989). In the case of meningococcal meningitis a number of C7 polymorphisms were found including a large deletion mutation in exon in an Korean 11 year-old girl (Ki et al. 2005) or a 11-base pair deletion of nucleotides leading to the generation of a downstream stop codon causing the premature truncation of the C7 protein product in a Bolivian and Czech origin patient and a two-base pair deletion of nucleotides leading again to the generation of a downstream stop codon that provokes the truncation of the C7 protein in Spanish family (Barroso et al. 2006). *Neisseria* recurrent infections due to single base deletion of nucleotide and a missense mutation resulting in a change of amino acid were found in individuals of Moroccan Sephardic Jewish ancestry, a Spanish family (Barroso et al. 2004) as well as among a highly inbred Arab population living in the lower Galilee region of Israel (Behar et al. 2002) or a two base-pair deletion of nucleotides leading to the generation of a downstream stop codon causing the truncation of the C7 protein product in another Spanish family (Barroso et al. 2004) or appearance of a silent allele in three C7-deficient families (Alvarez et al. 1995) or a homozygous G→T transversion leading to a stop codon in a Spanish family (Horiuchi et al. 1999) or a homozygous A→T transversion leading to a stop codon TGA and a homozygous 2-bp deletion caused a frameshift, generating a premature termination codon 4-6 nucleotides downstream in two unrelated Japanese males (Nishizaka et al. 1996). Other case reports describe ankylosing spondylitis conducting on the serum of a 44 year-old woman of French-Canadian descent (Delâge et al. 1977), systemic lupus erythematosus due to a single base mutation leading to a stop codon that causes the premature truncation of the C7 protein in Spanish family additionally a A→C transversion resulting in an amino acid change (Barroso et al. 2006).

### 2.5.3 The alpha and gamma chain of the eighth complement component (C8A-G)

The alpha-gamma chain extensive genetic polymorphisms were found in human (Zhang et al. 1995, Nakamura et al. 1986, Rogde et al. 1985, Rittner et al. 1984), in 203 Chinese

Han population in Liaoning Province of Northeast China (Tsunenari et al. 1992). Particularly, Dewald et al. (1996) described the first known polymorphisms in the human C8G gene, namely a polymorphic site in exon and two more ones in intron segments. In pig, an allele variation G→A (amino acid substitution Arg→His, codon 485CGC→CAC) was found in several different breeds Landrace, Large White, Duroc, Berkshire, Jinhua, Crown Miniature Pig, wild boar, and Meishan. (Nakajima et al. 1998).

Two kinds of inherited C8 deficiency have been reported in man. Type I, in which no C8 antigen is detected, was thought to represent deficiency of the whole molecule, whereas in type II, antigenically deficient C8, which apparently lacks only the beta chain, is found ([http://www.genome.ad.jp/dbget-bin/www\\_bget?omim+120950](http://www.genome.ad.jp/dbget-bin/www_bget?omim+120950)). Association between the hereditary deficiency of the eighth component of complement and inflammatory disease like systemic lupus erythematosus disease in human (Jasin 1977) was reported. In addition, C8 deficiency in the serum of a 23-yr-old woman resulting in gonococcal infection syndrome (Petersen et al. 1976), inherited deficiency of C8 in a patient with recurrent meningococcal infections (Densen et al. 1983) or in a family with xeroderma pigmentosum (Giraldo et al. 1977) was identified. Genetic deficiency of the alpha-gamma-subunit of the eighth complement component was found in a strain of the New Zealand White rabbit leading to the following physiologic characteristic changes such as the body weight at the first week of life, mature weight, litter size, survival rates (Komatsu et al. 1985).

#### 2.5.4 The beta chain of the eighth complement component (C8B)

Detection of C8B different genetic mutations was described in human (Dewald et al. 1994, Kaufmann et al. 1993, Herrmann et al. 1989, Rogde et al. 1985, Alper et al. 1983).

The mutated allele was also found in intron of the C8B gene (Barba et al. 1994). C8 beta deficiency is usually detected in individuals who survive meningococcal disease (Wulffraat et al. 1994). Association between C8B deficiency and juvenile chronic arthritis in a 13-year-old boy (Wulffraat et al. 1994) or recurrent neisserial infections re-

sulting an allele substitution leading to a codon stop in C8B human sequence (Kaufmann et al. 1993) was reported. The common C→T mutation in the C8B genes is the genetic basis of C8 beta-chain deficiency in two members of this Bosnian family (Kotnik et al. 1997). Causes for C8B deficiency can be due to presence of mutated null alleles in a healthy Italian population (Bellavia et al. 1996), in 34 unrelated families from the United States and the former Soviet Union (Saucedo et al. 1995) or a nucleotide exchange creating a stop codon (Rao et al. 2004, Barba et al. 1994, Kaufmann et al. 1993) or not a major allele deletion in Norwegian meningococcal disease patients (Rogde et al. 1990).

#### 2.5.5 The ninth complement component (C9)

Different single nucleotide polymorphisms have been reported in human (Witzel-Schloemp et al. 2001, Coto et al. 1991, Rogne et al. 1991, Coto et al. 1990). Deficiency of the ninth component of human complement is the most common complement deficiency in Japan but is rare in other countries (Kira et al. 1998). C9 deficiency can lead to meningococcal meningitis due to creating a point mutation leading to termination codon in four Japanese C9-deficient patients (Kira et al. 1998), dermatomyositis due to the low titre of serum hemolytic complement (CH50) and a non-sense mutation in a 28 year-old Japanese woman (Ichikawa et al. 2001), recurrent neisseria infections associating two different point mutations, both generating TGA stop codons in the coding sequence in Swiss family with inherited C9 deficiency (Witzel-Schloemp et al. 1997) or a 'silent' allele (Alvarez et al. 1995) and the long-surviving mismatched kidney allografts in an Caucasian group due to absence of C9 (Hobart et al. 1997). Absence of C9 can also happen in a person with good health (Inai et al. 1979) or in serum of a 47 year-old woman with paroxysmal nocturnal haemoglobinuria.

Human C9 contains a motif (48WSEWS52) common to a family of cytokine receptors that is similar to a tryptophan-rich motif (WEWWR) of the membrane pore formers, thiol-activated cytolytins. Mutation of this motif in C9 resulted in polymerized protein, consistent with this site keeping the N-terminus in a protected conformation and preventing premature self-polymerization (Taylor et al. 1997).



## 2.6 Position of candidate genes on chromosome

RH panels consist of rodent cell lines that contain fragments of porcine foreign DNA, produced by X-Ray breakage. DNA fragment size is inversely proportional to the radiation dose and breakage is largely sequence independent as is the retention frequency of fragments (Alexander et al. 1999). Since the 118 clones/7000Rad INRA-Minnesota porcine radiation hybrid (IMpRH) panel has been constructed by Yerle et al (1998) it rapidly becomes the wonderful physical mapping tool for assigning candidate markers on chromosomes and for linking cytogenetic maps further. Considering the porcine map, it appears necessary to have a more accurate tool to determine the position and order of genes and markers. The IMpRH method based on the presence or absence of specific porcine/hamster DNA sequences in RHs is mainly determined by PCR. Order and distance between markers is calculated using a statistical algorithm that estimates the frequency of breakage between markers.

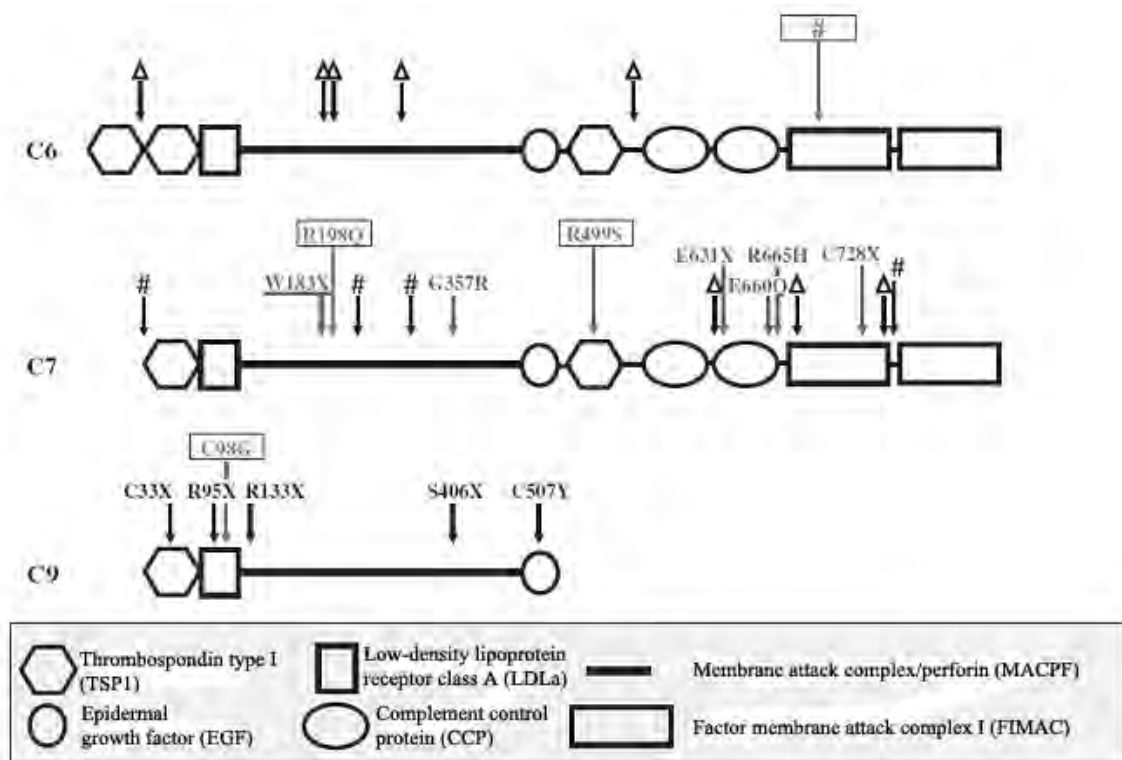


Figure 5 Locations of published human C6, C7 or C9 mutations leading to deficiencies in relation to the modules (indicated in the box). Boundary mutations, generating nonsense intron sequence till a stop codon is reached (#), and amino acid substitutions (A111B) are indicated. In addition, mutations leading to subtotal deficiency are boxed (Wuerzner 2003)

The assignment for human complement component C9 to chromosome 5 was confirmed by in situ hybridization to human metaphase chromosomes, giving a regional localization of 5p13 (Abbott et al. 1989) which is the same regional chromosomal position of human C6 and C7 (Hobart et al. 1993). The gene encoding the human C8A and C8B are physically linked on chromosome 1p32 (Platteborze et al. 1996, Michelotti et al. 1995, Theriault et al. 1992, Rogde et al. 1986) whereas the human C8G is closely linked to a series of marker loci located in the most telomeric region of chromosome 9q34.3. Remarkably, this chromosomal region contains at least four other lipocalin genes (Dewald et al. 1996). In pig, the C8A gene was physically mapped to chromosome 6q3.3-q3.5 by in situ hybridization using the porcine bacterial artificial chromosome (BAC) clone as a hybridization probe (Nakajima et al. 1998) whereas the C9 was mapped to 16q1.4 by fluorescence in situ hybridisation (Thomsen et al. 1998).

## 2.7 Hemolytic complement activity

The previous studies for hemolytic complement activity were conducted on many different animal species. Schwab et al. (1966) reported the effect of temperature on the hemolytic activity of lizard and toad sera was similar to the effect on bactericidal activity, but with pig serum hemolytic activity was at a maximum between 35 and 40°C. This has been suggested that fetal pigs are unable to produce antibodies, and, hence, their serum does not kill smooth strains of gram-negative bacteria (Sterzl et al. 1962). Classical pathway hemolytic complement activity was assayed in camel trypanosomiasis infected with *Trypanosoma evansi* (Ouma et al. 1997)

The hemolytic complement activity of pig C6 gene was conducted to test the lysis of plasma ability with sensitised sheep erythrocytes. No lysis occurred in C6-deficient samples, but lysis was observed in normal plasma. Shitaba et al (1993) summarized that the presence of a null allele in pig population was found in other animals including man. Inheritance of C6 deficiency has not been tested yet.

A high degree of biochemical and structural similarity between porcine and human C7 protein was observed by Agah et al. (2000). Hemolytic assays were performed to test

whether porcine C7 could functionally replace human C7 in C7-deficient human sera. The experiment was conducted by addition of purified porcine C7 that restored the hemolytic activity of C7-depleted human sera in a dose-dependent manner. A functionally inhibitory mAb (Monoclonal antibody) against porcine C7 attenuated the hemolytic activity of human, rabbit, or rat sera, suggesting an important conserved C7 epitope among species. The results demonstrated that porcine and human C7 are highly conserved, sharing structural and functional characteristics (Agah et al. 2000). Additionally, CH50 titer was used to analyze hemolytic complement activity in human having ankylosing spondylitis due to C7 component deficiency (Delâge et al. 1977).

Absence of the eighth complement component and a disease compatible with systemic lupus erythematosus was found in a 56 year-old black woman. After the addition of functionally pure C8 hemolytic activity could be restored. C8 deficiency may be associated with a subtle defect in the defense mechanisms to viral infection leading to viral persistence and perhaps to diseases such as systemic lupus erythematosus where chronic viral infections have been implicated (Jasin 1977). Total hemolytic complement activity (CH50) was tested for the purpose of developing hereditary deficiency of complement component of C8 alpha-gamma and C6 in a colony of New Zealand White rabbits (Komatsu 1985).

Association between C9 deficiency with CH50 titer and diseases such as meningococcal meningitis (Zoppi et al. 1990), systemic lupus erythematosus (Takeda et al. 1994, Kawai et al. 1989), long-surviving mismatched kidney allografts (Hobart et al. 1997) were found. Almost all the ninth complement components (C9) deficiency in Japan shows Arg<sup>95</sup> stop mutation of C9 gene but not in patients with systemic lupus erythematosus susceptibility (Kanemitsu et al. 2000). C9 deficiency is also described in serum complement levels (CH50) of a 29 year-old woman in good health (Inai et al. 1979). Supplement of complement component C9 in neonatal serum can enhance the capacity of *Escherichia coli* killing shown by a radiobinding assay and immunogold electron microscopy using a monoclonal anti-C9 antibody (Lassiter et al. 1994). Using purified components demonstrated that C5b-9 exerts a regulatory effect on the formation of the classical and alternative pathway C3 convertases and on the utilization of C5 by cell-

bound C5 convertase. C5b-9 complex was unable to inhibit the lysis of cells bearing C5b-7 by C8 and C9 (Bhakdi et al. 1988).

In recent years people are much more concerned about the sustainable animal production system including quality control of food for human and animal, animal welfare and health status, effects of animal production on the environment, the influences of new technology application in animal and plant production, future animal production sites and global trading strategy (<http://www.agriculture.de/acms1/conf6/pdf/sumsum.pdf>). With the economic development the income of many people is increasing quickly not only in developing countries but also in poor countries. The demand for both quantity and quality of food in the daily meal is more and higher. Thus in animal breeding programs people address to produce animals with less fat, much muscle, less water in meat, less toxin or antibiotic, etc. Development in biotechnology and farm animal breeding genetic technology contributes a great potential for this. It will bring producer and customer much more benefit because of decreasing in production cost and price of product. Since the complement system was known as a natural disease resistance mechanism of host body and can support to solve several problems as described above, many studies on the complement components have been conducted in various animal species. In pig some porcine genes encoding complement components have already been identified or even analyzed as candidate genes for disease resistance. So far, one by one, the porcine C1 composed of three subcomponents C1q (alpha chain (GenBank acc. no. AY349424) and beta chain (GenBank acc. no. AY349420)), C1r (GenBank acc. no. AY349421) and C1s (GenBank acc. no. AY349426) were sequenced (Trakooljul et al. 2004). The porcine C2 nucleotide (2574 bp) and deduced protein (752 amino acids) sequences were found in GenBank (acc. no. NM\_001101815, NCBI homepage) derived from the pig DNA sequence of clone PigI-707F1 (GenBank acc. no. AL773527) (Sehra 2007). The porcine C3 cDNA sequences including 5' and 3' flanking regions (GenBank acc. no. AF154933) were determined and polymorphisms were detected to carry out an association analysis between C3 and complement activity traits (Mekchay et al. 2003, Wimmers et al. 2003). Moreover, Firth et al. (2007) showed the porcine C3d sequence (GenBank acc. no. EU257630) with 999 bp coding 333 amino acid whereas the porcine C4 (GenBank acc. no. AY349423) consists of 1656 bp encoding 517 amino acids (Trakooljul et al. 2004). In a total of 5422 bp of cDNA sequence coding the 1677 amino

acid precursor, the porcine complete C5 cDNA (GenBank acc. no. AY332748) was sequenced, screened for single nucleotide polymorphisms, and an association analysis with various immunological parameters was conducted (Kumar et al. 2004). In addition, Yi and Kang (2007) also separated porcine C5a cDNA sequence with 723 bp coding 226 amino acids (GenBank acc. no. EF210575) from C5. Several polymorphisms were detected within porcine C6 gene (Shibata et al. 1996) while the porcine C7 (GenBank acc. no. AF162274) was isolated and characterized (Agah et al. 2000). The complete C8A cDNA (2145 bp) as well as deduced protein (589 amino acid) sequence (GenBank acc. no. ADB13968) was determined and mapped to porcine chromosome 6q3.3-q3.5 (Nakajima et al. 1998) whereas the porcine C9 was assigned to chromosome 16q14 (Thomsen et al. 1998). In this work, focus was on the characterization of the genes encoding the terminal pathway of the complement system. Their variation and association with hemolytic complement activity in the classical and alternative pathway were studied in a F<sub>2</sub> DUMI resource population immunized with *Mycoplasma*, Aujeszky and PRRS vaccination, and the genes were assigned to porcine chromosome.

### 3 Materials and methods

Hereafter a general schema, which allows following the experiment up easily, is displayed (Figure 6).

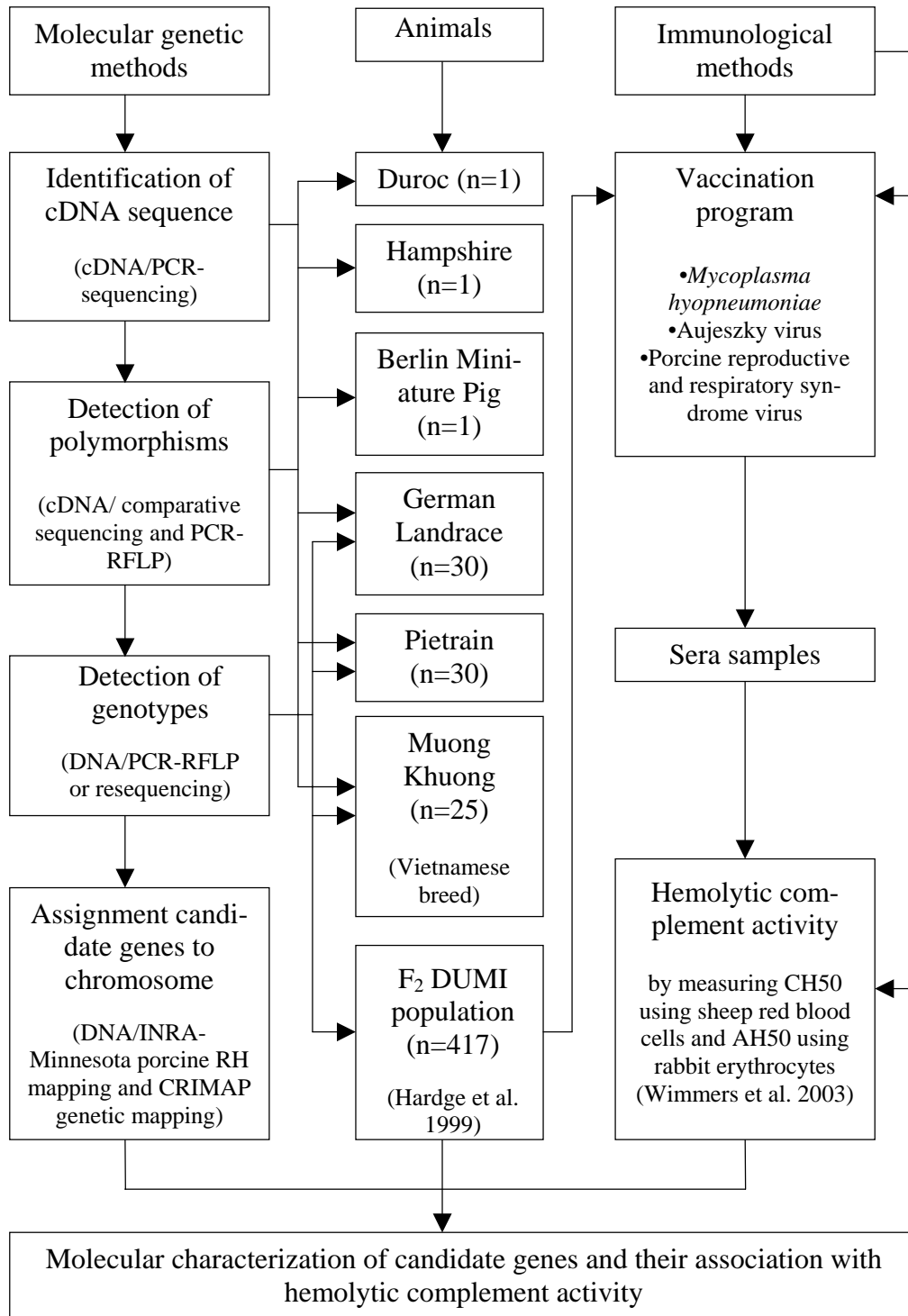


Figure 6 General schema of experimental approach

### 3.1 Materials

#### 3.1.1 Materials on station

##### 3.1.1.1 Experimental animals, tissue and data collection

In this study, one animal of each of the pig breeds Hampshire (HS), German Landrace (LR), Duroc (DR), Pietrain (PIE), Berlin Mininature Pig (BMP) and Muong Khuong (MK) (Vietnamese potbelly pig) was used to determine the entire sequence of the coding region (cds) of the candidate genes and to screen for polymorphisms. In addition, unrelated LR (n=30), PIE (n=30) and MK (n=25) animals as well as animals of the F<sub>2</sub> DUMI resource population were used for genotyping to determine allele frequencies and to perform linkage mapping and association analysis with hemolytic complement activity.

The Muong Khuong breed is a large resource population of pig reared by H'Mong householders who are living in the Northwest mountain area, mainly in Lao Cai province of Vietnam. These pigs are often kept at small-scale level (one-several animals per householder). The animals are allowed moving freely around the house/ garden and foraging for feed. They can enjoy anything they like in nature and in soil such as vegetables, worms, minerals, etc. They are also fed available simple feed resources or farm by-products such as rice bran, ground-yellow corn, vegetable, etc. The Muong Khuong pigs are divided in three different phenotypes (Figure 7). Birth weight can reach  $0.60\pm 0.04$  kg and weight at 12 months of age can weigh up to  $92.44\pm 2.13$  kg (average gain 4-6 kg per month). The adult weight at around 18 months can be up to 120 kg. The Muong Khuong pig, one of Vietnamese potbelly pig breeds, known as a breed, which can resist diseases and endoparasites and which is able to adapt well to harsh environmental conditions (temperature, humidity, nutrition, management) but not scientifically verified, represents a naturally selected population. In a recent report Thuy et al. (2006) indicated that the Vietnamese autochthonous breeds including Muong Khuong breed show higher degree of polymorphism, allelic diversity, and heterozygosity than several European pig breeds (German Landrace, Pietrain and Large White). Vietnamese local breeds are a source for promising alleles of unpredictable economic value (Lemke et al.

2005). Genetic distances show large differences among European-based, Chinese, and Vietnamese indigenous breeds and reflect the geographical distribution of breeds. In comparison with the European breeds, the Vietnamese indigenous pig breeds harbour a considerable amount of genetic diversity and therefore they will be of significance for livestock bioconservation (Thuy et al. 2006).



Figure 7 Three different phenotypes in Muong Khuong pig breed: ❶\_black colour with 6 white points (one in the middle of head, four in legs and one in the bottom of tail), ❷\_brown colour, and ❸\_dark black colour

Six adult Muong Khuong animals (three males and three females) selected at an average live weight of about 65 kg were bought for experimental sampling in Muong Khuong district of Lao Cai province. The fresh tissue specimens were submerged in *RNAlater* solution and stored at  $-20^{\circ}\text{C}$  according to Ambion's manufacture instruction (Figure 8). Total mRNA was isolated from the liver tissue and reverse transcribed to cDNA whereas DNA was extracted from tail/ ear samples of 25 Muong Khuong animals.

Liver tissue and ear/tail tissue samples from 30 unrelated individuals of each commercial breed LR and PIE raised at the Research and Performance Test Station, Frankenfurst, University of Bonn, Germany were collected for RNA isolation and for DNA isolation, respectively.

The  $F_2$  DUMI resource population was derived from a reciprocal cross of Duroc and Berlin Miniature Pig (Hardge et al. 1999). The Berlin Miniature Pig was bred from Vietnamese potbelly pig, Saddleback Pigs and German Landrace (Hardge et al. 1999). Five sows of Berlin Miniature Pigs were crossed with a Duroc boar and four Duroc sows were crossed with a Berlin Miniature boar to produce  $F_1$  animals (parental genera-



tion). In total 47 F<sub>1</sub> animals were used to produce the F<sub>2</sub> animals (n=902) called the F<sub>2</sub> Berlin-Bonn-DUMI resource population.



Blood sampling



Fresh tissue sampling



Carcass of Muong Khuong pig

Figure 8 Sampling in Muong Khuong animals on farm in Lao Cai province, Vietnam

About half of them (n=485) were reared at the Research Station of the Humboldt University, Berlin (F<sub>2</sub> Berlin DUMI) while the other F<sub>2</sub> animals (n=417) were reared and tested performance at the Research Farm of Frankenforst, Institute of Animal Breeding Science, University of Bonn (F<sub>2</sub> Bonn DUMI). The F<sub>2</sub> Bonn DUMI animals used in this study originated from 11 sows and three boars of the F<sub>1</sub> generation and they have been phenotyped regarding the complement activity and immune responsiveness as described below (Figure 9).

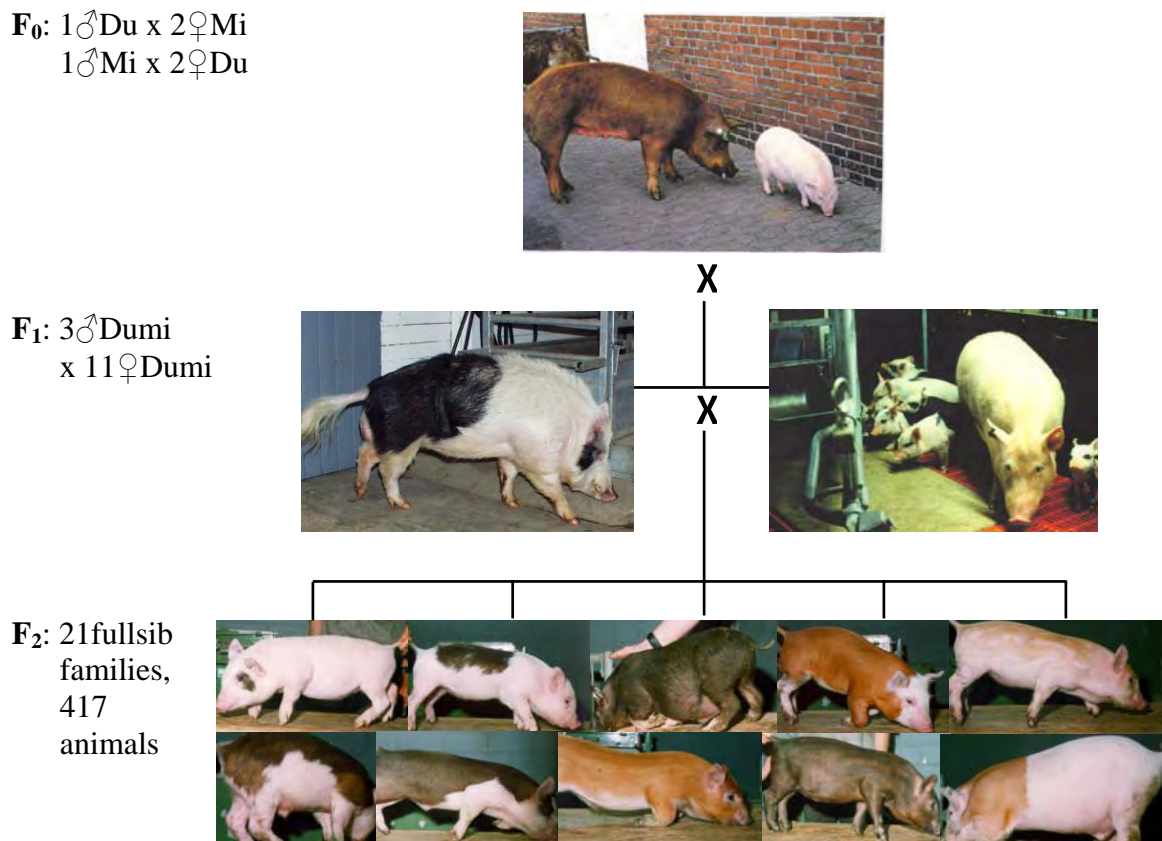


Figure 9 The structure of F<sub>2</sub> DUMI resource population based on reciprocal crossing of Duroc and Berlin Miniature Pig

### 3.1.1.2 Phenotyping of the immune responsiveness: sampling design

In order to investigate immune competence traits the following experiment was conducted. The experiment carried out at the Frankenforst Research Station of University of Bonn consisted of two trials: (1) The control group with 36 F<sub>2</sub> DUMI animals was unvaccinated. (2) The trial group containing the 381 remaining animals was immunized

with *Mycoplasma hyopneumoniae* vaccine (Mh) (Stellamune, Mycoplasma, Pfizer, Karlsruhe, Germany), Aujeszky virus vaccine (ADV) (Porcilis, Begonia Diluvac, Intervet, Tönisvorst, Germany) and Porcine Reproductive and Respiratory Syndrome Virus vaccine (PRRSV) (Ingelvac PRRS MLV, Boehringer Ingelheim, Germany) at six, 14 and 20 weeks of age for measuring immunological parameters and analysing association between the candidate genes and hemolytic complement activity in both the classical and alternative pathway. EDTA anticoagulated blood samples were collected from each animal before (day 0) and after Mh and ADV vaccination (day 4 and 10) but only at day 10 after PRRSV vaccination (eight sampling time points) and were then cooled immediately in ice. Sera and plasma isolations were obtained within 2 hours after blood taking by centrifugation at 4°C at a maximum speed of 14.000 rpm for 10 min and stored at -80°C for further analysis (Wimmers et al. 2003) (Figure 10).

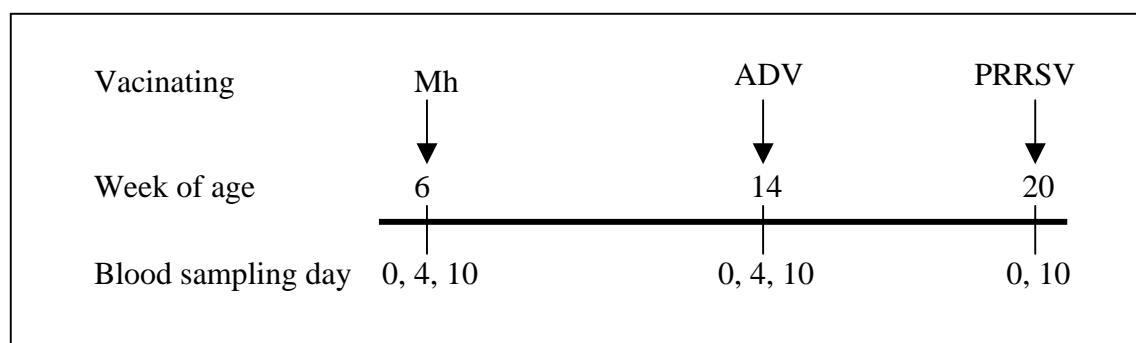


Figure 10 Vaccination program and time of blood sampling before and after immunization in F<sub>2</sub> DUMI population

### 3.1.2 Materials for laboratory analysis

#### 3.1.2.1 Chemicals, biological materials, kits and others

Applied Biosystems (Darmstadt)	:	RNAlater <sup>®</sup> -ICE
Biomol (Hamburg)	:	Phenol
Fermentas (Leon-Rot)	:	Glycogen, Restriction enzymes (AluI, BbvI, BseDI, BsrDI, Eco91I, FnuDII, Hin6I, HpaII, PstI,)

Invitrogen (Karlsruhe)	:	SuperScrip <sup>TM</sup> III reverse transcriptase
Macherey-Nagel (Düren)	:	NucleoSpin <sup>®</sup> Extract II
New England Biolabs <sup>®</sup> Inc. (Frankfurt am Main)	:	Restriction enzymes (AciI, BsrI, DdeI, HpyCH4III)
Promega (Mannheim)	:	Dithiothreitol (DTT), Rnasin Ribonuclease inhibitor, RQ1 RNase-free, DNase I and 10 x buffer, oligo (dT)13 VN primer, random hexamer primer
Roth (Karlsruhe)	:	Formaldehyde, Proteinase K, Sodium chloride, Sodium hydroxide, Tris, Acetic acid, Chloroform, dNTP, Ethylenediaminetetraacetic acid (EDTA), Ethanol, Ethidium bromide, Formaldehyde, Sodium dodecylsulphate (SDS), Sodium chloride
Sigma (Genosys, Steinheim)	:	Oligonucleotide primers, Agarose
Sigma-Aldrich (Taufkirchen)	:	Isopropanol, TRI-Reagent, Taq polymerase

### 3.1.2.2 Buffers and reagents

Digestion buffer	:	Tris HCl, pH 8.0	0.05 M
		EDTA	0.1 M
		SDS	0.5%
		Water added up to	1000 ml
SDS solution	:	Sodium dodecylsulphate	10%
Sodium acetate	:	Sodium acetate trihydrat (>99.5%, D~1.42, M 136.08)	3 M
		Acetic acid to adjust pH 5.2	
		Water added up to	1000 ml
1 x TE buffer	:	Tris	10 mM
		EDTA	1 mM
		Water added up to	1000 ml

---

50 x TAE buffer	:	Tris	2 M
		Acetic acid	57% (v/v)
		EDTA, pH 8.0	0.05 M
		Water added up to	1000 ml

### 3.1.2.3 Used softwares

CIRMAP version 2.4: Linkage mapping

(<http://linkage.rockefeller.edu/soft/crimap/>)

Chromas software

(<http://www.technelysium.com.au/chromas.html>)

ClustalW2: DNA/ protein alignment

(<http://www.ebi.ac.uk/clustalw/>)

Conserved Domains Search: Functional protein domain analysis

(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)

DiANNA 1.1 web server: Cysteine state and disulfide bond partner prediction

(<http://clavius.bc.edu/%7Eclotelab/DiANNA/>)

INRA-UMN porcine Radiation Hybrid (IMpRH) panel: RH mapping

(<http://www2.toulouse.inra.fr/lgc/pig/RH/IMpRH.htm>)

MEGA3: Phylogenetic analysis

(<http://www.megasoftware.net/>)

MultAlin : Multiple sequence alignment

(<http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html>)

NAGRP Blast Server

(<http://www.animalgenome.org/blast/>)

NCBI/BLAST

(<http://www.ncbi.nlm.nih.gov/BLAST/>)

NCBI/Map Viewer

(<http://www.ncbi.nlm.nih.gov/mapview/>)

NCBI/Trace archive: retrieving genomic DNA sequence

(<http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?>)

NEBcutter V2.0: Restriction digestion analysis

(<http://tools.neb.com/NEBcutter2/index.php>)

- NetCGlyc 1.0 Server: C-Manosylation site prediction  
(<http://www.cbs.dtu.dk/services/NetCGlyc/>)
- NetNGlyc 1.0 Server: N-Glycosylation site prediction  
(<http://www.cbs.dtu.dk/services/NetNGlyc/>)
- ORF Finder: deduced amino acid sequence  
(<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>)
- P O L Y V I E W - 3D: 3D Protein structures  
(<http://polyview.cchmc.org/polyview3d.html>)
- Primer 3: Primer design  
([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi))
- SAS version 9.3.1, SAS Institute Inc., Cary, NC: Statistical analysis  
(<http://support.sas.com/documentation/onlinedoc/91pdf/>)
- SignalP 3.0: Signal peptide analysis  
(<http://www.cbs.dtu.dk/services/SignalP/>)
- SMART mode: Functional protein domain analysis  
(<http://smart.embl-heidelberg.de/>)
- Webcutter: Restriction digestion analysis  
(<http://www.firstmarket.com/cutter/cut2.html>)
- 3D-Jigsaw comparative modelling: three-dimensional models for proteins  
(<http://www.bmm.icnet.uk/servers/3djigsaw/>)

#### 3.1.2.4 Equipments

##### Centrifuges

ABI 310 Genetic Analyzer (Applied Biosystems, Darmstadt)

MegaBACE 1000 (GE Healthcare, München)

PCR thermocycler (Biometra, Göttingen)

Power supply BIO-RAD Model 200/2.0 (Electrophoresis for agarose gel, Groton)

Spectrophotometer Nanodrop ND-1000 (Peqlab, Erlangen)

UV transilluminator (High performance 302 – 365 nm, Cambridge)

## 3.2 Methods

### 3.2.1 Immunological methods: complement activity measurement

Data of complement activity were obtained previously by the following procedure: for determining complement activity in the classical and alternative complement pathway hemolytic assays were applied according to the methods of Liu and Young (1988) and Demey et al. (1993). Hemolytic complement activity was expressed as the amount of serum that caused a 50% hemolysis of antibody sensitised sheep red blood cells (SRBCs) in the reaction mixture. Calculation was based on the change in light-scattering properties of erythrocytes upon lysis. The test conditions were evaluated and standardised for the assay of pig complement. SRBCs sensitised with hemolysin (anti-sheep red blood cell stroma from rabbit, Sigma, Taufkirchen, Germany) were collected from healthy animals, submerged into Alsever's solution and stored at 4°C until analysis. SRBCs were diluted 1:75 in GVBSS and used as a 2% cell-suspension. Serial dilutions of test sera (50 µl) were made in duplicate in flat-bottomed 96-well microtitre plates (1:2, 1:4, 1:8, 1:16, 1:32, 1:64, 1:128, 1:256). Fifty microlitter of sensitised SRBCs suspension were added to each serum dilution, respectively. The plates were incubated for 90 min at 37°C. The results were recorded through absorbance level of a microplate reader (ThermoMax ELISA reader and Soft Max Pro software, Molecular Devices, USA) at 650 nm. The readings were transformed by the method of von Krogh according to Mayer (1961) and the hemolytic titer was expressed as the titre that lysed 50% of the erythrocytes (CH50, U/ml) (Wimmers et al. 2003).

Rabbit erythrocytes (RRBCs) were used as foreign surfaces for testing alternative complement pathway activity in pig sera. Twenty-five microlitres of diluted sera (1:2, 1:3, 1:4 and 1:6) with GVBS-Mg-EGTA buffer and 25 µl of 0.5% (v/v) rabbit erythrocytes were incubated at 37°C for 90 min in duplicate wells of flatted bottom 96-well microplates (Roth). A 150 µl GVBSS-Mg-EGTA buffer was added in each well and the absorbance values at 650 nm were read. The activity was expressed in term of serum amount which causes 50% hemolysis of rabbit erythrocytes (AH50, U/ml) (Wimmers et al. 2003).

### 3.2.2 Molecular genetic methods

#### 3.2.2.1 DNA extraction

Preparation of genomic DNA from tail or ear samples was performed by standard procedures involving Proteinase K digestion followed by phenol-chloroform extraction and ethanol precipitation. Tail or ear tissues were cut into small pieces of 2-3 mm and weighed about 0.1 g. The sample was placed in 2 ml tube containing 700  $\mu$ l of digestion buffer. To lyse the cells and digest proteins, 35  $\mu$ l of Proteinase K solution (20 mg/ml) was added and the samples were then incubated at 55°C overnight with mixing. An equal volume of phenol-chloroform (1:1 v/v, 700  $\mu$ l for each) was added. The two phases were mixed until they formed a homogenous emulsion. The two phases were separated by centrifugation at 5000 rpm for 3 min at 4°C. The aqueous phase was collected in fresh tubes. Phenol-chloroform extraction was repeated and was followed by a chloroform extraction. One-tenths volume of sodium acetate (3 M, pH 6.0) and an equal volume of isopropanol (700  $\mu$ l) were added. The samples were shaken gently until precipitation of DNA. The DNA pellet was washed three times with 1 ml ethanol (70%) and dried at room temperature. Finally the DNA was resuspended in 1 $\times$  TE buffer and stored at 4°C for analysis.

#### 3.2.2.2 RNA isolation

Total RNA was isolated from liver tissue of six different breeds (HS, DR, LR, PIE, BMP and MK). About 100 mg of liver was ground into fine powder with a set of porcelain mortar and pestle, submerged with 1 ml TRI Reagent in 2 ml tubes, mixed well, homogenized by passing several times through 0,9 mm needle attached to a syringe, and incubated at room temperature for 5 min for complete cell lysis. The sample was centrifuged for 15 min at 4°C at 12,000 x g. The aqueous phase was transferred to a fresh tube and 500  $\mu$ l of isopropanol were added and gently mixed by inversion. After incubation at room temperature for 10 min the RNA was pelleted by centrifugation at 4°C for 10 min at 12,000 x g. The supernatant was discarded while the white RNA pellet at the bottom of the tube was kept and washed with 1 ml of chilled 70% ethanol by gentle vortexing lightly and centrifuging at 12,000 x g for 5 min at 4°C. The ethanol was re-



moved by pipetting and the RNA pellet was air dried at room temperature. The RNA pellet was resuspended in 30  $\mu$ l of RNase-free water. DNaseI was used to remove traces of genomic DNA. The DNaseI digestion was performed in 50  $\mu$ l solution, containing 30  $\mu$ l RNA solution, 5  $\mu$ l of 10 x buffer, 5  $\mu$ l DDT (0.1 M), 1  $\mu$ l RNase-free water and was then incubated for 1 hour at 37°C. For subsequent RNA clean-up the NucleoSpin<sup>®</sup> Extract II Kit was used according to manufacturer instructions. The final RNA concentration was measured by using Nanodrop spectrophotometer. The integrity was checked on 1% agarose gel/TAE containing formaldehyde and ethidiumbromide.

### 3.2.2.3 Complementary DNA (cDNA) synthesis

First-strand cDNA was synthesis from 1  $\mu$ g of total RNA using 500 ng of oligo (dT)13 VN primer and 500 ng random hexamer primers. For denaturing the RNA was mixed with the primers, brought to a final volume of 11  $\mu$ l and subsequently heated for 5 min at 68°C and immediately chilled on ice for at least 2 min. After addition of a reaction mix containing 4  $\mu$ l of 5x First Strand Buffer, 1  $\mu$ l of 0.1M dithiothreitol (DTT), 1  $\mu$ l of rRNasin (Rnase Inhibitor 40 U/ $\mu$ l), 25  $\mu$ M dNTPs and 1  $\mu$ l of Super Script<sup>™</sup>III Rnase H<sup>-</sup> and RNase free water to a final volume of 20  $\mu$ l, the cDNA synthesis reaction proceeded by incubation of the reaction mixture at 25°C for 5 min (Hexamer annealing), followed by 50°C for 1 hour (reverse transcription), 70°C for 15 min (reverse transcription inhibition) in a thermal cycler. The cDNA solution was diluted 1:5 and stored at –20°C for the subsequent PCR reaction. To confirm the complete removal of genomic DNA and successful cDNA synthesis, RNA and cDNA were used as template in a PCR reaction using intron spanning primers amplifying a part of  *$\beta$ -actin* gene.

### 3.2.2.4 Amplicon and primer design

In order to obtain whole cDNA sequence of the candidate genes, screening known human and murine sequences in public databases was conducted and the nucleotide-nucleotide BLAST tool (blastn) was used to retrieve available porcine orthologs (expressed sequence tags, ESTs) of the candidate genes. The retrieved porcine sequences from previous sequence information (GenBank acc. no. BP444694.1, BP444335.1, BP454270.1, BP445832.1, BP451787.1 for C6, BI336392, CF176130, AF162274 for

C7, BP448099.1, BP 449671.1, BX920949.1 for C8A, BP447190.1, BP446376.1, BX918195.1 for C8B, BP139629.1, BP443148.1 for C8G, BP139003.1, CJ007506.1, BX920671, CF363669 for C9) were assembled into contigs and the exon structure. Possible gaps were determined and their approximate sizes were estimated using the orthologous human and mouse sequences. For the porcine C7 gene the whole cds was already cloned by Agah et al. (2000, GenBank acc. no. AF162274). In order to bridge existing gaps and to amplify overlapping fragments of the whole cds, comparative sequencing primers amplifying amplicons suitable for direct sequencing were designed using Primer3 tool. The primer sequences, their position in the cDNA sequence and the lengths of corresponding amplicons are given in table 1.

#### 3.2.2.5 Polymerase chain reaction (PCR)

A standard PCR mixture for sequencing and genotyping contained 50 ng of liver cDNA or 100 ng of genomic DNA, 0.2 mM of each primer (forward and reverse primer), 50  $\mu$ M of each dNTP, 0.5 U of Taq polymerase and 1xPCR buffer containing 1.5 mM of MgCl<sub>2</sub> in a final volume of 20  $\mu$ l. Standard PCR thermal cycling program was set up with an initial denaturation step of 94°C for 4 min, followed by 40 cycles at 94°C for 30 sec, annealing at 60°C for 30 sec, elongation at 72°C for 1 min and a final extension at 72°C for 5 min. Individually, due to the combination condition the thermal program for primers C9.5 and C9.8.8a were conducted under the annealing temperature 50 and 55°C for 45 sec, respectively, followed by elongation step at 72°C for 75 sec. Additionally, for some primer pairs touchdown PCR profiles turned out to be optimal with conditions: initial denaturation at 94°C for 4 min, followed by eight cycles at 94°C for 30 sec, from 58-54°C for primers C6.1, C7.1 and C7.2, from 64-60°C for primers C7.3a and C7.4 or 68-64°C for primer C7.5 for 30 sec, at 72°C for 1 min (annealing temperature was step-downs 0.5°C for each cycle repeated two times), then followed by 35-40 cycles of 94°C for 30 sec, at 54, 60 or 64°C for 30 sec, at 72°C for 1 min, and ending with an extension step at 72°C for 5 min. Thermal cycling conditions are summarized in table 1. Three to 3-5  $\mu$ l of PCR products were analyzed on 1% agarose gel stained with ethidium bromide in 1 x TAE buffer and electrophoresed to evaluate specificity and efficiency of the amplification and the remainder was retained for sequencing or genotyping.

### 3.2.2.6 PCR product purification

The PCR products were purified by ethanol precipitation as follow: 42.5  $\mu$ l of 100% ethanol, 1.7  $\mu$ l of sodium acetate (pH 5.2, 3 M) and 1  $\mu$ l of glycogen (1  $\mu$ g/ $\mu$ l) were added into 17  $\mu$ l of the PCR product and the mixture was incubated overnight at  $-20^{\circ}\text{C}$ . The PCR products were pelleted by centrifugation at 14000 rpm at  $4^{\circ}\text{C}$  for 50 min, subsequently air dried and dissolved in 10  $\mu$ l of PCR-grade water. Alternatively the PCR-products were purified using the NucleoSpin<sup>®</sup> Extract II Kit according to manufacturer instructions. The concentration and purity of the purified PCR products was evaluated on the Nanodrop spectrophotometer.

### 3.2.2.7 Sequencing

In order to identify cDNA/ DNA sequences and polymorphisms of the candidate genes, the PCR products after purification were used for sequencing based on the Sanger dideoxy nucleotide triphosphate (ddNTP) terminator method. According to the protocol, a total of 10  $\mu$ l reaction mixture was prepared containing 5 ng of the purified PCR products for each 100 bp of the DNA fragment length, 5 pm (0.5  $\mu$ l) primer either up or down, 1.5  $\mu$ l BigDye<sup>R</sup> Terminator Ready Reaction Premix and water. The cycle sequencing reaction was done with an initial denaturation step of  $96^{\circ}\text{C}$  for 1 min followed by 25 cycles at  $96^{\circ}\text{C}$  for 10 sec, from  $50$ - $65^{\circ}\text{C}$  (according to the annealing temperature of each primer as given in table 1 and 4) for 5 sec,  $60^{\circ}\text{C}$  for 4 min and cooled down to  $6^{\circ}\text{C}$ . The products were purified using the Princeton Separations Centri Sep Spin Column Kit. An aliquot of 4  $\mu$ l of PCR products was analyzed on the MegaBACE 1000 or the ABI 310 Genetic Analyzer capillary sequencer. Each sequence chromatogram was manually inspected using Chromas and when need trimmed of poor-quality sequence.

### 3.2.2.8 Screening for single nucleotide polymorphisms (SNPs)

For SNP identification the processed individual sequences of each of the six breeds were aligned and compared together using the web-based program 'MultAlin'. The SNP sites are confirmed by either using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP), which can discriminate between the alleles, or by

resequencing. Individual sequence traces were inspected using Chromas software. Verification of SNP was performed using cDNA as template. The typical restriction digestion reaction mixture was performed in a total volume of 20  $\mu$ l containing 15  $\mu$ l aliquot of PCR product, 10 U of the appropriate restriction enzyme and 2  $\mu$ l supplied reaction buffer. Incubation was done overnight to ensure complete digestion and was finished by heat inactivation for 20 min according to Fermentas or New England Biolabs manufacturer's instruction. The enzymes and conditions used for RFLP analysis are summarized in table 3. Fragment length polymorphisms were analyzed on a 3-5% agarose gel stained with ethidium bromide for evaluation and documentation.

### 3.2.2.9 SNP genotyping

Once the general mutation picture was established, some of the SNP sites in each gene were selected for genotyping in unrelated healthy animals of breeds LR (n=30), PIE (n=30) and MK (n=25) and 417 animals of F<sub>2</sub> DUMI population by using PCR-RFLP analysis or by resequencing. Amplifications as well as methods for the restriction digestion reaction are described above. All of the PCR products were separated on 3-5% agarose gel for discriminated allele identification. For genotyping, genomic DNA samples were used. The exon structure of the individual genes, derived from pig-human and/or pig-mouse comparison, was considered when designing amplicons. Intron sequences were either determined by amplification of genomic DNA using suitable primer pairs and sequencing or the sequences were retrieved from public databases as described in paragraph 3.2.2.4. Therefore one of two primers located on one exon segment whereas another primer was designed in the neighbouring exon or intron sequence. A summary of PCR conditions, amplicon designs and primers used for genotyping is given in table 4.

### 3.2.2.10 Assignment of candidate genes to chromosomes

In order to perform physical mapping of the candidate genes the INRA-University of Minnesota porcine 7000Rad Radiation Hybrid panel (IMpRH) containing 118 pig/hamster DNA hybrid clones (Hawken et al. 1999, Yerle et al. 1998) was employed. Based on accession number sequence DQ333199, NM\_214282, 5' flanking region of clone XX-1C1, DQ333201, DQ333202 and DQ333198, specific primer pairs, which allowed amplifying DNA fragments of 159 to 707 bp for mapping C6, C7, C8A, C8B, C8G and

C9 were derived. Amplicon design, primer sequences and PCR conditions are summarized in table 2. Prior to mapping, PCR conditions were optimized so that specific amplification of porcine DNA but not DNA of the hamster parental lines or amplification of fragments unambiguously distinguishable between the two species was achieved. The expected porcine DNA fragments were then sequenced to verify their identity. PCR reaction was performed in a total 15  $\mu$ l reaction mixture containing 25 ng of hybrid DNA, 100  $\mu$ M of each dNTP, 0.1  $\mu$ M of each primer, 1 $\times$ supplied PCR buffer containing 1.5 mM MgCl<sub>2</sub> and 1 U Taq polymerase. PCR reactions were prepared for 118 DNA templates of the IMpRH panel (positive), a hamster DNA template (negative control), a blank-template without DNA (negative control) and a porcine genomic DNA (positive control). The PCR products were amplified using standardized thermal profile as follows: 4 min of initial denaturation at 94°C, 40 cycles at 94°C for 30 sec, at the annealing temperature 58°C or 60°C (depending on the specific primer pairs shown in table 2) for 30 sec, at 72°C for 1 min and a final extension at 72°C for 5 min. For mapping of the C7 gene a modified, touch-down thermal profile was employed: 94°C for 4 min, followed by 9 cycles at 94°C for 30 sec, annealing temperature from 60-51°C (-1°C per cycle) for 30 sec, 72°C for 1 min, followed by 40 cycles of 94°C for 30 sec with annealing temperature at 50°C for 30 sec, at 72°C for 1 min, ending with an extension step at 72°C for 5 min. The entire PCR reactions were separated on 3% agarose gel stained with ethidium bromide and amplification products were independently scored as present (1), absent (0), or ambiguous (?). The results translated into vector format for submitting to IMpRH mapping tool. The PCR-screening of the IMpRH panel was performed twice. For ambiguous data, the PCR was repeated to minimize genotyping errors and the remaining discrepancies scored as ambiguous. Two-point linkage analysis was done using the IMpRH mapping tool available at the IMpRH Web Server.

Linkage mapping of candidate genes was performed using CRIMAP version 2.4 (Green et al. 1990). Therefore genotype information of the respective complement genes and of microsatellite markers previously obtained of the DUMI resource population was used (Wimmers et al. 2002).

### 3.2.2.11 Phylogenetic analysis

In order to estimate evolutionary distances for amino acid sequences of the candidate genes among various species that are believed to have a common ancestor in interrelationships unrooted phylogenetic trees were constructed by using the neighbour-joining distance-matrix method with 500 bootstrap replicates on the reliability of internal branches. Amino acid sequences were extracted from reports in GenBank. In this study the computational phylogenetics program MEGA 3 (Kumar et al. 2004), which calculates genetic distance from multiple sequence alignments, was used for analysis.

### 3.2.2.12 Statistical analysis

In this study, the SAS's PROC MIXED procedure combined with REPEATED statement (The SAS software package, release 9.1) was used for analyzing the variance of experimental dataset to estimate the effect of genotypes of candidate genes on complement activity at eight different time points of vaccination. The analysis model address valid standard errors of the fixed effect factor estimates in order to identify other significant environmental and genetic effects apart from the factor of genotypes and its interaction by stepwise elimination of non-significant effects. The SAS mixed linear model was as given below:

$$y_{ijklmno} = \mu + \text{sire}_i + \text{dam}_j + \text{parity}_k + \text{treatment}_l + \text{genotype}_m + \text{time}_n + \text{sex}_o + \text{ANIMAL}_{ijklmno} + (\text{genotype} \times \text{time})_{mn} + \varepsilon_{ijklmno}$$

Where:

$y_{ijklmno}$	: hemolytic complement activity in classical pathway
$\mu$	: overall mean
$\text{sire}_i$	: fixed effect of sire; $i=1-3$
$\text{dam}_j$	: fixed effect of dam; $j=1-11$
$\text{parity}_k$	: fixed effect of parity; $k=1-5$
$\text{treatment}_l$	: fixed effect of treatment-vaccinated trial/unvaccinated control; $l=1-2$
$\text{genotype}_m$	: fixed effect of genotype; $m=1-3$
$\text{time}_n$	: fixed effect of time point of measurement prior and after vaccinations; $n=1-8$
$\text{sex}_o$	: fixed effect of sex; $o=1-2$
$\text{ANIMAL}_{ijklmno}$	: random effect of animal
$(\text{genotype} \times \text{time})_{mn}$	: interaction between genotype and time point
$\varepsilon_{ijklmno}$	: residual error

**Table 1** Oligonucleotide primer pairs used for determination of the coding sequence and 5' and 3' untranslated regions (UTR) and for screening SNP of the candidate genes

Gene name/ primer set	Primer sequence (localization on GenBank acc. no) <sup>(*)</sup>	Annealing temperature (°C)	Length (bp)
<b>C6 gene</b>	<b>GenBank acc.no. DQ333199</b>		
- C6.1	up (5'-3'): ttacaagaagatgaaaggagga (nt. 55-77, 5'UTR) down (3'-5'): tcatccgaattgtctccaca (nt. 634-653, exon 5)	touchdown 58-54	599
- C6.6	up (5'-3'): cagcaagttggaatgcaatg (nt. 603-622, exon 5) down (3'-5'): gacgtcagaaagctgcaaac (nt. 1135-1155, exon 8)	60	553
- C6.7	up (5'-3'): tcccacaaaaggattccag (nt. 1060-1079, exon 7&8) down (3'-5'): aatttgctgcgtattctcg (nt. 1678-1697, exon 11)	60	638
- C6.10	up (5'-3'): ctttgcatgggagaaaagg (nt. 1505-1523, exon 10) down (3'-5'): ctggcattccacatctcctc (nt. 2225-2244, exon 14)	60	740
- C6.8	up (5'-3'): catcggatgatgaggaca (nt. 2013-2032, exon 13) down (3'-5'): tggagtctgtatcaagcacaca (nt. 2546-2566, exon 17)	60	554
- C6.9	up (5'-3'): ttaaaggccattgtcaacc (nt. 2449-2468, exon 16) down (3'-5'): aactctggatgtggcagtc (nt. 3018-3037, 3' UTR)	60	589
<b>C7 gene</b>	<b>GenBank acc.no. AF162274</b>		
- C7.1	up (5'-3'): ttctgcctgaatgtttcc (nt. 96-115, 5'UTR) <sup>(**)</sup> down (3'-5'): gggattcataggttttgcac (nt. 827-847, exon 7)	touchdown 58-54	752
- C7.2	up (5'-3'): ttttgaacttgctgagtcac (nt. 806-826, exon 8) down (3'-5'): atgccaccgttttgacaag (nt. 1376-1394, exon 11)	touchdown 58-54	589
- C7.3a	up (5'-3'): gccctggaggagtatctgg (nt. 1330-1348, exon 11) down (3'-5'): ccaacaggaacgtggtctc (nt. 1756-1775, exon 14)	touchdown 64-60	446
- C7.4	up (5'-3'): agaattctgtccatcacctcct (nt. 1704-1725, exon 13) down (3'-5'): caggtctcccaacagaaag (2269-2288, exon 17)	touchdown 64-60	585
- C7.5	up (5'-3'): ccctaatgacggatcttc (nt. 2152-2171, exon 16&17) down (3'-5'): acgaggggtttgcgtttac (nt. 575-593, 3'UTR) <sup>(***)</sup>	touchdown 68-64	496
-C7pro2DNA	up (5'-3'): catggaaaatgaggctctgg (nt. 410-429, ti: 784743778; name: dpcxa0_039446.z1; mate:784159587, 5'UTR) down (3'-5'): gctgggactgttgaaatg (nt. 111-130 in reverse com- plement,ti: 848892623; name: rdpaxb0_005371.y1; mate: 854561309), 5'UTR)	58	531

Table 1 (continued)

Gene name/ primer set	Primer sequence (localization on GenBank acc. no) <sup>(*)</sup>	Annealing temperature (°C)	Length (bp)
-C7pro3DNA	up (5'-3'): gagccttatcccttgcttac (nt. 603-623, ti: 860292528; name: bye_78346.z1; mate:768266164, 5'UTR) down (3'-5'): tgttgggaaaacattcagg (nt.387-406 in reverse com- plement, ti: 808296981; name: rdpcxb0_216535.y1; mate: 785838286, 5'UTR)	58	403
<b>C8A gene</b>	<b>GenBank acc.no. DQ333200</b>		
- C8A.1	up (5'-3'): ggtcctggctctgtagacacctc (nt. 5-26, 5'UTR) down (3'-5'): gatcctgacatcttcacagtcac (nt. 485-505, exon 4)	60	501
- C8A.3	up (5'-3'): gagtgtctgaggcaagcaca (nt. 365-384, exon 3) down (3'-5'): ttgctgtaaatgggctacctg (nt. 870-890, exon 6)	60	526
- C8A.4	up (5'-3'): aagccctggcagacagtaaa (nt. 753-772, exon 6) down (3'-5'): gtggccacctccaagtttt (nt. 1306-1324, exon 8)	60	572
- C8A.5	up (5'-3'): catcatgctttggagggtct (nt. 1218-1237, exon 8) down (3'-5'): ccttggctcctctgactt (nt. 1700-1718, exon 10-11)	60	501
- C8A.7.8	up (5'-3'): tacgagatactgcgccacac (nt. 1499-1517, exon 10) down (3'-5'): agctggcaggacagagaaaa (nt. 1969-1988, 3'UTR)	60	490
<b>C8B gene</b>	<b>GenBank acc.no. DQ333201</b>		
- C8B.1	up (5'-3'): atgaagacctctgggacgtg (nt. 65-84, exon 1) down (3'-5'): tccatttctctgacacttt (nt. 544-564, exon 4)	60	500
- C8B.2	up (5'-3'): atcagtcggatgaggcaaac (nt. 507-526, exon 4) down (3'-5'): tgacctctgaaggaactcg (nt. 1009-1028, exon 7)	60	522
- C8B.3	up (5'-3'): agcgcgttacaagctgaaat (nt. 967-986, exon7) down (3'-5'): cggtgtcaccagttcatag (nt. 1474-1493, exon 10)	60	527
- C8B.4	up (5'-3'): gcagtgcagtacaaccaga (nt. 1427-1446, exon 9) down (3'-5'): gtggaagctggtgtggagat (nt. 1922-1941, 3'UTR)	60	515
<b>C8G gene</b>	<b>GenBank acc.no. DQ333202</b>		
- C8G.1	up (5'-3'): tcgtctgctctctgtctcc (nt. 2-21, 5'UTR) down (3'-5'): gcgtacagcttcaccgacag (nt. 487-506, exon 4)	60	505
- C8G.4	up (5'-3'): acctccgaaagctggatg (nt. 313-331, exon 2-3) down (3'-5'): ttgagaccgacagtgttaatga (nt. 790-812, 3'UTR)	60	500



Table 1 (continued)

Gene name/ primer set	Primer sequence (localization on GenBank acc. no) <sup>(*)</sup>	Annealing temperature (°C)	Length (bp)
<b>C9 gene</b>	<b>GenBank acc.no. DQ333198</b>		
- C9.1	up (5'-3'): cctttgcagtatgcattttaga (nt. 112-133, exon 1) down (3'-5'): ttctgaaattttgcactttgg (nt. 706-729, exon 5)	60	618
- C9.3.1	up (5'-3'): aaagaggcgcctctgtgta (nt. 428-447, exon 4) down (3'-5'): tgggtccataggttccaa (nt. 1116-1135, exon 7)	62	708
- C9.5	up (5'-3'): tattcctgcatgtcaaaggag (nt. 979-999, exon 7) down (3'-5'): ccaagtttttcttaggtgtg (nt. 1561-1584, exon 10)	50	606
- C9.8.8a	up (5'-3'): atgctccggctcataaat (nt. 1492-1511, exon 9) down (3'-5'): ccgcatattgactgctgac (nt. 1998-2017, exon 11)	55	526

<sup>(\*)</sup> The primer sequences followed by TI number, name and mate information collected from the pig genome sequencing project accessible via the Trace Archive tool (NCBI homepage)

<sup>(\*\*)</sup> according to GenBank acc. no. BI336392

<sup>(\*\*\*)</sup> according to GenBank acc. no. CF176130

Table 2 Oligonucleotide primers used for radiation hybrid mapping

Primer name	Primer sequence (genome localization)	Annealing tem- perature (°C)	Length (bp)
C6.4DNA	up 5'-3': ttctttttgcaaggatcaga (nt. 1437-1457, exon 10) down 3'-5': tcaatcacagcaggattttcc (nt. 1575-1595, exon 10)	58	159
C7.4.1DNA	up 5'-3': agttatcagttgttggttca (nt. 739-761, exon 8) down 3'-5': ctctcctaaggaccagac (nt. 915-931, exon 8)	50 <sup>(*)</sup>	196
C8Apro2DNA	up 5'-3': tgcttctggagggttctatt (clone xx-1c1) down 3'-5': cggttcaccttctctgtatg (clone xx-1c1)	60	707
C8B6DNA	up 5'-3': gaaacaagagaagcagcatgg (nt. 1302-1322, exon 9) down 3'-5': ttaattttgatgatgtctgggtg (nt. 1438-1461, exon 9)	60	160
C8G1DNA	up 5'-3': cctctgacgctgctct (nt. 75-92, exon 1) down 3'-5': gagccacgtgcagtgaagt (nt. 268-286, exon 2)	58	295
C9.2DNA	up 5'-3': ggagcattgagaccttgga (nt. 283-302, exon 3) down 3'-5': gccagctcagactctccac (nt. 528-547, exon 4)	60	511

<sup>(\*)</sup> A touch down PCR program was used to amplified a fragment of 196 bp under following thermal cycling conditions: 94°C for 4 min, followed by 9 cycles at 94°C for 30 sec, annealing temperature step-downs every 1 cycle of 1°C (from 60-51°C) for 30 sec, at 72°C for 1 min, then followed by 40 cycles of 94°C for 30 sec, at 50°C for 30 sec, at 72°C for 1 min, and ending with an extension step at 72°C for 5 min.

Table 3 List of restriction enzyme used for PCR-RFLP

Enzyme name	Specificity	(1)	(2)	(3)	Source
AciI	5'-C <sup>^</sup> CGC-3' 3'-GGC <sup>^</sup> G-5'	37	65	NEBuffer 3	<i>Arthrobacter citreus</i> (C. Polisson) (Biolabs)
AluI	5'-AG <sup>^</sup> CT-3' 3'-TC <sup>^</sup> GA-5'	37	65	Tango <sup>TM</sup>	<i>Arthrobacter luteus</i> (Fermentas)
BbvI (BseXI)	5'-GCAGC(N) <sub>8</sub> <sup>^</sup> -3' 3'-CGTCG(N) <sub>12</sub> <sup>^</sup> -5'	65	80	BseXI	<i>Bacillus stearothermophilus</i> Ra 3-212 (Fermentas)
BseDI	5'-C <sup>^</sup> CNNGG-3' 3'-GGNNC <sup>^</sup> C-5'	55	80	Tango <sup>TM</sup>	<i>Bacillus stearothermophilus</i> RFL1434 (Fermentas)
BsrDI (BseMI)	5'-GCAATGNN <sup>^</sup> -3' 3'-CGTTAC <sup>^</sup> NN -5'	55	80	R	<i>Bacillus stearothermophilus</i> Isl 15-111 (Fermentas)
FnuDII (Bsh1236I)	5'-CG <sup>^</sup> CG-3' 3'-GC <sup>^</sup> GC-5'	37	65	R	<i>Bacillus sphaericus</i> RFL1236 (Fermentas)
BsrI	5'-ACTGGN <sup>^</sup> -3' 3'-TGAC <sup>^</sup> CN -5'	65	80	NEBuffer 3	<i>Bacillus stearothermophilus</i> (C. Polisson) (Biolabs)
DdeI	5'-C <sup>^</sup> TNAG-3' 3'-GANT <sup>^</sup> C-5'	37	65	NEBuffer 3	A <i>E. coli</i> strain that carries the DdeI gene from <i>Desulfovibrio desulfuricans</i> (NCIB 83120) (Biolabs)
Eco91I	5'-G <sup>^</sup> GTNACC-3' 3'-CCANTG <sup>^</sup> G-5'	37	65	O <sup>+</sup>	<i>Escherichia coli</i> RFL91(Fermentas)
Hin6I	5'-G <sup>^</sup> CGC-3' 3'-CGC <sup>^</sup> G-5'	37	65	Tango <sup>TM</sup>	<i>Haemophilus influenzae</i> RFL6 (Fermentas)
HpaII	5'-C <sup>^</sup> CGG-3' 3'-GGC <sup>^</sup> C-5'	37	65	Tango <sup>TM</sup>	<i>Haemophilus parainfluenzae</i> (Fermentas)
HpyCH4III	5'-ACN <sup>^</sup> GT-3' 3'-TG <sup>^</sup> NCA-5'	37	80	NEBuffer 4	A <i>E. coli</i> strain that carries the HpyCH4III gene from <i>Helicobacter pylori</i> CH4 (S.A. Thompson) (Biolabs)
KpnI	5'-GGTAC <sup>^</sup> C-3' 3'-C <sup>^</sup> CATGG-5'	37	80	KnpI	<i>Klebsiella pneumoniae</i> OK8 (Fermentas)
MaeII (TaiI)	5'-ACGT <sup>^</sup> -3' 3'- <sup>^</sup> TGCA-5'	65	80	R	<i>Thermus aquaticus</i> Cc1-331 (Fermentas)
MboII	5'-GAAGA(N) <sub>8</sub> <sup>^</sup> -3' 3'-CTTCT(N) <sub>7</sub> <sup>^</sup> -5'	37	65	B	<i>E.coli</i> that carries the cloned <i>mboIIR</i> (Fermentas)
PstI	5'-CTGCA <sup>^</sup> G-3' 3'-G <sup>^</sup> ACGTC-5'	37	80	O	<i>Providencia stuarti</i> (Fermentas)
TaqI	5'-T <sup>^</sup> CGA-3' 3'-AGC <sup>^</sup> T-5'	65	80	TaqI	<i>Thermus aquaticus</i> YT-1 (Fermentas)

<sup>(1)</sup>\_Incubation temperature (°C), (2)\_Thermal inactivation in 20 min (°C), (3)\_Recommended buffer

Table 4 Oligonucleotide primers used to discriminate alleles for genotyping

Primer name	Primer sequence (genome localization) (*)	Gene: position of SNP in acc. no. as indicated in table 1	Annealing (°C)	Length (bp)
C6.2DNAmk	up 5'-3': ggagagcccagaggagaagt (nt. 748-767, exon 6) down 3'-5': cgagcctataatcagcattgg (nt. 566-586, ti: 1373546263; name: rplun0101_g6.y1)	C6: 862A→G	58	177
C6.6DNAmk	up 5'-3': tcctttttgcaaggatcagag (nt. 1438-1458, exon 10) down 3'-5': aacagcataaacaatggaggtg (nt. 754-775, ti: 1420506448; name:SS_WGS-957p20.p1k; mate:1420506449)	C6: 1557C→G	56	352
C7.1.BsrDI	up 5'-3': acaactgggtctctgggttc (ti: 848149583; name:bd_52849.z1; mate:863305860) down 3'-5': aacatctgaaacgctctccac (nt. 254-274, exon 4)	C7: 154A→G	56	301
C7.9DNAmk	up 5'-3': gagttatcagttggtgttcag (nt. 738-762, exon 8) down 3'-5': tgcgactttctctgagcca (nt. 956-975, exon 8)	C7: 870C→T	56	238
C7.4.1DNA	up 5'-3': agttatcagttggtgttc (nt. 859-881, exon 8) down 3'-5': ctctctctaaggaccagac (nt. 915-934, exon 8)	C7: 881A→G	51	196
C8A.5.Tfil	up 5'-3': cacctcgtgtgtaacggaga (nt. 434-453, exon 4) down 3'-5': gccaccagcgtatggtattt (clone xx-1c1)	C8A: 535A→G	56	322
C8A4DNAmk	up 5'-3': aagccatttacgagatactgc (nt.1490-1511, exon 10) down 3'-5': gtcaagtccctgtgggtt (nt. 1656-1673, exon 10)	C8A: 1544C→T	60	184
C8A.19.KnpI	up 5'-3': agcacagaggtcttgggtg (clone xx-1c1) down 3'-5': agctggcaggacagagaaaa (nt. 1969-1988, 3'UTR)	C8A: 1768C→T	56	426
C8B1DNA	up 5'-3': tgagaggccacactctcttg (nt. 160-179, exon 2) down 3'-5': ctctgacaggatcacacg (nt. 288-307, exon 2)	C8B: 222C→T	54 (**)	148
C8B13DNAmk	up 5'-3': caagagcacctgcttccaa (nt. 481-499, ti: 775597740; name: rbyc_22121.y1; mate:813000763) down 3'-5': cgagttccttcagagggtca (nt. 1009-1028, exon 7)	C8B: 935A→G→T	56	312
C8B.14.MaeII	up 5'-3': ttcttatcatcgggtctgtc (clone XX-1E1) down 3'-5': ctgccttctcttctctt (clone XX-1E1)	C8B: 1244A→G	56	528

Table 4 (continued)

Primer name	Primer sequence (genome localization) (*)	Gene: position of SNP in acc. no. as indicated in table 1	Annealing (°C)	Length (bp)
C9.2DNA	up 5'-3': ggagcattgagaccttgga (nt. 283-302, exon 3) down 3'-5': gccagctcagactcttccac (nt. 528-547, exon 4)	C9: 350A→G	60	511
C9.2bDNA	up 5'-3': gagccttgcaagaccttg (nt. 363-381, exon 3) down 3'-5': atagatgccccctttcac (intron detected)	C9: 407C→G	60	292

(\*)\_The primer sequences followed by TI numbers, name and mate information were collected from various large-scale sequencing projects of pig using the Trace Archive tool (NCBI homepage)

(\*\*)\_A touch down PCR program was used with initial denaturation 94°C for 4 min, followed by 8 cycles at 94°C for 30 sec, annealing temperature 58-54°C (step-downs of 0.5°C for each cycle), then followed by 30 cycles of 94°C for 30 sec, 54°C for 30 sec, 72°C for 1 min, and ending with an extension step at 72°C for 5 min.

## 4 Results

Information in this chapter is given in the following order: molecular characterization of the candidate genes and identification of polymorphisms, detection of genotypes and location of the candidate genes on the porcine chromosomes as well as genetic association of the genes to hemolytic complement activity in both the classical and alternative pathways.

### 4.1 Molecular characterization and detection of polymorphisms

#### 4.1.1 The sixth component of the porcine complement system (C6)

##### 4.1.1.1 Complementary DNA and predicted protein sequence

The porcine C6 gene was sequenced using six exon-primers as described in table 1. Full-length cDNA nucleotide sequence consists of 3306 bp including 144 and 354 nucleotides of 5' and 3' untranslated region, respectively (GenBank acc. no. DQ333199). Examination of the cDNA sequence revealed a single large open reading frame encoding 935 amino acids (GenBank acc. no. ABD13967), which commenced at the second available ATG codon. The porcine C6 protein sequence contains 64 cysteines. An analysis using the SignalP 3.0 tool has revealed a putative cleavage site after the signal peptide sequence between positions 21 and 22. The porcine C6 sequence has 18 exons predicted due to comparison with human intron-exon structure using the Map Viewer or alternatively, by collecting DNA genomic porcine sequences using the Trace Archive tool. The length and position of C6 exons are shown in table 22. Further detailed characterization of C6 molecule is in annex (Figure 40). The complete porcine C6 sequence shows higher homology ( $\geq 82\%$ ) to *B. taurus*, *C. familiaris*, *H. sapiens*, *P. troglodytes* than to other species at both DNA and protein levels (Table 5 and Figure 11). By using the SMART analysis tool it is possible to identify and delineate ten distinct structural units in the porcine C6 protein, which have the characteristic features of the modules TSP1, LDLa, MACPF, EGF, CCP and FIMAC. Hence, the C6 modules are expected to fold into stable structures, to exert a specific binding function in at least some cases, and

to be potentially encoded by a single exon (Haefliger et al. 1989). Location of the module domains in porcine C6 protein sequence is showed in figure 12.

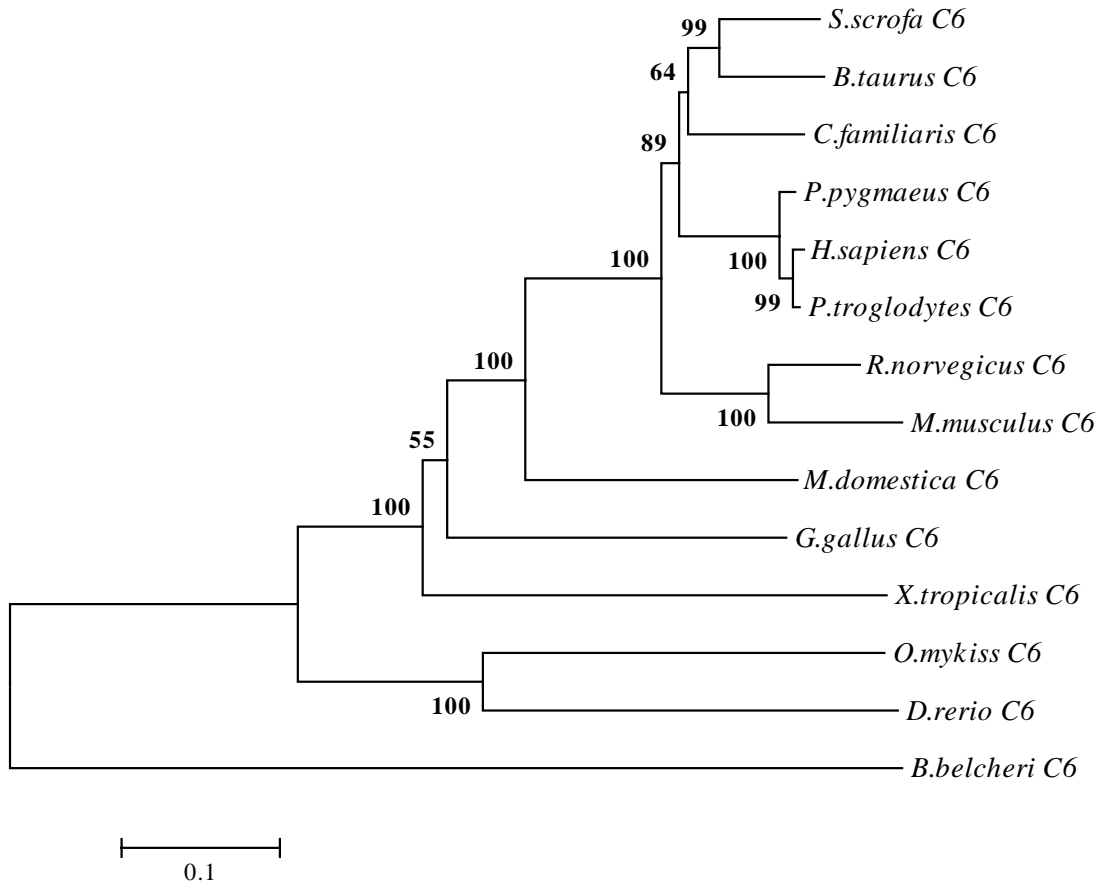


Figure 11 The phylogenetic tree of C6 proteins. An unrooted tree was constructed by the neighbor-joining method based on the alignment of amino acid sequences. It shows the phylogenetic relationship of C6 proteins among species. Bootstrap indices were used to verify the reliability of the branches. The numbers on interior branches refer to the bootstrap values with 500 replications; only values >50% are indicated. The numbers indicate the 'relative phylogenetic distance' between two branches of the tree. The relevant amino acid sequences were identified as GenBank accession numbers in table 5. The phylogenetic tree for evaluating evolutionary distances was computed using the MEGA3 software (Kumar et al. 2004)

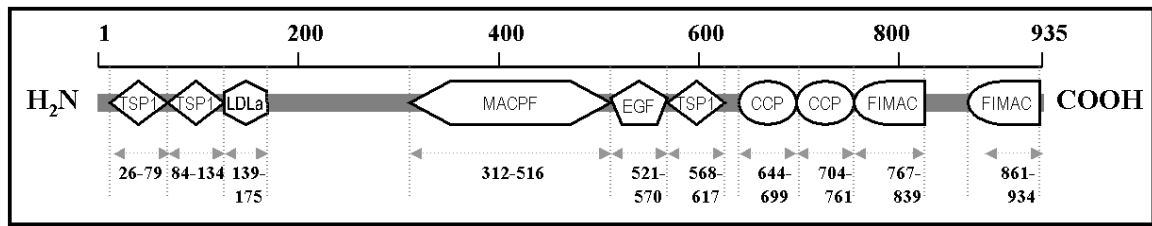


Figure 12 Location of protein domains in the porcine C6 sequence

Table 5 Percent nucleotide sequence and amino acid sequence identities of the porcine C6 with different species analysed by using ClustalW2 multiple sequence alignment

C6	Complementary DNA			Amino acids		
	GenBank acc. no	bp	%	GenBank acc. no	a.a	%
<i>S. scrofa</i>	DQ333199	3306	100	ABD13967	935	100
<i>C. familiaris</i>	AB126594	2805	85	XP_536488	936	83
<i>H. sapiens</i>	AB126593	2805	85	CAA50994	934	82
<i>B. taurus</i>	NM_001045979	3295	83	AAI14015	932	84
<i>X. tropicalis</i>	XM_536488	3166	82	AAH76972	934	53
<i>P. troglodytes</i>	NM_000065	3551	81	BAD02322	934	82
<i>P. pygmaeus</i>	NM_176074	3038	74	BAD02323	934	82
<i>R. norvegicus</i>	XM_001371501	2802	71	NP_788263	934	76
<i>M. musculus</i>	XM_429140	2808	65	AAH11251	769	74
<i>O. mykiss</i>	BC011251	2844	64	CAF22026	941	43
<i>M. domestica</i>	BC076972	3286	52	XP_001371538	933	65
<i>B. belcheri</i>	BC057429	3075	49	BAB47147	921	27
<i>G. gallus</i>	AJ622903	3216	39	XP_429140	935	58
<i>D. rerio</i>	AB050669	3508	27	AAH57429	885	45

#### 4.1.1.2 Single nucleotide polymorphisms

By sequencing an animal of each breed (HS, DR, LR, PIE, BMP, and MK), five single nucleotide polymorphisms (SNPs) were found in the coding region of the porcine C6 gene, respectively. The SNPs were determined by PCR-RFLP method using the restriction enzymes *AciI*, *TaqI*, *PstI*, *BsrI* and *DdeI*. Both the nucleotide variations at nt. 1138C→T and 1557C→G belong to the MACPF domain. Two SNPs at nt. 862A→G and 1557C→G have amino acid substitutions asparagines→asparagine acid and asparagine acid→glutamine acid, respectively. The fifth appears within the FIMAC module of amino acid sequence. Point mutations at position 843 and 1557, respectively, are transversion A→C and C→G while the others are transitions (Table 6).

Table 6 Characterization of SNPs in the porcine C6 candidate gene

Variation, bp	Variation, a.a	Codon	Exon	Enzyme	Module
843A→C	-	233CCA→CCC	6	<i>AciI</i>	-
862A→G	Asn →Asp	240AAC→GAC	6	<i>TaqI</i>	-
1138C→T	-	332CTG→TTG	8	<i>PstI</i>	MACPF
1557C→G	Asp→Glu	471GAC→GAG	10	<i>BsrI</i>	MACPF
2604C→T	-	820GCC→GCT	17	<i>DdeI</i>	FIMAC

#### 4.1.2 The seventh component of the porcine complement system (C7)

##### 4.1.2.1 Complementary DNA and predicted protein sequence

The 2532 bp full-length porcine C7 cDNA sequence (GenBank acc. no. AF162274) and an open reading frame encoding 843 amino acid residues (GenBank acc. no. NP\_999447) have been reported by Agah et al. (2000). The protein C7 precursor contains 56 cysteine residues. Here we report 807 (nt. -806→0) and 122 (nt. 2533-2654) nucleotides of new sequences in the 5' and 3' flanking region without any putative polyadenylation signal, respectively. The transcriptional initiation site was assigned to position -110 bp upstream of the start codon predicted by the DRAGON



GC+PROMOTER FINDER version 1.0 (<http://sdmc.lit.org.sg/ERE-V2/index>) (Figure 41). The first 28 amino acids represent the signal peptide as predicted using the SignalP 3.0 tool. The signal peptide predictions from both the SMART and the SignalP 3.0 tool are identical. This means that the residues numbered from 29 to 843 encode the mature porcine C7. Eighteen exons of the pig C7 gene were identified due to blasting the porcine and human genomic C7 sequence by using the Trace Archive and the Map Viewer tool, respectively. According to figure 41, the length and position of exon fragments are displayed in table 23. The porcine C7 sequence shows higher homology ( $\geq 83\%$ ) to *B. taurus* and *H. sapiens* than other species concerning nucleotide and protein sequences (Table 7 and Figure 13). In porcine C7 gene five classes of the modules TSP1, LDLa, MACPF, CCP, FIMAC were distinguished using the SMART analysis tool (Figure 14).

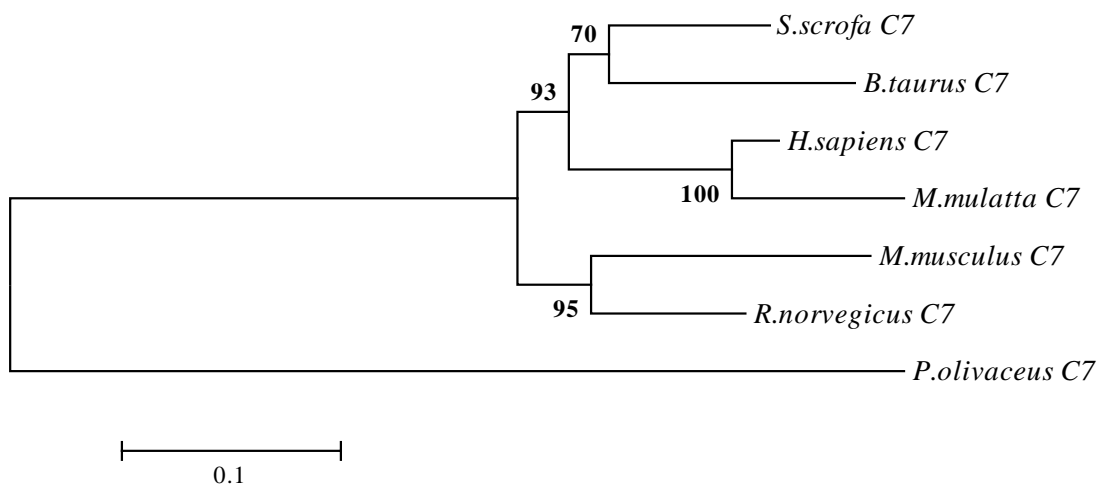


Figure 13 The phylogenetic tree of C7 proteins. An unrooted tree was constructed by the neighbor-joining method based on the alignment of amino acid sequences. It shows the phylogenetic relationship of C7 proteins among species. Bootstrap indices were used to verify the reliability of the branches. The numbers on interior branches refer to the bootstrap values with 500 replications; only values  $>50\%$  were indicated. The numbers indicate the 'relative phylogenetic distance' between two branches of the tree. The relevant amino acid sequences were identified as GenBank accession numbers in table 7. The phylogenetic tree for evaluating evolutionary distances was computed using the MEGA3 software (Kumar et al. 2004).

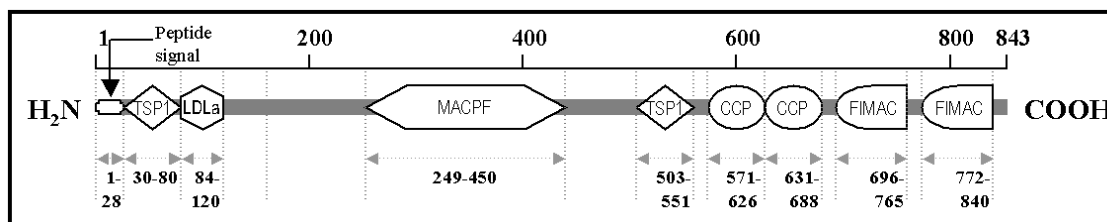


Figure 14 Location of protein domains in the porcine C7 sequence

Table 7 Percent nucleotide sequence and amino acid sequence identities of the porcine C7 with different species analyzed by using ClustalW2 multiple sequence alignment

C7	Complementary DNA			Amino acids		
	GenBank acc. no	bp	%	GenBank acc. no	a.a	%
<i>S. scrofa</i>	AF162274	2532	100	AAD45918	843	100
<i>B. taurus</i>	BC114076	3709	87	AAI14077	843	83
<i>H. sapiens</i>	BC063851	4015	86	AAH63851	843	83
<i>R. norvegicus</i>	XM_001054007	4727	80	XP_001054007	844	78
<i>M. mulatta</i>	XM_001085533	1659	79	XP_001085533	552	76
<i>M. musculus</i>	XM_973486	2547	77	XP_978580	848	72
<i>O. mykiss</i>	AJ622902	2777	54	CAF22025	845	43
<i>P. olivaceus</i>	AB020964	2618	39	BAA88899	805	43
<i>D. rerio</i>	XM_685854	2532	54	XP_690946	849	41

#### 4.1.2.2 Single nucleotide polymorphisms

Five primer pairs were designed along the body of the porcine C7 cDNA sequence to screen polymorphisms. Using PCR-RFLP method with digestion of specific restriction enzymes, which separate DNA fragments at the sites with different alleles, six polymorphic sites were detected. All they are transitions. The polymorphisms at nt 546, 870 and 1764 are silent mutations whereas the other remainders have amino acid substitutions. Five of six polymorphisms belong to the functional protein domains such as TSP1, MACPF and CCP (Table 8).

Table 8 Characterization of SNPs in the porcine C7 candidate gene

Variation, bp	Variation, a.a	Codon	Exon	Enzyme	Module
154A→G	Ile→Val	52ATT→GTT	4	BsrDI	TSP1
200C→T	Thr→Met	67ACG→ATG	4	MaeII	TSP1
546C→T	-	182AAC→AAT	6	MaeII	-
870C→T	-	290AGC→AGT	8	Hin6I	MACPF
881A→G	Lys→Arg	294AAA→AGA	8	MboII	MACPF
1764A→G	-	588ACA→ACG	17	MaeII	CCP

#### 4.1.3 The alpha chain of the eighth component of the porcine complement system (C8A)

##### 4.1.3.1 Complementary DNA and predicted protein sequence

The nucleotide sequence coding for the alpha-eighth component of porcine complement factors (C8A) has been determined by sequencing PCR products from five primer pairs, which amplified overlapping fragments of 490 to 572 nucleotides in length (Table 1). The full-length cDNA sequence consists of 2146 nucleotides (GenBank acc. no. DQ333200) encoding 589 amino acids in a single polypeptide chain (GenBank acc. no. ABD13968). Nakajima et al. (1998) reported the leader peptide sequence formed by the first 31 amino acids in the porcine C8A protein sequence. Validation in length and position of 11 porcine C8A exons was conducted by comparing with porcine and human genomic sequences using the Trace Archive and the Map Viewer tool, respectively (Table 21). A part of the first exon belongs to the 5' flanking region while the ending segment of the last exon is a part of the 3' flanking region. Further characterization of C8A sequence is found in figure 42. The porcine C8A sequence shows more homology with *B. taurus*, *C. familiaris*, *M. mulatta*, *H. sapiens*, *P. troglodytes* than with other species at both nucleotide and protein levels ( $\geq 77\%$ ) (Table 9 and Figure 16). The polypeptide chain of the gene encoding the porcine C8A complement component contains 31 cysteines. Small cysteine-rich modules in the porcine C8A protein are highly homologous with TSP1, LDLa, and MACPF modules using the SMART analysis tool (Figure 15).

#### 4.1.3.2 Single nucleotide polymorphisms

The same primer pairs were used to identify porcine C8A cDNA sequence as well as to screen polymorphisms, which were confirmed via restriction digestion by specific enzymes (Table 1). Results showed that seven SNPs were detected in the porcine C8A gene. Most SNPs were silent mutations except two of them at sites 1544C→T and 1674A→G leading to amino acid exchanges. The protein domains containing polymorphisms were TSP1 (358C→T, 1768C→T), and MACPF (1207C→T, 1510A→G, 1544C→T). At the nucleotide position 1545, variation of alleles G→A leading to amino acid substitution Arg→His at the second position of codon 485 (CGC→CAC) was confirmed by restriction digestion with HhaI enzyme (Nakajima et al. 1998). Further characterization of point mutations is given in table 10.

Table 9 Percent nucleotide sequence and amino acid sequence identities of the porcine C8A with different species analyzed by using ClustalW2 multiple sequence alignment

C8A	Complementary DNA			Amino acids		
	GenBank acc. no	bp	%	GenBank acc. no	a.a	%
<i>S. scrofa</i>	DQ333200	2146	100	ABD13968	589	100
<i>H. sapiens</i>	BC132913	1915	83	AAI32914	584	78
<i>C. familiaris</i>	XM_536695	1816	83	XP_536695	589	77
<i>B. taurus</i>	BC112635	2316	82	AAI12636	589	80
<i>P. troglodytes</i>	XM_513435	2392	81	XP_513435	584	78
<i>M. mulatta</i>	XM_001114403	2393	80	XP_001114403	584	77
<i>O. cuniculus</i>	NM_001082255	2070	72	NP_001075724	585	74
<i>R. norvegicus</i>	XM_001064004	3133	66	XP_001064004	587	71
<i>M. musculus</i>	BC027748	3665	65	CAM23655	587	72
<i>X. tropicalis</i>	BC074554	1973	50	AAH74554	584	48
<i>G. gallus</i>	XM_426667	2227	49	XP_426667	603	52
<i>O. mykiss</i>	AJ851931	2037	44	CAH65481	615	43

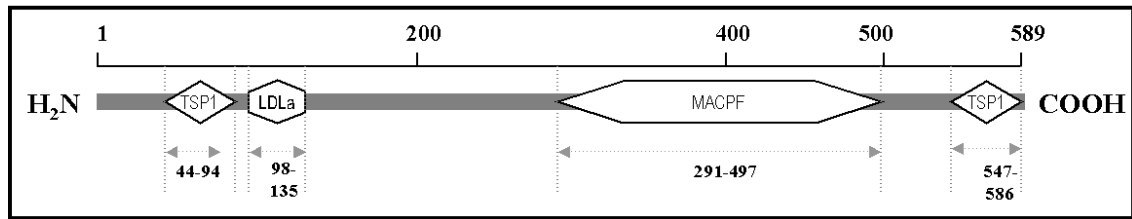


Figure 15 Location of protein domains in the porcine C8A sequence

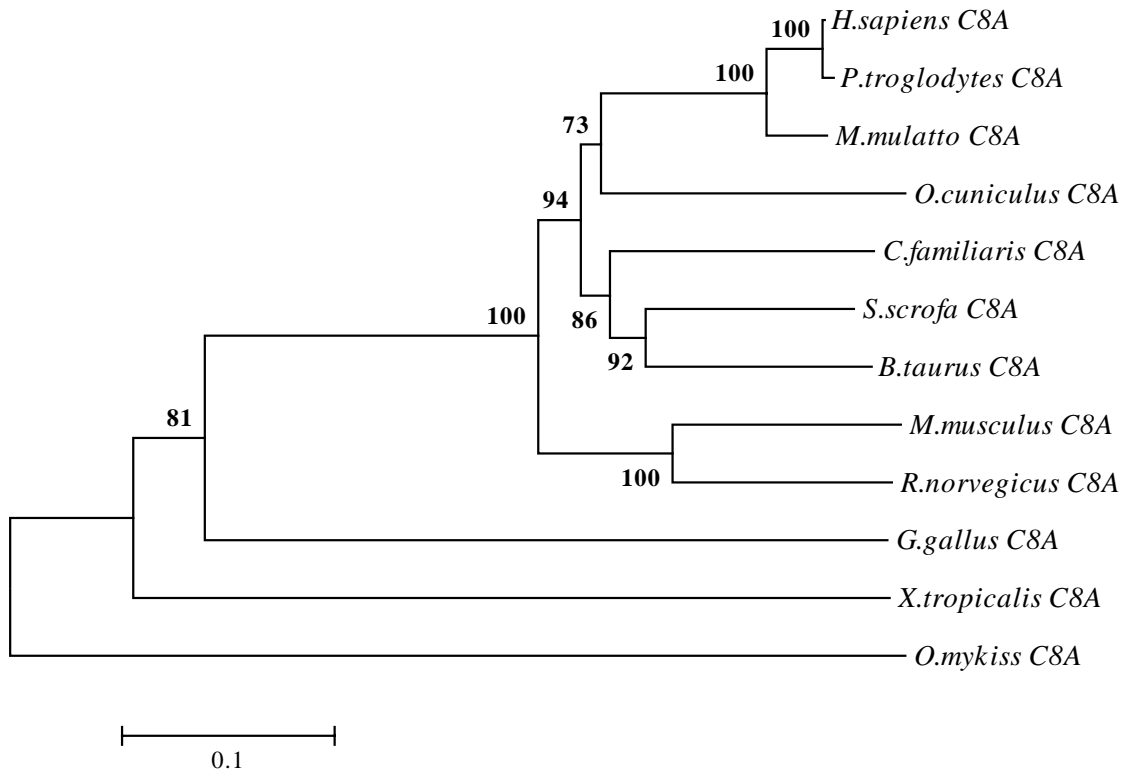


Figure 16 The phylogenetic tree of C8A proteins. An unrooted tree was constructed by the neighbor-joining method based on the alignment of amino acid sequences. It shows the phylogenetic relationship of C8A proteins among species. Bootstrap indexes were used to verify the reliability of the branches. The numbers on interior branches refer to the bootstrap values with 500 replications; only values >50% were indicated. The numbers indicate the 'relative phylogenetic distance' between two branches of the tree. The relevant amino acid sequences were identified as GenBank accession numbers in table 9. The phylogenetic tree for evaluating evolutionary distances was computed using the MEGA3 software (Kumar et al. 2004)

Table 10 Characterization of SNPs in the porcine C8A candidate gene

Variation, bp	Variation, a.a	Codon	Exon	Enzyme	Module
358C→T	-	89AGC→AGT	3	HpyCH4III	TSP1
535A→G	-	148CCA→CCG	4	TfiI	-
1207C→T	-	372AGC→AGT	8	BsrI	MACPF
1510A→G	-	473CTA→CTG	10	Hin6I	MACPF
1544C→T	Arg→Cys	485CGC→TGC	10	Hin6I	MACPF
1674A→G	Tyr→Cys	528TAC→TGC	10	Eco91I	-
1768C→T	-	559GGC→GGT	11	KpnI	TSP1

#### 4.1.4 The beta chain of the eighth component of the porcine complement system (C8B)

##### 4.1.4.1 Complementary DNA and predicted protein sequence

The complete porcine C8B cDNA sequence spans 1987 nucleotides, including 28 and 123 nucleotides of 5' and 3' untranslated region, respectively (GenBank acc. no. DQ333201). The predicted protein sequence of the gene is 661 amino acids (GenBank acc. no. ABD13969), which contains ten amino acids more than human protein. A total of 27 cysteines were found in protein structure. Figure 43 shows further information for porcine C8B characterization. The signal peptide was identified between amino acids the first 44 and 45 using the SignalP 3.0 tool. Nucleotide ( $\geq 80\%$ ) and protein ( $\geq 75\%$ ) sequence show higher similarity with *C. familiaris*, *H. sapiens*, *M. mulatta*, *O. cuniculus* than with other species (Table 11 and Figure 17). Alignment of the porcine C8B cDNA sequence along with collected porcine and human DNAs using the Trace Archive and the Map Viewer tool, respectively, has revealed length and position of twelve exons in the porcine C8B (Table 24). Using the SMART analysis tool it is possible to identify and delineate four distinct structural units in the C8B protein sequence, which have the characteristic features of the modules TSP1, LDLa, MACPF, and EGF (Figure 18).

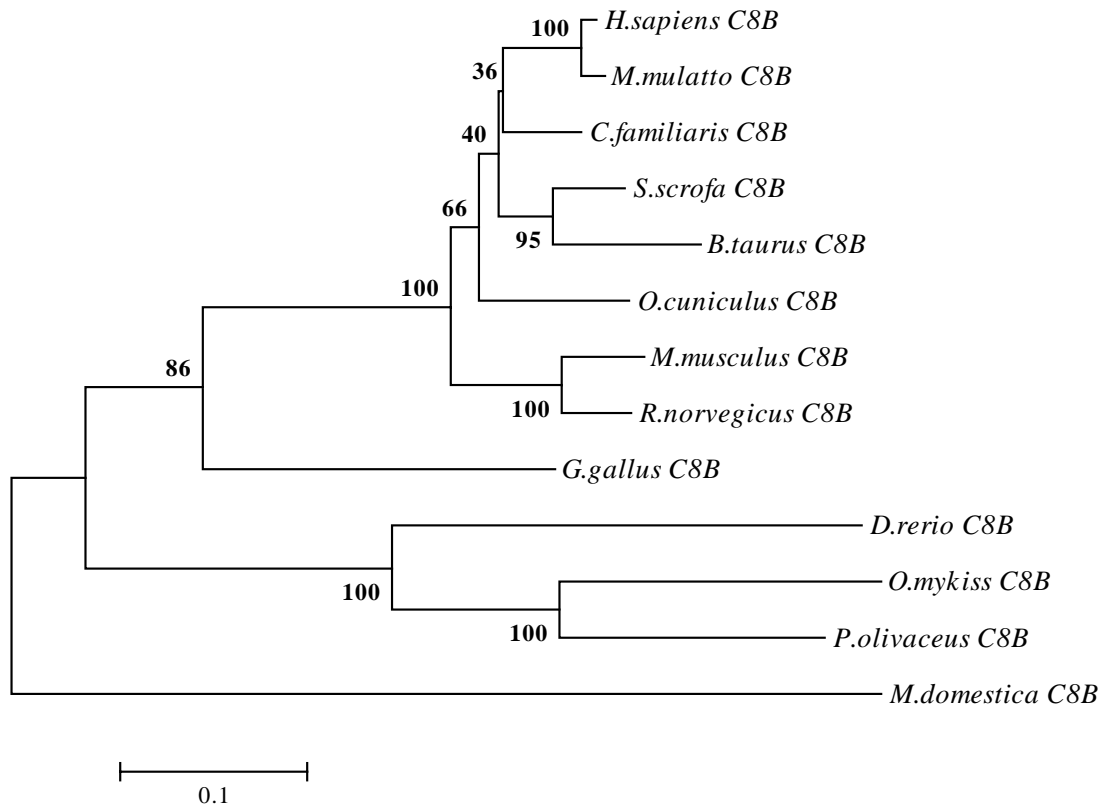


Figure 17 The phylogenetic tree of C8B proteins. An unrooted tree was constructed by the neighbor-joining method based on the alignment of amino acid sequences. It shows the phylogenetic relationship of C8B proteins among species. Bootstrap indexes were used to verify the reliability of the branches. The numbers on interior branches refer to the bootstrap values with 500 replications. The numbers indicate the 'relative phylogenetic distance' between two branches of the tree. The relevant amino acid sequences were identified as GenBank accession numbers in table 11. The phylogenetic tree for evaluating evolutionary distances was computed using the MEGA3 software (Kumar et al. 2004)

#### 4.1.4.2 Single nucleotide polymorphisms

Here nine polymorphisms were detected in porcine C8B gene. Most of them were identified using PCR-RFLP analysis except the polymorphic site at nucleotide position 935 with three different alleles using repeated comparative sequencing. The polymorphic

site at locus 935 with three different alleles was not only a transition but also a transversion while all remainders were transitions. Eight of these nine contain one amino acid exchange. Five of nine SNPs belong to the MACPF protein domain (Table 12).

Table 11 Percent nucleotide sequence and amino acid sequence identities of the porcine C8B with different species analysed by using ClustalW2 multiple sequence alignment

C8B	Complementary DNA			Amino acids		
	GenBank acc. no	bp	%	GenBank acc. no	a.a	%
<i>S. scrofa</i>	DQ333201	1987	100	ABD13969	611	100
<i>C. familiaris</i>	XM_536694	1773	87	XP_536694	590	79
<i>H. sapiens</i>	NM_000066	2053	84	AAI30576	591	79
<i>M. mulatta</i>	XM_001114456	2082	83	XP_001114456	591	78
<i>O. cuniculus</i>	NM_001082668	2052	80	NP_001076137	590	75
<i>B. taurus</i>	BC112600	1330	79	AAI12601	383	85
<i>M. musculus</i>	AB077306	1586	79	BAC41371	523	71
<i>R. norvegicus</i>	XM_001058539	2395	73	XP_001058539	589	73
<i>G. gallus</i>	XM_422502	2165	59	XP_422502	584	57
<i>D. rerio</i>	XM_001332783	1742	53	XP_001332819	562	44
<i>O. mykiss</i>	AF418597	1855	51	AAL16647	587	43
<i>P. olivaceus</i>	AB020962	1849	50	BAA86877	588	43
<i>M. domestica</i>	XM_001381244	1896	44	XP_001381281	631	34

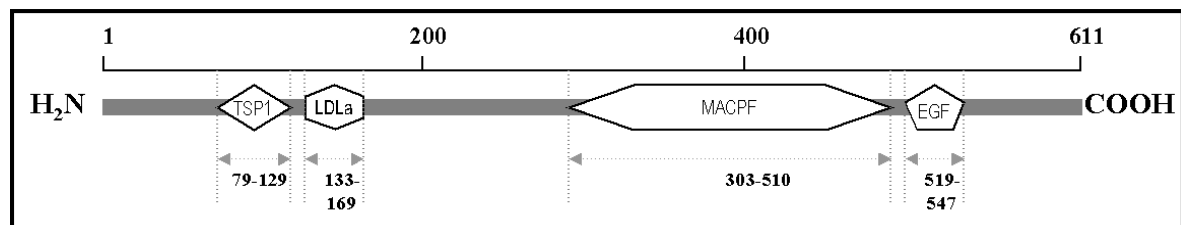


Figure 18 Location of protein domains in the porcine C8B sequence



Table 12 Characterization of SNPs in the porcine C8B candidate gene

Variation, bp	Variation, a.a	Codon	Exon	Enzyme	Module
99C→T	Pro→Leu	24CCG→CTG	1	HpaII	-
222C→T	Thr→Met	65ACG→ATG	2	FnuDII	-
935A→G→T	Thr→Gly→Ser	303ACA→GCA→TCA	7	(*)	MACPF
1244A→G	Ile→Val	406ATC→GTC	8	MaeII	MACPF
1259C→T	Pro→Ser	411CCG→GCG	8	HpaII	MACPF
1374C→T	Ala→Val	449GCC→GTC	9	BseDI	MACPF
1494C→T	Ala→Val	489GCG→GTG	10	AciI	MACPF
1797C→T	Ala→Val	590GCA→GTA	12	BbvI	-
1801C→T	-	591GCC→GTC	12	AluI	-

(\*)\_detected by resequencing

#### 4.1.5 The gamma chain of the eighth component of the porcine complement system (C8G)

##### 4.1.5.1 Complementary DNA and predicted protein sequence

The gene encoding the porcine C8G gene was sequenced and its structure was determined. The complete porcine C8G cDNA sequence (GenBank acc. no. DQ333202) contains 840 bp including 51 and 180 bp of the 5' and 3' UTR regions, respectively. The entire coding region is composed of 609 bp corresponding to an open reading frame encoding 202 amino acids (GenBank acc. no. ABD13970). Both human and pig C8G proteins contains three cysteine residues. Seven exons of the porcine C8G were identified based on alignment along length of porcine AK233484 genomic DNA sequence in GenBank (clone: LVRM10157D04, Uenishi et al. 2004) according to intron-exon junction rule (start with GT – end with AG) and length of human C8G DNA genomic sequence using the Map Viewer sequence alignment tool (Table 25). Further information about porcine C8G sequence is found in figure 44. The SignalP 3.0 analysis revealed, that the first exon contains the 5' UTR together with the signal peptide coding sequences between amino acids 20-21. The porcine C8G cDNA and polypeptide precursor sequence show a high degree of  $\geq 65\%$  identity with *M. musculus*, *O. cuniculus*, *C. familiaris* and *H. sapiens* sequences (Table 13 and Figure 20). A sequence, which is

similar to lipocalin module was found between amino acids 48-184 using Conserved Domains Search CDS/CCD protein domain analysis tool (Figure 19).

#### 4.1.5.2 Single nucleotide polymorphisms

The porcine C8G was screened for polymorphisms by comparative sequencing but no polymorphism was found in its cDNA sequence.

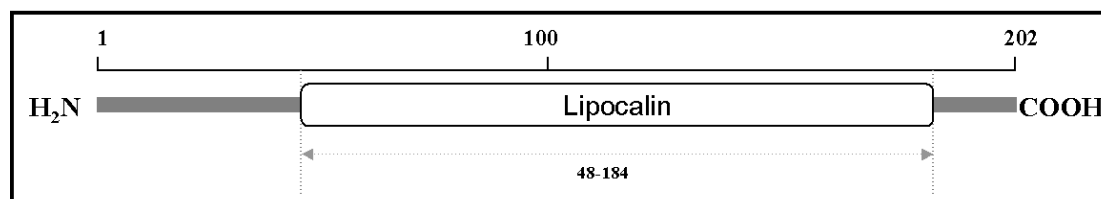


Figure 19 Location of protein domains in the porcine C8G sequence

Table 13 Percent nucleotide sequence and amino acid sequence identities of the porcine C8G with different species analyzed by using ClustalW2 multiple sequence alignment

C8G	Complementary DNA			Amino acids		
	GenBank acc. no	bp	%	GenBank acc. no	a.a	%
<i>S. scrofa</i>	DQ333202	840	100	ABD13970	202	100
<i>M. musculus</i>	BC019967	1063	71	AAH19967	202	71
<i>C. familiaris</i>	XM_537791	769	66	XP_537791	202	80
<i>O. cuniculus</i>	L26979	751	66	AAA31193	202	74
<i>H. sapiens</i>	M17999	857	65	AAI13627	202	82
<i>G. gallus</i>	XM_001234026	684	59	XP_001234027	225	40
<i>R. norvegicus</i>	XM_001073643	1212	55	XP_001073643	202	73
<i>M. mulatta</i>	XM_001117839	417	53	XP_001117839	138	73
<i>O. mykiss</i>	AJ622904	977	34	CAF22027	221	32
<i>D. rerio</i>	BC062865	1364	21	AAH62865	209	32

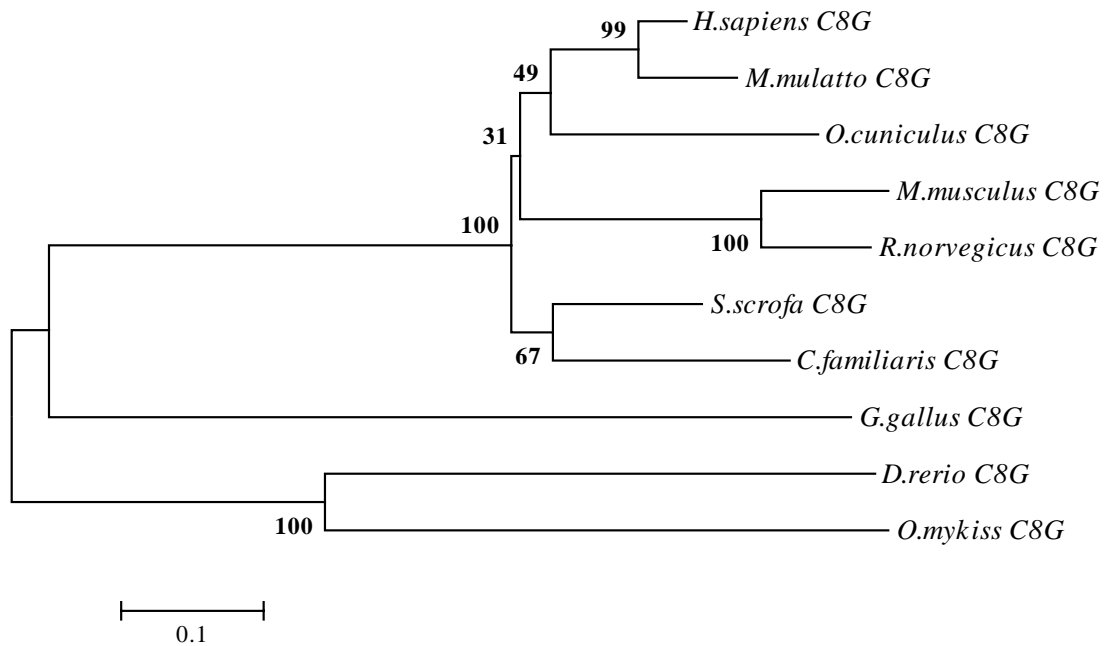


Figure 20 The phylogenetic tree of C8G proteins. An unrooted tree was constructed by the neighbor-joining method based on the alignment of amino acid sequences. It shows the phylogenetic relationship of C8G proteins among species. Bootstrap indexes were used to verify the reliability of the branches. The numbers on interior branches refer to the bootstrap values with 500 replications. The numbers indicate the 'relative phylogenetic distance' between two branches of the tree. The relevant amino acid sequences were identified as GenBank accession numbers in table 13. The phylogenetic tree for evaluating evolutionary distances were computed using the MEGA3 software (Kumar et al. 2004)

#### 4.1.6 The ninth component of the porcine complement system (C9)

##### 4.1.6.1 Complementary DNA and predicted protein sequence

The cDNA length of the porcine C9 candidate gene was obtained due to amplifying and sequencing PCR product fragments derived from four primer pairs given in table 1. Sequencing results established a continuous cDNA sequence of 2536 bp in length (GenBank acc. no. DQ333198) including 1632 bp of coding sequence. The 5' and 3' UTR contain 89 and 815 bp, respectively. The 3' UTR has a polyadenylation signal sequence

(ATTAAA) between nucleotides 2505-2510. Alignment to porcine and human C9 sequence using the Trace Archive and the Map Viewer tool, respectively, confirmed the length and position of 11 exon fragments in the porcine C9 as shown in table 26. Computer prediction revealed the first 21 amino acids in C9 protein sequence as being a signal peptide sequence using the SignalP 3.0 tool. Examination of the cDNA sequence revealed a single large open reading frame of 543 amino acids with 19 cysteine residues. The open reading frame is terminated by a TAG stop codon after nucleotide position 1718. The protein sequence of porcine C9 shows high similarity to that of *E. caballus*, *B. taurus*, *M. mulatta*, *H. sapiens* ( $\geq 67\%$ ) according to the ClustalW2 alignment (Table 14 and Figure 22). Using the SMART analysis tool we have detected porcine protein structure characterized by modules TSP1, LDLa and MACPF (Figure 21). Further characterization for porcine C9 sequence is showed in figure 45.

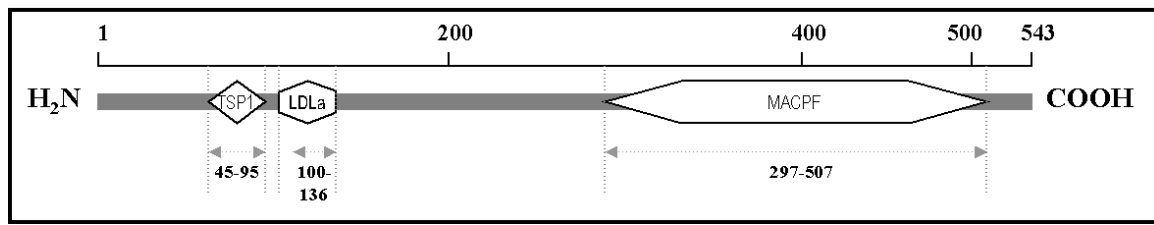


Figure 21 Location of protein domains in the porcine C9 sequence

#### 4.1.6.2 Single nucleotide polymorphisms

Two segments of the porcine C9 gene encompassing the presumably polymorphic codons from comparative sequencing were amplified using polymerase chain reaction. According to the restriction fragment patterns obtained after BsrDI and HpyCH4III digestion of the PCR products, two polymorphisms were detected at nucleotide 350A→G and 407C→G corresponding to codon 87CAA→CAG and 106CAC→CAG. The first SNP belongs to the TSP1 module and the second locates on the LDLa module. The second polymorphism is a nucleotide substitution C→A in the third position of codon 106 causing an amino acid exchange His→Gln, whereas the remainder is a silent mutation. Both located on the same exon 3.

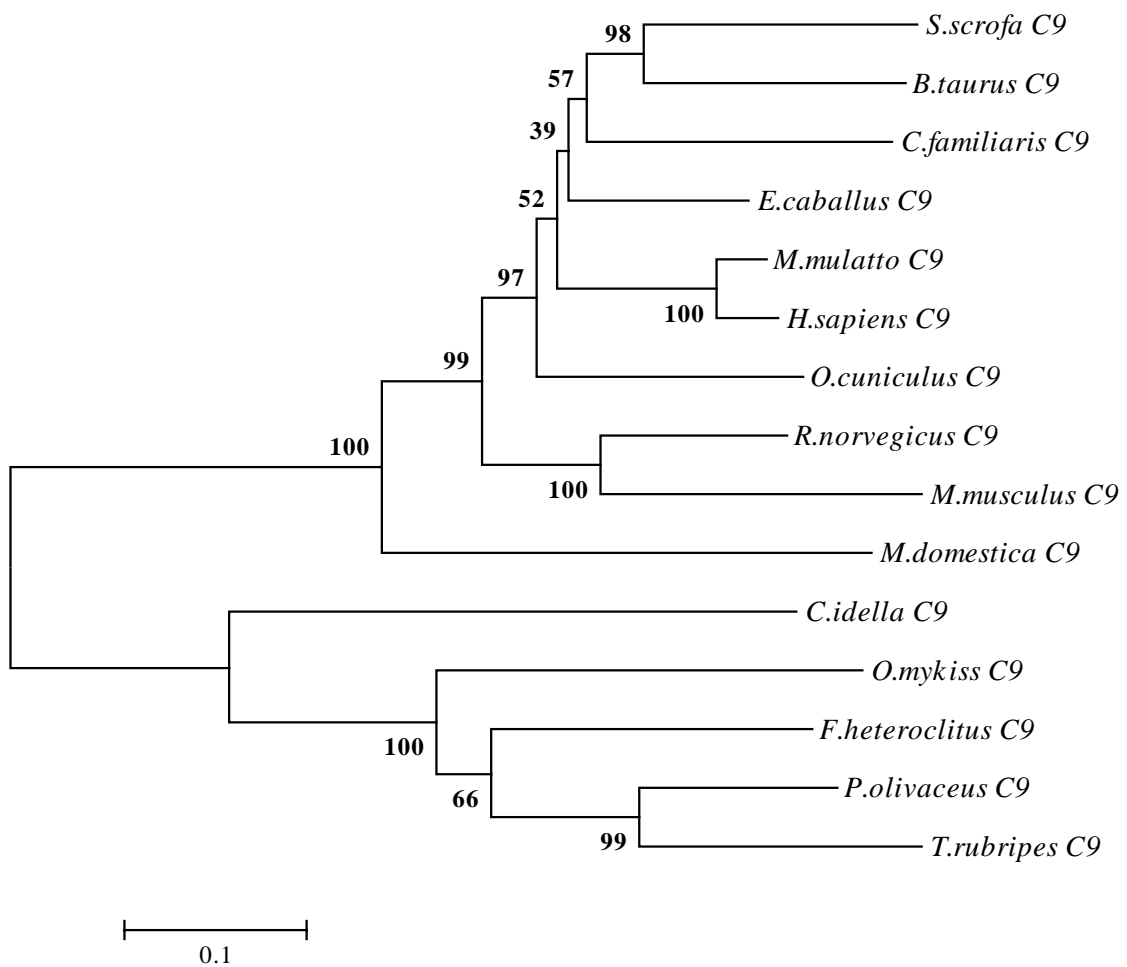


Figure 22 The phylogenetic tree of C9 proteins. An unrooted tree was constructed by the neighbor-joining method based on the alignment of amino acid sequences. It shows the phylogenetic relationship of C9 proteins among species. Bootstrap indexes were used to verify the reliability of the branches. The numbers on interior branches refer to the bootstrap values with 500 replications. The numbers indicate the 'relative phylogenetic distance' between two branches of the tree. The relevant amino acid sequences were identified as GenBank accession numbers in table 14. The phylogenetic tree for evaluating evolutionary distances were computed using the MEGA3 software (Kumar et al. 2004)

Table 14 Percent nucleotide sequence and amino acid sequence identities of the porcine C9 with different species analysed by using ClustalW2 multiple sequence alignment

C9	Complementary DNA			Amino acids		
	GenBank acc. no	bp	%	GenBank acc. no	a.a	%
<i>S. scrofa</i>	DQ333198	2536	100	ABD13966	543	100
<i>E. caballus</i>	NM_001081950	2661	79	NP_001075419	547	71
<i>B. taurus</i>	BC105174	2697	78	AAI05175	548	70
<i>H. sapiens</i>	BC020721	2094	78	AAH20721	559	67
<i>O. cuniculus</i>	NM_001082346	2018	77	NP_001075815	557	66
<i>M. mulatta</i>	XM_001084671	2774	73	XP_001084671	561	69
<i>C. familiaris</i>	XM_536494	1840	72	XP_536494	589	65
<i>R. norvegicus</i>	NM_057146	2083	66	NP_476487	567	61
<i>M. domestica</i>	XM_001372031	1572	66	XP_001372068	523	52
<i>M. musculus</i>	BC011137	1767	64	AAH11137	561	55
<i>F. heteroclitus</i>	AY495671	1844	45	AAR87007	577	35
<i>C. idella</i>	AY391781	2123	41	AAS76086	650	35
<i>O. mykiss</i>	AM039888	2104	40	CAJ01692	601	33
<i>P. olivaceus</i>	AB020963	2313	31	BAA86878	558	33
<i>T. rubripes</i>	FRU87241	2854	5	AAC60288	586	32

#### 4.2 Structural relationship analysis among the porcine homologous complement candidate sequences

From above obtained results, it is very possible to recognize that there is a relationship of functional domains in the protein sequences of the porcine terminal complement components. The amino acid sequences of the candidate gene are closely related to each other with respect to the percentage of residue identities ranging from 22-30% as revealed by ClustalW2 alignment tool (Figure 23). The porcine C8G protein with only 202 amino acids is a lipocalin. It was, therefore, not referred in this content.

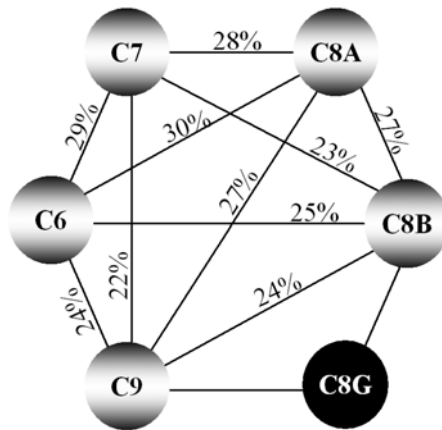


Figure 23 Similarity among amino acid sequences of the candidate genes shown by using ClustalW2 multiple sequence alignment

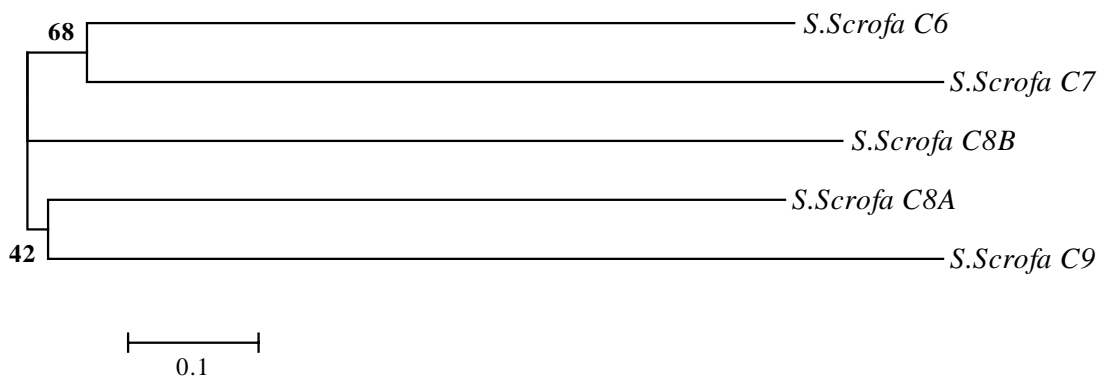


Figure 24 The phylogenetic tree of homologous proteins of the terminal lytic pathway. An unrooted tree was constructed by the neighbor-joining method based on the alignment of amino acid sequences. It shows the phylogenetic relationship among the terminal complement components. Bootstrap indexes were used to verify the reliability of the branches. The numbers on interior branches refer to the bootstrap values with 500 replications. The numbers indicate the 'relative phylogenetic distance' between two branches of the tree. The relevant amino acid sequences were identified as GenBank accession numbers ABD13967 (C6), NP\_999447 (C7), ABD13968 (C8A), ABD13969 (C8B) and ABD13966 (C9). The phylogenetic tree for evaluating evolutionary distances were computed using the MEGA3 software (Kumar et al. 2004)

Data in figure 24 re-emphasizes the great similarity between C6 and C7 while C9 has a structure more similar to C8A than to C8B. The protein sequences of the candidate

genes contain segments which are homologous to those of other proteins described as TSP1, LDLa, MACPF, EFG, CCP and FIMAC. Most of these modules are cysteine-rich sequences. The amino acid sequence of CCP module contains four cysteines in which two of them are at the first and last position of sequence. There are 58, 45, 17, 18, three, 13 of 64, 56, 31, 27, three, 19 cysteines found in C6, C7, C8A, C8B, C8G and C9 that located in small functional protein domains, respectively (Table 15 and Figure 25).

#### Type I protein segment (similar to TSP1 module)

```

C8A (a. a 44-94)   LSNWAEWTEC-FPCQDKKYRYSLLQPNKFG-GAICS--GNVWDQASCHSPTECL 51
C8B (a. a 79-129) LSSWSSWTTC-DPCQKKRYRHASLLRPSQFH-GEPNCN--FSDKEVEDCVSNRPCR 51
C6 (a. a 84-134)  LGDYGPWSDC-DPCVEKQFKVRSILRPNQFG-GQPCT--EPLMTFRPCIPSKLCK 51
C9 (a. a 45-95)   MSPWGEWSRC-DPCQKQMFRRSRIETFGQFN-GQKCV--DAVGDRRQCVPTEPE 51
C7 (a. a 30-80)   WDSYAPWSEC-NGCTKTQTRRRPVAVYGYG-GHPCV--GSTFETQSCEPTRGCP 51
C6 (a. a 568-617) WGWSSWSTC-DATYKRRTRENNPAPRQG-GKPCD--GERRQEEHCTFSIMQ- 50
C8A (a. a 547-586) WGWSSWSAC-RLG-TQERRRENNPAPQNG-GASCS--GHKVQT----- 40
C7 (a. a 503-551) WGWSSWGPC-AQG-KKTRSRKNNPPPSGG-GKSCI--GETSESRQCEDEDLE- 49
C6 (a. a 26-79)   HYPWTQWSSCSKTCNSGTQTRQRQITINQYYLDNFCDRLCITKQETRECN-WQTC 54

```

#### Type II protein segment (similar to LDLa module)

```

C6 (a. a 139-175) DQKNKFRFD-SGRCTIASKLEENGENDCG-DNSDERNCGR 37
C8B (a. a 249-450) RCEGFVCAQ-TGRCTINRRLLCNGDNDCG-DQSDLEANCKR 37
C8A (a. a 98-135)  QCGQDFQCKETGRCLKRHLVINGDRDCL-DGSDDEDDCED 38
C9 (a. a 133-169) DCGSDFQCG-TGRCTIKRRLLCNGDNDCG-DFSDEDDCDS 37
C7 (a. a 100-136) GGERFRF-SGQCISKSLVINGDSDCEEDSAEDRCE- 37

```

#### Type III protein segment (similar to EGF module)

```

C6 (a. a 512-570) CQAPCPNNGRPVLSGTECLCVQSGTYGDNERRAPDYKSNAVDGN 47
C8B (a. a 519-547) ----PCQNGVPLKESRDCICPAGFQGSACE----- 29

```

#### Type IV protein segment (similar to CCP module)

```

C6 (a. a 644-699) CPQPVPPENGFIRNEKTLYSVGEDVEIVCLTGFKTVGYQYFRCLPDRTW---RGDVEC 56
C7 (a. a 571-626) CPSPPALKDGFVQNEETTFPVGKNIVYSQNEGYSLVGDVPVARGEDLQWT---VGKMHCC 56
C7 (a. a 631-688) CVLPTLMRGLQSHFPQKPFYTVGEKVTFSQSSGMSLEGPSTFLCGSSLKWSP-EMKNVQC 58
C6 (a. a 704-761) CLLPVVPEVLKLSPPFORLYKIGDSIELSPKGFVIAGPSRYTCSGD-SWTPPISSSLTCC 58

```

#### Type V protein segment (similar to FIMAC module)

```

C7 (a. a 696-765) AKKVPECCQLWEK--LQNSKCVCKMP-YECCGSS--LDVCCARDERSKRILRLTVCKMHVLQC 55
C7 (a. a 772-840) --ACGACPLWEKDAQSSKCVCRRA-SECEEAG-FRVCV--EVNGREQTMTECEAGVLRCC 54
C6 (a. a 861-934) SCGYDTCYSWEKCSATTSKICLLP-FQQLKGGYQHYCVKMGSSTTTTRTMNICVGAIRC 59
C6 (a. a 767-839) TRLKGHCQPGQK--QLGSECVCLSPPEEDCSHHS-EDLQVLDTDSNHYFTSSACKFLAEKC 57

C7 (a. a 696-765) QG-RNYTLSVGETCTL 70
C7 (a. a 772-840) LG-LSITVTSIRPCAP 69
C6 (a. a 861-934) AL-RKMEILYPGRORS 74
C6 (a. a 767-839) LNNQQLQFLHIGSQD 73

```

Figure 25 Multiple alignments of homologous protein domains in both the amino terminus and carboxyl terminus of the candidate genes C6, C7, C8A, C8B and C9. Gaps (-) were added to maximize the similarities among sequences.



## MACPF module

```

C6 (a.a 312-516)   SFIRIHKVIKVLNFTMKT-KDLQLSDVFLKALNHLPLEYNAALYSRIFDDDFGTHYFTSGS 59
C8A (a.a 291-497) SFIRIFTKVQTASFMMRR-DNIMLDEVMLQSLMELPEQYNYGMYAKFIDYDGYTHYITSGS 59
C9 (a.a 297-507)   IFLHVKGVIHLGRFVMRK-RDVMLTKTFLDDVKYLPSTYEKGEYFAFLETYGYTHYSSSGS 59
C7 (a.a 249-450)   QLLVVQNTVEVAQFINNNPEFLQLAESFWKELSYLPPLYDYSAYRRLIDQYGYTHYLSQSGS 60
C8B (a.a 303-510) TFLHARSdleVARYKlKs-RNLMLHYEFLQRVKQLPLEYSYGEYRDLFRDFGTHYITEAV 59

C6 (a.a 312-516)   LGGVYDlLYQFSNEELKNSGLTQEEAKNIRIETKKRYFIVTKTKVEHRITNRMSEKYE 119
C8A (a.a 291-497) MGGVYeyILVLNKENMTKSGVTSDDVTSDFGGSGFIDYDyTDNLQITG-SLSGKHCKKLG 118
C9 (a.a 297-507)   LGGlyELIYVLDKATMTEKGIELRDVHRDLG--FNLDLSLNFgVEIKG-KIDSENCLKRG 116
C7 (a.a 249-450)   LGGEYKVLfYVDSEKVAESDLGSEDKkKcASSHISFLFKS-----SKHKCKAME 109
C8B (a.a 303-510) LGGVYeyTLIMNKEAMERADYSLKDVHACAOHGFKIGVAIEEVYVKLG--VPVHKCKDIL 117

C6 (a.a 312-516)   GSFLQgSEKSISLVKGRSEYAAALAWEKgSSGPG-----EKTYSdWLESVkenPAV 171
C8A (a.a 291-497) GGHREDEESNMAVEDIIISRVrgSSGwGGGLTQNGS-----IITyRAWGRSLKYNPAV 171
C9 (a.a 297-507)   DGKTENIMNDdFIDdVISFIRGGTRKYATELKEKLLKGAKMINVTDfVNWASSLNDAPVL 176
C7 (a.a 249-450)   EALKSASGTQSNVLRGVpFVRGGRpGFVSGLSyLELDNPDG-NKQRySSWAGSVTDLPQV 168
C8B (a.a 303-510) NEIKDRNKRSSMvNDLVVLRvGGASEHITAlAYKDLp-----TADLMQEWGDavQYNPDi 172

C6 (a.a 312-516)   IDfELAPITDLVR--NIPAVTRRNlRRAFREYAA 205
C8A (a.a 291-497) IDfEMKPIYEILRHTNLGPLEAKRQNLRRALDQYLM 207
C9 (a.a 297-507)   INQKLSPIYDLIP-VKLNDAHLKRQNLERAIEDYIN 211
C7 (a.a 249-450)   IKQKLTPLYELVK--EVPASVKRlyLKRALEeYLD 202
C8B (a.a 303-510) IKIKAEPLYELVTAADFAYSSTVKNMKRALEEFEK 208

```

Figure 25 (continued) The alignment was created using ClustalW2 program. Cysteines are shaded in grey. Amino acids are enumerated on the right hand margin. The TSP1, LDLa, EGF, CCP, FIMAC and MACPF domains contain 40-54, 37-38, 29-47, 56-58, 69-74 and 202-211 amino acids, respectively. The TSP1 region is repeated three, two, two times in C6, C7 and C8A protein sequence, respectively. Similarity among TSP1s, LDLas, EGFs, CCPs, FIMACs or MACPFs in the candidate genes is found in a range of 5-55%, 32-70%, 48%, 14-35%, 13-36% or 26-34%, respectively.

Table 15 Number of cysteine residues in each functional protein domain of the candidate genes. Each value in the same cell reflects an individual domain of the same type, i.e. C6 has three TSP1 domains

Gene	TSP1	LDLa	MACPF	EGF	CCP	FIMAC	Lipocalin
C6	6, 5 and 5	6	3	7	4 and 4	8 and 10	-
C7	5 and 5	6	3	-	4 and 4	8 and 10	-
C8A	5 and 4	6	2	-	-	-	-
C8B	5	6	2	5	-	-	-
C8G	-	-	-	-	-	-	3
C9	5	6	2	-	-	-	-

### 4.3. Detection of genotypes and assignment of candidate genes to pig chromosome

#### 4.3.1. Genotype analysis

Table 16 Genotypic frequencies of the candidate genes in different breeds. (Up to down) the SNPs were detected by restriction digestion enzyme TaqI, PstI, BsrDI, Hin6I, MboII, TfiI, Hin6I, KnpI, FnuDII, by repeated comparative sequencing, MaeII, BsrDI and HpyCH4III

Gene: SNP	LR	%	PIE	%	MK	%
C6: 862A→G (n)	30		30		25	
AA	14	0.47	1	0.03	0	0.00
AG	13	0.43	11	0.37	0	0.00
GG	3	0.10	18	0.60	25	1.00
C6: 1557C→G (n)	30		30		25	
CC	2	0.07	0	0.00	0	0.00
CG	5	0.17	0	0.00	0	0.00
GG	23	0.77	30	1.00	25	1.00
C7: 154A→G (n)	30		30		25	
AA	3	0.10	0	0.00	0	0.00
AG	12	0.40	5	0.17	0	0.00
GG	15	0.50	25	0.83	25	1.00
C7: 870C→T (n)	30		30		25	
CC	0	0.00	0	0.00	7	0.28
CT	13	0.43	12	0.40	16	0.64
TT	17	0.57	18	0.60	2	0.08
C7: 881A→G (n)	30		30		25	
AA	0	0.00	0	0.00	7	0.28
AG	13	0.43	12	0.40	16	0.64
GG	17	0.57	18	0.60	2	0.08
C8A: 535A→G (n)	30		30		25	
AA	2	0.07	9	0.30	0	0.00
AG	10	0.33	15	0.50	5	0.20
GG	18	0.60	6	0.20	20	0.80

Table 16 (continued)

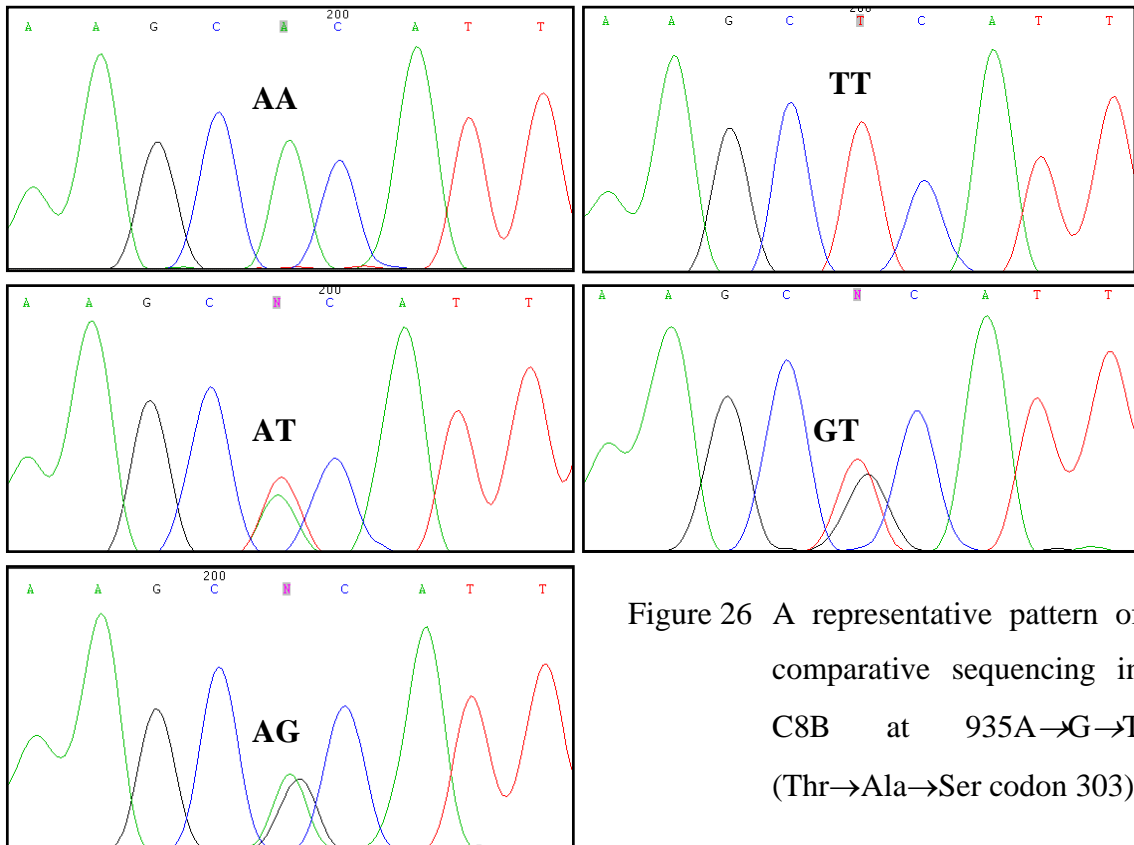
Gene: SNP	LR	%	PIE	%	MK	%
C8A: 1544C→T (n)	30		30		25	
CC	10	0.33	14	0.47	23	0.92
CT	12	0.40	12	0.40	2	0.08
TT	8	0.27	4	0.13	0	0.00
C8A: 1768C→T (n)	30		29		25	
CC	23	0.77	25	0.86	18	0.72
CT	7	0.23	4	0.14	7	0.28
TT	0	0.00	0	0.00	0	0.00
C8B: 222C→T (n)	30		30		25	
CC	30	1.00	30	1.00	18	0.72
CT	0	0.00	0	0.00	7	0.28
TT	0	0.00	0	0.00	0	0.00
C8B: 935A→G→T (n)	30		30		25	
AA	7	0.23	0	0.00	0	0.00
AG	6	0.20	0	0.00	0	0.00
AT	13	0.43	9	0.30	2	0.08
GT	2	0.07	4	0.13	7	0.28
TT	2	0.07	17	0.57	16	0.64
C8B: 1244A→G (n)	30		29		25	
AA	7	0.23	6	0.21	0	0.00
AG	19	0.63	17	0.58	4	0.16
GG	4	0.13	6	0.21	21	0.84
C9: 350A→G (n)	30		30		25	
AA	2	0.07	2	0.07	0	0.00
AG	18	0.60	8	0.27	0	0.00
GG	10	0.33	20	0.67	25	1.00
C9: 407C→G (n) (*)	30		30		25	
CC	0	0.00	0	0.00	0	0.00
CG	0	0.00	0	0.00	0	0.00
GG	30	1.00	30	1.00	25	1.00

(\*)\_ The alleles segregate in F<sub>2</sub> DUMI animals but not in LR, PIE or MK

Table 17 Allelic frequencies at several SNPs in different breeds of pig

Gene: SNP	LR	%	PIE	%	MK	%
C6: 862A→G						
A	41	0.68	13	0.22	0	0.00
G	19	0.32	47	0.78	50	1.00
C6: 1557C→G						
C	9	0.15	0	0.00	0	0.00
G	51	0.85	60	1.00	50	1.00
C7: 154A→G						
A	18	0.30	5	0.08	0	0.00
G	42	0.70	55	0.92	50	1.00
C7: 870C→T						
C	13	0.22	12	0.20	30	0.60
T	47	0.78	48	0.80	20	0.40
C7: 881A→G						
A	13	0.22	12	0.20	30	0.60
G	47	0.78	48	0.80	20	0.40
C8A: 535A→G						
A	14	0.23	33	0.55	5	0.10
G	46	0.77	27	0.45	45	0.90
C8A: 1544C→T						
C	32	0.53	40	0.67	48	0.96
T	28	0.47	20	0.33	2	0.04
C8A: 1768C→T						
C	53	0.88	54	0.93	43	0.86
T	7	0.12	4	0.07	7	0.14
C8B: 222C→T						
C	60	1.00	60	1.00	43	0.86
T	0	0.00	0	0.00	7	0.14
C8B: 935A→G→T						
A	33	0.50	9	0.15	2	0.04
G	14	0.21	4	0.07	7	0.14
T	19	0.29	47	0.78	41	0.82
C8B: 1244A→G						
A	33	0.55	29	0.50	4	0.08
G	27	0.45	29	0.50	46	0.92
C9: 350A→G						
A	22	0.37	12	0.20	0	0.00
G	38	0.63	48	0.80	50	1.00
C9: 408C→G						
C	0	0.00	0	0.00	0	0.00
G	60	1.00	60	1.00	50	1.00

Once mutations were established, genomic DNA from unrelated animals of Landrace (n=30), Pietrain (n=30), Muong Khuong (n=25) and F<sub>2</sub> DUMI population (n=417) were amplified using the primer pairs in table 1 and table 4 in order to examine genetic variability in the candidate genes by using PCR-RFLP or repeated comparative sequencing. Differences of allelic frequencies were observed among Vietnamese and European commercial breeds. The frequency of heterozygous animals in the European breeds is higher than in the Vietnamese breed Muong Khuong pig for C6, C8A and C9 candidate genes. However for C7 and C8B it is reverse. Especially, all genotypes of C8B at locus 935 with three different alleles were found in the Landrace breed but in Pietrain and Muong Khuong the genotypes AA and AG were not found (Table 16 and Table 17).



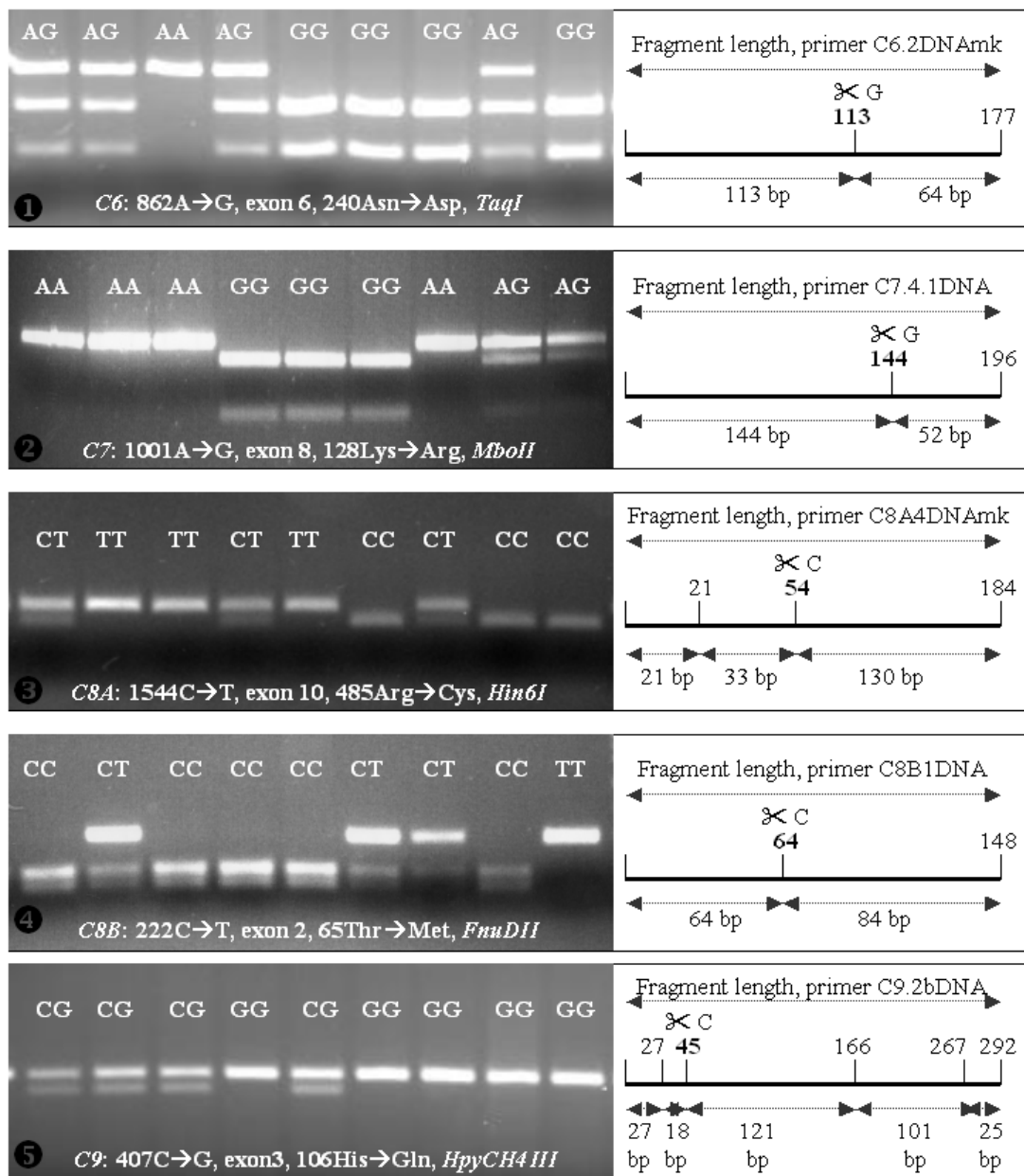


Figure 27 Representative patterns of detection of SNP using PCR-RFLP at positions ①\_862A→G, homologous AA (177 bp), GG (113 bp + 64 bp) and heterologous AG (113 bp + 64 bp + 177 bp) ②\_881A→G, homologous AA (196 bp), GG (144 + 52 bp) and heterologous AG (144 bp + 52 bp + 196 bp) ③\_1544C→T, homologous CC (33 bp), TT (163 bp) and heterologous CT (33 bp + 163 bp) ④\_222C→T, homologous CC (64 bp + 84 bp), TT(148 bp) and heterologous CT (64 bp + 84 bp + 148 bp) ⑤\_407C→G, homologous GG (139 bp) and heterologous CG (121 bp + 139 bp) in porcine C6, C7, C8A, C8B and C9 candidate genes, respectively. PCR products were digested with *TaqI*, *MboII*, *Hin6I*, *FnuDII* and *HpyCH4III*.

No significant differences between the observed and the expected numbers of genotypes were found in genotyped SNPs in Landrace, Pietrain and Muong Khuong populations in the candidate genes. It means that all genotypic patterns were found to be in Hardy-Weinberg equilibrium in commercial breeds. The restriction enzymes TaqI, MboII, Hin6I, FnuDII and HpyCH4III were used for detecting the variation of the nucleotides 862A→G, 881A→G, 1544C→T, 222C→T and 407C→G in C6, C7, C8A, C8B and C9 genes, respectively, which were segregating in the F<sub>2</sub> DUMI resource population. They were used to evaluate association between genotypes and hemolytic complement activity in both the classical and alternative pathway as well as for linkage mapping using the CRIMAP tool. Representative patterns for SNP genotyping using repeated comparative sequencing and PCR-RFLP are shown in figure 26 and figure 27, respectively.

#### 4.3.2 Location of the candidate genes on porcine chromosome

Genetic linkage and chromosomal localization studies previously established that C8A and C9 are located on chromosome 6q3.1-q3.3 (Nakajima et al. 1998) and 16q1.4 (Thomsen et al. 1998) using in situ hybridization, respectively. In this study, to determine the location and order of certain candidate genes, we used the IMpRH panel, which allows assigning unambiguously the six loci to porcine chromosomes from pig/hamster genomic DNA amplification products converted into vector sequences. Two-point RH analyses were used for the identification of linkage groups using LOD score threshold of 5.0. With all information for markers, including vectors summarized in table 18, it was demonstrated that both porcine C8A and C8B loci were physically assigned to the same chromosome 6q3.1-q3.5 whereas the porcine C6, C7 and C9 genes were closely linked on the same q-arm of chromosome 16 (16q1.4). The assignment of C9 agrees with previously obtained results of Thomson et al. (1998). C8G was assigned to chromosomes 1q2.13. Moreover, the porcine C6, C7, C8A, C8B, C8G and C9 significantly linked to markers S0077, SW1069, SW322, SSC10D08 and SW403, respectively (Figure 29). All LOD scores are greater than 6.

Genetic linkages using two-point analysis from CIRMAP 2.4 (Green et al. 1990) were confirmed between C7 and C6 with rec. frags. = 0.02, LOD = 50.45, between C9 and

C6 with rec. frags = 0.06, LOD = 31.50, between C9 and C7 with rec. frags = 0.06, LOD = 35.84, between C8B and C8A with rec. frags = 0.05, LOD = 65.72.

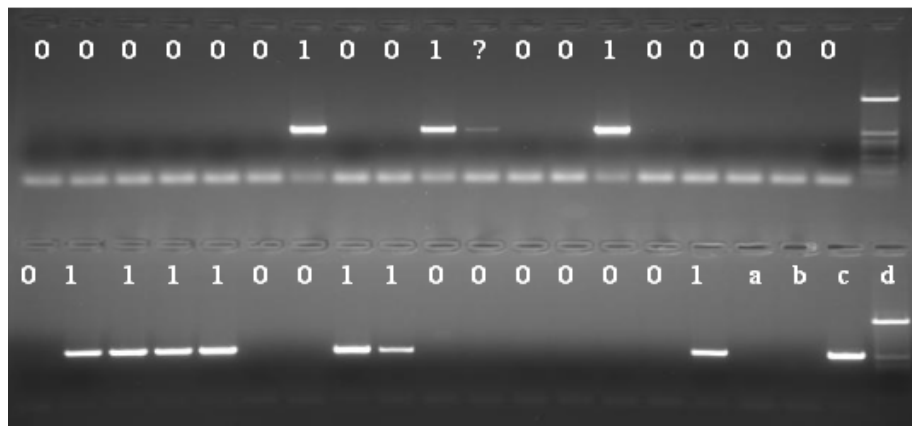


Figure 28 A representative pattern for PCR result of 118 DNA clones of the IMpRH. The symbols (0), (?), and (1) used for absent, ambiguous and present results, respectively. The letters a, b, c, and d show a hamster sample (negative control), a blank sample without DNA (negative control), a pig DNA (positive control) and a marker to estimate length of DNA fragments.

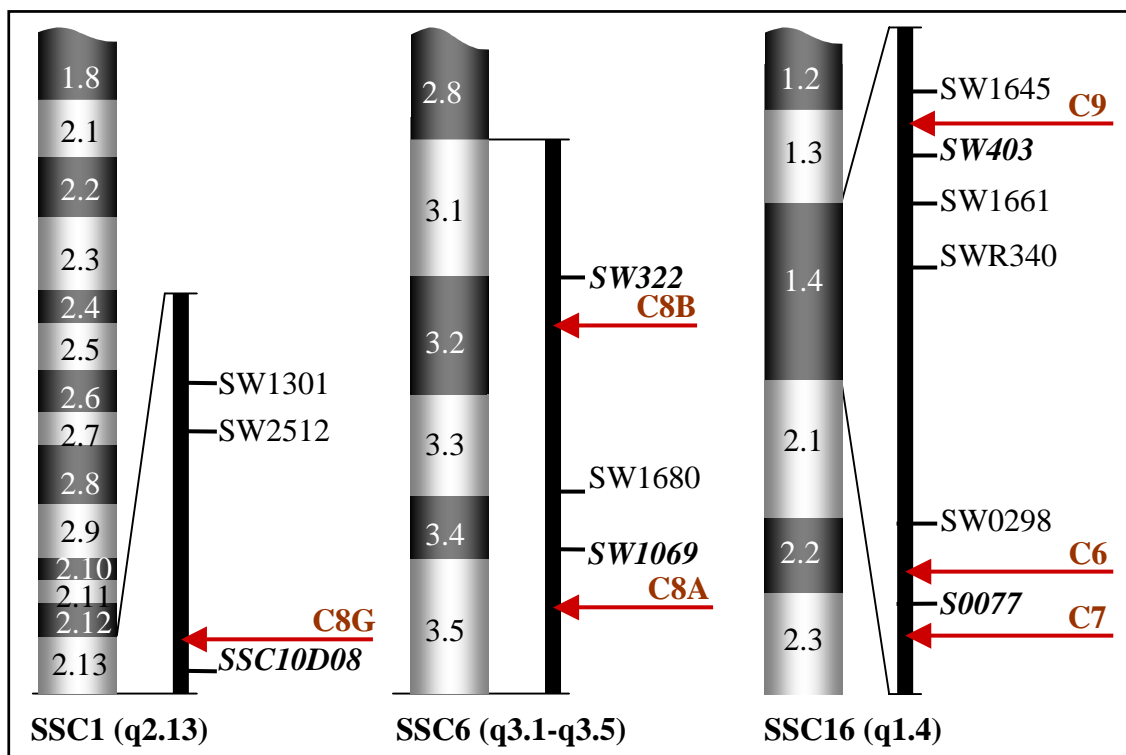


Figure 29 Position of the candidate genes on porcine chromosomes. Linkage is closed to markers in bold and italic



Table 18 Summary of RH mapping results of the candidate genes

	Result in vector	1	2	3	4
C6	000010000001001100001111000111 001001000110011010000011010010 011011011011011100000001110000 0001000010000101111111111100	18.13	21	42	16q1.4
C7	11011000000100010000011100010? 001001000110010010100000000000 011001011011011100000001110000 0001000010000101111111111011	16.65	23	38	16q1.4
C8A	001000010101100000001110000010 000000000000010001000000000000 001000011010000000010110100011 0000000100101100000000010001	7.00	56	22	6q3.1-q3.5
C8B	001000010101100000011110000010 000000010000010001000000000000 001000011000100000010110100011 0000000100101100000000010001	8.18	50	24	6q3.1-q3.5
C8G	000000000000100001100000000000 000000000100001000000010000001 1000000000000000000000110000001 0011000000000010100001001010	8.26	44	15	1q2.13
C9	110110000001000100000001000100 01100100010000001000000010000 010001011011010100000001000000 0001100010001101111111000101	14.45	27	31	16q1.4

<sup>(1)</sup>\_ LOD score, <sup>(2)</sup>\_ Distance in cR, <sup>(3)</sup>\_ Retention fraction (%), <sup>(4)</sup>\_ Chromosome

#### 4.4. Association of candidate genes with hemolytic complement activity

Genotype analysis results were applied to associate with hemolytic complement activity in both classical and alternative pathway in the F<sub>2</sub> DUMI resource population.

##### 4.4.1. The classical lytic pathway

The effect of different genotypes on hemolytic complement activity in the classical pathway was analyzed using repeated measures mixed models of analysis of variance. Genotypes of the porcine C7 (p=0.0080), and C9 genes (p=0.0488) revealed significant

effects on CH50 whereas the genotypes of C6 ( $p=0.0853$ ) and C8B ( $p=0.4410$ ) did not. The effect of C8A was close to significance ( $p=0.0650$ ). For the genotypes CC and TT of C8A least squares means were obtained differing at  $p=0.0522$  (Figure 30 and Table 19). Additionally, although there was no difference between C8B genotypes in CH50 ( $p=0.4410$ ) the interaction of genotypes and eight different time points before and after immunizations with Mycoplasma, Ausjezky and PRRSV was highly significant difference ( $p=0.0048$ ) (Table 31 and Figure 32). In general, CH50 activities were increased after each of complement activation by vaccination. The highest least squares means were observed at the fourth day after ADV vaccination in all candidate genes.

Table 19 Least squares means of hemolytic activity in both the classical (CH50) and alternative (AH50) complement pathways for different genotypes (LSM $\pm$ SE) (U/ml)

	CH50	AH50		CH50	AH50
Porcine C6			Porcine C7		
AA	59.72 $\pm$ 7.65	66.94 $\pm$ 5.87	AA	70.27 $\pm$ 3.62	61.11 $\pm$ 2.70
AG	64.98 $\pm$ 3.39	61.75 $\pm$ 2.57	AG	67.72 $\pm$ 3.60	61.42 $\pm$ 2.72
GG	70.42 $\pm$ 3.32	61.41 $\pm$ 2.49	GG	44.13 $\pm$ 8.55	59.69 $\pm$ 7.22
p-value	0.0853	0.6581	p-value	0.0080	0.8319
Porcine C8A			Porcine C8B		
CC	59.45 $\pm$ 4.07	61.94 $\pm$ 3.13	CC	65.58 $\pm$ 3.11	60.96 $\pm$ 2.33
CT	65.08 $\pm$ 3.49	60.64 $\pm$ 2.66	CT	62.07 $\pm$ 3.47	60.30 $\pm$ 2.61
TT	68.24 $\pm$ 3.29	60.90 $\pm$ 2.40	TT	64.27 $\pm$ 4.93	63.99 $\pm$ 3.54
p-value	0.0650	0.8983	p-value	0.4410	0.5110
Porcine C9					
GG	69.49 $\pm$ 3.17	62.73 $\pm$ 2.82			
CG	63.59 $\pm$ 3.65	61.52 $\pm$ 2.39			
p-value	0.0488	0.5968			

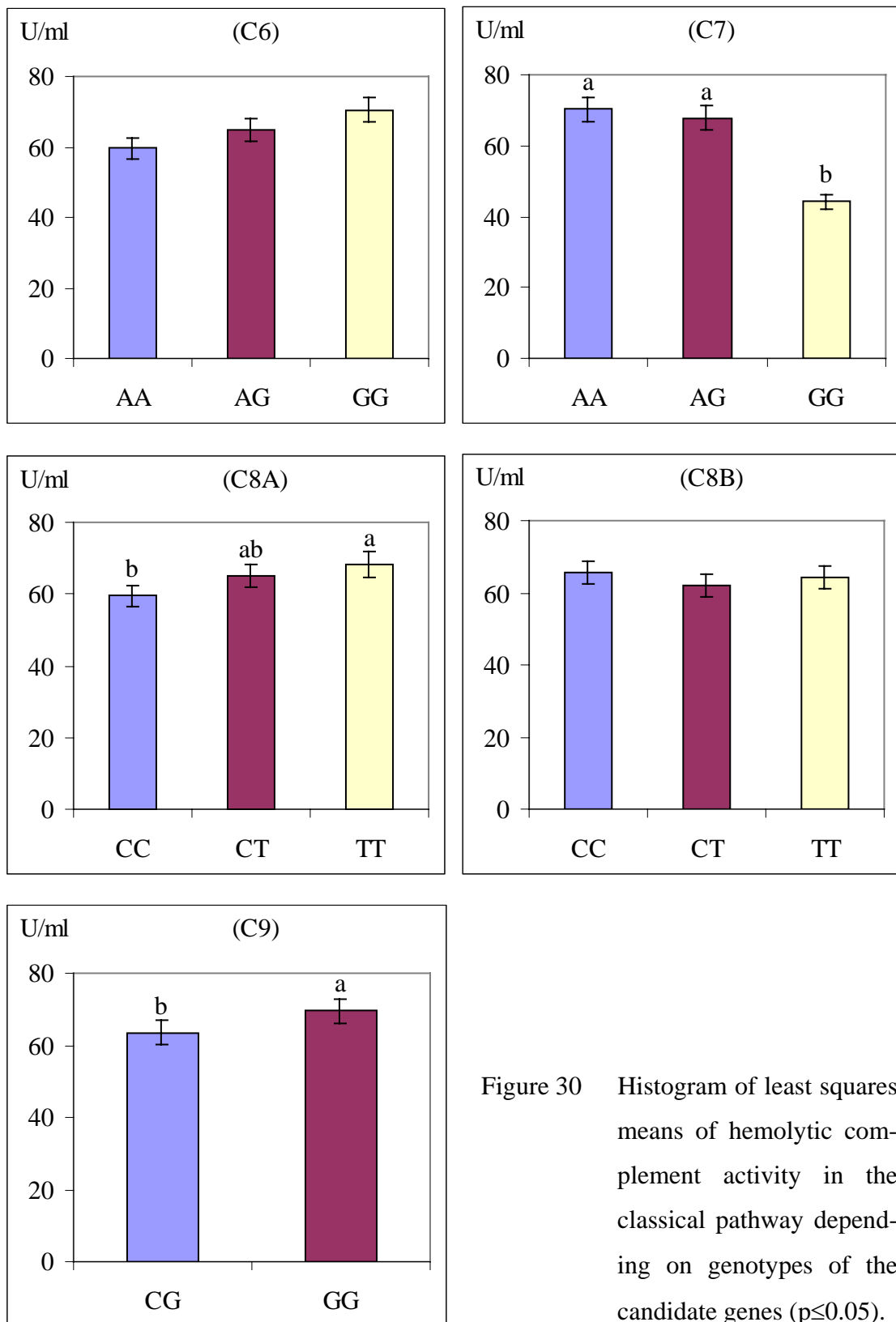


Figure 30 Histogram of least squares means of hemolytic complement activity in the classical pathway depending on genotypes of the candidate genes ( $p \leq 0.05$ ).

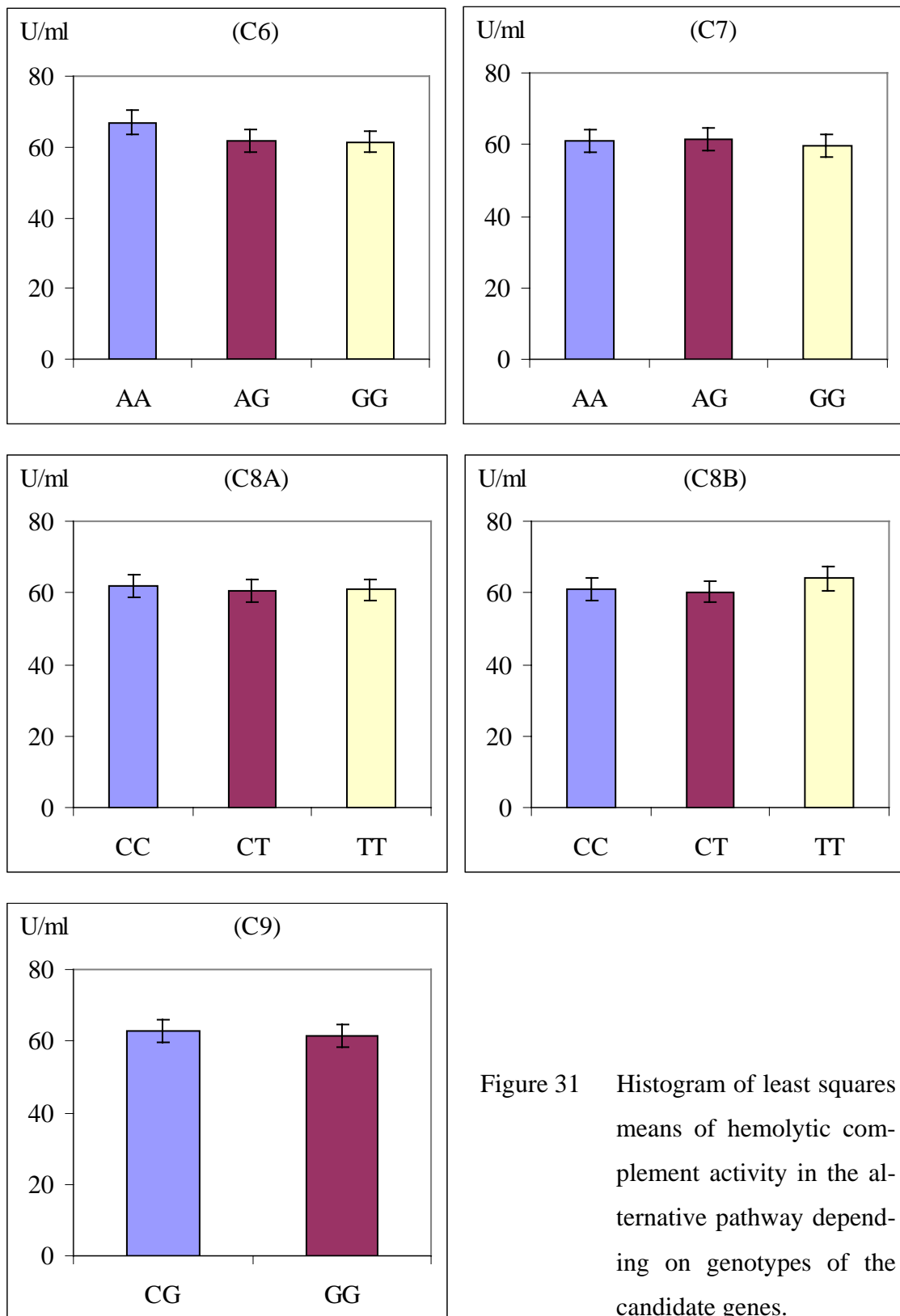


Figure 31 Histogram of least squares means of hemolytic complement activity in the alternative pathway depending on genotypes of the candidate genes.

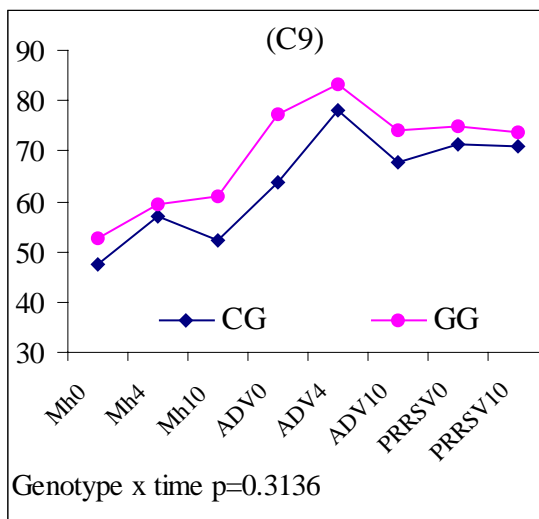
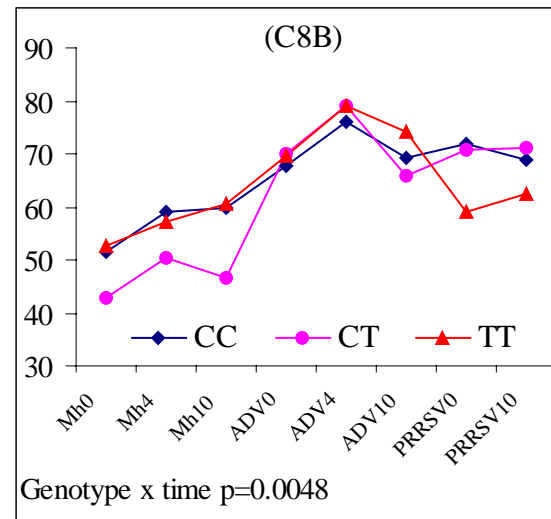
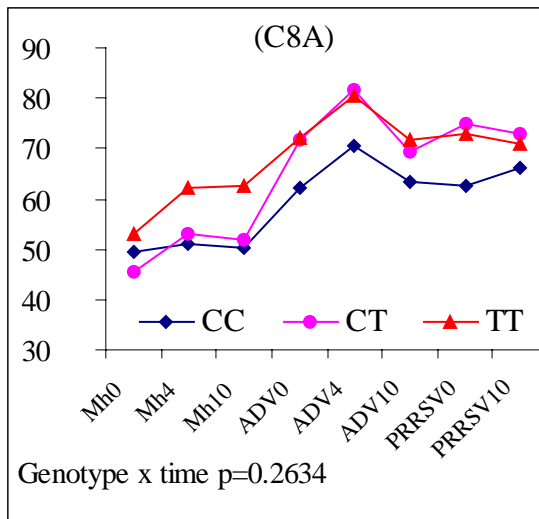
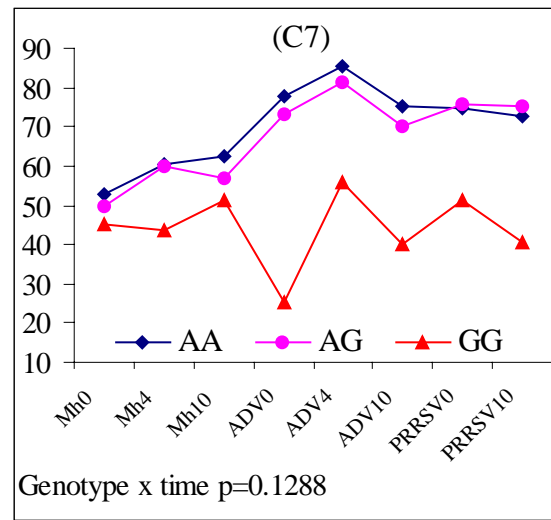
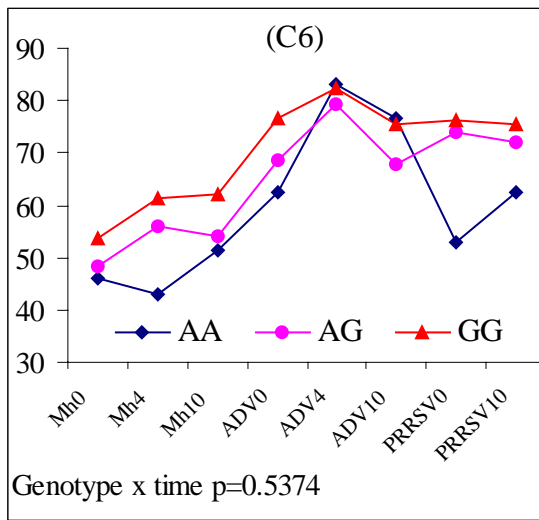


Figure 32 Histogram of least squares means of hemolytic complement activity in the classical pathway along vaccination for the interaction of genotypes of each candidate gene and eight different time points.

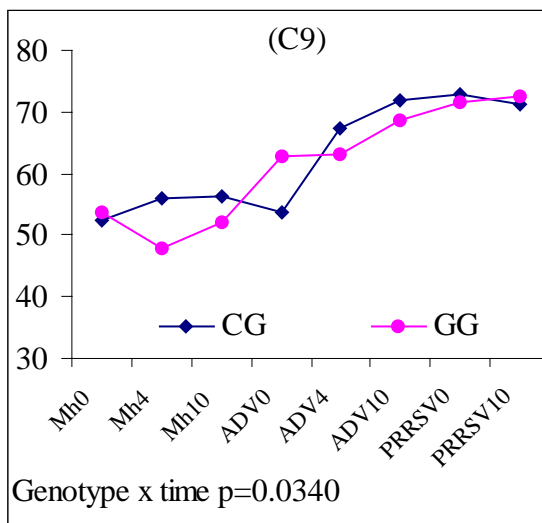
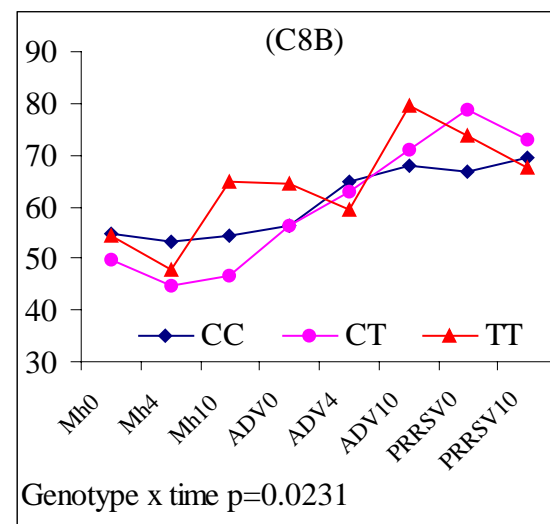
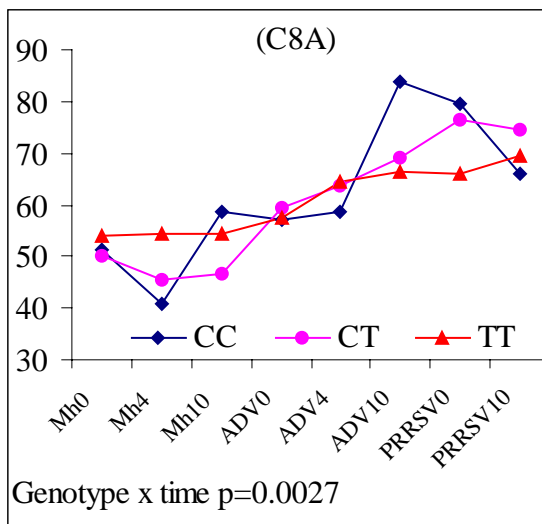
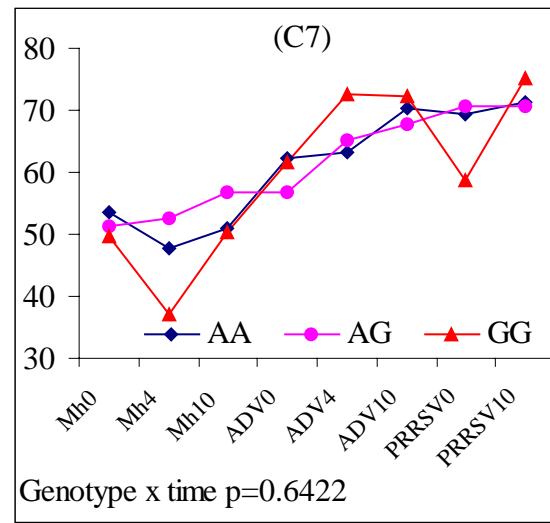
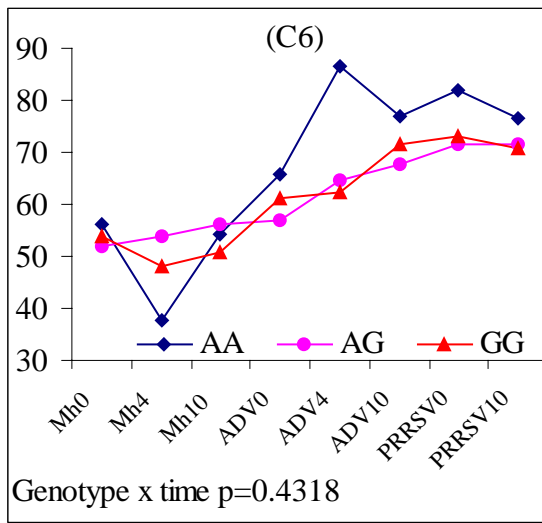


Figure 33 Histogram of least squares means of hemolytic complement activity in the alternative pathway along vaccination for the interaction of genotypes of each candidate gene and eight different time points.

#### 4.4.2 The alternative pathway

Studying the effect of genetic variation on hemolytic complement activity in the alternative pathway was conducted for all candidate genes except C8G because of no polymorphism was found within its sequence. AH50 values were used to evaluate lysis level of different genotypes in each gene. As a result, no difference of genotypes in AH50 test was found in all porcine candidate genes C6, C7, C8A, C8B and C9 (Figure 31 and Table 19). However, there was statistically significant interaction of genotypes and time points of immunization on AH50 for C8A ( $p=0.0027$ ), C8B ( $p=0.0231$ ) and C9 ( $p=0.0340$ ) (Figure 33). Homozygous genotypes AA, CC and TT in candidate genes C6, C8A and C8B, respectively, showed higher complement activity in the alternative pathway. AH50 increment of these genotypes was unlinear after vaccinations. Hemolytic activity in this pathway was highest after ADV or PRRSV immunization.

## Discussion

### 5.1 Structure of candidate genes

#### 5.1.1 Characteristic of the porcine C6 and C7 complement genes

Complement C6 is a glycosylated plasma protein (DiScipio et al. 1989). The porcine C6 protein sequence (GenBank acc. no. ABD13967) showed >80% identity to the human orthologue. With an atomic composition formula of  $C_{4565}H_{7108}N_{1288}O_{1429}S_{74}$  the porcine C6 molecular weight ( $M_r$ ) was predicted about 105.27 kDa; i.e. similar to human C6 ( $M_r = 105$  kDa) (Schreck et al. 2000) by using the ProtParam tool (<http://au.expasy.org/tools/protparam.html>). Both pig and human sequences for C6 contain 64 cysteines forming 32 disulfide bonds, which are important to stabilize the protein structure and to maintain its activity (DiScipio et al. 1989). The bonds could be formed between Cys<sup>1</sup>-Cys<sup>11</sup>, Cys<sup>2</sup>-Cys<sup>16</sup>, Cys<sup>3</sup>-Cys<sup>20</sup>, Cys<sup>4</sup>-Cys<sup>31</sup>, Cys<sup>5</sup>-Cys<sup>10</sup>, Cys<sup>6</sup>-Cys<sup>30</sup>, Cys<sup>7</sup>-Cys<sup>36</sup>, Cys<sup>8</sup>-Cys<sup>9</sup>, Cys<sup>12</sup>-Cys<sup>39</sup>, Cys<sup>13</sup>-Cys<sup>17</sup>, Cys<sup>14</sup>-Cys<sup>38</sup>, Cys<sup>15</sup>-Cys<sup>60</sup>, Cys<sup>18</sup>-Cys<sup>26</sup>, Cys<sup>19</sup>-Cys<sup>33</sup>, Cys<sup>21</sup>-Cys<sup>48</sup>, Cys<sup>22</sup>-Cys<sup>24</sup>, Cys<sup>23</sup>-Cys<sup>45</sup>, Cys<sup>25</sup>-Cys<sup>62</sup>, Cys<sup>27</sup>-Cys<sup>52</sup>, Cys<sup>28</sup>-Cys<sup>32</sup>, Cys<sup>29</sup>-Cys<sup>50</sup>, Cys<sup>34</sup>-Cys<sup>44</sup>, Cys<sup>35</sup>-Cys<sup>40</sup>, Cys<sup>37</sup>-Cys<sup>42</sup>, Cys<sup>41</sup>-Cys<sup>53</sup>, Cys<sup>43</sup>-Cys<sup>58</sup>, Cys<sup>46</sup>-Cys<sup>56</sup>, Cys<sup>47</sup>-Cys<sup>64</sup>, Cys<sup>49</sup>-Cys<sup>51</sup>, Cys<sup>54</sup>-Cys<sup>61</sup>, Cys<sup>55</sup>-Cys<sup>63</sup>, and Cys<sup>57</sup>-Cys<sup>59</sup> as predicted using the DiANNA 1.1 software (Ferre and Clote 2005). According to von Heijne's theory (1986), the signal peptide sequence is most likely cleaved after the alanine residue. Consequently, the first 21 residues of the open reading frame of human C6 show the typical features of a signal peptide, as one would expect for a secreted protein (Haefliger et al. 1989). In silico prediction for porcine C6 indicates existence of a complementary signal peptide. After removing signal peptide sequence, the biological active polypeptide chain is composed of 913 residues in human sequence (DiScipio et al. 1989) and 914 residues in pig sequence. The porcine amino terminal C6a fragment (a.a 22-631) starts just after the signal peptide followed by the carboxyl terminal C6b segment (a.a 632-934) (Haefliger et al. 1989, DiScipio et al. 1989, Ripoché et al. 1988, Goldberger et al. 1987, Kristensen et al. 1986). The whole porcine C6 protein sequence was most closely homologous to the porcine complement C7 with 29% identity. However, analysis of C6a revealed high similarity to C8A, C8B and C9. The C6b fragment only shows considerable similarity to C7. Corresponding



results were found in human (DiScipio et al. 1989, Haefliger et al. 1987, Rao et al. 1987, Stanley et al. 1985, DiScipio et al. 1984). Similarly to the human C6 (Ripoche et al. 1988, Goldberger et al. 1987, Kristensen et al. 1986) the porcine C6b had partial homology to the complement regulatory proteins, CCP and FIMAC. Consequently a number of predicted and/or experimentally proven functions and properties of particular human amino acid sequence motifs can be expected in the porcine protein too. Human C6 has two possible sites for N-glycosylation (Asn-X-Ser/Thr), one at the amino (a.a 324) and one at the carboxyl terminus (a.a 855) (DiScipio et al. 1989, Clamp et al. 1975). The first is conserved in the pig whereas at the second site at position 855 there is a Ser instead of Asn. Using the NetNGlyc 1.0 Server tool one more N-glycosylation site was found at a.a 860 in the carboxyl terminus of the porcine C6 (Figure 46). In the C6 amino-terminus there are three cysteine-rich segments, a pair of TSP1 domains followed by a LDLa domain, containing 11 disulfide bonds. This region has a predominance of  $\beta$ -sheet and  $\beta$ -turns (DiSipio et al. 1989). Further, several hydrophobic peaks are seen in the middle segment of C6 protein containing a MACPF module, analogous to C9. The membrane-binding region may reside here (DiSipio et al. 1989). The region consists of type EGF and TSP1 units containing seven disulfide bridges with a predominance of  $\beta$ -sheet and  $\beta$ -turns. In the putative carboxyl domain, each CCP and each FIMAC has two and nine disulfide bonds, respectively. The model for the CCP motif envisions two pairs of overlapping disulfide bonds with distorted  $\beta$ -sheet interspaced by turns (DiSipio et al. 1989). The third TSP1 of C6 contributes to the C5 binding site of C6 and involved in terminal complement complex assembly (Wuerzner et al. 1995)

Porcine and human C7 are highly conserved, sharing structural and functional characteristics (Agah et al. 2000). The porcine C7 molecular weight ( $C_{4043}H_{6287}N_{1123}O_{1268}S_{67}$ ) was 93.06 kDa, similar to human C7 ( $M_r = 92$  kDa) (Schreck et al. 2000). Similar to human C7, the predicted amino acid sequence of the precursor porcine C7 (GenBank acc. no. AAD45918) consists of 843 amino acid residues. The pig precursor protein contains a 28-residue hydrophobic signal peptide identified in the pig using the SMART and the SignalP analysis tools while it is 22 residues in human (DiScipio et al. 1998). Both pig and human C7 have 56 cysteines (6.8% of the total amino acid composition) giving rise to cysteine-rich domains (Agah et al. 2000) and share an overall identity of 83% at the protein level. The cysteine residues are presumed to form 28 disulfide bonds,

which are important for stabilizing the structure of this protein (DiScipio et al. 1988). Via calculation with the DiANNA 1.1 software the possible disulfide bonds in the porcine C7 protein were predicted at: Cys<sup>1</sup>-Cys<sup>6</sup>, Cys<sup>2</sup>-Cys<sup>32</sup>, Cys<sup>3</sup>-Cys<sup>22</sup>, Cys<sup>4</sup>-Cys<sup>23</sup>, Cys<sup>5</sup>-Cys<sup>18</sup>, Cys<sup>7</sup>-Cys<sup>36</sup>, Cys<sup>8</sup>-Cys<sup>55</sup>, Cys<sup>9</sup>-Cys<sup>50</sup>, Cys<sup>10</sup>-Cys<sup>14</sup>, Cys<sup>11</sup>-Cys<sup>16</sup>, Cys<sup>12</sup>-Cys<sup>45</sup>, Cys<sup>13</sup>-Cys<sup>34</sup>, Cys<sup>15</sup>-Cys<sup>56</sup>, Cys<sup>17</sup>-Cys<sup>54</sup>, Cys<sup>19</sup>-Cys<sup>43</sup>, Cys<sup>20</sup>-Cys<sup>21</sup>, Cys<sup>24</sup>-Cys<sup>25</sup>, Cys<sup>26</sup>-Cys<sup>40</sup>, Cys<sup>27</sup>-Cys<sup>28</sup>, Cys<sup>29</sup>-Cys<sup>31</sup>, Cys<sup>30</sup>-Cys<sup>39</sup>, Cys<sup>33</sup>-Cys<sup>44</sup>, Cys<sup>35</sup>-Cys<sup>38</sup>, Cys<sup>37</sup>-Cys<sup>46</sup>, Cys<sup>41</sup>-Cys<sup>42</sup>, Cys<sup>47</sup>-Cys<sup>49</sup>, Cys<sup>48</sup>-Cys<sup>52</sup>, and Cys<sup>51</sup>-Cys<sup>53</sup>. Most of cysteine residues are found in small modules like TSP1, LDLa, CCP or FIMAC. Mature human C7 consists of two protein domains at the amino (a.a 23-575) and carboxyl terminus (a.a 576-843) (Haefliger et al. 1989, DiScipio et al. 1988). Both human and porcine C7 have two N-glycosylation sites, one at amino acid 202 of amino terminus and another at amino acid 754 of carboxyl terminus (Agah et al. 2000, DiScipio et al. 1988) (Figure 47). The amino terminal area of human C7 (a.a 22-347) is hydrophilic, but the carboxyl-terminal polypeptide (a.a 348-843) contains several intermittent zones of hydrophobicity, with the region between positions 629 to 807 having a particularly strong hydrophobic tendency (DiScipio et al. 1988).

### 5.1.2 Characteristic of the porcine C8 complement gene

The eighth component of the terminal complement lytic pathway is a serum protein consisting of three amino acid chains, encoded by C8A, C8B, and C8G (Kaufmann et al. 1993). Human C8A and C8G are covalently linked while C8B is noncovalently associated with the C8A-G (Tschopp et al. 1981). The C8B is responsible for binding of C8 to C5b-7 (Monahan et al. 1980) whereas the C8A interacts with C9 (Stewart et al. 1985) to form the membrane attack complex (C5b-9) which can polymerize to form trans-membrane pores and causes osmotic cell lysis. The C8G may act as a retinol transporting protein in plasma (Haefliger et al. 1991) and bind to yet unidentified ligands (De-wald et al. 1996, Kaufman et al. 1994, Ng et al. 1987). The C8G is not involved in the lytic mechanism and its function is still unknown (Brickner et al. 1984).

The pig C8A cDNA sequence (GenBank acc. no. DQ333200) contains 2146 nucleotides encoding 589 amino acids (GenBank acc. no. ABD13968). This is in agreement with 558 amino acids and a 31 amino acid leader sequence (signal peptide) in porcine C8A

protein sequence previously reported by Nakajima et al. (1998). However, the authors published a cDNA sequence (GenBank acc. no AB008156) with 2145 nucleotides. With an adenine missing at position 825 this would lead to a frameshift with a stop codon causing a deficient C8A protein. In addition, using the SignalP 3.0 tool analysing human and porcine signal peptides revealed that their putative cleavage site is situated between a.a 19 and 20 whereas Nakajima et al. (1998) reported the first 31 residues belonging to the signal peptide corresponding to predictions in the human C8A (Rao et al. 1987). Moreover, compared with the protein sequence reported by Nakajima et al. (1998) we have found three different amino acids at positions 244 (K→X), 419 (M→L) and 483 (A→V). These could be polymorphic sites with amino acid substitutions between porcine breeds analysed. Porcine C8A has a molecular weight of approximately 66 kDa calculated from the atomic composition formula  $C_{2856}H_{4406}N_{808}O_{906}S_{45}$  using the ProtParam tool while it was 64 kDa in human C8A (Steckel et al. 1980). The whole porcine C8A protein sequence, which is similar to human and rabbit C8A in length (Nakajima et al. 1998), contains 31 cysteines. Possibly, the disulfide bonds are formed between Cys<sup>1</sup>-Cys<sup>7</sup>, Cys<sup>2</sup>-Cys<sup>26</sup>, Cys<sup>3</sup>-Cys<sup>5</sup>, Cys<sup>4</sup>-Cys<sup>20</sup>, Cys<sup>6</sup>-Cys<sup>23</sup>, Cys<sup>8</sup>-Cys<sup>25</sup>, Cys<sup>9</sup>-Cys<sup>19</sup>, Cys<sup>10</sup>-Cys<sup>14</sup>, Cys<sup>11</sup>-Cys<sup>15</sup>, Cys<sup>12</sup>-Cys<sup>18</sup>, Cys<sup>13</sup>-Cys<sup>21</sup>, Cys<sup>16</sup>-Cys<sup>27</sup>, Cys<sup>22</sup>-Cys<sup>31</sup>, Cys<sup>24</sup>-Cys<sup>28</sup>, and Cys<sup>29</sup>-Cys<sup>30</sup>. Cysteine-rich modules like TSP1 (one in amino terminus (a.a 43-95) and one in carboxyl terminus (a.a 542-598)), LDLa (a.a 111-133), and EGF (a.a 522-533) were found in porcine C8A protein (Nakajima et al. 1998). The content of these modules is similar to the computer prediction result here using SMART analysis tool although they showed some different nucleotide positions. There are one or two N-glycosylation sites in human C8A (Rao et al. 1987). These correspond to sequences Asn-Trp-Ala (a.a 46-48) and Asn-Gly-Ser (a.a 442-444) in the porcine C8A protein sequence. Examination of the porcine C8A revealed one more N-glycosylation site at a.a 364-366 (Asn-Met-Thr) using the NetNGlyc 1.0 Server tool (Figure 48). Nakajima et al. (1998) proposed candidate site for CD59 recognition between a.a 382-419. The CD59 is known as a 18-20 kDa GPI-anchored membrane protein that functions as a key regulator of the terminal step of the complement activation cascade. It restricts binding of C9 to the C5b-8 complex, thereby preventing the formation of the MAC (Qian et al. 2000). During biosynthetic process, C8A contains a unique insertion (residues 189-205), which includes Cys<sup>194</sup> that forms the disulfide bond to human C8G. With a small exchange from Thr<sup>193</sup> in human to Ala<sup>196</sup> in pig, both these segments are highly con-

served. Therefore C8G must contain a complementary binding site for the C8A (Plumb and Sodetz 2000). The C8A-G disulfide bond has a key role in binding to C8B through a site located on C8A (Plumb et al. 1999). Plumb et al. (1999) illustrated that the binding site for C8B and C8G reside within the MACPF region of C8A. C8A mediates binding and self-polymerization of C9 to form a pore-like structure on the membrane of target cells (Slade et al. 2006, Scibek et al. 2002). The amino-terminal modules in C8A are important for C9 binding and/or expression of C8 activity (Plumb et al. 1999).

The porcine C8B cDNA sequence (GenBank acc. no DQ333201) shows 99% identity to a 2455 bp region (GenBank acc. no AK232952) separated from clone LVRM10055B03 (Uenishi et al. 2004). The difference is due to (1) lack of one nucleotide at the first position in 5'UTR flanking region (2) abundance of one nucleotide after position 1910 in 3' UTR flanking region (3) deletion of two nucleotides at positions 1728 and 1737 in coding region of GenBank acc. no AK232952. The deletions would lead to deficiency of C8B protein due to the formation of a stop codon. As compared to the 2563 bp nucleotide sequence of GenBank acc. no AK233291 (Uenishi H et al. 2004) we added 474 bp of 3'UTR flanking region of porcine C8B cDNA sequence including a polyadenylation signal (AATAAA) (nt. 2408-2413). Therefore, full-length porcine C8B sequence as shown in figure 43 is 2461 bp. The signal peptide sequence is 54 residues in human (Howard et al. 1987) while it is only 44 amino acids in the pig according to the SignalP 3.0 analysis. Three asparagines as candidates for potential N-glycosylation sites are at sequences Asn-Phe-Ser (a.a 101-103), Asn-Val-Thr (a.a 243-245) and Asn-Trp-Ser (a.a 553-555) in human C8B (Howard et al. 1987) corresponding to sequences Asn-Phe-Ser (a.a 113-115), Asn-Val-Thr (a.a 255-257) and Asp-Trp-Phe (a.a 565-567) in the pig, respectively. Probably the porcine C8B has three N-glycosylation sites, which two of them are conserved with human C8B and one extra site at a.a 56-58 (Asn-Glu-Ser) in the amino terminus (Figure 49). One of three possible sites of human C8B contains the same Ser-Asn-Trp-Ser sequence found at the proposed glycosylation site in human C8A. Probably, the porcine C8B only contains two possible sites because the first amino acid in the third sequence is an asparagine acid (Asp) instead of an asparagine (Asn). Using the ProtParam tool the molecular weight of porcine C8B (C<sub>3040</sub>H<sub>4735</sub>N<sub>855</sub>O<sub>915</sub>S<sub>39</sub>) was predicted to be 69.15 kDa while that of human C8B was 64 kDa (Steckel et al. 1980). With a total 27 cysteines found in the C8B protein structure

disulfide bonds could be formed between Cys<sup>1</sup>-Cys<sup>15</sup>, Cys<sup>2</sup>-Cys<sup>9</sup>, Cys<sup>3</sup>-Cys<sup>26</sup>, Cys<sup>4</sup>-Cys<sup>24</sup>, Cys<sup>5</sup>-Cys<sup>13</sup>, Cys<sup>6</sup>-Cys<sup>7</sup>, Cys<sup>8</sup>-Cys<sup>27</sup>, Cys<sup>10</sup>-Cys<sup>12</sup>, Cys<sup>11</sup>-Cys<sup>23</sup>, Cys<sup>14</sup>-Cys<sup>16</sup>, Cys<sup>17</sup>-Cys<sup>22</sup>, Cys<sup>18</sup>-Cys<sup>21</sup>, Cys<sup>20</sup>-Cys<sup>25</sup>. Analysis of the derived amino acid sequence revealed several membrane surface seeking segments that may facilitate C8B interaction with target membranes during complement-mediated cytolysis (Howard et al. 1987). The soluble terminal complement complex can be formed in C8B-deficient sera that contain little or no C8B (Hogasen et al. 1998). The homology between the porcine C8B and C8A protein is 28% on the basis of amino acid identities.

The primary sequence of porcine the C8G (GenBank acc. no. DQ333202) has been elucidated with 840 bp encoding 202 deduced amino acid residues (GenBank acc. no. ABD13970). The first 20 residues of the open reading frame of the human C8G show the typical features of a signal peptide using signalP analysis tool. The length of the human and porcine C8G protein sequence is the same and both share 82% identity. Using the ProtParam tool for the C8G molecule (C<sub>999</sub>H<sub>1582</sub>N<sub>288</sub>O<sub>281</sub>S<sub>5</sub>) 22.29 kDa molecular weight was predicted, similar to human C8G with 22 kDa (Steckel et al., 1980). Porcine C8G gene was not homologous to the other candidate genes in the terminal lytic complement pathway because it only contains a lipocalin domain (aa 48-184), which binds small hydrophobic ligands (Ortlund et al. 2002, Schreck et al. 2000). Both the human and porcine C8G have three cysteine residues at a.a 60, 96 and 188. The position of these cysteines is conserved in pig and human. The DiANNA 1.1 computer prediction (Ferre and Clote 2005, Ortlund et al. 2002) revealed a disulfide bond between Cys<sup>1</sup>-Cys<sup>3</sup>. According to Haefliger et al. (1987) the first cysteine (the 60<sup>th</sup> amino acid) in C8G protein sequence forms the disulfide bridge with C8A whereas another internal disulfide bridge exists between the second (the 96<sup>th</sup> a.a) and the third cysteine residues (the 188<sup>th</sup> a.a). The possible N-glycosylation site is at a.a 173 (Asn-Leu-Thr) using the NetNGlyc 1.0 Server tool (Figure 50).

### 5.1.3 Characteristic of the porcine C9 complement gene

Complement component C9 binds to C5b-8 sites on target cells and polymerizes to form the membrane attack complex (Taylor et al. 1997). In this study, the deduced porcine C9 protein sequence contains 543 amino acids (GenBank acc. no. ABD13966) showing

67% identity to human C9 (559 a.a). The first 21 amino acids may represent the signal peptide according to the SignalP 3.0 computer prediction tool. The domain within the first 16 amino acids at the amino-terminus of C9 is crucial in preventing the self-polymerization of the globular protein (Taylor et al. 1997). Porcine C9 (C<sub>2748</sub>H<sub>4301</sub>N<sub>753</sub>O<sub>836</sub>S<sub>30</sub>) has a weight of 62.23 kDa using the ProtParam tool whereas human C9 is  $M_r = 72$  kDa (Schreck et al. 2000). The porcine C9 has 19 cysteine residues corresponding to disulfide bonds Cys<sup>1</sup>-Cys<sup>11</sup>, Cys<sup>2</sup>-Cys<sup>14</sup>, Cys<sup>3</sup>-Cys<sup>19</sup>, Cys<sup>4</sup>-Cys<sup>5</sup>, Cys<sup>6</sup>-Cys<sup>15</sup>, Cys<sup>7</sup>-Cys<sup>9</sup>, Cys<sup>8</sup>-Cys<sup>13</sup>, Cys<sup>10</sup>-Cys<sup>18</sup>, and Cys<sup>12</sup>-Cys<sup>16</sup> as predicted with the DiANNA 1.1 analysis tool. The C9 protein has a carbohydrate content of 7.8% (Biesecker et al. 1982, 1980). Carbohydrate is usually attached to plasma proteins through an asparaginyl side chain within the characteristic sequence Asn-X-Ser/Thr (Clamp et al. 1975). Examination of the sequence of human C9 reveals only one sequence Asn-Glu-Thr (a.a 277-279) with an asparaginyl residue at position 277 likely to be a carbohydrate attachment site (DiScipio et al. 1984). This N-glycosylation site was not found in the porcine C9 protein. However by using the NetNGlyc 1.0 Server tool it showed that the porcine C9 has two N-Glycosylation sites at a.a 35-37 (Asn-Gly-Thr) in the amino terminus and (Asn-Val-Thr) in the carboxyl terminus (Figure 51). C9 water-soluble glycoprotein will insert quickly into lipid bilayers after binding to the membrane bound C5b-8 complex (Steckel et al. 1983, Hu et al. 1981, Podack et al. 1981). The C8A, C8B and C9 are members of a family of structurally related proteins that are capable of interacting to produce a hydrophilic to amphiphilic transition and membrane association (Howard et al. 1987). The principal binding site for C9 lies within the MACPF domain of C8A (Slade et al. 2006). The C9 polypeptide chain assumes a 'U' shape, in which the TSP1 and LDLa modules are located on the upper rim. The EGF module is located on the lower edge of the upper rim, and midsection of the polypeptide chain constructs the barrel of the tubule. TSP1 and LDLa modules do not participate directly in polymerization but cover the hydrophobic central region of the polypeptide chain in the monomer (DiScipio and Berlin 1999). In vertebrates, C9 and perforin form oligomeric pores that lyse bacteria and kill virus-infected cells, respectively (Rosado et al. 2007)

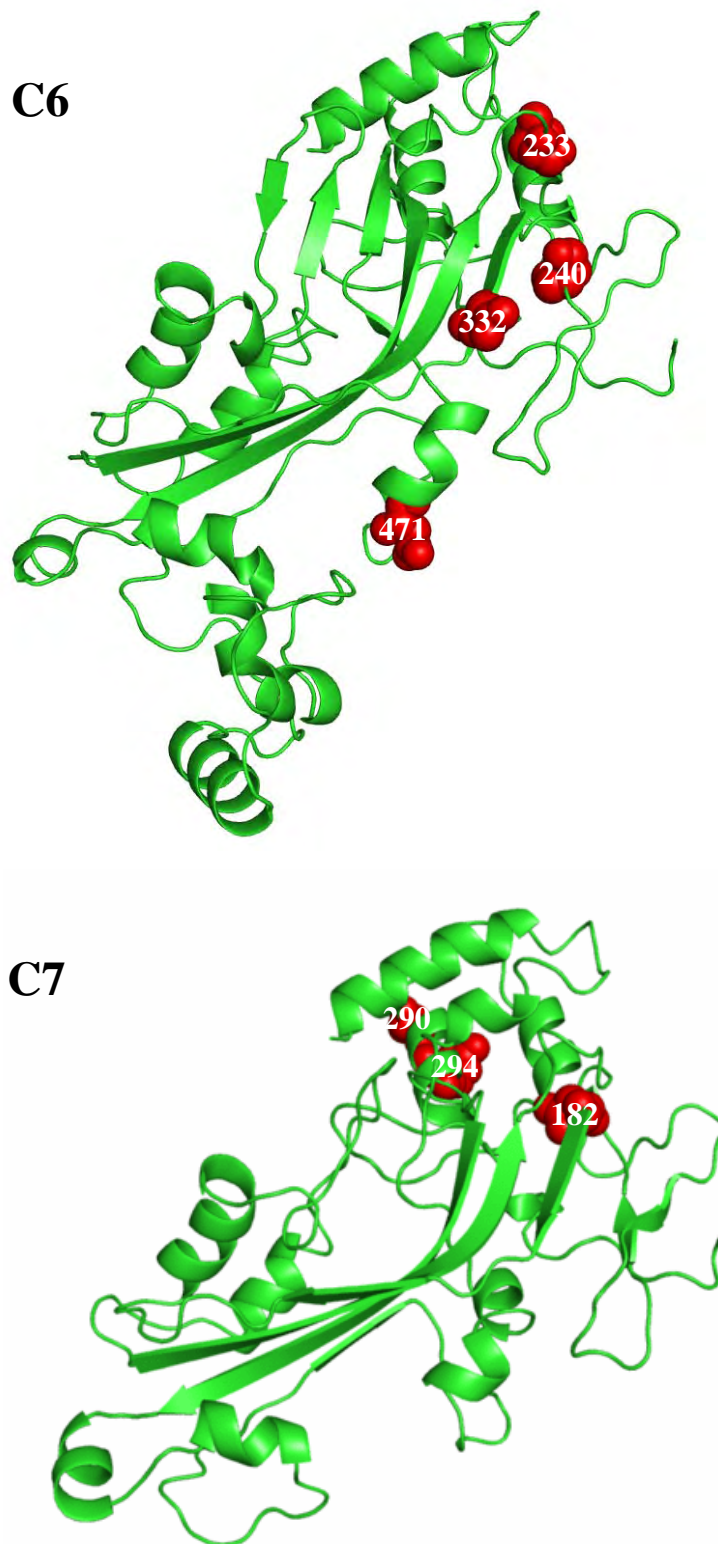


Figure 34 3D-structure of C6 (a.a 180-515), C7 (a.a 128-411), C8A (136-497), C8B (a.a 168-508) and C9 (122-506) homologous protein fragments built based on knowledge of structure features of homologues peptides and proteins with 3D-Jigsaw and displayed with Polyview 3D.

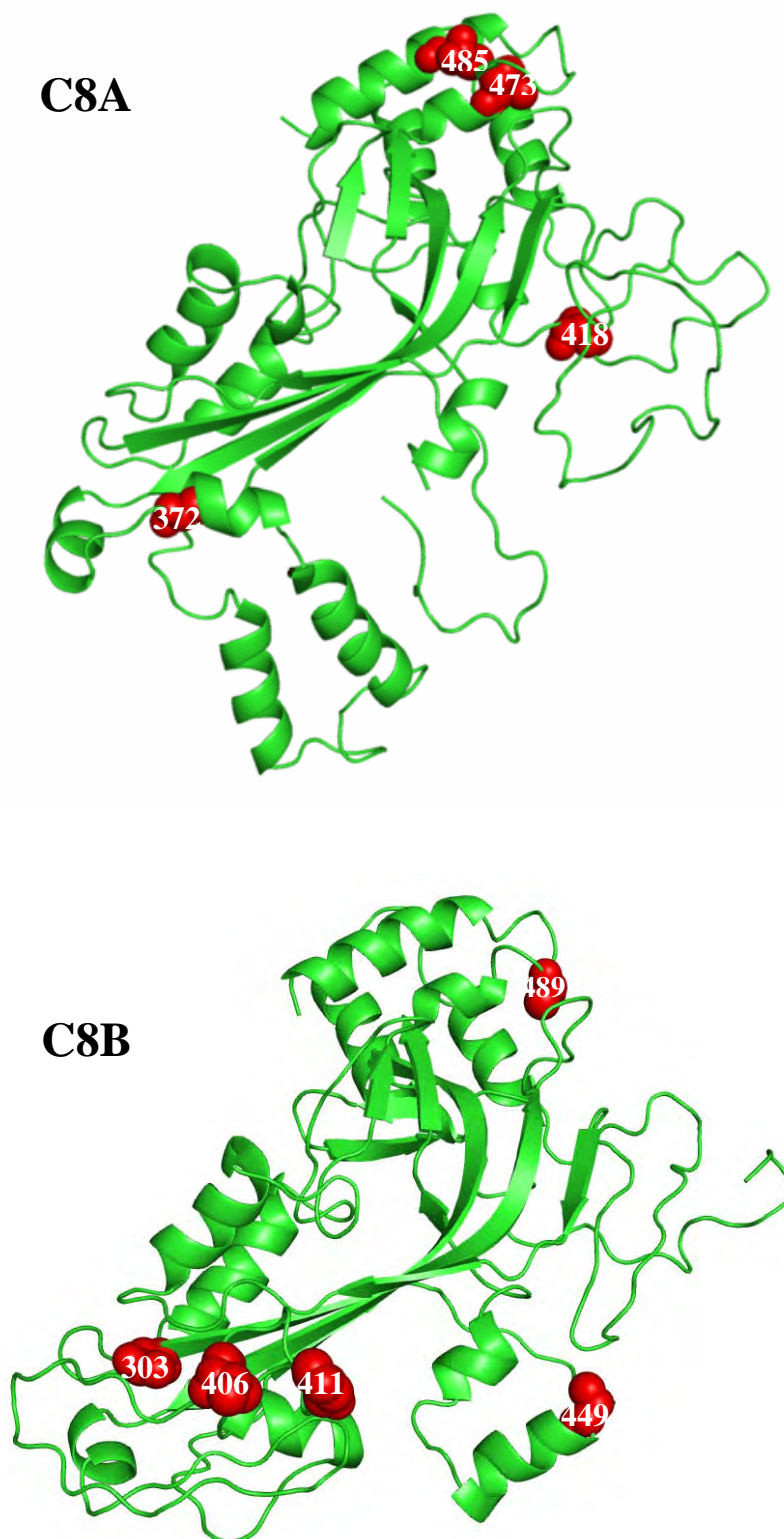


Figure 34 (continued) Position of amino acids at point mutations within 3D-protein regions is being numbered. Alpha and beta structures are in turns and sheets, respectively



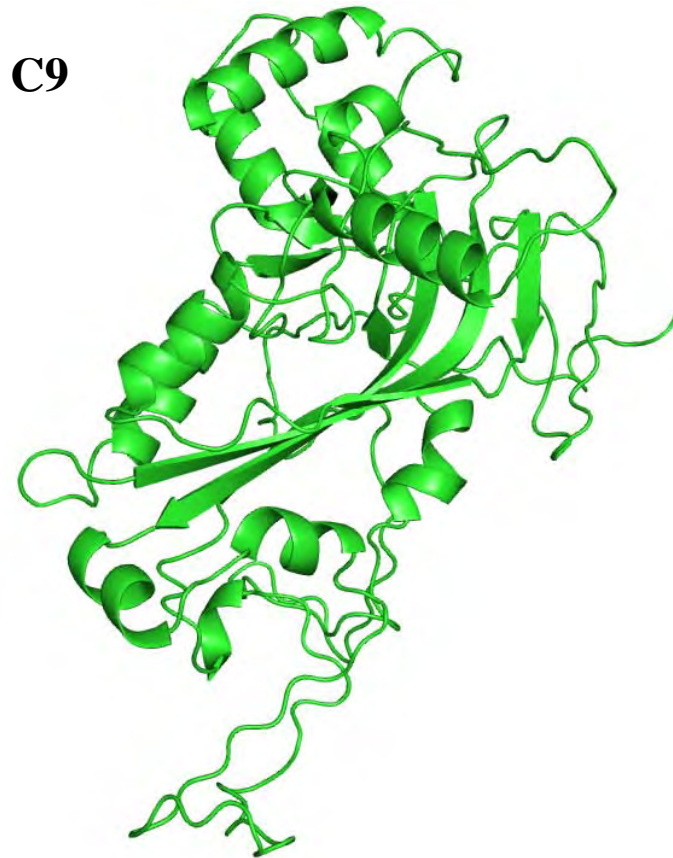


Figure 34 (continued)

#### 5.1.4 Common and specific structural and functional features of components of the terminal lytic complement pathway

It is really crucial to study the characteristic of the last complement proteins C6, C7, C8, C9 that act in concert to form the transmembrane pore at molecular level. The architecture of porcine C6, C7, C8A, C8G and C9 genes was identified and analyzed. Compared to human protein sequences the porcine C6, C8A, and C8B are one, five, and 20 amino acids longer, respectively, the porcine C9 is 16 amino acids shorter, whereas porcine C7 or C8G are of the same length as the human proteins. Actually, homology (22-31%) was found between amino terminus of C6 and C7 with whole mature C8A, C8B and C9 proteins (Figure 35 and Figure 36). The porcine C6, C7, C8A, C8B and C9 are mosaic proteins constructed from small cysteine-rich segments like TSP1, LDLa, EGF, CCP, FIMAC and MACPF whereas C8G individually composes of a lipocalin domain. The ubiquity of several types of cysteine-rich homology segments suggests that they have some utility in protein function (DiScipio et al. 1988). Respectively, 58, 45, 17, 18,

three, 13 of total 64, 56, 31, 27, three, 19 cysteines found in C6, C7, C8A, C8B, C8G and C9 are located in these functional protein domains. Each CCP (56-58 amino acids) contains four cysteines whereas six cysteines were found in each LDLa segment (37-38 amino acids). Within each CCP unit the two disulfide bonds are formed (1) between the first and the third cysteines, and (2) between the second and the fourth cysteines. Each short consensus unit is stabilized by small irregular groups of anti-parallel  $\beta$ -sheet interspaced by four or five tight turns (DiScipio et al. 1989). The FIMAC units are somewhat knotted by disulfide bonds but contain elements of conventional secondary structure (DiScipio et al. 1989). In both human and pig, the amino terminus of C6 has a tandem repeat of TSP1 whereas the carboxyl termini of C6 and C7 have a tandem of CCP and FIMAC. Each of five terminal complement mosaic proteins (C6, C7, C8A, C8B and C9) has a MACPF motif. Proteins containing MACPF domains play important roles in vertebrate immunity, embryonic development, and neural-cell migration (Rosado et al. 2007). The MACPF domain structure is similar with poreforming cholesterol-dependent cytolysins (CDCs) from gram-positive bacteria. These lytic MACPF proteins may use a CDC-like mechanism to form pores and disrupt cell membranes (Rosado et al. 2007). The CCP and FIMAC domains in C6 and C7 have the capacity to link specifically with C5 (DiScipio 1992, Haefliger et al. 1989, Ripoche et al. 1988, Catterall et al. 1987). Although the FIMAC is not required absolutely for C6 activity, this module promotes interaction of C6 with C5 enabling a more efficient bimolecular coupling ultimately leading to the formation of the C5b-6 complex (DiScipio et al. 1999, Haefliger et al. 1989). Binding of the FIMAC in C7 to the C345C domain in C5 is essential for incorporation of C7 into C5b-6 and hence for successful MAC assembly (Thai and Ogata 2005, 2004). The CCP modules mediate specific protein-protein and protein-carbohydrate interactions that are key to the biological function of the regulators of complement activation and, paradoxically, provide binding sites for numerous pathogens (O'Leary et al. 2004). The TSP1 conserved sequence motif has been suggested to play a key role in mechanisms by which malaria parasites avoid host defenses mediated by complement (Goundis et al. 1988). Both amino-terminal modules TSP1 and LDLa in C8A have a role in forming the principal binding site for C9 and that binding may be dependent on a cooperative interaction between these modules and the C8A MACPF domain (Scibek et al. 2002). However, Slade et al. (2006) indicated that principal binding site for C9 lies within the MACPF domain of C8A. The binding specificity between C8B and C8A-G

subunits is determined by a cooperative interaction of the asparaginyl-terminal TSP1 module and MACPF (Musingarimi et al. 2002). Site-specific antibodies directed to the LDLa homology unit in C9 cross-reacted with C6 and C7 (Tschopp et al. 1986). The porcine C8G is a lipocalin protein. According to King et al. (1994) some of the lipocalin proteins are allergens. Allergies are hypersensitivity reactions of the immune system to specific substances called allergens (such as pollen, stings, drugs, or food) that, in most people, result in no symptoms.

Based on obtained results it can be summarized that the complex of the porcine terminal complement components (C6-9) comprises of typical related mosaic proteins of high complexity and different size. These proteins contain modules being similar to segments of other proteins in the complex. The MAC can be regarded as a protein macromolecule from many smaller polypeptide sequences (C5b-9), which merge to each other due to interaction among their cysteine-rich homologous modules to form the membrane association. Homologous molecules of the terminal pathway (C6-9) originate from an ancestral gene. The porcine C6 is the largest protein in the terminal lytic pathway and most similar in structure to complement C7. Distance between porcine C8A and C9 is close on the phylogenetic tree. The C8A and C8B genes have correspondingly similar roles in MAC-mediated lysis of erythrocytes and bacterial killing and C8G is not required for expression of C8 activity (Schreck et al. 1998) and for complement-mediated killing of gram-negative bacteria (Parker et al. 2002). Presence of the MACPF module in all proteins C6 to C9 may have an important role in their combination to form in the stable complement complex as well as in lysis activity. The components of the MAC are physically formed in an associated arrangement C5b-C6-C7-C8B-C8A.G-C9. The characteristic pore-like appearance of the MAC is attributed to the association of as many as 12-18 C9 molecules that self-polymerize to form a circular structure in the membrane (Tschopp 1984). The individual components are hydrophilic proteins; however, when combined they form an amphipathic complex that is capable of binding to and disrupting local membrane organization (McCloskey et al. 1989). The increase in membrane permeability leads to osmotic lysis of simple cells such as erythrocytes or initiation of a variety of intracellular signaling events in the case of nucleated cells (Mold 1998). In bacteria, the MAC disrupts the outer membrane thereby increasing permeability and inducing lethal changes in the inner membrane (Esser 1994). Regulation of the mem-

brane attack pathway is essential to protect host cells from damage at sites of complement activation (Morgan 1999). In most tissues, ubiquitously expressed complement regulatory proteins prevent autologous destruction, protecting host cells from the powerful cytolytic activity of activated complement (Scolding et al. 1998).

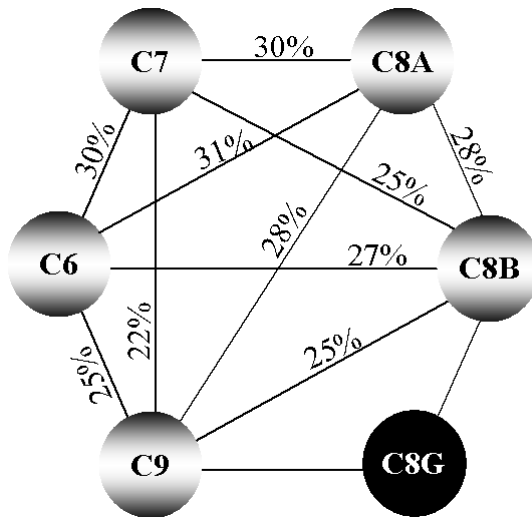


Figure 35 Comparison of the porcine C6 (a.a 22-631) and C7 (a.a 23-575) amino-termini with whole mature porcine C8A, C8B and C9 proteins using Clustal W alignment tool.

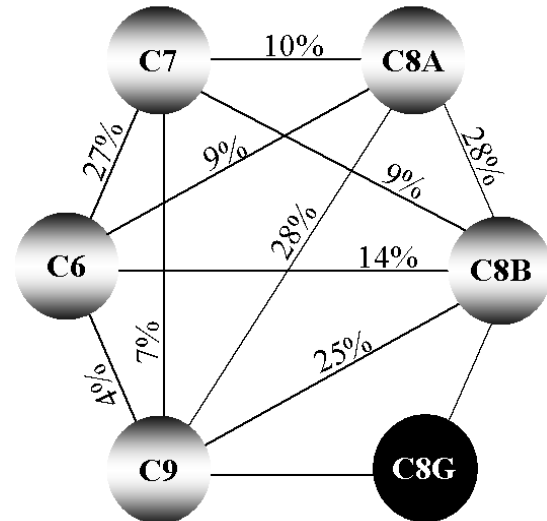


Figure 36 Comparison of the porcine C6 (a.a 632-935) and C7 (a.a 576-843) carboxyl-termini with whole mature porcine C8A, C8B and C9 proteins using Clustal W alignment tool.

## 5.2 Polymorphisms, genotyping and assignment of candidate genes to chromosome

Screening of the porcine C6, C7, C8A, C8B and C9, respectively, detected five, six, seven, nine, two single nucleotide polymorphisms (SNPs) in the coding regions, whereas there is no polymorphism within the porcine C8G when using one animal of each porcine breed (HS, DR, LR, PIE, BMP and MK). Most of the polymorphisms lie in cysteine-rich functional protein domains such as TSP1, LDLa, MACPF, CCP or FI-MAC. Especially, eight of nine SNP within C8B have amino acid substitutions including the SNP at nt 935 with three different segregating alleles (A→G→T). In the porcine C9 gene, the polymorphic site at nucleotide 407C→G corresponding to codon

106CAC→CAG (His→Gln) shows two different genotypes GG and CG in F<sub>2</sub> DUMI resource population but not in LR, PIE or MK. Allele 'C' maybe derived from Saddle-back Pigs, one of the breeds used to breed Berlin Miniature Pigs. These genetic variations may play an important role in the structural formation and function of the protein domains and therefore may affect the activity of the complement components in cell lysis. In order to evaluate genetic variation between Asian and European porcine breeds several SNP sites were selected for genotyping animals of porcine breeds LR, PIE, and MK. As a result, allele variety was found higher in LR and PIE (European breeds) than in MK (Asian breed) with different frequencies. Calculation and evaluation of all genotypic and allelic frequencies fit to Hardy-Weinberg equilibrium by Chi-square test. Genetic variations with amino acid substitution at loci that affect disease resistance traits found between European and Asian breed have great significance in breeding to improve pig health. Genetic variation in pigs has been shown in response to pathogens or immune system challenges (Henryon et al. 2002, Wilkie and Mallard 1999, Mallard et al. 1998) and in responses of pigs to infections with PRRSV (Petry et al. 2005).

In human the genes encoding the proteins of the membrane attack complex of the terminal lytic pathway (C5b-9) are required for complement bactericidal activity (Esser 1994, Schreiber et al. 1979). Genetic variations in C6 to C9 proteins associate with *Neisseria* infections (Ki et al. 2005, Witzel-Schloemp et al. 1997, Alvarez et al. 1995, Kaufmann et al. 1993, Figueroa et al. 1991). The polymorphic sites found in pig candidate genes are not conserved to ones associated with *Neisseria* infections in human.

Radiation hybrid mapping is known as a powerful tool for comparative gene mapping because gene assignments are made without the detection of genetic polymorphism as needed for linkage mapping. The assignment of genes using radiation hybrid (RH) panels is an efficient way to map genes and markers as well as to integrate the linkage and cytogenetic maps of a species (Hawken et al. 1999, Yerle et al. 1998). Application of the IMpRH panel allowed to unambiguously assigning the six loci to porcine chromosomes. The analysis revealed four new assignments and confirmed previous results. RH mapping result showed that C8A was assigned to 6q3.1-q3.5. This is also in agreement with previous result of Nakajima et al. (1998). The authors reported that the porcine

C8A gene physically maps to chromosome 6q3.3-3.5 according to in situ hybridization using the porcine bacterial artificial chromosome (BAC) clone as a hybridization probe.

The porcine C9 has been mapped to 16q1.4 by fluorescence in situ hybridization (Thomsen et al. 1998), which is confirmed by our results. Genetic mapping correspondingly revealed linkage of C6 and C7 to C9 and AGXT2 that have previously been assigned to the q arm of Ssc16 (Ponsuksili et al. 2001, Thomsen et al. 1998, Wintero et al. 1998) while on chromosome 6q3.1-q3.5 C8B links closed to C1q and C8A that have been reported by Jorgensen et al. (1997) and Nakajima et al. (1998), respectively. This is in agreement with the most recent human-porcine comparative map (Meyers et al. 2005).

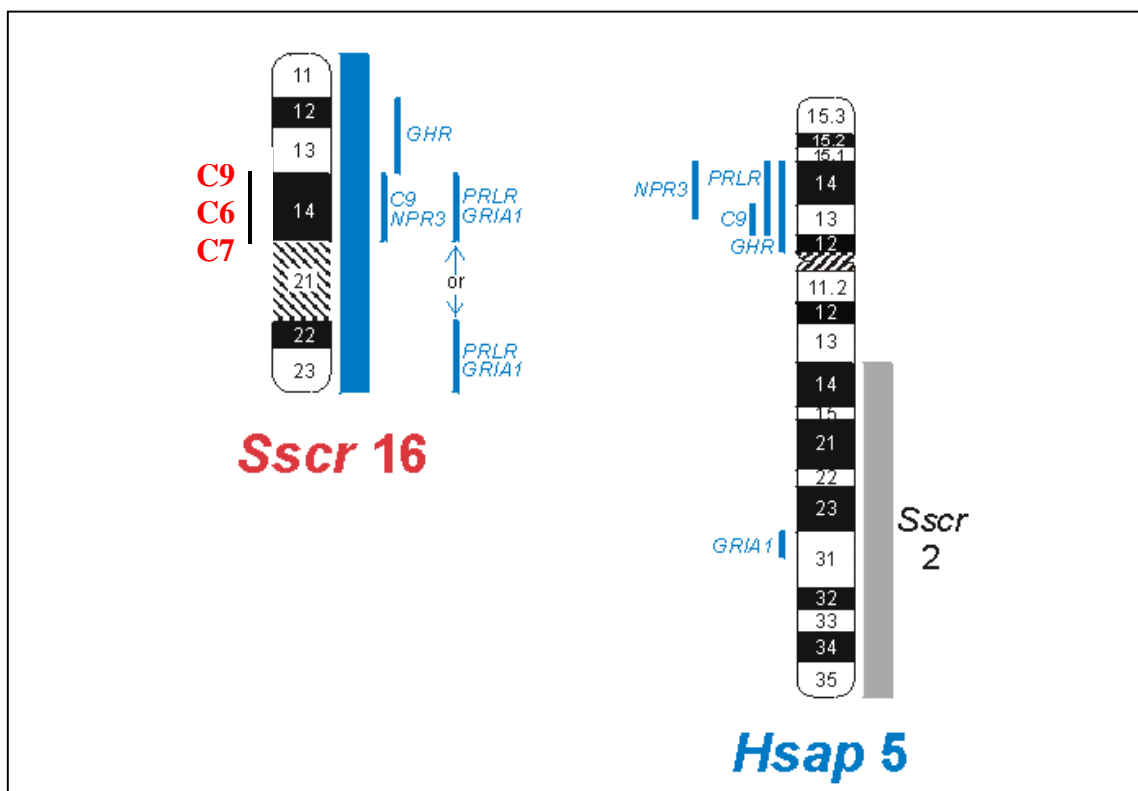


Figure 37 Correspondences between human and pig chromosomal segments. Location of the porcine C6, C7 and C9 gene on the q-arm of chromosome 16 (Ssc16q1.4) adapted from <http://www2.toulouse.inra.fr/lgc/pig/compare/SSCHTML/SSC1B.HTM>

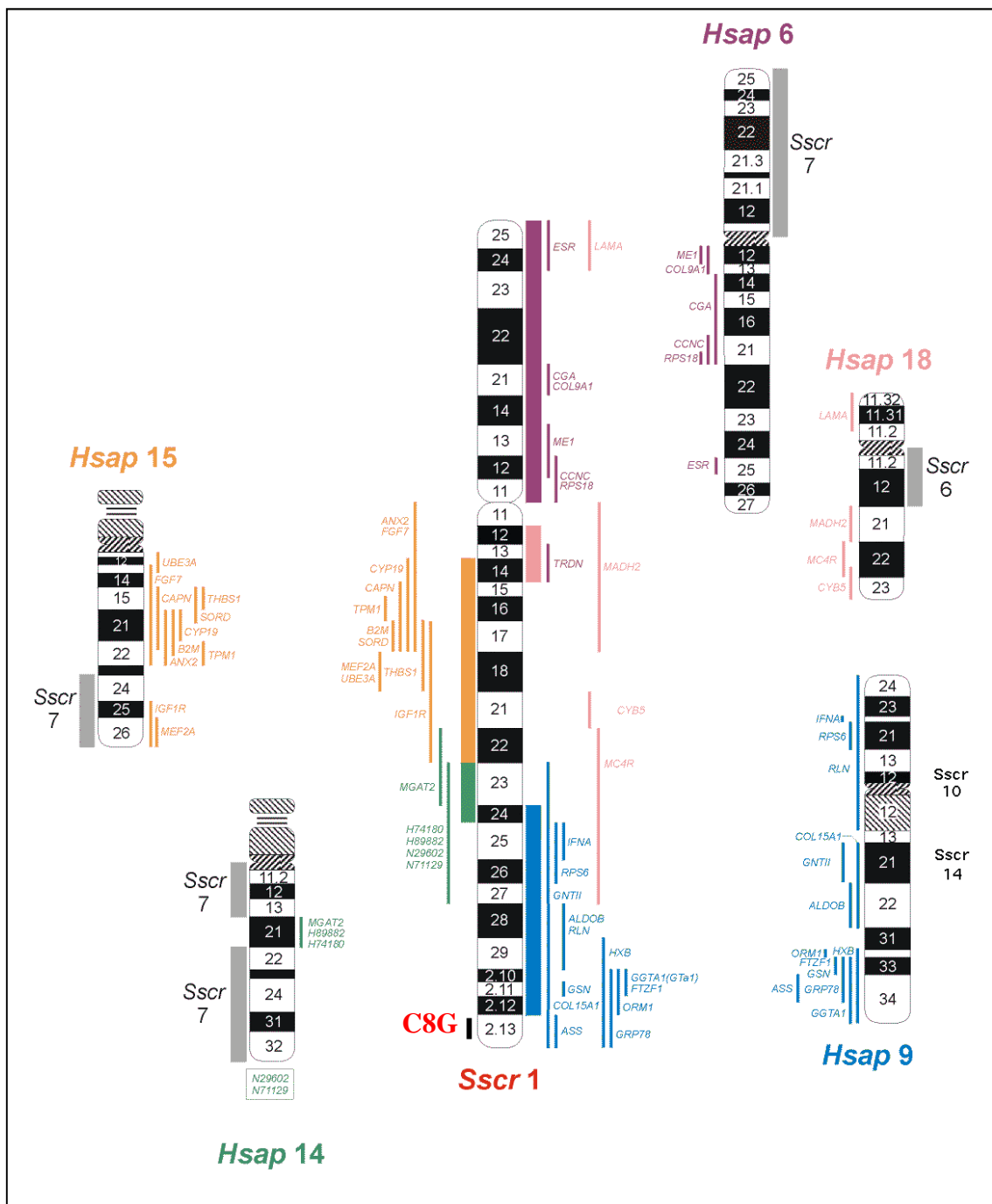


Figure 38 Correspondences between human and pig chromosomal segments. Location of the porcine C8G gene on the q-arm of chromosome 1 (Sscr1q2.13) adapted from <http://www2.toulouse.inra.fr/lgc/pig/compare/SSCHTML/SSC1B.HTM>

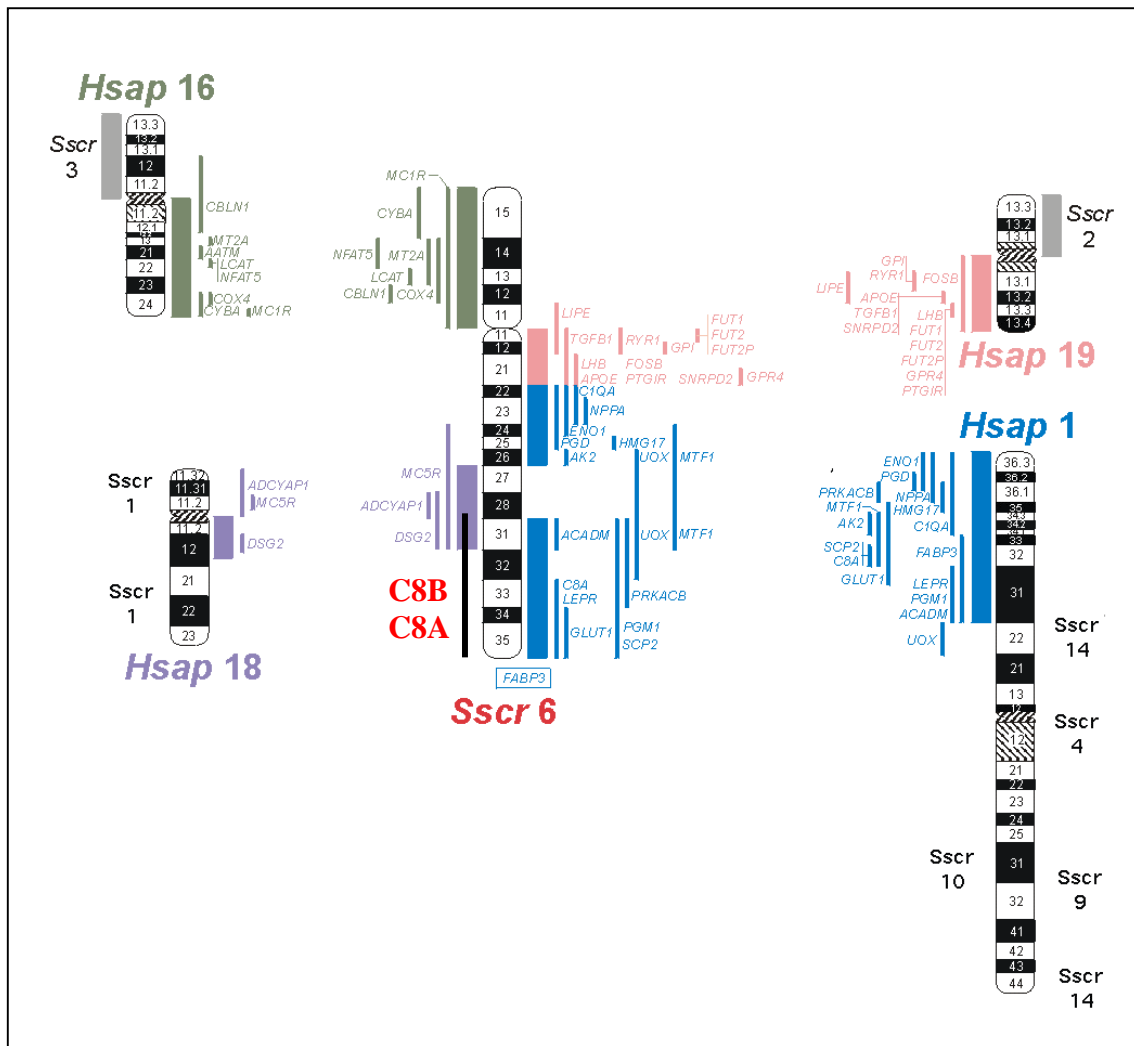


Figure 39 Correspondences between human and pig chromosomal segments. Location of the porcine C8A and C8B gene on the q-arm of chromosome 6 (Ssc 6q3.1-q3.5) adapted from <http://www2.toulouse.inra.fr/lgc/pig/compare/SSCHTML/SSC1B.HTM>

The mapping results also fit the current pig-mouse and pig-rat comparative maps as accessible via the MGI webpage ([http://www.informatics.jax.org/searches/homology\\_form.shtml](http://www.informatics.jax.org/searches/homology_form.shtml)). Further, the location of distinct C8 loci on chromosomes supports genetic evidence that C8 contains three separate genes encoding different proteins C8A, C8B and C8G.



Looking on the genetic maps of other species the sixth and seventh complement components are mapped to the same chromosome. Also C8A and C8B components are always found on the same chromosome. Additionally C9 complement component could be mapped to the same chromosome carrying C6 and C7 in all species addressed here except in *B. taurus* and *D. rerio*. The C8G gene is always located in a different chromosome (Table 20)

Table 20 Chromosomal assignments of the genes in difference species

	Mapping localization								
	<i>B. taurus</i>	<i>D. rerio</i>	<i>G. gallus</i>	<i>H. sapiens</i>	<i>M. mulatta</i>	<i>M. musculus</i>	<i>R. vegicus</i>	<i>nor-troglodytes</i>	<i>P.</i>
C6	20	21	-	5p13	6	15A1	2q16	5	
C7	20	21	Z	5p13	-	15A1	2q16	5	
C8A	3	-	-	1p32	1	4C6	5q34	1	
C8B	3	-	8	1p32	1	4C6	5q34	1	
C8G	-	24	-	9q34.3	-	2A3	3p13	-	
C9	1	5	-	5p14-p12	6	15A1	2q16	-	

### 5.3 Association analysis with hemolytic complement activity

#### 5.3.1 Hemolytic complement activity depending on genotypes of the candidate genes

In human, deficiencies of candidate genes C6, C7, C8 and C9 have been studied and associated with meningitis infection in many different populations (Takata et al. 1989, Cooke et al. 1987, Keller et al. 1987, Moolenaar et al. 1987, Vogler et al. 1979). The studies have provided important information about the role of complement in the defense against pathogens, agents of infectious diseases, especially meningococcal disease. Here we refer to the association and linkage of mutation points of the terminal complement components C6, C7, C8 and C9 with immune response of two kinds of pathogen, bacteria (*Mycoplasma*) and virus (Aujeszky and Porcine Reproductive and

Respiratory Syndrome) before and after vaccinations for hemolytic complement activity in both classical and alternative pathway in pigs.

In the porcine C6 candidate gene, studying genetic association with hemolytic complement activity was carried out for the SNP 862A→G, causing the 240Asn→Asp amino acid substitution. Difference of genotypes in CH50 test was close to significance ( $p=0.0853$ ) although along the vaccination program homozygous genotype GG always shows higher lysis levels. The allelic frequency of 'G' was found to be higher in MK than in other breeds. No evidence was found for significant association with hemolytic performance between the C6 genotypes and AH50. Comparing effects of different alleles on hemolytic complement activity in both pathways, one might tend to regard the 'G' allele as favorable and thus carriers of the allele or breed with high 'G' allele frequency as valuable genetic resources.

In porcine C7 gene, the SNP site at position 881A→G (294Lys→Arg) was genotyped and associated with hemolytic complement activity in the F<sub>2</sub> DUMI resource population. It showed that lysis activity reached maximum level at the day 4 post-ADV vaccination in CH50 and at day 10 after PRRSV immunization in AH50. There is highly significant difference between genotypes in CH50 with the allele 'A' being favorable in terms of complement activity ( $p=0.0080$ ) while there was no difference in AH50. It indicated that the time course of hemolytic complement activity in the classical pathway depends on the C7 genotypes. In the DUMI resource population the allele 'A' originates from the Berlin Miniature Pig grandparents. Berlin Miniature Pigs were bred from Vietnamese potbelly pig, Saddleback Pig and German Landrace with the first being the likely source of the 'A' allele. Interestingly this allele is more frequent in the Muong Khuong compared to German Landrace and Pietrain. According to Clapperton et al. (2005) levels of innate immune traits were shown to differ between different pig breeds although it is unknown whether this effect would lead to breed differences for resistance to infectious diseases.

In the porcine C8A gene, association of C8A genotypes at SNP site 1544C→T (485Arg→Cys) with hemolytic complement activity was studied. As a result, the association of C8A with CH50 was close to significance ( $p=0.0650$ ), however the difference

between the two homozygous genotypes CC and TT was significant at  $p=0.0522$ . Animals with genotype TT show strong lysis ability at the early time point before Mh vaccination and this hemolytic activity showed a linear increment during experiment. At later time points the performance of heterozygous animals is similar to animals of the TT genotype, whereas at the early time points the hemolytic complement activity of heterozygous animals was as low as for the CC genotype. This might indicate dominance effects of the allele `T` that are different in their direction depending on the vaccination. Additionally, interaction between genotypes and time points caused significant difference in AH50 test ( $p=0.0027$ ). This showed that hemolytic complement activity in alternative pathway depends on C8A during the time course of vaccination.

In the porcine C8B gene, analysis of CH50 hemolytic complement activity depending on C8B genotypes at SNP 222 (65Thr→Met) in animals of the DUMI population along the vaccination program demonstrated that at the beginning of the vaccination program, before and after Mycoplasma vaccination, the hemolytic complement activity in animals with genotype CT ( $42.8\pm 3.4$ ) was the lowest compared to CC ( $51.5\pm 3.4$ ) or TT ( $52.5\pm 5.7$ ). At time points around ADV vaccination differences depending on C8B genotypes were not obvious whereas at PRRSV vaccination genotype CT ( $71.2\pm 4.4$ ) vs CC ( $68.8\pm 3.9$ ) or TT ( $62.3\pm 7.2$ ) revealed highest value, i.e. `T` allele really exposes its own maximal hemolytic ability in the heterozygous pattern at this point time. Also the allele `T` performs its own optimum lysis ability compared with remainders at day 10 after Mh or ADV immunization in the alternative complement pathway. However, the difference for CH50 and AH50 activity among genotypes at particular time points was not statistically significant. The `T` allele at nucleotide 222 in the DUMI populations originates from the Berlin Miniature Pig grandparents. Berlin Miniature Pigs were bred from Vietnamese potbelly pig, Saddleback Pig and German Landrace (Hardge et al. 1999) with the first being the likely source of the `T` allele. The time course of hemolytic complement activity depends on the C8B genotypes. Different effects of C8B genotypes on hemolytic complement activity associated with particular vaccinations against bacterial and viral pathogens could be due to host-pathogen-interaction. Vietnamese local breeds are a source of promising alleles of unpredictable economic value (Lemke et al. 2005). In this case it can suggest that allele `T` presenting in Muong

Khuong population is a valuable genetic resource, which provides new perspectives to improve innate immune response in pigs.

Screening the full-length cDNA sequence of the porcine C9 (GenBank acc. no. DQ333198) revealed two SNP sites: 350A→G (codon 87CAA→CAG) and 407C→G (106CAC→CAG). The second SNP has an amino acid substitution His→Gln whereas the first is a silent mutation. Therefore, the second was investigated to associate with hemolytic complement activity in both classical and alternative pathway. For the porcine C9, significant difference in hemolysis was found between two genotypes GG (69.49±3.17) and CG (63.59±3.65) for the classical (p=0.0488) but not for the alternative pathway (62.73±2.82 vs 61.52±2.39, respectively). However, statistically significant difference was found in the interaction of genotypes and immunization point times (p=0.0340) in the alternative pathway. Appearance of allele 'C' in heterozygous CG, probably from Saddleback Pigs, has made decreasing lysis activity whereas homozygous GG always shows hemolysis at a higher level during experiment in the classical pathway. CH50 tests showed the highest lysis value at the fourth day after ADV vaccination in both homozygous GG genotype (83.09±3.85) and heterozygous CG (78.12±4.72), whereas AH50 can be seen as a linear increasing line during the experiment. The genotype GG shows high variation in the level and reactivity of the complement cascade along the experiment. It could come up to the highest value (83.09±3.85) at the fourth day post ADV immunization in CH50 and go down to one of the lowest values (47.81±3.10) at the fourth day post Mh immunization in AH50. Due to their different impact on the classical and alternative pathway the C9 genotypes are of high interest for further studies.

### 5.3.2 Other factors on hemolytic complement activity in both pathways

The complement system comprises over 30 distinguished proteins and is known as a biochemical cascade for lysis invading external pathogens. These proteins are primarily formed in the liver and circulate under inactive form in blood plasma. When there is, however, presence of attacking pathogens these proteins will become active. As an activated enzyme, prior protein bind to subsequent protein to form the membrane attack complex (C5b-9). Of course, concentration of MAC will be higher in blood plasma and

on membrane of target cell at that time. Then the complex inserts into the cell membrane, punches a hole, increases leakage of water and ions through the cell membrane, initiates osmotic cytolysis and kills target cells. In turn, C5b-9 complex exerts a regulatory effect on the formation of C3 convertases of the classical and alternative pathway and on the utilization of C5 by cell-bound C5 convertases. Feedback inhibition by C5b-9 represents a biologically relevant mechanism through which complement may auto-regulate its effector functions (Bhakdi et al. 1988).

Due to the important role of the complement system in natural immune regulation mechanism of host body against pathogens, many studies for the complement components have been carried out; for example in bacterial meningitis infections in which the C5-9 complex attacks the gram-negative bacteria (Vázquez-Bermúdez et al. 2003, Zhu et al. 2000, Eng 1980, Haeney et al. 1980). The morbidity of meningococcal infections may be attenuated in the presence of reduced levels of the membrane attack complex (Ross and Densen 1984). Deficiencies of complement components of the terminal pathway predispose to both autoimmune disease and infections were also reported. Causes of deficiency can be inherited (Toshiro et al. 1999, Inai et al. 1989, Fukumori and Horiuchi 1998) or acquired and complete or partial. Moreover, deficiency of membrane attack complex proteins of the complement system often appears common in countries where bacterial diseases are highly prevalent (Zimran et al. 1987, Orren et al. 1987). Most of deficiencies are associated to polymorphisms present in the molecular structure of at least one complement component. Vissher et al. (2002) also provided evidence for substantial genetic variation between pigs and concluded that there is the possibility for genetic improvement of the immune capacity and response to pathogens.

Since the complement system plays an important role in natural defense system of pigs, genetic variation of C3, C5 as well as mannose-lectin genes (MBL1 and MBL2) were associated with hemolytic complement activity (Phatsara et al. 2007, Kumar et al. 2004, Mekchay et al. 2003, Wimmers et al. 2003). On the same experimental materials, here it was focused on the porcine C6, C7, C8A, C8B, and C9 genes. Not as other complement components with dependently specific functions, the C6-9 assembly has the same function of cell lysis because all they are members of the MAC macromolecule. This means that reciprocal influence among the terminal complement components on hemolytic

complement activity is indispensable. Catalytic activity of previous components could affect not only the subsequent others but also the activity of whole MAC complex. The study showed significant differences between genotypes of several genes and hemolytic complement activity in the classical pathway but not in the alternative pathway for all candidate genes. Performance of genetic variation of the terminal complement components C6-9 on hemolytic complement activity in the classical and alternative pathway is complicated. It is really indirect to use CH50 and AH50 for evaluating genetic effect of the terminal candidate genes on hemolytic complement activity because direct products at the end of the complement activation from three pathways (classical, lectin, and alternative) are C5-convertases, the essential enzyme to split C5 into two smaller fragments C5a as a potent anaphylatoxin and C5b as an event center for binding to C6 followed by C7-9. However measurement of CH50 and AH50 is the best screening test for deficiency of the complement components of classical, alternative or terminal pathway. Beside the genotypes of candidate genes, a number of other factors that affect complement activity were evident along the experiment. These factors have been taken into account in the analysis using adequate statistical models. Some of these factors may contributed to the fact that the association of the candidate genes of the terminal lytic pathway of the complement cascade, that is common to the classical and alternative pathway of complement activation, was shown for CH50 but not for AH50 for all genes except C8G. Factors such as gender and age of the animals, the used vaccines, and interaction of genotypes and vaccinations, interaction among complement genes themselves and also other immune genes that are discussed here may increase the power to detect effects on hemolytic complement activity measured in the classical pathways but mimic these effects in the alternative pathway.

#### 5.3.2.1 Effect of gender

Higher hemolytic activity was found in male animals as compared to in the female animals in both pathways. Analysis of variance also indicated significant differences between genders. According to Olaho-Mukani et al. (1995) sex has an effect on complement activity.

### 5.3.2.2 Effect of age

The complement system is immature at birth. Repeated antigenic stimulation leads to the complete maturation of immunity during the first few years of life (Durandy 2003). Clearly, hemolysis showed low hemolytic values before primary immunization with Mh and reach higher ones after the following vaccinations with ADV and PRRSV. In human (Yonemasu et al. 1978) and in pigs (Olaho-Mukani et al. 1995, Tyler et al. 1988) age has an effect on complement component concentration, i.e. observation on serum concentration of C3 in healthy Japanese Yonemasu et al. (1978) reported that the concentration reaches the lowest value at 24 hours after birth ( $680.0 \pm 45.7 \mu\text{g/ml}$ ) and comes up the highest at one year old ( $1385 \pm 65.3 \mu\text{g/ml}$ ). The age of the patient affects interpretation of complement levels in various diseases (Yonemasu et al. 1978). Short term elevation of complement activity after immunizations is an (sub-)acute phase response of the immune system and the long term increase between vaccinations may be due to aging (Wimmers et al. 2003). Pig breed and age are important factors influencing the response to various stressors or infectious challenges (Sutherland et al. 2005)

### 5.3.2.3 Interaction of complement components

Due to aging as well as accumulation after various vaccinations the complement system is increasingly activated and responsive. Activities of the terminal complement components C6 to C9 depend on the activity of other components acting on higher levels of the cascade. Especially the formation of the C3 convertase and subsequently the C5 convertase are key to the final overall hemolytic complement activity. The release of C5b from C5 that binds to C6-9 to form the membrane attack complex is cause of interaction of different protein sequences in the three pathways of complement activation. The hemolytic complement activity in the alternative pathway partly depends on C3b formed after the complement activation of the classical pathway. The C3b plays an important role for accelerating complement activation in the alternative pathway. Hemolysis in subsequent time points was higher than the previous along the experiment for AH50 due to both antigen-specific and antigen-nonspecific immune responses and responsiveness that accumulated after each vaccination. Clearly, generation of C3-convertase has directly strong effect on the formation of the MAC (C5b-9). Of course,

complement activation also depends on other signals of the immune system like interleukins. Vaccinations are known to effect acute phase response through cytokines like IL1, IL6 and TGFB1 that are released during immune response and in particular trigger the expression of C3 (Gonzalez-Ramon et al. 2000, Mackiewicz et al. 1990, Castell et al. 1989).

CRP known as an acute-phase serum protein and a mediator of innate immunity binds to microbial polysaccharides and to ligands exposed on damaged cells. Surface bound CRP reduces deposition of and generation of C5b-9 by the alternative pathway and deposition of C3b (Mold et al. 1999). CRP is most efficient at early classical pathway activation and reacts with Fc  $\gamma$  receptors (Marnell et al. 1995, Crowell et al. 1991, Berman et al. 1986). CRP binding to C1q differs from IgG binding in being localized to collagen-like regions rather than the globular head groups of C1q (Jiang et al. 1991)

#### 5.3.2.4 Effect of vaccination

It was shown that vaccination leads to an acute phase response measurable by an increased serum level of haptoglobin in pigs (Rekitt et al. 2001, Asai et al. 1999). Increment of the C3c serum concentration, as observed in this experiment, is part of the acute phase response (Volanakis et al. 1995, Kusher et al 1982). At the same time it reflects in vivo complement activation, while hemolytic complement activity describes the in vitro capacity of the complement system (Storm et al. 1992). Also hemolytic complement activity is elevated after immunizations. Various structures of lipopolysaccharides (LPS) of gram-negative enterobacteria, influence the complement activation. The rough LPS of *Coxiella burnetii* variant activated complement via the alternative pathway, whereas the intermediate LPS variant activated the classical pathway (Vishwanath and Hackstadt 1988). LPS of various *Mycoplasma* species have been shown to interact with C1 and induce an antibody-independent activation of the complement system via the classical pathway (Rosendal et al. 1994, Brecht et al. 1977). A previous study showed that *Brucella abortus* do not activate the bovine alternative pathway (Hoffmann and Houle 1983). The rough (O-antigen deficient) *Brucella abortus* will bind C1 and be killed by the classical pathway (Eisenschenk et al. 1999) whereas the alternative pathway does not kill *Brucella abortus* (Eisenschenk et al. 1995). For *Salmonella minne-*



*sota*, it was shown that the bacteria avoided killing by complement because membrane attack complexes were rapidly sloughed off. Although C3b as well as the stable complexes C5b-7 on the surface of *Salmonella minnesota* was formed (Joiner et al. 1982a), subsequent additions of C8-9 caused release of C5b-7 complexes (Joiner et al. 1982b). Furthermore, the lipid A region of the LPS is responsible for classical pathway activation not to depend upon antibody to the lipid A whereas alternative pathway activation proceeds by a lipid A-independent mechanism (Morrison and Kline 1977). Lipid A of gram-negative bacteria binds and activates C1 to initiate the classical pathway (Betz et al. 1981). In the absence of anticapsular antibody, complement-mediated killing occurred but not by the alternative pathway (Steele et al. 1984). In reverse way, an encapsulated strain of *Escherichia coli* was shown to be killed primarily by antibody-dependent activation of the classical pathway (Taylor et al. 1983) whereas an unencapsulated rough strain of *Escherichia coli* activated the classical pathway independently of antibody (Betz et al. 1981). There also is evidence that killing of bacteria by the classical pathway system did not require specific antilipopolysaccharide antibodies and the chemical compositions of both the capsular polysaccharide and the O-antigen determine the degree of complement activation (Pluschke and Achtman 1984). O-antigen prevented both complement component C3 deposition on the surface and complement-mediated killing of *Bordetella parapertussis*. In addition, O-antigen was required for *Bordetella parapertussis* to systemically spread in complement-sufficient mice, but not complement-deficient mice (Goebel et al. 2008). This can cause deficiency of C3 for the alternative pathway activation. In other studies it is reported that porin protein of bacteria binds to C1q (Alberti et al. 1996, Loos and Clas 1987). The association of LPS and porins in the absence of specific antibodies is a prerequisite for the activation of the classical pathway of complement leading to the killing of serum-sensitive bacteria (Loos and Clas 1987). Clearly the mechanisms of complement activation by bacteria are quite different depending on characteristics of the bacterial surface such as lipopolysaccharides, lipid A, O-antigen of LPS, porin, ect. However, also viruses were shown to activate the complement cascade independent from existing antibodies via both the classical and the alternative pathway. Alphaherpesviridae, in particular Epstein-Barr Viruses, is an example for complement activation in both pathways (Mold et al. 1988, Mayes et al. 1984).

Furthermore, the glycoprotein C (gC) of HSV-1 and -2, bovine herpesvirus type 1, porcine herpesvirus 1 (ADV), and equine herpesvirus types 1 and 4 interact with C3b (Friedman et al. 2000). Glycoprotein C known as a receptor for C3bBb impedes decay-accelerating activity for the alternative pathway C3 convertase and could block the effects of complement on the infected cells in the absence of antibody (Fries et al. 1986, McNearney et al. 1987). The complement activity differs between animal species infected by HSV. It is due to the difference in affinity between the gC of HSV and either the complement component or glycoprotein structure (i.e. a difference in glycosylation level) in each animal (Hidaka et al. 1991). Thus the increment of complement activity in blood samples obtained after vaccinations might be at least partly due to specific interaction between the bacterial and viral vaccination strains and complement components. It may be conducted that activation of either the alternative or the classical pathway in the experiment differentially depends on the vaccinations. Moreover, the immunization with Mycoplasma is a booster vaccination with the first application-taking place in the first week of life. For ADV it has been shown that maternal antibodies are present until week 14 and 15 of age (Pensaert and Kluge 1989, Wittmann 1984). ADV immunization was done when maternal antibodies may still be present. Those mean that at the time point of Mh and ADV vaccination an immediate activation of the complement system via the classical pathway might have happened due to the formation of specific antigen-antibody complexes.

All factors discussed above push hemolytic complement activity in the classical pathway significantly stronger than in the alternative pathway. Consequently, the experiment may have more power to elucidate effects on the hemolytic complement activity in the classical pathway than in the alternative pathway.

In summary, in the present study the porcine genes encoding the terminal components of the complement cascade C6, C7, C8A, C8B, C8G, and C9 were characterized. The cDNA molecular structure was sequenced, polymorphisms detected, assignment to the porcine chromosome done, and association between SNPs and hemolytic complement activity evaluated. Genetic variations between European and Vietnamese porcine breeds are valuable. In the present study, Muong Khuong pig is also owner of genotypes which performed strong hemolytic complement activity. Therefore these genotypes can be

considered as the candidates to enhance natural resistance in pig in future. These genotypes might be selected through many generations of Muong Khuong pig raised in bad conditions (no vaccination, insufficiency of feed and nutrient unbalance ration, bad hygiene). Also these are result of natural selection of host against infectious environmental condition. Interestingly, variations of complement activities in Meishan breed were higher than in European breeds (Duroc, Landrace and Large White) (Komatsu et al. 1998). Moreover, high similarity in molecular structure of the candidate genes among other species may open a new way to treat animals and humans with deficiency of complement components even with newly designed proteins. The obtained results provide the means for further understanding the role of C6, C7, C8, and C9 in natural immune response of the host against pathogens. It also promotes the porcine C6, C7, C8, and C9 as candidate genes in efforts to genetically improve general animal health, a goal of breeding programmes for food animals. The results also enrich evidence for genetic variation in host responses to vaccinations. The age-old principle should be regarded in breeding and selection for resistance to diseases.

## 6 Summary

This study aimed to characterize the porcine genes encoding the terminal components of the complement cascade C6, C7, C8A, C8B, C8G, and C9. Therefore the entire cDNA sequences were identified; SNPs were detected within candidate genes; genetic linkage and radiation hybrid mapping was performed; and association between SNPs and hemolytic complement activity was analyzed.

The cDNA sequences of the candidate genes C6, C7, C8A, C8B, C8G and C9 was identified with 3306, 3561, 2146, 2461, 840, and 2536 bp, including 144, 807, 91, 28, 51, and 89 bp of 5' UTR and 354, 122, 285, 597, 180 and 815 bp of 3' UTR region, respectively. Alignment of cDNA sequences with human and porcine genomic sequences allowed deducing 18, 18, 11, 12, seven, and 11 exons of the porcine candidate genes C6, C7, C8A, C8B, C8G and C9 encoding 935, 843, 589, 611, 202 and 543 amino acids, respectively. The porcine cDNA and deduced protein sequences of the candidate genes showed high (65-86% and 67-83%, respectively) identities with human analogues. A number of functional domains such as TSP1, LDLa, MACPF, EGF, CCP or FIMAC holding different roles in the formation of membrane attack complex (MAC) and cell lysis were detected.

Screening the cDNA sequences using comparative sequencing revealed five, six, seven, nine and two single nucleotide polymorphisms (SNPs) within the candidate genes C6, C7, C8A, C8B and C9, respectively. These SNPs were confirmed using either PCR-RFLP or resequencing method. There are two (240Asn → Asp, 471Asp → Glu), three (52Ile → Val, 67Thr → Met, 294Lys → Arg), two (485Arg → Cys, 528Tyr → Cys), eight (24Pro → Leu, 65Thr → Met, 303Thr → Gly → Ser, 406Ile → Val, 411Pro → Ser, 449Ala → Val, 489Ala → Val, 590Ala → Val), and one (106His → Gln) SNPs with amino acid substitution in the porcine C6, C7, C8A, C8B and C9, respectively. No SNP was found in C8G among the breeds Hampshire (HS), Duroc (DR), German Landrace (LR), Pietrain (PIE), Berlin Miniature Pig (BMP) and Muong Khuong (MK). Most of the polymorphisms lie in cysteine-rich functional protein domains such as TSP1, LDLa, MACPF, CCP or FIMAC. These genetic variations may play an important role in the

structural formation and function of the protein domains and therefore may affect the activity of the complement components in cell lysis.

The polymorphic sites 862A→G (240Asn →Asp, TaqI) and 1557C→G (471Asp→Glu, PstI) for C6, 154A→G (52Ile→Val, BsrDI), 870C→T (Hin6I) and 881A→G (294Lys→Arg, MboII) for C7, 535A→G (TfiI), 1544C→T (485Arg→Cys, Hin6I), and 1768C→T (524Tyr→Cys, KnpI) for C8A, 222C→T (65Thr→Met, FnuDII), 935A→G→T (303Thr→Gly→Ser, resequencing), and 1244A→G (406Ile→Val, MaeII) for C8B, 350A→G (BsrDI) and 407C→G (106His→Gln, HpyCH4III) for C9 were genotyped using PCR-RFLP or resequencing in unrelated animals of the breeds German Landrace (n=30), Pietrain (n=30), and Muong Khuong (n=25). In the porcine C9 gene, the SNP 407C→G shows two different genotypes GG and CG in F<sub>2</sub> DUMI resource population but not in LR, PIE or MK animals. Allele 'C' maybe derived from Saddle-back Pigs, one of the breeds used to breed Berlin Miniature Pigs. Genetic and allelic frequencies fit to Hardy-Weinberg equilibrium as revealed by using Chi-square test.

Using 118 DNA clones of the INRA-Minnesota RH mapping panel the porcine C8A (LOD=7.00, 56 cR, linked close to marker SW1069), and C8B (LOD=8.18, 50 cR, linked close to marker SW322) were assigned to the same chromosome 6 (q3.1-q3.5), whereas the porcine C6 (LOD=18.13, 21 cR) and C7 (LOD=16.65, 23 cR) were linked to marker S0077 and C9 (LOD=14.45, 27 cR) linked to marker SW403 on the q-arm of chromosome 16 (q1.4). C8G (LOD=8.26, 44 cR, linked close to marker SSC10D08) was assigned to chromosome 1 (q2.13). Genetic mapping was performed in the DUMI F<sub>2</sub> resource population using a set of previously genotyped markers. Genetic linkage using two-point analysis from CRIMAP 2.4 (Green et al. 1990) was confirmed between C7 and C6 (rec. frags. = 0.02, LOD = 50.45), C9 and C6 (rec. frags = 0.06, LOD = 31.50), C9 and C7 (rec. frags = 0.06, LOD = 35.84) on Ssc16, C8B and C8A (rec. frags = 0.05, LOD = 65.72) on Ssc6.

Furthermore, once the polymorphisms were established, one SNP for each candidate gene C6 (862A→G), C7 (881A→G), C8A (1544C→T), C8B (222C→T), and C9 (407C→T), was used for genotyping 417 animals of F<sub>2</sub> DUMI resource population immunized with Mycoplasma (Mh), Aujeszky (ADV) and porcine reproductive and respi-

ratory syndrome (PRRSV) vaccine. These genetic variations were used to evaluate association with hemolytic complement activity in both the classical (CH50) and alternative pathway (AH50) for identifying potential effects on immune response ability of the candidate genes. Changes of CH50 and AH50 concentrations depending on genotypes of the candidate genes were observed prior (day 0) and after vaccinations (day 4 and 10 for Mh and ADV but only day 10 for PRRSV vaccination). (1) For the hemolytic complement activity in the classical pathway, statistically significant difference in CH50 hemolysis was found between genotypes of the porcine C7 ( $p=0.0080$ ) and C9 ( $p=0.0488$ ). Simultaneously, CH50 performance was close to significant difference between genotypes of the porcine C6 ( $p=0.0853$ ) and C8A ( $p=0.0650$ ). Along the vaccination program, the interaction between C8B genotypes and different immunization time points was found highly significant for the CH50 ( $p=0.0048$ ). (2) For the hemolytic complement activity in the alternative pathway, although no significant differences were found between genotypes of the porcine candidate genes, the interactions of genotypes and time points in the AH50 activity were significant for C8A ( $p=0.0027$ ), C8B ( $p=0.0231$ ), and C9 ( $p=0.0340$ ). (3) Also male animals always performed higher hemolysis than females in both pathways. Analysis of variance also indicated significant differences between genders. In summary, the CH50 shows the highest hemolysis value at the fourth day after ADV vaccination whereas the AH50 tends to increase linear during the experiment. Immune response in the short phases after each vaccinating was found in both pathways.

Genetics, sex, age, kind of vaccine, interaction of complement components are major factors modulating to complement activation, therefore causing the lysis of target cells. These obtained results promote C6, C7, C8 and C9 as candidate genes in efforts to improve the animal health in future.

## 7 Zusammenfassung

Diese Studie hatte das Ziel, porcine Gene zu charakterisieren, welche die terminalen Komponenten der Komplementkaskade, C6, C7, C8A, C8B, C8G und C9, kodieren. Dafür wurden die gesamten cDNA-Sequenzen identifiziert, SNPs wurden innerhalb von Kandidatengenen gesucht, 'genetic linkage mapping' und 'radiation hybrid mapping' wurden durchgeführt und Assoziationen zwischen den SNPs und der hämolytischen Komplementaktivität wurden analysiert.

Die cDNA-Sequenzen der Kandidatengene C6, C7, C8A, C8B, C8G und C9 wurde identifiziert, mit 3306, 3561, 2146, 2461, 840, bzw. 2536 bp Länge, einschließlich 144, 807, 91, 28, 51, und 89 bp Länge von der 5' untranslatierte Region (UTR) und 354, 122, 285, 597, 180 und 815 bp Länge der 3' UTR. Der Vergleich der porcinen cDNA-Sequenzen mit humanen genomischen Sequenzen erlaubte auf 18, 18, 11, 12, 7, und 11 Exons der porcinen Kandidatengene C6, C7, C8A, C8B, C8G und C9 zu schließen, welche für 935, 843, 589, 611, 202 bzw. 543 Aminosäuren kodieren. Die porcinen cDNAs und die abgeleiteten Proteinsequenzen der Kandidatengene zeigten hohe Homologie (65-86% bzw. 67-83%) mit den menschlichen Orthologen. Eine Reihe von Abschnitten mit Homologie zu funktionellen Domänen wie TSP1, LDLa, MACPF, EGF, CCP oder FIMAC wurden identifiziert, mit verschiedenen Rollen bei der Bildung von Membran-Angriffs-Komplexen (MAC) und Zellysis.

Untersuchung der cDNA-Sequenzen mit vergleichender Sequenzierung ergab, fünf, sechs, sieben, neun bzw. zwei Einzel-Nukleotid-Polymorphismen (SNPs) in den Kandidatengenen C6, C7, C8A, C8B bzw. C9. Diese SNPs wurden entweder mit Restriktions-Fragment-Längen-Polymorphismus (PCR-RFLP) Methoden oder mit anschließender Resequenzierung bestätigt. Es gibt zwei (240Asn→Asp, 471Asp→Glu), drei (52Ile→Val, 67Thr→Met, 294Lys→Arg), zwei (485Arg→Cys, 528Tyr→Cys), acht (24Pro→Leu, 65Thr→Met, 303Thr→Gly→Ser, 406Ile→Val, 411Pro→Ser, 449Ala→Val, 489Ala→Val, 590Ala→Val), oder ein (106His→Gln) SNPs mit Aminosäurenaustausch in den porcinen Genen C6, C7, C8A, C8B bzw. C9. Kein SNP wurde in C8G beim Vergleich von Individuen der Rassen Hampshire (HS), Duroc (DR), Deutsche Landrasse (LR), Pietrain (PIE), Berliner Miniaturschwein (BMP) und Muong

Khuong (MK) entdeckt. Die meisten der Polymorphismen liegen in der cysteinreichen funktionellen Proteindomäne wie z.B. TSP1, LDLa, MACPF, CCP oder FIMAC. Diese Variationen können eine wichtige Rolle für die Struktur und Funktion der Proteindomänen spielen und könnten daher die Aktivität der Komplementkomponenten in der Zellyse beeinflussen.

Die SNPs 862A→G (240Asn→Asp, TaqI) und 1557C→G (471Asp→Glu, PstI) für C6, 154A→G (52Ile→Val, BsrDI), 870C→T (Hin6I) und 881A→G (294Lys→Arg, MboII) für C7, 535A→G (TfiI), 1544C→T (485Arg→Cys, Hin6I), und 1768C→T (524Tyr→Cys, KnpI) für C8A, 222C→T (65Thr→Met, FnuDII), 935A→G→T (303Thr→Gly→Ser, Resequenzierung) und 1244A→G (406Ile→Val, MaeII) für C8B, 350A→G (BsrDI) und 407C→G (106His→Gln, HpyCH4III) für C9 wurden mittels PCR-RFLP oder Resequenzierung genotypisiert, in nicht verwandten Tieren der Rassen Deutsche Landrasse (n = 30), Pietrain (n = 30), und Muong Khuong (n = 25) und die Allel- und Genotyp-Frequenzen wurden geschätzt. In dem porcinen C9 segregierte der SNP 407C→G mit zwei Genotypen GG und CG in der F<sub>2</sub>-Generation der DUMI-Ressourcenpopulation, aber nicht bei LR, MK oder PIE Tieren. Das Allel 'C' könnte aus Sattelschweinen stammen, eine der Rassen, welche zur Züchtung der Berliner Miniaturschweine benutzt wurde. Die relativen Häufigkeiten der Allele und Genotypen passen zum Hardy-Weinberg-Gleichgewicht, wie mit Hilfe eines Chi-Quadrat-Tests gezeigt wurde.

Unter Benutzung von 118 DNA-Klonen des INRA-Minnesota RH-Mapping-Panel konnte das porcine C8A (LOD = 7,00, 56 cR, in der Nähe von Marker-SW1069), und C8B (LOD = 8,18, 50 cR, in der Nähe von Marker-SW322) auf dem Chromosom 6 (q3.1-q3.5) lokalisiert werden, während das porcine C6 (LOD = 18,13, 21 cR) und C7 (LOD = 16,65, 23 cR) in der Nähe von Marker S0077 und C9 (LOD = 14,45, 27 cR) in der Nähe von Marker SW403 auf dem Q-Arm von Chromosom 16 (q1.4) lokalisiert werden konnte. C8G (LOD = 8,26, 44 cR, in der Nähe von Marker SSC10D08) wurde Chromosom 1 (q2.13) zugewiesen. Genetische Kartierung wurde in der DUMI F<sub>2</sub> Ressourcenpopulation durchgeführt, unter Verwendung der zuvor genotypisierten Marker. Genkopplung konnte unter Verwendung der Zwei-Punkt-Analyse von CRIMAP 2.4 (Green et al. 1990) zwischen C6 und C7 (rec. Fracs. = 0,02, LOD = 50,45), C6 und C9



(rec. Fracs = 0,06, LOD = 31,50) bzw. C9 und C7 (rec. Fracs = 0,06, LOD = 35,84) auf Ssc16 und C8B und C8A (rec. Fracs = 0,05, LOD = 65,72) auf Ssc6 bestätigt werden.

Des Weiteren wurde, nachdem die Polymorphismen etabliert waren, ein SNP für jedes der Kandidatengene C6 (862A→G), C7 (881A→G), C8A (1544C→T), C8B (222C→T) bzw. C9 (407C→T) für die Genotypisierung von 417 Tieren der F<sub>2</sub> DUMI Ressourcenpopulation benutzt, welche mit Mycoplasma- (Mh), Aujeszky- (ADV) und dem 'Porcinen Reproduktiven und Respiratorischen Syndrom'- (PRRSV) -Impfstoff immunisiert worden waren. Um potenzielle Auswirkungen auf die Immunantwort durch die Kandidatengene zu identifizieren, wurde der Zusammenhang zwischen der hämolytischen Komplementaktivität mit dem klassischen Stoffwechselweg (CH50) als auch mit dem alternativen Stoffwechselweg (AH50) und der genetischen Variation an den Kandidatengenen untersucht. Änderungen der CH50 und AH50 in Abhängigkeit von Genotypen der Kandidatengene waren vor der Impfung (Tag 0) und nach der Impfung (Tag 4 und Tag 10 bei Mh und ADV, aber nur Tag 10 bei PRRSV) beobachtet worden. (1) Für die hämolytische Komplementaktivität im klassischen Reaktionsweg CH50 wurde ein statistisch signifikanter Unterschied zwischen Genotypen des porcinen C7 ( $p = 0,0080$ ) und C9 ( $p = 0,0488$ ) entdeckt. Effekte auf die CH50 waren statistisch nicht abzusichern für C6 ( $p = 0,0853$ ) und C8A ( $p = 0,0650$ ). Die Interaktion zwischen C8B-Genotypen und Messzeitpunkten für CH50 vor und nach Immunisierung war hochsignifikant ( $p = 0,0048$ ). (2) Für die hämolytische Komplementaktivität beim alternativen Reaktionsweg (AH50) waren die Wechselwirkungen von Genotypen und Zeitpunkten in der AH50-Aktivität für C8A ( $p = 0,0027$ ), C8B ( $p = 0,0231$ ), bzw. C9 ( $p = 0,0340$ ) signifikant. (3) Dazu zeigten männliche Tiere immer höhere Hämolyse im Vergleich zu weiblichen in beiden Stoffwechselwegen. Die Analyse der Varianz wies außerdem auf signifikante Unterschiede zwischen den Geschlechtern hin. Zusammenfassend lässt sich sagen, dass die CH50 Hämolyse den höchsten Wert am vierten Tage nach der ADV-Impfung zeigt, während der AH50 tendenziell zu einer linearen Zunahme während des Experimentes neigt. Immunreaktionen in der kurzen Phase nach jeder Impfung wurde in beide Stoffwechselwegen beobachtet.

Die Genetik, das Geschlecht, die Art des Impfstoffs und die Interaktion der Komplementkomponenten sind wichtige Faktoren, die das Komplementsystem modulieren, und

dadurch die Lyse von Zielzellen hervorrufen. Die erzielten Ergebnisse lassen darauf schließen, dass C6, C7, C8 und C9 Kandidatengene sind, um die Tiergesundheit in der Zukunft zu steigern.

## 8 References

- Abbott C, West L, Povey S, Jeremiah S, Murad Z, DiScipio R, Fey G (1989): The gene for human complement component C9 mapped to chromosome 5 by polymerase chain reaction. *Genomics* 4, 606-609
- Agah A, Montalto MC, Kiesecker CL, Morrissey M, Grover M, Whoolery KL, Rother RP, Stahl GL (2000): Isolation, characterization, and cloning of porcine complement component C7. *J Immunol* 165, 1059-1065
- Alberti S, Marques G, Hernandez-Alles S, Rubires X, Tomas JM, Vivanco F, Benedi VJ, (1996). Interaction between complement subcomponent C1q and the *Klebsiella pneumoniae* porin OmpK36. *Infect Immun* 64, 4719-4725
- Alexander LJ, Hawken R, Murtaugh J, Flickinger G, Beattie G, Schook L, Milan D, Robic A, Yerle M, Gellin J (1999): Characterization of a porcine radiation hybrid (RH) panel. Plant & Animal Genome VII Conference. Town & Country Hotel, San Diego, CA, 17-21
- Alper CA, Marcus D, Raum D, Petersen BH, Spira TJ (1983): Genetic polymorphism in C8 beta-chains: evidence for two unlinked genetic loci for the eighth component of human complement (C8). *J Clin Invest* 72, 1526-1531
- Asai T, Mori M, Okada M, Uruno K, Yazawa S, Shibata I (1999): Elevated serum haptoglobin in pigs infected with porcine reproductive and respiratory syndrome virus. *Vet Immunol Immunopathol* 70, 143-148
- Alvarez V, Coto E, Setién F, Spath PJ, López-Larrea C (1995): Genetic detection of the silent allele (\*Q0) in hereditary deficiencies of the human complement C6, C7, and C9 components. *Amer J Med Genet* 55, 408 – 413
- Anderson JE, Ladiges WC, Giblett ER, Weiden P, Storb R (1983): Polymorphism of the sixth component of complement (C6) in the dog. *Biochem Genet* 21, 155-160

- Barba GM, Kaufmann TJ, Schneider PM, Rittner C, Brai M (1994): Polymorphism of the complement C8A and -B genes in two families with C8 beta deficiency and neisserial infections. *Clin Immunol Immunopathol* 72, 83-89
- Barroso S, Rieubland C, José álvarez A, López-Trascasa M, Bart PA, Núñez-Roldán A, Sánchez B (2006): Molecular defects of the C7 gene in two patients with complement C7 deficiency. *Immunology* 118, 257-260
- Barroso S, Sanchez B, Alvarez AJ, Lopez-Trascasa M, Lanuza A, Luque R, Wichmann I, Nunez-Roldan A (2004): Complement component C7 deficiency in two Spanish families. *Immunology* 113, 518-523
- Behar D, Schlesinger M, Halle D, Ben-Ami H, Edoute Y, Shahar E, Kasis I, Shihab S, Elstein D, Zimran A, Mandel H (2002): C7 complement deficiency in an israeli arab village. *Am J Med Genet*, 25-29
- Bellavia D, Schneider PM, Rittner C, Malfitano G, Kaufmann TJ, Brai M (1996): Detection of heterozygous C8 beta deficiency by PCR in a healthy Italian population. *Exp Clin Immunogenet* 13, 173-180
- Berman S, Gewurz H, Mold C (1986): Binding of C-reactive protein to nucleated cells leads to complement activation without cytolysis. *J Immunol* 136, 1354-1359
- Betz JJ, Isliker H (1981): Antibody-independent interactions between *Escherichia coli* J5 and human complement components. *J Immunol* 127, 1748-1754
- Bhakdi S, Maillet F, Muhly M, Kazatchkine MD (1988): The cytolytic C5b-9 complement complex: Feedback inhibition of complement activation. *Proc Natl Acad Sci USA* 85, 1912-1916
- Bhakdi S, Tranum Jensen J (1991): Complement lysis: a hole is a hole. *Immunol Today* 12, 318-320
- Biesecker G, Gerard C, Hugli TE (1982): An amphiphilic structure of the ninth component of human complement. *J Biol Chem* 257, 2584-2590

- Biesecker G, Muller-Eberhard HJ (1980): The ninth component of human complement: purification and physicochemical characterization. *J Immunol* 124, 1291-1296
- Bhole D, Stahl GL (2004): Molecular basis for complement component 6 (C6) deficiency in rats and mice. *Immunobiology* 209, 559-568
- Bornstein P (2001): Thrombospondins as matricellular modulators of cell function. *J Clin Invest* 107, 929-934
- Bredt W, Wellek B, Brunner H, Loos M (1977): Interactions between mycoplasma pneumoniae and the first components of complement. *Infect Immun* 15, 7-12
- Brickner A, Sodetz JM (1984): Function of subunits within the eighth component of human complement: identification of a direct role for alpha subunit of C8. *Biochemistry* 23, 832-837
- Castell JV, Gomez-Lechon MJ, David M, Andus T, Geiger T, Trullenque R, Fabra R, Heinrich PC (1989): Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes. *FEBS Lett* 242, 237-239
- Catterall CF, Lyons A, Sim RB, Day AJ, Harris TJ (1987): Characterization of primary amino acid sequence of human complement control protein factor I from an analysis of cDNA clones. *Biochem J* 242, 242-849
- Chakravarti DN, Chakravarti B, Parra CA, Muller-Eberhard HJ (1989): Structural homology of complement protein C6 with other channel-forming proteins of complement. *Immunology* 86, 2799-2803
- Chakravarti DN, Muller-Eberhard HJ (1988): Biochemical characterization of the human complement protein C6: association with  $\alpha$ -thrombin-like enzyme and absence of serine protease activity in cytolytically active C6. *J Biol Chem* 263, 18306-18312
- Clark LK, Armstrong CH, Freeman MJ, Scheidt AB, Sands-Freeman L, Knox K (1991): Investigating the transmission of *Mycoplasma hyopneumoniae* in a swine herd with enzootic pneumonia. *Vet Med* 86, 543-550

- Clamp JR (1975): The plasma proteins: structure, function and genetic control. In: Putnam FW (Eds.). *The Plasma Proteins*. Academic, New York 2, 163-211
- Clapperton M, Bishop SC, Glass EJ (2005): Innate immune traits differ between Meishan and Large White pigs. *Vet Immunol Immunopathol* 104, 131-144
- Cooke RP, Zafar M, Haeney MR (1987): Recurrent meningococcal meningitis associated with deficiencies of C8 and anti-meningococcal antibody. *J Clin Lab Immunol* 23, 53-56
- Corbeil LB (1991): Immunity to infectious disease. In: Woolcock JB (Eds.). *Microbiology of animals and animal products*. ELSEVIER, Amsterdam, 115-140
- Corbo RM, Scacchi R, Cossu G, Brega A, Scozzari R (1994): Genetic studies on the Senegal population. II. Polymorphisms of the plasma proteins F13A, F13B, ORM1, AHSB, C6, C7, and APOC2. *Hum Biol* 66, 885-903
- Coto E, Martínez-Naves E, Domínguez O, DiScipio RG, Urra JM, López-Larrea C (1991): DNA polymorphisms and linkage relationship of the human complement component C6, C7, and C9 genes. *Immunogenetics* 33, 184-187
- Coto E, Martínez-Naves E, Domínguez O, López-Larrea C (1990): DNA polymorphism of the human complement component C7 gene in familial deficiencies. *Hum Genet* 85, 251-252
- Crowell RE, Du Clos TW, Montoya G, Heaphy E, Mold C (1991): C-reactive protein receptors on the human monocytic cell line U-937. Evidence for additional binding to Fc  $\gamma$ RI. *J Immunol* 147, 3445-3451
- Delâge JM, Bergeron P, Simard J, Lehner-Netsch G, Prochazka E (1977): Hereditary C7 deficiency. Diagnosis and HLA studies in a French-Canadian family. *J Clin Invest* 60, 1061-1069
- Demey F, Pandey VS, Baelmans R, Agbede G, Verhulst A (1993): The effect of storage at low temperature on the haemolytic complement activity of chicken serum. *Vet Res Comm* 17, 37-40

- Densen P, Brown EJ, O'Neill GJ, Tedesco F, Clark RA, Frank MM, Webb D, Myers J (1983): Inherited deficiency of C8 in a patient with recurrent meningococcal infections: further evidence for a dysfunctional C8 molecule and nonlinkage to the HLA system. *J Clin Immunol* 3, 90-99
- Dewald G (1988): Polymorphism of the seventh component of complement (C7) in a healthy Caucasian population: an immunoblotting study with neuraminidase-treated samples. *Ann Inst Pasteur Immunol* 139, 507-515
- Dewald G, Cichon S, Bryant SP, Hemmer S, Nothen MM, Spurr, NK (1996): The human complement C8G gene, a member of the lipocalin gene family: polymorphisms and mapping to chromosome 9q34.3. *Ann Hum Genet* 60, 281-291
- Dewald G, Hemmer S, Nothen MM Human complement component C8 (1994): Molecular basis of the beta-chain polymorphism. *FEBS Lett* 340, 211-215
- Dewald G, Nothen MM, Cichon S (1993): Polymorphism of human complement component C6: an amino acid substitution (Glu/Ala) within the second thrombospondin repeat differentiates between the two common allotypes C6 A and C6 B. *Biochem Biophys Res Commun* 194, 458-564
- DiPietro LA, Nissen NN, Gamelli RL, Koch AE, Pyle JM, Polverini PJ (1996): Thrombospondin 1 synthesis and function in wound repair. *Am J Pathol* 148, 1851-1860
- DiScipio RG (1992): Formation and structure of the C5b-7 complex of the lytic pathway of complement. *J Biol Chem* 267, 17087-17094
- DiScipio RG, Berlin C (1999): The architectural transition of human complement component C9 to poly(C9). *Mol Immunol* 36, 575-585
- DiScipio RG, Chakravarti DN, Muller-Eberhard HJ, Fey GH (1988): The structure of human complement component C7 and the C5b-7 complex. *J Biol Chem* 263, 549-560
- DiScipio RG, Daffern PJ, Schraufstatter IU, Sriramarao P (1998): Human polymorphonuclear leukocytes adhere to complement factor H through an interaction that involves  $\alpha$ M $\beta$ 2 (CD11B/CD18). *J Immunol* 160, 4057-4066

- DiScipio RG, Gagnon J (1982): Characterization of human complement components C6 and C7. *Mol Immunol* 19, 1425–1431
- DiScipio RG, Gehring MR, Podack ER, Kan CC, Hugli TE, Fey GH (1984): Nucleotide sequence of cDNA and derived amino acid sequence of human complement component C9. *Proc Nat Acad Sci* 81, 7298-7302
- DiScipio RG, Hugli TE (1985): The architecture of complement component C9 and poly(C9). *J Biol Chem* 260, 14802–14809
- DiScipio RG, Hugli TE (1989): The molecular architecture of human complement component C6. *J Biol Chem* 264, 16197-16206
- DiScipio RG, Linton SM, Rushmere NK (1999): Function of the Factor I Modules (FIMS) of Human Complement Component C6. *J Biol Chem* 274, 31811-31818
- Doenhoff MJ (2000): The immune system. In: Axford RFE, Bishop SC, Nicholas FW, Owen JB (Eds.). *Breeding for disease resistance in farm animals*. 2nd Edition, CAB International, Wallingford, 379-396
- Durandy A (2003): Ontogeny of the Immune System. *Transfus Med Hemother* 30, 222-227
- Eisenschenk FC, Houle JJ, Hoffmann EM (1995): Serum sensitivity of field isolates and laboratory strains of *Brucella abortus*. *Am J Vet Res* 56, 1592-1598
- Eisenschenk FC, Houle JJ, Hoffmann EM (1999): Mechanism of serum resistance among *Brucella abortus* isolates. *Vet Microbiol* 68, 235-244
- Eldridge PR, Hobart MJ, Lachmann PJ (1983): The genetics of the sixth and seventh components of complement in the dog: polymorphism, linkage, locus duplication, and silent alleles. *Biochem Genet* 21, 81-91
- Eng RHK (1980): Bactericidal screening test for late complement component deficiencies or defects. *J Clin Microbiol* 11, 631-634



- Esser AF (1987): C9-mediated cytotoxicity and the function of poly (C9). In: Bonavida B, Collier RJ (Eds.). Membrane-mediated cytotoxicity. UCLA Symposia on Molecular and Cellular Biology. New York: New Series, Liss 45, 411
- Esser AF (1991): Big MAC attack: complement proteins cause leaky patches. *Immunol Today* 12, 316–318
- Esser AF (1994): The membrane attack complex of complement: assembly, structure and cytotoxic activity. *Toxicology* 87, 229-247
- Fearon DT, Austen KF (1975): Properdin: binding to C3b and stabilization of the C3b-dependent C3 convertase. *J Exp Med* 142, 856-863
- Fernie BA, Finlay A, Price D, Chan E, Orren A, Joysey VC, Joysey KA, Hobart MJ (1996): Complement C6 and C7 DNA polymorphisms analysed by PCR in seven ethnic groups and characterisation of the C6 MspI RFLP. *Exp Clin Immunogenet* 3, 92-103
- Fernie BA, Hobart MJ (1997): An unusual combined insertion/deletion polymorphism in intron 10 of the human complement C6 gene. *Hum Genet* 100, 104-108
- Fernie BA, Hobart MJ (1999): Five new polymorphisms in the complement C7 gene and their association with C7 deficiency. *Exp Clin Immunogenet* 16, 150-161
- Fernie BA, Orren A, Schlesinger M, Wuerzner R, Platonov AE, Cooper RC, Williams YE, Hobart MJ (1997): DNA haplotypes of the complement C6 and C7 genes associated with deficiencies of the seventh component; and a new DNA polymorphism in C7 exon 13. *Ann Hum Genet* 61, 287-298
- Fernie BA, Wuerzner R, Unsworth DJ, Tuxworth RI, Hobart MJ (1995): DNA polymorphisms of the complement C6 and C7 genes. *Ann Hum Genet* 59: 163-181
- Fernie BA, Wuerzner R, Orren A, Hobart M (1996): A new intronic polymorphism in the C7 gene 36 bp from the common expressed C7 M/N polymorphism. *Ann Hum Genet* 60, 179-182

- Fernie-King BA, Seilly DJ, Willers C, Wuerzner R, Davies A, Lachmann PJ (2001): Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C5b67 onto cell membranes. *Immunology* 103, 390–398
- Ferre F and Clote P (2005): DiANNA: a web server for disulfide connectivity prediction. *Nucleic Acids Res* 33 (Web Server issue): 230-232
- Figueroa JE, Densen P (1991): Infectious diseases associated with complement deficiencies. *Clin Microbiol Rev* 4, 359-395
- Firth MA, Colyer S, Shewen PE, Hodgins DC (2007): Sus scrofa complement component C3d mRNA, partial cds: GenBank acc. no. EU257630 (NCBI homepage)
- Fraser IP, Koziel H, Ezekowitz RAB (1998): The serum mannose-binding protein and the macrophage mannose receptor are pattern recognition molecules that link innate and adaptive immunity. *Semin Immunol* 10, 363-372
- Friedman HM, Wang L, Pangburn MK, Lambris JD, Lubinski J (2000): Novel mechanism of antibody-independent complement neutralization of herpes simplex virus type 1. *Immunol* 165, 4528-4536
- Fries L F, Friedman HM, Cohen GH, Eisenberg RJ, Hammer CH, Frank MM (1986): Glycoprotein C of herpes simplex virus I is an inhibitor of the complement cascade. *J Immunol* 137, 1636-1641
- Fujita T (2002): Evolution of the lectin–complement pathway and its role in innate immunity. *Nature reviews immunology* 2, 346-353
- Fukumori Y and Horiuchi T (1998): Terminal complement component deficiencies in Japan. *Exp Clin Immunogenet* 15, 244-248
- Giraldo G, Degos L, Beth E, Sasportes M, Marcelli A, Gharbi R, Day NK (1977): C8 deficiency in a family with xeroderma pigmentosum: lack of linkage to HLA region. *Clin Immun Immunopath* 8, 377-384
- Goebel EM, Wolfe DN, Elder K, Stibitz S, Harvill ET (2008): O antigen protects *Bordetella parapertussis* from complement. *Infect Immun* 76, 1774-1780

- Goldberger G, Bruns GA, Rits M, Edge MD, Kwiatkowski DJ (1987): Human complement factor I: analysis of cDNA-derived primary structure and assignment of its gene to chromosome 4. *J Biol Chem* 262, 10065-10071
- Goldman MB, Cohen C, Stronski K, Bangalore S, Goldman JN (1982): Genetic control of C6 polymorphism and C6 deficiency in rabbits. *J Immunol* 128, 43-48
- Gonzalez-Ramon N, Hoebe K, Alava MA, Van Leengoed L, Pineiro M, Carmona S, Iturralde M, Lampreave F, Pineiro A (2000): Pig MAP/ITIH4 and haptoglobin are interleukin-6-dependent acute-phase plasma proteins in porcine primary cultured hepatocytes. *Eur J Biochem* 267, 1878-1885
- Goundis D, Reid KBM (1988): Properdin, the terminal complement components, thrombospondin and the circumsporozoite protein of malaria parasites contain similar sequence motifs. *Nature* 335, 82-85
- Granados J, Cramer DV, Caputo JB, Marcus D and Alper CA (1984): Genetic polymorphism of the sixth component (C6) of rat complement. *J Immunol* 133, 405-407
- Green P, Falls K, Crooks S (1990): Documentation for CRIMAP, version 2.4. Washington University, School of Medicine, St Louis, MO
- Hadders MA, Beringer DX, Gros P (2007): Structure of C8A-MACPF reveals mechanism of membrane attack in complement immune defense. *Science* 314, 1552 - 1554
- Haefliger JA, Jenne D, Stanley KK, Tschopp J (1987): Structural homology of human complement component C8 gamma and plasma protein HC: identity of the cysteine bond pattern. *Biochem Biophys Res Commun* 149, 750-754
- Haefliger JA, Peitsch MC, Jenne DE, Tschopp J (1991): Structural and functional characterization of complement C8 gamma, a member of the lipocalin protein family. *Mol Immunol* 28, 123-131
- Haefliger JA, Tschopp J, Nardelli D, Wahli W, Kocher HP, Tosi M, Stanley KK (1987): Complementary DNA cloning of complement C8 beta and its sequence homology to C9. *Biochemistry* 26, 3551-3556

- Haefliger JA, Tschopp J, Vial N, Jenne DE (1989): Complete primary structure and functional characterization of the sixth component of the human complement system: Identification of the C5b-binding domain in complement C6. *J Biol Chem* 264, 18041-18051
- Haeney MR (1998): The role of the complement cascade in sepsis. *J. Antimicrob Chemother* 41, 41-46
- Haeney MR, Thompson RA, Faulkner J, Mackintosh P, Ball AP (1980): Recurrent bacterial meningitis in patients with genetic defects of terminal complement components. *Clin Exp Immunol* 40, 16-24
- Haesebrouck F, Pasmans F, Chiers K, Maes D, Ducatelle R, Decostere A (2004): Efficacy of vaccines against bacterial diseases in swine: what can we expect? *Vet Microbiol* 100, 255-268
- Hall JE, Alper CA (1977): Genetic polymorphism of the sixth component of complement (C6) in the rhesus monkey. *J Immunol* 119, 253-255
- Hardge T, Koepke K, Reissman M, Wimmers K (1999): Maternal influences on litter size and growth in reciprocal crossed Miniature Pigs and Durocs. *Arch. Anim Breed* 1, 83-92
- Hayakawa J, Nikaido H, Koizumi T (1984): Genetic polymorphism of the sixth component of complement (C6) in mice. *Immunogenetics* 20, 633-638
- Hawken RJ, Murtaugh J, Flickinger GH, Yerle M, Robic A, Milan D, Gellin J, Beattie CW, Schook LB, Alexander LJ (1999): A first-generation porcine whole-genome radiation hybrid map. *Mammal Genome* 10, 824-830
- Henryon M, Juul-Madsen HR, Berg P (2002): Genetic variation for total and differential numbers of leukocytes in growing pigs. In: Proceedings of the 7<sup>th</sup> WCGALP, INRA, August 19-23, Montpellier, France
- Herrmann D, Sodetz JM, Rittner C, Schneider PM (1989): DNA polymorphism of the human complement C8B gene: formal genetics and intragenic localization. *Immunogenetics* 30, 291-295

- Heyworth PG (2001): Immunodeficiency, primary. In: Fullerlove G and Robertson S (Eds.). Affecting the innate immune system. . Encyclopedia of Life Sciences 9, Nature Publishing Group, London, 640-649
- Hidaka Y, Sakai Y, Toh Y, Mori R (1991): Glycoprotein C of herpes simplex virus type 1 is essential for the virus to evade antibody-independent complement-mediated virus inactivation and lysis of virus-infected cells. *J Gen Virol* 72, 915-921
- Hill JH, Ward PA (1971): The phlogistic role of C3 leukotactic fragments in myocardial infarcts of rats. *J Exp Med* 133, 885-900
- Hobart MJ, Fernie B, DiScipio RG (1993): Structure of the human C6 gene. *Biochemistry* 32, 6198-6205
- Hobart MJ, Fernie BA, DiScipio RG (1995): Structure of the human C7 gene and comparison with the C6, C8A, C8B, and C9 genes. *J Immunol* 154, 5188-5194
- Hobart MJ, Fernie BA, DiScipio RG, Lachmann PJ (1993): A physical map of the C6 and C7 complement component gene region on chromosome 5p13. *Hum Mol Genet* 2, 1035-1036
- Hobart MJ, Fernie BA, Fijen, KA, Orren A (1998): The molecular basis of C6 deficiency in the Western Cape, South Africa. *Hum Genet* 103, 506-512
- Hobart MJ, Lachmann PJ, Alper CA (1975): Polymorphism of human C6. In Peeters H (Eds.). *Protides of the Biological Fluids. 22nd Colloquium, Brugge 1974*. Oxford, Pergamon Press, 575-580
- Hobart MJ, Fernie BA, Wuerzner R, Oldroyd RG, Harrison RA, Joysey V, Lachmann PJ (1997): Difficulties in the ascertainment of C9 deficiency: lessons to be drawn from a compound heterozygote C9-deficient subject. *Clin Exp Immunol* 108, 500-506
- Hoffmann EM, Houle JJ (1983): Failure of *Brucella abortus* lipopolysaccharide to activate the alternative pathway of complement. *Vet Immunol Immunopath* 5, 65-76

- Hofsteenge J, Blommers M, Hess D, Furmanek A, Miroshnichenko O (1999): The four terminal components of the complement system (C6, C7, C8 and C9) are C-mannosylated on multiple tryptophan residues. *J Biol Chem* 274, 32786–32794
- Hogasen K, Mollnes TE, Nurnberger W, Pausa M, Fukumori Y, Tedesco F (1998): Characterization of soluble terminal complement complex assembled in C8beta-deficient plasma and serum. *Scand J Immunol* 48, 261-268
- Horiuchi T, Ferrer JM, Serra P, Matamoros N, López-Trascasa M, Hashimura C, Niho Y (1999): A novel nonsense mutation at Glu-631 in a Spanish family with complement component 7 deficiency. *J Hum Genet* 44, 215-218
- Horiuchi T, Nishimukai H, Okiura T, Nishimura K, Nishizaka H, Kojima T, Tsukamoto H, Hayashi K, Harada M (2002): Molecular bases for human complement C7 polymorphisms, C7\*3 and C7\*4. *Biochem Biophys Res Commun* 298, 450-455
- Horiuchi T, Nishizaka H, Tsukamoto H, Harashima S, Sawabe T, Morita C, Niho Y (1999): An NcoI polymorphism in the human complement component 7 (C7) gene. *J Hum Genet* 44, 270-271
- Howard OMZ, Rao AG, Sodetz JM (1987): Complementary DNA and derived amino acid sequence of the beta subunit of human complement protein C8: identification of a close structural and ancestral relationship to the alpha subunit and C9. *Biochemistry* 26, 3565–3570
- Hu VW, Esser AF, Podack ER, Wisnieski BJ (1981): The membrane attack mechanism of complement: photolabeling reveals insertion of terminal proteins into target membrane. *J Immunol* 127, 380–386
- Hunt LT, Elzanowski A, Barker WC (1987): The homology of complement factor C8 gamma chain and alpha-1-microglobulin. *Biochem Biophys Res Commun* 149, 282-288
- Ichikawa E, Furuta J, Kawachi Y, Imakado S, Otsuka F (2001): Hereditary complement (C9) deficiency associated with dermatomyositis. *Br J Dermatol*, 1080-1083

- Inai S, Akagaki Y, Moriyama T, Fukumori Y, Yoshimura K, Ohnoki S, Yamaguchi H (1989): Inherited deficiencies of the late-acting complement components other than C9 found among healthy blood donors. *Int Arch Allergy Appl Immunol* 90, 274-279
- Inai S, Kitamura H, Hiramatsu S, Nagaki K (1979): Deficiency of the ninth component of complement in man. *J Clin Lab Immunol* 2, 85-87
- Jasin HE (1977): Absence of the eighth component of complement in association with systemic lupus erythematosus-like disease. *J Clin Invest* 60, 709-715
- Jiang H, Siegel JN, Gewurz H (1991): Binding and complement activation by C-reactive protein via the collagen-like region of C1q and inhibition of these reactions by monoclonal antibodies to C-reactive protein and C1q. *J Immunol* 146, 2324-2330
- Joiner KA, Hammer CH, Brown EJ, Cole RJ, Frank MM (1982): Studies on the mechanics of bacterial resistance to complement-mediated killing. I. Terminal complement components are deposited and released from *Salmonella minnesota* S218 without causing bacterial death. *J Exp Med* 155, 797-808
- Joiner KA, Hammer CH, Brown EJ, Frank MM (1982): Studies on the mechanism of bacterial resistance to complement-mediated killing. II. C8 and C9 release C5b67 from the surface of *Salmonella minnesota* S218 because the terminal complex does not insert into the bacterial outer membrane. *J Exp Med* 115, 809-819
- Jorgensen CB, Wintero AK, Yerle M, Fredholm M. (1997): Mapping of twenty-two expressed sequence tags isolated from a porcine small intestine cDNA library. *Mamm Genome* 8, 423-427
- Kanemitsu S, Ihara K, Kira R, Kaku Y, Sakai K, Tsuzaka K, Takeuchi T, Hara T (2000): Complement component 9 deficiency is not a susceptibility factor for SLE. *Lupus* 9, 456-457
- Kang HJ, Ki CS, Kim YS, Hur M, Jang SI, Min KS (2006): Two mutations of the C7 gene, c.1424G > A and c.281-1G > T, in two Korean families. *J Clin Immunol* 26, 186-191

- Kaufmann T, Hansch G, Rittner C, Spath P, Tedesco F, Schneider PM (1993): Genetic basis of human complement C8-beta deficiency. *J Immunol* 150, 4943-4947
- Kaufmann T, Rittner C, Schneider PM (2004): The human complement component C8B gene: structure and phylogenetic relationship. *Hum Genet* 92, 69-75
- Kaufman KM, Sodetz JM (1994): Genomic structure of human complement protein C8 $\gamma$ : homology to the lipocalin gene family. *Biochemistry* 33, 5162-5166
- Kawai T, Katoh K, Narita M, Tani K, Okubo T (1989): Deficiency of the 9th component of complement (C9) in a patient with systemic lupus erythematosus. *J Rheumatol* 16, 542-543
- Keller H, Löke S, Hänsch GM, Jentschura D, Gerhard H, Heene DL (1987): Rezidivierende Meningitis bei familiärem Mangel der achten Komponente des Komplementsystems. *J Mol Med* 65, 387-390
- Ki CS, Kim JW, Kim HJ, Choi SM, Ha GY, Kang HJ, Kim WD (2005): Two novel mutations in the C7 gene in a Korean patient with complement C7 deficiency. *J Korean Med Sci* 20, 220-224
- King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TA, Thomas W (1994): Allergen nomenclature. WHO/IUIS Allergen Nomenclature Subcommittee. *Int Arch Allergy Immunol* 105, 224-233
- Kira R, Ihara K, Takada H, Gondo K, Hara T (1998): Nonsense mutation in exon 4 of human complement C9 gene is the major cause of Japanese complement C9 deficiency. *Hum Genet* 102, 605-610
- Kliebenstein JB (2004): Economic Impact of PRRS on the Cost of Pork Production. The official website of the national pork board funded by the pork checkoff. <http://www.pork.org/PorkScience/Research/Documents/02-223-KLIEBENSTEIN.ABS.3-5-04.pdf>
- Kolb WP, Kolb LM, Savary JR (1982): Biochemical characterization of the sixth component (C6) of human complement. *Biochemistry* 21, 294-310



- Komatsu M (1985): A method for developing hereditary deficiency of complement component in the rabbit. *Jikken Dobutsu* 34, 173-182
- Komatsu M, Yamamoto K, Kawashima T, Migita S (1985): Genetic deficiency of the alpha-gamma-subunit of the eighth complement component in the rabbit. *J Immunol* 134, 2607-2609
- Komatsu M, Yoshiwara S, Akita T (1998): Immunological characteristics of the Meishan pig: serum complement activity, serum C3 level, immune response and skin structure. In: *Proceedings of the 6<sup>th</sup> World Congress on Genetics Applied to Livestock Production*. Armidale, NSW, Australia, 323-326
- Komatsu N, Kido A, Oya M (1989): C7 polymorphism in Japanese: New allele C7\*8. *Hum Hered* 39, 49-51
- Kotnik V, Luznik-Bufon T, Schneider PM, Kirschfink M (1997): Molecular, genetic, and functional analysis of homozygous C8 beta-chain deficiency in two siblings. *Immunopharmacology* 38, 215-221
- Kristensen T, Tack BF (1986): Murine protein H is comprised of 20 repeating units, 61 amino acids in length. *Proc Natl Acad Sci USA* 83, 3963-3967
- Kucher AN, Puzyrev VP, Ivanova OF, Khu TsIu, Siu TsTs, Du ZhF (1993): Genetic polymorphism of the sixth (C6) and seventh (C7) components of complement in Russian rural inhabitants of the Tomsk region. *Genetika* 29, 1889-1894
- Kushner I (1982): The phenomenon of the acute phase response. *Ann N Y Acad Sci* 389, 39-48
- Kühnl P, Kreckel P (1980): C6 phenotypes and allele frequencies in a German population. *Immunobiology* 158, 50-54
- Kumar KG, Ponsuksili S, Schellander K, Wimmers K (2004): Molecular cloning and sequencing of porcine C5 gene and its association with immunological traits. *Immunogenetics* 55, 811-817

- Kumar S, Tamura K, Nei M (2004): MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform* 5, 150-163
- Kunstmann G, Mauff G (1980): Genetic polymorphism of rabbit C6. *Immunobiology* 158, 30-33
- Kunstmann G, Mauff G, Pulverer G (1980): C6 polymorphism and rare alleles in Western Germany. *Immunobiology* 158, 55-59
- Lassiter HA, Wilson JL, Feldhoff RC, Hoffpauir JM, Klueber KM (1994): Supplemental complement component C9 enhances the capacity of neonatal serum to kill multiple isolates of pathogenic *Escherichia coli*. *Pediatr Res* 35, 389-396
- Lemke U, Huyen LTT, Roeßler R, Thuy LT and Zárata V (2005): Impact of the use of exotic compared to local pig breeds on socio-economic development and biodiversity in Vietnam. Conference on International Agricultural Research for Development. Stuttgart-Hohenheim, October 11-13, 1-4
- Letson CS, Kaufman KM, Sodetz JM (1996): In vitro expression of the beta subunit of human complement component C8. *Mol Immunol* 33, 1295-1300
- Liu CC, Young JD (1988): A semi-automated microassay for complement activity. *J Immunol Methods* 114, 33-39
- Loos M and Clas F (1987): Antibody-independent killing of gram-negative bacteria via the classical pathway of complement. *Immunol Lett* 14, 203-208
- Loos M, Trinder PKE, Kaul (1998): Factors of the classical pathway. In: Rother K, Till GO, Haensch GM (Eds). *The complement system*. 2nd revised edition. Springer-Verlag Berlin Heidelberg, 1-23
- Mackiewicz A, Ganapathi MK, Schultz D, Brabenec A, Weinstein J, Kelley MF, Kushner I (1990): Transforming growth factor beta 1 regulates production of acute-phase proteins. *Proc Natl Acad Sci USA* 87, 1491-495
- Makrides SC (1998): Therapeutic Inhibition of the Complement System. *Pharmacol Rev* 50, 59-87

- Mallard BA, Wilkie BN, Kennedy BW, Gibson J, Quinton M (1998): Immune responsiveness in swine: eight generations of selection for high and low immune response in Yorkshire pigs. In: Proceedings of the 6th World Congress on Genetics Applied to Livestock Production. University of New England, Armidale, New South Wales, Australia, 257-64
- Marnell LL, Mold C, Volzer MA, Burlingame RW, Du Clos TW (1995): C-reactive protein binds to FcγRI in transfected COS cells. *J Immunol* 155, 2185-2193
- Mayes JT, Schreiber RD, Cooper NR (1984): Development and application of an enzyme-linked immunosorbent assay for the quantitation of alternative complement pathway activation in human serum. *J Clin Invest* 73, 160-170
- McCloskey MA, Dankert JR, Esser AF (1989): Assembly of complement components C5b-8 and C5b-9 on lipid bilayer membranes: visualization by freeze-etch electron microscopy. *Biochemistry* 28, 534-540
- McNearney TA, Odell C, Holers VM, Spear PG, Atkinson JP (1987): Herpes simplex virus glycoproteins gC-1 and gC-2 bind to the third component of complement and provide protection against complement-mediated neutralization of viral infectivity. *J Exp Med* 166, 1525-1535
- Mead RJ, Singhrao SK, Neal JW, Lassmann H, Morgan BP (2002): The membrane attack complex of complement causes severe demyelination associated with acute axonal injury. *J Immunol* 168, 458-465
- Medzhitov R (2001): Toll-like receptors and innate immunity. *Nat Rev Immunol* 1, 135-145
- Mekchay S, Ponsuksili S, Schellander K, Wimmers K (2003): Association of the porcine C3 gene with haemolytic complement activity in the pig. *Genet Sel Evol* 35, S83-96
- Meri S, Morgan BP, Davies A, Daniels RH, Olavesen MG, Waldmann H, Lachmann PJ (1990): Human protectin (CD59), an 18,000-20,000 MW complement lysis restricting factor, inhibits C5b-8 catalyzed insertion of C9 into lipid bilayer. *Immunology* 71, 1-9

Meyers SN, Rogatcheva MB, Larkin DM, Yerle M, Milan D, Hawken RJ, Schook LB, Beever JE (2005): Piggy-BACing the human genome II. A high-resolution, physically anchored, comparative map of the porcine autosomes. *Genomics* 86, 739-752

Michelotti GA, Snider JV, Sodetz JM (1995): Genomic organization of human complement protein C8 alpha and further examination of its linkage to C8 beta. *Hum Genet* 95, 513-518

Minion FC (2002): Molecular pathogenesis of mycoplasma animal respiratory pathogens. *Front Biosci* 7, 1410-1422

Monahan JB, Sodetz JM (1980): Binding of the eighth component of human complement to the soluble cytolytic complex is mediated by its beta subunit. *J Biol Chem* 255, 10579-10582

Mold C (1998): In: *The Human Complement System in Health and Disease*. In: Volanakis JE, Frank MM (Eds.). Cellular responses to the membrane attack complex. Marcel Dekker, New York, 309-325

Mold C, Bradt BM, Nemerow GR, Cooper NR (1988): Activation of the alternative complement pathway by EBV and the viral envelope glycoprotein, gp350. *J Immunol* 140, 3867-3874

Mold C, Gewurz H, Du Clos TW (1999): Regulation of complement activation by C-reactive protein. *Immunopharmacology* 42, 23-30

Moolenaar W, Pott Hofstede D, Hack CE, Hannema AJ (1987): A patient with meningococcal meningitis and deficiency of the 7th factor of the complement system. *Ned Tijdschr Geneeskd* 131, 866-868

Morgan BP (1989). Complement membrane attack on nucleated cells: resistance, recovery and non-lethal effects. *Biochem J* 264, 1-14

Morgan BP (1992): Isolation and characterization of the complement-inhibiting protein CD59 antigen from platelet membranes. *Biochem J* 282, 409-413

- Morgan BP (1999): Regulation of the complement membrane attack pathway. *Crit Rev Immunol* 19, 173-198
- Morgan BP, Vora JP, Bennett AJ, Thomas JP, Matthews N (1989): A case of hereditary combined deficiency of complement components C6 and C7 in man. *Clin Exp Immunol* 75, 396-401
- Morrison DC, Kline LF (1977): Activation of the classical and properdin pathways of complement by bacterial lipopolysaccharides (LPS). *J Immunol* 118, 362-368
- Müller-Eberhard HJ (1986): The membrane attack complex of complement. *Annu Rev Immunol* 4, 503-528
- Mueller-Eberhard HJ (1988): Molecular organization and function of the complement system. *Annu Rev Biochem* 57, 321-347
- Musingarimi P, Plumb ME, Sodetz JM (2002): Interaction between the C8 alpha-gamma and C8 beta subunits of human complement C8: role of the C8 beta N-terminal thrombospondin type 1 module and membrane attack complex/perforin domain. *Biochemistry* 41, 11255-11260
- Nakajima E, Itoh T, Suzuk K, Kawakami K, Takeda K, Onishi A, Komatsu M (1998): Characterization, chromosomal localization, and genetic variation of the [alpha] subunit of porcine eighth component of complement. *Animal Genetics* 29, 377-380
- Nakamura S, Ohue O, Abe K (1986): Genetic polymorphism of human complement component C81 in the Japanese population. *Hum Genet* 72, 344-347
- Nakamura S, Ooue O, Abe K (1984): Genetic polymorphism of the seventh component of complement in a Japanese population. *Hum Genet* 66, 279-281
- Ng SC, Rao AG, Howard OMZ, Sodetz JM (1987): The eighth component of human complement: evidence that it is an oligomeric serum protein assembled from products of three different genes. *Biochemistry* 26, 5229-5233
- Nicholson-Weller A, Halperin JA (1993): Membrane signaling by complement C5b-9, the membrane attack complex. *Immunol Res* 12, 244-257

Nishimukai H, Nishimura K, Orimoto C, Okiura T, Fujii K, Fukumori Y, Harihara S (2003): Single nucleotide polymorphisms in the human complement C6 and C7 genes. *Leg Med (Tokyo)* 5, 198-200

Nishimukai H, Tamaki Y (1986): Genetic polymorphism of the seventh component of complement: a new variant. *Vox Sang* 51, 60-62

Nishizaka, H., Horiuchi, T., Zhu, Z. B., Fukumori, Y., Nagasawa, K., Hayashi, K., Krumdieck, R., Cobbs, C. G., Higuchi, M., Yasunaga, S (1996): Molecular bases for inherited human complement component C6 deficiency in two unrelated individuals. *J Immunol* 156, 2309-2315

Nishizaka H, Horiuchi T, Zhu ZB, Fukumori Y, Volanakis JE (1996): Genetic bases of human complement C7 deficiency. *J Immunol* 157, 4239-4243

Norman DG, Barlow PN, Baron M, Day AJ, Sim RB, Campbell ID (1991): Three-dimensional structure of a complement control protein module in solution. *J Mol Biol* 219, 717-725

Olaho-Mukani W, Nyang'ao JNM, Kimani JK, Omuse JK (1995): Studies on the haemolytic complement of the dromedary camel (*Camelus dromedarius*). II. Alternate pathway haemolytic activity in serum. *Vet Immunol Immunopathol* 48, 169-176

Olaho-Mukani W, Nyang'ao JNM, Kimani JK, Omuse JK (1995): Studies on the haemolytic complement of the dromedary camel (*Camelus dromedarius*). I. Classical pathway haemolytic activity in serum. *Vet Immunol Immunopathol* 46, 337-347

O'Leary JM, Bromek K, Black GM, Uhrinova S, Schmitz C, Wang X, Krych M, Atkinson JP, Uhrin D, Barlow PN (2004): Backbone dynamics of complement control protein (CCP) modules reveals mobility in binding surfaces. *Protein Sci* 13, 1238-1250

Olving JH, Teisberg P, Olaisen B (1980): Polymorphism of the sixth component of complement (C6) in Norwegian Lapps. *Hum Hered* 30, 211-214

Orren A, Potter PC, Cooper RC, du Toit E (1987): Deficiency of the sixth component of complement and susceptibility to *Neisseria meningitidis* infections: studies in 10 families and five isolated cases. *Immunology* 62, 249-253

- Orren A, Wallace ME, Horbart MJ, Lachmann PJ (1989): C6 polymorphism and C6 deficiency in site strains of the mutation-prone Peru-Coppock mice. *Complement Inflamm* 6, 295-296
- Ortlund E, Parker C, Schreck SF, Minor W, Ginel S, Sodetz JM, Lebioda L (2002): Crystal structure of human complement protein C8 $\gamma$  at 1.2Å resolution reveals a lipocalin fold and a distinct ligand binding site. *Biochemistry* 41, 7030 -7037
- Ouma JO, Olaho-Mukani W, Wishitemi BEL, Guya SO (1997): Changes in classical pathway complement activity in dromedary camels experimentally infected with *Trypanosoma evansi*. *Vet Immunol Immunopathol* 57, 135-140
- Parham KL, Roberts A, Thomas A, Wuerzner R, Henderson HE, Potter PC, Morgan BP, Orren A (2007): Prevalence of mutations leading to complete C6 deficiency (C6Q0) in the Western Cape, South Africa and detection of novel mutations leading to C6Q0 in an Irish family. *Molecular Immunol* 44, 2756-2760
- Parish CR (2001): Innate Immune Mechanisms: Nonself Recognition. In : Fullerlove G and Robertson S (Eds.). *Encyclopedia of Life Sciences* 10, Nature Publishing Group, London 305-312
- Park KS, Yanagisawa M, Tokunaga K, Omoto K (1988): Genetic polymorphisms of complement components C6 and C7 in Korean. *J Hum Genet* 33, 325-331
- Parker CL, Sodetz JM (2002): Role of the human C8 subunits in the complement-mediated bacterial killing: evidence that C8 gamma is not essential. *Mol Immunol* 39, 453-458
- Phatsara C, Jennen DG, Ponsuksili S, Murani E, Tesfaye D, Schellander K, Wimmers K (2007) Molecular genetic analysis of porcine mannose-binding lectin genes, MBL1 and MBL2, and their association with complement activity. *Int J Immunogenet* 34, 55-63
- Platteborze PL, Hobart MJ, Sodetz JM (1996): Physical linkage and orientation of the human complement C8a and C8 $\beta$  genes on chromosome 1p32. *Human Genetics* 98, 443-446

- Plumb ME, Scibek JJ, Barber TD, Dunlap RJ, Platteborze PL, Sodetz JM (1999): Chimeric and truncated forms of human complement protein C8 alpha reveal binding sites for C8 beta and C8 gamma within the membrane attack complex/perforin region. *Biochemistry* 38, 8478-8484
- Plumb ME, Sodetz JM (1998): Proteins of the membrane attack complex. In: Volanakis JE, Frank MM (Eds.). *The Human Complement System in Health and Disease*. Marcel Dekker, New York, 119-148
- Plumb ME, Sodetz JM (2000): An indel within the C8 alpha subunit of human complement C8 mediates intracellular binding of C8 gamma and formation of C8 alpha-gamma. *Biochemistry* 39, 13078-13083
- Pensaert MB, Kluge JP (1989): Pseudorabiesvirus (Aujeszky's disease). In: Horzinek (Eds.), *Virus infections of vertebrates/ Volume 2; Virus infections of porcines*. Verlag Elsevier Science Publishers BV, Amsterdam, The Netherlands
- Petersen BH, Graham JA, Brooks GF (1976): Human deficiency of the eighth component of complement. The requirement of C8 for serum *Neisseria gonorrhoeae* bactericidal activity. *J Clin Invest* 57, 283-290
- Petry DB, Holl JW, Weber JS, Doster AR, Osorio FA, Johnson RK (2005): Biological responses to porcine respiratory and reproductive syndrome virus in pigs of two genetic populations. *J Anim Sci* 83, 1494-1502
- Pluschke G, Achtman M (1984): Degree of antibody-independent activation of the classical complement pathway by K1 *Escherichia coli* differs with O antigen type and correlates with virulence of meningitis in newborns. *Infect Immun* 43, 684-692
- Podack ER, Biesecker G, Muller-Eberhard HJ (1979): Membrane attack complex of complement: generation of high-affinity phospholipid binding sites by fusion of five hydrophilic plasma proteins. *Proc Natl Acad Sci* 76, 897-901
- Podack ER, Kolb WP, Esser AF, Mueller-Eberhard HJ (1979): Structural similarities between C6 and C7 of human complement. *J Immunol* 123, 1071-1077



- Podack ER, Kolb WP, Muller-Eberhard HJ (1976): The C5b-9 complex: subunit composition of the classical and alternative pathway-generated complex. *J Immunol* 116, 1431-1434
- Podack ER, Stoffel W, Esser AF, Mueller-Eberhard HJ (1981): Membrane attack complex of complement: Distribution of subunits between the hydrocarbon phase of target membranes and water. *Proc natn Acad Sci USA* 78, 4544-4548
- Podack ER, Tschoop J, Muller-Eberhard HJ (1982): Molecular organization of C9 within the membrane attack complex of complement. Induction of circular C9 polymerization by the C5b-8 assembly. *J Exp Med* 156, 268-282
- Ponsuksili S, Wimmers K, Yerle M, Schellander K (2001): Mapping of 93 porcine ESTs preferentially expressed in liver. *Mamm Genome* 12, 869-872
- Qian YM, Qin X, Miwa T, Sun X, Halperin JA, Song WC (2000): Identification and functional characterization of a new gene encoding the mouse terminal complement inhibitor CD59. *J Immunol* 165, 2528-2534
- Rao AG, Howard OM, Ng SC, Whitehead AS, Colten HR, Sodetz JM (1987): Complementary DNA and derived amino acid sequence of the alpha subunit of human complement protein C8: evidence for the existence of a separate alpha subunit messenger RNA. *Biochemistry* 26, 3556-3564
- Rao L, Li YB, Chen GD, Zhou B, Schneider PM, Zhang L (2004): Further study on heterogenic basis of complement C8 beta deficiency. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 21, 10-13
- Raum D, Balner H, Petersen BH, Alper CA (1980): Genetic polymorphism of serum complement components in the chimpanzee. *Immunogenetics* 10, 455-468
- Reid KBM, Porter RR (1981): The Proteolytic Activation Systems of Complement. *Annu Rev Biochem* 50, 433-464
- Rekitt M, Hiss S, Dickhöfer D, Wimmers K, Schellander K, Petersen B, Sauerwein H (2001): Haptoglobin responses towards vaccination with three different antigens in pigs.

In: Proceedings of the 11<sup>th</sup> International Conference on Production Diseases in Farm Animals. Frederiksberg, Denmark, 62

Ripoche J, Day AJ, Harris TJ, Sim RB (1988): The complete amino acid sequence of human complement factor H. *Biochem J* 249, 593-602

Rittner C, Hargesheimer W, Mollenhauer E (1984): Population and formal genetics of the human C81(alpha-gamma) polymorphism. *Hum Genet* 67, 166-169

Rogde S, Høiby EA, Teisberg P, Olaisen B (1990): Genetic aspects of complement component C8 in Norwegian meningococcal disease patients. *Scand J Infect Dis* 22, 673-679

Rogde S, Mevåg B, Teisberg P, Gedde-Dahl T, Tedesco F, Olaisen B (1985): Genetic polymorphism of complement component C8. *Hum Genet* 70, 211-216

Rogde S, Olaisen B, Gedde-Dahl T, Teisberg P (1986): The C8A and C8B loci are closely linked on chromosome 1. *Ann Hum Genet* 50, 139-144

Rogne S, Myklebost O, Olving JH, Kyrkjebo HT, Jonassen R, Olaisen B, Gedde-Dahl Jr T (1991): The human genes for complement components 6 (C6) and 9 (C9) are closely linked on chromosome 5. *J Med Genet* 28, 587-590

Rosado CJ, Buckle AM, Law RH, Butcher RE, Kan WT, Bird CH, Ung K, Browne KA, Baran K, Bashtannyk-Puhlovich TA, Faux NG, Wong W, Porter CJ, Pike RN, Ellisdon AM, Pearce MC, Bottomley SP, Emsley J, Smith IA, Rossjohn J, Hartland EL, Voskoboinik I, Trapani JA, Bird PI, Dunstone MA, Whisstock JC (2007): A Common Fold Mediates Vertebrate Defense and Bacterial Attack. *Science* 317, 1548-1551

Rosendal S (1994): Ovine and caprine mycoplasmas. In: Whitford, HW, Rosenbush, RF, Laueran, LH (Eds): *Mycoplasmosis in Animals: Laboratory Diagnosis*. Iowa State University Press, Ames, 84-96

Ross SC, Densen P (1984): Complement deficiency states and infection: epidemiology, pathogenesis and consequences of neisserial and other infections in an immune deficiency. *Medicine* 63, 243-273

- Saucedo L, Ackermann L, Platonov AE, Gewurz A, Rakita RM, Densen P (1995): Delineation of additional genetic bases for C8 beta deficiency. Prevalence of null alleles and predominance of C→T transition in their genesis. *J Immunol* 155, 5022-5028
- Schmiedt W, Kinscherf R, Deigner HP, Kamencic H, Nauen O, Kilo J, Oelert H, Metz J, Bhakdi S (1998): Complement C6 Deficiency Protects Against Diet-Induced Atherosclerosis in Rabbits. *Arteriosclerosis, Thrombosis, and Vascular Biology* 8, 1790-1795
- Schreiber RD, Morrison DC, Podack ER, Mueller-Eberhard HJ (1979): Bactericidal activity of the alternative complement pathway generated from 11 isolated plasma proteins. *J Exp Med* 149, 870-882
- Schwab GE, Reeves PR (1966): Comparison of the Bactericidal Activity of Different Vertebrate Sera. *J Bacteriol* 91, 106-112
- Sehra H (2007): Pig DNA sequence from clone PigI-707F1 on chromosome 7, complete sequence: GenBank acc. no. AL773527 (NCBI homepage)
- Schreck SF, Parker C, Plumb ME, and Sodetz JM (2000): Human complement protein C8G. *Biochim Biophys Acta* 1482, 199-208
- Schreck SF, Plumb ME, Platteborze PL, Kaufman KM, Michelotti GA, Letson CS, Sodetz JM (1998): Expression and characterization of recombinant subunits of human complement component C8: Further analysis of the function of C8A and C8G. *J Immunol* 161, 311-318
- Schwarzenholz B, Correns A, Schroeder H, Geserick G (2000): Complement Component C6 Polymorphism: Novel Protein-Typing Technique and Distribution of Allotypes in Germany. *Exp Clin Immunogenet* 17, 23-28
- Scibek JJ, Plumb ME, Sodetz JM (2002): Binding of human complement C8 to C9: role of the N-terminal modules in the C8 alpha subunit. *Biochemistry* 41, 4546-4551
- Shibata T, Abe T, Tanabe Y (1995): Genetic polymorphism of the sixth component of complement (C6) in dogs. *Anim Genet* 26, 105-106

- Shibata T, Akita T, Abe T (1993): Genetic polymorphism of the sixth component of complement (C6) in the pig. *Anim Genet* 24, 97-100
- Shiverg JW, Dankert JR, Donovan JJ, Esserfl AF (1986): The ninth component of human complement (C9): Functional activity of the b fragment. *J Biol Chem* 261, 9629-9636
- Slade DJ, Chiswell B, Sodetz JM (2006): Functional studies of the MACPF domain of human complement protein C8alpha reveal sites for simultaneous binding of C8beta, C8gamma, and C9. *Biochemistry* 45, 5290-5296
- Smith KF, Harrison RA, Perkins SJ (1992): Molecular modeling of the domain structure of C9 of human complement by neutron and X-ray solution scattering. *Biochemistry* 31, 754-764
- Soelder BM, De Stefano GF, Dierich MP, Wuerzner R (1996): The Cayapa Indians of Ecuador: a genetically isolated group with unexpected complement C7 M/N allele frequencies. *Eur J Immunogenet* 23, 199-203
- Soejuma S, Tachida H, Tsuneoka M, Takenaka O, Kimura H, Yoda Y (2005): Nucleotide sequence analyses of human complement 6 (C6) gene suggest balancing selection. *Ann Hum Genet* 69, 239-252
- Scolding NJ, Morgan BP, Compston DA (1998): The expression of complement regulatory proteins by adult human oligodendrocytes. *J Neuroimmunol* 84, 69-75
- Stanley KK, Kocher HP, Luzio JP, Jackson P, Tschopp J (1985): The sequence and topography of human complement component C9. *Embo J* 4, 375-382
- Steckel EW, Welbaum BE, Sodetz JM (1983): Evidence of direct insertion of terminal complement proteins into cell membrane bilayers during cytolysis: labeling by a photosensitive membrane probe reveals a major role for the eighth and ninth components. *J Biol Chem* 258, 4318-4324
- Steckel EW, York RG, Monahan JB, Sodetz JM (1980): The eighth component of human complement: purification and physicochemical characterization of its unusual subunit structure. *J Biol Chem* 255, 11997-12005

Steele NP, Munson RS, Granoff JrDM, Cumming JE, Levine RP (1984): Antibody-dependent alternative pathway killing of *Haemophilus influenzae* type b. *Infect Immun* 44, 452-458

Sterzl J, Kostka J, Lanc A (1962): Development of bactericidal properties against gram-negative organisms in the serum of young animals. *Folia Microbiol* 7, 155-161

Stewart JL, Kolb WP, Sodetz JM (1987): Evidence that C5b recognizes and mediates C8 incorporation into the cytolytic complex of complement. *J Immunol* 139, 1960-1964

Stewart JL, Sodetz JM (1985): Analysis of the specific association of the eighth and ninth components of human complement: identification of a direct role for the alpha subunit of C8. *Biochemistry* 24, 4598-4602

Storm KE, Arturson G, Nilsson UR (1992): Purification and characterization of porcine C3. Studies of the biologically active protein and its split products. *Vet Immunol Immunopathol* 34, 47-61

Sunyer JO, Lambris JD (2001): Complement. In: Fullerlove G and Robertson S (Eds.). *Encyclopedia of Life Sciences* 5, Nature Publishing Group, London 1-5

Sutherland MA, Rodriguez-Zas SL, Ellis M, Salak-Johnson JL (2005): Breed and age affect baseline immune traits, cortisol, and performance in growing pigs. *J Anim Sci* 83, 2087-2095

Takeda I, Igarashi S, Nishimaki T, Kasukawa R (1994): A case of systemic lupus erythematosus in late component (C9) complement deficiency. *Ryumachi* 34, 628-632

Takata Y, Moriyama T, Fukumori Y, Yoden A, Shima M, Inai S (1989): A biotin-avidin sandwich ELISA for quantification of intact complement component C9. The sera from hereditary C9 deficient individuals completely lack C9. *Immunol Methods* 117, 107-113

Taylor PW, Knoll HP (1983): Killing of an encapsulated strain of *Escherichia coli* by human serum. *Infect Immun* 39, 122-131

- Taylor KM, Trimby AR, Campbell AK (1997): Mutation of recombinant complement component C9 reveals the significance of the N-terminal region for polymerization. *Immunology* 91, 20-27
- Thacker EL (2003): Immunology - the innate immune system. *The Pig Journal* 52, 111-123
- Thai CT, Ogata RT (2004): Complement Components C5 and C7: Recombinant factor I modules of C7 bind to the C345C domain of C5. *J Immunol* 173, 4547-4552
- Thai CT, Ogata RT (2005): Recombinant C345C and factor I modules of complement components C5 and C7 inhibit C7 incorporation into the complement membrane attack complex. *J Immunol* 174, 6227-6232
- Theriault A, Boyd E, Whaley K, Sodetz JM, Connor JM (1992): Regional chromosomal assignment of genes encoding the alpha and beta subunits of human complement protein C8 to 1p32. *Hum Genet* 88, 703-704
- Thiel S, Vorup-Jensen T, Stover CM, Schwaeble W, Laursen SB, Poulsen K, Willis AC, Eggleton P, Hansen S, Holmskov U, Reid KBM, Jensenius JC (1997): A second serine protease associated with mannan-binding lectin that activates complement. *Nature* 386, 506-510
- Thomsen DP, Winterø AK, Fredholm M (1998): Chromosomal assignments of 19 porcine cDNA sequences by FISH. *Mamm Genome* 9, 394-396
- Thuy NTD, Melchinger-Wild E, Kuss AW, Cuong NV, Bartenschlager, and Geldermann H (2006): Comparison of Vietnamese and European pig breeds using microsatellites. *J Anim Sci* 84, 2601-2608
- Tokunaga K, Yukiyma Y, Omoto K (1983): Polymorphism of the complement component C6 in Japanese. *Eur J Immunogenet* 10, 419-424
- Toshiro H, Ryutaro K, Kenji I, Hidetoshi T (1999): Inherited deficiency of plasma complement component and complement control protein. The clinical finding and gene abnormality. *Japanese Journal of Clinical Immunology* 22, 53-62

- Trakooljul N, Ponsuksili S, Schellander K, Wimmers K (2004). *Sus scrofa* complement C1qA mRNA, complete cds: GenBank acc. no. AY349424 (NCBI homepage)
- Trakooljul N, Ponsuksili S, Schellander K, Wimmers K (2004). *Sus scrofa* complement C1qB mRNA, complete cds: GenBank acc. no. AY349420 (NCBI homepage)
- Trakooljul N, Ponsuksili S, Schellander K, Wimmers K (2004). *Sus scrofa* complement C1r mRNA, complete cds: GenBank acc. no. AY349421 (NCBI homepage)
- Trakooljul N, Ponsuksili S, Schellander K, Wimmers K (2004). *Sus scrofa* complement C1s mRNA, complete cds: GenBank acc. no. AY349426 (NCBI homepage)
- Trakooljul N, Ponsuksili S, Schellander K, Wimmers K (2004). *Sus scrofa* complement C4 mRNA, partial cds: GenBank acc. no. AY349423 (NCBI homepage)
- Tschopp J (1984): Circular polymerization of the membranolytic ninth component of complement. Dependence on metal ions. *J Biol Chem* 259, 10569-10573
- Tschopp J (1984): Ultrastructure of the membrane attack complex of complement. Heterogeneity of the complex caused by different degree of C9 polymerization. *J Biol Chem* 259, 7857-7863
- Tschopp J, Masson D, Stanley KK (1986): Structural/functional similarity between proteins involved in complement- and cytotoxic T-lymphocyte-mediated cytolysis. *Nature* 322, 831-834
- Tschopp J, Podack ER (1981): Membranolysis by the ninth component of human complement. *Biochem Biophys Res Commun* 100, 1409-1414
- Tsunenari S, Kibayashi K, Higashi T, Pang H, Ding M, Jia J (1992): Genetic polymorphism of complement component C6, C7 and C8(1) in Chinese Han population in northeast China. *Nihon Hoigaku Zasshi* 46, 469-473
- Tyler JW, Cullor JS, Osburn BI, Parker K (1988): Age-related variations in serum concentrations of the third component of complement in swine. *Am J Vet Res* 49, 1104-1106

- Uenishi H, Eguchi T, Suzuki K, Sawazaki T, Toki D, Shinkai H, Okumura N, Hama-sima N, Awata T (2004): PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries. *Nucleic Acids Res* 32, 484-488
- Van Dixhoorn MGA, Timmerman JJ, Van Gijlswijk-Janssen DJ, Muizert Y, Verweij C, Discipio RG, Daha MR (1997): Characterization of complement C6 deficiency in a PVG/c rat strain. *Clin Exp Immunol* 109, 387-396
- Vázquez-Bermúdez MF, Barroso S, Walter K, Alvarez AJ, Alarcón A, López-Trascasa M, Wichmann I, Aguilar F, Núñez-Roldán A, Sánchez B (2003): Complement component C7 deficiency in a Spanish family. *Clin Expl Immunol* 133, 240-246
- Vishwanath S, Hackstadt T (1988): Lipopolysaccharide Phase Variation Determines the Complement-Mediated Serum Susceptibility of *Coxiella burnetii*. *Infect Immun* 56, 40-44
- Vissher AH, Janss LL, Niewold TA, de Greef KH (2002): Disease incidence and immunological traits for the selection of healthy pigs. A review. *Vet Q* 1, 29-34
- Vogler LB, Newman SL, Stroud RM, Johnston RB Jr (1979): Recurrent meningococcal meningitis with absence of the sixth component of complement: an evaluation of underlying immunologic mechanisms. *Pediatrics* 64, 465-467
- Volanakis JE (1995): Transcriptional regulation of complement genes. *Annu Rev Immunol* 13, 277-305
- Von Heijne G (1986): A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* 14, 4683-4690
- Washio K, Tokunaga K, Omoto K, Misawa S (1986): Human C7 polymorphism: classification and association analysis with C6. *J Hum Genet* 31, 345-352
- Whitehouse DB (1982): Genetic variation of the sixth component of complement (C6) in the Manx shearwater (*Puffinus p. puffinus*). *Anim Blood Groups Biochem Genet* 13, 179-187



- Whitehouse DB (1984): Genetic polymorphism and linkage of the sixth and seventh complement components (C6 and C7) in the common marmoset. *Biochem Genet* 22, 51-63
- Wilkie B, Mallard B (1999): Selection for high immune response: an alternative approach to animal health maintenance? *Vet Immunol Immunopathol* 72, 231-235
- Wimmers K (2002): Kandidatengen- und QTL-Analyse für Merkmale der humoralen Immunabwehr beim Schwein. Habilitationsschrift, Rheinische Friedrich-Wilhelms-Universität Bonn
- Wimmers K, Mekchay S, Schellander K, Ponsuksili S (2003): Molecular characterization of the pig C3 gene and its association with complement activity. *Immunogenetics* 54, 714-724
- Wintero AK, Jorgensen CB, Yerle M, Fredholm M (1998): Improvement of the porcine transcription map: localization of 33 genes of which 24 are orthologous. *Mamm Genome* 9, 366-372
- Wittmann G. (1984): State and control of Aujeszky's disease in West Germany. *Tierärztl Prax* 12, 141-147
- Witzel-Schloemp K, Hobart MJ, Fernie BA, Orren A, Wuerzner R, Rittner C, Kaufmann T, Schneider PM (1998): Heterogeneity in the genetic basis of human complement C9 deficiency. *Immunogenetics* 48, 144-147
- Witzel-Schloemp K, Rittner C, Schneider PM (2001): The human complement C9 gene: structural analysis of the 5' gene region and genetic polymorphism studies. *Eur J Immunogenet* 28, 515-522
- Witzel-Schloemp K, Spaeth PJ, Hobart MJ, Fernie BA, Rittner C, Kaufmann T, Schneider PM (1997): The human complement C9 gene: identification of two mutations causing deficiency and revision of the gene structure. *J Immunol* 158, 5043-5049
- Wuerzner R (2003): Deficiencies of the complement MACII gene cluster (C6, C7, C9): is subtotal C6 deficiency of particular evolutionary benefit?. *Clin Exp Immunol* 133, 156-159

- Wuerzner R, Hobart MJ, Fernie BA, Mewar D, Potter PC, Orren A, Lachmann PJ (1995): Molecular basis of subtotal complement C6 deficiency. A carboxy-terminally truncated but functionally active C6. *J Clin Invest* 95, 1877–1883
- Wuerzner R, Joysey VC, Lachmann PJ (1994): Complement component C7: assessment of in vivo synthesis after liver transplantation reveals that hepatocytes do not synthesize the majority of human C7. *J Immunol* 152, 4624-4629
- Wuerzner R, Tokunaga K, Nitze R, Goetze O, Lachmann PJ (1991): C7\*N is a hypomorphic allele of the human complement C7 M/N protein polymorphism. *Exp Clin Immunogenet.* 8, 177-183
- Wulffraat NM, Sanders EAM, Fijen CAP, Hannema A, Kuis W, Zegers BJM (1994): Deficiency of the beta subunit of the eighth component of complement presenting as arthritis and exanthem. *Arthritis Rheum* 37, 1704-1706
- Yerle M, Pinton P, Robic A, Milan D, Gellin J (1998): Generation of a porcine radiation hybrid panel: a new tool for mapping the pig genome. *Plant & Animal Genome VI Conference*. Town & Country Hotel, San Diego, CA, January 18-22
- Yerle M, Pinton P, Robic A, Alfonso A, Palvadeau Y, Delcros C, Hawken R, Alexander L, Beattie C, Schook L, Milan D, Gellin J (1998): Construction of a whole-genome radiation hybrid panel for high-resolution gene mapping in pigs. *Cytogenet and Cell Genetics* 82, 82-188
- Yi KS, Kang YH (2007): *Sus scrofa* C5a anaphylatoxin receptor mRNA, partial cds: *Sus scrofa* C5a anaphylatoxin receptor mRNA, partial cds. GenBank acc. no. EF210575 (NCBI homepage)
- Yonemasu K, Kitajima H, Tanabe S, Ochi T, Shinkai H (1978): Effect of age on C1q and C3 levels in human serum and their presence in colostrums. *Immunology* 35, 523-530
- Yonemura Y, Kawakita M, Koito A, Kawaguchi T, Nakakuma H, Kagimoto T, Schichishima T, Terasawa T, Akagaki Y, Inai S (1990): Paroxysmal nocturnal haemo-

globinuria with coexisting deficiency of the ninth component of complement: lack of massive haemolytic attack. *Br J Haematol* 74, 108-113

Zhang L, Rittner C, Sodetz JM, Schneider PM, Kaufmann T (1995): The eighth component of human complement: molecular basis of C8A (C81) polymorphism. *Hum Genet* 96, 281-284

Zimran A, Rudensky B, Kramer MR, Tedesco F, Ehrenfeld M, Raz R, Greif Z, Gelber M, Lishner, Golan E, Hershko C (1987): Hereditary complement deficiency in survivors of meningococcal disease: high prevalence of C7/C8 deficiency in sephardic (Moroccan) Jews. *Q J Med* 63, 349-358

Zipfel PF (2001): Complement: alternative pathway. In: Fullerlove G and Robertson S (Eds.). *Encyclopedia of Life Sciences* 5, Nature Publishing Group, London 25-32

Zhu ZB, Atkinson TP, Hovanky KT, Boppana SB, Dai YL, Densen P, Go RCP, Jablecki JS, Volanakis JE (2000): High prevalence of complement component C6 deficiency among African-Americans in the South-eastern USA. *Clin Exp Immunol* 119, 305-310

Zhu ZB, Totemchokchayakarn K, Atkinson TP, Volanakis JE (1998): Molecular defects leading to human complement component C6 deficiency in an African-American family. *Clin Exp Immunol* 111, 91-96

Zoppi M, Weiss M, Nydegger UE, Hess T, Spaeth PJ (1990): Recurrent meningitis in a patient with congenital deficiency of the C9 component of complement. First case of C9 deficiency in Europe. *Arch Intern Med* 150, 2395-2359

## 9 Annex

Hereafter, the primer sequences followed by TI number, name and mate information collected from the pig genome sequencing project accessible via the Trace Archive tool (NCBI homepage) are given in table 21, 22, 23, 24, 25, and 26.

Table 21 Position and length of predicted porcine C8A exons

Description	Source
Exon 1: nt. 5'UTR-171	Based on start of exon 2
Exon 2: nt. 172-271 (100 bp)	ti: 862132317; name:rdpaxb0_150586.y1; mate:815435067
Exon 3: nt. 272-416 (145 bp)	ti: 863184494; name:rdpaxb0_077985.y1; mate:854478286
Exon 4: nt. 417-564 (148 bp)	ti: 852982643; name:cpg0_164359.z1 mate:815846657
Exon 5: nt. 565-754 (190 bp)	Based on end of exon 4 and start of exon 6
Exon 6: nt. 755-955 (201 bp)	ti: 1420494233; name:ss_WGS-24b13.p1k; mate:1420494234
Exon 7: nt. 956-1196 (241 bp)	ti: 1420494234; name:ss_WGS-24b13.q1k; mate:1420494233
Exon 8: nt. 1197-1328 (132 bp)	Based on end of exon 7 and start of exon 9
Exon 9: nt. 1329-1486 (158 bp)	ti: 1420545232; name:SS_WGS-667o09.q1k; mate:1420545231
Exon 10: nt. 1487-1709 (223 bp)	Based on end of exon 9 and start of exon 11
Exon 11: nt. 1710-3'UTR	ti: 1650055823 name:DurocWGS_10036_F12.g mate:1650055822

Table 22 Position and length of predicted porcine C6 exons

Description	Source
Exon 1: nt. 15 (5'UTR)-121	ti:1008570313; name:DurocWGS_20549_B01.b; mate:1008570314
Exon 2: nt. 122-287 (166bp)	ti: 812320666; name: rcpg0_152426.y1 ; mate: 858788103
Exon 3: nt. 288-444 (157 bp)	ti:1008617751; name:DurocWGS_20796_C15.b ; mate:1008617752
Exon 4: nt. 445-589 (145 bp)	ti: 815809106; name: byc_44258.z1; mate: 812837595
Exon 5: nt. 590-731 (142 bp)	ti: 808385472; name: dpcxb0_315815.z1; mate: 810707950
Exon 6: nt. 732-870 (139 bp)	ti: 1373546263; name: rplun0101_g6.y1
Exon 7: nt. 871-1071 (201 bp)	ti: 857624103; name: rbde_60509.y1; mate: 861251703
Exon 8: nt. 1072-1312 (241 bp)	ti: 1420614517; name: SS_WGS-857e09.p1k; mate:1420614518
Exon 9: nt. 1313-1435 (123 bp)	ti: 1420666160; name: SS_WGS-513i21.p1k mate:1420666161
Exon 10: nt. 1438-1602 (167 bp)	ti: 1420506449; name: SS_WGS-957p20.q1k; mate:1420506448
Exon 11: nt. 1603-1828 (226 bp)	ti: 1420506448 ; name: SS_WGS-957p20.p1k; mate:1420506449
Exon 12: nt. 1829-2000 (172 bp)	Based on end of exon 11 and start of exon 13
Exon 13: nt. 2001-2112 (112 bp)	ti: 821008954; name: rdpaxb0_030576.y1; mate:854483952
Exon 14: nt. 2113-2245 (134 bp)	ti: 847766479; name: rbdd_56378.y1; mate:768215414
Exon 15: nt. 2246-2434 (188 bp)	ti: 768215414; name: bdd_56378.z1; mate: 847766479
Exon 16: nt. 2435-2525 (91 bp)	ti: 774040088; name: bdc_90777.z1; mate:823852002
Exon 17: nt. 2526-2766 (242 bp)	Based on human genomic C6 sequence
Exon 18 : nt. 2767-3'UTR	Based on human genomic C6 sequence

Table 23 Position and length of predicted porcine C7 exons

Description	Source
Exon 1: nt. -91 (5'UTR)-6 (96 bp)	Based on human genomic C7 sequence and ti: 848892623; name:rdpaxb0_005371.y1; mate:854561309
Exon 2: nt. 7-62 (56bp)	ti: 774255673 name:rbda_68237.y1
Exon 3: nt. 63-138 (76 bp)	ti: 862200672; name:dpcxb0_048413.z1; ma- te:808192830
Exon 4: nt. 139-280 (142 bp)	ti: 848149583 ; name:bd_52849.z1; mate:863305860
Exon 5: nt. 281-428 (148 bp)	ti: 862214178; name:dpcxb0_129835.z1; mate:808226750
Exon 6: nt. 429-567 (139 bp)	Based on end of exon 5 and start of exon 7
Exon 7: nt. 568-738 (171 bp)	ti: 1420727546; name:SS_WGS-599c18.p1k; mate:1420727547
Exon 8: nt. 739-982 (244 bp)	ti: 848447315; name:rddb_77708.y1; mate:776645100
Exon 9: nt. 983-1093 (111 bp)	ti: 815376779 name:rdpaxb0_028437.y1 mate:857329704
Exon 10: nt. 1094-1260 (167 bp)	ti: 1420836647 ; name:SS_WGS-971f04.q1k mate:1420836646
Exon 11: nt. 1261-1489 (229 bp)	ti: 842736254; name:rdbf_90630.y1; mate:769812863
Exon 12: nt. 1490-1661 (172 bp)	ti: 311627814; name:bE40H18T7; mate:311627813
Exon 13: nt. 1662-1749 (88 bp)	ti: 853061954; name:rbdg_17201.y1; mate:776550611
Exon 14: nt. 1750-1882 (133 bp)	ti: 1420455996; name:SS_WGS-658h03.p1k; mate:1420455997
Exon 15: nt. 1883-2074 (192 bp)	ti: 854810180; name:rdpaxb0_227330.y1
Exon 16: nt. 2075-2165 (91 bp)	ti: 1657934248; name: DurocWGS_30149_C10.b mate:1657934249
Exon 17: nt. 2166-2350 (185 bp)	ti: 863233311; name:rdpba0_029753.y1 mate:808444829
Exon 18: nt. 2351-3'UTR	ti: 1658588611 ; name: DurocW- GS_20991_D07.g; mate:1658588610

Table 24 Position and length of predicted porcine C8B exons

Description	Source
Exon 1: nt. 5'UTR-156	Based on start of exon 2
Exon 2: nt. 157-313 (157 bp)	ti: 1420204053; name:SS_WGS-413k10.q1k; mate:1420204052
Exon 3: nt. 314-455 (142 bp)	ti: 823797500; name:bdf_14280.z1; mate:847963640
Exon 4: nt. 456-597 (142 bp)	Based on end of exon 3 and start of exon 5
Exon 5: nt. 598-730 (133 bp)	ti: 781569495; name:rbyb_35660.y1; mate:812924227
Exon 6: nt. 731-928 (198 bp)	Based on end of exon 5 and start of exon 7
Exon 7: nt. 929-1169 (241 bp)	ti: 857822592; name:rdpboxa0_158567.y1; mate:853343914
Exon 8: nt. 1170-1298 (129 bp)	Based on end of exon 7 and start of exon 9
Exon 9: nt. 1299-1462 (164 bp)	ti: 1420654355; name:SS_WGS-860j08.p1k; mate:1420654356
Exon 10: nt. 1462-1616 (154 bp)	ti: 860428303; name:rbye_44037.y1
Exon 11: nt. 1617-1684 (68 bp)	Based on human genomic C8B sequence
Exon 12 : nt. 1685-3' UTR	Based on human genomic C8B sequence

Table 25 Position and length of predicted porcine C8G exons

Description	Source
Exon 1: nt. 50 (5'UTR)-189 (140bp)	Based on human genomic C8G sequence and GenBank acc. no AK233484
Exon 2: nt. 190-326 (137 bp)	Based on GenBank acc. no AK233484
Exon 3: nt. 327-397 (71 bp)	Based on GenBank acc. no AK233484
Exon 4: nt. 398-505 (108 bp)	Based on GenBank acc. no AK233484
Exon 5: nt. 506-607 (102 bp)	Based on GenBank acc. no AK233484 and human genomic C8G sequence
Exon 6: nt. 608-646 (39 bp)	Based on human genomic C8G sequence
Exon 7: nt. 647-3'UTR	Based on human genomic C8G sequence

Table 26 Position and length of predicted porcine C9 exons

Description	Source
Exon 1: 5'UTR-166	Based on human genomic C9 sequence
Exon 2: nt. 167-272 (106 bp)	Based on human genomic C9 sequence and start of exon 3
Exon 3: nt. 273-417 (145 bp)	ti: 1382154540; name:DurocWGS_21175_M13.g; mate:1382154539
Exon 4: nt. 418-565 (148 bp)	ti: 1382154540; name:DurocWGS_21175_M13.g; mate:1382154539
Exon 5: nt. 566-704 (139 bp)	ti: 311779440; name:bT60L15SP6; mate:311779441
Exon 6: nt. 705-971 (267 bp)	ti: 863194696; name:cpg0_070876.z1; mate:811045924
Exon 7: nt. 972-1212 (241 bp)	ti: 767686602; name:DurocWGS_20407_A01.g; mate:767686601
Exon 8: nt. 1213-1338 (126 bp)	ti: 784826320 name:dpcxb0_271689.z1; mate:766131427
Exon 9: nt. 1339-1517 (179bp)	ti: 767686601; name:DurocWGS_20407_A01.b; mate:767686602
Exon 10: nt. 1518- 1747 (3'UTR) (230 bp)	ti: 854554002; name:rdpcxb0_001135.y1; mate:853317835
Exon 11: nt. 1748-3'UTR	Based on end of exon 10

Table 27 Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C6 gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	AA	AG	GG
Mycoplasma	1	45.87±9.32	48.44±3.87	53.57±3.78
	2	42.99±10.03	56.03±3.99	61.52±3.87
	3	51.46±11.02	53.92±4.15	62.28±3.99
Aujeszky	4	62.43±11.63	68.50±4.20	76.68±4.04
	5	83.05±12.00	79.24±4.29	82.21±4.12
	6	76.70±12.63	67.72±4.40	75.51±4.16
PRRSV	7	52.92±12.73	73.90±4.12	76.27±3.92
	8	62.38±12.22	72.12±4.27	75.33±4.07



Table 28 Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C7 gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	AA	AG	GG
Mycoplasma	1	52.89±4.08	49.87±4.08	45.21±10.28
	2	60.58±4.19	59.95±4.28	43.85±10.94
	3	62.35±4.32	56.90±4.42	51.17±11.86
Aujeszky	4	77.88±4.45	73.05±4.52	25.44±14.28
	5	85.61±4.45	81.13±4.54	55.65±13.48
	6	75.47±4.48	70.08±4.65	39.86±13.72
PRRSV	7	74.46±4.25	75.80±4.36	51.32±13.49
	8	72.93±4.42	74.99±4.55	40.56±13.53

Table 29 Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C8A gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	CC	CT	TT
Mycoplasma	1	49.51±4.67	45.41±4.01	52.96±3.72
	2	50.95±5.14	52.89±4.10	62.38±3.89
	3	50.39±5.34	51.94±5.26	62.59±4.02
Aujeszky	4	62.06±5.49	71.76±4.30	71.94±4.15
	5	70.59±5.67	81.61±4.37	80.35±4.17
	6	63.46±6.13	69.25±4.49	71.91±4.23
PRRSV	7	62.51±5.66	74.86±4.23	73.00±4.03
	8	66.11±5.80	72.91±4.43	70.78±4.19

Table 30 Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C8B gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	CC	CT	TT
Mycoplasma	1	51.54±3.45	42.83±4.02	52.53±5.69
	2	59.06±3.65	50.36±4.14	57.16±6.51
	3	59.91±3.75	46.65±4.27	60.38±6.73
Aujeszky	4	67.91±3.86	70.00±4.25	69.52±6.80
	5	76.15±3.89	79.06±4.33	79.04±7.05
	6	69.27±3.95	65.79±4.47	74.13±7.66
PRRSV	7	71.98±3.73	70.64±4.20	59.09±6.77
	8	68.84±3.89	71.21±4.39	62.28±7.24

Table 31 Least squares means of the classical complement hemolytic activity for the interaction of different genotypes x time point in porcine C9 gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	CG	GG
Mycoplasma	1	47.62±4.31	52.66±3.55
	2	56.91±4.45	59.36±3.63
	3	52.23±4.64	61.18±3.74
Aujeszky	4	63.73±4.63	77.24±3.79
	5	78.12±4.72	83.09±3.85
	6	67.60±4.85	73.92±3.91
PRRSV	7	71.48±4.53	74.96±3.70
	8	71.03±4.82	73.51±3.87

Table 32 Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C6 gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	AA	AG	GG
Mycoplasma	1	56.27±11.86	51.75±3.77	54.02±3.44
	2	37.72±11.31	53.66±3.64	48.00±3.30
	3	54.11±10.58	56.15±3.48	50.58±3.18
Aujeszky	4	65.60±10.31	56.89±3.44	61.07±3.20
	5	86.65±11.56	64.69±3.60	62.17±3.30
	6	76.79±15.87	67.59±4.57	71.59±4.05
PRRSV	7	81.87±12.75	71.72±4.07	72.94±3.67
	8	76.54±12.91	71.59±4.06	70.88±3.67

Table 33 Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C7 gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	AA	AG	GG
Mycoplasma	1	53.56±3.64	51.37±3.94	49.71±13.31
	2	47.80±3.55	52.67±3.86	37.21±12.95
	3	50.87±3.36	56.71±3.64	50.22±11.88
Aujeszky	4	62.33±3.46	56.65±3.66	61.58±12.75
	5	63.27±3.53	65.03±3.81	72.53±13.24
	6	70.35±4.16	67.84±4.62	72.27±16.98
PRRSV	7	69.51±3.64	70.55±3.99	58.73±13.99
	8	71.17±4.06	70.58±4.49	75.29±16.60

Table 34 Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C8A gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	CC	CT	TT
Mycoplasma	1	51.29±5.18	50.06±3.81	54.09±3.45
	2	40.85±4.97	45.31±3.68	54.56±3.35
	3	58.62±4.80	46.60±3.54	54.48±3.16
Aujeszky	4	56.99±4.80	59.39±3.52	57.46±3.23
	5	58.74±5.11	63.67±3.59	64.57±3.40
	6	83.69±7.34	69.07±4.65	66.50±4.42
PRRSV	7	79.50±6.04	76.56±4.15	66.02±3.81
	8	65.87±5.90	74.42±4.06	69.50±3.74

Table 35 Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C8B gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	CC	CT	TT
Mycoplasma	1	54.90±3.45	49.80±3.76	54.52±6.08
	2	53.29±3.27	44.56±3.57	47.81±5.66
	3	54.46±3.09	46.60±3.46	64.73±5.42
Aujeszky	4	56.44±3.11	56.14±3.40	64.55±5.56
	5	64.67±3.30	62.88±3.52	59.53±5.86
	6	67.96±4.24	70.97±4.48	79.46±8.53
PRRSV	7	66.59±3.71	78.58±4.03	73.83±7.01
	8	69.37±3.66	72.82±3.96	67.47±6.95

Table 36 Least squares means of the alternative complement hemolytic activity for the interaction of different genotypes x time point in porcine C9 gene (LSM±SE) (U/ml)

Vaccination	Blood sampling	CG	GG
Mycoplasma	1	52.52±4.32	53.72±3.21
	2	56.08±4.08	47.81±3.10
	3	56.19±3.96	51.98±2.98
Aujeszky	4	53.81±3.90	62.66±2.97
	5	67.46±4.15	63.00±3.12
	6	71.90±5.25	68.71±3.76
PRRSV	7	72.70±4.63	71.66±3.46
	8	71.17±4.77	72.61±3.44

<b>C6 seq</b>	agcttctg	9
<b>C6 seq</b>	agc <b>agg</b> acagaaggaaaaataagaaataaaagattcaaaatat	54
<b>C6 seq</b>	ttacaagaagatggaaaggaggatcctacagttgggaacaaacta	99
<b>C6 seq</b>	gagaatcttctgaacctgcc <b>agg</b> atctggagactctccaggcatg	144
<b>C6 seq</b>	<b>atg</b> gacaaacactctgtcttgtatTTTTatcttgctgagtggtctg	189
<b>C6 A.A</b>	M D K H S V L Y F I L L S G L	15
<b>C6 seq</b>	attgacaagagccaagcctgTTTctgtgatcactacccatggact	234
<b>C6 A.A</b>	I D K S Q A <u>C</u> F <u>C</u> D H Y P W T	30
<b>C6 seq</b>	cagtgggccagctgctcaaaaacctgtaattctggaaccagacc	279
<b>C6 A.A</b>	Q W S S <u>C</u> S K T <u>C</u> N S G T Q T	45
<b>C6 seq</b>	agacag <b>ag</b> acaaattactataaatcagtactatcttgacaacttt	324
<b>C6 A.A</b>	R Q R Q I T I N Q Y Y L D N F	60
<b>C6 seq</b>	tgtgaccggcTTTgcaccaagcaggagaccagagaatgtaactgg	369
<b>C6 A.A</b>	<u>C</u> D R L <u>C</u> T K Q E T R E <u>C</u> N W	75
<b>C6 seq</b>	caaacttgTcctatcaattgcctcctgggagattatggaccatgg	414
<b>C6 A.A</b>	Q T <u>C</u> P I N <u>C</u> L L G D Y G P W	90
<b>C6 seq</b>	tcagattgtgacccttgTgttgaaaaacag <b>gt</b> ttaaggttaggTcc	459
<b>C6 A.A</b>	S D <u>C</u> D P <u>C</u> V E K Q F K V R S	105
<b>C6 seq</b>	atcttgcgccccaatcagTttggaggacaaccatgtactgagcca	504
<b>C6 A.A</b>	I L R P N Q F G G Q P <u>C</u> T E P	120
<b>C6 seq</b>	ctcatgacctttcgaccatgtattccatctaaactctgcaaaatt	549
<b>C6 A.A</b>	L M T F R P <u>C</u> I P S K L <u>C</u> K I	135
<b>C6 seq</b>	gaagaggTtgactgcaagaataaattccgctgtgacagT <b>gg</b> tcgc	594
<b>C6 A.A</b>	E E V D <u>C</u> K N K F R <u>C</u> D S G R	150
<b>C6 seq</b>	tgcattgccagcaagTtggaatgcaatggagaaaatgattgtgga	639
<b>C6 A.A</b>	<u>C</u> I A S K L E <u>C</u> N G E N D <u>C</u> G	165
<b>C6 seq</b>	gacaattcggatgagaggaattgtgggagaaacaaaagcagTatgc	684
<b>C6 A.A</b>	D N S D E R N <u>C</u> G R T K A V <u>C</u>	180

Figure 40 Nucleotide and predicted protein sequence of the porcine C6 gene.

Nucleotide and amino acids are enumerated on the right side. The numbering of the nucleotide sequence starts with the first base of GenBank acc. no BP445061 and ends with the last base of GenBank acc. no BQ598214. The 5' flanking region contains 144 nucleotides (nt. 1-144). The start codon (ATG) is in the first bold box (nt. 145-147). The porcine boundaries between exons are in bold and underline. Positions of polymorphism present in shade blocks with white bold font. Cysteine residues are in underline. The open reading frame is terminated by a TGA stop codon in the second bold box (nt. 2950-2952). The 3'UTR (nt. 2953-3306) contains the putative polyadenylation signal aataaa (nt. 3240-3245) in bold and italic. The translation stop codon is identified by (\*)

C6 seq tcacggaagtataatcccatccctagtgtagcagttgatgggcgct 729  
 C6 A.A S R K Y N P I P S V Q L M G A 195  
 C6 seq gggtttcatattctggcaggagagcccagaggagaagtccttaggt 774  
 C6 A.A G F H I L A G E P R G E V L G 210  
 C6 seq aattctttcactggaggaatatgtaaaaccgtcaaaagcagtaaa 819  
 C6 A.A N S F T G G I C K T V K S S K 225  
 C6 seq gccagtaatccatttctgtgttccagccaatctagaaaacgtccac 864  
 C6 A.A A S N P F R V P A N L E N V N 240  
 C6 seq tttgagggtacaaactaaagaagatgattttcaaacagattttctat 909  
 C6 A.A F E V Q T K E D D F Q T D F Y 255  
 C6 seq gaggatttaattcctcttgaaaatagtaaagatcaacaagccaca 954  
 C6 A.A E D L I P L E N S K D Q Q A T 270  
 C6 seq ggtttcggccaagagaagagctcttttcatgtaccaattttttat 999  
 C6 A.A G F G Q E K S S F H V P I F Y 285  
 C6 seq tcctcaaagaaaagccaaaccagcagccataattctgccttcaaa 1044  
 C6 A.A S S K K S Q T S S H N S A F K 300  
 C6 seq caagctattcaagcctcccacaaaaaggattccagttttattagg 1089  
 C6 A.A Q A I Q A S H K K D S S F I R 315  
 C6 seq atccataaagtgataaaagtcttaaacttcacaatgaaaactaaa 1134  
 C6 A.A I H K V I K V L N F T M K T K 330  
 C6 seq gatctgcagctttctgacgtctttttgaaagcacttaaccatctg 1179  
 C6 A.A D L Q L S D V F L K A L N H L 345  
 C6 seq cctctagaatacaacgctgctttgtacagccgatatttgatgat 1224  
 C6 A.A P L E Y N A A L Y S R I F D D 360  
 C6 seq tttgggactcactatttcacctctggctccctgggaggcggtgac 1269  
 C6 A.A F G T H Y F T S G S L G G V Y 375  
 C6 seq gaccttctctatcagtttagcaatgaggaactgaagaactcagggt 1314  
 C6 A.A D L L Y Q F S N E E L K N S G 390  
 C6 seq ttaacacaggaagaagccaaaaactgtatccggattgaaacaaag 1359  
 C6 A.A L T Q E E A K N C I R I E T K 405  
 C6 seq aaacgttattttatagttacaaaaacaaaggtggaacaccgggtgc 1404  
 C6 A.A K R Y F I V T K T K V E H R C 420  
 C6 seq accactaacaggatgtccgagaaatacgaagggttcctttttgcaa 1449  
 C6 A.A T T N R M S E K Y E G S F L Q 435  
 C6 seq ggatcagagaaatccatataccctgggttaaaggtggaagaagtgag 1494  
 C6 A.A G S E K S I S L V K G G R S E 450  
 C6 seq tatgcagcggctttggcatgggagaaagggagttctggcccagga 1539  
 C6 A.A Y A A A L A W E K G S S G P G 465  
 C6 seq gagaagacatattctgactgggttggaaatcagtgaaggaaaatcct 1584  
 C6 A.A E K T Y S D W L E S V K E N P 480  
 C6 seq gctgtgattgactttgagccttgctcccatcacagacttggttaaga 1629  
 C6 A.A A V I D F E L A P I T D L V R 495  
 C6 seq aacatcccctgtgcggtgacaagacggaacaacctcaggagagct 1674  
 C6 A.A N I P C A V T R R N N L R R A 510  
 C6 seq ttccgagaatacgcagccaaatttgacccttgccagtggtgctccg 1719  
 C6 A.A F R E Y A A K F D P C Q C A P 525  
 C6 seq tgccccaacaacggccgcccctgtgctctcagggaccgaatgtcta 1764  
 C6 A.A C P N N G R P V L S G T E C L 540  
 C6 seq tgtgtgtgccagagcggcacctacggtgacaactgtgagaggcgt 1809  
 C6 A.A C V C Q S G T Y G D N C E R R 555

Figure 40 (continued)

C6 **seq** gccccggattacaaatctaatgccgtagatgggaactggggctgc 1854  
C6 **A.A** A P D Y K S N A V D G N W G C 570  
C6 **seq** tggcttcctggagcacatgtgacgctacgtacaagagatcaagg 1899  
C6 **A.A** W S S W S T C D A T Y K R S R 585  
C6 **seq** acccgagaatgcaacaaccctgcccccggaaggaggaaacc 1944  
C6 **A.A** T R E C N N P A P R Q G G K P 600  
C6 **seq** tgcgacggggagaggcggcaggaggaacactgcacattctccata 1989  
C6 **A.A** C D G E R R Q E E H C T F S I 615  
C6 **seq** atgcagaacactggacagccatgcatcggatgatgaggacatg 2034  
C6 **A.A** M Q N T G Q P C I G D D E D M 630  
C6 **seq** aaagaaatagaccttcctgaaatagaatcagattcagggtgtcct 2079  
C6 **A.A** K E I D L P E I E S D S G C P 645  
C6 **seq** caaccagtccctccagaaaatggatttattcggaatgaaaagaca 2124  
C6 **A.A** Q P V P P E N G F I R N E K T 660  
C6 **seq** ctgtattcagttggggaagatggtgaaattgtatgccttactgga 2169  
C6 **A.A** L Y S V G E D V E I V C L T G 675  
C6 **seq** ttcaaaactgttggctaccagtacttcagatgcttacctgacaga 2214  
C6 **A.A** F K T V G Y Q Y F R C L P D R 690  
C6 **seq** acctggaggcgaggagatgtggaatgccagcggactgactgcctc 2259  
C6 **A.A** T W R R G D V E C Q R T D C L 705  
C6 **seq** ttgccagtgggtgccagaagtactgaaactgtcaccatttcagaga 2304  
C6 **A.A** L P V V P E V L K L S P F Q R 720  
C6 **seq** ttgtataaaattgggtgactccattgagctaagctgtcccaaaggc 2349  
C6 **A.A** L Y K I G D S I E L S C P K G 735  
C6 **seq** tttgtcattgctgggccatcgaggtacacatgcagtggggattcc 2394  
C6 **A.A** F V I A G P S R Y T C S G D S 750  
C6 **seq** tggacaccaccatctcaagctcactcacctgtgaacaagatttt 2439  
C6 **A.A** W T P P I S S S L T C E Q D F 765  
C6 **seq** ctgacacgcttaaagggccattgtcaaccaggacaaaaacagttg 2484  
C6 **A.A** L T R L K G H C Q P G Q K Q L 780  
C6 **seq** ggatctgaatgtgtttgctgtctccagaagaagactgcaggccat 2529  
C6 **A.A** G S E C V C L S P E E D C S H 795  
C6 **seq** cattcagaagatctctgtgtgcttgatacagactccaatcattac 2574  
C6 **A.A** H S E D L C V L D T D S N H Y 810  
C6 **seq** tttacttcatccgcttgtaagtttttggctggagaaatgtttaaat 2619  
C6 **A.A** F T S S A C K F L A E K C L N 825  
C6 **seq** aatcagcaactccaatttctacatattggttcgtgccaaagatgg 2664  
C6 **A.A** N Q Q L Q F L H I G S C Q D G 840  
C6 **seq** ccacagttggaatggggtccttgaaaggataaaactttcatccagt 2709  
C6 **A.A** P Q L E W G L E R I K L S S S 855  
C6 **seq** agcacaagaatgaatcctgtggctatgatacctgctacagctgg 2754  
C6 **A.A** S T K N E S C G Y D T C Y S W 870  
C6 **seq** gaaaaatgttcagccaccacctccaagtgcattctgcctactgccc 2799  
C6 **A.A** E K C S A T T S K C I C L L P 885  
C6 **seq** ttccagtgtttgaagggtggataccaacactactgcgctcaaaatg 2844  
C6 **A.A** F Q C L K G G Y Q H Y C V K M 900  
C6 **seq** ggatcatcaacgaccacaagaaccatgaacatctgtgaagtagga 2889  
C6 **A.A** G S S T T T R T M N I C E V G 915  
C6 **seq** gccataagatgtgccctcaggaaaatggaaatactttatcctggg 2934  
C6 **A.A** A I R C A L R K M E I L Y P G 930  
C6 **seq** agatgtaggtctaactga 2952  
C6 **A.A** R C R S N \* 935

Figure 40 (continued)

```

C6 seq aatcttgataaacagatcttactgtccaaaaataatccctacagag 2997
C6 seq atgctcttgataaacagcagactgccacatccaggagttactaa 3042
C6 seq cataaattcttttgtgtagtttcatgtaattattctccctgct 3087
C6 seq cttgtgtcttctcctctatccaattctagtccacacccttaggta 3132
C6 seq ataactttcacatcagctctagtaaattattctgttcatgtggcac 3177
C6 seq taaaatgatggcacttctaatagaccttatgaagtacaaaaacctt 3222
C6 seq aatttcttcaatgaatcaataaacatagaagtctcgaactgctta 3267
C6 seq tcaaatacattcttgaaactgaaaaaaaaaaaaaaaaaaaaa 3306

```

Figure 40 (continued)

```

-806 catggaaaatgaggctctggcaaaaaaaagacaaggcaactg -765
C7 seq aggaagagaaaatggacagggataatgactgtgacaagaacttcc -720
C7 seq caaaagagagatctgaatggcataggattaagggaagttcttgagg -675
C7 seq aaaattagtgatgcctttaaataatgtttgctaaaaatgatgctct -630
C7 seq ttttgatcattctgtgaaagtccgacgtgaagggatgcttattat -585
C7 seq gttgtgttaggctatctctgtatattctggggggattccactcca -540
C7 seq gaatgttccacttacattttgggtaactgaactgcgtttgaattt -495
C7 seq ttttcctccacaaaagcgtttaactcaaaaggctgttgataaacac -450
C7 seq ccatttagagaatcatgaaaacatccacttagtacagtagatcga -405
C7 seq aggaggcttatcccttgccttacacagtctccttgggttgaaatc -360
C7 seq acttaaacattgaagtgagaacatttgctgcctttcttgatttat -315
C7 seq agatgtttttcttagaactacatttccaaacagtcccagctggct -270
C7 seq gtagctaaggcctgtgtaaattgataagatgcatttggcttcaag -225
C7 seq acacagactaagtacagggcagactcctgggtgtccccagagaact -180
C7 seq ccgttggggtagaagcagcatctcggaactgcttgcctctatatat -135
C7 seq gagtgccaagttaatccccagcagGgagggacaggcaaggaacag -90
C7 seq actgttgggctcttctgctgctgaaaattcactgggcactggag -45
C7 seq gaggaaatctaccttactcttctgccctgaatgttttcccaaac 0

```

Figure 41 Nucleotide and predicted protein sequence of the porcine C7 gene.

Nucleotide and amino acids are enumerated on the right side. The numbering of the nucleotide sequence starts with the first base of trace\_name:dpcxa 0\_039446.z1, ti: 784743778 and ends with the last base of GenBank acc. no CF176130.1. The start codon (ATG) is in the first bold box (nt. 1-3). The porcine boundaries between exons are in bold and underline. Positions of polymorphism present in shade blocks with white bold font. Cysteine residues are in underline. Transcription start site (**G**) is in uppercase and bold at position -110. The open reading frame is terminated by a TAG stop codon in the second bold box (nt. 2530-2532). The downstream noncoding region contains further termination codons but does not include a polyadenylation signal. The translation stop codon is identified by (\*)

C7 seq atgaaggcaatgagtttagtcttcttggtgggacttataggagag 45  
 C7 A.A M K A M S L V F L V G L I G E 15  
 C7 seq ttccaagttttttcagtgcctcctctcctgtcaattgtcagtgg 90  
 C7 A.A F Q V F S S A S S P V N C Q W 30  
 C7 seq gattcctatgctccttggtcagaatgcaatggtgtaccaagact 135  
 C7 A.A D S Y A P W S E C N G C T K T 45  
 C7 seq gactcgcaggagaccgttgctgtttatgggcagtatggcggg 180  
 C7 A.A Q T R R R P V A V Y G Q Y G G 60  
 C7 seq catccctgtgtcggaagtacgtttgaacacaatcatgtgaacct 225  
 C7 A.A H P C V G S T F E T Q S C E P 75  
 C7 seq acacgaggatgtccaacagaagaagggtgtggagagcgtttcaga 270  
 C7 A.A T R G C P T E E G C G E R F R 90  
 C7 seq tgtttttcaggtcagtgcacagcaaacttttggtttgcaacggg 315  
 C7 A.A C F S G Q C I S K S L V C N G 105  
 C7 seq gattctgactgtgaagaagacagtgtgatgaagaccggtgtgag 360  
 C7 A.A D S D C E E D S A D E D R C E 120  
 C7 seq gactcagaaaagcagaccttctcctgtgacctcagtaaacctcctccc 405  
 C7 A.A D S E S R P S C D L S K P P P 135  
 C7 seq aacatagaacttactggaaatggttacaatgcactcacgggccag 450  
 C7 A.A N I E L T G N G Y N A L T G Q 150  
 C7 seq tttaggaacagagtcctcaacactaaaagttttggtggtcaatgc 495  
 C7 A.A F R N R V L N T K S F G G Q C 165  
 C7 seq agaaaggtgttttagtggggatgggagagatttctacagactgagt 540  
 C7 A.A R K V F S G D G R D F Y R L S 180  
 C7 seq ggaaatgtcctctcctacacattccaggtgaaagtaaataatgat 585  
 C7 A.A G N V L S Y T F Q V K V N N D 195  
 C7 seq tttaattatgaattttacaatagtacctggtccttatgcaaaacat 630  
 C7 A.A F N Y E F Y N S T W S Y A K H 210  
 C7 seq acatctacagaacatacatcatccagtaaaggacgcgtcttcatt 675  
 C7 A.A T S T E H T S S S K G R V F I 225  
 C7 seq tttagttcttctcctcttctcctccagttattatgcaaaaacctat 720  
 C7 A.A F S S S S S S S S Y Y A K T Y 240  
 C7 seq gaaatccttaagaaaagagttatcagttggttggttgcagaac 765  
 C7 A.A E I L K K K S Y Q L L V V Q N 255  
 C7 seq actggtgaagtggctcaatttatcaataacaatccagaatttttg 810  
 C7 A.A T V E V A Q F I N N N P E F L 270  
 C7 seq caactgctgagtcattctggaaggaactctcctaccttctcct 855  
 C7 A.A Q L A E S F W K E L S Y L P P 285  
 C7 seq ctgtatgattacaggccctaccgaagattaattgaccagtatggg 900  
 C7 A.A L Y D Y S A Y R R L I D Q Y G 300  
 C7 seq acacattatctgcagctctgggtccttaggaggagaatacaaaagt 945  
 C7 A.A T H Y L Q S G S L G G E Y K V 315  
 C7 seq ttattttatgtggactcagagaaagtgcagaaagcgatcttggt 990  
 C7 A.A L F Y V D S E K V A E S D L G 330  
 C7 seq tcagaggataagaagaaatgtgcttctcacatattagttttcta 1035  
 C7 A.A S E D K K K C A S S H I S F L 345  
 C7 seq tttaaatcatcaaagcacaatgcaaggcgatggaagaggcctta 1080  
 C7 A.A F K S S K H K C K A M E E A L 360  
 C7 seq aatcagcttcaggaactcagagcaatgtggtgagggggtcccg 1125  
 C7 A.A K S A S G T Q S N V L R G V P 375  
 C7 seq tttgtcagagggggacgtcctggctttgtgtctggccttagttac 1170  
 C7 A.A F V R G G R P G F V S G L S Y 390

Figure 41 (continued)



C7 **seq** ctggagctggacaaccctgatggaaacaacaacgatattcttcc 1215  
 C7 **A.A** L E L D N P D G N K Q R Y S S 405  
 C7 **seq** tgggcaggatctgtgactgatcttccccaagtcataaaacaaaag 1260  
 C7 **A.A** W A G S V T D L P Q V I K Q K 420  
 C7 **seq** ctgacacctttgtatgagctggtaaaggaagtaccctgtgcttct 1305  
 C7 **A.A** L T P L Y E L V K E V P C A S 435  
 C7 **seq** gtgaaaagactgtacttgaaacgggccctggaggagtatctggat 1350  
 C7 **A.A** V K R L Y L K R A L E E Y L D 450  
 C7 **seq** gaatttgactcctgccattgccaaccttgtcaaaacggtggcatg 1395  
 C7 **A.A** E F D S C H C Q P C Q N G G M 465  
 C7 **seq** gccagtgtcgaggggaccagtgctcagtgccattgcaaaccaaac 1440  
 C7 **A.A** A S V E G T Q C Q C H C K P N 480  
 C7 **seq** acatttggtgtggcgtgtgaacaaggagtcctcgtaggggatcat 1485  
 C7 **A.A** T F G V A C E Q G V L V G D H 495  
 C7 **seq** gcaggaggcattgatggaggggtggagttgctggtcctcttggggc 1530  
 C7 **A.A** A G G I D G G W S C W S S W G 510  
 C7 **seq** ccctgtgcccagggaagaaaacaaggagccgaaaatgcaataac 1575  
 C7 **A.A** P C A Q G K K T R S R K C N N 525  
 C7 **seq** ccacccccagtggggggtgggaaatcctgtattggagaaacgtca 1620  
 C7 **A.A** P P P S G G G K S C I G E T S 540  
 C7 **seq** gaaagcaggcaatgcgaagatgaagatctggagcatcttccggttg 1665  
 C7 **A.A** E S R Q C E D E D L E H L R L 555  
 C7 **seq** cttgaaccacattgtttccctttgtctttggttccaacagaattc 1710  
 C7 **A.A** L E P H C F P L S L V P T E F 570  
 C7 **seq** tgtccatcacctcctgccttgaaagatggatttgttcaaaatgaa 1755  
 C7 **A.A** C P S P P A L K D G F V Q N E 585  
 C7 **seq** gagaccacggtttcctgttgggaaaaacatagtgtagcttgcaat 1800  
 C7 **A.A** E T T F P V G K N I V Y S C N 600  
 C7 **seq** gaaggatactctcttgttggagaccctggttggcagatgtggagaa 1845  
 C7 **A.A** E G Y S L V G D P V A R C G E 615  
 C7 **seq** gatttacagtggactgttgggaaaatgcattgtcagaaaattgcc 1890  
 C7 **A.A** D L Q W T V G K M H C Q K I A 630  
 C7 **seq** tgtgttctacctacactgatgcgtggcttacagagtcacccccag 1935  
 C7 **A.A** C V L P T L M R G L Q S H P Q 645  
 C7 **seq** aagcctttctacacagttggcgagaaggtgaccttttcctgttca 1980  
 C7 **A.A** K P F Y T V G E K V T F S C S 660  
 C7 **seq** agcggcatgtccttagaagggtccatcgacatttctctgtggatcc 2025  
 C7 **A.A** S G M S L E G P S T F L C G S 675  
 C7 **seq** agcctcaagtggagccctgagatgaaaaatgtccagtgtgtgcca 2070  
 C7 **A.A** S L K W S P E M K N V Q C V R 690  
 C7 **seq** aaagaggccccttggcaaagaaagtgcctgaatgtcagctctgg 2115  
 C7 **A.A** K E A P L A K K V P E C Q L W 705  
 C7 **seq** gagaaactgcagaattcaaatgtgtttgtaaaatgccctatgaa 2160  
 C7 **A.A** E K L Q N S K C V C K M P Y E 720  
 C7 **seq** tgcggaatcttccttagacgtgtgctcgagatgagagaagcaaa 2205  
 C7 **A.A** C G S S L D V C A R D E R S K 735  
 C7 **seq** aggatcctgcgactgacagtttgcaagatgcatgttctccaatgt 2250  
 C7 **A.A** R I L R L T V C K M H V L Q C 750  
 C7 **seq** cagggtagaaattacactctttctgttggggagacctgtactctg 2295  
 C7 **A.A** Q G R N Y T L S V G E T C T L 765  
 C7 **seq** cctggctccgctgagaaagcttgcggtgcatgtccactctgggag 2340  
 C7 **A.A** P G S A E K A C G A C P L W E 780

Figure 41 (continued)

```

C7 seq aaatgtgatgcccaaagcagcaaatgtgtctgcagagccgcctcg 2385
C7 A.A K C D A Q S S K C V C R A A S 795
C7 seq gagtgcgaggaagcggggttccgcgtctgcgtggaggtgaacggc 2430
C7 A.A E C E E A G F R V C V E V N G 810
C7 seq agggagcagacgatgaccgagtgtaggcccggcgtcctgagatgc 2475
C7 A.A R E Q T M T E C E A G V L R C 825
C7 seq ctaggactgagcatcactgtcaccagcatcaggccctgcgcgccc 2520
C7 A.A L G L S I T V T S I R P C A P 840
C7 seq gaagccccgtag 2532
C7 A.A E A P * 843

C7 seq caatcccggcgcgcacaaacctgccacgtaacagcttctcagcaca 2577
C7 seq caggcctcttctttctcctcccactcatgtccacttctcccagtc 2622
C7 seq cctgcgtaaacgcaaaccctcgttctccca 2654

```

Figure 41 (continued)

```

C8A seq c 1
C8A seq ggcacgagtggctctgtagacacctcttacttcaactccctaagta 46
C8A seq gtttttagtccttctttgtaatatagggtagtggcttctggatgag 91

C8A seq atgttggctgttgcttttttcatcttgtctttgatgacttgtcag 136
C8A A.A M L A V A F F I L S L M T C Q 15
C8A seq cctggggtaaccataccggagaaggtgaaccggagagtcagtcgg 181
C8A A.A P G V T I P E K V N R R V S R 30
C8A seq gctgttctgagctccacccccacagccgtttctcctgccagctaagc 226
C8A A.A A V L S S T P T A V S C Q L S 45
C8A seq aactgggcagagtgacagaatgcttcccatgccaggacaaaaag 271
C8A A.A N W A E W T E C F P C Q D K K 60
C8A seq tatagataccggagcctcctgcagccaaacaagtttgggggagcc 316
C8A A.A Y R Y R S L L Q P N K F G G A 75

```

Figure 42 Nucleotide and predicted protein sequence of the porcine C8A gene. Nucleotide and amino acids are enumerated on the right side. The numbering of the nucleotide sequence starts with the first base of GenBank acc. no AB008156 and ends with the last base of GenBank acc. no AB008156. The 5' flanking region contains 91 nucleotides (nt. 1-91). The start codon (ATG) is in the first bold box (nt. 92-94). The porcine boundaries between exons are in bold and underline. Positions of polymorphism present in shade blocks with white bold font. Cysteine residues are in underline. The open reading frame is terminated by a TGA stop codon in the second bold box (nt. 1859-1861). The 3'UTR (nt. 1862-2176) contains the putative polyadenylation signal attaaa (nt. 2110-2115) in bold and italic lowercase. The translation stop codon is identified by (\*)

C8A seq	atctgcagtgggAACGTCTGGGACCAAGCCAGCTGTCAcag <b>ccc</b>	361
C8A A.A	I <u>C</u> S G N V W D Q A S <u>C</u> H S P	90
C8A seq	acagagtgtctgaggcaagcacagtgtggacaggatttccagtgt	406
C8A A.A	T E <u>C</u> L R Q A Q <u>C</u> G Q D F Q <u>C</u>	105
C8A seq	aaggagaca <b>gg</b> gcgctgcctgaaacgccacctcgtgtgtaacgga	451
C8A A.A	K E T G R <u>C</u> L K R H L V <u>C</u> N G	120
C8A seq	gacagggactgcttggatggctctgacgaggatgactgtgaagat	496
C8A A.A	D R D <u>C</u> L D G S D E D D <u>C</u> E D	135
C8A seq	gtcaggatcttcgaagatgactgcagccagtagcacc <b>a</b> attccg	541
C8A A.A	V R I F E D D <u>C</u> S Q Y D P I P	150
C8A seq	ggatcggagagggcgaccttgg <b>gg</b> tacaatattctgaccagga	586
C8A A.A	G S E R A T L G Y N I L T Q E	165
C8A seq	gaaacgcagagtgtgtatgatgccaggtattatggggccagtgt	631
C8A A.A	E T Q S V Y D A R Y Y G G Q <u>C</u>	180
C8A seq	gagactgtctacaacggggagtgaggaggagcttcggtatgactct	676
C8A A.A	E T V Y N G E W R E L R Y D S	195
C8A seq	gcctgtgagcgtctctactatggagatgatgagaagtacttccgg	721
C8A A.A	A <u>C</u> E R L Y Y G D D E K Y F R	210
C8A seq	aacccttacaatttctgaagtaccacttcca <b>ag</b> ccctggcagac	766
C8A A.A	N P Y N F L K Y H F E A L A D	225
C8A seq	agtaaattttctcagagtcatac gatgatgcaaatgaccttctt	811
C8A A.A	S K F S S E S Y D D A N D L L	240
C8A seq	aaaaaagtaaaaaatgagaagtctgtgtcagctggagtgaccgtt	856
C8A A.A	K K V K N E K S V S A G V T V	255
C8A seq	ggtataggccctacaggtagcccatttacagcaaatgtgggctta	901
C8A A.A	G I G P T G S P F T A N V G L	270
C8A seq	tcagggtcacgagagtctgcgcttcttgaacaagctaagcacgat	946
C8A A.A	S G S R E S A F L N K L S T Y	285
C8A seq	aacgaga <b>aga</b> aatacagcttcatcaggattttcacaaggtacag	991
C8A A.A	N E K K Y S F I R I F T K V Q	300
C8A seq	actgctagttttatgatgaggagggacaatattatgctggatgaa	1036
C8A A.A	T A S F M M R R D N I M L D E	315
C8A seq	gttatgctgcagtcattaatggagcttccggagcagtacaattac	1081
C8A A.A	V M L Q S L M E L P E Q Y N Y	330
C8A seq	ggcatgtacgccaaagttcattgatgactacggcaccattatact	1126
C8A A.A	G M Y A K F I D D Y G T H Y I	345
C8A seq	acgtctggatc gatgggtggtgtttatgaatatatcctggtgctt	1171
C8A A.A	T S G S M G G V Y E Y I L V L	360
C8A seq	aacaaagaaaatatacacaatct <b>gg</b> tggtaccag <b>c</b> gatgatgtc	1216
C8A A.A	N K E N M T K S G V T S D D V	375
C8A seq	acgtcatgcttggagggtcttggcatcgactatgactataca	1261
C8A A.A	T S <u>C</u> F G G <b>S F G I D Y D Y T</b>	390
C8A seq	gataacttacaattacaggaagtttatcaggaaaacattgtaaa	1306
C8A A.A	<b>D N L Q I T G S L S G K H <u>C</u> K</b>	405
C8A seq	aaacttggaggtggccacagag <b>ga</b> agacgaagagagcaacatggct	1351
C8A A.A	<b>K L G G G H R E D E E S N M A</b>	420
C8A seq	gtggaagacatcatttctcgggtgaggggtggcagttctggctgg	1396
C8A A.A	V E D I I S R V R G G S S G W	435
C8A seq	ggcgggtggcttgacgcaaacggcagcatcattacataaccgtgcc	1441
C8A A.A	G G G L T Q N G S I I T Y R A	450
C8A seq	tgggggaggtcattaagataatcctgctgttattgattttgag	1486
C8A A.A	W G R S L K Y N P A V I D F E	465

Figure 42 (continued)

**C8A seq** atgaagcccatttacgagatact**g**cgccacacaaacctggggccc 1531  
**C8A A.A** M K P I Y E I L R H T N L G P 480  
**C8A seq** ctggaggccaag**c**gccagaacctgcgacgggccttgatcagtac 1576  
**C8A A.A** L E A K R Q N L R R A L D Q Y 495  
**C8A seq** ctgatggaattcaacgcctgccgctgtgggacctgcttcaacaat 1621  
**C8A A.A** L M E F N A C R C G P C F N N 510  
**C8A seq** ggcgagcccatacctcgtgggtaccagctgccggtgtcagtgcct 1666  
**C8A A.A** G E P I L V G T S C R C Q C P 525  
**C8A seq** gtgggtt**g**ccagggccttgctgtgagcaaataagtcagag**gga** 1711  
**C8A A.A** V G C Q G L A C E Q M K S E G 540  
**C8A seq** gccaaaggctgatggctcgctggagctgctggagctcctggctgct 1756  
**C8A A.A** A K A D G R W S C W S S W S A 555  
**C8A seq** tgcagattggg**c**accaggaaggagagagtgtacaaccct 1801  
**C8A A.A** C R L G T Q E R R R E C N N P 570  
**C8A seq** gcacccagaatggaggtgcctcgtgttcaggccacaaagtgcag 1846  
**C8A A.A** A P Q N G G A S C S G H K V Q 585  
**C8A seq** acccaggcatgc**tga** 1861  
**C8A A.A** T Q A C \* 589

**C8A seq** gggcctctgggcacaggctggaacccaccagcagacgtcacgc 1906  
**C8A seq** cggcaccgaccactggatacagacctcactccgctgagagaagat 1951  
**C8A seq** gccaatcaacgtggaccttttctctgtcctgccagcttccaggct 1996  
**C8A seq** gaagactgggcatgccagctatacccaactgttctgttactcac 2041  
**C8A seq** aaaactcagtcagtcacaactgggggttgagagttaactagtgtt 2086  
**C8A seq** aactgcttgcaacactttggatc**attaaag**aaaagaaaagagaaa 2131  
**C8A seq** aaaaaaaaaaaaaa 2146

Figure 42 (continued)

**C8B seq** attccctgtgcaggttcttcatggatcc 28

**C8B seq** **atg**agcctcagtcagatctcctgtcgcatagaaaaaatgaagacc 73  
**C8B A.A** M S L S Q I S C R I E K M K T 15  
**C8B seq** tctgggacgtgggcttgaggatgc**cg**gcaaggctgttcttctc 118  
**C8B A.A** S G T W A W R M P A R L F L L 30

Figure 43 Nucleotide and predicted protein sequence of the porcine C8B gene. Nucleotide and amino acids are enumerated on the right side. The numbering of the nucleotide sequence starts with the first base of GenBank acc. no AK232952 and ends with the last base of GenBank acc. no AK233291. The 5'flanking region contains 28 nucleotides (nt. 1-28). The start codon (ATG) is in the first bold box (nt. 29-31). The porcine boundaries between exons are in bold and underline. Positions of polymorphism present in shade blocks with white bold font. Cysteine residues are in underline. The open reading frame is terminated by a TGA stop codon in the second bold box (nt. 1862-1864).

C8B seq	tgtgctgcacctgggggtgtctcagtttacctggctccagagggtgag	163
C8B A.A	<u>C</u> A A L G <u>C</u> L S L P G S R G E	45
C8B seq	aggccacactctcttgagtcaaatgcagtcacgagagccttggc	208
C8B A.A	R P H S L E S N A V N E S L G	60
C8B seq	aagagtagacagaagcggagtgatgtcaccttgacgcctatt	253
C8B A.A	K S R Q T R S V D V T L T P I	75
C8B seq	gattgtgaactgtccagctggcctcctggaccacgtgtgatccc	198
C8B A.A	D <u>C</u> E L S S W S S W T T <u>C</u> D P	90
C8B seq	tgtcagaagaaaagggtacagacatgcctccttgctccggcctct	343
C8B A.A	<u>C</u> Q K K R Y R H A S L L R P S	105
C8B seq	cagttccatggggaaccgtgcaacttctctgacaaggaagttaa	388
C8B A.A	Q F H G E P <u>C</u> N F S D K E V E	120
C8B seq	gactgtggttccaacagaccatgccgaagtcaagtgcgatgtgaa	433
C8B A.A	D <u>C</u> V S N R P <u>C</u> R S Q V R <u>C</u> E	135
C8B seq	ggctttgtgtgtgctgcagacagggagatgcatcaaccgtagactt	478
C8B A.A	G F V <u>C</u> A Q T G R <u>C</u> I N R R L	150
C8B seq	ctttgcaatggtgacaatgactgtggagatcagtcggatgaggca	523
C8B A.A	L <u>C</u> N G D N D <u>C</u> G D Q S D E A	165
C8B seq	aactgtaaaaggatttataaaaagtgtcagcaggaaatggaccag	568
C8B A.A	N <u>C</u> K R I Y K K <u>C</u> Q Q E M D Q	180
C8B seq	tactgggccattggcagcctggccagtgaggataaattttgttcacg	613
C8B A.A	Y W A I G S L A S G I N L F T	195
C8B seq	aacaacttggagggcccagttctcgatcacaggtattacgccgggt	658
C8B A.A	N N L E G P V L D H R Y Y A G	210
C8B seq	gcattgctccccccattacataactcaacacgaggtttcggaaacca	703
C8B A.A	A <u>C</u> S P H Y I L N T R F R K P	225
C8B seq	tacaatgtggaaaacttcaaccacagaccccaaggaaaatatgac	748
C8B A.A	Y N V E N F N P Q T Q G K Y D	240
C8B seq	tttgactgacagaatacgaatcatactcagatttcgaacaaaat	793
C8B A.A	F A L T E Y E S Y S D F E Q N	255
C8B seq	gtcacaactgcagcaattagcaagtctagtttcagcttcggtttt	838
C8B A.A	V T T A A I S K S S F S F G F	270
C8B seq	aaaatatctggaatgtttgaatttggcatcagtagcacaagtgat	883
C8B A.A	K I S G M F E F G I S S T S D	285
C8B seq	aaaggcaaacgttttatttagcagaaccaagcgattctctcacacg	928
C8B A.A	K G K R F I S R T K R F S H T	300
C8B seq	aaaagcaccatttctgcacgcgcgctctgaccttgaggtagcgcgt	973
C8B A.A	K S T F L H A R S D L E V A R	315
C8B seq	tacaagctgaaatccagaaacctcatgctgcattacgagttcctt	1018
C8B A.A	Y K L K S R N L M L H Y E F L	330
C8B seq	cagagggtcaagcagctgcctctggagtacagctacggggagtag	1063
C8B A.A	Q R V K Q L P L E Y S Y G E Y	345
C8B seq	agagatctcttccgagattttgggaccactacatcacagaggcc	1108
C8B A.A	R D L F R D F G T H Y I T E A	360
C8B seq	gtgctcgggggtgtttatgagtacacactcatcatgaacaaagag	1153
C8B A.A	V L G G V Y E Y T L I M N K E	375
C8B seq	gccatggagagagcagattattctctaaaggatgtccatgcctgt	1198
C8B A.A	A M E R A D Y S L K D V H A <u>C</u>	390
C8B seq	gcccaacatggttttaaaattggtgtcgccatagaagagggtctac	1243
C8B A.A	A Q H G F K I G V A I E E V Y	405
C8B seq	gtcaagttgggggtgaccgggtgcacaagtgcaaggacattctcaat	1288
C8B A.A	V K L G V P V H K <u>C</u> K D I L N	420

Figure 43 (continued)

```

C8B seq gagataaaagacagaaacaagagaagcagcatggtgaatgacttg 1333
C8B A.A E I K D R N K R S S M V N D L 435
C8B seq gtggtccttgtgCGGGGaggtgcaagtgaacacatcactgcccctg 1378
C8B A.A V V L V R G G A S E H I T A L 450
C8B seq gcatataaggatctgccaacagcggacctgatgcaggaatggggga 1423
C8B A.A A Y K D L P T A D L M Q E W G 465
C8B seq gatgcagtgcagtacaacccagacatcatcaaaattaaggcagag 1468
C8B A.A D A V Q Y N P D I I K I K A E 480
C8B seq cctctctatgaactggtgacagccgcgggactttgcctattccagc 1513
C8B A.A P L Y E L V T A A D F A Y S S 495
C8B seq acagtgaagcagaacatgaagcgagccctggaggagtgcgagaag 1558
C8B A.A T V K Q N M K R A L E E F E K 510
C8B seq gaaatcagctcctgccactgtgctccctgccaggggaatggagtc 1603
C8B A.A E I S S C H C A P C Q G N G V 525
C8B seq cctgtcctgaaagaatcacgctgtgactgcatctgtcctgctgga 1648
C8B A.A P V L K E S R C D C I C P A G 540
C8B seq ttccaaggctcagcctgtgaggtcaccaatcggaaaaatgtcccc 1693
C8B A.A F Q G S A C E V T N R K N V P 555
C8B seq attgatgggaagtggaattgctggtcagactggttctccgtgttt 1738
C8B A.A I D G K W N C W S D W F S V F 570
C8B seq ctggaggacgtaaaacaagacgaaggcagtgcaacaatccacctc 1783
C8B A.A L E D V K Q D E G S A T I H L 585
C8B seq ctcaaaatgggggtagccccctgcttgggtcctgcttcagaaacac 1828
C8B A.A L K M G V A P A W V L L Q K H 600
C8B seq ttaactgttaaggaagggagagcttcttagcgggtga 1864
C8B A.A L T V K E G R A S S G * 611

C8B seq tgctactgtgggctgcacacagtgagagctctgagcccttaggaa 1909
C8B seq ccaggccagctcatctccacaccagcttccacctgggctggccc 1954
C8B seq aggggtgggaaggctgtgccattcagatttgaataaagatggta 1999
C8B seq tttgtaaaatgcacatggatttgaacaaatagcaagttaaatact 2044
C8B seq cattatagtctcctgaagggttaagcctcttagtaataattcact 2089
C8B seq ttgcttacctctccaatttttatttctacttctctctggaggggca 2134
C8B seq cactgtctcattcagtgccgtggttagttcatatttgttgaataa 2179
C8B seq ttagtacatgccagctacagtgttcagtgcatgaattcattccat 2224
C8B seq cctcatgacagtcttataagggtggatgctgtcttagccccatgaa 2269
C8B seq acagcaggggaaccagcttaggacagttaagtgacttgtccaagg 2314
C8B seq tgatgcaattagggagtgggggagcggggctgtgaacttgggcac 2359
C8B seq atggactccagagttttaaacgcttaattaggatactcccttctc 2404
C8B seq cttaataaatgggcacttgattaaaaaaaaaaaaaaaaaaaaaaaa 2449
C8B seq aaaaaaaaaaaaa 2461

```

Figure 43 (continued) Interestingly, alignment of sequences of GenBank accession DQ333201 (in normal) and AK233291 (in italic) in Genbank allowed merging 474 bp longer (nt. 1988-2461) in 3' flanking region of the porcine C8B cDNA sequence, which contains putative polyadenylation signal **ataaaa** in bold and italic. The translation stop codon is identified by (\*)

```

C8G seq                                     atcgtc      6
C8G seq ctgctctctgtctccaccggctctgtcgccgttgccgccgtgggcc 51

C8G seq atgctagtccccagggctgcacccctcttgacgctgctcctggcc 96
C8G a.a M L V P R A A P L L T L L L A 15
C8G seq acgggctccttgggcccagagggctcagagacccccctcgacgccc 141
C8G a.a T G S L G Q R A Q R P P R R P 30
C8G seq tcccccatcagcaccatccagcccaggccagctttgatgcccg 186
C8G a.a S P I S T I Q P K A S F D A Q 54
C8G seq cagtttgcagggacgtggctcctggtggccgtggcttcctcctgc 231
C8G a.a Q F A G T W L L V A V A S S C 60
C8G seq cgcttcctgcaagagcagggccaccgggctgaggccacttcactg 276
C8G a.a R F L Q E Q G H R A E A T S L 75
C8G seq cacgtggctcctcagggctgcagccatggccgtcagcaccttccga 321
C8G a.a H V A P Q G A A M A V S T F R 90
C8G seq aagctggatgggatctgctggcaggtgcggcagctcttcagagac 366
C8G a.a K L D G I C W Q V R Q L F R D 105
C8G seq acggggctcccaggtcgcttcctgctccagggcccgaggcgcccga 411
C8G a.a T G L P G R F L L Q A R G A R 120
C8G seq ggcgcgggtggatgtggtcgttggggagacggactaccggagcttt 456
C8G a.a G A V D V V V G E T D Y R S F 135
C8G seq gccatcctgtacctggagagggcccggcagctgtcggatgaagctg 501
C8G a.a A I L Y L E R A R Q L S V K L 150
C8G seq tacgcccgctcgctccccgtgagcgaatcagccctgagtgcttt 546
C8G a.a Y A R S L P V S E S A L S V F 165
C8G seq gagcagcgggtccagggggccaacctgaccgaggaccacatcctg 591
C8G a.a E Q R V Q G A N L T E D H I L 180
C8G seq ttcttcccccaagtacggcttctgcgacgctgcagaccagttccac 636
C8G a.a F F P K Y G F C D A A D Q F H 195
C8G seq gtcttgacgaagcaaagcagtga 660
C8G a.a V L D E A K Q * 202

```

Figure 44 Nucleotide and predicted protein sequence of the porcine C8G gene. Nucleotide and amino acids are enumerated on the right side. The numbering of the nucleotide sequence starts with the first base of GenBank acc. no BP139629.1 and ends with the last base of GenBank acc. no BP443148.1. The 5' flanking region contains 51 nucleotides (nt. 1-51). The start codon (ATG) is in the first bold box (nt. 52-54). The porcine boundaries between exons are in bold and underline. Positions of polymorphism present in shade blocks with white bold font. Cysteine residues are in underline. The open reading frame is terminated by a TGA stop codon in the second bold box (nt. 568-660). The 3' UTR (nt. 661-840) contains the putative polyadenylation signal attaaa (nt. 792-797) in bold and italic. The translation stop codon is identified by (\*)

**C8G seq** ggccggggccttgccgcccagagagagccacaagctggaagtgtg 705  
**C8G seq** ggcacccgccactcctggccgagcggggcctctcccattctgcc 750  
**C8G seq** tggatgctgccacccccccacgagctcggggcgccacctt**atta** 795  
**C8G seq** **aa**cactgtcgggtctcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 840

Figure 44 (continued)

**C9 seq** aaaaacaattgtggtgccccagtccttggcagataaatgtcttc 44  
**C9 seq** caggcactgactgccattccgctcggagctgcctcactagagcag 89  
  
**C9 seq** **atg**cggggtccgcccggaggttcacctttgcagtatgcattttagaa 134  
**C9 A.A** M R V R R R F T F A V C I L E 15  
**C9 seq** ataagcattctgagcgcggggcccacacc**agt**taccacccgaag 179  
**C9 A.A** I S I L S A G P T P S Y H P K 30  
**C9 seq** cccgtggactgcaatggcacaccatcgcccattgactgcaggatg 224  
**C9 A.A** P V D C N G T P S P I D C R M 45  
**C9 seq** agtccctggggcgaatggtcaagatgtgacccttgccctcaaaaa 269  
**C9 A.A** S P W G E W S R C D P C L K Q 60  
**C9 seq** **atgt**ttcgctcaaggagcattgagaccctttggacaatttaacggg 314  
**C9 A.A** M F R S R S I E T F G Q F N G 75  
**C9 seq** caaaagtgcgtggacgccgtgggggacaggcggca**gt**gcgtgccc 359  
**C9 A.A** Q K C V D A V G D R R Q C V P 90  
**C9 seq** acggagccttgcaagaccttgaggaggactgtggaagtgacttt 404  
**C9 A.A** T E P C E D L E E D C G S D F 105  
**C9 seq** ca**gt**gtggtac**agg**cagatgcataaagaggcgccttctgtgtaat 449  
**C9 A.A** Q C G T G R C I K R R L L C N 120  
**C9 seq** ggtgacaatgactgcccagacttttccagatgaggatgactgtgac 494  
**C9 A.A** G D N D C G D F S D E D D C D 135  
**C9 seq** agtgatccccgccccctgcccgtgagagagtgggtggaagagtct 539  
**C9 A.A** S D P R P P C R E R V V E E S 150  
**C9 seq** gagctggcacggacagcaggctac**ggg**atcaacatcttaggaatg 584  
**C9 A.A** E L A R T A G Y G I N I L G M 165

Figure 45 Nucleotide and predicted protein sequence of the porcine C9 gene. Nucleotide and amino acids are enumerated on the right side. The numbering of the nucleotide sequence starts with the first base of GenBank acc. no BP139003.1 and ends with the last base of GenBank acc. no CF363669. The 5' flanking region contains 89 nucleotides (nt. 1-89). The start codon (ATG) is in the first bold box (nt. 90-92). The porcine boundaries between exons are in bold and underline. Positions of polymorphism present in shade blocks with white bold font. Cysteine residues are in underline. The open reading frame is terminated by a TGA stop codon in the second bold box (nt. 1719-1721). The 3' UTR (nt. 1722-2536) contains the putative polyadenylation signal attaaa (nt. 2505-2510) in bold and italic. The translation stop codon is identified by (\*)



C9 seq gacccccctcaccacccccatttgacaacgagtactacaacggactc 629  
 C9 A.A D P L T T P F D N E Y Y N G L 180  
 C9 seq tgtgaccgagttcgggatggaaacaccttgacctactaccgcaa 674  
 C9 A.A C D R V R D G N T L T Y Y R K 195  
 C9 seq ccctggaacgtggctgccttgatctatgaaaaccaaagttgacaaa 719  
 C9 A.A P W N V A A L I Y E T K V D K 210  
 C9 seq aatttcagaactgaatatcatgaaagacagattcaagtactcaaa 764  
 C9 A.A N F R T E Y H E R Q I Q V L K 225  
 C9 seq actatcatcgaagagaagaaatcaaattttaatgcagatttaact 809  
 C9 A.A T I I E E K K S N F N A D L T 240  
 C9 seq ataaaattcacacctaccgaagcaattgaacagctgaaaagtaaa 854  
 C9 A.A I K F T P T E A I E Q L K S K 255  
 C9 seq aatgtagaacttgccaacgaagaaaactcaaactcctatgaataac 899  
 C9 A.A N V E L A N E E N S N P M N N 270  
 C9 seq aaggctcacttccgatttacatattccaaaacggaaacctacaaa 944  
 C9 A.A K A H F R F T Y S K T E T Y K 285  
 C9 seq ctgttgttgtcatattcttctaagaaggaaaaaatattcctgcat 989  
 C9 A.A L L L S Y S S K K E K I F L H 300  
 C9 seq gtcaaaggagtaattcacctgggaagatttgtgatgagaaagcgg 1034  
 C9 A.A V K G V I H L G R F V M R K R 315  
 C9 seq gatgtcatgctgacaaaaacttttcttgatgatgtaaagtatctg 1079  
 C9 A.A D V M L T K T F L D D V K Y L 330  
 C9 seq ccaagtacctatgaaaagggagaatattttgcatttttggaaacc 1124  
 C9 A.A P S T Y E K G E Y F A F L E T 345  
 C9 seq tatggaaccactacagtagctctgggtctctgggaggactctat 1169  
 C9 A.A Y G T H Y S S S G S L G G L Y 360  
 C9 seq gaactaatatagtcttgataaagctaccatgactgagaaaggt 1214  
 C9 A.A E L I Y V L D K A T M T E K G 375  
 C9 seq attgaactacgggatgtacacagatgccttgggtttaatctggat 1259  
 C9 A.A I E L R D V H R C L G F N L D 390  
 C9 seq ttatctctgaattttggagttgaaatcaaaggaaagattgattca 1304  
 C9 A.A L S L N F G V E I K G K I D S 405  
 C9 seq gaaaattgcttaaagaggggtgacggtaaaactgaaacatcatg 1349  
 C9 A.A E N C L K R G D G K T E N I M 420  
 C9 seq aatgacgacttcatagatgatgttatttcattcataagaggagga 1394  
 C9 A.A N D D F I D D V I S F I R G G 435  
 C9 seq accagaaaatacgcaactgaactgaaagagaagcttctcaaagga 1439  
 C9 A.A T R K Y A T E L K E K L L K G 450  
 C9 seq gccaaaatgattaacgtgactgactttgtaaactgggcctcttcc 1484  
 C9 A.A A K M I N V T D F V N W A S S 465  
 C9 seq ttaaagatgctccggtgctcataaatcaaaaactgctctccgata 1529  
 C9 A.A L N D A P V L I N Q K L S P I 480  
 C9 seq tatgatctgattccagtgaaactaaatgatgcacatctaaagaga 1574  
 C9 A.A Y D L I P V K L N D A H L K R 495  
 C9 seq caaaatttgaaagagccattgaagactacatcaatgaattcaat 1619  
 C9 A.A Q N L E R A I E D Y I N E F N 510  
 C9 seq gtaagaaaaatgccaactgtgccagaatggagggacagtgattct 1664  
 C9 A.A V R K M P T V P E W R D S D S 525  
 C9 seq gctggatggacagtgtttgtgttctgcccacaacaaatttcaggg 1709  
 C9 A.A A G W T V F V F L P K Q I S G 540  
 C9 seq aattgcttgtga 1721  
 C9 A.A N C L \* 543

Figure 45 (continued)

```

C9 seq aatccccaaaatagaaaatctctgaaggattgccagccccccacca 1766
C9 seq aaaaaaaaaacaaaacaaaacaaaacactgagctggttggttccctg 1811
C9 seq ggatcacatgggaagaaaaacaccaggactccaggactgtctaga 1856
C9 seq ggtagtccttgctgccgaatgaaaagcaacatgcttcatgaaaat 1901
C9 seq ccccccaacctctgaagtctcctctctcagatccacagagcttct 1946
C9 seq tctcttccctgaaactcctatgggttccagttttttattactaatg 1991
C9 seq aggagagtcagcagtcaaataatgccaagactgctttctcccacag 2036
C9 seq gcaatgccaatctcttgctaacaaaacaaaattaaattaaaaaga 2081
C9 seq aagtgttggttaaaaaagaccttaaagtcattgccaaaggctcttg 2126
C9 seq tatgacgaatacgttctgccaggtccatcaccacaagtctgtacc 2171
C9 seq atgcaatttcactcctgtttacatacttggtcatgttcgggtcat 2216
C9 seq ttgtagatacgtgcatgccttcgttctgaaagtacaaacaccac 2261
C9 seq cacaagcacagcactctgctttttaagtcttagctctctggttta 2306
C9 seq actctatgtgtccaaggaggaaaatgtattgtatttgaacccaa 2351
C9 seq aactattagattcagtaatgtgaaattacctttattcagtcatag 2396
C9 seq tttaccaaaaaatgaaaacaaaacaaaacaaaacccagcaatt 2441
C9 seq ccatatggtccccttaatacctagtgttacgagttgtataaagtc 2486
C9 seq taccatataagaagttaattaaaggagctggctgtacaattgca 2531
C9 seq aaaaa 2536

```

Figure 45 (continued)

```

S.scrofa      MDKH-SVLYFILLSGLIDKSQACFCCDHYPWTQWSSCSKTCNSGTQTRQRQITINQYYLDN 59
B.taurus     MARH-SVMYFILLSALIDKSQACFCCDHYPWSWSSCSKTCNSGTQTRQRRIVTDKYYFEN 59
H.sapiens    MARR-SVLYFILLNALINKGQACFCCDHYAWTQWTSCSKTCNSGTQSRHRQIVVDKYYQEN 59
P.troglodytes MARR-SVLYFILLNALINKGQACFCCDHYAWTQWTSCSKTCNSGTQSRHRQIVVDKYYQEN 59
P.pygmaeus   MARC-SVLYFILLSALINKGQACFCCDHYPWTQWTSCSKTCNSGTQSRHRQIVVDKYYQEN 59
C.familiaris MAGQ-SILYFILLSALIDKQPCFCCDHYPWSWSSCSKTCNFGTQSRQRKIVVDQYYLDN 59
R.norvegicus MTRH-LTLCFFILLIILIDKSEACFCCDHYPWTHWSSCSKSCNSGTQSRQRQIVVNDYYRDN 59
M.domestica  MDQN-ILLYSIVLSALISGSQACYCERYPWTSWSSSSSCNYGSQIRERHIVRDKYYEEN 59
G.gallus     MEKT-VLIQLILLSLVVGSQGGCYCEHYPWGSWSSCSKTCNYGTQTRHRQIKMDEYYNQN 59
X.tropicalis MGSR-PCIIFLLIVSLAGGALSCFCDHYPWTSWSSCSKTCDHGTQSRRTISYDDYYWKH 59
O.mykiss     MDRTNSLLVLLHVLGSVTVSLGCFCDHYAWSSWSTCTRTCNYGTQERRRSIRYDDYYWKN 60
*           :   :   :           *:::*.* *::*: :*: * * * * * :.***. :

```

Figure 46 Alignment of the C6 proteins among species according to the GenBank accession no. as given in table 5 using the ClustalW2 tool. The amino acids are enumerated above each line on the right side. The identical, similar and missing residues are indicated with asterisks (\*), points (.), and dashes (-), respectively. Signal peptides in human C6 protein (DiScipio et al. 1989) and in pig C6 protein are in bold and italic. The amino termini of C6 mature protein starts just after the signal peptide. The cysteine residues are highlighted. The cysteine-rich functional protein domains like TSP1, LDLa, MACPF, EGF, CCP, and FIMAC are in continuous underline. The glycosylation sites in human (Chakravarti et al. 1989, DiScipio et al. 1989) and in pig using the NetNGlyc 1.0 Server tool present in shade blocks with white bold font. The human C6 cleavage site into two fragments C6a (a.a 22-631) and C6b (a.a 632-934) by trypsin is in bold and double underline.

	TSP1 domain (26-79)	TSP1 domain (84-134)	
<i>S. scrofa</i>	FCDLCTKQETRECNWQTCPCINCLLGDYGPWSDCDPCVEKQFKVRSILRPNQFGGQPCTE	FCDLCTKQESRECNWQTCPCINCLLGDYGPWSDCDPCVQKRFKVRSILRPSQFGGQPCTE	119
<i>B. taurus</i>	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCIEKQSKVRSVLRPSQFGGQPCTE	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQSKVRSVLRPSQFGGQPCTE	119
<i>H. sapiens</i>	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCIEKQSKVRSVLRPSQFGGQPCTE	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQSKVRSVLRPSQFGGQPCTE	119
<i>P. troglodytes</i>	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQSKVRSVLRPSQFGGQPCTE	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQSKVRSVLRPSQFGGQPCTE	119
<i>P. pygmaeus</i>	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQSKVRSVLRPSQFGGQPCTE	FCEQICSKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQSKVRSVLRPSQFGGQPCTE	119
<i>C. familiaris</i>	FCDQLCTKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQFKVRSILRPSQFGGQPCTE	FCDQLCTKQETRECNWQRCPCINCLLGDYGPWSDCDPCVEKQFKVRSILRPSQFGGQPCTE	119
<i>R. norvegicus</i>	SCDQLCTKQETROCNVETCPINCLVLDYGTWSDCDPCIRKQVKVRSVLRPSQFGGQPCTE	SCDQLCTKQETROCNVETCPINCLVLDYGTWSDCDPCIRKQVKVRSVLRPSQFGGQPCTE	119
<i>M. domestica</i>	YCDQLCTKRESRACNQMQPCINCLMGMDFEPWSECDPCVKKQFRMRSILRPSQFGGQPCTE	YCDQLCTKRESRACNQMQPCINCLMGMDFEPWSECDPCVKKQFRMRSILRPSQFGGQPCTE	119
<i>G. gallus</i>	FCDLCTKQESRACNQQTCPINCLQGLDFGPWSECDPCIEKQFRVRKLLRPSQFGGQACTE	FCDLCTKQESRACNQQTCPINCLQGLDFGPWSECDPCIEKQFRVRKLLRPSQFGGQACTE	119
<i>X. tropicalis</i>	NCGGLCTMRETRSCNEQSCPINCLKLDYGFWSWSDCDPCLKKQFRTRSLRPSQFGGEGCTE	NCGGLCTMRETRSCNEQSCPINCLKLDYGFWSWSDCDPCLKKQFRTRSLRPSQFGGEGCTE	119
<i>O. mykiss</i>	SCEQLCQKHESRACNVQACFIHCQLTDWANWSGCSPCAQQLRTRSLRPSQFGGVECDTA	SCEQLCQKHESRACNVQACFIHCQLTDWANWSGCSPCAQQLRTRSLRPSQFGGVECDTA	120
	* : * * : * * * : * * : * * * : * * * : * * * : * * * : * * * : *		
	LDLa domain (139-175)		
<i>S. scrofa</i>	PLMTFRP CIPSKLCKIEEVDCKNFKRCDGSRGCIASKLENGENDCGDNSDERNCGRTKAV	PLMTFQP CIPSKLCKIEEIDCKNFKRCDGSRGCIASKLENGENDCGDNSDERNCGRKKTAV	179
<i>B. taurus</i>	PLMTFRP CIPSKLCKIEEIDCKNFKRCDGSRGCIASKLENGENDCGDNSDERNCGRKKTAV	PLMTFQP CIPSKLCKIEEIDCKNFKRCDGSRGCIASKLENGENDCGDNSDERNCGRKKTAV	179
<i>H. sapiens</i>	PLVAFQP CIPSKLCKIEEADCKNFKRCDGSRGCIARKLENGENDCGDNSDERDCGRTKAV	PLVAFQP CIPSKLCKIEEADCKNFKRCDGSRGCIARKLENGENDCGDNSDERDCGRTKAV	179
<i>P. troglodytes</i>	PLVAFQP CIPSKLCKIEEADCKNFKRCDGSRGCIARKLENGENDCGDNSDERDCGRTKAV	PLVAFQP CIPSKLCKIEEADCKNFKRCDGSRGCIARKLENGENDCGDNSDERDCGRTKAV	179
<i>P. pygmaeus</i>	PLVAFQP CIPSKLCKIEEADCKNFKRCDGSRGCIARKLENGENDCGDNSDERDCGRTKAV	PLVAFQP CIPSKLCKIEEADCKNFKRCDGSRGCIARKLENGENDCGDNSDERDCGRTKAV	179
<i>C. familiaris</i>	PLVTFQP CIPSKLCKIEEIDCKNFKRCDGSRGCIANKLENGENDCGDNSDERNCGRKAV	PLVTFQP CIPSKLCKIEEIDCKNFKRCDGSRGCIANKLENGENDCGDNSDERNCGRKAV	179
<i>R. norvegicus</i>	PLVTFQP CVPSELCKIEETDCKNFKFLCDGSRGCI PSKLLKNGENDCGDNSDERNCGRTKPV	PLVTFQP CVPSELCKIEETDCKNFKFLCDGSRGCI PSKLLKNGENDCGDNSDERNCGRTKPV	179
<i>M. domestica</i>	QLVTSQKCYSTKLCNIEDIDCKNFKFKCDGSRGCIAMNLLCNEENDCGDSSDEKNCDRKKKV	QLVTSQKCYSTKLCNIEDIDCKNFKFKCDGSRGCIAMNLLCNEENDCGDSSDEKNCDRKKKV	179
<i>G. gallus</i>	TLVQSRP CFPAPKLCNIVDCKNFKQCENGRCIAKKLENGDNDCGDNSDERNCGRKVV	TLVQSRP CFPAPKLCNIVDCKNFKQCENGRCIAKKLENGDNDCGDNSDERNCGRKVV	179
<i>X. tropicalis</i>	ALVESRICVPAKLCIEQLDCSGKFCVSGRCPFNLCNGDNDCGDNSDERNCGRKREPQ	ALVESRICVPAKLCIEQLDCSGKFCVSGRCPFNLCNGDNDCGDNSDERNCGRKREPQ	179
<i>O. mykiss</i>	VLTEDRACYPSTECKMETLNCK-EFQCGNGRCISSKLTCKNQDNDCGDNSDEKNDEFKIV	VLTEDRACYPSTECKMETLNCK-EFQCGNGRCISSKLTCKNQDNDCGDNSDEKNDEFKIV	179
	* : * . : * : * : * : * : * : * : * : * : * : * : * : * : *		
<i>S. scrofa</i>	CS--RKYNP IPSVQLMGAGFHILAGEPRGEVLGNSFTGGICKTVKSS---KASNPFRVPA	CS--RSHNPIPQVQLMGMGFHFHLAGEPRGEVLGNSFTGGVCRVTVKSS---RASNPYRVP	234
<i>B. taurus</i>	CS--RSHNPIPQVQLMGMGFHFHLAGEPRGEVLGNSFTGGVCRVTVKSS---RASNPYRVP	CS--RSHNPIPQVQLMGMGFHFHLAGEPRGEVLGNSFTGGVCRVTVKSS---RASNPYRVP	234
<i>H. sapiens</i>	CT--RKYNP IPSVQLMGNGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	CT--RKYNP IPSVQLMGNGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	234
<i>P. troglodytes</i>	CT--RKYNP IPSVQLMGNGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	CT--RKYNP IPSVQLMGNGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	234
<i>P. pygmaeus</i>	CT--RKYDP IPSVQLMGSGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	CT--RKYDP IPSVQLMGSGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	234
<i>C. familiaris</i>	CPGKRQWNP IPSVQLMGAGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---KASNPYRVP	CPGKRQWNP IPSVQLMGAGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---KASNPYRVP	236
<i>R. norvegicus</i>	CS--RRTYTP IPSVQLMGAGFHFHLAGEPRGDVDPDNSFTGGICKSVRSS---RTSNPHRVP	CS--RRTYTP IPSVQLMGAGFHFHLAGEPRGDVDPDNSFTGGICKSVRSS---RTSNPHRVP	234
<i>M. domestica</i>	CN--RVYQP IPSVQLMGNGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	CN--RVYQP IPSVQLMGNGFHFHLAGEPRGEVLGNSFTGGICKTVKSS---RTSNPYRVP	234
<i>G. gallus</i>	CS--RKFESIPGVHLIGSGFHILSGESRGEVLGNSFNNGECCRTVRRN---ETRKSYPVP	CS--RKFESIPGVHLIGSGFHILSGESRGEVLGNSFNNGECCRTVRRN---ETRKSYPVP	234
<i>X. tropicalis</i>	RS----FEP IPIQLMGNGFNYSGESRGEVLGNSFTGGICKTVKSS---KASNPYRVP	RS----FEP IPIQLMGNGFNYSGESRGEVLGNSFTGGICKTVKSS---KASNPYRVP	235
<i>O. mykiss</i>	CP--VEKRVPVPGADLIGNGFDAMAEAMRGAVALDNMFMGDTCNLNRSRGS-SYRLYYRVP	CP--VEKRVPVPGADLIGNGFDAMAEAMRGAVALDNMFMGDTCNLNRSRGS-SYRLYYRVP	236
	: * . . * : * * : * * : * * : * * : * * : * * : * * : * * : *		
<i>S. scrofa</i>	NLENVNFEVQTKEDDFQT--DFYEDLIPLENSKDQOATGFGQE-KSSFHVP I FYSSKKSQ	NLENVNFEVQTKEDDLEA--DFYDDLIPLEDNKQDQALGSGLA-TSSFVRV I FYSSKRSQ	291
<i>B. taurus</i>	NLENVNFEVQTKEDDLEA--DFYDDLIPLEDNKQDQALGSGLA-TSSFVRV I FYSSKRSQ	NLENVNFEVQTKEDDLEA--DFYDDLIPLEDNKQDQALGSGLA-TSSFVRV I FYSSKRSQ	291
<i>H. sapiens</i>	NLENVGFEVQTAEDDLKT--DFYKDLTSLGHNENQOQGSFSSQG-GSSFV I FYSSKRSE	NLENVGFEVQTAEDDLKT--DFYKDLTSLGHNENQOQGSFSSQG-GSSFV I FYSSKRSE	291
<i>P. troglodytes</i>	NLENVGFEVQTAEDDLKT--DFYKDLTSLGHNENQOQGSFSSQG-GSSFV I FYSSKRSE	NLENVGFEVQTAEDDLKT--DFYKDLTSLGHNENQOQGSFSSQG-GSSFV I FYSSKRSE	291
<i>P. pygmaeus</i>	NLENVGFEVQTAEDDLKT--DFYKDLTSLGHNENQOQGSFSSQG-GSSFV I FYSSKRSE	NLENVGFEVQTAEDDLKT--DFYKDLTSLGHNENQOQGSFSSQG-GSSFV I FYSSKRSE	291
<i>C. familiaris</i>	NLENVNFEVQTKEDDLET--DFYKDLISLTKNENQOGLFGGEE-QSSLY I P I FYSSKKTQ	NLENVNFEVQTKEDDLET--DFYKDLISLTKNENQOGLFGGEE-QSSLY I P I FYSSKKTQ	293
<i>R. norvegicus</i>	NLENVNFEVQTIEDDLKT--DFYKDLATIGKNKNEDRSLSGEK-KDSFV I P I FYSSKKE	NLENVNFEVQTIEDDLKT--DFYKDLATIGKNKNEDRSLSGEK-KDSFV I P I FYSSKKE	291
<i>M. domestica</i>	NIESINFVNEEDDLQT--DFYSNLIPLGNDHVLHKSDNLQG-NSHSG I PLLWSTKRKE	NIESINFVNEEDDLQT--DFYSNLIPLGNDHVLHKSDNLQG-NSHSG I PLLWSTKRKE	291
<i>G. gallus</i>	NLEAVSFQVIDEEDDVKS--DFYRDLTPLSDGDVSGSSSSSHS-QRRSG I PGLF SKKRV	NLEAVSFQVIDEEDDVKS--DFYRDLTPLSDGDVSGSSSSSHS-QRRSG I PGLF SKKRV	291
<i>X. tropicalis</i>	NLDTFFKELRNEEDDVP--TFYNSLIDFDKENSRTGSSQSSR-RGSAG I PLLFHRKSN	NLDTFFKELRNEEDDVP--TFYNSLIDFDKENSRTGSSQSSR-RGSAG I PLLFHRKSN	292
<i>O. mykiss</i>	NVESFEIKVEIPDDFKQEPQPVYSETVNLASPPVSTRSDSSSGDSSSIWV I FVFGSRRH	NVESFEIKVEIPDDFKQEPQPVYSETVNLASPPVSTRSDSSSGDSSSIWV I FVFGSRRH	296
	* : . : * : * : * : * : * : * : * : * : * : * : * : * : *		
	MACPF domain (312-516)		
<i>S. scrofa</i>	TSSHNSAFKQAIQASHKKDSSFIRIHKV I KVLNFTMKTKDLQLSDVFLKALNHLPLEYNA	TSSHNSAFKQAIQASHKKDSSFIRIHKV I KVLNFTMKTKDLQLSDVFLKALNHLPLEYNS	351
<i>B. taurus</i>	TSSHNSAFKQAIQASHKKDSSFIRIHKV I KVLNFTMKTKDLQLSDVFLKALNHLPLEYNS	TSSHNSAFKQAIQASHKKDSSFIRIHKV I KVLNFTMKTKDLQLSDVFLKALNHLPLEYNS	351
<i>H. sapiens</i>	NINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDVFLKALNHLPLEYNS	NINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDVFLKALNHLPLEYNS	351
<i>P. troglodytes</i>	NINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDVFLKALNHLPLEYNS	NINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDVFLKALNHLPLEYNS	351
<i>P. pygmaeus</i>	NINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDI FLKALNHLPLEYNS	NINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDI FLKALNHLPLEYNS	351
<i>C. familiaris</i>	STTHNSAFKQAIQASHKKDSSFIRIHKV I EVLNFTMKTKDLQLSDVFLKALNHLPLEYNF	STTHNSAFKQAIQASHKKDSSFIRIHKV I EVLNFTMKTKDLQLSDVFLKALNHLPLEYNF	353
<i>R. norvegicus</i>	NFQRNSGFKNAI EASHKKDSSVFR I HKV I KVLNFTMKTTDLQLSDVFLKAL I H L P L E Y N F	NFQRNSGFKNAI EASHKKDSSVFR I HKV I KVLNFTMKTTDLQLSDVFLKAL I H L P L E Y N F	351
<i>M. domestica</i>	RITQQYSFKKAIQASHKKDSSFIRVHKV I SVLNFTMKPTLQLSDVFLKALNHLPLQYNY	RITQQYSFKKAIQASHKKDSSFIRVHKV I SVLNFTMKPTLQLSDVFLKALNHLPLQYNY	351
<i>G. gallus</i>	QITSSSSFKKAI EASHEKNSNFIRIHKV I SVANFTMKESNLQLSDVFLKALNHLPLEYNY	QITSSSSFKKAI EASHEKNSNFIRIHKV I SVANFTMKESNLQLSDVFLKALNHLPLEYNY	351
<i>X. tropicalis</i>	RITSSSSFREAAQASRQSSKFI R I N K V I S V S D F T M K K N L W L S D V F L K A L N H L P L E Y N Y	RITSSSSFREAAQASRQSSKFI R I N K V I S V S D F T M K K N L W L S D V F L K A L N H L P L E Y N Y	352
<i>O. mykiss</i>	HQSN T D T F K Q A I S A S K K T D S Q F F R V H Q V L P V S R F R V K D S D L Y L T E P F L T S L P L E Y N Y	HQSN T D T F K Q A I S A S K K T D S Q F F R V H Q V L P V S R F R V K D S D L Y L T E P F L T S L P L E Y N Y	356
	* : * * : * * : * * : * * : * * : * * : * * : * * : * * : *		

Figure 46 (continued)

<i>S. scrofa</i>	ALYSRIFDDFGTHYFTSGSLGGVYDILLYQFSNEELKNSGLTQEEAKNCIRIETKKRYFIV	411
<i>B. taurus</i>	ALYSRIFDDFGTHYFTSGSLGGVYDILLYQFSKEELKNSGLTQEEAKNCIRIETKKRFLV	411
<i>H. sapiens</i>	ALYSRIFDDFGTHYFTSGSLGGVYDILLYQFSSEELKNSGLTEEEAKHCVR IETKKRVLFA	411
<i>P. troglodytes</i>	ALYSRIFDDFGTHYFTSGSLGGVYDILLYQFSSEELKNSGLTEEEAKHCVR IETKKRVLV	411
<i>P. pygmaeus</i>	ALYSRIFDDFGTHYFTSGSLGGVYDILLYQFSSEELKNSGLTEEEAKHCVR IETKKRVLFA	411
<i>C. familiaris</i>	ALYSRIFDDFGTHYFTSGSLGGVYDILLYQYSAEELRTSGLTEKEIQNCVSTETKKS VLF	413
<i>R. norvegicus</i>	ALYSRIFDDFGTHYFTSGSLGGVYDILLYQFSRQELQNSGLTEEEETRNCVRYETK RFLV	411
<i>M. domestica</i>	ALYSRIFDDFGTHYFTSGSMGGTYDILLYQYSREELKNSGLTEEEESQHCVRIETV K RRLV	411
<i>G. gallus</i>	ALYSRIFDDFGTHYFTSGMGGSYDILLYQYSSSEELKNSGLSVDESMECIRTETTRRV FFR	411
<i>X. tropicalis</i>	PLYSRIFDDFGTHYFTSGMGGSYDILLYQYSSSENKSSGLTNQESLVCVRTEITYRV FFR	412
<i>O. mykiss</i>	ALYREIFKHFGTHYFASGTLGGHYDILLYQYSRLELKNSGLTEEHKTKGCLKSESSM FII Y	416
	.* . : * . . . * .	
<i>S. scrofa</i>	TKTKVEHRCCTTNRMSEKYEYEGSFLQSEKISLVKGGRRSEYAAALAWEKGSSGPEKTYSD	471
<i>B. taurus</i>	KKTKVEHRCCTTNKLSKYEYEGSFLQSEKISLVQGGRRSAYAAALAWEKGSPVPEERVFS D	471
<i>H. sapiens</i>	KKTKVEHRCCTTNKLSKHEGSFIQGAEKISLIRGGRRSEYGAALAWEKGSSGLEEKTFSE	471
<i>P. troglodytes</i>	KKTKVEHRCCTTNKLSKHEGSFIQGAEKISLIRGGRRSEYGAALAWEKGSSGLEEKTFSE	471
<i>P. pygmaeus</i>	KKTKVEHRCCTTNKLSKHEGSFIQGAEKISLIRGGRRSEYGAALAWEKGSSGLEEKTFSE	471
<i>C. familiaris</i>	SKKEVEHRCCTTNKMSKYEYEGSFLQGAEKISLQGGRRSAYAAALAWEKGSSVPTKEKFSE	473
<i>R. norvegicus</i>	TKTYKEDRCCCTTNRLSEKYEYEGSFLQSEKISLVQGGRRSQAAAALAWEKGSSGPEANV FSE	471
<i>M. domestica</i>	KKKVENRCCCTTNKMLDRYEGSFLQGAEKISLVLRGGRRSEHAAALAWEKGSSPETVFTTE	471
<i>G. gallus</i>	KKKVVSTECITNKMTVKHDSGILESAAERSVSLVKGRRSEYAAALAWEKGGAFPGNTIFTN	471
<i>X. tropicalis</i>	KKRKVRESCCTRNKMSERYEGSFLQGAEKISLQGGRRSAYAAALAWEKGGTLPENTVFEE	472
<i>O. mykiss</i>	SQSSNVVRCSDNTMTQKHEGSFVQSSEKSFVSMVRGGRTGAAAALAWEKGAAPDSTTYKN	476
	. : * * : . : . * .	
<i>S. scrofa</i>	WLESVKENPAVIDFELAPIIDLVRNIPCAVTRRNLRRAFREYAAKFDPCCCAPCPNNGR	531
<i>B. taurus</i>	WLESVKENPSVIDFALAPIIDLVRNIPCAVTRRNLRRAFREYAAKFDPCCCAPCPNSGR	531
<i>H. sapiens</i>	WLESVKENPAVIDFELAPIVDLVRNIPCAVTKRNLKALQEYAAKFDPCCCAPCPNNGR	531
<i>P. troglodytes</i>	WLESVKENPAVIDFELAPIVDLVRNIPCAVTKRNLKALQEYAAKFDPCCCAPCPNNGR	531
<i>P. pygmaeus</i>	WLESVKENPAVIDFELAPIVDLVRNIPCAVTKRNLKALQEYAAKFDPCCCAPCPNNGR	531
<i>C. familiaris</i>	WLESVKENPVVIDFELAPIIDLVRNIPCAVTRRNLRRAFREYAAKFDPCCCAPCPNNGR	533
<i>R. norvegicus</i>	WLESVKENPAVVDYELAPIIDLVRNIPCAVTKRNLKALQEYAAKFDPCCCAPCPNNGR	531
<i>M. domestica</i>	WLQSVKENPAVVDYELAPIIDLVRNIPCAVTKRNLKALQEYAAKFDPCCCAHCPNNGK	531
<i>G. gallus</i>	WLESTKDNPVVIDFVSSIVDLVKNIPCAVTRRNLRRALREYAGRFDPCCCAPCPNNGR	531
<i>X. tropicalis</i>	WVKSTIDNPVVVDYELAPIIDLITGIPCAVTKRNLKALQKAFKYLETFDPCICAPCPNNGR	532
<i>O. mykiss</i>	WAKSLIDNPVVVDYELAPIIDLVRNIPCAVTKRNLKALQEYAAKFDPCCCAPCPNNGR	536
	* : * .	
	<b>EGF domain (521-570)</b>	<b>TSP1 (568-617)</b>
<i>S. scrofa</i>	PVLSGTECLCVCQSGTYGNCERRAPDYKSNVAVDGNWGCWSSWSTCDATYKRSRTRECN	591
<i>B. taurus</i>	PVLSGTECLCVCQSGTYGNCERRSPDYKSNVAVDGNWGCWSSWSTCDATYKRSRTRECN	591
<i>H. sapiens</i>	PTLSGTECLCVCQSGTYGNCERRKQSPDYKSNVAVDGNWGCWSSWSTCDATYKRSRTRECN	591
<i>P. troglodytes</i>	PTLSGTECLCVCQSGTYGNCERRKQSPDYKSNVAVDGNWGCWSSWSTCDATYKRSRTRECN	591
<i>P. pygmaeus</i>	PTLSGTECLCVCQSGTYGNCERRSPDYKSNVAVDGNWGCWSSWSTCDATYKRSRTRECN	591
<i>C. familiaris</i>	PTLSGTECLCVCQSGTYGNCERRSPDYKSNVAVDGNWGCWSSWSTCDATYKRSRTRECN	593
<i>R. norvegicus</i>	PRLSGTECLCVCQSGTYGNCERRSPDYKSNVAVDGNWGCWSSWSTCDATYKRSRTRECN	591
<i>M. domestica</i>	PTLSGTECLCVCQSGTYGNCERRAPDYTSNEVDGAWSCWSSWSTCDATYKRSRTRECN	591
<i>G. gallus</i>	PVLSGTECLCVCQAGTYGTNCEIRAPGYESVAVDGRWSCWSEWSTCDATYKRSRTRECN	591
<i>X. tropicalis</i>	VVLSGTECLCVCQPGTYGNCERRTPDYTSVVDGAWSCWKAWSSTCDATYKRSRTRECN	592
<i>O. mykiss</i>	AVLSGTECLCVCQPGTYGNCERRAPDYTSVVDGAWSCWGTWSTCDATYKRSRTRECN	596
	**** : * .	
<i>S. scrofa</i>	PAPRQGGKPCDGERRQEEHCTFSIMQNTGQPCIGDDEDMEIDLPEIES-DSGCPQPVPP	650
<i>B. taurus</i>	PAPRQGGKRCEGERRQEEHCTFSIMQNDGQPCISDDEDMEIDLPEIES-DSGCPQPVPP	650
<i>H. sapiens</i>	PAPRQGGKRCEGEKQEEHCTFSIMENNGQPCINDDEEMKEVDLPEIEA-DSGCPQPVPP	650
<i>P. troglodytes</i>	PAPRQGGKRCEGEKQEEHCTFSIMENNGQPCINDDEEMKEVDLPEIEA-DSGCPQPVPP	650
<i>P. pygmaeus</i>	PAPRQGGKHCEGEKQEEHCTFSIMENNGQPCINDDEEMKEIDLPEIEA-DSGCPQPIP	650
<i>C. familiaris</i>	PAPRQGGKPCGEGERQEEHCTFSIMENNGQACISDDEEVEKIDLPEIES-DSGCPRPVSP	652
<i>R. norvegicus</i>	PEPQGGQRCEGKHQEEHCTFSIMEKVGQPCISDDEEIKEVDLAEPEA-DSGCPQPP	650
<i>M. domestica</i>	PAPQNGGKPCVGEQEEDCSFIFENRGDLINDEEAEKEVDIDETNP-ELGCVKPIP	650
<i>G. gallus</i>	PSPINGGKPCKGEREEEEECYVSVFMDNGAPCINDDEAKSELDIVVGL-ETGCSRDPDP	650
<i>X. tropicalis</i>	PYPRNGGKPCGEATQEEHCTFSIFEDTGALCINEGD-KKETDVRVQPEH-DTGCPKPDLL	650
<i>O. mykiss</i>	PAPLRGGKLCQGPQEEGCFISIFQEQ-NVCVNDDEFATEGRAEGLPPGVEGCPRPKRP	655
	* * .	

Figure 46 (continued)

<b>CCP domain (644-699)</b>	
<i>S. scrofa</i>	ENGFIRNEKTLYSVGEDVEIVCLTGFKTVGYQYFRCLPDRTWRRGDVEQRTDCLLPVVP 710
<i>B. taurus</i>	ENGFIRNEKKQYSVGEEVEILCFGTGFKAVGYQYFRCLPDRSWRQGDVEQORTECLKPIVP 710
<i>H. sapiens</i>	ENGFIRNEKQLYLVGEDVEISCLTGFTVGYQYFRCLPDGTWRQGDVEQORTECIKPVVQ 710
<i>P. troglodytes</i>	ENGFIRNEKQLYSVGEDVEISCLTGFTVGYQYFRCLPDGTWRQGDVEQORTECIKPVVQ 710
<i>P. pygmaeus</i>	ENGFIRNEKKLYSVGEDVEILCLTGFTVGYQYFRCLPDGTWRQGDVEQORTECIKPVVQ 710
<i>C. familiaris</i>	ENGFIRNEKKLYSVGEEVEILCFGTGFKTVGYQYFRCLPDRTWRRGDVECRKTTQCLKPVVQ 712
<i>R. norvegicus</i>	ENAFVWNEKKLYSVGEEVEISCLTGFKAVGYQYFRCLPDRTWRQGDVEQORTECLKPVVQ 710
<i>M. domestica</i>	ENGFLRNEKRLYSVGEEAEVACMMGFNLVGYPYLRCLPDQTWRQEGVEQQRVGCICKPFVQ 710
<i>G. gallus</i>	ENGFIRNEKNQYAVGEEAEIACVSGHVLIGYQFLRCLPDQTWTQPPVEQPSLCLRPPTS 710
<i>X. tropicalis</i>	GI AFLTNEKNWYDVAEQIEIACVSGYELSGYPFLRCLPDGTWKQEDVECVKTTCPRPKAS 710
<i>O. mykiss</i>	ANSHLRKDKRYEFQDMEEFLCFGTGFEEMGYQFINCRPDGTWTPPTGMCIRKVCSPPAVP 715
	.. : : * * .. : * . * . * . * * : : . * * * : *
<b>CCP domain (704-761)</b>	
<i>S. scrofa</i>	EVLKLSPFQRLYKIGDSIELSCPK-GFVIAGPSRYTCSGD-SWTTPPISSSLTCEQDFLTR 768
<i>B. taurus</i>	EGLTLPSPFQTLKIGDSIELTICPR-GLVVGPSRYTCSGD-SWTTPPIDSLSCSEKDVLTG 768
<i>H. sapiens</i>	EVLTIPTFQRLYRIGESIELTICPK-GFVVAGPSRYTCQGN-SWTTPPISNSLTCEKDTLTK 768
<i>P. troglodytes</i>	EVLTIPTFQRLYRIGESIELTICPK-GFVVAGPSRYTCQGN-SWTTPPISNSLTCEKDTLTK 768
<i>P. pygmaeus</i>	EVLTIPTFQRLYRIGESIELTICPK-GFVVAGPSRYTCQGN-SWTTPPISNSLTCEKDTLTK 768
<i>C. familiaris</i>	EVLTIPTFQRLYRIGESIELTICPK-GFVVAGPSRYTCSGD-SWTTPPISNSLTCEKDVLTG 770
<i>R. norvegicus</i>	DVLTISPFPQSVYKIGESIELTICPR-GFVVAGPSRYTCKGD-SWTTPPIPNSLSCSEKDLTK 768
<i>M. domestica</i>	EGISITPYQSVYDIGNTIKLTCPY-GFAITGPPNYLQCGKQ-SWEPAIPSSSLTCRQDILP- 767
<i>G. gallus</i>	DSVEISPFKLHYNIGETVKLSRA-GFVVGTQTYTCGKDLQSWFSPILRSITCEKDEQAK 769
<i>X. tropicalis</i>	EDITIYHYKTDYKVGESIHVSCPR-DFVVVGSARYTCGRDLEWNPPIRLQLTCKGKETQKV 769
<i>O. mykiss</i>	EGMTLYPSKKEYKVDHSGVLECTERGMVPEQGFYTCASLTTWEPPLPKDLHCKIDKPFV 775
	: : : : * : : : : * . . . * * * * . * * . : * :
<b>FIMAC domain (767-839)</b>	
<i>S. scrofa</i>	LKGHCQPQKQLGSECVCLSPPEEDCSHHSSEDLVLDTDSNHYFTSSACKFLAEKCLNNQQ 828
<i>B. taurus</i>	LRGHCQPQKQLGSECVCMSPPEEDCGHYSEEICVLDTTSDDYFTSSACKLLAEKCLNNQQ 828
<i>H. sapiens</i>	LKGHCQLGQKQSGSECI CMSPPEEDCSHHSSEDLVDFDTSNDYFTSPACKFLAEKCLNNQQ 828
<i>P. troglodytes</i>	LRGHCQLGQKQSGSECI CMSPPEEDCSHHSSEDLVDFDTSNDYFTSPACKFLAEKCLNNQQ 828
<i>P. pygmaeus</i>	LKGHCQPQKQSGSECI CMYPEEEDCSHYSEDLVDFDTSNDYFTSPACKFLAEKCLNNQQ 828
<i>C. familiaris</i>	LKGHCQPQKQLGSECI CMSPPEEDCSHYSEDLVDFDTSNDYFTSACKFLAEKCLNNQQ 830
<i>R. norvegicus</i>	SKGLCQPQKQSGSECVCMSPPEEDCSSYSEDLVDFDEGSSQYFTSSACKFLAEKCLNSNQ 828
<i>M. domestica</i>	FPENCQPGHKLVESEKCVCMSPPEEDCSHYSEDLVFNSESNHYFTKPSCKFLAEKCLNAQ 827
<i>G. gallus</i>	IRGVCNPGQKQVSECVCMNPEEEDCGHYSEDI CVLHALSEQHVTKPSCQFSAEKCLNEQS 829
<i>X. tropicalis</i>	IQGNCKPGQKQMGSDCI CMSPETDCGHFTEDLCTFDVAVSQNPVTISRCKFLAENCLGHKK 829
<i>O. mykiss</i>	PDNQCGRGERHVGSKCICVARET-CLPYKTEFCIFNAEAVGSGVMMSYCGFHSGRCHG-DQ 833
	* * . : * . * : * * . . . * * . : : * . . .
<i>S. scrofa</i>	LQFLHIGSC-QDGPQLEWGLERIKLSSSSTKNESCQYDTCYDWEKCSATTSKCI CLLPQ 887
<i>B. taurus</i>	LHFVHIGSC-EEGPQLKWLGERIKLSSSSTKNESCQYDTCYNWEKCSATTSKCVLLPSQ 887
<i>H. sapiens</i>	LHFLHIGSC-QDGRQLEWGLERLRLSSNSTKKESCGYDTCYDWEKCSASTSKCVLLPPQ 887
<i>P. troglodytes</i>	LHFLHIGSC-QDGHQLEWGLERLRLSSNSTKKESCGYDTCYDWEKCSASTSKCVLLPPQ 887
<i>P. pygmaeus</i>	LHFLHIGSC-QDGRQLEWGLERARFSSNSTKKESCGYDTCYDWEKCSASTSKCVLLPPQ 887
<i>C. familiaris</i>	LHFLHIGSC-QDGPQLEWGLERRILSSNSTKKEPCGYDTCYDWEKCSASVSKCVLLPPQ 889
<i>R. norvegicus</i>	FHFVHAGSC-QEGPQLEWGLERLKLAMKSTKRVPCGYDTCYDWEKCSAHTSNVCLLPQ 887
<i>M. domestica</i>	LHFLSAGAC-QDGPQLERAIERIKLSYNSSTKKEPCGYDTCYDWEKCSASTSTCFCLMPYQ 886
<i>G. gallus</i>	FHFLHAGPC-HSTSSVHWAIERAKLSANSLKVKPCGYDTCYDWEKCSAHTSNVCLLPYQ 888
<i>X. tropicalis</i>	LQFLDNGPC-KRD-QLDWYRTRVSLAVSSTKKEPCGYDFCYDWEKCSG--SECFCLLPQ 885
<i>O. mykiss</i>	LFFMNVPQDEDVVSLDWAKFRVSMMAKSSVQEPDSDTCYEWESCTES-KQCLCKVPRD 892
	: * : * * . . . : * : : * . . . * * * * . * * * : :
<b>FIMAC domain (861-934)</b>	
<i>S. scrofa</i>	CLKGGYQHYCVKMGSSSTTRTMNICEVGAIRCALRKMELIYPGRCSN- 935
<i>B. taurus</i>	CTKGGDQLFCVQIGSSANGKTMNICEVGAVKCAKREMEILHSGR- 932
<i>H. sapiens</i>	CFKGGNQLYCVKMGSSSTSEKTLNICEVGTIRCANRKMELIHPGKCLA- 934
<i>P. troglodytes</i>	CFKGGNQLYCVKMGSSSTSEKTLNICEVGTIRCANRKMELIHPGKCLA- 934
<i>P. pygmaeus</i>	CFKGGNQLYCVKMGSSSTSEKTLNICEVGAIRCANRKMELIHPGKCLA- 934
<i>C. familiaris</i>	CFKGGSLYCVKMGSSSLEKTVNICEVGAIRCANRKMELIYPGRCSA- 936
<i>R. norvegicus</i>	CPKDENQLHCVKMGSSMRGKTVNICTLGAVRCANRKMELIHPGKCLA- 934
<i>M. domestica</i>	CPKNGSQVYVCI RTGSSKTEKTMNICALGTYLCAQKKVEIVHHGQCIA- 933
<i>G. gallus</i>	CPREESQLHCIRMESTGRRKRVSCHVLAAMKACAGIGLEVEEESCLG- 935
<i>X. tropicalis</i>	CPPNSEVQYCAKVGSAQKRTISLCSLGAIKCASIKVEVLYDGACSTTN 934
<i>O. mykiss</i>	CPKEGEHMFCKVLLKTSKRVSVNLCFMAAIKCRKMEFELINEGLCEEST 941
	* . * : . : : : * : : : * . . . : * *

Figure 46 (continued)

<i>M. musculus</i>	<u>MQVTSLLILVCFIAAFQV</u> -----FSSASSPVNCKWDSYGPWSECNGCTKTQTRRR	50
<i>R. norvegicus</i>	<u>MKATSLLLILVGFMTFQI</u> -----FSSASSPVNCEWDSYGPWSECNGCTKTQTRRR	50
<i>S. scrofa</i>	<u><b>MKAMSLVFLVGLIGEFQV</b></u> ----- <b>FSSASSPVNCQWDSYAPWSECNGCTKTQTRRR</b>	50
<i>B. taurus</i>	<u>MKAITLLFLVGFIFGEFQV</u> -----FSSASSPINCQWGSYAPWSECNGCTKTQTRRR	50
<i>H. sapiens</i>	<u><b>MKVISLFLILVGFIFGEFQS</b></u> ----- <b>FSSASSPVNCQWDFYAPWSECNGCTKTQTRRR</b>	50
<i>D. rerio</i>	<u>MKQVFLHGVFLLSLLSSAPIT</u> THAGSLRSIRSASEP <sup>VH</sup> CLWGSWSSWSACDRCSKTQTQIR	60
<i>O. mykiss</i>	<u>MKDRLSVSLICLSWLF</u> FG-----MFSPINCVEPVHCQWGSYGDWSECNGCTKTQSRAR	53
<i>P. olivaceus</i>	<u>MKLSLAVCSLLLLILLSP</u> -----VCCQQSVNCRWGPYGEWSECNGCTSTEARTR	49
	*: : . . . : * * . : * * * : * : * : * :	
	<b>TSP1 domain (30-80)</b> <b>LDLa domain (84-120)</b>	
<i>M. musculus</i>	<u>SVAVYGQYGGYPCEGS</u> AFETQSCKPERGCPTEEGCGDRFCFSGQCISKSLVCNGDPDCE	110
<i>R. norvegicus</i>	<u>SI</u> AVYGQYGGHRCVGS <sup>AF</sup> ETQSCNPERGCPTEEGCGDRFCFSGQCISKSLVCNGDSDCE	110
<i>S. scrofa</i>	<u>PVAVYGQYGGHPCV</u> GSTFETQSC <sup>EP</sup> TRGCPTEEGCGERFCFSGQCISKSLVCNGDSDCE	110
<i>B. taurus</i>	<u>SI</u> AVYGQYGGHSCVGS <sup>AF</sup> ETQPCQPTRGCPTEEDGCGERFCFSGQCISKSLVCNGDSDCE	110
<i>H. sapiens</i>	<u>SVAVYGQYGGQPCV</u> GNAFETQSC <sup>EP</sup> TRGCPTEEGCGERFCFSGQCISKSLVCNGDSDCD	110
<i>D. rerio</i>	<u>FVAVFSQF</u> GGQPCTGSSTR <sup>TQ</sup> TCISTQVCPLEEGCGGRFCQSGK <sup>CI</sup> SLSLVCNSDQDCE	120
<i>O. mykiss</i>	<u>AMVVYAQF</u> GGSPCSGGATQ <sup>TQ</sup> PCVTARGCPLKEGCGGRFCRS <sup>GK</sup> CISSQSMVCNGDQDCE	113
<i>P. olivaceus</i>	<u>HVEVYAQF</u> GGAAACSGEATQ <sup>TQ</sup> CPVPQKRCPLETGCGDRFRC <sup>TS</sup> GQCISRTLLCNGDFDCN	109
	: * : * : * * * * * * : * * : * * * * * * * * : * * * * * * * * :	
<i>M. musculus</i>	<u>EDGADEDKCENVAN</u> -PSCNID---KPPPNIELTGYGYNVVTGQGKKRVINTKSFGGQCRK	166
<i>R. norvegicus</i>	<u>EDGADEDRCEGA</u> ESRPSCDTV---KPPPNIELTGN <sup>GY</sup> NALTGQFRNRVINTKSFGGQCRK	167
<i>S. scrofa</i>	<u>EDSADEDRCE</u> DSESRPSCDLS---KPPPNIELTGN <sup>GY</sup> NALTGQFRNRVINTKSFGGQCRK	167
<i>B. taurus</i>	<u>EDGADEDRCE</u> DAESRPACDKD---KPPPNIELTGR <sup>GY</sup> NALTGQFRNQVLN <sup>TK</sup> SFGGQCRK	167
<i>H. sapiens</i>	<u>EDSADEDRCE</u> DSESRPSCDID---KPPPNIELTGN <sup>GY</sup> NELTGQFRNRVINTKSFGGQCRK	167
<i>D. rerio</i>	<u>-DG</u> SDEQRCD <sup>SK</sup> PICKISENTNLQKPPPNVEITGQGFDAAKREARGT <sup>VINT</sup> KSFGGLCQK	179
<i>O. mykiss</i>	<u>EDNQDELKCG</u> PDKTFPVCNND---KPPPNVEQLGLGFDAVTGKQ <sup>RG</sup> SVINTKSFGGQCRT	170
<i>P. olivaceus</i>	<u>-DGL</u> DERGCFQDDQ <sup>F</sup> SCDLD---KTPPNADFTGRGYDVL <sup>TG</sup> KLKLAGVINTQ <sup>S</sup> FSGGQCGK	165
	* . * * * : : * * * * * : * * : * : * : * : * * * * * * * * :	
<i>M. musculus</i>	<u>V</u> FGDGKDFYRLSGN <sup>IL</sup> SYTFQVKVDNDFNYEFYNSSWSYIKHTSTEQNTFY <sup>S</sup> -WKGLFS	225
<i>R. norvegicus</i>	<u>V</u> FGDGRDFYRLSGN <sup>IL</sup> SYTFQVKIDNDFNYEFYNSSWSYV <sup>KHT</sup> STDY <sup>TSS</sup> SSNRFLFFS	227
<i>S. scrofa</i>	<u>V</u> FGDGRDFYRLSGN <sup>VL</sup> SYTFQVKVNDFNYEFYN <b>ST</b> WSYAKHTSTEHTSSSKGRV <sup>FIF</sup> S	227
<i>B. taurus</i>	<u>V</u> YSGDGRDFYRLSGN <sup>IL</sup> SYTFQVKINDFNNEFYNSTWAV <sup>VK</sup> ETSTEHS <sup>SS</sup> SKGRFLFFS	227
<i>H. sapiens</i>	<u>V</u> FGDGKDFYRLSGN <sup>VL</sup> SYTFQVKINDFNYEFYN <b>ST</b> WSYV <sup>KHT</sup> STEHTSSSRKRS <sup>FF</sup> RS	227
<i>D. rerio</i>	<u>TF</u> SGDHKDFYRLPQSVLSY <sup>SF</sup> QVTAKNDF <sup>TDES</sup> FASSWHY <sup>LH</sup> HYEKHEK-----TT	230
<i>O. mykiss</i>	<u>V</u> LSGDNKVIYRLPQSTLRYNFEV <sup>KV</sup> QND <sup>F</sup> SDEFY <sup>TSS</sup> WSYAKDIVKRET-----TT	221
<i>P. olivaceus</i>	<u>V</u> FGD <sup>HK</sup> VYRLPQSVLRYNFEV <sup>KV</sup> DNED <sup>TDES</sup> YESSWSY <sup>TQ</sup> HIQAN-----	212
	. * * * : * * * . * * . * * . * * . * * : * * : * * * * * :	

Figure 47 Alignment of the C7 proteins among species according to the GenBank accession no. as given in table 7 using the ClustalW2 tool. The amino acids are enumerated above each line on the right side. The identical, similar and missing residues are indicated with asterisks (\*), points (.), and dashes (-), respectively. Signal peptides in human C7 protein (DiScipio et al. 1988) and in pig C7 protein are in bold and italic. The amino termini of C7 mature protein starts just after the signal peptide. The cysteine residues are highlighted. The cysteine-rich functional protein domains like TSP1, LDLa, MACPF, CCP, and FIMAC are in continuous underline. The glycosylation sites in human (DiScipio et al. 1988) and in pig (Agah et al. 2000) present in shade blocks with white bold font. Fragment of amino- (a.a 23-575) and carboxyl-terminus (a.a 576-843) is in bold and double underline (DiScipio et al. 1988). Continuously italic sequence in CCP domain is the internal peptide obtained by tryptic digestion and Edman degradation of native porcine C7 (Agah et al. 2000)

<i>M. musculus</i>	HSRNTYGHGSAKEEIDTKMKSYKLLVVQNTVEVAQFTNNNPEFLQVAEPFWKELSHLPTL	285
<i>R. norvegicus</i>	SSSNRHTQSSSTKEIYMK-KSYKLLVVQNTVEVAQFINNNPEFLQLAEPFWRELSHLPTL	286
<i>S. scrofa</i>	SSSSSSSYAKTYEILKK-KSYQLLVVQNTVEVAQFINNNPEFLQLAESFWKELSYLPLL	286
<i>B. taurus</i>	SSSSSHGYSSNTNLTGK-KSYQLLVVQNTVEVAQFINNNPEFLQLAESFWKELSYLPSL	286
<i>H. sapiens</i>	SSSSSRYSHTSHTNEIHKG-KSYQLLVVQNTVEVAQFINNNPEFLQLAEPFWKELSHLPSL	286
<i>D. rerio</i>	GTDYGHDDYVHFHDELSQS-KSKNLMIIKSDVEVGQFKNKDPEYLPSEDFWKALVALPVV	289
<i>O. mykiss</i>	GTTTGFNNYDLHQTEEKN-RNNHLLVVKNNVEVAQFQNPAPGYLSLSEEFWKVLATLPTV	280
<i>P. olivaceus</i>	-ALFGHRRRTFHKDLTEN-KASRLIILKNKVELAQFQNSAPQYLTLSSEGFWKALSSLPFT	270
	: * : : : : * : : * * * * * : * : * : * * : * * *	
<b>MACPF domain (249-450)</b>		
<i>M. musculus</i>	YDYSAYRRLIDQYGYTHYLQSGSLGGEYRVLVYVDSGSAKETGFQSDQDNACSSADFQFLF	345
<i>R. norvegicus</i>	YDYSAYRRLIDQYGYTHYLQSGSLGGEYRVLVYVDSGDKVHRGFGSVQEKACSSSDFKFI	346
<i>S. scrofa</i>	YDYSAYRRLIDQYGYTHYLQSGSLGGEYKVLVYVDSKVAESDLGSEDKKKACASSHLSFLF	346
<i>B. taurus</i>	YDYSAYRRLIDQYGYTHYLQSGSLGGEYKVIHMDSEKVKKFDHSEDKRKCASSHFQFLV	346
<i>H. sapiens</i>	YDYSAYRRLIDQYGYTHYLQSGSLGGEYRVLVYVDSKELKQNFVSEKFKKSSGWHFVV	346
<i>D. rerio</i>	YDYAAYRNVLERFTHYIIEGSLGGHFKLYLMASEDVISKLKSEKRDYEDCVVTSHSVVF	349
<i>O. mykiss</i>	YDYATYRMVVERFGTHYLSGTLGGYFQALLSIDQETATQMAKVWTKYNECTKTKHRILF	340
<i>P. olivaceus</i>	YDYSAYRQLLQTYGYTHYLSGSLGGEYQGLLELDRQAFTSTSTTDEYKRCWKVKRRFF	330
	***:*** : : : * : : * : : * : : * : : * : : * : : * : : * : : *	
<i>M. musculus</i>	TS-SADQRCKMQLTEKSTSGNKGRLLR--GKPLVRGGDSGFVADLSFLDLDNPAGNKQR	402
<i>R. norvegicus</i>	VS-SANQRCKELENIKFSNSGNQGRVLR--GEPFVIGGGPGFVADLSFLDLDNPAGNRQR	403
<i>S. scrofa</i>	KS-SK-HKCKAMEEALKSASGTQSNVLR--GVPFVRGGPRGFGVSLSYLLELDNPAGNKQR	402
<i>B. taurus</i>	TS-SK-QKCTTMEEVKLSVSENEGNNLR--GVPFVRGGHSGFGLAGLSYLDLNNPAGNKRR	402
<i>H. sapiens</i>	KF-SS-HGCKELENAKKAASGTQNNVLR--GEPFIRGGGAGFISGLSYLELDNPAGNKRR	402
<i>D. rerio</i>	FIRWSTKTKTDTIDETKTFYKSIPESD--MKTDIIGGDPGFIKLSMFYKYDVKENART	407
<i>O. mykiss</i>	VS-WTEKCRKDEKEYTLNPPSISRSDDTVKVDVEGGATAHIAALKALDLNTPGRNNDM	399
<i>P. olivaceus</i>	RK-KVKITCETFMNSISSRDGHNVNKMP--IKVNVFGGDPSPFIALSVLDLEKPEITNGEI	387
	* : : : : * : : * : : * : : * : : * : : * : : *	
<i>M. musculus</i>	YSSWAGSVTRLQVVIKEKLAPLYELVKEVPCASVKRLLYKRAIEEYFDEFDPCHCRPCQN	462
<i>R. norvegicus</i>	YSSWASSVTSPLQVIKQRLTPLYELVKEVPCASVKRLLYKRAIEEYFDEFDPCHCRPCQN	463
<i>S. scrofa</i>	YSSWAGSVTDLPQVIKQKLTPLYELVKEVPCASVKRLLYKRAIEEYFDEFDPCHCRPCQN	462
<i>B. taurus</i>	YSQWAGSVTDLPEVIKQKLTPLYELVKEVPCASVKRLLYKRAIEEYFDEFDPCHCRPCQN	462
<i>H. sapiens</i>	YSAWAESVTNLPQVIKQKLTPLYELVKEVPCASVKRLLYKRAIEEYFDEFDPCHCRPCQN	462
<i>D. rerio</i>	FSQWAGSLKYPRIKSKLRSHELVLVKEVPCAGLKRFLKRAIEYTYLTKHSCQCRECQN	467
<i>O. mykiss</i>	YKNWAESVRFPAVIAKRMPLVKEVPCAGMKRFLKRAIEQYLNHRHPCRCQPCRN	459
<i>P. olivaceus</i>	YDNWASSVKDFPNIIDQKLRPLHELVLVKEVPCAGLKKLHLERATEEYLSSEHPCHCRPCQN	447
	: . * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
<i>M. musculus</i>	GGLAIVVETQCQCLCKPYTFGSACEQGVLVGDQAGGVDGGWNCWSSWSPCVQGKTRRSRE	522
<i>R. norvegicus</i>	GGIASVVTQCQCHCKPYTYGMACEQGVLVGDQAGGVDGGWSCWSSWSPCVQGKTRRSRE	523
<i>S. scrofa</i>	GGMASVEGTQCQCHCKPNTFSGVACEQGVLVGDHAGGIDGGWSCWSSWSPCAQGGKTRSRK	522
<i>B. taurus</i>	GGMATVQGSQCQCCKPKTSGVACEQGVLVGDQAGGVDGGWNCWSSWSPCVQGKTRRSRQ	522
<i>H. sapiens</i>	GGLATVEGTHCLCHCKPYTFGAACEQGVLVGNQAGGVDGGWSCWSSWSPCVQGKTRRSRE	522
<i>D. rerio</i>	NGLRVLDGNVCKCCKPGTSGQACEYGTAYDEQPGVIHGDWACWSSWSSCSGGQKSRRSR	527
<i>O. mykiss</i>	NGLVVMAGDKCSCICKPGTDGLACEKGEVEGQEGVIHGSWSCWSSWSSCSGGQKSRTRA	519
<i>P. olivaceus</i>	NGQPLLSGSVCRCRPGTSGPACQGTGAVIGEQQPLIHGWSCWSSWSSCSGGGRMSRTRS	507
	. * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *	
<b>TPS1 domain (503-551)</b>		
<i>M. musculus</i>	CNNPPPRDDGKSCLGETTESKQCEDQD-LEKLRLLLEPHCFHSSLAPKEFCLSPALKDGF	581
<i>R. norvegicus</i>	CNNPPPRAGGKACIGETTESRQCEDQD-LENLRLLLEPHCFHLSLVKPGFCPSPPALKDGF	582
<i>S. scrofa</i>	CNNPPPSGGGKSCIGETTESRQCEDED-LEHLRLLLEPHCFPLSLVPTTEFCPSPPALKDGF	581
<i>B. taurus</i>	CNNPSPSAGGKSCIGETTESRQCEDEE-LEHLRLLLEPHCFPLSLVPTTKFCSSPPALKDGF	581
<i>H. sapiens</i>	CNNPPPSGGGRSCVGETTESRQCEDEE-LEHLRLLLEPHCFPLSLVPTTEFCPSPPALKDGF	581
<i>D. rerio</i>	CTRPAAPS-GGRDCIGNTEERTACEDEEELNHLRSMEPHCFDSDIKPRESCKTPPFVPPNGF	586
<i>O. mykiss</i>	CNSPAPQRGGHHCNGEVRETTGCDDDQDLQYLQTMPEQCFDLTVPPKETCRSPPPLPNGY	579
<i>P. olivaceus</i>	CNNPAPSRGQHCTGLRAEQKPCEDPQ-IQYLQTMPEQCFSLSVTPPKTCGPPNLRNGF	566
	* . . * . * : * * * * * * : * : * : * : * : * : * : * : * : * : * : * :	
<b>CCP domain (571-626)</b>		
<i>M. musculus</i>	VQEGTMVPVQSVVYACDEGYSLIGDPVARCGEDLQWLVGEMHCQKLACVLPGMNGMQ	641
<i>R. norvegicus</i>	VQSEGTMFPIGKNIVYACNEGYSLVGDPVARCGEDLQWLVGEMHCQKIAACVLPGMNGMR	642
<i>S. scrofa</i>	VQNEETTFVPGKNIVYSCNEGYSLVGDPVARCGEDLQWTVGKMHQKIAACVLPMLMDGIQ	641
<i>B. taurus</i>	VQDEGATFVPGKNIMYTCNEGYSLVGDPVARCGEDLQWLVGEMHCQKIAACVLPALMDGIQ	641
<i>H. sapiens</i>	VQDEGTMFVPGKNVYVTCNEGYSLIGNPVARCGEDLRWLVGEMHCQKIAACVLPVLMMDGIQ	641
<i>D. rerio</i>	VLYPKDVYVPGSKIEYTCIEGYHLIGNAIAKQEDLNWLQYVPECKKTQDPPQLPPDVT	646
<i>O. mykiss</i>	VLDPKDVYLVGSKIEYTCIEGYHLIGIRIAECTVALTWSTPSKECKSSRCHVPSLLNDVT	639
<i>P. olivaceus</i>	IQNPRDFYVVGNTVEFSCTEGHYLSGDAVTRCTENQWTWTPRTTVCKSTGIPQLGAEV	626
	: * . . : * : * : * * : * : * : * : * : * : * : * : * : * : * : * : * :	

Figure 47 (continued)

CCP domain (631-688)	
<i>M. musculus</i>	SHPQKPFYMVGEKVTLSVGGMSLEGPS-TLLCGSSLKWSPELKDAKCVQKENPATSTPP 700
<i>R. norvegicus</i>	SDPQKPFYMVGEKVTLSVGGMSLEGPS-TFLCGSSLKWSPELKSAQCQVQKEN----TLT 697
<i>S. scrofa</i>	<b>SHPQKPFYTVGEKVTFS</b> SSGMSLEGPS-TFLCGSSLKWSPEMKNVQCVRKEAP----LA 696
<i>B. taurus</i>	SHPHKPFYTI GEKVTISCSGGRSLEGPS-TFLCGSSLKWSPEVKDVQCVQREAP----LT 696
<i>H. sapiens</i>	SHPQKPFYTVGEKVTVSCSGGMSLEGPS-AFLCGSSLKWSPEMKNARCVQKENP----LT 696
<i>D. rerio</i>	ANPVKLNYNIGEAIALSCPEGKVKEGLD-EIQCNAGLSWSPQPKNTKCLTAVLKPT-VVP 704
<i>O. mykiss</i>	GSPWQPTYDIGERIPLSCKPKGRHIVGDK-EIICDSSLHWSPPDNTITCSQAPKTLT-HLD 697
<i>P. olivaceus</i>	ATPTKAAYQIGDGVLSLCPAGSLLDGEVSEVICSPSLQWSPSPAGARCKAAPTAPP-PPS 685
	. * : * :*: :..** * * *...* ***. *
FIMAC domain(696-765)	
<i>M. musculus</i>	PAAPACQPWEKLNKSKCICKMPYECGSSLDVCADES RNKILSLTICKMHVVHCHKRNYT 760
<i>R. norvegicus</i>	PAAPACQPWEKLNKSKCVCKMPYECGSSLDVCADES RNKILPLTICKMHVLQCCQGRNYA 757
<i>S. scrofa</i>	<b>KKVPECQLWEKLNKSKCVCKMPYECGSSLDVCADES</b> SKRILRLTVCKMHVLQCCQGRNYT 756
<i>B. taurus</i>	PKVPKCQPWEKLNKSRVCKMPYECGSSLDVCADES SKRILPLTVCKMHVLQCCQGRNYT 756
<i>H. sapiens</i>	QAVPKCQRWEKLNKSRVCKMPYECGSSLDVCADES SKRILPLTVCKMHVLHLCQGRNYT 756
<i>D. rerio</i>	---VKCQPWENLAKDKCVCKVPHQCKSSLEVCVTDEKRGRTQRLSVCKVQAMRCLGHQYS 761
<i>O. mykiss</i>	GPAGQCKPWEKLNKSKCICKMPYECTSSLVQVANNLENGRTRNRLSVCKMHTLNLCLGRSYN 757
<i>P. olivaceus</i>	G--LKCKPWETAGKNKCVCKMPYECGSSLDVCADES SASSLLLCAR--LGSAPKLLSTCQLGALQCLGRSFS 741
	*: ** . :*:***:.* ** :*. *: **: .:* :.:
FIMAC domain (772-840)	
<i>M. musculus</i>	LVGKESCTLPSPAIEKVCAGCPIWSKCDQAQNGKCI CREASECQTAGYRICVEVN----GKE 816
<i>R. norvegicus</i>	LVGKNSCTLPASTEKACGACPIWSKCDQAQNSKCV CREASECQKAGYRICVEVN----GKE 813
<i>S. scrofa</i>	<b>LSVGETCTLP</b> PSAEKACGACPLWEKCDQAQSSKCV CRAASECEEAGFRVCVEVN----GRE 812
<i>B. taurus</i>	VAGRESCTLPASAEKACGVCPLWEKCDAEGSECV CRAASECEEAGFSVCVEVN----GRE 812
<i>H. sapiens</i>	LTGRDSCCTLPASAEKACGACPLWGCDAESSKCV CREASECEEAGFSVCVEVN----GKE 812
<i>D. rerio</i>	LAEDSACQWTQPSTNNCTRCTLGETCDEQTNSCRKTS EDCSSPDTWIHVQVKQDKASAP 821
<i>O. mykiss</i>	LAEDSACEWPTNTTSPCTDCQLWETCDGQTNQCRCKDWAE CSDPGLSVQVRMGDDANSAT 817
<i>P. olivaceus</i>	LTSDADCQWPEG-----AFCQESVRECPTDSAPLQVS-----SGVV 777
	: * . *: . * *
<i>M. musculus</i>	ETMSECEAGILRCRGQSSISITAIKPCAEEAAQ 848
<i>R. norvegicus</i>	ETMSECEAGILRCRGQSSISITAIKPCAQEVQ- 844
<i>S. scrofa</i>	<b>QTMTECEAGVLRCLGLSITVTSIRPCAPEAP-</b> 843
<i>B. taurus</i>	<b>QTMTECEAGVLRCLGLSISVTGIRPCVAEAA-</b> 843
<i>H. sapiens</i>	<b>QTMSECEAGALRCRGQSSISVTIRPCAETQ-</b> 843
<i>D. rerio</i>	VTMTECEVAMRKCRGETVHIVSIQACQS---- 849
<i>O. mykiss</i>	QTLSECEAGLRRCKGEKVSVVSIPLPSA---- 845
<i>P. olivaceus</i>	TTMTECEVGARRCGGELMNVIGIDACPE---- 805
	*:***. : * * : : . * *

Figure 47 (continued)

<i>S. scrofa</i>	----- <b>MLAVAFFILSLMTCQP</b> ----- <b>GVTI</b> PEKVNRRVSR 30
<i>B. taurus</i>	-----MFAAAFFLLFLMTYQP-----GVTIQEKVNRRVSR 30
<i>C. familiaris</i>	-----MFAAAFFILSLMICQP-----GVTIQEKVNRRVSR 30
<i>H. sapiens</i>	----- <b>MFAVVFFILSLMTCQP</b> ----- <b>GVTAQEKVNQRVRR</b> 30
<i>P. troglodytes</i>	-----MFAVVFFILSLMTCQP-----GVTAQEKVNQRVRR 30
<i>M. mulatta</i>	-----MFAVVFFILSLMTCQP-----GVTAQEKVNQRVRR 30
<i>O. cuniculus</i>	-----MLVAFFTLFLVTCQP-----AVTAQEKVNQRVNR 30
<i>M. musculus</i>	-----MFVVAFFGLSLVAWHP-----GVTAQEKVNQRVTR 30
<i>R. norvegicus</i>	-----MFTVAFFGLSLMAWHS-----GVTAQEKVNQRVTR 30
<i>G. gallus</i>	-----MWSLQVSPILITICSLSITARWATAAA-CGEPAAVPAQRRSSR 44
<i>X. tropicalis</i>	-----MIVMLYHALALWGYTA-----SIGNMDLLFR 27
<i>O. mykiss</i>	MSSFSSRSSFISNMNRLICVLLGSYILLVNLKSPVDATEYSWNMAETRTGQSAIRRV 60
	: * :

Figure 48 Alignment of the C8A proteins among species according to the GenBank accession no. as given in table 9 using the ClustalW2 tool. The amino acids are enumerated above each line on the right side. The identical, similar and missing residues are indicated with asterisks (\*), points (.), and dashes (-), respectively. Signal peptides in human C8A protein (Rao et al. 1987) and in pig C8A protein are in bold and italic.



<b>TSP1 (44-94)</b>	
<i>S. scrofa</i>	<u>AVLSSTPTAVS</u> <u>CLSNWAEWTE</u> <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGAI</u> <u>CSGNVWDQAS</u> <u>CHSP</u> 90
<i>B. taurus</i>	<u>AVQSFTPTAVS</u> <u>CLDNWAEWTD</u> <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGNVWDQAS</u> <u>CHSP</u> 90
<i>C. familiaris</i>	<u>AVQSAPITAVP</u> <u>CLSNWSEWTD</u> <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGDVWDQAS</u> <u>CHSP</u> 90
<i>H. sapiens</i>	AAT---PAAVTCQLSNWSEWTD <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGDIWDQAS</u> <u>CSSS</u> 87
<i>P. troglodytes</i>	AAT---PAAVTCQLSNWSEWTD <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGDIWDQAS</u> <u>CSSS</u> 87
<i>M. mulatta</i>	AAT---PAAVTCQLSNWSEWTD <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGDVWDQAS</u> <u>CSSS</u> 87
<i>O. cuniculus</i>	AAT---PRAFDCQLSSWSEWTD <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGDIWDRAS</u> <u>CYSP</u> 87
<i>M. musculus</i>	AVT---PQAVSCLSDWYKWTDC <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGDIWDEAS</u> <u>CDSP</u> 87
<i>R. norvegicus</i>	AVT---PGAVACQLSDWSEWTD <u>CFPCQDKKRYRSL</u> <u>LLQPNKFGGTI</u> <u>CSGDIWDKANC</u> <u>DSPI</u> 87
<i>G. gallus</i>	DISS---PPVDCQLSQWSEWTD <u>CFPCQGRKRRHRT</u> <u>LLQPNAMFGGRCK</u> <u>GPLWDEQ</u> <u>SCSAW</u> 102
<i>X. tropicalis</i>	SIAA---PQPEDCQLDQWSQWTS <u>CFPCQDKKRYRSL</u> <u>LLQPNAMFGGRCK</u> <u>GPLWDEQ</u> <u>SCSAW</u> 85
<i>O. mykiss</i>	SVN---KPAPINCKMKMWSWSP <u>CDSCD</u> <u>TKKFRFR</u> <u>YMEKASQ</u> <u>FGGRQCL</u> <u>DSQWEELAC</u> <u>PTA</u> 118
	. *::: * .*: * . * :.*.* :.: : * * . * * :
<b>LDLa (98-135)</b>	
<i>S. scrofa</i>	<u>TE-CLRQAQCGQDF</u> <u>CKETGRCLKRHLV</u> <u>CNGDRD</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DVRI</u> <u>FEDD</u> <u>CSQYDPI</u> 149
<i>B. taurus</i>	<u>TA-CLSQAQCGQDF</u> <u>CKETGRCLKRHLV</u> <u>CNGDKD</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DVRI</u> <u>LEND</u> <u>CSQYDPI</u> 149
<i>C. familiaris</i>	<u>TA-CLHQAQCGQDF</u> <u>CKETGRCLKRHLV</u> <u>CNGDKD</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DVRA</u> <u>ENDD</u> <u>CSQYDPI</u> 149
<i>H. sapiens</i>	TT-CVRQAQCGQDF <u>CKETGRCLKRHLV</u> <u>CNGDQD</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DVRA</u> <u>IDED</u> <u>CSQYDPI</u> 146
<i>P. troglodytes</i>	TT-CVRQAQCGQDF <u>CKETDRCLKRHLV</u> <u>CNGDQD</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DVRA</u> <u>IDED</u> <u>CSQYDPI</u> 146
<i>M. mulatta</i>	TP-CIRQAQCGQDF <u>CKETGRCLKRHLV</u> <u>CNGDQD</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DVRA</u> <u>IDED</u> <u>CSQYDPI</u> 146
<i>O. cuniculus</i>	TA-CLRQAQCGQDF <u>CKETGRCLKRHLV</u> <u>CNGEND</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DIRATE</u> <u>SDCAQYDPI</u> 146
<i>M. musculus</i>	TP-CLRQAQCGQDF <u>CRETGRCLKRHLV</u> <u>CNGDND</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DVRE</u> <u>FEDD</u> <u>CHQYDPI</u> 146
<i>R. norvegicus</i>	TP-CLRQAQCGQDF <u>CRETGRCLKRHLV</u> <u>CNGDQD</u> <u>CLDGSDEDD</u> <u>CE</u> <u>DARV</u> <u>IEDD</u> <u>CRQYDPI</u> 146
<i>G. gallus</i>	ES-CARAPSCGND <u>FCKE</u> <u>SGRCIK</u> <u>QHLV</u> <u>CNGNTD</u> <u>CRDGSDEDD</u> <u>CE</u> <u>HE</u> <u>-E</u> <u>I</u> <u>EH</u> <u>PC</u> <u>DNLF</u> <u>PI</u> 160
<i>X. tropicalis</i>	QK-CVPENNCGND <u>FCKE</u> <u>SGRCIK</u> <u>RRLV</u> <u>CNGDLD</u> <u>CRDGSDEDD</u> <u>CE</u> <u>DAP</u> <u>-E</u> <u>HET</u> <u>FC</u> <u>TLF</u> <u>PI</u> 143
<i>O. mykiss</i>	<u>QAVCWEPDI</u> <u>CGERFT</u> <u>CNATGR</u> <u>CVSQ</u> <u>ALRN</u> <u>GEVDC</u> <u>DDE</u> <u>SD</u> <u>ETD</u> <u>CE</u> <u>QVDR</u> <u>QDK</u> <u>STLL</u> <u>PI</u> 178
	* ** : * * . :.*:: : * ** : ** * ** :.* : : * **
<i>S. scrofa</i>	<u>PGSERATL</u> <u>GYNILTQE</u> <u>ETQ</u> <u>SVYD</u> <u>ARYYG</u> <u>GCET</u> <u>VYNGEW</u> <u>REL</u> <u>RYD</u> <u>SAC</u> <u>ER</u> <u>LYY</u> <u>GDE</u> <u>KEYF</u> 209
<i>B. taurus</i>	<u>PGSEKAAL</u> <u>GYNILTQ</u> <u>EEAQ</u> <u>HVYD</u> <u>ARYYG</u> <u>GCET</u> <u>VYNGEW</u> <u>REL</u> <u>QYD</u> <u>PAC</u> <u>ER</u> <u>LYY</u> <u>GDD</u> <u>KEYF</u> 209
<i>C. familiaris</i>	<u>PGSESAAL</u> <u>GYNILTQ</u> <u>KEAQ</u> <u>HVYD</u> <u>PRYYG</u> <u>GCET</u> <u>VYNGEW</u> <u>REI</u> <u>RYD</u> <u>PTC</u> <u>ER</u> <u>LYY</u> <u>GDE</u> <u>KEYF</u> 209
<i>H. sapiens</i>	<u>PGSQKAAL</u> <u>GYNILTQ</u> <u>EDAQ</u> <u>SVYD</u> <u>ASYYG</u> <u>GCET</u> <u>VYNGEW</u> <u>REL</u> <u>RYD</u> <u>STC</u> <u>ER</u> <u>LYY</u> <u>GDE</u> <u>KEYF</u> 206
<i>P. troglodytes</i>	<u>PGSQKAAL</u> <u>GYNILTQ</u> <u>EDAQ</u> <u>SVYD</u> <u>ASYYG</u> <u>GCET</u> <u>VYNGEW</u> <u>REL</u> <u>RYD</u> <u>STC</u> <u>ER</u> <u>LYY</u> <u>GDE</u> <u>KEYF</u> 206
<i>M. mulatta</i>	<u>PGSQKAAL</u> <u>GYNILTQ</u> <u>EDAQ</u> <u>SVYD</u> <u>ATYYG</u> <u>GCET</u> <u>VYNGEW</u> <u>REL</u> <u>RYD</u> <u>PTC</u> <u>ER</u> <u>LYY</u> <u>GDE</u> <u>KEYF</u> 206
<i>O. cuniculus</i>	<u>PGSEKAAL</u> <u>GYNILTQ</u> <u>EEAQ</u> <u>SVYD</u> <u>ARYYGG</u> <u>RCET</u> <u>VYNGEW</u> <u>RHV</u> <u>RYD</u> <u>VP</u> <u>C</u> <u>ER</u> <u>LH</u> <u>H</u> <u>G</u> <u>E</u> <u>D</u> <u>D</u> <u>KEYF</u> 206
<i>M. musculus</i>	<u>PGSERAAL</u> <u>GYNILTQ</u> <u>EEAQ</u> <u>SVYD</u> <u>AKYYG</u> <u>GCET</u> <u>VYNGE</u> <u>WR</u> <u>K</u> <u>L</u> <u>R</u> <u>YD</u> <u>PTC</u> <u>ER</u> <u>LYY</u> <u>GDE</u> <u>KEYF</u> 206
<i>R. norvegicus</i>	<u>PGSERAAL</u> <u>GYNILTQ</u> <u>EEGQ</u> <u>SVYD</u> <u>AKYYG</u> <u>GCET</u> <u>VYNGD</u> <u>W</u> <u>R</u> <u>R</u> <u>L</u> <u>YD</u> <u>PTC</u> <u>ER</u> <u>LYY</u> <u>GDE</u> <u>KEYF</u> 206
<i>G. gallus</i>	<u>PGSEKAAR</u> <u>GYNILTQ</u> <u>EAKR</u> <u>YIYD</u> <u>PK</u> <u>FLG</u> <u>SH</u> <u>C</u> <u>S</u> <u>V</u> <u>Y</u> <u>N</u> <u>G</u> <u>E</u> <u>W</u> <u>R</u> <u>A</u> <u>L</u> <u>R</u> <u>YD</u> <u>A</u> <u>C</u> <u>E</u> <u>R</u> <u>L</u> <u>Y</u> <u>G</u> <u>D</u> <u>E</u> <u>K</u> <u>E</u> <u>Y</u> <u>L</u> 220
<i>X. tropicalis</i>	<u>PGA</u> <u>E</u> <u>K</u> <u>S</u> <u>V</u> <u>R</u> <u>G</u> <u>I</u> <u>N</u> <u>I</u> <u>L</u> <u>T</u> <u>H</u> <u>E</u> <u>D</u> <u>T</u> <u>R</u> <u>N</u> <u>V</u> <u>I</u> <u>D</u> <u>H</u> <u>N</u> <u>Y</u> <u>F</u> <u>G</u> <u>G</u> <u>C</u> <u>E</u> <u>Y</u> <u>I</u> <u>N</u> <u>G</u> <u>E</u> <u>W</u> <u>R</u> <u>E</u> <u>L</u> <u>Y</u> <u>E</u> <u>P</u> <u>V</u> <u>C</u> <u>E</u> <u>M</u> <u>Y</u> <u>S</u> <u>D</u> <u>E</u> <u>E</u> <u>K</u> <u>Y</u> <u>F</u> 203
<i>O. mykiss</i>	<u>PG</u> <u>A</u> <u>G</u> <u>R</u> <u>G</u> <u>T</u> <u>Q</u> <u>G</u> <u>F</u> <u>N</u> <u>I</u> <u>L</u> <u>T</u> <u>G</u> <u>E</u> <u>F</u> <u>V</u> <u>D</u> <u>H</u> <u>V</u> <u>L</u> <u>D</u> <u>P</u> <u>Q</u> <u>Y</u> <u>Y</u> <u>G</u> <u>G</u> <u>C</u> <u>E</u> <u>Y</u> <u>V</u> <u>N</u> <u>G</u> <u>E</u> <u>W</u> <u>R</u> <u>K</u> <u>L</u> <u>I</u> <u>Y</u> <u>D</u> <u>P</u> <u>F</u> <u>C</u> <u>E</u> <u>N</u> <u>L</u> <u>H</u> <u>Y</u> <u>N</u> <u>E</u> <u>D</u> <u>E</u> <u>K</u> <u>N</u> <u>Y</u> 238
	** : .. * **** : : * : * . :.* : ** : ** : * . ** :.*:: : **

Figure 48 (continued) Position starts for amino terminus sequence of the mature porcine C8A protein in bold and double underline (Nakajima et al. 1998). The cysteine residues are highlighted. The cysteine-rich functional protein domains like TSP1, LDLa, and MACPF are in continuous underline. The glycosylation sites in human (Rao et al. 1987) and in pig (Nakajima et al. 1998) present in shade blocks with white bold font. Different amino acid sites in our protein and Nakajima et al.'s (1998) are in bold and double underlined. Position of sequence forming the disulfide bond to C8G is continuously bold underlined in italic (Plumb and Sodetz 2000). The proposed candidate site for CD59 recognition is performed in continuously bold in MACPF domain of the porcine C8A (a.a 382-420) (Nakajima et al. 1998).

<i>S. scrofa</i>	RNPYNFLKYHFEALADSKFSSESYYDDANDLLKKV <sup>2</sup> NEKSVSAGVTGIGPTGSPFTANVG	269
<i>B. taurus</i>	RKPYNFLKYHFEAQADTKISSEIYNDANDLLTKVKNKSVSSGLTIGVIRGVPVTVTAG	269
<i>C. familiaris</i>	RKPYNFLKYHFEALADTRFSSELYDDAHDLLSKVKNNNFVSTGTVGVFTGSSVTVDDVG	269
<i>H. sapiens</i>	RKPYNFLKYHFEALADTGISSEFYDNANDLLSKVKKDKSDSFGVTIGIGPAGSPLLVGVG	266
<i>P. troglodytes</i>	RKPYNFLKYHFEALADTGISSEFYDNANDLLSKVKKDKSDSFGVTIGIGPAGSPLLVGVG	266
<i>M. mulatta</i>	RKPYNFLKYHFEALADTGISSELYDNANDLLSKVRKDKSDSFGVTIGIGPAGSPLMVGVG	266
<i>O. cuniculus</i>	RKPYNFLKYHFEARADTGISFELYVDGNDLFSKVKNDKSHSAGVTISAGLTGSPLLGTVG	266
<i>M. musculus</i>	RKPYNFLKYHFEALADTISSEFYDDANDLFFHIKNGKSHSAGVTGVPVSPVSIIEVT	266
<i>R. norvegicus</i>	RKPYNFLKYHFEALADTLISSEFYDDANDLFSKIQRDKSQSNSVTFGISPAKSPITLDAS	266
<i>G. gallus</i>	LKPYNFHVYQFLAHADSGFSSEFYDDSKDLIDALKSSKSEGGGFITIGIPKKIDFKLNLG	280
<i>X. tropicalis</i>	RKPYNFHLYQFLARADTKMSIEIYEDSKDVVNAVKRDFSNIGLTFGIVPEAPVGLGLG	263
<i>O. mykiss</i>	RKPYNFHTYRFMAQATSEGSSEYEDMATLLKARKTEDSFNLGVTGIR-----YVEFG	292
	:**** *:* * * : * * * : :. : . . . * . .	
<i>S. scrofa</i>	LSGSRESAFLNKLSTYNEKKYSFIRIFTKVQTA <sup>2</sup> SFMMRRDNIMLDEVMLQSLMELP-EQY	328
<i>B. taurus</i>	VSMSQDAFLKLLSKYHEKKYSFMRIFTKVQTAHFKMRRRENIVLDEGMLQSLMELP-ERY	328
<i>C. familiaris</i>	VSSSQNSSSLDELKKNKKYSFLRVFTKVQTAHFKMRRGNIVLDEGMLQSLMELP-EQY	328
<i>H. sapiens</i>	VSHSQDTSFLNELNKNYNEKKFIFTRIFTKVQTAHFKMRRDDIMLDEGMLQSLMELP-DQY	325
<i>P. troglodytes</i>	VSHSQDTSFLNELNKNYNEKKFIFTRIFTKVQTAHFKMRRDDIMLDEGMLQSLMELP-DQY	325
<i>M. mulatta</i>	VSKSEDASFLKELNKNYNEKKFIFMIRIFTKVQTAHFKMRRDNIMLDEGMLQSLMELP-DQY	325
<i>O. cuniculus</i>	VSGSEDASFLNKLQYNEKKYFMRIFTKVQTAHFKMRRDDIVLDEGMLQALVELP-EQY	325
<i>M. musculus</i>	GS <sup>2</sup> GSKASSFLNKLNKNYNEKRYGFMRVSTKIQTAFKMRNNIVLDEGMLQSLMELP-EQF	325
<i>R. norvegicus</i>	VSWSEDSFMKELSKYNEKKYSFMRVSTKVQTAHFKMRRHNIIVLDEGMLQSLMELP-EQF	325
<i>G. gallus</i>	FTLSRGGKSLKNFTEYTAKNLGFIRIATKVQTA <sup>2</sup> RFKMRNNIVLDEDMILSLRELPE-DTY	339
<i>X. tropicalis</i>	LNYGLKTSFLK <sup>2</sup> ITSFNQKNLEFVRMVTKIQTARFKMRRNLLTLDEDMQSLMELP-DEY	322
<i>O. mykiss</i>	VSGNVESALLTNLT <sup>2</sup> TKYTNQELGFI <sup>2</sup> RLQSKVQTAQFKMRSEGLMLHEDMYLSLMEPEEKY	352
	. . . : :. : . . * * : * * * : *	
	<b>MACPF domain (291-497)</b>	
<i>S. scrofa</i>	<u>NYGMYAKFIDYGYTHYITSGSMGGVYEYILV<sup>2</sup>LNKE<sup>2</sup>NMTKSGVTSDDVTS<sup>2</sup>CFGGSFGIDYD</u>	388
<i>B. taurus</i>	<u>HYGMYAKFINDYGYTHYITSGSMGGVYEYILV<sup>2</sup>LNREKMETAGVTSAEIQCFGVSLGIEYE</u>	388
<i>C. familiaris</i>	<u>NYGMYAKFINDYGYTHYITSGSMGGIYEHILV<sup>2</sup>LNKEEMESQRITSRDIEKCFGISVGIYEYD</u>	388
<i>H. sapiens</i>	<u>NYGMYAKFINDYGYTHYITSGSMGGIYEYILVIDKAKMESLGITSRDITTCFGGSLGIQYE</u>	385
<i>P. troglodytes</i>	<u>NYGMYAKFINDYGYTHYITSGSMGGIYEYILVIDKAKMESLGITSRDITTCFGGSLGIQYE</u>	385
<i>M. mulatta</i>	<u>NYGMYAKFINDYGYTHYITSGSMGGTYEYILVIDKAKMESLGITSRDITTCFGGSLGIQYE</u>	385
<i>O. cuniculus</i>	<u>NYGMYSKFINDYGYTHYITSGSMGGTYEYILV<sup>2</sup>LNTEKMESLGVTSEDISSCFGFGEIQYE</u>	385
<i>M. musculus</i>	<u>NYGMYAKFINDYGYTHYITSGTMGGIYEV<sup>2</sup>MLDKEKMKTEGTTVDEVQKICIGGGIGIGIK</u>	385
<i>R. norvegicus</i>	<u>NYGMYSKFINDYGYTHYITSGTMGGIYEV<sup>2</sup>MLDKEKMRIGHITVEDVKKICIGGVGLQFG</u>	385
<i>G. gallus</i>	<u>NYGMYAKFINDYGYTHYITSGTMGGDLEYILV<sup>2</sup>VNKEEMRRKDISYEEVTT<sup>2</sup>CFGLSLGVS<sup>2</sup>AK</u>	399
<i>X. tropicalis</i>	<u>NYGLYAQFINDFGTHYTTSGTMGGIYEV<sup>2</sup>MLDKEIMKQEI<sup>2</sup>TASMDVCFGASVGLV<sup>2</sup>SQ</u>	382
<i>O. mykiss</i>	<u>DFGLYSRFLNTYGYHYV<sup>2</sup>TQIMGGTLEYVAVVNT<sup>2</sup>TAMKTSKIDAEQLKGLGGSIGISSP</u>	412
	. : * : * : * : : * * * * * : * : * : * : * : *	
<i>S. scrofa</i>	<u>YTD--NLQITGSLSGKHCKK---LGGG<sup>2</sup>HREDEES<sup>2</sup>NMAVEDIISRV<sup>2</sup>RGGSSG<sup>2</sup>WGGLTQ<sup>2</sup>N</u>	442
<i>B. taurus</i>	<u>YSE--AIQIKGSSSLGPKCK---SGDGKLTENEKAMGVEDFI<sup>2</sup>SRV<sup>2</sup>RGDSSG<sup>2</sup>WGGLTQ<sup>2</sup>D</u>	442
<i>C. familiaris</i>	<u>YAN--WMKIGGSLGGKCKEN---IGGGDSK<sup>2</sup>GHRLTTAVEDIISL<sup>2</sup>V<sup>2</sup>RGSSG<sup>2</sup>WGAGLAEK</u>	442
<i>H. sapiens</i>	<u>D----KINVG<sup>2</sup>GLSGDHCKK---FGGK<sup>2</sup>TERARKAMAVEDIISRV<sup>2</sup>RGGSSG<sup>2</sup>WSGGLAQ<sup>2</sup>N</u>	437
<i>P. troglodytes</i>	<u>D----KINVG<sup>2</sup>GLSGDHCKK---FGGK<sup>2</sup>TETARKAMAVEDIISRV<sup>2</sup>RGGSSG<sup>2</sup>WSGGLAQ<sup>2</sup>N</u>	437
<i>M. mulatta</i>	<u>G----GINVDG<sup>2</sup>LSGDHCKK---FGDGK<sup>2</sup>TERARKAMAVEDIISRV<sup>2</sup>RGGSSG<sup>2</sup>WSGGLAQ<sup>2</sup>N</u>	437
<i>O. cuniculus</i>	<u>KG---KINAQ<sup>2</sup>GILSGKHCKK---SGSGDKEADKMGQAVKDIISRV<sup>2</sup>RGGSSG<sup>2</sup>WGGLSQ<sup>2</sup>N</u>	438
<i>M. musculus</i>	<u>DS----TIEGV<sup>2</sup>GISGFCEN---SGDGRDIRKKITGVEDIISRV<sup>2</sup>QGGSSV<sup>2</sup>WGSV<sup>2</sup>LTHN</u>	437
<i>R. norvegicus</i>	<u>EK----IIEG<sup>2</sup>DLSGESCVM---TGDGNQDKR<sup>2</sup>KDLAVEDIISRV<sup>2</sup>QGGSSR<sup>2</sup>WSTGLAHN</u>	437
<i>G. gallus</i>	<u>KL---MLHWEASVLSVCAEQ<sup>2</sup>LLNAVDFSDAGSN<sup>2</sup>SPVMEDIIVRIKGGDT<sup>2</sup>SY<sup>2</sup>SARPISS</u>	456
<i>X. tropicalis</i>	<u>TEFE-EIIPSLKLSGDFCTK---FERENQDNSSSSRAIKDVIT<sup>2</sup>YVIGGDSG<sup>2</sup>SAGGILNV</u>	437
<i>O. mykiss</i>	<u>IGKTKQVEVGGKLEVKGCKG---TGSYEKEMYGSSSLIKDIVL<sup>2</sup>VKGGSTG<sup>2</sup>SGGLLAI</u>	468
	* : * : * : * : *	
<i>S. scrofa</i>	<u>GSIIITYRAWGRSLKYNPAVIDFEMKPIYEILRHTNLG-<sup>2</sup>PLEAKRQNLRRALDQYLMEFNA</u>	501
<i>B. taurus</i>	<u>SSLVITYRSWGRSLKYNPAVIDFEMKPIHEILQHTNLG-SLETQRQNLRRALDKYLMEFNA</u>	501
<i>C. familiaris</i>	<u>RSTITYRSWGRSLKYNPVVIDFEMQPIHEVLRHTTLG-<sup>2</sup>PLETKRQNLRRALDQYLMEFNA</u>	501
<i>H. sapiens</i>	<u>RSTITYRSWGRSLKYNPVVIDFEMQPIHEVLRHTSLG-<sup>2</sup>PLEAKRQNLRRALDQYLMEFNA</u>	496
<i>P. troglodytes</i>	<u>RSTITYRSWGRSLKYNPVVIDFEMQPIHEVLRHTSLG-<sup>2</sup>PLEAKRQNLRRALDQYLMEFNA</u>	496
<i>M. mulatta</i>	<u>RSTITYRSWGRSLKYNPVVIDFEMQPIHEVLRHTSLG-<sup>2</sup>PLEAKRQNLRRALDQYLTEFNA</u>	496
<i>O. cuniculus</i>	<u>GSATTYRFWGRSLKYNPVVIDFEMQPIHEVLLHTNLG-HVEAKRQNLRRALDQYLMEFNA</u>	497
<i>M. musculus</i>	<u>SSAITYRSWGRSLKYNPVVIDFEMQPIYQLLRHTNLG-<sup>2</sup>PLETKRQNLRRALDQYLMEFNA</u>	496
<i>R. norvegicus</i>	<u>SSAITYRSWGRSLKYNPVVIDFEMQPIYQLLRHTNLG-<sup>2</sup>PLETKRRNLRRALDQYLMEFNA</u>	496
<i>G. gallus</i>	<u>WDSNTYRRWGRSLKYNPAIDFELQPIHEILPRSDAG-NMETKRQHLKQALDEYLLEFNA</u>	515
<i>X. tropicalis</i>	<u>FDGRMYRWGRSLKYNPAVIDFETQPIYEG<sup>2</sup>LQQTGLS-GIEAKRQNLKRAYNEYLSEFDP</u>	496
<i>O. mykiss</i>	<u>KDPD<sup>2</sup>TYRKWGLSLKYNPNLIEFETLPIFELVRLSTAGDHV<sup>2</sup>GARPAHLRAWEYLLQFNS</u>	528
	. * : * * * * * : * * * : : : : : * : * : * : * : *	

Figure 48 (continued)

<i>S. scrofa</i>	CRCGPCFNNGEPILVGTSCRCQCQPVGCGQLACEQMKSE---GAKADGRWSCWSSWSACRL	558
<i>B. taurus</i>	CRCGPCFNNGEPILEGTSCKCQCQPLGKGLACEQMQSE---GAQADGRWSCWSSWSACRS	558
<i>C. familiaris</i>	CRCGPCFNNGKPILEGTSCKCQCPLGRKGLSCEQMEQK---GAKADGHWSCWSSWSACRA	558
<i>H. sapiens</i>	CRCGPCFNNGVPILEGTSRCRCQRLGLSLGAACEQTQTE---GAKADGWSWSCWSSWSVCRA	553
<i>P. troglodytes</i>	CRCGPCFNNGVPILEGTSRCRCQRLGLSLGAACEQTQTE---GAKADGWSWSCWSSWSVCRA	553
<i>M. mulatta</i>	CRCGPCFNNGVPILEGTSRCRCQRLGLSLGPACEQIQTE---GAKADGWSWSCWSSWSVCRT	553
<i>O. cuniculus</i>	CRCGPCFNNGKPILEGTSRCRCQSLGLQGPACEQTEQQ---GAKADGHWSCWGSWSPCTA	554
<i>M. musculus</i>	CRCGPCFNNGEPILDGTNCRQCQCSMGRQGLACERTVIEGLKDFKAAGHWSCWSSWSSECRG	556
<i>R. norvegicus</i>	CRCGPCFNNGVPILEGTSRCRCQCSMGRQGPACESMVLEAMEGAKADGRWSCWSSWSSECRG	556
<i>G. gallus</i>	CRCGPCQNNGEPVVLVGDECSCQCQPSGYSGPACERSEHQ---GTEVDGRWSCWSSWTPCQS	572
<i>X. tropicalis</i>	CRCGPCHNNGMPLENNVCTCLCAAGFNGPSCENTLRK---DVKADGRWSCWSPWTQCQS	553
<i>O. mykiss</i>	CRCAPCRHDGIPVLSQTSCHCIKQGFGEACEETLRK---DSTTDGAWSCWGAWSSCQS	585
	***. ** : : * * : * * * * * * * * : * : . . * * * * * . * : *	
	<b>TSP1 domain (547-586)</b>	
<i>S. scrofa</i>	GTQERRRE <sup>C</sup> NNPAPQNGGASCSGHKVQ^TQAC	589
<i>B. taurus</i>	GTQERRRE <sup>C</sup> NNPAPQNGGASCPGHRVQ^TQAC	589
<i>C. familiaris</i>	GTQERRRE <sup>C</sup> NNP^TPQNGGASCPGWKAQ^TQAC	589
<i>H. sapiens</i>	GIQERRRE <sup>C</sup> DNPAPQNGGASCPGRKVQ^TQAC	584
<i>P. troglodytes</i>	GIQERRRE <sup>C</sup> DNPAPQNGGASCPGRKVQ^TQSC	584
<i>M. mulatta</i>	GTQERRRE <sup>C</sup> DNPAPQNGGASCPGRKVQ^TQAC	584
<i>O. cuniculus</i>	GTERRRRE <sup>C</sup> NNPAPQNGGAPCPGWRVQ^TQAC	585
<i>M. musculus</i>	GSQERRRQ <sup>C</sup> NNPPKNGGTPCLGRNLQ^TQAC	587
<i>R. norvegicus</i>	GSRRERRQ <sup>C</sup> NNPAPQNGGAPCLGKSLQ^TQAC	587
<i>G. gallus</i>	GSRRRSRQ <sup>C</sup> TNPAPQHGAPCMGRDVQ^SSSC	603
<i>X. tropicalis</i>	GKRQRTRE <sup>C</sup> NNPAPKNGGAWCLGKSLQ^SEPC	584
<i>O. mykiss</i>	GSKTRRRSC <sup>D</sup> NPQP-DGGAAACLGSSSQ^NQR	615
	* : * * . * * * * * . * : * * * * . . *	

Figure 48 (continued)

<i>O. mykiss</i>	-----MNHKLKP--TVGLGYCLLCAAL <sup>C</sup> LLLLRD-----	27
<i>P. olivaceus</i>	-----MFRVAIPRSALNLHSC <sup>L</sup> LHVTL <sup>L</sup> SLVLSK-----	29
<i>D. rerio</i>	-----MHSFLRVN-----	8
<i>H. sapiens</i>	----- <b>MKNSRTWAWRAPVELFLLCAALG<sup>C</sup>LSLPGSRGER</b> ----	34
<i>M. mulatta</i>	-----MKNSRTSAWRALVELFLLCAALG <sup>C</sup> LSLPGSRGER----	34
<i>O. cuniculus</i>	-----MKKSWTWTRVPAELLLLCAALG <sup>C</sup> LPVPGSRSE----	34
<i>C. familiaris</i>	-----MKTSRPQWRALAE <sup>L</sup> LILCAALG <sup>C</sup> LSLPGSRSE----	34
<i>S. scrofa</i>	----- <b>MSLSQIS<sup>C</sup>RIEKMTSGTWAWRMPARL<sup>L</sup>CAALG<sup>C</sup>LSLPGSRGER</b> ----	46
<i>R. norvegicus</i>	-----MKTG-AQVWRALAKS <sup>C</sup> L <sup>L</sup> CAALG <sup>C</sup> LHLPGARGEK----	33
<i>G. gallus</i>	-----MTVAVMAFTL <sup>C</sup> PIKLLLLCAAL <sup>C</sup> FLDVH <sup>C</sup> FFSSG----	33
<i>M. domestica</i>	MFVWQNRWKGSSSRDSETWQNGDWPKRTV <sup>R</sup> GLMDYVRAMLF <sup>T</sup> SSWRNLSNVDTVSLYNTDD	60

Figure 49 Alignment of the C8B proteins among species according to the GenBank accession no. as given in table 11 using the ClustalW2 tool. The amino acids are enumerated above each line on the right side. The identical, similar and missing residues are indicated with asterisks (\*), points (.), and dashes (-), respectively. Signal peptides in human (Howard et al. 1987) and in pig C8B protein are in bold and italic. The amino termini of C8B mature protein starts just after the signal peptide. The cysteine residues are highlighted. The functional protein domains TSP1, LDLa, MACPF, and EGF are in continuous underline. The glycosylation site in human (Howard et al. 1987) and in pig using the NetNGlyc 1.0 Server tool present in shade blocks with white bold font.



```

O.mykiss          EFFLRLRALPQSYNYGEYRQIYRDYGYTHYITEATLGGDYEYTVILDKEKLEKTYGSLEAY 371
P.olivaceus      EFLQRLRLSLPQAYVYGEYRQIYRDYGYTHYITEAALGGEYEHTIILDKEKLAKTDYSLEDY 371
D.rerio          EFLQRLRLSLPQAYVYGEYRQIYRDYGYTHYITEAALGGEYEHTIILDKEKLAKTDYSLEDY 347
H.sapiens        EFLQRVKRLPLEYSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVNMKEAMERGDYTLNNV 375
M.mulatta        EFLQRVKRLPLEYSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVNMKEAMERGDYTLNNV 375
O.cuniculus      DFLOQRVQRVPLEYSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVNMKEAMERADYSLNDV 375
C.familiaris     EFLQRVKRLPLEYSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVNMKEAMERADYSLKAI 375
S.scrofa         EFLQRVKRLPLEYSYGEYRDLFRDFGTHYITEAVLGGVYEYTLVNMKEAMERADYSLKDV 387
R.norvegicus     EFLQRVKSLPLEYSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVNMKEAMERGDYTLHDV 374
G.gallus         EFLQRLHLQPLEYSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVNMKEAMERGDYTLHDV 369
M.domestica      EFLMRVGRLPDSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVNMKEAMERGDYTLHDV 416
                  : * . * : * * * * * : : * * * : * * : * * : * * : * * : * * : * * : * * : * * :

```

MACPF domain (303-510)

```

O.mykiss          KNCEQIVLKVGANIKGVYVTVGLEGGCDGLLNEMGEDTVKGS---MVEDYVAVVSGGDS 428
P.olivaceus      KSCVQAGLKI GANLYGVYVSAGIEGSGCNGLLNEMGEDTAIGS---SVEDFVAVVSGGSS 428
D.rerio          KNCVQVGLKVGQVVKKIYVGLGLSGGSC EGLLKEIGDSTKRD---MVEDVFIIVKGGDS 404
H.sapiens        HACAKNDFKIGGAI EEVYVSLGVSVGKCRGILNEIKDRNKRDT---MVEDLVVLRVGGAS 432
M.mulatta        HACAKNDFKIGGAI EEVYVSLGVSVGKCRGILNEIKDRNKRDT---MVEDLVVLRVGGAS 432
O.cuniculus      QACAKNDFKIGGAI EEVYVSLGVSVGKCRGILNEIKDRNKRDT---MVQDLVLRVGGAS 432
C.familiaris     QTCANQDFKIGGAI EEVYVSLGVSVGKCRGILNEIKDRNKRDT---MVEDLVVLRVGGAS 432
S.scrofa         HACAQHGFKIGGAI EEVYVSLGVSVGKCRGILNEIKDRNKRDT---MVNDLVVLRVGGAS 444
R.norvegicus     SACAGGGFQIGGNVYKVLKLVSEKKSDILNEIKDRNKRRT---MVEDLVVLRVGGTS 431
G.gallus         QKCAQHGFKIGGTIKAVSLILGVNVEGCKSLKKEIGDSTSKKQ---YVEDFIALVVRGGAS 426
M.domestica      KSCVQVGLKVGQVVKKIYVGLGLSGGSC EGLLKEIGDSTKRD---MVEDVFIIVKGGDS 475
                  * : . * : * * * * * : : * * * : * * : * * : * * : * * : * * : * * :

```

```

O.mykiss          ESITWLAANKLPTPPLMRLWGEAVHYNLD F IRSVTRPLYELVTARD FSSANSLKKNLRR 488
P.olivaceus      ESITGLVSKKLPTPQLMRLWGEAVRFPNDF IRKTRTRPLYELVTSKDFSHD ATLKRNKRA 488
D.rerio          ETVSRLAALKLPTPDIMQMGEAVFYNPEF I SKKIEPIYELVPPRE-ANANILKKNLKR 463
H.sapiens        EHI T T L A Y Q E L P T A D L M Q E W G D A V Q Y N P A I I K V K V E P L Y E L V T A T D F A Y S S T V R Q N M K Q A 492
M.mulatta        EHI T T L A Y Q E L P T A D L M Q E W G D A V Q Y N P A I I K V K V E P L Y E L V T A T D F A Y S S T V R Q N M K Q A 492
O.cuniculus      EHI T T L A Y Q E L P T A D L M Q E W G D A V Q Y N P A I I K V K V E P L Y E L V T A T D F A Y S S T V R Q N M K Q A 492
C.familiaris     EHI T T L A Y Q E L P T A D L M Q E W G D A V Q Y N P A I I K V K V E P L Y E L V T A T D F A Y S S T V R Q N M K Q A 492
S.scrofa         EHI T T L A Y Q E L P T A D L M Q E W G D A V Q Y N P A I I K V K V E P L Y E L V T A T D F A Y S S T V R Q N M K Q A 504
R.norvegicus     EHI T T L A Y Q E L P T A D L M Q E W G D A V Q Y N P A I I K V K V E P L Y E L V T A T D F A Y S S T V R Q N M K Q A 491
G.gallus         EHI T T L A Y Q E L P T A D L M Q E W G D A V Q Y N P A I I K V K V E P L Y E L V T A T D F A Y S S T V R Q N M K Q A 486
M.domestica      PSITSLAYKRLPNKELMEKWGDAAKLFP E VLKIKVIPMHEMWTGQNF INRHLKKNM RSA 535
                  : : * . . * * . : * . * * : . : * * * * : : : * : : : * : : * :

```

EGF domain (519-547)

```

O.mykiss          LAEYLEESSSRCR CAPCRNNGLA VLKGRTRCE CVCPVSGYSGLGCEITQR P-DIGIDGSWSCW 547
P.olivaceus      LSEYLAESSSRCR CAPCHNNGVA VLKGRTRCDVCPVSGYSGLGCEITQR KQIATDGSWSCW 548
D.rerio          LSEYLAESSSRCR CAPCLNNGLA VLKGRMTRCICPAGVKGVSCEITQR K-GLAIDGNWSCW 522
H.sapiens        LEEFQKEVSSCHCAP CQGNVGVKGRSRDCDCIPVGSQGLACEVSYR K-NTPIDGKWNWCW 551
M.mulatta        LEEFQKEVSSCHCAP CQGNVGVKGRSRDCDCIPVGSQGLACEVSYR K-NIPTDGKWNWCW 551
O.cuniculus      LEEFQKEVSSCHCAP CQGNVGVKGRSRDCDCIPVGSQGLACEVSYR K-NTPIDGKWNWCW 551
C.familiaris     LEEFQKEVSSCHCAP CQGNVGVKGRSRDCDCIPVGSQGLACEVSYR K-NTPIDGKWNWCW 551
S.scrofa         LEEFQKEVSSCHCAP CQGNVGVKGRSRDCDCIPVGSQGLACEVSYR K-NTPIDGKWNWCW 563
R.norvegicus     LEEFQMEVSSCR CAPCRNNGVPIKGRTRCE CIPAGFQGVACEVTNR K-DIPIDGKWNWCW 550
G.gallus         LDEFQLETSSCR CAPCQGNVGVKGRTRCE CIPVGSQGLACEVSYR K-DAAINGNWGCW 545
M.domestica      VKEYYEETHVCHCAP CLGNFPPI IRDSYCE LCGMGTGICISCELTSE--GDATGSWSCW 593
                  : * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * :

```

```

O.mykiss          GSWSPCR-GRSKTRS RQC NNPPVSDGGLACRGLQMETTDC F----- 587
P.olivaceus      GAWSSCS-GRKMSRS RQC NNPPVSDGGLACRGLQESTDC F----- 588
D.rerio          SSWSSCS-GKIQRTR RQC NNPPVSDGGLACRGLQESTDC F----- 562
H.sapiens        SNWSSCS-GRKTRR RQC NNPPVSDGGLACRGLQESTDC F----- 591
M.mulatta        SSWSSCS-GGRKTR RQC NNPPVSDGGLACRGLQESTDC F----- 591
O.cuniculus      SRWSSCS-GGQKTR RQC NNPPVSDGGLACRGLQESTDC F----- 590
C.familiaris     SNWSPCS-GGHKTR RQC NNPPVSDGGLACRGLQESTDC F----- 590
S.scrofa         SDWFSVFLQDEGSATIHLLKMGVAPAWVLLQKHLTVKEGRASSG 611
R.norvegicus     SDWSPCS-GGRKTR RQC NNPPVSDGGLACRGLQESTDC F----- 588
G.gallus         ASWSPCS-GGRTRR RQC NNPPVSDGGLACRGLQESTDC F----- 584
M.domestica      TSWTECIEAKQK-RERLCNSPIPG-SIKT PGRQKEETS C----- 631
                  * . . . : . . :

```

Figure 49 (continued)

<i>S. scrofa</i>	----- <b>MLVPRAAPLLTLLLATGSLGQRAQRPPRRPSP</b> -----	32
<i>C. familiaris</i>	-----MLTPWTALLLTLTLAGGSVSQRARRPPRPASR-----	32
<i>H. sapiens</i>	-----MLPPGTATLLTLLLAAGSLGQKPQRPRRPASP-----	32
<i>O. cuniculus</i>	-----MVLGRVLLAVLLAAGSLGRWAQKPRGAPSA-----	32
<i>M. musculus</i>	-----MLSPGAVLFFTLTLLTASSLGQRTRKPIGSTSP-----	32
<i>R. norvegicus</i>	-----MLSPGAVLFFTLILMASSLGQRTRKPYGSTSP-----	32
<i>G. gallus</i>	MNPVARPAGARPAPLPGTGLRCAMAAPRALMLLSRAPRRAAGDGNSSGRP-LPTAL-----	54
<i>D. rerio</i>	-----MIRFWLYL-FVFLA <b>CL</b> SFWPEVETR-ARYKPEPPK-PKKTETQKA	43
<i>O. mykiss</i>	-----MTGVVQ <b>CV</b> TMLMV <b>CV</b> CLWGSAAEAIIGAKSRPRPQRPPKPKVNP	46
:		
<i>S. scrofa</i>	ISTI <b>QPKAS</b> FDAQQFAGTWLLVAVASS <b>CRFLQEQGHRAEATSLHVA</b> -PQGAA--MAVSTF	89
<i>C. familiaris</i>	ISTI <b>QPKAN</b> FDAHQFAGTWLLVAVASS <b>CRFLQEQGHRAEATLLHVA</b> -PQGAD--MAVSTF	89
<i>H. sapiens</i>	ISTI <b>QPKAN</b> FDAQQFAGTWLLVAVGSAC <b>CRFLQEQGHRAEATTLHVA</b> -PQGT--MAVSTF	89
<i>O. cuniculus</i>	ISAI <b>QPKAN</b> FDAQQFAGTWLLAAVGSACH <b>FLQEQGHRAEATALHVA</b> -PQGAA--MAVSTF	89
<i>M. musculus</i>	ISTI <b>QAQVNF</b> SAQKFAGTWLLVAVGSS <b>CRFLQEQGHRAEATTLHAA</b> -PQGAA--MAVSTF	89
<i>R. norvegicus</i>	ISTI <b>QAQAN</b> FNAQQFAGTWLLVAVGSAC <b>CRFLQEQGHRAEATTLHIA</b> -PQGAA--MAASTF	89
<i>G. gallus</i>	LRKVVTEGNLSL <b>GELVGRWFLVGVASRCS</b> SYLAENSHRLEATAMTVAVPDGQS--LAISTF	112
<i>D. rerio</i>	IDTLAPGQINIDQMSGKWHL <b>LTVASRCKNLL</b> ESGFKTESTSLTWNITAD---TVTVGTV	100
<i>O. mykiss</i>	IDDT <b>PPAQ</b> NIDIQQMGGPWYLVNAAS <b>KCNFLMKNGLKVEATVMTL</b> TSPSSQNPTLSVSTT	106
: . . . . . : : * * * * * * * . . . : * * :		
<b>Lipocalin domain (48-184)</b>		
<i>S. scrofa</i>	RKLDGICWQVRQLFRDTGLPGRFLLQARGARGAVDVVVGETDYRSF <b>AILYLERARQLSVK</b>	149
<i>C. familiaris</i>	QKLDGICWQVRQLYRDGEV <b>LG</b> GRFLLQARGARGAVNMVVG <b>ETDYQG</b> F <b>AILYLEQKRQLSVK</b>	149
<i>H. sapiens</i>	RKLDGICWQVRQLYGD <b>TG</b> VLGRFLLQARGARGAVNVV <b>VAETDYQS</b> F <b>AVLYLERAGQLSVK</b>	149
<i>O. cuniculus</i>	RKLDGICWQVSQRYGATGV <b>PGRFLL</b> PARGPRGAVHV <b>VAAETDYHS</b> F <b>AVLYLERARQLSVK</b>	149
<i>M. musculus</i>	RKLDGICWQVRQLFENTGV <b>PGRFL</b> FQVSRARGPVHM <b>VVAETDYQS</b> F <b>AILYLEQGRKLSVK</b>	149
<i>R. norvegicus</i>	RKLDGICWQVRQLYGITGV <b>PGRFLL</b> QAPRARGPVHV <b>VVAETDYQS</b> F <b>AILYLEQARRLSVK</b>	149
<i>G. gallus</i>	RKLDGQCWEIRQRV <b>YV</b> PEG <b>AHRRF</b> SVRGRG <b>YN</b> SKMEVV <b>VGEADPR</b> SY <b>AI</b> IYQDS <b>QGLSVK</b>	172
<i>D. rerio</i>	RKLN <b>FVC</b> WEIKQ <b>NYM</b> TKTP <b>Q</b> LF <b>LK</b> GRPSDNVDIM <b>VLETDY</b> STY <b>AMLVFKRAEKITMK</b>	160
<i>O. mykiss</i>	TRLNHQ <b>CWE</b> ILQAYTIT <b>PT</b> PGRLVLNGSR <b>PL</b> LNTD <b>IVIGES</b> SDYSS <b>YAVFYYQKQGLTMK</b>	166
: * : * * : * : : : : : * * : * * : : * * :		
<i>S. scrofa</i>	<b>LYARSLP-VSESALS</b> VFEQ <b>R</b> VQ <b>GANL</b> TEDHILFFPKYGF <b>CE</b> AA <b>DQ</b> FHV <b>LDEAKQ</b> -	202
<i>C. familiaris</i>	<b>LYARSLP-PSDSALS</b> SAFEQ <b>R</b> IQRV <b>N</b> L <b>TEDH</b> VLF <b>FP</b> KYGF <b>CE</b> AA <b>DQ</b> FHV <b>LDEAGR</b> -	202
<i>H. sapiens</i>	<b>LYARSLP-VSDSVLS</b> GFQ <b>R</b> VQ <b>E</b> AHL <b>TEDQ</b> I <b>F</b> Y <b>FP</b> KYGF <b>CE</b> AA <b>DQ</b> FHV <b>LDEVRR</b> -	202
<i>O. cuniculus</i>	<b>LYVRS</b> LP-VSDSV <b>LGA</b> FQ <b>R</b> V <b>AQ</b> AN <b>L</b> T <b>QD</b> V <b>LF</b> FP <b>TY</b> GF <b>CE</b> AA <b>DQ</b> FH <b>LDEVRR</b> -	202
<i>M. musculus</i>	<b>LYVRS</b> LP-VNDSV <b>L</b> DF <b>FERR</b> V <b>REAN</b> L <b>TEDQ</b> I <b>L</b> FF <b>FP</b> KYGF <b>CE</b> T <b>ADQ</b> LH <b>LNEVPR</b> -	202
<i>R. norvegicus</i>	<b>LYTR</b> TL <b>P-VSDS</b> AL <b>NA</b> F <b>EER</b> V <b>RGAN</b> L <b>TEDQ</b> I <b>F</b> FF <b>FP</b> KYGF <b>CE</b> T <b>ADQ</b> FH <b>ILNEMPK</b> -	202
<i>G. gallus</i>	<b>LYGRSSQ-LSNAI</b> V <b>DK</b> F <b>EQ</b> RA <b>VGL</b> SE <b>DV</b> TH <b>YFP</b> TYGF <b>CD</b> S <b>ADDF</b> H <b>ILDETEL</b> -	225
<i>D. rerio</i>	<b>LYGRS</b> GE-V <b>PDNI</b> V <b>DK</b> F <b>EDRA</b> KT <b>FN</b> L <b>GLD</b> V <b>VF</b> Q <b>PDY</b> GF <b>CE</b> SA <b>EK--</b> V <b>LDL</b> T---	209
<i>O. mykiss</i>	<b>LYGR</b> SK <b>DTL</b> SE <b>AI</b> L <b>DK</b> F <b>EDL</b> A <b>EKK</b> GL <b>GLAY</b> V <b>FA</b> FP <b>NY</b> SH <b>CE</b> SV <b>DK</b> D <b>HVIN</b> <b>CVPTC</b>	221
* * * : : : . * * * * * * * * * . . . : : :		

Figure 50 Alignment of the C8G proteins among species according to the GenBank accession no. as given in table 13 using the ClustalW2 tool. The amino acids are enumerated above each line on the right side. The identical, similar and missing residues are indicated with asterisks (\*), points (.), and dashes (-), respectively. Signal peptide in pig C8G protein is in bold and italic. The amino termini of C8G mature protein starts just after the signal peptide. The cysteine residues are highlighted. The lipocalin functional protein domain is in continuous underline. The glycosylation site presents in shade blocks with white bold font using the NetNGlyc 1.0 Server tool.

```

R.norvegicus      MLLRTPGLPRRSGMASGVTITLAI AIFALEINAQAPEPTPREEPSAD--ALLPIDCRMST 58
M.musculus        MLCKPPGLPRRSMASGMAITLALAI FALGVNAQMPIPVSRREEQEQH--YPIPIDCRMSP 58
S.scrofa          -----MRVRRRFTFAVCILEISILSAGPTPSYHPKPVDCGNGTSPSPIDCRMSP 47
B.taurus          -----MSAGQRF AFAICILEISLLRAGPTPSYD--PAERQGTPLPIDCRMSS 45
M.mulatta         -----MSACWSFAAAICILEISVLTA EYTPSYDPQPTE SRGSASHIDCRMSP 47
H.sapiens         -----MSACRSFAVAICILEISILTAQYTTSYDPEL TESSGSASHIDCRMSP 47
E.caballus        -----MSAGRTF AFAICILEIVSVLTA GPTPNYAPEPEQQSGTPLPIDCRMSS 47
O.cuniculus       -----MAASHSF AFAVCVLEIGALTA GPTPSYVHEPIQRSDPLQPIDCRMSP 47
M.domestica       -----MSLSETALN-----RPRREIETPAPIDCKLTS 27
                :::                               ***:::

                TSP1 domain (45-95)
R.norvegicus      WSQWSQCDPCLKQRFRRSRSM EVFGQFQGKSCADALGDRQHCEPTQEC EEEVQENCG-NDFQ 117
M.musculus        WSNWSECDPCLKQRFRRSRI LAFGQFNGKSCVDVVLGDRQGCEPTQEC EEI QENCG-NDFQ 117
S.scrofa          WGEWSRCDPCLKQMFRRSRI ETFGQFNGQKCVDAVGDRRQCVPT EPCEDLEEDCG-SDFQ 106
B.taurus          WSEWSKCDPCLKQMFRRSRI EIFGQFNGRKCVDVAVGDRRQCVPT EACEDPEEGCG-NDFQ 104
M.mulatta         WSEWSQCDPCLRQMFRRSRI EVFGQFNGKSC TDAVGDRRQCVPT EPCEDAEDDCG-NDFQ 106
H.sapiens         WSEWSQCDPCLRQMFRRSRI EVFGQFNGKRCTDAVGDRRQCVPT EPCEDAEDDCG-NDFQ 106
E.caballus        WSEWSECDPCLRQMFRRSRI EVFGQFNGQR CVDVAVGDRRQCVPT EACEEVEDDCG-NDFQ 106
O.cuniculus       WSEWSKCDPCLRQMFRRSRI EVFGQFNGKSCVDALGDRRACI PTEACEEVEDDCG-NDFQ 107
M.domestica       WSEWSSCEPCQKEMYRSRTI EAFGQFGGKRCLHSLGDRRSCEPFRACDDEEYDCE-NDFK 86
                *.:** *:** :: :****:  **** *: * . :****: * * . **: : . * .:.*

                LDLa domain (100-136)
R.norvegicus      CETGRCIKRLLCNGDND CGDFSDESDCESDPRLP CRDRVVEESELGR TAGYGINILGMD 177
M.musculus        CETGRCIKRLLCNGDND CGDYSDENDCDDDPRTPC RDRVAEESELGLT AGYGINILGME 177
S.scrofa          CGTGRCIKRLLCNGDND CGDFSDEDDCDS DPRPPCRERVVEESELARTAGYGINILGMD 166
B.taurus          CGTGRCIKRLLCNGDND CGDYSDENCEQDPRPP CRNRVVEESELARTAGYGINILGMD 164
M.mulatta         CGTGRCIKRLLCNGDND CGDFSDEDDCEGDPRPP CRDRVVEESELARTAGYGINILGMD 166
H.sapiens         CSTGRCIKRLLCNGDND CGDFSDEDDCESEPRPP CRDRVVEESELARTAGYGINILGMD 166
E.caballus        CGTGRCIKRLLCNGDND CGDFSDEDDCENDPRPP CRERVVEESELARTAGYGINILGMD 166
O.cuniculus       CGTGRCIKRLLCNGDND CGDFSDEDDCETEPRLT CRNREVQ ESELARTAGYGINILGMD 167
M.domestica       CETGRCIKRLLCNGDND CGDFSDEDNCEK DPRSPCHTD-VEMSELGR TAGYGMNILGMD 145
                * ***** :* ** *****:***.:*: :** .*: .: ***. ***:***:*****:

R.norvegicus      PLGTPFDNEFYNGLCDRVRD GNTLTYYRKPWNVAFLAYETKADKNFR TENYEEQFEMFKT 237
M.musculus        PLRTPFDNEFYNGLCDRVRDEK--TYRKPWNV VSLIYETKADKSFRTENYDEHLEVFKA 235
S.scrofa          PLTTPFDNEY YNGLCDRVRD GNTLTYYRKPWNVAAL IYETKVDKNFRTEYHERQIQVLKT 226
B.taurus          PLSTPFDNQY YNGLCDRVRD GNTLTYYRKPWNVA SLYDTKADKNFRTENHEESIQLRT 224
M.mulatta         PLSTPFDNEFYNGLCNRDRD GNTLTYYRRPWNVA SLYIYETKGEKNLRTEHYEEQIEAFKS 226
H.sapiens         PLSTPFDNEFYNGLCNRDRD GNTLTYYRRPWNVA SLYIYETKGEKNFRTEHYEEQIEAFKS 226
E.caballus        PLSTPFDNEY YNGLCDRVRD GNTLTYYRKPWNVA SLYETKADKNFRIEHYEQIQAFRS 226
O.cuniculus       PLATPFDNEY YHGLCDRVRD GNTLTYYRKPWNVA VLAYETKIDKNFRTEYEEQMQAFKS 227
M.domestica       PLDTPFDNEY FHLCDRVRD GNTGTYRKPWNVA TLN YDTKA EKRLRTE NYEEHVLQITD 205
                ** *****:***:***: * : * :***:***: * * :***: * * : * * : * . : . :

```

Figure 51 Alignment of the C9 proteins among species according to the GenBank accession no. as given in table 4 using the ClustalW2 tool. The amino acids are enumerated above each line on the right side. The identical, similar and missing residues are indicated with asterisks (\*), points (.), and dashes (-), respectively. Signal peptides in human C9 protein (DiScipio et al. 1984) and in pig C9 protein are in bold and italic. The amino termini of C9 mature protein starts just after the signal peptide. The cysteine residues are highlighted. The cysteine-rich functional protein domains TSP1 followed by LDLa, and MACPF are in continuous underline. The glycosylation sites in human (DiScipio et al. 1984) and in pig present in shade blocks with white bold font. The human C9 cleavage site into C9a and C9b by  $\alpha$ -thrombin is in bold and asparaginyl linked glycosylation site is boxed in bold (DiScipio et al. 1984)

<i>R. norvegicus</i>	IVDRRTTSFNANLALKFTITEAPIKK-VGVDEVSP--KNSSPKDSSVDFQFSYFKKEN	294
<i>M. musculus</i>	INREKTSNFDNADFAKFSATEVPEK---GAGEVSPA--EHSSKPTNISAKFKFSYFMGKN	290
<i>S. scrofa</i>	IIIEKKSNFNADLTIKFTPTAEIQLKSKNVELANE-ENSNP--MNNKAHFRFTYSKTET	283
<i>B. taurus</i>	IIIEKKLNFNAGLSVKYTPVEAIEKNCVDLEHSDKGSTSSPSKLAAEAKFRFTYSKDDI	284
<i>M. mulatta</i>	IVQEKTSNFDNADISLKFIPTEANK----VKTEKSSE-KQASSNSLRGQGSFRFSYSKNET	281
<i>H. sapiens</i>	IIQEKTSNFDNAAISLKFPTETN-----KAEQCCCE-ETASSISLHGKGSFRFSYSKNET	279
<i>E. caballus</i>	YIEERRSHFNADFTLKFTPTAEK-----KCKQPEPE-ESQNGTDSSENRIFRFAYSKNET	279
<i>O. cuniculus</i>	IIIEETSFNANLALKFTPTAEKAS---KAEEASPKNKSLLDNDKGFSSKQFSYSKNET	284
<i>M. domestica</i>	TFRERQKNFGFDISLKLSTEAPEG---GTPKRKPD---PKRNQQDAGLTFRFRYSKNES	259
	. : . * . : : * * . * . : : : * * * .	
<i>R. norvegicus</i>	FQRLSSYLSQTKKMFHLVHRGMIQLGRFVMNRNRMVLTTFVLDVVKALPVSYEKGEYFGFL	354
<i>M. musculus</i>	FRRLSSYFSQSKMFMVHLRGVQLGRFVMNRNDRVLRSTFLDDVKALPTSYEKGEYFGFL	350
<i>S. scrofa</i>	YKLLSSYSSKKEKIFLHVKGVIHLGRFVMRKRDRVMLTKTFLDDVKYLPSTYEKGEYFAFL	343
<i>B. taurus</i>	YRLSSYSYSAQEKMFHLVKGVIHLGRFVMRNRDVMVLTTFVLDVVKALPVSYEKGEYFAFL	344
<i>M. mulatta</i>	YQLFLSYSSKKEKMFHLVKGVIHLGRFVMNRNDRVLTTFVLDVVKALPVSYEKGEYFAFL	341
<i>H. sapiens</i>	YQLFLSYSSKKEKMFHLVKGVIHLGRFVMNRNDRVLTTFVLDVVKALPVSYEKGEYFAFL	339
<i>E. caballus</i>	YQLFLSYSSKKEKMFHLVKGVIHLGRFVMRNRDVMVLTTFVLDVVKALPVSYEKGEYFAFL	339
<i>O. cuniculus</i>	YQLFLSYSSQKEMFLLVKGVIHLGRFVMNRNRMVLTTFVLDVVKALPVSYEKGEYFAFL	344
<i>M. domestica</i>	LHLIKYYASDKTKMFLQVKGEIQLGRFHMNRNREFMLKSTFLDDVKALPTSYEKGEYFGFL	319
	: : * : . * : * : : * : : * : * * : * . * : * * : * : * : * : * : * : * : * : *	
<i>R. norvegicus</i>	ETYGTHYSSGSLGGLYELIYVLDKASMKKEKGVELSDVVKRCLGFNLDVSLYTPLOTALEG	414
<i>M. musculus</i>	ETYGTHYSTSGSLGGQYEIVYVLDKASMKKEKGVLDNDVKHCLGFNMD--LRIPLQDDLKD	408
<i>S. scrofa</i>	ETYGTHYSSGSLGGLYELIYVLDKATMTEKGIELRDVHRCLGFNLDLSLN-----FG	396
<i>B. taurus</i>	ETYGTHYSSGSLGGLYELIYVLDKKSMEQKDIELRDVQRCLGFNLDLSLN-----VG	397
<i>M. mulatta</i>	ETYGTHYSSGSLGGLYELIYVLDKASMNKRGVLEKDVVKRCLGYHLDVSLDF-----SK	395
<i>H. sapiens</i>	ETYGTHYSSGSLGGLYELIYVLDKASMKRKGVELKDIKRCLGYHLDVSLAF-----SE	393
<i>E. caballus</i>	ETYGTHYSSGSLGGLYELIYVLDKASMDQKGVLELRIQRCLGFNLDLSLN-----DK	392
<i>O. cuniculus</i>	ETYGTHYSSGSLGGRYELIYVLDKASMKKEKGIELNDIKKCLGFNLDLSLNIPGK--SAG	402
<i>M. domestica</i>	ETYGTHYSSGNIIGKYEELIYVLDKAEEMQRKGLEIQDVRKCLGLDLDLSYQS-----A	372
	*****:*. : * * : : * * * * * * . * . : : * : : * * * . : *	
<b>MACPF domain (297-507)</b>		
<i>R. norvegicus</i>	PSLTANVNHSDCLKTGDGKVNISRDHIIDDDVISFIRGGTRKQAVLLKEKLLRGAKTIDV	474
<i>M. musculus</i>	ASVTASVNADGCIKTDNGKTVNITRDNIIDDDVISFIRGGTREQAAILLKEKILRGDKTFDK	468
<i>S. scrofa</i>	VEIKGKIDSENCLKRGDGKTENIMNDDFIDDDVISFIRGGTRKYATELKEKLLKGAKMINV	456
<i>B. taurus</i>	VEVTGNFDSKLSCKKGMGQETETNPADLFDVITFIRGGTRKYATELKEKLLRGARMINV	457
<i>M. mulatta</i>	ISAGAKADKDDCVKRGEGRAVNITSDHLIDDDVISLIRGGTRQYAFELKEKLLRG-TMIDV	454
<i>H. sapiens</i>	ISVGAEFNKDDCVKRGEGRAVNITSENLIIDDDVSLIRGGTRKYAFELKEKLLRG-TVIDV	452
<i>E. caballus</i>	YEVTAIKIDKNDCLKRNEKEIVNIMDGSLLIDDDVISLIRGGTRKYAFELKEKLLGAKTVNV	452
<i>O. cuniculus</i>	LSLTGQANKNNCLKSGHNAVNITRANLIIDDDVISLIRGGTRQKFAFELKEKLLTKAKMVDV	462
<i>M. domestica</i>	VNFEANIKGSDCCSVN--WKHDNERKHIIDDDVISLIEGGTREYATNLKEKLLRGPVKVDV	430
	. . . . * . . . : : * * : : * . * * : * * * * * * : *	
<i>R. norvegicus</i>	NDFINWASSLDDAPALISQKLSPIYNLIPLTMKDAYAKKQNMKEAIEDYVNEFSARKCYP	534
<i>M. musculus</i>	TDFANWASSLANAPALISQRMSP IYNLIPLKIKDAYIKKQNLKAVEDYIDFSTKRCYP	528
<i>S. scrofa</i>	TDFVNWASSLNDAPVLINQKLSPIYDLIPVKLNDAHLKRQNLERAIEDYINEFNVKMP	516
<i>B. taurus</i>	TDFVNWASSLNHAPVLISQKLVPIYDLIPVKMKDAHLKQNLERAIEDYINEFSVRKQCP	517
<i>M. mulatta</i>	TDFVNWASSINDAPVLISQKLSPIYNLVPVKMKNAHLKQNLERAIEDYINEFSVRKCHS	514
<i>H. sapiens</i>	TDFVNWASSINDAPVLISQKLSPIYNLVPVKMKNAHLKQNLERAIEDYINEFSVRKCHT	512
<i>E. caballus</i>	TDFVNWASSLNDAPVLISQRLSPIYNLIIPVKMKDAHQQKQNLERAIEDYINEFSVRKCHP	512
<i>O. cuniculus</i>	TDFINWASSLSDAPVLINQKLSPIYNLIIPVKIKDAHQQKQNLERGIEDYINEFSTKCKSP	522
<i>M. domestica</i>	TDFLNWASSLNDAPVLINQKLLPIGNLVPVRMENAHEKKQNLQAINDYVDQFNHCKQCP	490
	. * * * : * : . * * . * * : * * : * * : * * : * * : * * : * * : * * : * * : *	
<i>R. norvegicus</i>	CQNGGTAILLDGQCMCSCSTIKFKGIACEISKQR-----	567
<i>M. musculus</i>	CLNGGTIILLDGQCLCSCPMFMFRGMACEIHQKI-----	561
<i>S. scrofa</i>	VPEWRSDSAGWTVFVFLPKQISGNCL-----	543
<i>B. taurus</i>	CQNGGTIVLLDGEVCVCSCPKFEKGVACEIKK-----	548
<i>M. mulatta</i>	CQNGGTAILMDGKCLCTCPFKFEGIAACEISKQKQVSEGLPALDFPREK	561
<i>H. sapiens</i>	CQNGGTIVILMDGKCLCACPFKFEGIACEISKQKISEGLPALEFPNEK	559
<i>E. caballus</i>	CQNGGTIVIQIDGQCLCSCPIAFEGIAACETGKKKIS-----	547
<i>O. cuniculus</i>	CQNGGTALLMDGQCLCTCPFMFEGIAACEISKRKL-----	557
<i>M. domestica</i>	CQNEGTTMLLDGECICACTGFQGVACQISPVV-----	523
	: . . . : * .	

Figure 51 (continued)



## Acknowledgements

It is really a great opportunity to express my deepest thank to all persons and organizations those handed me the chance and possibility to complete this thesis:

- Prof. Dr. Klaus Wimmers, my supervisor, for his valuable guidance and discussion, constructive suggestion and advice on working,
- Dr. Eduard Muráni and PD. Dr. Siriluck Ponsuksili for their recommendation and kind help,
- Hannelore Tychsen, Angela Garve and Marlies Deutscher for their excellent technical assistance in laboratory,
- Dr. Ronald Marco Brunner for supporting computer technique,
- Jana Masermann for her kind helps along the educational course,
- Anja Hartman and Dr. Alexander Rebl for reading manuscript,
- Mr. Philipp Westphal for supporting German language

All are working under the same house-roof of Molecular Biology Research Unit, Research Institute for the Biology of Farm Animals (FBN-dummerstorf), Germany

- Prof. Dr. Karl Schellander (Institute of Animal Science, Faculty of Agriculture, University of Bonn) and Prof. Dr. Jéns Léon (Institute of Crop Science and Resource Conservation, Faculty of Agriculture, University of Bonn) for their painstaking comments and advices on the manuscript,
- Dr. Le Viet Dung, Asso. Prof. Dr. Vo Van Son, Dr. Bui Xuan Men and Msc. Nguyen Minh Thong (Can Tho University, Viet Nam) for promoting me to PhD student position of the BMBF/DLR scholarship,
- Dr. Huynh Thanh Tung (Laboratory of Plants Molecular Genetics, Phytology Department, Laval University, Canada) for his fruitful discussions, encouragement and always willing to help,
- Prof. Dr. Wolfgang Büscher (Institute of Agricultural Engineering, Faculty of Agriculture, University of Bonn) and Prof. Dr. Karl-Heinz Südekum (Institute of Animal Science, Faculty of Agriculture, University of Bonn) for kindly being my co-examiners in animal production and animal nutrition,

- Ms. Iris Pützer and Ms. Matina Hohensee (Faculty of Agriculture, University of Bonn) for kindly help to finish necessary procedures for the PhD course,
- Dr. Nguyen Thi Dieu Thuy (Institute of Biotechnology, Viet Nam) for her helping to treat samples in laboratory,
- Dr. Nguyen Van Duc (National Husbandry Institute, Viet Nam) and Ms. Dao Thi Uyen (Nha Uyen Company, Lao Cai, Viet Nam) for their help to collect sample of Muong Khuong animals,
- Mr. Nguyen Hoang Liet 's family (Nielhagener Str.69, 22147 Hamburg, Germany) and Mr. Ngo Tuan Anh 's family (Doberaner Str 59, 19890 Rostock, Germany) for their help, support, interest as well as sharing all in life during my interesting PhD course in Germany,
- Dr. Nguyen Chi Ngon (Can Tho University, Viet Nam), Dr. Bui Viet Duc (University of Agriculture 1, Viet Nam), Vietnamese Rostockers and colleagues (FBN-dummerstorf) for sharing the great student life in Germany,
- The Federal Ministry of Education and Research (BMBF) and the Project Management Agency in the German Aerospace Center (DLR) for granting a full 3 year-scholarship,
- Research Institute for the Biology of Farm Animals (FBN-dummerstorf) for supporting living and study conditions,
- The Faculty of Agriculture, University of Bonn for supporting study and graduation conditions,
- Last not so much but certainly not least to send special thanks to beloved persons in my large family including my external grandmama, my parents, my parents-in-law, my elder brother as well as my younger sister, especially my wife for support, encouragement, everlasting love and everything great or small during these years.

AGAIN THANK YOU WITH ALL MY HEART  
ALWAYS REMEMBER YOU

## Curriculum Vitae

Family name : Do  
First name : Vo Anh Khoa  
Date of birth : 25 May 1975  
Place of birth : Tien Giang, Viet Nam  
Nationality : Vietnamese  
Status family : married, one daughter  
Address : 42 Ly Tu Trong street, Can Tho City, Viet Nam  
Email : dvakhoa@ctu.edu.vn, dvakhoa@yahoo.com



## Education

- 1981-1990 Tru Van Tho I Primary and Secondary School (Cai Lay, Tien Giang, Viet Nam)
- 1990-1992 Doc Binh Kieu High School (Cai Lay, Tien Giang, Viet Nam)
- 1992-1997 Bachelor of Science in Animal Husbandry and Veterinary Medicine, College of Agriculture and Applied Biology, Can Tho University, Viet Nam.
- 2001-2003 Master of Science in Animal Sciences with the thesis '*Evaluation on the reproductive capacity of cross-bred and pure-bred female pig herd at Phuoc Tho Breeding Enterprise*' carried out at College of Agriculture and Applied Biology, Can Tho University, Viet Nam.
- 2004-2008 PhD student in Animal Genetics with the thesis '*Molecular characterization of porcine genes encoding complement components of the terminal lytic pathway and their association with hemolytic complement activity*' conducted at Molecular Biology Unit, Research Institute for the Biology of Farm Animals (FBN-dummerstorf) and Faculty of Agriculture, University of Bonn, Germany.

## Professional Experiences

- 1997-1999 Researcher in Department of Animal Sciences, College of Agriculture and Applied Biology, Can Tho University, Viet Nam.

- 1999-2001 Soldier in Can Tho Army, Viet Nam
- 2003-present Lecturer in Department of Animal Sciences, College of Agriculture and Applied Biology, Can Tho University, Viet Nam.

## **Publications**

- Do HQ, Son VV, **Do VAK** and Khang NTK (1999): Urea supplementation of rice straw for Sindhi x Yellow cattle; sprayed in solution, as a soft cake or hard block. *The International Journal for Research into Sustainable Developing World Agriculture* 11
- Yamasaki S, Manh LH, Takada R, Men LT, Dung NNX, **Do VAK**, Taniguchi T (2002): Admixing synthetic antioxidants and sesame to rice bran for increasing pig performance in Mekong Delta, Viet Nam. In: *Proceedings of the 2002 Annual Workshop of JIRCAS Mekong Delta Project*, 176-179
- Yamasaki S, Manh LH, Takada R, **Do VAK**, Men LT, Dung NNX (2002) Effect of admixing synthetic antioxidants and sesame with rice bran on daily weight gain, feed intake and feed efficiency of swine in Mekong Delta, Viet Nam (Poster paper). Edited by Mori Y, Hayashi T, Highley E. *The 9th JIRCAS International Symposium Value -`Addition to Agricultural Products´*, Tsukuba, 16-17 October, 168-169
- Men LT, Yamasaki S, Caldwell JS, Thuan TK, Yamada R, **Do VAK**, Takada R (2003): Effect of rice based balanced resources and the water hyacinth (*Eichhornia crassipes*) in diets on pig production in the Mekong Delta of Viet Nam. In: *Proceedings of the 2003 Annual Workshop of Jircas Mekong Delta Project*, Viet Nam, 164-172
- Do VAK**, Ponsuksili S, Murani E, Huynh TPL, Wimmers K (2006): Molecular characterization of cDNA sequences of porcine genes encoding complement component of the terminal lytic pathway. In: *Proceedings of International Workshop on Biotechnology in Agriculture*. Nong Lam University, Ho Chi Minh city, Viet Nam, 20-21 October, 29-31

- Do VAK**, Ponsuksili S, Murani E, Wimmers K (2006): IMpRH mapping of the porcine complement genes of the terminal lytic pathway. In: Proceedings of International Workshop on Biotechnology in Agriculture. Nong Lam University, Ho Chi Minh city, Viet Nam, 20-21 October, 32-34
- Wimmers K, **Do VAK**, Murani E, Ponsuksili S (2006): Sus scrofa complement component C6 mRNA, complete cds: GenBank accession number DQ333199. NCBI homepage
- Wimmers K, **Do VAK**, Murani E, Ponsuksili S (2006): Sus scrofa complement component C8A mRNA, complete cds: GenBank accession number DQ333200. NCBI homepage
- Wimmers K, **Do VAK**, Murani E, Ponsuksili S (2006): Sus scrofa complement component C8B mRNA, complete cds: GenBank accession number DQ333201. NCBI homepage
- Wimmers K, **Do VAK**, Murani E, Ponsuksili S (2006): Sus scrofa complement component C8G mRNA, complete cds : GenBank accession number DQ333202. NCBI homepage
- Wimmers K, **Do VAK**, Murani E, Ponsuksili S (2006): Sus scrofa complement component C9 mRNA, complete cds: GenBank accession number DQ333198. NCBI homepage
- Do VAK**, Murani E, Ponsuksili S, Wimmers K (2007): Genetic variation in C8B complement gene between Muong Khuong and European pig breeds as a potential source to improve animal health. In: Proceeding of 5<sup>th</sup> Vietnamese-Hungarian Conference on Animal Production and Aquaculture. Can Tho University, 123-129
- Do VAK**, Murani E, Ponsuksili S, Wimmers K (2007): Polymorphism and association with hemolytic complement activity in the pig C7 gene. Aus der Arbeit der Forschungsstätten für Tierproduktion. Jahrestagung der DGfZ. Universität Hohenheim, 26-27 September, A01