DEVELOPMENT OF IMMUNOGLOBULIN DIVERSITY AT THE HEAVY CHAIN LOCUS OF CATTLE

SUBHASH VERMA
B.V.Sc & A.H., M.V.Sc. (Vet. Microbiology),
M.S. (Immunopharmacology)

Division of Infection and Immunity
Institute of Biomedical and Life Sciences
University of Glasgow

A thesis submitted for the degree of
Doctor of Philosophy
January 2007

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I express my deepest gratitude to my research supervisor, Dr. Rob Aitken for his excellent supervision and guidance throughout the course of this work. His trust, generosity, honesty and professionalism all inspired me deeply. I could not have wished for a better supervisor.

I am grateful to Dr. Mark Dagleish for providing the animal tissues, Dr. Norma Morrison, Dr. Tom Goldammer and Dr. Olwyn Byron for collaboration and Prof. Tim Mitchell for allowing the use the facilities in his lab.

Many thanks to all people in the lab and office for sharing good times, especially to Mohammad for taking part in theological and political discussions which helped me to relax from the stress of writing.

I owe special thanks to my wife, Geetanjali and our daughter, Suryakshi for providing a loving environment during all these years. Suryakshi for being so very loving and understanding; and Geetanjali for helping me get through the difficult times, for all the love and emotional support despite being busy with her own Ph.D.

I am forever indebted to my parents for their love and blessings and promoting good values in me. I dedicate this thesis to them.

Finally, I thank the University of Glasgow, Faculty of Biomedical and Life Sciences for providing University Postgraduate Research Scholarship and Universities UK for its Overseas Research Scholarship.

Subhash Verma

AUTHOR'S DECLARATION

I declare that this thesis and the results presented in it are entirely my own work, except where otherwise stated.

Subhash Verma

ABSTRACT

Immunoglobulins (Ig) are formed by combinatorial recombination of a set of germline genes and undergo mutation in B-lymphocytes as they develop and mature. Combinatorial assortment can result in enormous diversification of the primary Ig repertoire in species where the number and variability within these segments is high (e.g. humans and mice). Many other species are unable to diversify the primary repertoire through rearrangement. In these animals, somatic hypermutation and gene conversion play a major role in driving pre-immune diversity.

The bovine humoral immune system is not capable of generating a significant level of heavy chain Ig diversity through combinatorial assortment due to the small size of the IgH gene family and the low diversity apparent within the CDRs of the V_H segments of this family. To learn more of the molecular processes responsible for diversification, this study hypothesized that somatic hypermutation would introduce nucleotide substitutions throughout the Ig reading frame including the J_H-Cμ intron whereas if gene conversion were the dominant process behind diversification, modifications would be confined to the Ig reading frame. To distinguish these possibilities, Ig heavy chain sequences were recovered from the rearranged locus using lymphoid tissues of cattle of different ages for sequence analysis. This also allowed determination of the timing of Ig diversification and comparison of the extent of diversification.

Analysis revealed that single base substitutions predominated, with purines targeted more frequently than pyrimidines and transitions favoured over transversions. Seventeen deletions spanning 1 to 26 nucleotides and 5 insertions in the range of 1 to 3 bases were also observed. As would be expected, mutational hotspots were encountered in CDR1 (complementary determining region) and CDR2 and the sequences downstream of FR4. The modified region extended into unutilized parts of the J_H locus and downstream intron and with mutations occurring as frequently as in CDR2. The frequency of mutation decreased over the 579 bases lying 3' to the rearranged VDJ gene. Therefore, the data were consistent with the predicted pattern of somatic hypermutation. The number of potential donors for diversification of the reading frame by gene conversion is not known: formally, this could act in parallel with somatic

hypermutation but it is considered unlikely. Sequences recovered from a very young animal had considerably fewer nucleotide substitutions than those of the older age groups (7 weeks [2.1 substitutions per sequence], 13 weeks [3.3] and adult [3.43]) (7, 13 and adult) but low-level diversification was detectable indicative of early onset. The lower rate of substitutions observed in FRs compared to CDRs presumably reflects the need to conserve Ig folding and structural stability.

Somatic hypermutation, gene conversion and Ig class switch recombination have been shown to have a common dependence upon activation induced cytidine deaminase (AID). The study cloned and characterized AID from cattle. The sequence data confirmed the high degree of conservation of protein coding sequence between the bovine gene and multiple other species with much greater divergence apparent in introns. The bovine AID cDNA was found to consist of an open reading frame encoding 199 amino acids residues. Conservation was especially high within the cytidine deaminase and CSR motifs which have been shown to be functionally important for the activity of AID. AID expression was observed in a range of lymphoid tissues to levels higher that in non-lymphoid tissues, the pattern being similar to that reported for canines and humans. The enzyme was shown to have mutagenic activity through heterologous expression in *E. coli*. This increased the frequency of mutations in *lac*Z and *rpo*B target genes. In *rpo*B, rifampicin resistance appeared in AID-expressing *E. coli* with a frequency that was 2-6 times higher that of controls. These findings confirmed that the bovine AID cDNA encoded a protein with biological activity. In collaborative studies, using cattle-hamster somatic cell hybrids, AID was assigned to BTA 5 (*Bos taurus* autosome) close to the marker, MAGP2 (microfibril associated glycoprotein).

A further aspect of this study used a bovine lymphoma line (BL-3) in an attempt to establish as *in vitro* model of Ig diversification. BL-3 cells were shown to express functional, rearranged IgM and there was no evidence of constitutive diversification or class switch to IgG in these cells. The cells expressed CD40 marker but were negative for CD21, marker for mature B cells. Expression of IL-4Rα transcript was detected but RT-PCR failed to recover AID transcripts from unstimulated BL-3 cells. The exposure of BL-3 cells to LPS, IL-4,

and anti-CD40 either alone or in combination did not result in any upregulation of AID transcript. The cells were totally resistant to transfection by lentiviruses and adenoviruses and lipofectamine yielded only 2-3% transfectants. Hence, the introduction of an AID transgene into BL-3 cells to assess the effects on IgH locus appears problematic if not impossible. In collaborative studies, transcripts of the pattern recognition receptors TLR4 and TLR9 were detected. As activation through PAMPs is capable of stimulating B cells it may be possible to exploit these branches of the innate immune response to stimulate expression of AID and assess its effects on Ig diversity and synthesis.

Finally, phage display was used to develop reagents for discrimination of the allotypes of bovine IgM that arise from rearrangement of BTA11 and BTA21. This ability to form variants of an Ig from separate chromosomes appears a further unusual property of the bovine humoral system. Phage clones able to bind to Cμ1 were isolated but they appeared to lack the ability to distinguish the subtle difference between the IgM variants.

ABBREVIATIONS

aa	amino acid
AID	activation induced cytidine deaminase
BAC	bacterial artificial chromosome
BL-3	bovine lymphosarcoma
BRECs	B cell recombination excision circles
BSA	bovine serum albumin
BTA	*Bos taurus* autosome
cDNA	complementary DNA
CDR	complementarity determining region
CFU	colony forming units
CSR	class switch recombination
CV	column volumes
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DSB	double strand breaks
EDTA	ethylene diamine tetra-acetic acid
ELISA	enzyme linked immunosorbant assay
FACS	fluorescence activated cell sorting
FBS	foetal bovine serum
FISH	fluorescence in situ hybridisation
FITC	fluorescein isothiocyanate
FR	framework region
g	gravity
GALT	gut-associated lymphoid tissue
GC	gene conversion
GFP	green fluorescent protein
H chain	heavy chain
HRPO	horse radish peroxidase
Ig	immunoglobulin
IMAC	immobilised affinity chromatography
IPP	ileal's Peyer's patches
IPTG	Isopropyl-β-D-thiogalactopyranoside
JPP	jejunal peyer's patches
kb	kilobase
l/ml/µl	litres/millilitres/microlitres
LB	Luria-Bertini
LPS	lipopolysaccharide
M.wt.	molecular weight
M/mM/µM/nM	molar/millimolar/micromolar/nanomolar
mA	milliamps
MAB	monoclonal antibody
MAR	matrix attachment regions
min	minutes
MNCs	mono-nuclear cells
MPBS	milk phosphate buffered saline
mRNA	messenger RNA
PBS	phosphate buffered saline
PCR	polymerase chain reaction
RAG	recombination-activation genes

RNA	ribonucleic acid
rpm	revolutions per minute
RSS	recombination signal sequence
RT	room temperature
RT-PCR	reverse transcriptase polymerase chain reaction
scAb	single chain antibody
scFv	single chain antibody fragment
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SHM	somatic hypermutation
SNAP	synonymous nonsynonymous analysis program
SSB	single strand breaks
TAE	Tris-acetic acid EDTA
TB	terrific broth
Tdu/ml	transduction units/ml
U	units
UV	ultraviolet
V	variable
v/v	volume/volume
VH	variable heavy chain
VL	variable light chain
WGRH	whole genome radiation hybrids
w/v	weight/volume

Chapter: 1 Development of Immunoglobulin diversity

1.1: GENERAL INTRODUCTION

One of the most remarkable features of the vertebrate immune system is its ability to respond to apparently limitless foreign antigens by defensive measures that are both specific and non-specific in nature. The non-specific component, innate immunity, is a set of disease resistance mechanisms that are not specific to a particular pathogen and can be triggered by molecules such as lipopolysaccharides, lipoteichoic acids, peptidoglycan and hypomethylated CpG DNA. These products are perceived as foreign and their recognition by host the innate immune system can signal the presence of infection (Janeway, 1989, , 1992). On the other hand, the specific component, adaptive immunity, displays a high degree of antigenic discrimination, diversity, the ability to distinguish between closely-related molecules that are of foreign rather than self origin, and the unique property of memory. Adaptive and innate immune responses do not operate independently of each other but work in a highly interactive and co-operative manner to mount a response which is much more effective than either could do alone.

Adaptive immunity can be divided into humoral and cell-mediated responses. Antibody functions as the effector of humoral responses by binding to antigen and neutralising it or facilitating its elimination whereas effector T-lymphocytes generated in response to antigen are responsible for cell mediated immunity. Although there are intrinsic differences in the specific structure of immunoglobulins (Igs), generally speaking, most immunoglobulins have a four chain structure composed of two identical light chains and two identical heavy chains. Both the heavy and light chains are divided into two regions (variable and constant) based on variability in the amino acid sequences. As immunoglobulin (Ig) sequence data became available, it was found that virtually each individual antibody molecule carried a unique amino acid sequence in its variable region but only one of a limited number of invariant sequences in its constant region. The genetic basis of such tremendous variation coupled with consistency in a single protein molecule was found to lie in the organisation of the Ig genes.

1.1.1: Organisation of immunoglobulin germline gene segments

All Ig genes in mammals are assembled during B lymphocyte development from incomplete segments distributed at the Ig loci. Variable (V) genes from many families are present upstream of many diversity (D) (only present at the heavy chain locus) and joining (J) segments (Bengten *et al.*, 2000). The numbers of segments able to contribute directly to antibody formation (functional segments versus non-functional, or pseudogenes) varies between species.

The human IgH locus at 14q32.33 spans 1250 kilobases (kb) and consists of 123 to 129 V_H genes, depending on the haplotype, 27 D_H segments belonging to 7 subgroups, 9 J_H segments, and 11 C_H genes (Lefranc, 2001a; Lefranc, 2003). Eighty-two to 88 V_H genes belong to 7 subgroups, whereas 41 pseudogenes, which are too divergent to be assigned to subgroups, have been assigned to 4 clans. Seven non-mapped V genes have been described as insertion/deletion polymorphisms but have not yet been precisely located (Lefranc, 2001a; Lefranc, 2003). In chickens, there is extreme polarization and the numerous segments present at the heavy and light chain loci are almost all pseudogenes, there being just single functional segments present at the J proximal termini.

The light and heavy chains loci are situated on different chromosomes e.g. in mice λ, κ and heavy chain loci are located on chromosome 16, 6 and 12, respectively and in the case of humans λ and κ chains are located on chromosomes 22 and 2, whereas the heavy chain locus is found on chromosome 14 (Goldsby *et al.*, 2003). The numbers of gene segments and gene families varies in different species.

The human κ locus carries 76 Vκ genes (Scaviner *et al.*, 1999; Zachau, 1993) falling into 7 subgroups, 5 Jκ segments (Hieter *et al.*, 1982; Scaviner *et al.*, 1999) and only a single Cκ gene (Hieter *et al.*, 1980). The 76 Vκ genes are organized in two clusters separated by 800 kb. The Vκ distal cluster spans 400 kb and comprises 36 genes whereas the Vκ proximal cluster (in the 3' region of the locus, closer to Cκ, and in the most telomeric position) spans 600 kb and comprises 40 genes (Lefranc, 2001b). Vκ-gene segment codes first 95 amino acids for the V region of the antibody light chain; the Jκ gene segment codes for the rest of the region. The λ locus in humans contains a set of 73 to 74 V genes

(Pallares *et al.*, 1998), but each of seven C genes is accompanied by just one J gene (Figure 2). Rearrangement thus creates continuous sequence encoding the whole of the light chain variable region. In the light chain, the V gene segment encodes framework (FR) 1 to FR3 and the amino terminus of third antigen binding or complementarity-determining region (CDR3). The J segment encodes the carboxy terminal of CDR3 and FR4. The rearranged VJ and VDJ gene segments encode the variable region of the antibody which determines the binding specificity of the molecule. A leader exon sits at the 5' terminus of each V gene segment and codes for a signal peptide. This is not present in the mature immunoglobulin molecule (Figure 1).

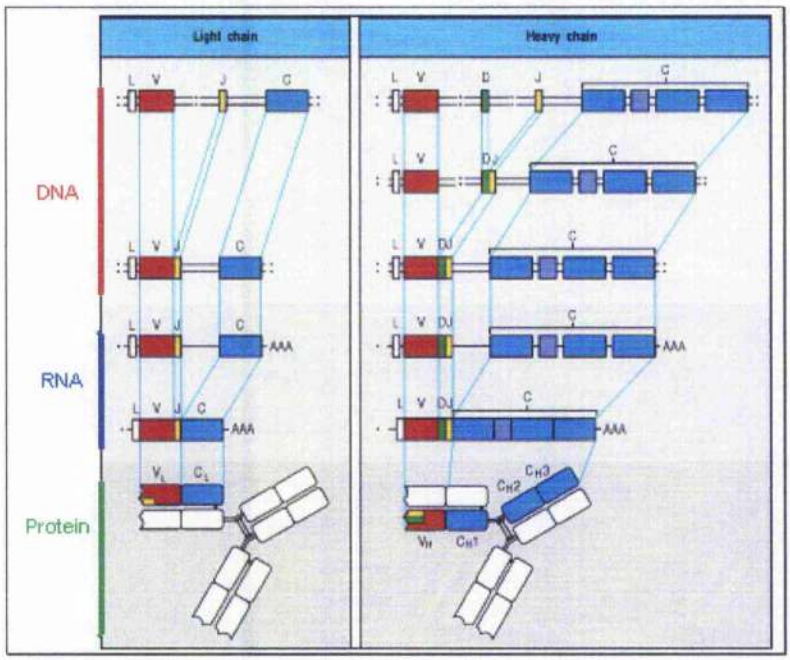


Figure 1: Schematic representation of Ig light and heavy chain rearrangement

Light chain V region genes are constructed from two segments (left panel). A variable (V) and a joining (J) gene segment in the genomic DNA are joined to form a complete light chain V region exon. A leader peptide (L) is cleaved off after directing the protein into the cell's secretory pathways. The light chain C region is encoded in a separate exon and is joined to the V region exon by splicing of the light chain RNA. Heavy chain V regions are constructed from three gene segments (right panel). First, the diversity (D) and J gene segments join, then the V gene segment joins to the combined DJ sequence, forming a complete VH exon. This figure has been adapted from Immunobiology, 5th ed. (Goldsby *et al.*, 2003).

1.1.2: Generation of immunoglobulin diversity

It has been estimated that the mammalian immune system can generate more than 10^{10} different antibodies (Goldsby *et al.*, 2003). It follows from this that antibody structure must arise from a genetic system comprised of

immunoglobulin genes capable of generating enormous diversity. These immunoglobulin genes are recombined, rearranged, and mutated by different molecular processes in B-lymphocytes as they develop and mature. Different events have been documented as generators of this diversity. These include: the selection of a gene from an initial set of multiple germline genes; rearrangement of these genes; junctional diversity arising from imprecise fusion and the addition of non-templated sequence; different independent combinations of heavy and light chain; gene conversion and somatic hypermutation. The principles by which functional genes are assembled are the same at the Ig heavy and light chain loci but as described, there are differences in the details of the organisation of the V and C gene segments and correspondingly of the recombination reaction between them. Most significantly, the heavy chain locus includes diversity (D) gene segments between the V and J segments (Figure 2). The diversity created by rearrangement is directly related to the number of non-identical segments available and the extent of imprecision in the rearrangement process.

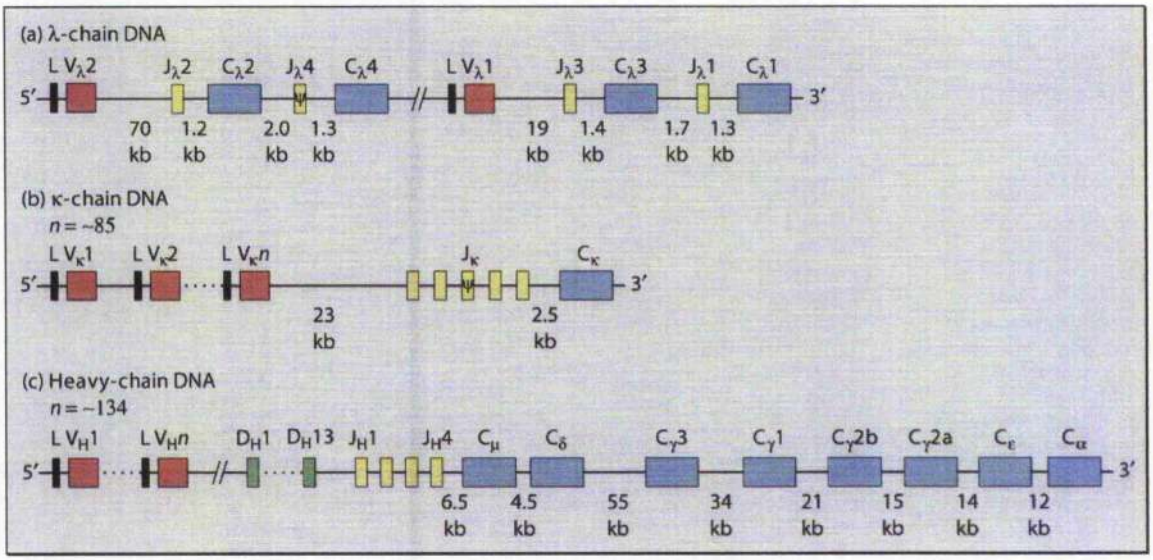


Figure 2: Multigene organisation of germline gene segments in mouse

The λ and κ light chains are encoded by V, J and C gene segments. The heavy chain family contain additionally D segments which encodes for additional CDR3 diversity. This figure has been adapted from Immunobiology, 5th ed. Janeway, Charles A.; Travers, Paul; Walport, Mark; Shlomchik, Mark.

1.1.3: Control of rearrangement and junctional flexibility

A unique recombination signal sequence (RSS) is located 3' to each V gene segment, 5' to each J gene segment, and on both flanks of each D segment. The function of the RSS is to guide the recombination process responsible for rearrangement. The RSS consists of 3 elements: a heptamer, spacer of 12-bp or 23-bp and a nonamer. The heptamer and nonamer are well-conserved whereas spacer is non-conserved (Tonegawa, 1983). The rearrangement occurs between segments with an RSS carrying a 12 bp spacer sequence, and an RSS with a 23 bp spacer sequence (the `12/23 rule') (Bogue and Roth, 1996). Some endogenous RSSs differ considerably from the optimal consensus RSS used in most biochemical analyses of V(D)J recombination and this sequence variation can have important effects on the efficiencies with which V, D, and J gene segments are used. However, not all recombination events actually occur with 12/23 rule and some recombinations do occur 'beyond 12/23 rule' (Oettinger *et al.*, 1990). The VJ and VDJ recombination takes place at the junctions between RSSs and coding segments and is brought about by a group of enzymes referred to as the V(D)J recombinase. V(D)J recombination is initiated when a complex containing the proteins of the recombination-activation genes (RAG)-1 and RAG-2 bind to RSS (Oettinger *et al.*, 1990). The initial DNA break is performed by the RAG1-RAG2 complex, which binds to the RSS and makes a single-strand cut at the 5' end of the heptamer at the precise border of the coding sequence and the RSS (McBlane *et al.*, 1995). This results in the formation of hairpin loops at the cut end of the coding sequence and a flush, 5'-phosphorylated, double-strand break at the signal sequence (Goldsby *et al.*, 2003). To proceed through the recombination process, the DNA hairpins formed by the coding ends have to be resolved into free ends, which are presumably done by double- or single strand nicking of the hairpin. The position of the nick(s) play a key role in the formation of junctional diversity (Nadel and Feeney, 1995). Nicking on only one strand of the hairpin results in a 3' or 5' protrusion carrying palindromic sequence known as P nucleotides (Lieber, 1991; Meier and Lewis, 1993). P regions are usually 1 or 2 nucleotides long, but P additions of 3 to 4 nucleotides are not unusual (Schuler *et al.*, 1991). Nuclease action can trim back the free DNA strands, deleting nucleotides from the coding sequence.

Variable-region coding joints in rearranged immunoglobulin genes have been shown to contain amino acid sequences that are not encoded in germline segments. The diversity generated as a result of V, D and J rearrangement is further increased by the editing of terminal deoxynucleotidyl transferase (TdT) (Landau *et al.*, 1987; Silverstone *et al.*, 1976). The RAG-1 and RAG-2 proteins and TdT enzyme are the only lymphoid-specific gene products that have been shown to be involved in V (D) J rearrangements. Once coding ends are opened up during recombination the addition of nontemplated (N) nucleotides by TdT can take place (Landau *et al.*, 1987). The potential diversity arising from N addition is quite large owing to the random nature of the process. This junctional flexibility leads to many non-productive rearrangements, but when the reading frame is retained alternative amino acids are created at coding joints thus augmenting diversity.

Immunoglobulin gene rearrangement occurs in an orderly fashion, IgH locus undergoing rearrangement first and then followed by light chain rearrangement. Allelic exclusion further ensures that the functionally mature B cell has a monospecific antigenic specificity. This means that once productive rearrangement has occurred at heavy and light chain loci, the recombination machinery is no longer functional in that particular B-cell and homologous chromosomes carrying the alternative allele are not expressed.

There are numerous aspects of immunoglobulin structure and function in vertebrates that are different from the conventional view derived from studies of mice and humans. The following sections highlight some differences that are of relevance to Ig diversification to cattle. Some of theses features are listed in Table 1.

Table 1: Features of Ig immune repertoire formation in vertebrates

Species	*Evolution (Mya)*	*Organ of diversification*	*Locus configuration*	*Diversification mechanism*	*Rearrangement throughout life?*	*Immunoglobulin isotype*	*Class switch*	*Germinal centre formation*
Mouse, human	100	Foetal liver. bone marrow	Translocon, V_H families	Rearrangement (functional), many V, (D), J genes (combinatorial)	Yes	IgM, IgD, IgG, IgE, IgA	Yes	Yes
Rabbit	100	Appendix (other GALT sites)	Translocon	Rearrangement*, single V_H modified by GC or SHM (V_L)	No	IgM, IgD, IgG, IgE, IgA	Yes	Yes
Sheep	100	IPPs	Translocon	Rearrangement*, few V_L modified by SHM (V_H?)	No	IgM, IgD, IgG, IgE, IgA	Yes	Yes
Chicken	200	Bursa of Fabricius	Translocon	Rearrangement*, single V_H and V_L modified by GC	No	IgM, IgY, IgA	Yes	Yes
Bony fish	400	Pronephros	Translocon (H), cluster (L), V_H families	Rearrangement, many V, (D) and J genes	Yes+	IgM, IgD	No	No
Cartilaginous fish	460	Leydig and/or epigonal organs	Cluster	Rearrangement, many V genes^	Yes+	IgM, IgW, IgNAR	No	No

* Only rearranged V genes become diversified by GC or SHM, but the junctional diversity is limited. + This feature is presumed because of the persistence of RAG expression throughout life in the primary lymphoid tissue. ^ Although the cartilaginous fish VH are of a single family, CDR1 and CDR2 are quite different in VH from different clusters (Shen *et al.*, 1996). Both IgG and IgE are believed to be derived from an IgY-like ancestor. IgNAR- immunoglobulin new antigen receptor; Mya- millions of year ago. This table has been adapted from Figure 1 and Table 1 of a review (Flajnik, 2002)

1.1.4: Somatic hypermutation (SHM)

SHM results in the alteration of a rearranged immunoglobulin sequence by the introduction of nucleotide changes during the lifetime of a B cell. Individual nucleotides are substituted resulting in the alteration of the specificity of the encoded immunoglobulin. This mechanism is in large part responsible for the modification of the antibody repertoire (Rudikoff *et al.*, 1984) and normally this process occurs within the germinal centres driven by the presence of antigen.

It was originally postulated that SHM contributed in part or entirely to antibody diversification (Brenner and Milstein, 1966; Lederberg, 1959). It has become clear over the years that this is not the case in humans and laboratory rodents. In these animals, rearrangement and associated processes seem to be sufficient to populate the immune system with B cells of adequate diversity. However, in animals such as sheep (Reynaud *et al.*, 1995), rabbits (Weinstein *et al.*, 1994), and frogs (Wilson *et al.*, 1992) SHM has emerged as an important process of diversification in the formation of the pre-immune (primary) repertoire. More primitive organisms such as members of the shark family also diversify their immunoglobulin genes by SHM (Hinds-Frey *et al.*, 1993). Sequence analysis of H chain cDNA derived from the spleen of catfish has also shown that somatic mutation occurs within both the V_H- and J_H-encoded regions (Yang *et al.*, 2006). Whether driven through antigen-dependent or independent mechanisms, analysis of somatic mutation has revealed that mutations are largely confined to the variable domains and are rarely found in the constant domain (Altenburger *et al.*, 1981; Ziegner *et al.*, 1994). Mutations can be observed throughout this part of the reading frame but are most pronounced in the CDRs (Berek and Milstein, 1987). The promoter-proximal boundary for mutation appears to be a site within the leader intron (Both *et al.*, 1990) although changes have been observed upstream of the transcriptional start site. The distal boundary is imprecise but mutations in the J-C intron occur at a lower frequency than in the V region (Lebecque and Gearhart, 1990). In the mouse, mutations have been observed in the $C_\lambda 1$ gene (Motoyama *et al.*, 1991). This might be due to the proximity of the J and C genes at the mouse λ locus. Mutations are typically single nucleotide substitutions, with transitions appearing more commonly than transversions. The mutation of G residues more often than

C, and A more than T residues indicates strand polarity. Furthermore, mutations are targeted to local DNA sequences are referred as hotspots, as compared to some regions that are rarely targeted and are called coldspots. Mutational hotspots are exemplified by DNA sequence motifs RGYW/WRCY where R represents A or G, Y represents C or T and W represents A or T. The hotspot in G: C is underlined. This motif appears to be targeted on both DNA strands. Another motif is WA/TW (the hotspot in A:T is underlined), but this appears to attract mutation on one DNA strand (Milstein *et al.*, 1998; Rogozin and Kolchanov, 1992). Very recently, DGYW/WRCH (D=A/G/T, H=T/C/A) has been shown to be a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif putting in context the involvement of AID (Rogozin and Diaz, 2004).

Hypermutation may first have arisen early in evolution to diversify the primary immunoglobulin repertoire when only limited combinatorial diversity was available. In species such as humans, repeated gene duplication and divergence has allowed diversification through rearrangement to emerge and hence the hypermutation mechanism has been reduced so as to play no role in primary diversification. Instead, SHM is controlled so that it intervenes in the germinal centre reaction to drive affinity maturation in the presence of antigen.

1.1.5: Gene conversion (GC)

This is a nonreciprocal transfer of genetic information from donor to recipient immunoglobulin gene(s) based on homologous recombination. During this process the donor segment remains unaltered and the recipient acquires a recombined genotype. This type of Ig diversifying mechanism is seen prominently in the chicken and results in the modification of rearranged V genes in the bursa of Fabricius. In turn this generates a diverse immunoglobulin repertoire. The mechanism is discussed in more detail in a later section.

1.1.6: Mechanisms of immunoglobulin gene diversification in sheep

Studies in sheep suggested that a single V_H family consisting of nine V_H segments was utilised in the formation of the IgH repertoire (Jenne *et al.*, 2006). More recent studies have identified a number of V_H segments that did not

match any of the published germline segments (Charlton *et al.*, 2000; White *et al.*, 1998) and were so divergent from V_H segments reported initially that they were classified into an additional eight families. Whatever the final position, the sheep V_H locus contains a limited number of V_H segments compared to those present at the V_H loci of mice and humans. None of the D_H segments have been sequenced in sheep, however from the analysis of rearranged CDR3 in sheep it appears that these segments consist of short and heterogeneous sequences. The J_H locus comprises six J_H segments, two of which are functional (Dufour and Nau, 1997). Both are utilised in sheep during IgH rearrangement.

The information on the Vκ locus is limited; only six germline Vκ genes have been reported. Kappa chains appear in only 20 to 25% of mature Ig molecules only (Griebel *et al.*, 1992). In contrast, the V_λ locus contains more than 100 germline segments divided into six V_λ families. Analysis has revealed that they possess little diversity and thus are unable to generate significant diversity to the combinatorial repertoire (Jenne *et al.*, 2006). The Jκ locus consists of three segments, two of which are functional, whereas J_λ possess two segments and both rearrange functionally. In summary, although a large number of Ig germline segments have been identified recently in sheep, they have very similar nucleotide sequences and thus overall capacity for Ig diversity through combinatorial rearrangement is limited (Jenne *et al.*, 2006).

The Ileal's peyer's patches (IPPs) were first described as aggregates of mucous secreting tissue by Johann Conrad Peyer in 1677 as cited by Griebel and Hein (Griebel and Hein, 1996). With the passage of time, the work of numerous researchers including that of Reynolds and associates identified these as a site of B cell proliferation and differentiation and the site where B cells undergo primary antibody repertoire expansion (Griebel and Hein, 1996).The exact nature of the B cell precursors which populate the organ in the sheep is not known. The IPPs follicles are oligocolonal and do not support ongoing Ig gene rearrangement (Reynaud *et al.*, 1991a). As indicated, combinatorial mechanisms do not contribute greatly to the overall diversity of the Ig gene repertoire in this species (Jenne *et al.*, 2006; Reynaud *et al.*, 1991c). The limited repertoire established through rearrangement is diversified through a process of somatic hypermutation (Reynaud *et al.*, 1997). In sheep, the neonatal repertoire

essentially represents what can be achieved through rearrangement (Reynaud *et al.*, 1995). The process of pre-immune repertoire expansion takes place in the IPPs and substitutions build up over a period of time often targeting hotspots in the CDRs resulting in the pattern of transitions over transversions consistent with the features of SHM. The analysis of rearranged Ig genes recovered from foetal sheep has indicated that SHM occurs at a low level before birth (Jenne *et al.*, 2003). Following birth and exposure to exogenous antigens, a rapid build up of mutations in the Ig genes is seen (unpublished work cited in (Jenne *et al.*, 2006)). However, experiments have revealed that the exclusion of external antigens from the system (for example, when a section of lamb intestine was removed and ligated to peritoneal cavity) does not reduce the extent of mutations (Reynaud *et al.*, 1995), suggesting that somatic hypermutation may not always be linked to antigen-driven selection. A review from Reynold's group (Jenne *et al.*, 2006) reported that B cells in the IPP undergo very high mutational frequencies and that in apoptotic B cells this frequency is 5 times less in the V gene segment as compared to positively selected cells. These authors hypothesised that for a B cell to survive and emigrate from the sheep IPP it must mutate its genes, but if it fails to do so adequately, it undergoes apoptosis and dies *in situ* (Jenne *et al.*, 2006). Whilst diversification in sheep shows many features of a SHM-mediated process, there are large number of Ig pseudogenes and a possible role for gene conversion in the diversification of the Ig gene repertoire has been proposed (Reynaud *et al.*, 1997; Reynaud *et al.*, 1995), but convincing evidence has yet to be gathered.

1.1.7: Immunoglobulin gene diversification in rabbits

Rabbit B cells are formed in the liver and omentum during foetal life. Shortly before birth, the site of production switches to bone marrow (Mage *et al.*, 2006). The evidence for Ig gene rearrangement in the foetus comes from detection of B cell recombination excision circles (BRECs) in bone marrow at 12 days gestation (Tunyaplin and Knight, 1995). A few months after birth, B lymphopoiesis appears to decrease rapidly as pro B or pre B cells, and BRECs were not found in any haematopoietic tissues tested (Jasper *et al.*, 2003). However, the demonstration of germinal centres containing B cells with

undiversified Ig genes in the spleen of adult rabbits may be an indication that B lymphopoiesis can occur in older animals (Sehgal *et al.*, 1998).

Although there are more than 100 V_H gene segments available within the rabbit Ig heavy chain locus, many of which appear to be potentially functional (Currier *et al.*, 1988), the 3'-most V_H gene segment, V_H1, is utilised in 80-90% of rearrangements (Knight, 1992). Most of the remaining 10-20% of rearrangements utilise only two other V_H gene segments, V_Hx and V_Hy (Friedman *et al.*, 1994). The information on the V_L gene usage in not available, but it is predicted that variety of Vκ and V_λ gene segments appear to occur in rabbit based on the presence of several V_L sequences found in expressed cDNA sequences (Mage *et al.*, 2006). It is thus apparent that the capacity of the system to generate diversity (at least at the IgH locus) through rearrangement in severely limited.

As in several other species, it appears that events in or around the intestine are crucial to Ig diversification. B cells leave the bone marrow to seed the gut-associated lymphoid tissue (GALT) sites including the appendix and sacculus rotundus, which then under the challenge of exoantigens and intestinal microbiota expand to form B cell follicles (Butler, 1997). Novel experiments in rabbits raised under germfree conditions or rabbits manipulated surgically to prevent the interaction between intestinal microbiota and GALT have shown the importance of intestinal bacteria in the development of GALT and VDJ gene diversification (Lanning *et al.*, 2000; Perey and Good, 1968). Diversification of rabbit Ig occurs in the GALT at 1-2 months of age through two, targeted mutational processes. Somatic hypermutation has been implicated in the diversification of the rabbit repertoire on the basis of appearance of point mutations in the V and D regions of the diversified V(D)J genes (Short *et al.*, 1991; Weinstein *et al.*, 1994). In addition to SHM, gene conversion-like mechanisms also generate antibody diversity in rabbits in the appendix. A somatic gene conversion-like mechanism inserts nucleotide sequence from V_H donors into the rearranged V_H gene segment (Becker and Knight, 1990; Lanning and Knight, 1997). The changes are referred to as gene conversion-like because the non-reciprocal nature of the alterations introduced has not yet been demonstrated; the large number of V_H segments that lie upstream of the functional V gene hindering this sort of analysis. In chickens (Carlson *et al.*,

1990) it has been shown that donor sequences were unaltered during diversification, providing formal confirmation of GC (Carlson *et al.*, 1990).

1.1.8: Generation of Ig diversity in chicken

The bursa of Fabricius is required in birds for B cell development (Cooper *et al.*, 2006). The B cells enter the bursa within a short time window and depletion of B cells from the organ a few weeks after hatching resulted in complete failure to restore the B cells population of chicks (Weill *et al.*, 1986). Thus, B cells move out from the bursa to seed the secondary lymphoid tissue early in life, after which the organ degenerates and further B lymphopoiesis in not possible. This is quite different to humans and mice where B cells are generated in the bone marrow throughout the life of the individual.

In chickens and possibly many other avian species, there are single functional V and J genes, thus limiting the capacity of combinatorial rearrangement to generate diversity. The chicken heavy chain locus contains a single J segment and a unique functional V gene (V_H1) 15 Kb upstream, with approximately 15 D elements in between. About 80 pseudo V genes similar to V-D joints exist upstream of V_H1(Reynaud *et al.*, 1989). The rearrangement takes place outside the bursa early in development and the expression of surface IgM is a pre-requisite for entry into the bursa (Reynaud *et al.*, 1992). Diversification of the unique rearranged V_H1 gene takes place during bursal ontogeny by a hyperconversion mechanism, with V_H pseudogene segments acting as donors (Reynaud *et al.*, 1987a). All the 15 D elements in chickens are very similar and there is lack of N terminal additions limiting the diversity of CDR3.

The features of the chicken λ locus are equally unsuited to the generation of combinatorial diversity. Upstream of the $V_\lambda1$ gene segment, 25 V_λ pseudogenes are found, organised in either orientation (Reynaud *et al.*, 1987a). All pseudo V genes lack promoter, leader exons and recombination signal sequences. Only a few of the pseudogenes contain stop codons or frameshift mutations, but rather more are truncated at their 5' or 3' termini. Diversification is characterised by clusters of nucleotide changes, generated mainly by gene conversion by templated replacement from the many V pseudogenes that exist to the 5' side flank.

One feature of the process confirming that gene conversion takes place is the observation that sequence identical to the diversifying substitutions is always to be found in one of the pseudogenes which themselves remain unchanged (Carlson *et al.*, 1990; Reynaud *et al.*, 1987a). Only pseudogenes on the same chromosomes are used as donors (Carlson *et al.*, 1990) and the pseudogenes that are commonly utilised are those with the best match to the recipient sequence, lie close by at the Ig locus or are in the opposite orientation to the rearranged V segment (McCormack and Thompson, 1990). Conversion tracts range from 8 bp to around 200 bp (McCormack and Thompson, 1990) with a 5' to 3' polarity in the gene conversion mechanism. Of the two sets of parental Ig alleles, only one is rearranged as a result of a strong silencer that makes rearrangement inefficient (Ferradini *et al.*, 1994; Lauster *et al.*, 1993). Strong silencing activity is important because gene conversion events otherwise may correct the out of frame rearrangements thus failing to preserve allelic exclusion (Sayegh *et al.*, 1999). Overall, the mechanism for generation of immunoglobulin diversity in chickens is radically different from that of non-avian vertebrate species.

1.1.9: Mechanisms to generate immunoglobulin diversity in cattle

In cattle, there are limited numbers of variable gene segments and the diversity in the CDRs of these V segments is restricted, as a result the combinatorial mechanism is unable to produce an Ig repertoire of significant diversity. Cattle must therefore utilize one or more post-rearrangement diversification processes to develop the repertoire to a position where an effective immune response can be mounted. From earlier descriptions, the most obvious candidates for these processes are SHM and GC. The current knowledge of development of the antibody repertoire in cattle is rather limited and will be discussed in detail in Chapter 3. Although they differ in many respects, SHM and GC are driven through a series of molecular steps with common dependence upon the recently identified master enzyme, activation induced cytidine deaminase (AID).

1.1.10: Activation induced cytidine deaminase and Ig modification

AID was identified as a novel gene that is specifically expressed in cells active for Ig class switch recombination. It is classified as a member of the

cytidine deaminase family, which includes cytidine deaminase, APOBEC-1, APOBEC-2, and phorbolin. Phylogenetic analysis has revealed that AID is structurally closer to an mRNA editing enzyme, APOBEC- 1, than the metabolic enzyme, cytidine deaminase (Muramatsu *et al.*, 1999). Inactivation of AID by mutation abolishes class switch recombination (CSR) and severely reduces the gene conversion in bursal B cells. These findings and others to be reviewed in chapter 4 have demonstrated that the AID controls somatic hypermutation in mouse (Muramatsu *et al.*, 2000) and human B lymphocytes (Revy *et al.*, 2000). In the chicken cell line DT40, the inactivation of GC by AID mutation could be reversed by reintroduction of AID cDNA (Arakawa *et al.*, 2002).

The common mechanistic basis to GC and SHM has been further illustrated by work with DT40 cells. By restricting or limiting the numbers of pseudogene donors available for GC (Sale *et al.*, 2001) diversification by SHM versus GC can be skewed in one direction or the other, depending upon the particular repair enzymes that are expressed in the cell. These studies have unified all the three B-cell specific DNA modification pathways of SHM, GC and CSR. How these modifications are brought about is still a question of investigation, one crucial question being whether AID is responsible for generating or repairing lesions in the DNA at the Ig loci (Figure 3).

A further uncertainty concerns the precise role of AID. Two models have emerged from studies into this issue, termed the DNA and RNA editing models (Arakawa and Buerstedde, 2004). According to the first model, AID catalyses cytosine deamination leading to guanine/uracil (G/U) mismatches. Putting it another way, AID is directly involved in lesion formation. In the second model, AID edits an mRNA that encodes the DNA-modifying activity and thereby acts indirectly. In both models, AID is responsible for modification to the DNA at the Ig loci. Presently, it is unknown how further processing leads to the process of gene conversion or somatic hypermutation. These models have been discussed at length in Chapter 4.

1.1.11: Rationale and objectives of this study

From the preceding text, it becomes apparent that the mechanisms responsible for the generation of Ig diversity vary considerably among different species of

vertebrates. Since in cattle, rearrangement generates little Ig diversity, post-rearrangement modification is required to establish the primary immune repertoire.

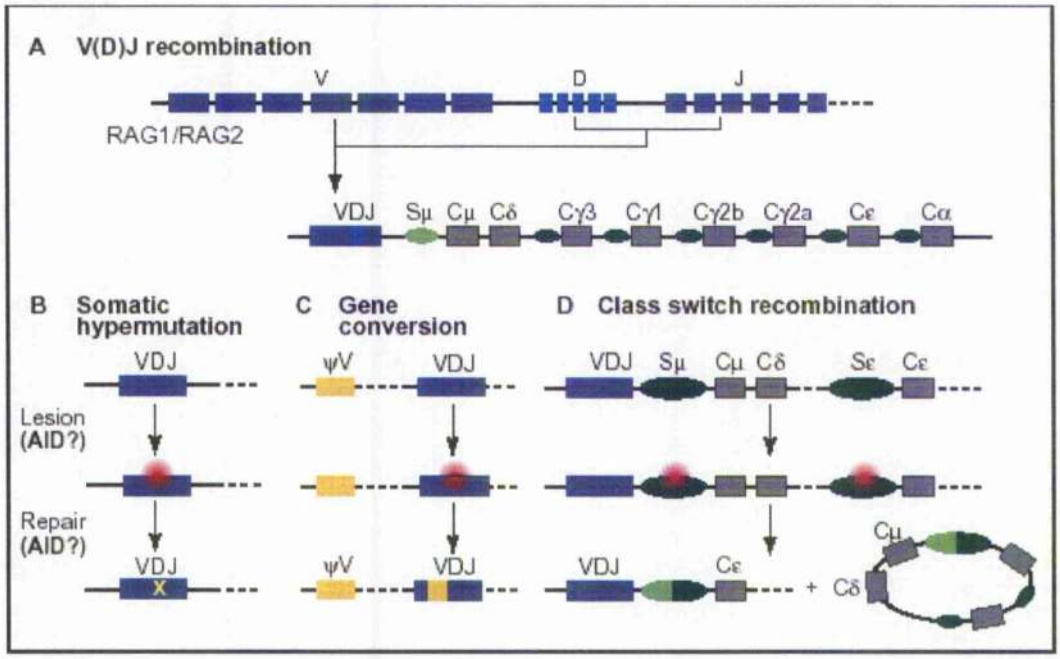


Figure 3: DNA shuffling events at the Ig heavy chain locus during B cell development and maturation

(A) Variable antigen binding domain of antibody results from V, D and J recombination resulting in initial diversity. Further diversity is brought about by (B) somatic hypermutation which introduces point mutations (yellow X) into the rearranged VDJ and sequences downstream through error-prone DNA repair mechanisms or (C) gene conversion, which results in stretches of nucleotide sequences (yellow boxes) inserted into VDJ by a non-reciprocal mechanism from pseudogene V segments (ψV). (D) During class switch recombination, the exons coding the constant region (C, gray boxes) of the antibody are swapped by recombination events between highly repetitive switch regions (S, green ovals). AID is required for SHM, GC and CSR, but it is not clear whether it is involved in the creation or repair of initial DNA lesion (red circles) (Fugmann and Schatz, 2002).

The most likely processes are gene conversion or somatic hypermutation or a combination of the two. One objective of the study was therefore to clarify the mechanistic basis for the Ig diversification at the bovine heavy chain locus. Since SHM and GC, the prime candidates, share a common dependence upon AID, a further goal is to characterise the AID gene from cattle.

Given the size of cattle, their cost and the difficulties of precise manipulation of their immune system, a characterised cell-line of lymphoid origin would be advantageous for detailed study of the mechanisms of Ig

diversification *in vitro*. The aim of one aspect of this study was therefore to identify and characterise a suitable cell-line suited to this purpose.

Whilst the goals outlined so far address the formation of a diverse Ig repertoire in cattle, there is also much to be learned about its expression. One aspect that has attracted comment from investigators has been the length of CDR3H. Many studies have noted that the length of this region of the IgH chain typically exceeds that of many other vertebrates (Saini *et al.*, 1999; Saini and Kaushik, 2002), and there has been speculation about the genetic basis for this (Saini *et al.*, 2003). A further peculiarity of the bovine Ig system identified recently is the duplication of bovine J_H and Cμ loci on BTA 11 and BTA 21. There are some suggestions from published data that while one locus dominates, both are capable of rearrangement (Berens *et al.*, 1997; Hosseini *et al.*, 2004). The final area for this study was therefore to seek confirmation of this finding and to develop reagents able to distinguish the alternative allotypes of IgM that may arise. It is proposed that phage display be used to develop these reagents. Potentially, allotypic specific antibodies could then be used to assess the presence of variants of IgM and also to isolate and compare B cells that have undergone rearrangement at BTA21 and / or BTA11.

Chapter: 2 Materials and Methods

2.1: Media, Buffers and Solutions

The composition of all the media, buffers and solutions is given in the Appendix I.

2.2: Growth of bacterial cultures

E. coli strains were grown from glycerol stocks on Luria-Bertini agar (LA) plates by incubating at 37°C overnight. Subsequent growth of *E. coli* was performed by inoculating single colonies into Luria-Bertini broth (LB) or Terrific broth (TB). Broths were incubated at 37°C with shaking at 220 rpm after inoculation.

2.3: Preparation of competent cells for heat shock

On most occasions, commercially prepared competent cells were used for transformations. However, in-house prepared cells were also used when particular strains of *E. coli* were required. The strain to be transformed was grown on LA overnight to obtain separate single colonies. A single colony was then inoculated into 2 ml LB and the culture grown at 37°C with shaking at 220 rpm. The following morning, 2 ml of overnight culture, was inoculated into 500 ml of LB in a 2 litre baffled flask and shaken vigorously at 37°C to an $OD_{600\ nm}$ of 0.4 - 0.5. Typically, this was achieved within 3-5 hours. The culture was transferred to two pre-chilled sterile centrifuge containers and cells were pelleted at 4000 g for 10 min at 4°C. Supernatants were discarded and pellets were kept on ice to be suspended in 5 ml ice-cold, filter-sterilised $CaCl_2$ solution. The suspensions were pooled together in another centrifuge tube and again pelleted at 1000 g for 10 min at 4°C. The supernatant was discarded and pellet resuspended in 10 ml ice-cold, filter-sterilised $CaCl_2$ solution and incubated on ice for 30 min. At the end of incubation, cells were pelleted as before and resuspended in 2 ml of $CaCl_2$ solution. Cells were used at this point for transformation or were stored in 100 µl aliquots at -70°C after adding sterile glycerol to a final concentration of 15%. Cells usually remained competent for several months when prepared in this way.

2.4: Preparation of electro-competent cells

A 500 ml culture of bacteria was prepared in early to mid-exponential phase as described in the previous section. The bacterial culture was chilled on ice for 30 min before centrifugation at 4°C. The centrifugation was performed at 13000 g for 10 min. The supernatant was discarded and the pellet was resuspended in 500 ml of ice-cold sterile distilled water (DW) and centrifuged as before. This time, the pellet was suspended in 250 ml DW and centrifuged. The supernatant was discarded and 10 ml of cold sterile 15% glycerol (v/v) in DW was added to the pellet and mixed gently to make a uniform suspension. The centrifugation was carried for another 10 min under the same parameters. Finally, the supernatant was discarded and pellet was resuspended in 1.5 ml of 15% of glycerol solution. 100 µl aliquots were dispensed into sterile cryovials, snap frozen in liquid nitrogen and stored at -70°C.

2.5: Genomic DNA extraction

QIAamp DNA Mini and DNeasy tissue kits from Qiagen were used to extract genomic DNA from cells and tissues as per the manufacturer's instructions. The DNA samples were stored at -70°C.

2.6: Design of PCR primers and polymerase chain reaction (PCR)

PCR primers were designed to contain balanced G:C composition and compatible melting temperatures and a GC clamp at 3' termini. They were screened for internal secondary structure like hair pin loops and for possible dimerisation using software available at the Sigma Genosys website (http://orders.sigma-genosys.eu.com/). Primers were used at 500 nM final concentration in PCR reactions. Other reaction components and cycling parameters like $MgCl_2$ concentration and annealing temperature were optimised by experimentation to obtain specific PCR amplification.

2.7: Purification of PCR or other DNA products

PCR products ranging from 100 bp to 5 kb were purified from primers, enzymes and salts either from agarose gels using QIAquick Gel extraction kit or directly after PCR using QIAquick PCR purification kits from Qiagen. The manufacturer's manuals provided protocols for purification.

2.8: Agarose gel electrophoresis

Agarose gels of different compositions were prepared by dissolving molecular grade agarose (Invitrogen, U.K.) into 1× Tris-Acetate EDTA (TAE) buffer (Invitrogen, U.K.). Agarose concentration was adjusted according to the sizes of fragments to be resolved. Agarose was heated in a microwave oven until completely dissolved. Ethidium bromide (BioRad) to a final concentration of 1 µg/ml or SYBR Safe DNA gel stain from (Invitrogen; 5µl/100ml) was added to the agarose solution, once the agarose was dissolved. The agarose was poured into a gel tray with comb and allowed to set. During this time, the DNA samples to be resolved were mixed with 6× DNA loading buffer prior to loading into the wells. A 1 kb or 100 bp molecular weight ladder (Invitrogen or Promega, U.K.) was prepared and used according to manufacturer's instructions. The set gel was placed into an electrophoresis tank, filled with 1× TAE, to a level approximately 1 cm above that of the gel. The comb was removed and samples were loaded into each well. The gel was run at 100 volts (applied 1-5 volts/cm-distance measured between the electrodes) until the dye front migrated to the other end of the gel such that single DNA bands could be visualised clearly on a UV transilluminator (Ultra Violet Products) attached to an image acquisition and analysis software package from Labworks.

2.9: Plasmid DNA extraction

Plasmids were extracted using QIAprep Miniprep kit (Qiagen) according to manufacturer's instructions with the following modification: plasmid DNA was extracted from a 2 ml of overnight bacterial culture in LB with appropriate antibiotic and was eluted from the miniprep columns with 30-50 µl of elution buffer and stored at -70°C.

2.10: Phenol extraction and ethanol precipitation of DNA

An equal volume of phenol was added to the reaction mixture containing DNA mixture and the sample was vortexed gently. An aqueous phase containing the DNA was separated from the organic phase by centrifugation in a microfuge, at 2000 rpm for 5 min. The aqueous phase was transferred with care to a fresh microfuge tube and an equal volume of 24:1 (v/v) chloroform-isoamyl alcohol was added. In order to precipitate DNA from the aqueous solution, 0.1 volume of

3 M sodium acetate, pH 5.5, was added followed by 2 volumes of absolute ethanol. The reaction mix was incubated at -20°C overnight or for up to 1 hour at -80°C. The precipitated DNA was recovered by centrifugation in a microfuge at 13 000 rpm for 15 min. Ethanol was removed with care and the pellet was allowed to dry. The dried pellet was then resuspended in sterile Tris-EDTA buffer, pH 8.0 and stored at 4°C for further manipulation or at -20°C for long-term storage.

2.11: DNA sequencing

DNA sequence was obtained by using vector or gene-specific primers. The sequencing was performed using conventional protocols based on Sanger chemistry using DYEnamic ET-Terminator chemistry (GE Healthcare) according to the manufacturer's instructions. Termination products were separated on a MegaBACE1000 (96 capillary) instrument. Sequencing was carried out by Ms. Julie Galbraith at the Molecular Biology Support Unit, Sir Henry Wellcome Functional Genomics Facility of IBLS, University of Glasgow. The sequences were read and analysed by VectorNTI software from Invitrogen.

2.12: Heat shock transformation

A frozen 100 µl aliquot of competent cells was thawed on ice and 2 µl (10 ng) of plasmid DNA was added and incubated on ice for a minimum of 30 min. Cells were then heat shocked in a waterbath at 42°C for 45 sec and returned to ice for another 2 min. After incubation, 900 µl of SOC medium (room temperature) was added to the cells and they were incubated at 37°C for 1 hr on a shaker-incubator. The cells were then plated out onto LA with appropriate antibiotics in different dilutions and incubated overnight at 37°C.

2.13: Electroporation

An aliquot of competent cells was removed from deepfreeze and allowed to thaw on ice. As before, 2 µl (10 ng) of plasmid DNA was added to the cell suspension. Electroporation was performed using a BioRad Gene pulser (BioRad laboratories, UK) under specific parameters of 2.5 kV, 25 mF, 200 Ω in cuvettes with 0.2 cm gap (BioRad). The voltage was decreased to 1.5 kV when using 0.1 cm cuvettes. Cells were placed on ice for 2 min followed by addition of 900 µl of pre-warmed LB or SOC. The mixture was pipetted to a sterile microfuge tube

and incubated at 37°C for 1 hr with shaking at 220 rpm before plating onto LA with the appropriate antibiotics. Plates were then incubated at 37°C overnight.

2.14: SDS-PAGE gel electrophoresis

Electrophoresis was carried out by the method of Laemmli (Laemmli, 1970) in a vertical electrophoresis tank using the Mini Protean 3 electrophoresis system from BioRad. Polyacrylamide gels were routinely cast with 5% stacking and 12% resolving layers (Sambrook *et al.*, 1989). Resolving gel mix was poured between the assembled plates allowing some space on top for the later addition of the stacking gel. Immediately, 100% isopropanol was added as a thin overlay to improve polymerisation and to create a smooth interface for the addition of a stacking gel. The assembly was placed on a leveled surface and the gel was allowed to set. Once set, the isopropanol was removed by gently soaking with absorbent tissue paper. Stacking gel mix was poured onto the polymerised resolving gel and an appropriate comb was placed into the solution during polymerisation. Once the reaction was complete the comb was removed gently to maintain the format of the wells. The wells were rinsed with distilled water and any residual gel on the assembled glass plates was removed. The electrophoresis tank and gels were assembled according to the manufacturer's instructions. The middle chamber was filled with 1× SDS running buffer. Twenty µl of protein sample was boiled with 10 µl of 3 × sample buffer for 5 min and samples were loaded into each well. A protein ladder (Sigma, UK) was included. The outside chamber was filled with 1× SDS running buffer. The gels were run at 200 volts for approximately 45 min or until dye from the sample buffer ran off the gel.

2.15: Coomassie Blue staining for the detection of protein

SDS-PAGE protein gels were stained with Coomassie Blue stain for a minimum of 30 min on a shaking platform. After incubation, the stain was decanted and the gel was destained several times, for a minimum of 30 min in between changes, on a rotating platform until protein bands could be visualised clearly. The gel was scanned immediately or immersed in water for later documentation.

2.16: Immunoblotting

After electrophoresis, proteins were transferred from SDS gels to nitrocellulose membrane (Amersham Life Science) as per standard protocol for Western blotting (Sambrook *et al.*, 1989). The stacking gel was cropped from the resolving gel before assembling into the transfer cassette. A nitrocellulose membrane (Amersham Pharmacia Biotech) and four pieces of Wattman filter paper were cut to the same size as the gel and soaked with the gel in transfer buffer for 5-10 min. For transfer, a fiber pad was overlaid with layer of two pieces of Wattman filters, followed by the gel on top of it, and then nitrocellulose membrane, and finally one more layer of filter papers. On top of this went the second fiber pad and the sandwich was assembled into a blotting cassette and loaded into a BioRad electrophoresis tank. Transfer buffer was added along with a frozen insert to avoid overheating during transfer. Care was taken to ensure that the gel lay to the anode (-) side of the assembly and the membrane lay to the cathode (+) side. Proteins were transferred from the gel to the membrane for 1 hr at 100 volts. After transfer, the membrane was carefully separated, washed three times with PBS and stained with Ponceau S solution for 2 min at room temperature to confirm that proteins had been transferred from the gel to the membrane. The position of each molecular weight marker was scored onto the membrane with indelible lead pencil. The membrane was then washed with PBS to remove the dye. The membrane was blocked with 3% milk-PBS (MPBS) for 1 hr at room temperature with shaking to block non-specific protein binding. After that, the membrane was washed and probed with an appropriate dilution of antibody or antibody-conjugate as the protocol demanded for 60 min at 37°C and 200 rpm shaking. The unbound antibody was washed out with three changes of PBS containing 0.05% Tween. If using an unconjugated antibody, the membrane was once again incubated in a 1:1000 dilution of anti-species-HRPO for 60 min at 37°C with 200 rpm shaking. The membrane was washed as before and developed using 0.05% 4-chloro-1-naphthol (Sigma, UK) with 0.015% hydrogen peroxide in 20% methanol in TBS for 20 min. The membrane was scanned for future record.

2.17: Enzyme linked immunosorbant assay (ELISA)

ELISA assays were performed in microtiter plates (Iwaki). The first coating comprised unlabelled antibody, antigen, and crude bacterial lysate or purified protein. Sample was coated to the designated wells in 50-100 µl solution. The plate was incubated overnight at 4°C to allow complete binding. The wells were washed with three washes of PBS, flicking the buffer out of the plate into a suitable container after each wash. The remaining unbound protein sites in the wells were saturated by incubating the plates for 2 hrs at room temperature with a suitable blocking buffer. The wells were washed 3 times with PBS containing 0.05% Tween-20 before adding the labelled or unlabelled antibody as dictated by the assay format. Appropriate antibody concentrations were determined by preliminary experiments. The plate was incubated for additional 1-2 hr at room temperature in a humid chamber. At the end of incubation, wells were washed with several changes of PBS containing 0.05% Tween-20. Ortho phenylene diamine (1mg/ml) in 0.1M citrate buffer (pH 4.5) was applied to each well if HRPO-labelled antibody had been applied in the previous step. When unlabelled antibody was used, an enzyme conjugated reagent was applied to each well in appropriate dilution and plate was incubated, washed and developed as before. The optical density was measured at 450 nm at the end of a brief incubation and the reaction was quenched by adding 50 µl of 2 N HCl.

2.18: Isolation of mononuclear cells (MNCs) by density gradient separation

Blood from the local slaughterhouse was collected into sterile universals containing acid dextrose citrate or heparin (10 U/ml) as anticoagulant. Blood was diluted with an equal volume of PBS and gently layered over double the volume of cool Histopaque (1.083g/L) (Sigma) in 15 ml conical centrifuge tubes and centrifuged at 20°C for 40 min at 400 g. The mononuclear cells located at the interface between plasma (upper layer) and Histopaque (lower layer). Cells were carefully removed and transferred to 15 ml conical tube and washed thrice with PBS and resuspended to 1 ml in 0.5% BSA in PBS (PBS-BSA). The cells were counted in a Neubar chamber after staining with trypan blue and their density adjusted before proceeding with cell staining procedures.

2.19: Single cell suspension preparation from calf tissue

Samples of calf spleen and prescapular lymph nodes were used to prepare single cell suspensions and remaining tissue was stored at -80° C. The connective tissue was removed and samples were homogenised separately in PBS. The suspension was transferred into a 50 ml conical tube and the large clumps and debris were allowed to settle to the bottom. The suspension was passed through a nylon mesh (70 µM pore size) to recover single cell suspension, which was then centrifuged for 5 min at 4°C and 400 g.

2.20: Lysis of calf spleen RBCs

Spleen cell preparations were resuspended in 10 ml of 1x RBC lysis buffer and incubated on ice for 5 min with occasional shaking. Lysis was stopped by diluting with 30 ml of PBS. Cells were spun down at 400 g and the pellet was resuspended in PBS-BSA. The cells were counted in a Neubar chamber after staining with trypan blue and their density adjusted with labeling buffer (PBS-BSA) before proceeding with cell staining procedures.

2.21: Isolation of CD21+ cells by magnetic separation

Single cell suspensions and isolated mononuclear cells were labelled with mouse anti-bovine CD21 IgG antibody at 1:50 dilutions in PBS-BSA. The cells were incubated at room temperature for 10 min or at 4°C for 30 min. Cells were then washed carefully by adding 10 times the volume of labeling buffer, centrifuged at 300 g for 10 min and the supernatant was removed. Washing was repeated and the cell pellet was resuspended in 80 µl of buffer per 10^7 or fewer total cells. 20 µl of MACS goat anti-mouse IgG microbeads (Miltenyi Biotech) per 10^7 or fewer cells was added, mixed well and incubated for 15 min at 6-12°C. Cells were further labelled with goat anti-mouse IgG-FITC conjugate at 1:200 dilution and incubated at room temperature or ice as before. Cells were washed and the pellet was resuspended in 500 µl of buffer per 10^8 or fewer cells before B cells were positively selected using MACS columns as per manufacturer's protocol. Briefly, the column was washed with 3 ml of buffer before separation by placing it in a magnetic field. The cell suspension was applied onto the column and unlabelled cells were allowed to pass through. The column was further rinsed with 3 ml of buffer and removed from the magnetic field. It was

then placed on a collection tube and 5 ml buffer was pipetted onto the column. CD21+ cells were eluted by firm pressure using a plunger.

2.22: Growth and maintenance of cell lines

Bovine lymphosarcoma cells (BL-3) and the African green monkey kidney cell line, COS-7 were purchased from the European Collection of Cell Cultures and maintained as per supplier's instructions. Both the cell-lines were cultured at 37°C with 5% CO_2 in 50% RPMI 1640 - 50% Leibovitz L-15 medium (Invitrogen) supplemented with 20% FBS, 2 mM L-glutamine, 100U penicillin, and 100 µg streptomycin. BL-3 cells were grown to a density of $1x10^9$ cells/ml and COS-7 cells to 100% confluency in 25 or 75 cm^2 tissue culture flasks and were sub-cultured and media replenished every 72 hours.

Since BL-3 cells grow in suspension, splitting and nutrient medium change was done after centrifugation at 300 g for 5 min. Cells were resuspended in a pre-warmed medium to a density of $1x10^5$ cells/ml for culture in fresh flasks as before.

COS-7 cells were harvested by trypsinisation. The old medium from confluent cells was aspirated off and the growth surface was rinsed with 5 ml sterilised PBS before adding 2 ml of trypsin-EDTA (Cambrex Biosciences, UK). The flask was gently rotated to make sure that the growth surface was covered with trypsin and it was incubated at 37°C in 5% CO2 incubator for 5 min. At the end of the incubation, the cells were quenched with 3 ml of RPMI-1640. The flask was gently tapped on the palm of the hand to dislodge the still loosely adherent cells and cells were pipetted repeatedly up-down to make a single cell suspension. The floating cells were observed under a phase contrast microscope to ensure single cell suspension before centrifugation at 300 g for 5 min. The trypsin was aspirated off. 5 ml RPMI-1640 was added and the centrifugation step was repeated. Finally the cells were suspended in complete medium and re-seeded at a pre-determined split ratio into fresh tissue culture flasks and incubated as before.

2.23: Flow cytometry

Flow cytometry was performed using a FACS Calibur and Cell Quest system (Becton Dickinson). Data were later analysed using WinMDI 2.8 software.

Chapter: 3 Mechanisms for IgH diversification in cattle

3.1: INTRODUCTION

Antibody diversity in vertebrates is generated by the recombination of separate germline gene segments both for the heavy (V, D and J) and light (V and J) chains and association of two identical heavy and light chains that are independently formed (Tonegawa, 1983). This diversity is further increased as a result of junctional flexibility from nucleotide deletions or additions during the recombination process. More antibody diversification is brought about by somatic hypermutation and insertions or deletions. While the recombination process is common across vertebrate immune systems, germline immunoglobulin gene sequence diversity differs considerably.

Most studies in the past have shown that in cattle both the light chain and heavy chain repertoires are founded upon the frequent expression of single gene families and sub-groups of segments, which are of conserved sequence. Bovine IgH chains possess CDR3 sequences that are frequently long (Berens et al., 1997)(Berens *et al.*, 1997; Saini *et al.*, 1997; Sinclair *et al.*, 1997) and sometimes in excess of 50 amino acids in length (Saini *et al.*, 1999). It is now known that all the antibody classes (IgM, IgD, IgG, IgE and IgA) that are present in human and mice are also found in cattle (Zhao *et al.*, 2002). Present understanding of bovine humoral immunity and the mechanisms generating the bovine immunoglobulin repertoire is far from complete but its structure and the levels of diversity which are generated through rearrangement reveal that this is another species which departs significantly from the murine/human paradigm (Aitken *et al.*, 1999). The following sections highlight the extent of these differences.

3.1.1: Limited VH genes

The bovine heavy chain locus is located on bovine chromosome 21q23-24 (Gu *et al.*, 1992; Tobin-Janzen and Womack, 1992). The exact size of the locus and the number of segments present is not known. However, studies have shown that a family of V_H segments homologous to mouse V_H Q-52 and the related human family are used in rearrangement. The gene family has been designated as BoV_H1 (Berens *et al.*, 1997; Lopez *et al.*, 1998; Saini *et al.*, 1997; Sinclair *et*

al., 1997). Data from Southern blotting, ssCP analysis and sequencing indicate that the size of the dominant gene family is small, comprising no more than 20 members. Further, according to Sinclair *et al.*(Sinclair *et al.*, 1997) it carries few distinct CDR sequences and diversity apparent in these sequences is extremely modest. Southern blot analysis of bovine genomic DNA have also demonstrated that homologues of at least three other murine families are present in the genome (Berens *et al.*, 1997; Lopez *et al.*, 1998). Transcripts from these families have not been detected and the reason why they do not undergo rearrangement is unknown. Gene rearrangement through combinatorial joining of limited V_H segments with D_H, and J_H gene segments thus seems unable to impart significant diversity to the primary heavy chain repertoire in cattle.

3.1.2: Unique Complementarity Determining Region (CDR) 3 (heavy) H region

The CDR3H region of an antibody is typically the most diversified of the CDRs and determines largely the antibody specificity. Various investigators have observed that cattle IgH chains carry CDR3s of exceptional length with an average size in bovine IgM and IgG antibodies being 22.18±10.86 and 19.56±6.64 amino acids, respectively (Berens *et al.*, 1997; Saini *et al.*, 1997; Sinclair and Aitken, 1995; Sinclair *et al.*, 1995). This sets apart the cattle immunoglobulin system from sheep which otherwise are very similar at the molecular level. Three different potential mechanisms have been proposed to explain the unusual length of CDR3s in cattle (Sinclair *et al.*, 1997). These include long bovine D segments, terminal transferase activity to expand this part of the reading frame to an exceptional degree, and rearrangement of D segments in pairs. Camels use long CDR3s in antibodies that lack light chains and it has been postulated that a long CDR3 produces a loop that compensates for this light chain deficiency (Desmyter *et al.*, 1996; Hamers-Casterman *et al.*, 1993). D-D gene fusions in humans have been shown to generate long CDRs (Sanz, 1991) as is the case in chickens (Reynaud *et al.*, 1989).

More recently, bovine germline D_H segments have been isolated from a bovine genomic DNA library using a radiolabelled probe. The partial characterisation of a 2.3 kb fragment led to the identification of three bovine germline D_H segments of 42 bp, 58 bp and 148 bp. Phylogenetic analysis

suggested that these are closest to rabbit and chicken D_H gene (Shojaei *et al.*, 2003). All the segments are flanked by classical RSSs with typical heptamer and nonamer sequences on either side though these differ significantly in one of the D_H segments. Rearrangement of a D segment of 148 bp could potentially generate a long CDR3H region. However, the isolation of 148bp bovine D_H gene does not fully explain the generation of CDR3H sequences as long as 61 codons as has been observed in some bovine antibodies (Shojaei *et al.*, 2003). There is a possibility that even larger germline segments exit or other mechanisms such as D-D fusion might contribute to the generation of long CDR3H (Meek *et al.*, 1989; Reynaud *et al.*, 1991b).

3.1.3: J_H locus

The analysis of BAC clone 66R4C11 isolated from a bovine library has shown that the J_H locus is located about 7 kb upstream of the Cμ gene (Zhao *et al.*, 2002). The bovine J_H locus is 1.8 kb in size and is composed of six genes but only two of these genes, J_H1 and J_H2 are expressed; the level of expression for J_H2 is much lower than J_H1 (Zhao *et al.*, 2002). These two gene segments are 862 bp apart and encode 15 and 17 amino acids, respectively. The first three of the remaining four J_H pseudogenes (ΨJ_H1, ΨJ_H2, ΨJ_H3, ΨJ_H4) are located upstream of J_H1 whereas the fourth one (ΨJ_H4) lies between J_H1 and J_H2. The two J_H segments encode a five amino acid motif, VTVSS, at their 3' ends. Recently, it was reported that duplicated copies of the J_H locus contribute to the bovine Ig repertoire (Hosseini *et al.*, 2004). These authors mapped a bovine J_H locus comprising a DQ52 segment, six J_H segments and sequence to a 5' H chain intronic enhancer to BTA 11 and evidence was found for rearrangement of the sixth J_H segment at a low but detectable frequency.

3.1.4: The bovine 5' Ig intronic enhancer (Eμ) and Switch (Sμ) region

Somatic hypermutation is restricted to the region around the rearranged V gene. Very little is known about the regulatory element which modulates the process of hypermutation or gene conversion but *cis*-acting sequences such as enhancer regions have been shown to be necessary for hypermutation. Several tightly clustered nuclear binding motifs, μE1-μE5-μE2-μA-μB-μE4-O are present

in a putative Eμ enhancer core region (Zhao *et al.*, 2002). AT-rich matrix attachment regions (MAR) are present on either side of the bovine enhancer core region as seen in humans and mice. The Sμ region is located upstream of the Cμ gene in mammals. This region is responsible for isotype switching through a non-homologous recombination process. The bovine J_H-Cμ intron has a 3-kb repetitive region abundant in switch region motifs (187 CTGGG and 127 CTGAG) similar to both the human and mouse Sμ regions. It is slightly shorter than the human Sμ (3.5 kb) but about twice as long as the mouse Sμ (1.5 kb). In addition, a 24-bp long sequence (GGCAGAGTGGGTGAGCTGGGCTGA) appears dispersed as 9 repeats within the bovine Sμ (Zhao *et al.*, 2002).

3.1.5: Mechanisms to diversify cattle antibody repertoire

Several investigations have studied the development of the bovine Ab repertoire. Berens and colleagues showed that sequences from 150-day-old foetuses showed little variability indicating that rearrangement in cattle generates minimal Ab diversity (Berens *et al.*, 1997). This is in agreement with the predictions based upon the sequencing of germline VH segments. The sequences from adult IgM transcripts were mutated as much as those from antigen-driven heterohybridomas indicating a post-rearrangement diversification process takes place before antigen encounter. In common with other reports, these authors showed the rearrangement of single family of V_H segments ($BovV_H1$), although the germline carries other V_H gene families. The reasons for the lack of expression of these families are not known. It has been suggested that chromosomal translocation of V_H genes from other families to different chromosomes (Matsuda *et al.*, 1990), closer proximity of $BovV_H1$ genes to the D gene cluster (Yancopoulos *et al.*, 1984) and/ or differences in the regulatory signals could explain the predominant expression of the $BovV_H1$ family. Little or no mutation was detected in FR4 from foetal and adult cattle, which resembles findings in other species.

Studies have indicated that the spleen is a likely site for Ig rearrangement in cattle (Meyer *et al.*, 1997). The expression of RAG-1 gene has been shown concurrently in thymus and spleen in 14 days old calves. No expression of RAG-1 was observed in the spleen of 32 week old cattle but it continued to be expressed in the thymus suggesting that Ig gene rearrangement occurs early in

the life of the animal only (Meyer *et al.*, 1997). By analogy with the sheep, it is likely that rearrangement in B cells begins during foetal development and that the lymphocytes then go on to populate the IPP before birth and for weeks thereafter (Reynolds and Morris, 1983). In chicken, and sheep, the post rearrangement diversification of the Ig occurs in the GALT. Both the bursa (chicken) and the IPP (ruminants) begin involuting 3 to 6 months after birth following the completion of diversification. The released B cells populate the periphery where they form a self-renewing pool (Griebel and Ferrari, 1994; Reynolds *et al.*, 1991).

In species that are unable to generate significant Ig diversity by rearrangement, gene conversion and somatic hypermutation have been shown to drive diversification, post-rearrangement. Chickens form the best characterised example of a species that undergoes frequent gene conversion during antibody diversification (Reynaud *et al.*, 1987b; Thompson, 1992). Gene conversion as a process is thought to have been responsible for the duplication of Ig genes during the evolution of the Ig locus in mammals (Baltimore, 1981). The first demonstration of gene conversion as a mechanism for generation of diversity in mammals came from studies in rabbit where it was shown to diversify the preferential V_H1 (D) J rearrangement (Becker and Knight, 1990).

Gene conversion has been suggested as one mechanism for diversification of Vλ genes of cattle (Lucier *et al.*, 1998; Parng *et al.*, 1996). This conclusion was based upon the identification in light chain cDNA of substitutions which matched the sequence of putative pseudogene donors. However, these experiments were unable to discriminate between differences that might have been the result of allelic variation and sequences donated from pseudogenes. The investigators (Parng *et al.*, 1996) compared cDNA sequences to a functional germline segment which was not typical of the majority and therefore might constitute an unlikely candidate for the rearranged segment. Moreover, many of the substitutions to which they drew attention were common to several cDNAs and the donors possessed near-identical sequences in the CDRs suggesting that little meaningful diversity could be generated in this way.

In sheep, somatic hypermutation is thought to generate diversity post-rearrangement in the IPP (Reynaud *et al.*, 1995; Reynaud *et al.*, 1991c). These

studies were centered on analysis of the lambda light chain repertoire and for many years, it was assumed that hypermutation was the dominant factor in diversification. More recently, it has been shown that rearrangement is capable of generating higher levels of light chain diversity than previously supposed (Jenne *et al.*, 2003). This study suggests that the sheep is not as heavily reliant upon mutation as might have been inferred from initial reports. However this work by Jenne and colleagues does not question the basic view that somatic hypermutation acts post-rearrangement to further diversify the ovine primary Ig repertoire. Studies of the bovine Ig heavy chain repertoire have generally favoured somatic hypermutation as a diversification process (Aitken *et al.*, 1999; Berens *et al.*, 1997; Kaushik *et al.*, 2002) but to date, definitive evidence has been unavailable.

From the preceding section, it becomes apparent that the mechanisms responsible for the generation of Ig diversity seem to vary considerably among different species of vertebrates. Since in cattle, rearrangement generates little Ig diversity, post-rearrangement modification is required to establish the primary repertoire. The main candidates for this process are gene conversion or somatic hypermutation. It can be predicted that somatic hypermutation would introduce nucleotide substitutions throughout the Ig reading frame and that these would extend into the intron separating the J_H locus from the $C\mu$ exons (Figure 4). This is because the process is non-templated and random by its nature. Work in other systems confirms that this is possible (Gonzalez-Fernandez *et al.*, 1994).

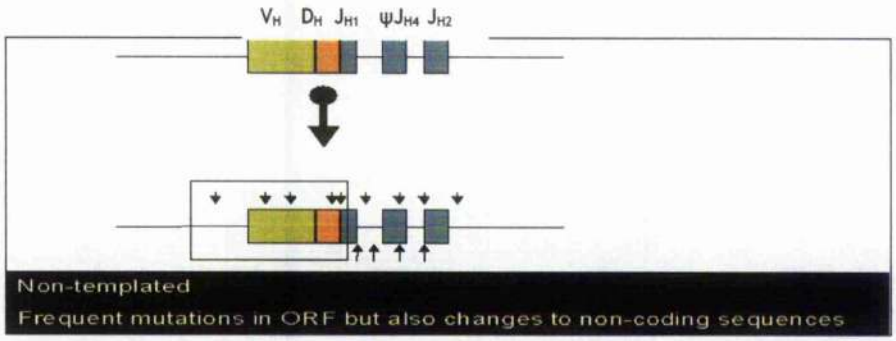


Figure 4: Targeting of the IgH locus by somatic hypermutation

Coding as well as non-coding region is targeted. Upper cartoon depicts the frequently rearranged Ig heavy chain locus (closed boxes indicating exons space between the boxes joined by line represents intronic region studied). Smaller black arrows indicate the same region as shown in upper row being targeted by point mutations.

In contrast, if gene conversion was the dominant process behind diversification it would be predicted that modifications would be confined to the Ig reading frame (Figure 5). This would arise through the use of pseudogene donors that in systems such as the rabbits are VH segments (Becker and Knight, 1990) or VHD fusions in the chicken (Weill and Reynaud, 1996).

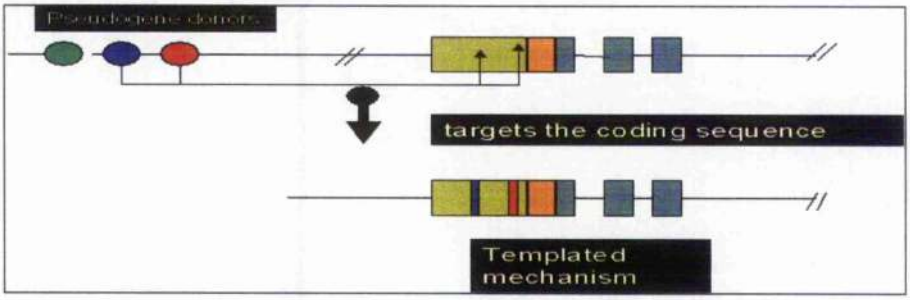


Figure 5: Targeting of the IgH locus by somatic hypermutation

In GC, pseudogene donors target the coding region in non-reciprocal and templated fashion. The rearranged locus shown here is same as in figure 4.

The aims of this part of the study were therefore to identify changes in the nucleotide sequence of the IgH reading frame and the J_H-Cμ intron during the gene diversification. This was to be accomplished by the following objectives:

1: To recover and sequence Ig heavy chain sequences for the rearranged locus from lymphoid tissues of cattle of different ages thereby determine the timing of Ig diversification.

2: To compare rates with which diversity becomes apparent in the CDRs, with the rates of nucleotide substitutions in FR domains and the J_H- Cμ intron.

3.2 MATERIALS AND METHODS

3.2.1: Animals

- **Source:** Local slaughter house (Glasgow) and Moredun Research Institute (Edinburgh) and Veterinary School, University of Glasgow
- **Breed:** Holstein Fresian Cross
- **Ages:** Day old, 7 weeks old, 13 weeks old and adult animals (25 to 30 months of age)

- **Tissues sampled at post-mortem:** Liver, spleen, various lymph nodes, thymus, muscle etc.

3.2.2: Preparation of DNA

Liver, spleen and lymph nodes were collected from cattle of various age groups and were stored at -80°C prior to isolation of DNA. The genomic DNA was extracted using a Qiagen DNA extraction kit as per manufacturer's instructions.

3.2.3: Recovery of the rearranged heavy chain variable locus and J_H intronic sequences

The sequence of the primers is given in Table 2. One microgram of genomic DNA was used as template to recover the rearranged heavy chain variable locus with forward VHF and reverse IR-0.5 primer set. Similarly, to recover J_H intronic sequences from a rearranged heavy chain locus, the primer set comprised forward FR-3 and reverse IR-1. The design of these primers is discussed in the "Results" section. Primers were used at final concentrations of 800nM in a reaction buffer containing 2mM magnesium chloride and 200µM dNTPs. Amplification was performed in final volumes of 25 µl with 2.5 U of recombinant Taq polymerase from Invitrogen, Paisley, UK. After initial denaturation at 94°C for 3 min, reactions were cycled 30 times through 94°C (1 min), 64°C (40 sec) and 72°C (1 min) with final 5 min incubation at 72°C. PCR products were resolved on a 1.2% agarose gel and extracted from the gel using an extraction kit from Qiagen. The extracted fragments were ligated into the plasmid vector pCR 2.1-TOPO and *E. coli* DH5α (Invitrogen, Paisley, UK) were transformed as per manufacturer's protocol.

The *E. coli* DH5α transformed with the ligation products were then selected on Luria-Bertini agar plates containing 100µg/ml ampicillin. Plasmid DNA was extracted from the candidate clones and characterised by restriction analysis. Inserts in the recombinant clones were sequenced using M13 forward and M13 reverse primers using conventional protocols based on Sanger chemistry using ET-Dye Terminator (Amersham Bioscience, UK) on MegaBACE1000 (96 capillary) instrumentation. Sequencing was carried out at the Molecular Biology Support Unit, Sir Henry Wellcome Functional Genomics Facility of IBLS, University of Glasgow.

Table 2: Oligonucleotides used in the recovery and analysis of the bovine heavy chain immunoglobulin locus

Primer	*Sequence(5' to 3')*
VHF	CCCTCCTCTTTGTGCTCTCA
FR3	CTCAGCATCACCAAGGACAA
IR0.5	GTGAGGAGAGAGGGCTGTTG
IR-1	CCTGGTCCTGCTCAAGTCA

3.2.4: Sequence Analysis

All DNA and amino acid sequence analyses were performed with *Vector* NTI Advance™ 9.0 from Invitrogen, UK. DNA alignments were manipulated and annotated visually and nucleotide differences were counted manually. The proportion of synonymous substitutions per potential synonymous site and the proportion of nonsynonymous substitutions per potential nonsynonymous site were calculated (Nei and Gojobori, 1986). Synonymous NonSynonymous Analysis Program (SNAP) calculates synonymous and non-synonymous substitution rates based on a set of codon-aligned nucleotide sequences, based on the method of Nei and Gojobori incorporating a statistic developed by Ota and Nei (Ota and Nei, 1994).

3.3: RESULTS

3.3.1: Origin of the analysed sequences

3.3.1.1: FR1-FR3 region

The location of various primers is shown in Figure 6. The 5' Forward VHF primer used to amplify rearranged heavy chain variable locus corresponds to bases 12 to 31 of *Bos taurus* immunoglobulin heavy chain variable region gene, accession number U55165 (Sinclair *et al.*, 1997). This primer targets a region of the V_H leader that does not vary between V_H segments characterised to date. The reverse IR-0.5 primer was complementary to bases 1073 to 1092 exactly 10 nucleotides downstream from J_H1 (Zhao *et al.*, 2003), the most frequently rearranged J_H segment in bovines.

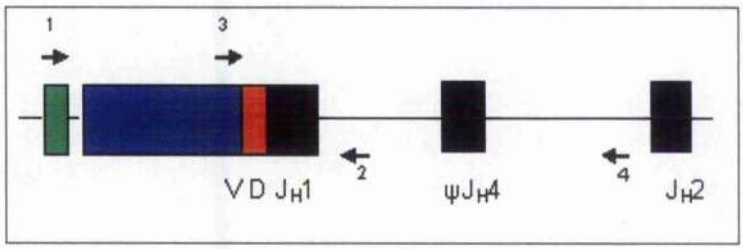


Figure 6: Organisation and the primer binding sites in the rearranged IgH locus

Of the two functional J_H segments (J_H1 and J_H2), the J_H1 on BTA 21 rearranges on most occasions. 1/2-VHF/IR 0.5; 3/4-FR3/IR1 are the primer pairs used to retrieve the rearranged IgH locus. Forward primers, Primer 1(VHF) targets conserved V_H leader and primer 3 (FR3) binds to conserved FR3 coding sequence ensuring successful amplicons irrespective of the V_H segment involved.

Amplification was therefore only possible from genomic DNA templates in which rearrangement at the Ig heavy chain locus had taken place. PCR products were on average about 470 bases, which correspond to the expected length of a rearranged VDJ locus accepting that the length of the CDR3 is widely variable. In total, 22 sequences were obtained and analysed from animals of one day old, 21 sequences from a 7 weeks old calf, 13 sequences from one of 13 weeks and 17 sequences from adult animals. The germline sequences for comparative analysis were obtained either previously in the laboratory (GenBank accession numbers U55164-69, 71-72, 74-75) by Sinclair (Sinclair *et al.*, 1997) or clones R21-1,2,3, LU-91-1,2 and NI-84-1 taken from the study of Berens *et al.*(Berens *et al.*, 1997).

3.3.1.2: $J_H1/\psi J_H4/J_H2$ intron

The forward primer FR3 used to recover J_H intronic sequences from a rearranged heavy chain locus was complementary to bases 333 to 552 of *Bos taurus* immunoglobulin heavy chain variable region gene, GenBank accession no. U55165 (Sinclair *et al.*, 1997). This targets conserved FR3-coding sequences of the V_H segments. The reverse primer IR-1 binds to bases 1636 to 1654 downstream of ψJ_H4 (Zhao *et al.*, 2003) and therefore recovers non-translated, intronic sequences only from the rearranged heavy chain locus. Most PCR products were on average about 790 bases, which correspond to the expected length of a rearranged Ig locus again factoring in CDR3 variability. For mutation analysis, the sequence spanning $J_H1/\psi J_H4/J_H2$ intron was used. In total, 19 sequences were obtained and analysed from a day old animal, 17 from a calf of 7 weeks, 9 from one of 13 weeks and 12 sequences from adult animals.

3.3.2: Comparison of germline and rearranged sequences

At the time of analysis, only 16 germline V_H segments were available at GenBank against which diversifying substitutions could be identified. Comparison of rearranged sequences against this data set might therefore score motifs from unidentified V_H segments as potential substitutions thereby over-estimating diversification frequency. Similarly, the IgH locus has not been analysed in sufficient detail to allow identification of the V_H alleles. To provide a secure foundation for the identification of substitutions, germline sequences were aligned with experimental data gathered from animals of all ages (data not shown). To determine whether the nucleotide patterns observed in rearranged sequences originated from the process of diversification or from unidentified germline IgH segments, a set of criteria was established (Figure 7). Individual nucleotides and motifs common to more than one entry in the alignment were identified and considered to represent germline segments or their alleles according to these criteria. Some of these segments are shown in Figure 8 and have been placed into five groups based on the identified motifs. The motifs occur throughout the rearranged genes but are prominent in the CDRs. The criteria were stringent, but it was argued that the likelihood of common nucleotides arising through diversification was substantially less than through rearrangement of a small but incompletely characterised pool of V_H segments. The germline sequences that were used for comparative analysis are shown in Figure 9. Hence when assessing diversification in animals of different ages, only unique nucleotides outside these criteria were counted. Using these criteria, the mutations that are scored and used for interpretation have been underlined in a series of alignment figures from variable region and downstream intronic sequences (Figures 10 to 13, 15 to 18).

The enzyme Taq polymerase is error prone and has a tendency to cause base-pair transitions (Barnes, 1992; Keohavong and Thilly, 1989). The possible contribution of polymerase errors was analysed by amplification from a plasmid containing a sequenced Ig insert. Under these experimental conditions, an error rate of 1.04×10^{-4} substitutions/base/pcr cycle was observed.

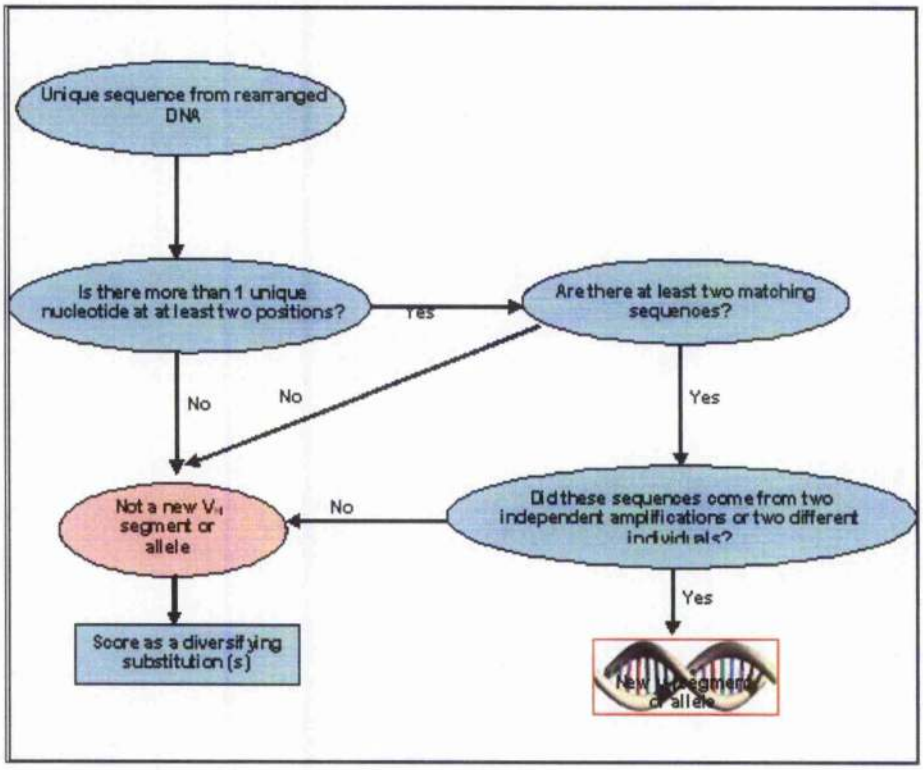


Figure 7: A flow chart depicting the criteria adopted to define a new gene segment

Taking this into consideration the diversifying frequencies calculated do not change significantly. In addition, PCR has the potential to generate PCR hybrid sequences, indistinguishable from the results of gene conversion events (Ford *et al.*, 1994). In this study, clone 16 recovered from a calf at one day of age matches to clone 25 up to nucleotide position 180 and thereafter it matches clone 1 in Figure 10. In this description position 1 represents the first nucleotide of FR1 based on accession number U55165. The probability of this being a chimeric sequence is further increased as both the clones were from a single PCR. No other candidate hybrids were detected.

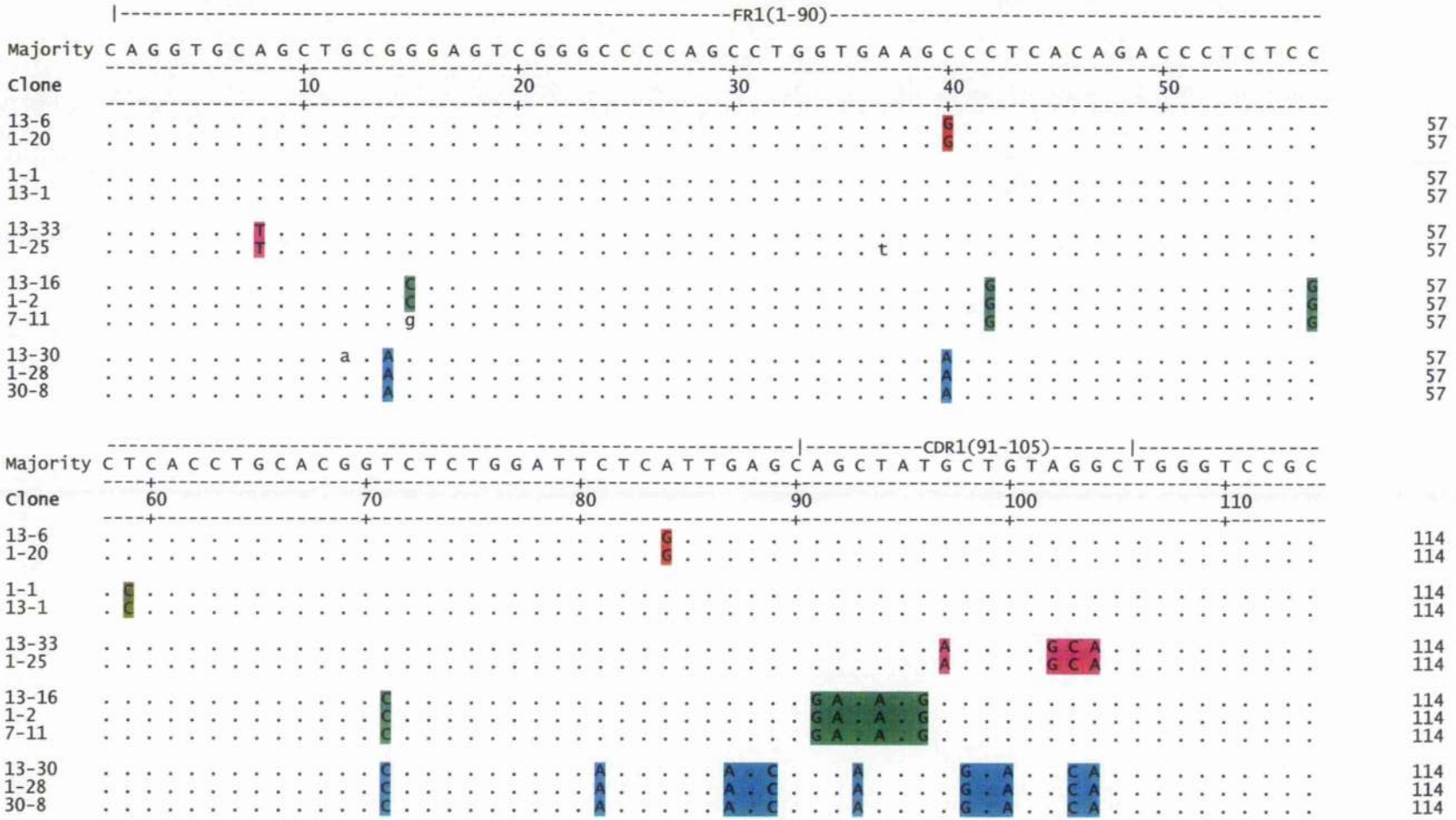


Figure 8: Unique IgH sequences recovered in this study are shown

IgH germline sequences taken from GenBank (Sinclair *et al.*, 1997)and a published source (Berens *et al.*, 1997) were aligned with experimental data from this study and individual nucleotides and motifs common to more than one entry from two different PCR or two individuals in the alignment were identified and considered to represent uncharacterised germline segments or its alleles. This provided the strict criteria for identification of diversifying substitutions. The bases in lower case letter were scored in the later mutation analysis from individual age groups. Nucleotide identity between the sequences is denoted by dot (“.”), differences indicated by letter code and gap by dash (“-”). The particulars of the clones are shown in the left-hand side of the Figure and clone 13-6 means clone 6 from a 13 weeks old animal. The numbering on the right-hand side represents the last position of the nucleotide. The same colour pattern within the individual block represents unique motifs.

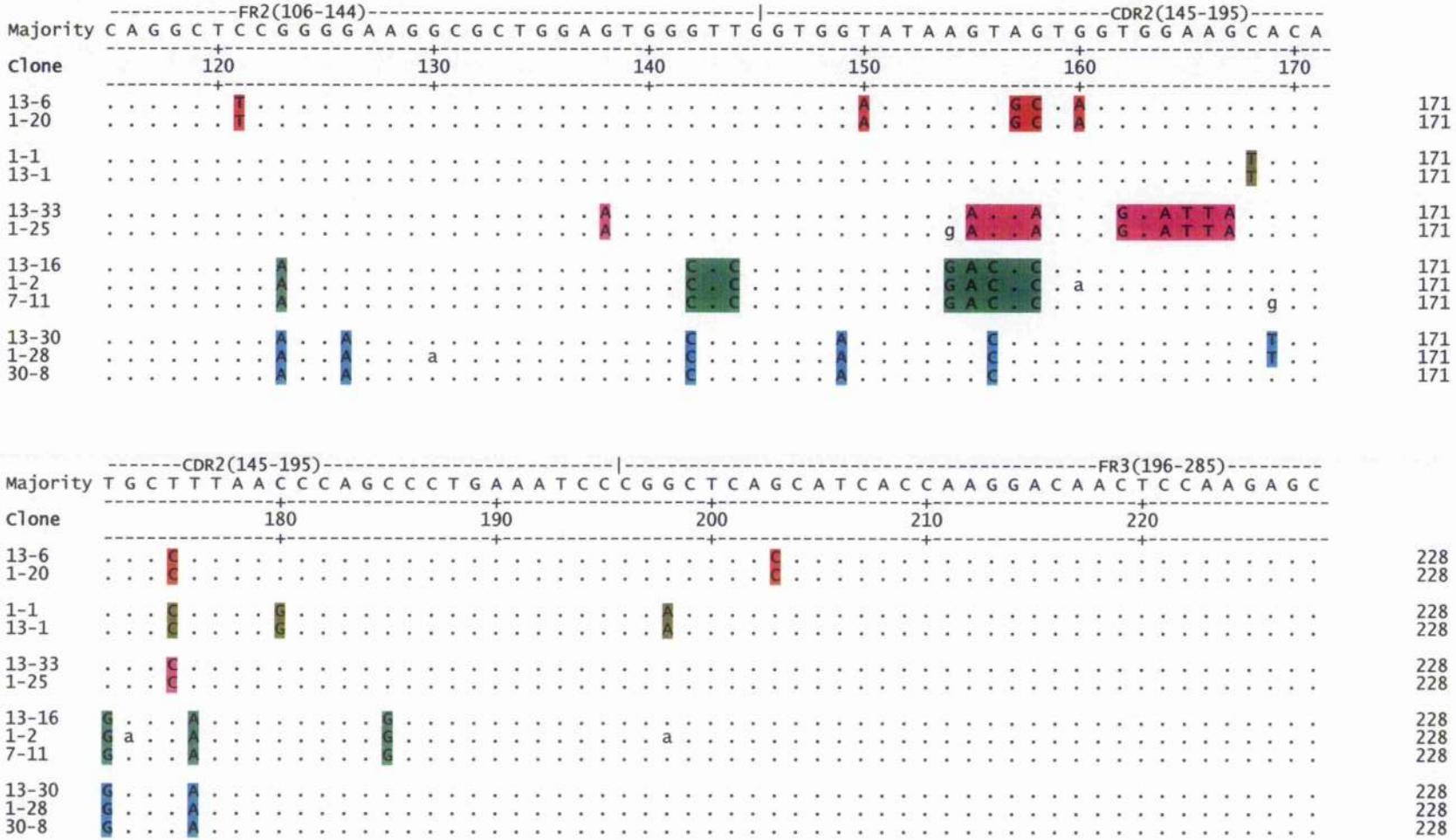


Figure 8 continued

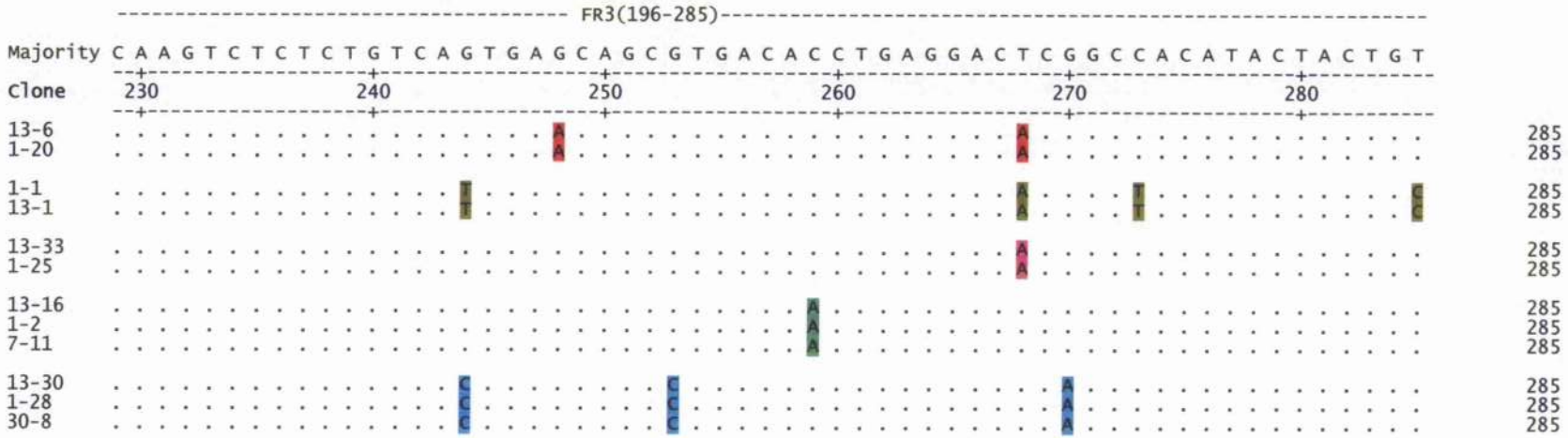


Figure 8 continued

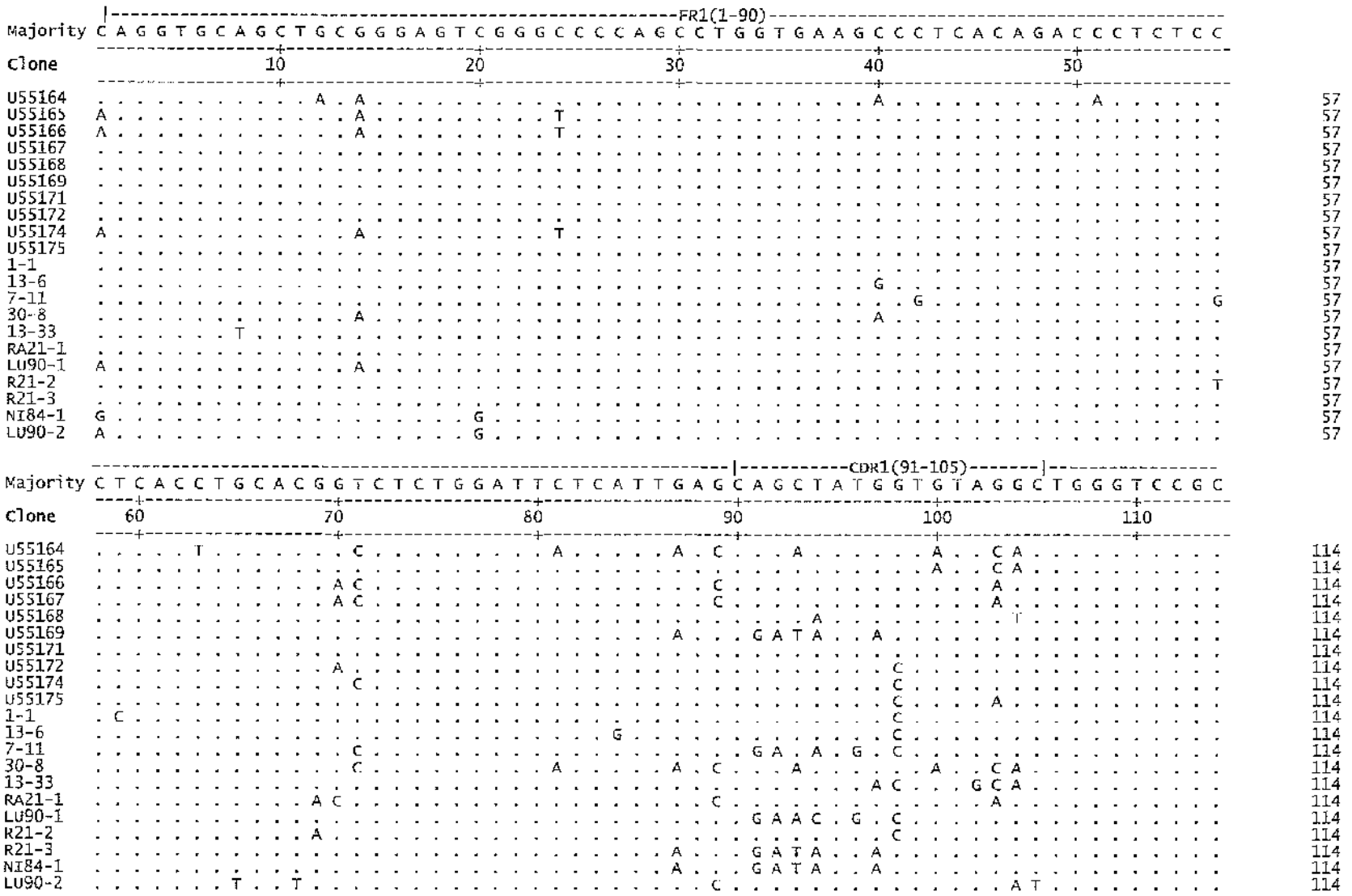


Figure 9: Alignment of the IgH germline sequences

IgH germline sequences taken from GenBank (Sinclair *et al.*, 1997) and published sources (Berens *et al.*, 1997) were aligned in order to identify the extent of diversity that exist amongst these sequences at the germline level. These diversifying nucleotides (in addition to ones in the newly identified unique sequences-Figure 8) were not counted when scoring mutations in the rearranged sequences. Nucleotide identity between the sequences is denoted by dot ("."), differences indicated by letter code and gap by dash ("-").The numbering on the right-hand side represents the last position of the nucleotide.

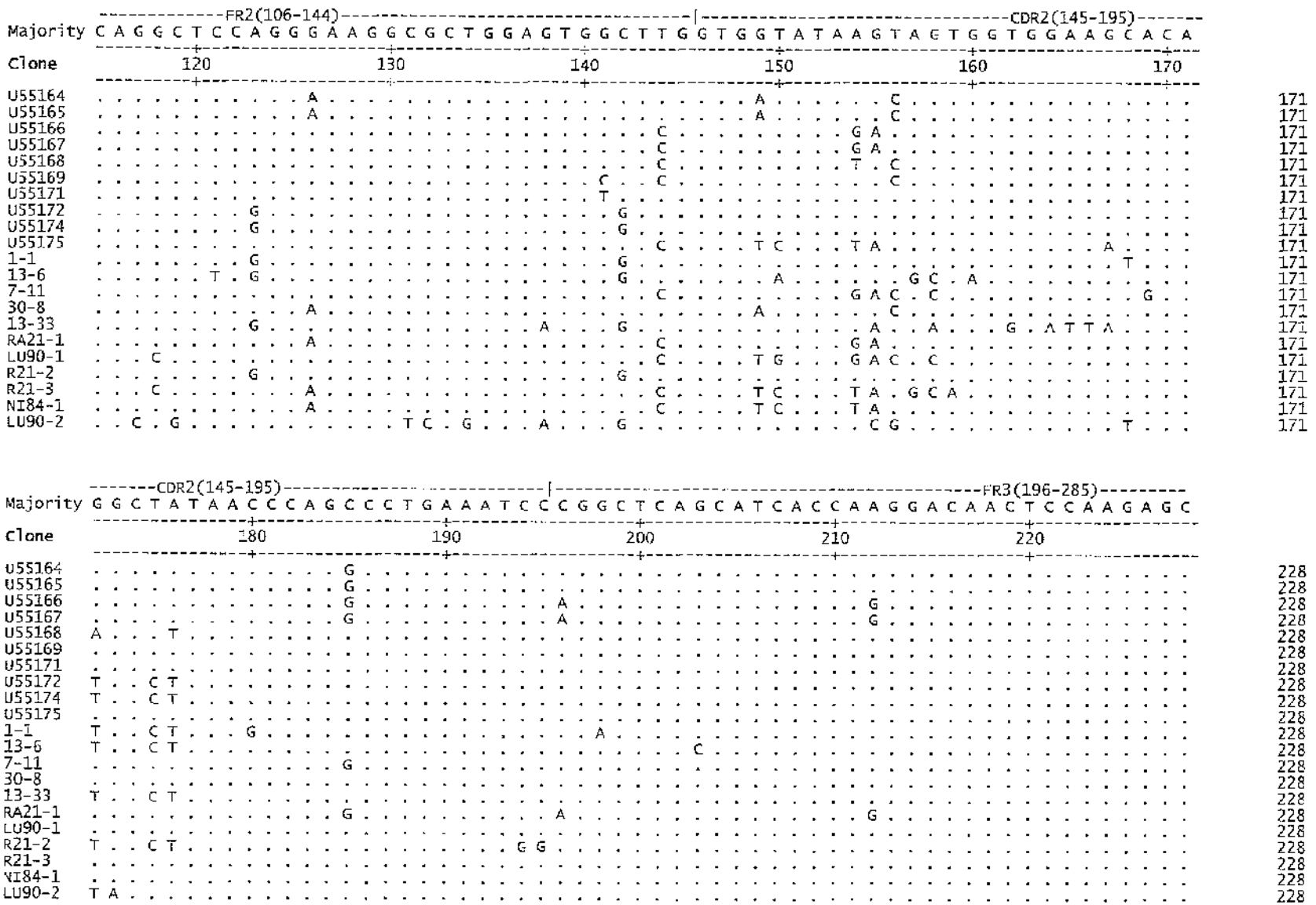


Figure 9 continued

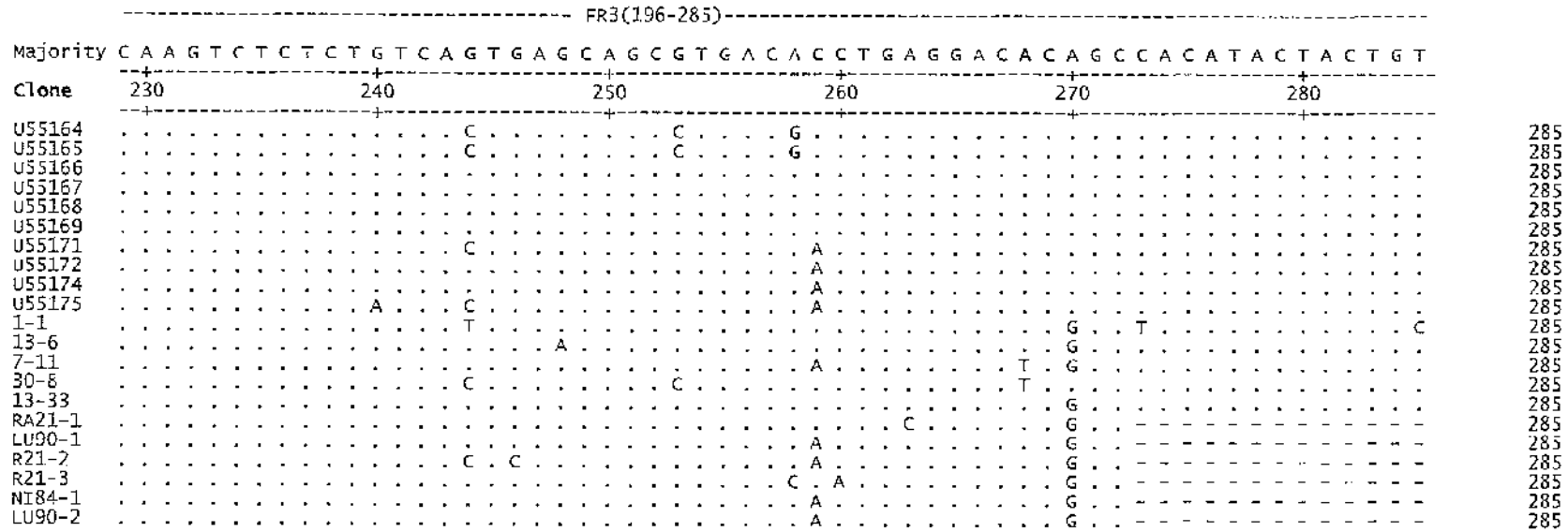


Figure 9 continued

3.3.3: Global analysis

Taking animals of all ages, mutations affected primarily purines (59%) and the ratio of transitions to transversions was 1.6 suggesting a bias towards transitions (Table 3). Moreover, the mutations from A to G and G to A accounted for 36.4% of all the changes. The most frequently targeted bases were G and T. Apart from substitutions, 17 deletions spanning 1 to 26 nucleotides and 5 insertions in the range of 1 to 3 were observed. Mutations were scored from the sense strand of the DNA so that an A to T replacement means that an A:T pair was replaced by a T:A pair.

Table 3: Bias of mutations in rearranged bovine IgH sequences

	TO				
FROM	**T**	**C**	**A**	**G**	**MUTATIONS***
T	NA	70(58%)	19(16%)	32(26%)	121(13%)
C	154 (61%)	NA	34(13%)	64(25%)	252(28%)
A	31(22%)	18(13%)	NA	95(66%)	144(16%)
G	54(14%)	104(26%)	237(60%)	NA	395(43%)

The region analysed spans a total of 946 nucleotides comprising 82 bases in the leader intron, 285 bases in FR1-FR3, and 579 bases in the $J_H2/C\mu$ intron. *Total mutations identified (except insertions and deletions). The percent of mutations appears in parentheses and has been corrected for the base composition of the region analysed.

3.3.4: FR1-FR3 rearranged region

The distribution of mutations over 285 bases comprising FR1 to FR3 of the heavy chain V region recovered from different ages are shown in Figure 10 (day old), Figure 11 (7 weeks), Figure 12 (13 weeks) and Figure 13 (adult animal < 30 months). The nucleotides that met the criteria outlined earlier were judged to result from mutation and were included for mutation analysis. In total, 22 sequences were analysed from an animal of one day old. In total, 31 quantifying substitutions were identified, equivalent to 1.1 per sequence. From an animal of seven weeks of age, 21 sequences encountered 44 nucleotide substitutions thus yielding overall 2.1 mutations per sequence. The mutations per sequence increased further to 3.3 in 13 weeks old animal where 13 sequences analysed encountered 43 substitutions.

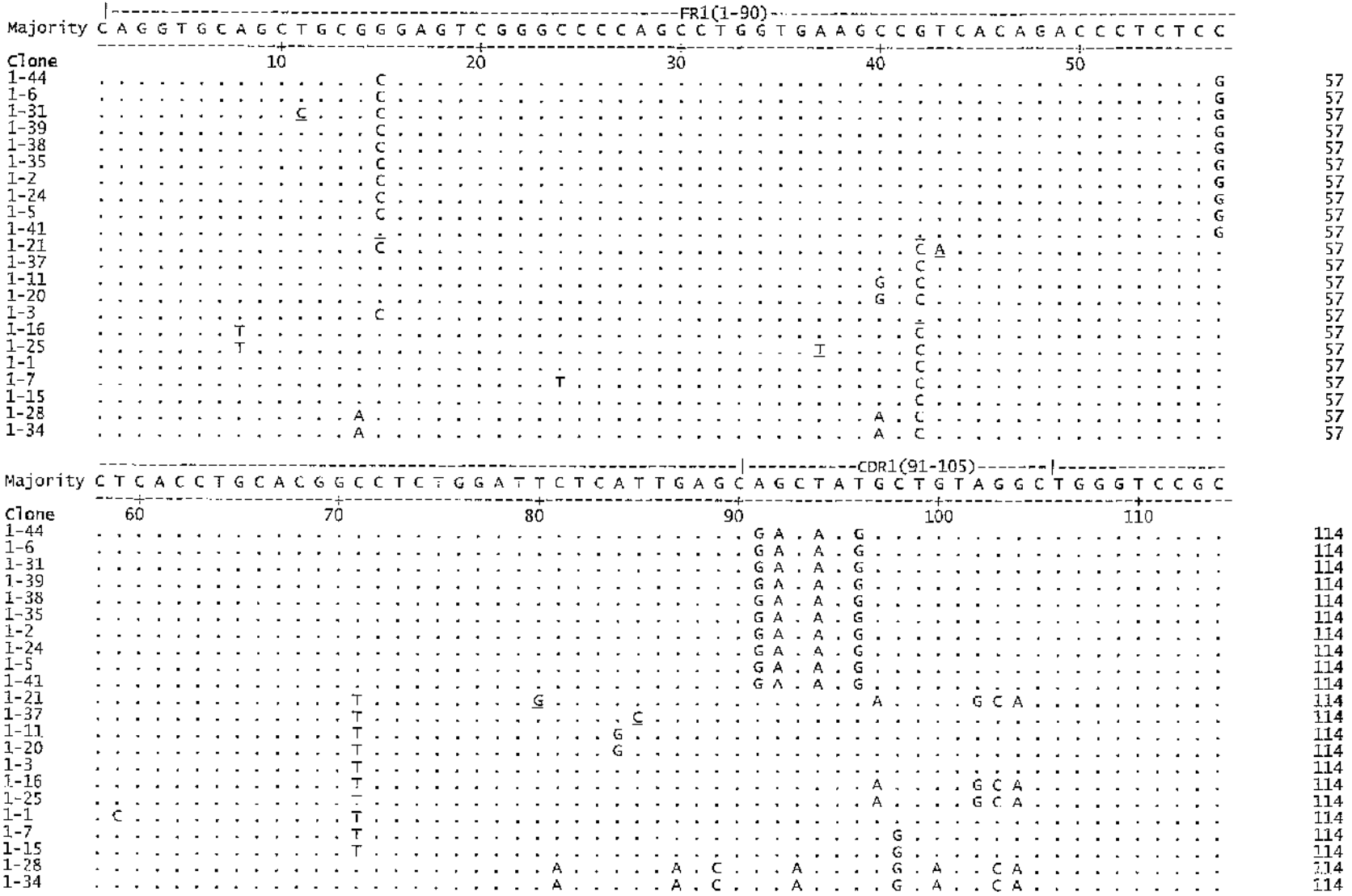


Figure 10: Nucleotide alignment of sequences recovered from a day old animal

The sequences were aligned against the germline and newly identified sequences. Nucleotide identity between the sequences is denoted by dot (“.”), differences indicated by letter code and gap by dash (“-”). The mutations scored are underlined. The numbering on the right-hand side represents the last position of the nucleotide.

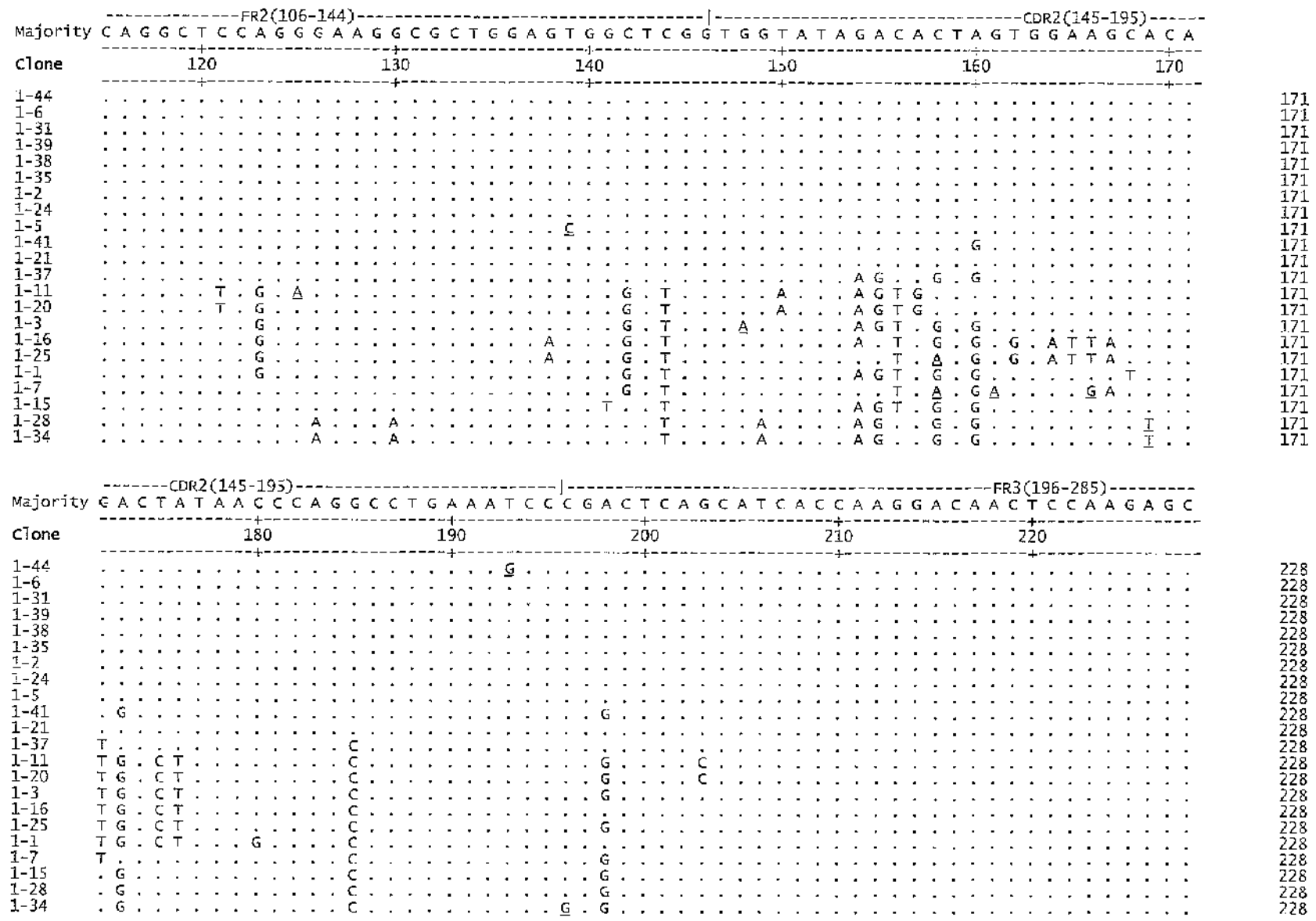


Figure 10 continued

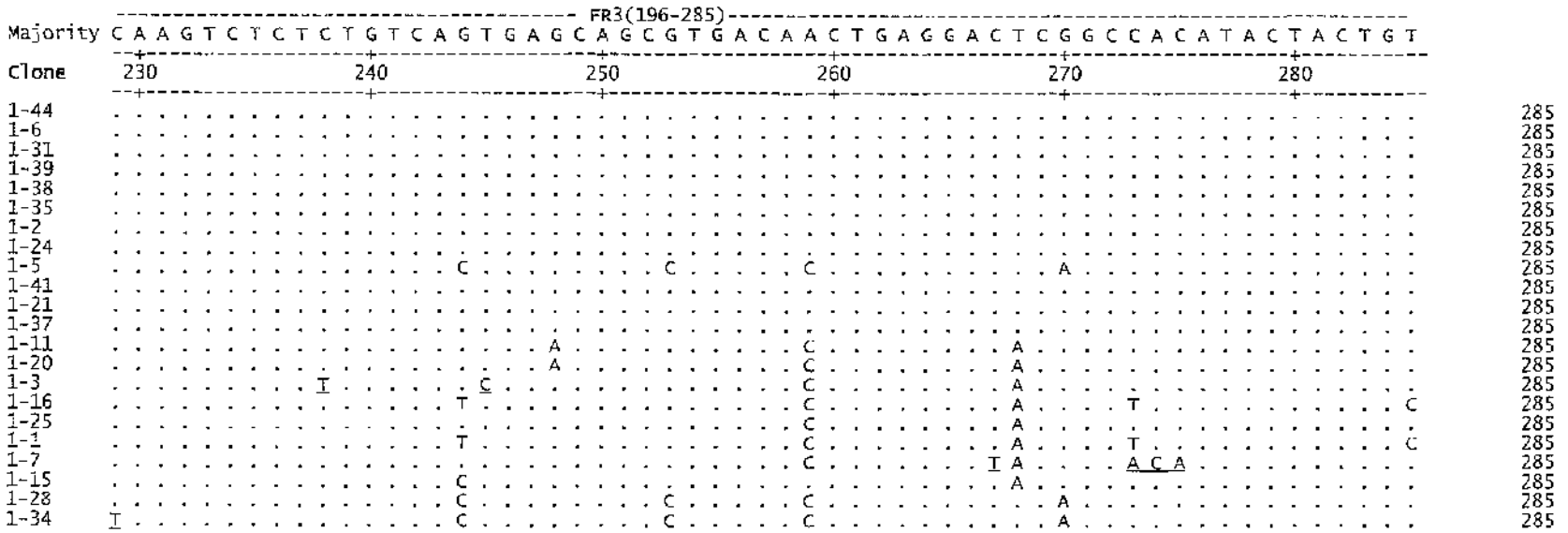


Figure 10 continued

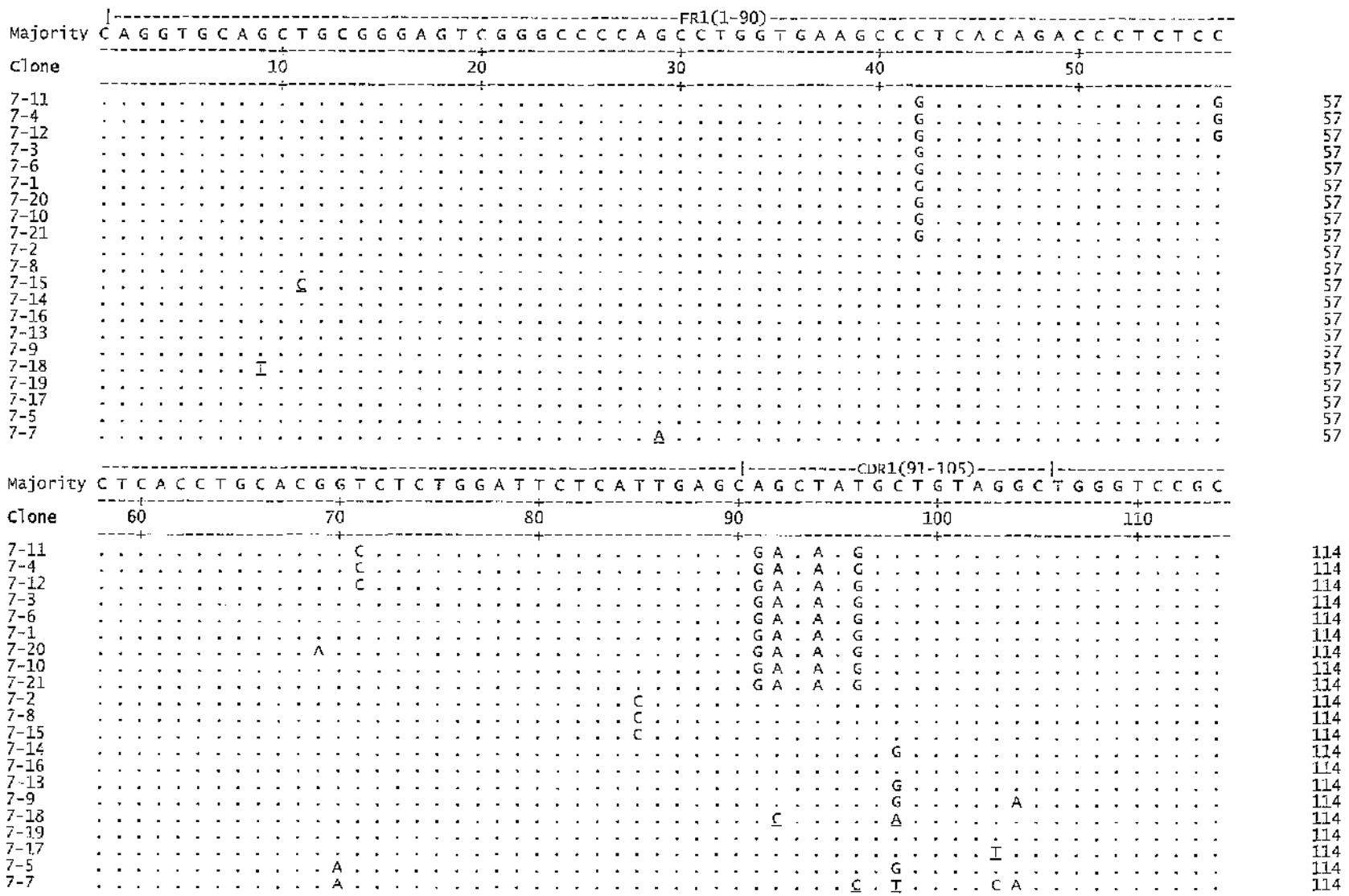


Figure 11: Nucleotide alignment of sequences recovered from seven weeks old animal

The sequences were aligned against the germline and newly identified sequences. Nucleotide identity between the sequences is denoted by dot (“.”), differences indicated by letter code and gap by dash (“-”). The mutations scored are underlined. The numbering on the right-hand side represents the last position of the nucleotide.

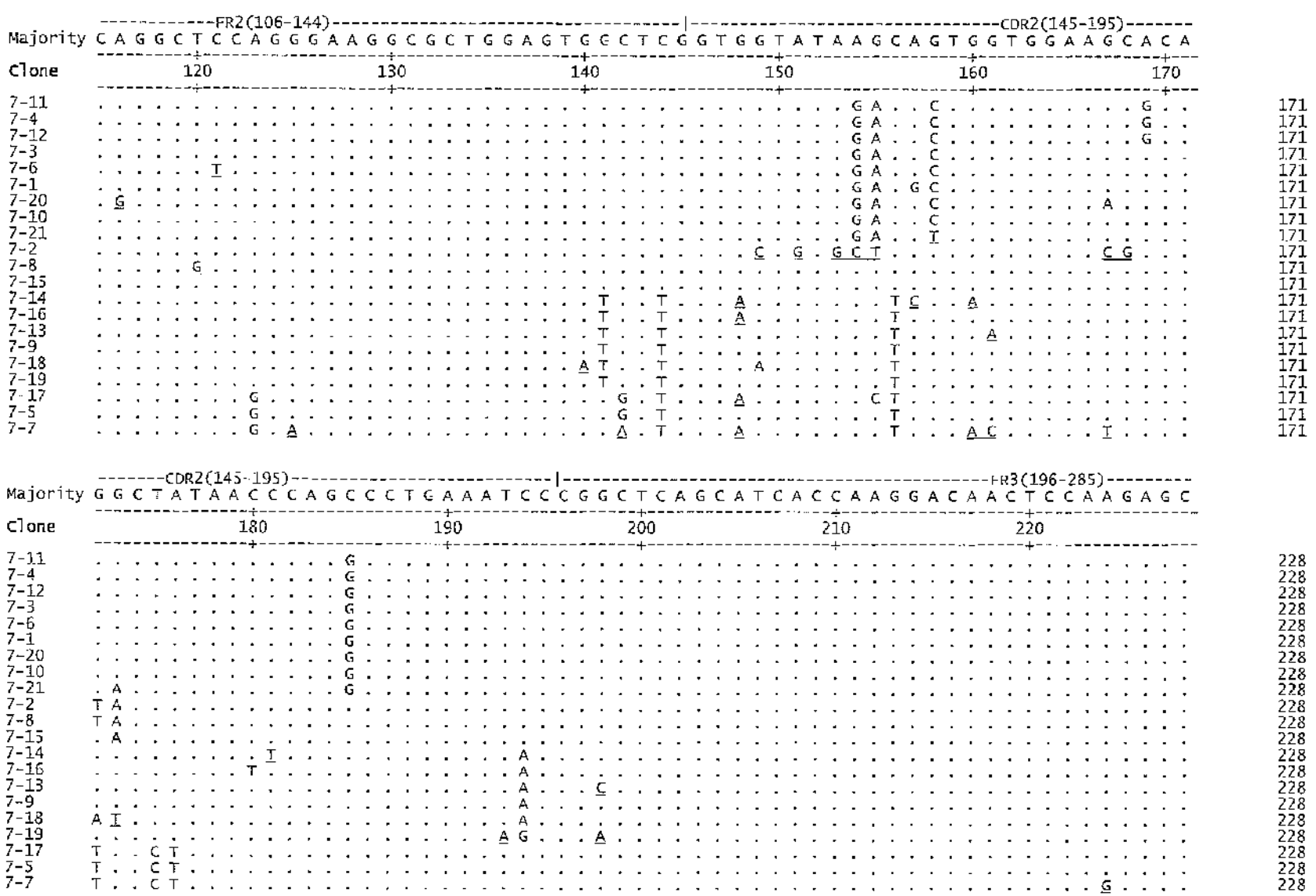


Figure 11 continued

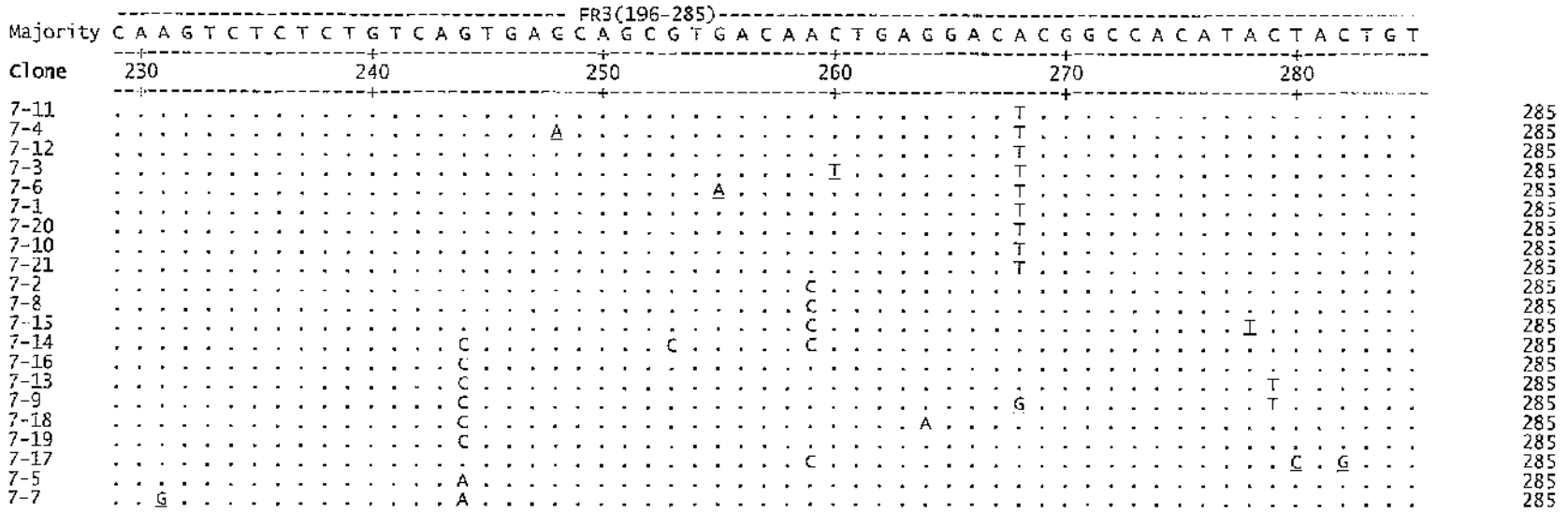


Figure 11 continued

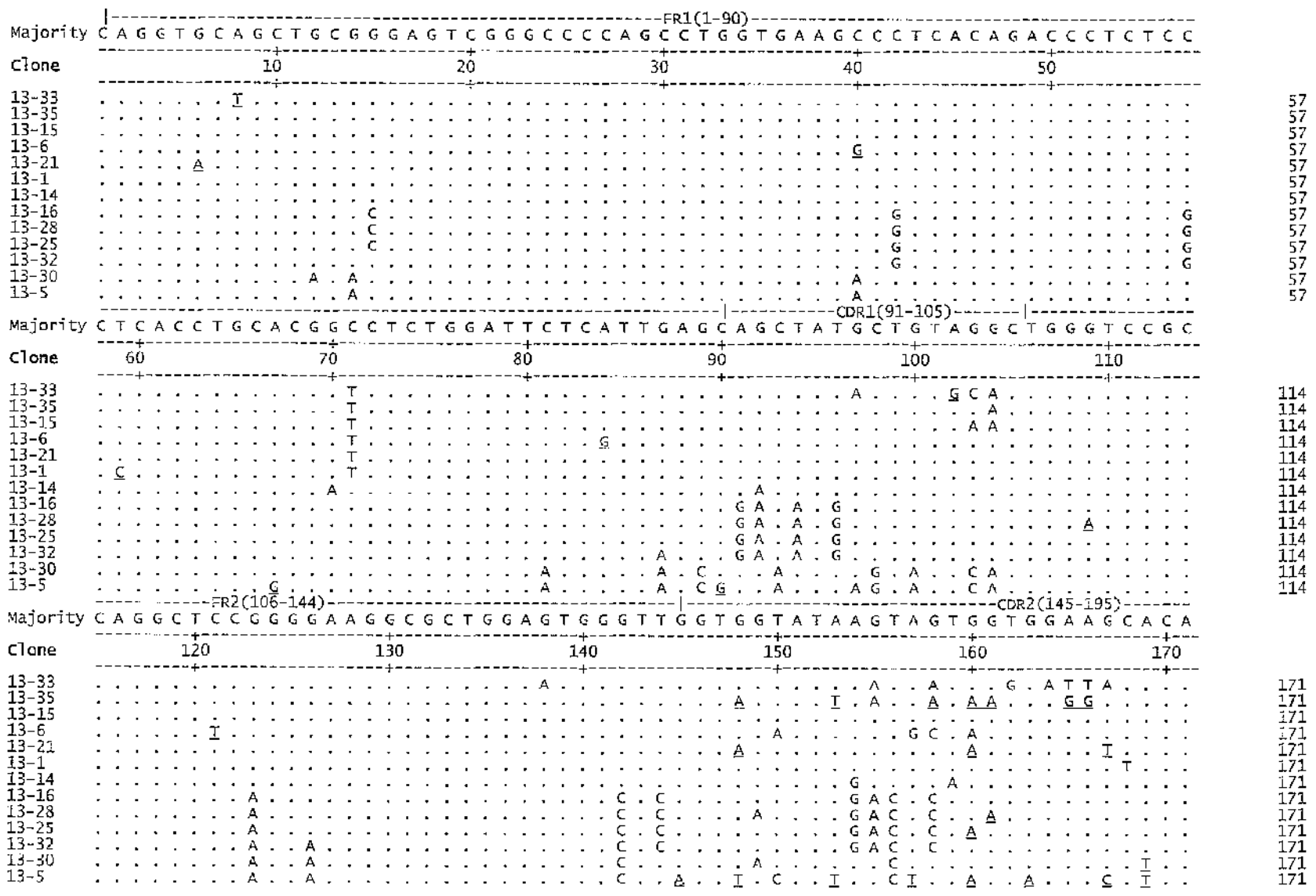


Figure 12: Nucleotide alignment of sequences recovered from 13 weeks old animal

The sequences were aligned against the germline and newly identified sequences. Nucleotide identity between the sequences is denoted by dot (“.”), differences indicated by letter code and gap by dash (“-”). The mutations scored are underlined. The numbering on the right-hand side represents the last position of the nucleotide.

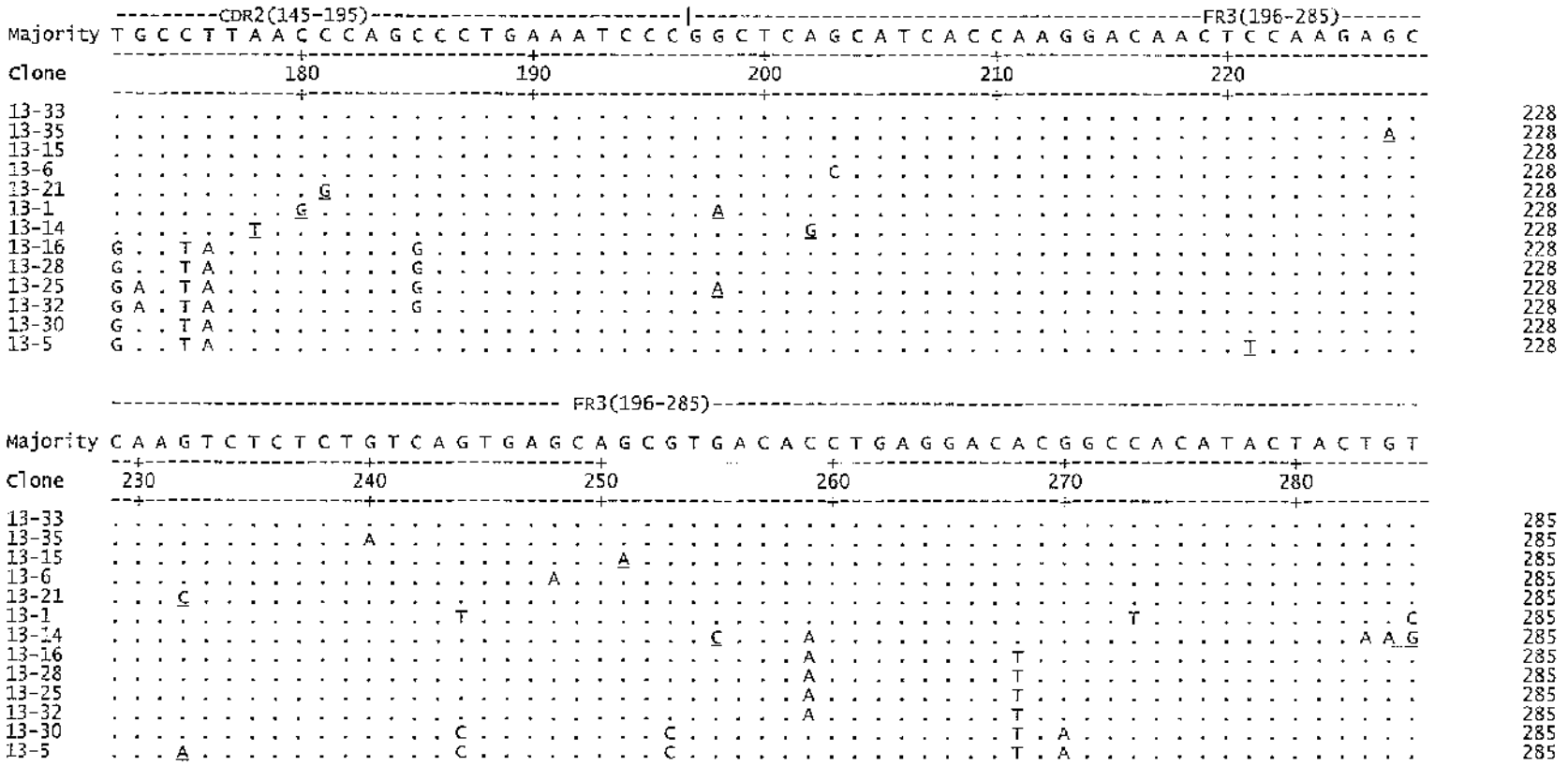


Figure 12 continued

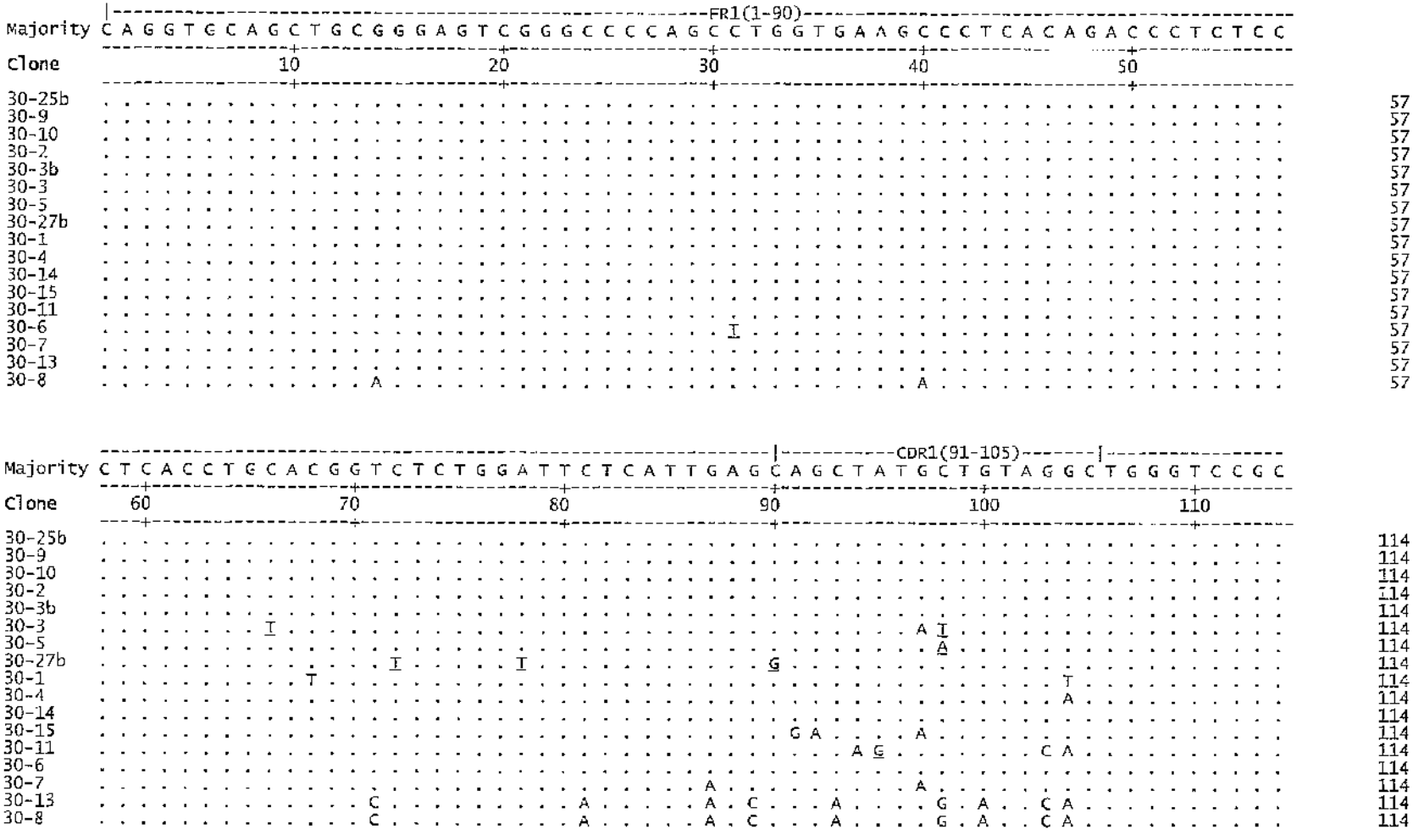


Figure 13: Nucleotide alignment of sequences recovered from adult animal (less than 30 months old)

The sequences were aligned against the germline and newly identified sequences. Nucleotide identity between the sequences is denoted by dot ("."), differences indicated by letter code and gap by dash ("-"). The mutations scored are underlined. The numbering on the right-hand side represents the last position of the nucleotide.

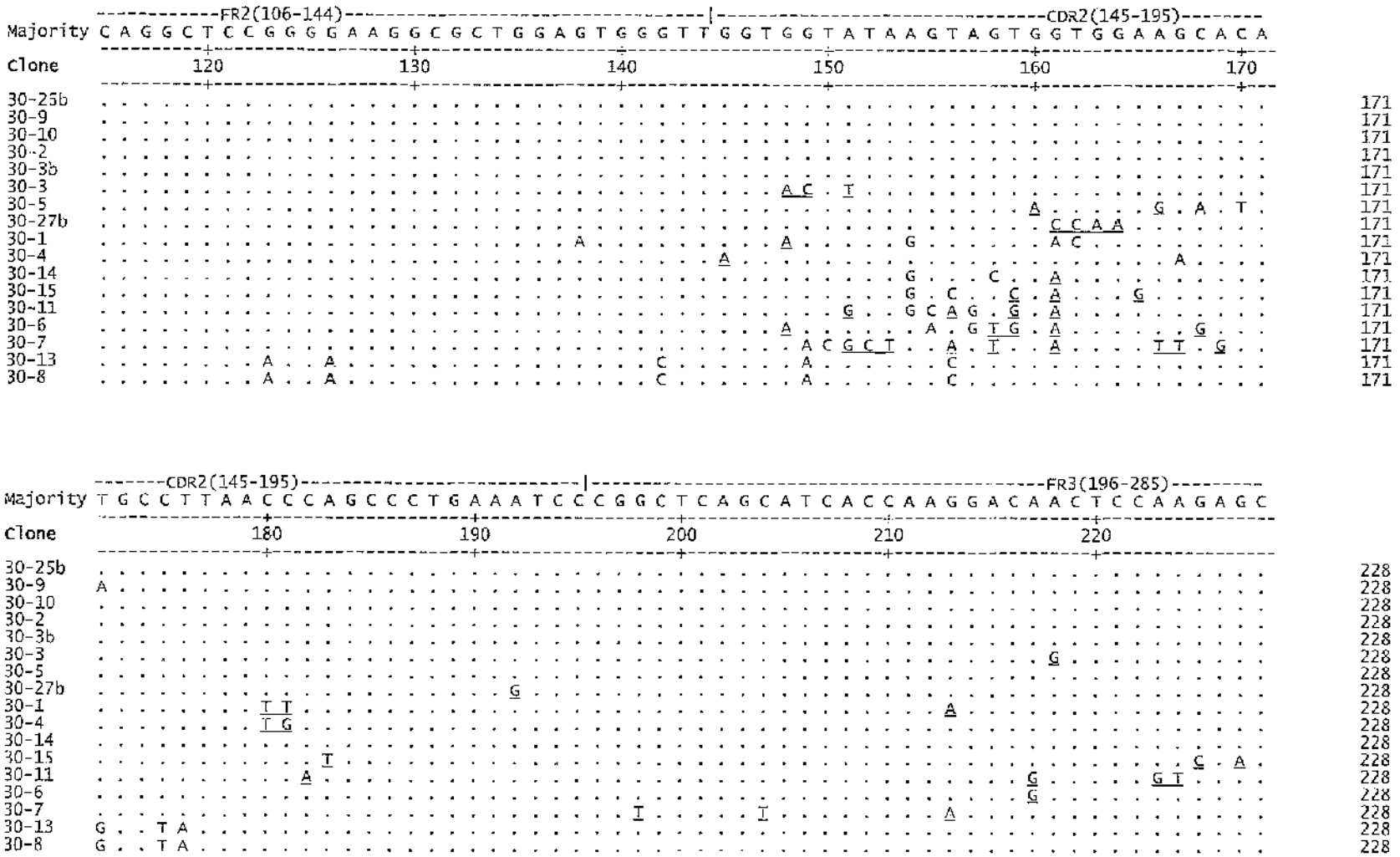


Figure 13 continued

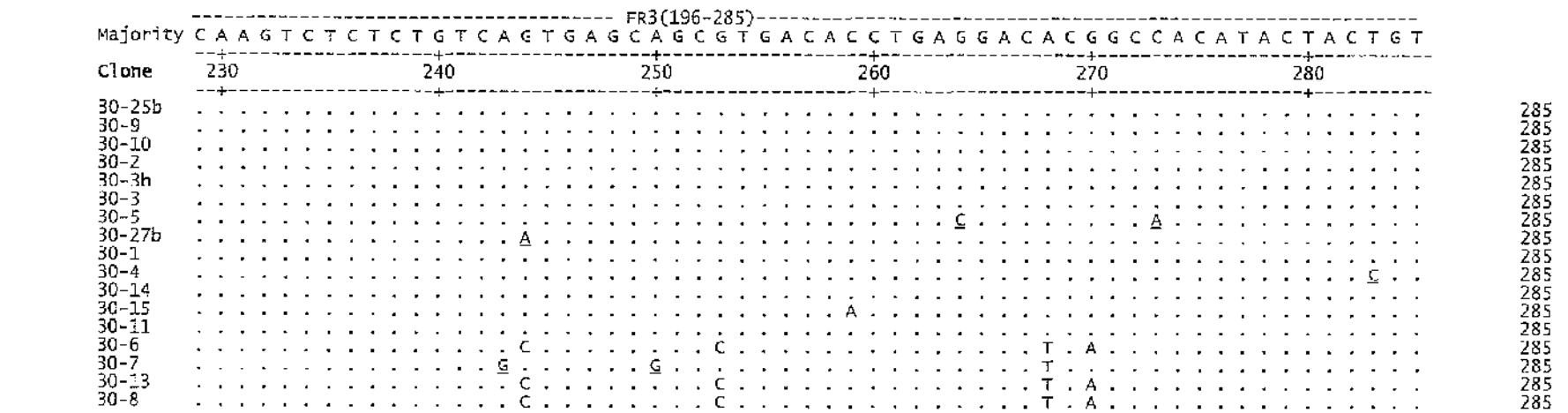


Figure 13 continued

The average number of mutations per sequence was 3.4 in 30 months old animal which was almost similar to as recorded for 13 weeks animal. A total of 57 mutations were encountered among 17 sequences from adults spanning region FR1- FR3 (Table 4). The clustering of nucleotide changes around the CDRs was evident in the different age groups (Table 4). Irrespective of the age group studied, mutated sequences were consistently observed more frequently in CDR2 than in CDR1. This was confirmed by Fisher's exact test. All values of p were less than 0.02 except in the 7 week age group in which significance was still evident (0.043). These values emphasize a significant difference from the null hypothesis that CDR2 undergoes similar level of mutation as CDR1.

Table 4: Region-wise distribution of mutations

REGION/LENGTH (BP)	*AGE GROUPS*							
	Day old		7 weeks		13 weeks		Adult	
	total	per kb	total	per kb	total	per kb	total	per kb
LEADER INTRON/82	7	3.9	3	1.7	10	9.4	6	4.1
FR1 /90	8	4.0	3	1.6	7	6.0	5	3.1
CDR1 /15	0	0	5	15.9	1	5.1	3	11.1
FR2 /39	2	2.3	6	7.3	2	3.9	0	0
CDR2 /51	8	7.1	18	16.8	23	35.0	33	36.0
FR3 /90	6	3.0	12	6.3	10	8.5	16	9.9

In total 22, 21, 13 and 17 sequences were analysed from day old, 7 weeks, 13 weeks and adult animals, respectively. CDRs encountered higher rate of muations as compared to FRs. The data shows an approximate doubling in substitution rate with CDR2 between birth and 7 weeks of age, and a comparable increase over the next six weeks.

Both increases were statistically significant when assessed with Fisher's exact test (day old vs 7 week, p=0.028; 7 week vs 13 week, p=0.014). No significant increase was evident in samples from animals of 13 weeks and maturity (p=0.419). Taking a closer look at the mutation rate in the CDRs and FRs, it is evident that there is a progressive increase in the diversity with age. This difference is very striking across the CDR2 and FR3 amongst all the age groups. The higher rate of substitutions in CDR2 noted earlier also showed an age dependency. In Figure 14, the substitution rate observed in CDR2 has been normalised for animals from the four age categories and plotted alongside data from FR3. The data shows an approximate doubling in substitution rate with CDR2 between birth and 7 weeks of age, and a comparable

increase over the next six weeks. No increase was evident in samples from animals of 13 weeks and maturity. Substitution rates were lower in FR3 than CDR2 in animals of one day old, but also doubled over the first seven weeks of life. Thereafter, no substantial rise in frequency was observed. The ratio of transitions to transversions in the region spanning FR1-FR3 was 1.5 across the entire dataset. Transitions were 4 times higher than transversions in sequences obtained from day old compared to only 1.5 times more in sequences obtained from 7, 13 and adult animals taken together.

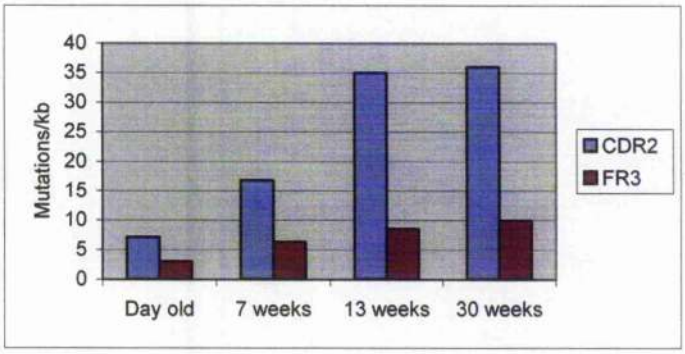


Figure 14: The normalised mutations/kilobase in CDR2 and FR3 from animals of different ages

Mutations encountered in FR3 were much less than in CDR3 irrespective of the age group studied. Higher mutations were encountered in older animals.

Purines were affected more (65%) than pyrimidines (35%), although purines and pyrimidines are almost equally represented in the germline genes (48.9% versus 51.1%).

Nucleotide substitutions in genes that result in change of amino acid are known as non-synonymous mutations, whereas the substitutions that do not change the amino acids are called synonymous. Most non-synonymous changes are eliminated by selection at the protein level, but under certain conditions they may be retained. Investigating the ratios of synonymous and non-synonymous substitutions in the different regions of the V_H segments may therefore provide information about the inherent bias in the system to undergo diversifying selection. Using a SNAP software, the ds:dn values expected in the case of a random mutation of the germline genes were found to be 2.5 and 0.6 for FRs and CDRs, respectively. This analysis revealed that CDR and FR sequences can differ significantly in their inherent susceptibility to amino acid replacement

given any single nucleotide change and that the CDR composition introduces an intrinsic bias toward nonsynonymous mutation. The germline ds:dn values could not be used as a reference base to evaluate the effect of ongoing selective pressure on the rearranged sequences as these sequences came from different animals raised in different locations and furthermore the variability in the CDRs and FRs was driven by both mutation and the use of VH segments which varied in nucleotide composition. The ds:dn analysis on the rearranged sequences recovered for this region from different age groups revealed values less than one for both CDRs and FRs. This suggests bias towards amino acid replacement. Overall, the mutations in FRs and CDRs resulted more often in amino acid replacements than silent substitutions. There was no specific trend that could be linked to the age of the animals. However in a day old and 7 weeks old animals, all but 2 mutations were replacements in the CDRs. Altogether, on average replacements were much more pronounced in the CDRs as compared to FRs. In FRs nonsynonymous mutations were recorded 1.1 to 2 times higher than synonymous mutations, where as in case of CDRs, there were 2.1 to 10.5 times more nonsynonymous mutations.

Sequences recovered from the day old animal had considerably fewer nucleotide substitutions than those of the older age groups (7, 13 and adult) although the distribution of mutations in sequences in different age groups was broadly similar with more mutations recorded in the CDRs. Substitutions ranged from 12.1% in the day old calf to 54.5% in adult animals within the CDRs. The proportion of substitutions in the FRs was less variable, ranging between 7.3% at day old to 9.6% at maturity (Table 5).

Table 5: Percent mutation and observed ds:dn values for FRs and CDRs

Region	*Day old*	*7 weeks*	*13 weeks*	*30 months*
FRs (% mutations)	7.3	9.6	8.7	9.6
FRs ds:dn	5.9	3.9	4.3	2.1
CDRs(% mutations)	12.1	34.8	36.4	54.5
CDRs ds:dn	0.4	0.4	0.5	0.8

Mutation percentages in the table are based on total mutations observed in that region. The ds:dn values are based on SNAP analysis of rearranged sequences. This analysis does not consider the contribution of different germline segments and targeting bias of SHM process and hence does not reflect the true picture of silent and replacement mutations at the protein level.

3.3.5: Nucleotide changes in the 5' flanking region of leader intron

The same criteria were applied for analysis of nucleotide changes in the partial leader intron recovered in the PCR. No leader sequences were available from clones R21-1,2,3, LU-91-1,2 and NI-84-1 for comparison. When comparing the sequences recovered from animals of the same age, the accumulation of mutations was comparable between leader intron and FR1 (Table 4). No evidence emerged of a consistent age-dependent pattern of changes in this region.

3.3.6: Mutations in $J_H1/\psi J_H4/J_H2$ intron

The bovine J_H locus exists on BTA 21 where rearrangement of J_H1 segment forms the majority of bovine IgH chains (Zhao *et al.*, 2003). The locus has also been duplicated onto BTA 11 where low frequency recruitment of J_H segment into the rearrangement process can provide an alternative FR4 sequence to that derived from J_H1. When characterising the sequences isolated from the rearranged IgH locus, only two germline sequences can therefore contribute to non-coding sequences downstream from the IgH reading frame. PCR products were cloned and sequenced as described earlier, and identical criteria applied for identification of substituting nucleotides. The distribution of mutations from animals of different ages is shown in Figure 15, 16, 17 and 18. Further analysis is provided in Table 6 and Figure 19.

Table 6: Distribution of mutations $J_H1/\psi J_H4/J_H2$ intron

REGION/SIZE(BP)	*AGE GROUPS*							
	Day old		7 weeks		13 weeks		Mature	
	total	per kb	total	per kb	total	per kb	total	per kb
INTRON1 /344	74	11.3	164	28.0	73	23.6	125	30.3
$\psi J_H4/48$	14	15.3	21	25.7	11	25.5	17	29.5
INTRON2 /187	33	9.3	43	13.5	24	14.3	46	20.5

This non-coding region was highly mutated, and on average, substitutions rates were higher than the rearranged V_H segment. For this region, 60.5% of changes were transitions as opposed to 39.5% transversions and purines (57%) mutated with higher frequency than pyrimidines (43%). The intron 1 encountered mutations with frequency comparable to CDR2. Higher incidence of deletions

and insertions was evident for this region as compared to coding region of IgH. There was increase in the number of mutations with age but the trend did not hold true for the analysis on sequences recovered from animals of 13 weeks of age.

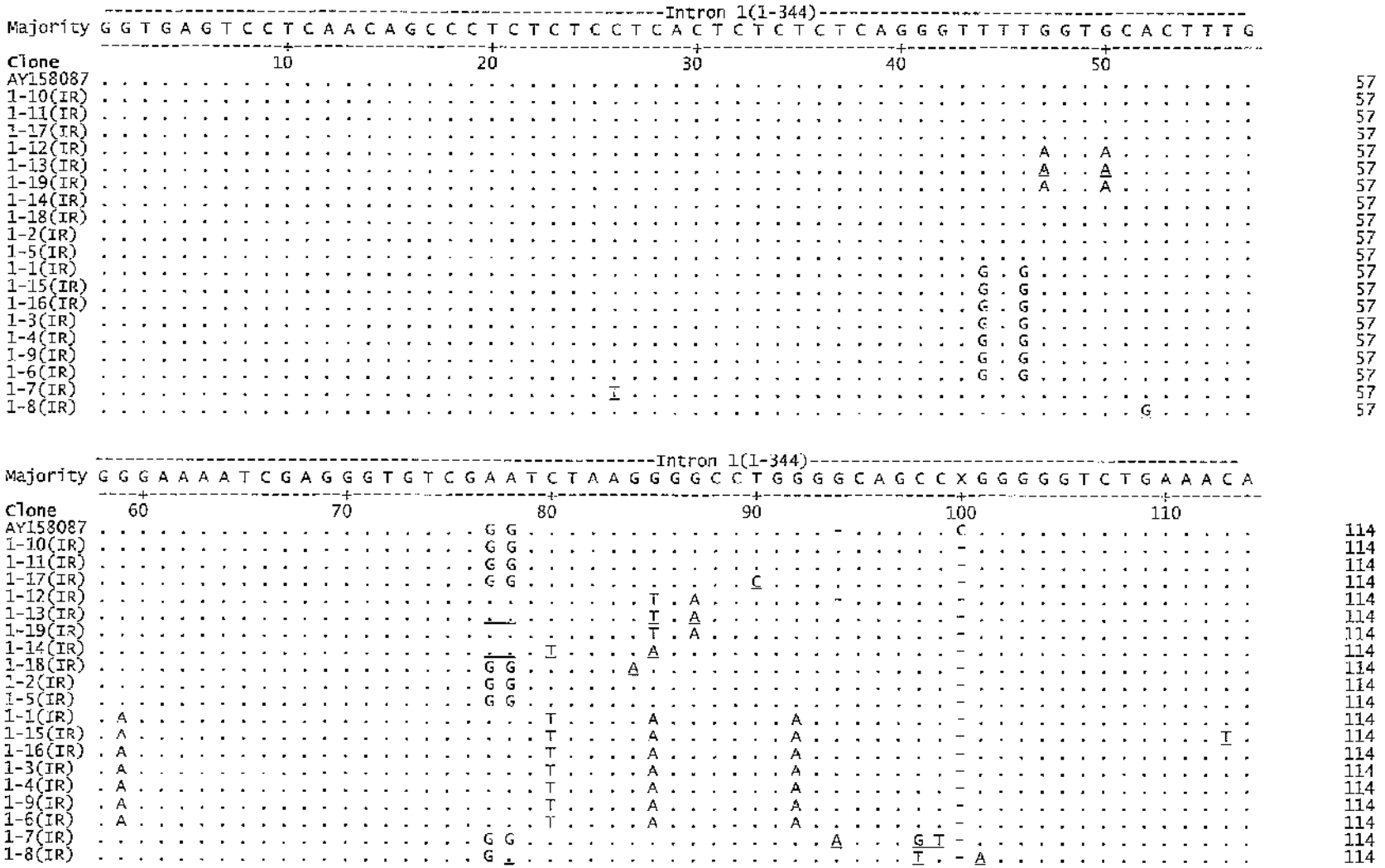


Figure 15:Nucleotide alignment of rearranged IgH locus sequences downstream of FR4 spanning regions of $J_H1/\psi J_H4/J_H2$ intron from a day old animal compared to most frequently rearranged bovine J_H locus on BTA 21 (AY158087)

The particulars of the clones are shown in the left-hand side of the Figure and clone 1-10(IR) means clone 10 from intronic region downstream of FR4 of a day old animal. Nucleotide identity between the secuences is denoted by dot (“.”), differences indicated by letter code and a gap by dash (“-”). Only the underlined bases have been considered for data analysis. The nucleotides which are not underlined belong to germline sequence (part of an either haplotype) or a multiple sequence from common PCR reaction. The numbering on the right-hand side represents the last nucleotide position.

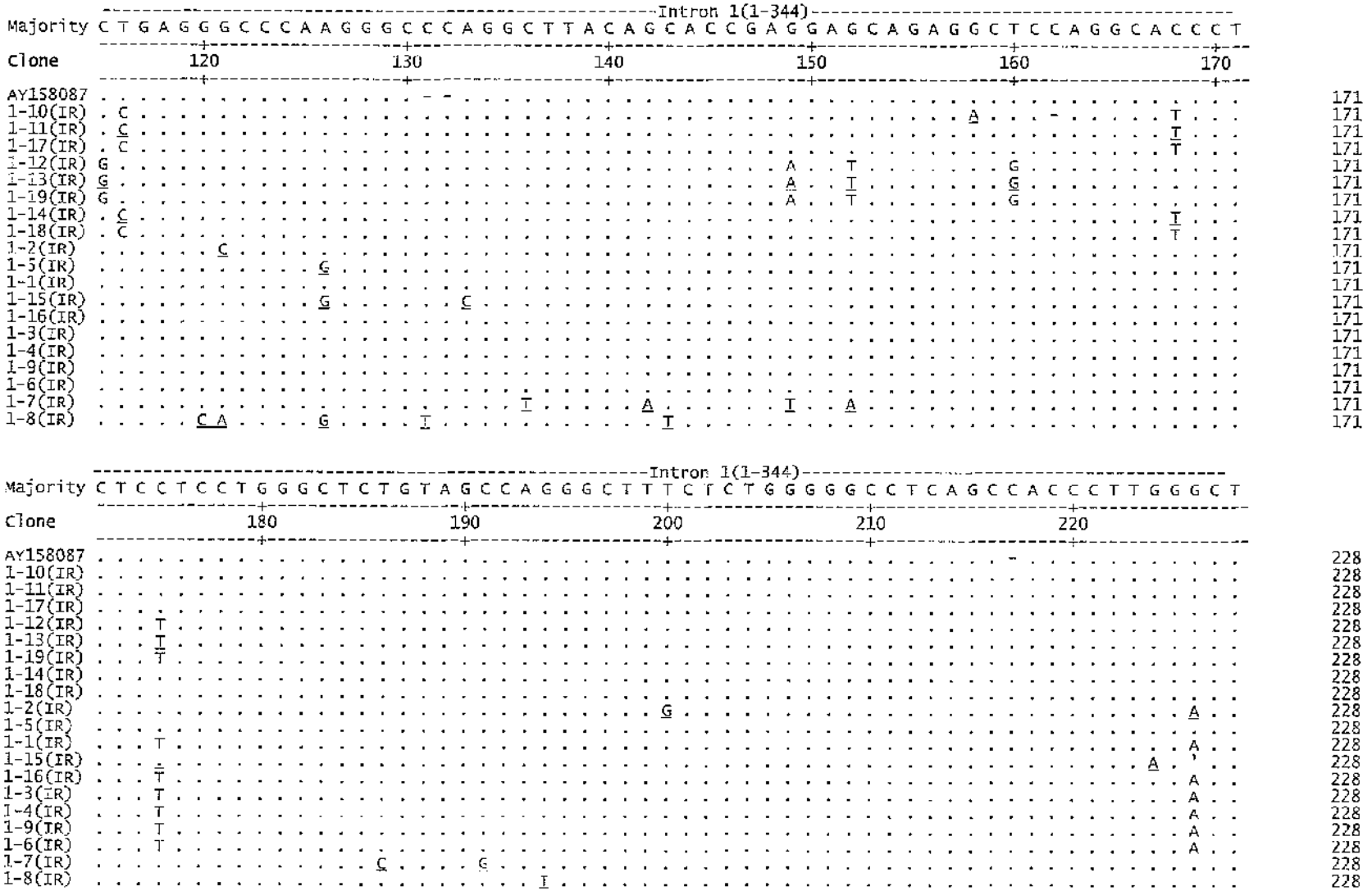


Figure 15 continued

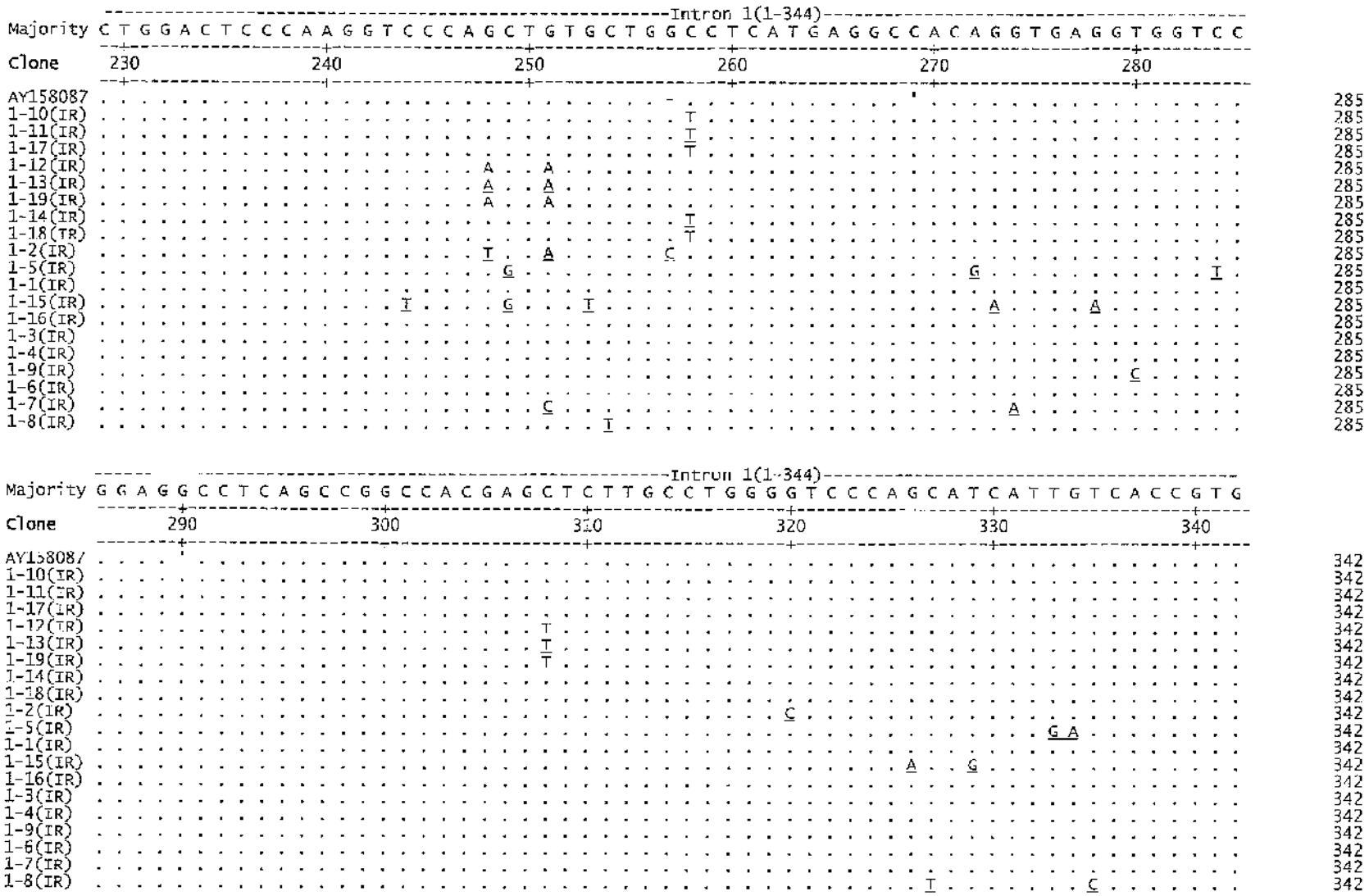


Figure 15 continued

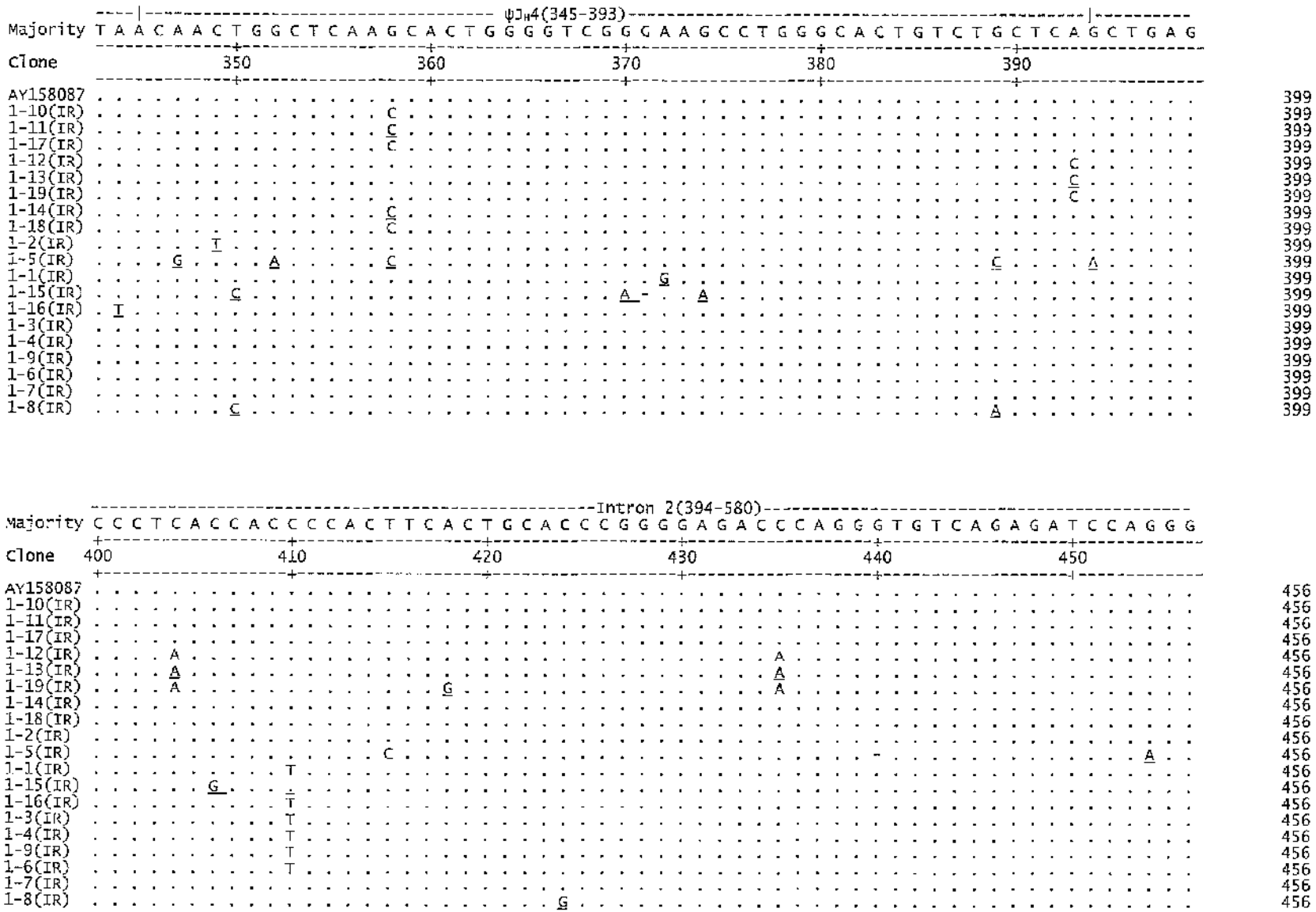


Figure 15 continued

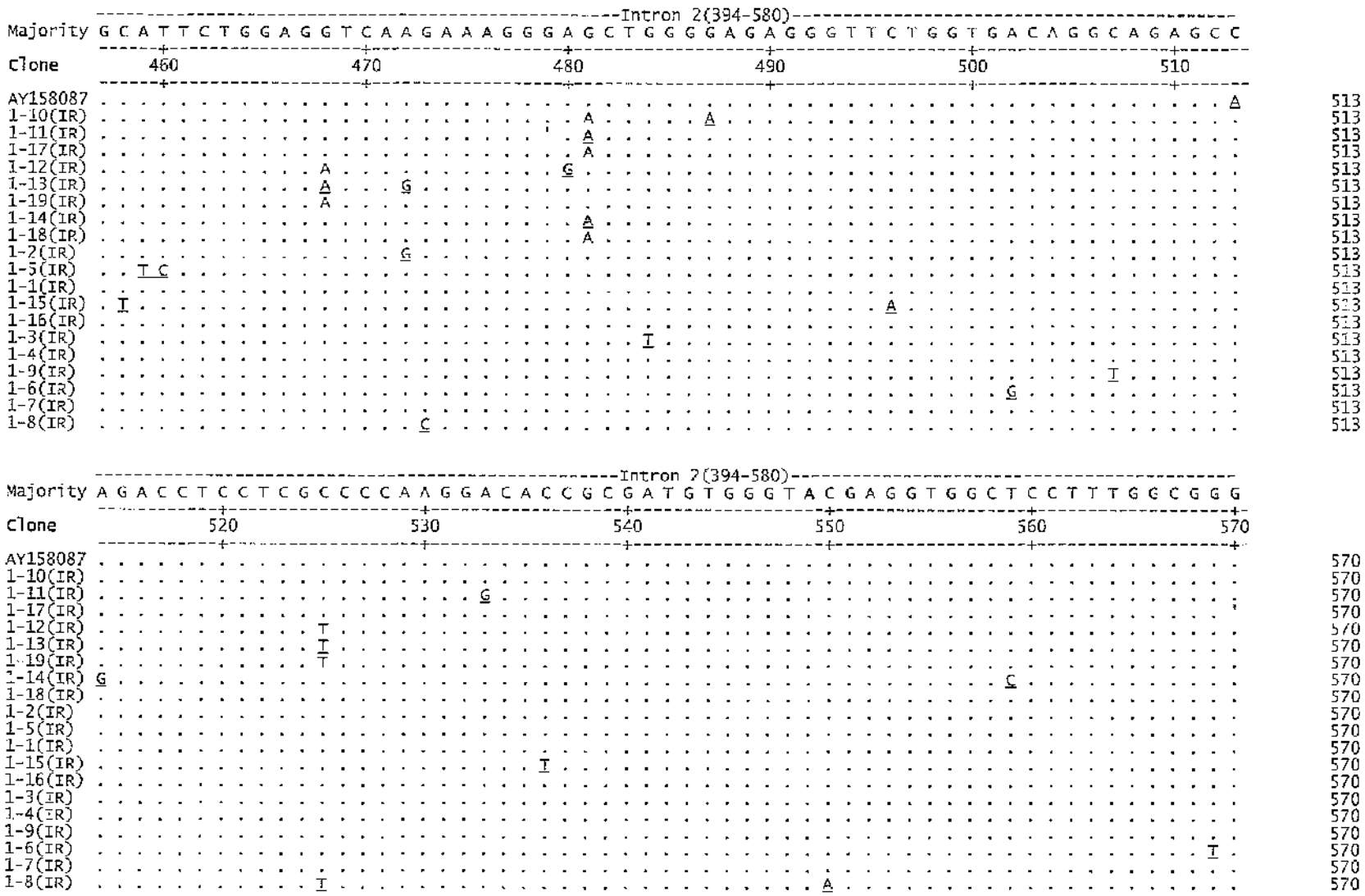


Figure 15 continued

Majority	T C T G G C T G C C	
Clone	580	
AY158087	. . . . .	580
1-10(IR)	. . . . A	580
1-11(IR)	. . . . .	580
1-17(IR)	. . . . .	580
1-12(IR)	. . . . .	580
1-13(IR)	. . . . .	580
1-19(IR)	. . . . .	580
1-14(IR)	. . . . .	580
1-18(IR)	. . . . .	580
1-2(IR)	. . . . .	580
1-5(IR)	. . . . .	580
1-1(IR)	. . . . .	580
1-15(IR)	. . . . .	580
1-16(IR)	. . . . .	580
1-3(IR)	. . . . .	580
1-4(IR)	. . . . .	580
1-9(IR)	. . . . .	580
1-6(IR)	. . . . .	580
1-7(IR)	. . . . .	580
1-8(IR)	. . . . .	580

Figure 15 continued

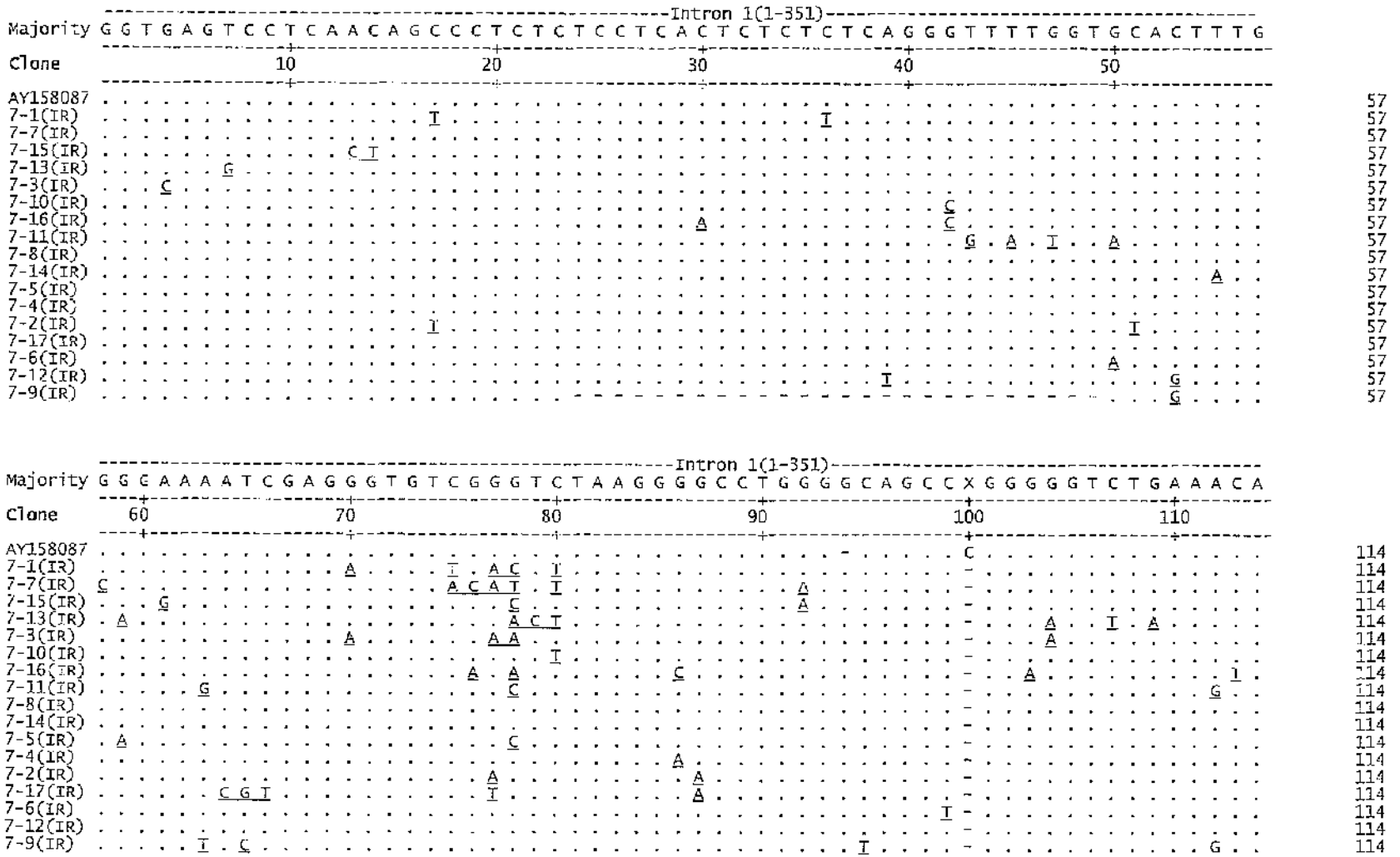


Figure 16: Nucleotide alignment of rearranged IgH locus sequences downstream of FR4 spanning regions of $J_H1/\psi J_H4/J_H2$ intron from a seven weeks old animal compared to most frequently rearranged bovine J_H locus on BTA 21 (AY158087)

The particulars of the clones are shown in the left-hand side of the Figure and clone 7-1(IR) means clone 1 from intronic region downstream of FR4 of seven weeks old animal. Nucleotide identity between the sequences is denoted by dot (“.”), differences indicated by letter code and a gap by dash (“-”).The nucleotides which are not underlined belong to germline sequence (part of an either haplotype) or a multiple sequence from common PCR reaction. The numbering on the right-hand side represents the last nucleotide position.

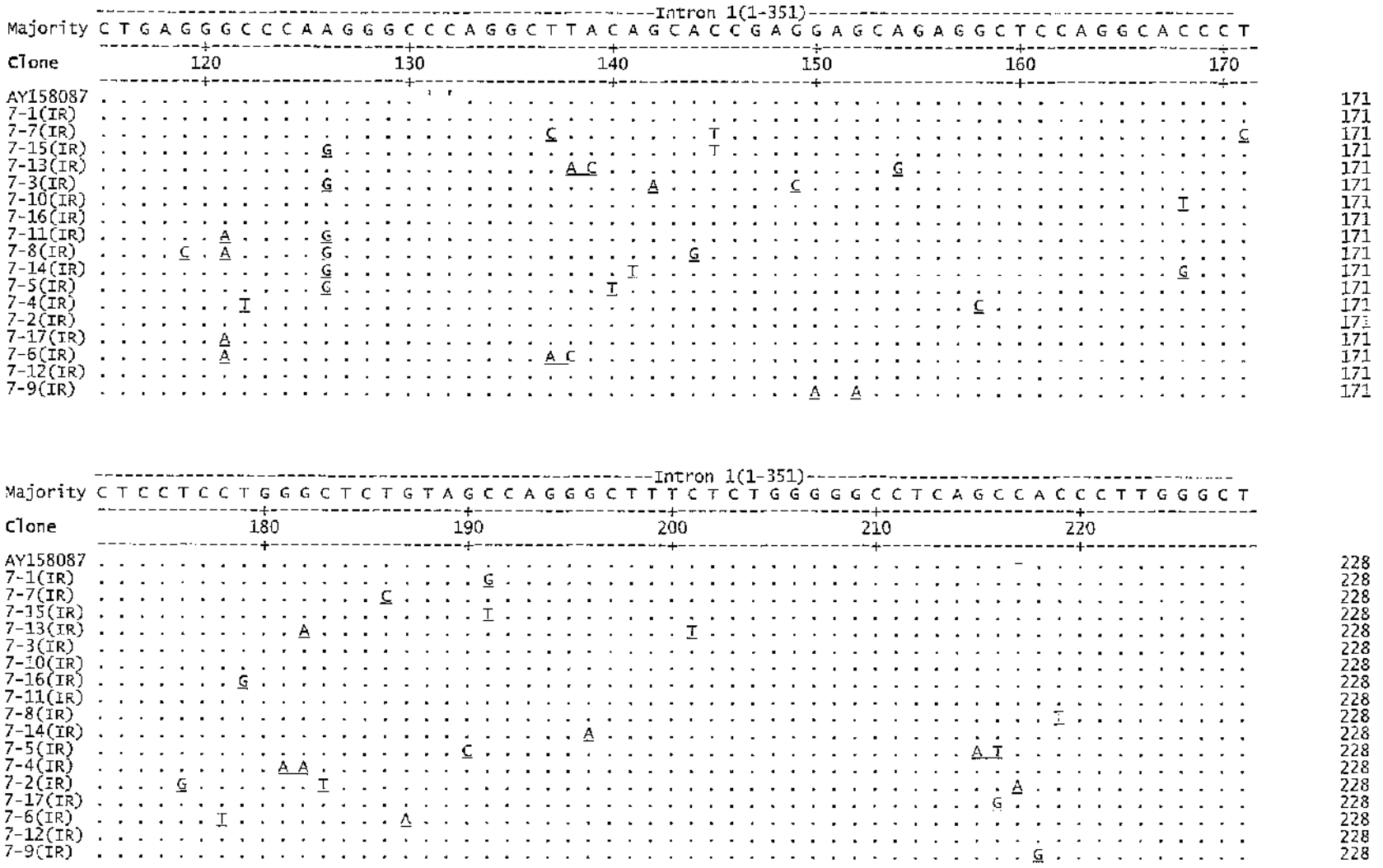


Figure 16 continued

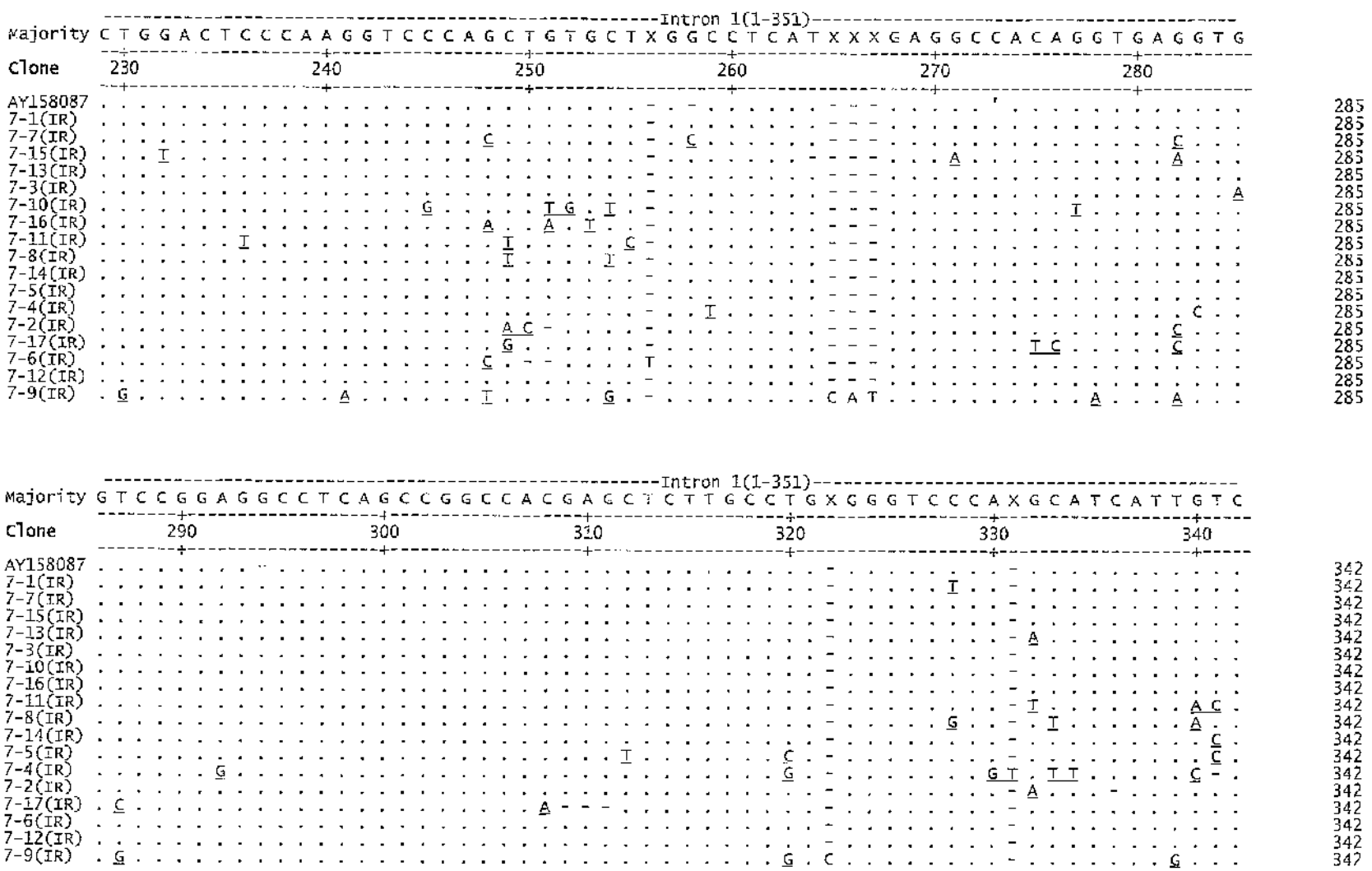


Figure 16 continued

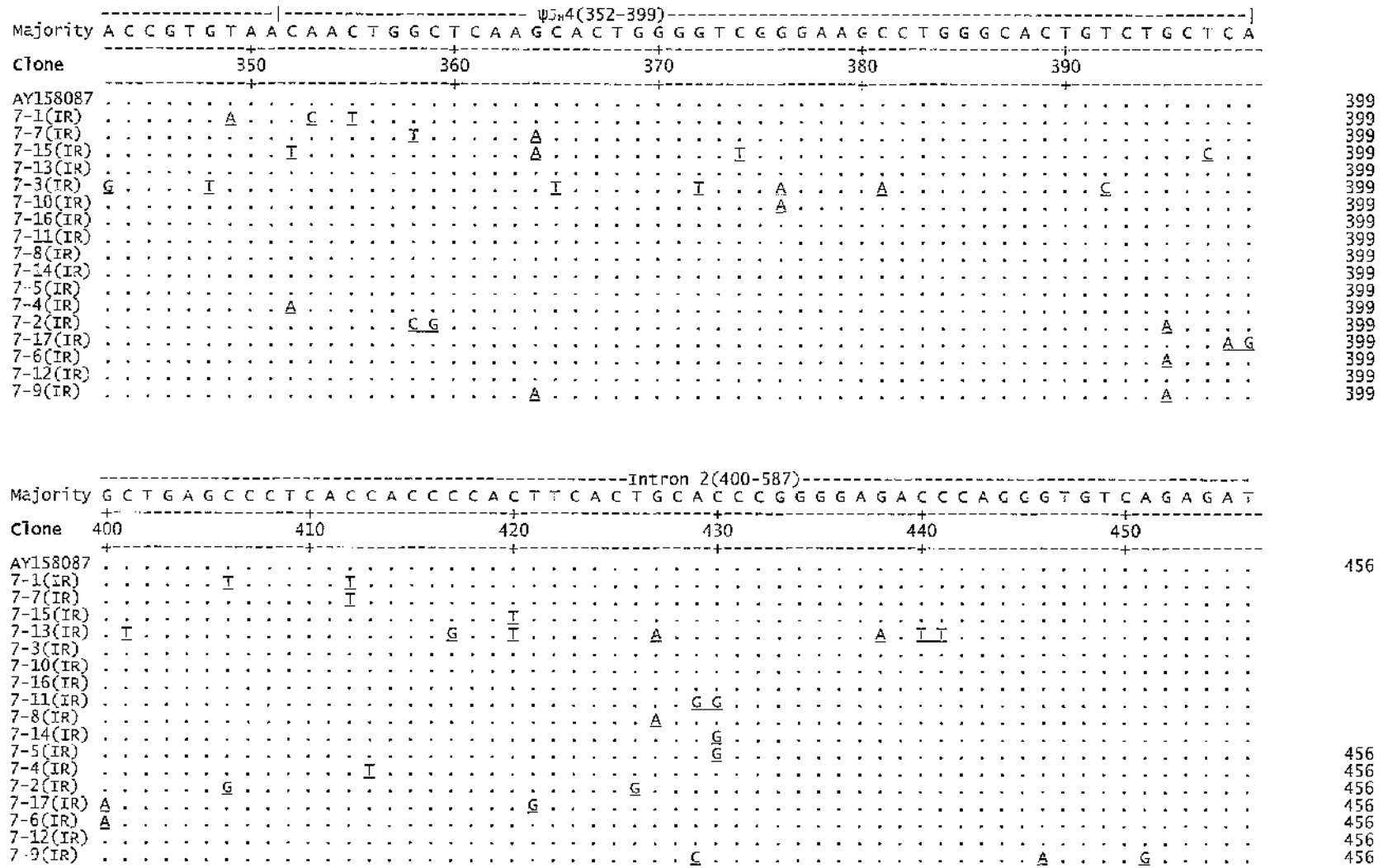


Figure 16 continued

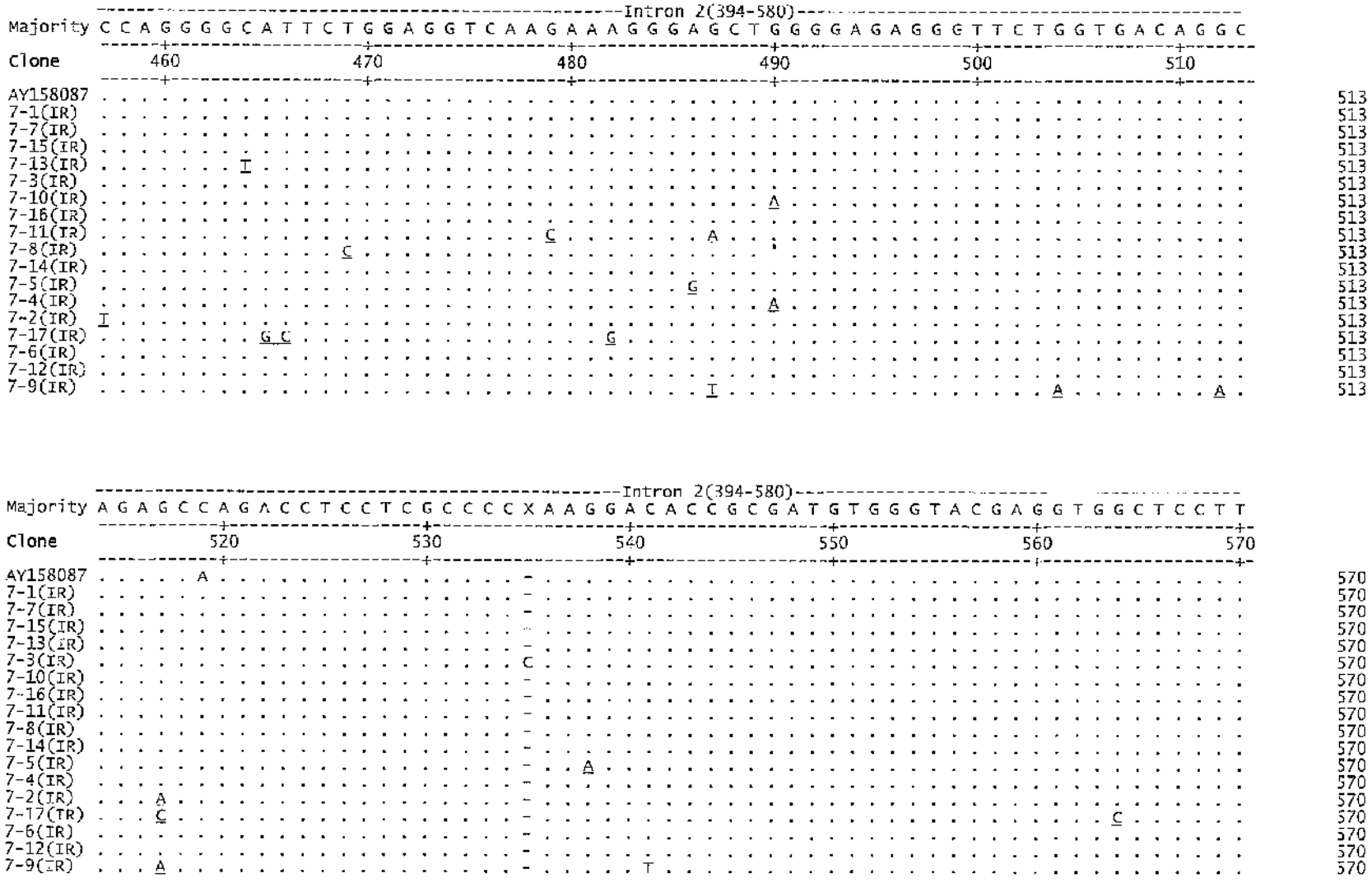


Figure 16 continued

Clone	580	
Majority	T G G C G G G T C T G G C T G C C	
AY158087	. . . . .	587
7-1(IR)	. . . . .	587
7-7(IR)	. . . . .	587
7-15(IR)	. . . . . T .	587
7-13(IR)	. . . . .	587
7-3(IR)	. . . . .	587
7-10(IR)	. . . . .	587
7-16(IR)	. . . . .	587
7-11(IR)	. . . . .	587
7-8(IR)	. . . . .	587
7-14(IR)	. . . . .	587
7-5(IR)	. . . . .	587
7-4(IR)	. . . . .	587
7-2(IR)	. . . . .	587
7-17(IR)	. . . . . C	587
7-6(IR)	. . . . .	587
7-12(IR)	. . . . .	587
7-9(IR)	. . . . .	587

Figure 16 continued

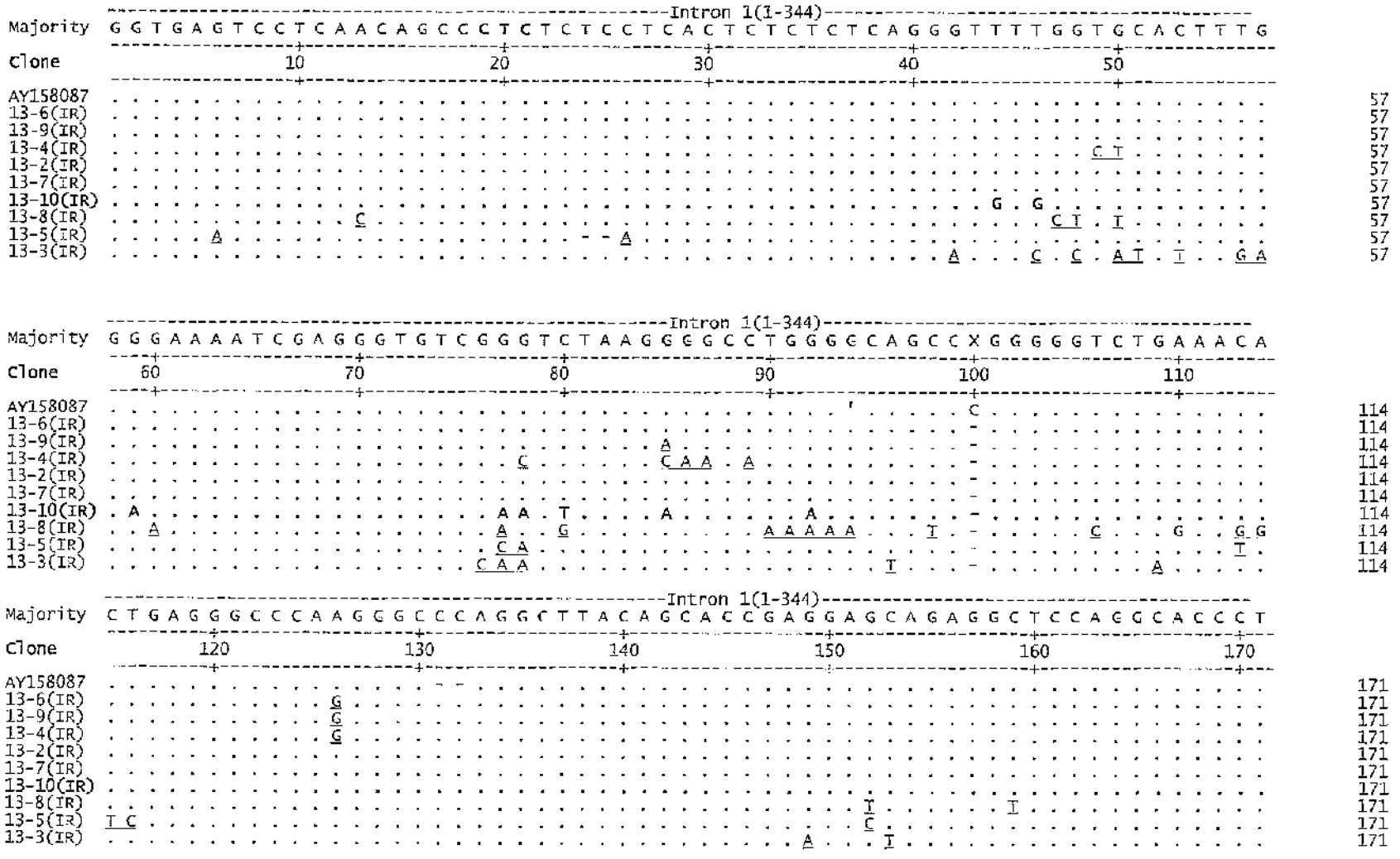


Figure 17: Nucleotides alignments of rearranged IgH locus downstream of FR4 spanning regions of $J_H1/\psi J_H4/J_H2$ intron from a 13 weeks old animal compared to most frequently rearranged bovine J_H locus on BTA 21 (AY158087)

The particulars of the clones are shown in the left-hand side of the Figure and clone 13-6(IR) means clone 6 from intronic region downstream of FR4 of 13 weeks old animal. Nucleotide identity between the sequences is denoted by dot ("."), differences indicated by letter code and a gap by dash ("-"). The sequence shown in bold has been identified is a haplotype and was also recovered from a day old animal. The nucleotides which are not underlined belong to germline sequence (part of an either haplotype) or a multiple sequence from common PCR reaction. The numbering on the right-hand side represents the last nucleotide position.

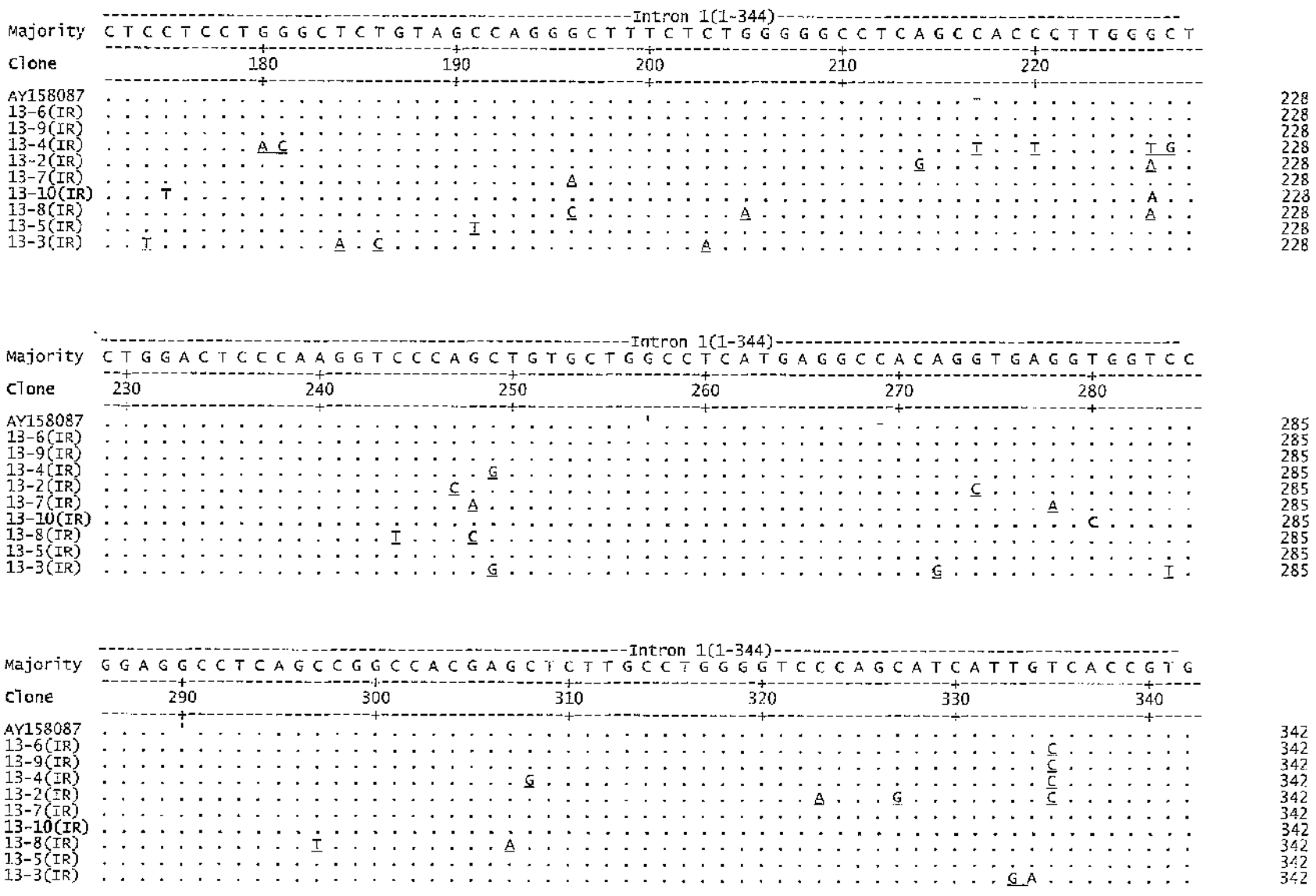


Figure 17 continued

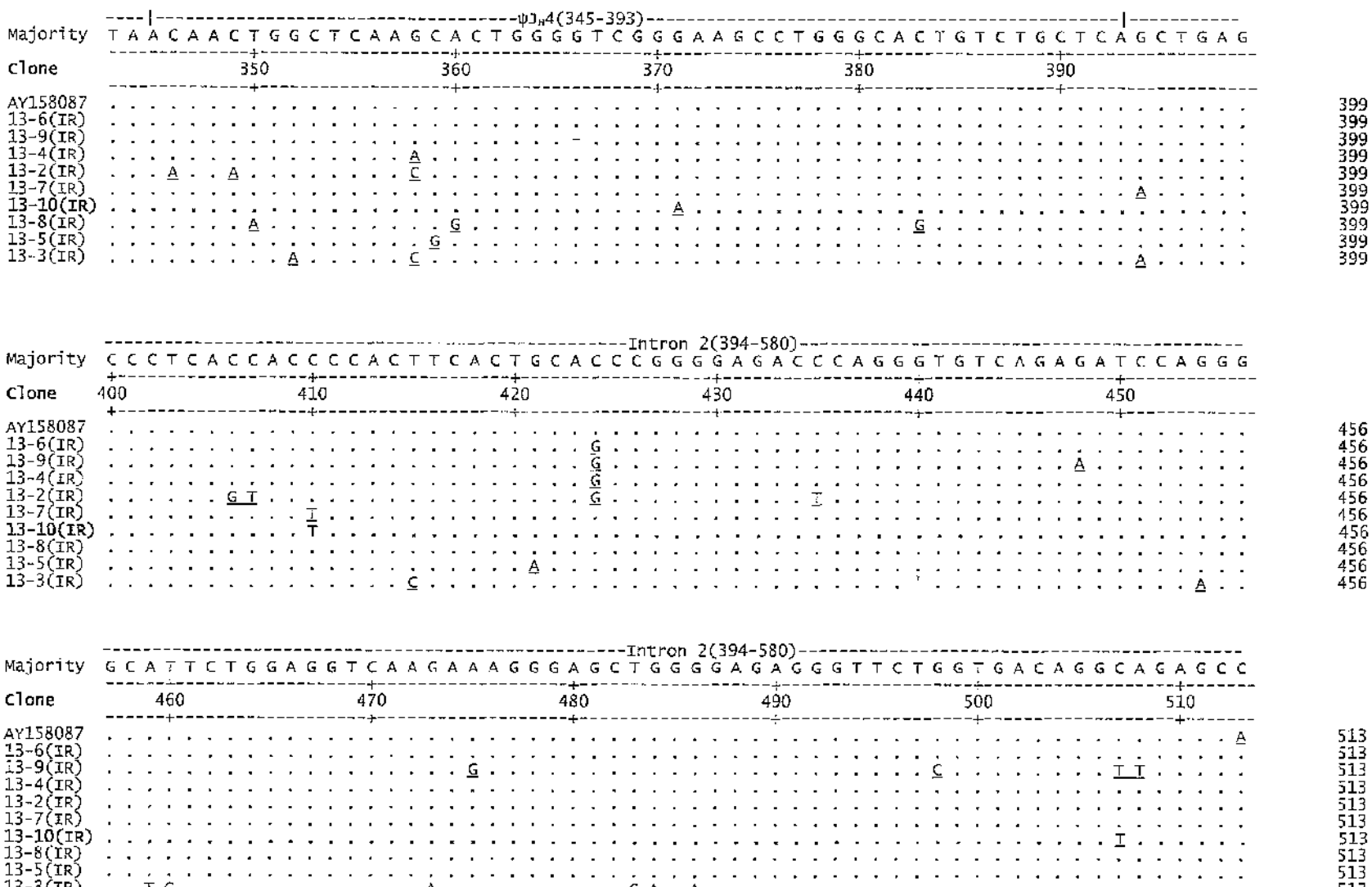


Figure 17 continued

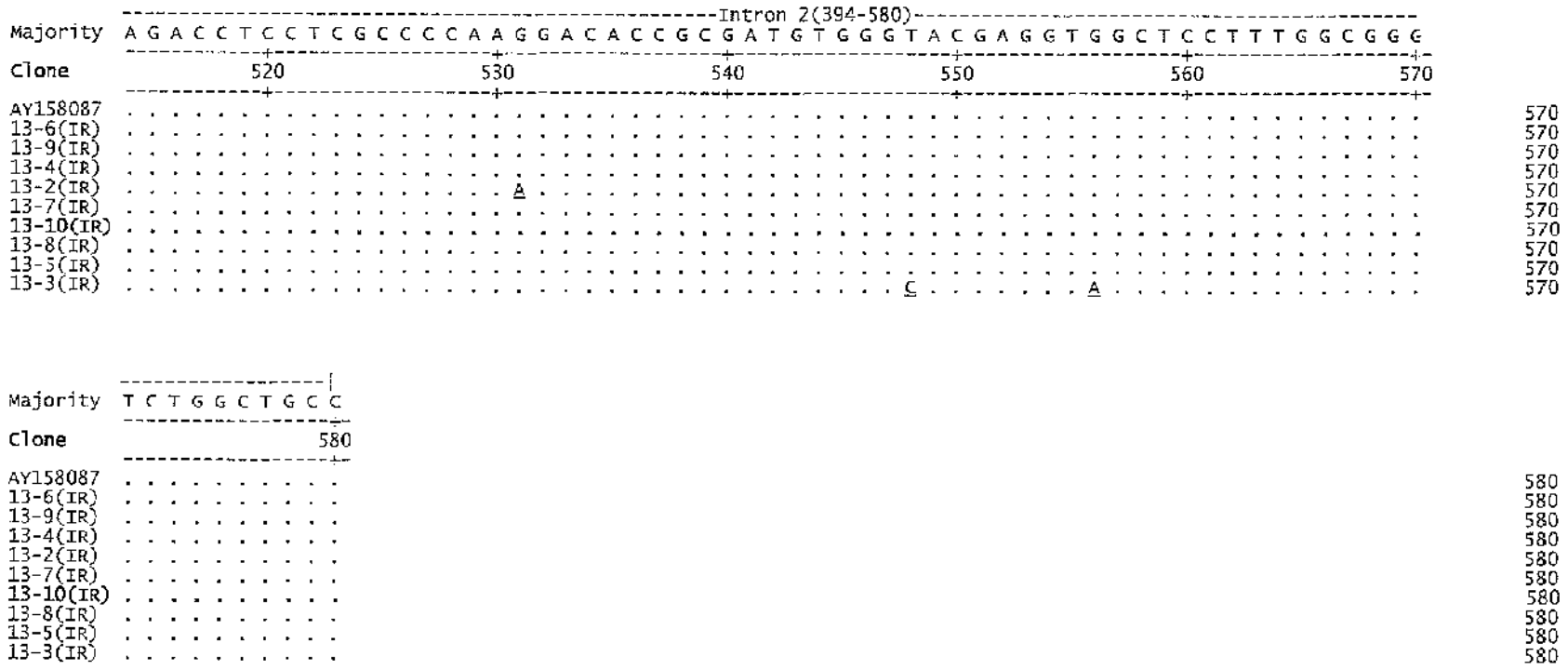


Figure 17 continued

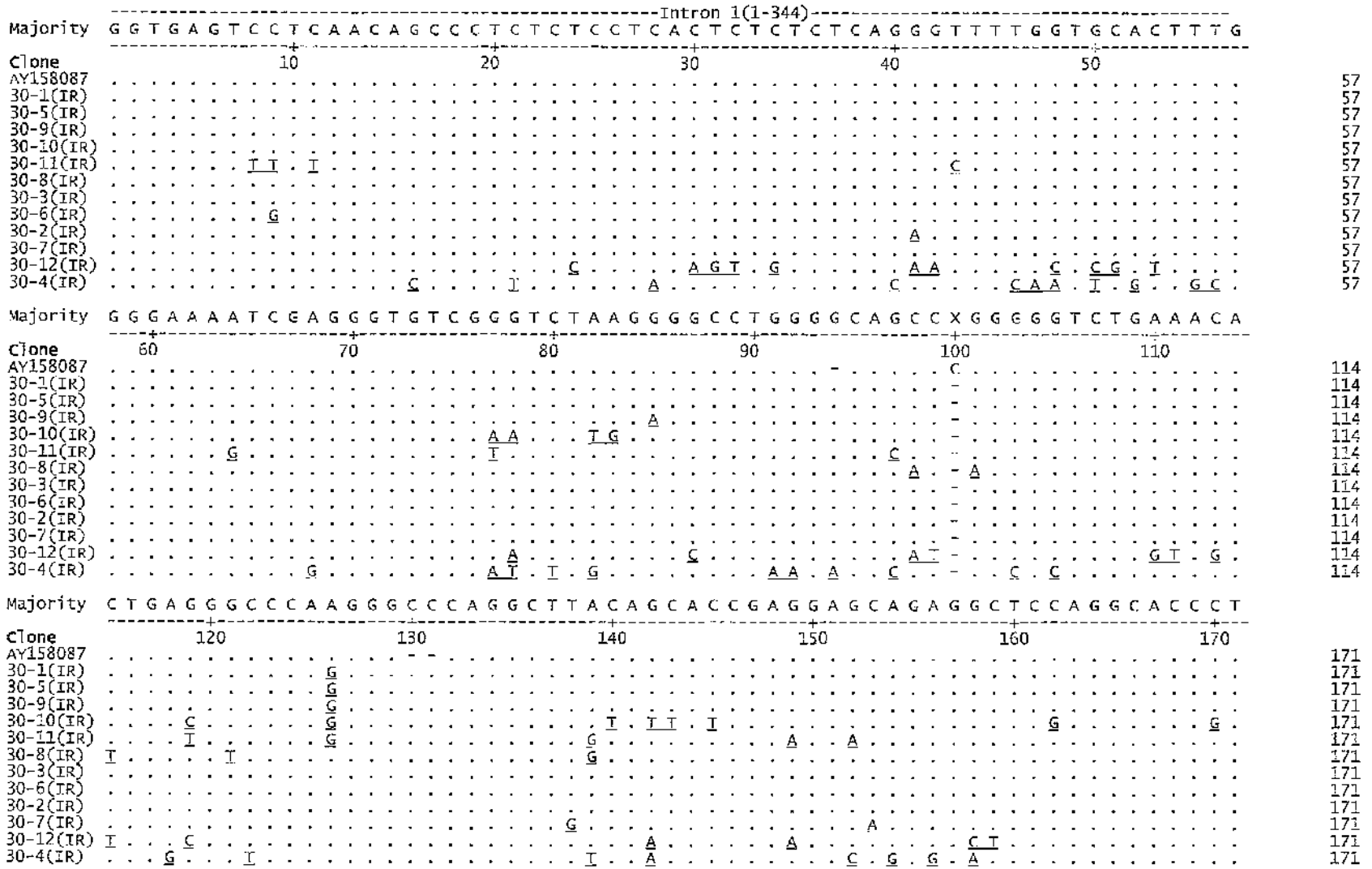


Figure 18: Nucleotides alignments of rearranged IgH locus downstream of FR4 spanning regions of $J_H1/\psi J_H4/J_H2$ intron from a 30 months old animal compared to most frequently rearranged bovine J_H locus on BTA 21 (AY158087)

The particulars of the clones are shown in the left-hand side of the Figure and clone 30-1(IR) means clone 1 from intronic region downstream of FR4 of adult animal less than 30 months old. Nucleotide identity between the sequences is denoted by dot ("."), differences indicated by letter code and a gap by dash ("-"). The nucleotides which are not underlined belong to germline sequence (part of an either haplotype) or a multiple sequence from common PCR reaction. The numbering on the right-hand side represents the last nucleotide position.

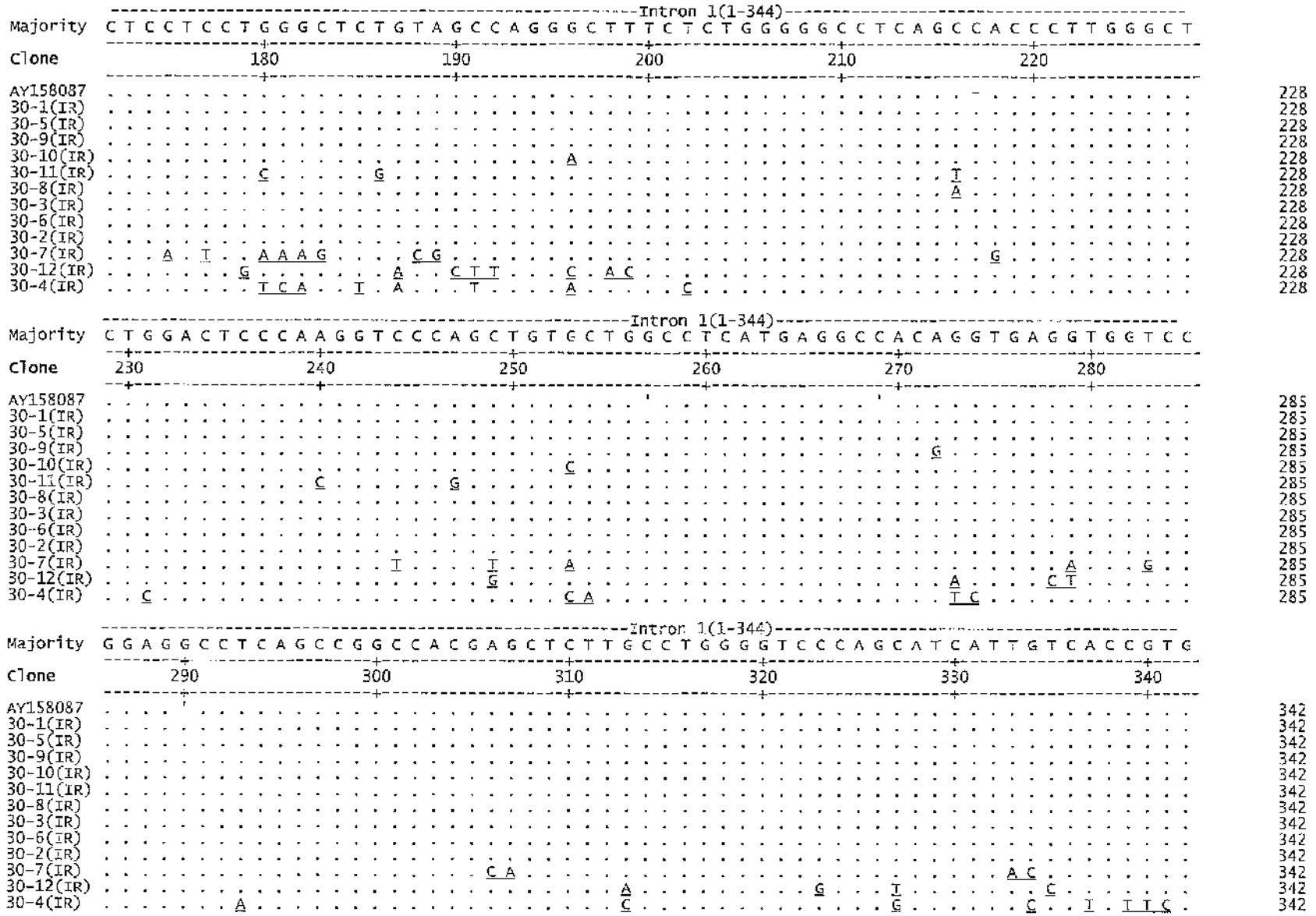


Figure 18 continued

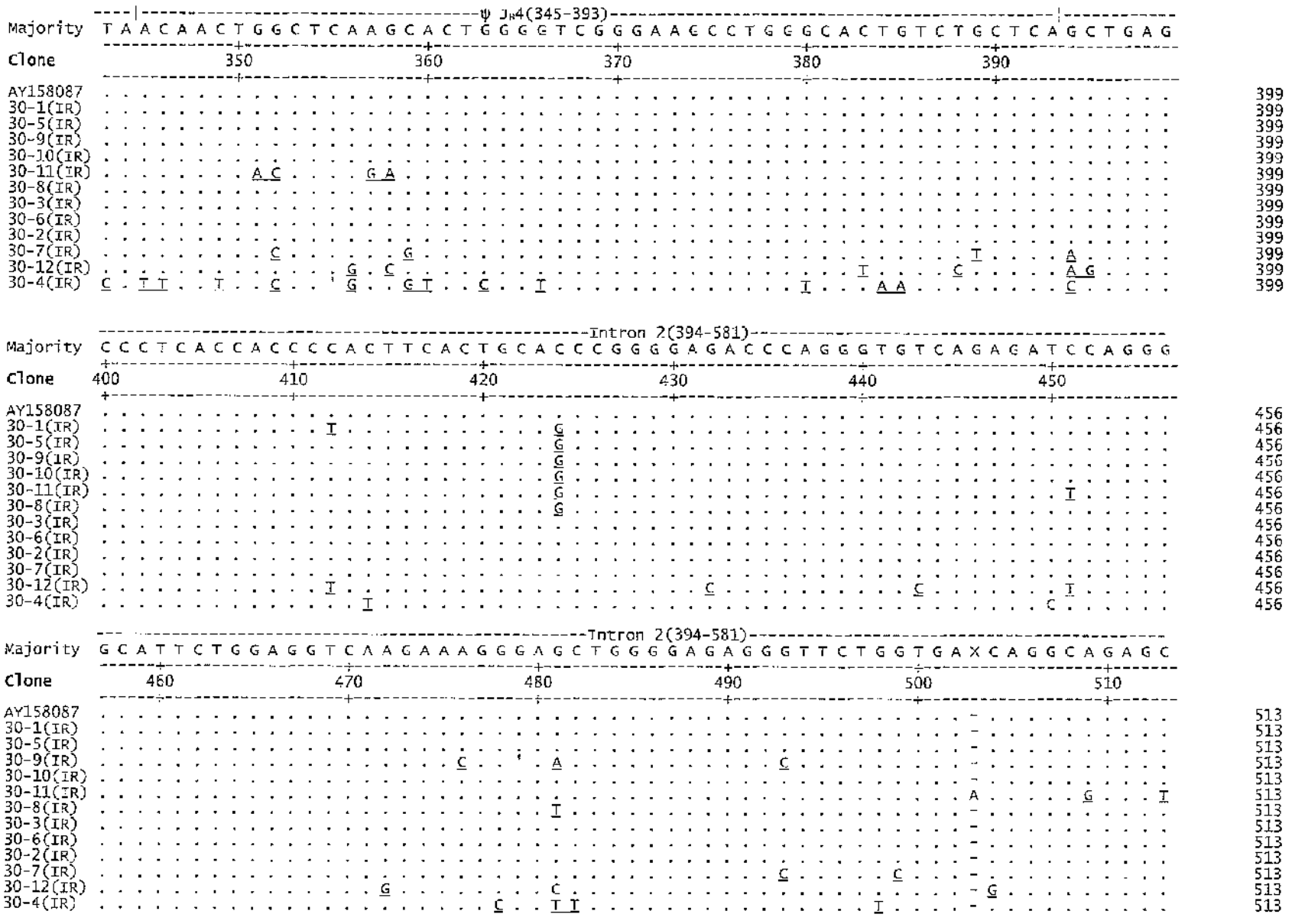


Figure 18 continued

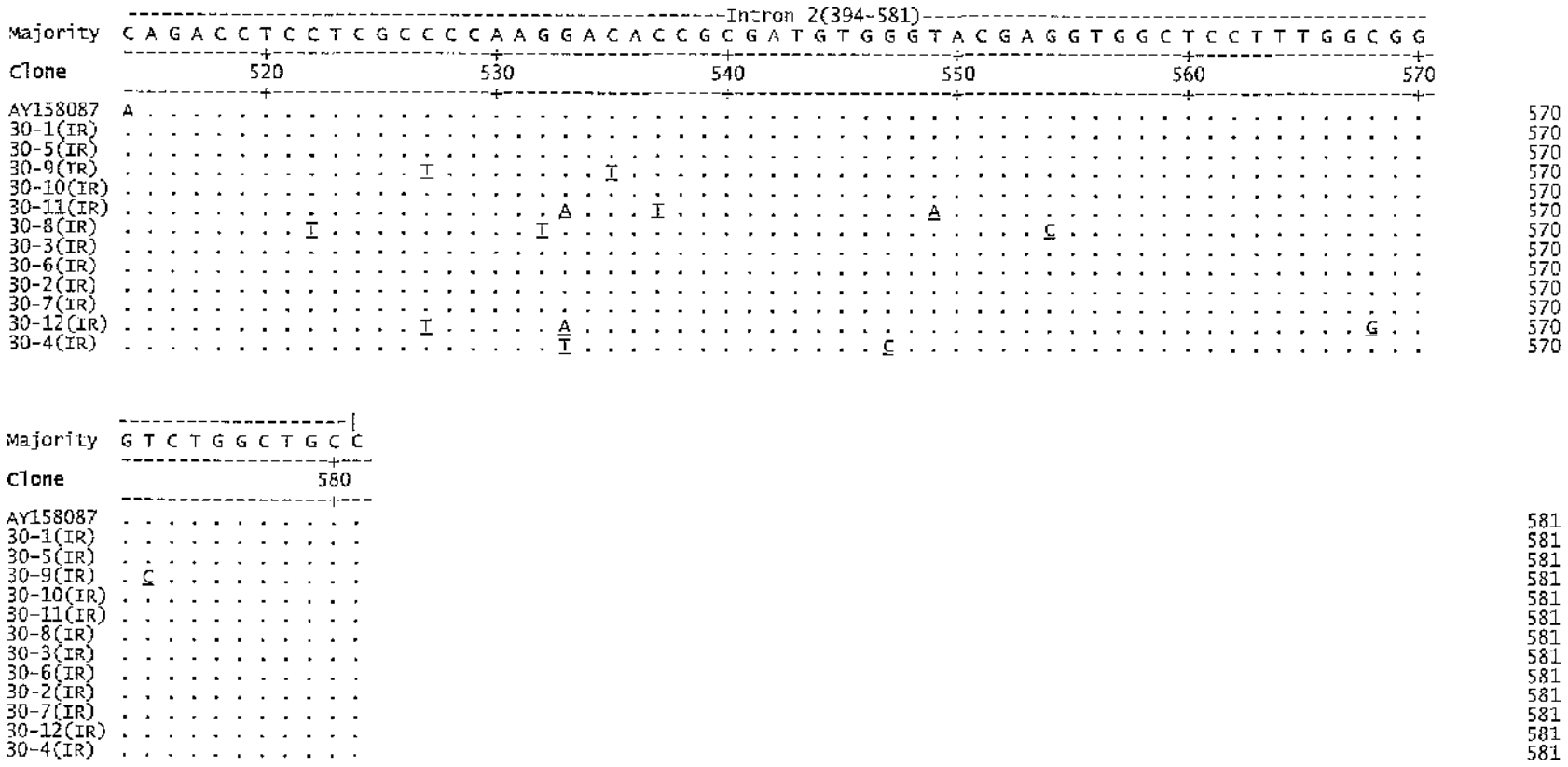


Figure 18 continued

Furthermore, the mutation rates are not identical but almost similar in specific intron in groups of animals from 7 weeks onwards. With the increase in distance away from CDR3 there was decline in rate of mutations as evidenced from Figure 19.

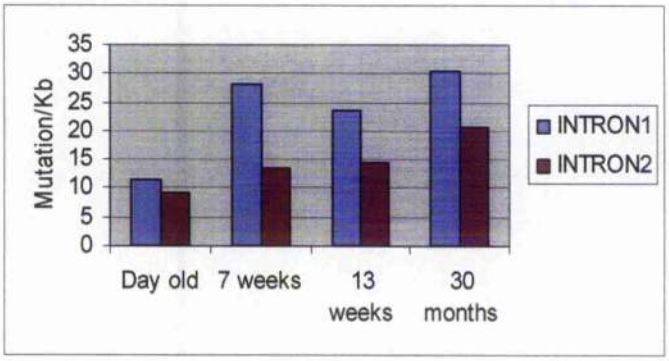


Figure 19: Mutations in J_H intron1 and intron 2

Introns were targeted for mutation with high rate. Within the rearranged IgH locus, intron 1 experienced higher mutations than intron 2.

The mutation frequency normalised per kilobase fell to 13.5 from 28 in 7 weeks old animal, 14.3 from 23.6 in 13 weeks old and 20.5 from 30.3 in adult animal. Overall, the pattern was broadly similar to that for the rearranged FR1-FR3 region. Mutations were unevenly scattered throughout the 5' region some stretches devoid of any mutations whereas in others mutations occurred in clusters.

3.4: DISCUSSION

The bovine humoral immune system is unable to generate a significant level of heavy chain diversity through combinatorial assortment of IgV_H segments. Data from Southern blotting, ssCP analysis and sequencing have consistently indicated that the size of the dominant gene family is small, it carries few distinct CDR sequences and the diversity apparent in these sequences is extremely modest (Sinclair *et al.*, 1997). The diversity generated through combinatorial means is thus minimal and further diversifying mechanisms operate to expand this primary repertoire.

Studies in the past have proposed roles for gene conversion (Parng *et al.*, 1996) or somatic hypermutation (Berens *et al.*, 1997) to drive this repertoire.

Both the processes are intrinsically different mechanistically. Gene conversion is templated in nature and results in non-reciprocal exchange of genetic information, whereas SHM is non-templated and introduce substitutions typically both in the coding and non-coding regions of the rearranged Ig gene in a random fashion. In cattle, it is very difficult if not impossible, to measure the distinct contribution of GC and/or SHM to the process of IgH diversification within the coding sequences of rearranged V region. The one reason for this is that not all the germline V_H segments are known and therefore donating sequences (if any) which may participate to diversify the rearranged V_H exon via GC cannot be traced to original V_H segments. The comparison between the two processes is easy and more rational within the rearranged non-coding region downstream of FR4 as this region has been shown to be targeted in SHM. This is due to the fact that only two donating J_H germline sequences exit for this region of the Ig gene and it would be easy to identify any donor sequences in the germline if GC was to happen. On the other hand, if SHM is driving the repertoire the diversifying substitutions will be noted both within the coding and non-coding region. Based on this existing information, the diversification process in cattle was studied at the IgH locus.

The IgH variable region and downstream $J_H1/\psi J_H4/J_H2$ intron sequences were isolated, sequenced and analysed for mutations from cattle of a range of ages. The sequences from animals of different ages were analysed as it was not apparent when exactly the process of diversification would commence. The recovered sequences came from range of different lymphoid tissues but the ones which were analysed and reported here all came from lymph nodes. Although lymph nodes are not the site for primary diversification, it is not known when and where the process of primary diversification begins in cattle. However, it is reasonable to expect diversification to become apparent in the peripheral lymph nodes shortly after its initiation so the animal is protected against infection.

All the studies to date have reported that the bovine immunoglobulin heavy chain repertoire utilises a single V_H family comprising no more than 20 segments that are at best only modest in their variability. This should provide a foundation against which, diversifying substitutions should be obvious. The present understanding about the bovine IgH locus is incomplete, and so

nucleotide variation could arise from the diversification process or the usage of uncharacterised germline genes or alleles. By applying the stringent criteria, as mentioned in this chapter, this study aimed to minimise this problem. The overall effect may be an underestimation of the contribution of somatic mutations to the process of diversification, but there are good reasons to believe that the substitutions that have been analysed genuinely arise through the diversification process. The uncertainties associated with this sort of work will clear away as the data become available from the cow genome project. The genome project will provide not only the better definition of number of total V_H segments available in the genome but also the understanding of their organisation in the IgH locus. It will also allow investigation of the reasons why only a single family is used in rearrangement when others appear to be present (Berens *et al.*, 1997).

Analysis of the data and features of diversification within the rearranged Ig V genes showed many of the characteristics of SHM as documented in other systems (Gonzalez-Fernandez *et al.*, 1994; Sinclair *et al.*, 1997; Wagner and Neuberger, 1996). The mutations were random in nature with some exception for certain regions which were highly mutated. Such hotspots were encountered not only in the hypermutating domains of CDR1 and CDR2 but also in the intronic sequences downstream of FR4, and the target region for the hypermutation extended well into the J_H-Cμ intron. Although this study did not analyse the region lying 5' upstream to the leader sequence for mutations, the earlier studies on mouse Ig genes have shown that this region is also targeted by SHM machinery. The distribution of 118 mutational events studied by Both *et al.*(Both *et al.*, 1990) clearly identified the target region for hypermutation which extended well into the 5' and 3'- flanking regions of the V_H genes. These authors reported that the 5' boundary of mutation was up to 550 bases upstream of cap site and that the 3' boundary extended 210 bases beyond J_H4 (last J segment). The statistical evaluation on data from this study (X^2_{test}) showed that CDR2 mutated at a much higher rate than CDR1. The independent studies in sheep have also revealed that CDR2 mutated more than CDR1. One study recorded an average of 4.63 mutations per sequence in CDR2 as opposed to only 2.25 in CDR1 in adult sheep (Gontier *et al.*, 2005). The variability observed at the nucleotide level is not always of the same magnitude at the amino acid level and this is due

to the degeneracy of codons. The substitutions at the nucleotide level in the gene, however, do result in amino acid replacement and this happens on most occasions when the base change is at 1st position of the codon. The observation on CDRs made at the cDNA level revealed that the variability index (a measure of occurrence of a different amino acid at a particular given position) in rearranged V_H genes was 17 in CDR1 as compared to 40 in CDR2 (Dufour *et al.*, 1996). The regional mutability in antibody V genes based on di- and trinucleotide sequence composition has predicted that V_H gene sequences preferentially direct somatic mutations to CDRs and away from FR regions in mice and humans (Shapiro *et al.*, 1999). Somatic hypermutation mechanism is not truly random and in a way it makes sense since it will be more useful to target mutations during antibody affinity maturation to CDRs rather than FRs. Some codons have low variability index whereas other accumulate a large number of mutations due to the intrinsic bias and there are suggestions that this is attributed to how the V region genes have evolved. During evolution Ig DNA sequence has been moulded to direct mutations to specific positions within codons in a manner that minimizes damage and maximizes the benefits of SHM (Shapiro *et al.*, 2002). This means that V gene sequences have evolved so that hotspots are strategically located. The amino acid serine is encoded by two types of triplet AGPy (Py is pyrimidine and is either C or T) and TCN (N is any nucleotide A, C, G or T). The sequence AGPy conforms to the hotspot consensus; TCN does not (Wagner *et al.*, 1995). Such a motif has been suggested by Wagner and Neuberger to provide a test of whether codon usage has evolved so as to favour local targeting of hypermutation (Wagner and Neuberger, 1996). These authors further reiterated that to ensure the hypothesis holds good i.e. that serine codons in CDRs are targeted more frequently for mutation than serine codons in frameworks would be to favour AGPy codons for CDR serines and TCN codons for FRs's serines. This was indeed the case when codon usage for serine was analysed separately for bovine IgH germline. All the serine codons but one in CDR1 and CDR2 were AGPy whereas in FR1, FR2 and FR3, TCN triplets coded for most serines. Apart from mice, biased serine codon usage is also in *Xenopus* V_H genes (Schwager *et al.*, 1989) and thus appears to be a general feature of antibody molecule design. It is not always the case that all the AGPy sequences are mutational hotspots and this consensus is not sufficient to define a mutational hotspot. For instance certain

serine codons in FRs use this triplet but are not targeted and are conserved. Certain other features of DNA are associated with high mutation rates peculiarities of local DNA sequence in the CDRs for example (Betz *et al.*, 1993; Steele *et al.*, 1992), such as palindromes or hairpin loops (Gonzalez-Fernandez *et al.*, 1994). Chang and Casali reported intrinsic sequence differences between CDR and FR codons favouring random replacements in the CDR over FR (Chang and Casali, 1994). Our analysis on the bovine germline genes suggested a lower ds:dn value of CDRs over FRs, meaning internal bias in CDRs towards replacement mutations. However, such analysis never takes into consideration the intricacies and targeting biases of the mutation mechanism. SHM, CSR and GC in mice, humans and chickens (Arakawa *et al.*, 2002; Muramatsu *et al.*, 2000; Okazaki *et al.*, 2002; Revy *et al.*, 2000) are dependent upon AID, and enzyme known to favour WRC motifs (W=A or T, R=A or G) as does the SHM machinery (Pham *et al.*, 2003; Sohail *et al.*, 2003).

In the CDRs, it appears that in a day old calf the sequences have the lowest level of diversification compared to other sampling times. This could be due to the usage of restricted V_H segments at that particular age, lack of exogenous antigen challenge, or the epitheliochorial nature of bovine placentation (Kruse, 1983). In human and mouse, there is a differential usage of germline genes in foetal, young and adult individuals. In the case of sheep, it was shown that no new rearrangement occurs after the initial colonization of the IPPs during the second half of foetal life (Reynaud *et al.*, 1991a) and IPPs act as a site to diversify the repertoire by non-templated SHM means under the pressure of exogenous antigens. But more recent studies in sheep suggest that rearrangements continue to occur several months after birth (Gontier *et al.*, 2005). It is not known whether the rearrangement mechanisms operate under a similar situation in cattle as in human and mouse or as in sheep. The site of B cell origin in cattle is highly unclear, but it is known that bone marrow is not the site of Ig gene rearrangement in bovines. Foetal liver or spleen have been suggested as sources of B cell proginators (Meyer *et al.*, 1997) from where B cells might migrate to ileal Peyer's patches (IPP), the most probable site of B-cell development in sheep and cattle (Yasuda *et al.*, 2006). The evidence of mutability in an animal of one day old in the $J_H1/\psi J_H4/J_H2$ intronic region is very convincing and it is highly unlikely that this could be attributed to anything

other than underlying diversification active at this stage. There is a marked elevation in diversity of CDRs at 7 weeks compared from day old group. The diversification of CDR1 then appears to level off whereas for CDR2 further increases could be detected to 13 weeks after which further diversification seemed to cease. It is also apparent that the rate of substitutions is much lower in FRs, presumably reflecting need to conserve Ig folding and structural stability.

There is a marked jump in diversification in the $J_H1/\psi J_H4/J_H2$ introns during the first 7 weeks after birth. In this region, many deletions and insertions were noted. Some substitutions were noted in this region which suggests rearrangement of second BTA21 haplotype. For instance, "G" at position 126 appears prominently in Figures 16, 17, and 18. A "G" at position 424 can also be noted. These were included in the mutation analysis, as the unique motifs were not observed throughout the sequence. The other reason for counting these substitutions was that these sequences vary from each other across many positions. Similar observations were made by Lebecque and Gearhart (Lebecque and Gearhart, 1990) in mice, reporting a high frequency of mutations for this region. Diversification of the region downstream of the J_H region has also been noted in rabbits (Sehgal *et al.*, 2002). The majority of mutations were found within the 300 bp of DNA flanking the 5' side of the VDJ gene and 850 bp flanking the 3' side. In this study of cattle, the frequency of mutations for this non-coding region was at least as high as seen for the variable coding region of IgH. The uniqueness of each mutation argues against the gene conversion as a mechanism for generating somatic diversity. Random point mutations apparently linked to the transcriptional state of the gene is more consisitent with SHM as the diversifying process.

Gene conversion process that uses a set of immunoglobulin variable (IgV) pseudogenes as template is the main mechanism of VDJ gene diversification in chicken (Reynaud *et al.*, 1989) and it has been suggested as one mechanism for diversification of V_λ genes of cattle (Lucier *et al.*, 1998; Parng *et al.*, 1996). The study on cattle reported by Parng and co-workes (Parng *et al.*, 1996) reported most mutations to be of single nucleotides and if these were introduced by gene conversion, this would imply very large number of donor sequences. This is incompatible with the properties of the bovine V_H system. Data from Southern

blotting of VH segments, does not suggest large numbers (Berens *et al.*, 1997; Sinclair *et al.*, 1997) even if conversion from one V_H family to another was possible. The observation of high frequency mutations in the 3' region flanking VDJ is also inconsistent with gene conversion as a diversifying mechanism. Studies with chicken VDJ genes have shown that 95% of the mutations were found in the coding region as opposed to 5% in the 3' flanking region. The existence of multiple V_H segments is unquestioned but there are no data to suggest that this is true of the J_H locus aside from the known duplication at BTA11. Gene conversion is therefore unlikely as a mechanism for introduction of mutations to the $J_H1/\psi J_H4/J_H2$ intron.

The pre-immune repertoire is defined as one which develops without bias towards specific environmental or microbial antigens. Primary and antigen driven diversification in new-born and older animals is difficult to separate. This could only be pursued using specific antibodies to isolate populations of naïve B cells from specific sites. Given the paucity of reagents to enable this and incomplete current understanding about the markers associated with specific B-cell developmental stages this is a formidable task, if not impossible. In a new born calf, the antibody repertoire present at birth can be regarded with reasonable confidence as representing pre-immune repertoire and it is unlikely that antigen-driven mutations would be observed the one day after birth. However, future studies would require the analysis from more samples from new-born calves to establish this claim. The impact of exogenous antigens on the immune repertoire is difficult to assess and to separate it further from the primary antibody repertoire is even more complex. Studies comparing germ-free and conventionally raised rodents suggest that normal gut flora plays a regulatory role in the development of the immune system (Thorbecke, 1959). The work on germ-free piglets by Butler and his collaborators (Butler *et al.*, 2000) might help to understand the role of external environmental, microbial and food antigens on immunoglobulin diversification. In this model system, the piglets are recovered by closed hysterectomy and maintained for 6 weeks in rigid tube isolators on sterile SPF-Lac food and the influence of factors transferred postnatally in colostrum/milk is then experimentally controlled. There are logistic and financial problems associated with adopting such a model and

performing similar experiments in cattle. Various *in vitro* models using human, mouse and chicken cell lines have provided useful information on these aspects.

Work reported later in this thesis aimed to resolve this issue. There are obvious advantages in using an appropriate cell line over purified B cells. Such a cell line would carry identical, rearranged IgH locus. Being a "closed" *in vitro* system, it would be possible to exclude external antigenic influences. More so, a cell line would provide cells in a uniform state of differentiation where (given appropriate stimuli) diversification might be activated simultaneously in the majority of cells. Would activation of such a cell line from cattle result in similar patterns of mutation as reported in this chapter using material from animals?

To summarise, many aspects of this analysis are consistent with somatic hypermutation as a mechanism for diversification of the bovine IgH repertoire. The substitutions occur with high frequency at DGYW/WRCH motifs (Rogozin and Diaz, 2004) where G:C is the mutable position. This mechanism has the potential to modify the V region and downstream $J_H1/\psi J_H4/J_H2$ intron, as observed. In order to generate and express functional Ig heavy chain, B cells with functional rearrangement are selected. It is not surprising that mutation rates appear to differ between coding and non-coding region. Since somatic hypermutation would be expected to drive affinity maturation in the presence of antigen, separation of primary diversification from Ag-driven processes is difficult. Again, appearance of mutations at Day 1 would argue that somatic hypermutation can drive the diversification of the primary repertoire. Perhaps repertoire diversification and Ag-driven mutations occur in parallel in older (e.g. week 7) animals. Finally, this study does not rule out the possibility of the possible involvement of gene conversion in antibody diversification in cattle but it seems unlikely that it plays a significant role to diversify the heavy chain repertoire.

Chapter: 4 Characterisation of *Bos taurus* Activation Induced Cytidine Deaminase (AID)

4.1: INTRODUCTION

B lymphocytes that have undergone rearrangement at the Ig loci are able to carry out three other types of modification reaction, depending on species, and circumstance. The first is CSR, when in the environment of the germinal centre; the IgM initially expressed by the B cell is changed to IgG, IgA or IgE. In this reaction, antigen specificity is retained. The second modification reaction that also takes place in the germinal centre is the introduction of nucleotide substitutions that attempt to elevate the affinity of the expressed antibody for antigen. Finally, in chickens and some other species, gene conversion can also take place to diversify the rearranged loci by non-reciprocal introduction of sequences from non-rearranging pseudogenes. In an effort to identify the molecular mechanisms responsible for the first modification reaction, cDNA from CH12F3-2, a murine B lymphoma line was analysed by subtractive cDNA hybridization to identify genes whose expression was increased in parallel with induction of CSR from class switch-stimulated and nonstimulated cells. The study was the first to show that AID was specifically expressed in class-switch recombination active cells (Muramatsu *et al.*, 1999). AID was classified as a member of the cytidine deaminase family, which includes cytidine deaminase, APOBEC-1, APOBEC-2, and phorbolin. The amino acid sequence of AID revealed closest similarity to apolipoprotein (apo) B mRNA editing catalytic polypeptide 1 (APOBEC-1), a well-characterized RNA editing enzyme. The apoB100 mRNA encodes the cholesterol carrier protein in low-density lipoprotein (Teng *et al.*, 1993). APOBEC-1 recognizes the structure of apoB100 mRNA through a cofactor, APOBEC-1 complementation factor (ACF) in which ACF binds to the mooring sequence in apo-B mRNA and docks APOBEC-1 allowing the deaminase to act upon its target cytidine (Mehta *et al.*, 2000). APOBEC-1 possesses a zinc binding catalytic region that carries out the specific deamination of the C6666 residue within apolipoprotein-B RNA. This generates a premature stop codon and enables the production of a shorter form of apolipoprotein B in the gastrointestinal tissues. AID possesses a putative catalytic domain that is near-identical to that of APOBEC-1.

4.1.1: Role of AID in modulating the immunoglobulin genes

The importance of AID to humoral immunity has been shown in a great many studies since its discovery. In transgenic mice that carry AID mutations, switch recombination is abolished as shown by accumulation of IgM in sera and faeces and severely reduced somatic hypermutation (Muramatsu *et al.*, 2000). Similar effects are seen in humans. Molecular studies have demonstrated that all patients with hyper IgM syndrome type 2 carry mutations in the AID gene (Revy *et al.*, 2000), and all these mutated AID cDNAs are defective in CSR, as determined by *in vitro* assays (Ta *et al.*, 2003).The enzyme is also crucial for the process of gene conversion in chickens and it has been shown that the disruption of the AID gene in the chicken B cell line DT40 completely blocked Ig gene conversion. This block could be complemented by reintroduction of AID cDNA (Arakawa *et al.*, 2002). Constitutive activation of the gene in the DT40 B cell line leads to uncontrolled gene conversion (Harris *et al.*, 2002). In addition, exogenous expression of this gene in fibroblasts (Okazaki *et al.*, 2002) and hybridomas (Martin and Scharff, 2002) is sufficient to induce isotype switching or mutation events on artificial constructs or endogenous rearranged Ig loci. The targeted deletion of nearby pseudogenes donors in the rearranged light chain locus had been shown to abolish Ig gene conversion in DT40 and activate SHM (Arakawa and Buerstedde, 2004). AID is therefore a pivotal player in the generation of antibody diversity being an essential mediator of somatic hypermutation, class switch recombination, and gene conversion.

4.1.2: Molecular mechanism for regulation of SHM, GC and CSR

All the three mechanisms of Ig modification have been intensively studied but the precise molecular steps remain unknown. Although the outcomes are distinct, the initiation step appears similar for all the three processes and is dependent upon DNA cleavage. As described earlier, SHM is a process that introduces point mutations in the rearranged Ig variable genes at high frequency and takes place in germinal centres. Characterization of SHM reveals several common features with CSR (Kinoshita and Honjo, 2001). Deletions and duplications are observed in addition to the introduction of point mutations in the V region, implicating DNA strand breaks in the mechanism of SHM (Bross *et al.*, 2000; Papavasiliou and Schatz, 2000). There is no doubt that CSR requires

double strand breakages (DSB). Study on an artificial inversion-type substrate of CSR has shown that two single-strand breakages (SSB) are triggering events in CSR (Chen *et al.*, 2001). A common initiating event for CSR and SHM was suggested from a study carried out in mice lacking MSH2, a mismatch recognition protein. In these mice, the distribution of mutations induced by SHM and the breaks generated by CSR revealed a common consensus AGCT motif. Petersen and colleagues (Petersen *et al.*, 2001) have shown that AID is involved in DNA cleavage rather than subsequent repair. This group detected the accumulation of a phosphorylated form of histone γ-H2AX into foci at the IgH locus indicating formation of DSB. Histone focus formation was dependent upon stimulation with LPS and IL-4 leading to CSR, and was absent in AID-deficient B cells stimulated with the same components. The work implicates AID in the formation of DSB at the IgH locus.

SHM has been explained by many authors as a result of SSB or DSB in the DNA followed by an error-prone DNA repair process (Diaz *et al.*, 2001; Faili *et al.*, 2002a). It has been proposed that SHM could arise from the formation of DSBs followed by homologous recombination (Papavasiliou and Schatz, 2000). However, ablation of the RAD51 paralogues XRCC2 and XRCC3- molecules essential for homologous recombination- resulted in reduced GC and enhanced SHM in DT40 cells (Sale *et al.*, 2001). The common genetic dependence upon AID of SHM and CSR in the human and mouse and GC in the chicken suggest that the different outcomes arise from different repair mechanisms after a common DNA cleavage event.

4.1.3: Models for modification of immunoglobulin genes by AID

AID has sequence similarity to the RNA-editing enzyme APOBEC-1, and exhibits deaminase activity on free deoxycytidine (dC) in solution, but how it modifies Ig genes remains uncertain. Two models currently exist, as shown in Figure 20. In the RNA editing model, it is proposed that AID's mRNA deaminase activity modifies an unidentified RNA messenger, leading to translation of a protein that is responsible for introducing mutations in DNA. This model originates from evolutionary conservation between AID and APOBEC-1. It is predicted that a complex of AID and its cofactors deaminates C to U at a specific position of the mRNA.

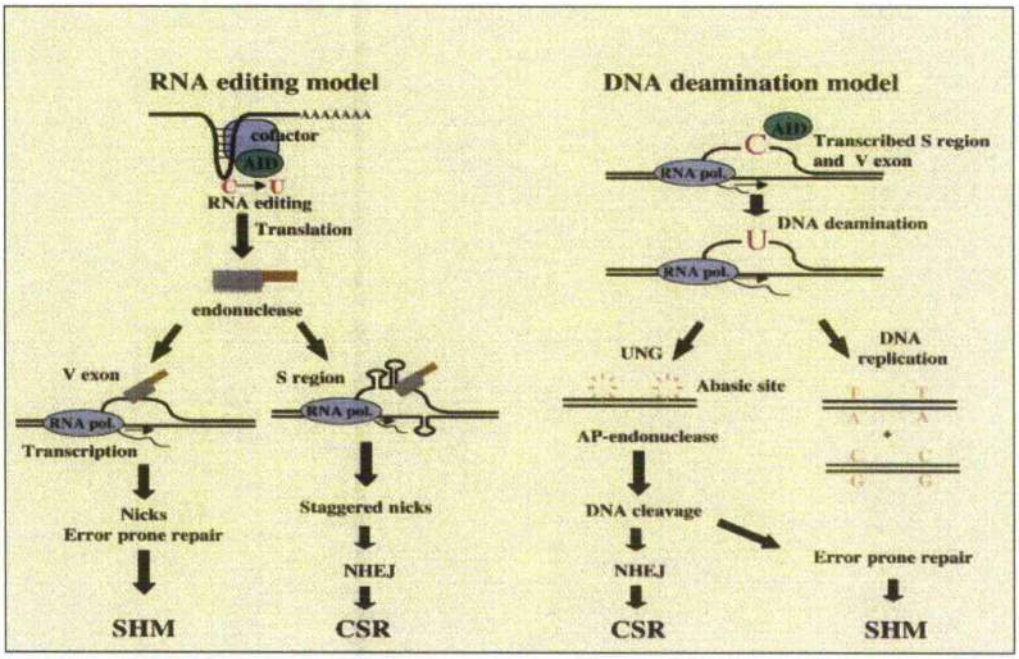


Figure 20: Models for modification of immunoglobulin genes

The figure has been taken from a review article (Honjo *et al.*, 2004).

The edited mRNA encodes a nicking DNA endonuclease that cleaves palindromic sequences at the Ig loci. Depending upon subsequent repair processes, SHM or CSR arises through error-prone repair or non-homologous end joining, respectively (Li *et al.*, 2003). Since exogenous expression of AID in fibroblasts can also initiate CSR and SHM (Okazaki *et al.*, 2002), this hypothesis postulates that AID's target mRNA and its cofactors are expressed in non-B cells. The experimental support for this idea comes from the observation that *de novo* protein synthesis after AID expression is required for CSR. Doi and his colleagues (Doi *et al.*, 2003) constructed AID fused with the hormone-binding domain of the estrogen receptor, which was introduced into AID-deficient spleen B cells. When such transfectants were treated with an estrogen analogue, 4-hydroxytamoxifen (OHT), CSR was induced within one hour. Cycloheximide or puromycin drastically suppressed OHT-induced CSR in AID- estrogen receptor expressing AID-/- B cells when added one hour before OHT but not after OHT, suggesting that *de novo* protein synthesis is required for an event downstream to AID expression in CSR.

Currently there is more experimental evidence to support the alternative DNA deamination model (Figure 20). The mutagenesis data in *E. coli* suggest that AID may deaminate dC on DNA (Petersen-Mahrt *et al.*, 2002), but its putative biochemical activities on either DNA or RNA remained a mystery. AID deaminates

ssDNA but not dsDNA *in vitro* or in *E. coli* (Bransteitter *et al.*, 2003; Chaudhuri *et al.*, 2003). The mutation produced *in vitro* shows preference for WRC motifs (W=A or T, R=A or G), a similar trend as seen in natural somatic mutation (Pham *et al.*, 2003). AID deaminates deoxy-cytidine (dC) into deoxy-uracil (dU) in DNA creating dU:dG mismatches (Di Noia and Neuberger, 2002). Such U:G pairs are assumed to be repaired by an evolutionary conserved base excision repair pathway, the members of which include uracil DNA glycosylase (UNG), apurinic/apyrimidinic (AP) endonuclease, and DNA polymerase beta. UNG activity removes uracil bases from the genome by hydrolysing them and generating an AP site. AP endonuclease then introduces SSB at the AP site. The subsequent repair process is error prone and hence has the capacity to introduce mutations. Apart from in *E. coli*, this hypothesis has also been tested in chicken DT40 cells for GC (Di Noia and Neuberger, 2002), mice for SHM and CSR (Rada *et al.*, 2002), and humans for SHM and CSR (Imai *et al.*, 2003). The idea that AID acts directly on DNA is further supported by chromatin precipitation assays showing that AID physically interacts with switch regions in cells stimulated to undergo CSR (Nambu *et al.*, 2003) and by the finding that replication protein A, a ssDNA binding protein, is a co-factor for AID (Chaudhuri *et al.*, 2004). There are still many issues that need to be solved, surrounding the role of AID in targeting the Ig loci and the molecular mechanisms involved. However, over the past 2-3 years it has become increasingly evident that cytidine deamination is the initiating event for SHM, CSR and GC (Petersen-Mahrt *et al.*, 2002).

4.1.4: AID in other organisms

AID has now been characterised from a variety of organisms including mouse, human, dog, chimpanzee, chicken, amphibian, and fish and despite showing considerable homology, it does show some differences between species. The putative cytidine deaminase motif in zebrafish, fugu and catfish encodes nine additional amino acids as compared to mammals. The identification and characterisation of *AID* from these organisms has helped to elucidate the mechanisms of immunoglobulin diversification. There are many differences between Ig loci of cattle and mice and the manner by which the bovine Ig repertoire is diversified remains unclear. Characterising *AID* from cattle would

provide an additional tool for studies of Ig diversification process and would also assist the *in vitro* analysis of repertoire development.

The aims of this aspect of the project were to clone and sequence bovine *AID*, to study its expression in a range of tissues, to determine its chromosomal localisation, and to assess its activity *in vitro*.

4.2: MATERIALS AND METHODS

4.2.1: Touch-down PCR

The touch down PCR can circumvent spurious priming during gene amplification. The following PCR cycling conditions were set up to amplify AID from bovine genomic DNA with AID F6 as forward and AID R5 as reverse primer. Conventional PCR failed to yield specific product. The oligonucleotides sequences are given in Table 7. After initial denaturation at 95°C for 5 min, the reaction used a high annealing temperature of 64°C for 1 min in the earliest cycles. For every subsequent 5 cycles, the annealing temperature was decreased by 2°C. The final 10 cycles were then performed at 54°C to make a total of 35 cycles. The cycling conditions were denaturation at 95°C for 1 min, annealing for 1 min and elongation at 72°C for 3 min followed by a final elongation step at 72°C for 10 min.

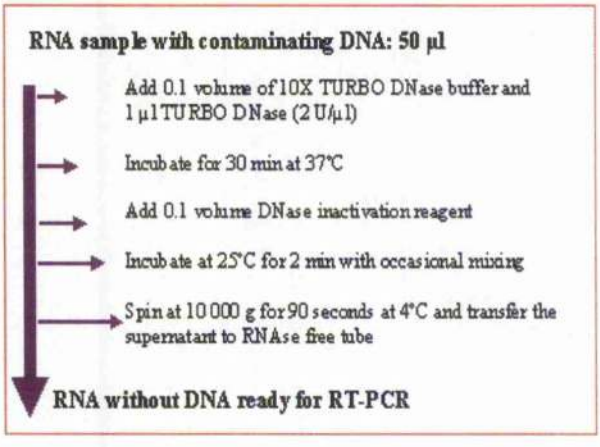
4.2.2: Extraction of total RNA

The tissues were stablised in RNAlater (Ambion) upon collection and later stored at -80°C. Total RNA was extracted with RNAeasy Midi Kit (Qiagen) according to the instructions of the manufacturer. No more than 250 mg of tissue was used in 10 ml of Buffer RLT as a starting material to extract the RNA by rotor-stator homogenisation. The contaminating DNA was removed from the RNA preparations using TURBO DNA-free (Ambion). This kit is designed to remove contaminating DNA from RNA preparations, and to subsequently remove the DNase and divalent cations from the sample. The flow chart below outlines the procedure. The quality and integrity of isolated RNA was assessed using an Agilent 2100 Bioanalyser (Agilent Technologies).

Table 7: Sequences of the primers used for the recovery and analysis of *B.taurus AID*

Primer	*Sequence (5' to 3')*
AID cDNA forward (Nco1)	TATCCATGGACAGCCTCTTGAAG
AID cDNA reverse (Spe1):	TTACTAGTCCCAAAGTACGAAATGC
AID F6	ACTTGTGCTACGTGGTGAAG
AID R5	ACCAAAAGGATGCGTGGAAGC
AID F1	GCCACGTGGAATTGCTCT
AID R4	ACGAAATGCGTCTCGTAAGTCATC
AID F2	GCGCGCCTCTACTTCTGT
AID R2	TCTACAAAGTATTCCAGCAG
AID Exon-1F	GGAGCCTAAAGGGCTGAGTT
AID Exon-5R	CTGGTGGGTCTGTGTTCCTT
AID cDNA F	ATGGACAGCCTCTTGAAG
AID cDNA R	TCAAAGTCCCAAAGTACG
AID (RH)-F	CCTCTTGAAGAAGCAGAGACAG
AID (RH)-R	GTGAGAAGGAGGTGGGACTG

The final concentrations were measured with a NanoDrop-1000 Spectrophotometer (NanoDrop Technologies).



4.2.3: First strand cDNA synthesis

In a small 0.2 ml thin-walled PCR tube, 1µg of total RNA plus RNase-free water was added to a total volume of 11 µl. To this, 100 µM of (5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGA-$(dT)_{24}$-3') was added and the contents of the tube were mixed and incubated at 70°C for 10 min. The tube was centrifuged briefly to collect the condensation after denaturation. The following were then added in the given order. First 4 µl of 5x first strand buffer, followed by 2 µl of 0.1M DTT and then 1 µl of 10 mM dNTP mix were added. All the

contents were mixed by flicking the tube and centrifuged to collect. The tube was incubated at 42°C, paused after 2 min to add 1µl of Superscript II (200U/µl; Invitrogen) was added before continuing the incubation for another 1 hour. The reaction was stopped by placing the tube in ice.

4.2.4: Real-time reverse transcription PCR

The fluorescence-based real-time reverse transcription PCR (RT-PCR) is widely used for the quantification of mRNA transcripts. The homology forward AID F1 and reverse AID R2 or forward AID F2 and reverse AID R2 primer pairs designed from human and mouse sequences were used in a real-time RT-PCR using a SYBR® Green PCR protocol from Applied Biosystem, UK. The primers were used on cDNA from BL-3 and spleen and melting curve and cT values were generated to analyse the results.

4.2.5: Conventional nested RT-PCR

Single-strand cDNA was synthesized from 1µg of total RNA with Superscript II from individual tissue specimens. This was used as a starting template to assess the presence of *AID* transcripts. Oligonucleotide primers were designed from the conserved sequences between human and mouse *AID* (Figure 21). The primers were used in a nested reaction to overcome the low representation of AID.

Figure 21: Proposed organisation of the bovine AID gene and location of primers for nested PCR

Exons (black boxes) are predicted from structure of mouse (NT_039353.6) and human (AB040430). Locations of primer annealing sites are shown with arrow. Outer (first round) primer set: 1- AID F6, 2- AID R5. Inner (second round) primer set: 3- AID F2, 4- AID R2.

The cDNA from lymph nodes, spleen, liver and muscle was analysed from day old and adult cattle. The touch down PCR with outer primer set was carried out exactly as before in section 4.3.1 except that timing for annealing and extension were reduced to 40 seconds each. This PCR product was then used as a template with the inner primer set. PCR amplification consisted of initial denaturation at 94°C for 5 min followed by 30 cycles of denaturation (94°C, 40

sec), annealing (56°C, 40 sec) and elongation (72°C, 40 sec). Finally the amplicons were allowed to extend for another 5 min at 72°C. As an internal control, a bovine glyceraldehydes-3-phosphate dehydrogenase primer pair (Endogen) was also used.

4.2.6: Cloning of full length of *AID* cDNA

Significant progress was made in characterising the bovine *AID* gene using primer designed from mouse and human sequences. However, it was predicted that an intron of several kilobases separated exons 1 and 2, complicating complete recovery of the gene. The deposition of sequence data at GenBank (Accession number NW_001001418) predicted to carry the bovine AID gene allowed design of primers against non-coding sequences upstream of exon 1 (AID Exon-1F) and downstream of exon 5 (AID Exon-5R). The sequence of primers is given in Table 7. Total cDNA was amplified with outer primer set (AID Exon-1F and AID Exon-5R) with pre-denaturation (94°C, 2 min), and 30 cycles of denaturation (94°C, 1 min), annealing (60°C, 1 min), extension (72°C, 1min), followed by final extension for 5 min. This PCR product from the outer pair of primers was then used as a template with inner primer set and the PCR amplification were performed in the same manner as described above except for denaturation (94°C, 40 sec), and annealing (56°C, 40 sec). In the second reaction, the forward primer (AID cDNA F) annealed to the entire predicted sequence of exon1 and part of exon 2 and reverse primer (AID cDNA R) bound to exon 5. The product was cloned into Topo 2.1 vector (Invitrogen) and three independent clones were sequenced. The sequences were aligned and annotated and deposited at GenBank. The *Bos taurus* AID cDNA sequence was translated *in silico* and compared with AID from other species. Phylogenetic relationships were established. Sequence editing, comparisons, alignments and phylogenetic construction were performed using Editseq and MegAlign from DNASTAR Lasergene software.

FISH was attempted in order to determine the chromosomal location of the bovine AID gene. This was carried out with help from Dr. Norma Morrison at the Cytogenetics Department, Yorkhill Hospital, Glasgow.

4.2.7: Synthesis of biotin-labelled DNA probe from *B. taurus* clone

The biotin-labelled probes were synthesised using a BioNick™ Labeling System from Invitrogen. This incorporates nick translation. The contents of the kit were mixed in 1.5 ml microcentrifuge tube. In total, 1µg of DNA template (partially cloned bovine AID) was used in a 50 µl total mix. The contents of the tubes were mixed well and centrifuged at 15000 g for 5 sec. The reaction was incubated at 16°C for 2 hours. Before stopping the reaction, the presence of nick translated products was checked by running 1/10 of the reaction on a 1% agarose gel and looking for a smear ranging from 200-500 nucleotides. The reaction was stopped with 5 µl of stop buffer. The unincorporated nucleotides were removed from the labelled DNA probe by using PCR purification kit from Qiagen. The probe was stored in TE buffer at -20°C for future use.

4.2.8: Synthesis of biotin-labelled DNA probe from Human-BAC clones

The human AID has been mapped to chromosome 12p13 using FISH (Muto *et al.*, 2000). In the bovine genome, BTA5 contains regions of synteny to human chromosomes 12 and 22 as revealed by Zoo-FISH experiments (Solinas-Toldo *et al.*, 1995). Radiation hybrid mapping of BTA5 (Ozawa *et al.*, 2000) supports this finding. Hence, BAC clones carrying a selection of human genes were obtained from BacPac Resource Centre at the Wellcome Trust Sanger Institure, UK (Table 8). These clones were used to make biotinylated probes as before.

Genes were chosen primarily on the basis of their known location on BTA5 (Liu *et al.*, 2003). NTS and DRIM lie far from AID in the human genome (HSA12q21 and HSA 12q23 versus HSA12p13, respectively) and are widely separated in the bovine genome (Liu *et al.*, 2003). However, it was argued that these genes would form clear markers for BTA5 in FISH and could preferentially be used in two-colour staining with bovine AID-derived probe. In contrast, A2M lies close to AID in the human genome (both HSA12p13). It was argued that if FISH revealed AID did not lie on BTA5, an A2M probe would control for hybridisation problems. All probes were derived from BAC clones carrying human sequence but given the success of ZOO-FISH procedures in comparative genomics (Solinas-Toldo *et al.*, 1995), the prospects of hybridisation to bovine chromosomal sequences were regarded as good.

4.2.9: Peripheral blood cultures for metaphase spreads

Bovine whole blood was collected in heparinised vacutainers from the local slaughterhouse. Using an aseptic technique, 1 ml of heparinised whole blood was added to 9.5 ml of blood culture medium. After mixing the contents of the culture tubes gently by inverting them a few times, these were incubated in a slanting position at 37°C for 72 hours before harvesting.

Table 8: The BAC clones that were used to develop probes along with their corresponding marker name and location

Marker	*Chromosomal Location*	*GenBank Accession (human gene)*	*Clone number*
NTS	HSA12q21 BTA5	AC016993	RP11-18J9
AID	HSA12p13	AC092184	RP11-438L7
A2M	HSA12p13 BTA5	AC007436	RP11-436I9
DRIM	HSA12q23 BTA5	AC063948	RP11-321F8

4.2.9.1: Cell harvesting from unsynchronised cultures

In total 200 µl of colcimid solution (10µg/ml) was added to 10 ml culture and mixed by gently shaking the tube and incubation for an additional 30 min. The culture tube was spun down at 300 g for 8 min and the supernatant was discarded into a container containing 10% chloros. The pellet was resuspended in 10 ml of hypotonic KCl solution (0.075M) pre-warmed to 37°C and thorough resuspension of the pellet was checked. The tube was incubated at 37°C for another 7 min. After the end of incubation the tube was spun at 300 g for 8 min and most of the supernatant was discarded as before. A little supernatant was left at the bottom of the tube to ensure cells were not lost and the pellet was thoroughly resuspended by gently tapping the side of the tube. Fresh fixative solution (3:1 methanol: glacial acetic acid) was added to tube initially drop-wise

with gentle agitation to a final volume of 10 ml. The tube was centrifuged at 300 g again for 8 min and the supernatant was pipetted off. The fixative was added again as before, the contents of the tube were mixed gently, spun down and the supernatant was discarded. The pellet was thoroughly re-suspended in 5 ml of fresh fixative and the tube was placed in refrigerator for a minimum of 30 min. After this it was spun down and supernatant removed. Finally, the pellet was re-suspended in 500-1000 µl of fresh, cold fixative ready to drop on to slides for metaphase spreads.

4.2.9.2: Cell harvesting from synchronised cultures

Bovine whole blood was collected in heparin from the local slaughterhouse. Using an aseptic technique 1 ml of heparinised whole blood was added to 9.5 ml of blood culture medium. After mixing the contents of the culture tubes gently by inverting them a few times, these were incubated in a slanting position at 37°C for 48 hours before harvesting. In total, 100 µl of thymidine (30 mg/ml) was added to each 10 ml culture and mixed by gently shaking the tube and incubation for an additional 8 hours. After this, 100 µl of 2-deoxycytidine was added to each 10 ml culture to achieve a final concentration of 10 µmol/ml. Cultures were incubated for a further 3 hours and 55 min at 37°C. After this, 100 µl of colcimid was added and cultures were incubated for an additional 20 min. Cultures were spun down at 300g for 8 min and the supernatant was discarded. Pellets were resuspended in 10 ml of hypotonoic KCl (0.0375M). The tubes were spun for 8 min and the supernatant was removed and fresh fixative was added to each culture as described above. Cells were then harvested in same manner as was done for unsynchronised cultures. Finally, the cell pellet was re-suspended in 500-1000 µl of fresh, cold fixative ready to drop on to slides for metaphase spreads.

4.2.9.3: Cytogenetic slide preparation

The slides were cleaned and stored at 4°C in distilled water before use. Three to four small drops of cell suspension were dropped from an optimal height onto the slide surface and allowed to spread out. The slides were then immediately placed on a hotplate and allowed to dry.

4.2.9.4: GTG-Banding

The slides were aged in a drying oven at 80°C for two hours and then immersed in trypsin solution for 15 seconds. Trypsin solution was prepared by adding 10 ml of fresh sterile saline solution to one vial of Difco trypsin (1:250) and then final volume was made up to 100 ml with sterile saline. Slides were rinsed in saline and stained with Leishman's stain for two min and washed with DW. They were then dried on a hotplate before examining for banding.

4.2.10: *In situ* hybridization and fluorescence detection

The probe was denatured immediately before hybridisation by incubating at 72°C for 5 min. The slides with metaphase spreads were denatured by incubating in pre-warmed formamide at 72°C for 2 min and thereafter quenched in ice-cold 70% ethanol for 2 min. The slides were then treated in a series of ethanol solutions, 70%, 90% and 100% for 2 min each, and dried.

4.2.10.1: Hybridization

The probe was applied to slides, glass coverslips were applied and air bubbles removed gently by pressing with forceps. The coverslips were sealed into position with Studio Gum and placed on a metal tray to be incubated overnight at 37°C in a water bath in the dark.

4.2.10.2: Post-hybridization wash and detection

The gum was taken off and the slides were placed in 2xSSC buffer for 2-3 min to remove the coverslip. The post-hybridization washes were carried at low or high stringency at 42°C or 45°C, respectively. Initially the slides were immersed in 50% formamide in 1XSSC solution at 42°C (45°C) for 5 min followed by washing in 2xSSC solution. Washes were repeated twice. Slides were removed and after applying 150 µl of Wash A solution; they were covered with a parafilm and incubated at 37°C waterbath for 10 min on a metal tray. All the steps from there on were carried out in reduced light. The parafilm covering was lifted and 150 µl FITC-avidin solution (Vector Labs, 2mg/ml) applied before returning the slides again to the 37°C waterbath for 15 min, again with parafilm covering. The parafilm was removed and slides were washed twice in 4xSSC for 5 min at room temperature. 150 µl of biotinylated anti-avidin solution was applied to the slides over a defined area and parafilm again used to seal the slides. They were

incubated for a further 15 min as before. The slides were washed again twice in 4xSSC and finally a third layer of 150 µl FITC-avidin solution was added before incubation at 37°C as before. After removing the parafilm, the slides were washed in 4xSSC buffer twice and excess fluid was removed. After applying 30 µl of DAPI counterstain mixture, a coverslip was applied and any air bubbles removed before inspecting for fluorescence signals.

4.2.11: Physical mapping of the bovine AID using radiation-hybrid (RH) panel

As an alternative to FISH, the chromosomal location of AID was sought by analysis of a bovine-hamster 5000 rad whole genome radiation cell hybrid panel ($WGRH_{5000}$) (Womack, 1997). The primers design was carried out in collaboration with Dr. Tom Goldammer and analysis was carried out in Dr. Goldammer laboratory at the Research Institute for the Biology of Farm Animals, Dummerstorf, Germany. Primers were designed based on the cDNA sequence for bovine AID (accession number: NM_001038682). The forward primer AID RH (F) spanned position 8-30 and AID RH (R) spanned position 111-130 of this sequence (Table 7) and hence both primers annealed to regions of exon 2. The primers were designed to maximise specificity for the bovine AID gene and to avoid amplification of the rodent sequence also present in the hamster hybrid cells. DNA from cells in the radiation hybrid panel was amplified through 35 cycles of 94°C for 30 s, 58°C for 30 s and 72°C for 45 s with an initial denaturation at 95°C for 15 min. A final extension step was performed at 72°C for 10 min. To establish synteny, statistical analysis was carried out with the data as per the method of Chevalet and Corpet (Chevalet and Corpet, 1986). Two-point linkage analysis for AID was done with all markers on the cattle $WGRH_{5000}$ gene map (Band *et al.*, 2000) using the software RHMAPPER 1.22 (Slonim *et al.*, 1997). This enabled assignment of AID to a bovine chromosome and an estimate to be made of its proximity to other markers.

4.2.12: Assessing the mutator phenotype of *AID*

The DNA-deamination model of AID action hypothesizes that the enzyme deaminates cytidine residues in the immunoglobulin loci, creating lesions that, through error prone repair, lead to diversification. To test if expression of

bovine AID in bacteria would create a mutator phenotype in a bacterial strain, bovine *AID* cDNA was cloned into the pBAD/*Myc*-His vector (Figure 22).

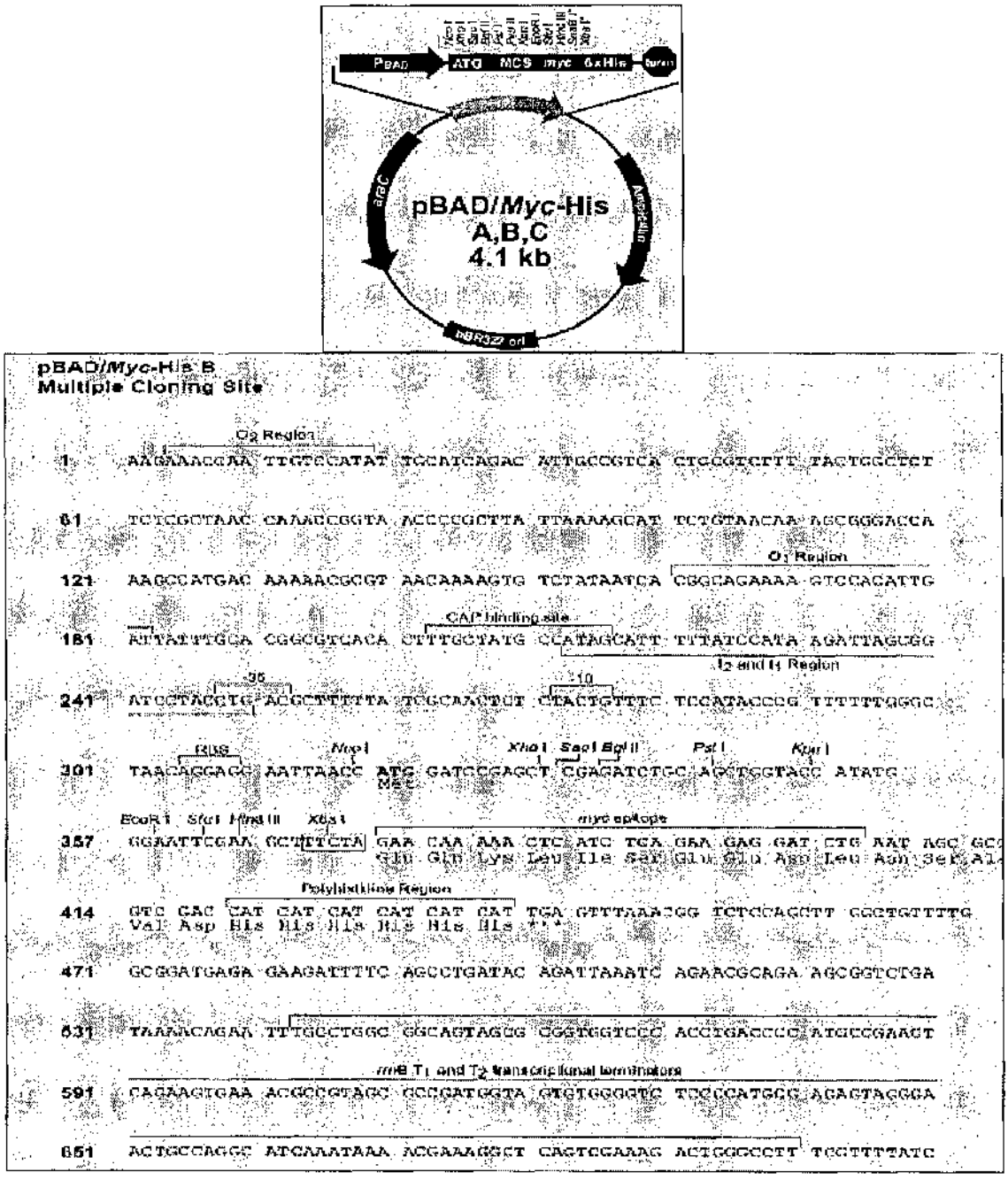


Figure 22: Summary of the features of pBAD/Myc-His vector

The multiple cloning site of version B are shown. Restriction sites are labelled to indicate cleavage sites. The figure has been taken from Invitrogen pBAD/Myc-His Manual.

This placed expression of the bovine gene under the control of the araBAD promoter. The recombinant plasmid was then transformed into different *E. coli* strains to assess its effects on the frequency of mutation in *Lac Z* and *RpoB*.

4.2.12.1: Cloning of bovine *AID* cDNA

Primers were designed to amplify AID cDNA from a TOPO-TA vector and introduce *Nco* I and *Spe* I sites at the 5' and 3' termini of the amplicon respectively. The sequences of the primers AID cDNA F (*Nco* I) and AID cDNA R (*Spe* I) are given in Table 7. For PCR, after initial denaturation at 94°C for 5 min, reactions were cycled 30 times through 94°C (30 sec), 58°C (30 sec) and 72°C (40 sec) with a final incubation at 72°C for 5 min. PCR products were resolved on a 1.2% agarose gel and extracted using a kit from Qiagen. Gel purified bovine AID cDNA was cloned into TOPO 2.1 as per manufacturer's protocol and checked for *Nco* I and *Spe* I restriction sites and sequenced. After digestion with these enzymes the bovine AID cDNA was ligated into the pBAD/*Myc*-His (version B) expression vector. The vector was prepared by digestion with *Nco* I and *Xba* I. In total, 100 ng of vector was mixed with a 3-fold molar excess of insert. For ligation, a Quick Ligation Kit from New England Biolab was used. Around 5 ng of ligation mix was transformed into competent TOP 10 *E. coli* and the bacteria were plated onto LA containing ampicillin. After overnight incubation at 37°C, clones were screened by colony PCR for the presence of insert. Briefly, around 10-12 colonies were picked, individually suspended in 7 µl of sterilized water and streaked to a master plate for future use. Individual bacterial suspensions were heated at 95°C for 5 min, chilled on ice and centrifuged to collect 4 µl of supernatant as a crude DNA template for PCR. The PCR was carried out with AID cDNA F (*Nco* I) and AID cDNA R (*Spe* I) as described earlier. Positive clones were further screened by restriction analysis and the reading frame again confirmed by sequencing.

4.2.12.2: Pilot expression experiments

To obtain optimal expression of bovine AID, different amounts of L-arabinose were used for induction over a 10,000 fold range (0.00002% to 0.2%). 2 ml aliquots of LB containing 50 µg/ml ampicillin were inoculated with single recombinant *E. coli* colonies and grown overnight at 37°C with shaking (225-250 rpm). The next day, five tubes each containing 10 ml of LB 50 µg/ml ampicillin were inoculated with 0.1 ml of overnight culture and were grown at 37°C with vigorous shaking to an OD_{600} of about 0.5. 1 ml samples were taken from each tube and centrifuged at 13000 rpm in a microcentrifuge for 30 seconds. The

supernatants were aspirated and cell pellets were frozen at -20°C awaiting analysis. L-arabinose was then added to the remaining cultures to final concentrations ranging from 0.00002% to 0.2% and the bacteria were grown with shaking for a further 4 hours. Finally, 1 ml samples from each tube were taken, centrifuged, and pelleted as before. Each pellet was resuspended in 100µl of SDS-PAGE sample buffer, boiled for 5 min and briefly centrifuged. 15 µl of each sample was loaded on an SDS-PAGE gel. After electrophoresis, the gel was stained with Coomassie Blue and scrutinized for a band of increasing intensity. Furthermore, samples were also analysed by Western blotting.

4.2.12.3: Western blotting

Cloning of the AID cDNA to the pBAD fused a c-myc tag to the carboxy terminal of the bovine protein. Hence, Western blots were probed with a mouse antibody to c-myc and detected with anti-mouse HRP reagents.

4.2.12.4: Effects of expression of AID on *lacZ*

Initial tests for a mutator phenotype arising from expression of AID sought changes in *lac Z*. *E. coli* G6, a strain with a functional *lac* operon with wild-type *lacZ* and expressing functional β-galactosidase was transformed with the empty pBAD plasmid or the recombinant construct carrying bovine AID. After selection on LA containing ampicillin, individual clones were cultured in LB medium and induced with 0.2% L-arabinose for 4 hours or overnight. The cultures were then plated on to LA plates containing isopropyl-β-D-thiogalactopyranoside (4 µl of a 20% solution), and X-Gal (40 µl of a 20mg/ml solution). Plates were grown at 37°C overnight and evidence sought for a switch of bacterial phenotype from blue to white.

4.2.12.5: Effects of expression of AID on *rpoB* in TOP10

The same plasmids were transformed into *E. coli* TOP10 (F' *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ*lac*×74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G) and after induction, the frequency of mutation to rifampicin resistance (Rif^R) was measured on rifampicin (100 µg/ml) enriched LA.

4.3: RESULTS

4.3.1: Partial recovery of *Bos Taurus AID* gene

Primers used to recover the partial sequence of *AID* were based on the high degree of conservation between the sequences of the *AID* gene in humans and mice. Furthermore, the cDNA sequences of chickens, dogs and zebra fish were sufficiently close in sequence to suggest that primer annealing to conserved exon sequences would successfully recover regions of the bovine gene (Figure 23).

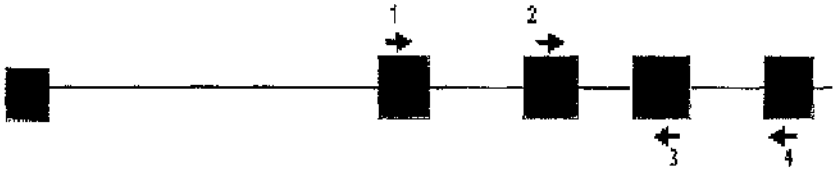


Figure 23: Strategy for recovery of bovine *AID* sequence by PCR

AID F6 (1) and AID R5 (3) were used for one PCR. AID F1 (2) and AID R4 (4) were used together. Primer sequences are shown in Table 7.

Based on their annealing position in the human gene, AID F6 and R5 were predicted to amplify a PCR product of around 2136 base pairs. AID F6 anneals to bases 6353 to 6373 in exon II of the human gene (GenBank accession number: AB040430) and AID R5 targets bases 8468 to 8488 in exon 4. Conventional PCR failed to yield a specific product and optimisation for different annealing temperatures and varying other parameters of the PCR including Mg++ concentration failed to improve the reaction. However, a touch down PCR protocol successfully yielded a specific product 3.4 kb (Figure 24). A conventional PCR strategy was used with primers AID F1 and AID R4 with annealing at 58°C for 40 sec. In this reaction, AID F1 anneals to exon III (bases 7814 to 7831 of the human gene) and AID R4 binds to exon V (8971 to 8994). A PCR product of 1495 bases was obtained as opposed to the predicted size of 1181 base pairs as expected from the human *AID* sequence.

4.3.2: Cloning of the PCR fragments

The 3.4kb amplicon retrieved with primers AIDF6/AIDR5 was cloned using a TOPO- XL PCR cloning kit and AIDF1/AIDR4 PCR product was cloned using TOPO-TA.

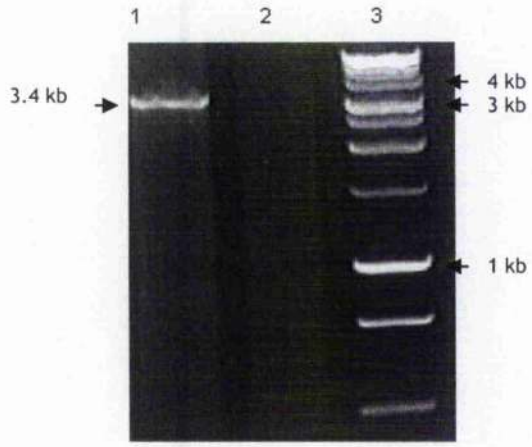


Figure 24: Amplification of part of bovine *AID* with primers AID F6 and R5

Lane 1 shows a product of 3412 base pairs amplified from bovine genomic DNA and resolved on 1% agarose in TAE buffer. Lane 2 is a negative control and lane 3 contains a standard 1 kb DNA ladder (Promega).

Clones were recovered from selective medium after overnight incubation and screened for the appropriate insert by digestion of plasmids with *EcoR* I. The clones were sequenced at MBSU, University of Glasgow. In order to get the complete sequence of the 3.4 kb AID insert, a primer walking strategy was employed. Based on initial data obtained with M13 forward and reverse primers, a custom oligonucleotide was designed, and then used to prime a second sequencing reaction. The data obtained from the second reaction overlapped with the initial data and extended the sequence further downstream. By repeated cycles of oligonucleotide synthesis and DNA sequencing, the cloned insert was sequenced completely in one direction. Since the products of the two reactions overlapped through part of exon III and IV and the intervening intron, in all 4253 bases of the bovine AID gene were sequenced by this approach (Appendix II). The sequence data confirmed the high degree of conservation of protein coding sequence through the majority of the AID across multiple species. Intron sequences were more highly divergent and given the very small size of exon I (8 nucleotides in human AID) and the long intron separating exon I and II (5753 nucleotides), it was felt that the strategy had come to its natural conclusion. Based on this information the genomic structure of AID gene was deduced as shown in Figure 25. The structure for human and chicken AID are derived from their respective publications (Arakawa *et al.*, 2002; Muto *et al.*, 2000). The map of the mouse AID gene was generated from the data in the NCBI GenBank Accession number NT_039353.6.

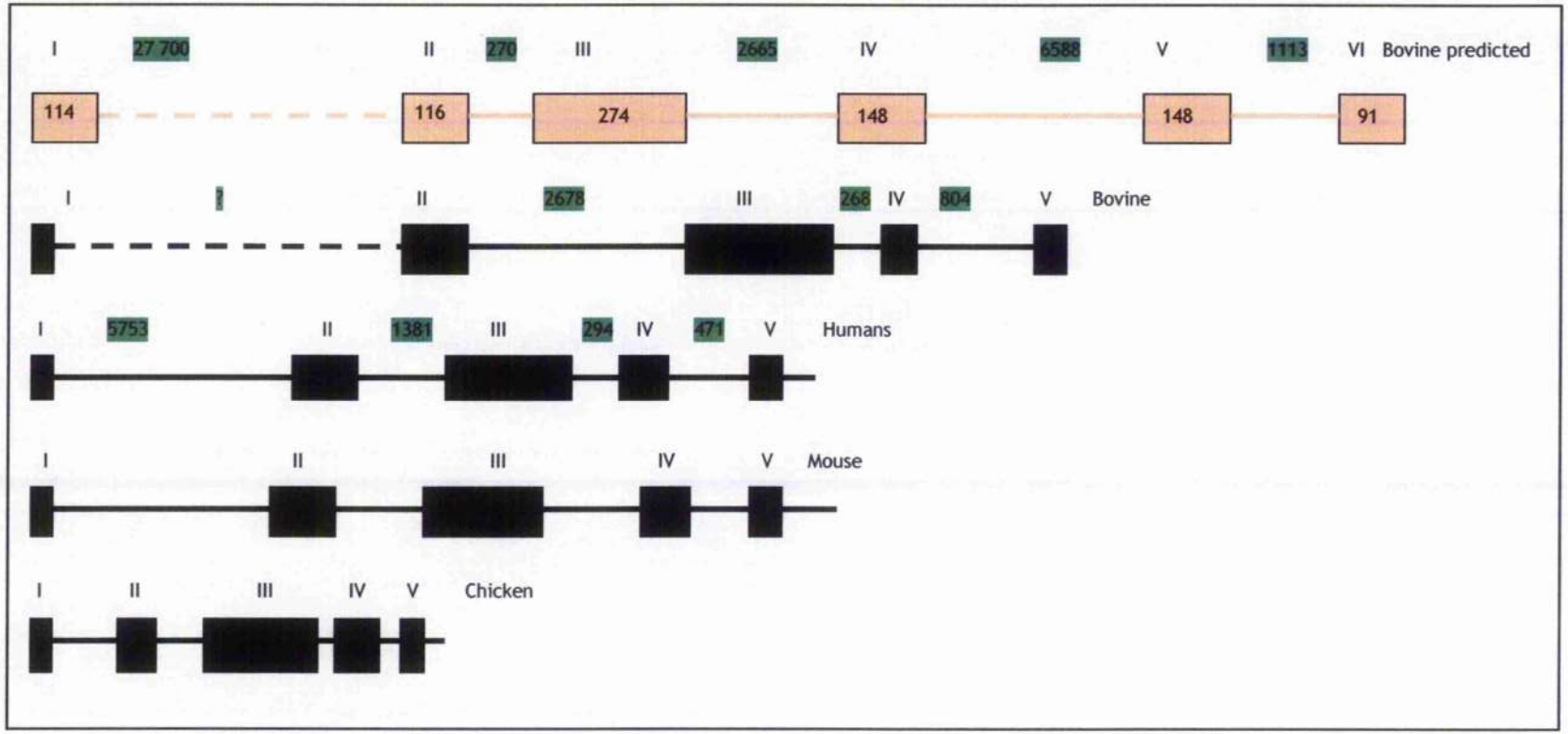


Figure 25: The genomic structure of *AID* genes in different species

The organization of *Bos taurus AID* is based on the sequencing data from this study. At this stage of the study, the size of exon I and parts of exon II and V could only be predicted and the length of the first intron (dashed) was unknown. The genomic organization in the top most line is based upon the predicted sequence of bovine *AID* (NW_001001418). Based on cDNA sequence of bovine AID and comparison with human AID the bovine exons I, II, III, IV and V are 8, 148, 274, 116 and 54, respectively. Exon III of bovine AID consists of 3 extra bases as compared to humans. The highlighted numbers in green indicate the size of introns in base pairs.

4.3.3: Recovery of full length of *AID* cDNA

During the course of this work a hypothetical cDNA sequence of bovine AID was submitted to GenBank under accession number XM_590093. This was derived by automated computational analysis from sequencing of a BAC clone forming a component of bovine genome sequencing project. This sequence was 891bp long with a predicted protein sequence of 296 amino acids. It did not match any of the reported AID sequences from mouse, human and dog especially towards 3' terminus. However, the sequence of the BAC clone (GenBank accession number NW_001001418) (see also Figure 25) provided the foundation for the synthesis of primers AID Exon-1F and AID Exon-5R against untranslated regions upstream of exon I and downstream of exon V, respectively. Further primers (AID cDNA F and AID cDNA R) were designed from a range of species, rather than by computational methods at GenBanK. Nested PCR with these primers yielded a 600 bp product which after sequencing and annotation has been deposited in GenBank under accession number NM_001038682 and is reproduced below in Figure 26.

Exon I

ATGGACAG (8)

Exon II

CCTCTTGAAGAAGCAGAGACAGTTTCTTTACCAGTTCAAAAACGTGCGCTGGGCTAAGGGCCGCCATGAGACCTACTTGTGCTACG
TGGTGAAGCGGCGGGACAGTCCCACCTCCTTCTCACTGGACTTCGGGCACCTTCGAAACAAG (148)

Exon III

GCCGGATGCCACGTGGAGTTGCTCTTCCTTCGCTACATCTCTGACTGGGATCTGGACCCTGGGCGGTGCTACCGCGTCACCTGGTT
CACGTCTTGGAGCCCCTGCTACGACTGTGCGCGGCACGTGGCCGACTTCCTGCGGGGGTACCCCAACCTGAGCCTGCGGATCTTCA
CGGCGCGCCTCTACTTCTGCGACAAGGAGCGCAAGGCCGAGCCAGAGGGGCTGCGGCGGCTGCACCGCGCTGGAGTCCAGATCGCC
ATCATGACGTTCAAAG (274)

Exon IV

ATTATTTTTATTGCTGGAATACTTTTGTGGAAAATCATGAAAGAACTTTCAAAGCCTGGGAGGGACTGCATGAAAATTCGGTTCGT
CTGTCTAGACAGCTTCGACGCATCCTTTTG (116)

Exon V

CCACTCTACGAGGTTGATGACTTGCGGGATGCATTTCGTACTTTGGGACTTTGA (54)

Figure 26: Full length *Bos taurus* AID cDNA

It consists of 600 nucleotides from five exons. Numbers in the brackets indicate the length of the contributing exon. Exon III contains one extra codon compared to humans and whole sequence has 87% identity at the nucleotide level with the human AID (Accession number: NM_020661).

The predicted protein sequence for bovine AID comprised 199 amino acids and was 94.4%, 93.9%, 90.9% and 86.9% identical with dog, human, mouse and

chicken sequences, respectively. The C-terminal leucine rich motif is conserved except for an amino acid replacement in mouse (Figure 27).

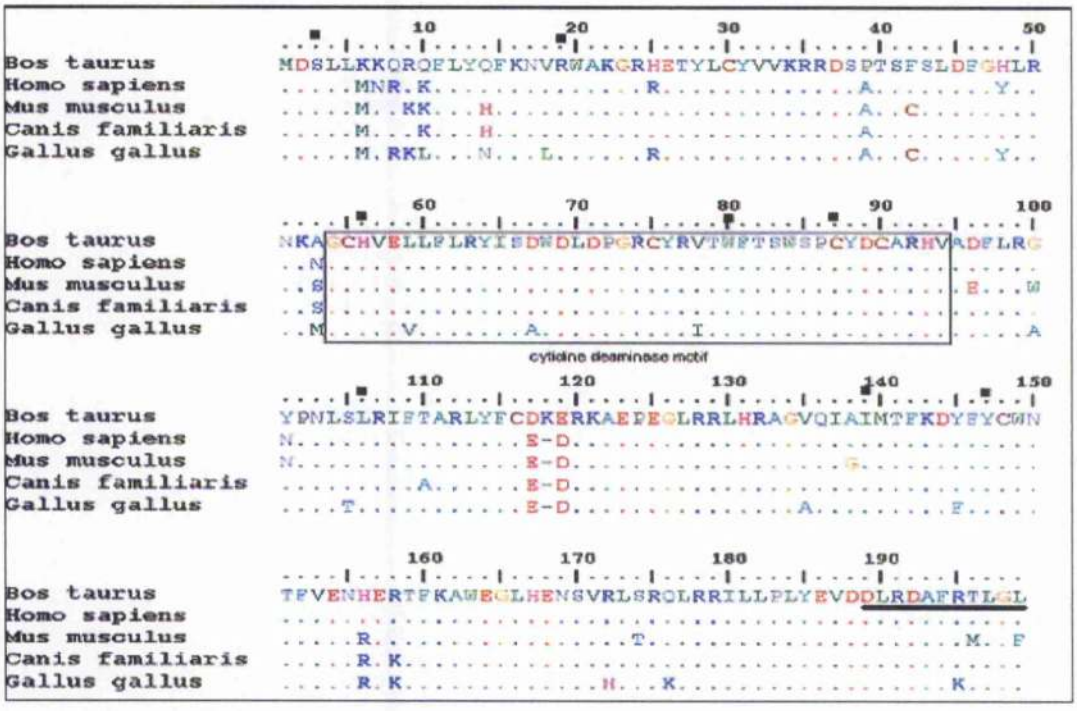


Figure 27: The alignment of bovine AID protein with other species

The sequences of the cytidine deaminase motif are boxed. Amino acids that are critical to AID mediated somatic hypermutation (Revy *et al.*, 2000) are indicated with filled square shape above the aligned sequences. The underlined sequence is a putative Nuclear Exporter Sequence (NES) and overlaps with a domain important for CSR that is also conserved (McBride *et al.*, 2004).

It was not possible to retrieve the full sequence of exon II by homology primers based on human and mouse intronic sequences as the preliminary sequence data of bovine introns from this study had revealed a great deal of variation. The partial genomic sequence of exon II and exon V were completed based on the AID cDNA sequence.

4.3.4: Expression of AID mRNA in bovine tissues

Total RNA was obtained from various tissues for cDNA synthesis. Analysis proved this to be of good quality and typical data is shown in Figure 28. The cDNA prepared from lymph nodes, spleen, liver and muscle was then PCR amplified from day old and adult cattle to recover AID transcripts. The primers sets AID F6/AID R5 and AID F2/AID R2 yielded reproducible amplicons from genomic DNA but failed to amplify product after conventional RT-PCR. This was probably due to the presence of very low levels of AID mRNA in the tissue

samples. The same primer sets were tested in quantitative RT-PCR. After optimization, quantitative RT-PCRs failed to yield reproducible results with AID specific primers, whereas consistently reproducible results were obtained

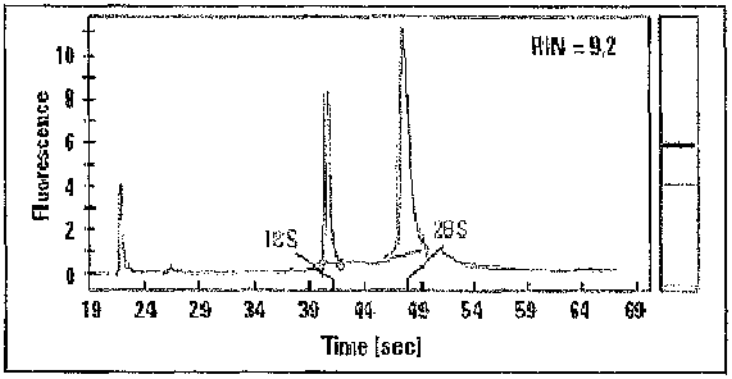


Figure 28: Electropherogram depicting the quality of RNA

For the Eukaryote Total RNA assay, two time windows (that are determined dynamically based on the ladder run) were used to assist in detecting the ribosomal RNA bands 18S and 28S for eukaryotic RNA. The area under the entire RNA electropherogram was determined and compared with the ladder area to determine the total concentration of the RNA using inbuilt software.

with a control primer set against GAPDH. To increase the sensitivity of the assay, a nested real-time RT-PCR was performed with AID F6/AID R5 as an outer primer set and AID F2/AID R2 as internal set. Although this generated amplicons of the correct size it was abandoned as it could not reflect accurately the levels of AID transcripts in the tissues. Instead, conventional nested PCR was used to simply determine the absence or presence of AID mRNA in the tissue samples. This protocol yielded a product of 131bp as expected. Transcripts of bovine *AID* were detected in lymph node, spleen and thymus in both day old and a 13 week old animal, no AID mRNA was detected in muscle tissues (Figure 29). In liver, yields of the PCR product were substantially lower than lymphoid samples suggesting much lower transcripts level were present. GAPDH transcripts were observed in all tissues examined and there was not any visible difference on the levels of expression.

4.3.5: Expression of AID in *E.coli*

The cDNA sequence derived for bovine AID was derived by comparison with genes from other organisms. To assess if the predicted protein was stable and possessed the properties expected, the cDNA sequence was cloned into an *E. coli* expression vector. pBAD was chosen because the araBAD promoter provides

a fine degree of control of expression level and at the outset, the toxicity of the translation product in a prokaryotic host could not be anticipated. The AID cDNA was fitted with restriction sites to enable cohesive ligation and to place insert in frame with the c-myc and hexahistidine tags, and stop codon present in the vector.

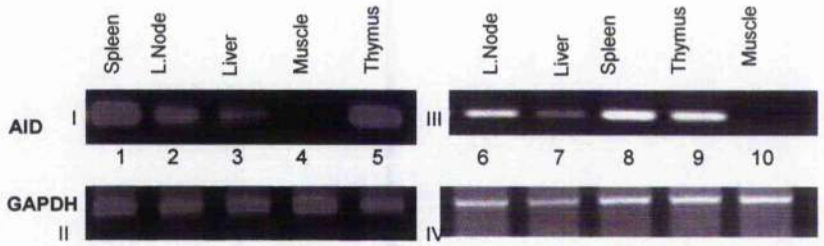


Figure 29: Expression of AID mRNA in various healthy bovine tissues

Samples (I, II-day old calf; III and IV-13 week cattle were analysed. Bovine AID transcripts of variable amounts were detected in a range of tissues from both day old and adult animal (upper lanes) using RT-PCR. Bovine GAPDH, a house keeping gene transcripts were found in all tissues studied (lower lanes).

E. coli TOP 10 cells were transformed with bovine AID cDNA cloned into the pBAD/*Myc*-His plasmid. AID expression was induced with varying concentrations of L-arabinose. Addition of L-arabinose successfully induced expression of protein of around 26 kDa at all the concentrations used (Figure 30). Future experiments were standardized at 0.2%.

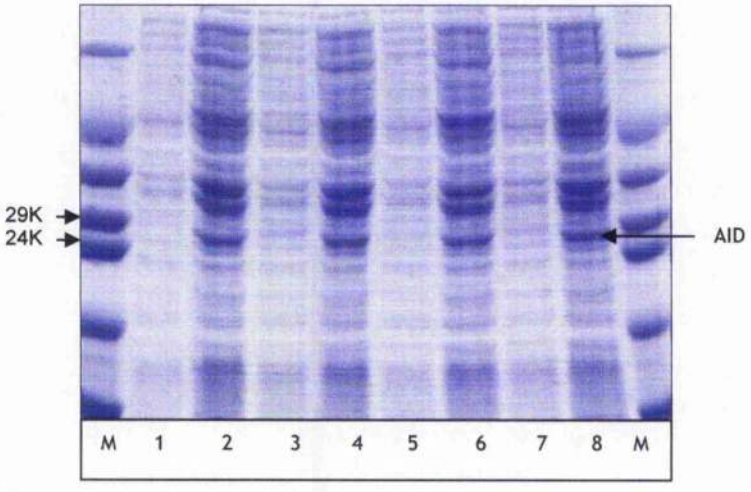


Figure 30: Analysis of AID expression by SDS-PAGE

E. coli TOP10 cells carrying a pBAD/*Myc*-His construct were induced with L-arabinose. M- Dalton marker VII-L from Sigma; 1,3,5,7 lanes-without induction; 2,4,6,8 lanes- different concentrations of L-arabinose 0.0002% to 0.2%. The SDS-PAGE gel was stained with Coomassie Blue and checked for a band of increasing intensity at 26 kDa expected for AID.

The complexity of the protein profile from whole-cell extracts prevented unambiguous identification of the induced translation product. To increase the specificity of the experiment, proteins were transferred from SDS-PAGE gels to nitrocellulose and probed with reagents against the c-myc tag. This identified a single reactive band at a molecular weight of 26 kDa (Figure 31) consistent with the size predicted for AID (24 kDa) with the fused detection and purification tags.

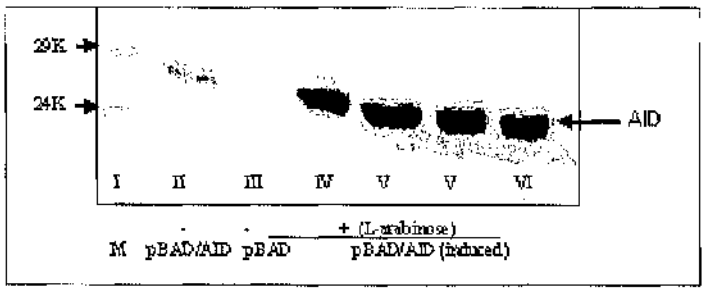


Figure 31: Detection of expression of AID on Western blots

Bovine AID in the induced lysates of *E. coli* TOP 10 was confirmed by probing the blot with anti-c-myc and anti-mouse-HRP antibodies. Lane I- protein marker(M); Lane II-Uninduced bacterial culture carrying recombinant plasmid; III-Uninduced TOP10 carrying empty vector; lanes IV-VI Induced TOP 10 with different concentration of L-arabinose ranging from 0.0002% to 0.2%.

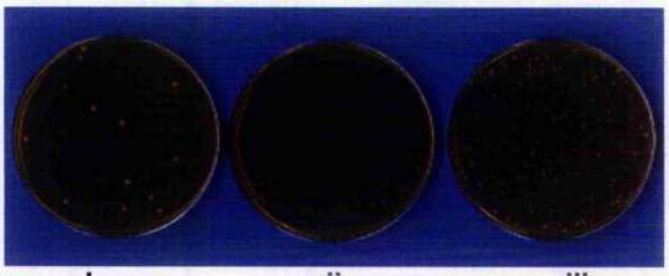
4.3.6: Effects of AID expression on *lacZ*

E. coli G6, a Lac-positive strain, was transformed with pBAD/myc-His and recombinant plasmid carrying AID. Individual clones were cultured in LB medium and induced with 0.2% L-arabinose for 4 hours and overnight. After the stipulated period bacteria were plated on LB plates containing the *lac* inducer isopropyl-β-D-thiogalactopyranoside and X-Gal as an indicator for expression of functional lacZ. Colonies with a range of phenotypes- white to intermediate to blue were observed. Clear evidence for an AID-dependent change in phenotype from blue to white proved to be difficult as the range of colony colours varied with many intermediate shades of blue. To improve the discrimination of the assay, a system based upon antibiotic resistance was developed.

4.3.7: Effects of expression of AID on r*poB*

The same plasmids were transformed into *E. coli* TOP10 (F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ*lac*×74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G). In several independent experiments, the inducer L-arabinose was added to the cultures of bacteria and the appearance of

rifampicin resistant colonies was assessed after 6-8 hours after induction. This revealed that rifampicin resistance appeared in AID-expressing *E. coli* with a frequency that was 2-6 times higher than that of controls i.e. pBAD vector only (Figure 32).



a i ii iii

E.coli strain	Plasmid	Rif^R (mutants/ 10^9 viable cells)
Top10	pBAD vector	24 (14-34)
Top10	pBAD/AICDA	75 (32-193)

b

Figure 32: AID expression and conversion of cells to rifampicin resistance

a. Appearance of rifampicin resistance in *E. coli* TOP 10 transformed with pBAD (i) and the recombinant plasmid in the absence (ii) or presence (iii) of L-arabinose inducer.b. Frequency of rifampicin-resistance in cells transformed with pBAD constructs after induction with L-arabinose.

4.3.8: Chromosomal localisation of *AID* using FISH

In order to map the *AID* gene to a bovine chromosome, biotinylated probes were prepared and hybridised onto metaphase spreads using a FISH protocol (Figure 33). The probe synthesised from a 3.4kb fragment of the AID gene failed to yield a specific signal on bovine chromosomal preparations. The signals emanated from many different chromosomes. Attempts to optimise the hybridisation using washes of various stringencies did not improve the specificity of the signal quality. The probes made from human BAC clones RP11-18J9 (NTS) and RP11-438L7 (AID) hybridised to HSA12 on metaphase spreads prepared from cultures of human blood and yielded site-specific fluorescence. Only the NTS probe yielded a specific signal from BTA5; the RP11-438L7 (AID) probe gave multiple fluorescent signals due to hybridisation to multiple chromosomes. The aim was to perform dual-labeling FISH with AID probe and a probe specific for BTA5 (e.g. NTS or DRIM). The other probes developed from BACS carrying A2M

and DRIM markers were not tested. In the absence of specific signal from either bovine or human AID probes, an alternative approach was sought.

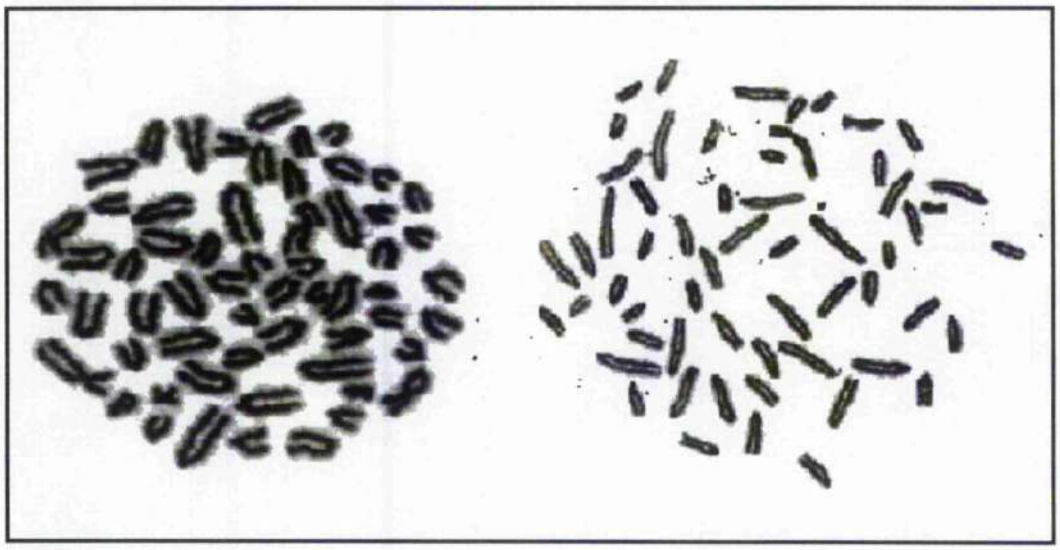


Figure 33: Typical non-synchronised (left) and synchronised (right) metaphase chromosome spreads

Bovine whole blood was cultured at 37°C in blood culture medium in the presence of colcimid. For synchronised chromosomes preparation thymidine and 2-deoxycytidine were used and the preps were used in FISH

4.3.8: Chromosomal localisation of Bovine AID using somatic hybrid panels

A cattle-hamster somatic hybrid panel consisting of 31 cell lines was screened by PCR with AID specific primers. Somatic cell hybrid panel contain small fragments of irradiated bovine chromosomes, each clone carrying a unique region. They are chosen to overlap one another to make a panel of radiation hybrids representing the whole genome and can be used to construct a high-resolution map by locating markers and genes of interest. The data obtained was 00000110000111010111110000001100, where 1 indicates positive PCR reaction with the cell line; and 0 a negative reaction. A statistically relevant concordance value of 0.97 was obtained with the BTA5 reference markers ACO2 and BM2830 both of which are located on the distal half of BTA 5 allowed the syntenic assignment of AID to BTA 5. RH mapping using same primer sets on a cattle-hamster $WGRH_{5000}$ panel consisting of 90 cell lines resulted in the output: 0000000000 0000000000 0000001010 0100000000 1000101000 0000002000 0000000100 0000000010 1000000001. The data evaluation, based on LOD score

which was greater than 10, places AID within a distance of 11.00 cR_{5000} from the marker MAGP2 on BTA 5 (Figure 34).

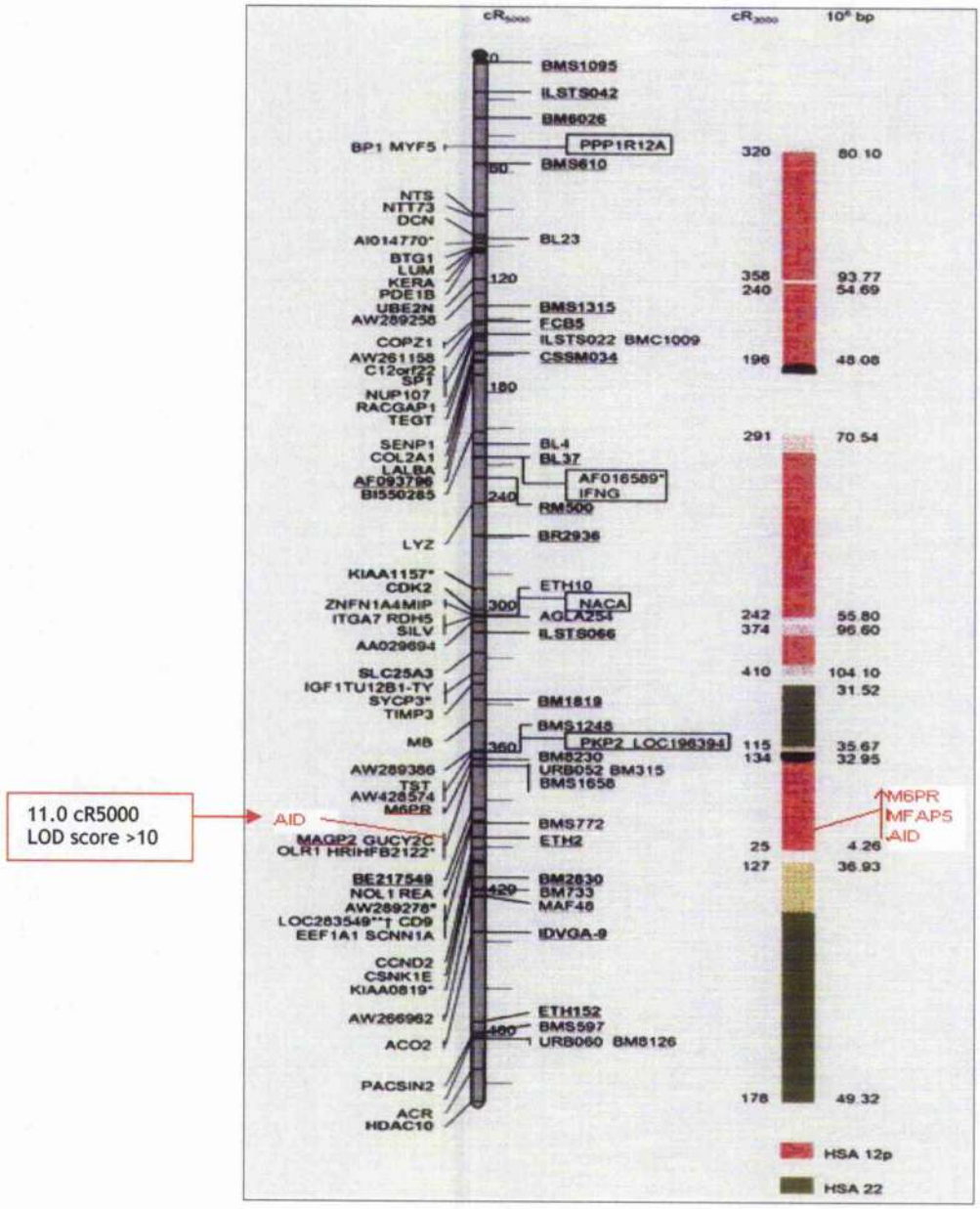


Figure 34: Chromosomal assignment of bovine AID using a cattle-hamster $WGRH_{5000}$ panel (Band *et al.*, 2000)

Human chromosome coordinates of genes mapped in both species have been used to align human chromosome segments from HSA 12p and HSA 22 (right panel) along the bovine BTA 5 (left panel). The figure has been adapted (Everts-van der Wind *et al.*, 2004).

4.4: DISCUSSION

This part of the project cloned and sequenced partial *AID* gene from cattle. Coding sequences of *AID* were found to be very similar to their counterparts from humans and mice, implying a very high level of stability over the eons since mammals emerged as a distinct branch of vertebrate evolution. Detailed phylogenetic analysis (Zhao *et al.*, 2005) shows that although the chicken AID orthologue has a radically different gene structure (Figure 25), it possesses a similar sequence. The *Xenopus* enzyme shows differences but it is also part of the same clade (Zhao *et al.*, 2005). AID from fish (fugu, catfish and zebrafish) are rather different in sequence (Zhao *et al.*, 2005), but are still able to catalyse CSR in mouse B cells (Barreto *et al.*, 2005), a further demonstration of the ancient origin of this enzyme. In contrast, little sequence similarity could be detected between intronic sequences of bovine AID and its counterparts from other mammals. Bovine intron sequences were also consistently longer than those of humans and mice. It is assumed that intron regions are not subject to selection, and are able to mutate at random thus accumulating mutations over a period of time. The size of intron I (5.7 kb) for the human gene (27.7 kb predicted for bovine) and the very small size of exon 1 (8 bp) prevented recovery of a full sequence of bovine AID by conventional PCR methods with homology primers.

The GenBank entry of a BAC clone predicting the sequence of bovine AID was inconsistent with data but helped to retrieve the full length bovine AID cDNA. Had this BAC sequence not been available, different approaches could have been employed to recover the AID sequence. One such approach would have been to screen a bovine genomic library with a probe constructed from the partly bovine cloned gene. Since this study used oligo (dT) primed reverse transcription this must have ensured the inclusion of 3'-end of the mRNA. To retrieve the 5' end of the mRNA, Rapid Amplification of cDNA Ends (RACE) could have been used. This involves performing a randomly-primed reverse transcription reaction, adding an adapter to the terminus of the cDNA by ligation or polymerase extension, and amplifying by PCR with a gene specific primer and a primer that recognizes the adapter sequence.

The bovine AID cDNA has an open reading frame of 199 amino acid residues. A comparison with its counterpart in mammals and chicken identified mostly conserved amino acid residues especially within the cytidine deaminase motif which has been shown to be important for the activity of AID (Papavasiliou and Schatz, 2002). Studies of hyper-IgM syndrome (HIGM) in human patients have identified the amino acid substitutions in AID leading to loss of CSR and SHM (Revy *et al.*, 2000). The mutation of these amino acids as indicated in Figure 6 resulted in abolition of SHM function of the protein. The phenotype observed in these patients demonstrated that there was an absolute requirement of AID for B cell terminal differentiation necessary for efficient antibody responses (Revy *et al.*, 2000). All these amino acids were conserved in bovine AID. The bovine AID gene carries extra codon in exon 3 which codes for lysine, the significance of which is unknown. This occurs in a part of the reading frame that is otherwise conserved across a wide range of vertebrates (Figure 7 and (Zhao *et al.*, 2005)). Interestingly, the AID proteins of fugu, zebrafish and catfish also possess an additional codon at this position (leucine in fugu, glutamic acid in zebrafish and catfish) (Zhao *et al.*, 2005). The data from this study showed that bovine AID is able to confer a mutator phenotype on *E. coli*, indicating the functionality of the protein; the protein in fish is similarly active (Barreto *et al.*, 2005). This suggests that a degree of variability can be accommodated without loss of function.

Other regions of the protein are clearly more sensitive. It has been shown that deletion of the carboxy-terminal 10 amino acids of AID uncouples CSR from SHM and GC (Barreto *et al.*, 2003). This mutant AID was more active in catalyzing cytidine deamination than wild-type in *E. coli* and resulted in high levels of SHM and GC when expressed in eukaryotic cells but failed to carry out CSR. These authors further observed that defective CSR in B cells expressing the AID mutant was not due to the absence of targeting to switch regions or decreased catalytic activity. This was revealed by the pattern of mutation in the Sμ region, a process that has shown to be absolutely dependent on cytidine deamination by AID (Nagaoka *et al.*, 2002; Petersen *et al.*, 2001). These findings suggest that AID has a CSR-specific function, perhaps recruiting or assembling CSR-specific DNA repair factors in a manner that is independent of its ability to deaminate C residues in ssDNA (Barreto *et al.*, 2003).

The other function that has been assigned to the C-terminal domain is translocation of AID from the nucleus to the cytoplasm. Conserved leucines and hydrophobic amino acids residues in the C-terminal domain of murine AID have been shown to be critical to this Nuclear Exporter Sequence (NES) function (McBride *et al.*, 2004). The alignment of the NES motifs from a range of important molecules is shown below in Table 9. A soluble shuttling receptor called chromosome region maintenance/exportin 1 has the ability to bind these leucine-rich NESs and translocate them out of the nucleus by interacting with the Ran GTPase and nuclear pore components (Mattaj and Englmeier, 1998; Pemberton *et al.*, 1998).

Table 9: Conservation of NES motifs

AID (cow)	189	D	L	R	D	A	F	R	T	L	G	L	199
AID (mouse)	188	D	L	R	D	A	F	R	M	L	G	F	198
MAPKK		A	L	Q	K	K	L	E	E	L	E	L	
PKI (alpha)		E	L	A	L	K	L	A	G	L	D	I	
STAT1		S	L	A	A	E	F	R	H	L	Q	L	
Cyclin B		D	L	C	Q	A	F	S	D	V	L	I	
HIV-Rev		L	Q	L	P	P	L	E	R	L	T	L	

The leucines and hydrophobic amino acids critical to NES function are shown in bold and are conserved across various enzymes especially in AID from cow and mouse. This Table has been adapted (McBride *et al.*, 2004).

Residues associated with the putative NES of bovine AID are identical to those of the murine enzyme and critical residues are conserved with other exported proteins. It is also not known whether the sequence difference in the AID proteins from different vertebrates accounts for different patterns of mutations.

Given the convenience of the mutation assay carried out as a part of this study, one possibility would be to assay the frequency with which *E. coli* was converted to rifampicin-resistant phenotype when AID from fish, chickens, mice and cattle were expressed. This kind of experiment has been tackled as a part of a phylogenetic analysis (Ichikawa *et al.*, 2006). Better still would be to express and purify each deaminase for assay *in vitro* (Larijani *et al.*, 2005). It is worth noting however that SHM and GC diversification processes with completely different foundations are closely linked at the molecular level. Arakawa *et al.* (2004) showed that progressive deletion of pseudogenes from the chicken line

DT40 could abolish AID-dependent gene conversion, switching the cell to a pattern of V-gene mutation similar to SHM (Arakawa and Buerstedde, 2004). A similar outcome could be obtained by inactivation of RAD51 paralogues in DT40 (Sale et al., 2001). Uracil-DNA glycosylase (UNG) also contributes to this process (Saribasak et al., 2006) although its activities apparently extend beyond removal of uracil from the deaminated target (Longerich and Storb, 2005). Although AID is clearly essential to SHM, GC and CSR, these studies indicate that the organisation of the Ig loci and the expression of other factors have a major influence in the overall outcome.

The frequency of somatic mutation appears to differ during ontogeny. The results from other components of this study revealed a great deal of variation in frequency of mutation in the CDRs of IgH from young and older animals. AID has a crucial role in the process of somatic hypermutation. Mice that express mutated AID do not exhibit somatic mutation and class switching (Muramatsu *et al.*, 2000) and mutation in AID causes hyper-IgM syndrome, as characterised by the inability to produce other immunoglobulin isotypes (Revy *et al.*, 2000). SHM was not detected in the germinal centre of aged mice immunized with hapten-conjugate T-dependent antigen (Miller and Kelsoe, 1995). To what extent might the diversification of bovine IgH therefore depend upon the level of AID expression in animals of different ages? The level of transcription of AID was compared in young and old animals. Whilst quantitative PCR proved problematic, there was no evidence from this study of dramatic variation in transcription of AID between young and mature animals. Interestingly, AID transcripts were present in a day old animal even though germinal centres could not be detected in the lymphoid tissue. This suggests that AID may be activated through T-independent mechanisms, perhaps to drive the diversification of the Ig repertoire during the later stages of gestation or immediately after birth. There is evidence that mouse B cells can be activated to express AID by exposure to papillomavirus and signalling through TLR4-MyD88 (Yang *et al.*, 2005). Similarly human B cells exposed to hypomethylated CpG DNA upregulate AID and undergo CSR (He *et al.*, 2004). This arises from signals transmitted via TLR9. An interesting area for future research would therefore be to identify the environmental or maternal trigger to AID expression in neonatal cattle and to assess the impact of these events upon Ig diversification. Experiments of this

kind have been carried out in rabbits, identifying the importance of interaction amongst the intestinal microflora in the development of GALT and diversification of the Ig repertoire (Rhee *et al.*, 2005b; Rhee *et al.*, 2004).

The expression pattern of AID mRNA in bovine tissues was similar to that reported for canines and humans. AID mRNA in these animals has been shown to be present in both lymphoid and non-lymphoid tissues. In dogs it is expressed in thymus, lungs, liver, small intestine and kidneys apart from lymphoid tissues (Keitaro *et al.*, 2004). In humans it is expressed strongly in lymph nodes and tonsils and to a lower level in thymus, kidney, pancreas and other tissues examined (Muto *et al.*, 2000). Honjo and colleagues (Muramatsu *et al.*, 1999) have reported that AID transcripts in mouse bone marrow are present at a level almost comparable to that seen in spleen. Transcripts of AID were observed at low abundance in bovine liver, and were undetected in muscle tissues. The low level of AID transcript in liver could be explained by the presence of activated B cells migrating between lymphoid organs. Since nested PCR was used another explanation for the apparent presence of AID transcripts in the liver and thymus would be that RT-PCR is detecting very low level ectopic expression of AID.

Although bovine AID cDNA was very similar in its nucleotide sequence to its counterparts in other vertebrates, evidence was sought of its biological activity. For this, the cDNA was cloned under the control of an *ara*BAD promoter in the vector pBAD. This plasmid is a pBR322-derived expression vector designed for regulated, dose-dependent expression of recombinant proteins in *E. coli*. The regulatory protein, AraC, is translated from the pBAD/*Myc*-His vectors allowing regulation of araBAD activity. In the presence of L-arabinose, expression from pBAD is turned on while in the absence of inducer, very low levels of transcription takes place (Guzman *et al.*, 1995; Lee *et al.*, 1987). Un-induced levels are repressed even further by growth of bacteria in the presence of glucose. Glucose reduces the levels of 3′, 5′-cyclic AMP, thus lowering expression from the catabolite-repressed pBAD promoter (Miyada *et al.*, 1984). *E. coli* from two different genetic backgrounds were then transformed to assess the mutator ability of bovine AID to cause mutation in *lacZ* and *rpoB*. In the first set of experiments, it was hypothesized that expression of AID would induce mutations at cytidine residues in *lacZ* leading to introduction of a range of

modifications including stop codons and thus disruption of the *lacZ* reading frame. The failure to metabolize X-gal would lead to growth of white colonies on plates containing X-gal and IPTG. The appearance of mutants was evident but the appearance of colonies with a range of intermediate phenotypes made quantitation of mutation frequency impossible and this approach was abandoned as data was too prone to error.

The second approach was based on the observation that nucleotide substitutions in *rpoB* produce rifampicin resistance which occurs constitutively at a low frequency (Miller, 1972). Petersen-Mahrt and colleagues (Petersen-Mahrt *et al.*, 2002) showed that expression of human AID in *E.coli* resulted in a mutator phenotype that yielded nucleotide transitions at dC/dG in a context-dependent manner in different genetic backgrounds. Almost all mutations arising upon expression of AID were G-A to C-T transitions, consistent with the DNA deamination model. They further observed enhanced AID- triggered mutations when they used a uracil-DNA glycosylase deficient *E.coli* mutant. These effects were also evident when mutation to nalidixic acid, valine or fucose resistance was monitored. Expression of human AID in yeast also induced mutation at G/C pairs in immunoglobulin genes, further evidence that the intrinsic substrate specificity of AID is a primary determinant of mutational hotspots during SHM (Mayorov *et al.*, 2005). The findings that bovine AID could elicit a mutator phenotype in *E. coli* confirmed that cDNA encoded a protein with biological activity.

One feature of the activation of AID is that it targets Ig genes for SHM without introducing mutations in most other genes, the exception being BCL6 in human B lymphocytes and Igα and Igβ in activated mouse B cells (Gordon *et al.*, 2003; Shen *et al.*, 1998). Recent work has shown that AID may be recruited to Ig genes during the process of transcription by molecules which bind to Ig enhancers or other cis-elements, providing highly regulated targeting. Among candidate cis-elements is a protein that binds the sequence CAGGTG which is present in all the Ig enhancers and is a target for E2A binding protein. This motif also occurs in Ig V regions and its presence has been shown to correlate directly with SHM (Michael *et al.*, 2003). BCL6, Igα and Igβ also have CAGGTG elements dispersed throughout their coding region as well as within transcriptional unit,

perhaps explaining their targeting by AID. Although CAGGTG occurs on average every 4 kb, it is possible that its position within or near to a transcribed gene determines whether AID is recruited as suggested by Storb and co-workers (Longerich *et al.*, 2006).

The RH mapping places bovine *AID* gene on BTA 5 and this finding is further supported by the fact that human *AID* also lies on the conserved segment of HSA12 which contains large fragments syntenic with BTA 5 (Liu *et al.*, 2003). The new second-generation RH maps have proved to be a powerful resource for enhancing the resolution of cattle-human comparative maps and allowing a more comprehensive analysis of chromosome organisation between the species. Taking a closer look at the comparative map (Figure 34), it appears that multiple internal rearrangements have occurred. Such rearrangements are common and most likely represents recent inversions that date after the divergence of the most recent ancestor of the primates and artiodactyls (Everts-van der Wind *et al.*, 2004). Comparative mapping is a useful tool for studying chromosome evolution and for the transfer of information between species especially to bovine where little information is available on this aspect. Many quantitative trait loci have been mapped to BTA 5 including ovulation rate (Kappes *et al.*, 1997), birth weight (Stone *et al.*, 1999) and twinning rate (Lien *et al.*, 2000). Since this chromosome also carries loci for important genes such as IFN-γ and AID, further analysis and integration of microsatellite markers on to this chromosome would be very helpful to identify the genes underlying economically important traits related to production and disease.

Chapter 5: Characterisation of BL-3: a bovine lymphosarcoma cell line to study Ig diversification *in vitro*

5.1: INTRODUCTION

The availability of lymphoid tissue is an essential requirement for studies of the mechanisms associated with antibody formation in cattle. Many impediments are associated with obtaining specimens from bovines. All bovine lymphoid tissues are subject to usage restriction intended to limit exposure to transmissible spongiform encephalopathies and their entry into the food chain. In particular, the availability of foetal and neonatal material is extremely limited and unpredictable in supply. In the studies described earlier, access to samples from young cattle could only be arranged through collaboration with colleagues at another Scottish research institute. The choice of the sampling time was therefore beyond our control. Similarly the size of calves, the limited availability of containment, and the high cost of the facilities dictates that animals cannot be sourced from "germ-free" conditions. Hence the contribution of environmental factors to Ig diversification cannot be easily assessed.

A further complication to studies of the biology of bovine B cells is the potential heterogeneity of cells isolated from tissues of animals. Lymphocytes may be present at different stages of development, and depending on the age of the animal and its antigenic exposure, cells may be in different states of activation. It is also inevitable that cells will have undergone independent processes of rearrangement at the Ig heavy and light chain loci. In mouse and human systems, some of these difficulties can be overcome by sorting cells to isolate lymphocytes that are at a common stage of development. This is not yet possible in cattle as stage-specific markers associated with developmental progression of B cells are not known and hence the reagents required for their isolation are lacking. To overcome these obstacles, the project would benefit from a well-characterised bovine B cell line of lymphoid origin for *in vitro* studies of Ig diversification. Such a system could not replace the information gained from fresh animal tissue; however it could be an important tool to investigate immunological, genetic and molecular mechanisms at cellular level and be a useful model to study the requirements for SHM, GC and CSR.

There is not a single well characterised cell-line that can be utilised to study the mechanisms of bovine immunoglobulin diversification. Many such cell lines of mammalian and avian origin are available to study repertoire development. Human Burkitt's lymphoma cell lines are widely studied and those that have already undergone hypermutation *in vivo* and are used as an *in vitro* model (Sale and Neuberger, 1998). The chicken B cell line DT40 has been used as a model to understand gene conversion and to test the contributions made by candidate genes through mutational approaches (Takeda *et al.*, 1992). This is possible because the cell line can be easily transfected and hence knock in and knock out constructs can be generated. The human cell line Ramos, a Burkitt's lymphoma line, has been similarly informative. Other useful cell lines are CL-01 and BL-2 in which hypermutation is inducible. Inducible systems are good options as they offer the prospect of tracking in time the events that ultimately lead to DNA modification at the Ig loci. Studies of these cell lines have identified processes that occur *in vivo* (e.g. specificity of SHM for rearranged V regions, the predominance of transitions to transversions, targeting of mutagenic lesions to specific motifs) and where novel activities have been identified the data serves as a foundation for studies *in vivo* (e.g. through transgenesis).

If a cell culture system comparable to DT40 or Ramos lines could be set up for studies of bovine Ig diversification, its potential benefits would be significant. Most obviously, it would remove many of the logistic and biological complications outlined earlier. It would also allow testing of the conclusion drawn earlier in this thesis that SHM is an important process of diversification for the bovine Ig repertoire. The aims set for this phase of study were therefore to identify and characterise a cell line that might serve as an *in vitro* model for Ig diversification. Characterisation should comprise study of the Ig locus, the expression of the functional Ig, characterization of CD markers and measurement of basal levels of AID expression. Depending upon outcome, further goals might be to attempt transfection of the line with a view to elevating or down-regulating AID, and then assessing impact upon the IgH locus.

5.2: MATERIALS AND METHODS

5.2.1: Growth and maintenance of cells

After a thorough survey, the only cell line of bovine origin available from commercial sources was the BL-3 lymphosarcoma. This line was originally derived from a case of sporadic bovine leucosis and data suggested the presence of surface IgM and MHC class II supporting the idea that the lymphoma was of B cell origin (Romano *et al.*, 1989). The cell line was purchased from the European Collection of Cell Cultures, U.K and was maintained as described earlier. Briefly, the cells were cultured at 37°C with 5% CO_2 in 50% RPMI 1640 - 50% Leibovitz L-15 medium (Invitrogen) supplemented with 20% FBS, 2 mM L-glutamine, penicillin (100 U/ml), and streptomycin (100µg/ml). Cells were grown to a density of $1x10^9$ cells/ml in 5 ml tissue culture flasks and were sub-cultured with fresh medium every 72 hours. The cells were screened for any mycoplasma contamination by PCR using a kit from Minerva Biolabs, the VenorGeM Mycoplasma Detection Kit.

5.2.2: Characterisation of BL-3 cells for cell surface markers

BL-3 cells were grown in complete BL-3 medium as before. The cells were centrifuged at 300 g for 5 min at 4°C and resuspended in RPMI-1640. They were further washed twice more to remove residual serum resuspended to a final concentration of $1x10^7$ cells/ml in PBS.

5.2.3: Staining for sIgM

A series of 500 µl aliquots of BL-3 cells were transferred into 1.5 ml microfuge tubes. A rabbit anti-bovine IgM antibody was added to a final dilution of 1:1000 and the sample was incubated for 30 min at 4°C on a revolving wheel. At the end of the incubation, 1 ml of PBS was added, tubes were vortexed and centrifuged at 300 g for 5 min at 4°C. The supernatant was discarded and cells were washed twice in PBS and finally resuspended in 500 µl PBS. After this, an anti-rabbit Ig-FITC conjugate was added to at a final dilution of 1:1000 and the sample incubated for 30 min as before. At the end of incubation, samples were washed twice and the cells were resuspended in 1ml of PBS. After transfer to 12 x 75 mm polypropylene FACS tubes, samples were kept on ice before analysis by flow cytometry. Controls included unstained cells and samples lacking the primary antibody.

5.2.4: Staining for other cell surface molecules

BL-3 cells were also characterised for the presence of other cell surface markers. These included CD21, CD40 and IgG. For CD21 staining, mouse anti-bovine CD21 antibody (Serotec, UK) was used at a working dilution of 1:50. For secondary labeling, anti-mouse-FITC conjugate was used. For IgG labeling, sheep anti-bovine IgG antibody (Serotec, UK) was used at 1:100 dilution followed by anti-sheep-FITC conjugate. The presence of CD40 on BL-3 was ascertained by labeling with the mouse monoclonal antibody IL-A156 (kindly provided by Dr. N. McHugh, Edinburgh) and detection with anti-mouse-FITC conjugate. The samples for FACS analysis were labelled and prepared at 4°C in the same manner as described for IgM staining.

5.2.5: Recovery of the rearranged heavy chain variable locus

Genomic DNA was extracted from cell pellets using a Qiagen DNA extraction kit as per the manufacturer's instructions. One microgram of genomic DNA was used as template to recover the rearranged heavy chain variable locus (Figure 35) with the primers VHF and IR-1. The sequences of these primers are shown in Table 1, Chapter 3.

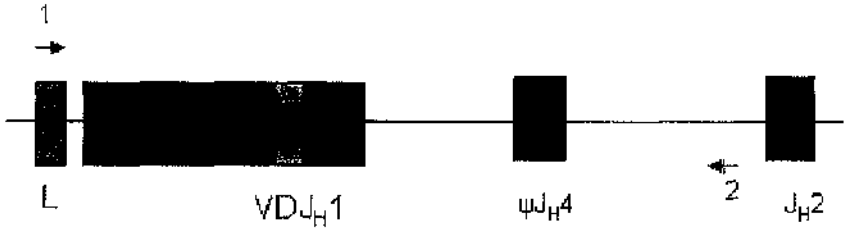


Figure 35: Organisation of the rearranged IgH locus showing primer binding sites

Forward primer (1, VHF) anneals to bases 12 to 31 of the leader sequence (GenBank accession number: U55165) and the reverse primer (2, IR-1) binds to bases 1636 to 1654 downstream of ψJ_H4 (Zhao *et al.*, 2003) allowing amplification of rearranged coding and non-coding regions of Ig heavy chain locus.

Primers were used at final concentrations of 800 nM in a reaction buffer containing 2mM magnesium chloride and 200µM dNTPs. Amplification was performed in final volumes of 25µl with 2.5 U of recombinant Taq polymerase (Invitrogen). After initial denaturation at 94°C for 5 min, reactions were cycled 30 times through 94°C (1 min), 64°C (40 sec) and 72°C (2 min) with final 5 min incubation at 72°C. PCR products were resolved on a 1.2% agarose gel to confirm

that the product was of the predicted size. The separated product was extracted from the gel using an extraction kit (Qiagen, UK). The extracted fragment was ligated into the plasmid vector TOPO pCR 2.1-TOPO and transformed into *E. coli* TOP10 cells. The clones were selected on LA plates containing 100µg/ml ampicillin by incubating at 37°C overnight. Clones were then picked randomly and liquid cultures were setup for the preparation of plasmid DNA to screen clones carrying the rearranged insert. Positive clones were sequenced using M13 forward and M13 reverse primers as before at the Molecular Biology Support Unit, Sir Henry Wellcome Functional Genomics Facility of IBLS, University of Glasgow.

5.2.6: Fingerprinting of IgH PCR products by a *Bst* N1 digestion

PCR products from the rearranged IgH locus were purified and treated with *Bst* N1 restriction endonuclease at 60°C that cuts Ig-coding sequences with high frequency. This allowed rapid assessment of sequence of multiple PCR products. In the course of study, PCR products were amplified from the rearranged IgH locus of BL-3 cells exposed to LPS and unstimulated cultures. By digesting with *Bst* N I, rapid assessment was possible of whether LPS stimulation resulted in significant changes to the IgH locus.

5.2.7: Detection of AID mRNA in BL-3 cells

Total RNA was extracted from BL-3 cells with RNAeasy Midi Kit (Qiagen) using manufacturer's protocol for isolation from growing cells. After removal of contaminating DNA, RNA quality and concentrations were determined as described previously. Single-stranded cDNA was synthesized from 1µg of total RNA with Superscript II (Invitrogen) as described before. The nested PCR method detailed earlier for assessing AID expression was used. Again, this method was used in the course of the study with samples from BL-3 cells treated with LPS and unstimulated cultures.

5.2.8: Stimulation of BL-3 Cells in culture

In experiments with BL-3 cells LPS was used at 30µg/ml. BL-3 cells were grown as previously described. Cells were transformed at a density of $1x10^7$/ml to a six well plate and allowed to grow either in the absence or presence of LPS from *Salmonella typhosa* (Sigma). Cells were treated in two sets and harvested

after 90 min, 24 hours and 48 hours of stimulation. After the end of the stipulated period, the cells were aspirated from the wells, pelleted and then washed twice with RPMI-1640 before snap freezing and storage at -80°C, pending extraction of genomic DNA or total RNA.

In other experiments, BL-3 cells were cultured for 2 and 4 days in medium alone or complete medium supplemented with anti-CD40 antibody (10 µg/ml; a kind gift from Dr. N. McHugh, Edinburgh, U.K.), recombinant bovine IL-4 (10 ng/ml; Endogen), IL-4 and anti-CD40, IL-4, anti-CD40 and LPS or a cocktail of IL-4, anti-CD40, LPS and anti-IgM (1:500 dilution). RNA was extracted from $1X10^6$ cells taken from untreated and treated cultures and processed as before to assess AID induction.

5.2.9: Detection of IL-4R mRNA in BL-3 cells

Single-strand cDNA prepared from BL-3 cells was used in RT-PCR with primers designed from the predicted sequence of *B. taurus* IL-4R (XM_877824). The forward primer (IL-4R-for) has a sequence of 5'TGCTTCTCGGACTACATCAG3' and reverse primer (IL-4R-rev) runs as 5'ATGTTGACCAGGTAGGTGAG3'. The reaction was carried out with the following parameters. The PCR amplifications on total cDNA consisted of initial-denaturation (94°C, 2 min), followed by 30 cycles of denaturation (94°C, 40 sec), annealing (58°C, 40 sec), extension (72°C, 40 sec), and by single final extension at 72°C for 5 min. As a control for cDNA quality, GAPDH primers purchased from Endogen were used. The PCR products were resolved on 1% agarose to confirm the presence and size of the amplicons.

5.2.10: Gene transfer into BL-3 cells

5.2.10.1: Electroporation

In total, 2×10^5 BL-3 cells were harvested and resuspended in 300 µl of RPMI-1640. Five micrograms of pEGFP-1 plasmid DNA was added, and the reaction mixtures were incubated on ice for 10 min. pEGFP-1 encodes a red-shifted variant of wild-type green fluorescent protein GFP (Chalfie *et al.*, 1994; Inouye and Tsuji, 1994; Prasher *et al.*, 1992) and has been optimized for brighter fluorescence and higher expression in mammalian cells. Cells were electroporated with a Bio-Rad Gene Pulser set at 260 V, 960 µF, and 200 Ω with cuvettes with a gap of 0.3 cm. After further 10 min incubation on ice, the cells

were plated into a six-well plate with 3 ml of complete medium and incubated at 37°C in 5% CO_2 for up to 72 hours. Cells were examined at 12 hours intervals to check for expression of GFP. Cells were viewed on an immunofluorescence microscope. COS-7 cells were used as a positive control for transfection efficiency.

5.2.10.2: Liposomal gene transfer in BL-3

BL-3 cells were transfected with pEGFP-1 plasmid using Lipofectamine 2000 (Invitrogen). The cells were seeded out in 24-well plates at densities of $5x10^5$ cells in 500 µl complete BL-3 medium without antibiotics. The Lipofectamine and pEGFP-1 were mixed in different ratios to optimize the transfection efficiency according to the manufacturer's recommendations. The resulting complexes were then added slowly and carefully to the cells. The cells were incubated at 37°C in 5% CO_2 for up to 72 hours after lipofection and examined at 12 hours intervals for the expression of GFP.

5.2.10.3: Gene transfer using viral vectors

As the transfection efficiency of BL-3 cells with Lipofectamine was not encouraging, it was decided to test viral-mediated DNA delivery.

5.2.10.4: Adenoviral-mediated gene transfer

Adenovirus carrying a GFP reporter construct was supplied by Prof. G. Gould (University of Glasgow). Viral stocks were non-CsCl-purified. Infections were carried out in 24-well plates, 5 x 10^5 cells in 50 µL of PBS supplemented with 1 mM $MgCl_2$. Viral stocks were diluted to generate multiplicities of infection (MOIs) ranging from 0 -200. After 2 hours of incubation at 37°C, 1 mL of complete culture medium was added to the cells. As before COS-7 cells were included as a positive control. Cells were observed at 12 hours intervals for the expression of GFP. In a revision of standard infection methods (Buttgereit *et al.*, 2000), adenovirus stocks were preincubated with Lipofectamine at a final concentration of 10ug/ml in 50 µl of the complete medium without serum. The virus-lipid mixture was incubated at 37°C for 30 min before addition to 5 x 10^5 growing BL-3 cells. The cells were examined for the expression of GFP at 12 hourly intervals for 72 hours post-infection.

5.2.10.5: Lentiviral transfection

Lentivirus carrying a GFP reporter construct was supplied by Prof. B. Whitelaw (Roslin Institute, Edinburgh) and infections were carried out in a tissue culture suite with the generous co-operation of Dr. L. Glass, Roslin Institute. BL-3 cells were grown in complete BL-3 medium and before transfection were suspended to $5x10^5$ cells/ml. As a positive control, Hapa 1 cells were also infected with lentivirus carrying the PGK-GFP cassette. Cells were examined at 12 hours intervals for 3 days before being discarded.

5.2.10.6: Growth patterns after transfection

Cell viability was determined by trypan blue exclusion. PBS was used for mock infection. Non-infected and infected cells were counted every 12 hours up to 3 days post-infection. Cell counts were determined at 24, 48 and 72 hours post-treatment.

5.3: RESULTS

The BL-3 cells grew and divided efficiently in a complete medium specially adapted for this cell line (C. Menge, University of Giessen, personal communication) and required subculture at 1:10 dilution every 3 days. The cells did not have an entirely rounded shape but appeared slightly polygonal perhaps due to the formation of surface protusions (Figure 36). Cells were screened for mycoplasma contamination using a PCR method and were found to be negative (Figure 37).

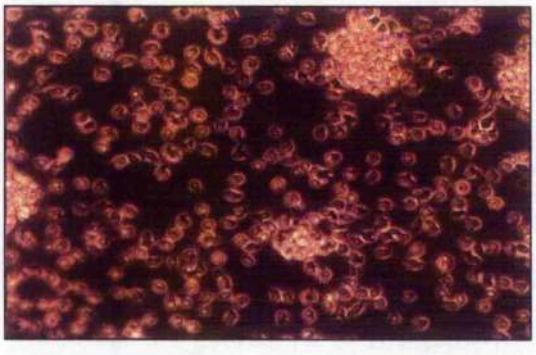


Figure 36: Growth of BL-3 cells in culture

Image captured on a Zeiss Axiovert inverted microscope at 20X magnification, with phase contrast illumination.

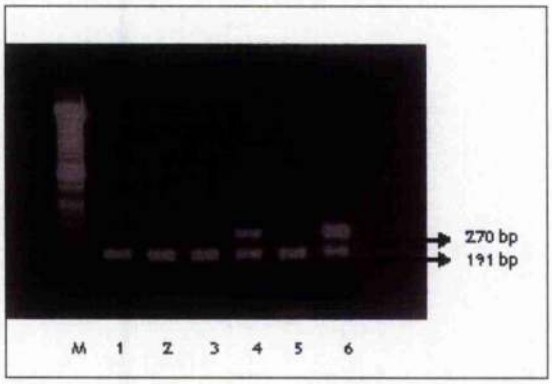


Figure 37: PCR analysis of BL-3 cultures for mycoplasma contamination

Lane M: 100 bp DNA molecular weight marker (Invitrogen); lanes 1-3; samples from BL-3 cultures; lanes 4-6: samples from unrelated cell lines tested on the same occasion. The band of 191 bp (indicated) is an internal control to check successful amplification. The upper band (270 bp; indicated) shows amplification from mycoplasma DNA and indicates contamination of the culture.

5.3.1: Functional rearrangement of the variable heavy chain locus in BL-3 cells

PCR with primers specific for rearranged heavy chain locus from BL-3 resulted in amplification of a fragment of approximately 1.1 kb products as expected. The product was cloned into a TA vector and sequenced. This confirmed that a rearranged heavy chain locus was present in the genome of BL-3 cells.

5.3.2: Conservation of IgH sequence in culture

Some cultured B cell lines display a constitutive mutator phenotype in which substitutions are introduced into the rearranged Ig loci without external stimulation e.g. Ramos cell line (Faili *et al.*, 2002b). Other cells are more stable but mutation of the Ig loci can be induced through stimulation in culture e.g. BL-2 cell line when stimulated with antibodies to CD19, CD21 and IgM simultaneously (Faili *et al.*, 2002b). LPS is a potent polyclonal activator of B cells. As a mitogen, LPS is routinely used in culture at 50 µg/ml (Andersson *et al.*, 1979). It has been shown to increase AID mRNA in murine spleen cells in dose-dependent fashion leading to class switching (Park *et al.*, 2005). To establish the behaviour of BL-3 cells, the rearranged IgH locus was recovered by PCR at pre-defined time points following addition of LPS (30 µg/ml) or from unstimulated cultures. Products were cloned into TA vector and inserts screened with *Bst* NI for rapid preliminary assessment of mutations. The Figure 38 below depicts the restriction profile of amplicons recovered from LPS-treated and

untreated BL-3 cells after 90 min, 24 hours and 48 hours. Two restriction patterns were observed but these appeared in both stimulated and unstimulated cultures. Identical patterns were observed at 48 hours (data not shown).

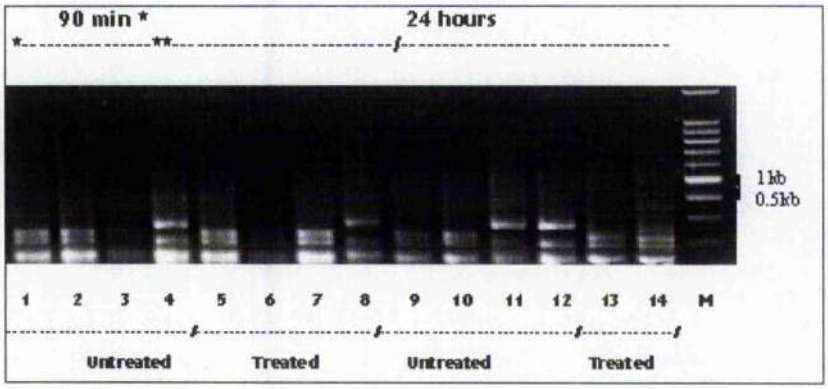


Figure 38: Restriction analysis of IgH amplicons from BL-3 cells

PCR products from the rearranged IgH locus were cloned, and individual inserts were re-isolated and digested with *Bst*NI. Digests were analysed by agarose gel electrophoresis. Two patterns (* and **) emerged. Lanes (1-8): samples from BL-3 cells after 90 minute of culture in the absence (1-4) or presence (5-8) of LPS. Lanes (9-14): samples after 24 hours culture in absence (9-12) or presence of LPS (13 and 14). M: DNA marker.

No evidence emerged of a shift from one pattern to the other, or emergence of new patterns during culture in the presence or absence of LPS, or when late time points were compared to the start of the experiment. Except for the appearance of a few large cells and some apoptotic bodies in the stimulated cultures, there were no obvious differences between the two treatment groups.

To provide more detailed analysis, several amplicons chosen on the *Bst* NI profile were sequenced (Figure 39). The mutations were recorded at a rate of $0.8x10^{-4}$ mutations/base/pcr cycle which was nearly same as Taq error rate of $1.04x10^{-4}$ mutations/base/pcr cycle recorded under present experimental conditions. The data thus did not suggest on going or LPS-dependent diversification was taking place with significant frequency.

5.3.3: Flow cytometry analysis

BL-3 cells were tested for the presence of B cell surface markers using flow cytometry. The cells stained successfully with anti-CD40 antibodies (Figure 40a) and bovine anti bovine-IgM (Figure 40b). The IgM staining was further verified by immunofluorescence microscopy (Figure 40c).

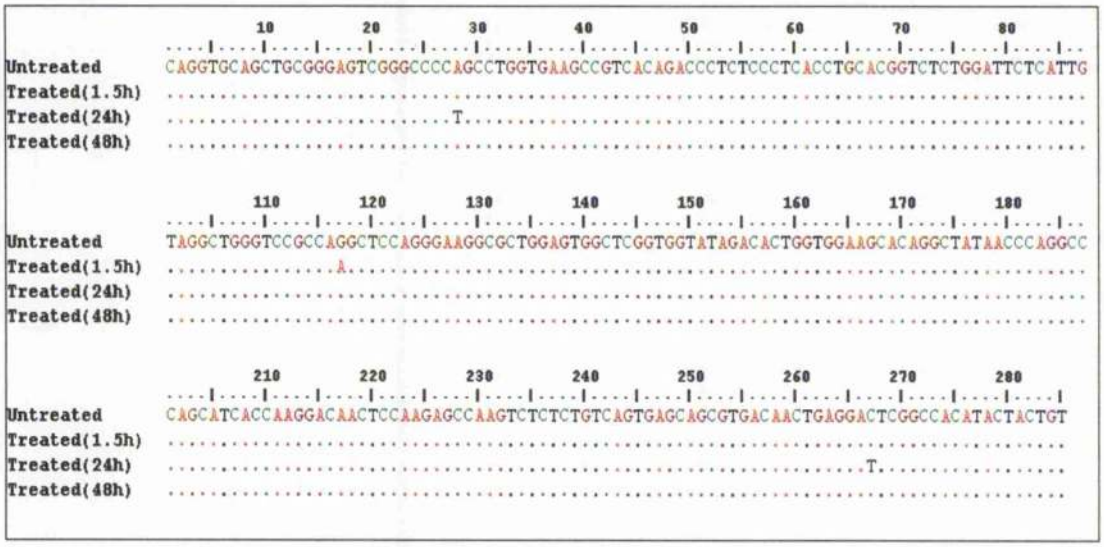


Figure 39: Sequence alignment of amplicons from rearranged IgH locus recovered from LPS-treated and untreated BL-3 cells

Samples were recovered from BL-3 cells after 90 min, 24 hours and 48 hours culture in the presence of LPS, as indicated. The region shown corresponds to FR1 to FR3.

This revealed a ring-like pattern suggestive of the aggregation of IgM receptor after binding of anti-IgM antibody. However, the binding of anti-CD21 and anti-IgG antibodies to BL-3 cells could not be detected in FACS analysis (data not shown).

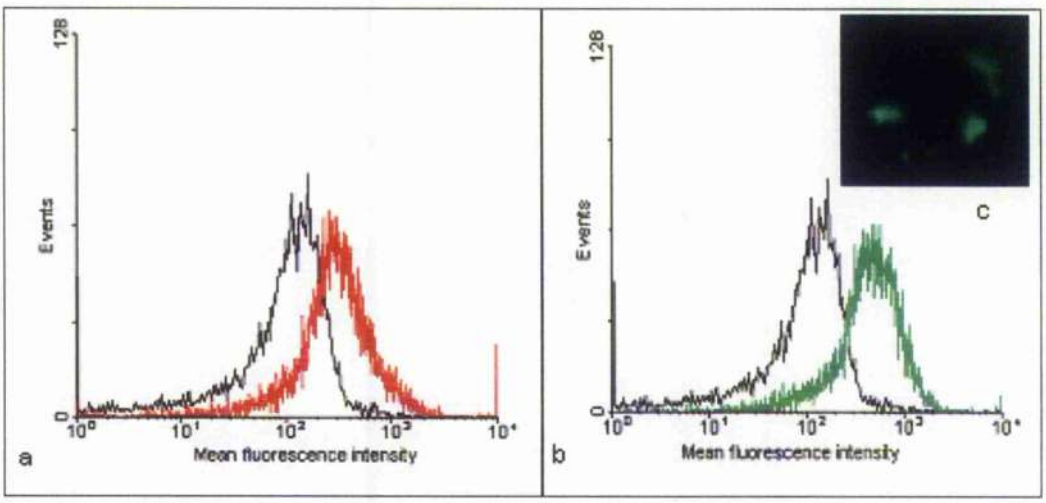


Figure 40: Detection of CD40 and IgM at the surface of BL-3 cells

BL-3 cells were labelled with antibodies to CD40 (a) or IgM (b) followed by FITC conjugate and analysed by flow cytometry. Control samples lacking primary antibody are indicated in black. The presence of surface IgM was also demonstrated in immunofluorescence microscopy; a single labelled BL-3 cell is shown in (c).

5.3.4: Expression of AID in BL-3 cells

The nested PCR reaction described earlier in this thesis provided a sensitive assay for the presence of AID transcripts in cDNA isolated from BL-3 cells. The amplicon obtained from the second round of amplification was 131 bp in size. An AID cDNA cloned into pBAD (Invitrogen) provided positive controls for the reaction. A GAPDH PCR reaction (description in Chapter 4) was used to confirm successful synthesis of cDNA from BL-3 cells. No AID mRNA transcripts were recovered from untreated BL-3 cells. The exposure of BL-3 cells to B-cell mitogen, LPS, IL-4, anti-CD40 alone or different combinations did not result in any upregulation of AID mRNA transcripts even after using nested PCR approach.

It was thus apparent that BL-3 cells did not express AID to a level detectable with this assay. Successful amplification from GAPDH transcripts confirmed that synthesis of cDNA from BL-3 cells had taken place and that inhibitory compounds were not present to an appreciable level. The result with AID primers, using spleen cDNA, demonstrated the success of the assay method. While the sensitivity of the AID assay was not determined, the complexity of a cDNA preparation from spleen (multiple cell types, multiple states of B cell activation) suggests it is high. The culture of BL-3 cells in the presence of LPS for up to 48 hours did not appear to activate the expression of AID, although all control reactions with GAPDH were successful. Similarly, addition of bovine IL-4 to cultures of BL-3 cells at 10 ng/ml, the cross-linking of surface IgM, or addition of anti-CD40 antibody, treatments that were applied singly and in combination, failed to activate AID expression.

Finally, primers against AID were redesigned to provide a perfect match to the bovine cDNA. The sequence of these modified primers was, Bt AID F1 (for): 5'CCAGTTCAAAAACGTGCGC3' and Bt AID R1 (rev): 5' CGCAAGTCATCAACCTCGTAGAG 3'. From bovine spleen cDNA, the predicted size of the amplicon was 547 bp. Using these reagents, AID transcripts could be detected in bovine spleen cDNA in a single PCR reaction and the need for a nested, second amplification was eliminated. In collaboration with a postgraduate student, Zhu Jie, cDNA samples from stimulated and non-

stimulated BL-3 cells were tested for evidence of AID expression but no amplification took place.

5.3.5: Introduction of transgenes to BL-3 cells

Since conditions could not be identified that would stimulate expression of the endogenous AID gene in BL-3 cells, an alternative strategy was planned: to transfect into cells an AID transgene under the control of a constitutive, viral promoter such as the CMV-IE promoter carried in vectors such as pcDNA. Assay for transgene expression with the AID PCR described earlier was possible, but inconvenient. To identify an appropriate, high efficiency transfection method, experiments with GFP were carried out since expression of the GFP reporter could be assessed in real time by fluorescence microscopy.

In initial experiments a chemical transfection method was tested with BL-3 cells using pEGFP-1, plasmid from which GFP is expressed. The BL-3 cells were transfected with pEGFP-1 plasmid using Lipofectamine 2000 and were examined at 12-hours intervals for 72 hours to check for the expression of GFP. The transfection efficiency was determined by counting the number of fluorescent cells as a proportion of total cell number at 24 and 48 hours post-transfection Around 2-3% in BL-3 cells were transfected (Figure 41a) as compared to 70-80% COS-7 cells (Figure 41b) which were included as a positive control. At 72 hours post-transfection, no fluorescent BL-3 cells were detectable.

The failure to introduce GFP into BL-3 cells with high efficiency prompted experiments with viral vector systems. Initial experiments used adenoviral transduction with non-CsCl-purified Ad-GFP. Inspection of BL-3 cells at pre-defined time points post-infection did not reveal any evidence of fluorescence. As before, COS-7 cells were included as a positive control and proved very easy to infect with adenovirus. Combined use of Lipofectamine and adenovirus did not increase the transfection efficiency for BL-3 cells.

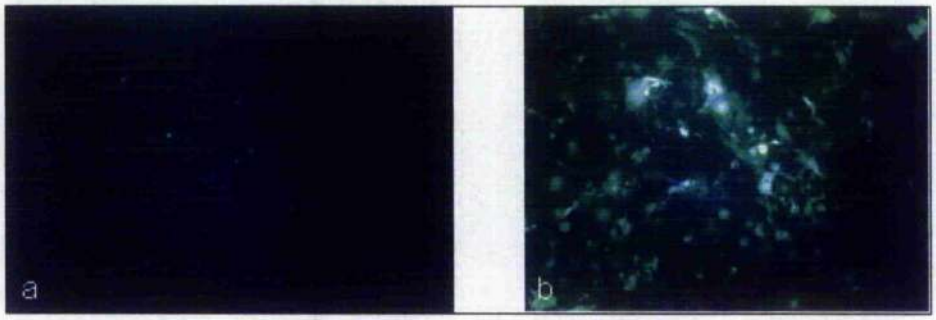


Figure 41: Lipofectamine- mediated transfection of cells in culture

pEGFP-1 was transfected in BL-3 (a) and COS-7 cells (b) using standard protocols. Images were gathered at 24 hours post-transfection using Inverted Zeiss Axiovert fluorescent microscope and 20 X magnifications.

Lentiviral vectors have attracted interest recently for the genetic modification of livestock animals (Clark and Whitelaw, 2003; Hofmann *et al.*, 2003; Whitelaw *et al.*, 2004) and it therefore seemed plausible that they might overcome problems encountered in transfection of BL-3 cells to this point. BL-3 cells were transfected with lentivirus carrying a PGK-GFP cassette and as a positive control; the Hapa 1 cell line was used. As before BL-3 cells proved resistant to the introduction of transgene as opposed to the Hapa 1 cells which were readily transduced (Figure 42).

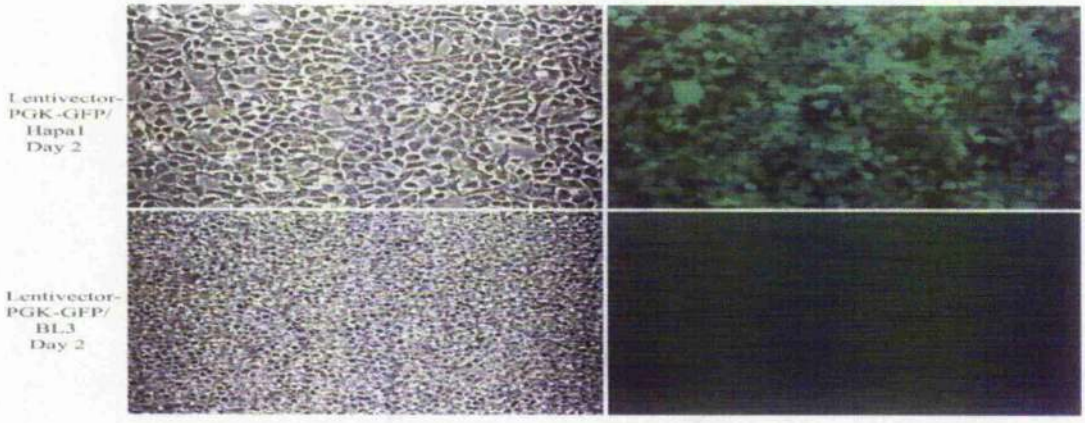


Figure 42: Lentivirus-mediated transfection of cells in culture

Viruses carrying a PGK-GFP cassette were mixed with Hapa 1 cells (upper row) or BL-3 cells (lower row). 48 hours post-infection, cell cultures were inspected by phase contrast microscopy (left panels) or by UV fluorescence microscopy (right panels) for evidence of GFP expression.

Cell viability was not significantly altered by lentiviral infection as determined by Trypan blue exclusion.

5.4: DISCUSSION

The goal of this component of the project was to characterize the bovine lymphosarcoma line BL-3 and establish whether it could be used for studies of Ig diversification, *in vitro*. The hope was that it would show evidence of a naïve or mature B cell phenotype and that SHM, GC or CSR could be activated by manipulation of the cellular levels of AID. Potentially, this might provide evidence in support of earlier conclusions that SHM is active during IgH diversification in Bovine B cells and is a major driving force in maturation of the humoral repertoire in cattle (Chapter 3). The advantage of an *in vitro* system would be that these processes could be studied in the absence of external antigen, maternal factors, and other confounding variables.

At the beginning of this project, there was little published information on the nature of markers associated with BL-3 cells. Dr. Christian Menge and colleagues from the Institute for Hygiene and Infections of Animals, Frankfurt, Germany reported that BL-3 did not bind antibodies to the T cell markers WC-1, CD2, CD8 and CD25. On the other hand the cells gave strong signals when stained with antibodies to MHC-I and MHC- II, CD11b. There is independent evidence that BL-3 line does not express CD25, the low affinity IL-2 receptor (Taylor *et al.*, 1992). Taken together, these general properties suggest the cells are unlikely to be of T cell origin. Evidence that the BL-3 line is of B cell origin can be taken from work by (Moreau et al., 1999); who showed expression of IgΛ transcripts and expression of surface IgM by flow cytometry. More recently Neill and colleagues at the National Animal Disease Centre in Iowa, USA have used Serial Analysis of Gene Expression (SAGE) to analyse BL-3 cells with and without infection with Bovine Viral Diarrohea Virus (BVDV) (personal communication). They were unable to detect tags for CD21, CD40 or IgD, and IL-4R was only positive after BVDV infection. The same group recently published a (SAGE) analysis of IPP B cells (Neill *et al.*, 2006). Communication with the authors indicated that the profile was very similar to that obtained for BL-3 cells but with some differences in expression levels. The inability to detect tags for IgD would suggest that BL-3 cells have a mature phenotype but this contradicted by the absence of SAGE tags for CD21 which would indicate the naïve phenotype of B lymphocytes. The results of this study substantiate and take forward this

profiling of the BL-3 cell line. Initial experiments addressed rearrangement at the IgH locus.

The successful recovery of the rearranged heavy chain sequence from BL-3 cells suggested that the cells had undergone successful recombination. The size and sequence of the product was consistent with amplicons from bovine lymphoid tissues indicating an open reading frame, and potentially functional translation product. These conclusions were substantiated by the presence of IgM at the cell surface as detected by flow cytometry. This and the presence of CD40 confirmed the B lymphocyte origins of the cell line. Cytometry was unable the detect IgG or CD-21 at the BL-3 surface. The presence of IL-4Rα was confirmed by RT-PCR. BL-3 cells also gave positive reaction when stained for IL-4Rα with recombinant bovine IL-4, albeit of a low intensity. Overall, the data do not provide conclusive evidence for the developmental stage of these cells.

There was no evidence that BL-3 cells express AID constitutively. A sensitive PCR assay for the AID transcript was consistently negative and restriction profiling of amplicons for the rearranged IgH locus showed no change during culture of the cells. This contrasts with the human Ramos B-cell line (a Burkitt's lymphoma cell line), in which hypermutation is constitutive. The cell lines CL-01 and BL-2 can be induced to express AID and undergo SHM (Denepoux *et al.*, 1997). Inducible systems are well-suited for studying Ig modification and diversification as they offer the prospects of studying the multitude of complex events that ultimately leads to mutation. Mutations observed in such cell lines have similar characteristics to the ones that occur *in vivo* i.e. specificity for rearranged V regions and with mutation frequencies close to background amounts in the constant region, more transitions as compared to transversions and targeting of specific motifs (Denepoux *et al.*, 1997).

The goal for further work with BL-3 cells was therefore to attempt to activate AID expression. Initial experiments exploited the biological activity of LPS. LPS is a T-independent type 1 antigen and a well known B-cell mitogen. This binds to LPS-binding protein and CD14. At high concentration, it causes the proliferation and differentiation of most B cells regardless of their antigenic specificity (Charles Janeway *et al.*, 2004; Janeway *et al.*, 2004; Kearney and Lawton, 1975). There are suggestions that LPS is indirectly involved in Ig CSR by

enhancing COX-2 expression through protein kinase C resulting in PGE2 synthesis (Lee *et al.*, 2003) which is known to supplement CSR (Roper *et al.*, 1990). LPS is capable of activating B cells in the absence of MHC class II-restricted T-cell help. Other antigens that are non-mitogenic but share this property with LPS activate B cells by cross-linking surface immunoglobulin via repeating epitopes (Mond *et al.*, 1995). Park and his colleagues (Park *et al.*, 2005) explored the effect of LPS on the expression of AID during B cell differentiation, and the role of AID in IgA isotype expression by normal spleen B cells. LPS was able to increase AID transcription up to 48 hrs post-stimulation, around the time that Ig CSR become evident. Their data indicated that LPS contributes to TGFβ1-induced IgA isotype expression by stimulating AID transcription before CSR and/or by enhancing the IgA secretion rate after CSR. This set a promising foundation for work with BL-3 cells. BL-3 cells were thus treated with high doses of LPS (30-50 μg/ml) to enable it to act as a polyclonal activator. Even at these levels, LPS did not induce AID expression and analysis of the rearranged IgH locus by *Bst*N1 restriction analysis and sequencing revealed no evidence of sequence variation between treated and untreated groups.

The generation of humoral responses to T-cell dependent antigens is thought not to require antigen cross-linking of surface immunoglobulin (Tony *et al.*, 1985). Instead, these responses are dependent upon direct interactions between cell-surface receptors like CD40-CD40 ligand and others (Foy *et al.*, 1996). Recent studies have shown that CD40 ligation (conveniently mimicked by the use of anti-CD40 antibodies that aggregate CD40 at the B cell surface) and IL-4 increased the expression of AID via NF-κB and the STAT pathway in human and mouse B cells (Dedeoglu *et al.*, 2004). CD40 ligation using anti-CD40 mAb induced weak or no AID mRNA expression; however CD40 ligation with IL-4 resulted in synergistic induction of AID mRNA. The mean fold induction of AID mRNA expression of three independent experiments was 8.4-fold with IL-4 alone, 2.5-fold with anti-CD40 alone and 61.5-fold through ligation of with CD40 in the presence of IL-4.

Although cross-linking of surface Ig is not a requirement for humoral responses to T-dependent antigen, it can provide a stimulus to B cells. A human B cell line BL-2, which closely resembles centroblasts, has been shown to

introduce mutations into the rearranged V_H gene at a frequency close to $7x10^{-4}$ /bp/round of cell division after aggregation of its IgM receptor with anti-IgM antibody and co-culture with T helper cells (Denepoux *et al.*, 1997). The effect was dependent upon the state of cell differentiation as the naïve B cell-derived line ML1 was unresponsive. This is consistent with events *in vivo*: the somatic mutation machinery is activated only at the centroblast stage in mice and humans (Liu *et al.*, 1996). BL-2 cells could be activated in this way with IL-4 and CD40 ligation, or by IL-4 alone, and it appears that AID gene expression in this line is independent of CD40 ligation (Dedeoglu *et al.*, 2004). The same cell line has been studied by Reynaud and colleagues who substituted the T cell-derived signals with antibodies to CD19, CD21 and IgM. They showed that mutation arose as a post-transcriptional event within 90 min of treatment and was absolutely dependent upon AID expression (Faili *et al.*, 2002b).

While these studies focused on SHM, it is equally apparent that CSR can be manipulated through cross-linking of markers at the B cell surface and the cytokine environment. It has been shown that precursor B cells from a range of immunodeficient mice can be induced to transcribe the IgE H chain and undergo Sμ-Sε switch recombination by treatment with anti-CD40 antibody and IL-4 (Rolink *et al.*, 1996). It has also been shown that a IgM+ lymphoma cell line, CH12.LX, can be activated to switch from surface IgM+ cells to surface IgA+ cells at a high frequency (50%) after 72 hrs stimulation with IL-4, TGF-β and CD40L (Nakamura *et al.*, 1996).

These studies provided the foundation for experiments with bovine BL-3 cells, the presence of IgM, CD40 and IL-4Rα having been confirmed. However, BL-3 failed to upregulate AID after treatment with multiple combinations of LPS, IL-4, anti-IgM, and anti-CD40. Since all appropriate receptors were shown to be present, and the concentrations of the reagents were appropriate, it appears that BL-3 has differentiated to a non-responsive state, perhaps as part of its transformed phenotype.

The failure of BL-3 cells to up-regulate AID expression after treatment with recombinant IL-4 and anti-CD40 and other reagents prompted an alternative strategy. This was to establish a method to transfect AID into the BL-3 cells. Lipid mediated transfection using Lipofectamine was selected due to its

simplicity, low toxicity and its ability to effectively deliver DNA to many cell types (Mannino and Gould-Fogerite, 1988). Once lipid-DNA complexes have formed, receptor-independent endocytosis has been shown to be a major pathway for entry into cells (Friend *et al.*, 1996). The limiting step for efficient gene transfer using this approach is the breakdown of DNA after fusion of lysosomes and the endosomes carrying the complexes (Zabner *et al.*, 1995). Although lipofectamine-mediated transfection of a GFP reporter into BL-3 cells was found to take place, the efficiency was very low, reaching only 2-3% of total cells. This was too low for meaningful analysis of the effects of introduction of AID.

Adenoviral-mediated gene delivery has the advantage that viruses can multiply to high titres and infect non-dividing cells. Adenoviral vectors have been shown to mediate high levels of gene transfer and expression in a wide range of target cells (Crystal *et al.*, 1994; Nielsen and Maneval, 1998). The success of the system is probably due in large measure to the common expression of the viral receptor, CD46 (membrane cofactor protein), a protein involved in the regulation of the complement cascade (Sirena *et al.*, 2004). BL-3 cells were found to be completely resistant when transduced with Ad-GFP at different MOI. Although adenovirus offers several potential advantages as gene transfer vectors, some haematopoietic cells, particularly lymphoid cells, have shown to be relatively resistant to adenovirus-mediated gene transfer. Stewart and his colleagues (Prince *et al.*, 1998) tested three lymphoma cell lines (OCI-Ly2, OCI-Ly13.2, and OCI-Ly17) and found them to be significantly less sensitive to adenovirus infection, with relatively low efficiencies of gene transfer even using high adenoviral titres.

The combination of adenovirus with lipofectamine has proved useful for the transfection of plasmid DNA into cultured adipocytes, muscle, hepatocytic cells and lymphoma cells which are almost impervious to efficient transfection by other means (Buttgereit *et al.*, 2000; Meunier-Durmort *et al.*, 1996). This approach exploits the ability of adenoviruses to lyse the endosomal membrane, directing DNA to the cytosol. Potentially, this could overcome breakdown of DNA in conventional lipid complexes following the fusion of endosomes and lysosomes. It has been demonstrated in adipose cells that a complex between

adenovirus, lipofectamine and plasmid DNA enters via an endocytic process which probably involves the adenovirus receptor and a cellular integrin (Meunier-Durmort *et al.*, 1997). Cationic liposomes had been previously shown to increase the delivery of adenoviral vectors, resulting in a 20-fold increase in vector-encoded transgene expression in smooth muscle cells (Qiu *et al.*, 1998). However, the combination of lipofectamine and adenoviral-mediated DNA delivery did not result in any improvement in the transfection efficiency for BL-3 cells.

Lentiviruses are attractive as vehicles for the delivery of transgenes, as they can bring about the transduction of proliferating and quiescent cells. The sustained expression of the transgene can be achieved by stable integration of the vector into the host cell genome. They can deliver a significant amount of genetic information into the genome of target cells and are one of the most efficient methods for gene delivery (Miyoshi *et al.*, 1997; Naldini *et al.*, 1996). The use of replication defective lentivirus vectors is ideal when the goal is stable transduction and genetic modification of the target cell. Again, BL-3 cells proved problematic and lentiviruses carrying a PGK-GFP cassette did not achieve transduction.

Finally, electroporation was attempted. During the process of electroporation, electric pulses create transient pores in the membrane structure of the treated cell, opening avenues for entry of DNA (Chu *et al.*, 1987). This method has been exploited for the introduction of cloned DNA molecules into a variety of cells in order to study the expression and analysis of gene products (Andreason and Evans, 1989; Baum *et al.*, 1994; Graham and van der Eb, 1973; Prausnitz *et al.*, 1993; Ramirez *et al.*, 1998). Given the ease with which electroporation can deliver DNA to cells growing in suspension, it was decided to attempt transfection of the GFP reporter into BL-3 cells.

Anderson and colleagues (Anderson *et al.*, 1991) have reported the electroporation efficiency of lymphoid cells was comparable to that reported for other cell types. These authors further observed that the two most important electrical parameters were high capacitance (960 μF) and moderate decay constants in the range of 10-15 ms. The optimal field strength depended on the cell line, but were within 0.6 to 1 kV/cm. For experiments with BL-3 and COS-7

cells, the parameters used were within this range. This was the only transfection protocol used in which gene transfer to the control cell line - COS-7 in this instance - was as unsuccessful as transfer to BL-3. Given the constraints of the time at this phase of the project, only one set of experimental conditions was tried and thus it is not possible to say with confidence that electroporation can be ruled out as a method to deliver transgenes to BL-3 cells. The approach merits further systematic investigation.

Amongst many newer approaches to deliver foreign DNA to target cells is one termed nanotube spearing, described by Cai and co-workers (Cai *et al.*, 2005). This is based on the penetration of nickel-embedded nanotubes into cell membranes by magnetic field driving. These authors were able to achieve exceptionally high transduction efficiencies in Bal17 B-lymphoma cells, *ex vivo* B cells, and primary neurons. Cell viability after transduction remained high. Given that this method does not require a surface receptor, or transit of the DNA through a potentially damaging intracellular compartment, it is also worth closer investigation.

Although attempts to activate endogenous expression of AID in BL-3 cells were unsuccessful, this may still be possible by exploiting the innate immune response (Medzhitov, 2001). Microbial products such as LPS, lipoteichoic acids and peptidoglycan are not made by eukaryotic cells and therefore their recognition by the innate immune system of the host can signal the presence of infection (Janeway, 1989, , 1992). T-cell independent (TI) antibody production can occur in response to microbial antigens with repetitive structure such as bacterial LPS (TI-type 1) and capsular polysaccharides (TI-type 2). Also it has been shown that pathogen associated molecular patterns (PAMPs) (Medzhitov, 2001) are capable of stimulating B cells by cross-linking BCRs encoded from a restricted set of V(D)J germline genes with few mutations (Bendelac *et al.*, 2001). B-cells can also be activated by PAMPs through pattern recognition receptors such as those belonging to the TLR family (Graham and van der Eb, 1973; Prausnitz *et al.*, 1993; Ramirez *et al.*, 1998). TLRs comprise a family of type I transmembrane receptors characterized by extra cellular leucine rich repeat domain (LRR) and an intracellular Toll/IL-1 receptor (TIR) domain (Hashimoto *et al.*, 1988; Janeway, 1992). LRRs are involved in ligand recognition

and signal transduction (Kobe and Deisenhofer, 1995). The TIR domain of Toll proteins is a conserved protein-protein interaction assembly found in many transmembrane and cytoplasmic proteins with a role in host defence (Aravind *et al.*, 2001). The requirement of TLR signaling in B cells during the response to T-cell dependent antigens (Pasare and Medzhitov, 2005) suggests the importance of these molecules for optimal production of antibody. TLR-induced dendritic cell maturation is required for activation of helper T-cells (Banchereau and Steinman, 1998), which in turn activate B cells through CD40-CD40L interaction and cytokine secretion (Bird, 2006). However, it was found that the activation and differentiation of helper T cells alone was not sufficient for induction of T-dependent B-cell responses - the process also required activation of TLRs in B cells (Pasare and Medzhitov, 2005). This was investigated by assessing the response of μMT mice (these animals lack mature B cells) to T-cell dependent and T-independent antigen. The μMT mice were reconstituted with B cells from MyD88 KO or TLR4 KO donor mice. Those μMT mice that received wild-type B cells responded strongly to both classes of antigen. Mice that received TLR4- or MyD88-deficient B cells responded much less successfully, even when challenged with T-dependent antigen. Further analysis of B cell responses to the T-cell dependent antigen HEL was also carried out to determine at which stage TLR signaling in B cells was involved. Using B cells expressing a transgenic BCR specific for HEL, it was shown that BCR-mediated uptake of HEL was increased in the presence of LPS and the mice that received TLR4- or MyD88-deficient B cell had fewer germinal centres than mice that received wild-type cells. These studies illustrate the overlap between innate and adaptive responses but also that B cell stimulation through TLRs is important for humoral immunity.

The cross-linking of TLR4 by LPS stimulates DCs to up-regulate B cell-activating factor (BAFF) (Nardelli *et al.*, 2001), which along with signals from the BCR, stimulates B cell survival, CSR, and Ab production (Litinskiy *et al.*, 2002; Mackay *et al.*, 2003). PAMPs stimulate B cells not only through BAFF but also directly via TLRs present on B cells. This has been shown by LPS engagement of TLR4 which subsequently led mouse B cells to undergo CSR and Ab production (Stavnezer, 1996). Human B cells have been shown to lack TLR4 but express TLR9 (Krieg, 2002) which is an intracellular pattern recognition receptor and has the ability to detect CpG DNA from viruses and bacteria (Hemmi *et al.*, 2000).

The studies have demonstrated that CpG DNA triggers B cell proliferation and not only TI IgM production (Krieg *et al.*, 1995), but also IgG CSR in mice in the presence of IL-10 (He *et al.*, 2004). Given the dependence of CSR upon AID, it thus seems likely that signalling via a range of TLRs may be able to activate AID expression in B lymphocytes. In collaborative studies with a Masters student, it was established that BL-3 cells express TLR4, and TLR9 but not TLR7 receptor. This was addressed by PCR with cDNA from BL-3 cells, and TLR-specific primer sets (Zhu Jie and Verma, unpublished data; Zhu Jie, MRes dissertation, University of Glasgow, 2006). It is not known if these molecules are present in BL-3 cells in a functional form. The PAMPs that activate cells through these receptors are well characterised and from previous sections, it seems likely that in other cells, AID induction is likely to take place. Whether these systems could be exploited in the future to induce AID and to mediate somatic hypermutation and CSR is still to be assessed.

In conclusion, these studies have established the rearranged heavy chain locus in BL-3 cells does not undergo constitutive mutation. Various approaches to induce AID in BL-3 cells were not fruitful thus limiting the potential use of this cell line to study immunoglobulin diversification *in vitro*. The cells also proved very hard to transfect, 2-3% transfection being the maximum that was achieved for eGFP if could be achieved at all. Although many lines of potential investigation with BL-3 cells remain, it may be that alternative strategies must be sought for the future. If reagents were available for the isolation of naïve and mature bovine B cells, they could then be cultured with appropriate stimuli to upregulate the molecules associated with affinity maturation and CSR including AID. Another approach would be to attempt to immortalise them using viral transformation e.g. bovine leukaemia virus infection (Murakami *et al.*, 1999) or through expression of telomerase reverse transcriptase to generate B cell lines as alternatives to BL-3. Another possible avenue would be to prepare heterohybridomas from bovine B lymphocytes taken from foetal lymphoid tissue or study the properties of B cells immortalised by infection with *Theileria* (Moreau *et al.*, 1999).

Chapter: 6 Recombinant antibodies against bovine IgM from phage display

6.1: INTRODUCTION

Following from the pioneering work of Smith (Smith, 1985), phage display technology has matured into a widely used technique for selecting peptides and proteins and its scope is expanding with the passage of time. It has become a powerful tool that can be used to map antibody epitopes, create vaccines and to engineer peptides, antibodies and other proteins as both diagnostic tools and as therapeutics. The display of millions of variants of ligands such as peptides and proteins on the surface of filamentous bacteriophage by fusion to pIII, a minor coat protein of phage (Smith, 1985), together with the generation of large macromolecular repertoires and use of powerful selection-amplification schemes, has led to the isolation of several classes of useful proteins. Filamentous bacteriophage like M13 replicate and assemble without killing the host cell in contrast to other phage (e.g., T4). The phage are genetically engineered so that a particular moiety (e.g. an antibody fragment) is fused to a phage capsid protein (e.g. pIII) and the gene encoding the displayed antibody is also contained within bacteriophage. This technology thus couples the displayed phenotype with its genotype, allowing the gene that codes for the displayed protein to be retrieved easily for downstream applications. To select the phage with affinity towards a target system, the target is attached to a solid surface and the phage are allowed to bind (Figure 43). Those phage that recognize the target molecule bind tightly and the non-binding phage are simply washed away. The DNA contained within the recovered phage then can be used to produce more of the displayed protein for use in research or diagnostics.

6.1.1: Antibody libraries and phage display

Antibody display libraries have been generated in phage vectors in many formats. **Immune libraries** are useful in isolating clinically useful antibodies against diseases experienced by the tissue donor. This kind of library is created by PCR using cDNA from immunised donors. Primer sets are designed to recover the active Ig repertoire by annealing to the coding sequences for framework

residues and, in the case of libraries of Fab fragments, the constant domains of Ig heavy and light chain.

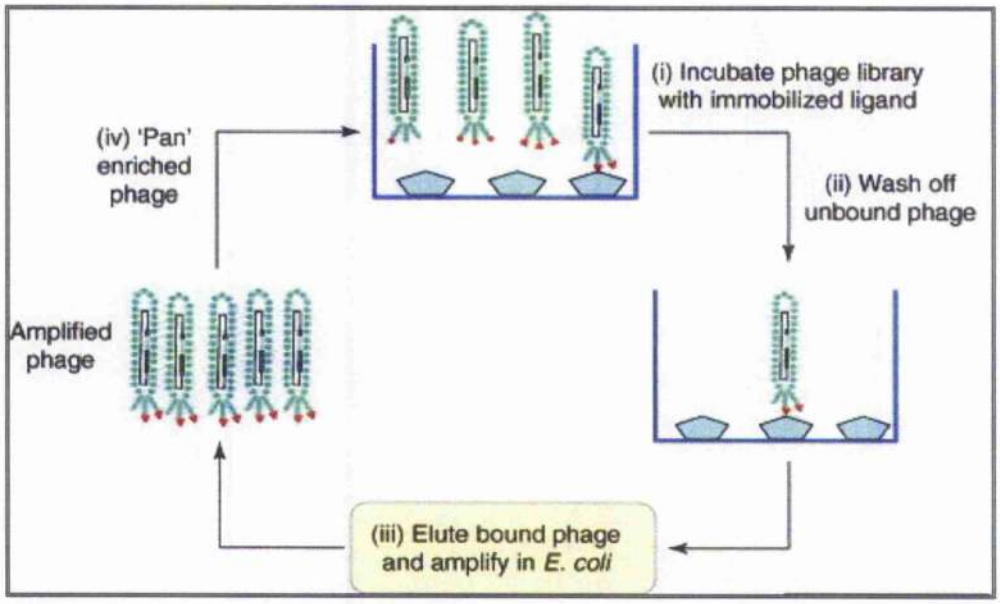


Figure 43: A cycle of affinity selection from a filamentous phage display library (Mullen *et al.*, 2006)

One advantage of immune libraries is that they are enriched for antibodies that bind to the immunogen of interest. In addition, the antibodies will have undergone affinity maturation in the immunised individuals hence may bind the target with high affinity. There are drawbacks associated with these libraries. As the composition of the library is heavily influenced by the host response, it may lack antibodies with the desired characteristics if, for example, the target is poorly immunogenic *in vivo* or the host response is biased by immunodominant antigens. Moreover, tolerance mechanisms make it difficult to isolate antibodies against self antigens. Additional complications impact upon the generation of libraries from immune human donors. The peripheral blood is easy to sample but contains small numbers of B cells; bone marrow and lymphoid tissues are better sources of B cells but are difficult to obtain. Another drawback is that screens against widely differing antigens require a new phage library to be made for each selection and libraries can take more than 3 months to construct.

Naive libraries are derived from naturally rearranged V genes from nonimmunised donors. The V genes can be amplified either from IgM mRNA (Marks *et al.*, 1991) or total mRNA (Vaughan *et al.*, 1996), libraries made from IgG mRNA being biased due to the donor immune responses (Marks *et al.*, 1991).

Naïve libraries provide antibody fragments with binding activities against many different antigens (Marks *et al.*, 1991). The obvious advantage over an immune library is that no immunised donor is needed and thus the time needed to generate the library is significantly reduced. If sufficiently large, naive libraries can yield antibodies against self antigens and toxic molecules. However, the exact nature of the V-gene repertoire contained in the library is poorly defined and poor expression and toxicity to the host bacteria is also a matter of concern since some antibodies express better than others. There is also a need to generate a large library if antibodies with high affinities are to be isolated (Vaughan *et al.*, 1996).

Semi-synthetic and **synthetic libraries** are constructed by *in vitro* manipulation of V-gene segments resulting in predetermined levels of randomisation in defined locations in CDRs and FRs. The location and degree of diversity may be chosen to correspond to areas of highest variability in a natural antibody repertoire. Synthetic antibody libraries have a significant advantage over libraries that use naturally rearranged V-genes. Importantly, the choice of V-gene segments may be guided by factors that will increase the overall performance of the library, such as good expression, successful folding and low toxicity in *E. coli,* factors that otherwise reduce the functional repertoire of the library. Large differences in V-gene usage both *in vivo* and in phage repertoires (Winter *et al.*, 1994) also suggest that some FRs may be better suited to the interaction with antigen than others. Second generation synthetic antibody libraries have been built using 'master' frameworks representing each of the Kabat subclasses to help choose only well expressed sequences. In preparing a library of this kind, the goal is usually to maximize the diversity of amino acids at pre-determined positions in the antibody molecule. However, the incomplete precision with which mononucleotides can be mixed, and the degeneracy of the genetic code means that only limited control over the distribution of amino acids is possible. Furthermore, undesired amino acids and stop codons often are difficult to avoid. To overcome the introduction of stop-codons during *in vitro* diversification, V-genes can be diversified by assembly from trinucleotides instead of single bases (Virnekas *et al.*, 1994). Tomlinson's laboratory has further enhanced the performance of such synthetic libraries by screening them for interaction with Ig-domain binding proteins (Protein A for V_H, Protein L for

Vκ, etc.) (Akerstrom *et al.*, 1994) to ensure the elimination of clones that carry stop codons and frameshifts. This simultaneously enriches for correctly folded V-domains.

6.1.2: Application of phage libraries

Phage libraries have been used to isolate high-affinity antibodies against very many targets without immunisation (Vaughan *et al.*, 1996; Winter *et al.*, 1994). Phage libraries also are useful in analysing natural humoral responses in patients suffering from autoimmune disease. In order to access the anti-human acetyl choline receptor (hAChR) repertoire, Graus and his colleagues (Graus *et al.*, 1997) constructed phage display Fab libraries of thymic lymphocytes from two Myasthenia Gravis (MG) patients. They were able to isolate four Fabs highly specific for huAChR which showed evidence of significant somatic mutations. Two of these Fabs were able to inhibit up to 90% of donor serum anti-huAChR Abs. Moreover, anti-huAChR Fabs were able to protect against AChR loss by antigenic modulation induced by MG serum Abs, suggesting a potential therapeutic role for these recombinant Fabs.

There are innumerable applications of this technology given its advantages of maintaining a link between phenotype and genotype, the enormous diversity of variant proteins displayed and the flexibility of selection that can be performed *in vitro*. These general principles need not benefit only the isolation of antibodies. DNA-binding proteins with altered DNA-binding specificity (Greisman and Pabo, 1997), improved hormones (Lowman and Wells, 1993) and enzyme inhibitors (Roberts *et al.*, 1992) have been isolated by phage display. Phage display is also used widely in various forms, including the use of fragment libraries of whole microbial genomes, to identify peptide-ligand and protein-ligand interactions that are of importance in infection (Mullen *et al.*, 2006). Antibody display remains however the area of primary application. In this study, the semi-synthetic Tomlinson I and J libraries (Center for Protein Engineering, Cambridge, UK) were used to select phage antibodies against bovine IgM variants coated on microtitre plates using a two-stage depletion / selection strategy.

6.1.3: Phage selection strategies

In order to select phage that carry antibodies fit for the intended purpose, choosing the right selection strategy is of equal importance as using the right type of library. The aim of all selection strategies is to remove unwanted phage antibodies and select those with highest specificity for the target. The conventional selection method of panning the phage on antigen-coated immunotubes requires purified antigen and works well for routine applications. However, antibodies generated through this technique may fail to bind to the native protein antigen and selecting for high affinity clones is difficult due to avidity effects and discriminating among clones with similar affinities is difficult, if not impossible (Griffiths and Duncan, 1998). Some alternative selection strategies are mentioned below.

6.1.4: Biotinylated antigens

The main problem observed when selecting for peptide-binding phage is the poor coating efficiency of some peptides and other small molecules to plastic and the reduced availability of potential phage binding sites. The high binding affinity of streptavidin to biotin can be exploited. The target is biotinylated, and interaction with phage takes place in solution. Complexes are then captured on streptavidin-coated paramagnetic beads (Hawkins *et al.*, 1992). The biotinylated Ags with bound phage are drawn out from the suspension by applying a magnet on the side of the tube. The beads are washed several times and the phage are eluted from the beads. To enable the selection of high affinity antibodies, this technique also allows precise control of the antigen concentration and the time of exposure of the antigen to the phage library. The interaction between antigen and phage antibody in the solution ensures that the full surface of the target is available for interaction and antibody fragments with low binding affinities and the tendency to form dimers are avoided (Schier *et al.*, 1996).

6.1.5: Competitive Deselection

In this procedure, the phage library is first pre-absorbed on the target of interest to remove phage that react with dominant sites. The unbound phage are then incubated a second time with target, eluted and amplified according to

conventional protocols. It has been noted that the removal of unwanted clones is only partial but the recovery of rare binding clones is enhanced considerably (Burioni *et al.*, 1998).

6.1.6: Epitope-masking

During phage selection, certain specificities may dominate the recovered population of phage as a result of immunodominance or because some clones have a selective advantage through high affinity interaction with the target during the panning process (Ditzel, 2002). Sometimes the goal of selection is to retrieve antibodies with specificities for minor epitopes and with different characteristics. Epitope masking relies on obstructing the interaction of phage with some epitopes on the target to drive the isolation of new antibody specificities. Monoclonal antibodies, or antigen-specific Fabs or scFv proteins retrieved by previous screens can be used to block dominant epitopes. Successive reselection of the phage library on the masked target refocuses the selection process towards other regions of the target.

6.1.7: Alternating selection

This protocol was described by Jun and Sloan (Lu and Sloan, 1999) to select phage displaying scFv specifically recognising a phosphorylated form of the transcription factor E47. In this method, the selection technique used in each round differs from the one used in the previous round. The authors claim that this technique removes phage selected against unwanted epitopes in one round of selection by providing a different basis for selection in the subsequent round of selection.

In this study, a **depletion and selection** protocol was used to isolate single chain antibodies recognizing minor epitopes unique to variants of bovine IgM. For this purpose the Tomlinson I and J semi-synthetic libraries of scFvs were used as a starting material.

6.1.8: Tomlinson scFv libraries

These libraries were created at Winter's lab at the MRC Laboratory of Molecular Biology and the MRC Centre for Protein Engineering in Cambridge. The libraries each comprise over 100 million different scFv fragments cloned in an

ampicillin resistant phagemid vector, pIT2 (Figures 44, 45). The vector was derived from pHEN1, one of the first purpose-made vectors for antibody phage display (Hoogenboom *et al.*, 1991). scFv fragments comprise a single polypeptide in which V_H and V_L domains are attached to one another by a flexible linker containing glycine and serine residues. Using these libraries, specific binders to target molecules can be selected by normal panning methods. Typically two or three rounds of selection are required to ensure that more than half of the scFvs in the selected population bind to the target molecule. The monoclonal scFvs can then be screened for binding and then used for further analysis of the target molecule. All the functional scFvs in the Tomlinson I and J libraries can bind Protein A from *Staphylococcus aureus*, an Ig-binding protein with five domains that mediate interaction with the Fc of most mammalian IgGs. Protein A also binds to the Fab region of a subset of Ig carrying heavy chains belonging to the V_HIII family (Harboe and Folling, 1974) through interaction with the V_H domain (Sasso *et al.*, 1989, , 1991). This explains the ability of Protein A to bind scFvs, providing the V_H component is derived from the V_H III family. The scFvs in the Tomlinson libraries can also interact with Protein L through the V_K domain (Bjorck, 1988). This allows either of these secondary reagents to be used for detection, purification or immobilisation. Alternatively, secondary reagents can be used that bind the myc or histidine tags fused to c-terminus of each scFv.

Both libraries are constructed on single human V_H (V3-23/DP-47 and J_H4b) and V_K (O12/O2/DPK9 and J_K1) frameworks. The canonical structures (V_H: 1-3, V_K :2-1-1) adopted by the chosen frameworks are by far the most common in the human antibody repertoire (de Wildt *et al.*, 2000). Diversity is incorporated at positions that make contacts to antigen in known structures and are naturally diverse in the mature antibody repertoire. This comprises a total of 18 residues - H50, H52, H52a, H53, H55, H56, H58, H95, H96, H97, H98, L50, L53, L91, L92, L93, L94 and L96. The CDR3 of the heavy chain was designed to be as short as possible yet still able to form an antigen binding surface. In library I, diversity is created through incorporation of DVT (D: A, G or T; V: A, C or G) codons at the designated positions. In library J, NNK (N: any base; K: G or T) codons are used instead of DVT. Both libraries are constructed in the vector pIT2 which provides his and myc tags. Both libraries contain about 1.4×10^8 clones and it is

estimated that 96% of clones in library I contain a complete insert whereas the value is somewhat lower (88%) for library J.

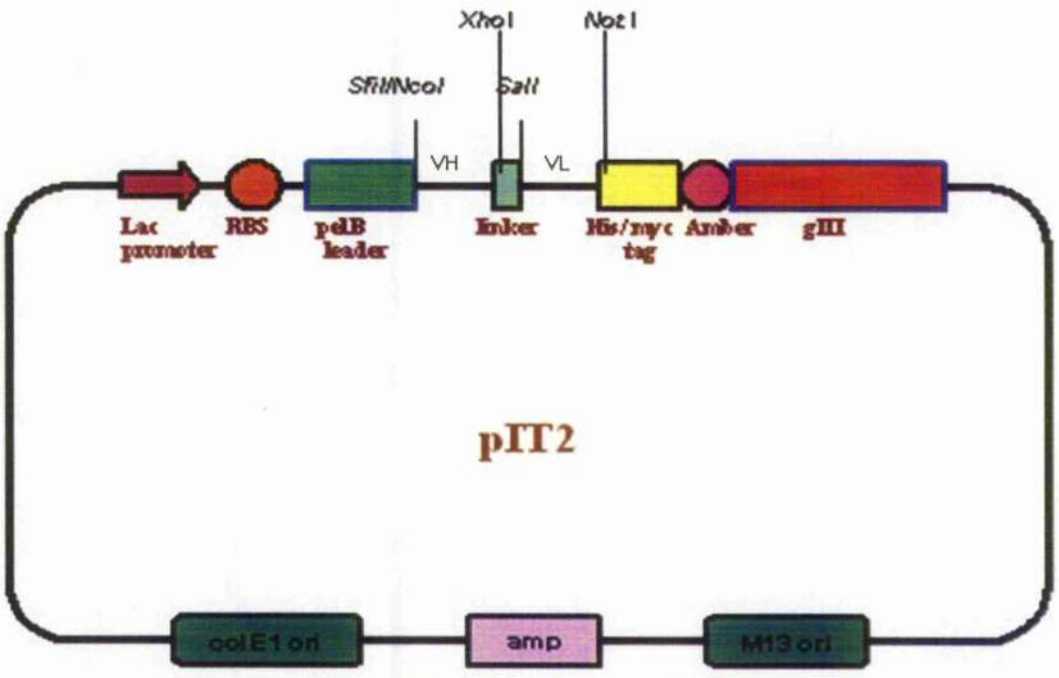


Figure 44: Structure of phagemid pIT2

The main features of the vector are indicated. Transcription takes place from a *lac* promoter and in a suppressor strain of *E. coli*, the translated product comprises a PelB leader, the single-chain antibody created by insertion of VH and VL sequences at the sites indicated, separated by a linker, peptide tags, and the phage pIII protein.

6.1.9: KM13 helper phage

The KM13 helper phage was developed by Kristensen and Winter (Goletz *et al.*, 2002) to overcome problems associated with other helper systems. In particular, K13 reduces the background of non-recombinant phage that are recovered during rounds of selection by phage display. Phage that are specifically bound to the selecting surface through the displayed scFv are recovered in a state infective to bacteria; phage that do not display scFv are inactivated and eliminated from further rounds of selection. In this helper phage, a trypsin cleavage site was introduced into the pIII between the second and third domains of the protein (Figure 46).

It has been shown that sequences can be attached to the amino terminus of pIII and that peptides can even be inserted between domains without abolishing infectivity (Endemann *et al.*, 1992; Krebber *et al.*, 1997; Smith,

1985). However, all three domains of pIII are essential for the successful infection of bacteria (Riechmann and Holliger, 1997).

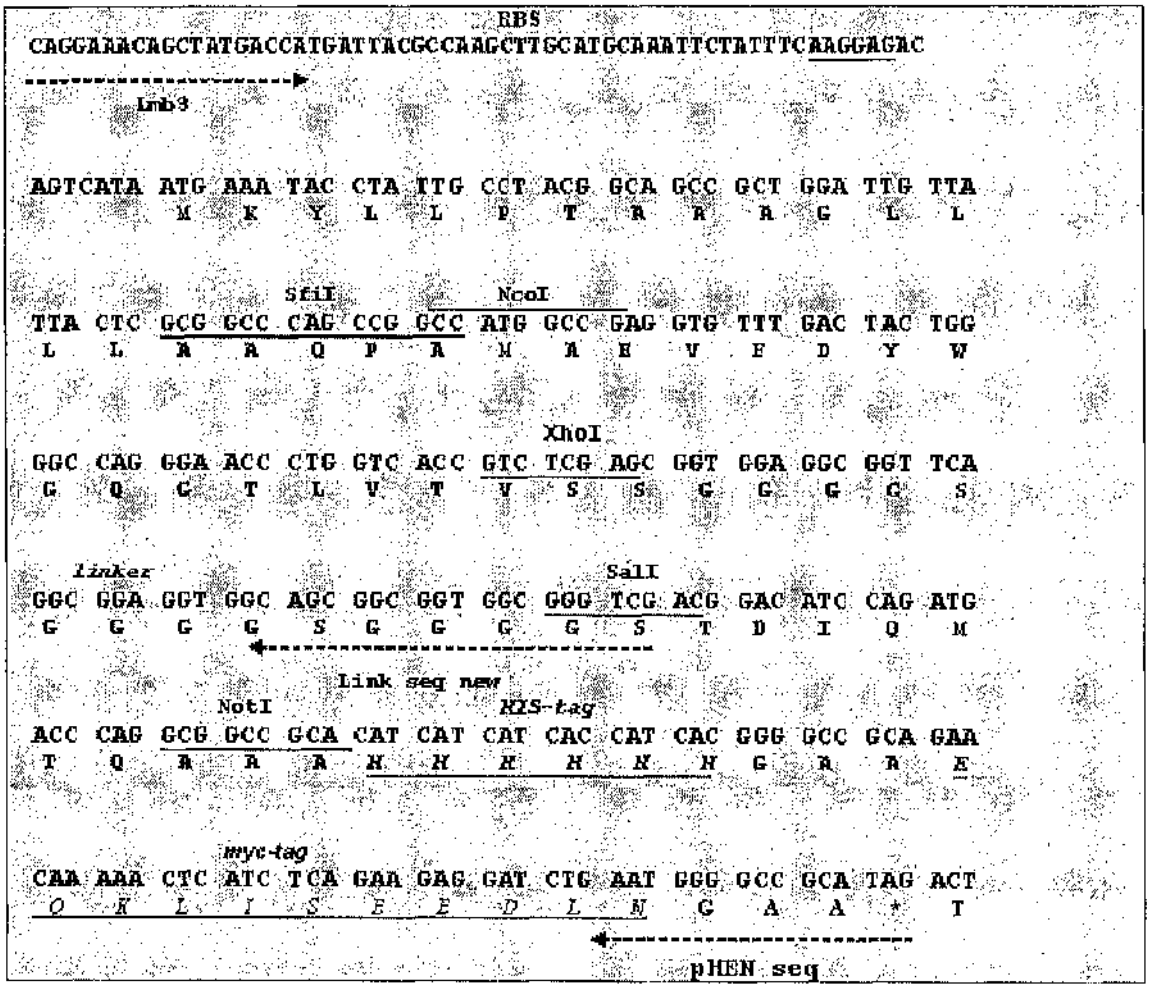


Figure 45: DNA sequence of scFv region of the pIT2 vector

The Figure presents part of the DNA sequence of the pIT2 vector. The restriction sites are underlined. The V_H sequence is located between *Nco*I and *Xho*I sites and the V_L sequence is located between *Sal*I and *Not*I sites. Primers for sequencing and recovery of inserts are shown by dotted arrows. The amber stop codon is presented as a star. The amino acid sequences of his and myc tags are underlined.

Therefore, trypsin treatment destroys the function of helper phage-derived pIII and hence phage carrying only pIII of this type are unable to infect bacteria. In contrast, the gene for pIII carried in pIT2 and to which scFv sequences are fused (Figure 46) does not encode trypsin cleavage sites. Trypsin treatment therefore does not alter the ability of this version of pIII to mediate infection of bacteria but the protease does cleave between the scFv and pIII, providing an easy and efficient method for release of infective phage from the target-coated surface used in selection. The method has been used in numerous

applications since its publication (de Wildt *et al.*, 2000; Demartis *et al.*, 1999; Goletz *et al.*, 2002; Riechmann and Winter, 2000).

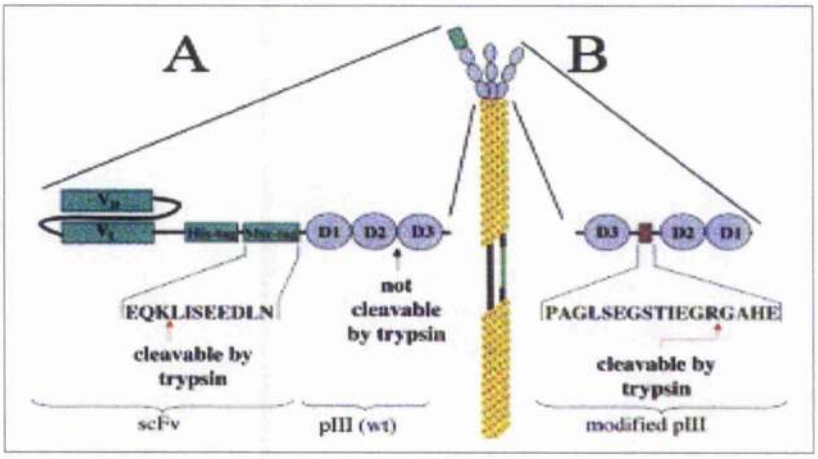


Figure 46: The proteolytic selection principle

The Figure shows scFv-phage in the middle (yellow) from a phagemid library prepared using KM13 as helper phage. During packaging and rescue of the library, the pIII from the helper phage (B) competes with the scFv-pIII fusion protein (A) for incorporation into phage particles. pIII, translated from the helper phage genome, carries trypsin-sensitive site between domains D2 and D3 (B). Trypsin treatment inactivates the function of these proteins. pIII translated from the phagemid (A) is insensitive to trypsin, but a cleavage site lies in the myc tag of the scFv-pIII (wt) fusion protein. This allows the proteolytic elution of infective phage from the selecting surface. This Figure has been adapted (Goletz *et al.*, 2002).

6.1.10: Application of Tomlinson I and J libraries

The Tomlinson I and J scFv libraries can be used to derive binders to almost any target molecule using phage display and selection techniques. These binders can be used in many different applications such as ELISA, Western blotting, FACS, immunochemistry etc. In one study, the libraries were used to isolate a scFv antibody against human fibrin clots. Soluble scFv was specific for human fibrin clots and did not show any reactivity against human fibrinogen in ELISA (Jun Peng Yan, 2004). Anti-idiotypic antibodies that mimic carbohydrate or conformational epitopes (Ab2β) are of considerable interest as surrogate immunogens for cancer vaccination. To investigate this area, scFvs from the Tomlinson libraries were isolated against an idiotypic antibody (Goletz *et al.*, 2002). Elution of binders was accomplished with the original antigen followed by trypsin treatment of the recovered phage before rescue with the protease sensitive helper phage KM13. The results showed that specific elution in combination with trypsin treatment of the eluted phage was far superior to

other conventional recovery methods, enabling the generation of a large variety of Ab2s after only two to three rounds of selection, thereby maintaining maximum diversity. Liu and colleagues (Liu *et al.*, 2005) have used Tomlinson I and J libraries to isolate highly specific scFvs B5 and B9 against human SARS corona virus.

The Tomlinson I and J libraries have also been used to develop a technique for high-throughput screening of recombinant antibodies which is based on the creation of antibody arrays (de Wildt *et al.*, 2000). This method used robotic picking and high-density gridding of bacteria containing antibody genes followed by filter-based ELISA screening to identify clones that expressed binding antibody fragments. This technique was used in different applications including the isolation of antibodies against impure proteins and complex antigens, areas in which several rounds of conventional phage display often fail. The recombinant antibodies can be used to rapidly identify their cognate antigens on high density arrays of denatured protein. To identify highly specific antibody-antigen interactions by protein array screening, 12 well-expressed antibody fragments were chosen from the Tomlinson I and J libraries. After capturing the scFv antibodies, their interaction with 27 648 human foetal brain proteins arrayed on a PVDF membrane was tested. The authors demonstrated that truly naive repertoires can be used to identify specific antibody-antigen interactions by screening high density arrays of target proteins (Holt *et al.*, 2000).

Using phage display, monoclonal scFv and Fab fragments have been produced from Tomlinson I and J libraries that bind to the 51-kDa subunit of bovine complex I, a large integral membrane protein complex from mitochondria. Conventional Western blotting and competition ELISA were then used to confirm the identity of the target subunit, showing that antibody bound to the native protein complex (Rubinstein *et al.*, 2003). The Brahma-related gene 1 (BRG1) protein-containing Swi/Snf complex (hSwi/Snf) is a potent tumor suppressor and mutations in BRG1 are associated with cancer (Cho *et al.*, 2004). A commercially available antibody has previously been used to image BRG1 (Wang *et al.*, 2005) but its non-specific interaction with chromatin makes it unsuitable for imaging hSwi/Snf in the presence of chromatin. Using the

Tomlinson I and J libraries, Marcus and co-workers (Marcus *et al.*, 2006) isolated anti-BRG1 scFv that bound more strongly and more specifically than the commercial antibody, thus making it more suitable for use in imaging to elucidate the mechanisms of hSwi/Snf and BRG1 action on chromatin.

Solid tumors are characterised by unique epitopes which arise from angiogenesis and metastasis. Targeting of solid tumors for radioimmunotherapy requires recognition of these unique features and the rapid accumulation of therapeutic agents at the tumor site. To identify features of the cell adhesion molecule laminin that might be unique to solid tumour sites, scFvs from the Tomlinson I and J phage display libraries were selected against laminin-1, *in vitro*. These scFv were then radioiodinated and injected into tumor-bearing mice. One particular scFv exhibited preferential accumulation at subcutaneous tumor sites compared to other anti-laminin scFvs or to a control scFv (Davern *et al.*, 2005).

Sorsa-Leslie et al (Sorsa-Leslie *et al.*, 2005) screened the Tomlinson J phage display library to select scFvs against partially purified gonadotrophin surge-attenuating factor (GnSAF). The initial screening with a simple binding immunoassay resulted in 8 clones that were further screened using a bioassay for GnSAF. Three of the antibodies recognised GnSAF bioactivity and one clone bound GnSAF in such as way as to block its bioactivity. This study concluded that phage display and alternative screening strategies may be combined to improve protein identification strategies.

Overall, these examples illustrate the high discriminating power that can be achieved through use of phage display selection. They also show the breadth of targets that can be recognised through use of the synthetic libraries developed by Tomlinson. The concept of a "single pot" antibody library was initiated by Winter's group in the mid-1990's (Griffiths *et al.*, 1994; Nissim *et al.*, 1994) - a single resource from which antibodies against any target could be extracted. The examples above show that the Tomlinson I and J libraries are sufficiently diverse to be considered in these terms.

6.1.11: Objectives

During the course of this study, evidence was found that two variants of IgM are transcribed in cattle. Earlier data had documented the presence of JH loci on bovine chromosomes BTA11 and BTA 21 (Hosseini *et al.*, 2004), showing that the duplication of IgM constant region exons reported by others (Hayes and Petit, 1993) was more extensive than previously supposed. Detailed analysis established that the locus on BTA11 is not an orphan. The locus on BTA11 carries several nonsynonymous sequence differences and a *Sma* I restriction site is created. PCR of IgM cDNA and restriction analysis confirmed that the locus on BTA11 undergoes rearrangement and transcription in normal lymphoid tissue. The observation raised a number of questions relevant to the diversification of the bovine Ig repertoire. Is the locus on BTA11 independent in its rearrangement and expression from the main locus on BTA21? To what extent do B lymphocytes expressing Ig heavy chains from BTA11 share properties with those (the majority) that express sequences at BTA21? In particular, are the Ig heavy chains they carry diversified to equal degrees?

Antibodies able to discriminate the two variants of IgM were not available, so the objective of this study was to prepare scFvs by phage display technology that were specific for epitopes encoded on BTA11 and BTA21. Such reagents could be then used to assess the presence of variants of IgM and to isolate B cells that have undergone rearrangement at BTA21 and / or BTA11. These studies utilised the Tomlinson I and J libraries and, as target, recombinant Cμ1 protein carrying sequences encoded on BTA11 and BTA21.

6.2: MATERIALS AND METHODS

6.2.1: Amplification of IgM constant domain 1 (Cμ1) from chromosome 11 (BTA11) and chromosome 21 (BTA21)

Cμ1 domain specific primers with internal restriction sites for *Afe* I and *Spe* I were designed based on the sequences available for this region in GenBank. The sequence accession number U63637 represents a sequence originating from BTA11, whereas the accession number AY158087 has the sequence from BTA21. Cloned fragments of the IgM loci were available (Aitken, unpublished) and were used as template. The PCR amplifications consisted of pre-denaturation (94°C, 5

min), and 30 cycles of denaturation (94°C, 1 min), annealing (56°C, 30s), extension (72°C, 30s), followed by final extension at 72°C for 5 min. The PCR products were termed Cμ1[11] and Cμ1[21] to reflect their origin. They were cloned by TA cloning into pCR TOPO 2.1 (Invitrogen) as described previously. Clones were screened for the inserts of interest by performing site specific restriction digestion with *Afe* I and *Spe* I and checking for the release of fragments of the expected size.

6.2.2: Homology models for Cμ1 domains:

Homology models of the bovine Cμ1 domains were built using SWISS-MODEL (Schwede *et al.*, 2003) based on the structure determined by protein crystallography for a complex between Protein A of *S. aureus* and three Fab fragments of a human IgM antibody (Protein Data Bank ID 1DEE) (Graille *et al.*, 2000b). Only the CH1 domain of the Fab heavy chain was used to obtain the homology model. Models of Cμ1 variants in different positions were generated using Rasmol 2.6. The lack of structural information for the last five residues of the C terminal of Cμ1 domain meant that it was not possible to model this section of Cμ1 using homology structure modelling.

Table 10: Sequences of the primers used to recover, clone and purify Cμ1 variants

Primer	*Sequence (5' to 3')*
BTA11 Spe1 shift1R	TTACTAGTCTACGCTGGAGTGACCACCCTCACGG
BTA21 Spe1 shift1R	TTACTAGTCTACGCAGTAGCGATCACCCTCACGG
M13 R	CAGGAAACAGCTATGAC
Cμ1-BTA11-Afe1F	AGCGCTGGTGAATCACACACCCGAGAGTC
Cμ1-BTA11-Spe1R	CTCACTAGTAGTGACCACCCTCACGG
Cμ1-BTA21-Afe1F	AGCGCTGGTGAATCACACCCGAGAGTC
Cμ1-BTA21-Spe1R	CTCACTAGTAGCGATCACCCTCACGG
His rpt f	CTAGCGCTCATCATCATCACCATCACTAAAA
His rpt R	CTAGTTTTAGTGATGGTGATGATGATGAGCG
pET-CF-BamHI	ACCAAGGATCCAATAAAACGTGCG
pET-CR-HindIII	GAGCGCTAAGCTTCGCAGTAGC
PelB F	TTCTTACTAGTGGTGGTGGTGGTGGTGTGAC
Skp R	GGTTATTAGCTGCAGGTC

6.2.3: Insertion of a stop codon into the cloned fragments

Primers were designed to relocate the *Spe* I site and insert a stop codon at the end of Cμ1 in clones shown to carry the intended inserts. The plasmid DNA was prepared from positive clones and PCR was performed with M13 R as a forward primer annealing to flanking vector sequence and specific reverse primers BTA11 *Spe* I shift1R and BTA21 *Spe* I shift1R (Table 10). The products were gel purified and cloned again by TA cloning. Plasmid DNA from positive clones was then digested sequentially with *Afe* I and *Spe* I restriction enzyme to isolate the $C\mu1^{11}$ and $C\mu1^{21}$ sequences for cloning into expression systems.

Two expression strategies were used. In the first, Cμ1 sequences were fused to a scFv against hen eggwhite lysozyme (HEL) to create a single chain antibody (scAb). Fusion of the human kappa constant domain to scFvs has been developed as a strategy by Porter's group (Hayhurst and Harris, 1999). It was envisaged that an anti-HEL scAb carrying Cμ1 constant region could be orientated on a HEL-coated surface for selection of anti-Cμ1 antibodies by phage display. Selection could be pushed towards the constant domain by depletion on surfaces coated with anti-HEL scAb carrying the human Kappa constant domain, thereby removing phage recognising the scFv component of the target. In the second strategy, Cμ1 proteins were expressed directly in a T7 expression plasmid.

6.2.4: scAb expression and purification

6.2.4.1: Cloning into pIMS147 vector

E. coli XL-1 Blue carrying an anti-HEL scAb construct (pIMS147, Figure 47) was grown overnight in 10 ml LB broth containing 100 μg/ml ampicillin and 1% glucose at 37°C and shaking at 200 rpm. The following day, plasmid DNA was extracted and purified using a Qiagen miniprep kit. Purified plasmid DNA was then digested using the restriction enzyme *Not* I from New England Biolabs in NEBuffer 3. Following digestion with *Not* I, the plasmid was run on an agarose gel and purified. In total around 2-3 μg of digested plasmid DNA was used in a filling-in reaction at the 5'-protruding termini with dNTPs and the Klenow fragment of DNA Polymerase (Promega). One unit of Klenow was used per microgram of DNA. The reaction was incubated at 25° C for 10 min and stopped

by heating at 75°C for 10 min. Following the creation of blunt termini, a second restriction digest was setup with *Spe* I to release the human Cκ sequence from the vector. The large fragment so generated was then purified and used in ligation with $C\mu1^{11}$ and $C\mu1^{21}$ inserts. The ligation reactions were setup using a Quick Ligation Kit from New England Biolabs. In total, about 100ng of vector was combined with a 3-fold molar excess of insert and the volume was adjusted to 10 µl with dH_2O. The amount of insert required was calculated using the following formula:

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

Then, 10 µl of 2X Quick Ligation Buffer was added, followed by 1µl of T4 DNA Ligase. The reaction mix was centrifuged briefly and incubated at 25°C in a thermocycler for 5 min. Following ligation, the mix was chilled on ice and 2 µl was transformed into DH5α competent cells. The controls - undigested vector, vector digested with a single enzyme, vector without insert - were routinely included in ligation-transformation experiments.

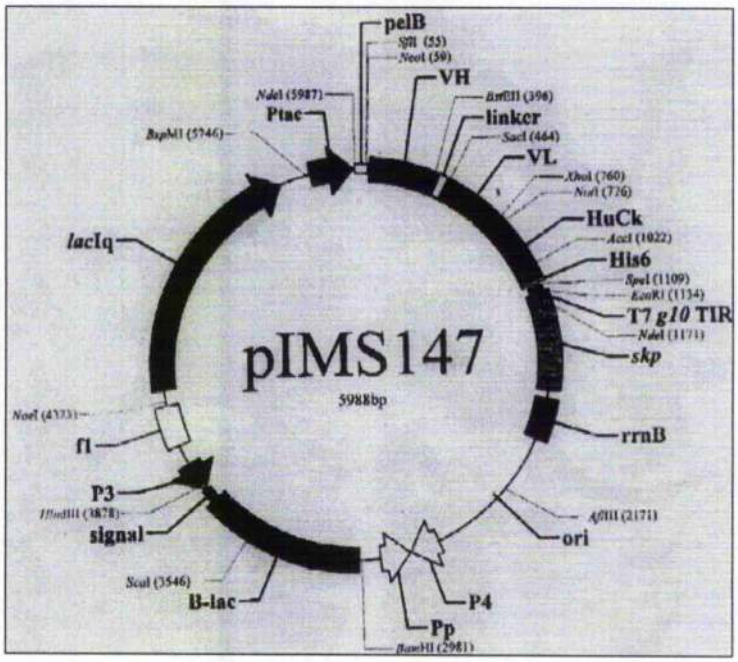


Figure 47: Restriction map of the scAb vector, pIMS 147

DH5α cells were transformed with 5-10 ng of DNA from each ligation reaction using heat shock transformation. Clones were selected on LA containing 100 µg/ml ampicillin and 1% glucose at 37°C overnight. Transformation with 0.1

ng of pUC18 plasmid was used to estimate the transformation efficiency. Clones were screened for successful ligation by colony PCR. The following day, plasmid DNA was extracted from PCR-positive clones digested with restriction enzymes to confirm the presence of inserts. Further analysis with *Sma* I confirmed if the Cμ1 insert was derived from BTA11 or BTA21. The clones were then sequenced before undertaking expression experiments.

6.2.4.2: Expression and isolation of soluble $C\mu1^{11}$ and $C\mu1^{21}$- fused scAbs:

Clones carrying scAb inserts with $C\mu1^{11}$ and $C\mu1^{21}$ sequences were grown in 5 ml of LB broth containing 100 μg/ml ampicillin and 1% glucose at 37°C with shaking at 220 rpm, overnight. 1 ml of each overnight culture was inoculated into 50 ml of LB broth containing 100 μg/ml ampicillin and 1% glucose and grown on at 37°C with shaking at 220 rpm until the OD at 600 nm was 0.8. At this point, IPTG was added to a final concentration of 1mM and growth continued overnight at the lower temperature of 30°C with shaking at 220 rpm. The cultures were then centrifuged at 13000 g for 10 min at 4°C. The supernatants were collected and after filtration, were stored at 4°C. Periplasmic extracts (PE) were prepared from the bacterial pellets by gently resuspending to 5% of the initial culture volume in cold 50 mM Tris-HCl, 20% sucrose, 1mM EDTA, pH 8.0. The cell suspensions were then incubated on ice for 60 min with occasional stirring and then were centrifuged in microfuge tubes at 13000 g for 30 min at 4°C. The supernatants - the soluble PEs - were carefully collected and filtered through filters with a pore size of 0.4 microns. Samples were stored at 4°C for future analysis.

6.2.4.3: Western blotting analysis

The expression of anti-HEL scAbs carrying Cμ1 domains was assessed by Western blotting analysis. In total, 15 μl of PE from the expressed cultures were run on a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane as described previously. After blocking with 3% MPBS and washing three times with PBS, the membrane was incubated in a 1:1000 dilution of bovine anti-IgM-HRPO conjugate in 2% MPBS and incubated at 37°C for 60 min with 200 rpm shaking. After that, the membrane was washed three times with PBS containing 0.05% Tween 20. The bound antibody was detected finally by

developing the reaction with 4-chloro-N-naphthol substrate in the presence of H_2O_2.

6.2.4.4: Purification of Cμ1^{11} and Cμ1^{21}

Since the clones did not carry a tag which could have been used in purification, it was decided to couple HEL with cyanogen bromide (CNBr)-activated Sepharose beads in order to generate a matrix for affinity capture through the scFv domain. One gram of Sepharose was suspended in 1 mM HCl for 30 min, allowed to swell to about 3.5 ml and then washed with 15 gel volumes of cold 1 mM HCl followed by washing with coupling buffer, pH 8.3. HEL was dissolved in coupling buffer at a recommended ratio of 5 mg HEL per ml gel. The OD_{280} reading of the HEL solution was taken at this stage. The washed gel was then added to the HEL solution in a tube and incubated overnight at 4°C. To determine the coupling efficiency, 500 μl of suspension was spun at 2000 rpm for 1 minute and the OD_{280} of the supernatant was measured. The gel was washed and resuspended in freshly prepared 1 M ethanolamine for 2-4 hours at room temperature to block activated sites that remained unoccupied. The gel was then washed 5 times with alternating 50 mM Tris, 1 M NaCl pH 8.0, and 50 mM glycine, 1 M NaCl pH 3.5 buffers before a final wash with 10 gel volumes of PBS. The coupling efficiency was measured by comparing the absorbance at 280nm of HEL before and after the process. The coupled gel was poured into a glass chromatography column in a single continuous motion to prevent the trapping of air bubbles. The filtered PE extract was diluted in 1:1 with PBS and applied to the affinity column under gravity flow. The unbound material was collected. The column was washed with 10 ml PBS. The bound material was eluted with a phosphate citrate buffer of low pH.

As an alternative approach, it was decided to insert his-tag to facilitate the purification of the recombinant proteins by immobilized metal affinity chromatography. The primers His rpt F and His rpt R (Table 10) were designed based on the sequences of Cμ1^{11} and Cμ1^{21} cloned fragments and the vector backbone. These were designed to form a synthetic duplex for insertion at the *Spe* I site, thereby adding a repeat of six histidine codons into the reading frame of the Cμ1. For annealing, 800 pmoles of each primer were mixed together at 65°C for 5 min in a reaction volume of 100 μl, and then $MgCl_2$ was added to a

final concentration of 2mM and the solution was allowed to cool slowly to room temperature. Once the reaction was at room temperature, the annealed his-tag adaptor was further purified by phenol extraction and ethanol precipitation and finally resuspended in 100 µl TE buffer pH 8.0. The duplex was then phosphorylated using T4 polynucleotide kinase (Promega) as per manufacturer's instructions. The reaction was assembled in a sterile tube as follows:

Contents	Quantity
Annealed His-tag (16 pmol/µl)	12 µl
Kinase buffer (10x)	4 µl
10mM ATP	4 µl
T4 PNK	2 µl
Deionised water	18 µl

The reaction was incubated at 37°C for 30 min and stopped by adding 2 µl of 0.5M EDTA. The phosphorylated His-tag adapter was phenol extracted and ethanol precipitated as before and finally suspended in 50µl TE buffer, ready for ligation. The plasmid DNA from individual recombinant clones was digested with *Spe* I and after inactivating the enzyme, the digested vector molecules were purified by spin-purification column from Qiagen and DNA quantified by nano-drop before dephosphorylation. The digested vectors were suspended in 1x buffer (0.5 µg/10µl) followed by addition of 0.5 units of calf intestinal alkaline phosphatase (CIP) per µg of vector DNA and incubation at 37°C for 60 min. The vectors were then purified by spin-column purification and quantified before setting up ligations.

Ligation was performed as before with 100 ng of dephosphorylated vector with a 3-fold molar excess of the phosphorylated his-tag adaptor. DH5α competent cells were transformed with 2µl of ligation mix. Controls (uncut vector, vector without insert) were included while performing ligation-transformation procedures. The clones were screened for successful ligation by colony PCR using SkpR/His-rpt F or PelB F/His-rpt R (Table 10) and confirmed by sequencing. In pIMS 147 (Figure 47) coding sequence for the Skp chaperone lies downstream of the scAb cassette. The scAb cassette carries a PelB leader sequence, providing an anchoring sequence on the upstream flank for diagnostic PCR. Plasmid DNA from pIMS147/Cµ1^{11}/His-tag and pIMS147/Cµ1^{21}/His-tag positive clones was then transformed into TG1 cells to assess expression levels.

6.2.4.5: Expression and isolation of His-tagged Cμ1^{11} and Cμ1^{21}-fused scAbs

Transformed TG1 clones carrying scAb sequences fused to Cμ1 from BTA11 and BTA21 were grown in 5 ml LB broth containing 100 μg/ml ampicillin and 1% glucose at 37°C and with shaking at 220 rpm. 1 ml of each overnight culture was inoculated into 50 ml of LB broth containing 100 μg/ml ampicillin and 1% glucose and grown on at 37°C with shaking at 220 rpm until the OD at 600 nm was 0.8. At this point, IPTG was added to a final concentration of 1mM and growth continued overnight at 30°C with shaking at 220 rpm. Conditions of lower IPTG concentration (0.5mM) and lower temperature (25°C) were also employed to optimize protein expression. The cultures were then centrifuged at 13000 g for 10 min at 4°C. The supernatants were collected and after filtration, were stored at 4°C. Periplasmic extracts (PE) were prepared from the bacterial pellets as described earlier by gently resuspending the cell pellet to 5% of the initial culture volume in cold 50 mM Tris-HCl, 20% sucrose, 1mM EDTA, pH 8.0. The cell suspensions were then incubated on ice for 60 min with occasional stirring and then were centrifuged in microfuge tubes at 13000 g for 30 min at 4°C. The supernatants - the soluble PEs - were carefully collected and filtered. Samples were stored at 4°C for analysis or purified using affinity chromatography.

6.2.4.6: Western blotting analysis

The PE extracts were analysed by SDS-PAGE and the presence of the scAbs was also assessed in PEs by Western blotting. 15 μl of PE from the expressed clones were run on a 12% SDS-acrylamide gel and transferred to a nitrocellulose membrane as described previously. After blocking with 3% MPBS and washing three times with PBS, the membrane was incubated in a 1:1000 dilution of bovine anti-IgM-HRPO conjugate in 2% MPBS at 37°C for 60 min with 200 rpm shaking. After that, the membrane was washed three times with PBS containing 0.05% Tween 20. The bound antibody was detected finally developing the reaction with 4-chloro-N-naphthol substrate in the presence of H_2O_2.

6.2.4.7: Screening by ELISA

Supernatants from induced cultures of *E. coli* were applied to ELISA plates coated with 100 μl of a 1:500 dilution of goat anti-human kappa antibody or 10 μg of hen egg lysozyme. After blocking with alternative blockers of 2% skimmed

milk, 3% BSA or rabbit serum (1:10 times diluted in PBS), the binding of proteins was detected by addition of a goat anti-human kappa HRP conjugate or anti-bovine IgM HRP conjugate and OPD substrate. Colour change was recorded at 450 nm.

6.2.4.8. Purification by immobilised metal affinity chromatography on HiTrap SP XL column

All the solutions, buffers and bacterial lysates were filtered before applying to a HiTrap SP XL column (Amersham Bioscience). The composition of buffers is given in Appendix I. Before running the sample, the column was stripped with 0.05 M EDTA in DW using a Pharmacia Biotech P-1 peristaltic pump (Sweden) to pump solutions at a flow rate of 5 ml/min. The column was washed with 5 Column Volumes (CV) i.e. 25 ml of DW followed by 0.5 CV of 0.1 M NiSO4. The column was again washed with 5 CV of water, followed by 5 CV of binding buffer. The column was washed with 5 CV of binding buffer containing 50 mM imidazole followed by 5 CV of binding buffer. The filtered protein sample was then applied and flowthrough was collected for reapplication. The column was washed with 5 CV of binding buffer containing 20 mM imidazole. Bound material was eluted by pumping 5 CV of binding buffer containing 70 mM, and then 200 mM imidazole. Fractions were stored at 4°C for further analysis by SDS-PAGE.

6.2.5: Subcloning Cμ1^{11} and Cμ1^{21} into pET-21a

Hind III and *Bam* HI sites were created in Cμ1^{11} and Cμ1^{21} clones to facilitate their cloning into pET-21a vector. Primers pET-CF-*Bam*HI and pET-CR-*Hind* III, carrying the restriction sites (Table 10) were used to amplify the Cμ1 coding sequences from clones used to this point for expression. The PCR amplifications consisted of pre-denaturation (94°C, 5 min), and 30 cycles of denaturation (94°C, 1 minute), annealing (56°C, 30s), extension (72°C, 30s), followed by final extension at 72°C for 5 min. The products were gel purified and cloned by TA cloning. Positive clones were digested with *Hind* III and *Bam* HI to retrieve the Cμ1^{11} and Cμ1^{21} inserts which were then subcloned into pET21-a. The map of the vector is given Figure 48. Ligation was performed as before with 100 ng of cut vector and with a 3-fold molar excess of each insert, and DH5α competent cells were transformed with 2 μl of ligation mix. The clones were screened by colony PCR and restriction analysis. The plasmid DNA from pET21-a/Cμ1^{11} and pET21-

a/Cμ1[21] clones was then transformed into BL21 (DE) cells (Stratagene) for protein expression.

6.2.5.1: Protein expression and solubility studies:

The composition of all the buffers used during protein purification is given in the Appendix. Overnight cultures were set up by inoculating LB broth supplemented with 100 μg/ml ampicillin with a single bacterial colony obtained from transformation. For protein expression, fresh medium was inoculated at a ratio of 1:100 with overnight bacterial culture. Bacteria were grown at 37°C to an absorbance at 600 nm of 0.8 and then induced with IPTG at a concentration of 500 mM for 14-16 hours at 18°C. Cells were harvested by centrifugation at 13000 g for 20 min. Bacterial pellets were stored at -80°C if not processed immediately. Protein solubility was tested by lysing the cells from 1 liter of culture in 60 ml of metal chelate binding buffer in the presence of an EDTA-free cocktail of protease inhibitors (Roche). A French press, operating at 750psi, was used for bacterial lysis. Cell debris was removed from the soluble fraction by centrifugation at 13000 g for 20 min.

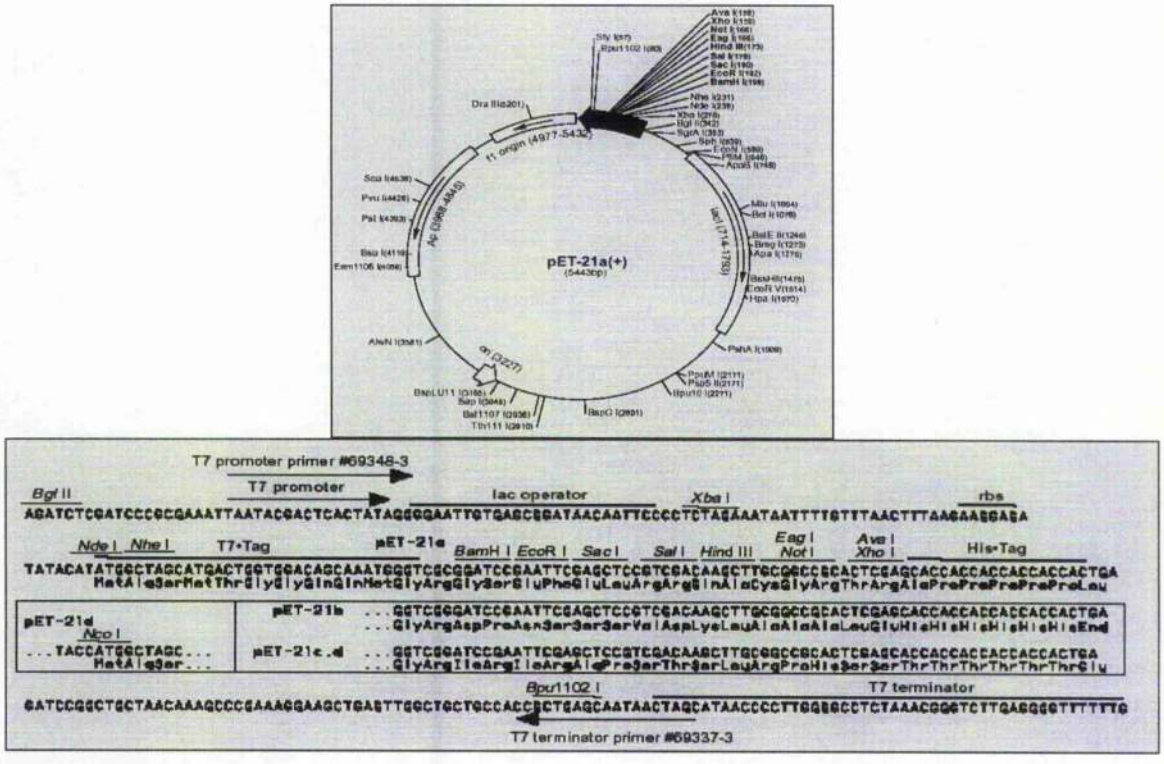


Figure 48: Restriction map and cloning/expression region of pET-21a(+)

Protein solubility was assessed by SDS -PAGE analysis of samples from the cytoplasmic extract, and the pellet of cell debris.

6.2.5.2: Protein purification

His-tagged recombinant proteins expressed from pET-21a/Cμ1[11] and pET21a/Cμ1[21] were first purified using metal chelate affinity chromatography on a BioCAD 700E workstation (Applied Biosystems, USA). A 20MC column was prepared at a flow rate of 10 ml/min. The column was initially stripped with 5 CV of stripping buffer to remove any metal ions from previous runs, followed by 5 CV of deionised water. 30 CV of 100mM $ZnCl_2$ was loaded onto the column at 5ml/min followed by further washing with 5 CV of deionised water at 10ml/min. The unbound zinc ions were removed with 5 CV of 500 mM NaCl and the column was then equilibrated with 5 CV of elution buffer followed by 10 CV of binding buffer. The sample was filtered before injection in 5 ml aliquots. Each injection was followed by a wash step of binding buffer for 3 CV. After the final injection, the column was washed with 9 CV of binding buffer. Bound protein was eluted from the column in a 0-100% gradient of elution buffer and the eluate was collected in 2 ml fractions. The peak fractions were then analysed by SDS-PAGE and fractions that showed greatest promise were pooled and subjected to gel-filtration chromatography.

A Sephacryl S-100 column (Amersham, USA) was equilibrated with 2 CV of buffer at 1ml/min. Fractions from the previous purification protocol were pooled, concentrated to less than 1 ml and injected onto the column. Protein was eluted over 1.2 CV of buffer run at 1ml/min while collecting 2 ml fractions.

6.2.6: Identification of recombinant Cμ1 by mass spectrometry

Bands predicted from their size to represent the recombinant Cμ1 proteins were excised from the SDS gel and submitted to the Sir Henry Wellcome Functional Genomics Facility (SHWFGF), University of Glasgow. This houses a high-performance bench top matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) system for protein identification. The analysis was carried out by Dr. Richard Burchmore.

6.2.7: Determination of protein concentration

The protein concentrations of purified recombinant $C\mu1^{11}$ and $C\mu1^{21}$ proteins were determined using their calculated extinction coefficients at 280 nm before using them in phage display.

6.2.8: Phage display-Tomlinson I and J libraries

The human single fold scFv libraries I and J, helper phage KM13, *Escherichia coli* TG1 and HB2151 were kindly provided by Greg Winter's lab at the MRC Laboratory of Molecular Biology and the MRC Centre for Protein Engineering (Cambridge, UK). The materials supplied comprised 500 µl of library I and 500 µl of library J in *E. coli* TG1 cells, glycerol stocks of positive control scFvs against BSA and ubiquitin in TG1, a glycerol stock of T-phage resistant *E.coli* TG1 for propagation of phage (K12 *Δ(lac-proAB) supE thi hsdD5/F' traD36 proA⁺B laclq lacZΔM15*), a glycerol stock of *E. coli* HB2151 for expression of soluble antibody fragments (K12 *ara Δ(lac-proAB) thi/F' proA⁺B laclq lacZΔM15*) and 100 µl of phage KM13 at 10^7 pfu/ml.

6.2.8.1: Growing the libraries

The 500 µl stocks of Tomlinson I and Tomlinson J were added to 200 ml of pre-warmed 2xTY containing 100 µg/ml ampicillin and 1% glucose and grown with shaking at 37°C until the OD_{600} nm reached 0.4. 150 ml of these cultures were used for making secondary stocks and the remaining cultures were infected with 2×10^{11} pfu of KM13 helper phage. Cultures were incubated without shaking in a 37°C water bath for 30 min, centrifuged at 3300 g for 10 min and the bacterial pellets were resuspended in 100 ml of 2xTY containing 100 µg/ml ampicillin, 50 µg/ml kanamycin and 0.1% glucose. The suspensions were then incubated overnight at 30°C with shaking. The next day, after spinning at 3300 g for 30 min, 20 ml PEG/NaCl was added to 80 ml of each supernatant. The mixtures were incubated for an hour on ice, spun at 3300 g for 30 min and the pellets resuspended in 4 ml PBS. After a further spin at 11600 g for 10 min, the supernatants containing the phage libraries were stored at 4°C.

6.2.8.2: Titration of phage

To titer the stocks of library phage, 1 µl of each library was diluted into 100 µl PBS. 1 µl of this was withdrawn and diluted in a further 100 µl PBS and so on until there were 5-6 dilutions in total. 900 µl of *E. coli* TG1 grown to an OD at 600 nm of 0.4 was added to each tube and incubated at 37°C in a water bath for 30 min. 10 µl of each dilution was then spread onto a TYE plate containing 100 µg/ml ampicillin and 1% glucose and grown overnight at 37°C. Phage titres were calculated by counting the number of ampicillin-resistant colonies, and correcting for dilution of the phage stock.

6.2.9: Selection strategy: Panning Tomlinson library I and library J against purified Cµ1^{11} and Cµ1^{21}

6.2.9.1: First round of selection

Wells of a microtitre plate were coated with 500 ng aliquots of purified Cµ1^{11} and Cµ1^{21} in 300 µl of PBS per well as shown in the Figure 49 below. The plates were sealed and incubated at 4°C overnight.

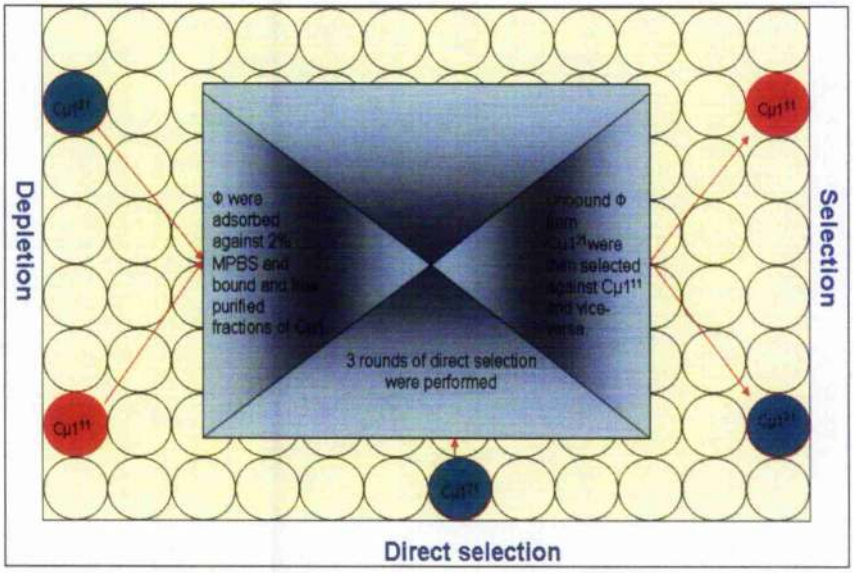


Figure 49: Depletion-selection strategy as detailed in text

Shaded wells were coated with specified amounts of purified fractions of Cµ1 for each round of selection to select phage (Φ). Direct selection was also performed against Cµ1 from BTA 21 as a control.

The following day, the wells were washed 3 times with PBS and were blocked by filling with 2% MPBS and incubating at room temperature for 2 hours. The wells were washed again 3 times with PBS. A depletion-selection strategy

was employed to remove most of the phage reactive to one form of the IgM domain and epitopes on skimmed milk, and then select the remaining virus mixture against the target protein. About 1.1×10^{12} phage with equal representation from libraries were adsorbed for one hour on $C\mu1^{11}$ in 2% MPBS, then were added in 300 µl volume to $C\mu1^{21}$ and allowed to bind for another hour at room temperature on a shaker. Similarly, another aliquot of phage were adsorbed for one hour on $C\mu1^{21}$ in 2% MPBS and were then transferred in 300 µl volume to $C\mu1^{11}$ and allowed to bind for one hour. After the selection step, the supernatants from wells were discarded and the wells were washed 10 times with PBS containing 0.1% Tween-20. After shaking out the excess PBS, phage were eluted by adding 300 µl of trypsin-PBS solution and shaking for 10 min at room temperature. As a control, an aliquot of phage was selected directly by application in MPBS to a well coated with $C\mu1^{21}$ protein.

6.2.9.2: Infecting TG1 cell with eluted phage antibodies

All the centrifugation steps were performed at 4°C during screening the libraries and various rounds of selection. The day before selection, one colony of *E. coli* TG1 was taken from a TYE agar plate and grown in 5 ml of 2xTY overnight at 37°C with shaking at 200 rpm. On the day of selection, 50 µl of this culture was inoculated to 50 ml of 2xTY and incubated at 37°C with shaking until the culture entered exponential growth and the OD_{600} was 0.4. Then, 150 µl of the eluted phage from each well was added to 1.75 ml of exponentially growing TG1 cells and incubated at 37°C in a water bath for 30 min. 10µl of infected TG1 cells, 10 µl of a 1 in 100 dilution, and 10 µl of a 1 in 10^4 dilution were spotted on TYE plates containing 100 µg/ml ampicillin and 1% glucose and grown overnight at 37°C to titer the phage recoveries. The remaining 150 µl of phage were stored at 4°C for later use. The following day, the output of phage from the first round of selection was calculated from the number of ampicillin-resistant colonies on titration plates. These plates were stored at 4°C for future analysis of first round phage antibodies.

The remaining cultures of infected TG1 were spun down in a microcentrifuge at 11600 g for 5 min. The pelleted bacteria were resuspended in 50 µl of 2xTY and plated on two TYE plates containing 100 µg/ml ampicillin and

1% glucose and grown overnight at 37°C. Colonies were then scraped from the plates with a sterilised spreader and inoculated into 2 ml of 2xTY containing 100 μg/ml ampicillin and 1% glucose. One ml of these bacteria was stored at -80°C in 15% glycerol and 50 μl of the remaining bacteria were used in phage rescue for the next round of selection.

6.2.9.3: Rescue of selected phage antibodies

50 μl of bacteria from the first round of selection were inoculated to 50 ml of 2xTY containing 100 μg/ml ampicillin and 1% glucose and grown at 37°C with shaking at 200 rpm until the OD_{600} reached 0.4. At this point, 5×10^{10} pfu of KM13 helper phage were added to 10 ml of culture and incubated without shaking at 37°C in a water bath for 30 min. The cultures were then centrifuged at 3000 g for 10 min. The pelleted bacteria were resuspended in 50 ml of 2xTY containing 100 μg/ml ampicillin, 50 μg/ml kanamycin and 0.1% glucose and incubated with shaking at 30°C overnight. The following day, overnight cultures were spun at 3300 g for 15 min. Phage were precipitated from the culture supernatants by addition of 10 ml of cold PEG/NaCl to 40 ml of each supernatant. After proper mixing, the mixtures were left for 1 hour on ice and then centrifuged at 3300 g for 30 min. Supernatants were poured away to remove PEG/NaCl and the precipitated phage were resuspended in 1 ml of PBS and spun at 11600 g for 10 min in a micro centrifuge to remove remaining bacterial debris. 600 μl of each rescued phage library, the supernatant, was stored at 4°C and the remaining 400 μl was used for next round of selection.

6.2.9.4: Titration of rescued phage:

To assay the titers of phage rescued from the first round of selection and thereby determine the phage input for the second round of selection, 1 μl of rescued phage from each library was diluted into 100 μl PBS and the dilution repeated to generate six serial 100-fold diutions in total. 900 μl of TG1 at an OD_{600} of about 0.4 was added to each tube and incubated at 37°C for 30 min. 10 μl of each dilution was spread onto a TYE plate containing 100 μg/ml ampicillin and 1% glucose and grown overnight at 30°C.

6.2.9.5: Second and third rounds of selection

In the second and third rounds of selection, wells of the microtitre plate were coated in the same format as for round one but with lower concentrations of target proteins. This was done to ensure isolation of phage antibodies with higher specificity and affinity. For the second and third rounds of selection, 400ng/ml and 300ng/ml of respective purified fractions of proteins were used for coating the wells. The next day coated wells were washed and blocked as described before. The number of phage used as input for each screen is given in Table 12. All the remaining procedures were done as described for the first round of screening.

6.2.10: Screening phage antibodies by monoclonal phage ELISA

After three rounds of selection, phage antibodies were screened for reaction against bovine IgM. Screening was done with individual phage clones recovered at the second and third rounds of selection. In total, 32 individual colonies picked at random from the second and third rounds of selection were inoculated into 100 µl of 2xTY containing 100 µg/ml ampicillin and 1% glucose in 96 well plates. Colonies were picked from plates used for titration of phage stocks. The plates were incubated overnight at 37°C with shaking at 200 rpm. The following day, approximately 2 µl of each overnight culture was transferred to wells of a fresh plate containing 200 µl of 2xTY, 100 µg/ml ampicillin and 1% glucose per well. A multi-channel pipette was used for transfer. Plates were incubated at 37°C for 2 hours to reach to an OD_{600} of about 0.4. Then, 25 µl of 2xTY containing 100 µg/ml ampicillin, 1% glucose and 10^9 helper phage was added to each well to initiate superinfection. Plates were incubated again at 37°C with shaking at 200 rpm for 1 hr to allow infection of the TG1 cells with helper phage. The plates were then centrifuged at 1800 g for 10 min and the supernatants were aspirated off. Each pellet was resuspended in 200 µl of 2xTY containing 100 µg/ml ampicillin and 50 µg/ml kanamycin. Cells were grown overnight at 30°C with shaking at 200 rpm to allow replication and rescue of the phage antibodies. The next day, plates were spun at 1800 g for 10 min. Each supernatant was then recovered for use as monoclonal phage antibody in ELISA assay.

Two ELISA plates (Nunc) were coated with 100 µl per well of 500ng/ml polyclonal anti-bovine IgM antibody. Plates were left overnight at room temperature to allow the antibody to bind to the plastic wells. The next day, unbound protein was discarded and the wells were washed three times with PBS. The wells were blocked by addition of 200 µl of 2% MPBS and incubated for 2 hours at 37°C.100 µl of normal bovine serum taken from healthy adult cattle was diluted 1:10 and added into each well as a source of bovine IgM. The serum was allowed to bind for 1 hr at 37°C. After the end of incubation the wells were washed three times with PBS containing 0.1% Tween 20. The culture supernatant containing phage antibodies from individual clones was diluted 1/2 in 2% MPBS and 100 µl was added to each well. Plates were left for 1 hr at room temperature for binding.

Unbound phage antibodies were removed by washing three times, wells with PBS containing 0.1% Tween 20. 100 µl of a 1:2500 dilution of HRP-anti-M13 monoclonal antibody (Amersham Pharmacia Biotech) in 2% MPBS was added to all wells. Plates were then incubated at room temperature for 1 hr. Wells were washed three times with PBS containing 0.1% Tween 20 to remove unbound antibodies. 100 µl of developer solution (composition as described previously) was added to each well. The colorimetric reactions were stopped after 3-10 min by adding 50 µl of 1 M sulphuric acid. Results were read at 450 nm in an ELISA reader.

6.3: RESULTS

6.3.1: Recovery and analysis of Cµ1 variants

In previous work, duplicated copies of the Cµ1 loci were located on BAC clones that could be mapped to BTA 11 and BTA 21 (Aitken, unpublished). Later, individual exons were recovered for comparison, to assess if significant differences existed. In preparation for phage display, the Cµ1 sequences were amplified with Cµ1-BTA11-Afe1F/BTA11-Spe1 shift R and Cµ1-BTA21-Afe1F/BTA21-Spe1 shift R primers mentioned in Table 10, cloned to a TA vector (Figure 50) and sequenced (Figure 51). A modest number of nucleotide differences were noted between sequences derived from BTA11 and BTA21. Some were silent (e.g. the third serine residue of the exon is encoded by TCG

[BTA11] or TCA [BTA21]), but in total 8 points of difference in the DNA sequence resulted in amino acid differences.

Figure 50: Restriction analysis of TOPO 2.1/Cμ1 clones

The random colonies were screened for insert of interest. The positive clones released around 321 bp Cμ1 fragments when digested with *Afe* I and *Spe* I restriction enzymes.

Some changes were conservative (e.g. valine/isoleucine at position 103) but other were of more significance (e.g. glutamine *versus* arginine at position 30). Potentially, recognition of some of these points of difference with specific reagents could enable discrimination of the variant forms of IgM produced in cattle. This strategy would only have the potential for success if variant residues were exposed at the surface of the Cμ1 domain.

```
                    4                                                  14
             G  E  S  L  P   R  V  F  P  L   V  S  C  M  S
Cmu1[11]  ggtgaatcgctcccg agagtcttcccccctg gtgtcctgcatgagc
Cmu1[21]  ggtgaatcacacccg agagtcttcccccctg gtgtcctgcgtgagc
             G  E  S  H  P   R  V  F  P  L   V  S  C  V  S

                                                               30
             S  P  S  D  E   S  T  V  A  L   G  C  L  A  Q
Cmu1[11]  tcccccatccgatgag agcacggtggccctg ggctgcctggcccag
Cmu1[21]  tcgccatccgatgag agcacggtggccctg ggctgcctggcccgg
             S  P  S  D  E   S  T  V  A  L   G  C  L  A  R

                   33
             D  F  M  P  N   S  V  S  F  S   W  K  F  N  N
Cmu1[11]  gacttcatgcccaat tcagtcagcttctcc tggaagttcaacaac
Cmu1[21]  gacttcgtgcccaat tcagtcagcttctcc tggaagttcaacaac
             D  F  V  P  N   S  V  S  F  S   W  K  F  N  N

                         49                                    58
             S  T  V  G  S   E  R  F  W  T   F  P  A  V  L
Cmu1[11]  agcacagtcggcagc gagagattctggacc ttccccgcagtcctg
Cmu1[21]  agcacagtcagcagc gagagattctggacc ttccccgaagtcctg
             S  T  V  S  S   E  R  F  W  T   F  P  E  V  L

             R  D  G  L  W   S  A  S  S  Q   V  V  L  P  S
Cmu1[11]  agggacggcttgtgg tcggcctcctctcag gtggtcctgccctcc
Cmu1[21]  agggacggcttgtgg tcggcctcctctcag gtggtcctgccctcc
             R  D  G  L  W   S  A  S  S  Q   V  V  L  P  S

             S  S  A  F  Q   G  P  D  D  Y   L  V  C  E  V
Cmu1[11]  tcaagcgcctttcaa gggccggatgactac ctggtgtgcgaagtc
Cmu1[21]  tcaagcgcctttcaa gggccggatgactac ctggtgtgcgaagtc
             S  S  A  F  Q   G  P  D  D  Y   L  V  C  E  V

                                                              104
             Q  H  P  K  G   G  K  T  V  G   T  V  R  V  V
Cmu1[11]  cagcaccccaaggga ggaaagaccgtcggc accgtgagggtggtc
Cmu1[21]  cagcaccccaaggga ggaaagaccgtcggc accgtgagggtgatc
             Q  H  P  K  G   G  K  T  V  G   T  V  R  V  I

          105 106
             T  P
Cmu1[11]  actcca
Cmu1[21]  gctaca
             A  T
```

Figure 51: Alignment of the nucleotide and protein sequences of Cμ1 domains derived from BTA11 and BTA21.

Nucleotide differences are underlined. Amino acid differences are numbered according to their position in the Cμ1 reading frame and the ones in red have been displayed using homology based Swiss-Model in Figure 52.

Using the crystal structure of a complex comprising *S. aureus* Protein A and a human IgM Fab (Protein Databank reference number 1DEE), models of the isolated Cμ1 domain were built. These were then adapted using Swiss-Model (Schwede *et al.*, 2003) to substitute bovine variant residues into the structure (Figure 52). Hybrid models were generated by substituting CH1 domain of Fab human IgM by Cμ1. The results confirmed that with the exception of residues 104-106, all variant amino acids were likely to adopt surface locations in the Cμ1 structure.

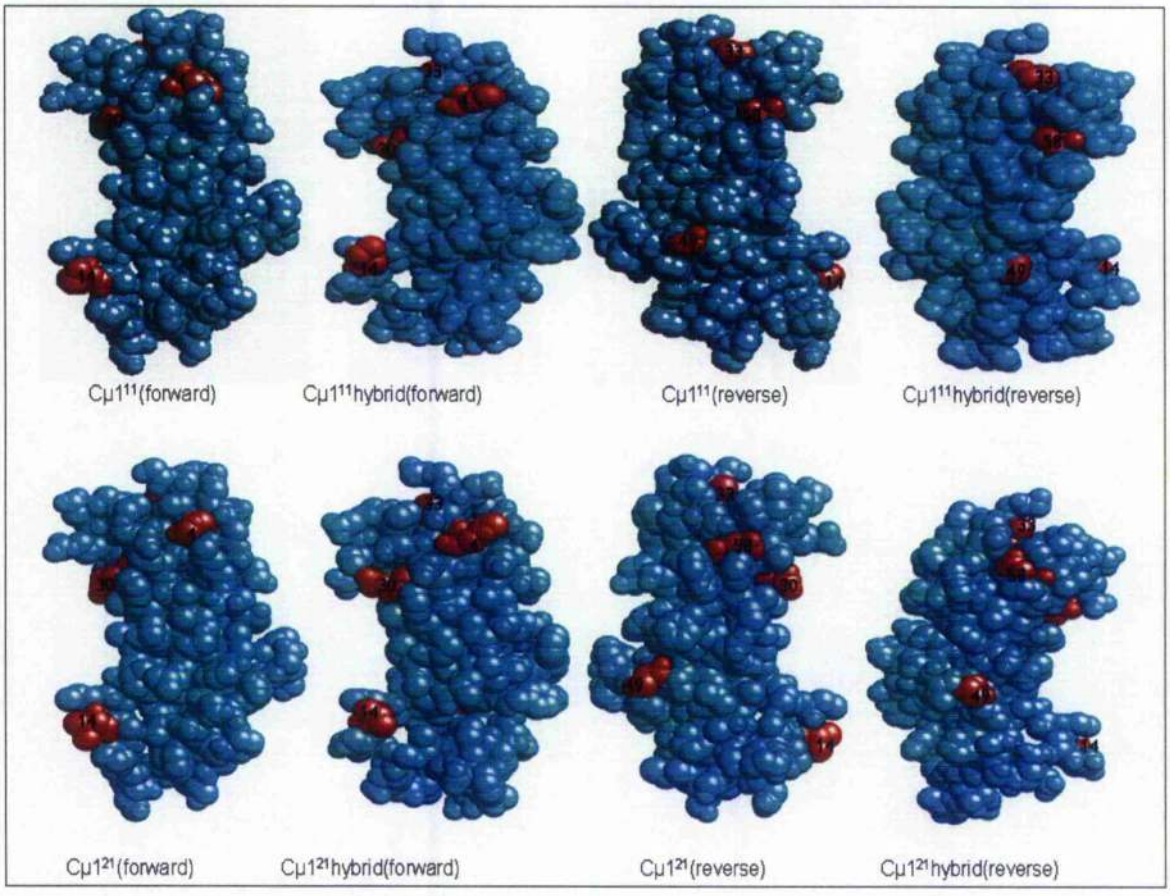


Figure 52: Homology models of the bovine Cμ1 domains.

Variant residues identified in Figure 50 are highlighted in red and labelled according to their position in the Cμ1 domain. Residues 4, 14, 30, 33, 49 and 58 correspond to L or H, M or V, Q or R, M or V, G or S and A or E in $C\mu1^{11}$or $C\mu1^{21}$, respectively.

The lack of structural information for the last five residues of the Cμ1 domains meant that this part of the protein could not be modeled satisfactorily. The isolation of reagents specific for the variants of bovine Cμ1 thus seemed a feasible goal.

6.3.2: Constructions of scAbs

Once the Cμ1 sequences had been successfully cloned and sequenced a stop codon was introduced at the 3' termini of the $C\mu1^{11}$ and $C\mu1^{21}$ sequences by PCR. Amplicons were cloned into pCR2.1-TOPO. Candidate TOPO clones were analysed and orientation of the insert was ascertained. Based on the results, it was decided to use M13 R with BTA11 Spe1 shift R and BTA21 Spe1 shift R in order to confirm the presence of the intended inserts by colony PCR. Successful amplification products of 410 bp identified clones for further characterisation. The sequencing of these PCR products in pCR 2.1 TOPO confirmed that the intended modifications had been created resulting in a stop codon and appropriate restriction sites in the correct reading frame. The inserts from these positive clones were then cut out by digestion with *Afe* I and *Spe* I and ligated into pIMS147 vector. In preparation for this, the pIMS147 vector was prepared by digestion with *Not* I, blunted, and further digested with *Spe* I. This removed a DNA fragment of about 330 bp encoding the human Cκ domain. The Cμ1-coding fragments were then ligated into place. The positive clones were screened by colony PCR (Figure 53) and further confirmed by restriction digest analysis.

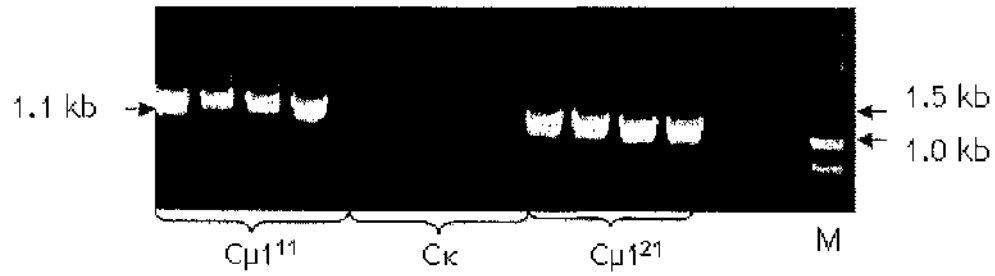


Figure 53: Colony PCR on clones screened for Cμ1 inserts

PelB forward and Cμ1-*Spe* I shift R primer on clones positive for Cμ1 inserts yielded amplicons 1.1Kb in length (clones marked $C\mu1^{11}$ and $C\mu1^{21}$). The middle panel is a control where Cκ was used as a reverse primer instead Cμ1-Spe I shift R to check for the specificity of PCR on such colonies. The positive clones were finally confirmed for reading frame by sequencing. Lane M is 1 kb DNA ladder from Invitrogen.

These manipulations created a single chain antibody (scAb) protein comprising a scFv moiety with affinity for HEL fused in frame to Cμ1 derived from either BTA11 or BTA21. This is shown schematically in Figure 54. In order to simplify the purification of the scAb proteins, a histidine purification tag was added to the constructs. DNA was digested with *Spe* I and dephosphorylated to prevent self-ligation. A synthetic duplex was then inserted. Colony PCR was

employed to identify successful ligation products using the primer pair PelB/SKPR.

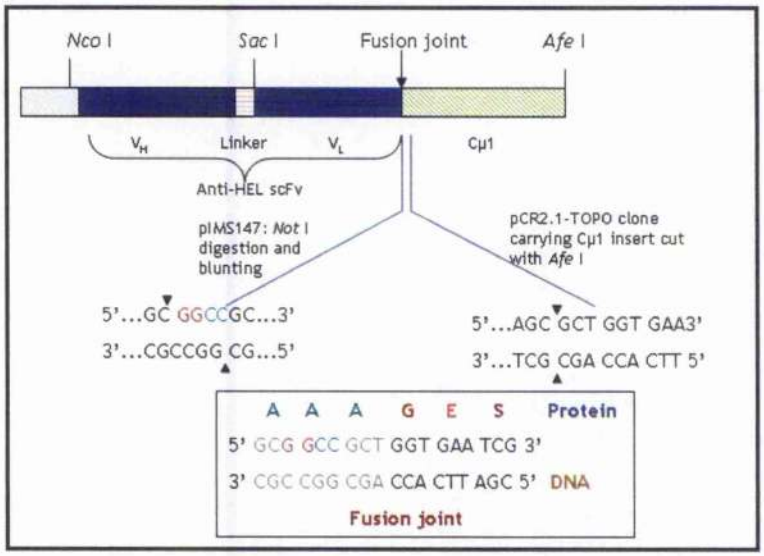


Figure 54: The scAb expression cassette showing features of the fusion joint

Coloured nucleotides were inserted as a result of 5'→3' polymerase activity of Klenow fragment.

The insertion of the histidine coding sequence yielded amplicons slightly bigger (>30bp) than those that lacked it. This analysis is shown in the Figure 55.

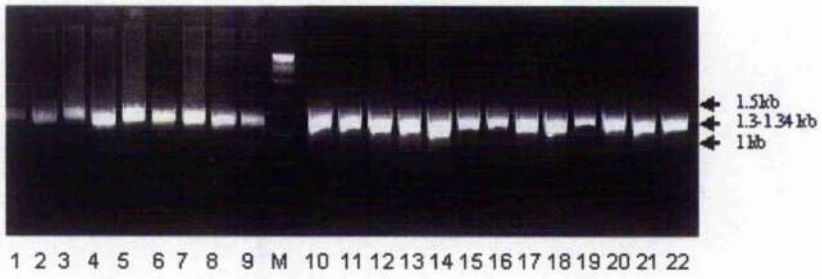


Figure 55: Colony PCR of clones screened for His-tag insert

PelB forward and SKP reverse primers were used with clones positive for his-tag adapter, yielding amplicons 31 bp bigger (e.g. clones 3 , 5, 19) than those without (e.g. 4, 18). Clones were confirmed by sequencing. Lane M is 1 kb DNA ladder from Invitrogen.

Candidates from this analysis were then sequenced, confirming the accuracy of the modification.

6.3.3: Expression of scAbs

Constructs were transformed into *E. coli* TG1 for expression analysis. This bacterial strain was chosen because in studies of the expression of recombinant bovine Fab fragments, it generated higher yields of protein than other strains

(O'Brien *et al.*, 1999). The analysis by SDS-PAGE of extracts from cultures induced with IPTG or grown without inducer failed to show appearance of protein of the expected size (~40 kDa) in an IPTG-dependent fashion. Attempts to optimise expression by varying the temperature of growth or the IPTG concentration did not improve the levels of $C\mu1^{11}$ and $C\mu1^{21}$ scAbs as determined by SDS-PAGE gels. Attempts to detect the recombinant protein on Western blots with an anti bovine IgM antibody resulted in non-specific signals (Figure 56).

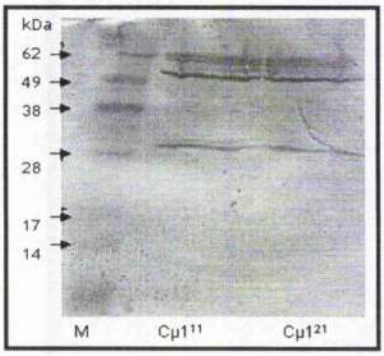


Figure 56: Western blotting of expressed Cμ1 scAbs

Periplasmic extracts from induced cultures were separated by SDS-PAGE, transferred to membrane and probed with polyclonal anti-bovine IgM antibody. M is protein molecular weight marker

In an attempt to improve the sensitivity and specificity of detection, an ELISA procedure was used. This made use of the specificity of the scFv component of the protein for HEL. Microtitre plates were coated with HEL, bacterial extracts were applied, and then detection attempted with antibodies against the constant domain of the scAb. It was hoped that the format of this assay would specifically capture scAb to the HEL-coated surface of the microtitre plate. The original anti-HEL scAb from pIMS147 was used as a positive control for expression, detection being made with antibody to its Cκ domain. A range of blocking reagents was employed as detailed in Table 11.

The analysis revealed that conditions employed for growth, induction and fractionation of bacteria were suitable for the detection of scAb carrying the human Cκ domain. However, it appeared that detection reagents for bovine IgM were incompatible with use of BSA or skimmed milk as blockers.

Table 11: Screening of different blocking reagents for application in ELISA

Detecting antibody	*Blocking solution*	*a*	*b*	*c*
Anti-human Cκ	3% BSA or 2% skimmed milk	Y	N	N
Anti-bovine IgM	3% BSA or 2% skimmed milk	Y	Y	Y
Anti-human Cκ	rabbit serum (1:10 diluted)	Y	N	N
Anti-bovine IgM	rabbit serum (1:10 diluted)	N	N	N

The wells were coated with hen egg lysozyme (10 µg/ml). Proteins (a-c) were detected using different detecting antibody-HRP conjugates under different blocking conditions. a. Original clone expressing Human Cκ and scFv against HEL b. Clone expressing bovine Cµ1 and scFv against HEL c. Negative control (clone expressing bovine AID in an unrelated vector)

Using blocking reagents that avoided non-specific binding of the anti-bovine IgM antibody, expression of scAb carrying bovine Cµ1 domain could not be detected (Figure 57).

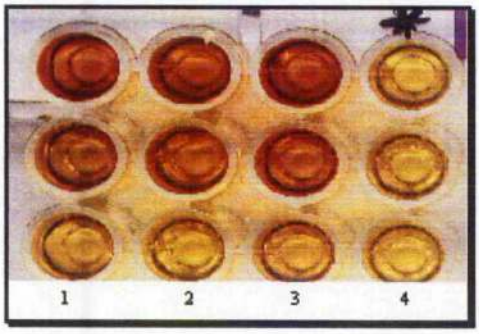


Figure 57: Screening of different blocking reagents for application in ELISA

All the wells were coated with hen egg lysozyme (10µg/ml). Wells in columns 1 to 3 received lysate from a clone expressing human Cκ and those in column 4 received lysate from a clone expressing Cµ1[21]. Wells in the column 1, 2, 3 and 4 were blocked with rabbit serum, 2% BSA, skimmed milk and rabbit serum, respectively. BSA and skimmed milk were previously tested as blocking agents in Cµ1 ELISA and were found to react non-specifically with anti-bovine polyclonal IgM. Rabbit serum was found to show a specific blocking and was used in future ELISA to detect the expression of Cµ1.

It remained formally possible that the anti-bovine IgM antibody could not recognise the Cµ1 domain of the scAbs. However, attempts to capture scAbs by affinity chromatography using media containing bound nickel failed to yield protein that was detectable in SDS-PAGE. It therefore, seemed more likely that a fundamental problem existed in expression of the recombinant proteins. This might have arisen through their toxicity to the host, low conformational stability and/or rapid proteolytic degradation.

6.3.4: Expression of Cμ1[11] and Cμ1[21] as independent proteins

Since Cμ1 domains failed to express as scAbs in *E. coli* TG1, an alternative strategy was adopted using expression in a T7 promoter-based vector. Potentially, this system allows the production of large quantities of cloned isolated domains. The expression with the pET system is cytoplasmic where as in case of pIMS147 vector system; the scAbs are transported to the periplasmic space by the Pel B leader. The PCR with PET cloning primers (section 6.2.4) recovered specific Cμ1 domains, which were then cloned into *Bam* HI and *Hind* III digested pET21-a vector. This provided in-frame fusion of the His-tag for purification which was confirmed by sequencing.

The constructs were then transformed into BL21 (DE3), an *E. coli* strain. This strain is genetically engineered to incorporate in its genome a gene for T7 RNA polymerase under the control of the *lac* promoter and the *lac* operator. When lactose or a molecule similar to lactose (IPTG) is present inside the cell, transcription of the T7 RNA polymerase is activated. Bacterial cultures were grown at 37°C to an absorbance at 600 nm of 0.6 to 0.8 and were then induced

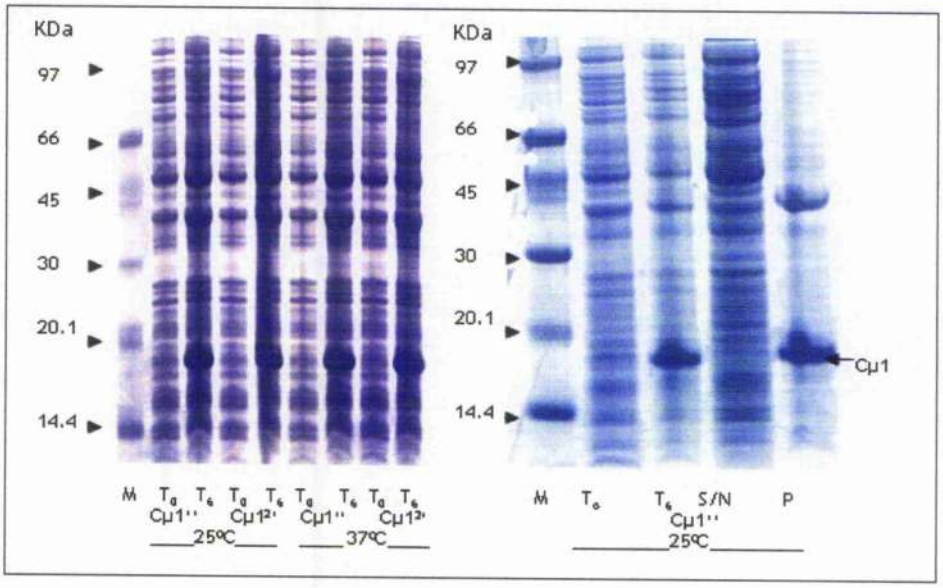


Figure 58: SDS-PAGE gel taken before and after the end of induction

After 6 hours of induction at 25°C and 37°C the appearance of strong bands towards the end of gel (16 kDa) are evident after equal amount of samples were analysed (left gel). Protein solubility was tested by lysing the cells in binding buffer with cocktail of protease inhibitors and assessed by SDS-PAGE analysis of supernatant and cell pellet (right gel). Even after the solubilisation, proteins remained in the inclusion bodies in the pellet and none is released in the supernatant. M is a protein marker.

with 500mM IPTG for 6 hours at 37°C, 30°C and 25°C. Efficient expression of recombinant Cμ1 domains was observed at all the temperatures employed, however the protein was expressed as insoluble inclusion bodies, shown in Figure 58 below.

To improve the solubility of the recombinant proteins the cultures were grown at 37°C to an absorbance at 600 nm of 0.8 and then induced with 500 mM of IPTG for 14-16 hours at 18°C. The Cμ1 domains were detected in soluble fractions by ELISA but Western blot analysis with anti-IgM reagent yielded non-specific bands as before. Given the complexity of cytoplasmic extracts, substantial purification was achieved with single IMAC step (Figure 59).

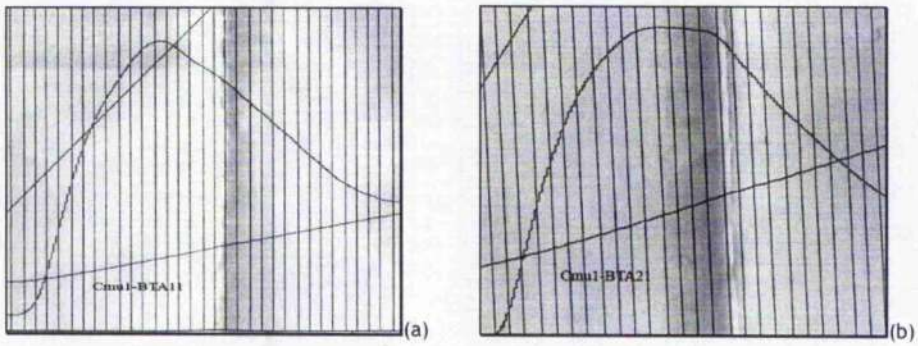


Figure 59: Strip chart recorder windows from BioCAD 700E workstation displaying the eluate fractions. The fractions were pooled together for $C\mu1^{11}$ (a) and $C\mu1^{21}$(b) and further refined by gel-filtration.

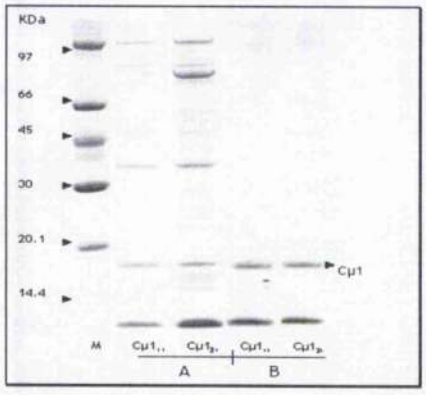


Figure 60: SDS-PAGE analysis of recombinant protein samples after purification

His-tagged pET-21a/Cμ1 variants were first purified using immobilised metal-chelate affinity chromatography (A) and then refined by gel-filtration chromatography (B).

In SDS-PAGE analysis of the eluates from IMAC, a prominent protein of approximately of 17 kDa is evident (Figure 60). The expected size of isolated Cμ1 protein was 15.15 kDa, suggesting that this band could be Cμ1. The eluted

fractions after IMAC (Figure 60, panel A) contained other proteins of both high and low molecular weight, most likely components of the *E. coli* cytoplasm that happened to be histidine-rich in composition. Therefore, an additional purification step of gel filtration chromatography was introduced. Substantial enrichment of species of interest was achieved (Figure 60, panel B).Excision of the band at 17 kDa and analysis by mass spectrometry confirmed it to be the Cμ1 protein (Figure 61).

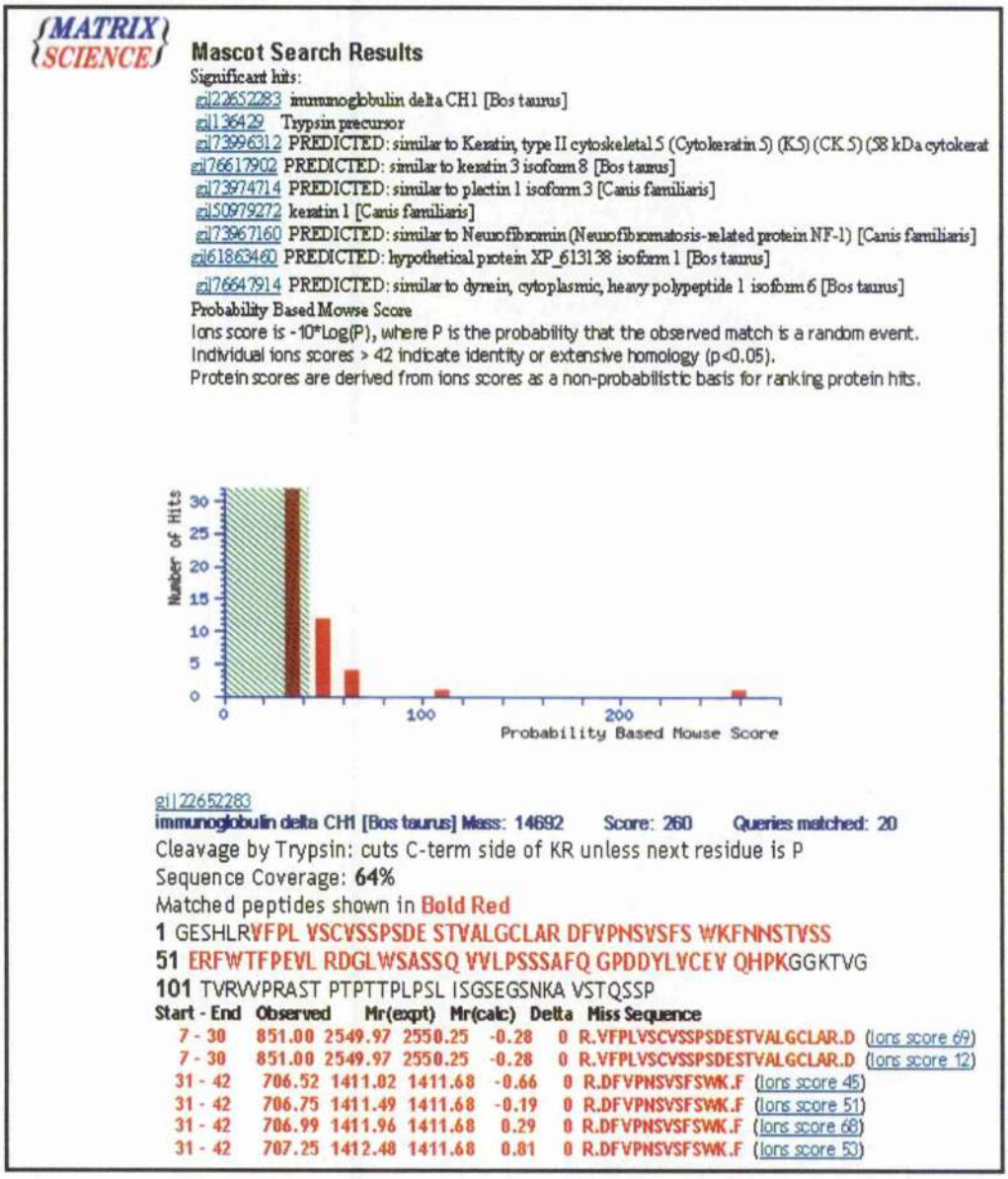


Figure 61: Modified screen shot from a Mascot search data base

Tryptic digest of a purified fraction excised from the gel was analysed by MALDI MS. The search returned a hit with a high score and matched to CHI constant domain with 64% sequence coverage. The protein was identified Cμ1 domain.

Separation from the other, unknown protein of around 11-12 KDa could not be easily accomplished and hence this material was used as target for phage display selection of anti-Cμ1 scFv.

6.3.5: Preparation and titration of phage library

TG1 bacteria comprising I and J libraries were grown in 2xTY medium and infected with KM13 helper phage. Phage displaying the scFvs were then isolated from culture supernatants by precipitation with PEG/NaCl and resuspended in PBS. The size of the Tomlinson combined I and J libraries were then determined by infection of *E. coli* TG1 cells with serial dilutions of phage and counting the number of ampicillin-resistant colonies that then grew on TYE plates. The titer of each phage library as calculated in this way was found to be $5x10^{15}$ tdu/ml. It was decided to use both the libraries together for selection to increase the chances of isolating specific binders. Although both the libraries are based on a single human framework for V_H and Vκ, in Tomlinson I library, side chains are diversified by DVT as compared to Tomlinson J library which has got NNK side chain diversity. Contamination of the target (i.e. Cμ1) with another unknown protein might have complicated the selection of phage specific for Cμ1. However, the selection strategy aimed to recover phage reactive with features present in selection stage but absent in depletion stage. This strategy envisaged bias away from sites shared between $C\mu1^{11}$ and $C\mu1^{21}$. It was also predicted to eliminate or at least reduce the isolation of phage reactive with the unknown contaminant. The selection strategy was heavily dependent upon the efficiency with which phage were removed in the initial depletion step. However, quantities of protein available for phage display were very limited (9 μg in total) so multiple depletion steps that might increase efficiency was not feasible. Inputs to each round of selection ranged from $5x10^{11}$ to $1.1x10^{12}$ tdu (Table 12). Given that library is thought to contain around $1 x10^8$ unique specificities, this represented several thousand-fold representation of each clone. Outputs from the selection stage were modest, representing at Round 1 about $4x10^{-8}$ to $1x10^{-7}$% of the numbers of virus that entered the round. Whilst low, these recoveries are similar to those noted in other studies in this lab (Golchin, 2004). It is also important to recall that some phage are likely to have been lost during the depletion phase of each round as indicated by the higher recovery from direct

selection on $C\mu1^{21}$. Whilst this direct selection protocol produced the most convincing evidence of enrichment, round on round (Table 12), this was also evident in the depletion/selection protocol for isolation of scFv binding to $C\mu1^{11}$-unique features. Here, a large enrichment was noted from round 1 to round 2, but this was not apparent in the final round of selection.

In order to establish whether phage recovered in this way were reactive and specific in interaction with the recombinant protein, monoclonal phage ELISA was conducted. Initial tests were carried out with a capture ELISA that recovered IgM from bovine serum. This was designed to test if the scFv displayed by individual phage clones would recognize epitopes on the native protein.

Table 12: Phage titres from different rounds of selection

Selection	*Input*			*Output*			*% recovery*		
	$C\mu1^{11*}$	$C\mu1^{21*}$	$C\mu1^{21**}$	$C\mu1^{11*}$	$C\mu1^{21*}$	$C\mu1^{21**}$	$C\mu1^{11*}$	$C\mu1^{21*}$	$C\mu1^{21**}$
Round 1	1.1×10^{12}	1.1×10^{12}	1.1×10^{12}	$1.1x10^{5}$	$3x10^{5}$	$7.2x10^{6}$	$1.0x10^{-7}$	$0.4x10^{-7}$	$0.6x10^{-6}$
Round 2	$5.0x10^{11}$	$8.0x10^{11}$	$1.2x10^{12}$	$9.8x10^{6}$	$1.5x10^{6}$	$2.5x10^{7}$	$0.5x10^{-5}$	$5.3x10^{-5}$	$0.5x10^{-5}$
Round 3	$5.4x10^{11}$	$3.9x10^{11}$	$7.6x10^{11}$	$1.2x10^{7}$	$9.2x10^{6}$	$1.7x10^{7}$	$4.5x10^{-4}$	$0.4x10^{-5}$	$4.4x10^{-4}$

Input and output phage titres were determined by transduction of *E. coli* to ampicillin resistance using serial dilutions of samples (*- represents depletion-selection strategy; **- represents direct selection).

A total of 64 bacterial colonies (32 from rounds 2 and 3 for $C\mu1^{11}$ and similarly for $C\mu1^{21}$) were picked at random during selection. These individual colonies were grown, infected with helper phage, and culture supernatants were then used in monoclonal phage ELISA. The phage samples were added to wells of ELISA plates coated with bovine serum bound to polyclonal anti-bovine IgM. Binding was tested with anti-M13 phage-specific reagents and OPD substrate (Figure 62). Figure 62 (a) shows that after a second round of selection in case of $C\mu1^{11}$, only one phage antibody selected from the libraries had activity against IgM. This is in contrast to third round of selection where three phage antibodies bound IgM with much more intensity. Figure 62 (b) shows all the 64 clones tested for $C\mu1^{21}$ on wells to which bovine IgM had been captured. A similar trend is evident. After second round of selection only 2 phage antibodies showed reactivity as compared to 4 phage antibodies after round three. This round also showed a strong reaction towards IgM as compared to phage from previous round. These results indicate that selection for recombinant phage antibodies against Cμ1 domain by phage display was successful.

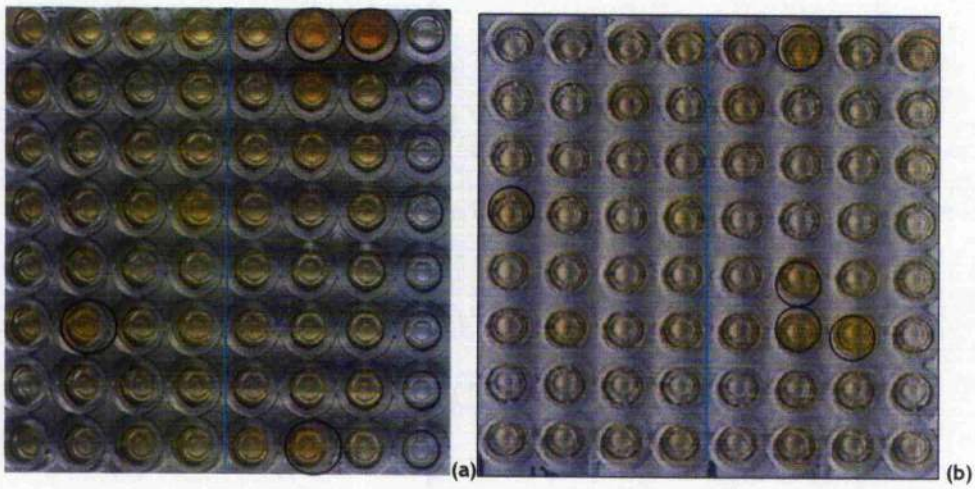


Figure 62: Monoclonal phage ELISA against $C\mu1^{11}$ and $C\mu1^{21}$ after depletion-selection

Thirty two clones picked at random from 2nd and 3rd round of selection (as indicated) were superinfected with helper phage to generate monoclonal phage antibodies. The phage were then applied to ELISA plates coated with bovine serum bound to polyclonal anti-bovine IgM and binding detected with anti-M13 reagents and OPD substrate. Final ODs were measured at 450 nm (data not shown) and images of plates were captured. The clones encircled were used for subsequent analysis on purified fractions of $C\mu1^{11}$ and $C\mu1^{21}$.

Having confirmed this property for 4 clones selected against $C\mu1^{11}$ and 4 clones from selection against $C\mu1^{21}$, the analysis moved to tests with the

recombinant proteins, since these could establish specificity for $C\mu1^{11}$ and $C\mu1^{21}$ -unique epitopes. The phage antibodies when tested for their respective specificity with $C\mu1^{11}$ domain gave cross-reaction with $C\mu1^{21}$ and vice-versa. Hence, no variant specificity could be detected.

6.4: DISCUSSION

The experiments described in this chapter were aimed at the development of antibodies specific to two variants of bovine IgM that differed in the amino acid sequences of the Cμ1 constant domains. Since the two target molecules shared amino acid sequence identity for the most part, the phage display screen aimed at pushing selection towards the areas of difference through alternating stages of depletion and selection. Given that the targets were very similar, conventional phage display techniques would likely have recovered recombinant antibodies reactive with shared features of the molecules. This would have yielded phage cross-reactive with the IgM variants. The approach taken in the present study thus involved the selection of phage antibodies from a synthetic display library against recombinant proteins displaying unique epitopes. This was achieved by allowing the phage reactive to shared epitopes to bind first a recombinant protein against which depletion was directed and then applying this mixture to select the phage to unique epitopes on a second recombinant protein against which selection was desired. Three such rounds were performed. This was performed in order to enrich phage antibodies binding to specific residues while avoiding those binding to common residues between the Cμ1 domains.

The serum levels of IgM from BTA11 are not known and without a convenient method to separate this protein from IgM derived from BTA21, recombinant methods were followed to generate target material for screening. Initially scAbs were constructed by fusing a scFv against HEL with the Cμ1 variants. Some scAbs express to high yields (Hayhurst and Harris, 1999). scAb against HEL carrying the human Cκ constant domain could be expressed and detected in this study, but scAbs carrying Cμ1 domains could not. Reports in the literature suggest that expression of recombinant Fab proteins, including those derived from IgM, is possible though yields may vary (Jahn *et al.*, 1995). In these proteins, chain interactions take place through the constant domains of the light

and heavy chains. scFv fusion to IgG constant domains have also been reported (Muller *et al.*, 1998); again interaction between chains likely occur through the CH2 and CH3 domains. The scAbs constructed here may have been destroyed by proteolysis, accumulated in a cellular compartment or formed aggregates through the hydrophobic surfaces exposed on the Cμ1 domains that would normally interact with the light chains constant domain. Having failed to detect expression, an alternative strategy was followed - expression of Cμ1 variants as isolated proteins.

This approach proved more successful and the identity of the purified protein was confirmed by mass spectrometry. The $C\mu1^{11}$ and $C\mu1^{21}$ proteins were then used as targets for phage display selection. On one level, the screen was successful and phage displaying scFv able to bind bovine IgM were successfully isolated. However, further analysis did not produce convincing evidence of variant specificity as had been anticipated from the design of the selection strategy.

Subtractive methods have been successful in isolating antibodies to unique self-epitopes which include phage antibodies against human leucocytes and erythrocytes (de Kruif *et al.*, 1995; Marks *et al.*, 1993), human melanoma cells (Cai and Garen, 1995) and many more. Phage antibodies were selected against immobilized crude lysates from cultured human keratinocytes that were specific for the differentiated state of the cells, *in vitro* (Stausbol-Gron *et al.*, 2001). This work employed a substraction - selection strategy. In another study, blockade of an apparently dominant, nonneutralizing epitopes led to the isolation of two new antibodies, one of which had neutralizing activity (Tsui *et al.*, 1996). This success suggested the possibility of mining diversity from libraries by successive panning in the presence of representative antibodies recovered in previous panning experiments.

Were these experiments to be repeated, more thorough depletion of phage reactive with shared epitopes might improve the chances of success. This was not possible as quantities of the recombinant Cμ1 proteins were limited and insufficient time was available for studies to increase yields. Were reagents available to distinguish the variants of IgM, it would be relatively easy to carry out capture ELISA assays to measure levels of the proteins in bovine serum. Using

the recombinant Cμ1 proteins as standards, levels of IgM could be quantified in absolute terms. The reagents could also be used to isolate and purify B cell populations expressing one or other of the variants of IgM. This would enable detailed investigation of the formation of these Ig molecules (for example, is rearrangement at BTA11 possible, or do transchromosomal processes operate), their relative levels of diversification, and the distribution of B cell subsets in lymphoid tissue.

An alternative strategy to meet the objectives set would have been to use short peptides derived from the variant-specific regions of the IgM Cμ1 domains. These peptides could then have been employed for selection. However, as shown in the Results section, the variant-specific residues are not clustered into contiguous amino acid sequence and thus it is questionable if peptides could be synthesized which would be long enough to select clones with the required specificity and yet sufficiently different to avoid the isolation of cross-reactive phage. Another complication is that shorter peptides have been shown to bind poorly to solid surfaces (Loomans *et al.*, 1997) that might be used for selection.

When aiming at subtraction between two components, the ideal situation is to have selection against all differences concurrently with the inhibition of selection against common features of the targets. The dilemma remains that attempts to improve the efficacy of the selection by using conditions of high stringency during depletion will tend to decrease the diversity available for positive selection because some binders are lost. As demonstrated by the results of this study, insufficient stringency during depletion affected the outcome of this depletion-selection strategy and the specificity of the recovered clones was a major issue of concern. The balance between these factors cannot be easily reached except by trial and error. Given the limited availability of the target protein and constraints of time at this point in the project, further attempts to achieve the optimum balance were not possible.

CONCLUSIONS

The work described in this thesis is the first definitive study of its to investigate the process of Ig diversification. Novelity of the experimental design provided the power to distinguish between SHM and GC by studying within the non-coding regions of the JH locus and downstream intronic sequences. This provided a level of discrimination for greater than could be achieved by study of coding regions of the heavy chain locus where information is fragmentary and incomplete. The stringent criteria laid down for the identification of diversifying mutations overcame problems arising from the incomplete understanding of the IgH locus, particularly with regard to the numbers and sequences of VH segments and their alleles. As with any scientific study, further improvements could be made. A high-quality Taq polymerase was used to recover sequence from genomic DNA, but it is acknowledged that proof-reading polymerases are better suited for this sort of work by virtue of their ability to correct mismatches between template and product, thereby reducing overall errors during amplification. This study attempted to measure the Taq error rate under the experimental conditions employed, over the modest sized amplicons that were generated. No attempt was made in this study to investigate the boundaries of the mutated region of the IgH locus. This could have been gathered from longer PCR products and potentially could have been of value in establishing the nature of cis-acting motifs and providing clues to the targeting mechanisms. The observations reported herein on the nature of the mutation process could be strengthened by the use of more animals per time point and by inclusion of other sampling points (e.g. foetal material). Samples were taken from outbred animals, but if biopsies from animals of defined pedigree had been available, this would also help minimise the complication of variation between haplotypes. Whilst these improvements would have enhanced the quality of the data, it is unlikely that they would have altered the central finding of the study, the implication of SHM as a process for diversification of the bovine Ig heavy chain repertoire.

One aspect of Ig diversification that was not studied for logistic reasons was the chemical or biological triggers to SHM that appear active late in gestation and immediately after birth. Given the strong likelihood that diversification is active in the IPPs, this area could be studied by the use of

sterile closed intestinal loops, with and without IPPs. Individual loops constructed in this way can act as a independent sites to which different antigens and triggers including food components maternal factors, and bacteria could be delivered. Activation of the diversification response could then be monitored in B cells of IPPs. This type of system has been applied in ovine immunology and has proven very informative. Experiments of a similar nature have also been conducted in pigs and rabbits and have provided a plethora of information on the role of microbes and GALT in the induction and development of immune responses (Butler *et al.*, 2000). The other approach could be the study animals raised in an SPF environment. The economics and logistic challenges of these experiments would however be substantial.

By characterizing AID, an enzyme known to play a crucial role in SHM, GC and CSR in all vertebrate systems studied to date, the study has provided a foundation for future studies of the processes. The analysis of AID reported here is first comprehensive study of the bovine enzyme. The bovine AID gene was partially sequenced and thus gave clues on its structural organisation within the bovine genome. Cloning of AID cDNA and definition of its functional and chromosomal localization has provided tools for future work on cytogenetics and understanding the role of enzyme in antibody diversification. Characterisation of AID- induced mutations in *E. coli* will provide more information on the nature of residues targeted and hot spots if any. However, better assessment of mutagenic activity of AID could be made in a mammalian system. This sort of investigation has been carried out by Barreto and colleagues (Barreto *et al.*, 2005) in a study of zebrafish AID, expressed in murine B cells. Although expression of bovine AID was tested in lymphoid and non-lymphoid tissues in our study, quantitation was not possible. With the availability of bovine AID cDNA that emerged later specific primers could now be designed for real time RT-PCR to assess levels of the AID expression. This could be used to measure where and when Ig diversification occurs in bovine tissues. To some extent, further dissection of the processes activated in cattle can take as foundation the clearer understanding of other veterinary species. For example, it is far from clear whether in cattle, the IPP serves as the site for Ig diversification, but studies in sheep have compared events in the IPP with events in jejunal Peyer's patches (Yasuda *et al.*, 2006). This work has shown that the IPP serves as a primary lymphoid tissue in which Ig

diversification takes place, whereas the JPP is more secondary in its characteristics forming a site for mucosal immune reaction in an antigen-specific fashion. With these findings in mind, formal confirmation of the role of the IPP in cattle could be made through experimental investigation helped by the data on AID that is now available.

Of equal importance are the processes that govern entry to and survival within the organ of diversification. This area has been dissected in chickens. Through transgenesis, Ratcliffe and colleagues have shown that entry to the bursa of Fabricius is dependent upon expression of Ig at the surface of B cells (Pike *et al.*, 2004). However, in recent studies, this group has shown that this property alone is insufficient to ensure continued development, instead survival and diversification is dependent upon ligation of receptor molecules at the B cell surface (Aliahmad *et al.*, 2005). This system benefits from the availability of retroviral vectors for transgene expression, easy access to embryonic material, and an experimental system that is better defined than cattle, but it provides a conceptual framework for assessing events in ruminants. The elegant separation of entry to the GALT and survival/diversification within the tissue highlights a further crucial question: what initiates diversification of B cells resident at this site? Again, other livestock systems are providing clues for further studies in cattle. The primary B cell repertoire in humans and mice is generated by rearrangement of multiple V, D, and J gene segments in the bone marrow throughout life. In species like the rabbit, chicken, sheep and pigs, the pattern of events is strikingly different. Firstly, the B cell repertoire develops by rearrangement of a limited number of V genes in primary lymphoid tissue that has been defined with variable precision amongst these species. Since little diversity is generated and rearrangement may be limited in its duration, post-rearrangement diversification is necessary and this takes place in the GALT. It has been shown in rabbits that surgical disruption of the interactions between GALT and the intestinal microflora prevented GALT development, B cell expansion, and somatic diversification of the B cell repertoire (Lanning *et al.*, 2000). Studies conducted on pigs have also provided similar observations (Butler *et al.*, 2000). A recent finding that the repertoire shift from V_Hn to V_Ha B cells in rabbits also depends on GALT-bacterial interactions (Rhee et al., 2005a) has led to the identification of sites on the exterior surface of the V_H region at which

a putative bacterial B cell superantigen or a bacterially induced GALT-derived superantigen might bind and thereby stimulate V_Ha B cells. A similar V_H binding site already exists for *S. aureus* Protein A on human V_H Ig molecules (Graille et al., 2000a) which binds to and preferentially stimulates B cells that use V_H gene segments of the V_H3 family (Silverman and Goodyear, 2002). Other studies have demonstrated that B cell expansion in GALT is specific-antigen independent (Casola *et al.*, 2004).

The potential for bacterial DNA to activate diversification through Toll-like receptors, and the role of maternal factors has been highlighted through investigations in pigs (Butler *et al.*, 2006). Studies from other species illustrate what has yet to be discovered about Ig formation in cattle whilst providing clues as to the events that may be taking place. The major questions are what sorts of B cells enter the GALT and when are they recruited? How is B cell development before and after birth regulated and in particular what are the contributions of exogenous antigens, maternal factors, developmentally-regulated events, and environmental influences? Of particular interest, does the intestinal microflora stimulate B cells in a BCR-dependent or independent manner and can the mutations that create diversity be distinguished from those that occur in response to specific antigens in the secondary lymphoid organs? Also, as cattle are unlikely to maintain continuous production of B cells as happens in mouse bone marrow, how can B cells in the periphery survive for the life-time of the animal to ensure humoral protection upon challenge late in life?

One major obstacle to resolution of these questions is the logistics of working with large animals that are expensive to obtain and house (compared with work using chickens or rabbits). Another problem is the lack of appropriate B cell markers and reagents for detecting their expression. For example, one can envisage studies of the differences between pre-term calves and full-term calves as to how Ig diversification changes during first few weeks after birth. The expense of this sort of experiments is obvious. Equally, experiments to study entry of B cells to the IPP and selection steps for expansion and proliferation are frustrated by the lack of markers.

Perhaps in compensation, the size of the animals would allow surgical procedures that are otherwise difficult to perform in other species (e.g.

construction of intestinal loops). The system also allows greater experimental control over exposure to maternal factors because of the differences in placentation between cattle and other animals and the fact that immunoglobulins are not transported across the bovine placenta from dam to fetus. Barring fetal infections, the young neonate therefore is born without circulating antibodies. Another facet of bovine immunology that provides novelity is the expression of heavy chain isotypes from duplicated loci on BTA21 and BTA11. If, as attempted in this study, reagents can be developed for separation of B cells that express these variants, better definition of the basis to this phenomenon could be made. This is of importance to the scientific community as it places the accessibility hypothesis of Yancopoulos and Alt in a novel context. Bovine isotypes of IgM also illustrate how understandings of biological processes in mammals are not always securely anchored when extrapolated from studies in rodents. All these issues need to be balanced and confronted in order to take forward our understanding of the bovine B cell immunology.

APPENDIX I

Luria-Bertini Agar (LA)
Tryptone, 10 g ;Yeast extract 5 g; NaCl 10 g; Agar 6 g, Distilled water up to 1 litre.

Purification on CnBr-activated sepharose

Coupling Buffer (for coating HEL)
0.1M NaHCO3 pH 8.3, 0.5M NaCl

Phosphate-Citrate Buffer
Add the following to create 100 ml of phosphate/citrate buffer solution. Stock solutions are 0.2 M dibasic sodium phosphate; 0.1 M citric acid (Pearse, 1980)

0.2 M Na_2HPO_4 (ml)	0.1 M citrate (ml)	pH
5.4	44.6	2.6
7.8	42.2	2.8
10.2	39.8	3.0
12.3	37.7	3.2
14.1	35.9	3.4
16.1	33.9	3.6
17.7	32.3	3.8
19.3	30.7	4.0
20.6	29.4	4.2
22.2	27.8	4.4
23.3	26.7	4.6
24.8	25.2	4.8
25.7	24.3	5.0
26.7	23.3	5.2
27.8	22.2	5.4
29.0	21.0	5.6
30.3	19.7	5.8
32.1	17.9	6.0
33.1	16.9	6.2
34.6	15.4	6.4
36.4	13.6	6.6
40.9	9.1	6.8
43.6	6.5	7.0

Immobilised Metal Affinity Chromatography on HiTrap SP XL $NiSO_4$ coated column:

Binding Buffer
20mM phosphate buffer pH 7.2, containing 0.5M NaCl and 20mM imidazole.

Elution Buffer
Same as binding buffer with different molarities of imidazole (50mM, 70mM, 100mM and 200mM)

Immobilised Metal Affinity Chromatography on Zn++ coated MC column (BioCAD 700E)

Binding Buffer
20mM Hepes, 100mM NaCl, 5mM Imidazole, pH8.0 (Metal chelate chromatography)
50mM Tris-HCl, 2mM EDTA, pH 7.5 (Gel filtration)

Elution Buffer
20mM Hepes, 100mM NaCl, 500mM Imidazole, pH7.0 (metal chelate chromatography)

SDS-PAGE

Running Buffer (SDS) 5X -500ml

Tris Base	-	4.5 gm
Glycine	-	21.6g
SDS	-	1.5gm
Water to	-	500ml

3 x SDS Sample Buffer (10ml)

1M Tris-HCl pH 6.8	2.4 ml
20% SDS	3.0 ml
Glycerol	3.0 ml
β-mercaptoethanol	1.6 ml
Bromophenol Blue	0.006 g

Trypan blue solution
Trypan blue 0.2 g; NaCl 4.25 g; dissolve each component individually in 100 ml distilled water, then mix 4 parts of trypan blue with 1 part saline

Western blot

Transfer buffer- 1 Litre
25mM Tris Base-3.03g
192 mM Glycine-14.4g
20% v/V Methanol 200ml
Volume to 1 litre

Blocking solution
3% skim milk (25 ml) 750mg in 25 ml PBS or 1.5g in 50 ml

Ponceau S solution
0.1 % w/v Ponceau S (Sigma); 5 % v/v acetic acid

Reagents for phage display

2xTY broth
16g Tryptone, 10g Yeast extract and 5g NaCl in 1 litre

TYE Agar
15 gm Bacto-Agar, 8 gm NaCl, 10 gm Tryptone, 5 gm Yeast extract in 1 litre

Trypsin stock solution
Trypsin (Fluka, UK) solution (10mg/ml) in 50mM Tris-HCl pH 7.4, 1mM $CaCl_2$

PEG/NaCl
20% polyethylene glycol 6000, 2.5 M NaCl

Reagents for FISH

20XSSC
3M NaCl, 0.3M Na Citrate, stored at room temperature

Hybridisation Buffer
4 ml of 50 % dextran sulphate (Sigma), 10 ml formamide (Fluka), 1 ml 20xSSC, 5 ml sterile DW, all mixed and aliquoted in 1 ml volumes and stored at -20°C

50% formamide
25 ml formamide (Fluka), 25 ml 2xSSC, mixed fresh before use

Wash A Solution (Blocking solution)
2.5 gm dried skimmed milk, 10 ml 20xSSC, 25 µl Tween 20. The solution is mixed with a magnetic stirrer for 10 min and centrifuged at 14000 rpm for 10 min in a 50 ml polypropylene centrifuge tube. 5 ml aliquots are stored in 5 ml plastic bijoux containers at -20°C until required.

DAPI Counterstain
40µg/ml stock solution of 4', 6-diamidino-2-phenylindole dihydrochloride (DAPI) in DW and stored in aliquots in -20°C. Before use dilute approximately (1:5) in Citifluor AF1

Biotinylated anti-avidin (Vector Labs)
2 mg/ml stock solution is prepared in DW, filtered and stored at 4°C.

Miscellaneous

Antibiotic	Stock conc.	Working conc.	Plasmid
Ampicillin(Amp)	100mg/ml(DW)	100µg/ml	TOPO 2.1, pBAD/Myc-His, pIMS147, pET21a
Kanamycin	50mg/ml(DW)	50µg/ml	TOPO-XL

SOC Medium
Tryptone 20 g; Yeast extract 5 g; NaCl 0.5 g; Distilled water up to 980 ml; after autoclaving, 10 ml of 1 M MgCl2 and 10 ml of 1 M MgSO4 added.

Phosphate Buffered Saline (PBS)
NaCl 10 g; KCl 0.75 g; Na2HPO4.12H2O 1.44 g; KH2PO4 0.125g; Distilled water up to 1 litre.

TAE Buffer (50 x)
Tris base 242 g; Acetic acid 57.1 ml; 0.5M EDTA 100ml; Distilled water up to 1 litre.

6 x DNA loading buffer
Tris base 60mM; EDTA 6mM; Sucrose 40 % (w/v); bromophenol blue).25 % (w/v)

0.1M Citrate Buffer
Na2HPO4.7 H2O 18.1 g; Citric acid 9.42 g; adjust pH to 5 and bring volume to 1 litre

APPENDIX II

Partial sequence of *B. taurus* AID

EXON II (PARTIAL)

ACTTGTGCTACGTGGTGAAGCGGCGGGACAGTCCCACCTCCTTCTCACTGGACTTCGGGCACCTTCGAAACAAG

INTRON II

GTATCAATTAATATCTTGCTATTAATTAATTAATGCCTGCTTCCCGAGCAGCTACACTGAGGGCACTTTTATAGCCACCTGCTGATGATG
CTACTCAGCCTGGTTCATTTGTCTCCATTAGATTTCTGAGATATCTTTTCTCTTTTATATGTCTCTGCACCCATAATAGACAGGTTCCAAA
GGCGGTCATTTAACATCCTCCCGAATCCTGCTGGGCACATCAAGGCCACACTTTGTTTCAGTGCTAAGAATCACTGCCCCAAACCTTCCT
CTCCTTTCCCCAGCATCCAAATTGCTTTTCAGATTAAGGAAGAAACCAAAAGAAAGAGATTTTTTATTTGGGGGAATCTTATCAGCTCATT
CAAGCTGAACTGTAAACGTGGAAAAACAAATTGTGGCTTTGGAATTTTAAATAATGAAGGAAGAAGAATTAGGAAAAATTAATGACTCAG
TTTTATCTTTCAAATGATTTCTATTCCTTCCTGCTCACCTAGATCATAAGCCAGCAAGTACATTCATCATAGGTGACCCTCAGCAAAATTC
AGGAAGCCATGAATTAAAAACTAGAAGATCAGAATTTCCCTGGGGCACAGTGGAGAAGAGTCTGCCTGCCAGTGCAGGGGACACAGGTT
CGATCCCTGGTGCAGGAAGATTTCACATGCCCTGGAGCAGCTAAGCCCACGTGCCACAACTACTGAGCTTTCATGCTGCAACTACTGAAC
CCTCGCACCCAAGAGCTCATGTTCTGCAACGAGAGGCCATGAATAGATAAATAATAAAAACTGCTTCTCAATAATATGGCAGCAGTGGTG
GTTTAGTCGCTCAGTTGTCTCCAACTCTTGGGACCCCATGGACTGTAGCCTGCCAGGCTCCTCTGTCCATAGGATTCTCCAGGCAAGAAT
ACTGGAGTGGGCAAGAATACTTGCCATTTCCTTTTCCAGGGAATCTTCCCAACCCAGGGATTGAACTCAGGTCTCCTACATTCCAGGCAG
ATTCTTCACCAACTGAGCAATAACACATCTAGCATTAAACAAAAAAAAAAAAACCTAGGAAGATGGGAGAATTACATTTTCAAGAGACTTCA
TCACCAAAATCCAGAAGAGTTTACATAACATGCAGTTGACATGTGGTCTTGGTGGAAACGCAGGGAAGAATCCCTGGGCAGAATGACAC
TGACGTATCTCACTTTTCCTTGGCCTCTAATTTGTGTGGTGTTCATTTGTCCCACTGCAATTGTTTTAGCTTTCTAATCTTTTCCCTATTT
GGTGATAAGACACATTCTAAACATAGAAAGCTTTTGTGTTCCTTTCCTTCTCTCTCCCTGAGTCTTCACAACAAAACTCCATTTAACTTAT
GAGGAATCAGAAGATCTAAGAGTTGGGTGACTTATGGAACTTAGGTCTTCAGATTCTGGGAGCAATGACTTCTCTGGCACTATGAAGTG
GACTTGCTATATCTCACTATGGTGATTACTATTACAATCATGGGGGTCATACAAGAGCAAACTCATTACAGGCACACCTCGGAGATTTT
GTANAATTTGATTCCAGAACATGGCAATAAAGCATATTGCAATAGAGAATATTGCAATAAAGTGAGTCAAATTTGGGGGTTTCCTGAAAG
TGCTAGTTTTTCAATCCTGTCTGACCCTTTGTGACCCCATGGACTATAGCCTGCCAGGCTCCTCTAGTCCATAGAATTCTCCAGGCAAGA
ATACTGGAGTGGGTTACCATGCCTTCTCCAGGGGATCTTCCCGATCCAGGGGATCAAACCTGGATCTCCCCATTGCAGGCAGATTCTTT
ACCATGTGAGCCACCAAGGAAGCCCTAGTACATACAAAAGTTATATGTTTACACTATACTACAGTTCTTTTAAGTATGCAATACAATTATA
TCTAAAACAAAAAGTGTATATCTTAATTTTAAAGCTTAATGCTAAAAATTGCTATCATCTAAGCCTTCAATGAGTCATACTAGTAACAAAG
ATCACTGACCACAGATAACCATAATAATGAAAAGTTTGAAATTTGCAAGAATCACGAACTGCGACACAGAGACACAAAGTGAGTGGATGC
TGCTGGAAAAAATGACACCAGGGTTGCCACAAACCTTCAGAACACAAAGTCTCTGCAAAGTCCAATATAGCCAAGCACAATAAAACAAGG
GACCCCTGTATTTAATTGTTCTTAAACCACAACCGGTGCCTACTACCTCTTTTACAACACCCACTTGTCAGTCTCAGACTCTCCTAACCCT
AACTTTTTATTCTCCACTTGGTGTAGTGGAGTCCAGTGTATTTTCACAACGGTTACATCAACAGCACTTCCAAGCCGTTTCTTCCTCTCCA
AGGAAGGAAAAGCAACTTCACAAACAAAATTAAATCTGAGATGGTGTGATGCTACTTGACCCACTCAACTCATTCCATCCTCCCACTGTTC
CACAGAAAATCTGCACCCTTTTCCCTTCCCCAAAGCCAAAAAGTGCAAAAGAGTTTAGGATCACCACTCATCATTATGTCCTTTCGGGCAG
CTGCTTAGCCTTTCTGTAACCTCGGACAAAACCAAAGACAGACACGCCAAAAAGGCTGACCGCTGGTGAGGGACACGAATTAATGACGA
TGAAGCTACCAGCAGGACAGTCAAGGTCGGTCACCATCCCGAGTCTCTGCTCTCTCCCCAG

EXON III

GCCGGATGCCACGTGGAGTTGCTCTTCCTCCGCTACATCTCTGACTGGGATCTGGACCCTGGGCGGTGCTACCGCGTCACCTGGTTCAC
GTCTTGGAGCCCCTGCTACGACTGTGCGCGGCACGTGGCCGACTTCCTGCGGGGGTACCCCAACCTGAGCCTGCGGATCTTCACGGCG
CGCCTCTACTTCTGCGACAAGGAGCGCAAGGCCGAGCCAGAGGGGCTGCGGCGGCTGCACCGCGCTGGAGTCCAGATCGCCATCATGA
CGTTCAAAG

INTRON III

GTGCGCGGGGAACTTTCGGCCCCGCTCTCCAGAGCAGGACTGGCCATTCAGATTGGGAAAAGCCAGGGTGAGAGAGAGCGGGCAGAAT
TCTAACTCGCCCCCAAAGTCCTAACTCGCCTTTTGGGGCTGAAAGCTGTCAAGGAGAATTGGAGGAAAAGAATCGGTGGACCCAAGACT
TTGGGGGAGGGGGTGGCAGGGTTTTCTTTCTTTCCTCATGATTTGAGCTCTGACCCATCCCCTCCCCTCTCCTTTTTAAATTTTCAAATA
G

EXON IV

ATTATTTTTATTGCTGGAATACTTTTGTGGAAAATCATGAAAGAACTTTCAAAGCCTGGGAGGGACTGCATGAAAATTCGGTTCGTCTGT
CTAGACAGCTTCGACGCATCCTTTTG

INTRON IV

GTAAGTGAGGGGGCTTTCTTTGCTTCTCATTTTCTTTCTCTCTCCAAAGTTATCTATCCCTGGCACTACTTATATTTCCTACATTTCCGCG
TGTGTGTTTGTGTGTGTTATATCATTCTCCCCTTATTTTTCCTGATGAAAACTTTCTCTTCTCTGTGTTTGGTTGTCATTTCTCCTACAAC
TGAGTTTGCTCTAATGTTGTTTCACAATTCAAAGCCCTCTACTAGCCCTCTGGGGTTTCCCAGGTGGCGCTAGTGGTAAAGAACACGCCA
GCCAATCCAGAAGACAAAAAAGATTCAGGTTCGTTCCCTGGGTTGGGAAGATCCCCTGGAGGAGGGCATGGCAACCCACTCCAGTATTC
TTGCTGGGAAATCCCACGGACAGAGGAGCCTGGTGGGCTACAGTCCATAGGGTCGCAAAGAGTCAGACACGACTTAGCCCAAGCACCAG
CCTTCCAGTAAGATTCCTTTCCCTTCCCCTCTTCCCCACCCTAATCCCCTCCCCTCTCTTTTCCTCACTCACATCTTTAACTTAGACTCTCC
TCAGAATCCATTTTACTGCTAGGAAAATTTCTTCCTCTCTTTCTCTTTCCTCTCCCTCTCCCCTTCTTTTCTTAAAAAATATTAATGAGAAC
GTCCCTGGTGGTCCAGTGGTTAAGACTCTGAGCTTCTGGGGCAGGGGACATGGATTCGACCCCTGGTCGGGGAGCTAAGATCCCAGAT
GTGGTGCAACCAAAAGATATAAATTTTAAAAAGAAAAAATTAACACAAAAATTACAAAGTCACCCACTGTCTTTCTCCACAG

EXON V (PARTIAL)

CCACTCTACGAGGTTGATGACTTACGAGACGCATTTCGT

REFERENCES

Aitken, R., Hosseini, A., MacDuff, R., 1999, Structure and diversification of the bovine immunoglobulin repertoire. Vet Immunol Immunopathol 72, 21-29.

Akerstrom, B., Nilson, B.H., Hoogenboom, H.R., Bjorck, L., 1994, On the interaction between single chain Fv antibodies and bacterial immunoglobulin-binding proteins. J Immunol Methods 177, 151-163.

Aliahmad, P., Pike, K.A., Ratcliffe, M.J.H., 2005, Cell surface immunoglobulin regulated checkpoints in chicken B cell development. Veterinary Immunology and Immunopathology 108, 3-9.

Altenburger, W., Neumaier, P.S., Steinmetz, M., Zachau, H.G., 1981, DNA-Sequence of the Constant Gene Region of the Mouse Immunoglobulin-Kappa Chain. Nucleic Acids Res. 9, 971-981.

Anderson, M.L., Spandidos, D.A., Coggins, J.R., 1991, Electroporation of lymphoid cells: factors affecting the efficiency of transfection. J Biochem Biophys Methods 22, 207-222.

Andersson, J., Coutinho, A., Melchers, F., 1979, Mitogen-activated B-cell blasts reactive to more than one mitogen. J. Exp. Med. 149, 553-564.

Andreason, G.L., Evans, G.A., 1989, Optimization of electroporation for transfection of mammalian cell lines. Analytical Biochemistry 180, 269-275.

Arakawa, H., Buerstedde, J.M., 2004, Immunoglobulin gene conversion: insights from bursal B cells and the DT40 cell line. Dev Dyn 229, 458-464.

Arakawa, H., Hauschild, J., Buerstedde, J.M., 2002, Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. Science 295, 1301-1306.

Aravind, L., Dixit, V.M., Koonin, E.V., 2001, Apoptotic molecular machinery: vastly increased complexity in vertebrates revealed by genome comparisons. Science 291, 1279-1284.

Baltimore, D., 1981, Gene Conversion - Some Implications for Immunoglobulin Genes. Cell 24, 592-594.

Banchereau, J., Steinman, R.M., 1998, Dendritic cells and the control of immunity. Nature 392, 245-252.

Band, M.R., Larson, J.H., Rebeiz, M., Green, C.A., Heyen, D.W., Donovan, J., Windish, R., Steining, C., Mahyuddin, P., Womack, J.E., Lewin, H.A., 2000, An ordered comparative map of the cattle and human genomes. Genome Res 10, 1359-1368.

Barnes, W.M., 1992, The fidelity of Taq polymerase catalyzing PCR is improved by an N-terminal deletion. Gene 112, 29-35.

Barreto, V., Reina-San-Martin, B., Ramiro, A.R., McBride, K.M., Nussenzweig, M.C., 2003, C-Terminal Deletion of AID Uncouples Class Switch Recombination from Somatic Hypermutation and Gene Conversion. Molecular Cell 12, 501-508.

Barreto, V.M., Pan-Hammarstrom, Q., Zhao, Y., Hammarstrom, L., Misulovin, Z., Nussenzweig, M.C., 2005, AID from bony fish catalyzes class switch recombination. J. Exp. Med. 202, 733-738.

Baum, C., Forster, P., Hegewisch-Becker, S., Harbers, K., 1994, An optimized electroporation protocol applicable to a wide range of cell lines. Biotechniques 17, 1058-1062.

Becker, R.S., Knight, K.L., 1990, Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits. Cell 63, 987-997.

Bendelac, A., Bonneville, M., Kearney, J.F., 2001, Autoreactivity by design: innate B and T lymphocytes. Nat Rev Immunol 1, 177-186.

Bengten, E., Wilson, M., Miller, N., Clem, L.W., Pilstrom, L., Warr, G.W., 2000, Immunoglobulin isotypes: structure, function, and genetics. Curr Top Microbiol Immunol 248, 189-219.

Berek, C., Milstein, C., 1987, Mutation Drift and Repertoire Shift in the Maturation of the Immune-Response. Immunol. Rev. 96, 23-41.

Berens, S.J., Wylie, D.E.a., Lopez, O.J., 1997, Use of a single VH family and long CDR3s in the variable region of cattle Ig heavy chains. Int Immunol 9, 189-199.

Betz, A.G., Neuberger, M.S., Milstein, C., 1993, Discriminating intrinsic and antigen-selected mutational hotspots in immunoglobulin V genes. Immunol Today 14, 405-411.

Bird, L., 2006, TLRs are crucial for B cells too. Nat Rev Immunol 6, 3-3.

Bjorck, L., 1988, Protein L. A novel bacterial cell wall protein with affinity for Ig L chains. J Immunol 140, 1194-1197.

Bogue, M., Roth, D.B., 1996, Mechanism of V(D)J recombination. Curr Opin Immunol 8, 175-180.

Both, G.W., Taylor, L., Pollard, J.W., Steele, E.J., 1990, Distribution of mutations around rearranged heavy-chain antibody variable-region genes. Mol Cell Biol 10, 5187-5196.

Bransteitter, R., Pham, P., Scharff, M.D., Goodman, M.F., 2003, Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. PNAS 100, 4102-4107.

Brenner, S., Milstein, C., 1966, Origin of antibody variation. Nature 211, 242-243.

Bross, L., Fukita, Y., McBlane, F., Demolliere, C., Rajewsky, K., Jacobs, H., 2000, DNA Double-Strand Breaks in Immunoglobulin Genes Undergoing Somatic Hypermutation. Immunity 13, 589.

Burioni, R., Plaisant, P., Bugli, F., Solforosi, L., Delli Carri, V., Varaldo, P.E., Fadda, G., 1998, A new subtraction technique for molecular cloning of rare antiviral antibody specificities from phage display libraries. Res Virol 149, 327-330.

Butler, J.E., 1997, Immunoglobulin gene organization and the mechanism of repertoire development. Scand J Immunol 45, 455-462.

Butler, J.E., Sun, J., Weber, P., Navarro, P., Francis, D., 2000, Antibody repertoire development in fetal and newborn piglets, III. Colonization of the gastrointestinal tract selectively diversifies the preimmune repertoire in mucosal lymphoid tissues. Immunology 100, 119-130.

Butler, J.E., Sun, J., Wertz, N., Sinkora, M., 2006, Antibody repertoire development in swine. Dev Comp Immunol 30, 199-221.

Buttgereit, P., Weineck, S., Ropke, G., Marten, A., Brand, K., Heinicke, T., Caselmann, W.H., Huhn, D., Schmidt-Wolf, I.G., 2000, Efficient gene transfer into lymphoma cells using adenoviral vectors combined with lipofection. Cancer Gene Ther 7, 1145-1155.

Cai, D., Mataraza, J.M., Qin, Z.H., Huang, Z., Huang, J., Chiles, T.C., Carnahan, D., Kempa, K., Ren, Z., 2005, Highly efficient molecular delivery into mammalian cells using carbon nanotube spearing. Nat Methods 2, 449-454.

Cai, X., Garen, A., 1995, Anti-melanoma antibodies from melanoma patients immunized with genetically modified autologous tumor cells: selection of specific antibodies from single-chain Fv fusion phage libraries. Proc Natl Acad Sci U S A 92, 6537-6541.

Carlson, L.M., McCormack, W.T., Postema, C.E., Humphries, E.H., Thompson, C.B., 1990, Templated insertions in the rearranged chicken IgL V gene segment arise by intrachromosomal gene conversion. Genes Dev 4, 536-547.

Casola, S., Otipoby, K.L., Alimzhanov, M., Humme, S., Uyttersprot, N., Kutok, J.L., Carroll, M.C., Rajewsky, K., 2004, B cell receptor signal strength determines B cell fate. Nat Immunol 5, 317-327.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Prasher, D.C., 1994, Green Fluorescent Protein As A Marker For Gene-Expression. Science 263, 802-805.

Chang, B., Casali, P., 1994, The CDR1 sequences of a major proportion of human germline Ig VH genes are inherently susceptible to amino acid replacement. Immunol Today 15, 367-373.

Charles Janeway, Paul Travers, Mark Walport, Shlomchik, M., 2004, The Adaptive Immune Response. Immunobiology : the immune system in health and disease Chapter 9.

Charlton, K.A., Moyle, S., Porter, A.J., Harris, W.J., 2000, Analysis of the diversity of a sheep antibody repertoire as revealed from a bacteriophage display library. J Immunol 164, 6221-6229.

Chaudhuri, J., Khuong, C., Alt, F.W., 2004, Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. Nature 430, 992-998.

Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E., Alt, F.W., 2003, Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature 422, 726.

Chen, X., Kinoshita, K., Honjo, T., 2001, Variable deletion and duplication at recombination junction ends: Implication for staggered double-strand cleavage in class-switch recombination. PNAS 98, 13860-13865.

Chevalet, C., Corpet, F., 1986, Statistical decision rules concerning synteny or independence between markers. Cytogenet Cell Genet 43, 132-139.

Cho, K.S., Elizondo, L.I., Boerkoel, C.F., 2004, Advances in chromatin remodeling and human disease. Current Opinion in Genetics & Development 14, 308-315.

Chu, G., Hayakawa, H., Berg, P., 1987, Electroporation for the efficient transfection of mammalian cells with DNA. Nucleic Acids Res 15, 1311-1326.

Clark, J., Whitelaw, B., 2003, A future for transgenic livestock. Nat Rev Genet 4, 825-833.

Cooper, M.D., Peterson, R.D., South, M.A., Good, R.A., 2006, The functions of the thymus system and the bursa system in the chicken. 1966. J Immunol 176, 6370-6404.

Crystal, R.G., McElvaney, N.G., Rosenfeld, M.A., Chu, C.S., Mastrangeli, A., Hay, J.G., Brody, S.L., Jaffe, H.A., Eissa, N.T., Danel, C., 1994, Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. Nat Genet 8, 42-51.

Currier, S.J., Gallarda, J.L., Knight, K.L., 1988, Partial molecular genetic map of the rabbit VH chromosomal region. J Immunol 140, 1651-1659.

Davern, S.M., Foote, L.J., Lankford, T.K., Macy, S.D., Wall, M.D., Kennel, S.J., 2005, Identification of an antilaminin-1 scFv that preferentially homes to vascular solid tumors. Cancer Biother Radiopharm 20, 524-533.

de Kruif, J., Terstappen, L., Boel, E., Logtenberg, T., 1995, Rapid selection of cell subpopulation-specific human monoclonal antibodies from a synthetic phage antibody library. Proc Natl Acad Sci U S A 92, 3938-3942.

de Wildt, R.M.T., Mundy, C.R., Gorick, B.D., Tomlinson, I.M., 2000, Antibody arrays for high-throughput screening of antibody-antigen interactions. Nat Biotech 18, 989.

Dedeoglu, F., Horwitz, B., Chaudhuri, J., Alt, F.W., Geha, R.S., 2004, Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NF{kappa}B. Int. Immunol. 16, 395-404.

Demartis, S., Huber, A., Viti, F., Lozzi, L., Giovannoni, L., Neri, P., Winter, G., Neri, D., 1999, A strategy for the isolation of catalytic activities from repertoires of enzymes displayed on phage. J Mol Biol 286, 617-633.

Denepoux, S., Razanajaona, D., Blanchard, D., Meffre, G., Capra, J.D., Banchereau, J., Lebecque, S., 1997, Induction of Somatic Mutation in a Human B Cell Line In Vitro. Immunity 6, 35.

Desmyter, A., Transue, T.R., Ghahroudi, M.A., Thi, M.H., Poortmans, F., Hamers, R., Muyldermans, S., Wyns, L., 1996, Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. Nat Struct Biol 3, 803-811.

Di Noia, J., Neuberger, M.S., 2002, Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. Nature 419, 43.

Diaz, M., Verkoczy, L.K., Flajnik, M.F., R. Klinman, N., 2001, Decreased Frequency of Somatic Hypermutation and Impaired Affinity Maturation but Intact Germinal Center Formation in Mice Expressing Antisense RNA to DNA Polymerase {{zeta}}. J Immunol 167, 327-335.

Ditzel, H.J., 2002, Rescue of a broader range of antibody specificities using an epitope-masking strategy. Methods Mol Biol 178, 179-186.

Doi, T., Kinoshita, K., Ikegawa, M., Muramatsu, M., Honjo, T., 2003, Inaugural Article: De novo protein synthesis is required for the activation-induced cytidine deaminase function in class-switch recombination. PNAS 100, 2634-2638.

Dufour, V., Malinge, S., Nau, F., 1996, The sheep Ig variable region repertoire consists of a single VH family. J Immunol 156, 2163-2170.

Dufour, V., Nau, F., 1997, Genomic organization of the sheep immunoglobulin JH segments and their contribution to heavy chain variable region diversity. Immunogenetics 46, 283-292.

Endemann, H., Bross, P., Rasched, I., 1992, The adsorption protein of phage IKe. Localization by deletion mutagenesis of domains involved in infectivity. Mol Microbiol 6, 471-478.

Everts-van der Wind, A., Kata, S.R., Band, M.R., Rebeiz, M., Larkin, D.M., Everts, R.E., Green, C.A., Liu, L., Natarajan, S., Goldammer, T., Lee, J.H., McKay, S., Womack, J.E., Lewin, H.A., 2004, A 1463 gene cattle-human comparative map with anchor points defined by human genome sequence coordinates. Genome Res 14, 1424-1437.

Faili, A., Aoufouchi, S., Gueranger, Q., Zober, C., Leon, A., Bertocci, B., Weill, J.-C., Reynaud, C.-A., 2002a, AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. Nat Immunol 3, 815.

Faili, A., Aoufouchi, S., Gueranger, Q., Zober, C., Leon, A., Bertocci, B., Weill, J.C., Reynaud, C.A., 2002b, AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. Nat Immunol 3, 815-821.

Ferradini, L., Reynaud, C.-A., Lauster, R., Weill, J.-C., 1994, Rearrangement of the chicken lambda light chain locus: a silencer/antisilencer regulation. Semin. Immunol. 6, 165-173.

Flajnik, M.F., 2002, Comparative analyses of immunoglobulin genes: surprises and portents. Nat Rev Immunol 2, 688-698.

Ford, J.E., McHeyzer-Williams, M.G., Lieber, M.R., 1994, Analysis of individual immunoglobulin lambda light chain genes amplified from single cells is inconsistent with variable region gene conversion in germinal-center B cell somatic mutation. Eur J Immunol 24, 1816-1822.

Foy, T.M., Aruffo, A., Bajorath, J., Buhlmann, J.E., Noelle, R.J., 1996, Immune regulationby CD40 and its ligand GP39. Annual Review of Immunology 14, 591-617.

Friedman, M.L., Tunyaplin, C., Zhai, S.K., Knight, K.L., 1994, Neonatal VH, D, and JH gene usage in rabbit B lineage cells. J Immunol 152, 632-641.

Friend, D.S., Papahadjopoulos, D., Debs, R.J., 1996, Endocytosis and intracellular processing accompanying transfection mediated by cationic liposomes. Biochimica et Biophysica Acta (BBA) - Biomembranes 1278, 41-50.

Fugmann, S.D., Schatz, D.G., 2002, Immunology: One AID to Unite Them All. Science 295, 1244-1245.

Golchin, M., 2004, Recombinant antibodies against K99 colonisation factor of *E. coli*. Ph.D Thesis, University of Glasgow.

Goldsby, R.A., Kindt, T.J., Osborne, B.A., Kuby, J., , 2003, Organisation and expression of immunoglobulin genes. Immunology, 110.

Goletz, S., Christensen, P.A., Kristensen, P., Blohm, D., Tomlinson, I., Winter, G., Karsten, U., 2002, Selection of large diversities of antiidiotypic antibody fragments by phage display. Journal of Molecular Biology 315, 1087.

Gontier, E., Ayrault, O., Godet, I., Nau, F., Ladeveze, V., 2005, Developmental progression of immunoglobulin heavy chain diversity in sheep. Veterinary Immunology and Immunopathology 103, 31-51.

Gonzalez-Fernandez, A., Gupta, S.K., Pannell, R., Neuberger, M.S., Milstein, C., 1994, Somatic mutation of immunoglobulin lambda chains: a segment of the major intron hypermutates as much as the complementarity-determining regions. Proc Natl Acad Sci U S A 91, 12614-12618.

Gordon, M.S., Kanegai, C.M., Doerr, J.R., Wall, R., 2003, Somatic hypermutation of the B cell receptor genes B29 (Igbeta, CD79b) and mb1 (Igalpha, CD79a). Proceedings Of The National Academy Of Sciences Of The United States Of America 100, 4126-4131.

Graham, F.L., van der Eb, A.J., 1973, Transformation of rat cells by DNA of human adenovirus 5. Virology 54, 536-539.

Graille, M., Stura, E.A., Corper, A.L., Sutton, B.J., Taussig, M.J., Charbonnier, J.-B., Silverman, G.J., 2000a, Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: Structural basis for recognition of B-cell receptors and superantigen activity. PNAS 97, 5399-5404.

Graille, M., Stura, E.A., Corper, A.L., Sutton, B.J., Taussig, M.J., Charbonnier, J.B., Silverman, G.J., 2000b, Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. Proc Natl Acad Sci U S A 97, 5399-5404.

Graus, Y.F., de Baets, M.H., Parren, P.W., Berrih-Aknin, S., Wokke, J., van Breda Vriesman, P.J., Burton, D.R., 1997, Human anti-nicotinic acetylcholine receptor recombinant Fab fragments isolated from thymus-derived phage display libraries from myasthenia gravis patients reflect predominant specificities in serum and block the action of pathogenic serum antibodies. J Immunol 158, 1919-1929.

Greisman, H.A., Pabo, C.O., 1997, A General Strategy for Selecting High-Affinity Zinc Finger Proteins for Diverse DNA Target Sites. Science 275, 657-661.

Griebel, P.J., Ferrari, G., 1994, Evidence for a stromal cell-dependent, self-renewing B cell population in lymphoid follicles of the ileal Peyer's patch of sheep. Eur J Immunol 24, 401-409.

Griebel, P.J., Hein, W.R., 1996, Expanding the role of Peyer's patches in B-cell ontogeny. Immunol Today 17, 30-39.

Griebel, P.J., Kennedy, L., Graham, T., Davis, W.C., Reynolds, J.D., 1992, Characterization of B-cell phenotypic changes during ileal and jejunal Peyer's patch development in sheep. Immunology 77, 564-570.

Griffiths, A.D., Duncan, A.R., 1998, Strategies for selection of antibodies by phage display. Current Opinion in Biotechnology 9, 102-108.

Griffiths, A.D., Williams, S.C., Hartley, O., Tomlinson, I.M., Waterhouse, P., Crosby, W.L., Kontermann, R.E., Jones, P.T., Low, N.M., Allison, T.J., et al., 1994, Isolation of high affinity human antibodies directly from large synthetic repertoires. Embo J 13, 3245-3260.

Gu, F., Chowdhary, B.P., Andersson, L., Harbitz, I., Gustavsson, I., 1992, Assignment of the bovine immunoglobulin gamma heavy chain (IGHG) gene to chromosome 21q24 by in situ hybridization. Hereditas 117, 237-240.

Guzman, L.M., Belin, D., Carson, M.J., Beckwith, J., 1995, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. 177, 4121-4130.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman, N., Hamers, R., 1993, Naturally occurring antibodies devoid of light chains. Nature 363, 446-448.

Harboe, M., Folling, I., 1974, Recognition of two distinct groups of human IgM and IgA based on different binding to staphylococci. Scand J Immunol 3, 471-782.

Harris, R.S., Sale, J.E., Petersen-Mahrt, S.K., Neuberger, M.S., 2002, AID is essential for immunoglobulin V gene conversion in a cultured B cell line. Curr Biol 12, 435-438.

Hashimoto, C., Hudson, K.L., Anderson, K.V., 1988, The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell 52, 269-279.

Hawkins, R.E., Russell, S.J., Winter, G., 1992, Selection of phage antibodies by binding affinity : Mimicking affinity maturation. Journal of Molecular Biology 226, 889-896.

Hayes, H.C., Petit, E.J., 1993, Mapping of the beta-lactoglobulin gene and of an immunoglobulin M heavy chain-like sequence to homoeologous cattle, sheep, and goat chromosomes. Mamm Genome 4, 207-210.

Hayhurst, A., Harris, W.J., 1999, Escherichia coliSkp Chaperone Coexpression Improves Solubility and Phage Display of Single-Chain Antibody Fragments. Protein Expression and Purification 15, 336-343.

He, B., Qiao, X., Cerutti, A., 2004, CpG DNA induces IgG class switch DNA recombination by activating human B cells through an innate pathway that requires TLR9 and cooperates with IL-10. J Immunol 173, 4479-4491.

Hemmi, H., Takeuchi, O., Kawai, T., Kaisho, T., Sato, S., Sanjo, H., Matsumoto, M., Hoshino, K., Wagner, H., Takeda, K., Akira, S., 2000, A Toll-like receptor recognizes bacterial DNA. Nature 408, 740-745.

Hieter, P.A., Maizel, J.V., Jr., Leder, P., 1982, Evolution of human immunoglobulin kappa J region genes. J Biol Chem 257, 1516-1522.

Hieter, P.A., Max, E.E., Seidman, J.G., Maizel, J.V., Jr., Leder, P., 1980, Cloned human and mouse kappa immunoglobulin constant and J region genes conserve homology in functional segments. Cell 22, 197-207.

Hinds-Frey, K.R., Nishikata, H., Litman, R.T., Litman, G.W., 1993, Somatic variation precedes extensive diversification of germline sequences and combinatorial joining in the evolution of immunoglobulin heavy chain diversity. J Exp Med 178, 815-824.

Hofmann, A., Kessler, B., Ewerling, S., Weppert, M., Vogg, B., Ludwig, H., Stojkovic, M., Boelhauve, M., Brem, G., Wolf, E., Pfeifer, A., 2003, Efficient transgenesis in farm animals by lentiviral vectors. EMBO Rep 4, 1054-1060.

Holt, L.J., Bussow, K., Walter, G., Tomlinson, I.M., 2000, By-passing selection: direct screening for antibody-antigen interactions using protein arrays. Nucl. Acids Res. 28, e72-.

Honjo, T., Muramatsu, M., Fagarasan, S., 2004, Aid: How does it aid antibody diversity? Immunity 20, 659.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., Winter, G., 1991, Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res 19, 4133-4137.

Hosseini, A., Campbell, G., Prorocic, M., Aitken, R., 2004, Duplicated copies of the bovine JH locus contribute to the Ig repertoire. Int. Immunol. 16, 843-852.

Ichikawa, H.T., Sowden, M.P., Torelli, A.T., Bachl, J., Huang, P., Dance, G.S., Marr, S.H., Robert, J., Wedekind, J.E., Smith, H.C., Bottaro, A., 2006, Structural phylogenetic analysis of activation-induced deaminase function. J Immunol 177, 355-361.

Imai, K., Slupphaug, G., Lee, W.-I., Revy, P., Nonoyama, S., Catalan, N., Yel, L., Forveille, M., Kavli, B., Krokan, H.E., Ochs, H.D., Fischer, A., Durandy, A., 2003, Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nat Immunol 4, 1023.

Inouye, S., Tsuji, F.I., 1994, Aequorea Green Fluorescent Protein - Expression Of The Gene And Fluorescence Characteristics Of The Recombinant Protein. Febs Letters 341, 277-280.

Jahn, S., Roggenbuck, D., Niemann, B., Ward, E.S., 1995, Expression of monovalent fragments derived from a human IgM autoantibody in E. coli. The input of the somatically mutated CDR1/CDR2 and of the CDR3 into antigen binding specificity. Immunobiology 193, 400-419.

Janeway, C.A., Jr., 1989, Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harb Symp Quant Biol 54 Pt 1, 1-13.

Janeway, C.A., Jr., 1992, The immune system evolved to discriminate infectious nonself from noninfectious self. Immunol Today 13, 11-16.

Janeway, C.A., Jr. , Paul Travers, Mark Walport, Shlomchik, M., 2004, The Adaptive Immune Response. Immunobiology : the immune system in health and disease Chapter 9.

Jasper, P.J., Zhai, S.K., Kalis, S.L., Kingzette, M., Knight, K.L., 2003, B lymphocyte development in rabbit: progenitor B cells and waning of B lymphopoiesis. J Immunol 171, 6372-6380.

Jenne, C.N., Kennedy, L.J., McCullagh, P., Reynolds, J.D., 2003, A new model of sheep Ig diversification: shifting the emphasis toward combinatorial mechanisms and away from hypermutation. J Immunol 170, 3739-3750.

Jenne, C.N., Kennedy, L.J., Reynolds, J.D., 2006, Antibody repertoire development in the sheep. Dev Comp Immunol 30, 165-174.

Jun Peng Yan, J.H.K.a.Y.P.Q., 2004, Generation and characterization of a novel single-chain antibody fragment specific against human fibrin clots from phage display antibody library.

Kappes, S.M., Keele, J.W., Stone, R.T., McGraw, R.A., Sonstegard, T.S., Smith, T.P., Lopez-Corrales, N.L., Beattie, C.W., 1997, A second-generation linkage map of the bovine genome. Genome Res 7, 235-249.

Kaushik, A., Shojaei, F., Saini, S.S., 2002, Novel insight into antibody diversification from cattle. Vet Immunol Immunopathol 87, 347-350.

Kearney, J.F., Lawton, A.R., 1975, B lymphocyte differentiation induced by lipopolysaccharide. I. Generation of cells synthesizing four major immunoglobulin classes. J Immunol 115, 671-676.

Keitaro, O., Sadatoshi, M., Taro, O., Kenichi, M., Koichi, O., Hajime, T., 2004, Molecular Cloning of Canine Activation-Induced Cytidine Deaminase (AID) cDNA and Its Expression in Normal Tissues. Journal of Veterinary Medical Science 66 739-741.

Keohavong, P., Thilly, W.G., 1989, Fidelity of DNA polymerases in DNA amplification. Proc Natl Acad Sci U S A 86, 9253-9257.

Kinoshita, K., Honjo, T., 2001, Linking class-switching recombination with somatic hypermutation. Nature Reviews Molecular Cell Biology 2, 493.

Knight, K.L., 1992, Restricted VH gene usage and generation of antibody diversity in rabbit. Annu Rev Immunol 10, 593-616.

Kobe, B., Deisenhofer, J., 1995, Proteins with leucine-rich repeats. Curr Opin Struct Biol 5, 409-416.

Krebber, C., Spada, S., Desplancq, D., Krebber, A., Ge, L., Pluckthun, A., 1997, Selectively-infective phage (SIP): a mechanistic dissection of a novel in vivo selection for protein-ligand interactions. J Mol Biol 268, 607-618.

Krieg, A.M., 2002, CpG motifs in bacterial DNA and their immune effects. Annu Rev Immunol 20, 709-760.

Krieg, A.M., Yi, A.-K., Matson, S., Waldschmidt, T.J., Bishop, G.A., Teasdale, R., Koretzky, G.A., Klinman, D.M., 1995, CpG motifs in bacterial DNA trigger direct B-cell activation. Nature 374, 546-549.

Kruse, P.E., 1983, The importance of colostral immunoglobulins and their absorption from the intestine of the newborn animals. Annales De Recherches Veterinaires. Annals Of Veterinary Research 14, 349-353.

Laemmli, U.K., 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Landau, N.R., Schatz, D.G., Rosa, M., Baltimore, D., 1987, Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. Mol Cell Biol 7, 3237-3243.

Lanning, D., Sethupathi, P., Rhee, K.J., Zhai, S.K., Knight, K.L., 2000, Intestinal microflora and diversification of the rabbit antibody repertoire. J Immunol 165, 2012-2019.

Lanning, D.K., Knight, K.L., 1997, Somatic hypermutation: mutations 3' of rabbit VDJ H-chain genes. J Immunol 159, 4403-4407.

Larijani, M., Frieder, D., Basit, W., Martin, A., 2005, The mutation spectrum of purified AID is similar to the mutability index in Ramos cells and in ung(-/-)msh2(-/-) mice. Immunogenetics 56, 840-845.

Lauster, R., Reynaud, C.A., Martensson, I.L., Peter, A., Bucchini, D., Jami, J., Weill, J.C., 1993, Promoter, enhancer and silencer elements regulate rearrangement of an immunoglobulin transgene. Embo J 12, 4615-4623.

Lebecque, S.G., Gearhart, P.J., 1990, Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene. J Exp Med 172, 1717-1727.

Lederberg, J., 1959, Genes and antibodies. Science 129, 1649-1653.

Lee, N., Francklyn, C., Hamilton, E.P., 1987, Arabinose-Induced Binding of AraC Protein to araI2 Activates the araBAD Operon Promoter. PNAS 84, 8814-8818.

Lee, S.C., Han, J.S., Seo, J.K., Cha, Y.N., 2003, Modulation of cyclooxygenase-2 expression by phosphatidylcholine specific phospholipase C and D in macrophages stimulated with lipopolysaccharide. Mol Cells 15, 320-326.

Lefranc, M.P., 2001a, Nomenclature of the human immunoglobulin heavy (IGH) genes. Exp Clin Immunogenet 18, 100-116.

Lefranc, M.P., 2001b, Nomenclature of the human immunoglobulin kappa (IGK) genes. Exp Clin Immunogenet 18, 161-174.

Lefranc, M.P., 2003, IGH (Immunoglobulin Heavy). Atlas Genet Cytogenet Oncol Haematol.

Li, Z., Woo, C.J., Scharff, M.D., 2003, Mutations in AID and UNG extend the function of AID. Nat Immunol 4, 945-946.

Lieber, M.R., 1991, Site-specific recombination in the immune system. FASEB J. 5, 2934-2944.

Lien, S., Karlsen, A., Klemetsdal, G., Vage, D.I., Olsaker, I., Klungland, H., Aasland, M., Heringstad, B., Ruane, J., Gomez-Raya, L., 2000, A primary screen of the bovine genome for quantitative trait loci affecting twinning rate. Mamm Genome 11, 877-882.

Litinskiy, M.B., Nardelli, B., Hilbert, D.M., He, B., Schaffer, A., Casali, P., Cerutti, A., 2002, DCs induce CD40-independent immunoglobulin class switching through BLyS and APRIL. Nat Immunol 3, 822-829.

Liu, Y.J., Arpin, C., de Bouteiller, O., Guret, C., Banchereau, J., Martinez-Valdez, H., Lebecque, S., 1996, Sequential triggering of apoptosis, somatic mutation and isotype switch during germinal center development. Semin Immunol 8, 169-177.

Liu, Z.-X., Yi, G.-H., Qi, Y.-P., Liu, Y.-L., Yan, J.-P., Qian, J., Du, E.-Q., Ling, W.-F., 2005, Identification of single-chain antibody fragments specific against SARS-associated coronavirus from phage-displayed antibody library. Biochemical and Biophysical Research Communications 329, 437-444.

Liu, Z., Hansen, M., Womack, J.E., Antoniou, E., 2003, A comparative map of interstitial bovine chromosome 5 with human chromosomes 12 and 22. Cytogenet Genome Res 101, 147-154.

Longerich, S., Basu, U., Alt, F., Storb, U., 2006, AID in somatic hypermutation and class switch recombination. Current Opinion in Immunology 18, 164-174.

Longerich, S., Storb, U., 2005, The contested role of uracil DNA glycosylase in immunoglobulin gene diversification. Trends Genet 21, 253-256.

Loomans, E.E., Petersen-van Ettekoven, A., Bloemers, H.P., Schielen, W.J., 1997, Direct coating of poly(lys) or acetyl-thio-acetyl peptides to polystyrene: the effects in an enzyme-linked immunosorbent assay. Anal Biochem 248, 117-129.

Lopez, O., Perez, C., Wylie, D., 1998, A single VH family and long CDR3s are the targets for hypermutation in bovine immunoglobulin heavy chains. Immunol Rev 162, 55-66.

Lowman, H.B., Wells, J.A., 1993, Affinity Maturation of Human Growth Hormone by Monovalent Phage Display. Journal of Molecular Biology 234, 564-578.

Lu, J., Sloan, S.R., 1999, An alternating selection strategy for cloning phage display antibodies. Journal of Immunological Methods 228, 109-119.

Lucier, M.R., Thompson, R.E., Waire, J., Lin, A.W., Osborne, B.A., Goldsby, R.A., 1998, Multiple sites of V lambda diversification in cattle. J Immunol 161, 5438-5444.

Mackay, F., Schneider, P., Rennert, P., Browning, J., 2003, BAFF AND APRIL: a tutorial on B cell survival. Annu Rev Immunol 21, 231-264.

Mage, R.G., Lanning, D., Knight, K.L., 2006, B cell and antibody repertoire development in rabbits: The requirement of gut-associated lymphoid tissues. Developmental & Comparative Immunology 30, 137-153.

Mannino, R.J., Gould-Fogerite, S., 1988, Liposome mediated gene transfer. Biotechniques 6, 682-690.

Marcus, W.D., Wang, H., Lohr, D., Sierks, M.R., Lindsay, S.M., 2006, Isolation of an scFv targeting BRG1 using phage display with characterization by AFM. Biochemical and Biophysical Research Communications 342, 1123-1129.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D., Winter, G., 1991, By-passing immunization : Human antibodies from V-gene libraries displayed on phage. Journal of Molecular Biology 222, 581-597.

Marks, J.D., Ouwehand, W.H., Bye, J.M., Finnern, R., Gorick, B.D., Voak, D., Thorpe, S.J., Hughes-Jones, N.C., Winter, G., 1993, Human antibody fragments specific for

human blood group antigens from a phage display library. Biotechnology (N Y) 11, 1145-1149.

Martin, A., Scharff, M.D., 2002, Somatic hypermutation of the AID transgene in B and non-B cells. Proceedings Of The National Academy Of Sciences Of The United States Of America 99, 12304.

Matsuda, F., Shin, E.K., Hirabayashi, Y., Nagaoka, H., Yoshida, M.C., Zong, S.Q., Honjo, T., 1990, Organization of variable region segments of the human immunoglobulin heavy chain: duplication of the D5 cluster within the locus and interchromosomal translocation of variable region segments. Embo J 9, 2501-2506.

Mattaj, I.W., Englmeier, L., 1998, Nucleocytoplasmic transport: The Soluble Phase. Annu. Rev. Biochem. 67, 265-306.

Mayorov, V.I., Rogozin, I.B., Adkison, L.R., Frahm, C., Kunkel, T.A., Pavlov, Y.I., 2005, Expression of human AID in yeast induces mutations in context similar to the context of somatic hypermutation at G-C pairs in immunoglobulin genes. BMC Immunol 6, 10.

McBlane, J.F., van Gent, D.C., Ramsden, D.A., Romeo, C., Cuomo, C.A., Gellert, M., Oettinger, M.A., 1995, Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell 83, 387-395.

McBride, K.M., Barreto, V., Ramiro, A.R., Stavropoulos, P., Nussenzweig, M.C., 2004, Somatic Hypermutation Is Limited by CRM1-dependent Nuclear Export of Activation-induced Deaminase. J. Exp. Med. 199, 1235-1244.

McCormack, W.T., Thompson, C.B., 1990, Chicken IgL variable region gene conversions display pseudogene donor preference and 5' to 3' polarity. Genes Dev 4, 548-558.

Medzhitov, R., 2001, Toll-like receptors and innate immunity. Nat Rev Immunol 1, 135-145.

Meek, K.D., Hasemann, C.A., Capra, J.D., 1989, Novel rearrangements at the immunoglobulin D locus. Inversions and fusions add to IgH somatic diversity. J Exp Med 170, 39-57.

Mehta, A., Kinter, M.T., Sherman, N.E., Driscoll, D.M., 2000, Molecular Cloning of Apobec-1 Complementation Factor, a Novel RNA-Binding Protein Involved in the Editing of Apolipoprotein B mRNA. Mol. Cell. Biol. 20, 1846-1854.

Meier, J.T., Lewis, S.M., 1993, P nucleotides in V(D)J recombination: a fine-structure analysis. Mol Cell Biol 13, 1078-1092.

Meunier-Durmort, C., Ferry, N., Hainque, B., Delattre, J., Forest, C., 1996, Efficient transfer of regulated genes in adipocytes and hepatoma cells by the combination of liposomes and replication-deficient adenovirus. European Journal Of Biochemistry / FEBS 237, 660-667.

Meunier-Durmort, C., Picart, R., Ragot, T., Perricaudet, M., Hainque, B., Forest, C., 1997, Mechanism of adenovirus improvement of cationic liposome-mediated gene transfer. Biochim Biophys Acta 1330, 8-16.

Meyer, A., Pamg, C.L., Hansal, S.A., Osborne, B.A., Goldsby, R.A., 1997, Immunoglobulin gene diversification in cattle. Int Rev Immunol 15, 165-183.

Michael, N., Shen, H.M., Longerich, S., Kim, N., Longacre, A., Storb, U., 2003, The E Box Motif CAGGTG Enhances Somatic Hypermutation without Enhancing Transcription. Immunity 19, 235-242.

Miller, C., Kelsoe, G., 1995, Ig VH hypermutation is absent in the germinal centers of aged mice. J Immunol 155, 3377-3384.

Miller, J.H., 1972, Experiments in Molecular Genetics. Cold Spring Harbor Press, Cold Spring Harbor.

Milstein, C., Neuberger, M.S., Staden, R., 1998, Both DNA strands of antibody genes are hypermutation targets. Proc Natl Acad Sci U S A 95, 8791-8794.

Miyada, C.G., Stoltzfus, L., Wilcox, G., 1984, Regulation of the araC Gene of Escherichia coli: Catabolite Repression, Autoregulation, and Effect on araBAD Expression. PNAS 81, 4120-4124.

Miyoshi, H., Takahashi, M., Gage, F.H., Verma, I.M., 1997, Stable and efficient gene transfer into the retina using an HIV-based lentiviral vector. PNAS 94, 10319-10323.

Mond, J.J., Lees, A., Snapper, C.M., 1995, T Cell-Independent Antigens Type 2. Annual Review of Immunology 13, 655-692.

Moreau, M.F., Thibaud, J.L., Miled, L.B., Chaussepied, M., Baumgartner, M., Davis, W.C., Minoprio, P., Langsley, G., 1999, Theileria annulata in CD5(+) macrophages and B1 B cells. Infect Immun 67, 6678-6682.

Motoyama, N., Okada, H., Azuma, T., 1991, Somatic mutation in constant regions of mouse lambda 1 light chains. Proc Natl Acad Sci U S A 88, 7933-7937.

Mullen, L.M., Nair, S.P., Ward, J.M., Rycroft, A.N., Henderson, B., 2006, Phage display in the study of infectious diseases. Trends Microbiol 14, 141-147.

Muller, K.M., Arndt, K.M., Strittmatter, W., Pluckthun, A., 1998, The first constant domain (CH1 and CL) of an antibody used as heterodimerization domain for bispecific miniantibodies. FEBS Letters 422, 259-264.

Murakami, K., Inumaru, S., Yokoyama, T., Okada, K., Sentsui, H., 1999, Expression of granulocyte-macrophage colony-stimulating factor (GM-CSF) receptor on B-1a cell from persistent lymphocytosis (PL) cows and lymphoma cell induced by bovine leukemia virus. Vet Immunol Immunopathol 68, 49-59.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., Honjo, T., 2000, Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell 102, 553-563.

Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., Honjo, T., 1999, Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J. Biol. Chem. 274, 18470-18476.

Muto, T., Muramatsu, M., Taniwaki, M., Kinoshita, K., Honjo, T., 2000, Isolation, Tissue Distribution, and Chromosomal Localization of the Human Activation-Induced Cytidine Deaminase (AID) Gene. Genomics 68, 85.

Nadel, B., Feeney, A.J., 1995, Influence of coding-end sequence on coding-end processing in V(D)J recombination. J Immunol 155, 4322-4329.

Nagaoka, H., Muramatsu, M., Yamamura, N., Kinoshita, K., Honjo, T., 2002, Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin S mu region: Implication of AID involvement in a common step of class switch recombination and somatic hypermutation. J. Exp. Med. 195, 529-534.

Nakamura, M., Kondo, S., Sugai, M., Nazarea, M., Imamura, S., Honjo, T., 1996, High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. Int Immunol 8, 193-201.

Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., Trono, D., 1996, In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. Science 272, 263-267.

Nambu, Y., Sugai, M., Gonda, H., Lee, C.-G., Katakai, T., Agata, Y., Yokota, Y., Shimizu, A., 2003, Transcription-Coupled Events Associating with Immunoglobulin Switch Region Chromatin. Science 302, 2137-2140.

Nardelli, B., Belvedere, O., Roschke, V., Moore, P.A., Olsen, H.S., Migone, T.S., Sosnovtseva, S., Carrell, J.A., Feng, P., Giri, J.G., Hilbert, D.M., 2001, Synthesis and release of B-lymphocyte stimulator from myeloid cells. Blood 97, 198-204.

Nei, M., Gojobori, T., 1986, Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3, 418-426.

Neill, J.D., Ridpath, J.F., Liebler-Tenorio, E., 2006, Global gene expression profiling of Bovine immature B cells using serial analysis of gene expression. Anim Biotechnol 17, 21-31.

Nielsen, L.L., Maneval, D.C., 1998, P53 tumor suppressor gene therapy for cancer. Cancer Gene Ther 5, 52-63.

Nissim, A., Hoogenboom, H.R., Tomlinson, I.M., Flynn, G., Midgley, C., Lane, D., Winter, G., 1994, Antibody fragments from a 'single pot' phage display library as immunochemical reagents. Embo J 13, 692-698.

O'Brien, P.M., Aitken, R., O'Neil, B.W., Campo, M.S., 1999, Generation of native bovine mAbs by phage display. PNAS 96, 640-645.

Oettinger, M.A., Schatz, D.G., Gorka, C., Baltimore, D., 1990, RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 248, 1517-1523.

Okazaki, I.-m., Kinoshita, K., Muramatsu, M., Yoshikawa, K., Honjo, T., 2002, The AID enzyme induces class switch recombination in fibroblasts. Nature 416, 340.

Ota, T., Nei, M., 1994, Variance and covariances of the numbers of synonymous and nonsynonymous substitutions per site. Mol Biol Evol 11, 613-619.

Ozawa, A., Band, M.R., Larson, J.H., Donovan, J., Green, C.A., Womack, J.E., Lewin, H.A., 2000, Inaugural Article: Comparative organization of cattle chromosome 5 revealed by comparative mapping by annotation and sequence similarity and radiation hybrid mapping. PNAS 97, 4150-4155.

Pallares, N., Frippiat, J.P., Giudicelli, V., Lefranc, M.P., 1998, The human immunoglobulin lambda variable (IGLV) genes and joining (IGLJ) segments. Exp Clin Immunogenet 15, 8-18.

Papavasiliou, F.N., Schatz, D.G., 2000, Cell-cycle-regulated DNA double-strand breaks in somatic hypermutation of immunoglobulin genes. Nature 408, 216.

Papavasiliou, F.N., Schatz, D.G., 2002, The Activation-induced Deaminase Functions in a Postcleavage Step of the Somatic Hypermutation Process. J. Exp. Med. 195, 1193-1198.

Park, S.-R., Hyun-A Kim, Sung-Ki Chun, and, J.-B.P., Kim, P.-H., 2005, Mechanisms Underlying the Effects of LPS and Activation-induced Cytidine Deaminase on IgA Isotype Expression. Mol. Cells 19, 445-451.

Parng, C.L., Hansal, S., Goldsby, R.A., Osborne, B.A., 1996, Gene conversion contributes to Ig light chain diversity in cattle. J Immunol 157, 5478-5486.

Pasare, C., Medzhitov, R., 2005, Control of B-cell responses by Toll-like receptors. Nature 438, 364-368.

Pearse, A., 1980, Histochemistry Vol 1.

Pemberton, L.F., Blobel, G., Rosenblum, J.S., 1998, Transport routes through the nuclear pore complex. Current Opinion in Cell Biology 10, 392-399.

Perey, D.Y., Good, R.A., 1968, Experimental arrest and induction of lymphoid development in intestinal lymphoepithelial tissues of rabbits. Lab Invest 18, 15-26.

Petersen-Mahrt, S.K., Harris, R.S., Neuberger, M.S., 2002, AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature 418, 99.

Petersen, S., Casellas, R., Reina-San-Martin, B., Chen, H.T., Difilippantonio, M.J., Wilson, P.C., Hanitsch, L., Celeste, A., Muramatsu, M., Pilch, D.R., Redon, C., Ried, T., Bonner, W.M., Honjo, T., Nussenzweig, M.C., Nussenzweig, A., 2001, AID is required to initiate Nbs1/[gamma]-H2AX focus formation and mutations at sites of class switching. Nature 414, 660.

Pham, P., Bransteitter, R., Petruska, J., Goodman, M.F., 2003, Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. Nature 424, 103.

Pike, K.A., Baig, E., Ratcliffe, M.J., 2004, The avian B-cell receptor complex: distinct roles of Igalpha and Igbeta in B-cell development. Immunol Rev 197, 10-25.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., Cormier, M.J., 1992, Primary Structure Of The Aequorea-Victoria Green-Fluorescent Protein. Gene 111, 229-233.

Prausnitz, M.R., Bose, V.G., Langer, R., Weaver, J.C., 1993, Electroporation of mammalian skin: a mechanism to enhance transdermal drug delivery. Proc Natl Acad Sci U S A 90, 10504-10508.

Prince, H.M., Dessureault, S., Gallinger, S., Krajden, M., Sutherland, D.R., Addison, C., Zhang, Y., Graham, F.L., Stewart, A.K., 1998, Efficient adenovirus-mediated gene expression in malignant human plasma cells: relative lymphoid cell resistance. Exp Hematol 26, 27-36.

Qiu, C., De Young, M.B., Finn, A., Dichek, D.A., 1998, Cationic liposomes enhance adenovirus entry via a pathway independent of the fiber receptor and alpha(v)-integrins. Hum Gene Ther 9, 507-520.

Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T., Neuberger, M.S., 2002, Immunoglobulin Isotype Switching Is Inhibited and Somatic Hypermutation Perturbed in UNG-Deficient Mice. Curr. Biol. 12, 1748.

Ramirez, L.H., Orlowski, S., An, D., Bindoula, G., Dzodic, R., Ardouin, P., Bognel, C., Belehradek, J., Jr., Munck, J.N., Mir, L.M., 1998, Electrochemotherapy on liver tumours in rabbits. Br J Cancer 77, 2104-2111.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A.G., Brousse, N., Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A., Durandy, A., 2000, Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell 102, 565-575.

Reynaud, C.-A., Mackay, C.R., Muller, R.G., Weill, J.-C., 1991a, Somatic generation of diversity in a mammalian primary lymphoid organ: The sheep ileal Peyer's patches. Cell 64, 995-1005.

Reynaud, C.A., Anquez, V., Grimal, H., Weill, J.C., 1987a, A Hyperconversion Mechanism Generates the Chicken Light Chain Preimmune Repertoire. Cell 48, 379-388.

Reynaud, C.A., Anquez, V., Weill, J.C., 1991b, The chicken D locus and its contribution to the immunoglobulin heavy chain repertoire. Eur J Immunol 21, 2661-2670.

Reynaud, C.A., Dahan, A., Anquez, V., Weill, J.C., 1989, Somatic Hyperconversion Diversifies the Single Vh-Gene of the Chicken with a High-Incidence in the D-Region. Cell 59, 171-183.

Reynaud, C.A., Dahan, A., Weill, J.C., 1987b, A Gene Conversion Program.During the Ontogeny of Chicken B- Cells. Trends Genet. 3, 248-251.

Reynaud, C.A., Dufour, V., Weill, J.C., 1997, Generation of diversity in mammalian gut-associated lymphoid tissues: restricted V gene usage does not preclude complex V gene organization. J Immunol 159, 3093-3095.

Reynaud, C.A., Garcia, C., Hein, W.R., Weill, J.C., 1995, Hypermutation Generating the Sheep Immunoglobulin Repertoire Is an Antigen-Independent Process. Cell 80, 115-125.

Reynaud, C.A., Imhof, B.A., Anquez, V., Weill, J.C., 1992, Emergence of committed B lymphoid progenitors in the developing chicken embryo. Embo J 11, 4349-4358.

Reynaud, C.A., Mackay, C.R., Muller, R.G., Weill, J.C., 1991c, Somatic Generation of Diversity in a Mammalian Primary Lymphoid Organ - the Sheep Ileal Peyers-Patches. Cell 64, 995-1005.

Reynolds, J.D., Kennedy, L., Peppard, J., Pabst, R., 1991, Ileal Peyer's patch emigrants are predominantly B cells and travel to all lymphoid tissues in sheep. Eur J Immunol 21, 283-289.

Reynolds, J.D., Morris, B., 1983, The evolution and involution of Peyer's patches in fetal and postnatal sheep. Eur J Immunol 13, 627-635.

Rhee, K.-J., Jasper, P.J., Sethupathi, P., Shanmugam, M., Lanning, D., Knight, K.L., 2005a, Positive selection of the peripheral B cell repertoire in gut-associated lymphoid tissues. J. Exp. Med. 201, 55-62.

Rhee, K.J., Jasper, P.J., Sethupathi, P., Shanmugam, M., Lanning, D., Knight, K.L., 2005b, Positive selection of the peripheral B cell repertoire in gut-associated lymphoid tissues. J Exp Med 201, 55-62.

Rhee, K.J., Sethupathi, P., Driks, A., Lanning, D.K., Knight, K.L., 2004, Role of commensal bacteria in development of gut-associated lymphoid tissues and preimmune antibody repertoire. J Immunol 172, 1118-1124.

Riechmann, L., Holliger, P., 1997, The C-terminal domain of TolA is the coreceptor for filamentous phage infection of E. coli. Cell 90, 351-360.

Riechmann, L., Winter, G., 2000, Novel folded protein domains generated by combinatorial shuffling of polypeptide segments. Proc Natl Acad Sci U S A 97, 10068-10073.

Roberts, B.L., Markland, W., Ley, A.C., Kent, R.B., White, D.W., Guterman, S.K., Ladner, R.C., 1992, Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. Proceedings Of The National Academy Of Sciences Of The United States Of America 89, 2429-2433.

Rogozin, I.B., Diaz, M., 2004, Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. J Immunol 172, 3382-3384.

Rogozin, I.B., Kolchanov, N.A., 1992, Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. Biochim Biophys Acta 1171, 11-18.

Rolink, A., Melchers, F., Andersson, J., 1996, The SCID but Not the RAG-2 Gene Product Is Required for S[mu]-S[var epsilon] Heavy Chain Class Switching. Immunity 5, 319-330.

Romano, M.J., Stewart, J.A., Lewin, H.A., 1989, Phenotypic characterization of bovine lymphoblastoid cell lines. Veterinary Immunology and Immunopathology 23, 293.

Roper, R.L., Conrad, D.H., Brown, D.M., Warner, G.L., Phipps, R.P., 1990, Prostaglandin E2 promotes IL-4-induced IgE and IgG1 synthesis. J Immunol 145, 2644-2651.

Rubinstein, J.L., Holt, L.J., Walker, J.E., Tomlinson, I.M., 2003, Use of phage display and high-density screening for the isolation of an antibody against the 51-kDa subunit of complex I. Anal Biochem 314, 294-300.

Rudikoff, S., Pawlita, M., Pumphrey, J., Heller, M., 1984, Somatic Diversification of Immunoglobulins. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 81, 2162-2166.

Saini, S.S., Allore, B., Jacobs, R.M., Kaushik, A., 1999, Exceptionally long CDR3H region with multiple cysteine residues in functional bovine IgM antibodies. Eur J Immunol 29, 2420-2426.

Saini, S.S., Farrugia, W., Ramsland, P.A., Kaushik, A.K., 2003, Bovine IgM antibodies with exceptionally long complementarity-determining region 3 of the heavy chain share unique structural properties conferring restricted VH + Vlambda pairings. Int Immunol 15, 845-853.

Saini, S.S., Hein, W.R., Kaushik, A., 1997, A single predominantly expressed polymorphic immunoglobulin VH gene family, related to mammalian group, I, clan, II, is identified in cattle. Mol Immunol 34, 641-651.

Saini, S.S., Kaushik, A., 2002, Extensive CDR3H length heterogeneity exists in bovine foetal VDJ rearrangements. Scand J Immunol 55, 140-148.

Sale, J.E., Calandrini, D.M., Takata, M., Takeda, S., Neuberger, M.S., 2001, Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic hypermutation. Nature 412, 921.

Sale, J.E., Neuberger, M.S., 1998, TdT-Accessible Breaks Are Scattered over the Immunoglobulin V Domain in a Constitutively Hypermutating B Cell Line. Immunity 9, 859.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989, Molecular Cloning: A Laboratory Manual. 2, 5.72.

Sanz, I., 1991, Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions. J Immunol 147, 1720-1729.

Saribasak, H., Saribasak, N.N., Ipek, F.M., Ellwart, J.W., Arakawa, H., Buerstedde, J.M., 2006, Uracil DNA glycosylase disruption blocks Ig gene conversion and induces transition mutations. J Immunol 176, 365-371.

Sasso, E.H., Silverman, G.J., Mannik, M., 1989, Human IgM molecules that bind staphylococcal protein A contain VHIII H chains. J Immunol 142, 2778-2783.

Sasso, E.H., Silverman, G.J., Mannik, M., 1991, Human IgA and IgG F(ab')2 that bind to staphylococcal protein A belong to the VHIII subgroup. J Immunol 147, 1877-1883.

Sayegh, C.E., Drury, G., Ratcliffe, M.J., 1999, Efficient antibody diversification by gene conversion in vivo in the absence of selection for V(D)J-encoded determinants. Embo J 18, 6319-6328.

Scaviner, D., Barbie, V., Ruiz, M., Lefranc, M.P., 1999, Protein displays of the human immunoglobulin heavy, kappa and lambda variable and joining regions. Exp Clin Immunogenet 16, 234-240.

Schier, R., Bye, J., Apell, G., McCall, A., Adams, G.P., Malmqvist, M., Weiner, L.M., Marks, J.D., 1996, Isolation of high-affinity monomeric human anti-c-erbB-2 single chain Fv using affinity-driven selection. J Mol Biol 255, 28-43.

Schuler, W., Ruetsch, N.R., Amsler, M., Bosma, M.J., 1991, Coding joint formation of endogenous T cell receptor genes in lymphoid cells from scid mice: unusual P-nucleotide additions in VJ-coding joints. Eur J Immunol 21, 589-596.

Schwager, J., Burckert, N., Courtet, M., Du Pasquier, L., 1989, Genetic basis of the antibody repertoire in Xenopus: analysis of the Vh diversity. Embo J 8, 2989-3001.

Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003, SWISS-MODEL: an automated protein homology-modeling server. Nucl. Acids Res. 31, 3381-3385.

Sehgal, D., Obiakor, H., Mage, R.G., 2002, Distinct clonal Ig diversification patterns in young appendix compared to antigen-specific splenic clones. J Immunol 168, 5424-5433.

Sehgal, D., Schiaffella, E., Anderson, A.O., Mage, R.G., 1998, Analyses of single B cells by polymerase chain reaction reveal rearranged VH with germline sequences in spleens of immunized adult rabbits: implications for B cell repertoire maintenance and renewal. J Immunol 161, 5347-5356.

Shapiro, G.S., Aviszus, K., Ikle, D., Wysocki, L.J., 1999, Predicting Regional Mutability in Antibody V Genes Based Solely on Di- and Trinucleotide Sequence Composition. J Immunol 163, 259-268.

Shapiro, G.S., Aviszus, K., Murphy, J., Wysocki, L.J., 2002, Evolution of Ig DNA Sequence to Target Specific Base Positions Within Codons for Somatic Hypermutation. J Immunol 168, 2302-2306.

Shen, H.M., Peters, A., Baron, B., Zhu, X., Storb, U., 1998, Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. Science 280, 1750-1752.

Shen, S.X., Bernstein, R.M., Schluter, S.F., Marchalonis, J.J., 1996, Heavy-chain variable regions in carcharhine sharks: development of a comprehensive model for the evolution of VH domains among the gnathanstomes. Immunol Cell Biol 74, 357-364.

Shojaei, F., Saini, S.S., Kaushik, A.K., 2003, Unusually long germline DH genes contribute to large sized CDR3H in bovine antibodies. Mol Immunol 40, 61-67.

Short, J.A., Sethupathi, P., Zhai, S.K., Knight, K.L., 1991, VDJ genes in VHa2 allotype-suppressed rabbits. Limited germline VH gene usage and accumulation of somatic mutations in D regions. J Immunol 147, 4014-4018.

Silverman, G.J., Goodyear, C.S., 2002, A Model B-Cell Superantigen and the Immunobiology of B Lymphocytes. Clinical Immunology 102, 117-134.

Silverstone, A.E., Cantor, H., Goldstein, G., Baltimore, D., 1976, Terminal deoxynucleotidyl transferase is found in prothymocytes. J Exp Med 144, 543-548.

Sinclair, M.C., Aitken, R., 1995, PCR strategies for isolation of the 5' end of an immunoglobulin-encoding bovine cDNA. Gene 167, 285-289.

Sinclair, M.C., Gilchrist, J., Aitken, R., 1995, Molecular characterization of bovine V lambda regions. J Immunol 155, 3068-3078.

Sinclair, M.C., Gilchrist, J., Aitken, R., 1997, Bovine IgG repertoire is dominated by a single diversified VH gene family. J Immunol 159, 3883-3889.

Sirena, D., Lilienfeld, B., Eisenhut, M., Kalin, S., Boucke, K., Beerli, R.R., Vogt, L., Ruedl, C., Bachmann, M.F., Greber, U.F., Hemmi, S., 2004, The human membrane cofactor CD46 is a receptor for species B adenovirus serotype 3. J Virol 78, 4454-4462.

Slonim, D., Kruglyak, L., Stein, L., Lander, E., 1997, Building human genome maps with radiation hybrids. J Comput Biol 4, 487-504.

Smith, G.P., 1985, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228, 1315.

Sohail, A., Klapacz, J., Samaranayake, M., Ullah, A., Bhagwat, A.S., 2003, Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic Acids Res 31, 2990-2994.

Solinas-Toldo, S., Lengauer, C., Fries, R., 1995, Comparative Genome Map of Human and Cattle. Genomics 27, 489-496.

Sorsa-Leslie, T., Mason, H.D., Harris, W.J., Fowler, P.A., 2005, Selection of gonadotrophin surge attenuating factor phage antibodies by bioassay. Reprod Biol Endocrinol 3, 49.

Stausbol-Gron, B., Jensen, K.B., Jensen, K.H., Jensen, M.O., Clark, B.F.C., 2001, De novo identification of cell-type specific antibody-antigen pairs by phage display subtraction: Isolation of a human single chain antibody fragment against human keratin 14. Eur J Biochem 268, 3099-3107.

Stavnezer, J., 1996, Antibody class switching. Adv Immunol 61, 79-146.

Steele, E.J., Rothenfluh, H.S., Both, G.W., 1992, Defining the nucleic acid substrate for somatic hypermutation. Immunol Cell Biol 70 (Pt 2), 129-144.

Stone, R.T., Keele, J.W., Shackelford, S.D., Kappes, S.M., Koohmaraie, M., 1999, A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. J Anim Sci 77, 1379-1384.

Ta, V.-T., Nagaoka, H., Catalan, N., Durandy, A., Fischer, A., Imai, K., Nonoyama, S., Tashiro, J., Ikegawa, M., Ito et, a., 2003, AID mutant analyses indicate requirement for class-switch-specific cofactors. Nat. Immunol. 4, 843.

Takeda, S., Masteller, E.L., Thompson, C.B., Buerstedde, J.M., 1992, RAG-2 expression is not essential for chicken immunoglobulin gene conversion. Proc Natl Acad Sci U S A 89, 4023-4027.

Taylor, B.C., Stott, J.L., Scibienski, R.J., Redelman, D., 1992, Development and characterization of a monoclonal antibody specific for the bovine low-affinity interleukin-2 receptor, BoCD25. Immunology 77, 150-154.

Teng, B., Burant, C.F., Davidson, N.O., 1993, Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science 260, 1816.

Thompson, C.B., 1992, Creation of Immunoglobulin Diversity by Intrachromosomal Gene Conversion. Trends Genet. 8, 416-422.

Thorbecke, G.J., 1959, Some histological and functional aspects of lymphoid tissue in germfree animals. I. Morphological studies. Ann N Y Acad Sci 78, 237-246.

Tobin-Janzen, T.C., Womack, J.E., 1992, Comparative mapping of IGHG1, IGHM, FES, and FOS in domestic cattle. Immunogenetics 36, 157-165.

Tonegawa, S., 1983, Somatic generation of antibody diversity. Nature 302, 575-581.

Tony, H.P., Phillips, N.E., Parker, D.C., 1985, Role of membrane immunoglobulin (Ig) crosslinking in membrane Ig- mediated, major histocompatibility-restricted T cell-B cell cooperation. J. Exp. Med. 162, 1695-1708.

Tsui, P., Tornetta, M.A., Ames, R.S., Bankosky, B.C., Griego, S., Silverman, C., Porter, T., Moore, G., Sweet, R.W., 1996, Isolation of a neutralizing human RSV antibody from a dominant, non-neutralizing immune repertoire by epitope-blocked panning. Journal Of Immunology (Baltimore, Md.: 1950) 157, 772-780.

Tunyaplin, C., Knight, K.L., 1995, Fetal VDJ gene repertoire in rabbit: evidence for preferential rearrangement of VH1. Eur J Immunol 25, 2583-2587.

Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J., Johnson, K.S., 1996, Human Antibodies with Sub-nanomolar Affinities Isolated from a Large Non-immunized Phage Display Library. Nat Biotech 14, 309-314.

Virnekas, B., Ge, L., Pluckthun, A., Schneider, K.C., Wellnhofer, G., Moroney, S.E., 1994, Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. Nucleic Acids Res 22, 5600-5607.

Wagner, S.D., Milstein, C., Neuberger, M.S., 1995, Codon bias targets mutation. Nature 376, 732.

Wagner, S.D., Neuberger, M.S., 1996, Somatic hypermutation of immunoglobulin genes. Annu Rev Immunol 14, 441-457.

Wang, H., Bash, R., Lindsay, S.M., Lohr, D., 2005, Solution AFM studies of human Swi-Snf and its interactions with MMTV DNA and chromatin. Biophysical Journal 89, 3386-3398.

Weill, J.C., Reynaud, C.A., 1996, Rearrangement/hypermutation/gene conversion: When, where and why? Immunol. Today 17, 92-97.

Weill, J.C., Reynaud, C.A., Lassila, O., Pink, J.R., 1986, Rearrangement of chicken immunoglobulin genes is not an ongoing process in the embryonic bursa of Fabricius. Proc Natl Acad Sci U S A 83, 3336-3340.

Weinstein, P.D., Anderson, A.O., Mage, R.G., 1994, Rabbit IgH sequences in appendix germinal centers: VH diversification by gene conversion-like and hypermutation mechanisms. Immunity 1, 647-659.

White, G.P., Roche, P., Brandon, M.R., Newton, S.E., Meeusen, E.N., 1998, Cloning and characterization of sheep (Ovis aries) immunoglobulin alpha chain. Immunogenetics 48, 359-362.

Whitelaw, C.B., Radcliffe, P.A., Ritchie, W.A., Carlisle, A., Ellard, F.M., Pena, R.N., Rowe, J., Clark, A.J., King, T.J., Mitrophanous, K.A., 2004, Efficient generation of

transgenic pigs using equine infectious anaemia virus (EIAV) derived vector. FEBS Lett 571, 233-236.

Wilson, M., Hsu, E., Marcuz, A., Courtet, M., Du Pasquier, L., Steinberg, C., 1992, What limits affinity maturation of antibodies in Xenopus--the rate of somatic mutation or the ability to select mutants? Embo J 11, 4337-4347.

Winter, G., Griffiths, A.D., Hawkins, R.E., Hoogenboom, H.R., 1994, Making antibodies by phage display technology. Annual Review Of Immunology 12, 433.

Womack, J.E., 1997, Mapping animal genomes. Adv Vet Med 40, 157-189.

Yancopoulos, G.D., Desiderio, S.V., Paskind, M., Kearney, J.F., Baltimore, D., Alt, F.W., 1984, Preferential utilization of the most JH-proximal VH gene segments in pre-B-cell lines. Nature 311, 727-733.

Yang, F., Waldbieser, G.C., Lobb, C.J., 2006, The nucleotide targets of somatic mutation and the role of selection in immunoglobulin heavy chains of a teleost fish. J Immunol 176, 1655-1667.

Yang, R., Murillo, F.M., Delannoy, M.J., Blosser, R.L., Yutzy, W.H.t., Uematsu, S., Takeda, K., Akira, S., Viscidi, R.P., Roden, R.B., 2005, B lymphocyte activation by human papillomavirus-like particles directly induces Ig class switch recombination via TLR4-MyD88. J Immunol 174, 7912-7919.

Yasuda, M., Jenne, C.N., Kennedy, L.J., Reynolds, J.D., 2006, The sheep and cattle Peyer's patch as a site of B-cell development. Vet Res 37, 401-415.

Zabner, J., Fasbender, A.J., Moninger, T., Poellinger, K.A., Welsh, M.J., 1995, Cellular and molecular barriers to gene transfer by a cationic lipid. The Journal Of Biological Chemistry 270, 18997-19007.

Zachau, H.G., 1993, The immunoglobulin kappa locus-or-what has been learned from looking closely at one-tenth of a percent of the human genome. Gene 135, 167-173.

Zhao, Y., Kacskovics, I., Pan, Q., Liberles, D.A., Geli, J., Davis, S.K., Rabbani, H., Hammarstrom, L., 2002, Artiodactyl IgD: the missing link. Journal Of Immunology (Baltimore, Md.: 1950) 169, 4408-4416.

Zhao, Y., Kacskovics, I., Rabbani, H., Hammarstrom, L., 2003, Physical mapping of the bovine immunoglobulin heavy chain constant region gene locus. J Biol Chem 278, 35024-35032.

Zhao, Y., Pan-Hammarstrom, Q., Zhao, Z., Hammarstrom, L., 2005, Identification of the activation-induced cytidine deaminase gene from zebrafish: an evolutionary analysis. Dev Comp Immunol 29, 61-71.

Ziegner, M., Steinhauser, G., Berek, C., 1994, Development of Antibody Diversity in Single Germinal-Centers - Selective Expansion of High-Affinity Variants. Eur. J. Immunol. 24, 2393-2400.