We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Immunoglobulin Polygeny:
An Evolutionary Perspective

J. E. Butler, Xiu-Zhu Sun and Nancy Wertz
Department of Microbiology & Interdisciplinary Immunology Program
Carver College of Medicine, University of Iowa, Iowa City,
USA

1. Introduction

The immune system of vertebrates is characterized by genes of the Ig-superfamily (IGSF) that encode the immunoglobulin (Ig) genes, genes that encode the T cell receptor (TCR), a portion of the structure of the genes encoding the major histocompatibility molecules (MHC), Ig cell surface and transport receptors, some families of cytokines and chemokines as well as numerous other proteins important to the immune system. IGSF genes also encode proteins in sponges, coelenterates and flatworms (Blumbach et al., 1998; Miller & Steele, 2000; Ogawa et al., 1998). While not a topic for this chapter, we acknowledge that the IGSF genes are not the only family of genes used to generate an antibody repertoire in vertebrates. The VLR-based receptors of jawless fishes that belong to the LRR family of receptors, have had a parallel evolution (Herrin & Cooper, 2010).

Figure 1 illustrates the signature features of proteins encoded by the IGSF genes. Highly diagnostic is the so-called "β-barrel" or "Ig fold". Anti-parallel β-pleated sheets form the staves of the barrel that are joined at each end by flexible polypeptide chains. These flexible polypeptides on the face of a heavy chain variable region domain (VH; Fig. 1A) contain three combinatorial determining regions (CDRs). The variable light chain domain (VL; not shown) also contributes three CDRs. CDRs from both VH and VL domains coalesce to form the antibody binding site (Fig. 1B). A striking feature of IGSF genes that encode the variable region domain of Igs is the degree of polygeny such that duplicated VH genes alone can occupy > three megabases (Matsuda et al., 1990). There are three such variable region loci in mammals: VH, Vλ and Vκ. The former encodes the variable heavy chain domain (Fig. 1A) while Vλ and Vκ encode the light chains variable region domains. All three loci are independent (non-linked) although a few orphan human VH genes can be found in other linkage groups (Matsuda et al., 1990). Popular textbooks suggest that this polygeny explains why antibodies can recognize >10^{10} different antigens. It is argued that if each specific antibody required a completely separate gene, more DNA would be needed than exists in the mammalian genome. To reduce the need for so many different germline encoded antibody binding sites, a system of somatic gene segment recombinations and later, somatic hypermutation (SHM) or somatic gene conversion (SCG), evolved.

The complete antibody molecule (and other proteins encoded by IGSF genes) is often composed of a tandem series of β-barrel domains as illustrated in Fig. 1B. Each domain in such multi-domain molecules differs slightly in structure and correspondingly, in function.

In this article we will focus on the genes encoding the VH domain of Ig (Fig. 1A) as well as those encoding the so-called "constant regions" (Cγ) of IgG, the mammalian flagship antibody isotype. We use as examples the sequences of duplicated VH genes in swine and bats (opposite extremes) and the duplicated Cγ genes encoding the subclasses of swine IgG, as evidence to suggest how this polygeny occurred.
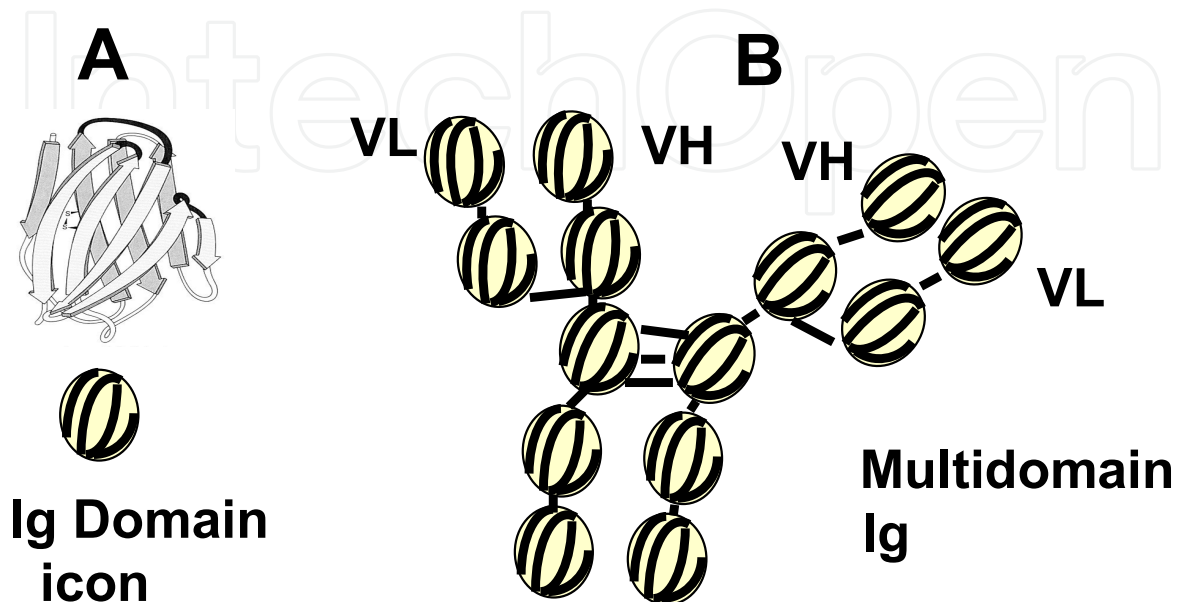

Fig. 1. Duplication/diversification of Ig genes resulted in macromolecules with repeating units. A. The variable heavy chain domain (VH) with its characteristic β-barrel or Ig fold. The dark polypeptides connecting the " barrel staves" contain the CDR regions. B. Complete Igs are multidomain molecules comprised of many Ig fold domains. The CDR-containing peptides occur only in the VH and VL domains. The remaining C-domains comprise the constant region of the Ig. The "monomeric Ig" shown in Fig. 1B is bivalent, with two identical VH/VL pairs that contain the antigen binding sites.

In the interests of those who are not immunologists, we describe the different Ig-loci, how they vary among vertebrates and the processes involved in the generation of the antibody repertoire (Section 2). Section 3 discusses the gene duplication phenomenon which resulted in the polygeny that characterizes the vertebrate Ig genome, while Section 4 reviews the somatic processes that lead to the synthesis and secretion of antibodies in higher vertebrates. Section 5 discusses the selection processes involved in gene usage. Finally, we provide data from studies in fetal/neonatal piglets, newborn rabbits and the chicken, to support the view that only a small number of the many duplicated V-region Ig genes are actually used. We provide examples in which only one or a few VH genes are needed to generate the antibody repertoire so long as the machinery for somatic recombination and somatic mutation is in place. Based on these examples and comparing them to antibody repertoire development in lower vertebrates, we hypothesize that the extensive polygeny in the Ig loci of higher vertebrates exists as an evolutionary vestige but is retained because of its redundancy value. While the recent duplication/diversification of Cγ allows for specialized effector function, IgG in rabbits did not diversify, yet few would argue against the success of this mammalian order. Thus, many of the Cγ duplicons in other mammals may also have been retained for

their redundancy value. This could explain why individuals with major deletions of certain Cγ genes remain healthy (see Section 6.3).

2. Organization of the Ig Loci

2.1 Translocon organization of gene segments characterizes higher vertebrates

Figure 2A shows the organization of the light and heavy chain loci. Each locus can be divided into subloci that, from 5′ to 3′, are known as the V, D, J and C regions. The light chain loci are similar but lack D subloci. As discussed above, the V, D, and J regions encode the antibody binding site for the heavy and light chain, and are comprised of a large number of duplicated gene segments that vary among species (Table 1). The VH and VL gene segments are the largest (~ 300 nucleotides) and encode both framework regions (FR) and CDR1 and CDR2. The FR regions encode the β-pleated sequences of the β-barrel (Fig. 1A). Displayed in linear fashion FR1, CDR1, FR2, CDR2 and FR3 comprise a VH (or VL) gene (Fig. 4 and 5). The 3′ portion of the JH segment (after the tryptophan codon; Fig. 8) encodes FR4 while CDR3 results from the recombination of V-D-J or V-J (see Section 4).
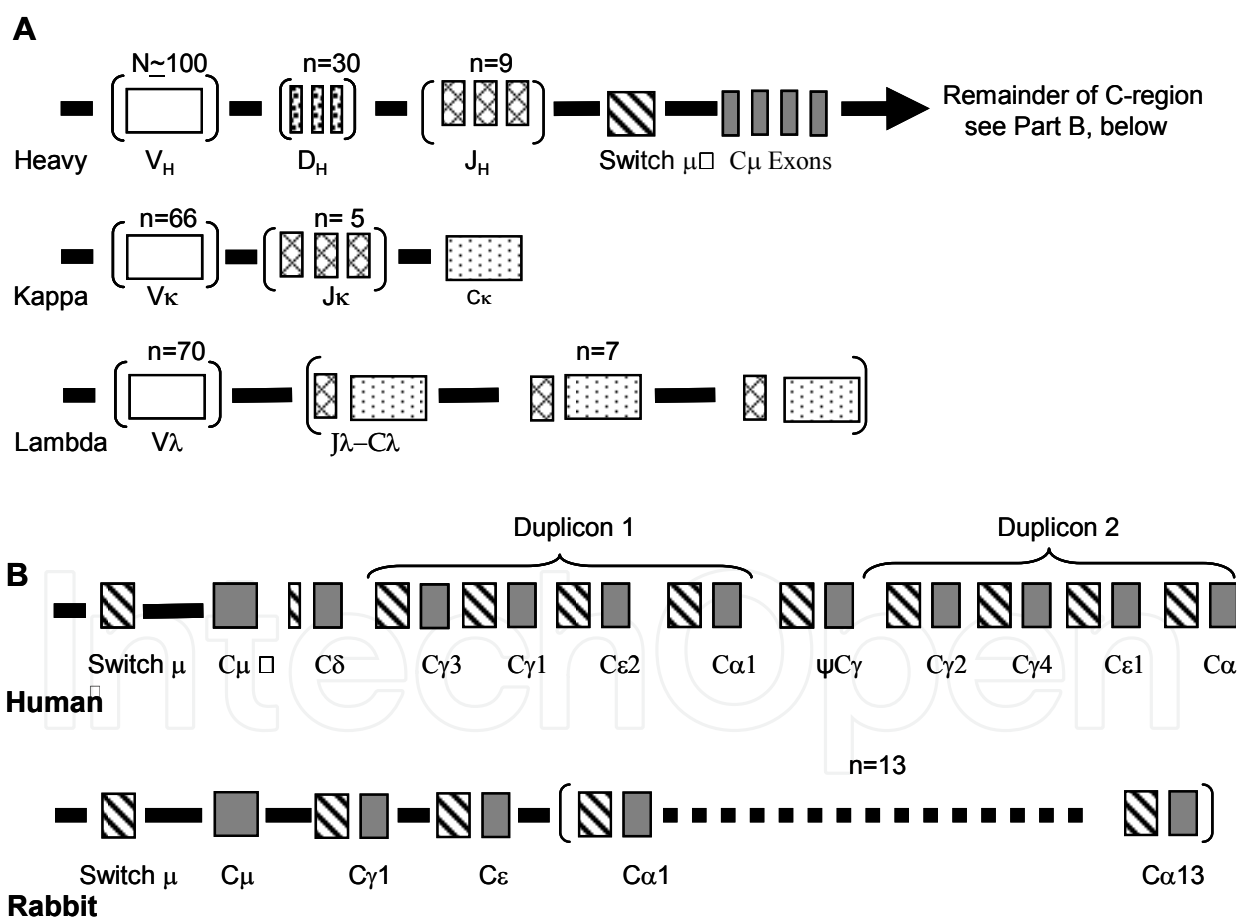
Fig. 2. The translocon organization of Ig genes of mammals. A. Organization of the variable region gene segments of the human heavy chain (VH), kappa (κ) and lambda (λ) loci. Brackets indicate the number (n) of gene segments of a particular type. Switch regions are depicted with diagonal strips. B. Organization of the constant region of the heavy chain locus of human and rabbit. The site of intralocus segment duplication in humans is indicated.

The C- region Sublocus is composed of exons of the genes that encode the "constant" domains of the antibody molecule (Fig. 1B). Fig. 2A illustrates the four exons that encode IgM (Cμ). Each exon encodes one of the constant region domains illustrated in Fig. 1B. Each domain possesses the β–barrel structures that characterizes the minimal structure of proteins encoded by IGSF genes (Fig. 1B). Within the C-region sublocus, sets of exons encode different antibody isotypes: e.g. IgM, IgD, IgG, IgE IgA (Fig. 2B). Two variations of the C-region sublocus are illustrated by the human and rabbit. The distribution of the encoded isotypes among common vertebrates is summarized in Table 2. Of interest to the theme of this review, the C-region sublocus contains a region that contains stretches of exons of duplicated Cγ genes encoding IgG subclasses (Fig. 2B; most mammals) or multiple IgA subclasses (Fig. 2B; in rabbit).

Species	V_H (F*)	D_H	J_H	V_λ (F*)	J_λ	C_λ	V_κ (F*)	J_κ	C_κ
Human	87 (7)	30	9	70 (7)	7	7**	66 (7)	5	1
Mouse	>100 (14)	11	4	3 (3)	4	4**	140 (4)	4	1
Rat	>100 (11)	?	5	15 (4)	1	1	18 (?)	6	?
Rabbit	>100 (1)	12	6	? (?)	2	2	>36 (?)	5	2
Swine	> 20 (1)	2	1	? (4)	>3	>3	60 (2)	5	1
Horse	>10 (2)	>7	>5	25 (3)	4	4	>20 (?)	5	1
Cattle	>15 (2)	3	5	30 (?)	>2	4	? (?)	?	1
Sheep	>10 (1)	?	6	>100 (3)	2	2	10 (4)	3	1
Camelid	VHH 42 (1) VH 50 (1)	10	6	? (?)	?	2	? (?)	?	?
Bat	>250 (5)	?	13	? (?)	?	?	? (?)	?	?
Oppossum	12	?	?	30 (3)	6	6	35 (4)	>2	1
Platypus	25 (1)	>5	7	15-25 (2)	6	4	? (4)	?	?

* Number of families (F) of variable region genes.

** J_λ-C_λ occurs as duplicons (see Fig. 2A).

Table 1. Variable region gene duplication among mammalian antibody genes

Species	IgM (Cμ)	IgD (Cδ)	IgG (Cγ)	IgE (Cε)	IgA (Cα)	C_λ	C_κ
Human	1	1	>4 1*	>1 1*	2	>4 3*	1
Mouse	1	1	4	1	1	>3 1*	1
Rat	1	1	4	1	1	1	?
Rabbit	1	0	1	1	13	8	2
Swine	1	1	6	1	1	>3	1
Horse	1	1	7	1	1	4	1
Cattle	1	1	3	1	1	4	1
Sheep	1	1	>2	1	1	>1	1
Camel	1	?	>3 4*	?	?	2	1
Cat	1	?	>2	1	?	>1	1
Dog	1	1	4	1	1	>1	1
Bat	1	1**	1 5**	1	1	?	?
Oppossum	1	0	1	1	1	6	1
Platypus	1	0	2	1	2	4	?

* Additional pseudogenes

** Varies with species

Table 2. Constant region gene duplication among mammalian antibody genes

2.2 The Ig loci in sharks and the chicken are differently organized

Figure 3 illustrates examples of the organization of V-D-J-C segments in three different shark species which are organized as repeating cassettes rather than in translocon fashion. Interestingly, Figure 2A also shows that an apparent evolutionary remnant of this form of organization is still found in the lambda light chain locus of mammals. In the shark, an entire cassette is used for encoding an antibody; recombination among cassettes is unusual. Furthermore, segments within the cassettes of certain sharks are fused in the genome so recombination (Section 4) does not occur. It is believed that the tandem repeat system of sharks later evolved into the translocon system (Marchalonis et al., 1998). In the translocon system, recombination among the various V, D, and J segments can occur and the rearranged VDJ is later spliced to a C region exon (Fig. 4A).
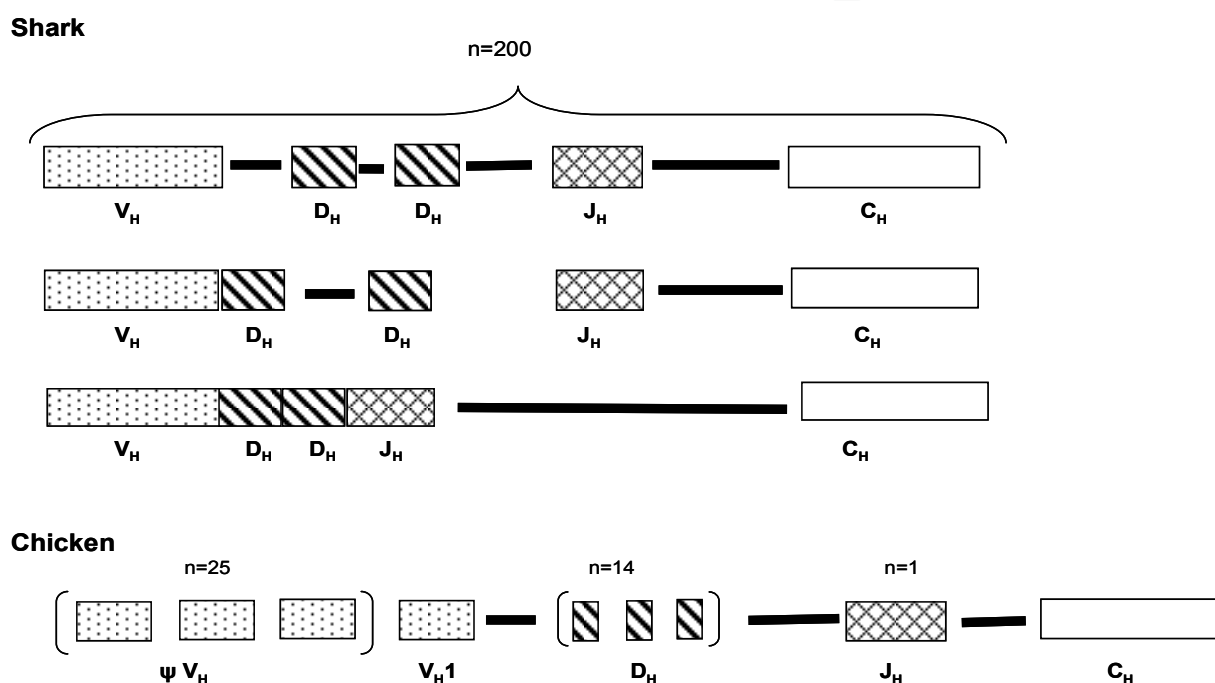
Fig. 3. Organization of heavy chain loci in sharks and chicken. Three different types of clusters are shown for sharks, some in which VDJs are fused in the genome; n= number of repeating clusters. Modified from Dooley & Flajnik, 2006. In the diagram for chicken, the number (n) of gene segments of each type is indicated. Only VH1 of chicken is a functional VH gene.

The chicken also displays a translocon system but there is only one functional VH (and one Vλ; not shown), multiple highly similar DH segments and only one JH. All VH genes upstream of VH1 in the chicken are pseudogenes (Fig. 3; Ratcliffe, 2006). These pseudo VH genes are used in SGC to create the chicken antibody repertoire (Reynaud et al., 1987; Ratcliffe, 2006).

3. Duplication and diversification of Ig genes

3.1 VH genes display evidence of duplication and genomic gene conversion

Genomic gene conversion was originally described in yeast (Meselson & Radding, 1975; Szostak et al., 1983) and is a form of non-homologous recombination in which the end result

is that a segment of one gene is "translocated" to another gene. When this process is combined with gene duplication, an array of modified duplicons results. Figure 4 shows the VH gene sequences for swine and Figure 5 the VH3 genes of the little brown bat (*Myotis lucifugus*). We have used color-coding to show that in swine, there are only four different FR1 sequences among the 24 known VH genes, only two FR2 sequences but a larger number of different FR3 sequences, although five are often shared (Fig. 4). Also shown is that CDR regions are often shared. Assuming these genes are the result of a combination of duplication and genomic gene conversion, VHT could be derived from VHE with CDR2 and FR3 translocated from VHF. A similar pattern of shared gene segments is seen among the VH3 germline gene repertoire of the little brown bat (Fig. 5). As shown, many of these share common FR1 sequences, a smaller number share CDR1, FR2 and CDR2 while the greatest diversity is seen in FR3 (Bratsch et al., 2011). This pattern of similarity among duplicated Ig genes is also seen in human and mouse, suggesting that after duplication and genomic gene conversion, the 3' segment was subjected to a higher rate of germline mutation and selection. We believe these examples support the hypothesis that the polygeny of VH took place by a combination of gene duplication and genomic gene conversion. It is well-documented that within a sublocus, intralocus duplication of segments containing several genes also occurs. This is illustrated for the C-region sublocus for humans and rabbits (Fig. 2B). The same phenomenon occurs in the VH sublocus of mice (Retter et al., 2007; Johnston et al., 2006) humans (Matsuda et al., 1990) and in swine (Eguchi-Ogawa et al., 2010). For example, the genomic segment in swine that contains VHA, VHB and VHE has been duplicated to yield VHA*, VHB* and VHF.

Genebank #	V_H gene	FR1	CDR1	FR2	CDR2	FR3
AF064686	V_HA	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	STYIN	WVRQAPGKGLEWLA	AISTSGG	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAT
AB513624	V_HA*	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	STYIN	WVRQAPGKGLEWLA	AISTSGG	STYYADSVKGRFTISRDDSQNTAYLQMNSLRTEDTARYYCAT
AF064687	V_HB	EEKLVESGGGLVQPGGSLRLSCVGSGFDFS	DNAFS	WVRQAPGKGLEWVA	AIASSDYDG	STYYADSVKGRFTISRDNSQNTVYLQMNSLRTEDTARYYCAI
AB513624	V_HB*	EEKLVESGGGLVQPGGSLRLSCVGSGFDFS	DYAFS	WVRQAPGKGLEWVA	AIASSDYDG	STYYADSVKGRFTISRDNSQNTVYLQMNSLRTEDTARYYCAI
AF064688	V_HC	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYEIS	WVRQAPGKGLEWLA	GIYSSGS	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR
	V_HD	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYEIS	WVRQAPGKGLEWVA	DICSGG	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR
AF064689	V_HE	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYAVS	WVRQAPGKGLEWLA	GIDSGSYSG	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR
AF064690	V_HF	EEKLVESGGGLVQPGGSLRLSCVGSGFDFS	SYGVG	WVRQAPGKGLESLA	SIGSGSYIG	STDYADSVKGRFTISSDNSQNTAYLQMNSLRTEDTARYYCAR
DQ886395	V_HG	EEKLVESGGGLVQPGGSLRLSCVGSGYTFS	SYGMS	WVRQAPGKGLEWLA	GIDSGSYSG	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR
DQ886392	V_HH	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYPIG	WVRQAPGKGLEWLA		STYYADYVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR
AY911501	V_HJ	EEKLVESGGGLVQPGGSLRLSCVGSGITFS	SYAVE	WVRQAPGKGLEWLA	SIGSGSYIG	STDYADSVKGRFTISSDDSQNTVYLQMNSLRTEDTARYYCAR
AF064692	V_HK	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SSPIG	WVRQAPGKGLEWLA	SIGSGSYSG	STYYADSVNGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR
AY911500	V_HL	EVKLVESGGGLVQPGGSLRLSCVGSVFDFS	SYAVS	WVRQAPGKGLEWLA	AIYSGGS	STYYADSVKGRFTISKDNSQNTAYLQMNSLRTEDTARYYCAT
AF321841	V_HN	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYSMS	WVRQAPGKGLEWLA	GIYSSGS	STYYADSVKGRFTISSDNSQNTAYLQMNSLRTEDTARYYCAR
AF321842	V_HO	EEKLVESGGGLVQPGGSLRLSCVGSGYTFS	SYPIG	WVRQAPGKGLEWLA	AISTSGS	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAT
AF321844	V_HQ	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYEIS	WVRQAPGKGLEWLA	AISTSGA	STVYADSVKDRFIYSRDNSQNTAYLQMNRNTAYLQMTYYCAR
AF321845	V_HR	EVKLVESGGGLVQPGGSLRLSCVGSGYTFS	SYPIG	WVRQAPGKGLEWLA	CIYSSGS	STYYADSVKGRFTISKDNSQNTAYLQMNSLRTEDTARYYCAR
AF321846	V_HS	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYNMI	WVRQAPGKGLEWLA	YITSSGG	STYYADSVKGRFTISSDNSQNTAYLQMTRNTAYLQMTYYCAR
AF321847	V_HT	EEKLVESGGGLVQPGGSLRLSCVGSGITFS	SYAVS	WVRQAPGKGLESLA	SIGSGSYIG	STDYADSVKGRFTISSDDSQNTVYLQMNRNTAYLQMTYYCAR
AF321848	V_HU	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYEIS	WVRQAPGKGLEWLA	AIGCGSYSG	STYYADSVKGRFTISSDNSQNTAYLQMTRNTAYLQMTYYCAR
AF321849	V_HV	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	STYIN	WVRQAPGKGLEWVA	AIASSDYDG	STYYADSVKGRFTISSDNSQNTAYLQMTRNTAYLQMTYYCAR
AY911502	V_HX	EEKLVESGGGLVQPGGSLRLSCVGSGYTFS	SYGIG	WVRQAPGKGLEWLA	GIYSGG	STYYADSVKGRFTISKDNSQNTAYLQMNSLRTEDTARYYCAR
AY911504	V_HZZ	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYSMS	WVRQAPGKGLEWLA	CIYSSGS	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR
DQ886393	V_HY	EEKLVESGGGLVQPGGSLRLSCVGSGFTFS	SYEIS	WVRQAPGKGLEWLA	AISTSGG	STYYADSVKGRFTISRDNSQNTAYLQMNSLRTEDTARYYCAR

Fig. 4. Deduced amino acid sequences for the framework (FR) and combinatorial determining regions (CDR) of germline porcine VH genes. Regions of FR and CDR regions that are shared among genes are color-coded. Those sequences that are not colored indicate segments with sequences that differ by one or a few changes that are not shared by other sequences.

Genebank #	group	FR1	CDR1	FR2	CDR2	FR3
GQ923685	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	DYSMN	WVRQAPGKGLEWVA	YTSYGSGNPI	YYADSVKGRFTISRDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923622	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	EYGMN	WVRQAPGKGPEWVS	YISSPSGSNI	YYAASVKGRFTISRENAKNSLYLQMSSLRAEDTAVYYCAR
GQ923683	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	EYGMN	WVRQAPGKGLEWVS	YISGDSSSDI	YYAASVKGRFTISRDNAKNMVYLQMNSLRAEDTALCYCVR
GQ923616	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	EYGMN	WVRQAPGKGLEWVS	YISSPSGSNI	YYAASVKGRFTISRENAKNSLYLQMSSLRAEDTAVYYCAR
GQ923644		LVESGGGLVQPGGSLRLSCAASGFTFS	NYDMH	WIRQAPGKELEWVA	HIWTDGSQK	YYAESVKGRFTISRDNTKNMAYLQMNSLRVKDTALYYCAR
GQ923675	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	NYGMN	WVRQAPGKGLEWIA	YTSSGDGNPI	YYADSVKGRFTISRDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923628		LVESGGGLVQPGGSLRLSCAASGFTFS	NYYMN	WVRQAPGKGLEWVA	SISDGSSYI	YYGEAVKGRFTISRDNTKNMLYLQMNSLRAEDSAVYYCAR
GQ923647	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	NYYMN	WVRQAPGKGLEWVA	YISSDGSSYI	NYANAVKGRFTISRDNAKNMVYLQMSSLRAEDTAMYYCAR
GQ923618	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SNSMN	WVRQAPGKGLEWVA	LISSSGGGST	YYAASVKGRFTISRDNAKNSLYLQMNSLRAEDMALYYCAR
GQ923650		LVESGGGLVQPGGSLRLSCAASGFTFS	SNWMS	WVRQAPGKGLEWVG	IISTDGGTT	NYADSVKGRFTISRDNAKNTLYLQMNSLTAKDTAVYYCAK
GQ923681	3-1	LVESGGGLVQPGGSLRLSCAASGFTFS	SSWMV	GVRQAPGKGLEWVS	LINPDGSIT	NYANSVKGRFTISSDNAKNMLYLQMNSLRAEETAMYYCAR
GQ923662	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYDMN	WVRQAPGKGLEWVA	LISTDGGST	YYANSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYFCAR
GQ923669	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYDMN	WVRQAPGKGLEWVS	LISPSGGST	YYADSVKGRFTISRDNAKNMVYLQMSSLKAEDKAVYFCAR
GQ923625	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYDMN	WVRQAPGKGLEWVS	LISPSGGST	YYADSVKGRFTISRDNAKNMVSLQMSSLRAEDTAVYYCAR
GQ923612	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYDMN	WVRQAPGKGLEWVA	YISSASNTI	YYANSVKGRFTISIDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923621	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYDMS	WVRQAPGKGLEWVS	AISNGGGST	YYAASVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCAR
GQ923679		LVESGGGLVQPGGSLRLSCAASGFTFS	SYGMH	WVRQAPGKGLEWVA	YQYISSDGRNYI	NYAASVKGRFTISRDNAKNTAYLQMNSLKAEDTAVYYCAR
GQ923617		LVESGGGLVQPGGSLRLSCAASGFTFS	SYGMH	WVRQAPGKGLEWVA	YQYISSDGRNYI	NYAASVKGRFTISRDNAKNMLYLQMSSLRAEDMAVYYCAR
GQ923658	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	SYGMH	WIRQAPGKGLEWVS	RIGSDGRSYI	HYADSVKSRFTISRDNAKNMLYLQMSSLRSEDTALYYCAR
GQ923672	3-4	LVESGGGLVQPGGSLRLSCAASGFTFS	SYGMN	WVHQVLGKGLECVS	GVSSIGGTT	YYADSVKGRFTVSRDNTTSMLYLQMNSLRTEDMAVYYCAR
GQ923613	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	SYGMY	WVGQVPGKGLEWVS	LISSDGSSTI	YYANSVKGRFTISRDNAKNTLYLQMNSLRAEDTAVYYCAR
GQ923657	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYHMN	WVRQAPGKGLEWVA	FISNGGSTI	YYAASVKGRFTISRDNAKNMLYLQMNSLRAEDTALYYCAR
GQ923638		LVESGGGLVQPGGSLRLSCAASGFTFS	SYQMH	WVRQAPGKGLEWVE	LISSSGGTI	YYADSVKGRFTISRDNAKNTLFLQMSSLRADDTAMYYCAR
GQ923629	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYSMD	WVRQAPGKGLEWVA	YISSASSTI	YYANSVKGRFTISRDNAKNTLYLQMSSLRAEDTAMYYCAR
GQ923642	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYSMN	WVRQAPGKGLEWVA	VISSSGGTI	YYADSVKGRFTISRDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923684	3-1	LVESGGGLVQPGGSLRLSCAASGFTFS	SYWMD	WVRQAPGKGLEWLC	RMNPDGSTT	HYANSVKGRFTISRDNAKNMLYLQMNSLRAEETAMYYCAR
GQ923619	3-3	LVESGGGLVQPGGSLRLSCAASGFTFS	SYWMH	WVRQAPGKGLEWVS	RISSSGSTI	SYAASVKGRFTISRDNTKNTLYLQMSSLRAEDTAVYYCAR
GQ923632	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	SYWMH	WVRQAPGKGLEWVS	LISSDGSSTI	YYANSVKGRFTISRDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923668	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	SYWMH	WVRQAPGKGLEWVS	LISSDGSSYI	YYAASVKGRFTISRDNAKNTLYLQMSSLKAEDTAVYYCAR
GQ923640		LVESGGGLVQPGGSLRLSCAASGFTFS	SYWMN	WVRQAPGKRLEWVS	AISSSGGST	YYADSVKGRFTISRDNTLFLQMSSLRVEDTAVYYCAR
GQ923627		LVESGGGLVQPGGSLRLSCAASGFTFS	SYWMS	WVRQAPGKGLEWVA	HINSGGST	YYADSVKGRFTISRHNAKNSLYLQMSSLRAEDTVGYYCVR
GQ923663	3-1	LVESGGGLVQPGGSLRLSCAASGFTFS	SYWMY	WVRQAPGKGLEWLC	RMNPDGSTT	NYANSVKGRFTISRDNAKNTLYLQMNSLSTQDTAMYYCST
GQ923623	3-2	LVESGGGLVQPGGSLRLSCAASGFTFS	SYYMN	WVRQAPGEGLEWVA	SISSGGGSYI	YYAASVKGRFTISRDNTKNTLYLQMNSLRAEDTAVYYCAR
GQ923680		LVESGGGLVLPGGSLRLSCAASGFTFS	DYWVH	WVCQAPWKGLEWVS	DFRGDGGTT	YYADSVKGRSTLSRDNVKNSLYLQMSSLRAKDTAIYYCAR
GQ923635		LVESGGGLVLPGGSLRLSCAASGFTFS	GYWIS	WARQAPGKKLEWVS	DISGDSSIT	YYAASVKGRFTISRDNAKNTLYLQMSSLRAEDTALYYCAR
GQ923636		LVESGGGLVLPGGSLRLSCAASGFTFS	NYDMH	WIRQAPGKGLEWVA	HIWTDGSQK	YYAESVKGRFTISRDNAKNSLYLQMNSLKAEDSALYYCAR
GQ923673	3-1	LVESGGGLVLPGGSLRLSCAASGFTFS	SNWMH	WVRQAPGKGLEWLC	RMNPDGSTT	NYANSVKGRFTISRDSAKNMLYLQMNSLRAEETAMYYCTT
GQ923655	3-1	LVESGGGLVLPGGSLRLSCAASGFTFS	SYWMH	WVRQAPGKGPEWLC	QINSDGNTI	YYANSVKGRFTISRDNAKNMLHLQMNSLRAEESALYYCAR
GQ923615	3-1	LVESGGGLVLPGGSLRLSCAASGFTFS	SYWMH	WVRQAPGKGLEWLC	QINSDGNTI	YYANSVKGRFTISRDNAKNMLHLQMNSLRAEESALYYCAR
GQ923676	3-3	LVESGGGLVQPGGSLRISCAASGFTFS	SYWMS	WVRQAPGKGPEWVS	HISDGGGST	YYANSVKGRFTNSRDNAKNSLSLQMNSLKPEDTALYYCAR
GQ923631		LVESGGGLVQPGGSLRLSCAASGFIFS	SYGMS	WVRQAPGNGLEWVS	GVSSIGGTTG	YYADSVKGRFTVSRDNGKNMLFLQMNSLRAKDTAVYYCAR
GQ923611	3-2	LVESGGGLVQPGGSLRLSCAASGFSFS	SYWMG	WVRQAPGKGLEWVA	LISTGGGGNT	YYATSVKGRFTISRDNAKNSLYLQMSSLRAEDTAVYYCAR
GQ923667	3-3	LVESGGGLVPPGGSLRLSCAASGFSFS	IYGMN	WVRQAPGKGLEWVS	GISTGGGST	YYAASVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCAR
GQ923654	3-2	LVESGGGLVQPGGSLRLSCAATGFTFS	SYWMH	WVRQAPGKGLEWVS	LISSDGSSYI	YYAASVKGRFTISRDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923645	3-3	LVESGGGLVQPGGSLRLSCSGSGFTFS	SYSMD	WVRQAPGKGLEWVA	YISSASNTI	YYANSVKGRFTISIDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923660	3-3	LVESGGGLVQPGGSLRLSCSGSGFTFS	SYSMD	WVRQTPGKGLEWVA	YISSASNTI	YYANSVKGRFTISIDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923677	3-6	LVESGGGVVPPGGSLSLSCKASGFTFT	NYSMD	WVSQAPGKGLQWVT	RVSKPKGMTQ	WYAPAVQGRFTIFRDNPMSTASLEITKLTSEDMAMYYCAR
GQ923648	3-6	LVESGGGVVPPGGSLSLSCKASGFTFT	NYSMD	WVSQAPGKGLQWVT	RVSKPKGMTQ	WYAPAVQGRFTIFRDNPMSTVSLEITKLTSEDMAMYYCTR
GQ923671	3-6	LVESGGGVVPPGGSLSLSCKASGFTFT	NYSMD	WVRQAPGKGLQWVA	RVSQPKGTTQ	WYAPAVRGRFTISRDNPTSTVSLKMTKLTSEDTAVYYCTR
GQ923626		LVESGGGDLVQPGGSLRLSCAASGFTFS	SYWMH	WVRQAPGKGLEWVS	LVNPAGSST	YYANSVKGRFTISRDNAKNTLYLQMSSLRAEDTGVYYCAR
GQ923653		LVESGGGDLVQPGGSLRLSCAASGFTFS	SYWMH	WVRQAPGEGLEWVS	LINPAGSST	YYANSVKGRFTISRDNAKNMLYLQMNSLRAEDTALYYCSR
GQ923674	3-1	LVESGGGFVQPGGSLRLSCAASGFTFS	SSSMD	WVRQAPGKGLEWLC	RINPDGSTT	YYANSVKGKFTTSRDNAKNLLYLQMNSLRAQDMAVYYCVT
GQ923652		LVESGGGLGKPEGSLRLSCAASGFASS	SYYMN	WIRQTPGKGLEWMA	VISYNGNNT	YYADSVKGRFTISRDNAKNMVYLQMSSLRAEDTALYYCVR
GQ923661	3-5	LVESGGGLGKPEGSLRLSCAASGYASS	SYYMN	WVRQTPGKGLEWIC	AITANGDST	YYADSVKGRFTISRDNAKNTIYLQMSSLKSEDTAVYYCST
GQ923624	3-5	LVESGGGLGKPEGTLRLSCAASGYASS	SYYMN	WVRQTPGKGLEWIC	AITGNSDST	YYADSVKGRFTISRDNAKNTIYLQMNSLRAEDMTVYYCAT
GQ923633	3-5	LVESGGGLGKPEGTLRLSCAASGYASS	SYYMN	WVRQTPGKGLEWIC	AITGNSDST	YYADSVKGRFTISRDNAKNTIYLQMNSLRAEDTGVYYCAK
GQ923637	3-3	LVESGGGLVKPGGSLRLSCAASGFTFS	SYSMD	WVCQAPGKGLEWVA	YISSASSTI	YYANSVKGRFTISIDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923614	3-2	LVESGGGLVKPGGSLRLSCAASGFTFS	EYGMN	WVRQAPGKGLEWVS	YISGDSSSDI	YYAASVKGRFTISRDNAKNMVYLQMNSLRAEDTAVYYCAR
GQ923651		LVESGGGLVPHGGSLRLSCAASGFTFN	DYEMN	WVRQAPGKGLEWVS	RITSTGGST	HYAASVKGRFTISRDNAKNTLFLQMNSLRAEDQAMYYCTA
GQ923659	3-1	LVESGGGLVPPGGSLRLSCAASGFTFS	SYWNT	WVRQAPGKGLEWLG	EINPDGSTT	NYANAVKGRFTISRDNTKNMLYLQMNSLSAQDMAVYYCVR
GQ923646		LVESGGGLVPPGGSLRLSCAASGFTFS	SYWMG	WVRQAPGKGLEWVS	FVTDYGGSI	YYADSVKGRFSISRDNAKNTLYLQMTSLRATDTAVYYCAR
GQ923630		LVESGGGLVPPGGSLRLSCATSGFTFS	GYWIS	WARQAPGKKLEWVS	DINGDSSTT	YYAASVKGRFTTSRDNAKNMLFLQMSSLRAEDTAVYYCAR
GQ923682		LVESGGGLVPPGGSLRLSCATSGFTFS	GYWIS	WARQAPGKKLEWVS	DINGDSSTT	YYAASVKGRFTISRDNAKNTLYLQMSSLRAEDTALYYCAR
GQ923649	3-2	LVESGGGLVQPGGSLRLSCAASGFSFW	SYPMN	WVRQAPGKGLEWVA	LISSGRDGNT	YYATSVKGRFTISRDNAKNSLYLQMSSLRSDDTALYYCAR
GQ923634		LVESGGGLVQPGGSLRLSCAASGFTFD	DYYMH	WVRQAPGKGLEWVT	SISEGGSYI	YYANAVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCAR
GQ923641	3-2	LVESGGGLVQPGGSLRLSCAASGFTFG	SYWMH	WVRQAPGKGLEWVS	RIGSDGSSYI	YYADSVKGHFTISRDNAKNMLYLQMSSLRAEDTAVYYCAR
GQ923665	3-3	LVESGGGLVQPGGSLRLSCAASGSTFS	SYSMD	WVRQAPGKGLEWVA	YISSASSTI	YYANSVKGRFTISRDNAKNSLYLQMSSLRAEDTAVYYCAR
GQ923670	3-3	LVESGGGLVQPGGSLRLSCTGSGFSFS	SYSMD	WVRQAPGKGLEWVA	YISSASNTI	YYANSVKGRFTISIDNAKNTLYLQMSSLRAEDTAVYYCAR
GQ923639	3-4	LVESGGGLVQPGGSLRLSFAASGFTFS	SYGMN	WAHQFPGKGLEWVS	GVSSIGGTT	YYADSVKGRFTVSRDNAKNTLYLQMSLRTEDMAVYYCAR
GQ923656	3-4	LVESGGGLVQPGGSLRLSFAASGFTFS	SYGMN	WAHQFPGKGLEWVS	GVSSIGGTT	YYADSVKGRFTVSRDNAKNTLYLQMSLRTEDMAVYYCAR
GQ923620		LVESGGGLVQPWGSLRLSCKGSGFTFS	DYYMN	WIRQTPGKGLEWVG	LIRNKANGHTT	EYAASVKGRVAISRDDSKSTADLQMNSLIAEDTAMYNYAR
GQ923664		LVESGGGLVQTGGSLRLSCAASGFTFS	SYWMN	WVRQAPGKGLEWVS	TITHTDGST	YYPDSVKGRFTISRDNTKNMLYLQMNSLRSEDTAVYYCAR
GQ923643	3-6	LVESGGGMVPPGGSLSLSCKASGFTFT	NYSMY	WVRQAPGKGLQWVA	GVSKPTGKNQ	WYAPAVQGRFMISRDNPTSTVSLEITKLTSEDTALYYCSR
GQ923678		LVESGGGLVQPGGSLRLSCAASGYTFS	SYWMG	WVCQAPCKGLEWVS	LITGNGGST	YYANTAKVRFTISRDNTKNMLYLQMNSLTAEDTAVYYCAR
GQ923666	3-5	LVESGGSLVQPEGSLRLTCAASGFTSN	SYYMN	WFHQTPGKGLEWIC	AITANGDST	YYSDSVKGSFTISRDTAKNPIYLQMSSLRSKDTAVYYCAR

Fig. 5. The deduced amino acid sequences for the framework (FR) and combinatorial determining regions (CDR) for 75 members of the VH3 family from the little brown bat (*M. lucifugus*). Shared sequences are color-coded as in Fig. 4. From Bratsch et al., 2011.

3.2 Duplication and deletion of Ig genes in the C-region sublocus

Figure 2B illustrates that in humans, a major segment of the C- region sublocus has been duplicated resulting in two IgEs, two IgAs and two pair of IgG genes (Lefranc et al., 1982; Flanagan & Rabbitts, 1982). A Cγ pseudogenes separates them. In mammals, the target of duplications in the C-region has been those genes encoding IgG (or IgA in rabbits; Fig. 2B; Table 2). The duplication process in the CH region suggests it also occurred together with genomic gene conversions to produce an array of modified Cγ genes (Fig. 6). In the example provided the IgG1 and IgG2 alleles share a common CH1 domain that is also found in the IgG4 alleles. The allelic variants of IgG1 and IgG4 differ only in their hinge exons. However, IgG1[a] and IgG4[a] have the same hinge as do IgG1[b] and IgG4[b]. The difference between IgG1 and IgG4 is in the CH3 domain which is not shared with any other Cγ subclass gene. Another example is IgG5[a] and IgG6 which share a common CH1 and hinge exons. IgG5[a] also has the same CH2 exon as IgG6[b], but the difference is in CH3. Thus Fig. 6 also shows that apparent genomic gene conversion events also involve allelic variants as well as entire genes. The pattern shows that the CH1, hinge and CH2 of

IgG1[a] and IgG4[a] were derived from the same ancestral Cγ gene but the IgG1[b] and IgG4[b] were derived from another ancestral gene. The reverse effects of genomic gene conversion may explain certain heterozygous Cγ deletions (Migone et al., 1984; Keyeux et al., 1989). Some swine lack certain Cγ genes (Butler et al., 2009a) and deletions in the human Cγ sublocus are well documented (LeFranc et al., 1983a; Rabbani et al., 1995).
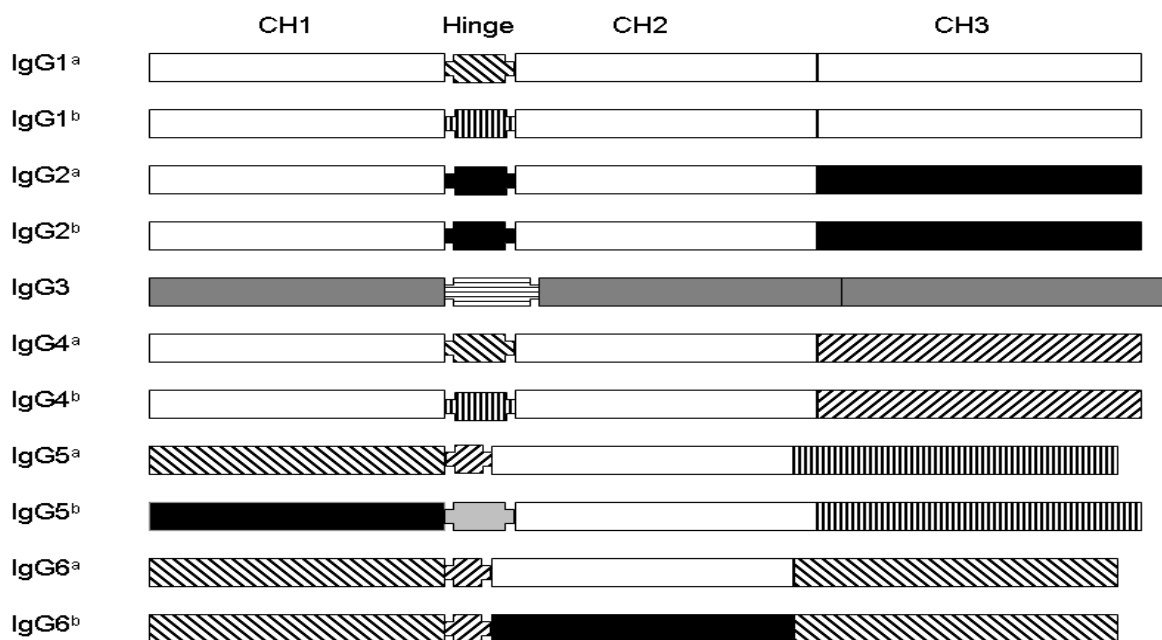
Fig. 6. Alignment of the constant region domains of the porcine Cγ genes. Regions of >95% homology are designated with the same texture. Superscripts in the gene designation denote allotypic variants. From Butler et al., 2009a.

As we have shown elsewhere, porcine IgG3 has a gene structure which is most similar to the consensus Cγ genes of other mammals and therefore is closest to the ancestral Cγ gene of all mammals (Butler et al., 2009a). IgG3 in humans, mice andswine all occupy the same 5′ position which is immediately downstream of Cδ (Eguchi-Ogawa-et al., 2010). Our studies indicate that the remainder of the porcine Cγ genes were derived from an ancestral Cγ that

diverged early from IgG3 (Butler et al., 2009a). We hypothesize that the duplication/diversification of Cγ subclass genes in mice and humans followed the same pattern.

4. Mammalian antibody repertoires result from somatic events

4.1 Somatic gene segment recombination characterizes B cell lymphogenesis

B lymphocytes, named because they form in the **B**one marrow or the chicken **B**ursa of Fabricius, are the cells that synthesize and secrete antibodies. This developmental process occurs in what are called "primary lymphoid tissues". These include the bone marrow, the chicken bursa, fetal liver, yolk sac and according to some, certain hindgut lymphoid tissues of artiodactyls. Among lower vertebrates, other tissues like the "head kidney"(pronephros), epigonal organ and Leydig organ are involved in this process (Solem & Stenvik, 2006; Rumfelt et al., 2002; Dooley & Flajnik, 2006) .

Somatic recombination is illustrated in Figure 7A. This process is mediated by Recombinase Activation Genes (RAGs) as well as a variety of DNA repair and ligation enzymes. In the heavy chain locus this process first involves recombination of one J region gene segment and one D region gene segment. This event also produces a circular DNA product containing the intervening DNA sequence that is excised and is known as a signal joint circle (Fig. 7B). Single joint circles are diagnostic evidence that B cell lymphogenesis has recently occurred since this nuclear product is rapidly degraded. The rearrangement process then proceeds to the rearrangement of the DJ unit with some V gene segment and generation of another signal joint circle (Fig. 7A). The selection of the J, D, and V gene segments is poorly understood and will be discussed in Section 5. A similar series of events occurs among segments in the light chain loci except that there are no D segments involved.
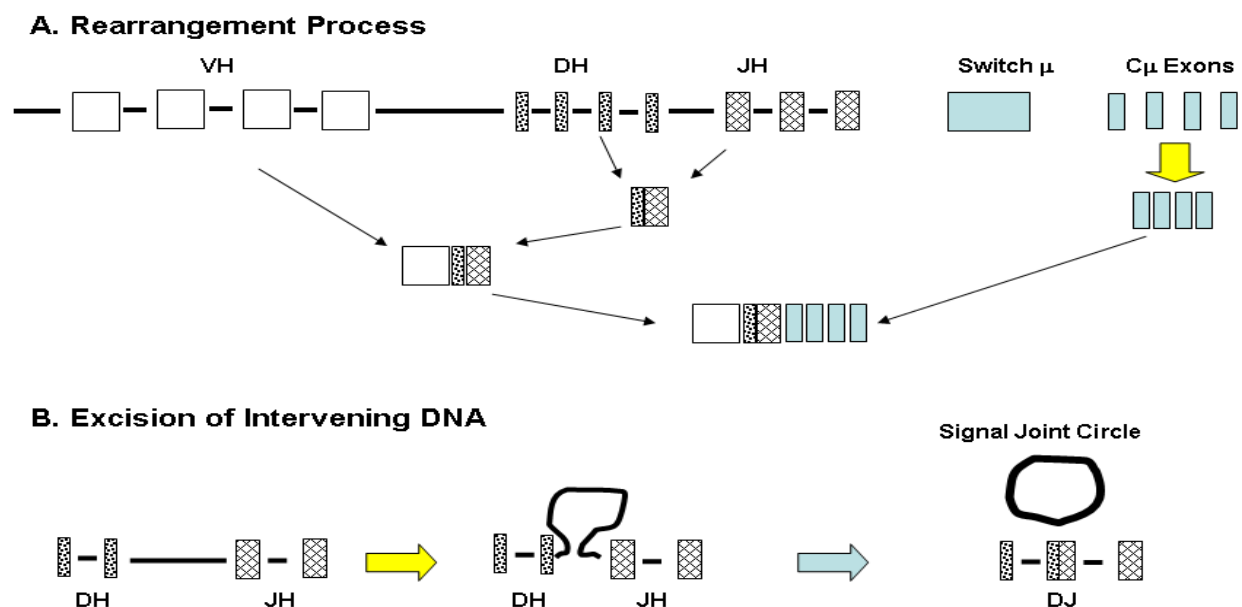
Fig. 7. The somatic rearrangement process among the gene segments of the variable heavy chain locus of mammals. A. Sequential rearrangement of D to J, then DJ to V and finally splicing of the primary transcript for VDJ to the exons encoding the C-region of IgM. B. Generation of a signal joint circle during the excision of intervening DNA during recombination of D and J.

The rearranged VDJ (heavy chain) and VJ (light chains) rearrangements are then transcribed and the primary transcript spliced to some set of C-region exons in the heavy and light chain loci respectively. For the heavy chain this is initially IgM in all higher vertebrates (Fig. 7A). The resulting VDJ-C and VJ-C transcripts are then translated into the light and heavy polypeptide chains that combine to form the complete antibody molecule (Fig. 1B). As shown in Fig. 1A and discussed above, the antigen-binding site is located in the peptide loops from the VH and VL domains that coalesce at the end of these V-domain and that contain the CDRs; three from the VDJ and three from the light chain VJ rearrangements. CDR1 and CDR2 of the VH and VL are encoded within the germline genes whereas the CDR3 region is the result of the combining of V-D-J segments (heavy chain) and V-J (light chains; Fig. 8). The CDR3 region of the heavy chain, hereafter designated HCDR3, is considered most important to the specificity of the binding site (Amit et al., 1986; Padlan, 1996; Xu & Davis2000; Mageed et al., 2001). In fact the same set of V-genes encoding CDR1 and CDR2 can theoretically and actuallycontribute the binding site for antibodies of quite different specificity in the context of different HCDR3 regions (Thomson et al., 2011; Ichiyoshi & Casali, 1994) as might be envisioned from a comparison of the CDR3 sequences shown in Fig. 8.

FR3	5' N-add	D$_H$A	3' N-add	J$_H$	FR4
TGNGCNAGN		GAATTGCTATAGCTATGGTGCTAGTTGCTATATGATGAC			ATTACTATGTTATGGATCTCTGGGGCCCA
TGTGCAAGT		TGCTATAGCTATGGTGCTAGTTGC	TTTTGGACAAGATCA		TACTATGCTATGGATCTCTGGGGCCCA
TGTGCAAGA	GGCTGTTTTC	GCTATAGCTATGGTGCTAGTTGCTAT	GATGTCG		ACTATGCTATGGATCTCTGGGGCCCA
TGTGCAA	CAGGCGAT	TGCTATAGCTA	GGTGCTAGTTGCACCGGGATG		GCTATGGATCTCTGGGGCCCA
TGTGCAA	TT	GCTATAGCTATGGTGCTAGTTG	T		TATGGATCTCTGGGGCCCA
TGTGCAA	CAGAG	TGCTATAGCTATGGTGCTAGTTGCTATAT	GTATGC		TATGGATCTCTGGGGCCCA
TGTGCAAGA	G	ATAGCTATGGTGCTAGTT	ACCCCTC		TATGGATCTCTGGGGCCCA
TGTGC	CCAG	GCTATAGCTATGGTGCTAGT	CCAGGATG		TGGATCTCTGGGGCCCA
TGTGCAA	CAGGC	ATAGCTATGGTGCTAGTTGCTAT	GAAGA		TGGATCTCTGGGGCCCA
TGTGCAAG	GTCC	AATTGCTATAGC	TCCGGTGGTGAGTGCTATGGTTACCCTTGGGGTTATGTTGCTG		TGGATCTCTGGGGCCCA
TGTGCAA	TT	GCTATAGCTATGGTGCTAGTT	AGATC		GGATCTCTGGGGCCCA

Fig. 8. The diversity of HCDR3 sequences resulting from the recombination of the same VH, DH and JH segments. Remnants of the DH germline segments are underlined. The 5′ and 3′ nucleotide additions are indicated. TGG is the codon for the invariant tryptophan found in the JH gene segments of all mammals while TGT is the codon for C that is nearly invariant in the FR3 of all VH3 family genes.

4.2 Maturation of the antibody repertoire involves class switch and somatic hypermutation

All immunologists, immunopathologist and physicians in specialties such as rheumatology know that most Igs are IgG (serum) or IgA (secretions). This means that the rearrangements involved in B cell lymphogenesis that initially favors the expression of IgM (Fig. 7A) switch to these isotypes. After environmental exposure, the concentration of the major Igs in serum is elevated 100-300 fold compared to newborn piglets or those reared in germfree isolators (Fig. 9A). The transition from newborn to conventionally-reared young adults favors IgG in serum (Fig. 9A) and IgA in secretions (Butler et al., 2011a). This change involves class switch recombination (CSR) which is mediated by activation-induced cytidine deaminase (AID) of the APOBEC family which facilitates the splicing of RNA encoding the rearranged VDJ to transcripts encoding IgG and IgA rather than IgM. This maturation process typically occurs in tandem with somatic hypermutation (SHM) of the rearranged VDJs or VJs prior to their transcription. SHM is another mechanism for repertoire diversification and is triggered

when the neonate encounters environmental antigen (Fig. 9B). Since both CSR and SHM occur simultaneously, it is not surprising that both are mediated by AID and that AID is also correlated with SGC (Withers et al., 2005; Arakawa et al., 1996). These events occur in germinal centers (GCs) of secondary lymphoid tissues after exposure to environmental antigen. GCs are found only in mammals and birds (Yasuda et al., 2003; Vigliano et al., 2006; Du Pasquier et al., 2000). Although lacking GCs, there is SHM and CSR in *Xenopus* (Marr et al., 2007) although it may be less efficient. However, for these events to occur, the naïve immune system must first or simultaneously be exposed to Pathogen Associated Molecular Patterns (PAMPs) that are recognized by a variety of innate immune system receptors. This dependence was demonstrated using the isolator piglet model (Butler et al., 2002; 2005; 2009b; Butler & Sinkora 2007).

SHM is not random across the entire rearranged VDJ-C transcript. Rather it is largely concentrated in the CDR regions of the rearranged VDJ or VJ segments (Fig. 9C). This is generally believed to result from selection of B centrocytes in GCs rather than specific targeting. Although Fig. 9C only shows the accumulation of somatic mutations in CDR1 and CDR2, the same occurs in CDR3. As discussed previously, the CDRs are those segments of the encoded protein that coalesce to form the antibody binding site (Fig. 1A; Fig 8). There is little evidence to suggest that SHM proceeds downstream from segments of transcript that begins with the codon for the invariant tryptophan in FR4 (Fig. 8) or to sequences further downstream in the C-sublocus.

4.3 The association of VH genes and VH- VL pairing in generation of specific antibodies

Much of the early studies on antibody specificity that appeared when VH or VL polygeny became known, attempted to correlate particular response to the use of certain VH or VL genes. We do not review that literature here but do provide a few examples. Cerato et al., (1997) studied hybridomas to show a lack of correlation between VH usage and specificity while Mo and Holmdahl (1996) show that mAbs to different epitopes used the same VH/Vk combinations. Boffey et al., (3004) showed that only 6/15 anti-LPS mAbs used the same VH gene (VH7183.3b). These observations should not be surprising considering the importance of HCDR3 in the specificity of antibodies (see Sectiion 4.3; 6.2). Lavoie et al., (1997) showed that nearly all mAbs to HEL use VH36-60 but differ in affinity because of SHM or HCDR3 differences. The antibody binding site involves CDRs (including CDR3) of both H and L chains (Fig 1B); this has been shown by separation and reassociation experiments. These experiments show that binding site specificity depends on both H and L chains even for antibodies specific for the same hapten since heterologous light chains seldom restore the full binding site (Kranz & Voss, 1981). This mutual dependence is also demonstrated by the non-random pairing found in antibodies of certain specificity such as to the capsular polysaccharides of *S. pneumoniae* (Thomson et al., 2011). Further evidence for the effect of H-L pairing comes from studies of autoantibodies in a phenomenon called "receptor editing" (deWildt et al., 1999). This *in vivo* phenomenon involves reactivation of recombinase activity in lymph nodes resulting in the replacement of the light chain with a new one. In this way, B cells expressing autoreactive BCRs acquire a new light chain which alters their specificity and removes or diminishes their autoreactivity apoptotic elimination (Tiegs et al., 1993; Gay et al., 1993).

While L-H pairing is important for binding site specificity, there are situations in which light chains are not needed to form an antibody binding site. The best known examples are the

naturally occurring single chain antibodies of the camelid group and some sharks (Hamers-Casterman et al., 1993; De Genst et al., 2006; Dooley et al., 2003; Diaz et al., 2002; Nguyen et al., 2002). Based on the convenience of producing single chain antibodies from these species for therapy and the evidence that the HCDR3 domainplays the major role in forming the antibody binding site (see Section 6.2) there have also been various attempts to developsynthetic single chain antibodies or "camelized" antibodies (Janssens et al., 2006; Reiter et al., 1999; Davies & Riechmann,1995). Among the camelids, single chain antibodies use a separate set of VH genes (called VHH) that encode a much largerportion of the binding sites than the conventional VH genes which compensates for the lack of a light chain. This topic has been recently reviewed (Muyldermanns et al., 2009). We mention these single chain antibodies here because we believe they further support the role played by HCDR3 in binding antigen and diminishes the value of polygeny of conventional VH genes (Section 6.2). It also shows that the extensive and universal Vκ and Vλ polygeny among mammals (Table 1) is unnecessary.
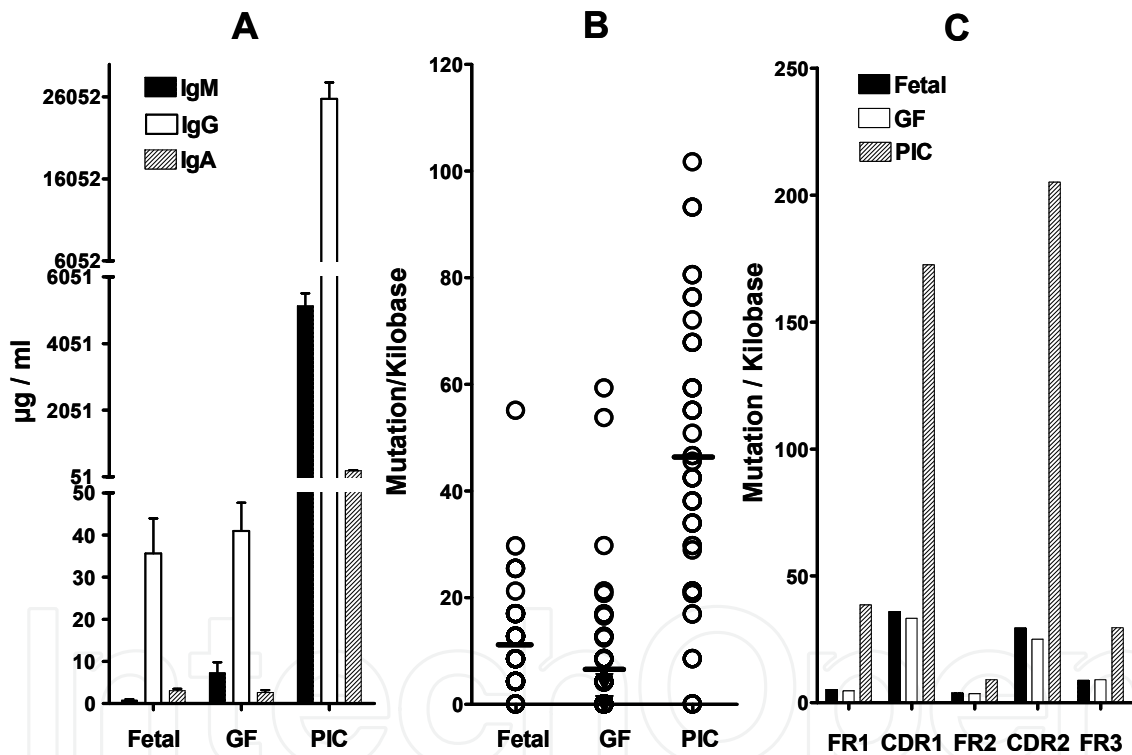
Fig. 9. The effect of antigen exposure on: A. Serum Ig levels; B. Frequency of SHM and C. Accumulation of somatic mutations in various segments of the VH genes. Germfree piglets are reared in isolators for 5 weeks and their only contact with potentially foreign antigen is food protein. PIC=conventionally-reared young pigs that are heavily antigenized through colonization and also infected with nematodes. The horizontal line (9B) in the scattergram is the mean frequency of SHM. SHM is significantly greater in PIC piglets than in fetal and germfree piglets. In 9C SHM accumulates in CDR regions as opposed to FR regions that encode the β-pleated "staves" of the β-barrel (Fig. 1A).

5. Patterns of V, D and J gene segment usage

5.1 VH usage is biased to favor certain VH genes

Table 1 shows that higher vertebrates have many duplicated V-region gene segments available for use in the formation of their antibody repertoire using the recombinatorial process illustrated in Figure 7. Humans have available ~ 100 VH segments, ~30 DH segments and 9 JH segments (Fig. 2A). By contrast, swine have <30 VH genes belonging to a single family (Fig. 4), only two functional DH segments and like the chicken (Fig. 3) one functional JH segment (Sun et al., 1994; Butler et al., 1996; Eguchi-Ogawa et al., 2010). While the ancestral VH3 family (Schroeder et al., 1990) dominates the V-region loci of many species, the ~100 VH genes of mice and human belong to 14 and 7 different families respectively (Table 1).

Usage of VH genes in rabbit is biased to the most 3′ VH gene, which accounts for 90% of VH usage in the pre-immune repertoire although there are >100 VH genes in the rabbit repertoire (Currier et al., 1988; Table 2). In humans there is bias for V3-23,V3-30,V3-33 and V4-34 (Glas et al., 2000). While some suggest that VH usage is random in mice (Dildrop et al., 1985) studies on J558 usage (one-half of the mouse genome) indicates that usage in unequal and rather scattered across the entire J558 genome even in the pre-immune repertoire (Gu et al., 1991) and that usage is not affected by SHM or CSR. Foster et al., (1997) showed that while most Vk genes were used, usage was non-random and the same was true for Jκ. Sheehan et al., (1993) showed that fetal VH usage can differ from 0.1 to 1.0 but that most 5′ VH genes are underrepresented. In swine VHA (IGHV4) and its near duplicate (IGHV10; see Figs. 4 & 10) account for one-third to one-half of the pre-immune repertoire (Butler et al., 2006; Eguchi-Ogawa et al., 2010; Butler et al., 2011b). Interestingly, the majority of these preferred genes in all these species belong to the ancestral VH3 family (Schroder et al., 1990; Brezinschek et al., 1997).

5.2 Variable region gene segment usage is not position dependent

Early studies suggested that VH usage was biased during early stages of B cell lymphogenesis to favor the most JH proximal DH segments and the most 3′ VH genes (Schroeder et al., 1987; Yancopoulos et al., 1984) but that this pattern became "normalized" in adults (Malynn et al., 1987). This concept gained support when it was found that young rabbits use their 3′ most VH gene > 90% of the time and then further diversfied their repertoire using upstream VH genes and SGC; perhaps a type of "developmental normalization (Knight 1992; Becker & Knight 1990). However, additional studies in humans neither substantiated the positional "3′ bias" (Matsuda et al., 1993) nor have our studies in swine (Eguchi-Ogawa et al., 2010; Fig. 10). The most 3′ functional VH in swine (IGHV2) is almost never used while upstream VH15 (IGHV15) can account for ~13% of VH usage (Fig. 10). Thus, the "position hypothesis" to explain VH usage has not been universally fulfilled.

5.3 VH gene usage remains constant in fetal and young pigs

Vertical studies on VH usage in especially humans and mice are difficult because: (a) the V-D-J repertoire of these species is complex (Table 2) and could require up to 56 primer sets to recover all VDJ rearrangements in mouse and 42 sets for human (b) maternal regulatory factors transmitted *in utero* or via colostrum/milk can influence pre-natal and postnatal development (Wikler et al., 1980; Rodkey & Adler, 1983; Klobasa et al., 1981; Wang & Shlomchik 1998; Yamaguchi et al., 1983) and (c) control of environmental and maternal

effects is difficult or impossible in species with altricial offspring. Therefore we addressed this issue using a piglet model in which there is no transfer of maternal factors *in utero* and the influence of environmental factors postnatally on their precosial offspring can be controlled by the experimenter (Butler & Sinkora 2007; Butler et al., 2009b). Use of this model revealed that VH usage was constant during fetal life and that seven major genes accounted for >90% of the repertoire (Fig. 10; Butler et al. 2011b) while four can explain >80% of the repertoire Interestingly in piglets exposed to viral infection, gut colonization or nematode parasites after birth, ~ 75% of the mutated genes used were the same seven (Fig. 10; extreme right). Furthermore, proportional usage of these genes was similar to what was seen in the pre-immune repertoire, albeit somatically mutated. Some modest changes were observed such as an increase in VHY and decreases in VHA* and VHN. In other words, swine seldom select other genes from their repertoire after exposure to environmental antigen, but continue to use the same VH genes that comprise ~93% of the pre-immune repertoire.
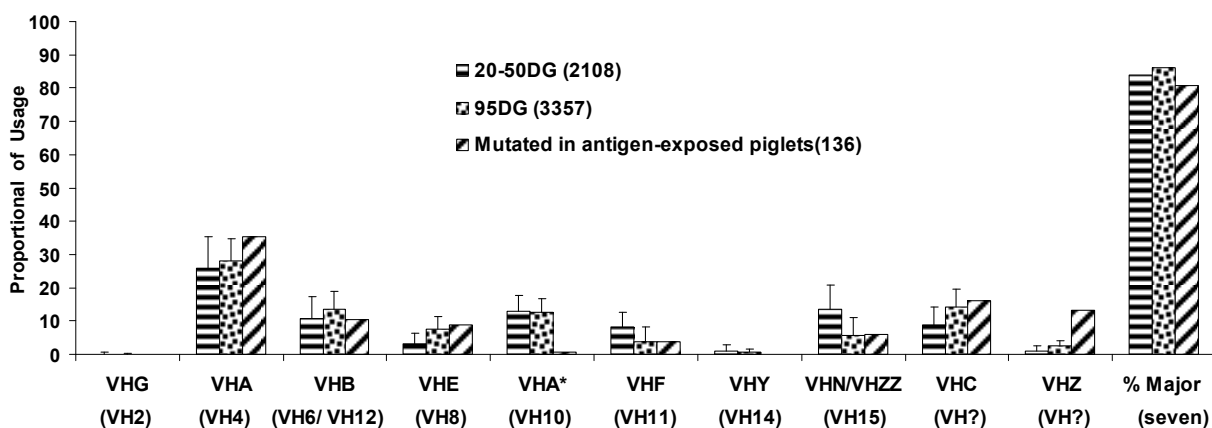
Fig. 10. VH gene usage in fetal piglets and among neonates that are antigen-exposed remains relatively constant. DG= days of gestation. The number of VH gene clones tested is given in the legend. The mutated VH genes are no longer recognized because they do not hybridize with VH gene-specific probes. Their identity must then be determined by sequencing. The bar graph on the extreme right gives the proportion of all mutated VH genes that are accounted for by the major seven genes used in the pre-immune repertoire by the fetus.

These observations should not be surprising considering that the specificity of binding site is heavily dependent on HCDR3 (Section 6.2). The HCDR3 repertoire in swine is diverse when only one VH, one DH and one JH segment are used (Fig. 8). Furthermore, the limited use of VH genes should not limit specificity, since their CDR regions can also be somatically mutated (Fig. 9B, C). As discussed in 4.1, the variants from a single VH (Fig. 8) can pair with different light chains, which can further reduce the need for large numbers of different VH genes for repertoire formation.

5.4 Why is the germline VH repertoire large but usage of the repertoire limited?

Use of the piglet model demonstrates that most of the piglet VH repertoire (Fig. 4 versus Fig. 10) is seldom used to form the antibody repertoire. While corresponding vertical studied are lacking in mice and human because of the logistic and experimental reasons

discussed in the last section, the bias usage of certain VH genes reported for these species (Section 5.1) suggests the outcome might be similar if such studies could be efficiently performed. The answer to the question may reside in understanding the evolution of the genes encoding antibody specificity that generated the vast array of Ig variable region genes first seen in more primitive vertebrates and the later evolution of somatic events that would appear to have made the original polygeny unnecessary.

6. The value of gene duplication to adaptive immunity

6.1 Adaptive versus innate immunity

The innate immune system is the "first responder" element of immune protection for higher vertebrates and may be the sole system for most invertebrate. Innate immunity works through phagocytic and epithelial cells that bear so-called "innate immune receptors", e.g. Toll-like (TLR) that recognize bacterial and viral entities that are not produced by eucaryotic cells of the host and are therefore considered foreign and dangerous. For vertebrates these include lipopolysaccharide, flagellin, bacterial DNA (non-methylated) and double-strained RNA. These are referred to as PAMPs (**P**athogen **A**ssociated **M**olecular **P**atterns). Products liberated from cells stimulated when their innate receptors recognize these PAMPs, then stimulate lymphocytes that leads to development of the adaptive immune system. This is demonstrated in studies using the isolator piglet model in which colonization or purified PAMPs are required for an adaptive immune response (Butler et al., 2002; 2005) and by infection with influenza virus which generates double-stranded RNA during replication (Butler, Lager, Vincent, Sun unpublished).

Unlike antibodies, the ligand binding sites of innate immune receptors do not change their specificity by any somatic process when they encounter a PAMP. The capacity for receptor modification after antigen encounter is the property of the adaptive immune system, as implied by the name. This adaptive capacity is illustrated in Figure 9 by showing there is a profound increase in Ig secretion and a shift in isotype usage (Fig. 9A) and an increase in SHM of the adaptive immune system antigen receptors after antigen exposure (Fig. 9B). As illustrated, these somatic hypermutations accumulate in those regions of the VH genes that encode the antibody binding site, i.e. the CDRs (Fig. 9C; Fig. 1A).

6.2 Are multiple V$_H$ genes required for host immune protection?

Since swine use a very small number of VH genes to generate a VDJ repertoire capable of protecting the host at all ages (Sun et al., 1998; Butler et al., 2006), we did a statistical itineration to suggest that >95% of the adaptive VDJ repertoire was the result of diversity within HCDR3 (Butler et al., 2000). HCDR3 is not encoded by any particular V-region gene segment but rather by the recombination of VH-DH-JH (Fig. 7 & 8). Joining of VH-DH-JH involves exonuclease removal of nucleotides from the gene segments involved as well as nucleotide additions using deoxynucleotide transferase. The types of variations generated are illustrated in Fig. 8 which shows different HCDR3 sequences found among the recombinants that use just one VH, DH and JH segment. The importance of the diversity generated in HCDR3 was empirically shown by Xu and Davis (2000) to be sufficient to allow most adaptive immune response using a transgenic mouse given only on VH gene but an intact DH and JH genome. Studies in rabbit also show that a single VH gene is primarily used in the establishment of the antibody repertoire (Knight 1992) but following antigen encounter this VH gene can be modified by SHM as well as by SGC

(Winstead et al., 1999; Schiaffella et al., 1999). We show that in swine seven VH genes can account for 93% of the entire pre-repertoire and that the two duplicated VHA genes (which have idetntical CDR1 and CDR regions; Fig. 4) can alone account for 30-50% (Fig. 10). Collectively these studies raise the question as to why the VH repertoire has been so heavily duplicated while so few of these duplicons are used. The answer may be found among bats, or at least in the insectivorous microbats. *Myotis lucifugus* has >250 VH3 family genes (Fig. 5) and probably 350 total VH genes, including all families (Bratsch et al., 2011). However, SHM occurs at the frequency seen in fetal piglets (Fig. 9B). Perhaps some older mammalian orders like the Chiroptera, rely more heavily on VH polygeny than somatic modification.

6.3 How does duplication/ diversification in the C-region effect protective immunity

The survival value of C-region polygeny can be appreciated because antibody isotypes encoded by the exons for IgM, IgG etc have distinct biological functions (Janeway et al., 2005). Additional duplication of Cγ and Cα genes has generated modified duplicons encoding Ig subclasses also with different important biological functions. In the case of IgG subclasses, these involve features like the ability to be recognized by different Fc receptors on phagocytic cells, transport across epithelial barriers, serum half-life differences involving FcRn and an association with antibodies of certain specificities. Duplication of Cγ genes in cattle (and other domestic ruminants) has lead to subclass IgG1 that is selectively transported by acinar epithelial cells of the mammary gland to provide essential antibodies for the survival of the newborn (Butler 1974; 1998). Neither IgG2 nor IgG3 function in this capacity. While there have been no functional studies on the many IgG subclass antibodies in horse or swine, it is possible that each of these subclass antibodies are best suited for particular activities in the same manner as described for human IgG subclasses.

The duplicated human Cα genes also differ in a number of features. Most striking is the susceptibility of IgA1 to IgA proteases produced by many common Gram positive bacteria while IgA2 is resistance due to the lack of the 13 amino acid hinge which is the substrate for these proteases in IgA1 (Plaut et al., 1974). Differential tissue expression of rabbit IgA subclasses also suggests a division of labor (Spieker-Polet et al., 1993). IgA in swine lack the long hinge of human IgA1 and is therefore not susceptible to the classical bacterial IgA1 proteases although a protease from *H. suis* can cleave the porcine α-chain (M. Mullins, K. Register, D.O. Bayles, J.E. Butler, unpub).

The "experiment of nature" is whether mammals with a deficiency in their Cγ sublocus are immunologically impaired. There are no known ruminates that like lack ruminat IgG1, so there are no data; perhaps such a deficiency would be a developmental lethal. The mammals best-studied for Cγ subclass deficiencies are humans. For example, humans lacking IgG2 have a deficiency in their response to bacterial polysaccharide antigens (Hammarstrom et al., 1986). Additional deletions of Cγ genes have been described, including individuals lacking Cγ1, Cγ2, Cγ4 and Cα1 (the 5'duplicon shown in Figure 2B). However, such individuals remain healthy and asymptomatic (Lefranc et al., 1983b; Notoaramgelo et al., 2009). Selective IgG1 deficicieny which accounts for the major portion of serum IgG, is not correlated with lower serum igG levels (Olsson et al., 1993) while some appear to be (Rabbani et al., 1995). In swine, the IgA[b] allotypic variant lacks a major portion of its hinge (Brown et al., 1995), yet this "defect" has not been correlated with any risk of disease; (Navarro et al., 2000).

6.4 Polygeny is widespread in other loci important to immunity

There are various examples of polygeny in the immune system but in these loci SHM is not important in repertoire formation. The loci encoding the various T cell receptors are similar to those encoding the Igs. Recombination of gene segments occurs in the same manner as for Ig genes, i.e. recombination requires RAGs, DNA excision, and splicing and DNA repair enzymes(see Fig. 7). However there is no convincing evidence for SHM after segment recombination so polygeny in the TCR is theoretically more important in repertoire generation than in the Ig loci where SHM can further diversify the repertoire.

The genes encoding the major histocompatibility genes (MHC) are partially encoded by IGSF genes and determine: (a) "individualism" and self recognition as demonstrated in tissue typing and (b) recognition of peptides generated and presented by antigen-presenting cells. The MHC gene system is highly conserved among mammals with three genes encoding MHC I and 6-8 encoding MHC II. The enormous diversity of MHC is not achieved by polygeny or somatic processes but rather by an enormous degree of polymorphism of the MHC genes within the population (Janeway et al., 2005). There can be as many as 300 allelic variants of any one MHC gene.

The selective advantage of polygeny for genes encoding the MHC, innate immune receptors, the CH and Cγ subclasses and even the TCR is easy to appreciate. However, such polygeny in loci encoding the variable Ig genes in higher vertebrates is more difficult to explain if these species can generate an effective antibody response using a single VH gene (Section 6.2). This creates an enigma for VH and VL polygeny that we shall attempt to explain from an evolutionary perspective.

7. Conclusions: Ig polygeny and redundancy in higher vertebrates is an evolutionary vestige

7.1 The case of mammalian IgD

Vestiges of genes are not unusual and mammalian IgD is an example. IgD was discovered as a myeloma protein nearly 50 years ago and its function has remained an enigma since that time (Rowe & Fahey, 1965). The considerable research funding invested to determine the function of IgD has largely generated only hypotheses. IgD is the least homologous isotype among mammals, e.g. <40%,(in part the reason it was overlooked in some mammals; Butler et al., 1996) whereas most other isotypes share 70-90% homology (Butler, 2006). IgD is even missing from the genome of some mammals (Table 2; Fig. 2B) and perhaps all birds. In mice and humans, IgD and IgM occur as dual B cell receptors but IgD-deficient mice have normal immune responses (Nitschke et al., 1993) although IgD can compensate for the loss of functional IgM (Lutz et al., 1998). While numerous studies have attempted to define a unique role for mammalian IgD, most of these have not been very convincing (Monroe et al., 1983; Liu et al., 1996; Roes et al., 1993; Vitetta et al., 1977).

Comparative immunologists have put the role of IgD into perspective beginning with the observation in catfish of a seven domain Ig with distant homology to mammalian IgD (Wilson et al., 1997; Bengten et al., 2002). This was followed by the discovery of a similar multi-domain IgD in *Xenopus* (Zhao et al., 2006) and in other teleosts (Hordvik et al., 1999; Srisapoome et al., 2004; Stenvik & Jorgensen, 2000). IgD has subsequently been found in the genome of many other lower vertebrates and in protherian mammals (reviewed by Edholm et al., 2010). Collectively these studies would morph into the realization that IgD and IgM are the primordial vertebrates Ig isotypes (Ohta & Flajnik, 2006; Bengten et al., 2006).

Interesting among the studies in catfish and swine, is that IgD can be produced as a chimeric Ig using the Cμ1 domain of IgM and the various domain exons of Cδ (Zhao et al., 2002; 2003). Thus, the expression of IgD in mice and humans by RNA splicing rather than classical CSR (Maki et al., 1981) has a primitive history.

Compared to the IgD of fish and *Xenopus*, the exons encoding for mammalian IgD appear to be relics. Depending on the mammal selected, the number of hinge and domain exons is highly variable (Butler et al., 2010). While switch regions are typically >3kb in length, mammalian Sδ (when present) exists as a short remnant of <0.5 kb (Zhao et al., 2003). Nevertheless it appears to function in some cases in CSR in humans and swine (Zhao et al., 2002; 2003; Arpin et al., 1998; Koelsch et al., 2007; Zheng et al., 2004).

Recent findings show that basophils have abundant membrane IgD but not T cells, NK cells, dendritic cells or monocytes (Chen et al., 2009; Dawichi & Marshall, 2007) although this observation was not supported by the recent report by Karasuyami et al., (2009). Assuming the case for adventious IgD on basophils is true, it agrees with studies in catfish showing surface IgD on granulocytes (Edholm et al., 2010). These observations would support a unique role of this ancestral Ig. In spite of the observation of IgD on human basophils, the "experiment of nature" that IgD deficiency does not impair mammalian adaptive immunity (Nitschke et al., 1993) has lead us to conclude that IgD is an evolutionary relic for mammals but persists because of its redundant value including its role as a BCR and its presence as an adventious Ig on basophils (Lutz et al., 1998; Chen et al., 2009). Thus, the relic remains because there has been no negative selection to completely remove IgD from most mammalian genomes.

7.2 CH duplication/diversification provides antibodies with specialized effector function but also redundancy

Isotype diversity in sharks and bony fish is limited to IgM and IgD (IgW in sharks). In tetrapods the CH repertoire diversified; 3 in birds and reptiles, 5-6 in amphibians and 5 in mammals (Table 2). In mammals this includes subclasses of IgG and IgA. Each major isotypes in mammals, perhaps with the exception of IgD, has some specialized effector function (Janeway et al., 2005). Perhaps the terrestrial environment offered a new challenge to survival and with the addition of homeothermia, the need for a more specialized adaptive immune system that could respond more quickly and lead to the evolution of GCs (Sections 3.2; 6.3). As discussed in Section 3-2, there is also evidence that the polygeny of Cγ genes resulted from a combination of gene duplication and genomic gene conversion. This duplication event was restricted to mammals that appeared in the "last minute before 12PM" on the evolutionary clock, appearing after mammalian speciation (Butler et al., 2009a). The subclass duplication/diversification in mammals resembles the pattern that produced V region polygeny. In sharks and bony fishes, isotype diversity is limited to IgM and IgD, while in higher vertebrates duplication/ diversification extends downstream into the CH sublocus, and in mammals, especially to the Cγ sublocus. However, IgG subclass deficiencies are not lethal defects, suggesting that even in late-evolving sites of Ig gene duplication, such duplication seem unnecessary.

7.3 Somatic mechanisms render much of the polygeny in VH, Vλ and Vκ loci to relics.

As we show in the piglet model, very few VH genes are needed and after antigen encounter, further repertoire diversification is by SHM of the same genes (Fig. 10). Furthermore, the

fact that >90% of the repertoire is generated by junctional diversity in HCDR3 (Section 6.2; Fig. 8) and that a transgenic mouse with only one VH is fully immune competent, strengthens the argument. Studies in rabbit also support this view in which one VH gene account for 90% of the early repertoire which can be diversified SHM and SGC after antigen encounter (Winstead et al., 1999; Schiaffella et al., 1999). Detractors from this view may argue that while few (or only one) VH gene are needed, polygeny in the DH and J_H regions is still necessary (Table 1). Again studies in swine counter this argument since they have only two functional DH segments and one J_H (Butler et al., 1996; Eguchi-Ogawa et al., 2010). In addition, the chicken has only a single JH, DH segments that are nearly identical and only one functional VH and one function Vλ gene (Fig.3). This species uses the "relics" of upstream pseudogenes for use in repertoire diversification by SGC (Reynaud et al., 1987; Ratcliffe 2006), a mechanism also available to rabbits (Becker & Knight, 1990). Thus we propose that the evolution of gene segment recombination and SHM rendered polygeny in the VH and VL loci of higher vertebrates unnecessary. We believe this polygeny is derived from the tandem array of V-D-J-C clusters in sharks that do not require somatic recombination (Fig. 3). Sharks lack germinal centers as do bony fished and amphibians which makes SHM a less efficient process (DuPasquir et al., 2000). While SHM had been described in these vertebrate classes, we cannot cite head-to-head studies on the frequency of SHM like that presented in Fig. 9B.

We believes that like IgD, the VH, Vλ and Vκ polygeny of the most evolved mammals and birds, remain in the genome largely as relics and because of the lack of negative selection. Early vertebrates have as many as four light chain loci but with evolution the number diminished and birds have lost all but lambda. That different Ig isotypes and certain IgG subclass have specific biological functions suggests a selective advantage for polygeny which we believe was true at the dawn of VH and VL duplicatation that lead to the polygeny in V-region loci. The alternative explanation for especially VH and VL polygeny is that this polygeny occurs in a "hot spot" of RAG-dependent gene segment recombination. For such recombination events to occur, opening of the chromatin structure to provide access to nuclear enzymes, is considered necessary. Perhaps such exposure to repeated recombination activity explains the instability of the locus (Lefranc et al 1983a; Matsuda et al 1990) which renders genes in the locus vulnerable to the molecular machinery involved in gene duplication and genomic gene conversion.

Mainstream immunology has invested almost entirely in studies of mouse and human immune systems and therefore seems to have missed the evolutionary significance of Ig polygeny. To avoid a similar criticism our analysis of polygeny reviewed the major elements of the Ig genes of all vertebrates that have been seriously studied and the mechanisms they use to generate their antibody repertoire. From these comparisons we have offered a hypothesis to explain the polygeny of the major Ig loci in mammals and the reason why at the "highest levels" of vertebrate development, this polygeny appears to be an evolutionary vestige.

8. Conclusions

Gene duplication is a common feature of eucaryotic genomes although the degree varies among gene families and loci. It is estimated that ~5% of the human genome is comprised of duplicated genes (Lewin, 2004). Among these are genes of the immunoglobulin supergene family (IGSF). This polygeny is widespread in loci important to the immune system as well as

encoding proteins with only indirect roles in immunity. The β-barrel encoded by genes of the IGSF has obviously been a successful structural motif which can explain its conservation during evolution and diversification into many variants. While IGSF polygeny is widespread, the degree of polygeny is especially pronounced among those that encode the variable heavy and light chain genes of antibodies and the T cell receptor (TCR). Early estimates suggested there were as many as 1000 variable heavy (VH) genes in mice and hundreds in humans. While subsequently studies, including genome projects, have lowered the number of VH genes to ~100-150 in these species, this is still a very large number of homologous genes to occupy a single locus. Similar duplication is seen among genes encoding the variable light chain genes, i.e. Vλ and Vκ. However, there are large variations in the numbers and features of duplicated VH and VL genes among mammals and other vertebrates.

This article surveys the duplicated Ig genes in a number of species and uses examples indicating that Ig polygeny resulted from a combination of duplication and genomic gene conversion. Since understanding the evolutionary forces at work in this process requires some understanding of the role played by these duplicated genes in humoral immunity, we review the processes involved in the generation of the antibody repertoire such as Ig gene segment recombination, junctional diversity, somatic hypermutation (SHM) and somatic gene conversion (SGC). We review these processes in various vertebrates but focus on data obtained using the neonatal and newborn piglet model to suggest that evolutionary improvements in somatic processes have reduced the need for the Ig polygeny that evolved among lower vertebrates. We also describe the more recent duplication of the Cγ genes of mammals that indicates the process was similar. Cγ genes encode the subclasses of mammalian IgG, the "flagship mammalian antibody" that is unique to this vertebrate class. Since this duplication event occurred more recently, we thought it could provide insight into the advantages conferred by gene duplication.

We propose that the extensive polygeny of VH, Vλ and Vκ genes among vertebrates gave adaptive advantage to the earliest vertebrates for generating a diverse repertoire of antibody specificities much as the more recent evolutionary diversification of Cγ genes resulted in IgG subclass antibodies with different effector functions. We suggest that the evolutionary appearance of mechanisms to somatically alter V-region genes reduced the importance of polygeny in V-region loci for certain mammals and birds. In higher mammals these mechanisms make it possible for a complete functional repertoire to be generated using just one or a few VH genes. This hypothesis can explain why so many of the V-region genes of higher mammals are seldom used, and why deletions of VH genes and Cγ genes have no effect. We propose that these genes remain as evolutionary vestiges or redundant back-ups in the genome in a manner that parallels the retention of IgD in most mammals. An alternative hypothesis is that the extensive somatic recombination which characterizes the variable region loci (Section 4) creates instability that promotes duplication and genomic gene conversion. In any case, these hypotheses challenge the existing paradigm that random VH, DH and JH recombination among the many gene segments is necessary for survival (presented in immunology textbooks) by placing Ig polygeny into evolutionary perspective.

9. Acknowledgement

10. References

Amit., A.G., R. A. Mariuzza, S. E. Phillips, & R. J. Pliak. (1986). Three –dimensional structure of an antigen-antibody complex at 2.8 A resolution. *Science* 233 pp. 747-753.

Arakawa, H., Furusawa S., Ekino, S., & Yamaguchi, H. (1996). Immunoglobulin gene hyperconversion ongoing in chicken splenic germinal centers. *EMBO J.* 15 pp. 2540-2546.

Arpin, C., de Bouteillier, O., Razanajaona, D., Briere, F., Banchereau, J., Lecque, S. & Liu, Y. J. (1997). Human peripheral B- cell development. sIgM-IgD + CD38+ hypermutatied germinal center centroblasts preferentially express lambda light chains and have undergone mu-to-delta switch. *Ann. N.Y. Acad. Sci.* 815 pp. 193-198.

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

Bengten, E., Quiniou, S., Hikima, J., Waldbleser, G., Warr, G.W., Miller, N.W. & Wilson, M. (2006). Structure of the catfish IGH locus: analysis of the region including the single functional IGHM gene. *Immunogenetics* 58 pp. 831-844.

Bengten, E., S.M. Quiniou, T.B. Stuge, T. Katagirai, N.W. Miller, L.W. Clem, G.W. Warr and M. Wilson. (2002). The IgH locus of the channel cafish, *Ictalurus punctatus*,contains multiple constant region gene sequences: different genes encode heavy chains of membrane and secreted IgD, *J. Immunol.* 169: 2488-2497.

Blumbach, B., Diehl-Seifer, B., Seack, J., Steffen,R., Muller, I.M. & Muller, W.E.G. (1999). Cloning and expression of new receptors belonging to the immunoglobulin superfamily from the marine sponge *Geodia cydonium. Immunogenetics* 49 pp. 751-763.

Boffey, J., Nicholl, D., Wagner, E.R., Townson, K., Goodyear, C., Furukawa, K., Furukawa, K., Conner, J. & Willison, H.J. (2004). Innate murine B cells produce anti-disialosyl antibodies reactive with Camylobacter jejuni LPS and gangliosides that are polyreactive and encoded by a restricted set of unmutated V genes. *J. Neuroimmunol.* 152 pp. 98-111.

Bratsch, S., Wertz, N., Chaloner, K., Kunz, T. H. & Butler, J. E. (2011). The little brown bat displays a highly diverse V_H, D_H and J_H repertoire but little evidence of somatic hypermutation. *Develop. Comp. Immunol.* 35 pp. 421-430.

Brezinschek, H-P, Foster, S. J., Brezinschek, R. I., Doerner, T., Domiati-Saad, R. & Lipsky P. E. (1997). Analysis of the human VH gene repertoire. *J. Clin. Invest.* 99 pp. 2488-2501.

Brown, W. R., Kacskovics, I., Amendt, B. Shinde, R., Blackmore, N., Rothschild, M. & Butler, J.E. (1995). The hinge deletion variant of porcine IgA results from a mutation at the splice acceptor site in the first Cα intron. *J. Immunol.* 154 pp. 3836-3842.

Butler, J. E. (1998). Immunoglobulins and immune cells in animal milks. In: *Mucosal Immunology,* Ogra, P. L., Mestecky, J., Lamm, M. E., Strober, W., McGhee, J. R.& Bienenstock, J. (Eds.), , Chapter 98. Academic Press, New York pp. 1531-1554.

Butler, J.E., Santiago-K. Mateo, K., Sun, X-Z, Wertz, N., Sinkora, M., Harvey, R. & Francis, D.L. (2011a). Antibody repertoire development in fetal and neonatal piglets. XX. The ileal Peyers patches are not a site of B cell lymphogenesis and are not required for systemic B cell proliferation and Ig synthesis.

Butler, J. E., (1974). Immunoglobulins of the mammary secretions. In: *Lactation, a Comprehensive Treatise*, (Larson, B. L. & Smith, V., eds.), Vol. III, Chapter V, pp. 217-255. Academic Press, New York.

Butler, J. E., & Kehrle Jr., M. E. (2005). Immunocytes and immunoglobulins in milk. In: *Mucosal Immunology*, (J. Mestecky, M. E. Lamm, W. Strober, J. R. McGhee, L. Mayer and J. Bienenstock, eds.), 3rd Edition, Academic Press, NY. pp. 1763-1793.

Butler, J. E., Francis, D., Freeling, J., Weber, P., Sun, J. & Krieg, A. M. (2005). Antibody repertoire development in fetal and neonatal piglets. IX. Three PAMPs act synergistically to allow germfree piglets to respond to TI-2 and TD antigens. *J. Immunol.* 175 pp. 6772-6785.

Butler, J. E., Sun, J. & Navarro, P. (1996). The swine immunoglobulin heavy chain locus has a single J$_\text{H}$ and no identifiable IgD. *International Immunology* 8 pp. 1897-1904.

Butler, J. E., Weber, P., Sinkora, M., Baker, D., Schoenherr, A., Mayer, B. & Francis, D. (2002). Antibody repertoire development in fetal and neonatal piglets. VIII. Colonization is required for newborn piglets to make serum antibodies to T-dependent and type 2 T-independent antigens. *J. Immunol.* 169 pp. 6822-6830.

Butler, J. E., Weber, P., Sinkora, M., Sun, J., Ford, S. J. & Christenson, R. (2000). Antibody repertoire development in fetal and neonatal piglets. II. Characterization of heavy chain CDR3 diversity in the developing fetus. *J. Immunol.* 165 pp. 6999-7011.

Butler, J. E. & Sinkora, M. (2007). The isolator piglet: A model for studying the development of adaptive immunity. *Immunol. Res.* 39 pp. 33-51.

Butler, J. E., Sun, X-Z , Wertz, N., Lager, K. M., Urban Jr., J. , Nara, P. & Tobin, G. (2011b). Antibody repertoire development in fetal and neonatal piglets. XXI. VH usage remains constant during development in fetal piglets and postnatally in pigs exposed to environmental antigen.

Butler, J. E., Zhao, Y., Sinkora, M., Wertz, N. & Kacskovics, I. (2009a). Immunoglobulins, antibody repertoire and B cell development. *Devel. Comp. Immunol.* 33 pp. 321-333

Butler, J. E., Lager, K. M., Splichal, I., Francis, D., Kacskovics, I., Sinkora, M., Wertz, N., Sun, J., Zhao, Y., Brown, W. R., DeWald, R., Dierks, S., Muyldermanns, S., Lunney, J.K., McCray, P. B., Rogers, C. S., Welsh, M. J., Navarro, P., Klobasa, F., Habe, F. & Ramsoondar, J. (2009b). The Piglet as a Model for B cell and Immune System Development. *Vet. Immunol. Immunopath.* 128 pp. 147-170.

Butler, J. E., Wertz, N., Zhao, Y., Kunz, T. H., Bratsch, S., Whitaker, J. & Schountz, T. (2010). Two suborders of bats have the canonical isotypes repertoire of other eutherian mammals. *Devel. Com. Immunol.* 35 pp. 272-284.

Butler, J. E., Weber, P. & Wertz, N. (2006). Antibody repertoire development in fetal and neonatal pigs. XIII. "Hybrid VH genes" and the pre-immune repertoire revisited. *J. Immunol.* 177 pp. 5459-5470.

Cerato, E., Birkle, S., Portoukalian, J., Mezazigh, A., Chatal, J.F. & Aubry, J. (1997). Variable region gene segments of nine monoclonal antibodies specific to disialgangliosides (GD2, GD3) and their O-acetylated derivatives. *Hybridoma* 16 pp. 307-316.

Chen, K., Xu, W., Wilson, M., He, B., Miller, N. W., Begnten, E., Edholm, E-S, Santini, P.A., Rath, R., Chiu, A., Cattalinei, M., Litzman, J., Busseel, J., Huang, B., Meini, A., Riesbaeck, K., Cunningham-Rudles, C., Plebani, A. & Cerutti, A. (2009). Immunoglobulin D enhances immune surveillance by activating antimicrobioal, proinflammatory and B cell-stimulating programs in basophils. *Nat. Immunol.* 10 pp. 889-898.

Currier, S. J., Gallara, J. L. & Knight K. L. (1998). Partial genetic map of the rabbit VH chromosomal region. *J. Immunol.* 140 pp. 1651-1659.

Davies, J. & Riechmannn, L. (1995). Antibody VH domains as small recognitions units. *Biotechnol.* 13 pp. 475-479.

Dawichki, W. & Marshall, J.S. (2007). New and emerging roles for mast cells in host defense. *Curr. Opin. Immunol.* 19 pp. 31-18.

De Genst, E., Saerens,D., Muyldermans, S. & Conrath, K. (2006). Antibody repertoire development in camelids. *Dev. Comp. Immunol.* 30 pp. 187-198.

deWildt, R. M., Hoet, R. M., van Venrooig, W. J., Tomlinson, I. M. & Winter, G. (1999). Analysis of heavy and light chain pairing indicates that receptor editing shapes the human antibody repertoire. *J. Mol. Bio.* 285 pp. 895-901.

Diaz, M., Stanfield, R.L., Greenberg, S. & Flajnik, M. F. (2002). Structural analysis, selection and ontogeny of the shark new antigen receptor (IgNAR): identification of a new locus preferentially expressed in early development. *Immunogenetics* 54: 501-512.

Dildrop, R, Krawinkel, U., Winter, E. & Rajewsky, K. (1985). VH gene expression in murine lipopolysaccharide blasts distribute over the nine known VH-groups and may be random. *Eur. J. Immunology* 15 pp. 1154-1156.

Dooley, H., Flajnik, M.F. & Porter, J. (2003). Selection and characterization of naturally–occurring single domain (IgNAR) antibody fragments from immunized sharks by phage display. *Mol. Immunol.* 40 pp. 25-33.

Dooley, H. & Flajnik, M.F. (2006). Antibody repertoire development in cartilagenous fish. *Devel. Comp. Immunol.* 30 pp. 43-56.

Du Pasquier, L., Robert, J., Courtet, M., & Mussmann, R. (2000). B cell development in the amphibian Xenopus. *Immunol. Rev.* 175 pp. 201-213.

Edholm, E-S, Bengton, E., Staffor, J. L., Sahoo, M., Taylor, E. R. , Miller, N. W. & Wilson, M. (2010). Identification of two IgD+ B cell populations in channel catfish, Ictalurus punctatus. *J. Immunol.* 185 pp. 4082-4094.

Eguchi-Ogawa, T, Sun, X-Z., Wertz, N., Uenishi, H., Puimi, F., Chardon, P., Wells, K., Tobin, G. J. & Butler, J. E. (2010). Antibody repertoire development in fetal and neonatal piglets. XI. The relationship of VDJ usage and the genomic organization of the variable heavy chain locus. *J. Immunol.* 184 pp. 3734-3742.

Flanagan, J.G. & Rabbitts, T. H. (1982) Arrangement of human immunoglobulin heavy chain constant region genes implies evolutionary duplication of a segment containing γ, ε, γ and α genes. *Nature* 300 pp. 709-713.

Foster, S. J., Brezinschek, H. P., Brezinschek, R. I. & Lipsky, P. E. (1997). Molecular mechansism and selective influences that shape the kappa gene repertoire of IgM+ B cells. *J. Clin. Invest.* 99 pp. 1614-1627.

Gay, D., Saunders, T., Camper, S. & Weigert, M. (1993). Receptor editing: an approach by autoreactive B cells to escape tolerance. *J. Exp. Med.* 177 pp. 999-1008.

Glas, A. M., van Monfort, E. H. N. & Milner, E. C. B. (2000). The human antibody repertoire: Old notions, current realities and VH gene-dependent biases. In: *The antibodies*, (Zanetti, M. &Capra, J.D., eds). Vol 6 , pp 63-79. Harwood Academic Publishers , Amsterdam.

Gu, H., Tarlinton, D., Muller, W., Rajewsky, K. & Forster, I. (1991). Most peripheral B cells in mice are ligand restricted. *J. Exp. Med.* 173 pp. 1357-1371

Hamers-Casterman, C., Atarhouch, T., Muylermans, S., Robinson, G., Hamers, C., Songa, E. B., Bendahman, N. & Hammers, R. (1993). Naturally occurring antibodies devoid of light chains. *Nature* 363 pp. 446-448.

Hammarstrom L., Lefranc G., Lefranc M-P, Persson, M.A.A. & Smith, C.I.E. (1986). *Monogr. Allergy* 20 pp. 18-25.

Herrin, B. R. & Cooper, M. D. (2010). Alternative adaptive immunity in jawless vertebrates. *J. Immunol.* 185 pp. 1367-1374.

Hordvik, I., Thevarajan, J., Sandal, I., Bastani, N. & Krossoy, B. (1999). Molecular cloning and phylogenetics analysis of the Atlantic salmon immunoglobulin D gene. *Scand. J. Immunol.* 50 pp. 202-210.

Ichiyoshi, Y. & Casali, P. (1994). Analysis of the structural correlates for antibody polyreactivity by multiple reassortments of chimeric human immunoglobulin heavy and light chain V segments *J. Exp. Med.* 180 pp. 885-895.

Janeway, C.A., Travers, P., Walport, M., & Shlomchik, M.J. (2005). Immunobiology, 6th Edition, Garland Science, New York.

Janssens, R., Dekker, S., Hendriks, R. W., Panayotou, G., van Remoortere, A., San, J. K., Groveld, F. & Drabek, D. (2006). Generation of heavy-chain-only antibodies in mice. *Proc. Nat'l. Acad. Sc*i. USA 103 pp. 15130-15135.

Johnston, C. M., Wood, A. L., Bolland, D. J. & Corcoran, A. E. (2006). Complete sequence assembly and characterization of the C57BL/6 mouse heavy chain V region. *J. Immunol.* 176 pp. 4221-4234.

Karasuyama, H., Mukai, K., Tsujimura, Y., & Obata, K. (2009). Newly discovered roles for basophils : a neglected minority gains new respect. *Nature Rev. Immunol.* 9 pp. 9-13.

Keyeux, G, Lefranc, G. & Lefranc, M. P. (1989). A multigene deletion in the human IGH constant region locus involves highly homologous hot spots of recombination. *Genomics* 5 pp. 432-441.

Klobasa, F., Werhahn, E., & Butler, J.E. (1981). Regulation of humoral immunity in the piglet by immunoglobulins of maternal origin. *Res. Vet. Sci.* 31 pp. 195-206.

Knight, K. L. (1992). Restricted VH usage and generation of antibody diversity in rabbit. *Ann. Rev. Immunol.* 10 pp. 593-616.

Koelsch, K., Zheng, N. Y., Zhang, Q., Duty, A., Helms, C., Mathias, M. D., Jared, M., Smith, K., Capra, J. D., & Wilson, P. C. (2007). Mature B cells class switch to IgD are autoreactive in healthy individuals. *J. Clin Invest.* 117 pp. 1558-1565.

Kranz, D. M. & Voss, Jr., E. W. (1981). Restricted reassociation of heavy and light chains from hapten-specific monoclonal antibodies. *Proc. Nat'l Acad. Sci.* 78 pp. 5807-5811.

Lavoie, T. B., Mohan, S., Lipschults, C. A., Grivel, J-C., Li, Y. l., Mainhart, C. R., Kam-Morgan, L. N. W., Drohan, W. N. & Smith-Gill, S. J. (1999). Structural differences among monoclonal antibodies with distinct fine specificities and kinetic properties. *Mol. Immunol.* 36 pp. 1189-1205.

Lefranc, M. P., Lefranc, G., de Lange, G., Out ,T. A., van den Broek, P. J., van Nieuwkoop, J., Radl, J., Hela, A. N., Chaabani, H., van Loghem, E. & Rabbitts, T. H. (1983a) Instability of the human immunoglobulin heavy chain constant region locus indicated by different inherited chromosomal deletions. *Mol. Biol Med.* 1 pp. 207-217.

Lefranc, G., Chaabani, H., van Loghem, E., Lefranc, M. P., de Lange, G. & Helal, A. N. (1983b). Simultaneous absence of the human IgG1, IgG2, IgG4 and IgA1

subclasses: immunological and immunogenetical considerations. *Eur. J. Immunol* 13 pp. 240-244.

Lefranc, M. P., Lefranc, G. & Rabbitts, T. H. (1982). Inherited deletion of immunoglobulin heavy chain constant region genes in normal individuals. *Nature* 300 pp. 760-762.

Lewin, B., (2004). *Genes VIII*. Pearson-Prentice Hall, Upper Saddle River, N.J. p. 87.

Liu, Y.J., de Bouteiller, O., Arpin, C., Briere, F., Gailbert, L., Ho, S. , Martinez-Valdez, H., Banchereau, J. & Lebecque, S. (1990). Normal IgD+IgM- germinal center B cells can express up to 80 mutations in the variable region of their IgD transcripts. *Immunity* 4 pp. 603-613.

Lutz, C., Ledermann, B., Kosco-Vilbois, M. H., Ochsenbein, A. F., Zingernagel, R. M., Kohler, G. & Brombacher, F. (1998). IgD can largely substitute for loss of IgM function in B cells. *Nature* 393 pp. 797-801.

Mageed, R. A., Marmer, I. J. , Wynn, S. L., Moyes, S. P, Maziak, B. B., Bruggemann, M. & MacKworth-Young, C. G. (2001). Rearrangement of the human heavy chain variable region gene V3-23 in transgenic mice generates antibodies reactive with a range of antigens on the basis of VHCDR3 and residues intrinsic to the heavy chain variable region. *Clin. Exp. Immunol.* 123 pp. 1-8.

Maki, R., Roeder W., Trawnecker, A., Sidman, C., Wabl, M., Raschke, W. & Tonegawa, S. (1981). The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta. *Cell.* 24 pp. 353-365

Malynn, B. A., Berman, J. E., Yancopous, G. D., Bona. C. A. & Alt, F. W. (1987). Expression of the immunoglobulin heavy-chain variable gene repertoire. *Curr. Top. Microbiol. Immunol.* pp. 135: 75-94.

Marchalonis, J. J., Schluter, S. F., Bernstein, R. M. & Edmundson, A. B. (1998). Phylogenetic emergence and molecular evolution of the immunoglobulin family. *Adv. Immunol.* 70 pp. 417-506.

Marr, S., Morales, H., Bottaro, A., Cooper, M., Flajnik, M. & Robert, J. (2007). Localization and differential expression of activation-induced cytidine deaminase in the amphibian Xenopus upon antigen stimulation and during early development. *J. Immunol.* 179 pp. 6783-6789.

Matsuda, F., Shin, E. K., Nagaoka, H., Matsumura, R., Haino, M., Fukita, Y., Taka-ishi, S., Imai, T., Riley, J. H., Amand, R., Soeda, E. & Honjo, T. (1993). Structure and physical map of 64 variable segments in 3' 0.8-megabase region of the human immunoglobulin heavy-chain locus. *Nature Genet.* 3 pp. 88-94.

Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, K., Hayashida, H., Miyata, T. & Honjo, T. (1998). The complete nucleotide sequence of the human immunoglobulin heavy chain region locus *J. Expt. Med.* 188 pp. 2151-2161.

Matsuda, F., Sin, 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: pp. 2501-2506.

Meselson, M. S. & Radding, C. M. (1975). A general model for genetic recombination. *Proc. Nat'l Acad. Sci.* USA 72 pp. 358-361.

Migone, N., Oliviero, S., de Lange, G., Delacroix, D. L., Boschis, D., Altruda, F., Silengo, L., Demarchi, M., & Carbonaro, A. O. (1984). Multiple gene deletions within the human heavy-chain cluste*r. Proc. Nat'l Acad. Sci.* USA 81 pp. 5811-5815.

Miller, M. A., & Steele, R.E. (2000). Lemon encodes an unusual receptor protein-tyrosine kinase expressed during gametogenesis in Hydra. *Dev. Biol.* 224:286-298.

Mo, J.A., & Holmdahl, R. (1996). The B cell response to autologous type II collagen: biased V gene repertoire with V gene sharing and epitope shift. *J. Immunol.* 157 pp. 2440-2448.

Monroe, J. G., Havran, W. L., & Cambier, J. C. (1983). B lymphocyte activation entry into cell cycle is accompanied by decrease expression of IgD but not IgM. *Eur. J. Immunol.* 13 pp. 208-213.

Muyldermans, S., Ghassabeh, G.H., & Saerens D. (2009). Single-domain antibodies. In : *Recombinant antibodies for immunotherapy*, Little, M. (ed.), Cambridge University Press. Chapter 16, pp. 216-230.

Navarro, P., Christenson, R. Ekhardt, G, Lunney, J.K., Rothschild,M., Bosworth, B, Lemke, J. & Butler, J.E. (2000). Genetic differences in the frequency of the hinge variants of porcine IgA is breed dependent. *Vet. Immunol. Immunopath.* 73 pp. 287-295.

Nguyen, V., Su, C., Muyldermans, S. van der Loo, W. (2002). Heavy chain antibodies in Camelidea; a case of evolutionary innovation. *Immunogenetics* 54 pp. 39-47.

Nitschke, L., Kosco, M. L., Kohler, G., & Lamers, M. C. (1993). Immunoglobulin D deficient mice can mount normal immune responses to thymus –independent and – dependent antigens. *Proc. Nat'l Acad Sci.* USA 90 pp.1887-1891.

Notarangelo, L. D., Fischer, A.,Geha, R.S., Casanova, J-L., Chapel, H, Conley, M.E., Cunningham-Rundles, C., Etzioni, A., Hammarstrom, L., Nonoyama, S., Ochs, H.D., Puck, J., Roifman, C., Seger, R., & Wedgwood, J. (2009). Primary immunodeficiencies: 2009 update. *J. Allergy Clin. Immunol.* 124: 1161-1178.

Ogawa, K., Wakayama, A., Kunsisada, T., Oril, H., Watanabe, K. & Agata, K. (1998). Identification of a receptor tyrosine kinase involved in germ cell differentiation in planarians. *Biochem. Biophys Res. Commun.* 248 pp. 204-209.

Ohta, Y. & Flajnik, M. (2006). IgD, like IgM , is a primordial immunoglobulin class perpetuated in most jawed vertebrates. *Proc. Nat'l. Acad. Sci.* USA 103 pp. 10723-10728.

Olsson, P.G., Rabbani, H., Hammarstromn, L. & Smith, C.I.E. (1993). Novel human immunoglobulin heavy chain constant region gene deletion haplotypes characterized by pulsed -field electrophoresis. *Clin. Exp. Immunol.* 94: pp. 84-90.

Padlan, E. A. (1994). Anatomy of the antibody molecule. *Mol. Immunol.* 31 pp. 169-217.

Plaut, A.G., Wustar, R. Jr. & Capra, J.D. (1974). Differential susceptibility of human IgA immunoglobulins to streptococal IgA proteases. *J. Clin. Invest.* 54 pp. 1295-1300.

Rabbani, H., Kondo, N., Smith, C.I., Hammarstrom, L. (1995). The influence of gene deletion and duplication within the IGHC locus on serum immunoglobulin subclass levels. *Clin. Immunl. Immunopathol.* 76:S pp. 214-218.

Ratcliffe, M. J. H. (2006). Antibodies, immunoglobulin genes and the bursa of Fabricius in chicken. B cell development. *Devel. Comp. Immunol.* 30 pp. 101-118.

Reiter, Y., Schuck, P., Boyd, L. F. & Plaksin, D. (1999). An antibody single –domain phage display library of a native heavy chain variable region: Isolation of functional single-domain VH molecules with a unique interface. *J. Mol. Bio.* 290 pp. 685-698.

Retter, I., Chevillard, C., Scharfe, M., Conrad, A., Hafner, M., Im, T-H , Ludewig, M., Nordsied,G., Severitt, S., Thies, S., Mauhar, A., Bloecker, H., Mueller, W. & Riblet, R. (2007). Sequence and characterization of the Ig heavy chain constant and partial variable region of the mouse strain 129S1. *J. Immunol.* 179 pp. 2419-2427.

Reynaud, C.A., Anquez, V., Daher,A. & Weill, J. (1987). A hyperconversion mechanism generates the chichen pre-immune light chain repertoire. *Cell.* 48 pp. 379-388.

Rodkey, L. S. & Adler, F. L. (1983). Regulation of natural anti- allotypic antibody responses by network induced auto-anti-idiotypic responsiveness of their offspring. *J. Exp. Med.* 152 pp. 1024-1035.

Roes, J. & Rajewsky, K. (1993). Immunoglobulin D (IgD)-deficient mice reveal an auxiliary receptor function for IgD in antigen-mediated recruitment of B cells. *J. Exp. Med.* 177 pp. 45-55.

Rowe, D. S. & Fahey, J. L. (1965). A new class of human immunoglobulins. *J. Expt. Med.* 121 pp. 171-199.

Rumfelt, L. L., McKinney, E. C., Taylor, E. & Flajnik, M. F. 2002). The development of primary and secondary lymphoid tissues In the nurse shark *Ginglymostoma cirratum.* B cell zones precede dendritic cell immigration and T cell cell zones formation during ontogeny of the spleen. *Scand. J. Immunol.* 56 pp. 130-148.

Schiaffella, E., Sehgal, D., Anderson, A. O. & Mage,R. G. (1999). Gene conversion and hypermutation during diversification of VH sequences in developing germinal centers of immunized rabbits. *J. Immunol.* 162 pp. 3984-3995.

Schroeder, H. W. Jr., Hillson, H. L. & Perlmutter, R. M. (1987). Early restriction of human antibody repertoire. *Science* 238 pp. 791-793.

Schroeder. H.W. Jr, Hillson, H.L. & Perlmutter, R.M. (1990). Structure and evolution of mammalian VH families. *Int'l Immunol* 2 pp. 41-45.

Sheehan, K. M., Mainville, C. A., Willert, S. & Brodeur, P. H. (1993). The utilization of individual V_H exons in the primary repertoire of adult BALB/c mice. *J. Immunol.* 151 pp. 5363-5375.

Solem, S.T. & Stenvik, J. (2006). Antibody repertoire development in teleosts--a review with emphasis on salmonids and *Gadus morrhua* L. *Develop. Comp. Immunology* 30 pp. 57-76.

Spieker-Polet, H., Yam, P-C., & Knight K. L. (1993). Differential espression of 13 IgA-heavy chain genes in rabbit lymphoid tissues. *J. Immunol.* 150 pp. 5457-5465.

Srisapoome, P., Ohira, T., Hirona, I. & Aoki, T. (2004). Genes of the constant regions of functional heavy chain of Japanese flounder. *Parealichthys olivaceus. Immunogenetics* 56 pp. 292-300.

Stenvik, J. & Jorgensen, T. O. (2000). Immunoglobulin D (IgD) of Atlantic cod has a unique structure. *Immunogenetics* 51 pp. 452-461.

Sun, J., Hayward, C., Shinde, R., Christenson, R., Ford, S.P. & Butler, J.E. (1998). Antibody repertoire development in fetal and neonatal piglets. I. Four V_H genes account for 80% of V_H usage during 84 days of fetal life. *J. Immunol.* 161 pp. 5070-5078

Sun, J., Kacskovics, I., Brown, W. R. & Butler, J. E. (1994). Expressed swine V_H genes belong to a small V_H gene family homologous to human V_H III. *J. Immunol.* 153 pp. 5618-5627.

Szostak, J. W., Orr-Weaver, T. L. & Rothstein, R. J. (1983). The double-strand break repair model for recombination. *Cell.* 33 pp. 25-35.

Thomson, C. A., Little, K. Q., Reason, D. C. &Schrader, J. W. (2011). Somatic diversity in CDR3 loops allows single V-genes to encode innate immunological memories for multiple pathogens. *J. Immumnol.* 186: pp. 2291-2298.

Tiegs, S. L., Russell, D. M. & Nemazee, D. (1993). Receptor editing in self-reactive bone marrow B cells. *J. Exp. Med.* 177 pp. 1009-1020.

Vigliano, F. A., Bemudez, R., Quiroga, M. I. & Nieto, J. M. (2006). Evidence for melamno-macrophage centers of telosts as evolutionary precursors of germinal centers of higher vertebrates: An immunohistochemical study. *Fish and Shellfish Immunology* 21 pp. 467-471.

Vitetta, E. S., Cambier, J. C., Ligler, F. S., Kettman, J. R. & Uhr, J. W. (1977). B cell tolerance. IV. Differential role of surface IgM and IgD in determining tolerance susceptibility of murine B cells. *J. Exp. Med.* 146 pp. 1804-1808.

Wang, H. & Shlomchik., M. J. (1998). Maternal Ig mediates neonatal tolerance in rheumatoid factor transgenic mice but tolerance breaks down in adult mice. *J. Immunol*. 160 pp. 2263-2271.

Wikler, M., Demeur, C., Dewasne, G. & Urbain, J. (1980). Immunoregulatory role of maternal idiotypes. Ontogeny of immune networks. *J. Exp. Med.* 152 pp. 1024-1035.

Wilson, M., Bengten, E., Miller, N. W., Chen, L. W., Du Pasquier, L. & Warr, G. W. (1997). A novel chimeric Ig heavy chain from a teleost fish shares similarities to IgD. *Proc. Nat'l. Acad. Sci.* USA 94 pp. 4593-4597.

Winstead, C. R, Zhai, S. K., Sethupathi, P. & Knight, K. L. (1999). Antigen-induced somatic diversification of rabbit IgA genes: gene conversion and point mutation. *J. Immunol* 162 pp. 6602-6612.

Withers, D. R., Davison, T. F. & Young, J. R. (2005). Developmentally programmed expression of AID in chicken B cells. *Devel. Comp. Immunol.* 29 pp. 651-662.

Xu, J. L. & Davis., M. M. (2000). Diversity in the CDR3 region of V_H is sufficient for most antibody specificities. *Immunity* 13 pp. 37-45.

Yamaguchi, N., Shimizu, S., Hara, T. & Saito, T. (1983). The effector maternal antigenic stimulation upon the active immune responsiveness of their offspring. *Immunology* 50 pp. 229-238.

Yancopoulos, G. D., Desiderio, S. V., Paskind, M., Kearney, J. F., Baltimore, D. V. & Alt, F. W. (1984). Preferential usage of the most JH proximal VH gene segments in pre-B cell lines. *Nature* 311 pp. 717-733.

Yasuda, M., Kajiwara, E., Ekino, S., Taura, Y., Hirota, Y., Horiuchi, H., Matsuda, H., Furusawa, S. (2003). Immunobiology of chicken germinal center: I. Changes in surface Ig class expression in the chicken splenic germinal center after antigenic stimulation. *Develp. Comp. Immunol.* 27 pp. 159-166.

Zhao, Y, Kacskovics, I., Pan-Hammarstrom, Q, Liberies, D. A., Geli, J., Davis, S. K., Rabbani, H. & Hammarstrom, L. (2002). Artiodactyl IgD: the missing link *J. Immunol.* 169 pp. 4408-4416.

Zhao, Y., Pan-Hammarstrom, Q., Kacskovics, I. & Hammarstrom, L. (2003). The porcine Ig delta gene: unique chimeric splicing of the first constant region domain in its heavy chain transcripts. *J. Immunol.* 171: pp. 1312-1318.

Zhao, Y., Pan-Hammarstrom, Q., Yu, S., Wertz, N., Zhang, X., Li , N., Butler, J.E. &. Hammarstrom, L. (2006). Identification of IgF, a hinge-region containing Ig class and IgD in *Xenopus tropicalis*. *Proc. Natl. Acad. Sci* (USA) 103 pp. 12087-12092.

Zheng, N. Y., Wilson, K., Wang, X., Boston, A., Kolar, G., Jackson, S. M, Liu, Y. I., Pascual, V., Capra, J.D. & Wilson, P.C. (2004). Human immunoglobulin selection associated with class switch and possible tolerogenic origin for C delta class-switched B cells. *J. Clin Invest.* 113 pp. 1188-1201.
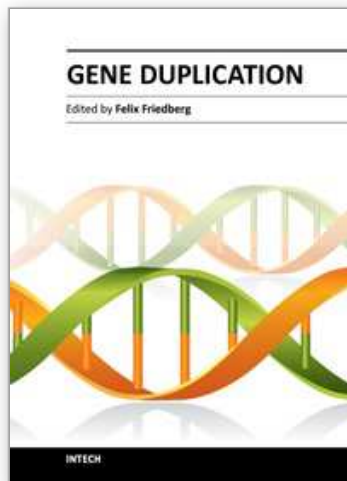
Gene Duplication

Edited by Prof. Felix Friedberg

The book Gene Duplication consists of 21 chapters divided in 3 parts: General Aspects, A Look at Some Gene Families and Examining Bundles of Genes. The importance of the study of Gene Duplication stems from the realization that the dynamic process of duplication is the "sine qua non" underlying the evolution of all living matter. Genes may be altered before or after the duplication process thereby undergoing neofunctionalization, thus creating in time new organisms which populate the Earth.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

J. E. Butler, Xiu-Zhu Sun and Nancy Wertz (2011). Immunoglobulin Polygeny: An Evolutionary Perspective, Gene Duplication, Prof. Felix Friedberg (Ed.), ISBN: 978-953-307-387-3, InTech, Available from: http://www.intechopen.com/books/gene-duplication/immunoglobulin-polygeny-an-evolutionary-perspective

INTECH
open science | open minds