LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

STRUCTURAL AND FUNCTIONAL CHARACTERISATION OF A NOVEL BOVINE β-DEFENSIN GENE CLUSTER



This thesis is submitted to the University of Dublin for the degree of Doctor of Philosophy

August 2013

Fernando Narciandi Martin

School of Biochemistry and Immunology
Trinity College Dublin

Supervisor: Professor Cliona O'Farrelly

TABLE OF CONTENTS

Declaration 5

Acknowledgements 6

Publications arising from this project: 7

Associated publications 7

Abbreviations 8

List of Figures 11

List of Tables. 13

Abstract 14

CHAPTER 1: General introduction 16

1.1. Innate immunity 17

1.2. Antimicrobial peptides 18

1.2.1. AMP gene and protein structure 22

1.2.2. Antibacterial activity of AMPs 23

1.2.3. Antiviral activity of AMPs 28

1.3. Emerging roles for Antimicrobial Peptides 29

1.4. Defensins 33

1.5. β-Defensins and Reproduction 37

1.8 Bovine breeding 41

1.9 Bovine AMPs 42

1. 10 Aims of the study 49

CHAPTER 2: Materials and methods 50

2.1. Sample collection 51

2.2. RNA isolation 52

2.3. cDNA synthesis 53

2.3. Primer design 53

2.4. Quantitative real-time polymerase chain reaction (qRT-PCR) 59

2.5. Polymerase chain reaction (PCR) 59

2.6. PCR product isolation 60

2.7. gDNA extraction, DNA Sequencing and SNP discovery 60

2.8. SNP genotyping and cluster analysis 61

2.9. Prediction of glycosylation sites 61

2.10. Promoter region bioinformatic analysis 62

2.11. Restriction enzyme digestion 62

2.12. Ligation reaction 63

TRINITY COLLEGE
- 4 MAR 2014
LIBRARY DUBLIN

Thesis 10313

2.13. Bacterial cell transformation 64
2.14. Maxiprep and Miniprep plasmid preparations 64
2.15. Luciferase reporter gene assay for *BBD126* and *LAP* 65
2.16. Sperm penetration assay 68
2.17. Statistical analysis 70
2.18. Antigen preparation for Alpaca (*Vicugna pacos*) immunisation 70
2.18. Alpaca (*Vicugna pacos*) immunisation 71
2.19. Alpaca (*Vicugna pacos*) white cell isolation 71
2.20. Alpaca (*Vicugna pacos*) single domain antibodies (VHH) gene isolation and cloning 72
2.21. Phage rescue 73
2.22. Phage selection by panning 74
2.23. VHH production 76
2.24. VHH Isolation 77
2.25. Sortase production 79
2.26. Sortase reaction 79
2.27. Immunoprecipitation 80
2.28. Silver staining 81
2.29. Construction of artificial VHH gene 81
2.30. Bovine TNF alpha ELISA 83
2.31. Custom production of anti-BBD126 monoclonal antibody 85
2.32. Immunocytofluorescence of BBD126 in sperm cells 86
2.33. Recombinant production of BBD126 87
2.34. Western Blot analysis of recombinant BBD126 89
2.35. Localization of BBD126 in the bovine male reproductive tract tissue. 89
CHAPTER 3: mRNA expression profiling of Defensin gene cluster 91
3.1. Introduction 92
3.2. Results 96
3.2.1. RNA extraction 96
3.2.2. Normaliser gene selection qPCR 96
3.2.3. Novel β-defensin gene expression 98
3.2.4. Regional specific β-defensin gene expression in male reproductive tract 100
3.2.5. Regional specific β-defensin gene expression in female reproductive tract 108
3.2.6. Overview of novel β-defensins in the reproductive tract 112
3.2.7. Elongated tail of β-defensins as target for glycosylation 115
3.2.8. Principal components analysis of β-defensins promoter region 118
3.2.9. β-Defensin gene expression response to LPS and hormones. 121

3.3. Discussion 124
CHAPTER 4: Genetic analysis of a bovine β-defensin cluster 129
4.1. Introduction 130
4.2. Results 132
4.2.1. SNP discovery 132
4.2.2. Genotyping of Holstein-Friesian and Norwegian red populations 133
4.2.3. Haplotype analysis of bovine Chr13 beta-defensin gene cluster 136
4.2.5. Effect of SNPs on sperm penetration of cervical mucus. 143
4.3. Discussion 147
CHAPTER 5: Development of research tools 151
5.1. Introduction 152
5.2. Results 159
5.2.1. Comercial production of anti-BBD126 monoclonal antibody 159
5.2.2. Anti-BBD126 clone selection. 159
5.2.3. BBD126 location in sperm cell surface. 160
5.2.4. BBD126 location in male reproductive tract. 162
5.2.5. Anti-BBD126 antibody characterisation by immunoprecipitation. 163
5.2.6. Recombinant BBD126 production in an eukaryotic expression system 165
5.2.7. Production of heavy chain-only antibodies from a complex protein preparation 169
5.2.8. Production of sortase 174
5.2.9. Sortase labelling of yeast library VHH 176
5.2.10. Characterisation of specificity of yeast library VHH by immunoprecipitation. 177
5.2.11. PCR generation of a heavy chain antibody library. 179
5.2.12. Analysis of anti-bovine TNFα VHH by ELISA. 183
5.2.13. Characterisation of specificity of anti-bovine TNFa VHH by immunoprecipitation. 185
5.3. Discussion 189
CHAPTER 6: Final discussion 196
6.1. Gene expression and regulation of epididymal β-defensins 197
6.2. Genetic variation on epididymal β-defensins 200
6.3. Tool development for functional analysis of Epididymal β-defensins 202
Appendix 206
References 220

DECLARATION

I hereby certify that this thesis has not been previously submitted for examination to this or any other University. The work described in this thesis has been carried out by the author alone except where stated.

The author gives permission for the library to lend or copy this work upon request

ACKNOWLEDGEMENTS

I would like to acknowledge and thank everybody involved in one way or another with this project. Firstly, I thank my supervisor Professor Cliona O'Farrelly; this project is the result of Cliona's enthusiasm and energy. I could not have gone this far without the opportunity she gave me so many years ago. I would also like to express my gratitude to Dr Kieran Meade for his guidance, support and, like Cliona, for the opportunities that have allowed me to work in science. My thanks as well to past and present members of the Comparative Immunology group. To Dr Andrew Lloyd, for being one of the most instructive persons I have ever met; a single conversation with Andrew will teach you more than ten lectures. Special thanks to Andrew for reading and correcting my terrible Spanglish over all these years. I also want to thank Dr Aspinas Chapwanya, Dr Ronan Shaughnessy and Dr Cormac Whelehan, the three musketeers of the bovine/chicken section of our group, for sharing the experience of doing research with me. I want to thank Dr Nollaig Bourke, Dr Aideen Collins, Aoife Kelly, Ronan Fahey, Sarah Whelan, Catherine Keogh, Paul Cormican, Tatyana Devine-Lysakova, Nigel Stevenson, Liz Ryan, Emma McGrath, Sarah Cahalane, Anne Barry-Reidy, Uma Thiruchelvam and Dr Troy Hibbard. I will not have enough paper to thank everyone for all their help over the years; I just want to you to know how grateful I am to all of you. I would like to thank as well Prof Hidde Ploegh of MIT, Boston, for all the direction, help and his patience and to Dr Max Popp and Dr Juan José Cragnolini for their expertise and help with the VHH project and mass spectrometry.

I want to thank my parents Fernando and Lidia for their support over many years as well as my brother Sergio, my aunt Teresa, my grandparents Fernando, Aurora, Pilar and Sergio.

To Neil, I could not have done any of this without you. Thanks for your support, patience and for making me laugh. Gracias.

PUBLICATIONS ARISING FROM THIS PROJECT:

- Reproductive tissue-specific expression profiling and genetic variation across a 19 gene bovine β-defensin cluster. **Narciandi F, Lloyd AT, Chapwanya A, O' Farrelly C, Meade KG**. Immunogenetics. 2011 Oct.
- A novel subclass of bovine β-defensins link reproduction and immunology. **Narciandi F, Lloyd AT, Meade KG, O' Farrelly C.** Reproduction, fertility, and development. **2013 Aug.**

ASSOCIATED PUBLICATIONS

- Divergent antimicrobial peptide (AMP) and acute phase protein (APP) responses to Trypanosoma congolense infection in trypanotolerant and trypanosusceptible cattle. **Meade KG, O'Gorman GM, Hill EW, Narciandi F, Agaba M, Kemp SJ, O'Farrelly C, MacHugh DE**. Mol Immunol. 2009 Dec.
- Directed alteration of a novel bovine beta-defensin to improve antimicrobial efficacy against methicillin-resistant Staphylococcus aureus (MRSA). **Meade KG, Cahalane S, Narciandi F, Cormican P, Lloyd AT, O'Farrelly C**. Int J Antimicrob Agents. 2008 Nov.
- Evolution, expression and effectiveness in a cluster of novel bovine beta-defensins. **Cormican P, Meade KG, Cahalane S, Narciandi F, Chapwanya A, Lloyd AT, O'Farrelly C**. Immunogenetics. 2008 Apr.

ABBREVIATIONS

Actb	B-actin
Amp	Antimicrobial peptides
Ap-1	Activator protein 1
Apc	Antigen presenting cell
Aqn1	Spermadhesins aqn1
Atpase	Adenosine triphosphate transferase
Bbd	Bovine beta defensin
Bend	Bovine endometrium cells
Bin 1b	sperm associated antigen 11e
Blat	Blast-like alignment tool
Bsp	Seminal plasma family proteins
Ca	Caput
Card	caspase activation and recruitment domains
Ccr2	Chemokine (c-c motif) receptor 2
Ccr6	Chemokine (c-c motif) receptor 6
Cd	Cluster of differentiation
Cdna	Complementary deoxyribonucleic acid
Cdr	Complementary determining regions
Ch	Constant heavy chain
Chr	Chromosome
Chr13	Bovibe chromosome 13
Cl	Constant light chain
Co	Corpus
Cu	Cauda
Dft	Distal Fallopian tube
Dht	Dihydrotestosterone
Dsrna	Double strand ribonucleic acid
Dt	Oligo thymine
Edta	Ethylenediaminetetraacetic acid
Elisa	Enzyme-linked immuno sorbent assay
Enh	Enhancer gfp specific vhh
Esp	Epididymal secretory protein e1
Est	Expressed sequence tag
Fc	Fragment crystallizable region
Fplc	Fast protein liquid chromatography
Fr	Framework regions
G	Gram
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase
Gdna	Genomic dna
Goi	Gene of interest
H3f3a	Histone h3.3

Hcamp18	Human cathelicidin, ll-37 precursor
Hdp	Host defense peptide
Hf	Holstein friesian
Hiv	Human immunodeficiency virus
Hnp1	human neutrophil peptide 1
Idc	Immature dendritic cells
Il-1	Interleukin-1
Il-1b	Interleukin-1 β
Il-6	Interleukin-6
Iptg	Isopropyl β-d-1-thiogalactopyranoside
Klh	Keyhole limpet hemocyanin
Lap	Lingual antimicrobial peptide
Ll37	Human cathelicidin
Lps	Lipopolysaccharide
Lrr	Leucine-rich repeats
Maf	Minor allele frequency
Mhc	Major histocompatibility complex
Min	Minute
Mm	Millimollar
Nf-kb	Nuclear factor kappa-light-chain-enhancer of activated b cells
Nm	Nanomolar
Nr	Norwegian red
°c	Celcious
Ov	Ovary
PAGEe	Polyacrylamide gel electrophoresis
Pbl	Peripheral blood leukocyte
Pbmc	Peripheral blood mononuclear cells
Pcr	Polymerase chain reaction
Pft	Proximal Fallopian tube
Ph	Power of hydrogen
Pmsf	Phenylmethanesulfonylfluoride
Ppia	Peptidylprolyl isomerase a
Qrt-pcr	Quantitative real time polymerase chain reaction
Rbbd126	Recombinant bbd126
Rna	Ribonucleic acid
Rps15	Ribosomal protein s15
Rps9	Ribosomal protein s9
Rt	Rete testis
Sds	Sodium dodecyl sulphate
Snp	Single-nucleotide polymorphism
Sp17	Sperm protein 17
Spag11e	sperm associated antigen 11e
Spam1	Hyaluronidase ph-20

Spla2iii	Group iii secreted phospholipase a2
Ssrna	Single strand ribonucleic acid
St	Seminal tubule
Tap	Tracheal antimicrobial peptide
Temed	Tetramethylethylenediamine
Tfbs	Transcription factor binding site
Tgf	Transforming growth factor
Th1	T helper cell 1
Tnfβ	Tumor necrosis factor β
U	Uterus
Vd	Vas deferens
Vh	Variable heavy chain
VHH	Heavy chain only antibody
Vl	Variable light chain
Mg	Microgram
Ml	Microlitre

LIST OF FIGURES

Chapter 1

Figure 1. 1 Unrooted phylogenetic tree of bacteriocidin from Gram positive bacteria. 19
Figure 1. 2 Unrooted tree of cysteine rich AMPs from plants. 20
Figure 1. 3. α-Helical peptides found in invertebrates and vertebrates. 21
Figure 1. 4. Cysteine rich peptides with 2 or less disulphide bonds. 21
Figure 1. 5. Cysteine rich peptides with 3 or more disulphide bonds 22
Figure 1. 6. Proposed mechanisms of action of AMPs 27
Figure 1. 7. Summary of intracellular effects and examples of AMPs causing such effects of cell metabolism. 28
Figure 1. 8. Roles and function described for antimicrobial peptides 30
Figure 1. 9. Models of AMPs mechanism of action as immune regulators. 32
Figure 1. 10. Protein alignment of Chromosome 13 bovine β-defensin cluster 34
Figure 1. 11. Defensin subclasses 36
Figure 1. 12. Tissue expression profile of bovine β-defensins. 39
Figure 1. 13. Graph representing the percentage of selection of a particular trait in NR breeding programmes since 1962. 42
Figure 1. 14. Genomic maps alignment of β-defensin cluster A from dog, human and bovine genome 44
Figure 1. 15. Genomic maps alignments of β-defensins gene clusters B and C 45
Figure 1. 16. Genomic map alignments of β-defensins gene cluster D in dog, human and bovine genome 46

Chapter 2

Fig. 2. 1 Male and female reproductive tract 52
Fig. 2. 2 Sperm penetration assay. 69
Fig. 2. 3 Phage display selection process. 75
Fig. 2. 4 Schematic of the structure of the VHH gene, with the highly variable CDR regions. 82
Fig. 2. 5 Protein sequence alignment of peptide inoculated in mice for antibody production and BBD126. 86

Chapter 3

Fig. 3. 1. Roles of DEFB126 in fertility. 94
Fig. 3. 2. Bovine reproductive trats RNA quality assessment 96
Fig. 3. 3. Average expression of stability of housekeeping genes on the male reproductive tract samples. 97
Fig. 3. 4. Average expression of stability of housekeeping genes on the female reproductive tract samples. 98
Fig. 3. 5. Agarose gel of qPCR products of the different genes in a panel of tissues. 99
Fig. 3. 6. Bovine Chr13 β-defensins gene expression across the male reproductive tract 107
Fig. 3. 7. Bovine Chr13 β-defensins gene expression across the female reproductive tract 112
Fig. 3. 8. Analysis of the β-defensins tail length 117
Fig. 3. 9. De-glycosylation effect on BBD126 molecular weight. 118
Fig. 3. 10. Scatter plot of component 1 vs component 2 and component 3 vs component 4 for each promoter region. 120
Fig. 3. 11. Scatter plot of component 1 vs component 2 and component 3 vs component 4 for each TFBS. 122
Fig. 3. 12. Relative luminescent units of pGL4.23-LAP plasmid 123
Fig. 3. 13. Relative luminescent units of pGL4.23-126 plasmid 124

Chapter 4

Fig. 4. 1. Haplotype maps for Chr13 β-defensin gene cluster between SNPs in Norwegian Red and Holstein-Friesian 137
Fig. 4. 2. Cell counts of sperm cells for BBD126 E1 (b) genotypes 144
Fig. 4. 3. Cell counts of sperm cells for BBD117 (b) genotypes 145
Fig. 4. 4. Cell counts of sperm cells for BBD123 E2 (a) genotypes 146

Chapter 5

Fig. 5. 1. Anti-BBD126 hybridoma test by western-blot. 160
Fig. 5. 2. Immunocytofluorescence localisation of BBD126 in sperm cells. 161
Fig. 5. 3. Immunohistochemistry localisation of BBD126 in cauda of bovine epididymis. 162
Fig. 5. 4. Immunohistochemistry localisation of BBD126 in seminal tubule of bovine testis. 163
Fig. 5. 5. Immunoprecipitation of anti-BBD126 antigen 164
Fig. 5. 6. Agarose gel of pEAK8-126F and pEAK8-126R PCR product. 165
Fig. 5. 7. Map of pEAK8-126 plasmid 167
Fig. 5. 8. DNA sequence and protein translation of pEAK8-126 vector. 168
Fig. 5. 9. Western blot of rBBD126 with antiBBD126 antibody. 168
Fig. 5. 10. Sequence alignment of protein sequences of pre- and post-round of selection for yeast library. 172
Fig. 5. 11. Protein production of yeast library selected VHH from the third round of selection. 173
Fig. 5. 12. Protein production of sortase enzyme. 174
Fig. 5. 13 Sortase purification by gel filtration column in a FPLC system. 175
Fig. 5. 14. Sortase labelling of yeast specific VHH with biotin. 176
Fig. 5. 15. Immunoprecipitation of yeast specific VHH with yeast lysate. 178
Fig. 5. 16. Consensus sequence for VHH sequence alignment 179
Fig. 5. 17. PCR generated VHH library construction. 180
Fig. 5. 18. Map of JSC-PCR-VHH vector 181
Fig. 5. 19. Sequence alignment of sequence of colonies for PCR-VHH library. 182
Fig. 5. 20. Protein production of PCR-VHH, B1 and B2. 183
Fig. 5. 21. Use of B1 and B2 for TNFα ELISA 184
Fig. 5. 22. Immunoprecipitation of B1 and B2 antigens from a bovine TNFα standard. 186
Fig. 5. 23. Immunoprecipitation of B1 and B2 antigens from human TNFα standard. 187
Fig. 5. 24. Immunoprecipitation of B1 and B2 antigens from human serum. 188

Chapter 6

Fig. 6. 1. Expression and hormonal regulation of β-defensins in the male reproductive tract. 199

LIST OF TABLES.

Chapter 1

Table 1. 1. List of bovine β-defensins and their features 46

Chapter 2

Table 2. 1. List of primers and conditions used for gene expression analysis by qPCR in this study. 53
Table 2. 2. List of primers and conditions used for PCR the coding region of genes in this study. 56
Table 2. 3. List of primers used for commercial genotyping of SNPs in both NR and HF breeds. 57
Table 2. 4. List of primers used for cloning, including the restriction enzymes sites incorporated in the primers 58
Table 2. 5 Conditions for the double digestion restriction enzymes. 63
Table 2. 6 Schematic for the distribution of the different vectors transfected into BEND cells. 67
Table 2. 7. Final concentration of the treatments used to stimulate luciferase production. 67
Table 2. 8. Schematic for the distribution of the different stimulations into BEND cells. 68
Table 2. 9. Ingredients of 12% acrylamide resolving gel, 4 % stacking gel and 4x loading buffer. 78
Table 2. 10. Description of capture antibody loaded in the different wells of the ELISA plate. 83
Table 2. 11. Description of standards loaded in the different wells of the ELISA plate 84
Table 2. 12. Description of detection loaded in the different wells of the ELISA plate. 85
Table 2. 13 Immunocytofluorescence experimental design. 87

Chapter 3

Table 3. 1. Summary representation of patterns of expression for Chr13 β-defensins in reproductive tract. 113
Table 3. 2. Length of peptide tail of the novel defensins and glycosylation sites predictions 115

Chapter 4

Table 4. 1. Genotyping results for 5 HF and 5 NR samples. 133
Table 4. 2. Comparison of novel SNP allelic frequencies between NR and HF populations. 135
Table 4. 3. Genotypic, genotype frequency, Ho1, He or Ho2, and Hardy-Weinberg analysis results for NR and HF populations 136
Table 4. 4. Haplotype frequencies detected in each breed. 139
Table 4. 5. Genotyping results for population of 19 bulls and pre- and post-freezing pregnancy frequencies 141
Table 4. 6. Pearson correlation results between 19 bulls' genotype data and frequency of pregnancy pre- and post-semen freezing 142

ABSTRACT

Antimicrobial peptides (AMPs) have been characterised by their ability to neutralise pathogens directly by disrupting their membranes. Recently described roles for these peptides include chemotactic activity, wound healing and regulation of the inflammatory response. AMPs have also been shown to have non-immune functions, such as control of coat colour and roles in embryonic development, sperm maturation and fertility. Using combined bioinformatics and molecular approaches, our group found a new cluster of β-defensin genes in the bovine genome on chromosome 13. Expression of the gene cluster was exclusive to the reproductive tract with little or no expression in any other tissue. The aims of this study are to further investigate this cluster of genes, to characterise their expression, to examine their genetic variability, population structure and to explore a possible role in fertility. We also set out to create the tools required for their functional characterisation.

Male and female reproductive tracts from slaughtered cattle at a local abattoir were retrieved and dissected to obtain several regions: seminal tubule, rete testis, caput, corpus and cauda of the epididymis and vas deferens from the male samples; ovary, proximal Fallopian tube, distal Fallopian tube and uterine body in the female tracts. RNA extracted from the dissected samples was used for cDNA synthesis and analysed for β-defensin gene expression by qRT-PCR. In the male samples, we found three distinctive patterns of gene expression. Bovine β-defensin (*BBD*) *115, 116, 117, 118, 124, 125a, 125, 126, 127, 128, 129* and *132*) were expressed in the distal epididymal region of the male reproductive tract; *BBD122, 122a* and *123* were expressed mostly in the testicular region while *BBD142, 119, 120* and *121* were expressed across all the tissues. Levels of expression in the female tissues were similar in all regions analysed. Several of these β-defensins were not found to be expressed in either female or immature male reproductive tracts (*BBD125, 125a, 125, 126,127, 128, 129* and *132*) which suggested a specific role in the mature male reproductive tract.

Regulation of expression of one of the members of the cluster, *BBD126*, was examined using a luciferase reporter gene assay. We found gene expression for BBD126 to be under hormonal control, in contrast with immune signalling regulation found in other β-defensins.

The genetic variation of the coding regions of these defensins between two bovine breeds, Holstein-Friesian (HF) and Norwegian Red (NR), was examined. We found 17 novel single nucleotide polymorphisms (SNP), of which 12 had a frequency higher than 10%. Of these 12

SNPs, 3 were found in non-coding regions surrounding the exons; the other 9 were divided between 5 non-synonymous and 4 synonymous SNPs. These 12 SNPs were genotyped in a larger population of 30 NR and 30 HF. Four were present with significantly different frequencies between the two breeds ($P < 0.05$). Two haplotypes were found, haplotype A (*BBD128 E1, BBD126 E1 (a), BBD126 E1 (b)* and *BBD115 E2*,) and haplotype B (*BBD121 E1, BBD122 E2, BBD123 E2 (a), BBD123 E2 (b)* and *BBD124 E1*) with different frequencies in HF and NR.

For localisation and functional studies of a defensin protein, a monoclonal antibody for BBD126 was commercially created and tested using different technologies. Western-blot analysis showed that the antibody bound to recombinant BBD126. Western-blot sperm cell lysate demonstrated binding of the antibody to a protein which was 25 kDa when glycosylated and 7 kDa when de-glycosylated as BBD126 would be expected to behave. Analysis of native protein by immunocytofluorescence and immunohistochemistry revealed that the antibody could detect a protein on the sperm cell surface and in the lumen of the epididymis. In contrast, in an immune precipitation assay, the antibody captured a 15 kDa protein identified by mass spectrometry as spermadhesin Z13. These results suggest different specificity for the anti-BBD126 antibody for native and denatured protein samples.

An alternative method for cost effective antibody production using recombinant heavy chain only antibodies (VHH) was tested for the creation of β-defensin reagents. As proof of concept, the technique was optimised using bovine TNFα as target. The PCR-generated VHH library resulted in creation of a VHH with specificity for bovine TNFα. This technology could have major functional implications for the generation of reagents for detecting and analysing bovine defensin proteins.

In conclusion, our demonstration of reproductive tract specific expression of the Chr 13 β-defensin cluster genes suggests a role for these proteins in fertility. Demonstration of the presence of polymorphism variation in these genes with significant different frequencies between HF and NR suggests that this may be a novel breeding target.

CHAPTER 1: GENERAL INTRODUCTION

1.1. INNATE IMMUNITY

The defense of the organism against harm has been established since the appearance of life. One of the earliest mechanisms of defense was the production of effector molecules with lethal or detrimental effect on invading organisms. Some examples of this chemical welfare are the antimicrobial substances found in bacteria with the ability to block growth of competitors, from the non-peptide antibiotics, such as penicillin, gentamicin to antimicrobial peptides, such as nisin. Antimicrobial peptides and other antimicrobial substances can be found in found in every stage of evolution indicating an extremely successful adaptation.

An additional mechanism of defense found widely distributed in the living world, that introduces an extra level of complexity, is phagocytosis, the ability for cells to engulf particles or smaller organism into an intracellular compartment where antimicrobial substances, target and neutralise the pathogen with higher concentrations of antimicrobial substance consuming considerably less resources compared with their secretion to the exterior of the cell. Phagocytosis is found in primitive unicellular eukaryotes such as amoebas. In these organisms, phagocytosis fulfils a double role in defense and nutrition. In social amoebas, a stress induced multicellular state, most cell lose their phagocytic activity, but a small group of cells retains the ability to phagocytose and defend the multicellular cluster against bacteria (Chen, Zhuchenko et al. 2007). In more complex multicellular eukaryotes, phagocytosis is associated with specialised cell types with an immune defense role.

These mechanisms are highly conserved and present in most organisms, and are generally known as part of the innate immunity, the first line of defense against pathogens. Innate immunity relies in the detection of pathogens by recognition receptors (PRRs). These are membrane and intracellular receptors activated by highly conserved pathogen molecules or pathogen-associated molecular patterns (PAMPs). The binding of a PAMP to a PRR initiates a signalling cascade that triggers the synthesis and release of different effector molecules that eliminate directly or indirectly, by immune cell recruitment, the pathogen in a process known as inflammation. The signs of inflammation are dolor (pain), calor (heat), rubor (redness), tumor (swelling) and functio laesa (loss of function). All the signs are induced by the production and secretion of molecules by activated cells. Inflammation is activated at local and systemic levels by a danger signal; this danger signal can be induced by direct cellular or tissue damage (by physical injury or pathogenic origin) or through pathogen detection by the

organism using surface or intracellular receptors. Inflammation has evolved to be an acute short-term process, quickly regulated and brought under control, to induce and facilitate the elimination of the danger signal inducer (pathogen or tissue damage). If poorly regulated, the inflammatory process is prolonged in time producing a clinical process called chronic inflammation. Most living organism relies completely upon the innate immunity with evolutionary success. Only in vertebrates, has second type of immune response, known as adaptive immune response evolved. Adaptive immune response relies on the detection and neutralisation of previously un-encounter pathogens by molecules with variable specificity that matures with the organism; in contrast with the innate immunity with highly conserved PRR for highly conserved PAMPs present in pathogens.

During inflammation, once the pathogen or tissue damage is detected by a cell, a signalling cascade is initiated, generating the induction of gene expression of a series of inflammatory molecules and release of the newly produced molecules into the extracellular environment. We can divide these molecules into two groups according to their role:

Signalling molecules include cytokines and chemokines. Their roles are to alert surrounding cells to damage or pathogen presence; to modify cell behaviour in an autocrine and/or paracrine manner (cytokines) and to attract immune cells to the site of inflammation (chemokines).

Effector molecules are small proteins with antimicrobial properties, produced by the cells to eliminate pathogens. Generally called antimicrobial peptides (AMPs), the best known are cathelicidins and defensins.

1.2. ANTIMICROBIAL PEPTIDES

Until the early 1960s, antimicrobial activity was assumed to consist solely of the phagocytic activity of macrophages and neutrophils. It was also known that neutrophil granules disappeared during the phagocytosis process. In 1963, guinea-pig poly-morphonuclear leukocytes (PMN) granules were isolated by gradient centrifugation, the granules were then lysed and the protein content was separated by electrophoresis. Proteins were divided into different fractions according with their migration speed and were tested for antimicrobial properties, it was found that a group of cationic peptides had a significant antimicrobial

activity, suggesting that the granules were fusing with the phagocytized pathogen and were essential for the elimination of the pathogen (Zeya and Spitznagel 1963). Similar proteins have been found in all analysed organisms, from bacteria, (Gao, Rodriguez et al. 2009) to humans (Ganz, Selsted et al. 1985) and plants (Thomma, Cammue et al. 2002). These antimicrobial agents, when of protein origin, were classified into a group known as antimicrobial peptides (AMP). Peptides with toxic activity can be found as early in evolution as prokaryotes, bacteriocins are small peptides with the ability to inhibit cell growth of surrounding cells. A recent classification found 12 distinctive groups of bacteriocins according to their 3 dimensional structure (Fig. 1. 1) (Zouhir, Hammami et al. 2010).

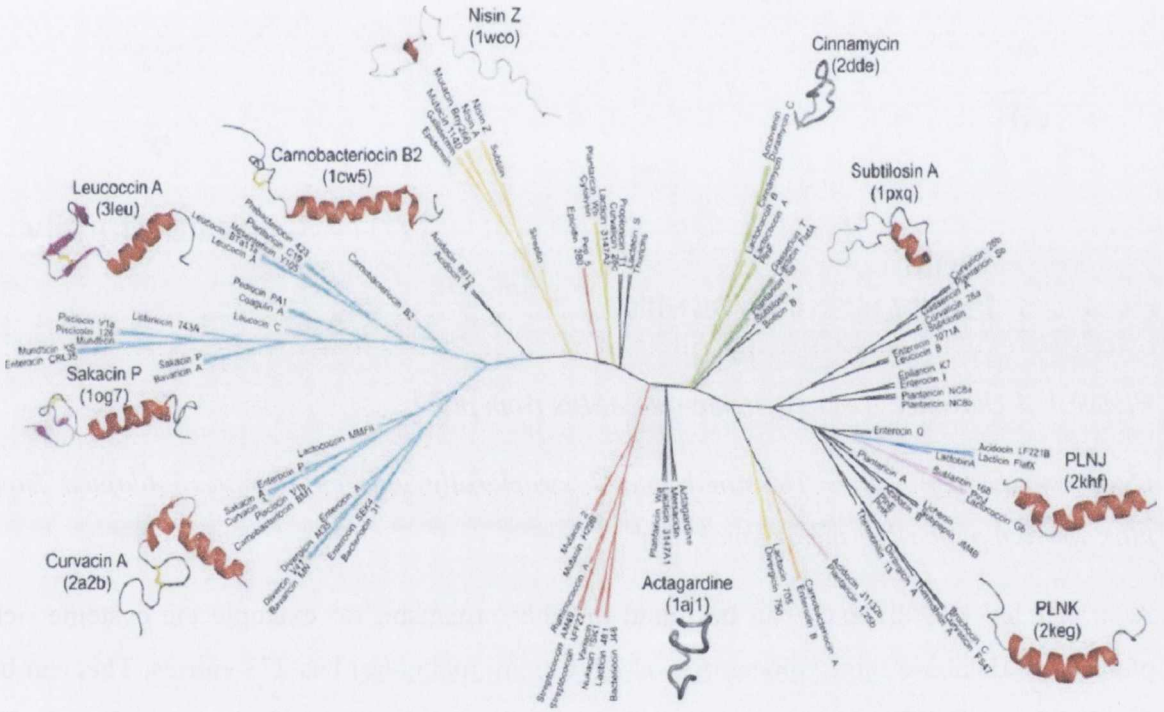


Figure 1. 1 Unrooted phylogenetic tree of bacteriocidin from Gram positive bacteria. Examples of 3 dimensional structures are shown for each group, in red α-helix, in purple β-sheet structures .Adapted from (Zouhir, Hammami et al. 2010).

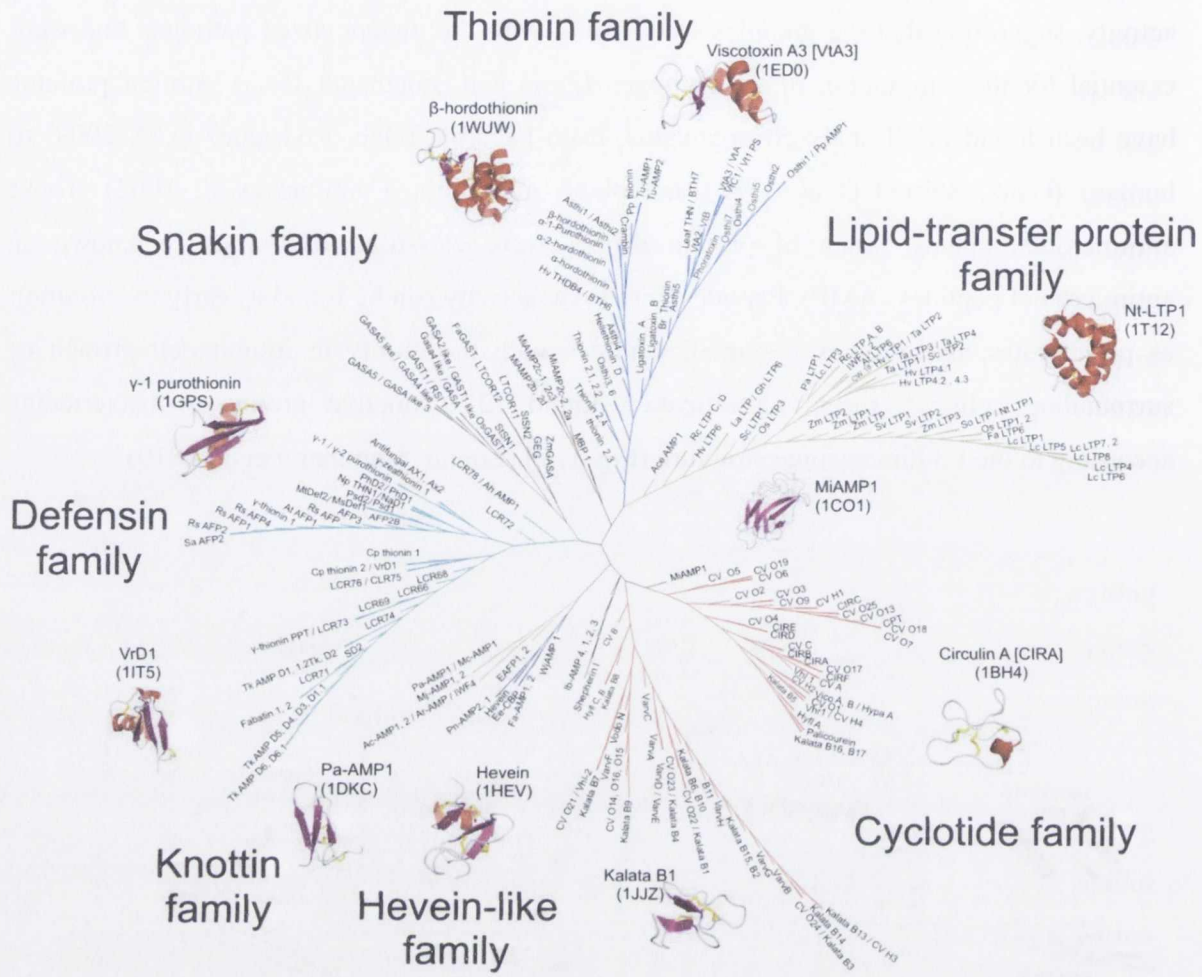


Figure 1. 2 Unrooted tree of cysteine rich AMPs from plants.

Cysteine rich AMP were classified into 7 structurally different families. Adapted from http://phytamp.pfba-lab-tun.org/main.php.

A similar level of diversity can be found in other organism, for example the cysteine rich plant AMP database (http://phytamp.pfba-lab-tun.org/main.php) has 273 entries. This can be divided in to 7 different groups according with their 3 dimensional structure (Fig. 1. 2). AMPs are essential components of plants, due the lack of specialised immune cells such as macrophages which are present in animals.

In animals, AMPs can be found in both invertebrates and vertebrates. A classification of AMPs in animals found 2 main groups (Bulet, Stöcklin et al. 2004). The first group includes α-helical AMPs, includes peptides such as cecropin, found in insects (Lauwers, Twyffels et al. 2009) or cathelicidins, found in mammals (Tomasinsig and Zanetti 2005) (Fig. 1. 3)

	Origin	Organism	Name	Activity	Primary structure
Invertebrates	Arthropods	Silk moth *Hyalophora cecropia*	Cecropin A	B	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK*
		Mosquito *Anopheles gambiae*	Cecropin A	B, F, Y	GRLKKLGKKIEGAGKRVFKAAEKALPVVAGVKAL*
		Stable fly *Stomoxys calcitrans*	Stomoxyn	B, F, Y, T	RGFRKHFNKLVKKVKHTISETAHVAKDTAVIAGSGAAVV AAT*
		Fire ant *Pachycondylas goeldii*	Ponericin G1	B, H, I	GWKDWAKKAGGWLKKKGPGMAKAALKAAMQ
		Termite *Pseudacanthotermes spiniger*	Spinigerin	B, F, Y	HVDKKVADKVLLLKQLRIMRLLTRL
		Spider *Oxyopes kitabensis*	Oxyopinin 1	B, H, I	FRGLAKLLKIGLKSFARVLKKVLPKAAKAGKALA KSMADENAIRQQNQ
		Spider *Cupiennius salei*	Cupiennin 1	B, H, I	GFGALFKFLAKKVAKTVAKQAAKQGAKYVVNKQME*
	Procordates	Ascidian *Styela clava*	Clavanin A	B	VFQFLGKIIHHVGNFVHGFSHVF*
			Styelin D	B, H, C	$GW^1LR^2K^2AAK^2SVGK^2FY^3Y^3K^2HK^4Y^3Y^3IK^4AAWQIGKHAL*$
Vertebrates	Fishes	Sole *Pardachirus pavoninus*	Pardaxin P-I	B	GFFALIPKIISSPLFKTLLSAVGSALSSSGEQE
		Mudfish *Misgurnus anguillicaudatus*	Misgurin	B, F, Y	RQRVEELSKFSKKGAAARRRK
		Flounder *Pleuronectes americanus*	Pleurocidin	B	GWGSFFKKAAHVGKHVGKAALTHYL
		Catfish *Parasilurus asotus*	Parasin I	B, Y	KGRGKQGGKVRAKAKTRSS
		Rainbow trout *Oncorhynchus mykiss*	Oncorhyncin II	B	KAVAAKKSPKKAKKPATPKKAAKSPKKVKKPAAAA KKAAKSPKKATKAAKPKAAKPKAAKAKKAAPKKK
		White bass *Morone chrysops*	Moronecidin	B, F, Y	FFHHIFRGIVHVGKTIHKLVTG
	Amphibians	Clawed frog *Xenopus laevi*	Magainin 1	B, F	GIGKFLHSAGKFGKAFVGEIMKS
		Common frog *Rana temporaria*	Temporin A	B, F, H	FLPLIGRVLSGIL*
	Mammals	Human *Homo sapiens*	Cathelicidin hCAT-18/LL-37	B, L	LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES
		Bovine *Bos taurus*	Indolicidin	B	ILPWKWPWWPWRR*

3D structure

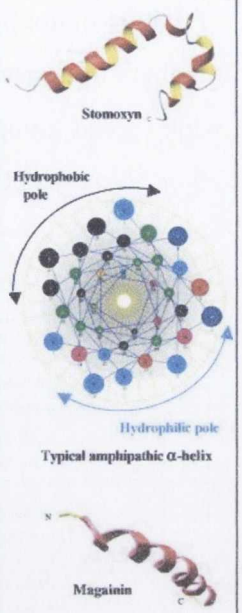


Figure 1. 3. α-Helical peptides found in invertebrates and vertebrates.

Adapted from (Bulet, Stöcklin et al. 2004).

A second group of AMPs includes cysteine rich peptides. The presence of cysteines induces the formation of disulphide bonds. The number of disulphide bonds generates two different structural motifs. In the presence of 1 or 2 disulphide bonds, the peptide acquires a structure known as β-hairpin, this can found in peptides such as thanatin in invertebrates or protegrin in mammals (Fig. 1. 4.).

	Origin	Organism	Name	Activity	Primary structure
Invertebrates	Arthropods	Insect *Podisus maculiventris*	Thanatin	B, F, Y	GSKKPVPIIYCNRRTGKCQRM
		Horseshoe crab *Tachypleus tridentatus*	Tachyplesin-I	B, Y, V, H	KWCFRVCYRGICYRRCR*
		Horseshoe crab *Limulus polyphemus*	Polyphemusin	B, Y	RRWCFRVCYRGFCYRKCR*
		Scorpion *Androctonus australis*	Androctonin	B, F	RSVCRQIKICRRRGGCYYKCTNRPY
		Spider *Acanthoscuria gomesiana*	Gomesin	B, F, Y, P, H	ZCRRLCYKQRCVTYCRGR*
Vertebrates	Fishes	White bass *Morone chrysops x saxalitis*	Hepcidin	B, Irh	GCRFCCNCCPNMSGCGVCCRF
	Amphibians	Japanese frog *Rana brevipoda*	Brevinin-1	B, Y	FLPVLAGIAAKVVPALFCKITKKC
		Edible frog *Rana esculenta*	Esculentin-1	B, Y, H	GIFSKLGRKKIKNLLISGLKNVGKEVGM DVVRTGIDIAGCKIKGEC
		Indian bull frog *Rana tigerina*	Tigerinin-1	B, Y	FCTMIPIPRCY*
	Mammals	Bovine *Bos taurus*	Bactenecin	B	RLCRIVVIRVCR
		Porcine *Sus scrofa*	Protegrin-1	B, F, Y, V	RGGRLCYCRRRFCVCVGR*
		Mouse *Mus musculus*	Hepcidin	B, Irh	DTNFPICIFCCKCCNNSQCGICCKT

3D structure

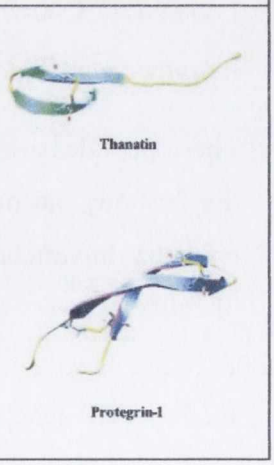


Figure 1. 4. Cysteine rich peptides with 2 or less disulphide bonds.

Adapted from (Bulet, Stöcklin et al. 2004).

In the presence of 3 or more disulphide bonds, a second motif, characterised by the presence of 3 antiparallel β-sheets is formed (Fig. 1. 5.). The β-hairpin peptides could be found in high numbers in amphibians, but is rarely represented in insects or mammals. In contrast the most wide spread group are β-sheet peptides, in particular a group known as defensins, more than 70 orthologous peptides have been found in invertebrates. The number in vertebrates exceeds 100.

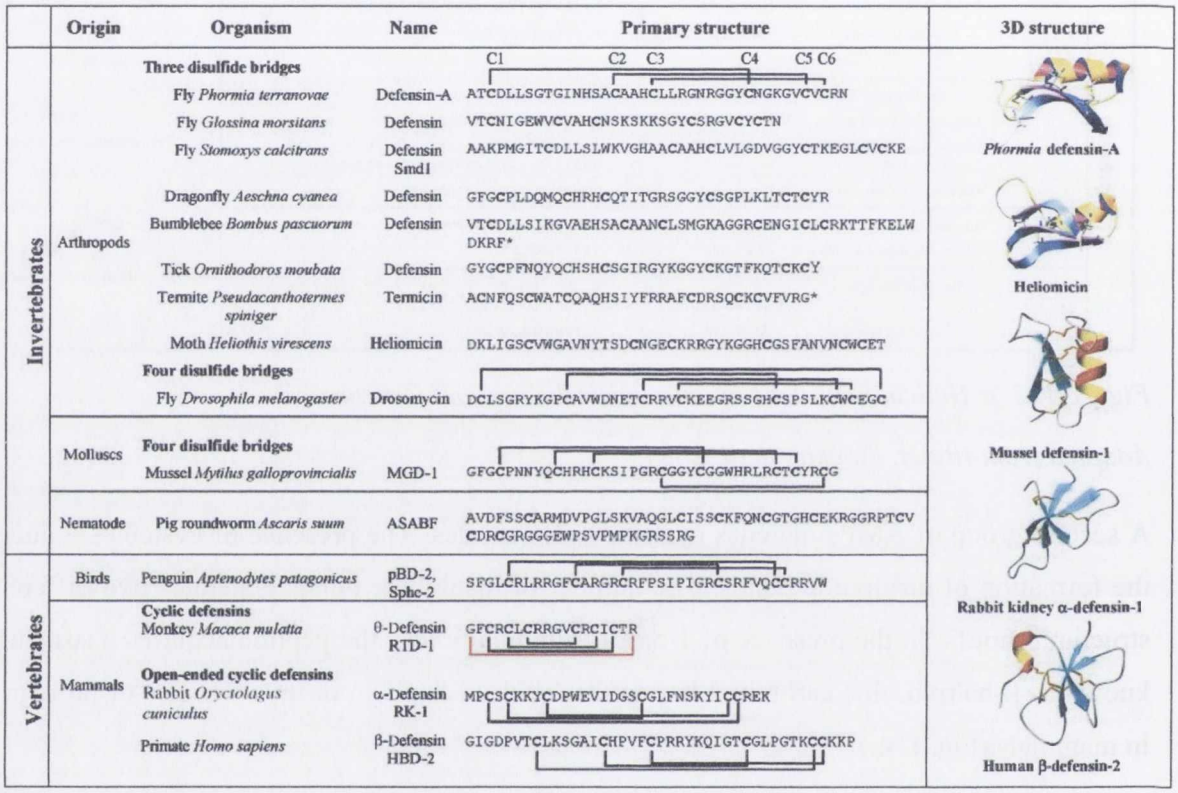


Figure 1. 5. Cysteine rich peptides with 3 or more disulphide bonds.

Adapted from (Bulet, Stöcklin et al. 2004).

These peptides are found in many tissues. In arthropods most AMPs are found expressed in the fat body, an organ with similar roles to the liver in vertebrates, blood cells and some epithelia. In vertebrates, AMPs can be found in the granules of neutrophils and in a multitude of epithelia.

1.2.1. AMP gene and protein structure

Most AMPs are encoded in clusters of genes in different areas of the genome, most probably originating from multiple rounds of gene duplication; these duplicates have eventually

diverged and assumed new functions and roles through positive selection (Cormican, Meade et al. 2008) but also through genetic drift. The multiplicity of genes allows diversity in time and location for gene expression; in some cases this regulation is highly specific.

AMPs in general have two very different patterns of expression; some are expressed constitutively in one or several cell types; for example the peptides normally produced in the intestinal epithelium. It has been hypothesised that these help to maintain the equilibrium of the intestinal micro-flora (Shi, Aono et al. 2007), and a lack or excess in production is correlated with some intestinal disorders (Wehkamp, Koslowski et al. 2008). The other expression pattern has been correlated with the immune response, these AMPs are expressed through the activation of PRRs like TLRs. Evidence of alternative mechanisms of activation have been suggested, such as hormone regulation, showed by the down-regulation of mouse defensins in the epididymis after castration (Yenugu, Chintalgattu et al. 2006). An example of tight regulation of expression of antimicrobial peptides is the mouse cathelicidin CRAMP which is normally expressed in foetal intestine but not in the adult (Menard, Forster et al. 2008).

Due to their potential toxic effect, AMPs are generally synthesised as precursor proteins that requires processing to become functional, which generally involves cleavage of a portion of the protein. In the case of defensins, the peptide is first synthesised as a pre-propeptide. The pre-propeptide includes the leucine rich signal peptide in the N-terminal region, required for the correct transport and secretion of the peptide. The propeptide is formed by the mature peptide in the C terminal region and the pro-sequence (Ganz 2003). This sequence is cleaved by a protease: metalloproteinase 7 (MMP7) deficient mice showed inhibition of antimicrobial peptide processing and susceptibility to *S. aureus* infection (Wilson, Ouellette et al. 1999). The precursor peptide is digested by proteases that free a smaller peptide from the C-terminal region that will become the functional peptide. An example is the only cathelicidin gene discovered so far in the human genome, which produces a precursor protein called hCAMP18, and subsequently processed into a smaller peptide known as LL-37 (Tomasinsig and Zanetti 2005).

1.2.2. Antibacterial activity of AMPs

Several models have been described to explain the mechanisms by which AMPs function. Because of the small number of interactions between AMPs and their specific target described to date, and due to the lack of resistance appearing against AMPs, it has been

concluded that they must act by interacting with an essential structure of the pathogen; the most studied being the lipid bilayer. These models are generally generated by experiments using one or a small number of peptides, although no model has been generated that can explain the mechanism of action of most AMPs. Their structural variability gives us a hint of the mechanism of action. All the models described are based on two requirements:

- A positive charge (a minimum of +2), which is one of the most interesting characteristics of AMP, and accounts for their specificity towards bacteria and encapsulated viruses. This phenomenon has been explained by the presence of a negative charge in prokaryotes in contrast with the neutral or positive charge found in most eukaryotic cells. The prokaryotic external lipid layer is rich in negatively charged lipids (phosphatidylglycerol and phosphatidyl-serine) in contrast with eukaryotic membranes which are rich in neutral and positively charged lipids (phosphatidylcholine, phosphatidylethanolamine, sphingomyelin). In addition, the transmembrane potential differences between mammalian cells (-90 / -110 mV) and bacteria in logarithmic growth phase (-130 / -150 mV) increases the differential charge between them (Yeaman and Yount 2003). It has been hypothesised that the presence of specific molecules in eukaryotes such as cholesterol, have a role in protecting their cells against the effect of AMPs. The AMP, sarcotoxin IA, has a reduced effect on liposomes containing cholesterol compared with liposomes without it (Nakajima, Qu et al. 1987). Due to the cholesterol role of condensation over the phospholipids, it could prevent incorporation of AMPs into the lipid bilayer, but this role has not been fully described.
- A certain degree of amphipathicity of the peptide, which allows interaction with the lipid bilayer (Yeaman and Yount 2003).

Not all AMPs fulfil both conditions, many are anionic or neutral and some have no amphipathic properties. No full explanation of how these non-cationic peptides act has been described. Antimicrobial peptides show a level of selection between Gram positive and Gram negative bacteria. There are two models of first contact of peptide and membrane according to the membrane structure:

- In Gram positive bacteria, the first contact is based in the electrostatic interaction between the positively charged peptide and the negatively charged membrane, which is enriched in negatively charged teichoic and teichuronic acids. One study showed that a modification of the components increasing the negative surface charge increased the killing ability of cationic peptides (Peschel, Otto et al. 1999).

- In Gram negative bacteria, an alternative model for the initial interaction has been presented, named self-promoted uptake, similar to the mechanism of aminoglycoside antibiotics. The positively charged peptides compete against LPS binding cations that stabilise the outer membrane (Robert Ew 1997). Several studies support this model that modifications of LPS induce the blocking of AMPs activity against Gram negative bacteria (Helander, Kilpeläinen et al. 1994).

There are some reports of receptor-mediator interactions, for example seminalplasmin, a peptide present in seminal plasma in bulls, which inhibits the synthesis of peptidoglycan (Chitnis and Prasad 1990). Other peptides exert their effects in the cytoplasm of the pathogens. Pyrrhocoricin has been shown to inhibit the ATPase actions of DnaK and prevents chaperone-assisted protein folding (Kragol, Lovas et al. 2001). Despite the presence of these and other examples of specific targets, the main line of research focuses on the interaction of AMPs and membranes and their disruption as a mechanism of action.

Several possible models of the mechanism of action of AMPs have been proposed over the years, none of which adequately explain the action of all AMPs (Fig. 1. 6). All models are based on experimental results obtained using a small number of AMPs:

- *Barrel-stave mechanism*, (Rapaport, Peled et al. 1996, Shai and Oren 1996) a variable number of peptides form a water filled channel in which the peptides acquire the conformation of a barrel, with the hydrophilic sections forming the interior of the pore and the hydrophobic regions in contact with the lipids of the membrane. According to this model, peptides act as monomers in solution but tend to form polymers once in contact with the membrane, where specific peptide-peptide interaction is required for the channel formation.
- *Toroidal* or *wormhole mechanism* (Ludtke, He et al. 1996), this model has been proposed for α-helical peptides. It proposes the formation of a water filled channel like the barrel model, but in this model the hydrophilic heads of the membrane lipids form part of the channel, forming a structure described as a supramolecular complex. Initially, the peptide has the α-helical conformation when it gets in contact with the membrane, adopting a parallel disposition to the membrane (Hori, Demura et al. 2001). The increasing concentration of the peptide on the membrane induces the displacement of the phospholipids' heads, in a determinate lipid/peptide ratio; the peptide acquires a perpendicular disposition to penetrate the membrane. At this point, the peptides began to interact with each other and with the hydrophilic heads of the phospholipids forming a channel. Some studies suggest that this mechanism is part of a larger mechanism in which peptides get transported into the cytoplasm

of the pathogens and act against internal targets (Uematsu and Matsuzaki 2000, Yang, Weiss et al. 2000).

- No pore formation models. The best known and described is the *carpet mechanism* (Shai and Oren 1996), which presupposes that an initial electrostatic interaction is required for the accumulation of the peptide on the surface of the membrane. When peptide concentration passes a threshold, the peptides cause the destabilisation of the membrane and its rupture. An alternative model to the *carpet mechanism* is the *detergent model* in which the peptides cause the formation of micelles and the fragmentation of the membrane (Bechinger and Lohner 2006). Other more recent models include the *molecular shape model* (Bechinger and Lohner 2006), in which the tertiary structure of the peptide determines the disruption of the membrane caused by the peptide; the *sinking raft model* (Pokorny and Almeida 2004) and the *interfacial activity model* (Rathinakumar and Wimley 2008).

All these models are limited by the experimental conditions used; synthetic vesicles of a limited lipid composition to simulate cell membranes, greater peptide concentration than found in physiological conditions and chemically synthesised peptides that raise questions about the correct folding of these peptides. Some of these issues could have important effects on peptide activity, changing the way the peptide interacts with cell membrane. AMPs are an extremely diverse group described together for their antimicrobial activity and small size; this variability will most likely produce diversity in the mechanisms by which they achieve bactericidal activity. There are several phenomena that do not fit with any of the models and we require more study to have a complete understanding. The most intriguing fact is that AMPs used in *in vitro* assays are in general not functional under physiological salt conditions, where the electrostatic nature of AMPs attracts anions that will affect the AMPs' ability to bind bacterial membranes (Boman 2003).

In vivo AMPs are produced by the cell in response to pathogens and the presence of AMPs produces a reduction in the pathogen load in proximity to the cell. The antimicrobial role of AMPs has been confirmed *in vivo* by the use of knockout mice. A cathelicidin knockout mouse (*Camp* -/-) demonstrated greater susceptibility to Group A *Streptococcus* (Lee, Ohtake et al. 2005). Mice knockouts for β-defensin 1 produce a slower clearing rate of *Haemophilus influenza* (Moser, Weiner et al. 2002). This contradiction may be explained by the existence of a charge independent alternative mechanism to the direct interaction with the pathogen membrane mechanism in which the AMP can eliminate pathogens in the host.

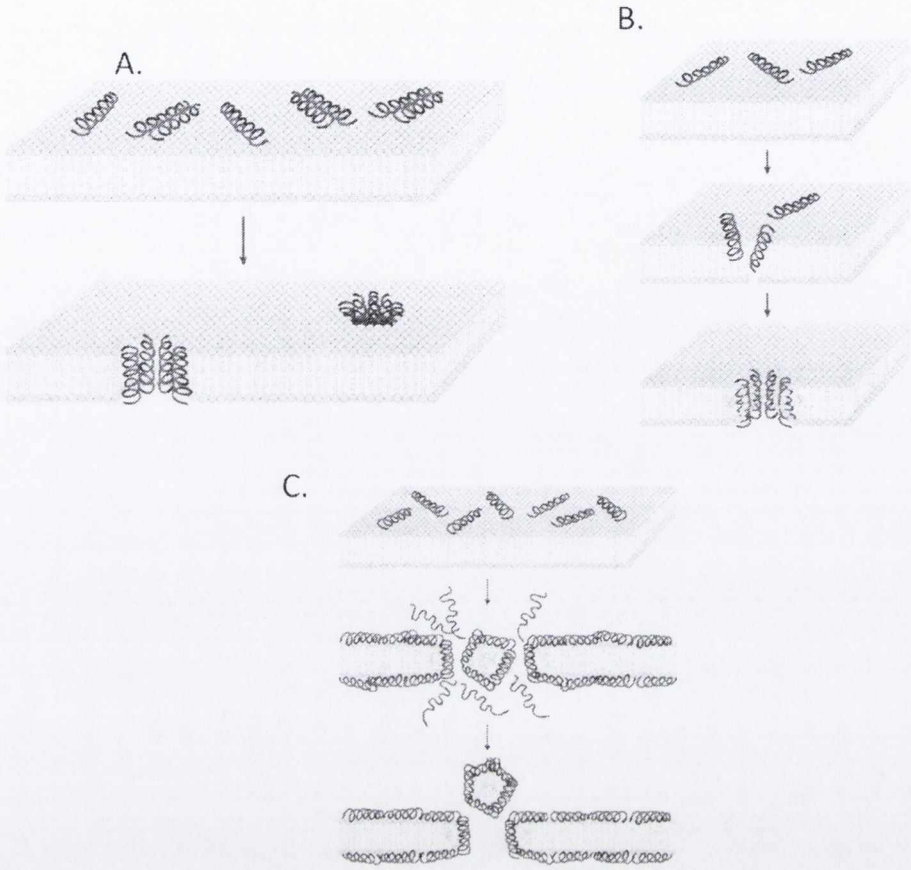


Figure 1.6. Proposed mechanisms of action of AMPs

A, Barrel-stove model peptides polymerise in contact with the membrane, hydrophilic regions form the lumen of the pore; B, toroidal or wormhole model, peptide forms pores in combination with the hydrophilic heads of the lipids part of the cellular membrane and C, carpet model, peptide accumulates in the membrane surface until a threshold is passed in which the peptide destabilises the membrane. Adapted from (Brogden 2005).

In addition to the effect on the cellular membrane there is evidence of intracellular effects of AMPs at different levels of the cellular metabolism, inhibiting cell growth and reducing cell viability. Several studies have demonstrate the ability of AMPs to block protein synthesis at different stages: for example pig PR-39 it is able to block DNA and RNA synthesis and RNA translation blocking protein synthesis (Boman, Agerberth et al. 1993). Other AMPs are able to block the synthesis of other essential components such as cell-wall, for example mersacidin block the synthesis of peptidoglycan (Brötz, Bierbaum et al. 1998).

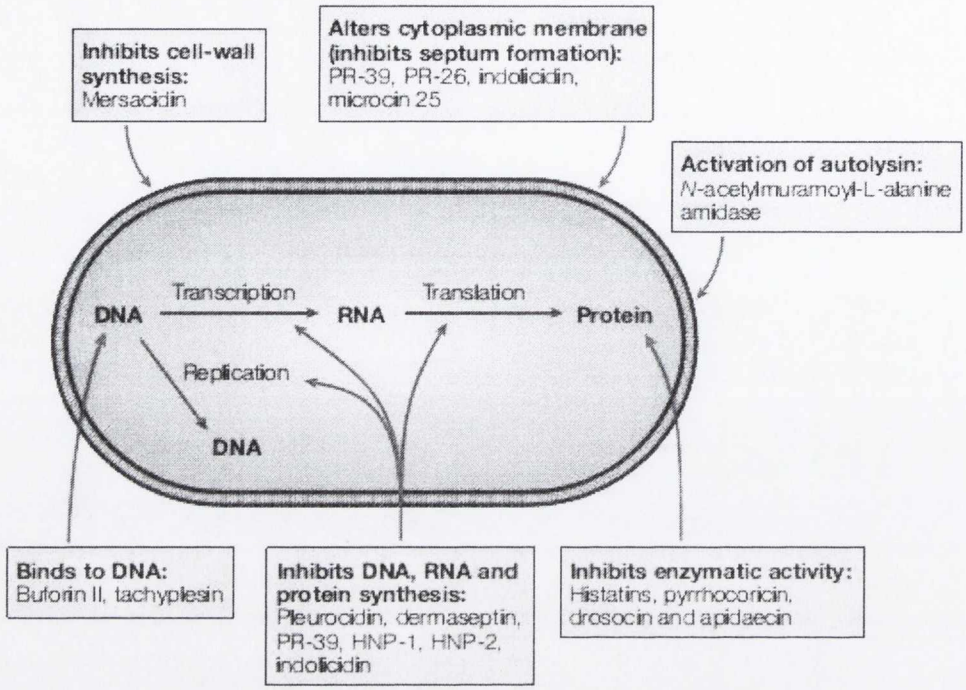


Figure 1. 7. Summary of intracellular effects and examples of AMPs causing such effects of cell metabolism.

Antimicrobial peptides can act over cell metabolism at different levels, binding to DNA, blocking DNA, RNA and protein synthesis at the replication, transcription and translation levels. AMPs can act on cell wall and cytoplasmic membrane synthesis and block enzymatic activity.

1.2.3. Antiviral activity of AMPs

There are several reports of viral induction of AMP gene expression, in the absence of viral replication (Quinones-Mateu, Lederman et al. 2003), and in the presence of viral replication (Proud, Sanders et al. 2004); similar effects have been achieved by TLR3 activation with Poly I:C (the TLR3 ligand) (Duits, Nibbering et al. 2003, Proud, Sanders et al. 2004). It is disputed whether this activation is a direct effect of TLR signalling or if AMP genes are indirectly activated by cytokines induced by TLR activation. Human neutrophil α-defensin 1 (HNP1) showed a direct effect against enveloped viruses but not other viruses (Daher, Selsted et al. 1986). It has been proposed that, similar to their antibacterial activity, AMPs disrupt the membrane of enveloped viruses, but the exact mechanism is still unknown. An alternative model hypothesised an interaction between AMPs and glycol proteins in the virus surface

which may block viral replication. There are reports of HNP1 acting in a lectin-like manner, binding to gp120 glycoprotein of HIV in *in vitro* studies, but the presence of serum reduces the antiviral activity (Wang, Owen et al. 2004).

1.3. Emerging roles for Antimicrobial Peptides

Since their discovery, AMPs research has focused on the antimicrobial role. In recent years, however, new roles for AMPs have been identified (Fig. 1. 8). There are many studies describing how the AMPs interact with the membranes of bacteria and viruses and how this interaction induced cell death of the pathogens as described in section 1.4.3. However, similar effects could be achieved by an indirect mechanism; for example, the regulation of the inflammatory process to promote the elimination of the pathogens. There are several reports of AMPs modulating the production of cytokines and chemokines in different cell lines; for example the single human cathelicidin, LL-37, can enhance the production of IL-6 and IL-10 in IL-1β induced peripheral blood mononuclear cells (PBMC) (Yu, Mookherjee et al. 2007); α-defensins produced by apoptotic macrophages can inhibit pro-inflammatory cytokines in chronic inflammation (Miles, Clarke et al. 2009) and in other cases can induce pro-inflammatory cytokines such as TNFα and IFNγ (Soehnlein, Zernecke et al. 2008).

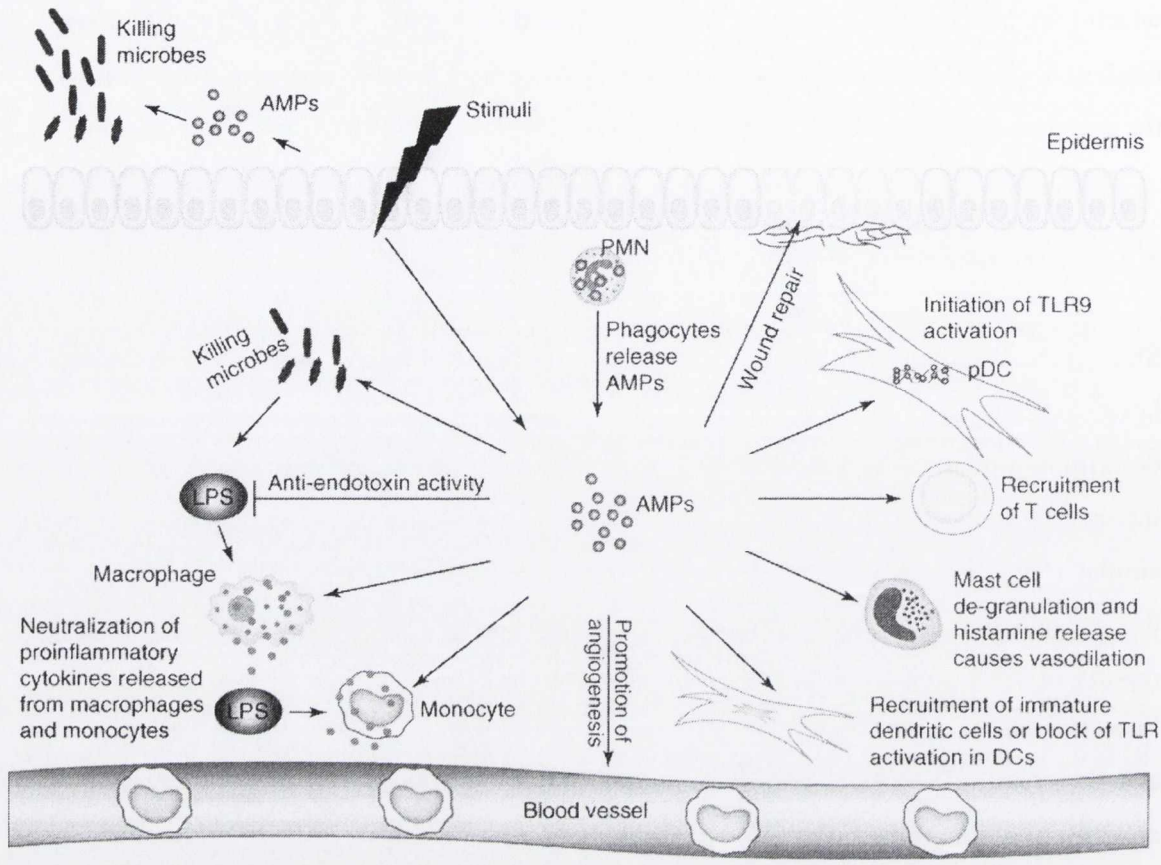


Figure 1. 8. Roles and function described for antimicrobial peptides

The number of roles given to AMP include, wound repair, cytokine functions such as blocking cytokine release by TLR activation, TLR9 activation, chemokine functions such as cell recruitment of T cells or immature dendritic cells. Other roles include angiogenesis and antimicrobial function. Adapted from (Lai and Gallo 2009).

Chemotactic properties for AMPs have also been described; β-defensins may be potent chemoattractants of immature dendritic cells (iDC). Human β-defensins 2 and 3 and their mouse orthologues interact with the membrane receptor CCR2 to induce chemotaxis of iDCs (Rohrl, Yang et al. 2010). Other examples are the effect of mouse cathelicidin, CRAMP, as chemoattractants of human monocytes, neutrophils, macrophages and mouse peripheral blood leukocytes (PBL) (Kurosaka, Chen et al. 2005). The chemotactic effect can be highly specific, even between closely related cell types as shown in (Agerberth, Charo et al. 2000), suggesting the presence of a specific receptor.

In contrast with the direct antimicrobial role of AMPs, immune regulation and chemotaxis seem to require a specific receptor and a specific protein folding. Several studies have tried to

identify the common receptor for AMPs, but with little success. Different studies show different targets for AMPs in the membrane of cells; the effects of some of the peptides have been reduced by the use of blocking antibodies for CCR6 (Yang, Chertov et al. 1999), in many cases the suggested receptor is a G-protein coupled receptor, such as formyl peptide receptor-like 1 (FPR1-1) (Yang, Weiss et al. 2000). Due to the importance of PRR signalling for AMP expression, it has been suggested that some AMPs could regulate PRR activity (Bowdish, Davidson et al. 2005, Mookherjee, Wilson et al. 2006, Morioka, Yamasaki et al. 2008). One of the first examples of this new role for the AMPs is (Mookherjee, Brown et al. 2006). This study aimed to determine if the presence of an antimicrobial peptide affected the outcome when cells are exposed to LPS that will activate the TLR pathway response through TLR4.

Previous research showed how AMPs could modulate the cell response to LPS, but the most widely accepted explanation was the direct interaction of AMPs and LPS; the peptides were able to bind to the LPS, blocking its ability to activate TLR4 (Motzkus, Schulz-Maronde et al. 2006). But the previously mentioned study (Mookherjee, Brown et al. 2006) was able to prove that AMPs have a more complex role than the direct interaction and blocking of the LPS. In a first experiment, cells exposed to a small amount of LPS, enough to produce a response, were treated with different doses of the cathelicidins, to determine whether there is any dose-dependent effect, the outcome of the experiment is determined by the levels of TNFα produced by the cells. In normal conditions, LPS produces an increase of TNFα production but in the presence of the cathelicidin, the production is reduced compared with the LPS-only positive control, showing as well a dose-dependent effect; the more peptide presence in the media, the bigger the reduction in TNFα production. This raises a key question of how AMPs regulate the TLR pathway activated by the LPS? Later studies have confirmed these results showing the regulatory abilities of AMPs over the immune response (Funderburg, Lederman et al. 2007, Morioka, Yamasaki et al. 2008). There are several models of interaction to explain these observations suggested in (Lai and Gallo 2009) (Fig. 1. 9):

- In trans-activation, the AMP can interact with the membrane inducing the removal of molecules originally attached there; removed molecules could interact with a membrane receptor. These models have been proposed for the epidermal growth factor (EGF) receptor (Filardo 2002).

- Membrane disruption, in which, in a similar fashion, the AMPs interact with the membrane; in this case, removing or inactivating the receptor, avoiding its activation and blocking the signalling pathway. This could explain the inhibition of TLR4 by LL-37.
- Alternate ligands, AMPs act as ligands for membrane receptors, as suggested for CCR6 activation by LL-37 (Mookherjee, Brown et al. 2006).

Another recently discovered role of AMPs is their ability to promote wound healing, for example human β-defensin 2 (hBD2) enhances cell migration and cell proliferation inducing the tube formation on endothelial cells in wounds and accelerating the wound closure (Baroni, Donnarumma et al. 2009). All these roles and functions suggest the importance of AMPs as part of the immune response at different levels: antimicrobial, signalling and wound repair.

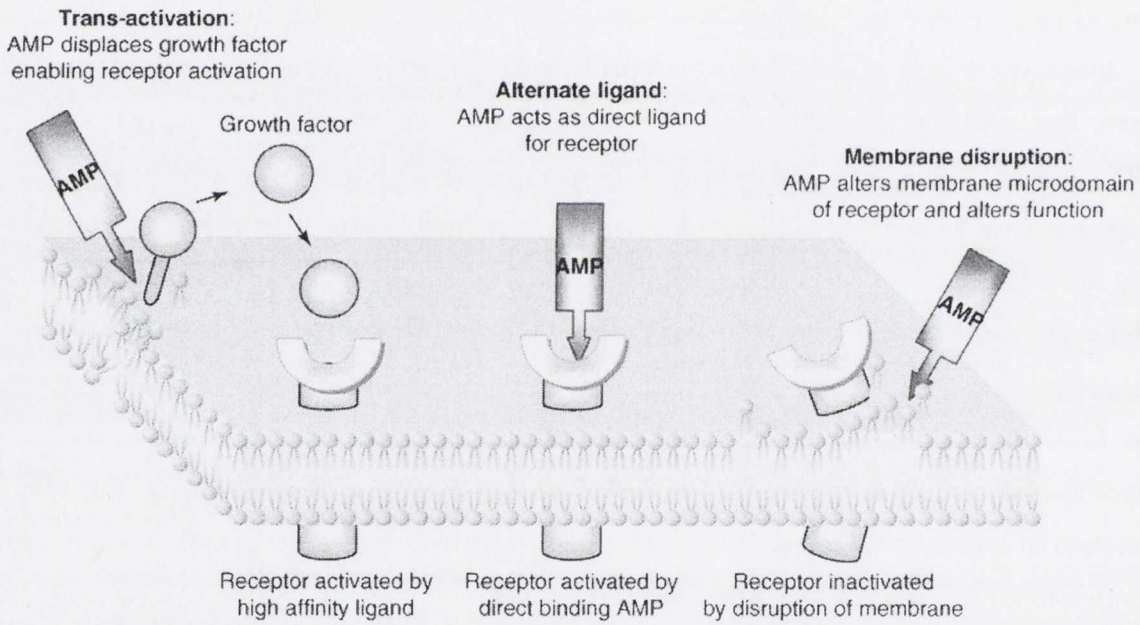


Figure 1. 9. Models of AMPs mechanism of action as immune regulators.

Different models of how an AMP might act as immune regulator have been proposed: AMPs could act by removing a surface molecule. Then, once free could activate a high affinity receptor in the cell surface that will transmit a signal cascade into the cell; a second hypothesis suggests a direct interaction of the AMP with a surface cell receptor; a third hypothesis, suggest the AMP could remove surface receptors blocking signalling into the cell. Adapted from (Lai and Gallo 2009).

1.4. DEFENSINS

Defensins are small cationic peptides, less than 100 amino acids in length, with broad spectrum antimicrobial activity that includes Gram positive and Gram negative bacteria, fungi and some viruses (Ganz, Selsted et al. 1985, Jenssen, Hamill et al. 2006). Defensins are characterised as highly variable in sequence among members of the group with the exception of 6 highly conserved cysteines with a conserved and characteristic length of peptide sequence between them (Fig. 1. 10).

These cysteines form 3 disulphide bonds that generate the so-called defensin motif: a triple stranded β-sheet structure held together by the bonds between cysteines (Ganz and Lehrer 1995). These secondary and tertiary structures are highly conserved across species despite the differences in protein sequence, defensins and defensin like proteins can be found in bacteria (Gao, Rodriguez et al. 2009); in plants, originally named thionines (Mendez, Moreno et al. 1990); invertebrates (Lambert, Keppi et al. 1989) and vertebrates such as chicken (Lynn, Higgs et al. 2004) and cattle (Cormican, Meade et al. 2008). The origin of defensins is not completely understood but a recent hypothesis suggested vertebrate defensin could be originated from arthopod "big defensins" (Zhu and Gao 2013). "Big defensins" contain the classical defensins motif as well as an additional N-terminus antimicrobial motif not present in vertebrate defensins. This motif could have been removed from the ancestral vertebrate defensin by exon shuffling or intronization of that coding region. The evolution of defensins in vertebrate seems to respond to episodes of gene duplication and episodes of positive selection. Differences in selection pressures, such as contact with pathogen not in contact for other lineages could explain the difference in the defensins repertoire between the rodents and other groups such as primates.

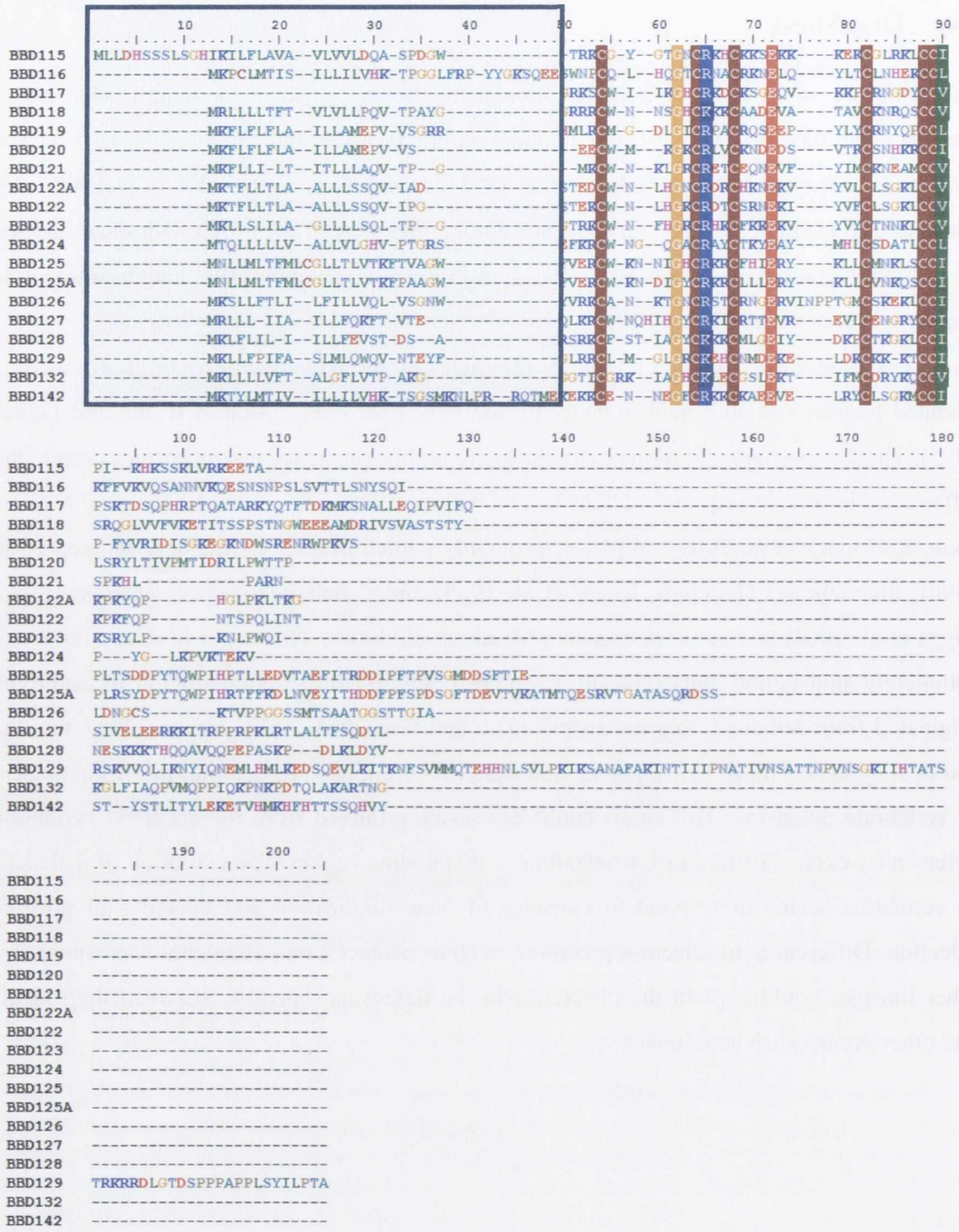


Figure 1. 10. Protein alignment of Chromosome 13 bovine β-defensin cluster

Highly conserved amino acids are marked with a colour background. This includes the 6 cysteins part of the defensin motif. Blue square mark the leader or signal peptide sequence that will be remove in the mature peptide after during the secretion process. Gaps between sequences are introduced to adjust for the different peptide lengths. Sequence alignment was carried out using T-coffe and graph was obtained using Bioedit software.

The pattern in which these cysteines form the disulphide bonds allows us to classify the defensins in three groups (Fig. 1. 11):

- **α-defensins**: are generally produced by neutrophils and in the intestinal epithelium by Paneth cells (Ayabe, Satchell et al. 2000, Cunliffe 2003). There are indications of their function as regulators of the intestinal micro-flora equilibrium (Shi, Aono et al. 2007). They are also reported to have antiviral functions (Zhang, Yu et al. 2002, Salvatore, Garcia-Sastre et al. 2007). α-Defensins have been found in primates, glires and more recently in most mammals (Lynn and Bradley 2007). Some α-defensin like peptides can be found in invertebrates (Oren 2005). The 6 cysteines form bonds in a 1-6, 2-4, 3-5 pattern (Fig. 1. 11). In addition to the cysteines, several amino acids (an arginine, a glutamic acid and two glycines) and the length of peptide sequence are found to be well conserved: $X_{1\text{-}2}$ C X C R $X_{2\text{-}3}$ C X_3 E X_3 G X C X_3 G X_5 CC $X_{1\text{-}4}$. Gene structure of α-defensins is formed by 2 or 3 exons; in the case of 3 exons, the first exon is transcribed as an untranslated region. The first exon, in the two exons case, or the second exon, in the three exon case, is translated into the signal peptide and an anionic pro-segment (Ganz 2003). The final exon is translated as the mature peptide. The enzyme that cleaves the signal peptide varies among different species; matrilysin has been identified as the enzyme responsible for cleavage of α-defensins in mice: in humans some reports have identified a metalloproteinase as responsible for the processing (Wilson, Ouellette et al. 1999); in other cases trypsin has been suggested (Ghosh, Porter et al. 2002).

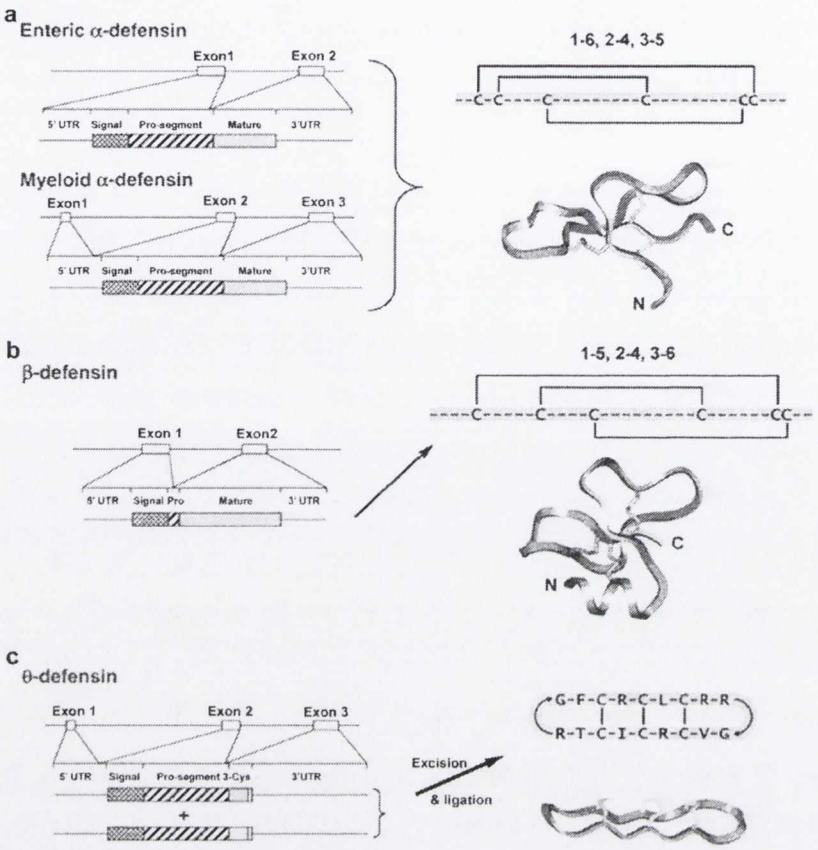


Figure 1. 11. Defensin subclasses

Defensins can be divided into subclasses: α-defensins genes can be formed by 2 or 3 exons; cysteines of the defensin motif are bound in a 1-6, 2-4, 3-5 pattern and are found mostly in the gut. β-defensins genes are formed by 2 exons and are the most abundant and broadly distributed; the cysteines are bound in a 1-5, 2-4, 3-6 pattern. And finally θ-defensins genes are formed by 3 exons of truncated α-defensins; 2 α-defensins ligate to each other through their cysteines forming a cycled peptide. Adapted from (Selsted and Ouellette 2005).

- **β-defensins:**the disulfide bonds have a pattern 1-5, 2-4, 3-6 (Fig. 1. 11). The peptide length between cysteines is conserved and can be described in a formula: X_{2-10} C X_{5-6} G/A X C X_{3-4} C X_{9-13} C X_{4-7} CC X_N. β-Defensin genes have a conserved structure of only two exons. The first exon is translated into the signal and most of the pro-segment; the second exon is translated into the mature peptide and, in some cases, a small section of the pro-segment. As with α-defensins, the immature peptide has to be processed after translation, but no evidence for a common enzyme for all β-defensins has been found. β-Defensins are found in most vertebrate species, mostly expressed by epithelial tissues, including airway,

nasolacrimal, mammary gland and urogenital tract. β-defensin gene expression is found at particularly high levels in epididymal regions in the male reproductive tract. The number of β-defensins genes varies in different species. At the moment of writing, 35 genes have been described in the human genome, 33 in chimps, 38 in dogs, 45 in mouse (Patil, Cai et al. 2005) and 57 in the bovine genome (Cormican, Meade et al. 2008). Despite the variation in gene numbers, all species share a similar gene distribution. In (Patil, Cai et al. 2005), a nomenclature was established for the different clusters to allow a better comparison among species; cluster A is found in Chr 8 in mouse, Chr 8 in humans and Chr 27 in bovine; cluster B is found in Chr14 in mouse, Chr 8 in humans and Chr 8 in bovine; cluster C is found in Chr 1 in mouse, Chr 6 in human and Chr 23 in bovine and finally cluster D is found in Chr 10 in humans, Chr 2 in mouse, Chr13 in bovine and in contrast with the 4 clusters found in most mammals, in primates, cluster D is divided into two separate clusters (Cormican, Meade et al. 2008).

- **θ-defensins**: these rare peptides were discovered in the rhesus macaque, and are present only in primates; in humans the only gene found is truncated by a stop codon creating a pseudo gene (Fig 1. 11). It has been suggested that they originated from a mutated α-defensin with a premature stop codon. These truncated peptides could form dimmers by the formation of 3 disulphide bonds among the cysteines of both peptides forming a cyclic structure. They are known to have an antiviral role (Cole and Lehrer 2003).

1.5. β-Defensins and Reproduction

For many years, our group has had a major interest in the innate immunity response in several species. One approach has been the *in-silico* discovery of innate immunity related genes such as defensins; this approach was used to investigate the avian immune response, in particular in chicken (Lynn, Higgs et al. 2004, Higgs, Lynn et al. 2005, Higgs, Lynn et al. 2007) and in cattle (Cormican, Meade et al. 2008, Meade, Cahalane et al. 2008, Meade, Higgs et al. 2009). In addition, we found high levels β-defensin expression during uterine infection postpartum (Chapwanya, Meade et al. 2009), embryonic development in chicken eggs (Meade, Higgs et al. 2009) and in the bovine male reproductive tract under normal healthy conditions (Cormican, Meade et al. 2008), which has suggested a possible role in development and reproduction. In the last two cases, the absence of infection suggests that the AMPs have other functions independent of their antimicrobial effect. There is some evidence that these

peptides have an additional function in the reproductive tract, several of them have been found to be expressed in the male reproductive tract of mice and rats with a very specific pattern through the different sections of the tract (Com, Evrard et al. 2003, Jelinsky, Turner et al. 2007). In primates BBD126 is required for the correct maturation, known as capacitation, for the sperm cell motility through the female reproductive tract and adherence into the female tract, a necessary step during fertilisation (Yudin, Generao et al. 2005, Tollner, Yudin et al. 2008, Yudin, Tollner et al. 2008). A SNP in human Defb126 was correlated with reduction in fertility indicating the importance of the protein in reproduction (Tollner, Venners et al. 2011). As previously discussed, our group has found a new cluster, in chromosome 13 comprised of 19 new defensins, in the bovine genome (Cormican, Meade et al. 2008). In this study, open reading frames similar to the characteristic β-defensin domain formed by six cysteines, were used to find potential new genes in the bovine genome. 57 β-defensins were predicted; 19 of them were found forming a cluster in chromosome 13. The expression of these new genes was confirmed by qPCR in a panel of tissues, finding preferential expression of these AMPs in the reproductive tract of a bovine male in the absence of infection (Fig. 1. 12.).

We propose that these novel defensins could have several possible roles in the male reproductive tract:

- Antimicrobial function: a recent study (Com, Bourgeon et al. 2003) characterised the presence of β-defensins expression in the male reproductive tract and their ability to eliminate pathogens. AMPs act as a line of defence against pathogens in the male tract. The AMPs characterised in the study are expressed in different tissues outside the testis. Protection of the reproductive tract and gonads against pathogens in both male and female reproductive tract is essential for the successful survival of a species. The male reproductive tract can be susceptible to retrograde infection from pathogens present in the vas deferens or via blood and lymphatic vessels. In any case, an infection could cause damage for the endocrine and reproductive functions of the male reproductive tract leading to temporary or permanent sterility.

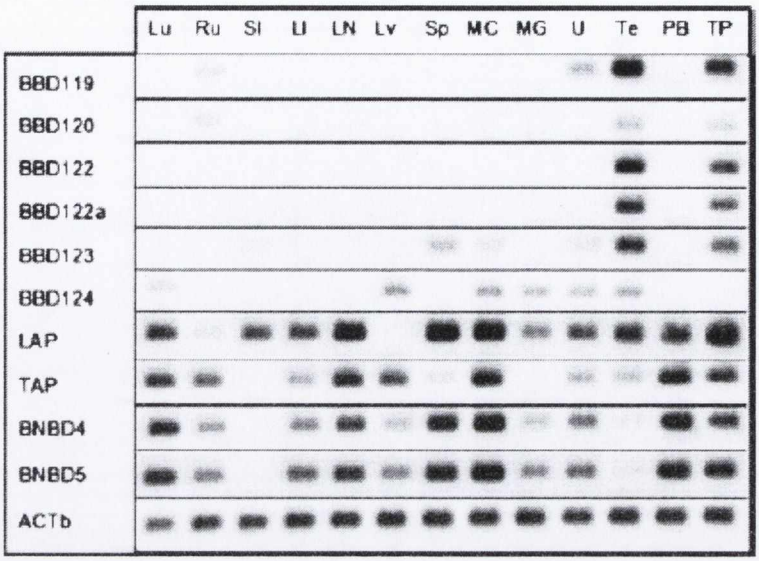


Figure 1. 12. Tissue expression profile of bovine β-defensins.

The expression of six novel (BBD119–BBD124) and four known (BNBD4, BNBD5, LAP and TAP) AMPs are shown. Lu Lung, Ru rumen, SI, small intestine, LI large intestine, LN lymph node, Lv liver, Sp spleen, MC mammary cells, MG mammary gland, U uterus, Te testis, PB peripheral blood mononuclear cells and TP tissue pool reference cDNA from healthy cattle. Adapted from (Cormican, Meade et al. 2008).

- Immune regulatory role: seminal plasma contains a great number of components, some of which have a direct effect on the female tract when they come into contact. It is known that the presence of seminal plasma accelerates ovulation (Waberski, Claassen et al. 1997) and creates an influx of cells into the uterine walls, increasing the number of MHC class II molecules (Waberski, Dohring et al. 2006). However, the exact pathway or mechanism is not known. Seminal plasma contains several chemokines, such as IL-8, that induce the influx of neutrophils into the uterus, a hypothesis which suggests that these phagocytic cells are required for the selection of healthy sperm cells (Schuberth, Taylor et al. 2008). The role of AMPs in this process is at the moment unknown but the recently described ability of AMPs to modify and regulate the immune response indicates that defensins will probably have a role in this process.
- Sperm maturation, β-defensins promote the maturation of the sperm through the male tract. This role has been already described in a mouse model (Zhou, Zhang et al. 2004). A defensin-like peptide known as Bin 1b is expressed in the initial region of the epididymis, called the head or caput and shows the ability to initiate the sperm maturation process by

inducing progressive motility in previously immotile sperm. Bin 1b activates Ca2+ channels in the sperm membrane, and the accumulation of Ca2+ signals the initiation of sperm motility. Other defensins could share a similar role in the maturation of the sperm during the transition along the epididymis.

- Sperm capacitation is a critical step in the development of sperm to fully activate it for the fertilization of the oocyte. In mammals, post-coitus, sperm temporarily attach to the oviduct epithelium creating the so-called sperm reservoir; the exact location varies according to the species. It is located in the anterior vagina in cattle, sheep, rabbits and primates and in the uterus in pigs, horses, dogs and some rodents. The mechanism that induces the binding of the sperm is not fully understood but several studies have found proteins in different species involved in the process; in cattle, the most common proteins coating the sperm are bovine seminal plasma family proteins (BSP), such as BSP-109 (Gwathmey, Ignotz et al. 2003) and small cationic proteins; in pigs, it is members of the spermadhesin family such as AQN1. In both cases, the interaction with the oviduct epithelium occurs through recognition of polysacharides in the female tract, fucose in cattle and mannose-galactose in pigs. In both cases disulfide-reductans will induce the release of the sperm. The formation of the sperm reservoir has a great importance at two different levels. In the first place, capacitation of the sperm has been described as a destabilization process; the sperm binding will delay the capacitation increasing the sperm life span, maintaining the sperm viable for a longer period of time. Oviduct sperm binding combined with a gradual release of the sperm induced by ovulation creates a mechanism that reduces the risk of polyspermia by reducing the amount of sperm that can have access to the oocyte. These mechanisms are key for sperm selection, allowing only the higher quality cells to reach the oocyte. During capacitation the sperm will be released from the reservoir form in the oviductal epithelium and the acrosomal head of the sperm will be modified enhancing its fusigenic potential towards the oocyte. The modification of the membrane changes the Ca++ influx, enhancing the motility of the cells. This function has been linked to the presence β-defensin 126 in the sperm cell surface and plasma in macaque (Tollner, Yudin et al. 2008). This defensin is present in different mammals and could conserve a similar function as shown for macaque.
- Egg penetration: the exact mechanism by which the sperm penetrate the egg membrane is not fully understood but we know that AMPs have the ability of direct interaction with membranes by pore formation and detergent like processes and interaction

with cell membrane receptors. In both cases, AMPs could interact with the egg to aid the sperm during fertilisation.

Any of these roles could have important implications in the reproductive potential and fertility of an individual. Could genetic variability explain differences in fertility? Most studies for reproductive immunology in animals emphasise the female perspective and tend to ignore the male but the presence of these AMPs could indicate that components of the male immunity could affect how the female immune system responds during fertilisation. From an animal stock perspective, this is a very important question. If these genes have a role in fertility, there is the possibility of using them to improve the reproductive potential of animal stock. These defensins could be used as markers for fertility in breeding programmes.

1.8 BOVINE BREEDING

Bovine domestication has been dated back to 10,000 BC, the advantages being a constant and reliable source of nutrients and tractive power, etc which drove forward the revolution in early Neolithic populations. With the spreading of the practice to new environments and populations with different requirements, the selective breeding process began. Classically, breeding was based on the selection of phenotypically desirable individuals for reproduction which introduced an increase in the selected trait over multiple generations. This process was enhanced after World War II by the incorporation, first of modern fertility techniques, artificial insemination, embryo and semen storage, and secondly by the incorporation of genetic information for the selection of individuals.

Breeding programmes have aimed to increase production traits like milk yield, growth or meat properties, depending of the breed. The use of modern techniques accelerated the increase in these traits, for example milk production has increased in average of 1.5% per year in the last 50 years (D.A 1993). But a negative effect has been observed, as there is a strong correlation between high milk yield and poor fertility performance. As an example, in the UK, two populations of HF separated by 15 years showed a reduction of calving rates from 55.6% in 1975-1982 population to 39.7% in a 1995-1998 population (Royal, Darwash et al. 2000). Different reasons have been suggested to explain this effect, most notoriously that an energy imbalance has been caused by the metabolic stress which high production cows suffer which reduces the ability to successfully impregnate the cows at a desired time. In

addition, the effects caused by the hormone levels that control both lactation and pregnancy may have an important role in the reduction of fertility.

In contrast with the intensive production breeding model, alternative models have emerged over the years, which use a more comprehensive approach; an example is the breeding programmes started in the early 1970's in Scandinavia for dairy breeds. In addition to milk yield, a set of traits were introduced in the programmes, including fertility, immune response and others (Fig. 1. 13). This programme has generated an improvement of fertility as shown by a 0.04% increase per year for 56-days non-return rate, defined as the proportion of heifers not seen to come back into estrus after breeding, over a 20 years period (IM, Heringstad et al. 2003).

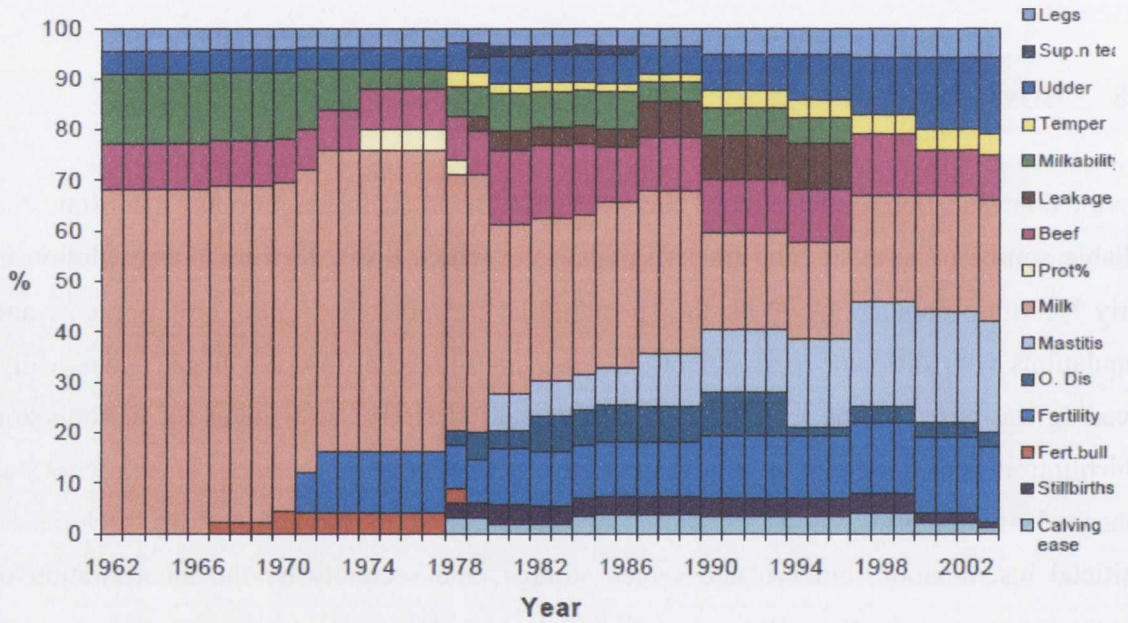


Figure 1. 13. Graph representing the percentage of selection of a particular trait in NR breeding programmes since 1962.

Number of selection traits in classical programmes included a reduced number of traits, generally related to yield production. Modern programmes include numerous traits that include animal health, fertility and others. This approach has successfully reversed the reduction in fertility and immune response observed with classical programmes. Adapted from http://www.genoglobal.no/.

1.9 BOVINE AMPS

Bos taurus was one of the first species studied in the search for β-defensins and β-defensins-like genes (Ryan, Rhodes et al. 1998). By using classical EST-based bioinformatics techniques, 18 genes were eventually discovered. This method has the negative effect of

under-representing lowly expressed genes as is the case with β-defensins in EST libraries from healthy samples. In other species, where a whole or partial genome was available, a higher number of genes was found, for example 35 in humans or 33 in *Pan troglodytes* (Patil, Cai et al. 2005).

After the bovine genome was released, more sensitive *in-silico* screening allowed searches for previously unknown genes (Cormican 2009). Two bioinformatics models were used in parallel, first homology Blast search to find homologous proteins to the previously discovered in other species were used. A second approach used the Hidden Markov model, to find new bovine defensins. Using this approach 57 open reading frames with defensin like structure were found. A similar genomic distribution was found as in previous species; 4 regions in different chromosomes contain all the defensin genes, in bovine these regions are found in chromosome 8, 13, 23 and 27. In most mammalian species the same numbers of regions or clusters were found; there are some exceptions in humans and chimpanzee, in which 5 regions were found. Using the previously described nomenclature for other species (Schutte, Mitros et al. 2002, Patil, Cai et al. 2005), the bovine clusters were named A for the cluster in chromosome 27, B for the cluster on chromosome 8, C for cluster on chromosome 23 and D for the cluster on chromosome 13.

Homology for cluster A genes have been found on the only β-defensin cluster found in birds (chicken and zebra finch) (Lynn, Higgs et al. 2004), suggesting this was the original cluster, and that it has evolved by several cycles of duplication and translocation across different chromosomes. Despite the similarities, the bovine cluster A presents some differences when compared to other species; the bovine genome forms a larger cluster spanning for 1.9 Mbases compared with the shorter one in humans (639 Kbases) or the dog (356 Kbases). Bovines also possess a larger number of genes, 30, similarly to rat (24) or mouse (29), compared with the more compact human (11 genes) or dog (15 genes). In humans, rats and mice, the only α-defensins gene cluster is inserted in cluster A (Fig. 1. 14). At least 16 of the bovine cluster A genes are species-specific and only described in bovine but not in humans, rats or mice; though some of these genes have been found in other artiodactyla, suggesting a common origin in a common ancestor (Cormican 2009).

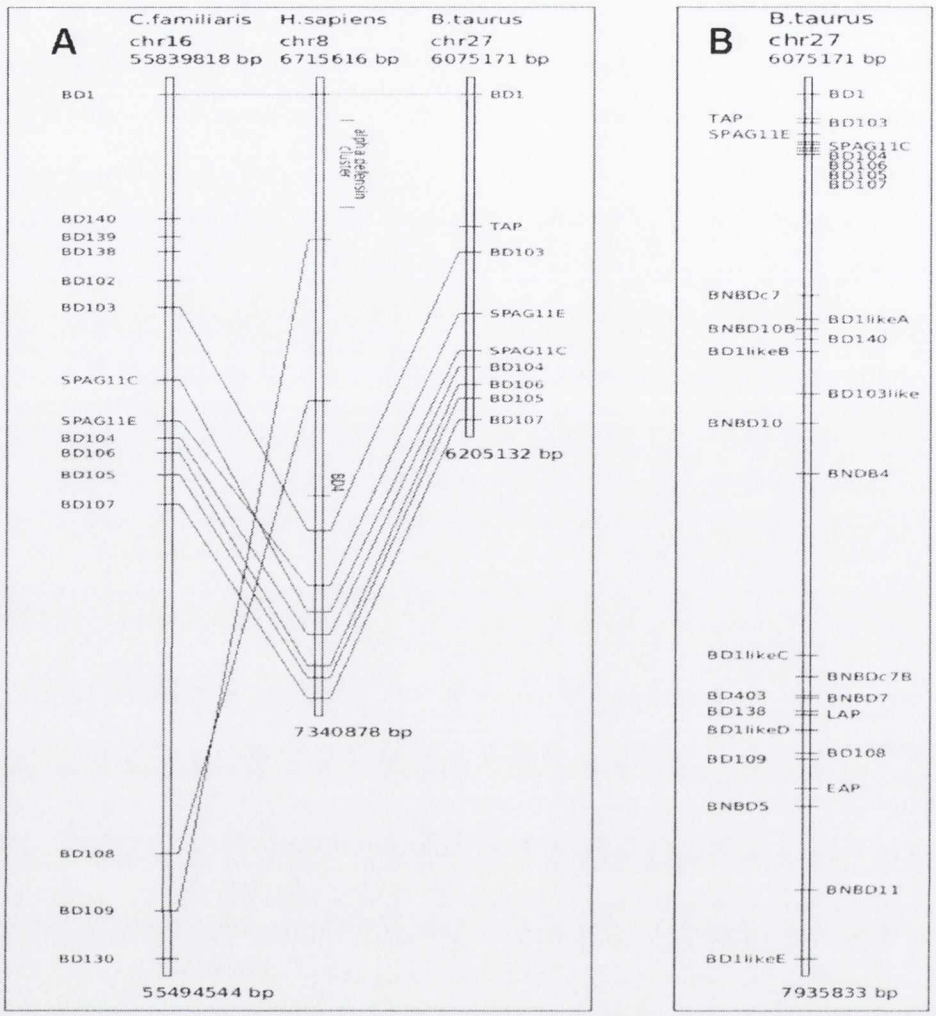


Figure 1. 14. Genomic maps alignment of β-defensin cluster A from dog, human and bovine genome

Bovine cluster is represented partially, only genes with orthologous genes in the other two species are represented. B, full bovine cluster representing all β-defensin genes including genes not present in other species. Adapted from (Cormican, Meade et al. 2008).

Clusters B and C are formed by 4 and 5 β-defensin genes each (Fig. 1. 15). These clusters most probably originated from the split of a single cluster into two. This is suggested by the presence of a single cluster in opossum, indicating that the split happened after the separation of marsupials from the rest of mammals. Cluster B and C present the lowest variability between species, showing neither gain nor loss in the species studied so far.

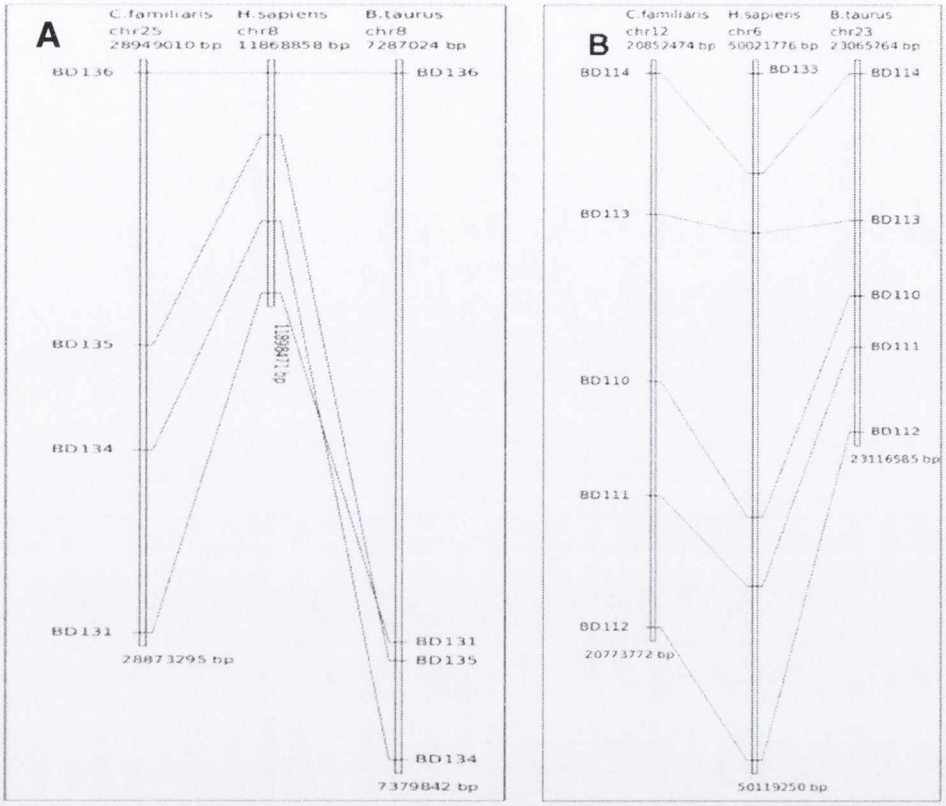


Figure 1. 15. Genomic maps alignments of β-defensins gene clusters B and C.

A, Alignments of genomic maps of cluster B from dog, human and bovine genome; B, Alignments of genomic maps of cluster C from dog, human and bovine genome. Adapted from (Cormican, Meade et al. 2008).

It is suggested that this cluster was established after the separation of metatheria and eutheria (Cormican 2009). The cluster seems established in most analysed species with the exception of a gene gain in the bovine genomes (*BBD125a*) and a gene loss, by the presence of a STOP codon in the middle of one gene (*BBD122a*) and the split and change of orientation of part of the human cluster (Cormican, Meade et al. 2008) (Fig. 1. 16). In bovine, eighteen full length β-defensin genes (typical 2 exon structure) and one partial gene (with only a second exon) were found. Similar gene numbers are found in human and dog with the exception of (*BBD125*a and *BBD122a*).

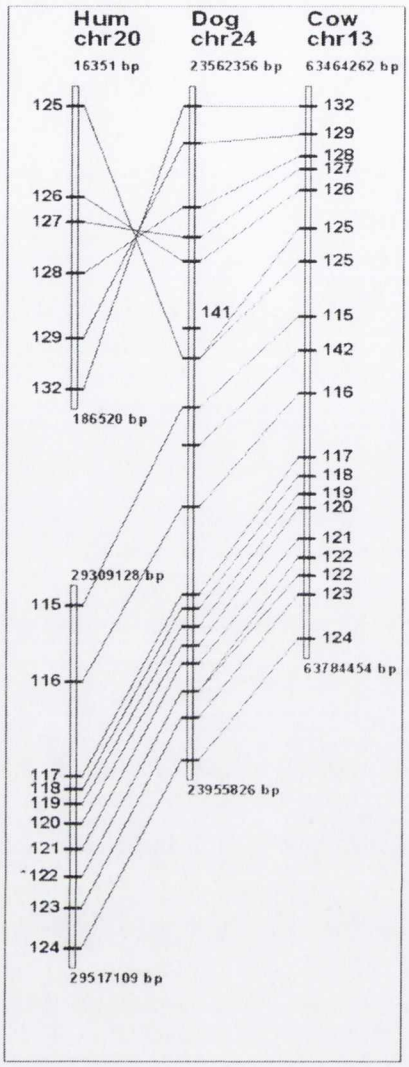


Figure 1. 16. Genomic map alignments of β-defensins gene cluster D in dog, human and bovine genome.

Cluster D is the second largest by number of genes. In primates, this cluster has split into two different clusters in the same chromosome. In addition, the splinted cluster E has suffered a change in orientation. Genes of this cluster are found almost exclusively in the reproductive tract. Adapted from (Cormican, Meade et al. 2008).

Earlier studies by our group had confirmed the antimicrobial activity of at least one cluster D gene, *BBD123*. This peptide was commercially synthesised and used in antimicrobial assays and showed strong antibacterial activity (Cormican, Meade et al. 2008). Primers were designed for a selected group of members of cluster D (*BBD119, BBD120, BBD122,*

BBD122a and *BBD123*). qRT-PCR analysis of gene expression showed a different pattern of expression compared with other β-defensin genes localised in the testis (Fig. 1. 10), in contrast with the broad distribution of most defensins. These results indicate that these genes may have functions in addition to the antimicrobial activity normally suggested for β-defensins; in particular, we aim to explore the possible role of these defensins in reproduction.

Table 1.1. Summary of all bovine β-defensins, molecular weight and isoelectric point for the full peptide and the mature peptide (after the removal of the signal peptide). In addition the length of the peptide tail, measure as the number of amino-acids present in a peptide between the last cysteine part of the defensin motif and C terminus. Grey background marks the Chr.13 β-defensin cluster members

	Full peptide		Mature peptide		
	kDa	PI	kDa	PI	Tail
104	7.2	9.81	5.2	9.64	8
BBD105	9.0	9.06	6.8	8.73	5
BBD106	7.5	8.53	5.8	8.24	13
BBD107	8.0	9.12	6.0	8.92	6
BBD108	5.9	5.13	5.9	5.13	17
BBD109A	7.9	9.3	6.4	9	27
BBD109	7.9	9.63	6.1	9.6	27
BBD110	7.5	7.56	5.3	6.35	1
BBD111	7.8	9.63	5.6	9.5	3
BBD112	10.7	9.19	8.9	9.05	10
BBD114	7.6	8.17	5.0	7.75	8
BBD113	10.2	8.81	8.6	8.71	28
BBD115	9.2	9.7	6.5	9.73	17
BBD116	11.4	9.06	9.7	9.01	34
BBD117	8.5	9.57	8.5	9.57	41
BBD118	10.5	9.07	8.5	8.86	40
BBD119	9.8	9.05	7.7	9.07	28
BBD120	8.3	8.63	6.2	8.64	22
BBD121	6.9	8.88	5.1	8.64	10
BBD122A	7.7	9.02	5.9	8.84	16
BBD122	7.7	9.24	5.9	9.08	16
BBD123	7.9	9.94	6.0	9.87	14
BBD124	7.6	8.84	5.8	8.88	13
BBD125	12.4	5.86	10.4	5.85	47
BBD125A	14.6	8.17	12.7	8.22	67
BBD126	9.4	9.33	7.3	9.18	30
BBD127	10.3	9.54	8.3	9.24	32

	Full peptide		Mature peptide		
	kDa	PI	kDa	PI	Tail
BBD128	9.3	9.22	7.6	9.24	29
BBD129	19.3	9.56	17.2	9.49	116
BBDB130	8.8	9.54	7.1	9.38	19
BBDB131	7.8	9.36	6.3	9.26	12
BBD132	9.3	9.61	7.7	9.5	32
BBD134	7.6	4.8	6.3	5.06	7
BBD135	8.6	9.75	7.0	9.59	12
BBD136	8.3	9.1	6.7	8.64	14
BBD138A	6.9	9.54	5.2	9.54	2
BBD138	7.0	9.65	5.3	9.42	2
BBD140	7.4	9.35	6.0	9.07	5
BBD142	11.2	9.2	9.4	9.1	30
SPAG11E	9.2	8.22	7.7	7.77	20
spag11C	15.5	9.73	13.7	9.66	31
BT402	7.1	10.03	5.3	9.63	25
EAP	7.1	10.66	5.3	9.99	3
TAP	7.0	10.26	5.2	10	3
BNBD10	6.9	10.74	5.3	10.21	1
BNDB4	7.2	11.46	5.4	11.17	2
BNBD11	6.5	10.66	4.7	9.99	3
LAP	7.0	11.24	5.3	10.85	3
sbtbd1	7.5	9.69	5.9	9.3	2
BNBD7	7.0	11.47	5.2	11.2	1
BBDB403	11.0	10.79	9.2	10.49	32
BNBD8	6.9	11.2	5.2	10.66	1
BNBD12	6.5	9.99	4.7	9.56	3
BNBD9	7.0	11.2	5.3	10.66	1
BNBD6	7.2	11.17	5.4	10.74	3
BT300	7.6	9.93	5.4	9.92	4
BNBD14	4.1	9.32	4.1	9.32	25

1.10 AIMS OF THE STUDY

This study aims to characterise bovine Chromosome 13 β-defensin cluster in bovine the male and female reproductive tract. To this end, several approaches are being explored:

- Further understanding of the gene expression patterns of all members of the gene cluster in several regions of both male and female reproductive tracts
- Characterisation of the mechanism of gene regulation of these genes.
- Characterisation of the genetic variability of the cluster by identifying polymorphism between two commercial bovine breeds and effect of these in the reproduction.
- Creation of tools for functional characterisation of β-defensins and their possible role in reproduction.

CHAPTER 2: MATERIALS AND METHODS

2.1. SAMPLE COLLECTION

Tissues were collected at a local abattoir from cattle within 30 minutes of slaughter. Six male genital tracts were retrieved from 18 month old Holstein-Friesian bulls and dissected to obtain 3 mm sized tissue samples of seminal tubule, rete testis, caput, corpus and cauda of the epididymis and vas deferens (Fig. 2. 1). Similar samples (without vas deferens) were retrieved for comparative purposes from a 43 day old bull calf after castration. The genital tracts from 10 commercial beef breed cows were retrieved after examination of several reproductive tracts to acquire 5 tracts in the luteal phase and 5 tracts in the follicular phase of ovulation. 3 mm sized tissue samples were collected from the ovary, proximal Fallopian tube, distal Fallopian tube and uterine body of each animal (Fig. 2. 1). The same dissection technique was used to retrieve samples from different organs for the tissue expression panel (lung, rumen, small intestine, large intestine, liver, spleen, lymph node and mammary gland). Several tissue fragments were collected for each sample. All samples were immediately flash frozen in liquid nitrogen and subsequently stored at -80°C. In addition extra material was preserved in 37% formaldehyde solution (Sigma Aldrich, St. Louis, MO, USA) for immunohistochemistry.

Bovine sperm samples were provided by Dr. Bernard Eivers, National Cattle Breeding Centre (NCBC), Enfield, Co. Meath. Fresh samples for protein analysis and laser scanning confocal microscopy were transported in a water bath at 37° C to minimise cell death. Frozen sperm straws were used for DNA extraction and sperm penetration assay. Straws were stored and transported in liquid nitrogen.

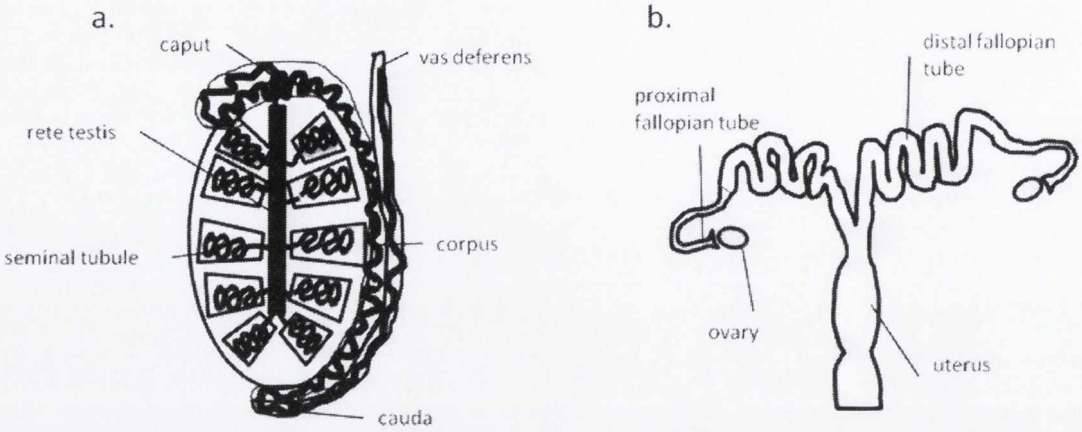


Fig. 2. 1 *Male and female reproductive tract*

a. Schematic representation of sampled areas across the male reproductive tract: seminal tubule, rete testis, caput, corpus and cauda of the epididymis and vas deferens. b, Schematic representation of sampled areas across the female reproductive tract: ovary, proximal Fallopian tube, distal Fallopian tube and uterine body.

2.2. RNA ISOLATION

Total RNA was extracted using a phenol-chloroform extraction protocol. Frozen tissue samples were immediately homogenised using a mechanical homogenator in 600 µl of Trizol (Invitrogen Ltd., Paisley, UK). It was important to keep samples frozen at all times until homogenised, to minimise degradation. 400µl of Trizol were added to the homogenised sample. Samples were then incubated for 5 minutes at room temperature. Tubes were spun down at 12,000g for 10 minutes at 4°C to precipitate any unhomogenised particles. The supernatant was transferred to a clean tube and 100 µl of chloroform (Sigma Aldrich, St. Louis, MO, USA) were added. Samples were vigorously mixed for 15 seconds; tubes were incubated for 3 minutes at room temperature. Samples were then centrifuged at 12,000g for 15 minutes at 4°C. Samples were separated into 3 phases: an aqueous phase containing RNA and some DNA, an interphase containing DNA and an organic phase containing mostly proteins and lipids. 400µl of the aqueous phase were recovered into a new tube. To this sample 500µl of isopropanol (Sigma Aldrich, St. Louis, MO, USA), 10µl of glycogen (Sigma Aldrich, St. Louis, MO, USA) and 10µl of 100mM sodium acetate (Sigma Aldrich, St. Louis, MO, USA) were added. Samples were then vortexed for 30 seconds and incubated for 30

minutes at -80°C. After incubation, samples were centrifuged at 12,000g for 10 minutes at 4°C. A small white pellet was formed, and the supernatant was removed without disrupting the pellet. The pellet was then resuspended in 1ml of 75% Ethanol (Sigma Aldrich, St. Louis, MO, USA). The samples were centrifuged again at 10,000g for 5 minutes at 4°C. Supernatant was removed and the pellet was air dried for 2 minutes. The RNA pellet was dissolved in 30µl of RNAse DNAse H_2O. Samples were then incubated at 70°C for 5 minutes to dissolve RNA and stored at -80°C. RNA quality was evaluated by agarose gel electrophoresis where the presence of ribosomal RNA, 28s and 18s, was detected. RNA quantity was assessed using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

2.3. cDNA SYNTHESIS

Omniscript (Qiagen, Crawley, UK) First Strand Synthesis kit and oligo dT primers (Qiagen, Crawley, UK) were used for cDNA synthesis. 2µl of 10x reaction buffer, 8 ng/µl of Oligo dT, 0.25 mM dNTP and 1µl of Omniscript reverse transcriptase enzyme were used in a final volume reaction of 20 µl with 1 µg of RNA. Volumes were calculated using the RNA quantity determined by the Nanodrop (Appendix: Tables 1 and 2). The reaction was incubated at 37°C for 1 hour enzyme was then deactivated at 95 ° C for 5 minutes. The cDNA was then quantified using a Nanodrop spectrophotometer. Stock cDNA solutions were diluted to prepare 40ng/µl working stocks.

2.3. PRIMER DESIGN

The qPCR primers were designed using Primer3 software (Rozen and Skaletsky 2000) and commercially synthesised (Invitrogen Ltd., Paisley, UK). Where possible, (*BBD117* has only 1 exon), primers were designed to be intron spanning in order to avoid the effect of genomic DNA presence on the qPCR results (Table 2.1). DNA sequences for primer design for the novel *β*-defensins were obtained by BLAT, at the UCSC genome Browser, http://genome.ucsc.edu/cgi-bin/hgBlat (Kent 2002), against the Bovine genome using the protein sequences retrieved by our group (Cormican, Meade et al. 2008).

Table 2. 1. List of primers and conditions used for gene expression analysis by qPCR in this study.

Gene Symbol	Gene name		Primer Sequence	Amplicon Length (bp)	Genbank accession number	Concentration nM
ACTB	Actin, beta	F-	AGATGACCCAGATCATGTTCGA	126	NM_173979	600
		R-	TGACCCCGTCACCGGAGTCCATCACGAT			
GAPDH	Glyceraldehyde-3-phosphate Dehydrogenase	F-	CTCCCAACGTGTCTGTTGTG	222	NM_001034034	900
		R-	TGAGCTTGACAAAGTGGTCG			
PPIA	Peptidylprolyl isomerase A	F-	CCACCGTGTTCTTCGACAT	155	NM_178320	600
		R-	TCTGTGAAGCAGGAACCTTT			
H3F3A	H3 histone, family 3A	F-	CATGGCTCGTACAAAGCAGA	136	NM_001014389	900
		R-	ACCAGGCCTGTAACGATGAG			
MRPS6	Mitochondrial ribosomal protein S6	F-	CGCACGCTTCCCTATAAGAT	176	NM_001040584	600
		R-	TTCACTTCCTGGGTCAGAGG			
RPS9	Ribosomal protein S9	F-	GATTACATCCTGGGCCTGAA	161	NM_001101152	300
		R-	ATGAAGGACGGGATGTTCAC			
MRPS15	Mitochondrial ribosomal protein S15	F-	GCAGCTTATGAGCAAGGTCGT	151	XM_585783	900
		R-	GCTCATCAGCAGATAGCGCTT			
BBD115	Bovine β-defensin 115	F-	ATGCTGCTGGATCATTCCTC	151	HQ423194	100
		R-	TTTTGCAATGTTTCCTGCAA			
BBD116	Bovine β-defensin 116	F-	TGACCATTTCCATCCTTCTGA	184	HQ423193	100
		R-	AGGCAGCACTTTTCGTGATT			
BBD117	Bovine β-defensin 117	F-	GGCCGAAAATCTTGTTGGAT	153	HQ423192	100
		R-	TCTGGCAGTTGCTTGAGTTG			
BBD118	Bovine β-defensin 118	F-	TGAGACTCCTGCTCCTGACTT	246	HQ423191	100
		R-	AGCCTCTTCTTCCCAACCAT			
BBD119	Bovine β-defensin 119	F-	GTTTCTTGCCATCCTTCTGG	188	EU036209	100
		R-	CATAGAAAGGGAGGCAGCAG			
BBD120	Bovine β-defensin 120	F-	TTTCCTGTTTCTTGCCATCC	118	EU036210	100
		R-	GCAGCGTGTGACACTGTCTT			
BBD121	Bovine β-defensin 121	F-	GTTCCTTCTGATTTTGACCATTACC	150	HQ423190	100
		R-	AGATACACAGCACATAGCCTCATTCT			
BBD122	Bovine β-defensin 122	F-	CTCTCATCCCAGGTCATTCC	155	EU036211	100
		R-	GGCGATGTATTTGGCTGAAA			
BBD122a	Bovine β-defensin 122a	F-	TGCTAACTTTGGCTGCACTG	179	EU036212	100
		R-	TGGTAAGCCATGTGGTTGGTAC			
BBD123	Bovine β-defensin 123	F-	AGCTCCTCTCGCTGATTCTG	152	EU036213	100
		R-	CAGCACAGTTTATTATTTGTGCAGTA			
BBD124	Bovine β-defensin 124	F-	TGACACAGCTGCTTCTGCTT	191	EU036214	100
		R-	TTGACAGGCTTGAGTCCGTA			
BBD125	Bovine β-defensin 125	F-	CTGATGCTGACCTTCATGCT	227	HQ423184	100
		R-	TCCAGTAATGTGGGGTGGAT			
BBD125a	Bovine β-defensin 125a	F-	TTGATGCTGACCTTCATGCT	300	HQ423185	100
		R-	ACCTCATCAGTAAACCCAGAATCG			
BBD126	Bovine β-defensin 126	F-	CCTTCTGGTCCAGTTGGTCT	145	HQ423186	100
		R-	CAGGAGGAACGGTTTTA			
BBD127	Bovine β-defensin 127	F-	TTGCAATTCTGCTGTTCCAG	210	HQ423189	100
		R-	GACGAGGTGGTCTGGTGATT			
BBD128	Bovine β-defensin 128	F-	TCCTCATTATCCTGCTGTTTGA	149	HQ423187	100
		R-	GTTGTTGTACGGCTTGCTGA			

Gene Symbol	Gene name		Primer Sequence	Amplicon Length (bp)	Genbank accession number	Concentration nM
BBD129	Bovine β-defensin 129	F-	GCTACAGTGGCAGGTGAACA	188	HQ423188	100
		R-	GCATATGGAGCATTTCATTTTG			
BBD132	Bovine β-defensin 132	F-	GCTCCTGCTCCTGGTCTTTA	243	HQ423183	100
		R-	TTGCTTTGGCCAACTGTGTA			
BBD142	Bovine β-defensin 142	F-	CCATTGTCATCCTTCTGATCC	150	HQ423182	100
		R-	AGCGCAGTTCAACTTCTTCG			
LAP	Lingual antimicrobial peptide	F-	GACAGCATGAGGCTCCATC	151	NM_173979	100
		R-	CTCCTGCAGCATTTTACTTGGGCT			

A first set of primers for single nucleotide polymorphism (SNP) discovery were designed to amplify the coding regions of the genes and were also determined from the predicted protein sequences using BLAT. A separate set of primers were designed for each exon of the two exons present in β-defensin, with the exception of *BBD117* having one exon, in order to be amplified separately due to the distance between the exon introduced by the intron. Primers designed for this study, and conditions used for the PCR, can be found in the Table 2. 2. A second set of primers used for the genotyping of the novel SNPs was designed according to Sequenom (www.sequenom.com) instructions for their use in their iPlex Gold assay (Table 2. 3).

Table 2. 2. List of primers and conditions used for PCR the coding region of genes in this study.

Gene Symbol		Primer Sequence	Amplicon Length (bp)	MgCl2 Concentration mM	Annealing Temperature
BBD115-E1	F-	TGGAACATGCTGTCCTTGTC	217	2.5	65
	R-	CCCTCCTCTTCATGGTCTTG			
BBD115-E2	F-	GCTTATTCATCTCTTTGTTTATTTTGA	250	2.5	55
	R-	GCTGCCAAAATCCTCTTGTC			
BBD116-E1	F-	GCAACACAACCCTCCATTCT	152	2.5	65
	R-	GAAATTAATATGGCTGTGTTTTTCC			
BBD116-E2	F-	TCTGGAGATAATTGCCCTTTTT	297	2.5	65
	R-	CCAGAGAGGCCAGAACAATC			
BBD117-E2	F-	TCTGGAGATAATTGCCCTTTTT	295	2.5	55
	R-	TCCCCAGTTTATTCTACTTTCTGC			
BBD118-E1	F-	CCACTCCAGAGAGGCCAATA	190	1.5	55
	R-	TACCTCCGTTTTTGCAGCTT			
BBD118-E2	F-	CAGTGATCTTTCCTTTTATTTTTCA	300	1.5	55
	R-	GGCTGGGTCAGACTTTTAATCTT			
BBD119-E1	F-	CACAACCACGTCCTTCTCCT	217	1.5	55
	R-	TGAGAAAACGTCCCAAAACC			
BBD119-E2	F-	TGCTCCCTGACTTGTCAGATT	298	1.5	55
	R-	GGCAGCAGGTTGAGGATAGT			
BBD120-E1	F-	AACCACGTCCTTCTCCTGTG	250	2.5	65
	R-	GGGAAGAGGAGGGAGAGACA			
BBD120-E2	F-	TGGGTGTTCCTATGTCCAGA	226	2.5	65
	R-	CCCTGTGTTGACAAGTTTGG			
BBD121-E1	F-	ACTGCTCAGCATGGAGTCCT	215	2.5	65
	R-	ATTCTCCCAGCCCAAGCTAT			
BBD121-E2	F-	CACCCCTCTCCTTTTCCTTT	248	2.5	55
	R-	ACCATAGGGTCCAACGATGT			
BBD122-E1	F-	TGTCTCTTCGCCAGAACTGA	245	2.5	65
	R-	GAACTGGGGGCTTTTTCTTC			
BBD122-E2	F-	CTGTCCTTTCAAGCCAGGAT	248	2.5	65
	R-	CTTCATTTTTGTGCCCTTGG			
BBD122a-E1	F-	TCGCCAGAACCGAGACTACT	238	2.5	65
	R-	AATTGGGAGCTTTGTCCTCTT			
BBD122a-E2	F-	GGCTGTTAGTTCTCTCAAACTAGGAT	249	2.5	65
	R-	TTTGTGCTCTTGGGCTTCA			
BBD123-E1	F-	CCAGAGAGGCCAGAACAATC	172	2.5	65
	R-	CCTGAAGAGGTCCGGGTTAC			
BBD123-E2	F-	AGGGCTGACACTGTCTCTGTT	238	2.5	65
	R-	CTGGGGTCAGACTCAGGGTA			
BBD124-E1	F-	TGTGCTCTCCCCAGTGGTTCCC	197	2.5	65
	R-	AGGCCTCCTCTATCCCCCAGC			
BBD124-E2	F-	ACCCCCTGGGTGGCTCTGTG	300	1.5	65
	R-	CCCTGAGGGTGGCGAGGCAA			
BBD125-E1	F-	CCTTCCTCTCTCCCAGAAGC	168	2.5	65
	R-	GGACCCTCTTTCTCCCTCAA			
BBD125-E2	F-	AGGGTTTGTCATCTTGTTGCT	343	2.5	65
	R-	ATCTGGCCCTAATTGCTGTG			
BBD125a-E1	F-	CCTTCCTCTCCTCCCAGAAG	169	2.5	65
	R-	GGACCCTCTTTCTCCCTCAA			
BBD125a-E2	F-	TTGCTGTTTGTAATCTTTTGCTAGA	400	2.5	65
	R-	TGCTCTCAGGAGTAGGTGTCA			
BBD126-E1	F-	CCACCTTGGGAGGACAATTA	353	2.5	55
	R-	CACTCCCATGGATGTTTGGT			

Gene Symbol		Primer Sequence	Amplicon Length (bp)	MgCl2 Concentration mM	Annealing Temperature
BBD126-E2	F-	TTTTTGTGATAGATCCCTGGC	329	2.5	55
	R-	ACCAAAATAAGGCCCTTTGC			
BBD127-E1	F-	TCTGAGGAAGGTGACATAGTGG	197	2.5	55
	R-	TTGATTTGAGTGAGCCCCTG			
BBD127-E2	F-	ACAAACATCCCCTGTCTTCC	329	1.5	65
	R-	GGCTTTCTGTGGGAGTGTGT			
BBD128-E1	F-	CAAAGCCGTGTCTCATCTTG	227	2.5	65
	R-	ACTATCTTGAAGGCCCTGCC			
BBD128-E2	F-	GGGATGAGTGTCAAGGAGGA	410	2.5	65
	R-	GAGGAGACCAGTGGAACAGC			
BBD129-E1	F-	CAGAGCCTAAGGCCATGAAG	261	2.5	55
	R-	TGGGCATCTCATTTTCAGTG			
BBD129-E2	F-	TACAAAGACGCTCTTGGCCT	423	2.5	55
	R-	AACTGGAGCTGGAGAAGCAG			
BBD132-E1	F-	GTAGACCCTATAAAGGCAGTAGCTCCA	195	1.5	65
	R-	ATCTGCTGGTGTTCCTGGCTCCC			
BBD132-E2	F-	CCCAACCCCTTGGCAGTCCC	346	1.5	65
	R-	AGATGCAGAGAGACAAAACAGGAGC			
BBD142-E1	F-	GCAGGACACTTCCCTGCTACAAAA	211	2.5	65
	R-	AGCTGTCTGACACTAGGTCACCAAAA			
BBD142-E2	F-	TGGTTGTATAGCGATTAACCATAGTC	361	1.5	65
	R-	ATCAACTTTGCCCCTTTTGA			

Promoter region analysis primers were designed using the predicted protein sequences in BLAT for the bovine genome. The University of California, Santa Cruz (UCSC) genome Browser was used to determine the orientation of both *LAP* and *BBD126* genes and the 1000 upstream region was retrieved. Primers were designed to amplify these particular regions of the upstream region from DNA. In addition, restriction enzyme digestion sites for BglII and XhoI were added for the cloning of the DNA fragments into the pGL4.23 luciferase vector (Table 2. 3).

Table 2. 3. List of primers used for commercial genotyping of SNPs in both NR and HF breeds.

SNP ID	Location in sequence
BBD115 E2	CATAAATCATCAAAACTTGTCAGAAAAGAAGAGACA[G/A]CATATAGGGTTATGGCAAGTACCACTAAG
BBD116 E2	AAGCAGGAGTCTAACTCCAACCCCAGTTTGTCAGTT[G/A]CAACGCTTTCAAACTATTCTCAGATTTGAAT
BBD117 A	TGTTGGATCATTAAAGGACACTGCAGAAAAGACTGC[A/T]AATCTGGTGAACAGGTTAAAAAGCCATGT
BBD117 B	AGGACACTGCAGAAAAGACTGCAAATCTGGTGAACAG[G/A]TTAAAAAGCCATGTAGAAATGGTGACTAT
BBD121 E1	GGAACCTGATGCATATTTAGGGAATTCTGGTGGCT[C/G]TAATAGCTTGGACTGGGAGAAT
BBD122 E2	TTTCTCTTTGTATCAGGCAGCACTGAAAAATGCTGGAA[T/C]CTTCATGGCAAGTGCCGTGACACATGTTCC
BBD123 E2 A	TCAGGTGGCACCCGAAAATGCTGGAATTTTCATGGCA[G/A]GTGCCGCCATAAATGCTTCAAGAAGGAAA
BBD123 E2 B	ACCCGAAAATGCTGGAATTTTCATGGCAGGTGCCGCCA[C/T]AAATGCTTCAAGAAGGAAAAGGTCTATG
BBD124 E1	CTTCTGCTTCTTGTGGCTCTCCTGGTCCTGGGTCATGTGCC[A/G]ACAGGTAACATCACTTGTGTGTGTGTG
BBD126 E1 A	ACCCGTGTTCAGAGTCACACTATATAGAGTGTTCTGGATTC[C/T]GTAGCACCCACTGCCTACTGTCTTTCAGAA
BBD126 E1 B	ATTCCATAGCACCCACTGCCTACTGTCTTTCAGAAGCAATG[A/G]AGTCTCTATTGTTCACTCTTATACTCTTTAT
BBD128 E1	ACAGATCACTGGACAAAGAAATTCTGCTGCCTAGTTGTCTTGGGA[G/A]GTCTGTAAACATGAAGCTGTTTCTGAT

Cloning primers for protein expression of BBD126 were designed using Ape plasmid editor (Table 2. 4).

Table 2. 4. List of primers used for cloning, including the restriction enzymes sites incorporated in the primers.

Gene symbol	Use	Primer name	Primer Sequence	Restriction enzyme site
LAP	promoter analysis	LAPpLF	GATCAGATCTAGCTGATGCTGGAGAAGAGG	BglII
		LAPpLR	GATCCTCGAGCCTGTATCTCCAAGCCCTGA	XhoI
BBD126	promoter analysis	BBD126pLF	GATCAGATCTTTGCTTCTGAAAGACAGTAGGC	BglII
		BBD126pLR	GATCAAGCTTTTATGTCTTCCAGTCACAATCCA	XhoI
BBD126	production BBD126 in eukaryotic cell line	pEAK126 F	GATCAAGCTTAATTGGTATGTGAGAAAGTG	Hind III
		pEAK126 R	GATCTCTAGAAGCAATGCCTGTTGTAGATC	XbaI
VHH	VHH isolation from white cells	AlVHH-shR1	GATCGGCGCGCCGAGGGGTCTTCGCTGTGGTGCG	AscI
		AlVHH-ihR1	GATCGGCGCGCCGGTTGTGGTTTTGGTGTCTTGGG	AscI
		AlVHH-F1	CTTGCGGCCGCTCAGKTGCAGCTCGTGGAGWCNGGNGG	NotI
VHH	cloning of VHH into pHEN6 vector	pHEN-VHH F	GCT TGC GGC CGC GTT CAG KTG CAG CTC GTG GAG WCN GGN GG	NotI
		pHEN-VHHshR	GAT CCT CGA GGA GGG GTC TTC GCT GTG GTG CG	XhoI
		pHEN-VHHlhR	GAT CCT CGA GGG TTG TGG TTT TGG TGT CTT GGG	XhoI
Art. VHH	construction of artifitial VHH library	P1 F	CGATGCGGCCGCATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCT	NotI
		P2 R	ACCTGCTACAAACTCACGCTCCTTCCCTGGAGCCTGGCGGAACCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGAGGCTGCACAGGA	
		P3 R	CTCTGGAGATGGTGAATCGGCCCTTCACGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACCTGCTACAAACTCACGC	
		P4 R	GTAATAAACGGCCGTGTCCTCAGGTTTCAGGCTGTTCATTTGCAGATCTACCGTGTTCTTGGCGTTGTCTCTGGAGATGGTGAAT	
		P5aR	CCCCTGGCCCCAGTANNNNNNNNNNNNNNNNNNNN

2.4. QUANTITATIVE REAL-TIME POLYMERASE CHAIN REACTION (QRT-PCR)

Quantitative real time PCR (qPCR) samples were prepared using a 25µl reaction mix protocol: 2 µl of 40 ng/µl cDNA, 12.5 µl of 2x qPCR SYBR green master mix (Stratagene Corp, La Jolla, CA, USA) and 10.5 µl of a mix of primers and water. This was optimised for each primer mix by titration of 100, 300, 600 and 900 nM final concentrations (Table 2.1). All products were run on an agarose gel to confirm the presence of a single PCR product of the correct size. A MX3000P quantitative PCR system (Stratagene Corp, La Jolla, CA, USA) was utilised to perform qPCR with the following cycle parameters: 95°C for 10 minutes, 40 cycles of 95°C for 30 seconds, 60°C for 1 minute and 72°C for 30 seconds, with a final step for amplicon dissociation. A non-template control (NTC) was run for each primer used in each 96-well plate to confirm absence of gDNA contamination. Levels of gene of interest (GOI) expression were determined as a ratio of abundance compared with GAPDH, [ratio=$2^{(-CtGOI)}/2^{(-CtGAPDH)}$]. *GAPDH* has been selected as the most stable gene across all samples from a panel of possible housekeeping genes (*ACTB, GAPDH, MRPS6, PPIA, RPS9, RPS15* and *H3F3A*) and was analysed using geNorm software version 3.4 (Vandesompele, De Preter et al. 2002). Due to the high number of samples, (6 different tissues for the 6 male tracts used and 4 tissue regions for the 10 female tracts obtained; a total of 76 samples) samples from each region were used to create pools. Each normalised gene was tested against these pools. Raw Ct values were used for the geNorm analysis.

2.5. POLYMERASE CHAIN REACTION (PCR)

SNP discovery was carried out by the sequencing of the coding region of the β-defensins genes. PCR was used to amplify the coding region to obtain enough PCR products to be sent for sequencing, and 100µl reactions were set up for each fragment. Primers were optimised for $MgCl_2$ concentration and annealing temperature, using Platinum Taq DNA polymerase enzyme (Invitrogen Ltd., Paisley, UK). The mixture consisted of 11 µl of MgCl2 (optimal concentration was primer dependent), 11 µl of 4mM dNTP mix, 11 µl of 10mM primer mix, 11 µl of 10x reaction buffer, 2.5 µl of Platinum Taq DNA polymerase and 10ng of DNA with a final reaction volume of 100 ul. A PTC-200 thermocycler (MJ Research) was set up with the following protocol: 95°C for 10 minutes and 40 cycles of 95°C for 30 seconds, corresponding annealing temperature for 1 minute and 72°C for 60 seconds.

PCR products for cloning were obtained using a HotStart PCR kit (Qiagen, Crawley, UK) with 50 µl reactions of 5 µl 10x reaction buffer containing MgCl2, 2 µl 5 mM dNTP, 1 µl of 50 mM forward primer, 1 µl the 50 mM reverse primer, 10 ng DNA and 0.25 µl of HotStart Taq polymerase enzyme in a final volume of 50 ul. The thermocycler was set up with the following protocol: 95°C for 15 minutes, 38 cycles of 94°C for 1 minute, 60°C for 1 minute, 72°C for 1 minute and a single step of 72°C for 10 minutes.

2.6. PCR PRODUCT ISOLATION

PCR reactions were run on 1.5% agarose gel with 0.05 ug/ml of ethidium bromide at 120v for 35 minutes. DNA was visualised using an ultraviolet lightbox, KODAK Gel Logic 100 System (Kodak, Rochester, NY, USA). Bands were then cut using a sterile scalpel and transferred into sterile 1.5 ml tubes. DNA isolation was carried out using a Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA) according to the manufacturer recommendations. DNA was eluted into 40 µl of DNAse/RNAse free water. DNA was then quantified using a Nanodrop spectrophotometer.

2.7. GDNA EXTRACTION, DNA SEQUENCING AND SNP DISCOVERY

A panel of genomic DNA samples from two breeds, Holstein-Friesian and Norwegian Red, was provided by Dr Kieran Meade (Teagasc). Five samples from each breed were used to discover novel SNPs. Primers designed to amplify each of the 2 exons found in β-defensin genes were independently used to obtain PCR products corresponding to each exon. PCR products were then commercially sequenced by Macrogen (Macrogen Inc., Rockville, MD, USA). Retrieved sequences were analysed using BioEdit Software version 7.0.9 (Hall 1999). Sequences were aligned to look for SNPs and to find differences among them. Once the possible SNPs were detected, each sequence trace was analysed manually to look for double peaks that would indicate heterogeneity.

Semen straws (N= 19) were used for DNA extraction. These samples included the records of fertility for each bull. Sperm cells were pelleted at 5000g and washed twice with PBS. The pellet was lysed by adding 450µl of extraction buffer (10mM Tris pH 8, 10mM EDTA pH 8, 0.1% SDS and 100mM NaCl) and warmed to 60°C, during overnight incubation with 50 µl of

dithiothreitol (Sigma Aldrich, St. Louis, MO, USA) and 10 µl of Proteinase K (20mg/ml) (Sigma Aldrich, St. Louis, MO, USA). After incubation, samples were pelleted at 5000g for 30 minutes and supernatant was recovered for use with for the Maxwell DNA system (Promega, Madison, WI, USA) following the supplier recommendations. Samples were eluted in 300µl of elution buffer and concentrations were determined using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) (Appendix: Table 3).

2.8. SNP GENOTYPING AND CLUSTER ANALYSIS

Twelve of the newly discovered SNPs, with a minor allele frequency (MAF) >0.05, were commercially typed by Sequenom (www.sequenom.com) using iPlex Gold assay in DNA from 30 Norwegian Red and 30 Holstein-Friesian cattle. Analysis of the typing results was carried out using Genepop for the population genetics (Fisher's exact test was used for genomic and genotypic frequencies, Hardy-Weinberg and linkage disequilibrium) (Medzhitov, Preston-Hurlburt et al. 1997) and HAP (Radhakrishnan, Hamil et al. 2005) and Arlequin software (Excoffier, Laval et al. 2005) for haplotype reconstruction.

In a second experiment, DNA samples from 19 bulls, extracted from sperm samples as described in section 2.7, were used to amplify *BBD126 E1, BBD115 E2, BBD117* and *BBD123 E2*. PCR products were commercially sequenced by Macrogen (Macrogen Inc., Rockville, MD, USA). Fertility data from these animals, including frequency of pregnancy pre- and post-semen freezing, was available from the National Cattle Breeding Centre. SNP frequency and fertility scores were analysed by Pearson's correlation analysis using SPSS statistical software.

2.9. PREDICTION OF GLYCOSYLATION SITES

To determine the presence of possible glycosylation sites, the coding region of the genes was retrieved and applied in the online tools of two glycosylation sites. NetOGlyc 3.1 (http://www.cbs.dtu.dk/services/NetOGlyc/) and NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) servers (Julenius, Molgaard et al. 2005) were

used for glycosylation sites prediction. Threshold of the analysis was set following the recommendations of the manufacturer.

2.10. PROMOTER REGION BIOINFORMATIC ANALYSIS

Bioinformatically predicted protein sequences were retrieved from were used in BLAT searches against the bovine genome (Cormican, Meade et al. 2008). From the contiguous upstream region 1000bp for each bovine Chr13 β-defensin cluster genes, several testis and epididymal specific genes (*SPAM1*, *SPAG11E*, *ArA*, *Esp*, *Ace*, *sPLA2III*, *Sp17* and *Catsp)*, and 4 bovine Chr27 β-defensin cluster genes (*BBD105*, *BBD106*, *BBD107* and *LAP)* were retrieved and added into the online transcription factor prediction tool MatInspector (http://www.genomatix.de). In our analysis, the threshold was set to 0.9 to reduce the number of false positives. The transcription factors predicted were grouped according to their family matrix, and a table with the number of times each transcription factor binding site was predicted per upstream region was created. The data was used for principal components analysis. Transcription factor binding site were used as variables. The analysis was carried out using SPSS statistical software.

2.11. RESTRICTION ENZYME DIGESTION

All restriction enzymes were purchased from New England Biolabs (Ipswich, MA, USA). Double digestion reactions were carried out following the recommendations of the manufacturer: 1000 ng of DNA, PCR product or vector, were digested using 2 µl of each of the restriction enzymes, 0.5 µl of the supplied BSA solution, 5 µl of the correct 10x buffer and H2O up to a final volume of 50 ul. Reaction was incubated overnight at 37 °C. Reaction was then stopped at 95 °C for 5 minutes.with the exception of the pair AscI-NotI that required individual reactions due to incompatibility with the buffers (Table 2. 5).

Table 2. 5 Conditions for the double digestion restriction enzymes.

All reactions were carried out at 37 °C, the correct 10x buffer for the pair was determined using the double digestion finder in www.neb.com.

Restriction enzime pair	10x Buffer	BSA	Temperature	Reaction Stop
BglII-XhoI	3	Yes	37	95
HindIII-XbaI	2	Yes	37	95
AscI	4	No	37	95
NotI	3	Yes	37	95
NotI-XhoI	3	Yes	37	95
NcoI-XhoI	4	Yes	37	95

Restriction enzyme reactions, once deactivated, were run on 1.5% agarose gel with 0.05 ug/ml of ethidium bromide at 120v for 35 minutes. DNA was visualised using an ultraviolet lightbox. Bands were then cut using a clean scalpel and transferred into a clean 1.5 ml tube. DNA isolation was carried out using a Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA) following the manufacturer's recommendations. DNA was eluted into 40 µl of DNAse/RNAse free water. DNA was then quantified using a Nanodrop spectrophotometer.

2.12. LIGATION REACTION

Digested PCR and plasmids were ligated using the presence of complementary sticky ends, as a result of the restriction enzymes digestion. Ligation reaction was carried out in 1.5 ml tubes combining 10 ng of the PCR product with 100 ng of the digested vector, 1 µl of T4 DNA ligase (New England Biolabs, Ipswich, MA, USA), 2 µl of 10x reaction buffer and up to a final volume of 20 µl with DNAse-free water. The tubes were incubated for 1 hour at room temperature and the reaction was stopped heating up the reaction to 95°C for 5 minutes. If used for electroporation into competent cells, the ligation reaction was cleaned using a Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA). For chemically competent cells, ligation product was used without clean-up.

2.13. Bacterial cell transformation

Chemically competent cells, BL21 Gold DE3 pLysS (Agilent Technologies, Inc, Santa Clara, CaA USA) and One Shot TOP10 Chemically Competent E. coli (Invitrogen Ltd., Paisley, UK), were used for chemical transformation. 10 µl of the ligation reaction were incubated with 40 µl of competent cells on ice for 30 minutes. Cells were then heat shocked at 42°C for 30 seconds. 150 µl of SOC medium (Invitrogen Ltd., Paisley, UK) were added to the cells and cells were incubated at 37°C for 1 hour and shaking (200 rpm). Cells were then spread in LB agar plates complemented with the corresponding antibiotic for the vector used (Kanamycin for pET28, Ampicillin for pEAK8, pGL4.23, pHEN6C and pD). BL21 cells were used for protein expression using the pET28 vector. One Shot cells were used to propagate the plasmids.

Electrocompetent cells TG1 (Agilent Technologies, Inc, Santa Clara, CA, USA) and WK6 (gift of Prof. Hidde Ploegh, Whitehead Institute, MIT, MA, USA) were used for electroporation. In both cases 1 ng of DNA from a pre-clean ligation reaction were added to 40 µl of cells. The cells were then transferred into a electroporation chamber (Bio-Rad, UK) and used in a electroporator set up to 10uF, 600 Ohms and 1800 volts. Cells were recovered using 150 µl of SOC medium and incubated for at 37°C for 1 hour and shaking (200 rpm). Cells were then spread in LB agar plates complemented with the corresponding antibiotic for the vector used. TG1 cells were used during the phage display protocol for their ability to be infected by the M13 helper phage. WK6 were used for the protein production of VHH.

2.14. Maxiprep and Miniprep plasmid preparations

PureYield Plasmid Miniprep (Promega, Madison, WI, USA) was used to isolate plasmid from small cell culture, 3 ml, for sequencing. Selected colonies from LB agar plates, supplemented with the selective antibiotic for each vector, were grown overnight at 37°C with agitation, 200 rpm, in 3 ml of LB broth supplemented by the correct antibiotic. 2 ml of the overnight culture were centrifuged at 12,000 rpm in a table-top centrifuge; the supernatant was then removed. The cell pellet was resuspended in 600 µl of DNAse-free water. Cells were lysed by the addition of 100 µl of lysis buffer by inverting the tube 6 times. The buffer

was neutralised with 200 µl of neutralisation buffer. Tubes were then centrifuged at 12,000 rpm for 10 minutes and supernantant was loaded into the spin column supplied with the kit. The sample was loaded through the column by centrifugation at 12,000 rpm for 2 minutes. Buffer was discarded and column was washed using 300 µl of endotoxin removal wash buffer by centrifugation at 12,000 rpm for 1 minute. The buffer was then discarded. The column was then washed with 600 µl of washing buffer by centrifugation at 12,000 rpm for 1 minute and buffer was then discarded. Column was transferred to a new tube and40 µl of DNAse-free water were added to the column. The water was left for 1 minute before elution by centrifugation at 12,000 rpm for 1 minute. DNA was sent to be commercially sequenced at GATC Biotech (Konstanz, Germany).

PureYield Plasmid Maxiprep (Promega, Madison, WI, USA) was used to isolate larger quantities of plasmid from for 100 ml bacterial cultures. Cells were cultured overnight in 100 ml of LB broth supplemented with the correct antibiotic. Cells were centrifuged at 4000g for 10 minutes in a Sorval RC 5C plus centrifuge. Supernatant was eliminated and cells were resuspended in 12 ml of resuspension solution. Cells were lysed by the addition of 12 ml of lysis buffer. Lysis was induced by gently inverting the tube five times. Lysis was neutralised by adding 12 ml of neutralisation buffer and inverting gently five times. Sample was centrifuged at 14,000g for 20 minutes in a Sorval RC 5C plus centrifuge. Supernant was loaded into the column and vacuum was applied until all buffer passed through the column. 5ml of endotoxin removal wash buffer were loaded in the column and vacuum was applied. Once the column was dry, 20 ml of washing buffer were loaded and vacuum was applied until the column was completely dry. Column was then assembled into the 1.5 ml tube collection adaptor. 500 µl of DNAse-free water were loaded into the column and incubated for 1 minute at room temperature. Vacuum was immediately applied and the elution was recovered in the 1.5 ml tube. DNA concentration was determined using a Nanodrop spectrophotometer.

2.15. Luciferase reporter gene assay for *BBD126* and *LAP*

In order to determine the effect of TLR signalling, and hormones such as dihydrotestosterone and oestrogen, in the regulation of gene expression of *LAP* and *BBD126*, firefly luciferase gene assay was used. The promoter regions of both genes (1,000 bases upstream) were cloned

into a pGL4.23 reporter gene vector. Promoter regions were amplified using LAPpLF and LAPpLR for *LAP*, and BBD126pLF and BBD126pLR for *BBD126*, by using a HotStart PCR kit as previously described in Section 2.5. Isolated PCR products and vector were digested using BglII and XhoI as described in Section 2.11. Isolated digestion products were then ligated with digested pGL4.23 vector as described in Section 2.12, and Top10 cells were transformed and spread in LB plates supplemented with ampicillin following the protocol described in Section 2.13. LB broth cultures, 3 ml, were used for plasmid isolation using PureYiel Plasmid Miniprep as described in Section 2.14. The isolated DNA was sent for sequencing. Once the sequence confirmed the presence and correct orientation of the cloned PCR product, 100 ml of LB broth (supplemented with ampicillin) were inoculated with the selected clone, and grown overnight at 37°C with agitation (200 rpm). This was then used for PureYiel Plasmid Maxiprep plasmid isolation as described in Section 2.15. DNA concentration was determined using a Nanodrop spectrophotometer.

Bovine endometrial (BEND) cell line was used as a target for the transfection and stimulation due to the presence of liposacharide (LPS), dihidrotestosterone (DHT), and oestrogen receptors. BEND cells were grown in DMEM supplemented with 10% FCS. No antibiotics were used in the culture of the cells at any point. Cells with a count of 40,000 were plated in 96 well tissue culture treated plates (Greiner Bio-One, Frickenhausen, Germany), and incubated overnight at 37°C with 5% CO_2. Transfections were carried out using 100 ng of the Luciferase vector, pGL4.23, pGL4.23-LAP, pGL4.23-126 and 10 ng of the TK renilla control vector (Table 2.3). Transfection was carried out using 0.2 µl of GeneJuice Transfection Reagent (Merk Millipore, Billerica, MA, USA) per well. DNA and Genejuice were mixed with serum-free DMEM, and incubated for 30 minutes at room temperature. The mix was then added into the wells with growing BEND cells (Table 2. 6).

Table 2. 6 Schematic for the distribution of the different vectors transfected into BEND cells.

	1	2	3	4	5	6	7	8	9	10	11	12
A	cells	pGL4.23 + TK renilla	pGL4.23-LAP + TK renilla									
B												
C												
D												
E			pGL4.23-126 + TK renilla									
F												
G												
H												

Cells were incubated overnight at 37°C with 5% CO_2 in a total volume of 100 ul. From each treatment, 1 µl was added for the final concentration shown in Table 2. 7.

Table 2. 7. Final concentration of the treatments used to stimulate luciferase production.

Treatment	Concentration
LPS 1	1.5 ng/ml
LPS 2	15 ng/ml
DHT 1	0.1 uM
DHT 2	10 uM
Seminal plasma (SP) 1	1:10,000 dilution
Seminal plasma (SP)2	1:100 dilution
Oestrogen 1	0.1 uM
Oestrogen 2	10 uM

The different treatments were added to the 96 well plates following the distribution shown in Table 2. 8. All treatments were done in quadruplicate.

Table 2. 8. Schematic for the distribution of the different stimulations into BEND cells.

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E			control	LPS1	LPS2	DHT1	DHT2	SP 1	SP 2	E 1	E 2	
F												
G												
H												

Cells were incubated for 48 hours at 37°C with 5% CO_2. Supernatant was removed from all wells by aspiration and cells were lysated with 40 µl of 1x passive lysis buffer (Promega, Madison, WI, USA), and agitated for 15 minutes. In order to determine the activity of firefly luciferase, 40 µl of luciferase assay mix (20 mM tricine, 1.07 mM (MgCO3)4Mg(OH)2*5H2O, 2.67 mM MgSO4, 0.1 M EDTA, 33.3 mM DTT, 270 mM coenzyme A, 470 mM luciferin and 530 mM ATP) were added to 20 µl of the cell lysate. For the renilla luciferase control, 40 µl of 2 ug/ml coelenterazine in PBS were mixed with 20 µl of the cell lysate. Luminescence was determined by a Luminoskan Ascent Microplate Luminometer (Thermo Scientific, Waltham, MA, USA). Firefly luminescence values were normalised relative to the renilla control and expressed as fold changes compared to the unstimulated control. Three independent biological replicates were done.

2.16. SPERM PENETRATION ASSAY

Flat capillary tubes (Vitrotubes, Mountain Lakes, NJ, USA) were marked at 20, 40, 70 and 80 mm intervals. These tubes were filled with the cervical mucus collected from a single cow and kept at 37°C. Single semen straws from three bulls of each group were thawed at 37°C for 30 minutes and stored in 15 ml tubes in a 37°C water bath. Cell numbers for each straw were determined by cell count on a haemocytometer. All samples were diluted in BWW media to a final concentration of 20 million cells/ml, and a volume of 250 µl, in pre-heated tubes at 37°C. Hoescht stain solution, 2.5 ul, was added to each sample. Two capillary tubes

per sample were incubated at 37°C in the dark for 10 minutes. After that time, the capillary tubes were incubated at 50°C for 30 seconds, and incubated in the -20°C freezer for 3 minutes to kill all of the cells. Sperm cells were counted in one field of vision at the previously mentioned marks using a fluorescent microscope. Each sample was independently replicated a minimum of four times.

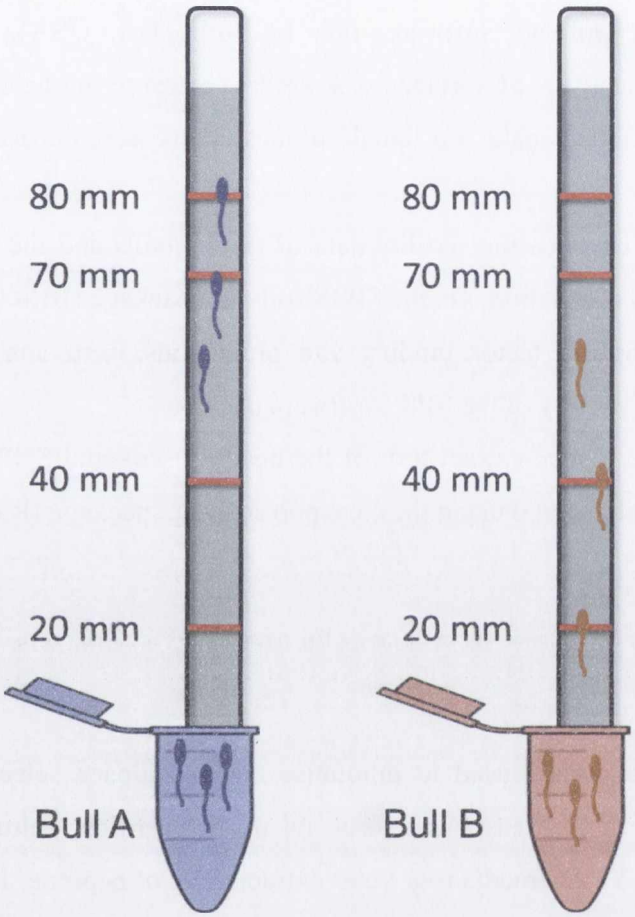


Fig. 2. 2 *Sperm penetration assay.*

Capillary tubes were filled with cervical mucus, marked at 20, 40, 70 and 80 mm and kept at 37°C. Live sperm cells were stained with Hoesch stain. Capillary tubes were introduced into a tubes containing the stained sperm cells and incubated at 37°C for 10 min. Capillary tubes were then heat to 50°C for 5 min. Death sperm cells were then counted at each mark with a fluorescence microscope.

2.17. STATISTICAL ANALYSIS

Mann-Whitney U tests were used to analyse the difference in length of peptide tail between Class A and other members of Chromosome 13 cluster as well as between Chromosome 13 cluster and Chromosome 27, 23, and 8 clusters. Tests were carried out using the Prism software package (GraphPad Software, Inc. La Jolla, CA, USA). Kruskal-Wallis non-parametric one-way analysis of variance was used to measure the heterogeneity of the gene expression in the different male and female tissues. Tests were carried out using the Prism software package.

Correlation analysis between the fertility data of the 19 bulls and the genotyped SNPs was done using Pearson's correlation, via the SPSS software package (IBM Corporation, Armonk, NY, USA). Transcription factor binding site predictions were analysed using principal components analysis by way of the SPSS software package.

Frequencies analysis, Fisher's exact test, of the newly discovered SNPs for the NR and HF populations were implemented using the Genepop software package (Rousset 2008).

2.18. ANTIGEN PREPARATION FOR ALPACA (*VICUGNA PACOS*) IMMUNISATION

Two different antigens were used to immunise the two alpaca selected: yeast lysate and Jurkat cells lysate. For the yeast lysate, 100 ml of an overnight culture of *Saccharomyces cerevisiae*, on sterile YPAD media (6gr yeast extract, 12 g of peptone, 12 g of glucose and 60 mg of adenosine henisulphate in 600 ml of H20), was centrifuged at 6,000g for 15 minutes in a Sorval RC 5C plus centrifuge. Supernatant was then removed, and cells were resuspended in 10 ml 50 mM TrisHCl, and 50 ml NaCl at pH 7.5. Cells were centrifuged again followed by elimination of the supernatant. Cells were resuspended in 5 ml of lysis buffer (50 mM TrisHCl, 50 mM NaCl, 1 mM EDTA, 1 mM PMSF, 1ug/ml leupeptin, 1 ug/ml aprotinin and pH 8). Aliquots consisting of 1 ml were transferred into 1.5 ml plastic tubes, and 200 µl of acid-washed glass beads (425-600 um) (Sigma Aldrich) were added while vortexing for 20 minutes at 4°C. Cell debris and glass beads were separated from the supernatant by centrifugation at 14,000 g for 15 minutes at 4°C in a table top centrifuge. Supernatant was transferred into new tubes, and sterilised by filtration with a 0.2 um cellulose acetate syringe filter (Merk-Millipore). Protein concentration was determined by Bradford protein assay

(Thermo Fisher Scientific, Waltham, MA, USA). Sample, consisting of 3 ul, was mixed with the buffers supplied in accordance with the manufacturer's instructions. Mix was incubated at 37°C for 30 minutes, and colour change was measured in a plate reader at 595 nm. A standard curve is included with the samples to determine the protein concentration.

For the Jurkat cell lysate, two 172 cm^2 tissue culture flasks containing growing Jurkat cells, an immortalised T lymphocyte cell line grown in RPMI (Invitrogen Ltd., Paisley, UK) and supplemented with 10% FCS, were utilised. They were centrifuged at 1,000g for 10 minutes. Cells were washed twice with 5 ml of 50 mM TrisHCl and 50mM NaCl, pH 8. Finally, cells were resuspended in 2 ml of lysis buffer (50 mM TrisHCl, 50mM NaCl and pH 8 supplemented with 100ug/ml of DNAse I (Roche, Penzberg, Germany). Cells were lysed by three rounds of freezing and thaw cycles. After the third cycle, cell debris was pelleted by centrifugation in a table top centrifuge at 13,000 rpm for 15 minutes at 4°C. Supernatant was transferred into new tubes, and sample was sterilised by filtration with 0.2 um cellulose acetate syringe filters. Protein concentration was determined by Bradford protein assay.

2.18. Alpaca (*Vicugna pacos*) immunisation

Two alpacas acquired from the Burren Alpaca farm (Co. Clare, Ireland) were inoculated with 0.5 ml subcutaneous injections in the prescapular region containing 2 mg of one of the antigens and 13 mg/ml of aluminium hydroxide gel as adjuvant. Immunisation was repeated three times at two week intervals. Blood samples, 50 ml, were collected two weeks after the final immunisation from each animal using heparinised tubes to avoid coagulation of the blood. No euthanisation was required.

2.19. Alpaca (*Vicugna pacos*) white cell isolation

White blood cells were isolated using 50 ml Leucosep tubes (Greiner Bio-One, Frickenhausen, Germany). Tubes were loaded with 15 ml of Histopaque-1077 (Sigma-Aldrich) and centrifuged at 290 g for 2 minutes. Blood samples of 10 ml were loaded into the tubes and sample was centrifuged at 1580 g for 15 minutes. A plastic Pasteur pipette was used to recover the upper phase containing the peripheral blood mononuclear cells (PBMC) and transferred to a new tube. Cells were washed twice with sterile PBS and pelleted at 650 g

for 10 minutes. Pellet was resuspended in 1 ml of 25% NaCl to lyse any red blood cells. Sample was again centrifuged at 650 g for 10 minutes. Supernantant was removed and cells were washed with 1 ml of PBS. Cell suspension was transferred into a 1.5 ml tube and centrifuged at 12,000 rpm for 10 minutes. Supernantant was eliminated and cells were snap frozen in liquid nitrogen where they were stored at -80°C unless used immediately for RNA extraction (protocol described in Section 2.2). The isolated RNA was the used for cDNA synthesis as described in section 2.3.

2.20. Alpaca (*Vicugna pacos*) single domain antibodies (VHH) gene isolation and cloning

During maturation of the B cells, the genes coding for the antibodies undergo a phenomenon known as gene rearrangement that determine the specificity of the produced antibody. These rearranged genes can be isolated by PCR (Maass, Sepulveda et al. 2007). The primers, AlVHH-shR1, AlVHH-lhR1, AlVHH-F1, HotStart Polimerase kit (Qiagen), and the cDNA obtained from the white cells isolated from the immunised alpacas were used following the protocol described in Section 2.5 in 50 µl reactions. After, the PCR products were run in a 1.5% agarose gel stained with ethidium bromide. DNA was isolated using the Wizard SV Gel and PCR Clean-Up (Promega, Madison, USA) as described in section 2.6, and phagemid vector JSC (gift of Prof. Hidde Ploegh, Whitehead Institute, MIT, MA, USA) was used for preparation of pIII phage display libraries (Sepulveda et al. 2008). PCR products and JSC vector were digested by AscI and NotI in individual reactions due to their buffer incompatibility. For AscI, 1,000 ng of DNA were digested using 2 µl of AscI restriction enzyme, 5 µl of 10x buffer 4, and up to 50 µl with DNAse-free water and incubated overnight at 37°C. The reaction was then run in a 1.5% agarose gel. DNA was isolated using a Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA). The result product was digested with NotI using 2 µl of NotI enzyme, 5 µl of 10x buffer 3, 0.5 µl of BSA, and up to 50 µl with DNAse-free water. Reaction was carried out overnight at 37°C, and was run in a 1.5% agarose gel. DNA was isolated using a Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA), and quantification took place using a Nanodrop spectrophotometer. Digested PCR and vector were used for ligation reaction. A total of ten reactions were set up following the protocol described in Section 2.12. The reactions were cleaned up using Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA) and 40 µl of DNAse free water was used for the elution. From the resulting sample, 1 µl was used for electroporation in

Electrocompetent TG1 cells as described in Section 2.13, and plated into YTAG agar plates (0.8% tryptone, 0.5% yeast extract, 0.5% NaCl, 1.5% agarose, autoclaved at 120 °C for 15 minutes and supplemented with 100 ug/ml ampicillin and 1% glucose) with subsequent overnight incubation at 37°C. This was repeated ten times. Serial dilutions up to $1:10^6$ were plated as well to determine the number of clones found in the VHH library. Confluent colonies were scraped into 10 ml of 2YTAG broth (1.6% tryptone, 1% yeast extract, 1% NaCl, autoclaved at 120oC for 15 minutes and supplemented with 100 ug/ml ampicillin and 1% glucose) and supplemented with 10% glycerol. Cells were stored at -80°C in 0.5 ml aliquots.

2.21. PHAGE RESCUE

The isolated VHH cloned into the JSC vector was transcribed as a fusion protein. It was formed by linkage of the VHH peptide through its C terminus to the phage surface protein pIII. In the presence of a helper phage, the completed phage particle is formed by incorporating the fusion protein onto its surface, and the JSC-VHH vector as genetic material. The process by which the VHH-phage particles are formed is known as phage rescue.

2YTA broth of 10 ml (1.6% tryptone, 1% yeast extract, 1% NaCl, autoclaved at 120 °C for 15 minutes and supplemented with 100 ug/ml ampicillin) was inoculated with 10 µl of the JSC-VHH TG1 stock library. Cells were incubated at 37 °C with 200 rpm agitation until OD600 = 0.4-0.5, 800 µl of VCS-M13 helper phage (Strategene, La Jolla, CA, USA) were then added to the cell culture. The culture was incubated at room temperature for 20 minutes with no agitation. The flask was then transferred into a 37°C shaker, with agitation at 200 rpm, for 1 hour. 2YTA broth, 90 ml, was added and supplemented with 25 ug/ml of kanamycin. The culture was then incubated overnight at 37 °C with agitation, 200 rpm.

The sample was centrifuged at 4,000 g for 15 minutes at 4°C in a Sorval RC 5C plus centrifuge. The supernatant was recovered and phage particles were precipitated by adding 20 ml 20% PEG and 2.5 M NaCl. This was followed by incubation of the sample for 1 hour on ice. The sample was then centrifuged at 10,000 g for 30 minutes at 4°C. The supernatant was removed, and the precipitated phage particles were resuspended in 1 ml of sterile PBS. The

phage solution was sterilised using a 0.2 um cellulose acetate syringe filter. Unless immediately used for panning, phages were stored at -80°C in 50% glycerol.

2.22. PHAGE SELECTION BY PANNING

Phages displaying VHH on their surfaces were used for a selection process known as panning. This is based on the affinity of some of the VHH for proteins attached to a surface matrix. TG1 cells were transformed with JSC-VHH phagemid. The resulting cells were infected with VCS-M13 helper phage producing phage particles which display VHH on their surface and contain the matching JSC-VHH phagemid as genetic material. A surface with antigen attached was used for affinity selection. All non-binders were washed away, and binders were eluted and used to infect TG1 cells. The cycle was repeated up to three times. In the final round of selection, eluted phages were re-cloned into pHEN6 vector for protein production (Fig. 2.2).

Two wells in an ELISA plate (Greiner Bio-One, Frickenhausen, Germany) were loaded with 50 µl of 2 mg/ml antigen solution, and incubated overnight at room temperature. After incubation, the wells were rinsed twice with RNAse-free water. Wells were filled with 200 µl of blocking buffer (3% BSA in PBS, pH 7.5) and incubated for 1 hour at 37°C. Wells were washed twice with RNAse-free water, and 50 µl of the VHH-gIII displaying phage solution was loaded into each well where they were incubated for two hours at 37°C. Buffer was eliminated by aspiration, and wells were filled with 400 µl of washing buffer (0.05% Tween-20 (Sigma Aldrich) in PBS) followed by incubation at room temperature for 2 minutes. Buffer was then removed by aspiration. This process was repeated five times. Bound phage were eluted with 50 µl of low pH elution buffer (0.1 M HCl, 1 mg/ml BSA, pH 2.2) and incubated for 10 minutes at room temperature. Eluted phages were transferred into a 1.5 ml tube and pH was neutralised with 50 µl of 2 M TrisHCl, pH 8. Exponentially growing TG1 cells, 2 ml, (OD600 = 0.5) were immediately infected with 50 µl of the eluted phage solution and incubated at room temperature for 30 minutes with no agitation. 2YTA broth of 10 ml was added and the sample was incubated at 37°C for 1 hour with agitation, 200 rpm. After incubation, serial dilutions of the cell culture were plated onto 2YTAG agar plates and cultured overnight at 37°C so as to determine the culture colony number. VCS-M13 helper phage (Stratagene), 800 ul, were inoculated into the 10 ml cell culture. 2YTA broth of 90 ml

was added to the culture and incubated at 37°C for two hours with agitation, 200 rpm. Cell culture was then supplemented with 70 ug/ml kanamycin and cells were incubated overnight at 37 °C with agitation, 200 rpm. Phages were recovered using the same protocol described in Section 2.21. The selection cycle was repeated up to three times to increase enrichment in high affinity binders.

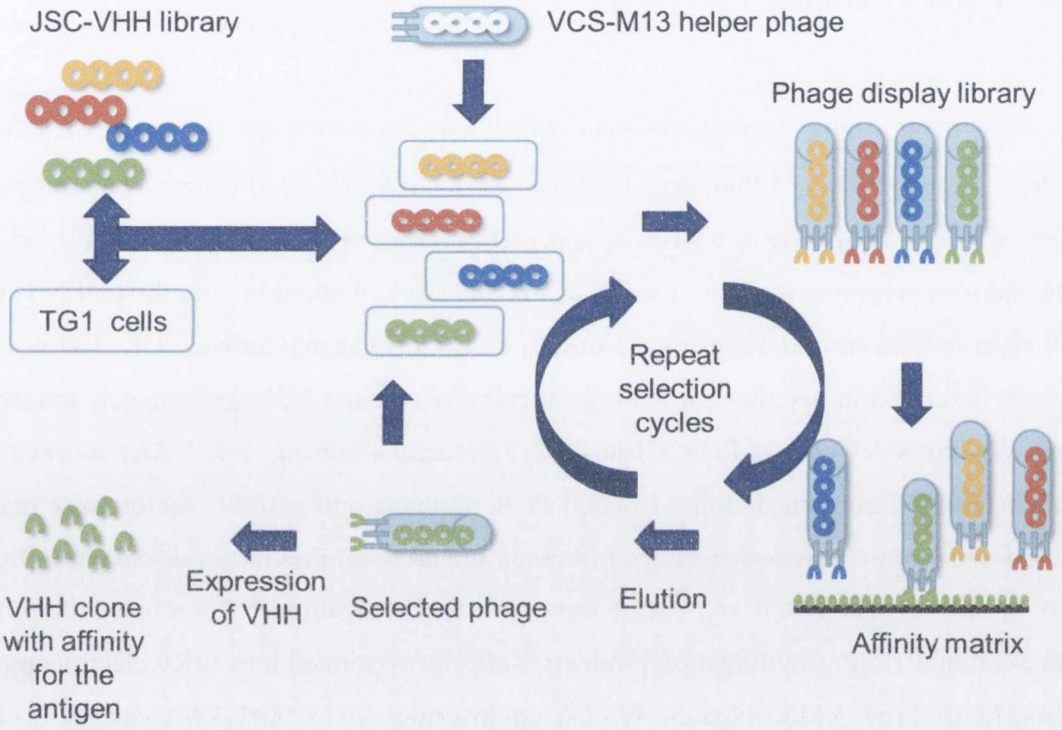


Fig. 2. 3 *Phage display selection process.*

Phage display technology is based on the formation of fusion proteins between the VHH and the phage surface protein gIII. TG1 cells containing the JSC-VHH phagemid are infected with VCS-M13 helper phage, creating all the machinery required to form whole phage particles displaying the fusion VHH-gIII protein in the surface and incorporating the JSC-VHH phagemid inside the phage. The resulting phage can be selected by affinity to an antigen; the eluted phage can infect TG1 cells but lacks the machinery for replication and require VCS-M13 helper phage to form phage particles. Several rounds of panning enrich the population with high affinity binders,that can be isolated and cloned into a protein production vector.

Colonies from the serial dilutions were cultured in 3 ml of 2YTAG broth at 37°C overnight for each round of selections. From the culture, 2 ml were used to isolate the JSC-VHH vector by way of a PureYiel Plasmid Miniprep (Promega, Madison, WI, USA). The rest of the culture was placed in storage with 50% glycerol at -80°C. The isolated vectors were commercially sequenced (GATC Biotech, Konstanz, Germany). JSC-seq primer was used for the sequencing reaction.

2.23. VHH PRODUCTION

The selected JSC-VHH clones were cloned into a periplasmic expression vector, pHEN6 (gift of Prof. Hidde Ploegh, Whitehead Institute, MIT, MA, USA). Primers containing NotI, pHEN-VHH F, XhoI, pHEN-VHHshR, and pHEN-VHHlhR were used to amplify the VHH gene and to incorporate the new restriction enzyme sites compatible with the pHEN6 vector. PCR reaction was carried out using a HotStart PCR kit (Qiagen, Crawley, UK) following the protocol described in Section 2.5. PCR products were run in a 1.5% agarose gel, and isolated using Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA) following the manufacturer's recommendations. Isolated PCR products and pHEN6 vector were digested by NotI and XhoI restriction enzymes following the protocol previously discussed in Section 2.11. Isolated digested PCR and vector were used for the ligation reaction using the protocol from Section 2.12. Purified ligation products were electroporated into WK6 electrocompetent cells (gift of Prof. Hidde Ploegh, Whitehead Institute, MIT, MA, USA) and spread into 2YTAG agar plates. Five colonies were selected for plasmid isolation. Colonies were grown in 3 ml of 2YTAG broth overnight at 37 °C with agitation, 200 rpm. Vectors were isolated using a PureYiel Plasmid Miniprep (Promega, Madison, WI, USA), as previously described, and sent to be commercially sequenced (GATC Biotech, Konstanz, Germany). A pHEN-seq primer was used for the sequencing reaction. Once the presence of the VHH and correct reading frame were confirmed, colonies were grown in 1 L of terrific broth (TB) (100 ml of 0.17 M KH2PO4 and 0.74 M K2HPO4, 12 g tryptone, 24 g of yeast extract, 4 ml of glycerol in a litre of water supplemented with 100 ug/ml ampicillin, 0.1% glucose and 2 mM MgCl2) until OD600 = 0.5. Protein production was then induced by the addition of 1 ml of 1 M IPTG (Sigma-Aldrich). Culture was incubated overnight at 28°C with agitation, 200 rpm.

Periplasmic extraction was used to extract the VHH. Cell culture was centrifuged at 4,000 g for 15 minutes at 4°C in a Sorval RC 5C plus centrifuge. Supernatant was removed and the

cell pellet was resuspended vigorously in 12 ml of TES buffer (0.5 M sucrose (Sigma Aldrich), 0.5 mM EDTA, 0.2 M TrisHCl, pH8). The cell suspension was incubated for 1 hour at 4°C with agitation. After the incubation, 24 ml of TES/4 buffer (1 in 4 dilution of TES buffer in water) was added to the cell suspension and incubated overnight at 4°C with agitation. After incubation, 100 µl of 2 M MgCl2 were added and cell suspension was centrifuged at 8,000 g for 30 minutes at 4°C. Supernatant containing the VHH protein was transferred into a new tube and used for further purification.

2.24. VHH Isolation

The vector pHEN6 incorporates a six histidine sequence into the produced VHH at the C terminus, known as His-Tag. The His-Tag has been used as a method for isolation of recombinant proteins due to its high affinity for nickel and cobalt. Nickel particles are bound to a matrix that can be put in contact with the protein solution. After affinity binding, and several washes, a purified solution of the His-Tag protein can be obtained.

The solution containing the target protein, with a C terminus His-Tag, was incubated with one volume of nickel agarose beads (Ni-NTA) (Qiagen, Crawley, UK) for 20 volumes of sample. The mix was incubated at 4°C for 1 hour with rotation. Sample was then loaded into separation columns containing filters. Gravitational force was allowed to drain the supernatant leaving the Ni-NTA beads captured in the column. The column was washed with 20 volumes of Ni-NTA washing buffer (50 mM ThisHCl, 200 mM NaCl, 20 mM imidazol, pH 8) in order to remove all unbound proteins. Bound proteins were removed with three volumes of Ni-NTA elution buffer (50 mM ThisHCl, 200 mM NaCl, 500 mM imidazol, pH 8), and 0.5 ml fractions were collected in 1.5 plastic tubes.

Small aliquots of the fractions were used in SDS-PAGE protein gels stained with Coomassie stain to confirm the presence of purified protein. From each fraction, 20 µl were mixed with 5 µl of gel loading buffer (Table 2. 9). Samples were denatured at 100°C for 5 minutes and loaded into hand cast 4%-12% SDS-PAGE acrylamide gels (Table 2. 9). A protein molecular marker of 3 µl was loaded into the outside well.

Table 2. 9. Ingredients of 12% acrylamide resolving gel, 4 % stacking gel and 4x loading buffer.

12% resolving gel		4% stacking gel		4X SDS Gel Loading Buffer
DDI H2O	4 ml	DDI H2O	6 ml	10% SDS
1.5 M TrisHCl, 0.1 % SDS, pH 8.8	3 ml	1.0 M TrisHCl, 0.1% SDS pH 6.8	2.5 ml	500mM TrisHCl pH 6.8
37% Acrylamide Stock	5 ml	37% Acrylamide Stock	1.25 ml	0.05% Bromophenol Blue
10% Ammonium Persulfate	100 ul	10% Ammonium Persulfate	100 ul	10 mM DTT
TEMED	10 ul	TEMED	10 ul	20% glycerol

Samples were run at 120V until the dye front migrated to the bottom of the gel. Gels were removed from the cast and stained with Coomassie stain buffer (50% methanol, 0.05% Coomassie brilliant blue (Sigma Aldrich), and 10% acetic acid) in a horizontal rocker for 1 hour at room temperature. The gel was then washed with Coomassie de-stain solution (50% methanol, 10% acetic acid) for at least 2 hours until the stain was removed from the background. Gels were then scanned using the Odyssey Infrared Imaging System (Li-cor, Lincoln, NE, USA).

Once the presence of protein was confirmed by SDS-PAGE, the fractions containing the protein were pooled and further purified by gel filtration in a fast protein liquid chromatography system (FPLC). This allowed for separation of the available protein fractions by size obtaining a higher level of purity, and simultaneous buffer exchange for a buffer that can more adequately be used with the VHH. Pooled sample was injected into an AKTA gel filtration system set up with a Hiprep 16/60 Sephacryl S100HR column (GE Healthcare, Little Chalfont, UK). The mobile phase was 50 mM TrisHCl, 150 mM NaCl, pH 7.5, with a flow of 0.5 ml/min. An automatic fraction collector collected 0.4 ml fractions. As before, 20 µl of the fractions with protein content were in 4-12% SDS-PAGE gel and stained with Coomassie stain. Selected fractions were pooled and loaded into a centrifugal concentrator with a cut-off point of 3 kDa (Merk Millipore, Billerica, MA, USA). Concentrator was centrifuged at 4,000 rpm for 1 hour in an Eppendorf 5810R refrigerated bench top centrifuge (Eppendorf, Hamburg, Germany).

2.25. SORTASE PRODUCTION

Sortase are prokaryotic proteins used by bacteria to modify other proteins. Sortase recognise a particular sequence of amino acids, LPXTG, and catalyse the cleavage between the T and G residues. In addition, a transpeptidation takes place that attaches a nucleophile to the peptide. In nature, the nucleophile is a component of the membrane that binds the protein covalently to the membrane, but this can also be used to attach other components such as fluorophores, biotin, or any other components with accessible glycine residues.

BL-21 *E.coli* cells (Stratagene Corp, La Jolla, CA, U.S.A.) were transfected with pET30-sortase A vector (gift of Prof. Hidde Ploegh, Whitehead Institute, MIT, MA, USA). Once the presence of the vector was confirmed by sequencing, cells were grown in 1 L of LB broth until OD600 = 0.7. Protein production was induced with 1 ml 1 M IPTG. Cells were then incubated overnight at 37°C with agitation, 200 rpm. After incubation, cells were centrifuged at 4,000g for 15 minutes at 4°C in a Sorval RC 5C plus centrifuge. Cell pellet was re-suspended in 20 ml of lysis buffer (50 mM TrisHCl, 50mM NaCl, pH 8 supplemented with DNAse I (Roche, Penzberg, Germany)). Cell suspension was sonicated on ice with four short bursts of 30 seconds followed by intervals of 30 seconds for cooling. Cell debris was pelleted by centrifugation at 10,000g for 30 minutes at 4°C in a Sorval RC 5C plus centrifuge. Supernatant was transferred to a new tube and protein was purified by Ni-NTA isolation and FPLC following the protocol described in Section 2.24.

2.26. SORTASE REACTION

Recombinant VHH were produced incorporating a sortase motif LPETG in their C-terminus, recognised by sortase A. Purified VHH of 10 µg were incubated with 150 uM of sortase A enzyme, 1 mM oligoglycine nucleophile (G-G-G-biotin or G-G-G-TAMRA, both supplied by Prof. Hidde Ploegh, Whitehead Institute, MIT, MA, USA), and 1x sortase reaction buffer (5o mM TrisHCl, 0.15 M NaCl, 10 mM CaCl2, pH 7.5). The reaction was incubated overnight at room temperature.

Sortase A enzyme and unreactive VHH were removed using their His-Tag. Reactive VHH lose their His-Tag during the transpeptidation catabolised by the sortase. The sample was incubated with one volume of Ni-NTA agarose beads for 20 volumes of sample for 1 hour at

4°C with rotation. Sample was centrifuged, and supernatant containing the reactive VHH was transferred into a new tube.

2.27. IMMUNOPRECIPITATION

Immunoprecipitation helps to determine the antigen recognised by an antibody by immobilising the antibody in a matrix and incubating a solution with the potential antigen. The antigen recognised by the antibody will be bound to the antibody allowing for the wash of unbound components. In this thesis, two forms of Immunoprecipitation were used.

VHH antibodies were biotinylated using the sortase labelling protocol described in Section 2.26. Biotinylated VHH of 1µg was incubated with 20 µl of Streptavidin agarose beads (Merk Millipore, Billerica, MA, USA) for 1 hour at 4°C with rotation. Beads were washed with 500 µl of Streptavidin agarose washing buffer (50 mM TrisHCl, 150 mM NaCl, 20 mM imidazol, pH 8) by centrifugation and elimination of the supernatant. This washing was repeated three times. Of the total protein from a cell lysate (yeast or Jurkat cells), 100 µg were added into the sample and incubated for 2 hours at 4°C with rotation. Agarose was washed five times with 500 µl of Streptavidin agarose washing buffer followed by addition of 50 µl Streptavidin agarose washing buffer with 10 µl of SDS-PAGE loading buffer to the sample. Protein was denatured at 100°C for 5 minutes.

In order to determine the antigen recognised by the monoclonal anti-BBD126 antibody obtained by a custom antibody production service, immunoprecipitation protocol using protein A/G beads was used. From 1 mg/ml monoclonal antibody, 2 µl were incubated with 100 µg of total protein lysate (sperm cells) overnight at 4°C with rotation. Mix was then incubated with 20 µl protein A/G beads (Santa Cruz Biotechnology, Santa Cruz, CA, U.S.A.) for 1 hour at 4°C with rotation. Beads were washed with 500 µl of protein A/G agarose washing buffer (0.05% Tween-20, PBS) by centrifugation. This washing was repeated five times. Protein A/G agarose washing buffer (50 ul), with 10 µl of SDS-PAGE loading buffer, were added to the sample. Protein was denatured at 100°C for 5 minutes. Samples were then loaded into a 4-12% 18 x 18 cm SDS-PAGE gel, and run at 120V until the dye front migrated to the bottom of the gel.

2.28. SILVER STAINING

A 4-12% 18 x 18 cm SDS-PAGE gel was fixed for 1 hour with a 40% ethanol, 10% acetic acid solution. Gel was then washed twice for 20 minutes with 30% ethanol solution, and once with deionised water for 20 minutes. The gel was sensitised for 1 minute with a 0.02% Na2S203 solution and washed three times with deionised water for 20 seconds. The gel was then incubated with cold 0.1% AgNO3 solution at 4°C for 20 minutes, and washed three times with deionised water for 20 seconds. The gel was developed in 3% Na2CO3, 0.05% formaldehyde until visible bands appeared in the gel. The reaction was stopped using 5% acetic acid solution. The gel was stored in 1% acetic acid solution at 4°C. Bands were cut using a scalpel and sent for mass spectrometry protein identification to Professor Ploegh, Whitehead Institute, MIT, MA, USA, or commercially to Fingerprints (Dundee, UK).

2.29. CONSTRUCTION OF ARTIFICIAL VHH GENE

Sequences of the VHH genes retrieved during the panning for yeast, and Jurkat libraries, were aligned using BioEdit Software version 7.0.9 (Hall 1999). A consensus sequence for the highly conserved FR regions, and the number of amino acids in each of the CDR regions, was built. Primers were built using Ape. Plasmid editor was used for the design of primers for primer walk. One forward primer (P1F) and five reverse primers (P2R, P3R, P4R, P5R, P6R) were designed. Primers P2R, P3R and P5R contained random nucleotide sequences in the same length of the corresponding CDR1, CDR2 and CDR3 (Fig. 2. 3). CDR1 and CDR2 have a conserved number of amino acids, but CDR3 presents high variability and we observe from 9 to 19 amino acids. To replicate this effect, a total of 12 P5R with different lengths in the CDR3 region, in increments of 3 nucleotides from 27 up to 57 nucleotides, were used. In addition, P1F and P6R incorporate the restriction sites for AscI and NotI for cloning into the JSC vector.

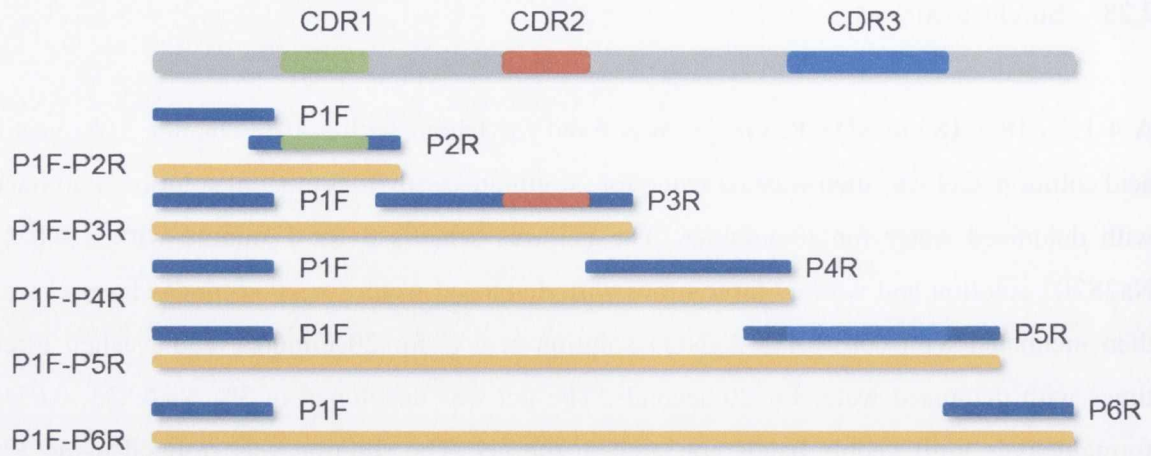


Fig. 2. 4 Schematic of the structure of the VHH gene, with the highly variable CDR regions. Seven primers were designed; P1F, P2R, P3R, P4R, 12 P5R and P6R. Of these P2R, P3R and P5R contain random sequences in the same position of the CDR regions. The gene was constructed by sequential PCR reactions. P1F and P2R generate the P1F-P2R PCR product that was used as sample for the reaction with P1F and P3R. With each PCR reaction, the product was enlarged until the final product was obtained with the P1F and P6R reaction.

PCR reactions were carried out using a HotStart PCR kit (Qiagen, Crawley, UK), following the protocol described in Section 2.5. PCR reaction, 50 ul, using P1F and P2R primers, was loaded into a 1.5% agarose gel. PCR product was cut from the gel and DNA was isolated using a Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA). The resulting PCR product was then used as a template for the PCR reaction using P1F and P3R. The purified PCR product of the reaction was used as a template for the P1F and P4R PCR reaction. P1F-P4R products were purified and used for the reaction with P1F and a pool of all 12 versions of P5R. In the final reaction, the clean P1F-P5R was used for the P1F and P6R reaction. The P1F-P6R product was loaded into a 1.5% agarose gel. PCR product was cut from the gel, and DNA was isolated using a Wizard SV Gel and PCR Clean-Up (Promega, Madison, WI, USA). Following the protocol from Section 2.20, the P1F-P6R product was cloned into the JSC vector. Selection and VHH production was carried out as previously discussed in Sections S2.21 to 2.24.

2.30. BOVINE TNF ALPHA ELISA

Bovine TNFα Duo Set and mouse IL4 Duo Set ELISA kit (R&D Systems, Minneapolis, MN, U.S.A.) were used to test the ability of two VHH, B1, and B2 to bind and detect bovine TNFα. ELISA plates (Greiner Bio-One, Frickenhausen, Germany) were loaded, in duplicate, with 20 µl of VHH preparation as shown in Table 2.10, and incubated overnight at room temperature.

Table 2. 10. Description of capture antibody loaded in the different wells of the ELISA plate.

	TNFα capture Ab	B2	ENH	B1
Blank	-	-	-	-
TNF control	+	-	-	-
TNFα	-	+	-	-
IL4	-	+	-	-
TNFα	-	-	+	-
IL4	-	-	+	-
TNFα	-	-	-	+
IL4	-	-	-	+

Buffer was removed by aspiration, and wells were washed three times with 400 µl of PBS buffer. The wells were loaded with 300 µl of blocking buffer (2% BSA in PBS) and incubated for 1 hour at room temperature. TNFα (8,000 pg/ml; 100 ul) and IL4 standard were loaded in the wells as shown in Table 2.11 and incubated for two hours at room temperature.

Table 2. 11. Description of standards loaded in the different wells of the ELISA plate.

	TNFα standard	IL4 standard
Blank	+	-
TNF control	+	-
TNFα	+	-
IL4	-	+
TNFα	+	-
IL4	-	+
TNFα	+	-
IL4	-	+

Wells were subsequently washed four times with 400 µl of ELISA washing buffer (0.05% Tween-20 in PBS). Following this, the wells were loaded with 100 µl of detection antibody, and TNFα or IL4 from the commercial ELISA kits as shown in Table 2.12. Antibodies were incubated for two hours at room temperature.

Table 2. 12. Description of detection loaded in the different wells of the ELISA plate.

	TNFα detectionAb	IL4 detection Ab
Blank	+	-
TNF control	+	-
TNFα	+	-
IL4	-	+
TNFα	+	-
IL4	-	+
TNFα	+	-
IL4	-	+

Wells were washed four times with 400 µl of ELISA washing buffer (0.05% Tween-20 in PBS), and 100 µl of Streptavidin-HRP were added to each well followed by incubation (4°C) for 20 minutes. After four rounds of washing with 400 µl of ELISA washing buffer (0.05% Tween-20 in PBS), 100 µl of substrate solution (KPL, Gaithersburg, MD, U.S.A.) were added to the wells and incubated in the dark for 20 minutes at room temperature. Stop solution ($2NSO_4H_2$; 50 ul) was added and the plate was read at 450 nm.

2.31. CUSTOM PRODUCTION OF ANTI-BBD126 MONOCLONAL ANTIBODY

A custom monoclonal antibody specific for BBD126 was ordered from GenScript (Piscataway, NJ, USA). A 14 amino acid peptide, RNGERVINPPTGMC (Fig. 2. 4), was chemically produced and conjugated to KLH. The conjugate was then inoculated in five BALB/c mice. After immune response was confirmed, cells were isolated for cell fusion and hybridoma production. Four hybridomas were selected and tested in ELISA against the peptide. This work was performed by GenScript as part of the custom antibody production service.

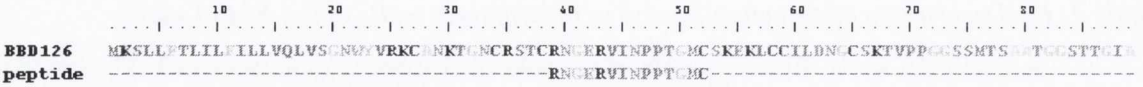


Fig. 2. 5 Protein sequence alignment of peptide inoculated in mice for antibody production and BBD126.

2.32. IMMUNOCYTOFLUORESCENCE OF BBD126 IN SPERM CELLS

Fresh sperm sample (1 ml) was centrifuged at 12,000 rpm for 2 minutes in a table top centrifuge. Supernatant was removed and cell pellet was re-suspended in 500 µl of sterile PBS. Cells were centrifuged again at 12,000 rpm for 2 minutes and supernatant was removed. Pellet was re-suspended in 500 µl of sterile PBS. Cells were diluted 1x to 100x with sterile PBS, and 500 µl of the cell dilution were loaded into the chambers of a EZ culture slide (Merk Millipore, Billerica, MA, USA). Cells were cultured overnight at 37°C. Supernatant was removed by aspiration and replaced with 200 µl of a 4% formaldehyde solution. The slide was incubated on ice for 30 minutes. Chambers were washed three times for 5 minutes with 400 µl of sterile PBS. Blocking buffer (0.01% BSA, 0.002% Triton X, PBS) of 200 µl was added to each chamber and incubated for 1 hour at room temperature. Supernatant was removed by suction, and the primary antibody, mouse monoclonal anti-BBD126, was loaded into different chambers in 1 ug, 100 ng, and 10 ng of total antibody as shown in Table 2.13. This was then incubated overnight at 4°C.

Table 2. 13 Immunocytofluorescence experimental design.

Distribution of sample in the EZ culture slide, b. Antibody amounts used for each well. Cells chamber was untreated with any antibody. Negative control was treated with anti-mouse IgG secondary to measure fluorescence background. 1, 2 and 3 were treated with an increasing amount of anti-BBD126 antibody and 1/500 dilution of anti-mouse IgG secondary antibody.

a.

cells	1	2	3
negative control	1	2	3

b.

	anti-BBD126	anti mouse IgG
cells	-	-
negative control	-	1/500
1	10 ng	1/500
2	100 ng	1/500
3	1 ug	1/500

Chambers were washed five times for 5 minutes with sterile PBS. From a 1/500 dilution of an anti-mouse IgG Alexa Fluor 568 conjugated antibody (Invitrogen Ltd., Paisley, UK), 100 µl were added to the corresponding chambers and incubated for 1 hour at room temperature in the dark. Chambers were washed five times for 5 minutes with sterile PBS, and chambers were removed from the slide. The slide was covered with 10 µl of Prolong Gold Dapi (Invitrogen Ltd., Paisley, UK) and covered with a cover glass. Pictures were taken using an Olympus Fv100 point scanning confocal microscope.

2.33. Recombinant production of BBD126

The expression system used was eukaryotic using the pEAK8 vector (a gift from Prof. Dirk Werling, Royal Veterinary College, London, UK.) in human embryonic kidney derived cells (HEK293T). The coding region of BBD126, including both exon 1 signal peptide and exon 2 mature peptide, was amplified using the HotStart PCR kit (Qiagen) as described in Section 2.5. The primers used for the reaction were pEAK8-126F and pEAK8-126R. These primers incorporate HindIII and XbaI restriction enzyme sites into the PCR product. The resulting PCR product and the pEAK8 plasmid were digested with HindIII and XbaI restriction

enzymes following the protocol from Section 2.11. The products of the digestion were ligated, and the ligation product was transfected into One Shot TOP10 Chemically Competent E. coli (Invitrogen Ltd., Paisley, UK) as described in Section 2.13. Cells were spread and incubated overnight in LB agar plates supplemented with 100 ug/ml of ampicillin. Of the growing colonies, five were selected to be grown for mini-prep plasmid isolation as described in Section 2.14. Sequences and correct reading frame were confirmed by commercial DNA sequencing (GATC Biotech, Konstanz, Germany) using pEAK8-seq primer. A selected colony was used to inoculate 100 ml of LB broth supplemented with 100 ug/ml of ampicillin. The culture was grown overnight at 37°C with agitation, 200 rpm. PureYiel Plasmid Maxiprep (Promega, Madison, WI, USA) was used to isolate a large quantity of plasmid. The DNA was eluted into 500 µl of DNAse free water. DNA concentration was determined using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

HEK293T cells were grown in 5 ml of Dulbecco's modified Eagles medium (DMEM) supplemented with 10% foetal calf serum. Cells (10^6) were seeded into 75 cm^2 tissue culture flasks and incubated overnight at 37°C with 5% CO_2. Of the pEAK8-126, 6 µg were incubated with 10 µl of GeneJuice Transfection Reagent (Merk Millipore, Billerica, MA, USA) and serum-free DMEM up to a total volume of 250 µl for 30 minutes at room temperature. As negative control, 6 µg of pEAK8 were incubated with 10 µl GeneJuice Transfection Reagent (Merk Millipore, Billerica, MA, USA), and serum free DMEM, up to a total volume of 250 µl for 30 minutes at room temperature. The mixes were then added to different tissue culture flasks with the growing HEK293T cells. Cells were grown for 72 hours at 37°C with 5% CO_2. Media and cells were collected for protein analysis. Media was used for Ni-NTA isolation of the His-Tag recombinant BBD126 as described in Section 2.24.

Cell pellets were lysated using 100 µl of protein analysis lysis buffer (2.5 ml of RIPA buffer consisting of50 mM Tris HCl, 150 mM NaCl, 1% Triton X, 0.5% sodium deoxycholate and 0.1% SDS), 25 µl PMSF (phenylmethanesulfonylfluoride), 16 µl aprotinin, 5 µl sodium orthovanadate and 2.5 µl of leupeptin). Cells were incubated for 30 minutes on ice and the centrifuge at 13,000 rpm at 4°C in a table top centrifuge. Supernatant was transferred into new tubes and stored at -20°C until used. Protein concentration was determined by Bradford protein assay (Thermo Fisher Scientific, Waltham, MA, USA).

2.34. Western Blot analysis of recombinant BBD126

From the total protein, 20 µg were loaded in 4-12% SDS-PAGE gel and run at 120V until the dye front reached the end of the gel. Protein was then transferred into a 0.45 um PMSF membrane (Merk Millipore, Billerica, MA, USA). A piece of PMSF membrane was first activated with methanol for 20 seconds and then washed with deionised water for 2 minutes. Activated membrane and gel were then equilibrated with transfer buffer (2 mM TrisBase, 200 mM Glycine, 20% methanol). A sandwich made with a cellulose pad wet in transfer buffer, PMSF membrane, gel, and wet cellulose pad was prepared in the semi-dry transfer cell (Bio-Rad, UK). Transfer was carried out at 15V for 30 minutes. After transfer, membrane was blocked using blocking buffer (0.5% BSA, 0.05% Tween-20 in PBS) for 1 hour while rocking.

Of the 1 mg/ml anti-BBD126 mouse antibody, 20 µl were added to 30 ml of washing buffer (0.05% Tween-20 in PBS) and incubated with the membrane overnight at 4°C while rocking. Membrane was then washed with washing buffer for 5 minutes. This was repeated three times. Membrane was then incubated with 3 µl of anti-mouse IgG rabbit secondary antibody in 30 ml of washing buffer. Secondary antibody was conjugated to the 680LT fluorophore (Li-cor, Lincoln, NE, USA). Membrane and secondary antibody were incubated for 1 hour at room temperature while rocking. Membrane was then washed three times with washing buffer for 5 minutes. Membrane was scanned using the Odyssey Infrared Imaging System (Li-cor, Lincoln, NE, USA).

2.35. Localization of BBD126 in the bovine male reproductive tract tissue.

Samples of bovine cauda of the epididymis and seminal tubule were preserved in 37% formaldehyde solution since collection in the abattoir. Fragments of the samples were imbedded in paraffin wax. Once solid, 0.7 um sections were obtained using a microtome and transferred onto glass slides. Slides were processed on an automated slide preparation system (Leica, Wetzlar, Germany). Once processed, slides were heated at 60°C for 10 minutes until the wax was softened. Sections were deparaffinised with Histoclear (National Diagnostics, Charlotte, NC, U.S.A) for 10 minutes. Sections were placed in absolute alcohol for 10

minutes and then 95% industrial mutilated solvent for 5 minutes. Finally slides were washed with tap water for 5 minutes before rinsed with deionised water. A pressure cooker was used to boil 1,500 ml of epitope retrieval solution (0.02% citric acid, 0.08% trisodium citrate pH 6 in H2O). The sections were loaded into the solution and the samples were heated for 5 minutes until boiling. After cooling down, sections were washed with running water. Staining was carried out using a Novolink Max Polymer detection Kit (Novacastra, Newcastle, UK) which utilised 5 µl of 1 mg/ml anti-BBD126 antibody following the manufacturer's instructions.

CHAPTER 3: mRNA EXPRESSION PROFILING OF DEFENSIN GENE CLUSTER

3.1. INTRODUCTION

β-Defensins are relatively small, cationic proteins produced by multiple cell types, including epithelial and immune cells in mammals (Ganz 2003), which act as important effector molecules of innate immunity with significant antimicrobial activity (Motzkus, Schulz-Maronde et al. 2006, Meade, Cahalane et al. 2008). Traditionally thought to be simple antimicrobial effector molecules, recent studies have revealed a wide range of immune regulatory functions for these peptides (Yang, Chertov et al. 1999, Biragyn, Ruffini et al. 2002, Bowdish, Davidson et al. 2005, Mookherjee, Wilson et al. 2006, Funderburg, Lederman et al. 2007, Lai and Gallo 2009). Cross-species comparison of defensin genes revealed them to be highly variable in number (Semple, Gautier et al. 2006, Cormican, Lloyd et al. 2009), presumably due at least partly to differences in pathogenic selective pressures among the different species but also due to genetic drift and historical contingency. The number of clusters also varies among species with four β-defensin gene clusters reported in dog, rat, mice and cattle and five in primates (Patil, Cai et al. 2005). In primates, including humans, one of the clusters is split in two as a result of a primate-specific episode of perocentric inversion producing five gene clusters (Lehrer and Ganz 2002, Schutte, Mitros et al. 2002). The four β-defensin gene clusters that have been identified in the bovine genome are located on chromosomes 27, 23, 13 and 8 (Cormican, Meade et al. 2008). The most studied of all is the cluster on bovine chromosome 27 where the tracheal antimicrobial peptide *(TAP),* lingual antimicrobial peptide *(LAP)* and other well-known defensins are found (Caverly, Diamond et al. 2003, Swanson, Gorodetsky et al. 2004, Lopez-Meza, Gutierrez-Barroso et al. 2009). Our group, using homology searches with the BLAST family of programmes (Altschul, Gish et al. 1990) and Hidden Markov Models (Rodriguez-Jimenez, Krause et al. 2003), identified a cluster of β-defensin genes in the bovine genome on chromosome 13, showing a similar gene distribution to that already described in other genomes (Schutte, Mitros et al. 2002). After analysis of the expression of selected genes across multiple tissues, we found predominant expression in the male testis in the absence of infection, suggesting an additional non-defence role (Cormican, Meade et al. 2008).

β-Defensin gene expression can be divided between the constitutively expressed pattern found with, for example, bovine neutrophils defensins 4 and 5 (BNBD4, BNBD5) (Selsted, Szklarek et al. 1984), expressed in neutrophils in the absence of a inflammatory process, and the inducible gene expression pattern found with, for example, bovine LAP and TAP (Diamond, Russell et al. 1996), stimulated by the detection of LPS by TLR4 and subsequent

activation by NFkB. Independently of the regulation of the gene expression, most β-defensins have broad expression patterns which are found expressed in multiple locations through the body, generally in epithelia. An exception is the group of β-defensins expressed in a very specific pattern through the different sections of the male genitalia found in human, mouse, rat and macaque (Com, Bourgeon et al. 2003, Tollner, Yudin et al. 2008). Hypotheses of the possible role of these β-defensins in the reproductive tract include a role in reproduction and fertility for some of these genes or a role in the prevention of urogenital infections. The presence of multiple genes expressed in specific locations in the reproductive tract suggest that each defensin could have a specific function in each section (Com, Evrard et al. 2003, Jelinsky, Turner et al. 2007). Members of this group of β-defensins have been shown to facilitate sperm maturation (Zhou, Zhang et al. 2004), sperm capacitation (Tollner, Yudin et al. 2008) and to protect the sperm against the female immune response (Yudin, Generao et al. 2005, Taylor, Rath et al. 2008) in other species.

A particular case is human DEFB126. Recently, a SNP in human DEFB126 has been highly correlated with fertility reduction in men (Tollner, Venners et al. 2011). The authors hypothesised that the mutation caused an effect in the glycocalyx structure of DEFB126 (Tollner, Bevins et al. 2012). Synthesised by the epithelial cells of the epididymis, DEFB126 was found coating the surface of sperm cells. β-Defensins are known for containing hydrophobic motifs that could interact with lipid membranes. This interaction grants a strong bond to the membrane that is not easily broken (Zanich, Pascall et al. 2003). DEFB126 presents 20 potential sites for O-linked glycosylation. Glycosylation accounts for more than half the weight of DEFB126, from 32-35 kDa to 10 kDa when glycosylation is removed. The addition of carbohydrates increases the cationic charge of the sperm surface. This can increment the ability of sperm to swim through the negative change cervical mucus. A mutation affecting the formation of the glycocalyx would reduce the surface charge added by the carbohydrates, impacting on the swimming ability of the sperm cells. In addition, DEFB126 has been shown to block the recognition of sperm cell surface protein by antibodies (Yudin, Generao et al. 2005). The enzymatic removal of glycosylation allowed the recognition of the surface epitopes by the antibodies. The presence of DEFB126 is essential in the formation of the oviductal reservoir. Sperm cells, after ejaculation, bind to the epithelium of the oviduct until ovulation occurs. This process is required to reduce polyspermy: multiple sperm cells fertilising a single egg. The formation of the reservoir is reduced in the absence of DEFB126 (Tollner, Yudin et al. 2008).

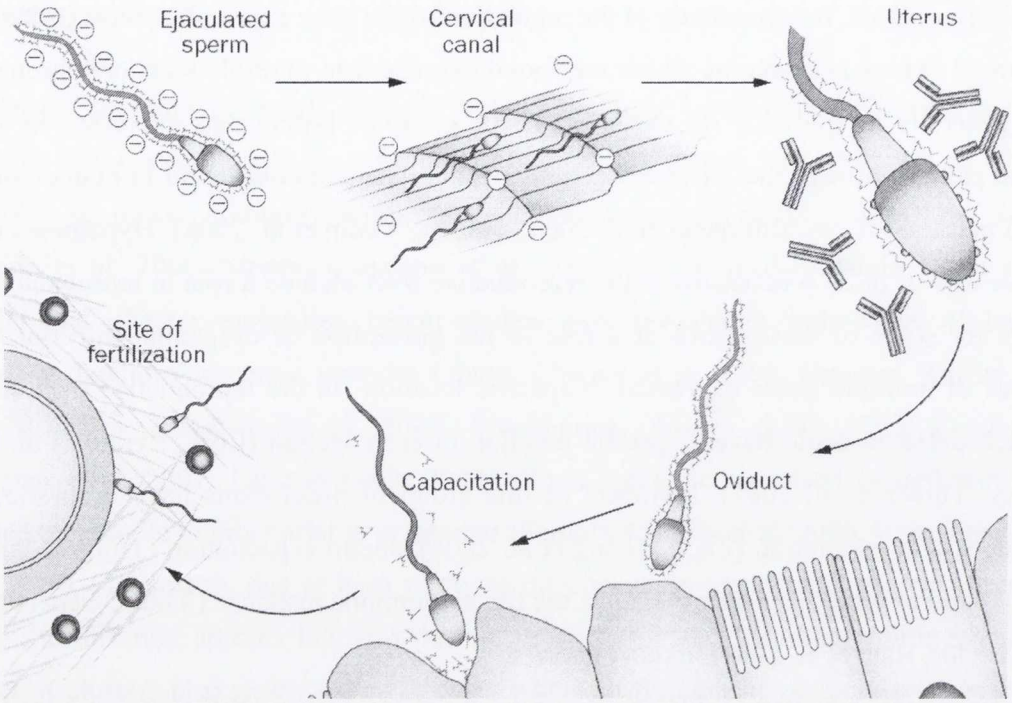


Fig. 3. 1. Roles of DEFB126 in fertility.

DEFB126 coats the ejaculated sperm, adding a negative change to the sperm surface. The cationic charge increased the ability of the sperm cell to penetrate the cervical mucus present in the cervical canal. Another role is the immunoprotection of surface sperm molecules that could be recognised as foreign by the female immune system. DEFB126 is able to block the access of antibody to the sperm cell surface molecules. The final role is the formation of the oviductal sperm reservoir after ejaculation sperm cells bind to the oviductal epithelium and are only free after ovulation occurs in a highly synchronised process that reduces the occurrence of polyspermy. DEFB126 is lost during the final stage of sperm maturation known as capacitation once the sperm cells are free from the oviductal reservoir. Figure adapted from (Tollner, Bevins et al. 2012).

Our initial study of the newly discovered bovine Chr13 β-defensin gene cluster showed an absence of gene expression of the analysed genes. *BBD119, BBD120, BBD122, BBD122a, BBD123* and *BBD124* showed strong expression in bovine testis with no or low expression in other organs (Cormican, Meade et al. 2008). A further analysis of the gene expression of the members of bovine Chr13 β-defensins in the male and female reproductive tract is required. In addition, an understudied aspect is the regulation of these genes. The dogma suggest that β-defensins as effector molecules of the immune systems are regulated by detection of danger

signals such as pathogens by PRR's, specifically TLR signalling through NF-kB activation (Diamond, Russell et al. 1996). This seems to be the case with some β-defensins. However, constitutive expressed defensins are expressed in the absence of immune response, suggesting the presence of alternative methods of gene regulation. Hormones have been known to stimulate antimicrobial peptide gene expression for some time (Wang, Nestel et al. 2004). The effect of castration in the gene expression of epididymal β-defensins is also known (Yenugu, Chintalgattu et al. 2006), suggesting a role of sexual hormones in the regulations of β-defensin genes in the epididymis.

3.2. RESULTS

3.2.1. RNA EXTRACTION

RNA extraction from reproductive tissue was optimised for high yield and quality required for qPCR. After different protocols and modifications the protocol, described in the Materials and Methods section was used for RNA extraction (the concentration values can be found in the Appendix: Table 1 and 2).

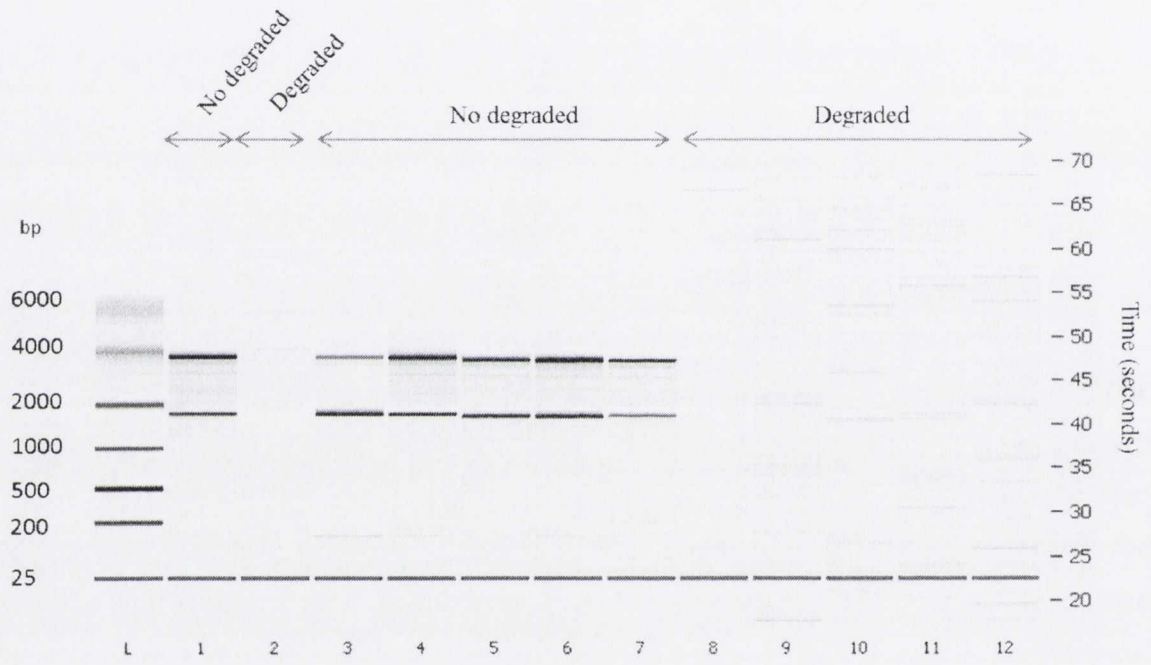


Fig. 3. 2. Bovine reproductive tissue RNA quality assessment

Example of RNA quality assessment by Agilent Bioanalyzer. L, RNA ladder. Degraded samples 2, 8, 9, 10, 11 and 12 do not present the two ribosomal RNA bands 28s and 18s. Non degraded samples, 1, 3, 4,, 5, 6 and 7 present the two ribosomal bands indicating no degraded RNA. Only samples with no degraded ribosomal RNA were used for gene expression analysis.

3.2.2. NORMALISER GENE SELECTION QPCR.

Ge-Norm software was used to determine the most stable reference gene for the samples analysed. Several normaliser genes (*ACTB, GAPDH, MRPS6, PPIA, RPS9, RPS15* and

H3F3A) were used to amplify using the same samples under the same conditions. The expression stability is calculated based on the geometric mean. Male and female tract samples were analysed independently to determine the most stable among them for both female and male samples. *GAPDH* was the most stable gene among the female and male samples and it was selected as normaliser for the qPCR analysis (Fig. 3. 3 and Fig. 3. 4). In contrast *ACTB* was the least stable gene in both male and female samples.

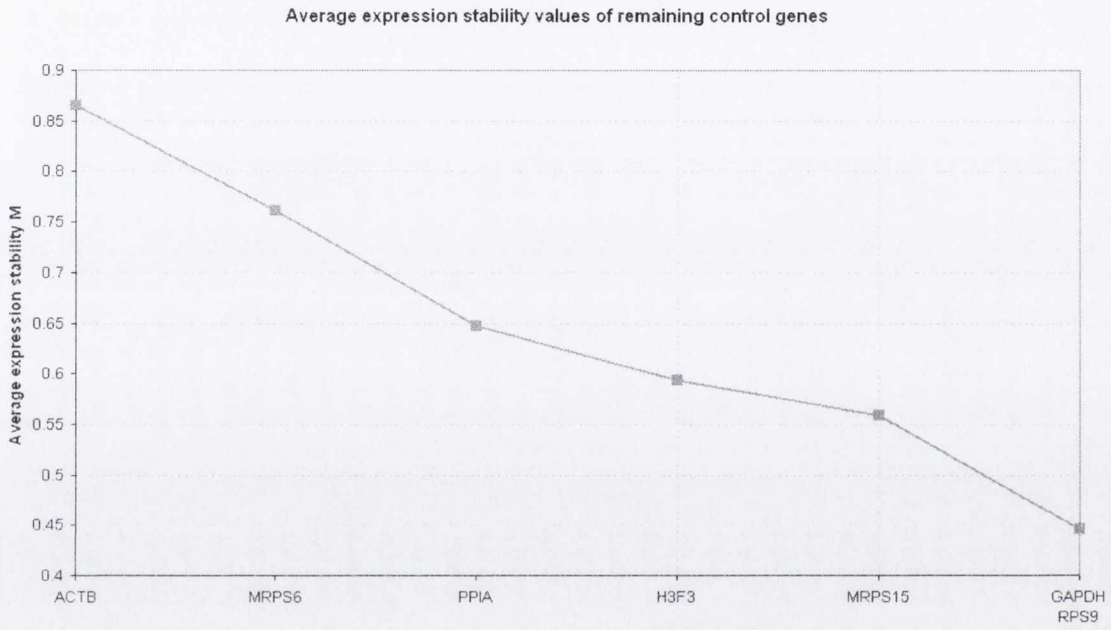


Fig. 3. 3. Average expression of stability of housekeeping genes on the male reproductive tract samples.

A set of male samples were used to amplify ACTB, GAPDH, MRPS6, PPIA, RPS9, RPS15 and H3F3A. The Ct values were analysed by GeNorm software (Vandesompele, De Preter et al. 2002). The geometrical mean was used to determine GAPDH as the most stable gene to use as reference for the male tract samples.

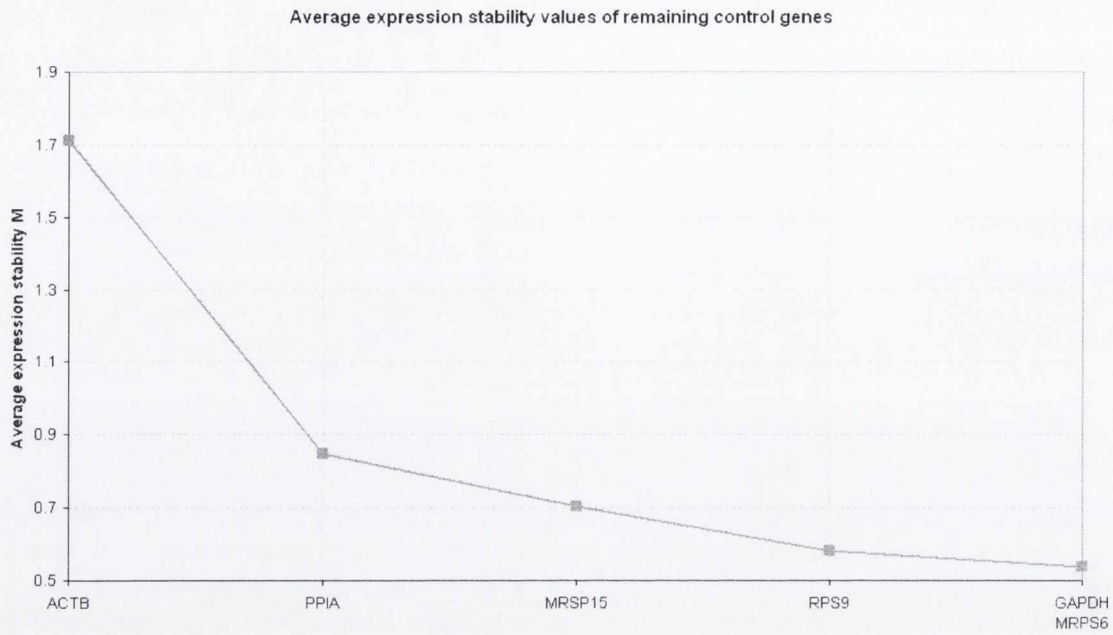


Fig. 3. 4. Average expression of stability of housekeeping genes on the female reproductive tract samples.

A set of male samples were used to amplify ACTB, GAPDH, MRPS6, PPIA, RPS9, RPS15 and H3F3A. The Ct values were analysed by Ge-Norm software (Vandesompele, De Preter et al. 2002). The geometrical mean was used to determine GAPDH as the most stable gene to use as reference for the female tract samples.

3.2.3. Novel β-defensin gene expression

Expression for the 19 novel β-defensins, the widely studied defensin *LAP* and the normaliser gene *GAPDH* was comprehensively examined in a panel of samples from different organs and in genitalia from males using quantitative real-time polymerase chain reaction (qPCR). Post-mortem tissue samples were taken from lung, rumen, small intestine, large intestine, liver, spleen, peripheral blood, lymph node, mammary gland and universal reference tissue pool from one healthy animal. A pool of cDNA from all the mentioned tissues as well as uterine and testis and 6 regions of the male genital tract (seminiferous tubules, rete testis, caput, corpus and caudal epididymis and vas deferens). All samples were retrieved from healthy animals with no obvious infection or inflammation. Results show reproducible regional tissue specificity in β-defensin gene expression across tissues of the male genital tract and an absence of expression of the novel β-defensins in other organs (Fig. 3. 5). The

exception was *β-defensin 117* (*BBD117*), which is present in several organs. Weak amplicons were also detected in the universal bovine cDNA pool (UB) which contains uterine cDNA for *BBD122* and *BBD122a*. In contrast, LAP was used as positive control to show the most common distribution of β-defensin gene expression in the tissues used. LAP shows gene expression in most of the tissues analysed.

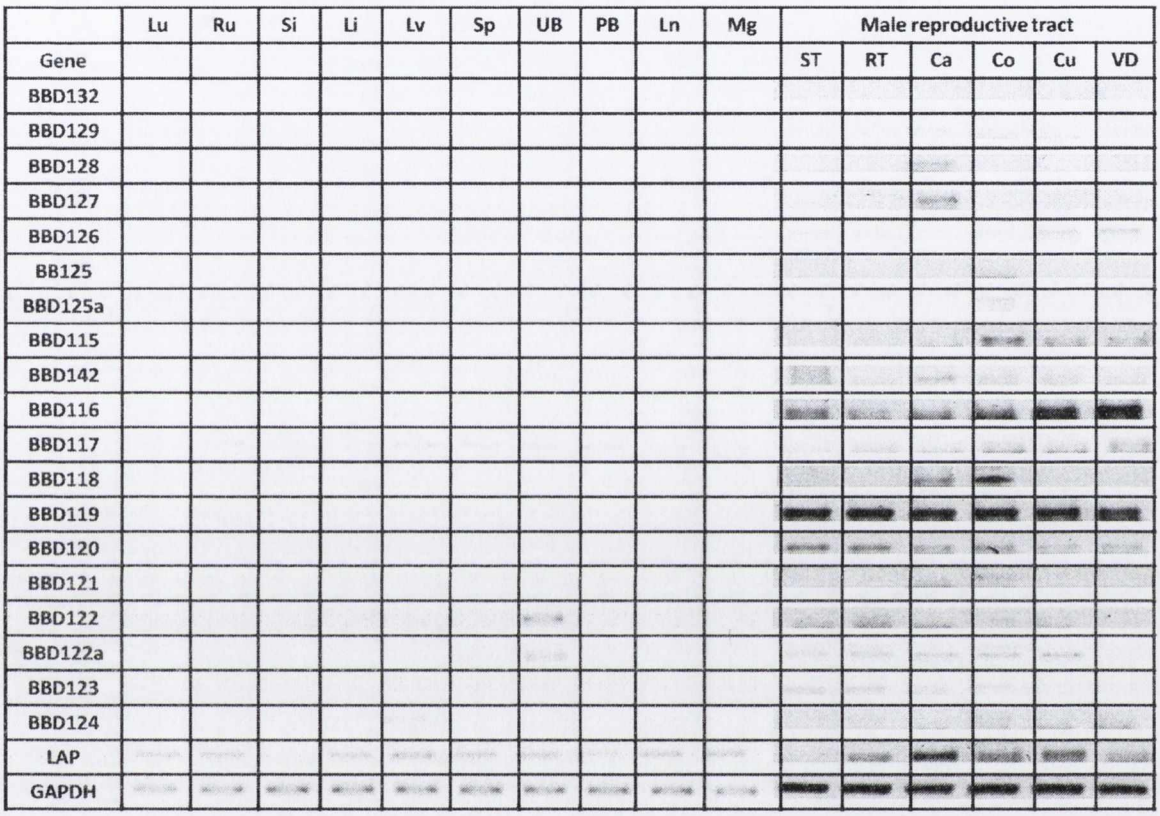
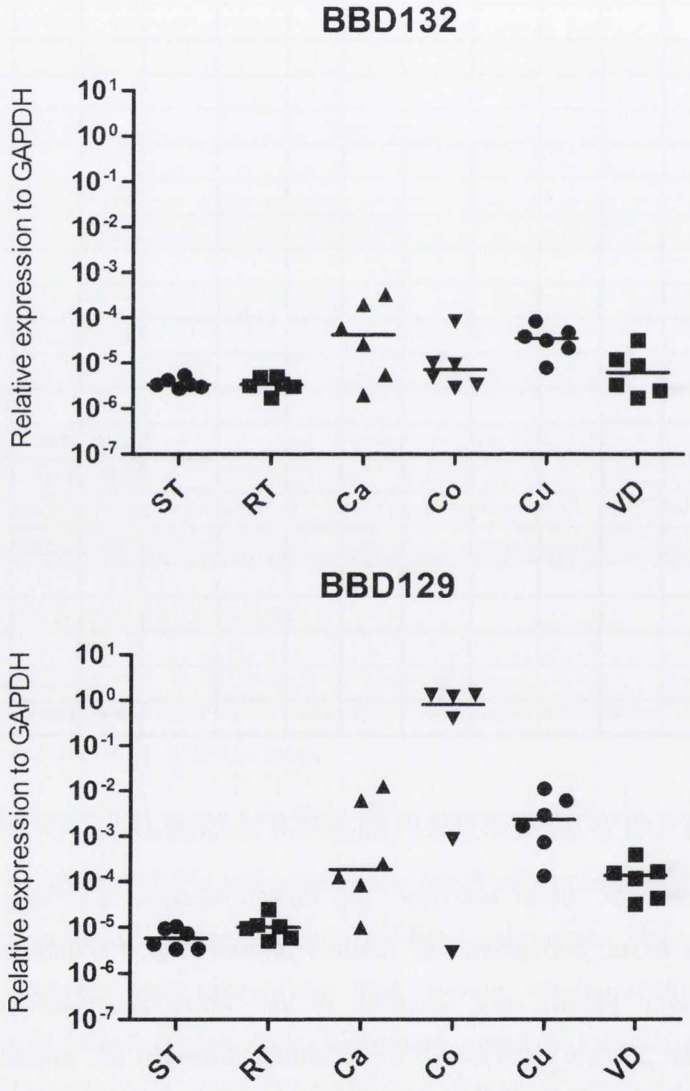


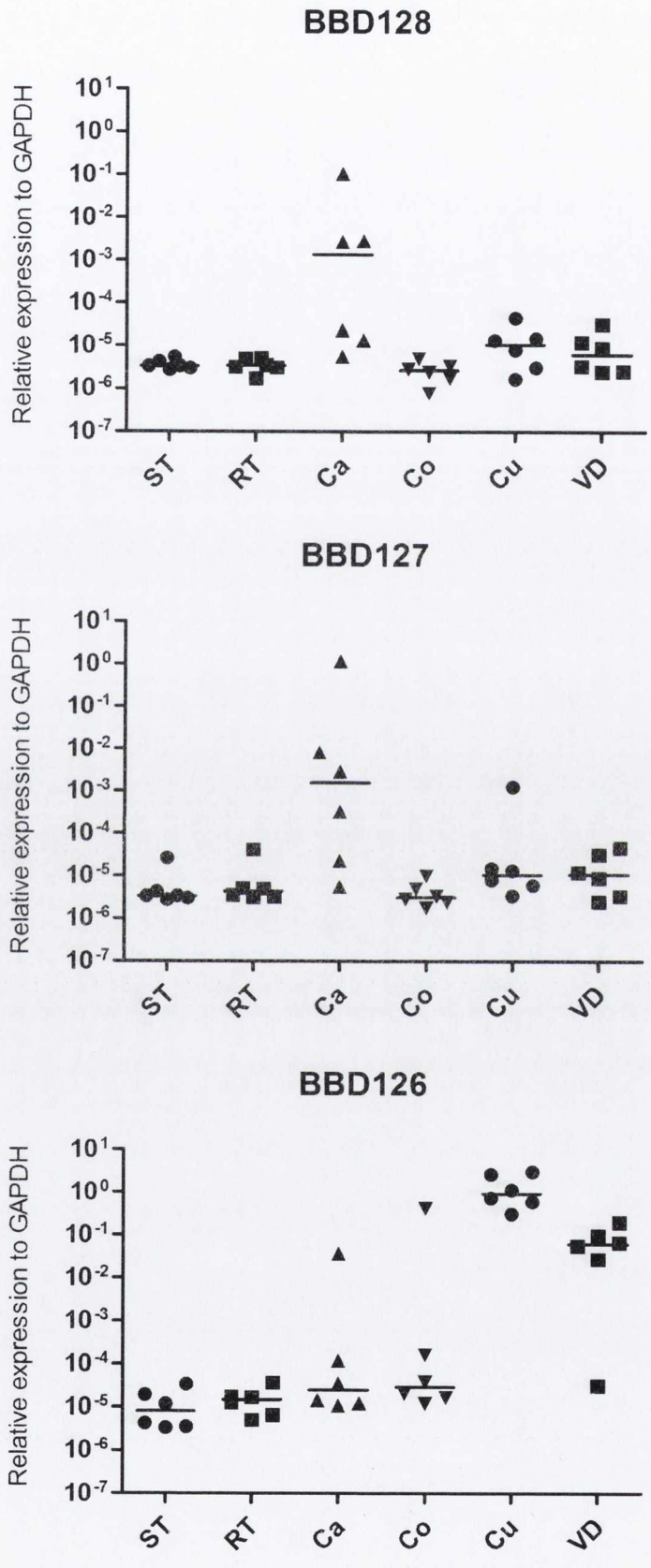
Fig. 3. 5. Agarose gel of qPCR products of the different genes in a panel of tissues.

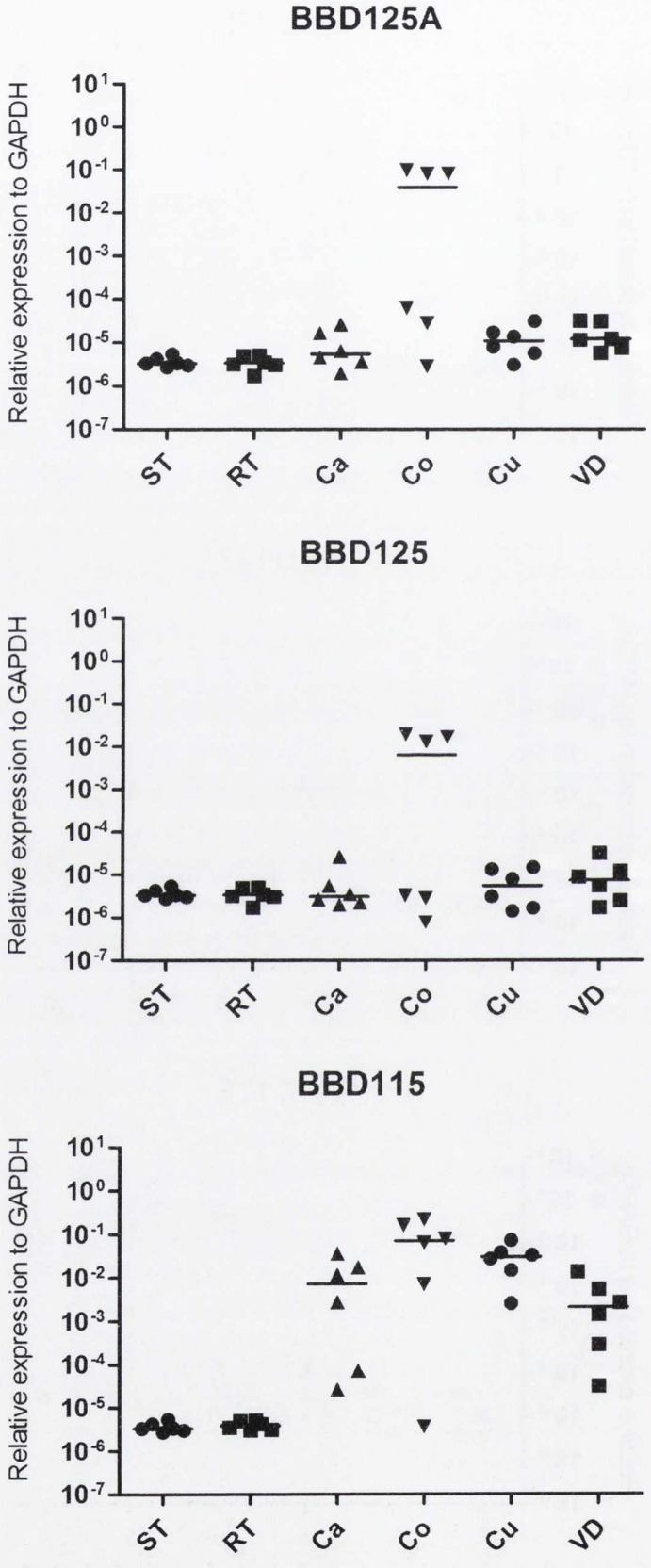
Lu, lung; Ru, rumen; Si, small intestine; Li, large intestine; Lv, liver; Sp, spleen; UB, universal reference tissue pool from one healthy animal; PB, Peripheral blood; Ln, lymph node; Mg, mammary gland; and sections of the male reproductive tract of a bull representative of the pattern detected in 6 individual animals, ST, seminal tubule; RT, rete testis; Ca, caput; Co, corpus; Cu, cauda and VD, vas deferens were analysed by qPCR to detect gene expression for BBD115, BBD116, BBD117, BBD118, BBD119, BBD120, BBD121, BBD122, BBD122a, BBD123, BBD124, BBD125, BBD125a, BBD126, BBD127, BBD128, BBD129, BB132, BBD142, LAP and GAPDH. Gene expression for Chr13 cluster β-defensins was found mostly in the reproductive tract section, in contrast with LAP which was expressed almost in all tissue.

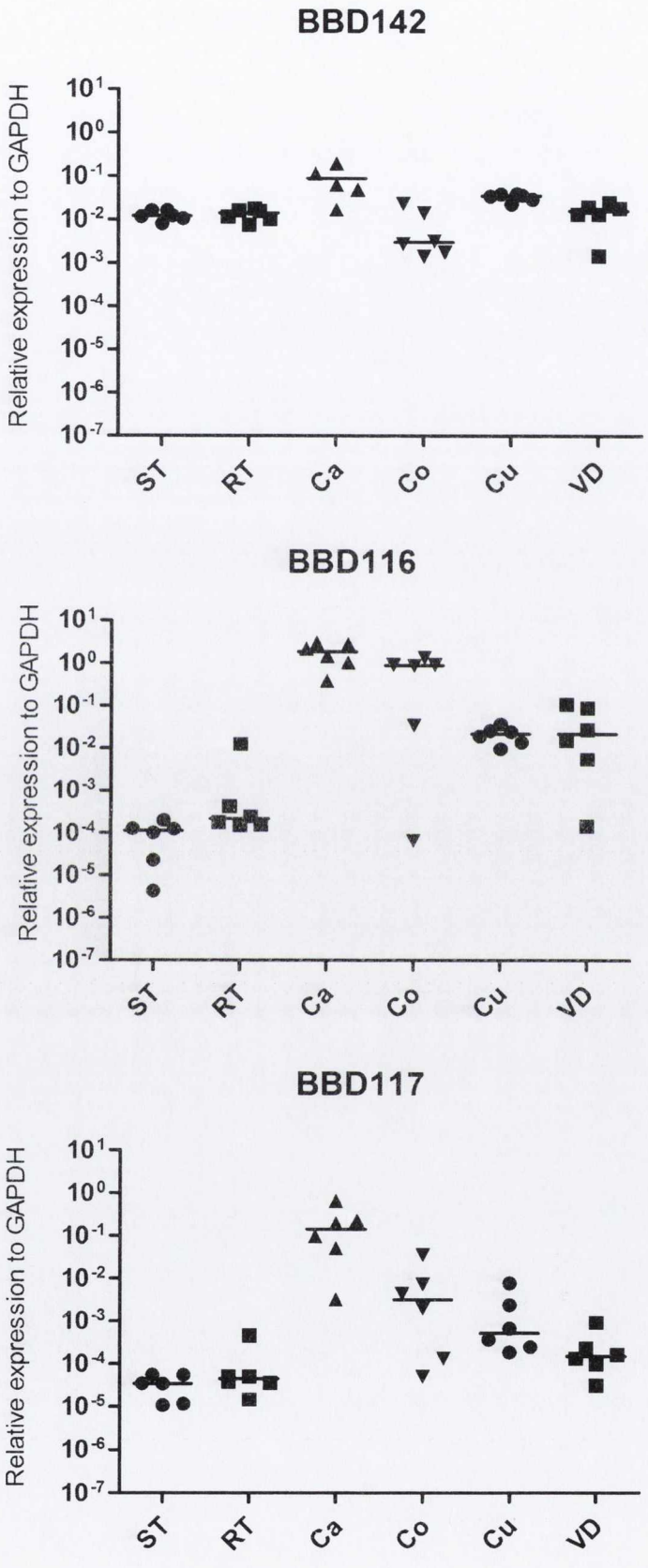
3.2.4. Regional specific β-defensin gene expression in male reproductive tract

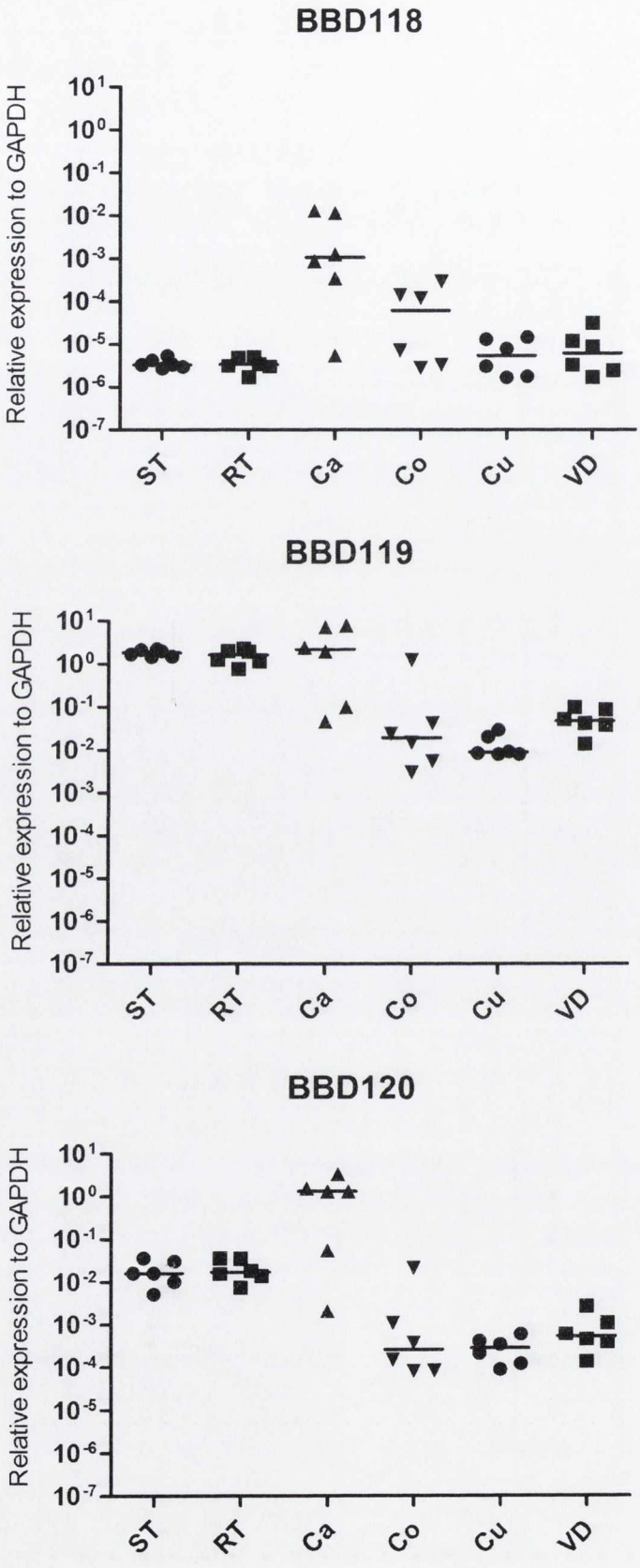
Gene expression analysis of the 19 β-defensins genes of the cluster in chromosome 13 plus LAP in the 6 regions dissected of the male reproductive showed that all genes analysed are expressed but in different patterns of expression (Fig. 3. 6).

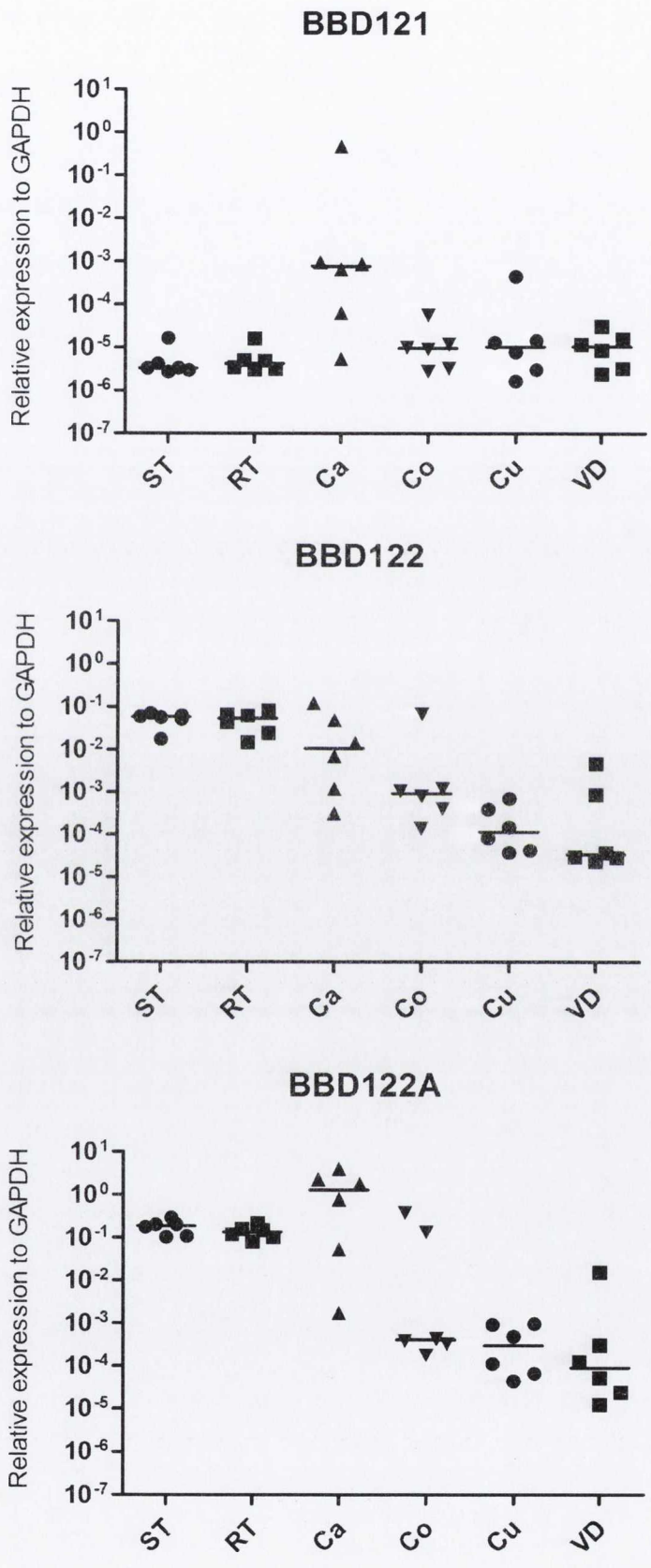



BBD128
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD127
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD126
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD


BBD125A
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD125
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD115
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD


BBD142
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD116
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD117
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD


BBD118
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD119
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD120
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD


BBD121
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD122
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD122A
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD

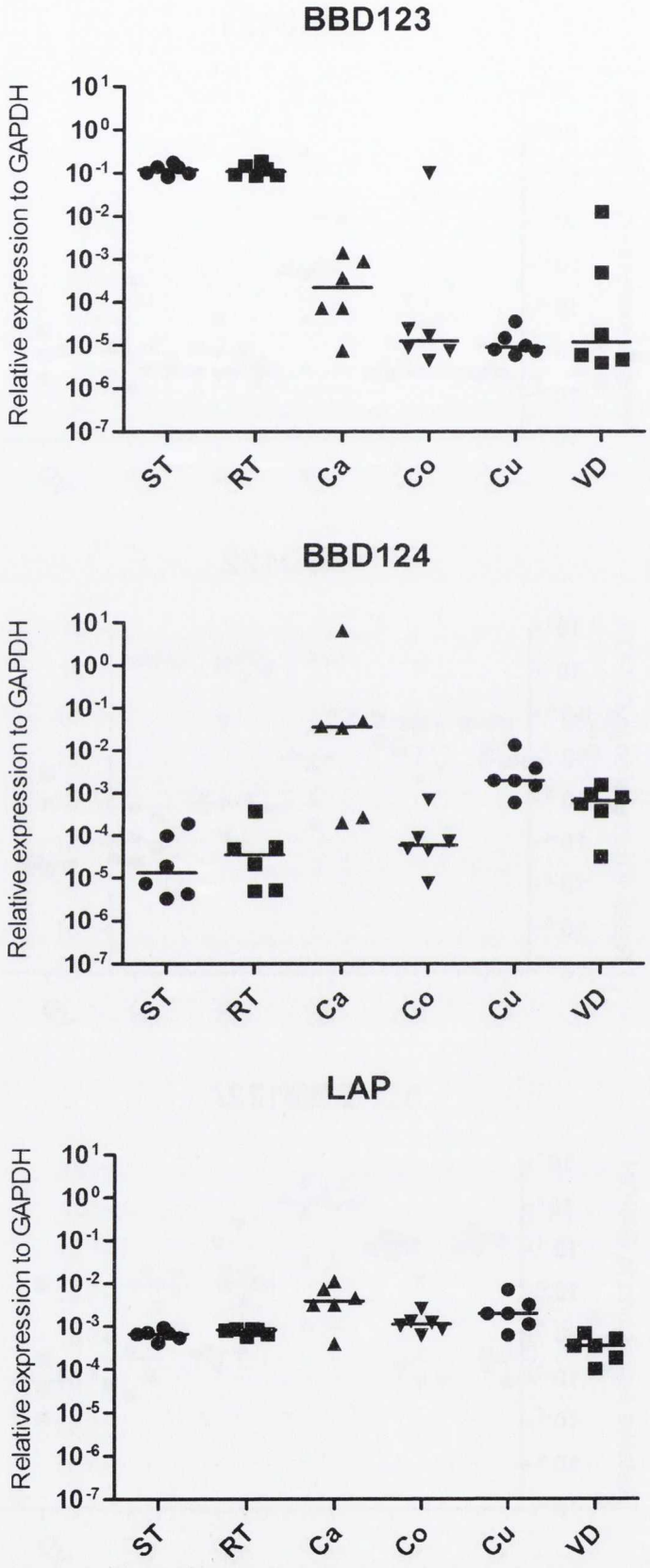

BBD123
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
BBD124
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD
LAP
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
ST
RT
Ca
Co
Cu
VD

Fig. 3. 6. Bovine Chr13 β-defensins gene expression across the male reproductive tract

Gene expression of bovine Chr13 β-defensins plus LAP was analysed in sections of the male reproductive tract of 6 Holstein-Friesian mature bulls. Expression values are normalised to GAPDH gene expression. ST, seminal tubule; RT, rete testis; Ca, caput; Co, corpus; Cu, cauda and VD, vas deferens. Horizontal bars represent median. Levels under 10^{-4} are considered under limit of detection.

All values are represented as gene expression relative to the normaliser gene *GAPDH*. Whereas some β-defensin genes are expressed homogeneously across the entire genital tract (*BBD119* and *132*), others show more site specificity in the testis or in the epididymis regions (*BBD122* and *123*). In general, the site with highest expression for almost all of these novel genes is in the caput of the epididymis. Closer examination of the data reveals predominant patterns of β-defensin gene expression; some of them with high expression in the epididymal regions and lower in the testis regions such as *BBD132, 129, 128, 127, 126, 125, 125a, 115, 116, 117, 118* and *124*; others with high expression in both testis sections and caput and lower in the epididymal regions, such us *BBD120, 122, 122a* and *123* and finally *BBD142, 119* and *121*, a homogeneous distribution in most tissues except for a peak of expression in the caput of the epididymis.

Variability in gene expression levels is evident among individual animals, although distinct patterns of preferential expression can be seen in a site specific manner. For example, with *BBD128*, clear preferential expression is seen in the caput of the epididymis, whereas for *BBD126*, preferential expression in the caudal region and vas deferens was observed. In most cases smaller variation among animals can be found as shown in *BBD119* and in other cases, little or no variation is found among animals as in the case of *LAP*. We found gene expression in the testis to have small variation among different animals compared with the regions in the epididymis.

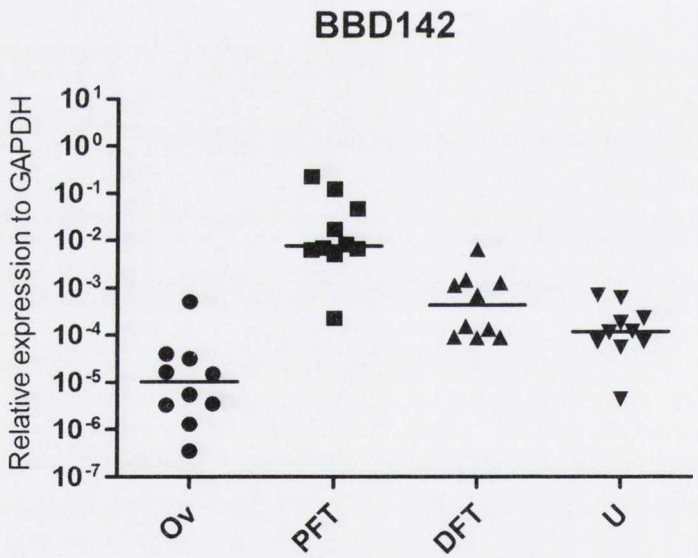
An extra set of tissues was retrieved during the castration of a 43 day-old animal as control for sexual development. We found no expression in any of the regions of the reproductive tract of *BBD132, 129, 128, 127, 126, 125, 125a* or *123*. Eleven of the β-defensin genes - *BBD115, 142, 126, 117, 118, 119, 120, 121, 122, 122a* and *124* were expressed in the 43 day-old bull (Table 3. 1.). In general, the gene expression of the defensins in the immature bull was lower compared with the expression in the adult males.

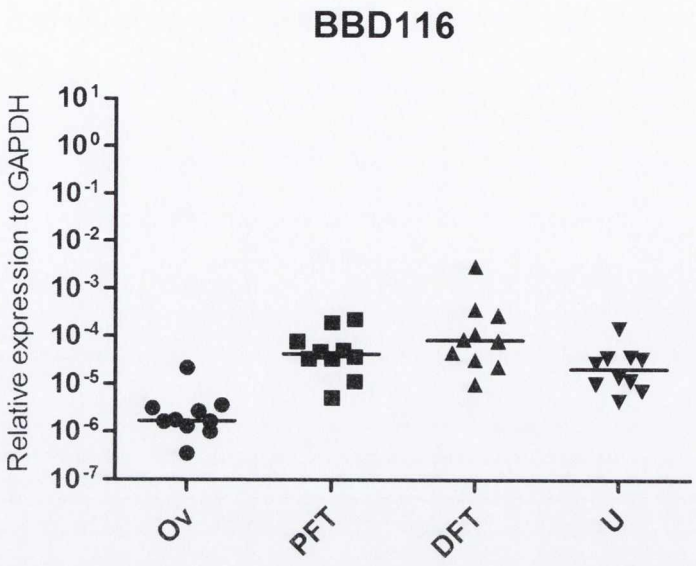
3.2.5. Regional specific β-defensin gene expression in female reproductive tract

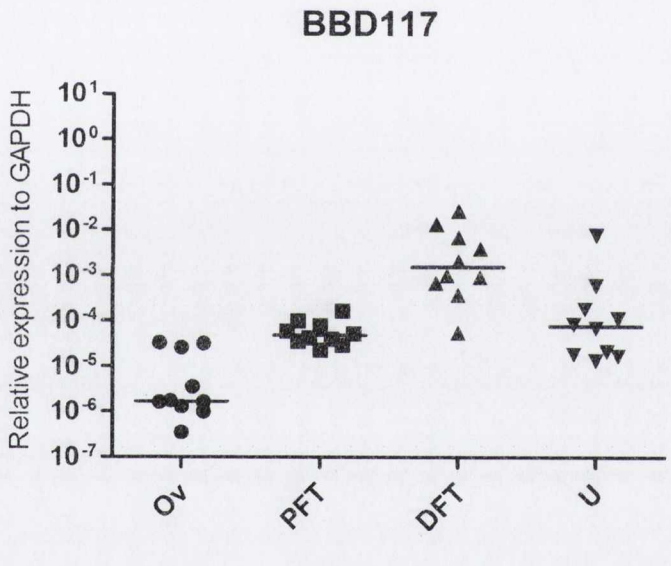
Gene expression analysis of the 19 novel β-defensins is carried out in samples of the female reproductive tract (Fig. 3. 7). In total, 9 of the genes of the cluster, *BBD116, 117, 119, 120, 122, 122a, 123, 124* and *142,* were expressed in the female tract. We found the distal Fallopian tube and uterus are the sites of predominant β-defensin expression in female reproductive tissues.

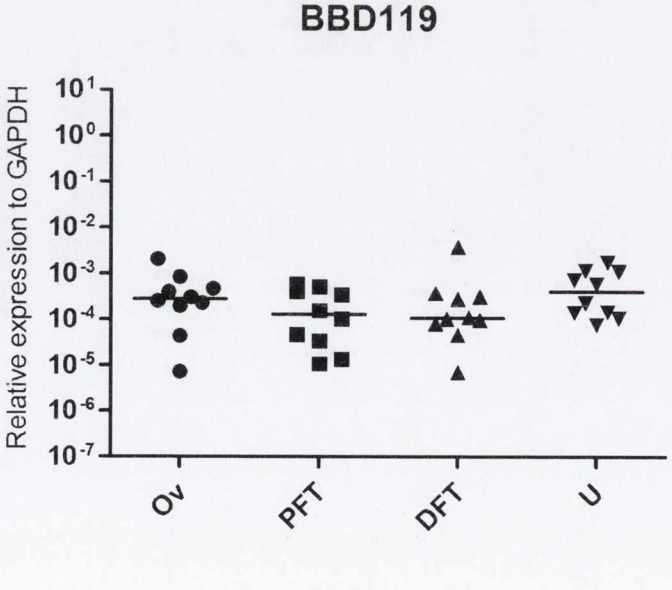
Individual animal variation is again evident as is site specific expression. *BBD117, 120, 122, 122a, 123, 124* and *LAP* are preferentially expressed in the distal Fallopian tube, whereas *BBD142* is expressed in the proximal Fallopian tube. *BBD119* showed little variation in expression levels between tissues.

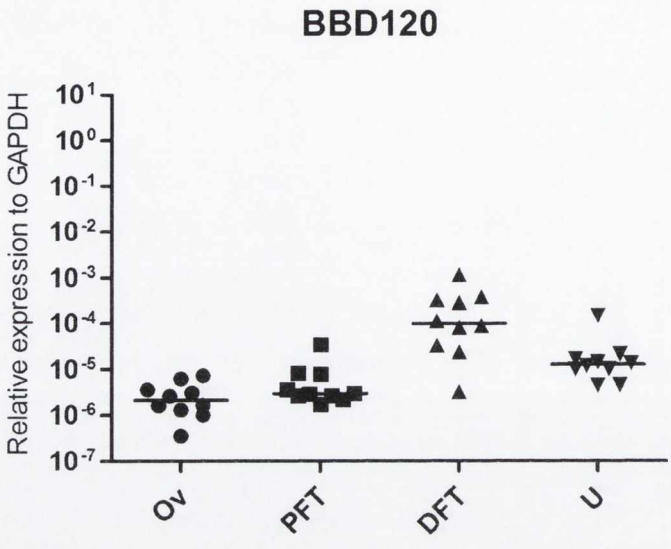
To measure the possible variation of gene expression of the hormone levels during the different phases of the ovulation, tissue samples from cows sampled with a predominant corpus luteum on either ovary (n=5) or cows with growing follicles (n=5) were collected. No significant differences in gene expression levels between the two phases were found.

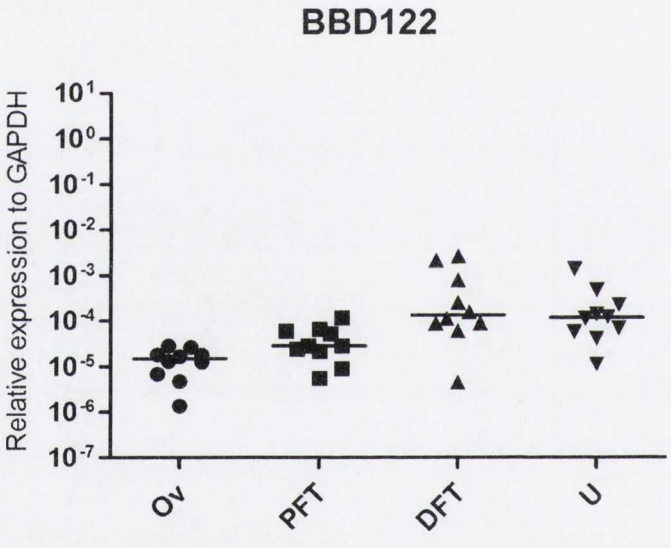


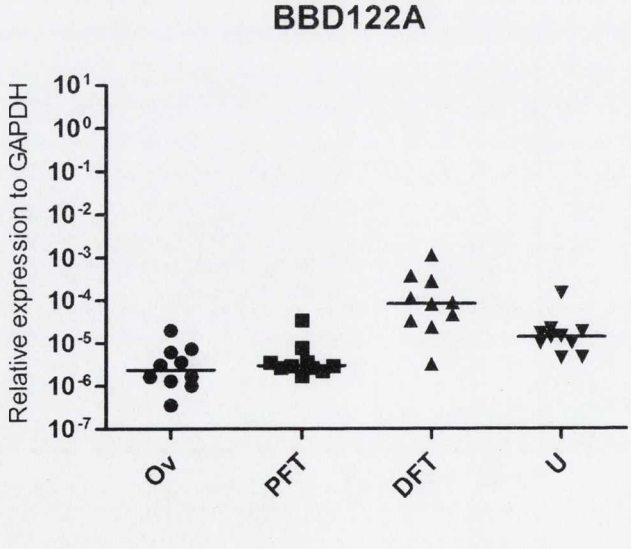

BBD116
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Ov
PFT
DFT
U


BBD117
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Ov
PFT
DFT
U


BBD119
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Ov
PFT
DFT
U


BBD120
Relative expression to GAPDH
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7
Ov
PFT
DFT
U


BBD122
Relative expression to GAPDH
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7
Ov
PFT
DFT
U


BBD122A
Relative expression to GAPDH
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7
Ov
PFT
DFT
U

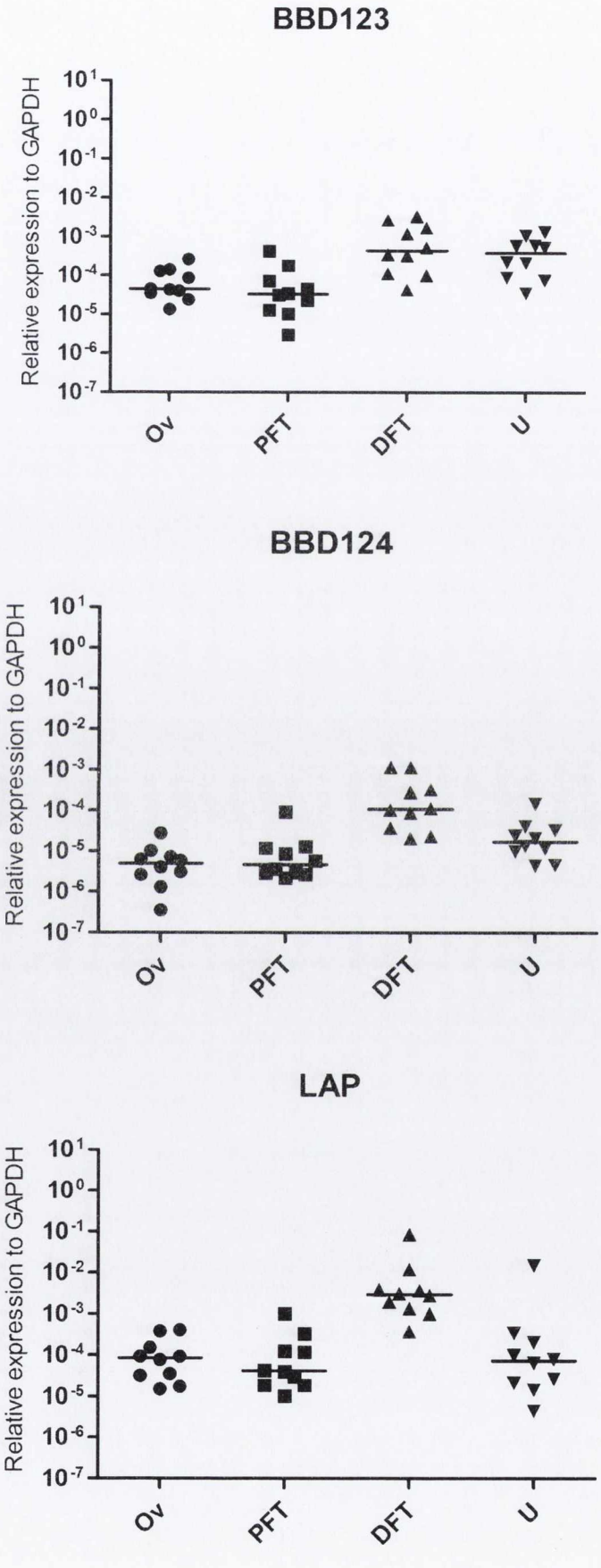

BBD123
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Ov
PFT
DFT
U
BBD124
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Ov
PFT
DFT
U
LAP
Relative expression to GAPDH
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Ov
PFT
DFT
U

Fig. 3. 7. Bovine Chr13 β-defensins gene expression across the female reproductive tract

Gene expression of bovine Chr13 β-defensins plus LAP was analysed in 4 regions of the reproductive tract of 10 cows. Expression values are normalised to GAPDH gene expression. Ov, ovary; PFT, proximal Fallopian tube; DFT, distal Fallopian tube and U, uterus. Horizontal bars represent the median. Levels under 10^{-4} are consider zero or under limit of detection.

3.2.6. Overview of novel β-defensins in the reproductive tract

All expression data is summarised in Table 3.1. In addition, a Kruskal-Wallis non-parametric analysis of variance was used to test the variance equality of each groups' medians of the data in the adult male and female samples. In males, all genes but *BBD125* were found to be significantly different. In the female samples, all genes expressed are found to be heterogeneous.

Closer examination of the data reveals a predominant pattern of β-defensin gene expression (Table 3. 1), which we call Class A, with low and preferential expression in the epididymis and vas deferens and with no expression in the immature bull or the female genital tract (*BBD132, 129, 128, 127, 126, 125* and *125a*).

Table 3. 1. Summary representation of patterns of expression for Chr13 β-defensins in reproductive tract.

Different regions of the reproductive tract of 6 mature (18 months old) bulls, a 43 days old immature bull and 10 cows were processed and used for gene expression analysis of Chr13 β-defensins. Each gene was compared with the control gene (GAPDH) across reproductive tracts. -, indicates no detectable gene expression; +, gene expression 10^{-4} fold lower than GAPDH (levels under 10^{-4} are consider zero or under limit of detection); ++, gene expression between 10^{-4} to 10^{-2} fold change lower than GAPDH; +++, gene expression between 10^{-2} to 1 fold change of GAPDH; ++++, gene expression higher than GAPDH. ST, seminal tubule; RT, rete testis; Ca, caput; Co, corpus; Cu, cauda; VD, vas deferens; T, testis; Ov, ovary; PFT, proximal Fallopian tube; DFT, distal Fallopian tube and U, uterus. Kruskal-Wallis non-parametric analysis of variance was used to determine heterogeneity of data. Several patterns of expression were observed: Class A genes, BBD132, BBD129, BBD128, BBD127, BBD126, BBD125, BBD125a, marked with a grey background, are found highly expressed in the epididymis of mature bulls, with low or no expression in testis, and no expression in immature bull or any of the female regions. A second group formed by BBD115, 116, 117, 118 and 124 was found highly expressed in the epididymis but lowly expressed in the testis but had expression in the female tract. A third group, BBD120, 122, 122a and 123, showed high levels of expression in testis, but low in the epididymis. Finally, BBD142, 119 and 121 presented a homogeneous expression pattern along the male reproductive tract.

	Adult bulls							Immature bull					Adult cows			
	ST	RT	Ca	Co	Cu	VD	P<0.05	T	Ca	Co	Cu	Ov	PFT	DFT	U	P<0.05
BBD132	+	+	+	+	+	+	Yes	-	-	-	-	-	-	-	-	N/A
BBD129	+	+	++	+++	++	++	Yes	-	-	-	-	-	-	-	-	N/A
BBD128	+	+	++	+	+	+	Yes	-	-	-	-	-	-	-	-	N/A
BBD127	+	+	++	+	+	+	Yes	-	-	-	-	-	-	-	-	N/A
BBD126	+	+	+	++	++++	+++	Yes	-	-	-	-	-	-	-	-	N/A
BBD125	+	+	+	++	+	+	No	-	-	-	-	-	-	-	-	N/A
BBD125a	+	+	+	++	+	+	Yes	-	-	-	-	-	-	-	-	N/A
BBD115	+	+	++	+++	+++	++	Yes	+	+	+	++	-	-	-	-	N/A
BBD142	+++	+++	++++	+	+++	+++	Yes	+++	++	++	+++	+	+++	++	++	Yes
BBD116	+	++	++++	++++	+++	+++	Yes	+	++	+	+++	+	+	++	+	Yes
BBD117	+	+	++++	+++	+++	+++	Yes	+	++	++	++	+	+	++	+	Yes
BBD118	+	+	+++	++	+	+	Yes	+	++	++	++	-	-	-	-	N/A
BBD119	++++	++++	++++	+++	+++	+++	Yes	+++	+++	++	++	++	++	++	++	No
BBD120	+++	+++	++++	++	++	++	Yes	++	+++	++	++	+	+	++	++	Yes
BBD121	+	+	++	+	+	+	Yes	+	+	++	++	-	-	-	-	N/A
BBD122	+++	+++	+++	+	+	+	Yes	++	++	+	+	+	+	++	++	Yes
BBD122a	+++	+++	++++	+	+	+	Yes	++	++	+	+	+	+	++	+	Yes
BBD123	+++	+++	++	+	+	+	Yes	-	-	-	-	+	+	++	++	Yes
BBD124	+	+	+++	+	++	++	Yes	+	++	++	++	+	+	++	+	Yes
LAP	++	++	++	++	++	++	Yes	++	+++	+++	+++	+	+	+++	++	Yes

3.2.7. ELONGATED TAIL OF B-DEFENSINS AS TARGET FOR GLYCOSYLATION

The NetOGlyc and NetNglyc server tools were used to determine the presence of glycosylation sites in the novel β-defensins peptides (Table 3. 2). Only four peptides were found to have N-glycosylation sites. In contrast all but five were found to have O-glycosylation sites. The second discovery was that the predicted sites of all but two N-glycosylation sites in BBD126 and BBD118 were found in the tail region of the defensins.

Table 3. 2. Length of peptide tail of the novel defensins and glycosylation sites predictions.

β-Defensin tail is the peptide sequence between the last cysteines of the defensin motif and the stop codon. Bovine Chr13 β-defensins are longer than other β-defensins clusters. The longer tail, allows for an increment in the number of possible glycosylation sites, particularly O-linked glycosylation, we found Class A genes to have an average of 4 sites; the rest of member of the cluster have an average of 1 site per gene. N-linked and O-linked glycosylation sites were predicted by NetOGlyc and NetNglyc servers. Class A genes in grey background.

	Length of Tail	O-Gly			N-Gly		
		No Tail	Tail	Total	No Tail	Tail	Total
BBD132	32	0	2	2	0	0	0
BBD129	116	0	8	8	0	3	3
BBD128	29	0	0	0	0	0	0
BBD127	32	0	1	1	0	0	0
BBD126	30	0	8	8	1	0	1
BBD125	47	0	9	9	0	0	0
BBD125a	67	0	3	3	0	0	0
BBD115	17	0	1	1	0	0	0
BBD142	30	0	0	0	0	0	0
BBD116	34	0	3	3	0	2	2
BBD117	41	0	2	2	0	0	0
BBD118	40	0	4	4	1	0	1
BBD119	28	0	0	0	0	0	0
BBD120	22	0	2	2	0	0	0
BBD121	10	0	0	0	0	0	0
BBD122	16	0	1	1	0	0	0
BBD122a	16	0	1	1	0	0	0
BBD123	14	0	0	0	0	0	0
BBD124	13	0	0	0	0	0	0
LAP	3	0	0	0	0	0	0

In order to assess the potential for glycosylation of β-defensin peptides, the length of the tail of each peptide (amino acid sequence between the last cysteine of the defensin motif and the stop codon of the protein) was compared between Class A genes expression patterns and the rest of Chr13 genes, another comparison was between the 19 genes of Chr13 cluster and the remaining defensins in the bovine genome (n=38). Class A genes were found to be significantly longer than the rest of defensins in the cluster ($P<0.05$). Similarly, Cluster 13 β-defensins were found to be significant longer ($P<0.01$) than the rest of defensins in the bovine genome (Fig. 3.8).

Anti-BBD126 antibody was used to detect the presence of BBD126 in sperm cell lysate. Samples treated with a protein de-glycosylation mix (New England Biolabs, Cambridge, MA, USA) showed a decrease in the molecular weight of the protein detected for anti-BBD126, from the 25 kDa of the untreated sample to 8 kDa, similar to the predicted size of a non-glycosylated BBD126 protein (Fig. 3. 9). The protein de-glycosylation mix contains PNGase F, O-Glycosidase, Neuraminidase, β1-4 Galactosidase and β-N-acetylglucosaminidase. These enzyme will remove all carbohydrates and will not determine the nature of the glycosylation.

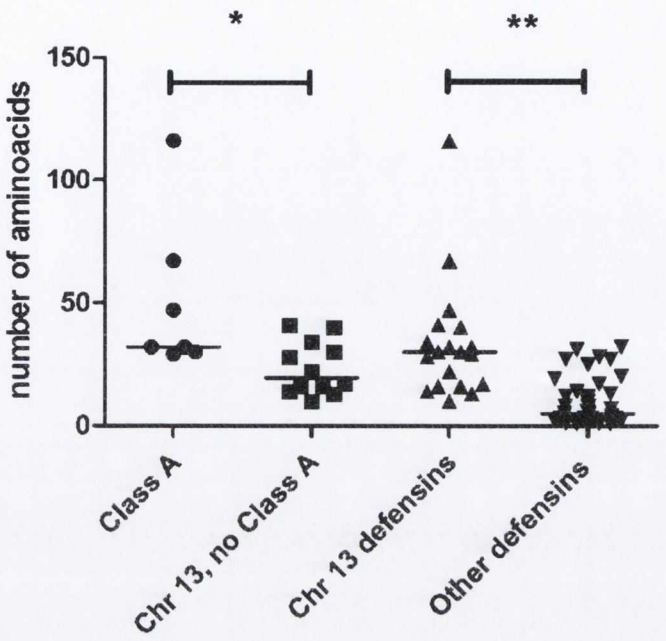


Fig. 3. 8. Analysis of the β-defensins tail length

*Tail length, measured from the last cysteines of the defensin motif to the stop codon, for Class A Chr13 genes. Other members of the Chr13 cluster and β-defensin genes of other clusters were analysed. Mann-Whitney U test was used to determine the significance of the difference between groups. Class A genes are significantly longer than other defensins in Chr13 cluster or any other cluster. P value denotes significant difference *<0.05, **<0.01.*

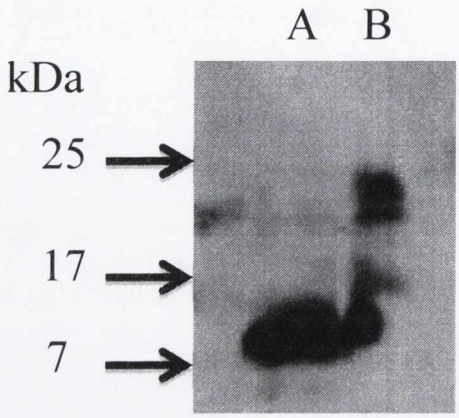


Fig. 3. 9. De-glycosylation effect on BBD126 molecular weight.

Anti-BBD126 antibody was used to measure the effect of glycosylation removal enzymes on the molecular weight of BBD126. A, Sperm cell lysate was treated with protein de-glycosylation mix (New England Biolabs, Cambridge, MA, U..); B, non-treated sperm cell lysate. Samples were run in a 4-12% SDS-PAGE and transferred into PMSF membrane and incubated with anti-BBD126 mouse monoclonal antibody. BBD126 weight is reduced from the 25 kDa glycosylated into the 8 kDa of the predicted protein without glycosylation. This result was replicated in a second reaction.

3.2.8. Principal components analysis of β-defensins promoter region

The online tool, MatInspector, was used to predict transcription factor binding sites in the upstream region of chromosome 13 cluster β-defensin genes. For comparative purposes, four defensins from different clusters (*BBD105, BBD106, BBD107* and *LAP*) and eight specific epididymal expressed genes *(SPAM1, SPAG11E, ArA, Esp, Ace, sPLA2III, Sp17* and *Catsper*) were included in the analysis. The matrix was populated with the number of times a transcription factor binding site was repeated in each promoter region and analysed using SPSS principal components analysis, as described in Materials and Methods section. The total variance explained by the different detected components (Appendix: Table 4) showed the presence of 32 different components, each of which explains a fraction of the variance found in the data. To decide which of the factors are statistically important a scree plot was done. The score (Eigen value) for each component is graphed against itself and the cutting

point marked by the point of inflexion of the curb graphed (Appendix: Fig. 1). This method identified the first four components as the most explicative of the data set variance. Most of variance will be explained by the first component, with a reduction of the amount of explained variance with each new component. Each gene promoter region is given a value for each component; the higher the value, the more relevant is that component to explain the variance seen in that promoter region. To visualise these data, plots of values for components 1 and 2 and components 3 and 4 were used (Fig. 3. 10).

Component 1 explains 36% of the variance present in the data set; component 2 just 13%. Both present similar distributions with promoter regions with low component 1 values presenting low component 2 values; for example, almost all epididymis specific genes, *ArA, Esp, Ace, sPLA2III* and *Sp17*, with the exception of *SPAM1*and *SPAG11E,* have low scores for both component 1 and 2, indicating the low importance of components 1 and 2 to explain the variance present on those promoter regions. On the other hand, component 3, 7.5% of the variance, and component 4, 4% of the variance, presented different distribution. High values for component 3 are found in epididymal specific genes, *ArA, Esp, Ace, sPLA2III* and *Sp17*, indicating the importance of component 3 to explain these genes promoter region variance. The highest values of component 4 are present in Class A genes, *BBD127, BBD128, BBD125, BBD125A, BBD132* and *BBD129,* and Chr13 cluster β-defensins, *BBD115* and *BBD124.*

The composition of each component was determined by the Eigen values of each transcription binding sites found in each promoter region with a high value indicated the importance in creating each component. The plot of components 1 and 2 showed transcription factor binding sites HOMF, gene cascade such as TLR, embryogenesis, and ARID, involved in cell development, were important for component 1. GATA, large family with roles in embryogenesis, cancer, development, and FKMD, involved in apoptosis, as important for Component 2 (Fig. 3. 11).. Component 3 was defined by high values for SP1F, early development regulation; EGRF, growth factors and KLFS, cell proliferation, cell survival and differentiation. Component 4 had high values for SORY, SOX/SRY sex/testis determining factors and STAT, a large family with roles in signal transduction and cytokine production

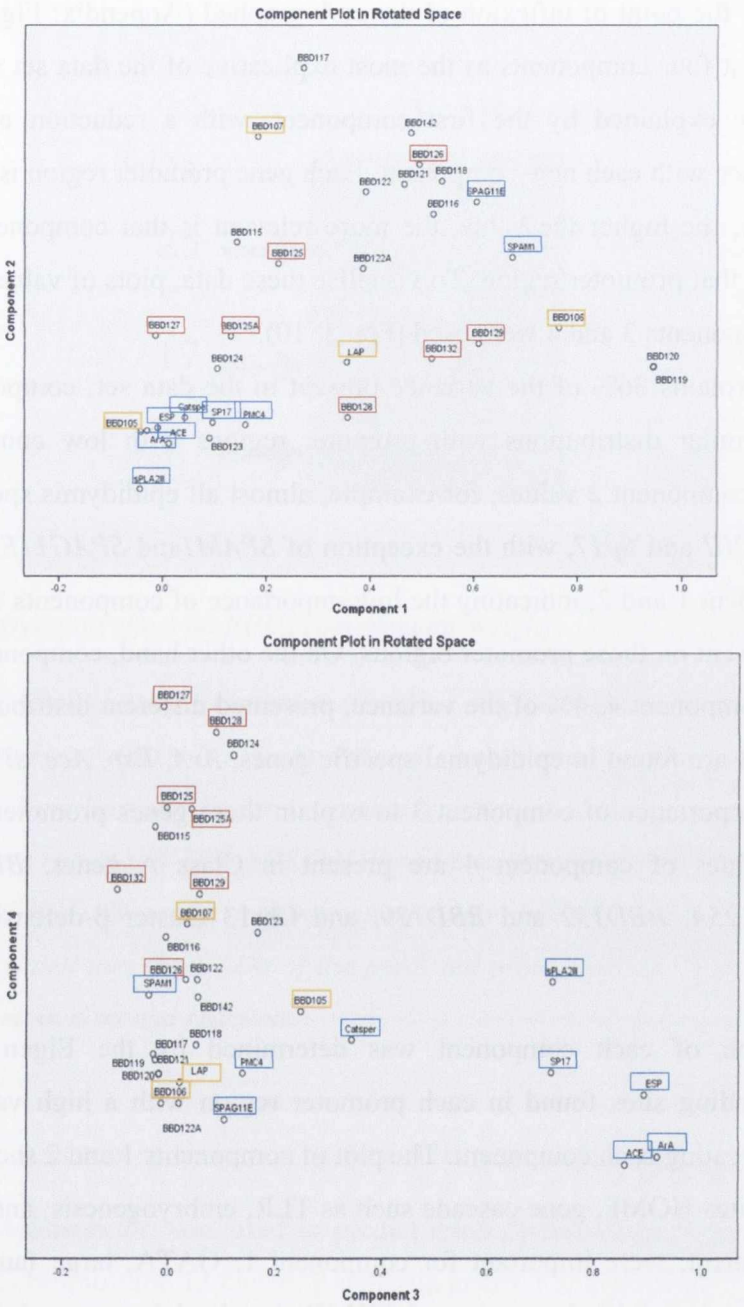


Fig. 3. 10. Scatter plot of component 1 vs component 2 and component 3 vs component 4 for each promoter region.

Promoter region of bovine Chr13 β-defesin cluster, bovine Ch 27 β-defesin cluster, BBD105, BBD106, BBD107 and LAP, and 8 specific epididymal expressed genes, SPAM1, SPAG11E, ArA, Esp, Ace, sPLA2III, Sp17 and Catspe. Eigen values for each of the 4 components were plotted. In red, Class A genes; in blue, epididymal specific genes and in orange, Chr 27 defensins. Apididymal genes had low values for components 1, 2 and 4 but high for

component 3. Class A genes had a broad distribution with components 1 and 2 but high values for component 4. The rest of defensins have broad distribution for components 1 and 2 and low values for components 3 and 4.

3.2.9. B-DEFENSIN GENE EXPRESSION RESPONSE TO LPS AND HORMONES.

Promoter region, 1000 bp upstream region of the genes, of *LAP* and *BBD126*, were cloned into a pGL4.23 vector. The plasmid incorporates the coding sequence for a firefly luciferase. Firefly luciferase enzyme catabolises a reaction generating light by consuming luciferin and ATP as substrates. The assay is based on the transfer of the promoter region of a gene into the vector. The expression of the firefly luciferase gene, incorporated in the vector, is then regulated for the promoter region cloned into the plasmid. The luminescent produced is proportional to the activator signal received by the promoter region used. Both pGL4.23-LAP and pGL4.23-126 constructs were transfected into a bovine endometrial cell line (BEND) as described in the Materials and Methods sections. Cells were then stimulated with LPS, dihidrotestosterone (DHT), seminal plasma or oestrogen. Luminescent units were normalised to the renilla control signal and shown as relative change to the untreated transformed cells.

Luminescent for LAP was up-regulated in the presence of all treatments. In particular with highest LPS concentration, 15 ng/ml of LPS, luminescent was dose-dependent (Fig. 3. 12). Luminescent for BBD126 construct was significantly up-regulated by DHT1, 0.1 uM DHT, and EST1, 0.1 uM of oestrogen (Fig. 3. 13). Higher DHT and oestrogen concentration did not show an increase in luminescent.

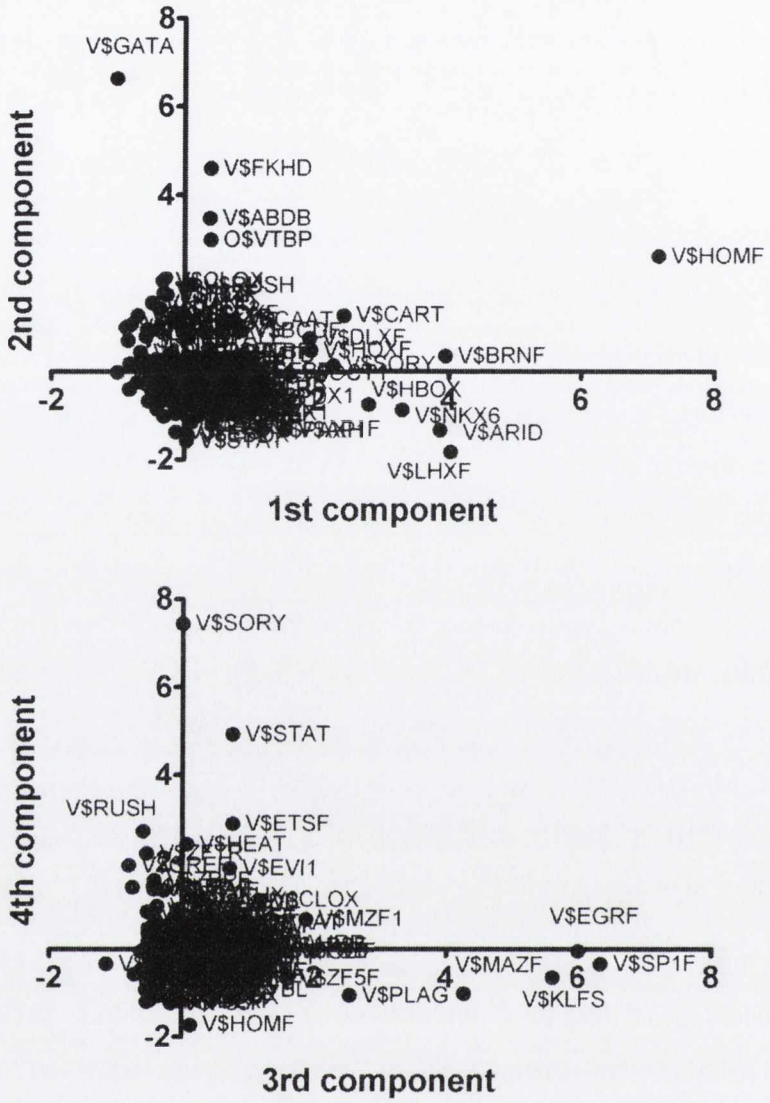


Fig. 3. 11. Scatter plot of component 1 vs component 2 and component 3 vs component 4 for each TFBS.

Eigen values of each TFBS for each component. Higher values indicate the importance of the TFBS in the creation of the component. GATA, common large family of TFBS involved in embryogenesis, cancer and other roles, and FKMD, involved in apoptosis, are the highest values for component 1. HOMF, another large family of TFBS involved in gene cascades such as TLR and embryogenesis, and ARID, important in cell development are the highest values for component 2. SP1F, a large family involved in multiple roles such as early development, EGRF, growth factor and KLFS, cell proliferation, cell survival and differentiation are the highest values for component 3. SORY, SOX/SRY sex/testis determing genes, and STAT, involved in cytokine signalling, are the highest values for component 4.

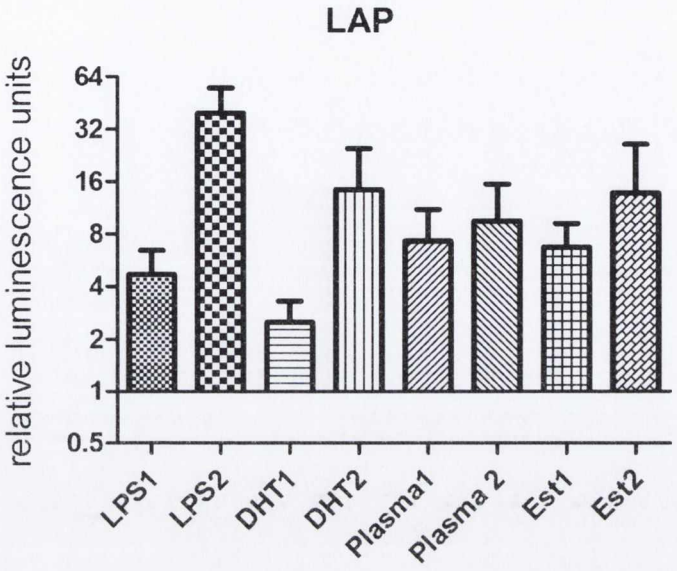


Fig. 3. 12. Relative luminescent units of pGL4.23-LAP plasmid

pGL4.23-LAP plasmid was transfected into BEND cells.Firefly luciferase production was stimulated by LPS1, 1.5 ng/ml; LPS2, 15 ng/ml,; DHT1, 0.1 uM; DHT2, 10 uM; Plasma1, 1:10,000 dilution of seminal plasma;Plasma2, 1:100 dilution of seminal plasma; Est1, 0.1 uM of oestrogen and Est2, 10 uM of oestrogen. Luminescent is up-regulated for all the treatments in a dose-dependent manner. Graph represents mean of 3 biological replicates, error bar indicate the standard error.

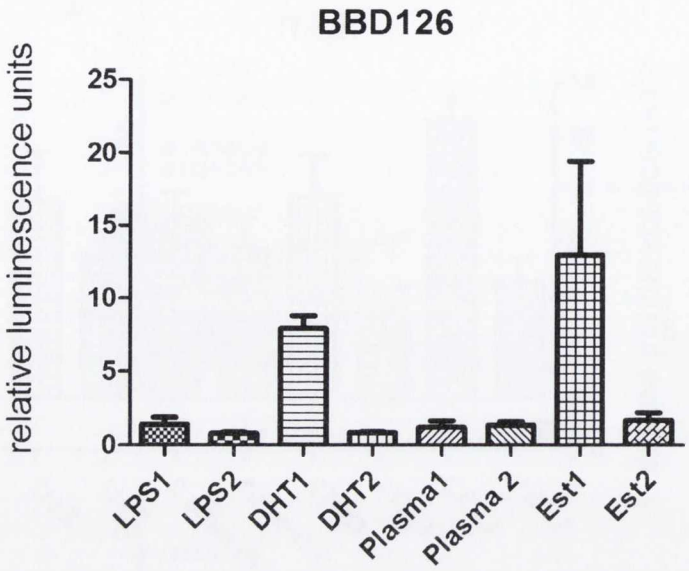


Fig. 3. 13. Relative luminescent units of pGL4.23-126 plasmid

pGL4.23-126 plasmid was transfected into BEND cells. Firefly luciferase production was stimulated by LPS1, 1.5 ng/ml; LPS2, 15 ng/ml; DHT1, 0.1 uM: DHT2, 10 uM: Plasma1, 1:10,000 dilution of seminal plasma; Plasma2, 1:100 dilution of seminal plasma; Est1, 0.1 uM of oestrogen and Est2, 10 uM of oestrogen. Luminescent is up-regulated by DHT1 and Est1. Graph represents mean of 3 biological replicates, error bar indicate the standard error.

3.3. Discussion.

In this study, we comprehensively profiled the expression of genes in this 19 gene cluster across the genital tracts in mature Holstein-Friesian bulls, a 43 day old male calf, a cohort of mature cows (in both luteal and follicular stages of ovarian growth) and a panel of tissues from different organs. No significant expression of members of the studied gene cluster was found in tissues outside the genital tract with the exception of *BBD117*. In contrast, clear and distinct patterns of β-defensin gene expression were uncovered in the reproductive tracts. Class A genes (*BBD132, 129, 128, 127, 126, 125* and *125a*) were expressed in the mature bull epididymis and vas deferens but were completely absent from the immature male sampled and also the female samples. This suggests that they may be developmentally

regulated or there may be a potential role in sperm maturation and/or attachment for these genes. Expression analysis of β-defensins in the genital tract of other animal models showed quantitative variation in gene expression in different sections of the reproductive tract, suggesting that site-specificity of gene expression may reflect differences in the biological role of these defensins (Com, Bourgeon et al. 2003, Zhou, Zhang et al. 2004, Jelinsky, Turner et al. 2007).

The Class A expression pattern has two interesting implications - the first is that the co-regulated genes may be under the influence of similar promoter elements and/or exposed to RNA polymerase during the same chromatin-unpacking event. Secondly, the activation of these promoter elements may be under hormonal control and are therefore developmentally regulated (Radhakrishnan, Hamil et al. 2005, Yudin, Tollner et al. 2008), suggesting that these genes may have a role in the maturation of sperm. Preferential expression of these genes was detected in the epididymis – the area where sperm are stored until maturation before ejaculation as previously described in other species (Tollner, Yudin et al. 2008). Non-class A genes exhibited different patterns of expression along the mature male genital tract, also suggesting different roles for each group. *BBD115, 116, 117, 118* and *124* showed predominant epididymal expression with low expression in the testis. *BBD120, 122, 122a* and *123* showed predominant expression in the testes and caput of the epididymis but were almost absent in the distal epididymal regions. *BBD142, 119* and *121* presented a homogeneous expression pattern along the male reproductive tract.

Based on their testicular or epididymal location, and in light of associated studies (Tollner, Yudin et al. 2008), defensins could mediate the binding of the sperm to the female epithelia. On the other hand, β-defensins could be important for sperm maturation. It is known that this process occurs by the attachment of a multitude of proteins to the sperm cell surface during its transit through the epididymis. Alternatively, the regional distribution of the defensins could be explained by each group targeting specific pathogens only present in each region or to prevent infection in the female reproductive tract after mating (Sorensen, Gram et al. 2003) but the lack of inflammation and/or presence of pathogens suggest a role outside of solely defence.

The Class A genes possess additional characteristics other than expression pattern which differentiates them from the other genes in the cluster such as significantly extended C-termini, Class A genes have a tail (measured from the last cysteine to the predicted stop codon) of 50 amino acids on average, which is significantly different in length ($P<0.05$) from the 24 amino acids found on average in the other genes in the cluster. Potential glycosylation

sites were predicted by an online tool, showing the presence of multiple potential glycosylation targets in the tail of the β-defensins. In (Yudin, Tollner et al. 2008), mouse *β-defensin 122*, orthologous to bovine *BBD126*, is characterised as one of the most abundant proteins of the mouse sperm glycocalyx. Recently, new evidence for the role of *BBD126* in reproduction has been found; a mutation in human β-defensin 126 was associated with a reduction in sperm penetration of vaginal mucus and reduced fertility (Tollner, Venners et al. 2011). Other publications suggested the importance of glycosylation of these defensins for their activity in the mammalian reproductive tract (Tollner, Yudin et al. 2008). Interestingly, most of the predicted sites are found in the tail outside the defensin motif; this elongated tail is a characteristic of Chr13 defensins cluster. Some of these defensins are potentially highly glycosylated. This was confirmed by BBD126. The molecular weight of the peptide is more than triple by the addition of carbohydrates, as the de-glycosylation treatment showed. It is important to mention that due the nature of the de-glycosylation enzymatic cocktail, containing PNGase F, O-Glycosidase, Neuraminidase, β1-4 Galactosidase and β-N-acetylglucosaminidase; it was not possible to determine if the glycosylation was N or O in origin. It is worth noticing that O-glycosylation is typically associated with protein secreted in mucous membranes. A selective glycosylation digestion should be able to determine the nature of the carbohydrates in BBD126 surface. In addition, orthologous of *BBD126*, a member of our Class A, have been found to have a reproductive role in coating macaque sperm which may mean that expression occurs after puberty (Radhakrishnan, Hamil et al. 2005, Yudin, Generao et al. 2005, Tollner, Yudin et al. 2008). Finally the Class A genes, which in bovine are found in a contiguous group inside the gene cluster, form a separate cluster in primate species like human, chimpanzee and macaque (Rodriguez-Jimenez, Krause et al. 2003, Patil, Cai et al. 2005, Radhakrishnan, Hamil et al. 2005). The retention of the grouping despite genomic translocation may indicate group regulation and/or synergistic activity.

The different expression patterns found for the chromosome 13 β-defensins cluster suggest a tight regulation of different nature compared to the classic antimicrobial β-defensins. The role of sex hormones, like testosterone, in regulating β-defensins has been previously described in castrated mice (Yenugu, Chintalgattu et al. 2006). In a similar fashion, we had found gene expression of *BBD125, 125a, 126, 127, 128, 129* and *132* missing in a pre-pubertal bull, suggesting the lack of the stimuli required for their expression.

Our approach to the question of how these genes are regulated was the analysis of the promoter region of these genes and comparison with other immune and fertility related genes.

We used the transcription factor binding site (TFBS) prediction tool Matinspector. With all prediction tools there is an intrinsic risk of false positives; to compensate for this issue, we raised the minimum score from the default 0.80 to 0.90. We decided to analyse not only presence or absence of the transcription factor binding site but the number of repetitions found in each sequence also , this having been shown to have an important role in the likelihood of the prediction (Holloway, Kon et al. 2005). Using factor analysis we are able to reveal factors or components, which are defined as clusters of correlated variables that explain the variability of the sample. Each component is able to explain part of the total variability and could be explained as different patterns found in all variables (transcription factor binding sites). All genes are scored for all four components with a high score indicating a closer similarity to the statistical pattern. These results showed dramatic differences between non-defensin epididymal genes and the rest of genes analysed. These genes clustered with low Eigen values for component 1, 2 and 4 but high values for component 3. Defensins seem to distribute broadly for components 1 and 2 but interestingly Class A present high values for component 4. These results seem to indicate the presence of differences between the promoter region of most β-defensins and Class A β-defensins. Most interestingly, the highest score TFBS responsible for component 4 is SORY, suggesting sexual maturation and development as important for the gene regulation. We hypothesise that hormones and not immune signals could regulate Class A genes

To test our hypothesis, the promoter of *LAP*, an antimicrobial defensin, and *BBD126*, were cloned into luciferase reporter gene vectors. The plasmids, pGL4.23-LAP and pGL4.23-126 were used to transfect BEND cells. These cells were then treated with LPS, DHT, seminal plasma and oestrogen. LPS is a known activator of the TLR pathway by TLR4 activation. This is the classical activation present in β-defensins as effector molecules of the immune system. As sexual hormone DHT, active forms of testosterone and oestrogen were used. In addition, seminal plasma as pool of different sexual and growth hormones and factors were used to stimulate gene expression. The results showed that LAP was stimulated in a dose-dependent manner by all four treatments. LPS activation of LAP was expected and previously shown (Diamond, Russell et al. 1996). The activation of LAP by DHT, oestrogen and seminal plasma were unexpected. An explanation could be the presence of small amounts of LPS in these treatments and/or secondary activation of LAP by internal signals activated by hormones. BBD126, on the other hand, was not activated by LPS stimulation; only DHT and oestrogen had an effect in BBD126 gene expression. We observed as well that the high

concentrations of DHT and oestrogen had no effect on BBD126 gene expression. We hypothesised that the concentrations used, 10 uM, are above the physiological concentrations in epididymis (Diba, Vinkler et al. 1994) and can induce danger signals switching off some genes like BBD126 but up-regulating others like LAP.

CHAPTER 4: GENETIC ANALYSIS OF A BOVINE B-DEFENSIN CLUSTER

4.1. INTRODUCTION

Genetic variation studies have identified the function of unknown genes and their role in different diseases. These studies have proved essential for the identification of novel single nucleotide polymorphisms (SNP) that account for up to most of genetic variation between individuals (Collins, Brooks et al. 1998) and that have been found to be the origin of many diseases and syndromes. SNPs can be found in the non-coding regions of the DNA or in the coding region. These coding region SNPs can be classified as either synonymous, in which the nucleotide change does not change the amino acid sequence of the cognate protein because of the degenerate nature of the genetic code or non-synonymous, in which the nucleotide change does change the protein sequence. These changes in the protein sequence can have at least two consequences. In most cases, if it is a similar amino acid or if the change occurs in a non-essential region or motif of the protein, no phenotypic effect is observed. Rarely a detectable change in the protein function is produced, whichin most cases results in a deleterious effect.

A single mutation can be the cause of a disease (Folli, Viglione et al. 2005) but this is uncommon. In most cases, several SNPs in different coding or non-coding regions are implicated and found to be correlated with the presence of a disease. SNP frequencies are determined by partly by their effect, positive or negative, on the individual; but are also subject to random genetic drift particularly in small populations including those where, because of a gross discrepancy in the sex-ratio of breeding animals the *effective* population size is diminished. If a mutation is deleterious the SNP will be found at low frequencies. Selection will eliminate individuals with disadvantages from the population, if a SNP provides some form of evolutionary advantage; this could increase its frequency over several generations and eventually lead to the fixation of the SNPs in the population. Domesticated animals present some interesting paradoxes; deleterious SNPs for animal survival in the wild may be fixed in the population if they are linked to or increase the positive traits used for the selection of the animals.

It is not uncommon to observe groups of adjacent SNPs being inherited together; this effect being referred to as a haplotype. Haplotypes are defined as fragments of DNA that lack recombination events, allowing several alleles, the length of the fragment depending on the local rate of recombination, to be transmitted together. Recombination events between homologous chromosomes occur at low frequencies during meiosis, as part of gamete

formation. Outside recombination hotspots, these events can be considered random. The chances of a recombination event between two points in the genome is determined by the distance between them. Another parameter is time; haplotypes are temporal constructs that will tend to disappear in a sufficient amount of time as recombination events break up the haplotype. Haplotypes, even as artificial constructs, are useful tools and provide us with valuable information. haplotype, It is possible to infer the composition of the rest of SNPs in the haplotype by genotyping a single SNP member. In genome-wide association studies, numerous SNPs and haplotypes have been correlated with a disease or an effect (Ge, Fellay et al. 2009). However, on many occasions, the correlated SNP is found in a non-coding region or in the coding region of a gene unrelated to the effect observed. An explanation is that the SNP found forms part of a haplotype with an unknown SNP responsible for the effect.

Traditional breeding programmes concentrate their efforts on a small number of traits, mostly milk production in dairy breeds and growth rate or feed efficiency in beef cattle breeds. This approach has increased the desirable traits but on many occasions have had negative effects in other non-selected traits, for example the decrease in fertility suffered by a high yield dairy breed like Holstein-Friesian, (Mee 2004). In contrast, other breeding programmes have included a greater number of traits, including reproduction, immunity and others in addition to the classic production traits, to develop more balanced breeds. An example is the Norwegian-Red breed which does not show the decrease in fertility present in HF. This situation gives us an opportunity to analyse the genetic variation of the chromosome 13 β-defensin cluster in two distinctive breeds with different fertility potential. The role of these defensins has been previously discussed and described in multiple species (Tollner, Yudin et al. 2008, Tollner, Venners et al. 2011).

Our aim is to analyse the genetic variability present in the chromosome 13 β-defensin cluster and compare this with two commercial breeds, the high production and low fertility HF and the lower production but higher fertility NR, and determine the effect of the different breeding programmes in the genetic variability of the cluster.

4.2. RESULTS

4.2.1. SNP DISCOVERY

Primers designed to amplify the full coding sequence of the Chr13 β-defensin cluster genes were used to amplify the regions from genomic DNA of five Norwegian Red and five Holstein-Friesian DNA samples. Separate primer sets were designed for each of the two exons present in β-defensin genes, because of the length of the intron; the exception was BBD117. No first exon, including the signal peptide, has been found for this gene in the bovine genome. Each animal's DNA was amplified independently for each primer set. PCR products were then cleaned up and sent for commercial DNA sequencing. The sequences corresponding to each exon were aligned to identify possible SNPs and sequencing traces were carefully scrutinised in order to confirm the presence of SNPs. Sequence analysis was carried out using BioEdit software (Ibis Biosciences, Carlsbad, CA, USA). A total of 17 novel polymorphisms across the 19 gene cluster were found (Table 4. 1) Seven SNPs were located in the coding regions of the β-defensin genes introducing non-synonymous changes of amino acid; six located in the coding region are synonymous changes and four appeared in non-coding regions.

Table 4. 1. Genotyping results for 5 HF and 5 NR samples.

Primers designed to amplify each of the two exons present in β-defensin genes were used to amplify these regions by PCR in five HF and five NR cattle. PCR products were commercially sequenced and analysed by BioEdit software. A total of 17 SNPs were found in different genes of which seven were non-synonymous, amino acid changing mutations. Six were synonymous and four were found in non-coding region adjacent to either exon. Loci are named according to their location in the first, E1, or second exon, E2.

Locus	HF					NR						
115 E2	GG	GA	GG	GA	GG	AA	GG	GA	GA	GG	G=Ala	A=Thr
116 E2	AA	AA	GA	AA	AA	AA	GA	GG	AA	GA	A=Leu	G=Leu
118 E2 (a)	AA	AA	GA	AA	AA	AA	AA	AA	AA	AA	A=Gly	G=Gly
118 E2 (b)	GG	GG	GA	GG	GG	GG	GG	GG	GG	GG	G=Gly	A=Arg
117 (a)	TT	TT	TA	TT	TT	TT	TT	TA	AA	TA	T=Stop codon	A=Lys
117 (b)	GG	GA	GG	GG	GA	GG	GA	GG	GG	GG	G=Val	A=Ile
121 E1	CG	CG	CG	CC	GG	GG	GG	CG	GG	CG	Non-coding	Intron
122 E2	CT	CT	CT	CC	TT	TT	TT	CT	TT	CT	T=Asn	C=Asn
122a E1	TT	TT	TT	CC	TT	TT	TT	TT	TT	TT	T=Val	C=Val
122a E2	AA	AA	AA	CC	AA	AA	AA	AA	AA	AA	A=Arg	C=Ser
123 E2 (a)	GA	GA	GA	GG	AA	AA	AA	GA	AA	GA	G=Arg	A=Lys
123 E2 (b)	CT	CT	CT	TT	CC	CC	CC	CT	CC	CT	T=His	C=His
124 E1	AA	AG	AG	AA	GG	AG	AG	AG	AG	AA	A=Pro	G=Pro
126 E1 (b)	AA	AA	AA	AA	AA	AA	AG	AA	AA	AA	G=Glu	A=Lys
126 E1 (a)	CT	TT	CC	TT	TT	TT	CT	TT	CT	TT	Non-coding	Promoter region
126 E1 (c)	GA	GG	AA	GG	GG	GG	GA	GG	GA	GG	Non-coding	Promoter region
128 E1	GA	GG	GA	GG	GG	GG	GA	GG	GA	GG	Non-coding	Promoter region

4.2.2. Genotyping of Holstein-Friesian and Norwegian red populations

In order to determine if the differences in selection carried out in the breeding programmes for HF and NR could have an effect on the selection of these SNPs, the alleles for the newly identified SNPs were determined for a larger sample size. Genotyping by sequencing is limited to the number of samples that can be processed in a cost-effective manner. For the larger sample size, we used iPLEX Gold genotyping (Sequenom, San Diego, CA, USA). This technology allows us to test multiple SNPs using custom designed labelled probes specific

for each SNP, reducing the cost of sequencing. We reduced the number of SNPs from 17 down to 12 and removed all SNPs present in a frequency smaller to 0.1 or present in a single animal to avoid possible sequencing error. The SNPs analysed were: *BBD128 E1, BBD126 E1 (a), BBD126 E1 (b), BBD115 E2, BBD116 E2, BBD117 (a), BBD117 (b), BBD121 E1, BBD122 E2, BBD123 E2 (a), BBD123 E2 (b)* and *BBD124 E1*. The SNPs were genotyped in a larger sample size of 30 NR and 30 HF DNA samples. Genotyping results were analysed using Genepop software package (Table 4. 2) (Rousset 2008). Allele frequencies were analysed by Fisher's exact test. Four SNPs showed significant differences in allele frequencies in the genetic analysis carried between the two breeds, two of the SNPs are non-synonymous: in genes *BBD115* and *BBD117* ($P<0.05$). The largest difference in allele frequencies was detected for the non-synonymous SNP in the *BBD115* gene ($P=0.002$) where the G allele, in an alanine codon, was present at 0.52 frequency in Norwegian Red but with a frequency of 0.80 in the Holstein-Friesian breed.

In addition, genotype frequency differences between HF and NR, and Hardy-Weinberg equilibrium were analysed using Genepop software (Table 4. 3). The genotypic analysis determines the differences of the different genotypes using Fisher's exact test. Genotype frequencies were found to be significantly different between the two populations for *BBD128 E1, BBD115 E2, BBD117 (b), BBD121 E1, BBD122 E2, BBD123 E2 (a)* and *BBD123 E2 (b)*. Hardy-Weinberg equilibrium indicates whether the population is randomly mating and is not affected by mutation, migration, or other influences that could affect the frequency distribution. No significant departures from H-W equilibrium were found.

Table 4. 2. Comparison of novel SNP allelic frequencies between NR and HF populations.

BBD128 E1, BBD126 E1 (a), BBD126 E1 (b), BBD115 E2, BBD116 E2, BBD117 (a), BBD117 (b), BBD121 E1, BBD122 E2, BBD123 E2 (a), BBD123 E2 (b) and BBD124 E1. SNPs were genotyped in 30 NR and 30 HF using iPlex Gold genotyping assay. Genepop software package was used to analyse the difference in frequencies between the two groups by Fisher's exact test. BBD115 E2, BBD117 (b), BBD121 E1 and BBD122 E2 were found to have significantly different frequencies between population. P values <0.05 are considered significant and highlighted in bold.

Locus	Region	SNP	Amino Acid changing	Allele Frequencies Norwegian Red N=30	Allele Frequencies Holstein-Friesian N=30	*P* value
BBD128 E1	Non-coding	G/A	-	0.82 0.18	0.90 0.10	0.293
BBD126 E1 (a)	Non-coding	C/T	-	0.82 0.18	0.87 0.13	0.617
BBD126 E1 (b)	E1	A/G	Lys/Glu	0.82 0.18	0.87 0.13	0.622
BBD115 E2	E2	G/A	Ala/Thr	0.52 0.48	0.80 0.20	**0.002**
BBD116 E2	E2	A/G	-	0.75 0.25	0.83 0.17	0.371
BBD117 (a)	Only 1 exon	T/A	Stop/Lys	0.57 0.43	0.73 0.27	0.087
BBD117 (b)	Only 1 exon	G/A	Val/Ile	0.97 0.03	0.80 0.20	**0.007**
BBD121 E1	Non-coding	G/C	-	0.70 0.30	0.50 0.50	**0.038**
BBD122 E2	E2	C/T	-	0.70 0.30	0.50 0.50	**0.039**
BBD123 E2 (a)	E2	A/G	Lys/Arg	0.72 0.28	0.53 0.47	0.058
BBD123 E2 (b)	E2	C/T	-	0.72 0.28	0.53 0.47	0.062
BBD124 E1	E1	A/G	-	0.72 0.28	0.53 0.47	0.062

Table 4. 3. Genotypic, genotype frequency, Ho1, He or Ho2, and Hardy-Weinberg analysis results for NR and HF populations

Genotyping analysis determines the presence of a significantly different distribution of genotypes, Ho1, He and Ho2, for each of the locus analysed between the two populations. We found that BBD128 E1, BBD115 E2, BBD117 (b), BBD121 E1, BBD122 E2, BBD123 E2 (a) and BBD123 E2 (b) have significantly different frequencies between NR and HF. In addition, we found that all loci are in Hardy-Weinberg equilibrium. Significant P values (<0.05) are marked in bold.

		Hardy-Weinberg	
	Genotypic analysis	NR	HF
Locus	P-Value	P-value	P-value
BBD128	0.247	0.549	1
BBD126 (a)	0.618	0.550	0.416
BBD126 (b)	0.607	0.550	0.412
BBD115	**0.004**	1	0.307
BBD116	0.403	0.331	1
BBD117 (a)	0.068	1	0.636
BBD117 (b)	**0.008**	1	1
BBD121	**0.032**	0.687	1
BBD122	**0.034**	0.689	1
BBD123 (a)	**0.046**	0.376	1
BBD123 (b)	**0.047**	0.373	1
BBD124	0.050	0.3783	1

4.2.3. Haplotype analysis of bovine Chr13 beta-defensin gene cluster

Haplotypes are artificial constructs with no physiological function and are defined as a fragment of genetic material not affected by recombination events. The effect of this is the migration of the alleles present in the haplotype as a unit. This effect has proven useful in gene-wide association studies where phenotypes have been linked to the presence of a particular haplotype. In order to determine the presence of haplotypes in the Chr13 β-defensin gene cluster, linkage disequilibrium, the non-random association between two loci, was calculated by Genepop, using the genotyping results. We found several loci to have a significant linkage which suggests the existence of a haplotype among them (Appendix: Table 5). The analysis was carried out individually in each population. We found differences between the values observed in NR and HF. *BBD115-BBD11E2, BBD117 (a)-BBD121, BBD117 (a)-BBD122, BBD117 (a)-BBD123 (a), BBD117 (a)-BBD123 (b)* and *BBD117 (a)-*

BBD124 were found to be in linkage in NR but not in HF. Similarly, *BBD116-BBD117 (a)*, *BBD117 (b)-BBD121* and *BBD117 (b)-BBD122* were found in linkage for HF but not for NR.

Haplotype reconstruction was carried out in parallel using Arlequin and HAP software. Subsequent cluster analysis of the genotype data identified two clear haplotypes across the gene cluster: cluster A and cluster B. The two analytical tools identified slightly different haplotypes for the two breeds under analysis. There are smaller differences as well in the predictions from each programme. The precise SNP composition of those haplotypes is not consistent. The consistent SNPs present in cluster A are *BBD128* to *BBD115* and in cluster B are *BBD121* to *BBD124* (Fig. 4. 1).

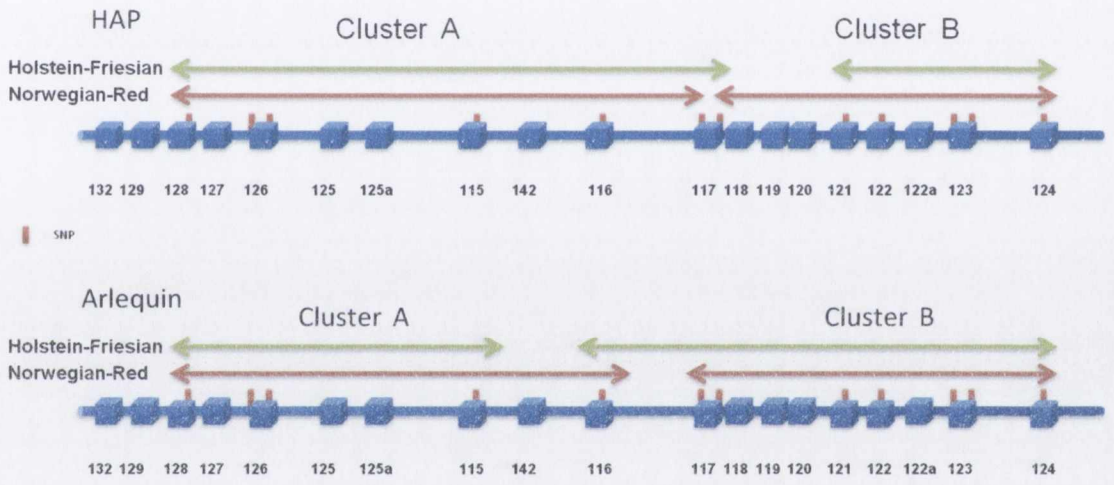


Fig. 4. 1. Haplotype maps for Chr13 β-defensin gene cluster between SNPs in Norwegian Red and Holstein-Friesian

Haplotypes were identified using the linkage disequilibrium data. Data for each population was analysed in parallel with two software packages: Arlequin and Hap. The predictions obtained differ in the two populations and between programmes. The two clusters, A and B, were built as the minimum SNPs common in each prediction , Cluster A was from BBD128 E1, BBD126 E1 (a), BBD126 E1 (b) and BBD115 E2. Cluster B was formed by BBD121 E1, BBD122 E2, BBD123 E2 (a), BBD123 E2 (b) and BBD124 E1. The Genes are displayed according to their location on Chr13. Small red vertical bars indicate the location of SNPs.

Three major haplotypes, with frequencies bigger than 5%, were detected in both breeds across both cluster A and B. The predominant haplotype in cluster A (GCAG) was found

with a frequency of 0.8 within the Holstein-Friesian breed but only 0.52 in Norwegian Red (Table 4. 4). However, 0.3 of NR are carrying a haplotype in this cluster that is segregating at a very low frequency in Holstein-Friesian (0.07). This second haplotype in cluster A differs from the predominant haplotype by only one SNP, i.e. the A/G (Thr/Ala) SNP in *BBD115*. Three major haplotypes were also found segregating within cluster B (Table 4.4). Two of these, which are clear haplotypes carrying different alleles at each locus, represented the majority of animals genotyped. GAACA was found with a frequency of 0.7 in Norwegian Red, but 0.47 in Holstein-Friesian; CTGTG was present at a frequency of 0.28 in Norwegian Red but at 0.43 in Holstein-Friesian.

Table 4. 4. Haplotype frequencies detected in each breed.

Of all the predicted haplotypes, only three in each cluster were found in a frequency higher than 0.05. In Cluster A, the highest frequency haplotype, GCAG, was found in 0.8 of all HF analysed. The same haplotype was only found in 53 of NR. In contrast, GCAA was found in only 0.07 of HF but was present in 0.3 of NR. For cluster B, 0.47 of HF were GAACA for 0.7 of NR, 0.43 of HF were CTGTG for 0.28 of NR and only 0.07 HF and 0.02 NR were CTACA.

Cluster A

BBD128 E1	*BBD126 E1 (A)*	*BBD126 E1 (B)*	*BBD115 E2*	HF	NR	HF Freq	NR Freq	Chi-Square P=0.001
G	C	A	G	48	31	0.8	0.52	
A	T	G	A	6	11	0.1	0.18	
G	C	A	A	4	18	0.07	0.3	
Others				2		0.03	0	

Cluster B

BBD121 E1	*BBD122 E2*	*BBD123 E2 (A)*	*BBD123 E2 (B)*	*BBD124 E1*	HF	NR	HF Freq	NR Freq	Chi-Square P=0.054
G	A	A	C	A	28	40	0.47	0.7	
C	T	G	T	G	26	17	0.43	0.28	
C	T	A	C	A	4	1	0.07	0.02	
Others					2	00	0.03	0	

4.2.4. SNPs' correlation with fertility data

Physiological relevance of SNPs is determined by correlation with a particular phenotype. The presence of an SNP does not necessarily have consequences in the protein synthesised from the gene. Even non-synonymous SNPs that induce an amino acid change could be neutral in terms of producing a different phenotype. Our priority is to determine if the presence of any of the alleles for the newly discovered loci has an effect in fertility. The initial genotyped population were female; our hypothesis is that these genes have a role in fertility from the male perspective, sperm maturation, sperm penetration, etc. The presence of different alleles could have an effect on these roles. In order to test the hypothesis, a new population of 19 bulls was genotyped for seven of the newly discovered SNPs *BBD126E1*

(a), BBD126 E1 (b), BBD115E2, BBD117a, BBD117b, BBD123E2 (a) and *BBD123E2 (b).* Primers previously used were applied for PCR amplification of *BBD126 exon 1, BBD115 exon 2, BBD117* and *BBD123 exon 2.* PCR products were then sent for commercial sequencing and retrieved sequences were analysed using Bioedit software package. In addition, fertility data such as frequency of pregnancy pre- and post-freeze semen were available for these bulls (Table 4. 5).

Genotyping results and fertility data were analysed using the SPSS statistical software package. Pearson correlation coefficients were used to detect the existence of associations between the presence of a particular allele and high or low pregnancy rate (Table 4. 6). No significant correlations were found between the presence of a particular allele and a difference in pregnancy with pre- and post-freezing sperm. No correlation could be calculated for *BBD126b* due to the lack of genetic variability in the sample analysed, all animals having the same allele composition. Strong correlation between the two fertility scores was found as expected.

Table 4. 5. Genotyping results for population of 19 bulls and pre- and post-freezing pregnancy frequencies

BBD126E1 (a), BBD126 E1 (b), BBD115E2, BBD117a, BBD117b, BBD123E2 (a) and BBD123E2 (b) were genotyped by PCR amplification and commercial sequencing. Fertility data was given by the National Cattle Breeding Centre which supplied the sperm samples for DNA extraction. BBD126 E1 (b) genotypes were identical in all bulls making impossible any time of correlation analysis for this locus.

Sample	BBD126a	BBD126b	BBD115	BBD117a	BBD117b	BBD123a	BBD123b	freq pregnancy frozen	freq pregnancy
1	CT	AA	AG	TT	GG	AA	CC	0.536	0.507
2	CC	AA	GG	TT	AG	AA	CC	0.551	0.552
3	CC	AA	AG	AT	GG	GG	TT	0.564	0.557
4	CC	AA	AG	AT	GG	GA	CT	0.529	0.491
5	TT	AA	AA	AA	GG	GG	TT	0.562	0.555
6	CC	AA	GG	TT	GG	GA	CT	0.553	0.508
7	CC	AA	GG	AT	GG	GA	CT	0.642	0.552
8	CC	AA	AG	AT	GG	GG	TT	0.556	0.531
9	CC	AA	GG	TT	GG	GA	CT	0.56	0.574
10	CT	AA	AG	AT	AG	GA	CT	0.56	0.567
11	CC	AA	GG	TT	AG	GA	CT	0.518	0.41
12	CC	AA	AG	AT	GG	GA	CT	0.575	0.551
13	TC	AA	AG	AT	GG	GA	CT	0.57	0.521
14	CC	AA	AG	TT	GG	GA	CT	0.558	0.489
15	CC	AA	GG	TT	GG	GA	CT	0.549	0.528
16	CC	AA	GG	TT	GG	GA	CT	0.454	0.438
17	CC	AA	GG	TT	GG	AA	CC	0.582	0.582
18	CC	AA	GG	TT	GG	AA	CC	0.578	0.581
19	CT	AA	AG	AT	GG	GA	CT	0.582	0.543

Table 4. 6. Pearson correlation results between 19 bulls' genotype data and frequency of pregnancy pre- and post-semen freezing

Genotyping results for BBD126E1 (a), BBD126 E1 (b), BBD115E2, BBD117a, BBD117b, BBD123E2 (a) and BBD123E2 (b). SNPs in 19 bulls and frequency of pregnancy pre- and post-semen freezing for each bull were analysed by Pearson correlation using SPSS software package. No correlation was found between the genotyping and fertility data. No analysis could be carried out for BBD126 E1 (b) due to lack of variability.

Correlations

		freq_pregn	freq_preg_frozen
BBD126 E1a	Pearson Correlation	-0.083	-0.161
	Sig. (2-tailed)	0.734	0.509
BBD126 E1b	Pearson Correlation	[a] .	[a] .
	Sig. (2-tailed)	.	.
BBD115 E2	Pearson Correlation	-0.075	-0.111
	Sig. (2-tailed)	0.76	0.652
BBD117a	Pearson Correlation	-0.341	-0.275
	Sig. (2-tailed)	0.153	0.254
BBD117b	Pearson Correlation	0.172	0.179
	Sig. (2-tailed)	0.482	0.465
BBD123 E2a	Pearson Correlation	-0.021	-0.098
	Sig. (2-tailed)	0.933	0.69
BBD123 E2b	Pearson Correlation	0.021	0.098
	Sig. (2-tailed)	0.933	0.69
freq_preg	Pearson Correlation	1	.706**
	Sig. (2-tailed)		0.001
freq_preg_frozen	Pearson Correlation	.706**	1
	Sig. (2-tailed)	0.001	

a. Cannot be computed because at least one of the variables is constant.

4.2.5. Effect of SNPs on sperm penetration of cervical mucus.

Fertility data previously used was obtained from artificial insemination i. This technique could bypass some of the potential β-defensins roles in fertility such as sperm penetration on cervical mucus. In order to determine the effect of the newly discovered SNPs on the ability of sperm cells to penetrate the cervical mucus, a swim-through assay or sperm penetration assay was used. The SNPs selected for this analysis were *BBD126 E1 (b), BBD123 E2 (a) and BBD117 (b).* Capillary tubes were filled with cervical mucus from two cows during ovulation. The mucus was then pooled. Sperm cells were stained with Hoescht stain solution and were then allowed to swim through the capillary tube. After 10 minutes, the cells were heat-killed and were counted at 20 mm, 40 mm, 70 mm and 80 mm marks in the capillary tube. Three different comparisons were carried out in the sperm penetration assay. In the first, for *BBD126 E1 (b)* SNP, cell count data for AA genotype were compared with AG genotype, due the lack of GG bulls. No significant differences were found between the two alleles (Fig. 4. 2). In a second comparison, cell counts numbers for TT genotype and AA genotype for *BBD117 (b)* were analysed. Significant differences were found between them for all distance marks (Fig. 4. 3). Finally, cell counts for genotype AA and genotype GG for BBD123 E2 (a) were analysed. No significant differences were found at any distance mark (Fig. 4. 4).

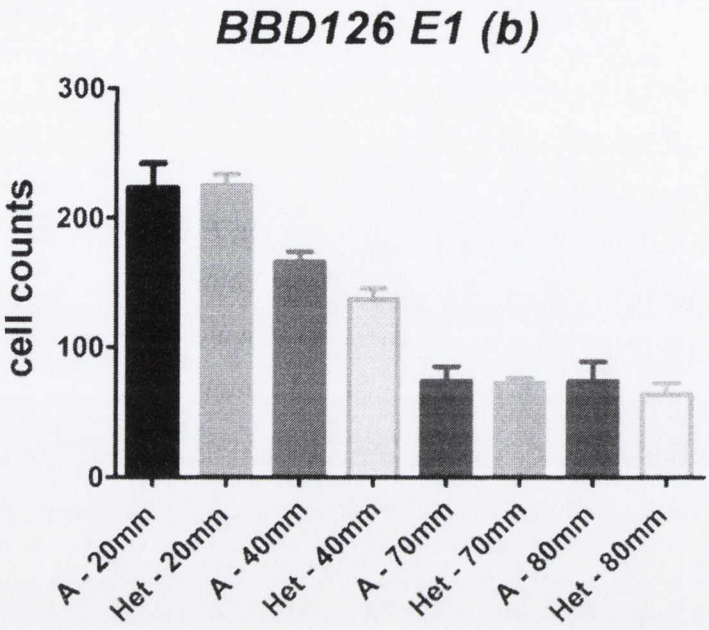


Fig. 4. 2. Cell counts of sperm cells for BBD126 E1 (b) genotypes

Sperm penetration assay results for two groups of bulls with different genotypes for SNP BBD126 E1 (b). Cells were counted at 20, 40, 70 and 80 mm marks in a capillary tube of two groups of bulls; one group with AA genotype and a second one with AG genotype for SNP BBD126 E1 (b). No significant differences were found between the two groups. Graph represents mean of 5 technical replicates of 3 bulls, error bars indicate the standard error. Data were analysed using Mann-Whitney U test.

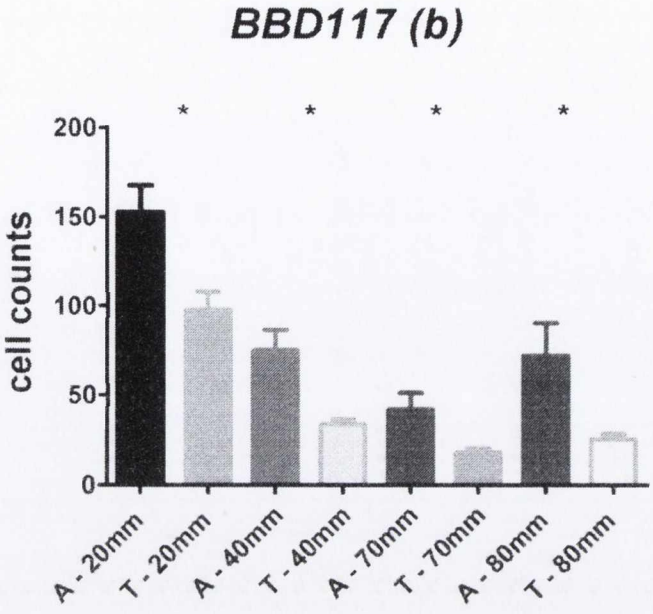


Fig. 4. 3. Cell counts of sperm cells for BBD117 (b) genotypes

Sperm penetration assay results for two groups of bulls with different genotypes for SNP BBD117 (b). Cells were counted at 20, 40, 70 and 80 mm marks in a capillary tube of two groups of bulls; one group with AA genotype and a second one with TT genotype for SNP BBD117 (b). Significant differences were found between the two groups for all marks. Graph represents mean of 5 technical replicates of 3 bulls, error bars indicate the standard error. Data was analysed by Mann-Whitney U test.

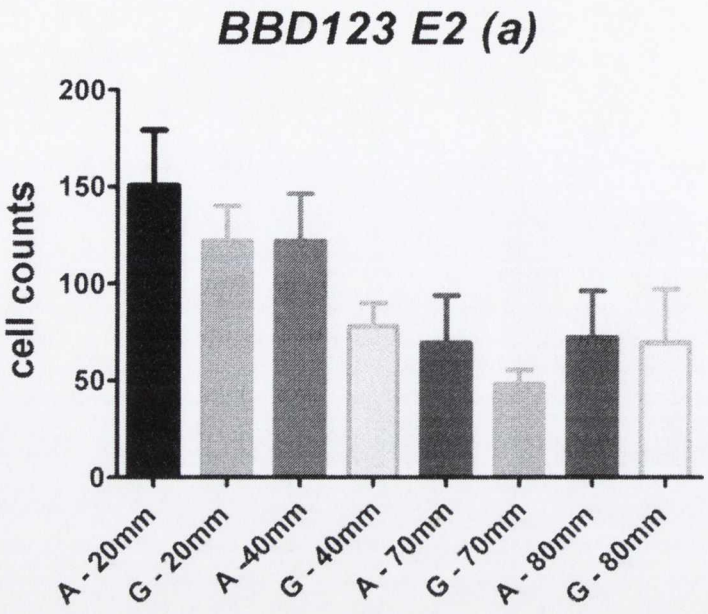


Fig. 4. 4. Cell counts of sperm cells for BBD123 E2 (a) genotypes

Sperm penetration assay results for two groups of bulls with different genotypes for SNP BBD123 E2 (a). Cells were counted at 20, 40, 70 and 80 mm marks in a capillary tube of two groups of bulls; one group with AA genotype and a second one with GG genotype for SNP BBD123 E2 (a). No significant differences were found between the two groups. Graph represents mean of 5 technical replicates of 3 bulls, error bars indicate the standard error. Data was analysed by Mann-Whitney U test.

4.3. Discussion

In a preliminary study the coding region of all members of the chromosome 13 β-defensins cluster were amplified by PCR and the resulting products sequenced. All sequences were aligned and analysed, resulting in the discovery of 17 novel SNPs (Table 4. 1). Several of these SNPs were discarded for further studies due to their presence in one animal, suggesting a possible sequencing error. The remaining 12 SNPs were commercially genotyped in a bigger population, 30 HF and 30 NR by iPlex Gold technology, based on the identification of the different alleles by mass spectrometry MALDI-TOF. The analysis of the data showed at least four of them with significant different frequencies between the two breeds: *BBD115 E2*, non-synonymous; *BBD117 (b)*, non-synonymous; *BBD121 E1*, non- coding region and *BBD122 E2*, synonymous. *BBD115 E2* SNPs introduce an alanine to threonine change from a non-polar to a highly polar amino acid. This type of change could result in a change in the protein conformation and/or activity. In *BBD117 (b)*, the change from a valine to isoleucine, both non-polar amino acids, is most likely to not introduce a drastic change in the protein structure. *BBD117* is a special case; initially thought to be a pseudogene due to the lack of a first codon and lacking the signal peptide required for protein secretion, nevertheless gene expression was detected. An explanation most likely explanation is that the first exon is present but has not been found. The exon could be as well shared with other genes by alternative splicing episodes from the nearest defensin.

BBD121 E1 was found in the immediate upstream region of *BBD121* gene. This region is part of the promoter region which controls gene expression regulation. SNPs in these areas are of interest because of their potential effect in the gene expression regulation by removing or creating new transcription factor binding sites. An initial analysis found no transcription factor binding sites predicted in the region of the SPNs. *BBD122 E2* does not incorporate a change in amino acid sequence in the protein. *BBD122 E2* and *BBD121 E1* do not have a direct effect on protein structure but due to their presence in one of the haplotypes found in gene cluster, they could act as markers for other polymorphisms in the same haplotype. The differences observed in SNP frequencies could have two origins. Firstly, differences could be the result of differences between the two primordial populations from which each breed originated. A second explanation could be the distinctive selective pressure used in the breeding programmes that have been implemented in HF and NR breeds.

In the previous chapter we discussed the presence of a distinctive group of genes we called Class A. Based on the gene expression data we hypothesised that these genes have a role in fertility. Three SNPs were found in members of these groups. Two of them, *BBD128 E1* and *BBD126 E1 (a,)* were synonymous; the third, *BBD126 E1 (b)*, introduces a non-synonymous change from isoleucine to methionine but no frequency differences were found between the two breeds. The lack of significant differences suggests that the breeding programme had no effect on the selection process for these SNPs. However, it does not discard a role in fertility for these genes.

Linkage disequilibrium analysis showed the presence of differentially linked SNPs between NR and HF. The SNPs form two clusters at each edge of the β-defensin cluster. For a better visualisation of the results, two programmes (HAP and Arlequin) were used to reconstruct the haplotypes found. The small differences in the predictions, could be caused by recombination events in the region between both haplotypes, the noise of the data and the difference in mathematical approach for the prediction. Ignoring the equivocal data, we can assert that two clusters exist: Cluster A formed by *BBD128 E1*, *BBD126 E1 (a)*, *BBD126 E1 (b)* and *BBD115 E2* and Cluster B formed by *BBD121 E1*, *BBD122 E2*, *BBD123 E2 (a)*, *BBD123 E2 (b)* and *BBD124 E1*. Linkage values change between NR and HF in the central region. This suggests small differences in composition between the two analysed populations, possibly due to recombination events that occurred after the populations were separated. Predominant frequencies of the different haplotypes found in each cluster were compared between breeds. We found in Cluster A that the predominant haplotype (GCAG) accounted for 80% of HF individuals. In contrast, the same haplotype is only found in 50% of NR individuals. It is interesting the presence of GCAA, found in 7% of HF and 30% of NR. The origin of this haplotype was probably a recombination event that affects the last SNP in the haplotype. The origin of the different frequencies could be in the original population used for the breeding programme or selective pressures of the breeding programme. In Cluster B, a predominant haplotype (GAACA) is found for 70% of individuals in NR. The same haplotype only represents 47% of individuals in HF. The selective pressures of divergent breeding programmes have resulted in different phenotypes for high milk production (Holstein-Friesian) or high fertility (Norwegian Red).

Differences in frequencies of SNPs and haplotypes between breeds could have two possible explanations. Firstly, with the breeds having been derived from two distinctive populations, these differences could have originated before the selective pressures of the breeding

programmes or be caused by the breeding programmes. This dichotomy could only be resolved by analysing the original population or a population derived from the same ancestral group as NR or HF but under a different breeding programme. A second unresolved question is the biological meaning of these SNPs and haplotypes. As mentioned before, the evidence for the role of these genes in fertility is strong, not just in bovine but other species. However, the exact role and function has not been worked out, making the evaluation of the biological importance of these SNPs impossible.

No correlation between the novel SNPs and fertility data was found. Low fertility bulls are rarely preserved in a population. This phenomenon biases all available samples towards high fertility bulls, therefore affecting the result of any possible correlation study. In addition, all fertility data is obtained by artificial insemination in which the sperm is injected into the uterus. This bypasses important regions in which the defensin could be important in natural conditions by the selection and outcome of the insemination. As a preliminary study, the ability of penetrating cervical mucus of different genotypes for some of the discovered alleles was analysed in a reduced number of bulls (N=3) for each genotype (Fig. 4. 2, Fig. 4. 3 and Fig. 4. 4). These showed promising results, in particular for the *BBD117 (a)* comparison; the T allele, that introduces a STOP codon, showed a lower penetration rate that the A allele. An added issue is the lack of homozygote bulls for some of the alleles analysed, in particular *BBD126 E (b)*. The bulls analysed in this study were samples donated by the National Cattle Breeding Centre, and have been subject to motility and survival tests. This could in fact bias our sample size, removing from our sample population the bulls with lower fertility that could correspond with the presence of a particular genotype. In addition, with a bigger sample size, discarded bulls will be genotyped to determine the possible correlation between a genotype and poor performance. Finally, differences between breeds should be analysed at gene regulation level. In chapter 3, we discussed the differences in the promoter region of the different genes. A population study of the promoter region of HF and NR could shed some light on these issues.

The negative association between production traits and fertility in cattle has created economic and welfare issues for the agricultural industry as a result of intensive selection for milk production. This has spawned immense interest in developing breeding strategies to overcome the limitations of poor fertility including genome-wide association studies (Hoglund, Guldbrandtsen et al. 2009, Huang, Kirkpatrick et al. 2010, Sahana, Guldbrandtsen et al. 2010) and other methods for the identification of fertility biomarkers (Peddinti, Nanduri

et al. 2008). The reproductive tissue specific expression profile of this cluster of β-defensins indicates that these genes could become an important tool as markers of fertility that could be incorporated into future cattle breeding programmes (Dekkers 2004, Berry, G. et al. 2010, Pimentel, Bauersachs et al. 2010). Furthermore, the novel SNPs detected could become useful tools in resolving the causative SNPs, underlying associations detected between fertility and genes on chromosome 13 in cattle (Sahana, Guldbrandtsen et al. 2010). Further functional analysis of the role of these potentially epistatic genes in reproduction and/or immunology is required.

CHAPTER 5: DEVELOPMENT OF RESEARCH TOOLS

5.1. INTRODUCTION

Previous work carried out on a human homologue of bovine defensin 126 has shown the role of these defensins in male fertility (Tollner, Bevins et al. 2012). Our own work with the bovine defensin cluster in Chr13 shows that these peptides have important functional differences compared to the other defensins. However, all our observations have been based on the DNA and RNA sequences of β-defensins. To determine the role and mechanism of action of these novel defensins, a series of tools are required, specifically antibodies to be used for the detection of β-defensins on the sperm cell surface, detection in epididymal epithelium and other functional assays.

A great barrier in bovine research is the lack of commercially available reagents, such as antibodies, due to the smaller market compared with human or mouse research. Each group has to invest time and resources in the generation of antibodies for their protein of interest. The technological usage of antibodies has continuously improved from the use of a mix of antibodies with different specificities for molecules (polyclonal antibodies) to the use of hybridoma technology to produce large quantities of a specific antibody of a particular specificity (monoclonal antibodies). Molecular biology techniques were used for further improvement of production by allowing the isolation of the genes responsible of the antibody formation. This enabled the synthesis of recombinant peptides (Fab, scFv) that conserved the specificity of the original antibody but incorporated new properties (smaller size, higher thermal and structural stability, higher solubility, multiple specificity by protein fusion and attachment of other molecules such as drugs for clinical usage or dyes for research use. All these new techniques share a common problem: the structural requirement of the correct combination of heavy and light chains to conserve the specificity observed in the original antibody. This adds a particular degree of difficulty, and has been a limiting step in the use of recombinant antibodies and antibody fragments at their full potential as research and clinical tools. The discovery of heavy chain antibodies in camelids (Hamers-Casterman, Atarhouch et al. 1993), opened a new field of possibilities for molecular engineering of antibodies, the lack of a light chain in the functional antibody allowed easier manipulation and use *in-vitro*; reducing time require to produce monoclonal antibodies.

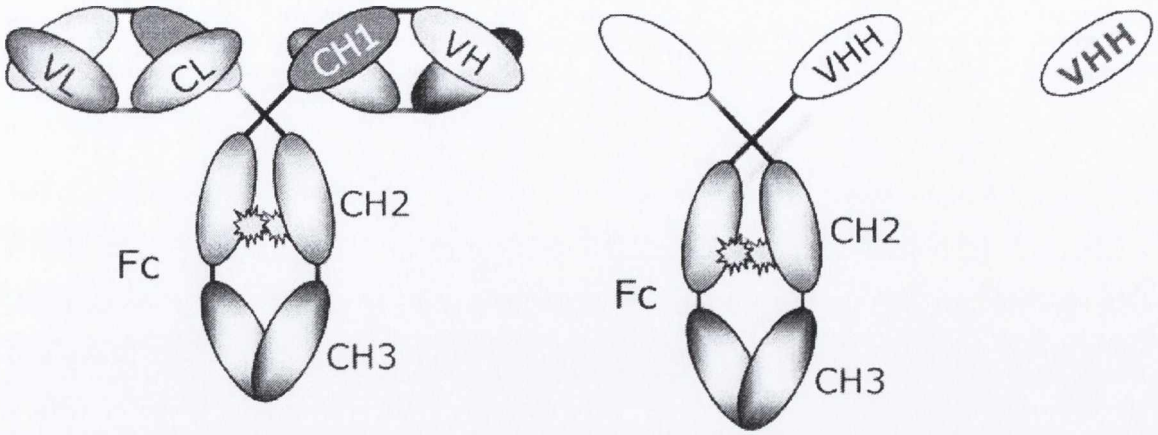


Fig. 5. 1. Comparison of classical antibody structure and heavy chain only antibody.

Classical IgG structure is formed from two heavy chains and two light chains linked by disulphide bonds. The heavy chains are divided in four regions, a variable region VH, and three constant regions CH. The light chain is formed by one variable region VL and one constant region CL. By contrast heavy chain only antibodies (VHH) lack light chains, and the CH1 region of the heavy chain. The VH region is known as VHH and is the smallest functional fragment of a naturally occurring antibody. Figure adapted from (Muyldermans, Baral et al. 2009).

Immunoglobulins are members of a family of proteins for which there are a few isotypes. Isotypes can share the same antigen specificity but they possess some differences in their structural skeleton that determine their function. Antibodies are generally represented as a Y-like structure (Fig. 5. 1.), based in the IgG structure. They are formed by the binding of two heavy chains and two light chains. Each heavy chain is formed by a variable immunoglobulin domain (VH) and three constant immunoglobulin domains (CH); the light chain is formed by a single variable (VL) and a single constant domain (CL). Both heavy chains are bound to each other by two disulphide bonds in a region between the first and second CH domains known as the hinge. Each light chain is bound to a heavy chain by a disulphide bond between the heavy chain constant domain 1 (CH1) and the light chain constant domain 1 (CL1). In addition, the variable domain from the light and heavy chain (VL and VH respectably) interact by hydrophobic amino acids. The C domains, as their name suggests, are conserved regions and form the crystallisable fragment (Fc), named for their ability to form crystals when cleaved from the variable region. This region is recognised by immunoglobulin receptors and complement molecules. The V regions hold the specificity for each antibody and are the regions that bind to the epitope. Each IgG contains two V domains, allowing the

binding of two antigens epitopes by a single antibody. This phenomenon is enhanced by the presence of a flexible hinge between CH1 and CH2, allowing the antibody to accommodate two antigens simultaneously.

The immunoglobulin structure was accepted as canonical and present in all vertebrates but the discovery of functional antibodies lacking the light chain in camels and later in sharks challenges this notion (Hamers-Casterman, Atarhouch et al. 1993, Greenberg, Avila et al. 1995). Three different IgG isotypes have been described in camelids: the classic heavy and light chain antibody (H2L2), heavy chain only antibody (H2) with a long hinge or with a short hinge. Similar antibodies were found in the other members of the camelid family suggesting their origin in a common ancestor (Flajnik, Deschacht et al. 2011). The most striking difference between H2L2 and the long and short hinge H2 is the lack of the CH1 domain in the heavy chain; as described before, CH1 anchors the light chain by a disulphide bond. It was found that a mutation in the splice signal region of the CH1 domain was responsible for the lack of CH1 in the protein (Nguyen, Hamers et al. 1999). The substitution of a G for an A disables the splice site, converting *de facto* the CH1 into an intron. The origin of this modification is known but the discovery of another mutation in the CH1 domain that induces the loss of a Cys has encouraged the formulation of a hypothesis; the mutated Cys is required for the formation of a disulphide bond that stabilised the folding of the immunoglobulin domain. If a mutation occurs in one of the Cys, the stability of the domain could be compromised, producing non-functional antibodies. In this scenario, a mutation which would lose the damaged domain would be beneficial. The loss of the CH1 domain was completed with a series of changes that allowed the resultant antibody to be functional; more specifically a reduction in the hydrophobicity of the VH domain. A similar process has been tested in human antibodies with the end of obtaining functional human VHH, this process has been was named "camelising" by which the CH1 domain was removed from human antibodies. However the resultant product showed poor solubility and tendency to precipitate (Davies and Riechmann 1994).

The classic VH domain is formed by four highly conserved regions known as framework regions (FR) and three highly variable regions known as complimentary determining regions (CDR) (Fig. 5. 2.). These CDR are responsible for the antibody selectivity and are the regions in contact with the correct epitope. The function of the FR is the formation of the scaffold required for the correct display of the CDR. Structurally, FR regions form two parallel β-sheets (one with four and the second with five strands) opposing each other (Muyldermans,

Cambillau et al. 2001). The CDR in this structure forms the loops connecting the strands of the β-sheets, displaying them to the exterior. In addition to the correct folding of the immunoglobulin domain, FR2 is essential for the correct binding of the VH and VL domains. This region is particularly rich in hydrophobic amino acids, promoting the interaction with the rich hydrophobic surface present in the VL domain (Davies and Riechmann 1994). In camelids this region has been subject to a series of mutations that produced several amino acids substitutions. The substitutions found in camelids FR2 are: Val42Phe (Tyr), Gly49Glu (Gln), Leu50Arg(Cys) and Try52Gly(Ser, Leu or Phe) (Muyldermans, Cambillau et al. 2001); these substitutions introduce the change of a hydrophobic amino acid for a hydrophilic one, in brackets are other possible substitutions found for that location. These substitutions reduce the solubility problem observed in the classic antibodies in which the light chain does not interact with the heavy chain. In addition to these differences found in all camelids, it was found that camels and dromedaries (Old World camelids) but not alpacas and llamas (New World camelids) have a disulphide bond between CDR1 and CDR3. This bond has several effects. Firstly, it increases the stability of the molecule and secondly introduces a structural motif that helps compensate for the loss of variability in the missing light chain.

The discovery of heavy chain-only antibodies in camelids raises the question of how the loss of diversity by the lack of light chain and the subsequent combinatorial diversity is compensated in these animals; especially, if we consider that in camels and dromedaries, the heavy chain-only antibodies (VHH) represent 50% of the total IgG and between 25-40% in New World camelids (Hamers-Casterman, Atarhouch et al. 1993, van der Linden, de Geus et al. 2000). By studying the immunoglobulin gene clusters in camelids, we have gained a greater understanding of how VHH can deal with apparent reduction of diversity. Thanks to sequence alignment and analysis, we know that the V segment of the VH domain in VHH antibodies (VHH3) are members of the same family and are related to V segment III (VH3) found in classic antibodies (Nguyen, Hamers et al. 2000). Around forty V segments exclusive to VHH antibodies have been found in dromedaries. They were found following the characteristic translocon organisation of immunoglobulin genes: VH3 + VHH3+ D + J (Flajnik, Deschacht et al. 2011). This organisation suggests the origin of the VHH3 gene fragment from an original VH3 segment and further segment duplication. Despite the significant level of homology between VH3 and VHH3 gene segments, the latter possesses some characteristics that replace some of the lost diversity. The most obvious is the increased

length of the CDR regions; in particular, CDR1 and CDR3 are the longest CDR regions found in vertebrates with seventeen amino acids in camels on average compared with an average of twelve in mouse and nine in humans (Nguyen, Muyldermans et al. 1998). Longer CDR regions mean a potential increase of contact surface between the epitope and the antibody. A second difference is the discovery of novel somatic hypermutation hotspots, in particular in a region outside the CDR1. Somatic hypermutation is the process by which an enzyme known as activation-induced (Cytidine) deaminase (AID) deaminates cytosine nucleotide into uracil, the uracil is then removed and the gap is filled by DNA polymerase introducing point mutations, this modifies the affinity of the antibody. These hotspots are not found to be variable in germ line segments; it is only when the hypermutation process is activated that these regions undergo mutation. The region outside the CDR1 is particularly important as it was found to be in direct contact with the antigen, further enhancing the antigen-antibody contact area (Nguyen, Hamers et al. 2000). The third change from the VH segments is the presence of a Cys in CDR3 that creates a disulphide bond between CDR1 and CDR3. This bond has several effects. Firstly, it stabilises the structure of the immunoglobulin domain and secondly, introduces the formation of a diversity of novel structures for the loops. These structures differ from the structures encountered in classic antibodies.

The lack of light chain is compensated by a higher diversity of tertiary structures. These novel structures do not follow the canonical structures described for classic antibodies. All these modifications are responsible for one of the most interesting properties of VHH antibodies: their ability of recognise epitopes not recognised by classic antibodies (Nguyen, Hamers et al. 2000). It is known that camelid VHH are able to recognise a complete new set of antigens compared with classic antibodies. Both types of antibodies do not compete for antigen recognition; the two groups complement each other. On one hand, we have the classic antibodies in which CDR region forms a flat or slightly concave region. On the other hand, VHH possess a protuberance because of the previously discussed modifications. Several publications have described an interesting effect of this structure, namely the ability of VHH to recognise enzymatic active sites as epitopes (Desmyter, Transue et al. 1996).

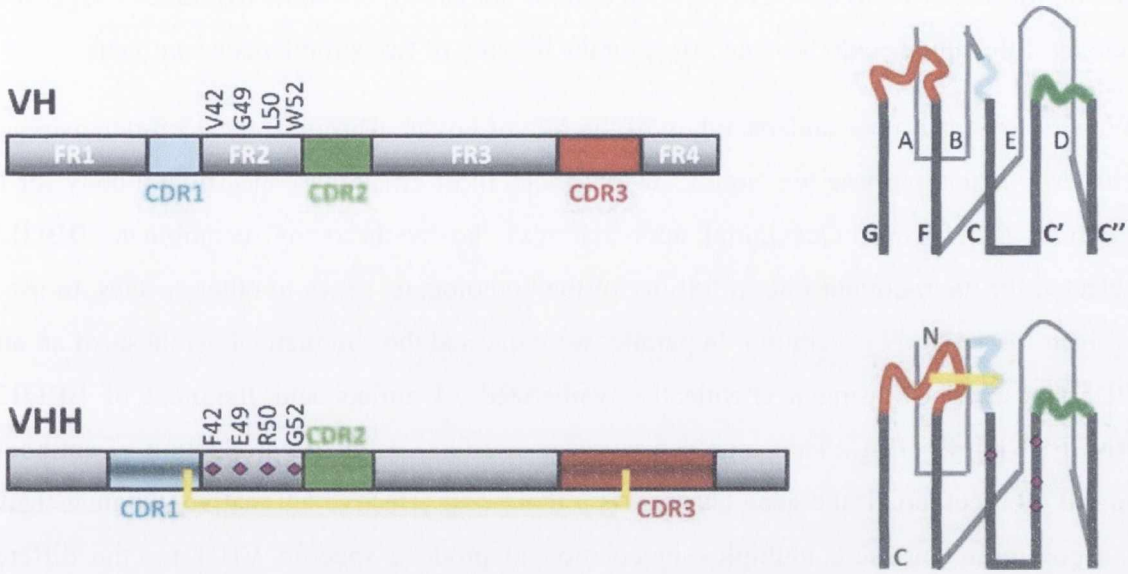


Fig. 5. 2. Structure of variable region of classic antibodies (VH) and heavy chain only antibodies (VHH).

Variable regions of immunoglobulins are formed by two types of domains, four constant framework regions (FR) and three highly variable complimentary determining regions (CDR). VHH have several modifications from the VH canonical structure to compensate for the loss of the light chain. The reduction of variability is compensated by the increase of size of the CDR, particularly CDR3. A series of mutations in FR2 reduce the hydrophobicity of that particular area that in VH is cover by the light chain. In addition some camelids such as camels (Camelus sp.), but not alpacas (Vicugna pacos), acquired a inter loop disulphide bond to increase the structural integrity and stability of the molecule. Di-sulphide bond between CDR1 and CDR3 mark in yellow in the VHH. Figure adapted from (Flajnik, Deschacht et al. 2011).

Camelid VHH differs from classic antibodies in other ways. The only described isotype for VHH is IgG and despite an active search, no IgM or other isotype VHH has been found (De Genst, Saerens et al. 2006). It is possible that the loss of the CH1 domain was only accomplished in a single γ chain. In camelids, nine γ chains are found but only five in camels and six in alpacas are functional (Nguyen, Hamers et al. 2000). Of the functional chains, three, in both species, are responsible for the VHH synthesis. In camels, one of the chains is responsible for the long hinge antibodies and the other two for short hinge antibodies. The

reason for the different in length between them is not known but some hypotheses suggest the length of the hinge could have an effect on the binding of two simultaneous antigens.

We aim to develop our understanding of the role of bovine Chromosome 13 β-defensin gene cluster. For this purpose we require several tools, most crucially a specific antibody for the different β-defensins. Our initial approach was the synthesis of recombinant BBD126, selected for its important role in fertility of the homologous genes in other species, to use as antigen for antibody generation. In parallel we requested the commercial synthesis of an anti-BBD126 antibody using a chemically synthetized 14 amino acid fragment of BBD126, RNGERVINPPTGMC. This approach is not cost-effective for the production of antibodies for all 19 members of the gene cluster. As a more cost effective alternative, we investigated the possibility of using multiplex inoculation to produce specific VHH for the different peptides members of the cluster. In the absence of a recombinant β-defensin protein, we used recombinant bovine TNFα as a proof of principle for this novel method of antibody production. The absence of light chains gives VHH several significant advantages over classical antibodies. The most important is the simplification of their generation by molecular biological methods. Their single chain structure allows for the rearranged gene to be isolated, cloned into bacteria, produced and used without requirement for a matching light chain. In theory, several antigens could be inoculated into a single animal and individual re-arranged genes for each of them could be isolated and produced reducing the animal cost. During the course of the study we devised a novel method to generate VHH antibodies by creating an artificially generated VHH antibody library. The gene structure of VHH is highly conserved with 3 highly variable CDR regions and 3 highly conserved FR regions. We studied the possibility of generating multiple copies of an artificial gene with the conserved sequence for the FR regions and random sequences for the CDR regions. This approach allowed generating a library of VHH antibodies with potentially $1.76e^{+72}$ specificities that could be isolated for the specific antigen by phage display technology.

VHH antibodies present the possibility for the synthesis of cost-effective antibodies essential for small markets such as bovine, porcine and other domestic animals

5.2. RESULTS

5.2.1. COMERCIAL PRODUCTION OF ANTI-BBD126 MONOCLONAL ANTIBODY.

Genscript monoclonal antibody service was described as follows: a 14 amino acids peptide (RNGERVINPPTGMC) was conjugated with KLH as immunogen and 5 BALB/c mice were immunized with the conjugated peptide. All mice showed immune response and the mice with satisfied immune response were used for cell fusion and hybridoma production. A total of four hybridoma cell lines derived from two parental clones are produced.

Table 5. 1. ELISA results for Supernatant samples for different clones.

Quality control carried out by Genscript, ELISA plates were coated with 100ul of 1ug/ml with antigen, RNGERVINPPTGMC.

	Supernatant Dilution								
Clones	1:10	1:30	1:90	1:270	1:810	1:2,430	Blank	Titer	Isotype
A	2.790	2.779	2.728	2.368	1.647	1.129	0.065	>1:2,430	IgG1,κ
B	2.864	2.797	2.704	2.664	2.383	0.883	0.065	>1:2,430	IgG1,κ
C	3.019	2.876	2.876	2.825	2.768	2.241	0.065	>1:2,430	IgG1,κ
D	3.040	2.739	2.491	1.769	1.233	0.610	0.065	>1:2,430	IgG1,κ

5.2.2. ANTI-BBD126 CLONE SELECTION.

GenScript provide us with four different hybridoma clones. The first step was to determine the best candidate as anti-BBD126 antibody for larger scale production. We had previously shown gene expression of BBD126 in the epididymis. We hypothesise, based on the findings in other species with orthologue to BBD126, that the peptide should be present in sperm. To test this hypothesis, a frozen bovine semen straw was separated into cells and supernatant. Sperm cells were lysed as described in section 2.33. Supernatant containing the lysate was recovered and finally the cell lysate pellet was boiled and used as sample (Fig. 5. 1). The western-blot showed clones A and D as best candidates for production and indicate the presence of the defensin in the cell lysate. Clone A was then chosen for larger production. From this point, all reference to anti-BBD126 antibody is based on work done with clone A.

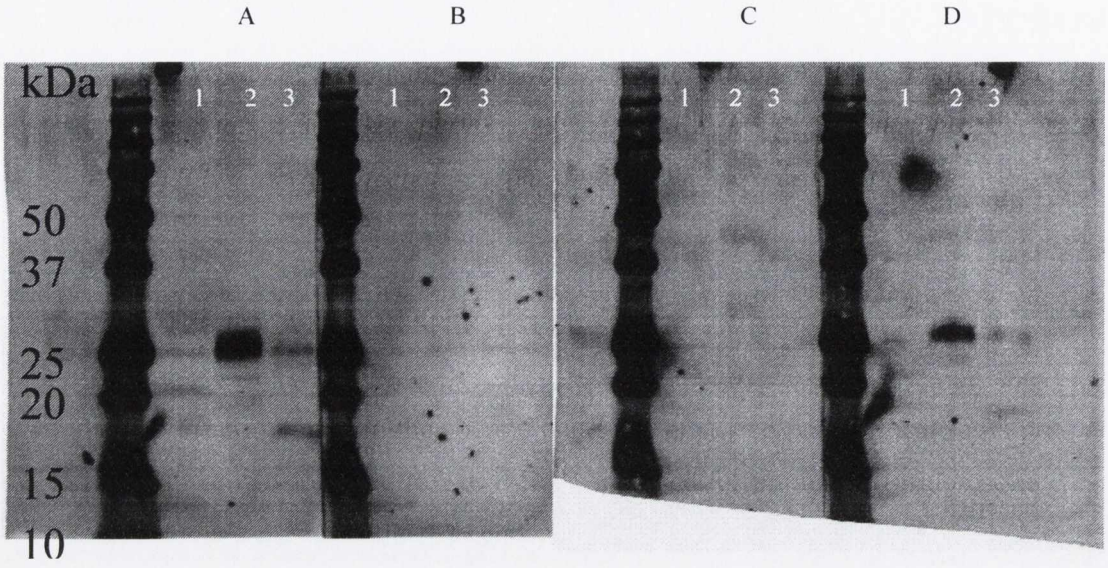


Fig. 5. 1. Anti-BBD126 hybridoma test by western-blot.

Four hybridoma clones producing anti-BBD126 antibodies (A, B, C and D) were tested against: 1, seminal plasma; 2, sperm cell lysate and 3, boiled cell membranes. A 25 kDa band in the cell lysate was detected by clones A and Din the sperm cell lysate.

5.2.3. BBD126 LOCATION IN SPERM CELL SURFACE.

Once the peptide was found in the sperm lysate, the next step was to localise its presence in the cell. Sperm cells were fixed to glass slides and used for immunocytofluorescence. Sperm cells were cultured with anti-BBD126 and stained with an anti-mouse IgG Alexa Alexa Fluor 568 conjugated antibody (Fig 5.2). BBD126 was found in high concentration in the tail surface of sperm cells and in lower quantities in the head of the sperm cells.

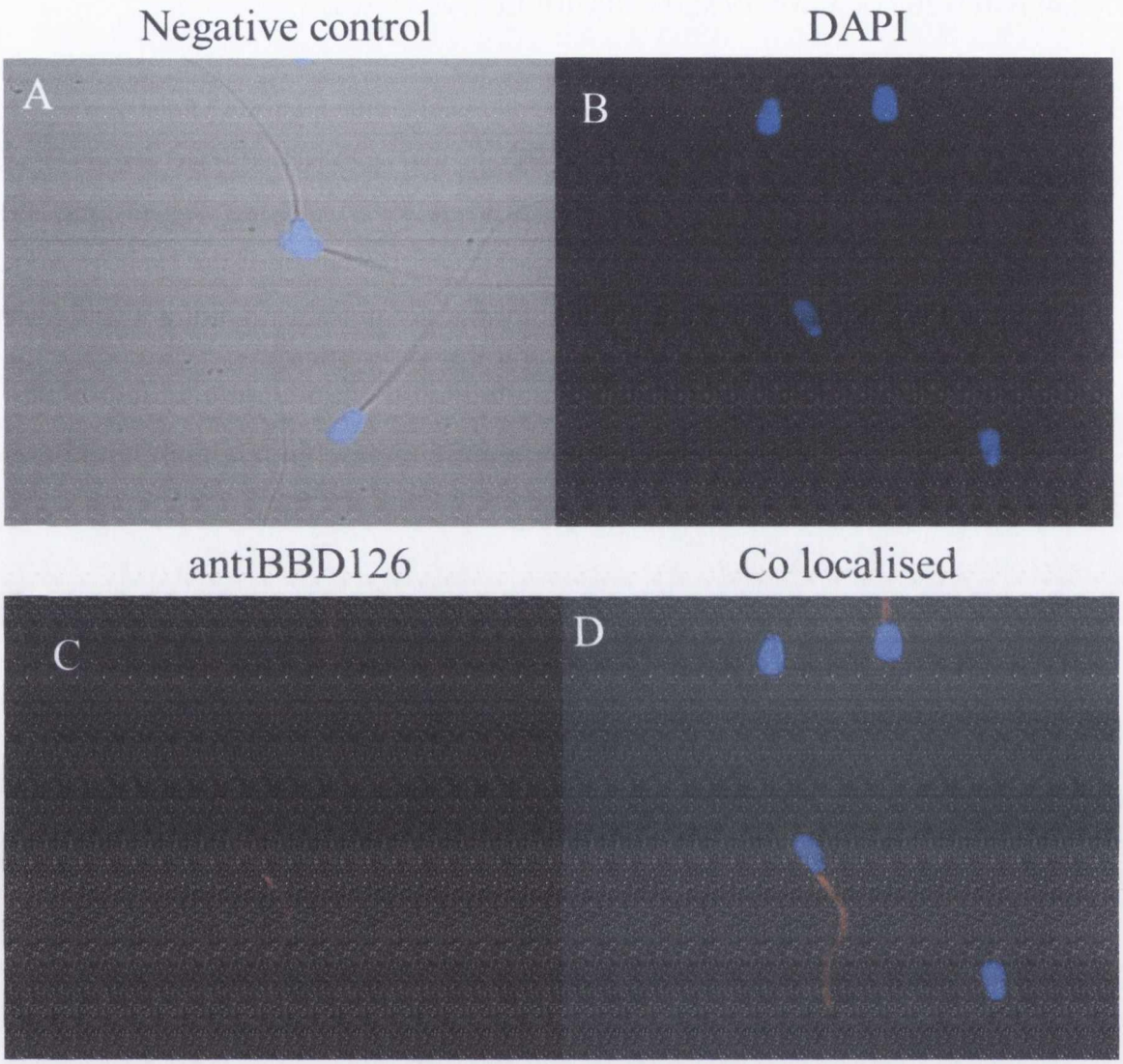


Fig. 5. 2. Immunocytofluorescence localisation of BBD126 in sperm cells.

Live sperm cells were fixed into glass slides and used for immunocytofluorescence with antiBBD126 antibody. A, cells incubated with secondary anti mouse Ab only; B, C and D, cells incubated with both antiBBD126 and secondary Ab. B shows DAPI stain for DNA; C, Alexa Fluor 568 (red) stain for BBD126 and D, co-localisation of A and B. BBD126 was detected in both head and tail of sperm cells.

5.2.4. BBD126 LOCATION IN MALE REPRODUCTIVE TRACT.

Gene expression of *BBD126* was found in the epididymis. Immunohistochemistry was used to characterise the presence of BBD126 protein in the reproductive tract tissue. Tissue samples of cauda of the epididymis and seminal tubule were fixed and stained using anti-BBD126 (Fig. 5. 3 and Fig. 5. 4). Peptide was found in the lumen of the epididymis and absent in the testis tissue, matching the results of the gene expression profiling from Chapter 3.

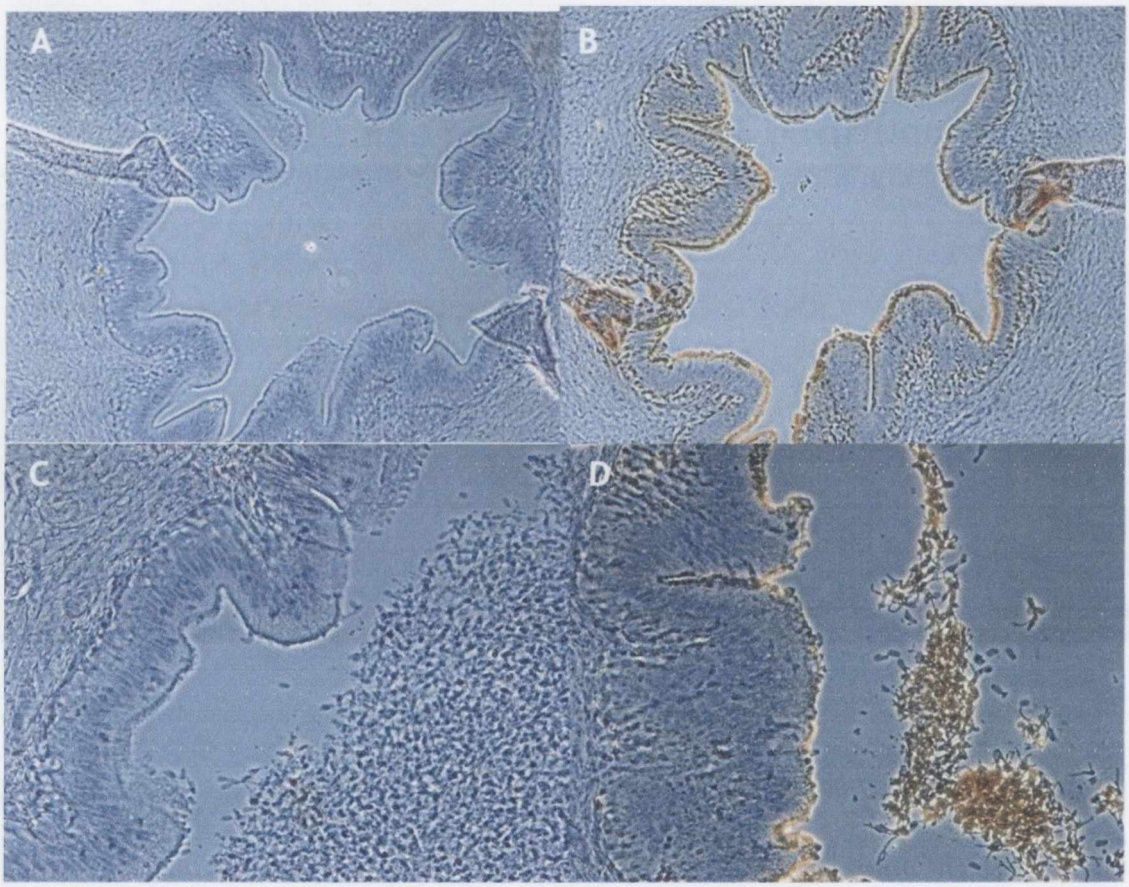


Fig. 5. 3. Immunohistochemistry localisation of BBD126 in cauda of bovine epididymis.

A and C, negative control without anti-BBD126, secondary antibody only; C and D, anti-BBD126 Ab. A and B, 10x; C and D 20x. BBD126 was found in the surface of epididymal epithelium and on the surface of sperm cells. High protein levels match the high gene expression levels found in Chapter 2.

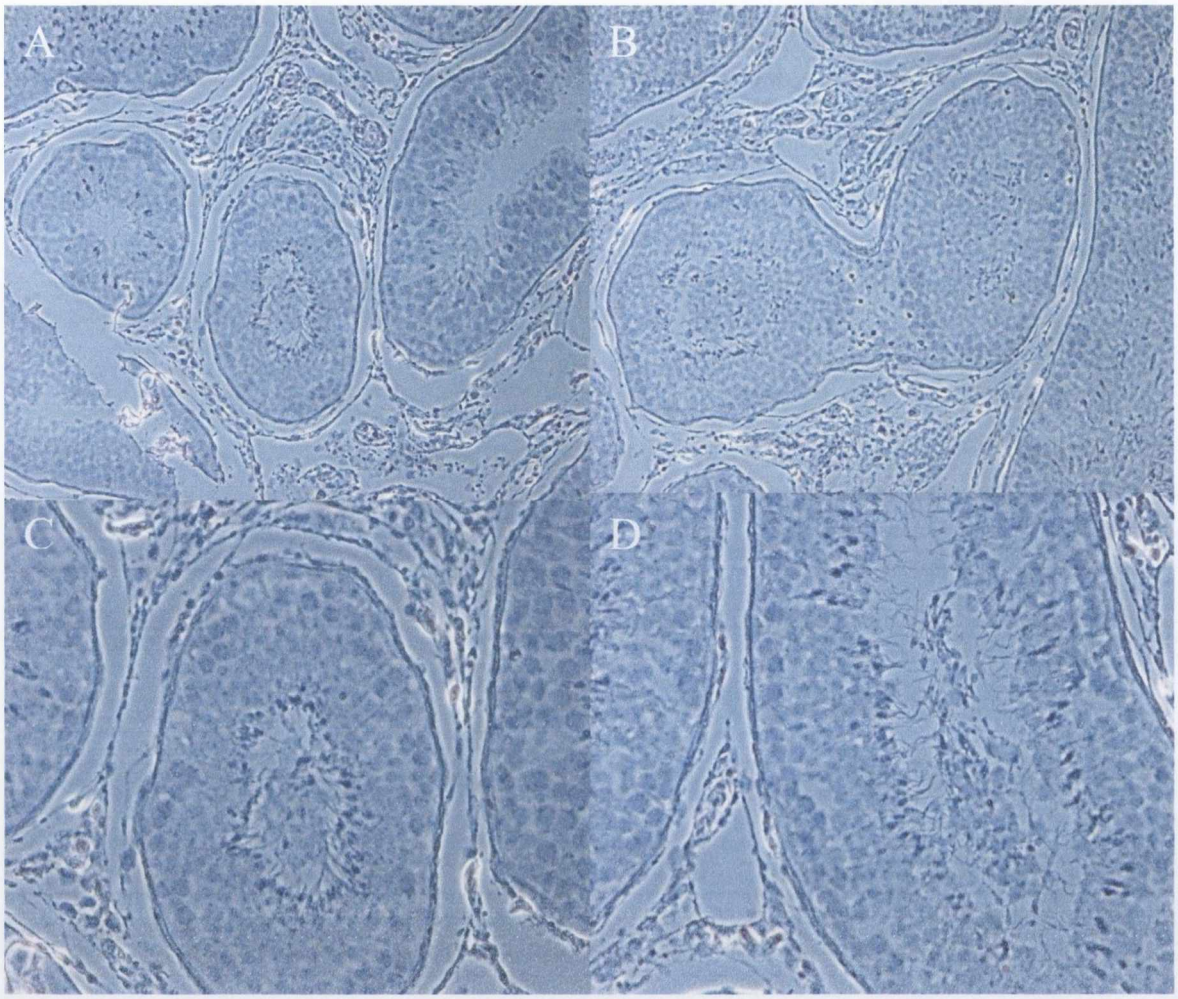


Fig. 5. 4. Immunohistochemistry localisation of BBD126 in seminal tubule of bovine testis.

A and C, negative control without anti-BBD126 Ab, secondary anti-mouse antibody only; B and D anti-BBD126 Ab. A and B, 10x; C and D, 20x. BBD126 was not found in the seminal tubule, matching the low gene expression of BBD126 found in testis in Chapter 2.

5.2.5. ANTI-BBD126 ANTIBODY CHARACTERISATION BY IMMUNOPRECIPITATION.

Further identification of anti-BBD126 specificity was carried out by immunoprecipitation of the antigen. 100 µg of total protein from seminal plasma sample and sperm cell lysate were incubated with anti-BBD126 antibody coated protein A/G agarose beads. Samples were run in a 4-12% SDS-PAGE gel and silver stained (Fig. 5. 5). A clear band was uniquely found in

the precipitation of seminal plasma (B and C). This band was cut and sent to be identified by mass spectrometry protein identification to Fingerprints (Dundee, UK). The band was identified as spermadhesin Z13 gi|12585540 MW: 13.3 kDa.

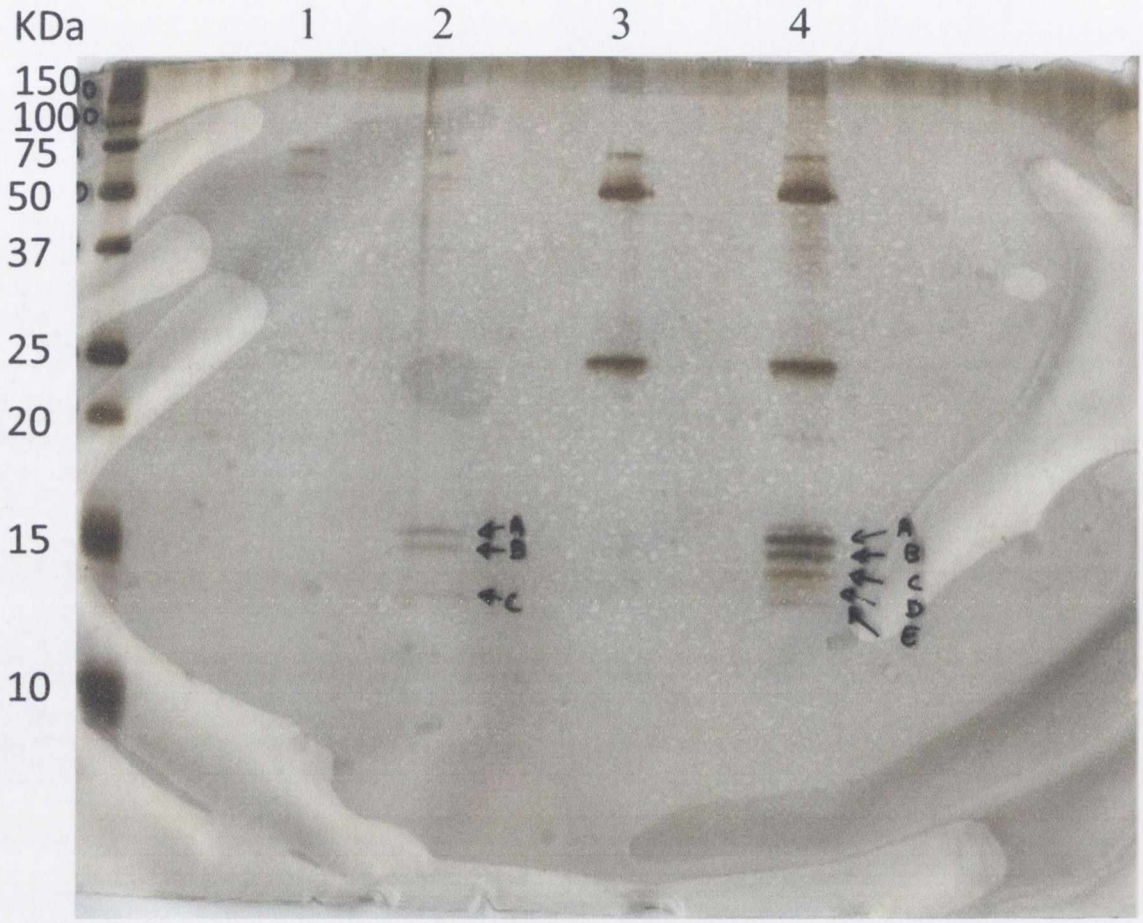


Fig. 5. 5. Immunoprecipitation of anti-BBD126 antigen

Anti-BBD126 antibody was bound to A/G agarose beads and incubated with sperm plasma and sperm cell lysate. After several washes, the elution was run in a 4-12% SDS-PAGE gel and stained with silver staining. 1 and 2, cell lysate and seminal plasma incubated with protein A/G beads. 3 and 4, cell lysate and seminal plasma incubated with anti-BBD126 coated protein A/G beads. 2 bands at 25 and 50 kDa bands for the heavy and light chain of BBD126 were found in lanes 3 and 4. No other bands were found in the 3 lane. Several bands were found in the 4 lane. When compared with the negative control lane 2, the 15 kDa, band C, was selected for mass spectrometry identification.

5.2.6. Recombinant BBD126 production in an eukaryotic expression system

BBD126 mRNA was used as target for the PCR to include both exons as the first exon contains the signal peptide required for the protein secretion into the culture media. Primers pEAK8-126F and pEAK8-126 were used to amplify the cDNA with HotStart PCR kit (Qiagen, Crawley, UK). The resulting product was 280 bp (Fig. 5. 6) and incorporates the restriction digestion sites for HindIII and XbaI. The PCR product was cleaned up and used for the double digestion with HindIII and XbaI; the vector pEAK8 was digested as well. The "sticky" end products of the digestion were used for the ligation reaction, resulting in the formation of the new vector pEAK8-126 (Fig. 5. 7).

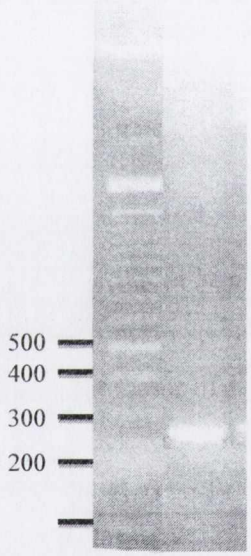


Fig. 5. 6. Agarose gel of pEAK8-126F and pEAK8-126R PCR product.

Coding region of BBD126 was amplified by PCR. A single PCR product of 280 bp was obtained and cloned into the pEAK8 expression vector.

One Shot TOP10 Chemically Competent *E. Coli* cells were chemically transformed with the ligation product. Cells were spread onto LB agar plates supplemented with 100 ug/ml of ampicillin. Only cells incorporating the plasmid were able to grow. In order to confirm the presence of BBD126 coding sequence in the plasmid and its correct reading frame, several colonies were grown overnight in 3 ml of LB broth supplemented with 100 ug/ml ampicillin. Plasmid DNA was extracted using a PureYiel Plasmid Miniprep (Promega, Madison, WI, USA). Isolated DNA was commercially sequenced. Once we confirmed the presence of

BBD126 coding sequence in a clone (Fig. 5. 8), this clone was grown overnight in 100 ml of LB broth supplemented with 100 ug/ml ampicillin. The resulting culture was used for plasmid isolation using PureYiel Plasmid Maxiprep (Promega, Madison, WI, USA). The resulting pEAK8-126 DNA concentration was determined using a Nanodrop spectrophotometer. HEK-293T cells were transfected with the pEAK-126 vector following the protocol described in Section 2.31. HEK293T cells were transfected with 6 μg of pEAK8-126 or pEAK8 vectors. Cells were grown for 72 hours. The recombinant BBD126 was expected to be secreted into the culture media. In order to confirm this, both cells and media were kept. Cells were retrieved from the flask, pelleted by centrifugation, and were lysed using protein analysis lysis buffer. Media was used for protein purification by nickel-agarose; pEAK8-126 incorporates a 6 histidine tag in the C terminus of the recombinant protein. Following the protocol in Section 2.24, 5 ml of tissue culture media was incubated with 250 μl of Ni-NTA beads for 1 hour at 4 ^{o}C with rotation. Preparation was then loaded into separation columns. The Ni-NTA beads were washed to eliminate non-binding proteins. Recombinant protein was eluded in a single fraction with 200 μl of NI-NTA elution buffer.

Protein concentration was determined by a Bradford assay for all samples. 20 μg of total protein were loaded into 4-12% SDS-PAGE gels. Samples were run in the gel and transferred into a PMSF membrane. The membrane was incubated with mouse monoclonal anti-BBD126 produced by GenScript (Piscataway, NJ, USA). The membrane was washed and incubated with an anti-mouse IgG 680LT antibody (Li-cor, Lincoln, NE, USA). Infrared fluorescence was measured using the Odyssey Infrared Imaging System (Li-cor, Lincoln, NE, USA) (Fig. 5. 9).

Anti-BBD126 antibody was then used to test the presence of the rBBD126 produced in HEK293T cells. The HEK293T cells were transfected with pEAK8-126 vector designed for the secretion of the peptide into the culture media. Cell lysate, culture media and culture media purified with Ni-NTA beads were run in a western-blot for antiBBD126 antibody (Fig. 5. 9). The antibody detected the presence of rBBD126 in the purified media of the pEAK8-126 transfected cells.



Fig. 5. 7. Map of pEAK8-126 plasmid.

BBD126 coding region, including signal peptide for the secretion of the peptide and mature peptide, were cloned into the pEAK8 vector. Protein production is regulated by the EF1-α promoter, generating a constitutive protein expression. The recombinant protein includes a His-Tag label for protein purification in the C terminus. pEAK8-126 vector is equipped with an ampicillin resistance gene.



Fig. 5. 8. DNA sequence and protein translation of pEAK8-126 vector.

Plasmid preparation of the transformed One Shot TOP10 Chemically Competent E. coli cells with the ligation resulting in the new VECTOR. pEAK8-126 were commercially sequenced to confirm presence of BBD126 coding region and correct reading frame inclusion in the vector.

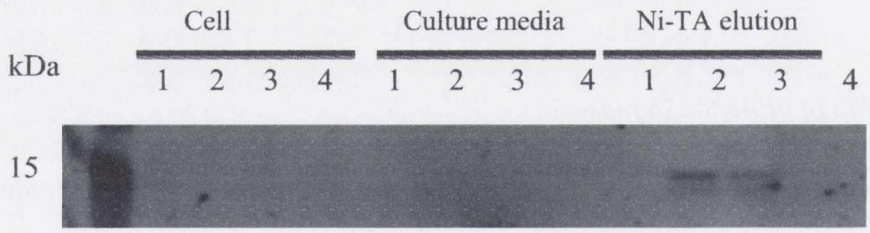


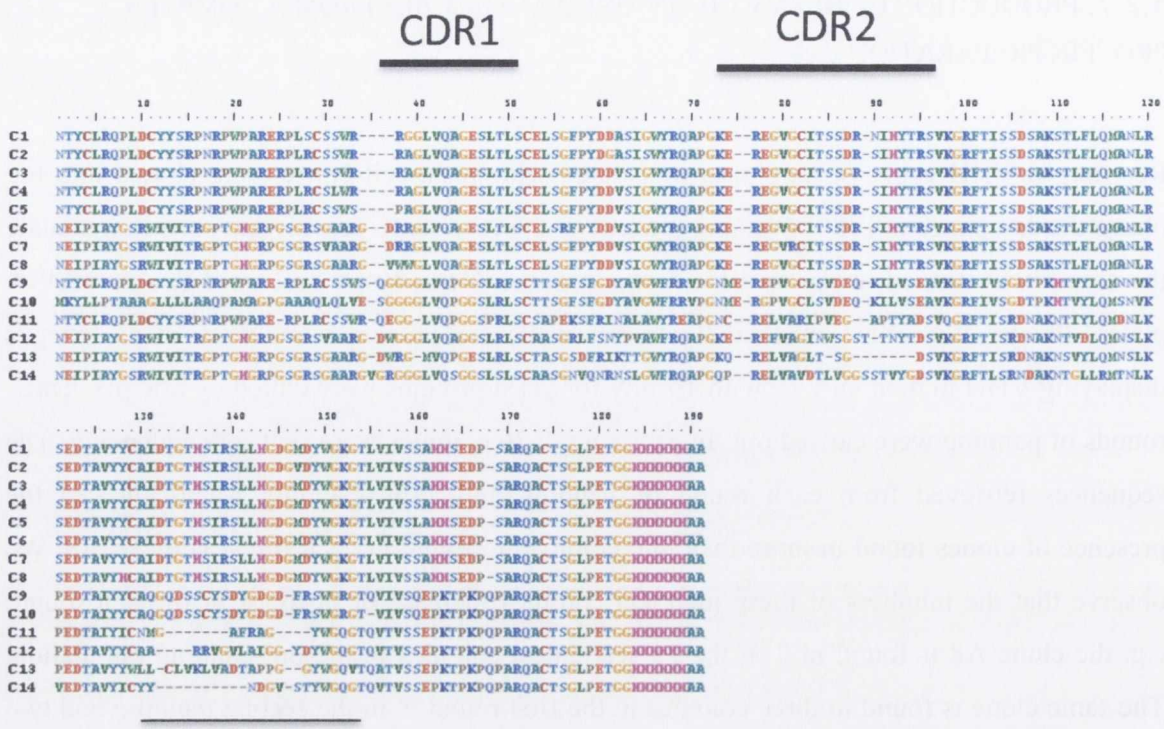
Fig. 5. 9. Western blot of rBBD126 with antiBBD126 antibody.

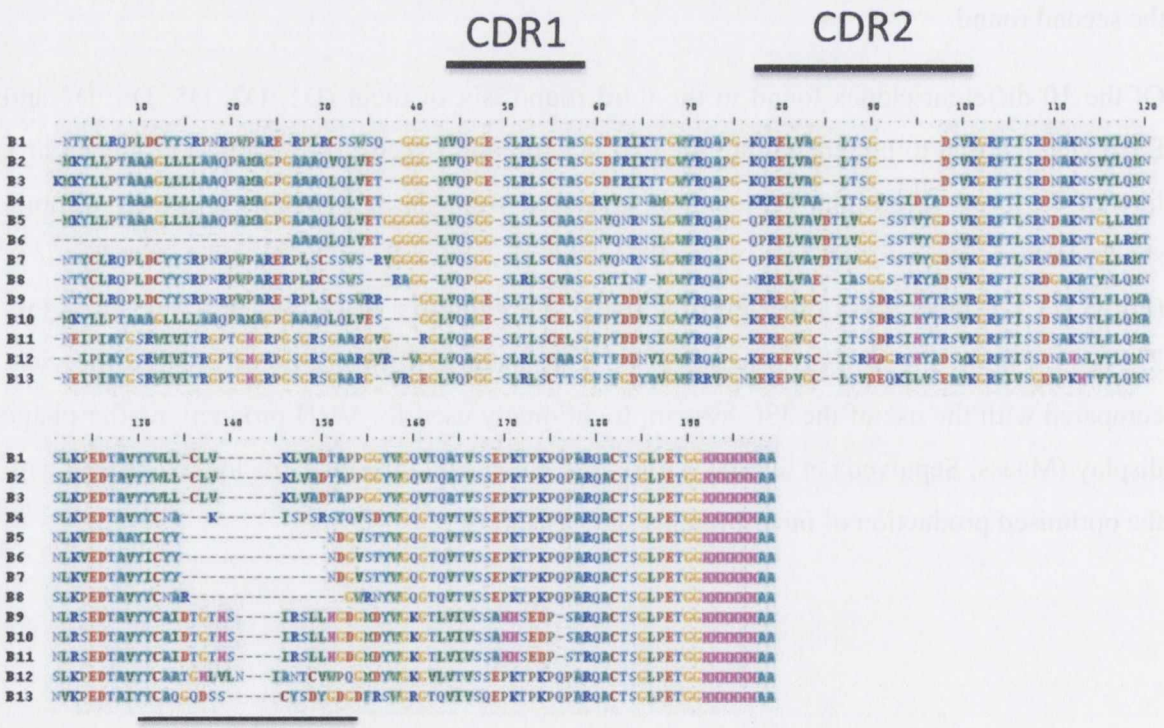
HEK293T cells lysate, culture media pre-Ni-NTA isolation and Ni-NTA elution were loaded into 4-12% SDS-PAGE for western-blot analysis with anti-BBD126 antibody. 1, empty cell without vector, 2 and 3, replicate transfection with pEAK8-126 and 4, cells transfected with empty vector pEAK8. Positive detection of rBBD126 was found in the Ni-NTA elution of culture media samples from pEAK8-126 transfected cell.

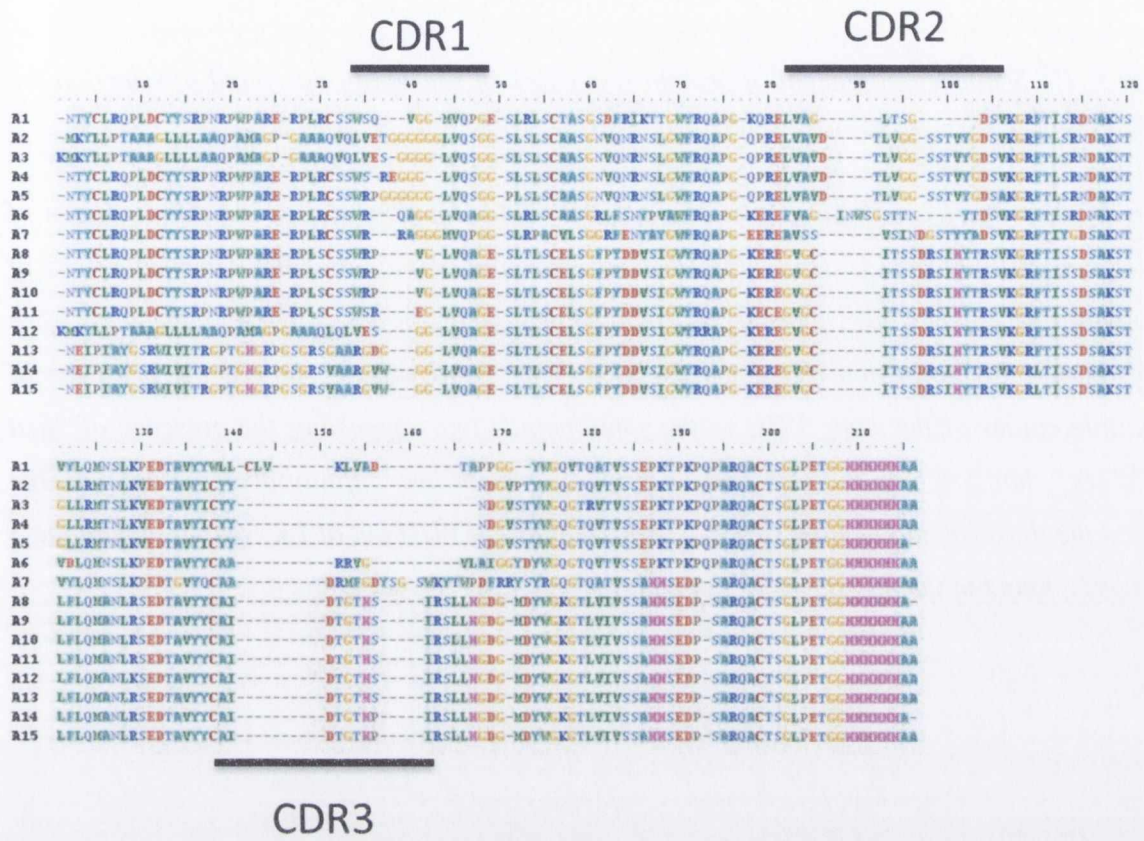
5.2.7. PRODUCTION OF HEAVY CHAIN-ONLY ANTIBODIES FROM A COMPLEX PROTEIN PREPARATION

Rearranged VHH genes of an alpaca inoculated with yeast cell lysate were cloned into the phage display phagemid JSC. Panning process against the same yeast lysate used to inoculate the alpaca was carried out in an ELISA-like process. Yeast lysate was used to coat ELISA plates. Phages were then incubated into the wells and after several wash steps, phages displaying VHH in their surface with affinity for yeast proteins were eluted by low pH. Three rounds of panning were carried out. In each round, 20 colonies were sent for sequencing. The sequences retrieved from each round of panning and a pre-panning round showed the presence of clones found in more than one colony and found in each round (Fig 5. 10.). We observe that the numbers of these identical colonies increase or decrease with each round, e.g. the clone A8 is found in 8 of the 15 sequenced colonies in the pre-panning population. The same clone is found in three colonies in the first round, 8 in the second panning and two in the third panning. The final third round presents more variability in the number of different colonies: ten compared with the five found in the pre-panning, seven in the first, and six in the second round.

Of the 10 different clones found in the third round, six of them (D1, D2, D5, D6, D7 and D10) were chosen to be cloned into the pHEN expression vector. Protein was produced in 1 litre of TB broth overnight. Protein was extracted from the periplasmic space by osmotic shock. Protein was purified first by Ni-NTA and was eluted from the Ni-NTA in 5 fractions (Fig. 5. 11.). Protein was further purified by gel filtration in a FPLC system (Appendix: Fig. 2; Fig. 3; Fig. 4; Fig. 5; Fig. 6 and Fig. 7). This method of production increased the yield compared with the use of the JSC system, traditionally used for VHH production after phage display (Maass, Sepulveda et al. 2007) but protein concentration remains low compared with the optimised production of other proteins such as sortase.


CDR1
CDR2
CDR3


CDR1
CDR2
CDR3


CDR1
CDR2
CDR3

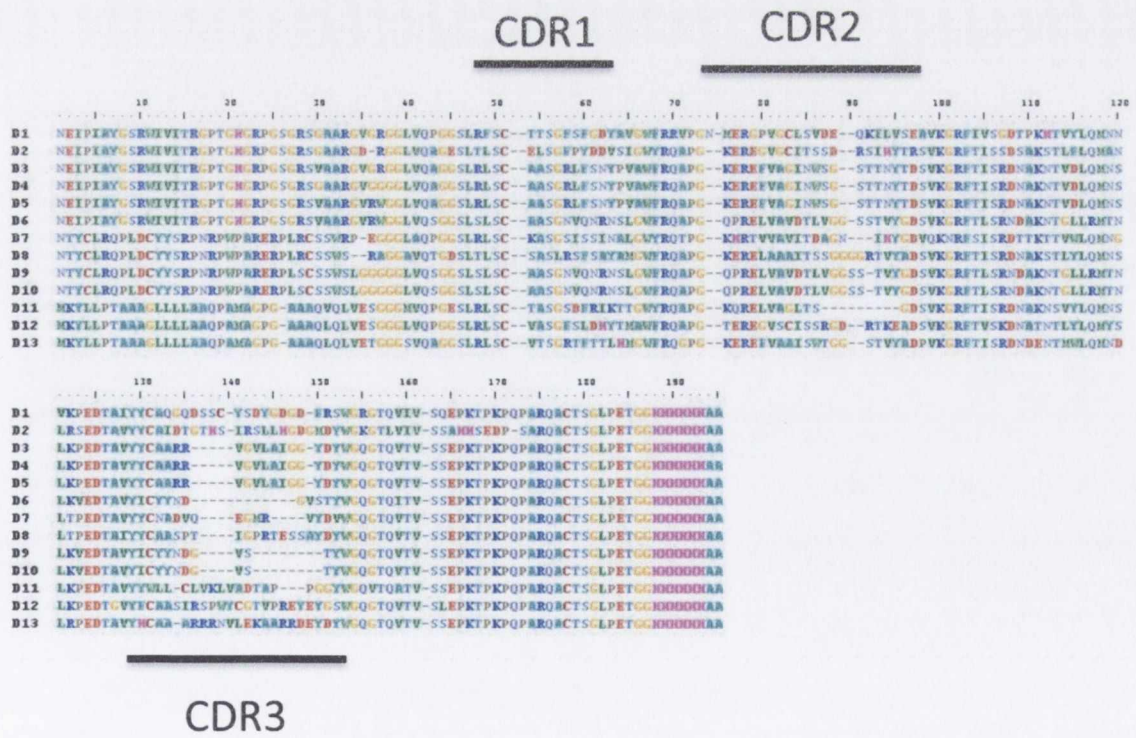

CDR1
CDR2
CDR3

Fig. 5. 10. Sequence alignment of protein sequences of pre- and post-round of selection for yeast library.

Twenty colonies from the pre-selection population, A, and the successive three rounds of selection, B, first round of panning, C, second round of panning and D, third round of panning, were sent for commercial sequencing. The retrieved DNA sequences were translated into the protein sequence encoded. This small sample showed the presence of multiple copies of the same VHH in the same population, suggesting the presence of "fast growers" that can bias the selection in terms of growth over affinity over the antigen. We found the third round the most variable with 10 different VHH out of 13, compared with the 5 out of 15 found in the pre-selection population.

kDa

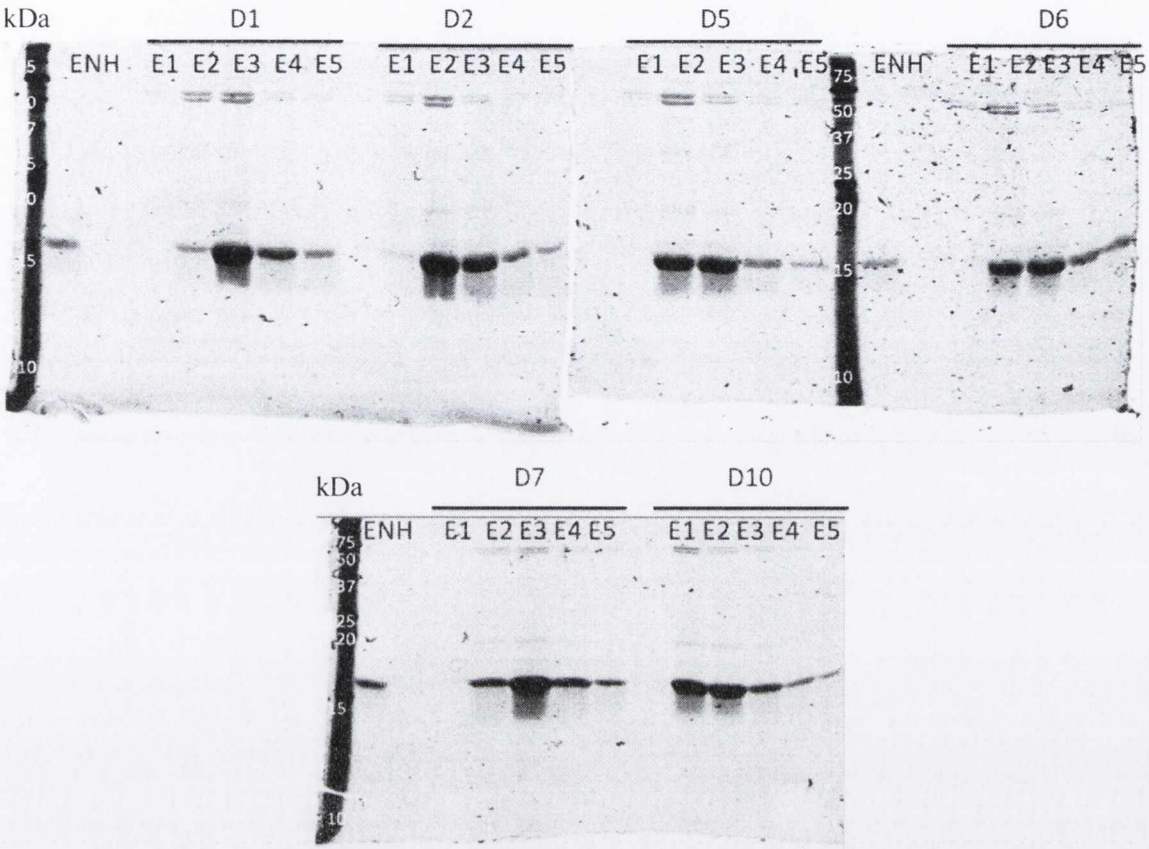


Fig. 5. 11. Protein production of yeast library selected VHH from the third round of selection.

Elution fractions of yeast specific clones D1, D2, D5, D6, D7 and D10 were run in 4-12% SDS-PAGE gel, and stained with Coomassie Brilliant Blue stain to confirm the production of VHH. A 15 kDa band, matching with predicted size of VHH, could be found in some of the fraction of all clones. E3 and E4 for D1; E2 and E3 for D2; E2 and E3 for D5; E2 and E3 for D6; E3 and E4 for D7 and E1 and E2 for D10 were individually pooled for each clone and used for further purification by FPLC.

5.2.8. PRODUCTION OF SORTASE

Overnight culture cell lysate of pET30-sortase A vector (gift of Prof. Hidde Ploegh, Whitehead Institute, MIT, MA, USA) transfected BL-21 *E.coli* cells (Stratagene Corp, La Jolla, CA, U.S.A.) were used for protein purification by Ni-NTA. Sortase was eluted from the Ni-NTA beads into five fractions. 10 µl of these fractions were run in 4-12% SDS-PAGE gels and stained with Coomassie Brilliant Blue stain (Fig. 5. 12.). Fractions E2, E3 and E4 were pooled and further purified in a gel filtration FPLC system (Fig. 5. 13). Fractions 46, 47, 48, 49 and 50 were collected and pooled. The sortase preparation was then stored in 200 µl aliquots at -80°C until use.

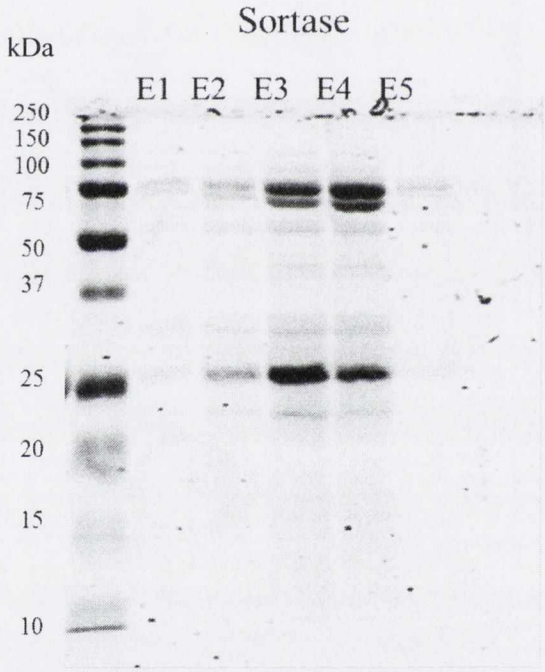


Fig. 5. 12. Protein production of sortase enzyme.

Ni-NTA elution fractions from cell lysate transfected with pET30-sortase vector were run in a 4-12% SDS-PAGE and stained with Coomassie Brilliant Blue stain. Sortase, 25 kDa, was found in most fractions. Fractions E2, E3 and E4 were pooled for further FPLC purification.

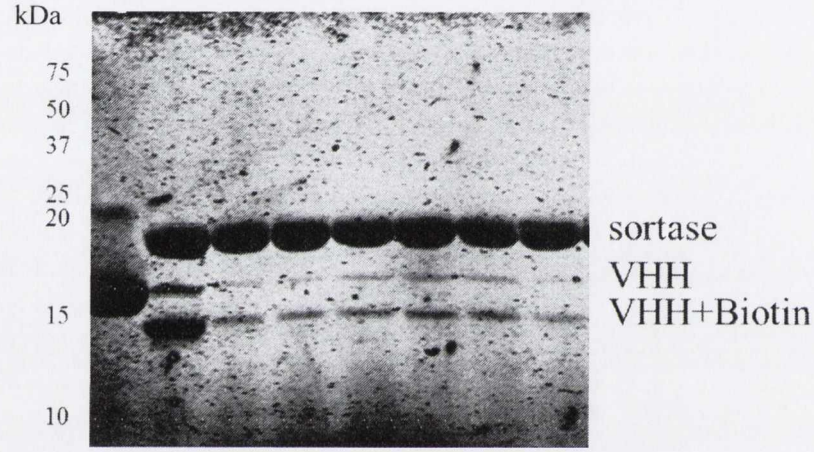


Fig. 5. 13 Sortase purification by gel filtration column in a FPLC system.

Pooled fractions of the Ni-NTA purification were injected into a Hiprep 16/60 Sephacryl S100HR column (GE Healthcare, Little Chalfont, UK). The trace showed a peak of 280 nm absorbance, proportional to protein concentration, at fractions 46 to 50. These factions were collected, pooled and stored at -80°C until needed.

5.2.9. SORTASE LABELLING OF YEAST LIBRARY VHH

Purified D1, D2, D5, D6, D7 and D10 VHH proteins were labelled with biotin, using the sortase transpeptidase enzyme and G3-biotin. Enhancer VHH (ENH), a previously described VHH with affinity for green fluorescence protein and gift of Prof. Hidde Ploegh, was used as positive control for the sortase reaction (Fig. 5. 14.). The reaction was cleaned of non-reactive substrates by the addition of Ni-NTA. Non-reactive VHH and sortase conserved the His-Tag. However, VHH labelled with biotin lose their His-Tag during the transpeptidation, eliminating their affinity for Ni-NTA.



	Enh	Enh Sort	D1 Sort	D2 Sort	D5 Sort	D6 Sort	D7 Sort	D10 Sort
Sort	-	+	+	+	+	+	+	+
G3 Biot	+	+	+	+	+	+	+	+

Fig. 5. 14. Sortase labelling of yeast specific VHH with biotin.

Labelling of VHH clones was carried out with the help of the sortase transpeptidase. Sortase enzyme recognises the LPETG motif found in the C terminus of the VHH, breaks the bond between T and G amino acids and creates a new bound between the T and the G present in the G-G-G-biotin label. In addition, this reaction removes several amino acids including the His-Tag from the VHH, generating a new protein of lower molecular weight. This can be observed as a new lower band in the gel. ENH VHH was used as positive control of the reaction.

5.2.10. CHARACTERISATION OF SPECIFICITY OF YEAST LIBRARY VHH BY IMMUNOPRECIPITATION.

Biotin labelled VHHs were then used for immunoprecipitation with a yeast lysate preparation. VHH-biotin preparation was incubated with streptavidin agarose beads and the VHH-biotin-streptavidin agarose were then incubated with 100 μg of yeast lysate. After several washing steps, the beads were boiled and the sample was run in a 4-12% SDS-PAGE gel and the gel was then silver stained (Fig 5. 15.). As negative control, biotinylated ENH was used to measure non-specific binding of yeast proteins to the Streptavidin agarose beads. No specific bands were found equivalent to any of the clones studied, suggesting a lack of affinity to any yeast proteins.

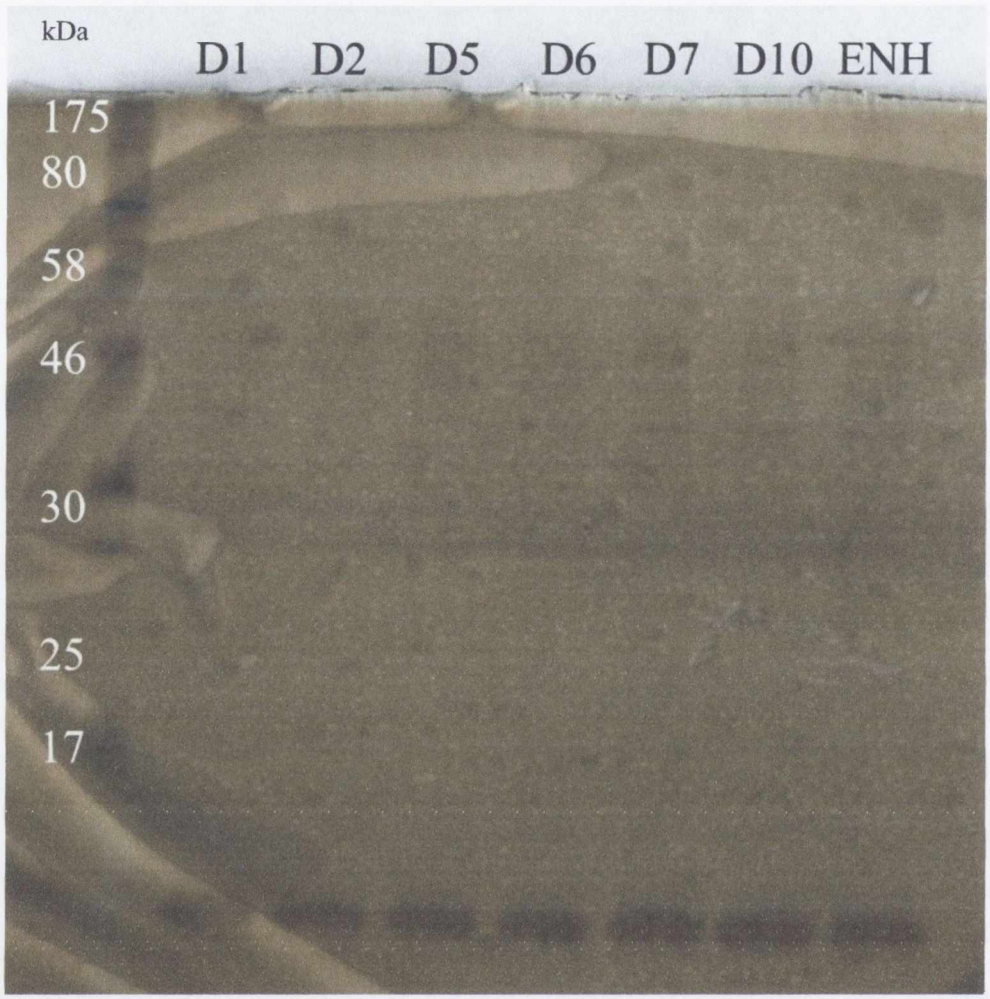


Fig. 5. 15. Immunoprecipitation of yeast specific VHH with yeast lysate.

Biotin labelled, D1, D2, D5, D6, D7 and D10, were covalently linked to streptavidin agarose beads. The VHH-biotin-streptavidin agarose beads complex was incubated with yeast lysate. After several washes, it was boiled and run in a 4-12% SDS-PAGE gel. Protein was then visualised by silver staining. No specific band for any of the clones was found. ENH was used as negative control as unspecific binder.

5.2.11. PCR GENERATION OF A HEAVY CHAIN ANTIBODY LIBRARY.

Using the sequences obtained from the yeast lysate library, a consensus sequence was built (Fig 5. 16.). This sequence was then used to create a series of primers for a primer walking reaction. The primers were designed so that the DNA sequence coding for the highly variable regions, CDR1, CDR2 and CDR3, were random nucleotides. In addition, we incorporated several primers coding for the CDR3 region, P5R, with different lengths; 12 primers with 3 nucleotides increasing from 27 up to 57 nucleotides. These primers were used for the following PCR reactions: P1F-P2R, P1F-P3R, P1F-P4R, P1F-P5R and P1F-P6R; the sample of one reaction was the product of the previous reaction. With each reaction, a larger sequence was generated (Fig. 5. 17).

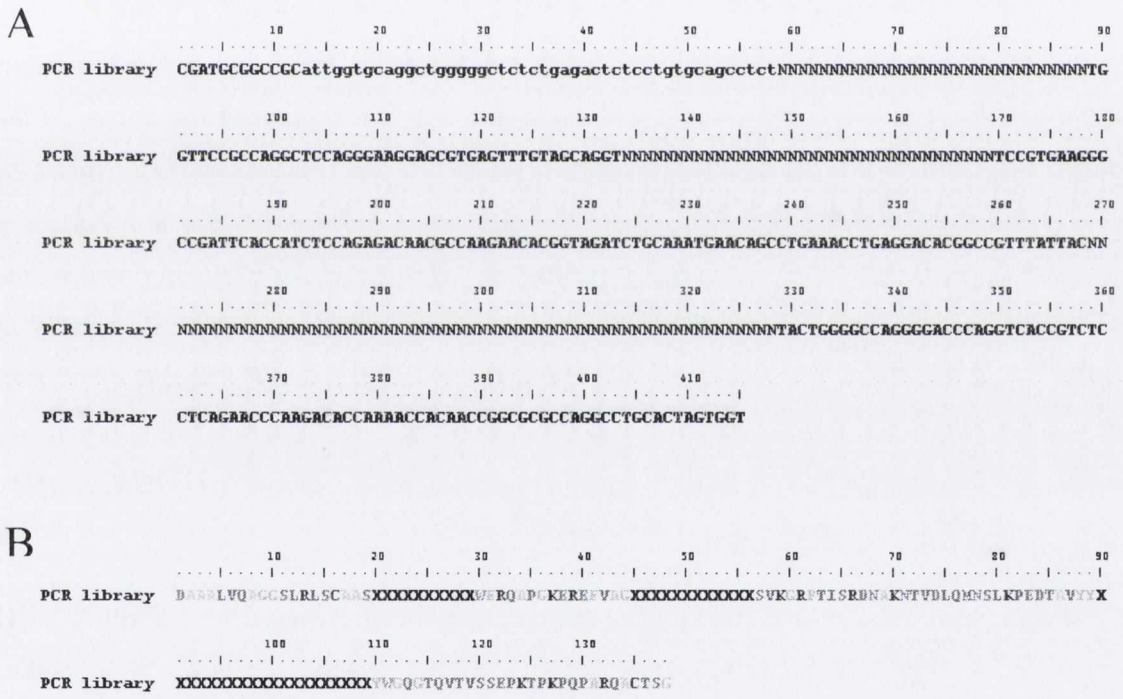


Fig. 5. 16. Consensus sequence for VHH sequence alignment.

VHH sequences from yeast library were aligned, using Bioedit software. Bioedit software was used to create a consensus sequence of a hypothetical VHH. This sequence was used as a template for the design of the primers for the construction of the PCR generated library. A, DNA sequence of the PCR consensus sequence of VHH. CDR regions have been labelled as

random nucleotides, N; B, protein sequence of the consensus sequence of VHH. CDR regions are labelled as X.

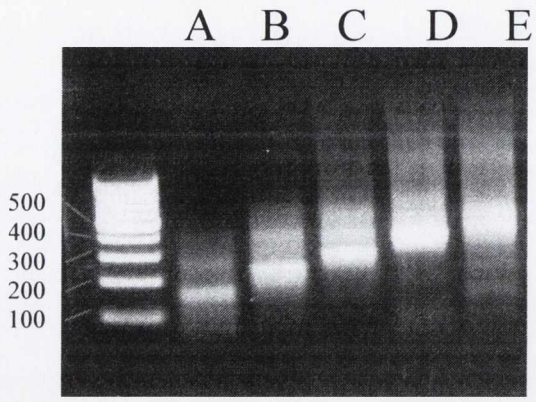


Fig. 5. 17. PCR generated VHH library construction.

PCR library was created by succession of PCR reactions using the product of the previous reaction as template for the next one, in which each reaction increased the length of the product until the complete product was built. The first reaction product, using P1F and P2R primers, was used as template for the second reaction with P1F and P3R. The products of this reaction was then used as template for the P1F and P4R reaction. Next reaction used the P1F-P4R product as template for the PCR reaction using P1F and P5R, a mix of P5R of different lengths in order to duplicate the variability of CDR3 length. Finally, P1F-P6R reaction used the P1F-P5R as template. The products of each reaction were run in an agarose gel showing the increment of length with each PCR reaction. A, P1F-P2R; B, P1F-P3R; C, P1F-P4R; D, P1F-P5R and E, P1F-P6R.

P1F-P6R product was isolated and cloned into the JSC vector creating the JSC-PCR-VHH vector (Fig. 5. 18.). PCR-VHH library was then used for panning against bovine TNFα. ELISA plates were coated with bovine TNFα. Two rounds of panning were carried out. In each round, 10 colonies were selected and sent for sequencing. In the first round, 7 sequences were obtained; in the second round, only 2 sequences were obtained (Fig. 5. 19).

B1 and B2 clones were amplified by PCR and cloned into a pHEN vector, creating the new pHEN-B1 and pHEN-B2. These vectors were used to transform WK6 cells. The resulting cells were used for protein production in 1 litre of TB broth. Osmotic shock was used to retrieve the VHH accumulated into periplasmic space in the WK6 cells. VHH was further

purified by Ni-NTA isolation and protein was eluted in 5 fractions (Fig. 5. 20). Due to low yield, no purification by FPLC was used. The sample was instead concentrated using a centrifugal concentrator, cut point for 3 kDa, (Merk Millipore, Billerica, MA, USA

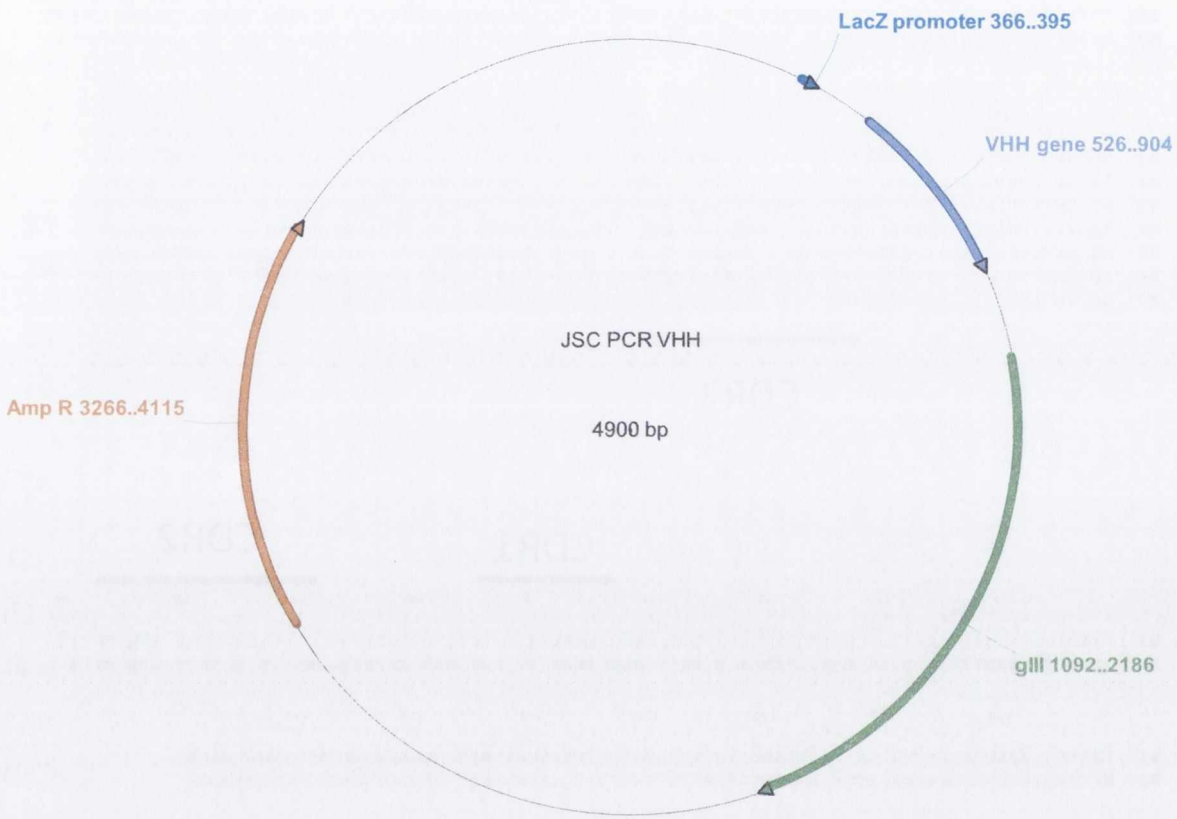


Fig. 5. 18. Map of JSC-PCR-VHH vector.

PCR generated VHH genes were cloned into the JSC phage display vector. VHH cloned into this vector will generate a fusion protein of VHH with gIII, phage surface protein, forming a VHH-pIII complex. gIII is one of the proteins forming the capsule of M13 phage. Bacteria transfected with the JSC-VHH vector, are not able to produce viral particles but if co-infected with M13 phage, the viral particles generated will display the VHH-gIII on their surface. In addition, the JSC-VHH will incorporate into the phage particle. A VHH-gIII displaying phage preserves the ability to infect bacteria but contains only the JSC-VHH vector as genetic material, preventing the formation of phage particles.

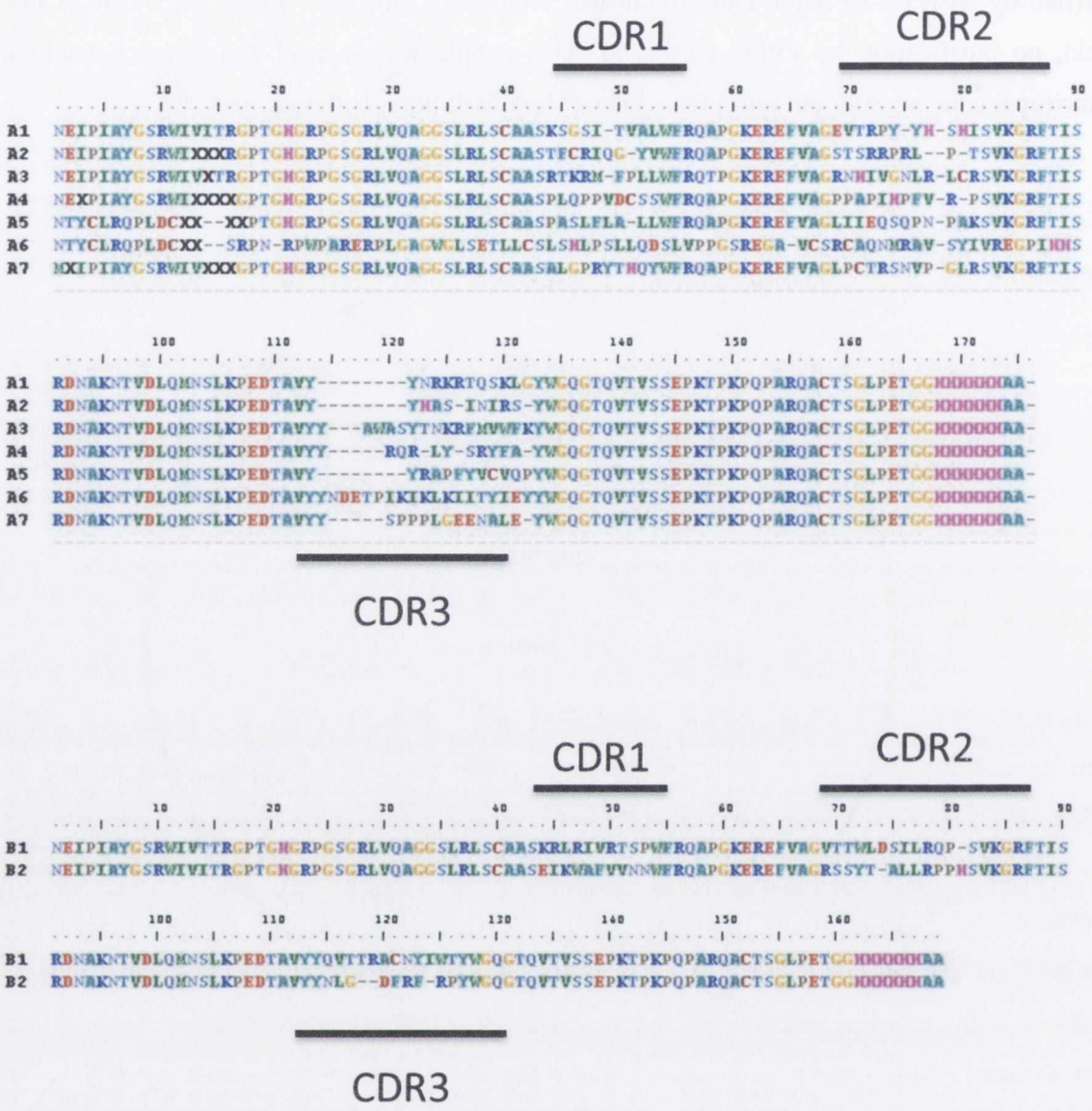


Fig. 5. 19. Sequence alignment of sequence of colonies for PCR-VHH library.

Ten colonies from each round of selection were selected and sequenced. The sequences were then aligned to determine differences between them. In the first round (A), seven different sequences were retrieved. In the second round (B), generate only two VHH sequences, which were used for protein production.

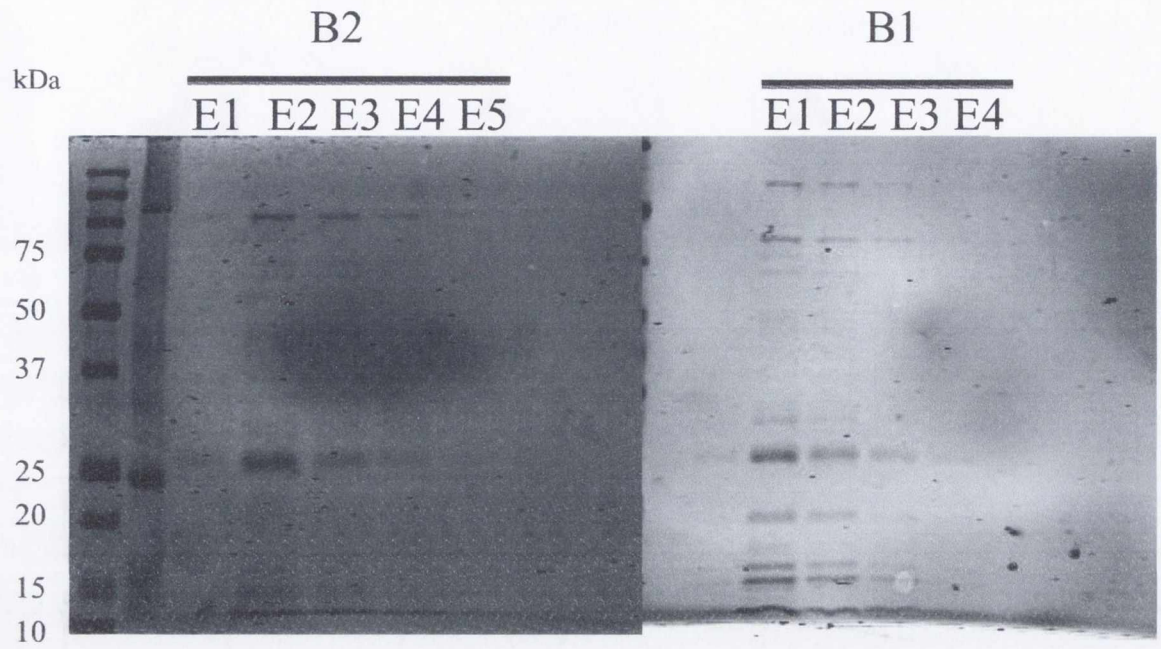


Fig. 5. 20. Protein production of PCR-VHH, B1 and B2.

JSC-B1 and JSC-B2 were cloned into the pHEN6 vector. The resulting pHEN-B1 and pHEN-B2 were transformed into WK6 cells for protein production. Protein was extracted by osmotic shock and was then purified by Ni-NTA affinity to His-Tag present in the recombinant VHH. Fractions from Ni-NTA elution for production of VHH B1 and B2 were run in a 4-12% SDS-PAGE gel stained with Coomassie Brilliant Blue stain. Due to low production yields, protein was not further purified by FPLC in order to avoid losing valuable VHH protein during the process.

5.2.12. Analysis of anti-bovine TNFa VHH by ELISA.

In order to test the ability of the VHH to detect bovine TNFα. B1, B2 and ENH VHHs were used to substitute the captured antibody in a bovine ELISA kit (Fig. 5. 21.). Wells with B1 changed colour when TNFα standard was added; no change was observed with the mouse IL4 standard. B2 and ENH induced no colour change with TNFα or IL4 which suggests that B1 was able to specifically bind to bovine TNFα.

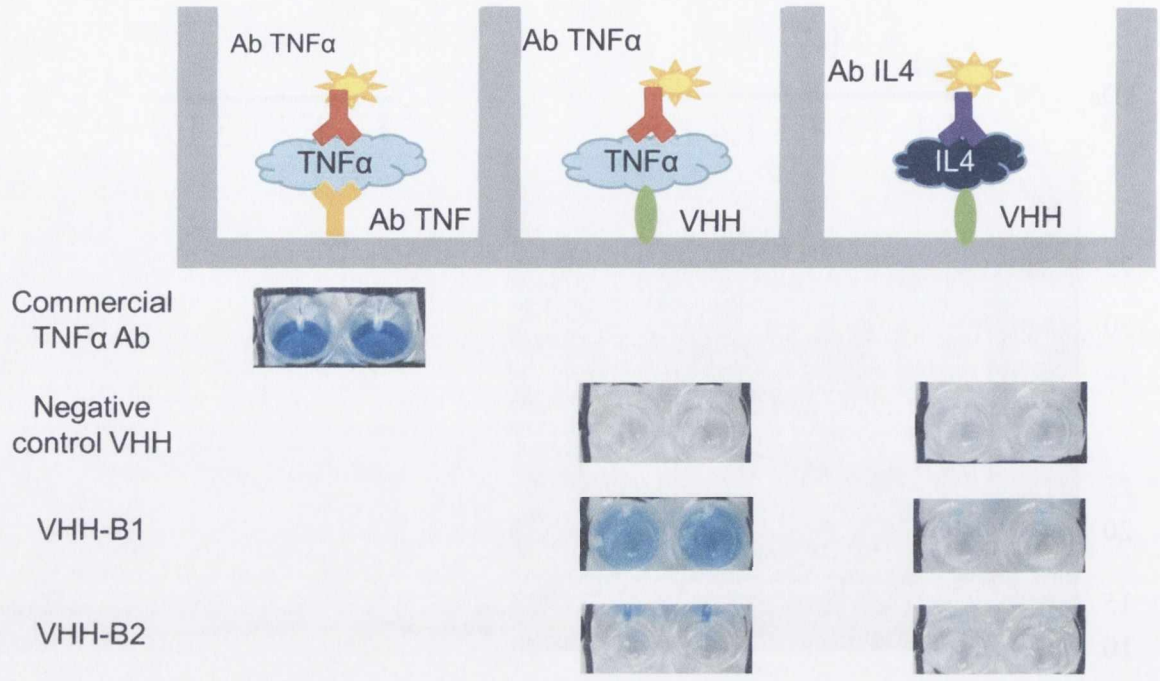


Fig. 5. 21. Use of B1 and B2 for TNFα ELISA

A, Schematic of loaded components in each well. For TNFα control, both capture and detection antibodies from the commercial kit were used. For TNFα-VHH, capture antibody was substituted by a VHH. TNFα substract was then added and detection TNFα antibody was used. In the IL4-VHH, capture antibody was substituted by a VHH. IL4 subtract was then added and detection IL4 antibody was used. B, readout of ELISA plate. Only the TNFα control and B1 for TNFα produced a colour change indicating the affinity of B1 for TNFα.

5.2.13. Characterisation of specificity of anti-bovine TNFa VHH by immunoprecipitation.

B1, B2 and ENH were biotinylated using a sortase reaction. The product of the reaction was used for immunoprecipitation of the antigen recognised by the VHH. Biotin-B1, biotin-B2 and biotin-ENH were incubated with streptavidin agarose. The complex form was then incubated with 5 ng of TNF-α standard. After several washes, the Streptavidin agarose-VHH complex was boiled and run in a 4-12% SDS-PAGE gel. The gel was then silver stained (Fig. 5. 22). A band of 17 kDa was found in the B1 lane, not present with B2 or ENH (recombinant bovine TNFα has a size of 17.3 kDa). The band was cut and sent for identification by mass spectrometry at Prof Hidde Ploegh's lab, MIT, MA, USA. The results suggest the band is bovine histone H4 (11.1 kDa). In order to confirm this result and to investigate the possibility of cross reactivity with human TNFα, the immunoprecipitation was repeated using human TNFα (Fig 5. 23). The same band of the band size was found with B1. Finally, human serum was used as sample for the immunoprecipitation. 30 µl of human serum were added to the VHH-streptavidin agarose complex and a similar band as previously described was found for B1 (Fig. 5. 24).

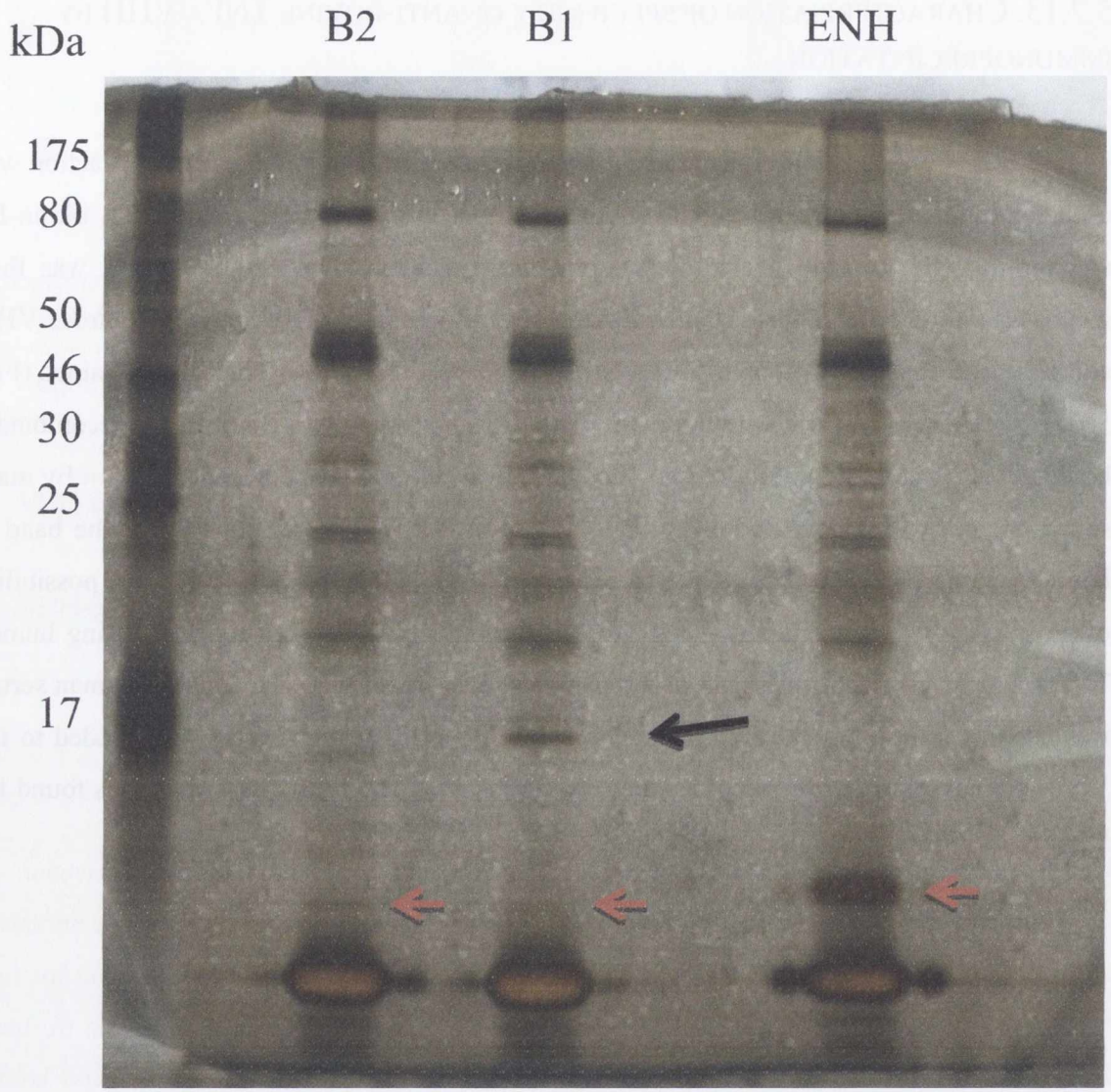


Fig. 5. 22. Immunoprecipitation of B1 and B2 antigens from a bovine TNFα standard.

Biotinylated B1, B2 and ENH were linked to streptavidin agarose beads. The VHH-biotin-streptavidin agarose complex was then incubated with bovine TNFα standard. After removing non-specific binding with several washes, antigens were eluted by denaturing at 95°C for 5 minutes. The elution sample was then run in a 4-12% SDS-PAGE gel and proteins were visualised by silver staining. B1 presented a specific band of 17kDa. Red arrows indicate the precipitated VHH fragments.

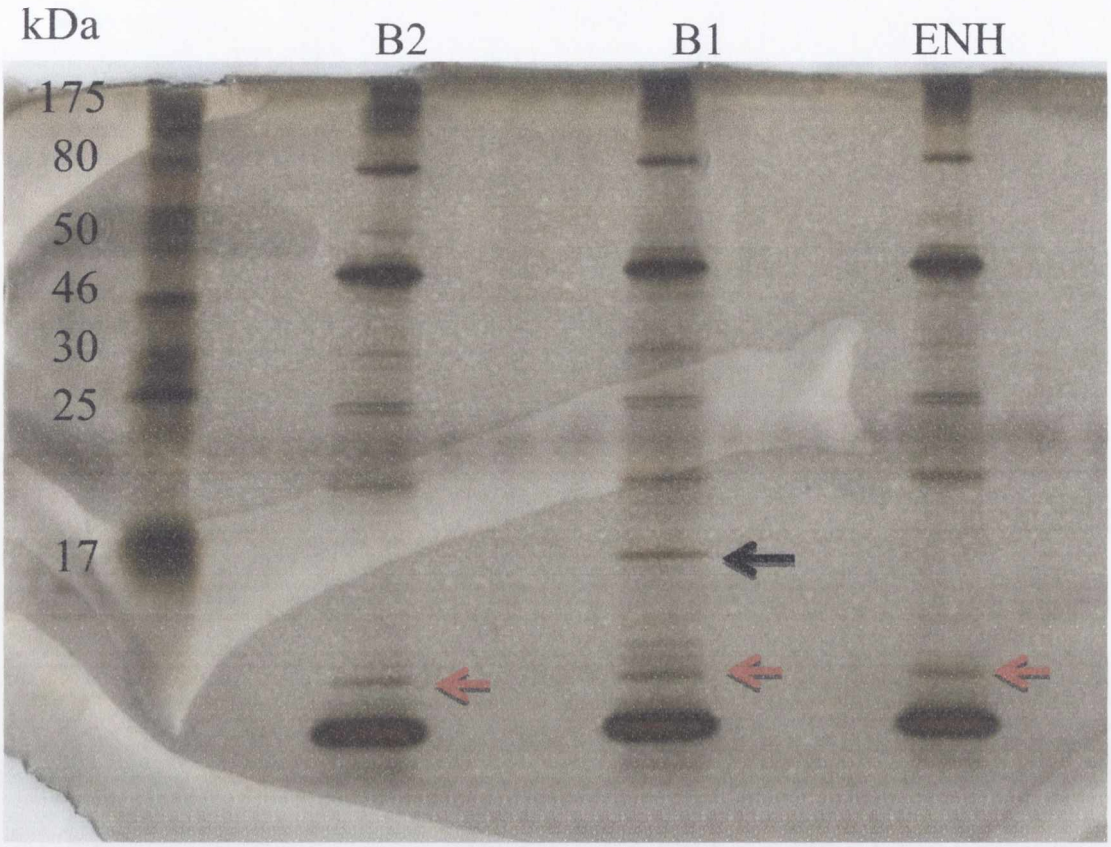


Fig. 5. 23. Immunoprecipitation of B1 and B2 antigens from human TNFα standard.

Biotinylated B1, B2 and ENH were linked to streptavidin agarose beads. The VHH-biotin-streptavidin agarose complex was then incubated with human TNFα standard. After removing non-specific binding with several washes, antigens were eluted by denaturing at 95°C for 5 minutes. The elution sample was then run in a 4-12% SDS-PAGE gel and proteins were visualised by silver staining. B1 presented a specific band of 17 kDa. This band was isolated and sent for mass spectrometry identification. Red arrows indicate the precipitated VHH fragments.

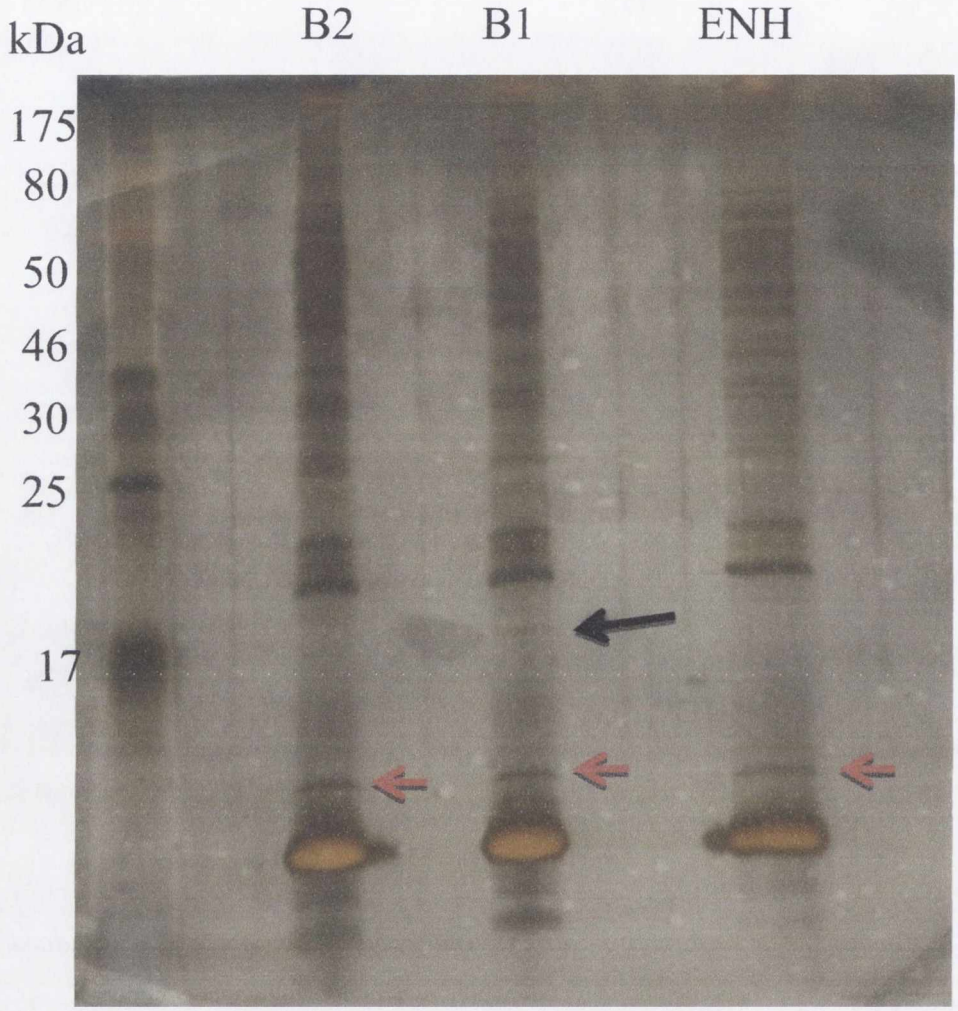


Fig. 5. 24. Immunoprecipitation of B1 and B2 antigens from human serum.

Biotinylated B1, B2 and ENH were linked to streptavidin agarose beads. The VHH-biotin-streptavidin agarose complex was then incubated with human serum. After removing non-specific binding with several washes, antigens were eluted by denaturing at 95°C for 5 minutes. The elution sample was then run in a 4-12% SDS-PAGE gel and proteins were visualised by silver staining. B1 presented a specific band of 17kDa. This band was isolated and used for mass spectrometry identification. Red arrows indicate the precipitated VHH fragments.

5.3. Discussion

We have confirmed exclusive gene expression of the Chr13 β-defensin gene cluster in the reproductive tract and the presence of SNPs in Chr13 β-defensin genes correlated with lower sperm penetration. These results suggest a different role to the classic antimicrobial function, at least for some of these β-defensins. To confirm this hypothesis, functional analysis of these defensins is required. However we lack the tools to carry this work. To face this problem, we decided to build some of these tools.

Our first target was the synthesis of a recombinant bovine Chr13 β-defensin. Several approaches can be taken for this purpose. Prokaryotic expression systems have been previously used for β-defensin protein synthesis (Bruhn, Regenhard et al. 2007, Huang, Leong et al. 2009). This provides a bigger production yield compared with other systems. However, it has several disadvantages. Most important is the lack of post-translational modifications, due to the lack of the molecular machinery such as glycosylation or protein folding machinery which is found in eukaryotes, but not in prokaryotes. β-Defensins are particularly sensitive to the effects of mis-folding. β-Defensins possess the defensin motif formed by three β-sheets held together with 3 di-sulphide bonds. The pattern of the disulphide bond formation is highly conserved in β-defensins C1-C5, C2-C4 and C3-C6 and different from the bridges in α-defensins. The incorrect folding of the protein will reveal the hydrophobic segments of the β-defensin, inducing the precipitation of the peptide in an aqueous solution and the formation of inclusion bodies. This issue can be solved by creating fusion proteins of β-defensins with other components in order to avoid the precipitation (Huang, Leong et al. 2009). This has not been an issue for the production of β-defensins as antimicrobial peptides. Incorrectly folded β-defensins are able to maintain their antimicrobial potency. Most likely, the antimicrobial role is derived from the physical properties of the peptides, such as electrostatic charge, hydrophobicity and amphipathicity which allowed the β-defensins interact with the prokaryotic membrane. Other β-defensin roles such as chemotaxis (Biragyn, Ruffini et al. 2002), chemokinesis (Soruri, Grigat et al. 2007) or fertility (Yudin, Generao et al. 2005, Tollner, Yudin et al. 2008, Tollner, Venners et al. 2011, Tollner, Bevins et al. 2012), involve protein-protein interactions and receptors, which require a correct three dimensional structure. In addition, glycosylation has been identified as essential for β-defensin role in sperm maturation.

An alternative system for protein production is eukaryotic expression systems. These are based on the incorporation of a plasmid that includes elements such as promoter regions, and signal peptides for the production of the recombinant protein into a eukaryotic cell. Eukaryotic cells possess all the machinery required for the correct protein folding and incorporate the enzyme required for glycosylation of proteins. However these systems have some disadvantages. Firstly, eukaryotic expression systems are more complex and are more laborious but more importantly, protein production yield can be considerably lower than that generated by prokaryotic systems. We considered that eukaryotic systems created a more physiological peptide that could help to understand the role of β-defensins in the reproductive tract.

Among the 19 different genes found in bovine Chr13 β-defensin gene cluster, BBD126 was considered the most relevant target for functional work, based on previous work done in macaques and humans (Yudin, Generao et al. 2005, Tollner, Yudin et al. 2008, Yudin, Tollner et al. 2008, Tollner, Venners et al. 2011). For this purpose, the coding region of BBD126 was amplified by PCR and cloned into a pEAK8 vector in order to create pEAK8-126 plasmid. The pEAK8 vector incorporates the human elongation factor 1 alpha (EF1α) promoter region that delivers a constitutive expression of the protein incorporated into the vector. This vector induces transient transfection. The incorporation of BBD126 signal peptide signaled the peptide for secretion out the cell into the media. In addition, the recombinant protein will incorporate a His-Tag in the C-terminus. Using monoclonal antiBBD126, recombinant BBD126 (rBBD126) could not be found by western blot in the cell lysate of transfected cells or in the culture media. However, when the media was incubated with Ni-NTA beads, the monoclonal anti-BBD126 was able to detect a 15 kDa protein in the eluted fractions. The size of detected protein surpasses the predicted 8.3 kDa for the rBBD126 protein. The difference in size could be explained by glycosylation of the protein by the HEK293T cells.

The second target was the synthesis of a specific antibody for some of these β-defensins. An important limitation for proteomic work in the bovine field is the lack of reagents, in particular antibodies specific for bovine proteins. For this reason, we requested the commercial synthesis of a monoclonal mouse antibody for BBD126. The antibody was created by the inoculation of several mice with a KLH conjugated 13 amino acid fragment of BBD126. Of the four clones resulting of hybridoma fusion, two clones, A and D, showed promising results by western-blot analysis of seminal cells and seminal plasma. A 25 kDa

band was detected in the sperm cell lysate sample. As described before, the BBD126 sequence contains several possible targets for protein glycosylation. This glycosylation may cause the increment in the peptide weight that can be observed in the western-blot. This effect has been observed in the homologous macaque protein, DEFB126, where the predicted 14 kDa protein is detected at a 36 kDa band by western-blot (Yudin, Generao et al. 2005). The antibody was used to detect the presence of rBBD126 in the media of transfected HEK293T cells with pEAK8-126, suggesting the ability of the antiBBD126 to detect the denatured BBD126. Our next step was to test the antibody's ability to detect the correctly folded BBD126. Initially, we used the antibody to determine the location of BBD126 in sperm cells by immunocytofluorescence of fixed sperm cells. The result showed the presence of BBD126 in the tail and head of the sperm; human DEFB126 presents similar distribution (Tollner, Venners et al. 2011). Using immunohistochemistry on fixed sections of bovine epididymis and testis, we found strong presence of BBD126 in the epithelium and lumen of the cauda of the epididymis with no presence on the testis. This matched the results obtained by qPCR, whereby gene expression was found in the cauda of the epididymis. The final test for the antibody was the identification of the antigen by mass spectrometry. The mass spectrometry results indicate bovine spermadhesins z13 as the antigen recognised by the anti-BBD126 antibody. The precipitation of this protein could be a result of unspecific binding and more testing is required. Moreover, the real antigen could be masked by the antibody light chain band, particularly if we consider that they have similar molecular weight. A method to solve this problem would be to run the sample in a 2D gel that would separate the proteins not just by size but also by isoelectric point. The antibody was able to detect the presence of rBBD126 in the media culture of the transfected cell and a 25 kDa band in the cell lysate, both by western-blot. The immune localisation of the defensin in the sperm cell and tissue section matched the expected distribution of BBD126, considering the macaque and human homologous. In contrast, the mass spectrometry suggests, as the antigen recognised by the antibody, a 15 kDa protein, differing from the 25 kDa protein detected by western-blot. The antibody was synthesised by immunisation using a short peptide. Antibodies produced in this manner are used generally for de-nature detection of the antigen where the antigen is accessible, such as SDS-PAGE western-blot. Only in some instances if the antigen is present in the protein surface, can the antibody be used to detect the native protein. It is possible that this antibody is able to detect the denatured BBD126, as shown by the western-blot, and to detect a different protein in a native sample. All tests to corroborate the specificity of the antibody should incorporate a competitive control using recombinant protein. If the binding

of the antigen is reduced in the presence of soluble recombinant BBD126, that will corroborate the specificity of the antibody for BBD126.

The synthesis of monoclonal antibodies is a long and costly process. The cost of generating antibodies for all the β-defensins found in Chr13 cluster would be prohibitive. An alternative to the classic method of antibody generation was the use of camelid VHH. VHH presents a series of advantages over classic antibodies. The most important for our purposes is the lack of a light chain allowing the rearranged IgG gene to be more manageable for molecular biology than if both light and heavy chains are required. As a result of this, in contrast with mouse or rabbit models, where the antibody producing cell has to be isolated, and cell is immortalised using hybridoma technology. The VHH model only required the isolation of the rearranged IgG genes. These can be cloned into different expression vectors to be produced. Our hypothesis was that it is possible to immunise a single animal with a mix of proteins and obtain an immune response to all the proteins. This approach could significantly reduce the cost and time required to obtain antibodies for a large group of proteins. As proof of concept, we inoculated two alpacas with two different mixes of proteins: the first was inoculated with a yeast cell lysate and the second one, inoculated with Jurkat cells lysate. The work carried out for this thesis focused on the yeast lysate library. Using phage display technology, we created a library of phages displaying the VHH in their surface. Three rounds of selection by panning were carried out to select for yeast protein specific VHH. In the population analysed before the panning process, we can observe the enrichment of the population in a particular clone caused by a disproportionate presence in the original blood sample or a competitive advantage over other VHHs. After three rounds of selection, the VHH population seems stable. The same clones appear several times in all rounds of selection. Finally, the third round presents somewhat more variability in the number of different VHH clones. The six clones selected from the third round of panning were synthesised and used to immunoprecipitate the possible antigen from a yeast lysate sample. The immunoprecipitation showed no specific antigen captured by any of the VHH. This could be caused by an excess of non-specific binders in the population with affinity for the matrix in which the antigen is fixed and/or due to the complexity of the protein mix, the amount of the possible antigen present is under the limit of detection by the immunoprecipitation. A more convenient approach would be to carry out a selection process for a single protein sample present in the yeast cell lysate. This addresses the problem of low antigen amounts in the tested sample.

However, the use of phage display technology presents several issues. The most important is the presence of the so-called "fast growers". The phage display technology used bacteria transformed by a vector containing the coding sequence for a VHH. Expression of the VHH in this vector is regulated by a lacZ operon. The basis of this regulatory system is a protein that blocks the expression of the genes regulated by the operon unless a specific signal is given; for the lacZ, generally isopropyl β-D-1-thiogalactopyranoside (IPTG) is used. IPTG binds to the lac repressor and removes it from the operon, effectively inducing protein production. Under normal circumstances the transformed cells are grown in suppressor conditions so that VHH is not expressed. However, under the extremely competitive environment residual production or production of a VHH that introduces disadvantages to the cell growth creates disequilibrium in the cell population, introducing a bias in the selection process towards fast growers, not for the higher affinity binders.

During the work carried out with the VHH antibodies, we developed a new method for VHH library construction. We designed a series of primers to build a gene by PCR walking but with the addition of random nucleotides in the regions coding CDR1, CDR2 and CDR3. During primer synthesis, each individual nucleotide in the sequence is added individually. One by one, we requested all four nucleotides to be added into the primers; only one will occupy the position. This is simultaneously done in every sequence for a particular primer, introducing a huge variability in the possible primers. A total of 126 positions were added, with a theoretical $7.2e^{+75}$ possible VHH specificities. In addition, in order to simulate difference in CDR3 length observed in other libraries, ten different primers with different length were designed as well. The PCR generated library was cloned into the phage display vector, and as suggested before, a simple antigen mix, bovine TNFα, was used for two rounds of selection. In contrast with the sequences from the yeast lysate library, no VHH was found twice in the population, suggesting a more variable population. In the second round, only two sequences out of ten selected clones were obtained. It is possible that the random approach generates VHH incompatible with the phage display system. The two VHH isolated in the second round were cloned into a vector for protein production. B1 and B2 were then tested, first by ELISA. B1 was able to act as the capture antibody for a bovine TNFα ELISA kit. No reaction was found with B2 or negative control, ENH. The B1 was negative when tested with mouse IL4. Biotinylated B1, B2 and ENH by sortase reaction were used to immunoprecipitate the antigen recognised by the VHH. When incubated with bovine standard TNFα, B2 precipitated a 17 kDa protein; Recombinant bovine TNFα is 17.6 kDa. The band

was sent for mass spectrometry identification. The mass spectrometry identified the band as bovine Histone H4 (11 kDa). In order to test that the prediction was not an issue with a poor bovine protein annotation library or a contamination due the presence of Histone H4 in the sample and to test possible cross reaction with human TNFα, we repeated the immunoprecipitation using human TNFα standard. A similar size protein was found with B1.

Strong evidence, such as the ELISA test, suggests B2 recognises TNFα. The immunoprecipitation tests are encouraging due to the similar protein size of the precipitated band compared to the expected size for TNFα. However the conflicting mass spectrometry results indicate more tests are required. A more definitive way to measure protein-protein interaction will be by using a Biacore (Fivash, Towler et al. 1998). This equipment is able to measure protein-protein interactions and kinetics. The PCR generated library was able to generate random VHH as expected, and has the potential to generate monoclonal antibodies for desired proteins without animal inoculation. This could be an alternative for proteins or even whole pathogens that require special ethical approval or are not approved because of the toxicity. In addition, this method reduces the time required for the production of the antibodies and incorporates the benefits of VHH such as simplification for molecular techniques. However, a PCR generated VHH library has some disadvantages. Use of random sequences can introduce STOP codons in the coding region that will not be present in a natural VHH library, proteins with folding issues or toxic protein for the bacteria. All these issues can be resolved with an adequate selection process. An intrinsic issue with VHH is the type of antigen recognised. The loss of the light chain considerably reduced the variability of the VHH. In order to compensate this loss, the CDR regions and more specific CDR3 suffer an increase in the number of amino acids. The antigen binding regions in VHH are considerably larger and form protuberances in contrast with the flatter morphology found in classic IgG. The larger and more 3-dimensional antigen binding regions increased the variability. A secondary effect is that VHH seems to originate more and more efficient blocking antibodies; the longer CDR regions are able to block active sites more efficiently than IgG. This has been used as an evolutionary advantage that allowed the conservation of VHH which due to their different morphology could identify new epitopes. 3-dimensional structure seems to be important for VHH recognition, suggesting that VHH will not be ideal for denaturing protein techniques such as SDS-PAGE western-blot.

As discussed before the phage display can be problematic due to the presence of fast growers. An alternative could be the use of ribosomal display. This technique is based on the creation

of complex formed by ribosome, RNA and the synthesised protein. The technique is completely *in vitro*, reducing the competition between VHH and the appearance of bias of the population to other factors outside the affinity to the antigen.

CHAPTER 6: FINAL DISCUSSION

6.1. Gene expression and regulation of epididymal β-defensins

Guided by the preliminary work, and research published on this defensin gene cluster in other species (Yudin, Generao et al. 2005, Tollner, Yudin et al. 2008, Yudin, Tollner et al. 2008, Tollner, Venners et al. 2011, Tollner, Bevins et al. 2012), gene expression analysis focused on specific regions of the reproductive tract, collected from animals post-mortem. The resulting gene expression profile across the male and female reproductive tract showed diverse patterns of expression. We found that, with the exception of *BBD142*, *119* and *121* which were expressed homogeneously in all analysed tissues, the Chr13 β-defensin gene cluster genes are found to be specifically localised along the reproductive tract. A group formed by *BBD120, 122, 122a* and *123* genes were found highly expressed in the testis but lowly expressed in epididymal tissue. The remaining genes have high expression levels in the epididymal regions but low in the testis. This group can be further divided into genes expressed also in the female tract (*BBD115, 116, 117, 118* and *124*) and those genes exclusively expressed in the mature male reproductive tract (*BBD132, 129, 128, 127, 126, 125* and *125a*). We named this latter group Class A. In addition no expression of Class A genes was found in an immature bull.

The Class A expression pattern has two interesting implications - preferential expression of these genes was detected in the epididymis – the area where sperm are stored until maturation before ejaculation as previously described for β-defensins in other species (Tollner, Yudin et al. 2008). The second is that in the absent of inflammation, an alternative pathway for gene expression, should be considered. The lack of expression on female reproductive tract and on immature bull suggest the activation of these promoter elements may be under hormonal control and are therefore developmentally regulated (Radhakrishnan, Hamil et al. 2005, Yudin, Tollner et al. 2008), suggesting that these genes may have a role in the fertility (Yudin, Generao et al. 2005, Tollner, Yudin et al. 2008, Yudin, Tollner et al. 2008, Tollner, Venners et al. 2011, Tollner, Bevins et al. 2012). Some of the roles in fertility include mediation of the sperm binding to the female epithelia. On the other hand, β-defensins could be important for sperm maturation. It is known that this process occurs by the attachment of a multitude of proteins to the sperm cell surface during its transit through the epididymis. Alternatively, the regional distribution of the defensins could be explained by each group targeting specific pathogens only present in each region or to prevent infection in the female reproductive tract after mating (Sorensen, Gram et al. 2003).

The Class A genes possess additional characteristics other than expression pattern which differentiates them from the other genes in the cluster such as significantly extended C-termini, Class A genes have a tail (measured from the last cysteine to the predicted stop codon) of 50 amino acids on average, which is significantly different in length ($P<0.05$) from the 24 amino acids found on average in the other genes in the cluster, increasing the chances for glycosylation of the peptide. In (Yudin, Tollner et al. 2008), mouse *β-defensin 122*, orthologous to bovine *BBD126*, is characterised as one of the most abundant proteins of the mouse sperm glycocalyx. Recently, new evidence for the role of *β-defensin 126* in reproduction has been found; a mutation in human β-defensin 126 was associated with a reduction in sperm penetration of vaginal mucus and reduced fertility (Tollner, Venners et al. 2011). Other publications suggested the importance of glycosylation of these defensins for their activity in the mammalian reproductive tract (Tollner, Yudin et al. 2008). Interestingly, most of the predicted sites are found in the tail outside the defensin motif; this elongated tail is a characteristic of Chr13 defensins cluster. Some of these defensins are highly glycosylated; BBD126 molecular weight is more than doubled by the addition of carbohydrates, as the de-glycosylation treatment showed. In addition, orthologues of *BBD126* have been found to have a reproductive role in coating macaque sperm (Radhakrishnan, Hamil et al. 2005, Yudin, Generao et al. 2005, Tollner, Yudin et al. 2008). Finally the Class A genes, which in bovine are found in a contiguous group inside the gene cluster, form a separate cluster in primate species like human, chimpanzee and macaque (Rodriguez-Jimenez, Krause et al. 2003, Patil, Cai et al. 2005, Radhakrishnan, Hamil et al. 2005). The retention of the grouping despite genomic translocation may indicate group regulation and/or synergistic activity.

The presence of such site specific patterns of gene expression and furthermore, expression in the reproductive tissues of apparently healthy animals hinted at a primary reproductive role for these molecules. This led us to investigate the regulation of the expression of these genes in an *in vitro* cell culture model. Whereas most β-defensins which are TLR mediated (Diamond, Kaiser et al. 2000) our analysis of the BBD126 gene promoter region by luciferase reporter assay, showed the activation of gene expression by hormones, dihydrotestosterone (DHT) and oestrogen, but not for a TLR ligand, LPS. Despite DHT being considered a male hormone and oestrogen a female hormone, both hormones are found in both sexes this suggest an additional unknown factor present only in adult male that induces the gene

expression. To our knowledge this is the first analysis of steroid hormones affecting the β-defensins promoter region (Fig. 6. 1.).

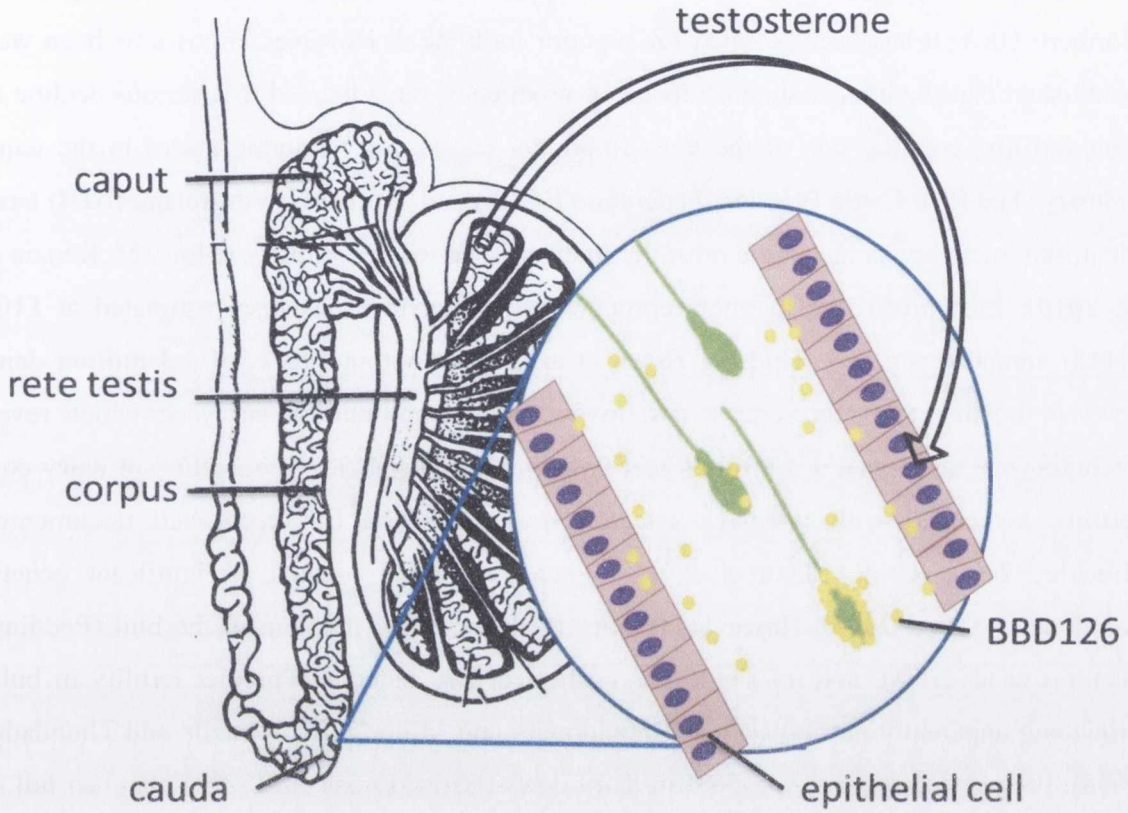


Fig. 6. 1. Expression and hormonal regulation of β-defensins in the male reproductive tract.

Hormones produced locally in the testis and in other organs such as the adrenal glands, activate gene expression of β-defensin in the epithelial cells of the epididymis. The β-defensins are then secreted into the lumen of the epididymis where they cover the surface of the sperm cell in transit.

To this point only *BBD126* and its human and macaque orthologous have been analysed in detail, our hypothesis is that in these genes, particularly Class A genes, expression will be up-regulated by DHT and oestrogen but not TLR ligands and therefore the future work should include the analysis of the promoter region of other members of Chr13 β-defensin gene cluster by luciferase reporter gene to determine the role of hormone regulation in expression of these genes. The roles of the other members of the cluster remain unknown.

6.2. GENETIC VARIATION ON EPIDIDYMAL B-DEFENSINS

The Holstein-Friesian breed of cow is the predominant breed in Ireland, Europe, and Northern USA. It has been selected for superior milk yield. However, it has now been well established that the almost singular focus on production traits has led to a serious decline in their fertility, creating one of the most important causes for economic losses in the dairy industry. The Irish Cattle Breeders Federation (ICBF) and Animal Health Ireland (AHI) have identified infertility as the single most detrimental issue for the Industry (More, McKenzie et al. 2010). Lost profit due to poor reproductive performance has been estimated at £100 (€118) annually per cow (Hudson, Breen et al. 2010). Within excess of 1.3 million dairy cows in Ireland, this represents a cost in excess of €153 million per year, which (even excluding the additional 1.3 million beef cows) is unsustainable. The decline in dairy cow fertility associated with intensive selection for production has been well documented (Buckley 2000, Royal, Mann et al. 2000), and is known to have a significant genetic component (Shook 2006). However, fertility is also a major problem in the bull (Peddinti, Nanduri et al. 2008), and no single diagnostic test can accurately predict fertility in bulls producing apparently normal semen (Braundmeier and Miller 2001, Kastelic and Thundathil 2008). Pregnancy rates are ~45-50% in dairy cows (Berry, Evans et al. 2011) but can fall as low as 25% with low fertility bulls (Larson and Miller 2000, Al Naib, Hanrahan et al. 2011), and this variation is a major impediment to the use of high genetic merit bulls and limits genetic gain. In contrast, Norwegian Red cattle have been selected with less emphasis on milk yield and include fertility traits. As a result fertility is a lot higher in this breed. Most of the research to identify the origin of the problem has concentrated on the role of the cow, in many cases ignoring the male side. Research in other species has shown the importance of β-defensins in fertility. We aimed to characterise the presence of genetic variation, such as single nucleotide polymorphisms (SNP), in the coding region of these genes that could explain the difference in fertility between bovine breeds. To address this we genotyped the coding region of Chr13 β-defensin gene cluster in two bovine breeds. We compared Holstein-Friesian (HF), high yield milk producers but with low fertility indexes with Norwegian Red (NR), a dairy breed with lower milk production yields but that has been selected with a broader range of parameters such as fertility and immune response to infection. We found initially 17 possible novel SNPs in a small sample, of which 12 were genotyped in 30 NR and 30 HF. We found four SNPs with significantly different frequencies between the two breeds.

In addition the genotyping allowed us to identify two haplotypes with significantly different frequencies between the two breeds for one of the haplotypes.

To determine the biological significance of these SNPs an initial study comparing the presence of the different SNPs and the frequency of pregnancy for artificial insemination obtained for the sperm produced by those bulls. This preliminary study found no correlation between SNPs and fertility indices. However, a second study tried to account for the bias introduced by the widespread phenomenon of artificial insemination that bypasses some of the processes in fertility in which β-defensins have been implicated such as sperm motility through cervical mucus (Tollner, Yudin et al. 2008). A sperm penetration assay using cervical mucus was used to determine the motility of sperm cells from bulls with different genotypes. We compared the motility for *BBD117 (a)*, *BBD123* and *BBD126 (b)*. We found a significant reduction in sperm motility for the T allele of the *BBD117 (a)*, that introduces a STOP codon, indicating a role for this gene in sperm motility. However, without a clear biological role it is not possible to determine the real significance of the genetic variation. *BBD126 (b)* despite the lack of any correlation with fertility indexes is particularly interesting. Work in other species with *BBD126* orthologues; suggest that *BBD126* is the β-defensin with a role in fertility. However, it was not possible to find bulls homozygous for the T allele although several TT females were found in the population. For our sperm penetration assay we compared AT vs. AA bulls, this could explain the lack of any effect in sperm motility. All samples were obtained from the National Cattle Breeding Centre. This centre collects, analyses and stores sperm from the national herd. As a result, bulls with low motility, low viability or poor quality sperm are eliminated from the breeding pool. Accordingly, the sample pool we analysed is highly biased towards high fertility bulls. We hypothesised that *BBD126 (b)* TT genes are eliminated from the sperm bank, but the T allele remain in the population in both females and heterozygous males. To test this hypothesis a large unselected population of bulls will be genotyped for the different SNPs and the sperm will be used for sperm penetration assay to determine the effect of these SNPs. To this point only the coding region of β-defensins has been analysed. However, as the gene expression characterisation showed us, these genes are highly regulated and localised in expression pattern. Genetic variation in the promoter region of these genes could have an important effect on the expression, production and secretion and subsequently in fertility of these bulls. Future work will characterise the genetic variability of the promoter region for these genes that could explain variation in fertility between breeds.

The genetic differences between breeds could have been originated by variation in the ancestral population that originated each of the breeds. However, the differences in frequencies could be a result of the differences in the breeding programme selection that were used to generate the analysed breeds. Further work should include the analysis of other breeds subject to different types of selection pressures.

6.3. Tool development for functional analysis of Epididymal β-defensins

The final aim of the study was to functionally characterise the role of these novel β-defensins. However, to do so several tools not available to us are required. To continue this work, these tools have to be developed. The most important of these tools is the production of antibodies specific for the detection of the different β-defensins. Due to the work done in other species, BBD126 was selected as the target for antibody production. We took several approaches to increase the chances to obtain an antibody.

We decided to use BBD126 as an antigen to immunise mice to produce polyclonal antibodies and to use in functional studies. Due the difficulty to isolate BBD126 on a sperm sample, we decided to synthesise recombinant BBD126 (rBBD126) as antigen. Production of rBBD126 proved more challenging than expected. Prokaryotic expression systems were tested without success. This was probably due first the toxic effect of β-defensins on bacterial growth and second the incorrect folding of the protein which will induce its precipitation and formation of inclusion bodies. As an alternative, eukaryotic expression systems were used. The pEAK8 vector system resulted in low levels of expression of a 15 kDa protein. This was significantly heavier than the 8 kDa molecular weight expected from its amino acid sequence. The differences in weight can be explained if we consider the several predicted glycosylation sites found in the sequence of BBD126. The low production yields could be caused by a toxic effect of β-defensins on the metabolism of eukaryotic cells. To solve this problem we propose to use a fusion protein system. A similar approach has been previously described for the production of human defensin 26 and 27 as fusion with thioredoxin (Huang, Leong et al. 2009). This approach increases the solubility of the resulting protein and reduces the toxic effect of the β-defensin. To this purpose a pBAD TOPO Thiofusion expression system will be used to include the coding region of BBD126 to produce a BBD126-thioredoxin protein. The resulting protein can be then cleaved by enterokinase to obtain a purified rBBD126.

The second approach was the commercial synthesis of a monoclonal antibody for BBD126. A couple of mice were inoculated with a chemically synthesised 14 amino acid peptide fragment of BBD126 conjugated with KLH. Antibody production cells were isolated and used to create immortalised cells by hybridoma technology. We were supplied with four clones resulting of the process to test with our samples. The four clones were tested in western-blot analysis of seminal plasma and sperm cell lysate. We found that clones A and D were able to detect a 25 kDa protein when used to detect protein in denature samples. Clone A was chosen for further studies. First we tried to determine the ability of anti-BBD126 antibody to detect native BBD126 first by immunocytofluorescence in sperm cells and second in immunohistochemistry in testis and epididymis tissue. In both cases we found the protein located as expected from work in other species, in sperm cell surface and in epididymis epithelia but not in testis. To further characterise the antibody an immunoprecipitation assay using sperm cell lysate and seminal plasma captured a 15 kDa protein. Mass spectrometry identified the protein as spermadhesin a 13 kDa common protein of seminal plasma with the same pattern of expression as found in the immunocytofluorescence and immunohistochemistry experiments. However, the protein obtained from the native samples did not match the protein weight found with western-blot. Our hypothesis to explain this effect is that the antibody recognises different protein in denatured and native samples. The antibody was generated using a linear antigen. This antigen might not be accessible in the native protein, but only in the denatured protein. To test our hypothesis a denatured sample immunoprecipitation assay will be required, to determine if a different protein is precipitated in these conditions. Heat-denatured seminal plasma will be incubated with anti-BBD126 in the presence of protein A/G agarose. Protein A/G binds with high affinity to the Fc region of IgG molecules incorporating it into the matrix. Then the IgG will bind to the antigen. The complex form can be then boiled and run out on an acrylamide gel.

A third approach was the use of heavy chain only antibodies. Monoclonal antibody production for even just some of the defensins is not economically viable to us. We tested the possibility to create monoclonal antibodies by this alternative method. As described before, the characteristic of these immunoglobulins could allow for a multiple inoculation, a single animal immunised with several antigens at once, thus reducing cost and time required to produce these antibodies. As proof of concept we used a highly complex mix of protein such as yeast cell lysate and Jurkat cells lysate to generate an immune response in the alpaca. The

antibody-antigen matching was done by phage display technology. This initial study resulted in the isolation of several VHH but it was not possible to match the antibody with the specific antigen. This was caused by the complexity of the protein mix during the selection process. A more sensible approach would be inoculation with a reduced number of proteins followed by an individual panning process for each specific antigen. In addition phage display presents some issues that could have a negative effect in the selection process. As with many display technologies based in live organisms, such as yeast display, phage display is based on the assumption that the selection process for the antigen is the only selection occurring in the library and all cells grow at the same rate. This is rarely the case; the fusion protein created by the VHH and pIII phage protein can have deleterious and beneficial effects on cell metabolism. For example, the process might induce an increase in the growth rate to produce "fast growers". These clones can overgrow and out-compete the selection effect of the panning thus dramatically reducing the chances of finding a good antibody. To reduce this effect, methods have been generated such as droplet growth, in which a single cell is grown in a droplet to reduce the competition for resources from other cells. However, alternative *in-vitro* systems may overcome these problems, in particular ribosomal display. This technology is based on the formation of a mRNA-ribosome-protein complex during protein translation. The mRNA is modified with a loop at the 3' end. When this loop reaches the ribosome it is able to block the translation creating the complex that will only by disrupted by radical changes in temperature, or pH. The complex form is the used for display rounds as in phage display. But due *in-vitro* nature of the system, no fast growers are produced. Our intention is to use ribosomal display in any future panning process to remove the effect of fast growers in the selection process.

In addition, we developed a novel method of VHH generation by creating a PCR generated library without requiring immunising an animal. As proof of concept we used an ELISA bovine TNFα standard as antigen for our panning process. The PCR generated library was then cloned into the previously used phage display vector. After two rounds of selection, two VHH were selected for production: B1 and B2. Once produced, both were used for an immunoprecipitation assay using the bovine TNFα standard. The result was the precipitation of a 17 kDa band in the B1 lane, recombinant TNFα has a weight of 17 kDa. When the same experiment was replicated for human TNFα standard and same protein band, 17 kDa, was found. These bands were sent for mass spectrometry identification to Prof. Hidde Ploegh lab in MIT. The results obtained could not be trusted as contamination and carry over issues were

discovered in the MS facility used. A third immune precipitation experiment using human serum showed the same protein at a lower concentration. In addition, to test the possibility of B1 as a specific anti-TNFα antibody; both B1 and B2 were used in an ELISA assay. The commercial capture antibody was substituted by B1, B2 or enhancer (ENH), a GFP specific VHH. We found that the well s containing B1 were able to detect the TNFα, but no detection was found in wells containing B2 or ENH. No reaction was found for any of them when used to detect IL-4. This result allows us to be quite confident in the specificity of B1 as TNFα antibody. However, further study is required. Some of the work we propose involve the use of a piece of equipment known as Biocore. This equipment is routinely used to measure protein-protein interactions and its kinetics. In addition future work should try to transfer the PCR generated library into a ribosomal display system, to reduce some of the issues previously described.

VHH antibodies have an immense potential for the generation of cost effective research tools for market not commercially attractive for companies by traditional means. These methods could allow to a single lab to generate the monoclonal antibodies for their own work without relaying in commercial sources and reducing the time frame normally associated with classical monoclonal antibodies.

Fertility in cattle, particularly in the dairy industry constitutes one the mayor economical. Most studies concentrate in the female side of reproduction to solve the problem, ignoring the male side of the equation. The discovery of Chr13 β-defensins presence in reproductive tract both and female, and their roles described in other species, suggest these genes a novel target for research involving fertility in cattle and the importance of the male in fertility.

Appendix

Table 1. Concentration values of extracted RNA of male reproductive tract samples and values used for cDNA synthesis.

ST, seminal tubule; RT, rete testis; Ca, caput; Co, corpus; Cu, cauda; VD, vas deferens.

Sample ID	*ng/ul*	*ng/ul*	*ug/ul*	*VDolume required for 1ug*	*water to add*
ST1	728.08	728.08	0.73	1.37	10.63
RT1	769.56	769.56	0.77	1.30	10.70
Ca1	245.82	245.82	0.25	4.07	7.93
Co1	285.73	285.73	0.29	3.50	8.50
Cu1	332.92	332.92	0.33	3.00	9.00
VD1	251.50	251.50	0.25	3.98	8.02
ST2	829.85	829.85	0.83	1.21	10.79
RT2	861.92	861.92	0.86	1.16	10.84
Ca2	494.59	494.59	0.49	2.02	9.98
Co2	1206.99	1206.99	1.21	0.83	11.17
Cu2	438.72	438.72	0.44	2.28	9.72
VD2	214.35	214.35	0.21	4.67	7.33
ST3	334.21	334.21	0.33	2.99	9.01
RT3	435.94	435.94	0.44	2.29	9.71
Ca3	449.13	449.13	0.45	2.23	9.77
Co3	718.69	718.69	0.72	1.39	10.61
Cu3	491.48	491.48	0.49	2.03	9.97
VD3	450.00	450.00	0.45	2.22	9.78
ST4	1493.12	1493.12	1.49	0.67	11.33
RT4	1253.85	1253.85	1.25	0.80	11.20
Ca4	2847.29	2847.29	2.85	0.35	11.65
Co4	2019.01	2019.01	2.02	0.50	11.50
Cu4	1113.76	1113.76	1.11	0.90	11.10
VD4	402.91	402.91	0.40	2.48	9.52
ST5	1198.55	1198.55	1.20	0.83	11.17
RT5	925.02	925.02	0.93	1.08	10.92
Ca5	1106.85	1106.85	1.11	0.90	11.10
Co5	1003.08	1003.08	1.00	1.00	11.00
Cu5	1575.87	1575.87	1.58	0.63	11.37
VD5	958.68	958.68	0.96	1.04	10.96
ST6	444.70	444.70	0.44	2.25	9.75
RT6	1491.02	1491.02	1.49	0.67	11.33
Ca6	1339.84	1339.84	1.34	0.75	11.25
Co6	1238.93	1238.93	1.24	0.81	11.19
Cu6	1853.57	1853.57	1.85	0.54	11.46
VD6	1012.85	1012.85	1.01	0.99	11.01

Table 2. Concentration values of extracted RNA of female reproductive tract samples and values used for cDNA synthesis for each sample

Ov, ovary; PFT, proximal Fallopian tube; DFT, distal Fallopian tube and U, uterus.

Sample ID	*ng/ul*	*ng/ul*	*ug/ul*	*Volume required for 1ug*	*water to add*
Ov1	679.37	679.37	0.67937	1.5	5.5
PFT1	1145.5	1145.5	1.14545	0.9	6.1
DFT1	705.95	705.95	0.70595	1.4	5.6
U1	1043.9	1043.9	1.04391	1.0	6.0
Ov2	727.45	727.45	0.72745	1.4	5.6
PFT2	1101.5	1101.5	1.10149	0.9	6.1
DFT2	653.22	653.22	0.65322	1.5	5.5
U2	1032.8	1032.8	1.0328	1.0	6.0
Ov3	1385.9	1385.9	1.38593	0.7	6.3
PFT3	1091.4	1091.4	1.09136	0.9	6.1
DFT3	698.69	698.69	0.69869	1.4	5.6
U3	675.28	675.28	0.67528	1.5	5.5
Ov4	1285.7	1285.7	1.2857	0.8	6.2
PFT4	1048	1048	1.04797	1.0	6.0
DFT4	1048.6	1048.6	1.04861	1.0	6.0
U4	1131.7	1131.7	1.1317	0.9	6.1
Ov5	529.92	529.92	0.52992	1.9	5.1
PFT5	411.15	411.15	0.41115	2.4	4.6
DFT5	596.67	596.67	0.59667	1.7	5.3
U5	560.42	560.42	0.56042	1.8	5.2
Ov6	1352.4	1352.4	1.35244	0.7	6.3
PFT6	922.14	922.14	0.92214	1.1	5.9
DFT6	1122.6	1122.6	1.12257	0.9	6.1
U6	846.4	846.4	0.8464	1.2	5.8
Ov7	1431.6	1431.6	1.43157	0.7	6.3
PFT7	1178.2	1178.2	1.17822	0.8	6.2
DFT7	449.99	449.99	0.44999	2.2	4.8
U7	629.46	629.46	0.62946	1.6	5.4
Ov8	1781.8	1781.8	1.78181	0.6	6.4
PFT8	1707.1	1707.1	1.70712	0.6	6.4
DFT8	419.41	419.41	0.41941	2.4	4.6
U8	506.14	506.14	0.50614	2.0	5.0
Ov9	932.7	932.7	0.9327	1.1	5.9
PFT9	1013.2	1013.2	1.0132	1.0	6.0
DFT9	881	881	0.881	1.1	5.9
U9	538.5	538.5	0.5385	1.9	5.1
Ov10	736.9	736.9	0.7369	1.4	5.6
PFT10	957	957	0.957	1.0	6.0
DFT10	688.3	688.3	0.6883	1.5	5.5
U10	522	522	0.522	1.9	5.1

Table 3. Genomic DNA concentrations isolated from semen sample for genotyping.

Sample ID	*ng/ul*
MTW	140.45
VMG	128.88
IRP	129.72
SJI	139.34
KOZ	134.78
OLG	79.06
RXR	64.89
BHZ	122.61
SOK	79.14
FLT	102.62
BQB	118.82
HMY	107.35
SIZ	145.11
GZL	112.61
GYK	143.74
BLANK	0.56
RDU	90.99
EDV	158.72
BEI	175.97
LBO	178.72

Table 4. Eigen values for the different components found in the analysis, and % of variance and accumulative variance explained for each component.

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	11.547	36.083	36.083
2	4.254	13.292	49.376
3	2.415	7.546	56.922
4	1.380	4.312	61.233
5	1.138	3.556	64.789
6	1.041	3.254	68.043
7	1.010	3.155	71.198
8	.912	2.850	74.048
9	.778	2.432	76.480
10	.739	2.309	78.789
11	.688	2.151	80.940
12	.610	1.906	82.846
13	.581	1.817	84.662
14	.535	1.673	86.335
15	.471	1.473	87.808
16	.413	1.292	89.100
17	.408	1.275	90.375
18	.376	1.174	91.549
19	.331	1.035	92.584
20	.320	1.000	93.584
21	.295	.920	94.505
22	.274	.855	95.359
23	.237	.741	96.100
24	.226	.707	96.807
25	.210	.658	97.465
26	.188	.588	98.053
27	.184	.575	98.628
28	.134	.418	99.046
29	.108	.337	99.383
30	.104	.326	99.709
31	.093	.291	100.000
32	2.776E-17	8.674E-17	100.000

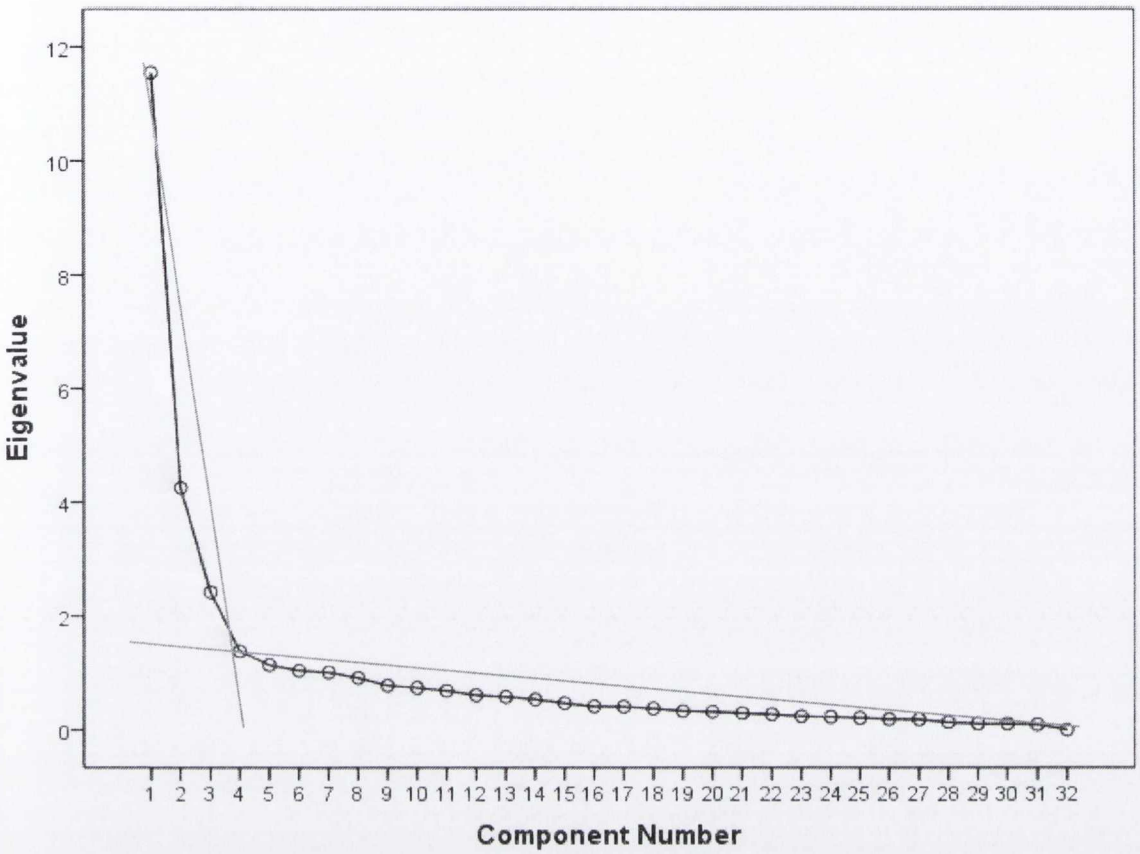


Fig. 1. Scree plot of eigenvalues and components numbers; inflexion point mark the cut off point of statistically important components.

Table 5. Linkage disequilibrium P values of NR, HF and combined populations.

Linkage disequilibrium analysis carried out by Genepop software package for the genotyped NR and HF population. <0.05 values indicate a significant linkage between the compared SNPs.

		P-value		
Locus par		NR+HF	HF	NR
BBD128 (B)	BBD126 (A)	**<0.001**	**0.000**	**0.000**
BBD128 (B)	**BBD126 (B)**	**<0.001**	**0.000**	**0.000**
BBD126 (A)	BBD126 (B)	**<0.001**	**0.000**	**0.000**
BBD128 (B)	**BBD115**	**0.000**	**0.000**	**0.023**
BBD126 (A)	BBD115	**<0.001**	**0.000**	**0.021**
BBD126 (B)	**BBD115**	**<0.001**	**0.000**	**0.022**
BBD128 (B)	BBD116	0.301	0.717	0.122
BBD126 (A)	**BBD116**	0.328	0.797	0.124
BBD126 (B)	BBD116	0.320	0.799	0.120
BBD115	**BBD116**	**0.002**	0.146	**0.002**
BBD128 (B)	BBD117 (A)	0.422	1.000	0.144
BBD126 (A)	**BBD117 (A)**	0.388	0.867	0.146
BBD126 (B)	BBD117 (A)	0.381	0.866	0.142
BBD115	**BBD117 (A)**	0.226	0.310	0.191
BBD116	BBD117 (A)	**0.001**	**0.000**	0.854
BBD128 (B)	**BBD117 (B)**	0.859	1.000	0.519
BBD126 (A)	BBD117 (B)	0.860	1.000	0.521
BBD126 (B)	**BBD117 (B)**	0.858	1.000	0.518
BBD115	BBD117 (B)	0.225	0.509	0.115
BBD116	**BBD117 (B)**	0.525	0.321	0.630
BBD117 (A)	BBD117 (B)	0.277	0.729	0.107
BBD128 (B)	**BBD121**	0.692	0.382	0.855
BBD126 (A)	BBD121	0.988	1.000	0.851
BBD126 (B)	**BBD121**	0.989	1.000	0.855
BBD115	BBD121	0.393	0.309	0.418
BBD116	**BBD121**	0.218	0.056	1.000
BBD117 (A)	BBD121	0.072	0.449	**0.030**
BBD117 (B)	**BBD121**	**0.010**	**0.002**	0.545
BBD128 (B)	BBD122	0.695	0.385	0.854
BBD126 (A)	**BBD122**	0.989	1.000	0.853
BBD126 (B)	BBD122	0.989	1.000	0.856
BBD115	**BBD122**	0.393	0.307	0.420
BBD116	BBD122	0.212	0.054	1.000
BBD117 (A)	**BBD122**	0.078	0.451	**0.033**
BBD117 (B)	BBD122	**0.008**	**0.002**	0.550
BBD121	**BBD122**	**<0.001**	**0.000**	**0.000**
BBD128 (B)	BBD123 (A)	0.529	0.316	0.648

Locus par	NR+HF	HF	NR	
BBD126 (A)	BBD123 (A)	0.829	0.729	0.652
BBD126 (B)	BBD123 (A)	0.830	0.729	0.653
BBD115	BBD123 (A)	0.113	0.096	0.250
BBD116	BBD123 (A)	0.321	0.096	1.000
BBD117 (A)	BBD123 (A)	0.089	0.804	**0.022**
BBD117 (B)	BBD123 (A)	0.082	0.058	0.275
BBD121	BBD123 (A)	**<0.001**	**0.000**	**0.000**
BBD122	BBD123 (A)	**<0.001**	**0.000**	**0.000**
BBD128 (B)	BBD123 (B)	0.520	0.307	0.648
BBD126 (A)	BBD123 (B)	0.829	0.728	0.653
BBD126 (B)	BBD123 (B)	0.831	0.732	0.653
BBD115	BBD123 (B)	0.117	0.097	0.257
BBD116	BBD123 (B)	0.324	0.097	1.000
BBD117 (A)	BBD123 (B)	0.097	0.803	**0.025**
BBD117 (B)	BBD123 (B)	0.079	0.056	0.273
BBD121	BBD123 (B)	**<0.001**	**0.000**	**0.000**
BBD122	BBD123 (B)	**<0.001**	**0.000**	**0.000**
BBD123 (A)	BBD123 (B)	**<0.001**	**0.000**	**0.000**
BBD128 (B)	BBD124	0.521	0.309	0.646
BBD126 (A)	BBD124	0.830	0.732	0.652
BBD126 (B)	BBD124	0.824	0.727	0.646
BBD115	BBD124	0.119	0.100	0.256
BBD116	BBD124	0.334	0.102	1.000
BBD117 (A)	BBD124	0.089	0.807	**0.022**
BBD117 (B)	BBD124	0.084	0.058	0.282
BBD121	BBD124	**<0.001**	**0.000**	**0.000**
BBD122	BBD124	**<0.001**	**0.000**	**0.000**
BBD123 (A)	BBD124	**<0.001**	**0.000**	**0.000**
BBD123 (B)	BBD124	**<0.001**	**0.000**	**0.000**

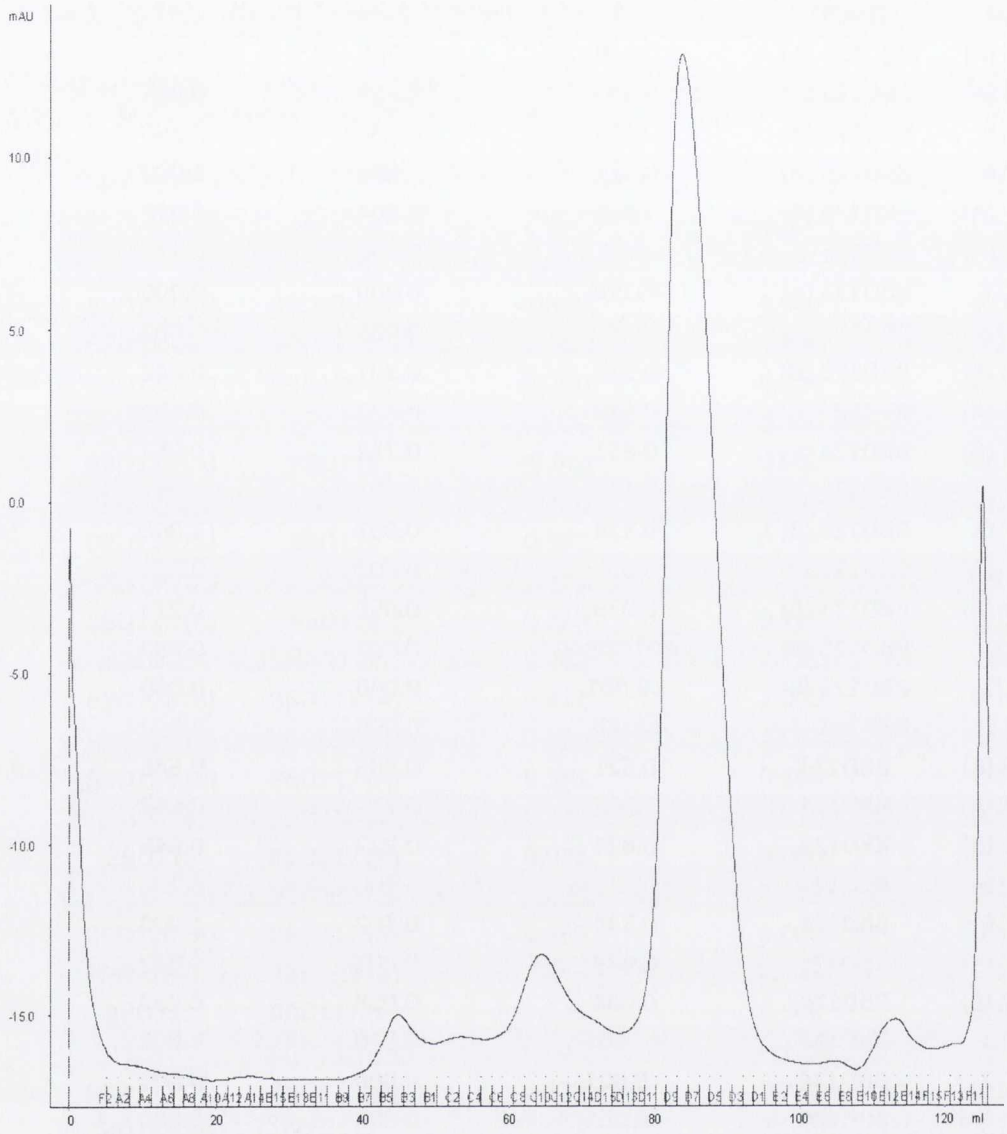


Fig. 2. FPLC trace for recombinant D1.

Protein was measure by absorption at 280nm. A peak at 80-90 ml correspond with the VHH.

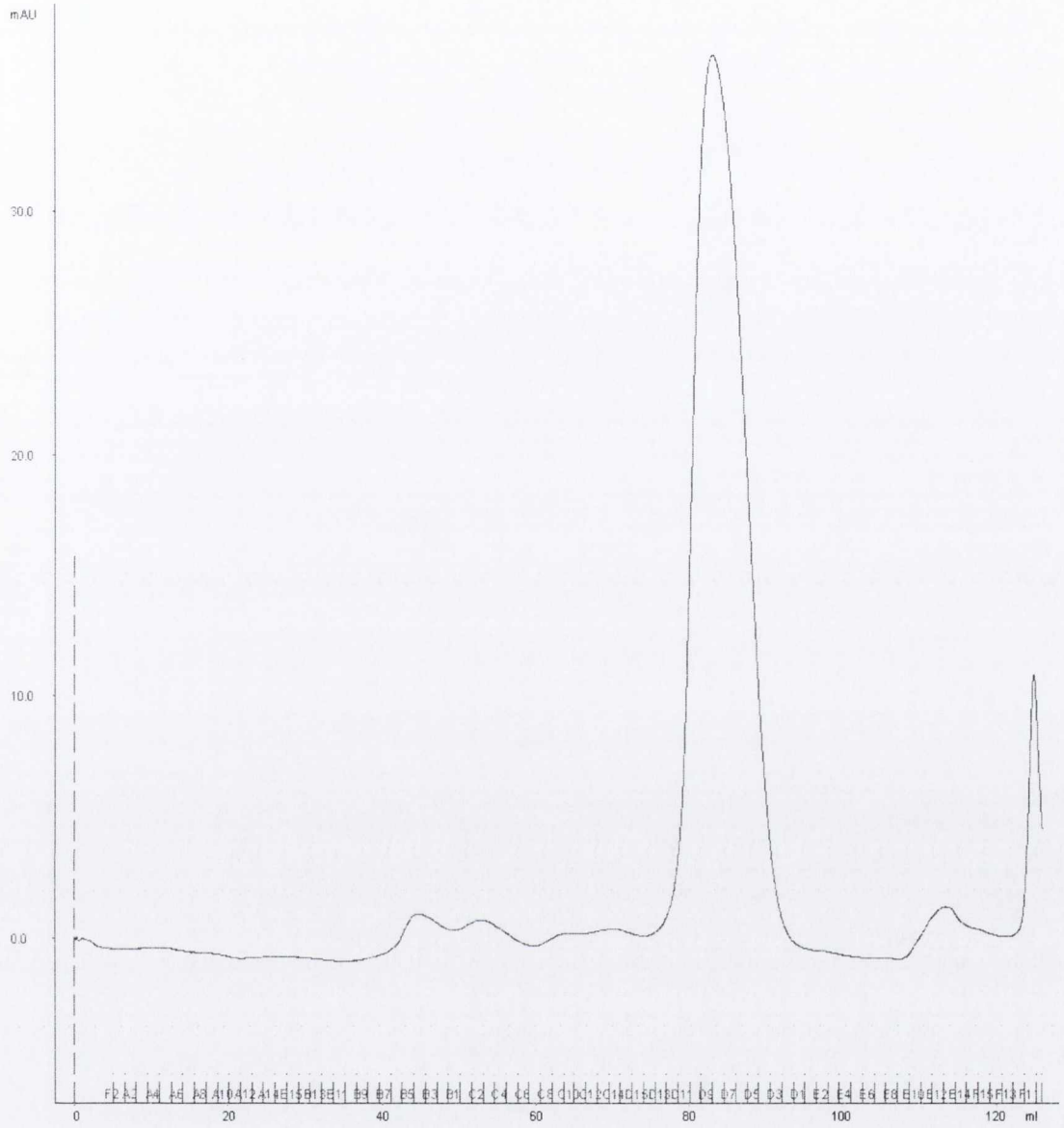


Fig 3. FPLC trace for recombinant D2.

Protein was measure by absorption at 280nm. A peak at 80-90 ml correspond with the VHH.

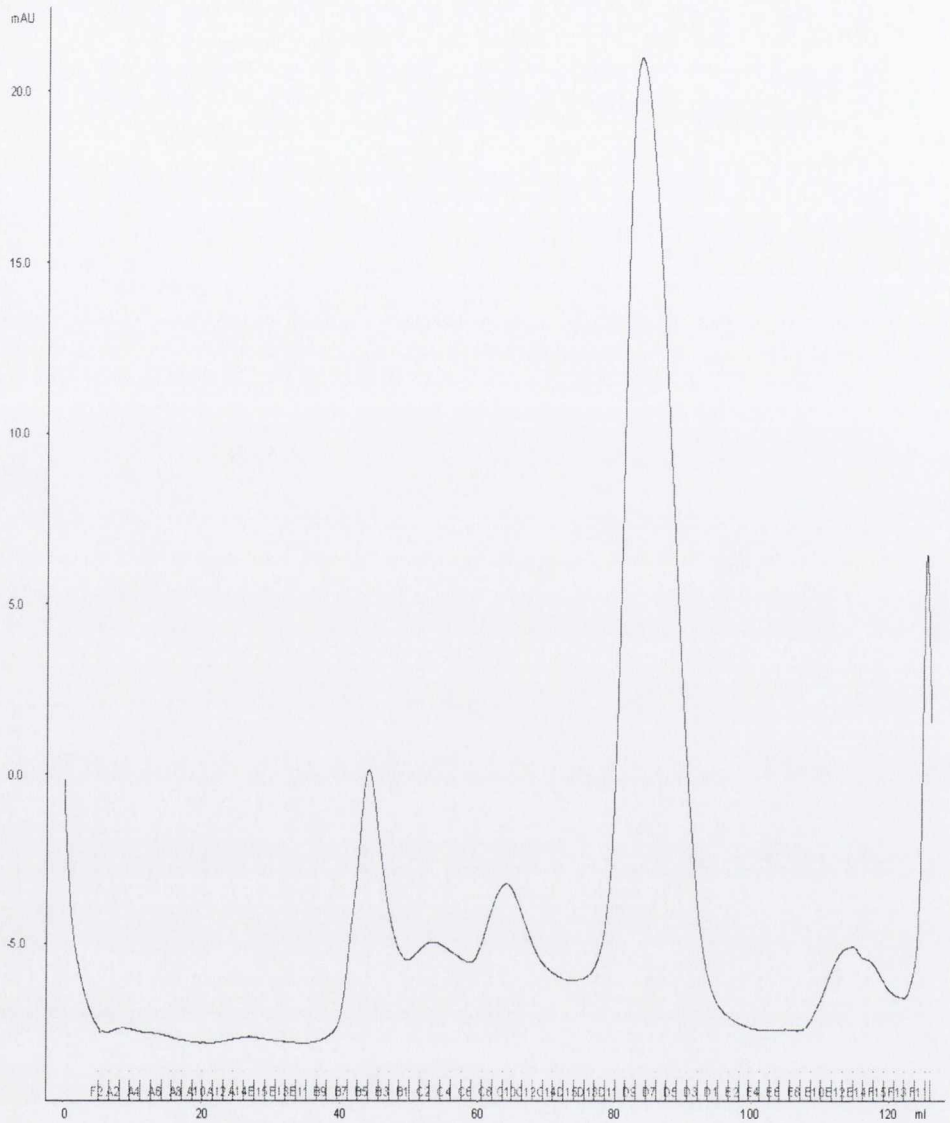


Fig 4. FPLC trace for recombinant D5.

Protein was measure by absorption at 280nm. A peak at 80-90 ml correspond with the VHH.

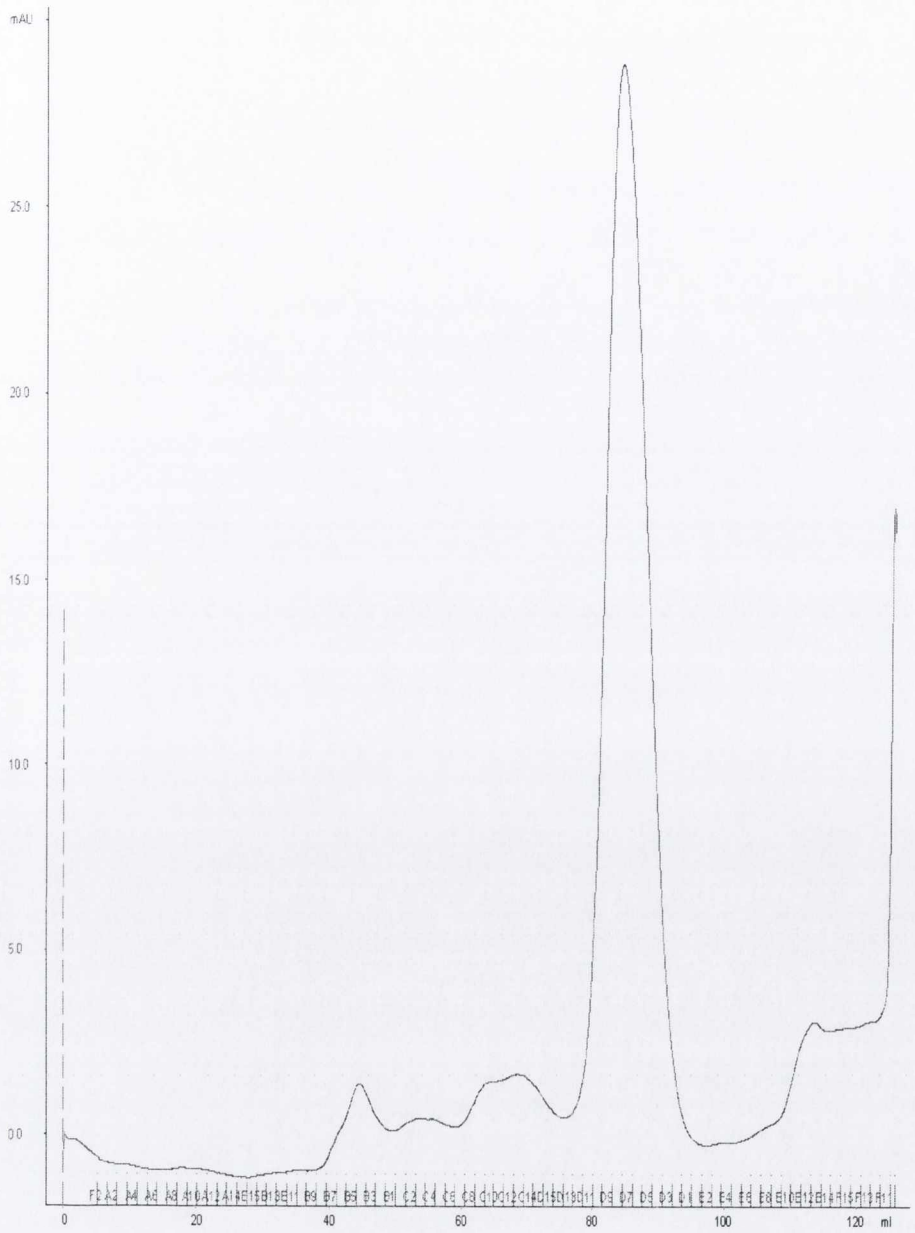


Fig 5. FPLC trace for recombinant D6.

Protein was measure by absorption at 280nm. A peak at 80-90 ml correspond with the VHH.

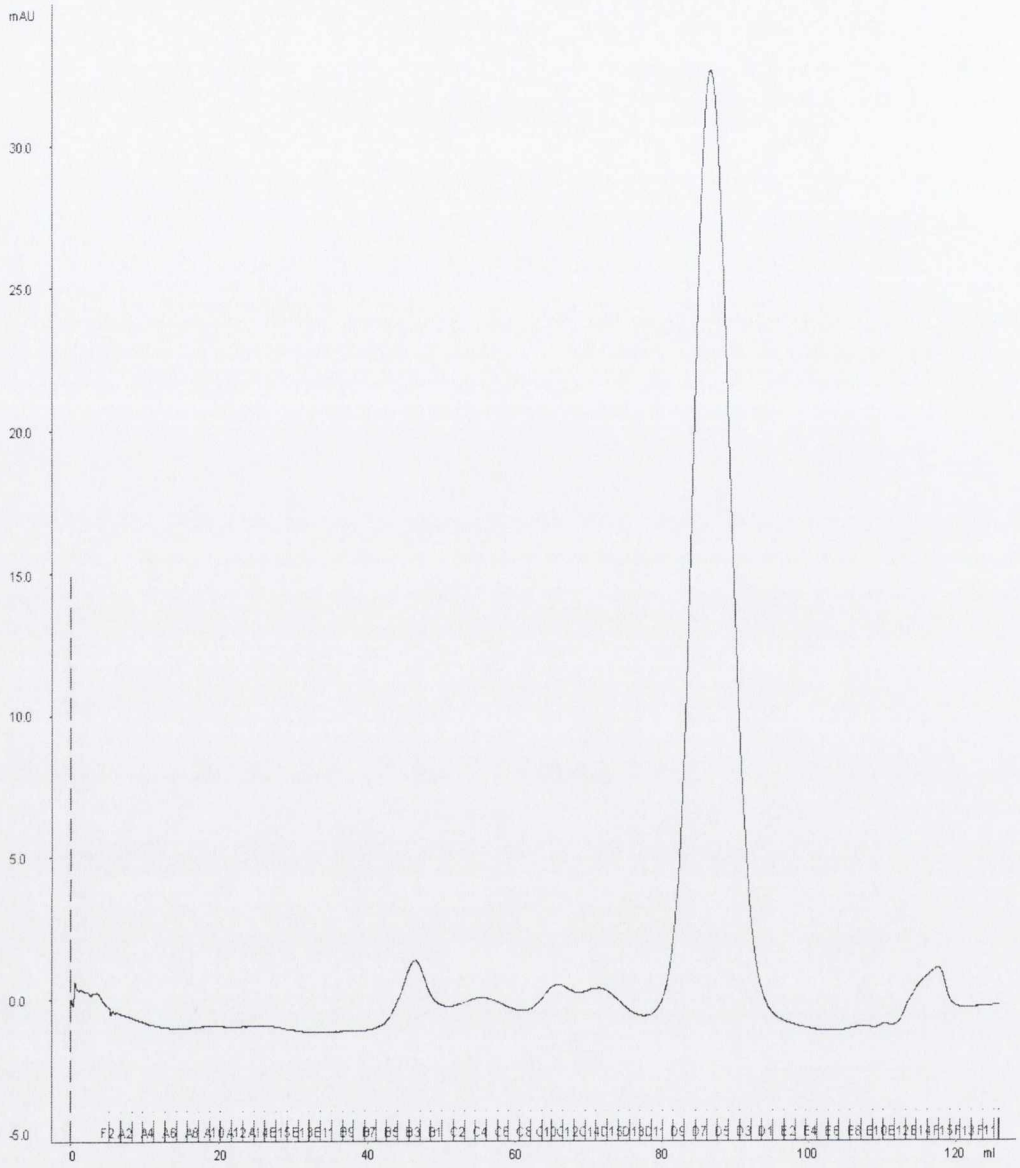


Fig 6. FPLC trace for recombinant D7.

Protein was measure by absorption at 280nm. A peak at 80-90 ml correspond with the VHH.

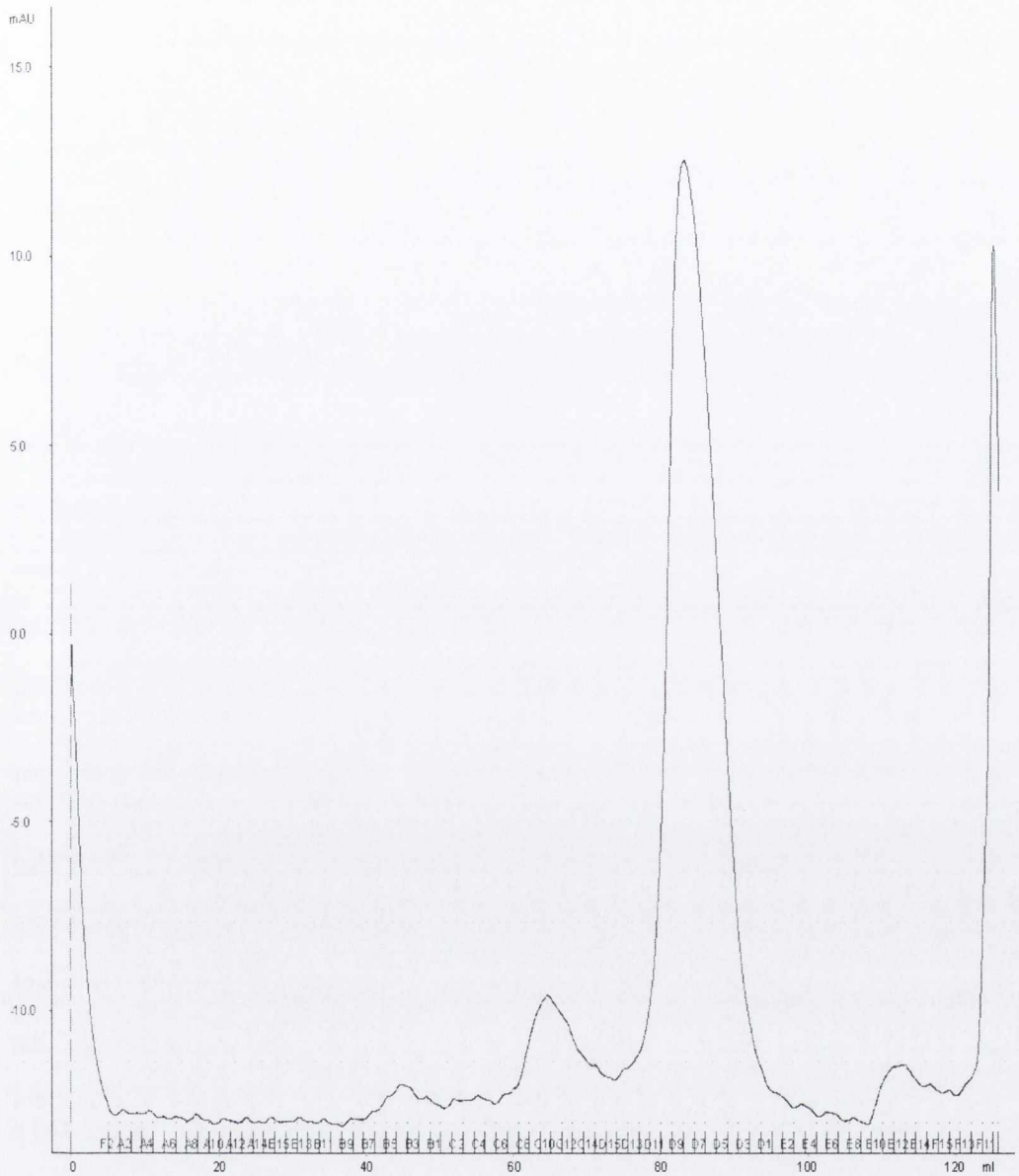


Fig 7. FPLC trace for recombinant D10.

Protein was measure by absorption at 280nm. A peak at 80-90 ml correspond with the VHH.

REFERENCES

- Agerberth, B., J. Charo, J. Werr, B. Olsson, F. Idali, L. Lindbom, R. Kiessling, H. Jornvall, H. Wigzell and G. H. Gudmundsson (2000). "The human antimicrobial and chemotactic peptides LL-37 and alpha-defensins are expressed by specific lymphocyte and monocyte populations." Blood **96**(9): 3086-3093.
- Al Naib, A., J. P. Hanrahan, P. Lonergan and S. Fair (2011). "In vitro assessment of sperm from bulls of high and low field fertility." Theriogenology **76**(1): 161-167.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.
- Ayabe, T., D. P. Satchell, C. L. Wilson, W. C. Parks, M. E. Selsted and A. J. Ouellette (2000). "Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria." Nat Immunol **1**(2): 113-118.
- Baroni, A., G. Donnarumma, I. Paoletti, I. Longanesi-Cattani, K. Bifulco, M. A. Tufano and M. V. Carriero (2009). "Antimicrobial human beta-defensin-2 stimulates migration, proliferation and tube formation of human umbilical vein endothelial cells." Peptides **30**(2): 267-272.
- Bechinger, B. and K. Lohner (2006). "Detergent-like actions of linear amphipathic cationic antimicrobial peptides." Biochim Biophys Acta **1758**(9): 1529-1539.
- Berry, D. P., R. D. Evans and S. Mc Parland (2011). "Evaluation of bull fertility in dairy and beef cattle using cow field data." Theriogenology **75**(1): 172-181.
- Berry, D. P., M. K. G., M. M. P., B. S., D. M. G., M. D. and C. C. J. (2010). "The integration of 'omic' disciplines and systems biology in cattle breeding." Animal.
- Biragyn, A., P. A. Ruffini, C. A. Leifer, E. Klyushnenkova, A. Shakhov, O. Chertov, A. K. Shirakawa, J. M. Farber, D. M. Segal, J. J. Oppenheim and L. W. Kwak (2002). "Toll-like receptor 4-dependent activation of dendritic cells by beta-defensin 2." Science **298**(5595): 1025-1029.
- Boman, H., B. Agerberth and A. Boman (1993). "Mechanisms of action on Escherichia coli of cecropin P1 and PR-39, two antibacterial peptides from pig intestine." Infection and immunity **61**(7): 2978-2984.
- Boman, H. G. (2003). "Antibacterial peptides: basic facts and emerging concepts." J Intern Med **254**(3): 197-215.
- Bowdish, D. M., D. J. Davidson and R. E. Hancock (2005). "A re-evaluation of the role of host defence peptides in mammalian immunity." Curr Protein Pept Sci **6**(1): 35-51.
- Braundmeier, A. G. and D. J. Miller (2001). "The search is on: finding accurate molecular markers of male fertility." Journal of Dairy Science **84**(9): 1915-1925.
- Brogden, K. A. (2005). "Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?" Nat Rev Microbiol **3**(3): 238-250.
- Brötz, H., G. Bierbaum, K. Leopold, P. Reynolds and H. Sahl (1998). "The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II." Antimicrobial agents and chemotherapy **42**(1): 154-160.
- Bruhn, O., P. Regenhard, M. Michalek, S. Paul, C. Gelhaus, S. Jung, G. Thaller, R. Podschun, M. Leippe, J. Grotzinger and E. Kalm (2007). "A novel horse alpha-defensin: gene transcription, recombinant expression and characterization of the structure and function." Biochem J **407**(2): 267-276.
- Buckley (2000). Trends in genetic merit for milk production and reproductive performance. National dairy conference.
- Bulet, P., R. Stöcklin and L. Menin (2004). "Anti-microbial peptides: from invertebrates to vertebrates." Immunological reviews **198**: 169-184.
- Caverly, J. M., G. Diamond, J. M. Gallup, K. A. Brogden, R. A. Dixon and M. R. Ackermann (2003). "Coordinated expression of tracheal antimicrobial peptide and inflammatory-response elements in the lungs of neonatal calves with acute bacterial pneumonia." Infect Immun **71**(5): 2950-2955.
- Chapwanya, A., K. G. Meade, M. L. Doherty, J. J. Callanan, J. F. Mee and C. O'Farrelly (2009). "Histopathological and molecular evaluation of Holstein-Friesian cows postpartum: toward an improved understanding of uterine innate immunity." Theriogenology **71**(9): 1396-1407.
- Chen, G., O. Zhuchenko and A. Kuspa (2007). "Immune-like phagocyte activity in the social amoeba." Science **317**(5838): 678-681.
- Chitnis, S. N. and K. S. Prasad (1990). "Seminalplasmin, an antimicrobial protein from bovine seminal plasma, inhibits peptidoglycan synthesis in Escherichia coli." FEMS Microbiol Lett **60**(3): 281-284.

- Cole, A. M. and R. I. Lehrer (2003). "Minidefensins: antimicrobial peptides with activity against HIV-1." Curr Pharm Des **9**(18): 1463-1473.
- Collins, F. S., L. D. Brooks and A. Chakravarti (1998). "A DNA polymorphism discovery resource for research on human genetic variation." Genome Res **8**(12): 1229-1231.
- Com, E., F. Bourgeon, B. Evrard, T. Ganz, D. Colleu, B. Jegou and C. Pineau (2003). "Expression of antimicrobial defensins in the male reproductive tract of rats, mice, and humans." Biol Reprod **68**(1): 95-104.
- Com, E., B. Evrard, P. Roepstorff, F. Aubry and C. Pineau (2003). "New insights into the rat spermatogonial proteome: identification of 156 additional proteins." Mol Cell Proteomics **2**(4): 248-261.
- Cormican, P. (2009). "A comparative genomics analysis of the vertebrate immune system: genes, pathways and evolution." Doctoral dissertation,Unpublished doctoral dissertation, Trinity College, University of Dublin.
- Cormican, P., A. T. Lloyd, T. Downing, S. J. Connell, D. Bradley and C. O'Farrelly (2009). "The avian Toll-Like receptor pathway--subtle differences amidst general conformity." Dev Comp Immunol **33**(9): 967-973.
- Cormican, P., K. G. Meade, S. Cahalane, F. Narciandi, A. Chapwanya, A. T. Lloyd and C. O'Farrelly (2008). "Evolution, expression and effectiveness in a cluster of novel bovine beta-defensins." Immunogenetics **60**(3-4): 147-156.
- Cunliffe, R. N. (2003). "Alpha-defensins in the gastrointestinal tract." Mol Immunol **40**(7): 463-467.
- Funk, D.A, (1993). "Optimal Genetic Improvement for the High Producing Herd." Journal of Dairy Science **76**(10): 3278-3286.
- Daher, K. A., M. E. Selsted and R. I. Lehrer (1986). "Direct inactivation of viruses by human granulocyte defensins." J Virol **60**(3): 1068-1074.
- Davies, J. and L. Riechmann (1994). "'Camelising' human antibody fragments: NMR studies on VH domains." FEBS Letters **339**(3): 285-290.
- De Genst, E., D. Saerens, S. Muyldermans and K. Conrath (2006). "Antibody repertoire development in camelids." Dev Comp Immunol **30**(1-2): 187-198.
- Dekkers, J. C. (2004). "Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons." J Anim Sci **82 E-Suppl**: E313-328.
- Desmyter, A., T. R. Transue, M. A. Ghahroudi, M. H. Thi, F. Poortmans, R. Hamers, S. Muyldermans and L. Wyns (1996). "Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme." Nat Struct Biol **3**(9): 803-811.
- Diamond, G., V. Kaiser, J. Rhodes, J. P. Russell and C. L. Bevins (2000). "Transcriptional regulation of beta-defensin gene expression in tracheal epithelial cells." Infect Immun **68**(1): 113-119.
- Diamond, G., J. P. Russell and C. L. Bevins (1996). "Inducible expression of an antibiotic peptide gene in lipopolysaccharide-challenged tracheal epithelial cells." Proc Natl Acad Sci U S A **93**(10): 5156-5160.
- Diba, F., A. Vinkler and L. Groch (1994). "[Testosterone levels in testicular and epididymal tissue during histopathologic changes in the testes in bulls]." Vet Med (Praha) **39**(10): 579-587.
- Duits, L. A., P. H. Nibbering, E. van Strijen, J. B. Vos, S. P. Mannesse-Lazeroms, M. A. van Sterkenburg and P. S. Hiemstra (2003). "Rhinovirus increases human beta-defensin-2 and -3 mRNA expression in cultured bronchial epithelial cells." FEMS Immunol Med Microbiol **38**(1): 59-64.
- Excoffier, L., G. Laval and S. Schneider (2005). "Arlequin (version 3.0): an integrated software package for population genetics data analysis." Evol Bioinform Online **1**: 47-50.
- Filardo, E. J. (2002). "Epidermal growth factor receptor (EGFR) transactivation by estrogen via the G-protein-coupled receptor, GPR30: a novel signaling pathway with potential significance for breast cancer." J Steroid Biochem Mol Biol **80**(2): 231-238.
- Fivash, M., E. M. Towler and R. J. Fisher (1998). "BIAcore for macromolecular interaction." Curr Opin Biotechnol **9**(1): 97-101.
- Flajnik, M. F., N. Deschacht and S. Muyldermans (2011). "A Case Of Convergence: Why Did a Simple Alternative to Canonical Antibodies Arise in Sharks and Camels?" PLoS Biol **9**(8): e1001120.
- Folli, C., S. Viglione, M. Busconi and R. Berni (2005). "Biochemical basis for retinol deficiency induced by the I41N and G75D mutations in human plasma retinol-binding protein." Biochem Biophys Res Commun **336**(4): 1017-1022.
- Funderburg, N., M. M. Lederman, Z. Feng, M. G. Drage, J. Jadlowsky, C. V. Harding, A. Weinberg and S. F. Sieg (2007). "Human -defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2." Proc Natl Acad Sci U S A **104**(47): 18631-18635.

- Ganz, T. (2003). "Defensins: antimicrobial peptides of innate immunity." Nat Rev Immunol **3**(9): 710-720.
- Ganz, T. and R. I. Lehrer (1995). "Defensins." Pharmacol Ther **66**(2): 191-205.
- Ganz, T., M. E. Selsted, D. Szklarek, S. S. Harwig, K. Daher, D. F. Bainton and R. I. Lehrer (1985). "Defensins. Natural peptide antibiotics of human neutrophils." J Clin Invest **76**(4): 1427-1435.
- Gao, B., M. Rodriguez, H. Lanz-Mendoza and S. Zhu (2009). "AdDLP, a bacterial defensin-like peptide, exhibits anti-Plasmodium activity." Biochemical and biophysical research communications **387**(2): 393-398.
- Ge, D., J. Fellay, A. J. Thompson, J. S. Simon, K. V. Shianna, T. J. Urban, E. L. Heinzen, P. Qiu, A. H. Bertelsen, A. J. Muir, M. Sulkowski, J. G. McHutchison and D. B. Goldstein (2009). "Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance." Nature **461**(7262): 399-401.
- Ghosh, D., E. Porter, B. Shen, S. K. Lee, D. Wilk, J. Drazba, S. P. Yadav, J. W. Crabb, T. Ganz and C. L. Bevins (2002). "Paneth cell trypsin is the processing enzyme for human defensin-5." Nat Immunol **3**(6): 583-590.
- Greenberg, A. S., D. Avila, M. Hughes, A. Hughes, E. C. McKinney and M. F. Flajnik (1995). "A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks." Nature **374**(6518): 168-173.
- Gwathmey, T. M., G. G. Ignotz and S. S. Suarez (2003). "PDC-109 (BSP-A1/A2) promotes bull sperm binding to oviductal epithelium in vitro and may be involved in forming the oviductal sperm reservoir." Biol Reprod **69**(3): 809-815.
- Hall, T. A. (1999). BioEdit, Nucl. Acids.Symp.Ser.41:95-98., 41, 95-98.
- Hamers-Casterman, C., T. Atarhouch, S. Muyldermans, G. Robinson, C. Hamers, E. B. Songa, N. Bendahman and R. Hamers (1993). "Naturally occurring antibodies devoid of light chains." Nature **363**(6428): 446-448.
- Helander, I. M., I. Kilpeläinen and M. Vaara (1994). "Increased substitution of phosphate groups in lipopolysaccharides and lipid A of the polymyxin-resistant pmrA mutants of Salmonella typhimurium: a 31P-NMR study." Molecular Microbiology **11**(3): 481-487.
- Higgs, R., D. J. Lynn, S. Cahalane, I. Alana, C. M. Hewage, T. James, A. T. Lloyd and C. O'Farrelly (2007). "Modification of chicken avian beta-defensin-8 at positively selected amino acid sites enhances specific antimicrobial activity." Immunogenetics **59**(7): 573-580.
- Higgs, R., D. J. Lynn, S. Gaines, J. McMahon, J. Tierney, T. James, A. T. Lloyd, G. Mulcahy and C. O'Farrelly (2005). "The synthetic form of a novel chicken beta-defensin identified in silico is predominantly active against intestinal pathogens." Immunogenetics **57**(1-2): 90-98.
- Hoglund, J. K., B. Guldbrandtsen, G. Su, B. Thomsen and M. S. Lund (2009). "Genome scan detects quantitative trait loci affecting female fertility traits in Danish and Swedish Holstein cattle." Journal of dairy science **92**(5): 2136-2143.
- Holloway, D. T., M. Kon and C. DeLisi (2005). "Integrating genomic data to predict transcription factor binding." Genome Inform **16**(1): 83-94.
- Hori, Y., M. Demura, M. Iwadate, A. S. Ulrich, T. Niidome, H. Aoyagi and T. Asakura (2001). "Interaction of mastoparan with membranes studied by 1H-NMR spectroscopy in detergent micelles and by solid-state 2H-NMR and 15N-NMR spectroscopy in oriented lipid bilayers." Eur J Biochem **268**(2): 302-309.
- Huang, L., S. S. Leong and R. Jiang (2009). "Soluble fusion expression and characterization of bioactive human beta-defensin 26 and 27." Appl Microbiol Biotechnol **84**(2): 301-308.
- Huang, W., B. W. Kirkpatrick, G. J. Rosa and H. Khatib (2010). "A genome-wide association study using selective DNA pooling identifies candidate markers for fertility in Holstein cattle." Animal genetics **41**(6): 570-578.
- Hudson, C., J. Breen, A. Bradley and M. Green (2010). "Fertility in UK herds; preliminary findings of a large-scale study." Cattle Practice(18): 89-94.
- IM, A. R., B. Heringstad, G. Klemetsdal, M. Svendsen and T. Steine (2003). "Heifer fertility in Norwegian dairy cattle: variance components and genetic change." J Dairy Sci **86**(8): 2706-2714.
- Jelinsky, S. A., T. T. Turner, H. J. Bang, J. N. Finger, M. K. Solarz, E. Wilson, E. L. Brown, G. S. Kopf and D. S. Johnston (2007). "The rat epididymal transcriptome: comparison of segmental gene expression in the rat and mouse epididymides." Biol Reprod **76**(4): 561-570.
- Jenssen, H., P. Hamill and R. E. Hancock (2006). "Peptide antimicrobial agents." Clin Microbiol Rev **19**(3): 491-511.
- Julenius, K., A. Molgaard, R. Gupta and S. Brunak (2005). "Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites." Glycobiology **15**(2): 153-164.

- Kastelic, J. P. and J. C. Thundathil (2008). "Breeding soundness evaluation and semen analysis for predicting bull fertility." Reproduction in domestic animals = Zuchthygiene **43 Suppl 2**: 368-373.
- Kent, W. J. (2002). "BLAT--the BLAST-like alignment tool." Genome Res **12**(4): 656-664.
- Kragol, G., S. Lovas, G. Varadi, B. A. Condie, R. Hoffmann and L. Otvos, Jr. (2001). "The antibacterial peptide pyrrhocoricin inhibits the ATPase actions of DnaK and prevents chaperone-assisted protein folding." Biochemistry **40**(10): 3016-3026.
- Kurosaka, K., Q. Chen, F. Yarovinsky, J. J. Oppenheim and D. Yang (2005). "Mouse cathelin-related antimicrobial peptide chemoattracts leukocytes using formyl peptide receptor-like 1/mouse formyl peptide receptor-like 2 as the receptor and acts as an immune adjuvant." J Immunol **174**(10): 6257-6265.
- Lai, Y. and R. L. Gallo (2009). "AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense." Trends Immunol **30**(3): 131-141.
- Lambert, J., E. Keppi, J. Dimarcq, C. Wicker, J. Reichhart, B. Dunbar, P. Lepage, A. Van Dorsselaer, J. Hoffmann and J. Fothergill (1989). "Insect immunity: isolation from immune blood of the dipteran Phormia terranovae of two insect antibacterial peptides with sequence homology to rabbit lung macrophage bactericidal peptides." Proceedings of the National Academy of Sciences of the United States of America **86**(1): 262-266.
- Larson, J. L. and D. J. Miller (2000). "Can relative spermatozoal galactosyltransferase activity be predictive of dairy bull fertility?" J Dairy Sci **83**: 2473-2479.
- Lauwers, A., L. Twyffels, R. Soin, C. Wauquier, V. Kruys and C. Gueydan (2009). "Post-transcriptional Regulation of Genes Encoding Anti-microbial Peptides in Drosophila." Journal of Biological Chemistry **284**(13): 8973-8983.
- Lee, P. H., T. Ohtake, M. Zaiou, M. Murakami, J. A. Rudisill, K. H. Lin and R. L. Gallo (2005). "Expression of an additional cathelicidin antimicrobial peptide protects against bacterial skin infection." Proc Natl Acad Sci U S A **102**(10): 3750-3755.
- Lehrer, R. I. and T. Ganz (2002). "Defensins of vertebrate animals." Curr Opin Immunol **14**(1): 96-102.
- Lopez-Meza, J. E., A. Gutierrez-Barroso and A. Ochoa-Zarzosa (2009). "Expression of tracheal antimicrobial peptide in bovine mammary epithelial cells." Res Vet Sci **87**(1): 59-63.
- Ludtke, S. J., K. He, W. T. Heller, T. A. Harroun, L. Yang and H. W. Huang (1996). "Membrane pores induced by magainin." Biochemistry **35**(43): 13723-13728.
- Lynn, D., R. Higgs, S. Gaines, J. Tierney, T. James, A. Lloyd, M. Fares, G. Mulcahy and C. O'Farrelly (2004). "Bioinformatic discovery and initial characterisation of nine novel antimicrobial peptide genes in the chicken." Immunogenetics **56**(3): 170-177.
- Lynn, D. J. and D. G. Bradley (2007). "Discovery of α-defensins in basal mammals." Developmental & Comparative Immunology **31**(10): 963-967.
- Lynn, D. J., R. Higgs, S. Gaines, J. Tierney, T. James, A. T. Lloyd, M. A. Fares, G. Mulcahy and C. O'Farrelly (2004). "Bioinformatic discovery and initial characterisation of nine novel antimicrobial peptide genes in the chicken." Immunogenetics **56**(3): 170-177.
- Maass, D. R., J. Sepulveda, A. Pernthaner and C. B. Shoemaker (2007). "Alpaca (Lama pacos) as a convenient source of recombinant camelid heavy chain antibodies (VHHs)." J Immunol Methods **324**(1-2): 13-25.
- Meade, K. G., S. Cahalane, F. Narciandi, P. Cormican, A. T. Lloyd and C. O'Farrelly (2008). "Directed alteration of a novel bovine beta-defensin to improve antimicrobial efficacy against methicillin-resistant Staphylococcus aureus (MRSA)." Int J Antimicrob Agents **32**(5): 392-397.
- Meade, K. G., R. Higgs, A. T. Lloyd, S. Giles and C. O'Farrelly (2009). "Differential antimicrobial peptide gene expression patterns during early chicken embryological development." Dev Comp Immunol **33**(4): 516-524.
- Medzhitov, R., P. Preston-Hurlburt and C. A. Janeway, Jr. (1997). "A human homologue of the Drosophila Toll protein signals activation of adaptive immunity." Nature **388**(6640): 394-397.
- Mee, J. F. (2004). "Temporal trends in reproductive performance in Irish dairy herds and associated risk factors." Ir Vet J **57**(3): 158-166.
- Menard, S., V. Forster, M. Lotz, D. Gutle, C. U. Duerr, R. L. Gallo, B. Henriques-Normark, K. Putsep, M. Andersson, E. O. Glocker and M. W. Hornef (2008). "Developmental switch of intestinal antimicrobial peptide expression." J Exp Med **205**(1): 183-193.
- Mendez, E., A. Moreno, F. Colilla, F. Pelaez, G. Limas, R. Mendez, F. Soriano, M. Salinas and C. de Haro (1990). "Primary structure and inhibition of protein synthesis in eukaryotic cell-free system of a novel thionin, gamma-hordothionin, from barley endosperm." European journal of biochemistry / FEBS **194**(2): 533-539.

- Miles, K., D. J. Clarke, W. Lu, Z. Sibinska, P. E. Beaumont, D. J. Davidson, T. A. Barr, D. J. Campopiano and M. Gray (2009). "Dying and necrotic neutrophils are anti-inflammatory secondary to the release of alpha-defensins." J Immunol **183**(3): 2122-2132.
- Mookherjee, N., K. L. Brown, D. M. Bowdish, S. Doria, R. Falsafi, K. Hokamp, F. M. Roche, R. Mu, G. H. Doho, J. Pistolic, J. P. Powers, J. Bryan, F. S. Brinkman and R. E. Hancock (2006). "Modulation of the TLR-mediated inflammatory response by the endogenous human host defense peptide LL-37." J Immunol **176**(4): 2455-2464.
- Mookherjee, N., H. L. Wilson, S. Doria, Y. Popowych, R. Falsafi, J. J. Yu, Y. Li, S. Veatch, F. M. Roche, K. L. Brown, F. S. Brinkman, K. Hokamp, A. Potter, L. A. Babiuk, P. J. Griebel and R. E. Hancock (2006). "Bovine and human cathelicidin cationic host defense peptides similarly suppress transcriptional responses to bacterial lipopolysaccharide." J Leukoc Biol **80**(6): 1563-1574.
- More, S. J., K. McKenzie, J. O'Flaherty, M. L. Doherty, A. R. Cromie and M. J. Magan (2010). "Setting priorities for non-regulatory animal health in Ireland: results from an expert Policy Delphi study and a farmer priority identification survey." Preventive Veterinary Medicine **95**(3-4): 198-207.
- Morioka, Y., K. Yamasaki, D. Leung and R. L. Gallo (2008). "Cathelicidin antimicrobial peptides inhibit hyaluronan-induced cytokine release and modulate chronic allergic dermatitis." J Immunol **181**(6): 3915-3922.
- Moser, C., D. J. Weiner, E. Lysenko, R. Bals, J. N. Weiser and J. M. Wilson (2002). "beta-Defensin 1 contributes to pulmonary innate immunity in mice." Infect Immun **70**(6): 3068-3072.
- Motzkus, D., S. Schulz-Maronde, A. Heitland, A. Schulz, W. G. Forssmann, M. Jubner and E. Maronde (2006). "The novel beta-defensin DEFB123 prevents lipopolysaccharide-mediated effects in vitro and in vivo." FASEB J **20**(10): 1701-1702.
- Muyldermans, S., T. N. Baral, V. C. Retamozzo, P. De Baetselier, E. De Genst, J. Kinne, H. Leonhardt, S. Magez, V. K. Nguyen, H. Revets, U. Rothbauer, B. Stijlemans, S. Tillib, U. Wernery, L. Wyns, G. Hassanzadeh-Ghassabeh and D. Saerens (2009). "Camelid immunoglobulins and nanobody technology." Vet Immunol Immunopathol **128**(1-3): 178-183.
- Muyldermans, S., C. Cambillau and L. Wyns (2001). "Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains." Trends Biochem Sci **26**(4): 230-235.
- Nakajima, Y., X. M. Qu and S. Natori (1987). "Interaction between liposomes and sarcotoxin IA, a potent antibacterial protein of Sarcophaga peregrina (flesh fly)." J Biol Chem **262**(4): 1665-1669.
- Nguyen, V. K., R. Hamers, L. Wyns and S. Muyldermans (1999). "Loss of splice consensus signal is responsible for the removal of the entire C(H)1 domain of the functional camel IGG2A heavy-chain antibodies." Mol Immunol **36**(8): 515-524.
- Nguyen, V. K., R. Hamers, L. Wyns and S. Muyldermans (2000). "Camel heavy-chain antibodies: diverse germline V(H)H and specific mechanisms enlarge the antigen-binding repertoire." EMBO J **19**(5): 921-930.
- Nguyen, V. K., S. Muyldermans and R. Hamers (1998). "The specific variable domain of camel heavy-chain antibodies is encoded in the germline." Journal of Molecular Biology **275**(3): 413-418.
- Oren, F. (2005). "Convergent evolution of invertebrate defensins and nematode antibacterial factors." Trends in Microbiology **13**(7): 314-319.
- Patil, A. A., Y. Cai, Y. Sang, F. Blecha and G. Zhang (2005). "Cross-species analysis of the mammalian beta-defensin gene family: presence of syntenic gene clusters and preferential expression in the male reproductive tract." Physiol Genomics **23**(1): 5-17.
- Peddinti, D., B. Nanduri, A. Kaya, J. M. Feugang, S. C. Burgess and E. Memili (2008). "Comprehensive proteomic analysis of bovine spermatozoa of varying fertility rates and identification of biomarkers associated with fertility." BMC Systems Biology **2**: 19.
- Peschel, A., M. Otto, R. W. Jack, H. Kalbacher, G. Jung and F. Götz (1999). "Inactivation of the dlt Operon inStaphylococcus aureus Confers Sensitivity to Defensins, Protegrins, and Other Antimicrobial Peptides." Journal of Biological Chemistry **274**(13): 8405-8410.
- Pimentel, E. C., S. Bauersachs, M. Tietze, H. Simianer, J. Tetens, G. Thaller, F. Reinhardt, E. Wolf and S. Konig (2010). "Exploration of relationships between production and fertility traits in dairy cattle via association studies of SNPs within candidate genes derived by expression profiling." Animal genetics.
- Pokorny, A. and P. F. Almeida (2004). "Kinetics of dye efflux and lipid flip-flop induced by delta-lysin in phosphatidylcholine vesicles and the mechanism of graded release by amphipathic, alpha-helical peptides." Biochemistry **43**(27): 8846-8857.
- Proud, D., S. P. Sanders and S. Wiehler (2004). "Human rhinovirus infection induces airway epithelial cell production of human beta-defensin 2 both in vitro and in vivo." J Immunol **172**(7): 4637-4645.

- Quinones-Mateu, M. E., M. M. Lederman, Z. Feng, B. Chakraborty, J. Weber, H. R. Rangel, M. L. Marotta, M. Mirza, B. Jiang, P. Kiser, K. Medvik, S. F. Sieg and A. Weinberg (2003). "Human epithelial beta-defensins 2 and 3 inhibit HIV-1 replication." AIDS **17**(16): F39-48.
- Radhakrishnan, Y., K. G. Hamil, S. Yenugu, S. L. Young, F. S. French and S. H. Hall (2005). "Identification, characterization, and evolution of a primate beta-defensin gene cluster." Genes Immun **6**(3): 203-210.
- Rapaport, D., R. Peled, S. Nir and Y. Shai (1996). "Reversible surface aggregation in pore formation by pardaxin." Biophys J **70**(6): 2502-2512.
- Rathinakumar, R. and W. C. Wimley (2008). "Biomolecular engineering by combinatorial design and high-throughput screening: small, soluble peptides that permeabilize membranes." J Am Chem Soc **130**(30): 9849-9858.
- Robert Ew, H. (1997). "Peptide antibiotics." The Lancet **349**(9049): 418-422.
- Rodriguez-Jimenez, F. J., A. Krause, S. Schulz, W. G. Forssmann, J. R. Conejo-Garcia, R. Schreeb and D. Motzkus (2003). "Distribution of new human beta-defensin genes clustered on chromosome 20 in functionally different segments of epididymis." Genomics **81**(2): 175-183.
- Rohrl, J., D. Yang, J. J. Oppenheim and T. Hehlgans (2010). "Human beta-defensin 2 and 3 and their mouse orthologs induce chemotaxis through interaction with CCR2." J Immunol **184**(12): 6688-6694.
- Rousset, F. (2008). "genepop'007: a complete re-implementation of the genepop software for Windows and Linux." Mol Ecol Resour **8**(1): 103-106.
- Rousset, F. (2008). "genepop'007: a complete re-implementation of the genepop software for Windows and Linux." Molecular Ecology Resources **8**(1): 103-106.
- Royal, M., G. E. Mann and A. P. Flint (2000). "Strategies for reversing the trend towards subfertility in dairy cattle." Vet J **160**(1): 53-60.
- Royal, M. D., A. O. Darwash, A. P. F. Flint, R. Webb, J. A. Woolliams and G. E. Lamming (2000). "Declining fertility in dairy cattle: Changes in traditional and endocrine parameters of fertility." Animal Science **70**(3): 487-501.
- Rozen, S. and H. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers." Methods Mol Biol **132**: 365-386.
- Ryan, L. K., J. Rhodes, M. Bhat and G. Diamond (1998). "Expression of beta-defensin genes in bovine alveolar macrophages." Infect Immun **66**(2): 878-881.
- Sahana, G., B. Guldbrandtsen, C. Bendixen and M. S. Lund (2010). "Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle." Animal genetics **41**(6): 579-588.
- Salvatore, M., A. Garcia-Sastre, P. Ruchala, R. I. Lehrer, T. Chang and M. E. Klotman (2007). "alpha-Defensin inhibits influenza virus replication by cell-mediated mechanism(s)." J Infect Dis **196**(6): 835-843.
- Schuberth, H. J., U. Taylor, H. Zerbe, D. Waberski, R. Hunter and D. Rath (2008). "Immunological responses to semen in the female genital tract." Theriogenology **70**(8): 1174-1181.
- Schutte, B. C., J. P. Mitros, J. A. Bartlett, J. D. Walters, H. P. Jia, M. J. Welsh, T. L. Casavant and P. B. McCray, Jr. (2002). "Discovery of five conserved beta -defensin gene clusters using a computational search strategy." Proc Natl Acad Sci U S A **99**(4): 2129-2133.
- Selsted, M. E. and A. J. Ouellette (2005). "Mammalian defensins in the antimicrobial immune response." Nat Immunol **6**(6): 551-557.
- Selsted, M. E., D. Szklarek and R. I. Lehrer (1984). "Purification and antibacterial activity of antimicrobial peptides of rabbit granulocytes." Infect Immun **45**(1): 150-154.
- Semple, C. A., P. Gautier, K. Taylor and J. R. Dorin (2006). "The changing of the guard: Molecular diversity and rapid evolution of beta-defensins." Mol Divers **10**(4): 575-584.
- Shai, Y. and Z. Oren (1996). "Diastereoisomers of cytolysins, a novel class of potent antibacterial peptides." J Biol Chem **271**(13): 7305-7308.
- Shi, J., S. Aono, W. Lu, A. J. Ouellette, X. Hu, Y. Ji, L. Wang, S. Lenz, F. W. van Ginkel, M. Liles, C. Dykstra, E. E. Morrison and C. O. Elson (2007). "A novel role for defensins in intestinal homeostasis: regulation of IL-1beta secretion." J Immunol **179**(2): 1245-1253.
- Shook, G. E. (2006). "Major advances in determining appropriate selection goals." Journal of Dairy Science **89**(4): 1349-1361.
- Soehnlein, O., A. Zernecke, E. E. Eriksson, A. G. Rothfuchs, C. T. Pham, H. Herwald, K. Bidzhekov, M. E. Rottenberg, C. Weber and L. Lindbom (2008). "Neutrophil secretion products pave the way for inflammatory monocytes." Blood **112**(4): 1461-1471.
- Sorensen, O. E., L. Gram, A. H. Johnsen, E. Andersson, S. Bangsboll, G. S. Tjabringa, P. S. Hiemstra, J. Malm, A. Egesten and N. Borregaard (2003). "Processing of seminal plasma hCAP-18 to ALL-38 by

gastricsin: a novel mechanism of generating antimicrobial peptides in vagina." J Biol Chem **278**(31): 28540-28546.

- Soruri, A., J. Grigat, U. Forssmann, J. Riggert and J. Zwirner (2007). "beta-Defensins chemoattract macrophages and mast cells but not lymphocytes and dendritic cells: CCR6 is not involved." Eur J Immunol **37**(9): 2474-2486.
- Swanson, K., S. Gorodetsky, L. Good, S. Davis, D. Musgrave, K. Stelwagen, V. Farr and A. Molenaar (2004). "Expression of a beta-defensin mRNA, lingual antimicrobial peptide, in bovine mammary epithelial tissue is induced by mastitis." Infect Immun **72**(12): 7311-7314.
- Taylor, U., D. Rath, H. Zerbe and H. J. Schuberth (2008). "Interaction of intact porcine spermatozoa with epithelial cells and neutrophilic granulocytes during uterine passage." Reprod Domest Anim **43**(2): 166-175.
- Thomma, B. P., B. P. Cammue and K. Thevissen (2002). "Plant defensins." Planta **216**(2): 193-202.
- Tollner, T. L., C. L. Bevins and G. N. Cherr (2012). "Multifunctional glycoprotein DEFB126--a curious story of defensin-clad spermatozoa." Nat Rev Urol **9**(7): 365-375.
- Tollner, T. L., S. A. Venners, E. J. Hollox, A. I. Yudin, X. Liu, G. Tang, H. Xing, R. J. Kays, T. Lau, J. W. Overstreet, X. Xu, C. L. Bevins and G. N. Cherr (2011). "A Common Mutation in the Defensin DEFB126 Causes Impaired Sperm Function and Subfertility." Sci Transl Med **3**(92): 92ra65.
- Tollner, T. L., A. I. Yudin, A. F. Tarantal, C. A. Treece, J. W. Overstreet and G. N. Cherr (2008). "Beta-defensin 126 on the surface of macaque sperm mediates attachment of sperm to oviductal epithelia." Biol Reprod **78**(3): 400-412.
- Tomasinsig, L. and M. Zanetti (2005). "The cathelicidins--structure, function and evolution." Curr Protein Pept Sci **6**(1): 23-34.
- Tomasinsig, L. and M. Zanetti (2005). "The cathelicidins--structure, function and evolution." Current protein & peptide science **6**(1): 23-34.
- Uematsu, N. and K. Matsuzaki (2000). "Polar angle as a determinant of amphipathic alpha-helix-lipid interactions: a model peptide study." Biophys J **79**(4): 2075-2083.
- van der Linden, R., B. de Geus, W. Stok, W. Bos, D. van Wassenaar, T. Verrips and L. Frenken (2000). "Induction of immune responses and molecular cloning of the heavy chain antibody repertoire of Lama glama." Journal of Immunological Methods **240**(1-2): 185-195.
- Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe and F. Speleman (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." Genome Biology **3**(7): research0034.0031 - research0034.0011.
- Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe and F. Speleman (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." Genome Biol **3**(7): Research0034.
- Waberski, D., R. Claassen, T. Hahn, P. W. Jungblut, N. Parvizi, E. Kallweit and K. F. Weitze (1997). "LH profile and advancement of ovulation after transcervical infusion of seminal plasma at different stages of oestrus in gilts." Journal of Reproduction and Fertility **109**(1): 29-34.
- Waberski, D., A. Dohring, F. Ardon, N. Ritter, H. Zerbe, H.-J. Schuberth, M. Hewicker-Trautwein, K. Weitze and R. Hunter (2006). "Physiological routes from intra-uterine seminal contents to advancement of ovulation." Acta Veterinaria Scandinavica **48**(1): 13.
- Wang, T.-T., F. P. Nestel, V. Bourdeau, Y. Nagai, Q. Wang, J. Liao, L. Tavera-Mendoza, R. Lin, J. H. Hanrahan, S. Mader and J. H. White (2004). "Cutting Edge: 1,25-Dihydroxyvitamin D3 Is a Direct Inducer of Antimicrobial Peptide Gene Expression." The Journal of Immunology **173**(5): 2909-2912.
- Wang, W., S. M. Owen, D. L. Rudolph, A. M. Cole, T. Hong, A. J. Waring, R. B. Lal and R. I. Lehrer (2004). "Activity of alpha- and theta-defensins against primary isolates of HIV-1." J Immunol **173**(1): 515-520.
- Wehkamp, J., M. Koslowski, G. Wang and E. F. Stange (2008). "Barrier dysfunction due to distinct defensin deficiencies in small intestinal and colonic Crohn's disease." Mucosal Immunol **1 Suppl 1**: S67-74.
- Wilson, C. L., A. J. Ouellette, D. P. Satchell, T. Ayabe, Y. S. López-Boado, J. L. Stratman, S. J. Hultgren, L. M. Matrisian and W. C. Parks (1999). "Regulation of Intestinal α-Defensin Activation by the Metalloproteinase Matrilysin in Innate Host Defense." Science **286**(5437): 113-117.
- Yang, D., O. Chertov, S. N. Bykovskaia, Q. Chen, M. J. Buffo, J. Shogan, M. Anderson, J. M. Schroder, J. M. Wang, O. M. Howard and J. J. Oppenheim (1999). "Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6." Science **286**(5439): 525-528.
- Yang, L., T. M. Weiss, R. I. Lehrer and H. W. Huang (2000). "Crystallization of antimicrobial pores in membranes: magainin and protegrin." Biophys J **79**(4): 2002-2009.

- Yeaman, M. R. and N. Y. Yount (2003). "Mechanisms of antimicrobial peptide action and resistance." Pharmacol Rev **55**(1): 27-55.
- Yenugu, S., V. Chintalgattu, C. J. Wingard, Y. Radhakrishnan, F. S. French and S. H. Hall (2006). "Identification, cloning and functional characterization of novel beta-defensins in the rat (Rattus norvegicus)." Reprod Biol Endocrinol **4**: 7.
- Yu, J., N. Mookherjee, K. Wee, D. M. Bowdish, J. Pistolic, Y. Li, L. Rehaume and R. E. Hancock (2007). "Host defense peptide LL-37, in synergy with inflammatory mediator IL-1beta, augments immune responses by multiple pathways." J Immunol **179**(11): 7684-7691.
- Yudin, A. I., S. E. Generao, T. L. Tollner, C. A. Treece, J. W. Overstreet and G. N. Cherr (2005). "Beta-defensin 126 on the cell surface protects sperm from immunorecognition and binding of anti-sperm antibodies." Biol Reprod **73**(6): 1243-1252.
- Yudin, A. I., T. L. Tollner, C. A. Treece, R. Kays, G. N. Cherr, J. W. Overstreet and C. L. Bevins (2008). "Beta-defensin 22 is a major component of the mouse sperm glycocalyx." Reproduction **136**(6): 753-765.
- Zanich, A., J. C. Pascall and R. Jones (2003). "Secreted epididymal glycoprotein 2D6 that binds to the sperm's plasma membrane is a member of the beta-defensin superfamily of pore-forming glycopeptides." Biol Reprod **69**(6): 1831-1842.
- Zeya, H. I. and J. K. Spitznagel (1963). "Antibacterial and Enzymic Basic Proteins from Leukocyte Lysosomes: Separation and Identification." Science **142**: 1085-1087.
- Zhang, L., W. Yu, T. He, J. Yu, R. E. Caffrey, E. A. Dalmasso, S. Fu, T. Pham, J. Mei, J. J. Ho, W. Zhang, P. Lopez and D. D. Ho (2002). "Contribution of human alpha-defensin 1, 2, and 3 to the anti-HIV-1 activity of CD8 antiviral factor." Science **298**(5595): 995-1000.
- Zhou, C. X., Y. L. Zhang, L. Xiao, M. Zheng, K. M. Leung, M. Y. Chan, P. S. Lo, L. L. Tsang, H. Y. Wong, L. S. Ho, Y. W. Chung and H. C. Chan (2004). "An epididymis-specific beta-defensin is important for the initiation of sperm maturation." Nat Cell Biol **6**(5): 458-464.
- Zhu, S. and B. Gao (2013). "Evolutionary origin of β-defensins." Developmental and comparative immunology **39**(1-2): 79-84.
- Zouhir, A., R. Hammami, I. Fliss and J. Hamida (2010). "A new structure-based classification of gram-positive bacteriocins." The protein journal **29**(6): 432-439.