Antibody Repertoire Development in Fetal and Neonatal Piglets. XI. The Relationship of Variable Heavy Chain Gene Usage and the Genomic Organization of the Variable Heavy Chain Locus

Tomoko Eguchi-Ogawa, Nancy Wertz, Xiu-Zhu Sun, Francois Puimi, Hirohide Uenishi, Kevin Wells, Patrick Chardon, Gregory J. Tobin and John E. Butler

J Immunol 2010; 184:3734-3742; Prepublished online 5 March 2010;
doi: 10.4049/jimmunol.0903616
http://www.jimmunol.org/content/184/7/3734

Supplementary Material
http://www.jimmunol.org/content/suppl/2010/03/05/jimmunol.0903616.DC1

References
This article cites 52 articles, 24 of which you can access for free at:
http://www.jimmunol.org/content/184/7/3734.full#ref-list-1

Subscription
Information about subscribing to The Journal of Immunology is online at:
http://jimmunol.org/subscription

Permissions
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts

Errata
An erratum has been published regarding this article. Please see next page or:
/content/190/3/1383.full.pdf
Antibody Repertoire Development in Fetal and Neonatal Piglets. XI. The Relationship of Variable Heavy Chain Gene Usage and the Genomic Organization of the Variable Heavy Chain Locus

Tomoko Eguchi-Ogawa,* Nancy Wertz,† Xiu-Zhu Sun,† Francois Puimi,‡ Hirohide Uenishi,* Kevin Wells,§ Patrick Chardon,‡ Gregory J. Tobin,§ and John E. Butler†

In this study, we have mapped the 3′ H chain V region (VH) genes and those in the H chain diversity, H chain joining, and 5′ portion of the H chain constant locus. We show that swine possess only two functional H chain diversity segments and only one functional H chain joining segment. These data help to explain more than a decade of observations on the preimmune repertoire of this species and reveal the vulnerability of swine to natural or designed mutational events. The results are consistent with earlier studies on the region containing Enh, Cμ, and Cδ while revealing that the ancestral IgG3 is the most 5′ Cγ gene. We also observed a recent duplication (~1.6 million years ago) in the VH locus that contains six of the seven VH genes that comprise 75% of the preimmune repertoire. Because there are no known transfers of immune regulators or Ags that cross the placenta as in mice and humans, fetal VH usage must be intrinsically regulated. Therefore, we quantified VH usage in fetal piglets and demonstrated that usage is independent of the position of VH genes in the genome; the most functional VH gene (IGHV2) is rarely used, whereas certain upstream genes (IGHV14 and IGHV15) are predominately used early in fetal liver but seldom thereafter. Similar to previous studies, three VH genes account for 40% of the repertoire and six for ~70%. This limited combinatorial diversity of the porcine VH repertoire further emphasizes the dependence on CDR3 diversity for generating the preimmune Ab repertoire of this species. The Journal of Immunology, 2010, 184: 3734–3472.

The H chain Ig variable locus is organized similarly in all mammals that have been studied. As few as 12 to >1000 H chain V region (VH) genes are thought to be arrayed 5′ of multiple H chain diversity (DH) segments that in turn are found 5′ of up to nine H chain joining (JH) segments (1, 2). VH genes characterized to date include many pseudogenes (3), such as those with nonconsensus signal sequences, stop codons, frameshift mutations, or the absence of functional leader sequences. In humans, for example, only 38–46 of the >100 VH genes (haplotype dependent) are functional, whereas 27 DH segments can be expressed, but only six of the nine JH segments are functional (2, 4). The VH genes of vertebrates belong to different families that have been grouped into three clans based on sequence homologies (5). There are large differences in the number of VH genes and VH gene families among species. In swine, <30 genes have been described, all belonging to the ancestral VH3 family (6, 7) and have identical or nearly identical framework regions but differ in their CDR1 and CDR2 regions (8, 9). However, CDR regions are often shared among different porcine VH genes, suggesting they could be a consequence of gene duplication combined with a type of genomic gene conversion (10, 11).

The distribution of VH gene segment usage in fetal and newborn mice, humans, rabbits, and swine is nonrandom, and usage patterns vary during ontogeny. In the fetal liver of mice, members of the VH5 and VH2 families (especially 7183 and Q52, respectively) are preferentially used (12, 13), but after birth, splenic B cells use the VH1 family (J558) (14). The VH2 and VH5 family genes are the most 3′ in the locus, whereas the J558 family is midway (more 5′) in the locus, suggesting that VH3 usage in the preimmune repertoire favors 3′ VH genes but that usage shifts upstream, presumably due to exposure to environmental Ag because this shift does not occur in germ-free mice (14, 15). A preferential usage of the most proximal VH3 family genes has also been reported in humans (16, 17). Early VH usage in rabbits favors the most 3′ VH gene (VH3-4) in 90% of VDJ rearrangements, but this preferential usage declines during development and with exposure to environmental Ag (18). However, developmental repertoire changes in rabbits are compounded by somatic gene conversion, a phenomenon rarely seen in mice (19).

The pattern of preferential usage of 3′ VH genes is not supported in all studies. In humans, V3-23 and V3-30 are preferentially used in the...
preimmune repertoire, but these map to the middle of the locus (20). Alternatives to the locus position concept include: 1) negative selection for B cells utilizing VH genes that recognize certain self-Ags; 2) positive selection for B cells by stromal ligands based on features of their BCR; and 3) the evolutionary conservation of BCRs bearing VH genes that encode binding sites that recognize bacteria that threaten the newborn. The basis for 2) and 3) above could be the same. For example, human V3-23, mouse V\textsubscript{H}283 (a V\textsubscript{H}2 family gene), and shark VH genes share >80% sequence similarity at the protein level (7, 21) and encode BCRs that are polyreactive and recognize bacterial polysaccharides (22). In humans and mice, such Abs are often associated with IgG3 (23, 24) and class switch that occurs in the absence of environmental Ag (25, 26) and in CD40/CD40L-deficient mice (27). In swine, IgG3 expression accounts for >60% of the total IgG expression in the ileal Peyer’s patches (IPP) of fetal piglets (28). The IPP of swine is the anatomical homolog of the IPP of sheep, in which Ab repertoire development is considered to be Ag-independent (29). In both the human and mouse genome, IgG3 is the most \( \text{C}_6 \) proximal functional \( \text{C}_\gamma \) gene (reviewed in Refs. 30, 31).

In contrast to rabbits, mice, and humans, fetal piglets present a different pattern of fetal/neonatal development that could affect VH gene usage. In swine, there is no known protein transfer from mother to fetus including IgG (32–34). Because maternal IgG can influence fetal immune responses (35, 36) and presumably B cell selection and development, usage of certain VDJ rearrangements is independent of this influence in fetal piglets. Although it is known that certain porcine viruses can cross the placenta, the mechanism remains unknown. In any case, our studies on VDJ usage in fetal piglets were done in animals free of all detectable viral pathogens. Offspring of swine are precocial and routinely reared in isolators separate from their mothers, allowing the experimenter to control environmental factors that might influence the development of adaptive immunity (37). In swine, VH, DH, and JH usage is highly restricted. In the expressed preimmune repertoire, four to seven VH genes account for 75–95% of the repertoire (9, 38) (Table I). Although the mechanism is unknown, swine use only a single JH and two DH segments to generate their preimmune Ab repertoire (38–40). We believed these observations might be better understood by mapping the VH, DH, and JH loci of swine. We were especially interested to know if the predominant use of a few VH genes in the preimmune repertoire was confined to those at the 3' end of the VH region of the H chain locus that includes 15 DH proximal VH genes and the 5' portion of the H chain constant (\( \text{C}_\text{\text{H}} \)) gene. Our map shows that swine possess only one functional JH gene segment, apparently two functional DH segments, and that preferential VH usage in primary B cell tissues of fetal animals cannot be explained as a stochastic process dependent on the location of VH genes in the locus. In contrast, the dominant early expression of porcine IgG3 in late-term fetuses in an organ believed to develop an Ag-independent repertoire appears correlated with its position in the H chain C region of the locus.

**Materials and Methods**

**Nomenclature**

The porcine VH genes were originally named according to their order of discovery and given vernacular names like VH\textsubscript{A} and VH\textsubscript{B} (Table I) (9). The intention was that once their location in the genome was known, the nomenclature would be changed to the familiar VH\textsubscript{1}, VH\textsubscript{2}, etc., system or the ImMunoGeneTics (IMGT) system (IGHV1, IGHV2). In this report, we have maintained the vernacular nomenclature in certain cases and used it in parenthesis in the text. Both the vernacular and IMGT system are used in Table II. Because the IMGT system contains an unfortunate redundancy (IGHD, IGHJ and IGHK regions and 5' IGHK genes (\( \text{C}_\mu \), \( \text{C}_\delta \), \( \text{C}_\gamma \), and \( \text{C}_\delta \)) diversity segments), we have used the conventional \( \text{C}_\mu \), \( \text{C}_\delta \), and \( \text{C}_\gamma \) for \( \text{C}_\mu \) regions.

**Screening of bacterial artificial chromosome clones and construction of bacterial artificial chromosome contigs**

All bacterial artificial chromosome (BAC) clones were isolated from a genomic library constructed by using a Large White boar resulting from intramurine matings at Institut Nationale Recherche Agronomique (Jouy en Josas, France) (42). BAC clones containing the porcine IGHV, IGHJ, and IGHK regions and 5' IGHK genes (\( \text{C}_\mu \), \( \text{C}_\delta \), \( \text{C}_\gamma \), and \( \text{C}_\delta \)) were identified by using a PCR-based screening system. Selected BAC clones were cultured overnight in Luria-Bertani broth containing 12.5 \( \mu \)M chloramphenicol. BAC clones 850A09 and 747A01 that hybridized with polynucleotide probes specific for VH genes, \( \text{J}_H \), and \( \text{C}_\mu \) (Table I) were selected for further study. DNA from relevant clones was extracted with a Qiagen Plasmid Midi Kit (Qiagen, Hilden, Germany).

**Sequencing of BAC clones**

BAC DNAs were extracted by using a Qiagen Large Construction Kit (Qiagen). The DNA were cleaved into 2- to 4-kb fragments by sonication and then subcloned into the plasmid vector pUC18 (Takara, Otsu, Japan) at the HincII site to construct a shotgun library. The inserts obtained were sequenced by an ABI 3730xl sequencer with a Big Dye Terminator version 3 cycle sequencing kit (Applied Biosystems, Foster City, CA) using universal primers annealing to vector sequences. Sequence data were processed with Phred base caller and assembled with Phrap (CodonCode, Dedham, MA) (43, 44). Compiled sequences of the two BAC clones have been uploaded in GenBank (AB513624 and AB513625). Restriction enzyme digestion was performed to confirm the sequence assembly. BAC clones were digested

---

**Table I. Probe sequences used to identify common porcine VH genes, \( \text{J}_H \), and \( \text{C}_\mu \)**

<table>
<thead>
<tr>
<th>VH Gene</th>
<th>CDR1</th>
<th>CDR2</th>
<th>3' FR1</th>
<th>CDR1</th>
<th>5' F2</th>
<th>CDR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>VH\textsubscript{A}</td>
<td>A</td>
<td>A</td>
<td>CAGT</td>
<td>AGTAATGCTAGTACTAGT</td>
<td>TGGCA</td>
<td>GCATTAGTACTAGT</td>
</tr>
<tr>
<td>VH\textsubscript{B}</td>
<td>B</td>
<td>B</td>
<td>GACACCTTCATTCAAC</td>
<td>CGGCATCCTGACTAGTAC</td>
<td>TGGAC</td>
<td>GCCATTGCTAGTACTAGT</td>
</tr>
<tr>
<td>VH\textsubscript{G}</td>
<td>G</td>
<td>E</td>
<td>TTCACTG</td>
<td>AGTTATGCTAGTACTAGT</td>
<td>TGGAC</td>
<td>GCCATTGCTAGTACTAGT</td>
</tr>
<tr>
<td>VH\textsubscript{E}</td>
<td>E</td>
<td>E</td>
<td>TCAGT</td>
<td>AGTTATGCTAGTACTAGT</td>
<td>TGGAC</td>
<td>GCCATTGCTAGTACTAGT</td>
</tr>
<tr>
<td>VH\textsubscript{F}</td>
<td>F</td>
<td>F</td>
<td>GTCAGT</td>
<td>AGTTATGCTAGTACTAGT</td>
<td>TGGAC</td>
<td>GCCATTGCTAGTACTAGT</td>
</tr>
<tr>
<td>VH\textsubscript{C}</td>
<td>C</td>
<td>C</td>
<td>GTCAGT</td>
<td>AGTTATGCTAGTACTAGT</td>
<td>TGGAC</td>
<td>GCCATTGCTAGTACTAGT</td>
</tr>
<tr>
<td>VH\textsubscript{Z}</td>
<td>E</td>
<td>C</td>
<td>TCAGT</td>
<td>AGTTATGCTAGTACTAGT</td>
<td>TGGAC</td>
<td>GCCATTGCTAGTACTAGT</td>
</tr>
<tr>
<td>VH\textsubscript{Y}</td>
<td>C</td>
<td>A</td>
<td>AGTTATGCTAGTACTAGT</td>
<td>TGGAC</td>
<td>GCCATTGCTAGTACTAGT</td>
<td></td>
</tr>
<tr>
<td>VH\textsubscript{A*}</td>
<td>A</td>
<td>A</td>
<td>AGAAGGACTTCCAGCCCAACAC</td>
<td>TGGA</td>
<td>GCCATTGCTAGTACTAGT</td>
<td></td>
</tr>
<tr>
<td>Pa9</td>
<td>CAGAGCTCCTCCAGGAGGA</td>
<td>CGAGAGCTCCTCCAGGAGGA</td>
<td>FR3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JH</td>
<td>TGGAGACCAACGACTTCCAA</td>
<td>TGGAGACCAACGACTTCCAA</td>
<td>FR2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{C}_\mu )</td>
<td>GCACATGCCCCGCTCGTGTAA</td>
<td>GCACATGCCCCGCTCGTGTAA</td>
<td>( \text{C}_\mu )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Columns labeled CDR1 and CDR2 list the vernacular designation used in the laboratory for the porcine VH genes (see also Table II). The upstream VH\textsubscript{A} (IGHV101) is designated as VH\textsubscript{A*} and requires a special probe that binds in the FR3 region for identification. Detection of VH\textsubscript{A*} requires a combination of the probes described in Table I. This method of recognition was confirmed by sequence analysis.
with HindIII, PvuI, or FspI, and the fragment sizes were compared with those predicted by a Genetyx sequence editor (Genetyx, Tokyo, Japan). The number of IGHV genes contained in the BAC clones was confirmed by PCR methods using each BAC clone DNA as a template. Each IGHV gene was amplified using a primer set annealing to sequences in 5’ framework region (FR) 1 (5’-GAGGAGAAGCTGGTGGAGT-3’) and 3’ FR3 (5’-GCCGTTCTTCCTGTA-3’) that are common to all known porcine V H genes except VHB. To amplify the VH B, 5’ FR1 (5’-GAGGAGAAGCTGGTGGAGT-3’) and FR3 IGHVB (5’-GCCGTTCTTCCTGTA-3’) were used. The products were cloned into the TOPO 2.1 vector (Invitrogen, Carlsbad, CA). More than 80 clones were selected at random from each PCR product for sequence analysis.

Analysis of mapping data

Sequence similarity was examined using the BLAST2 program (45) (www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi), and interspersed repetitive sequences in the pig genome were detected by using RepeatMasker (A. Smit and P. Green, unpublished observations; www.repeatmasker.org/) with Repbase (46) (www.girinst.org/repbase/index.html). Homologous regions between genomic sequences were identified by using the PipMaker program (47). Sequence alignment and phylogenetic analysis were performed by the CLUSTALW (1.83) program (48). The detected IGHV genes were analyzed by the BLAST program with public nucleotide databases (DNA Data Bank of Japan, EMBL, and GenBank). Only sequences that showed ≥98% identity to previously identified porcine IGHV sequences were adopted.

Quantitation of V H gene usage in fetal liver and late-gestation bone marrow

DNA from fetal liver at 30 d of gestation (DG30) from four piglets each from four unrelated gilts and RNA from the bone marrow of five fetuses at DG95 were used to recover VDJ rearrangement from DNA and cDNA, respectively, as previously described (9, 39, 40). The VDJs recovered by PCR were cloned into pcR²-TOPO (Invitrogen), and 632 clones from fetal liver and 432 clones from bone marrow transcripts were then sequentially hybridized with pan-specific FR2 probe to identify all clones with a plasmid containing a VDJ insert (Table I). These clones were also hybridized with a pan-specific FR2 probe to identify all clones with a plasmid containing a VDJ insert (Table I).

Statistical analysis

V H usage was compiled in Excel spreadsheets (Microsoft, Redmond, WA) and displayed as proportional usage (mean and SD) in GraphPad Prism (GraphPad, San Diego, CA). Mean differences were compared by Student t test routines embedded in the Prism program (GraphPad).

Results

Genomic structure of the region containing porcine IGHD, IGHI, and the 3’ IGHV genes

We screened several BAC clones containing IGHD and IGHI segments by PCR to identify clones that encompass the full locus. Clone 851A09 contained 118,088 bp with sequences from the IGHD3 segment to IGHV15. Clone 747A01 overlapped 851A09 by ∼29.6 kbp and encompassed IGHG2 through IGHV1P (Fig. 1). Together the two clones represent 229,148 bp of genomic sequence. The two BAC clones were subjected to shotgun sequencing as described in Materials and Methods. The sequence overlap between 747A01 and 851A09 was 100% homologous, indicating that they were derived from the same allele. By analyzing open reading frames and using BLAST against GenBank sequences, we identified 15 IGHV, 4 IGHD,
and 5 IGHJ gene segments as well as the constant regions of IgM (Cμ) and IgD (Cδ) and the two most 5’ IgG genes (Cy3 and Cy5) (Fig. 1, Tables II–IV). The sequence of the switch region and C region of IgM had been reported previously (49); this sequence had a high level of identity to the sequence determined in this study, showing that these sequences were derived from the same region in the locus.

Previous studies showed that functionally rearranged VDJ segments used only two IGHD segments and a single IGHJ segment (38, 40). Although only transcripts carrying IGHD1, IGHD2, and IGHJ5 are in the public databases, we identified four IGHD segments and five IGHJ segments. The porcine genomic locus carrying the IGHD and IGHJ genes was much shorter than its human counterpart (9 IGHJ and 27 IGHD are detected in the corresponding region of the human genome: NT_026437.11, 86930000 to 87630000). In the case of JH, only reading frame III of IGHJ5 translates to the sequence of the genome: NT_026437.11, 86930000 to 87630000). In the case of JH, only one JH region of swine (38) and IGHD1, IGHD2, and IGHJ5 are in the public databases, we identified four IGHD segments and five IGHJ segments. The porcine genomic locus carrying the IGHD and IGHJ genes was much shorter than its human counterpart (9 IGHJ and 27 IGHD are detected in the corresponding region of the human genome: NT_026437.11, 86930000 to 87630000).

As shown in Fig. 2B, the IGHJ segments have noncanonical heptamer sequences (Table IV). A sequence encoded by IGHJ3 has yet to be found in a rearranged or expressed VDJ, although its recombination signal sequence (RSS) seems functional, and no stop codons are present.

Analysis of the D J shows that IGHD1 (D HA) and IGHD2 (D HB) have canonical RSSs, whereas both the heptamer and nonamer of IGHD3 are noncanonical (Table III). The RSS of IGHD4 should have canonical RSSs, whereas both the heptamer and nonamer of IGHD5 are noncanonical (Table III). The RSS of IGHD4 should have canonical RSSs, whereas both the heptamer and nonamer of IGHD5 are noncanonical (Table III).

The sequence of the switch region and C region of IgM was identical to the one previously reported (49); this sequence had a high level of identity to the sequence determined in this study, showing that these sequences were derived from the same region in the locus.

Although only transcripts carrying IGHD1, IGHD2, and IGHJ5 are highly expressed in newborn piglets and amniotic fluid (38, 40), no stop codons are present.

The IMGT system is based on the position in the genome as shown in Fig. 1, whereas the IMGT system is based on the position in the genome as shown in Fig. 1. The IMGT system is based on the position in the genome as shown in Fig. 1.

Table II. IgH genes identified from the genome sequence determined in this study

<table>
<thead>
<tr>
<th>IMGT</th>
<th>Vernaculara</th>
<th>No. of Exons</th>
<th>Start</th>
<th>End</th>
<th>CDS Length (bp)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGHV15</td>
<td>V H N</td>
<td>2</td>
<td>3032</td>
<td>3473</td>
<td>535</td>
<td></td>
</tr>
<tr>
<td>IGHV14</td>
<td>V H Y</td>
<td>2</td>
<td>9306</td>
<td>9747</td>
<td>535</td>
<td></td>
</tr>
<tr>
<td>IGHV13P</td>
<td></td>
<td>1</td>
<td>12,154</td>
<td>12,366</td>
<td>213</td>
<td>Exxon 1 and part of exon 2 were collapsed</td>
</tr>
<tr>
<td>IGHV12</td>
<td>V H B*</td>
<td>2</td>
<td>15,678</td>
<td>16,126</td>
<td>359</td>
<td></td>
</tr>
<tr>
<td>IGHV11</td>
<td>V F</td>
<td>2</td>
<td>27,618</td>
<td>28,065</td>
<td>359</td>
<td></td>
</tr>
<tr>
<td>IGHV10</td>
<td>V A*</td>
<td>2</td>
<td>34,235</td>
<td>34,676</td>
<td>353</td>
<td></td>
</tr>
<tr>
<td>IGHV9P</td>
<td></td>
<td>1</td>
<td>37,094</td>
<td>37,306</td>
<td>213</td>
<td>Exxon 1 and part of exon 2 were collapsed</td>
</tr>
<tr>
<td>IGHV8</td>
<td></td>
<td>2</td>
<td>42,141</td>
<td>42,589</td>
<td>359</td>
<td></td>
</tr>
<tr>
<td>IGHV7P</td>
<td></td>
<td>1</td>
<td>45,002</td>
<td>45,214</td>
<td>213</td>
<td>Exxon 1 and part of exon 2 were collapsed</td>
</tr>
<tr>
<td>IGHV6</td>
<td></td>
<td>2</td>
<td>48,526</td>
<td>48,974</td>
<td>359</td>
<td></td>
</tr>
<tr>
<td>IGHV5</td>
<td></td>
<td>2</td>
<td>60,198</td>
<td>60,645</td>
<td>359</td>
<td></td>
</tr>
<tr>
<td>IGHV4</td>
<td></td>
<td>2</td>
<td>66,808</td>
<td>67,249</td>
<td>353</td>
<td></td>
</tr>
<tr>
<td>IGHV3P</td>
<td></td>
<td>1</td>
<td>69,661</td>
<td>69,873</td>
<td>213</td>
<td>Exxon 1 and part of exon 2 were collapsed</td>
</tr>
<tr>
<td>IGHV2</td>
<td></td>
<td>2</td>
<td>76,020</td>
<td>76,470</td>
<td>359</td>
<td></td>
</tr>
<tr>
<td>IGHV1P</td>
<td>Psg</td>
<td>2</td>
<td>90,845</td>
<td>91,288</td>
<td>352</td>
<td>Stop codon in exon 2</td>
</tr>
<tr>
<td>IGHD1</td>
<td>D H A</td>
<td>1</td>
<td>108,609</td>
<td>108,646</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>IGHD2</td>
<td>D H B</td>
<td>1</td>
<td>109,219</td>
<td>109,246</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>IGHD3</td>
<td></td>
<td>1</td>
<td>113,937</td>
<td>113,973</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>IGHD4</td>
<td></td>
<td>1</td>
<td>131,317</td>
<td>131,327</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>IGHJ1</td>
<td></td>
<td>1</td>
<td>132,104</td>
<td>132,157</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>IGHJ2</td>
<td></td>
<td>1</td>
<td>132,277</td>
<td>132,329</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>IGHJ3</td>
<td></td>
<td>1</td>
<td>132,587</td>
<td>132,634</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>IGHJ4</td>
<td></td>
<td>1</td>
<td>132,979</td>
<td>133,029</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>IGHJ5</td>
<td>J S</td>
<td>1</td>
<td>133,523</td>
<td>133,576</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>IGHM</td>
<td>IgM (Cμ)</td>
<td>4</td>
<td>139,705</td>
<td>141,406</td>
<td>1216</td>
<td></td>
</tr>
<tr>
<td>IGHD</td>
<td>IgD (Cδ)</td>
<td>5</td>
<td>147,270</td>
<td>152,588</td>
<td>1060</td>
<td></td>
</tr>
<tr>
<td>IGHG1</td>
<td>IgG3* (Cy3)</td>
<td>4</td>
<td>188,628</td>
<td>190,144</td>
<td>1001</td>
<td></td>
</tr>
<tr>
<td>IGHG2</td>
<td>IgG5* (Cy5)</td>
<td>4</td>
<td>209,664</td>
<td>211,155</td>
<td>974</td>
<td></td>
</tr>
</tbody>
</table>

*aV H gene names were classified according to CDR1 and CDR2 sequence specificities (9). The vernacular terminology is based on order of discovery, whereas the IMGT system is based on the position in the genome as shown in Fig. 1.

*bSwitch region for IgM

*cIgG nomenclature based on recent studies of Butler et al. (41).

dIgG5* is an allele of IgG5 (41).

Psg, pseudogenes discovered in genomic DNA prior to mapping studies (see Ref. 9).
from each other ∼1.6 million y ago. This calculation is based on the estimated mutation rate in pseudogenes: 4.6 × 10⁻⁸ (50).

Evidence for recent duplication of the IGHV region is further supported by the observation of conserved HindIII and Pvu1 restriction fragments in Fig. 1. This duplication unit contained the duplicated V<sub>H</sub>A and V<sub>H</sub>B genes, which, along with V<sub>H</sub>C and V<sub>H</sub>E, comprise >75% of the preimmune repertoire in piglets (39). The exon-intron structures of the detected V<sub>H</sub>A and V<sub>H</sub>B genes were also conserved, perhaps a clue that these genes are important in generating the Ab repertoire.

**V<sub>H</sub> gene usage in primary lymphoid tissues at DG30 and DG95**

Fig. 3 summarizes V<sub>H</sub> usage in VDJ rearrangement in 632 clones from DG30 fetal liver and 432 cloned transcripts from DG95 bone marrow. V<sub>H</sub> genes are ordered on the x-axis according to their position in the locus, 3’ to 5’ (Fig. 1, Table II). The results confirm earlier findings that IGHV4, IGHV6, IGHV10, and IGHV12 dominate the early repertoire (∼30%) but also show that IGHV14 (V<sub>H</sub>Y) and IGHV15 (V<sub>H</sub>N) play an equally dominant role at DG30. Noteworthy is that the first functional V<sub>H</sub> gene IGHV2 (V<sub>H</sub>G) (Fig. 1) is seldom used. Thus, early usage in a primary lymphoid tissue cannot be ascribed to the position of the V<sub>H</sub> genes in the locus. In older fetuses (i.e., DG95), usage of IGHV14 and IGHV15 nearly disappears, whereas V<sub>H</sub>C and both IGHV4 (V<sub>H</sub>A) and IGHV10 (V<sub>H</sub>A*) are significantly increased. V<sub>H</sub> usage at DG95 DNA rearrangements or transcripts from bone marrow, IPP, and spleen does not significantly differ (J. E. Butler, X. Z. Sun, N. Wertz, K. M. Lager, and G. Tobin, unpublished observations).

These data reject the notion that V<sub>H</sub> usage during development follows a continuous progression of V<sub>H</sub> usage from 3’ to 5’. In both DG30 liver and DG95 bone marrow, usage of unidentified V<sub>H</sub> genes (UNK) does not change. Previous studies (9) indicate that many VDJ clones identified as UNK in conventionally reared young pigs use the same common genes that account for 70% of the preimmune repertoire (Fig. 3) but have mutated CDRs so they no longer hybridize with gene-specific probes (Table I). This low frequency of somatic hypermutation occurs in utero in the complete absence of environmental Ag stimulation, although there is no accumulation of mutations in CDR regions as is the case in Ag-stimulated piglets (S. Bratsch, N. Wertz, T. Kunz, and J.E. Butler, unpublished observations).

As indicated above, Enh, C<sub>p</sub>, and C<sub>d</sub> map to the same region as previously described (40, 51). Among the six C<sub>y</sub> subclasses and eleven allotypes (41), C<sub>y</sub>3 and C<sub>y</sub>5 are the most C<sub>d</sub> proximal C region genes.

**Discussion**

We present the first partial map of the porcine H chain locus that covers the 5’ portion of the C<sub>j</sub> region, J<sub>H</sub> region, D<sub>H</sub> region, and the first 15 of the most 3’ V<sub>H</sub> genes. These results confirm previous data on C<sub>p</sub> and C<sub>d</sub> (51) and S<sub>p</sub> and C<sub>p</sub> (49). Data presented in this study also explain >15 y of observations that swine express only two D<sub>H</sub> segments, IGHD1, formally called D<sub>λ</sub>A, and IGHD2, formally called D<sub>λ</sub>B (40). Our mapping data also explain why swine express only a single J<sub>H</sub> sequence (38). Evidence that only IGHJ5 (J<sub>H</sub>5) is functional has recently been confirmed in studies in which this segment was mutated, resulting in B cell knockout piglets (M. Mendicino, J. Ramsoondar, C. Phelps, T. Vaught, S. Ball, T. Lerooth, J. Moonahan, S. Chen, A. Dandro, J. Boone, P. Jobst, A. Vance, N. Wertz, Z. Bergman, X-Z Sun, I. Polejaeva, J. Butler, Y. Dai, D. Ayares, and K. Wells, personal communication). Collectively, the map and sequence data we provide for D<sub>H</sub>, J<sub>H</sub>, S<sub>p</sub>, C<sub>p</sub>, and C<sub>d</sub> are consistent with earlier partial mapping studies (49, 51) and explain why swine express only two D<sub>H</sub> segments and a single J<sub>H</sub> (38, 40).
Table IV. IGHJ genes detected in this study

<table>
<thead>
<tr>
<th>Gene</th>
<th>9-mer Sequence</th>
<th>23-mer Sequence</th>
<th>7-mer Reading Frame</th>
<th>Sequence and Deduced Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGHJ1</td>
<td>GCCATGGCTACTTAGATTCGTGGGGCCAGGGCATCCTGGTCACCGTCTCCTCAG</td>
<td>I IAMAT*IRGARASWSPSPQ</td>
<td>I II PWLLRFVGPGHPGHRLL</td>
<td>I II I HGYLDSWGQGILVTVSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGHJ2</td>
<td>CCCCTGGAAACTTGACCACTGGGGCAGGGGCGTCCTGGTCACCGTCTCCTCAG</td>
<td>I PLET*PLGQGRPGHRLL</td>
<td>I II PWKLDHWGRGVLVTVSS</td>
<td>I II I PGNLTTGAGASWSPSPQ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGHJ3</td>
<td>ACAACGGGCTCGAAAGCTGGGGCCAGGGGACCCTAGTCTACGACGCCTCGG</td>
<td>I TIFTAGAEESRSPSPQ</td>
<td>I II PSSQLGPRSRGHRLL</td>
<td>I II I HLHSWGRGVEVTVSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGHJ4</td>
<td>ATTACTATGCTATGGATCTCTGGGGCCCAGGCGTTGAAGTCGTCGTGTCCTCAG</td>
<td>I TTGSKAGARGP*STTPR</td>
<td>I II QRARKLGPGDPSLRRL</td>
<td>I II I NGLESWGQGTLVYDAS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGHJ5</td>
<td>CTGGGATCCGGGCTTAGTTGTCGCAATGTG</td>
<td>I ITMLWISGAQALKSSCPQ</td>
<td>I II LLCYGSLGPRR*SRRVL</td>
<td>I II YYAMDLWGPGVEVVVSS</td>
</tr>
</tbody>
</table>

Boldface text indicates only expressed J region.

Canonical heptamers and nonamers.

p

Table II).

Ab repertoire development during fetal life in piglets argues that the pattern of V\_H usage that we observed is intrinsic and therefore may differ from earlier studies in mouse and human fetuses. The mapping studies provided in this paper eliminate gene position as an explanation for V\_H, D\_H, and J\_H expression. Although it is a mistake to extrapolate data from one mammal to another, our findings and the controversy in the literature make us skeptical of any arguments that V\_H usage in early development is exclusively positionally dependent. The studies of Pospisil et al. (33) using the Alicia rabbit would support our skepticism.

Our study does offer several insights that may be related. First is the evidence for recent duplication within the porcine V\_H locus. IGHV4 (V\_HA) and IGHV10 (V\_HA*) share identical CDR1 and CDR2 regions (Table I) and differ only in one nucleotide in FR3. The latter difference allowed us to prepare a probe that could distinguish the expression of these two genes by probe hybridization (Fig. 3, Table I). Likewise, IGHV6 (V\_HB) and IGHV12 (V\_HB+) differ by only one nucleotide in CDR1. For this reason, gene-specific probes for V\_HB do not distinguish between IGHV6 and IGHV12 so that usage of both is included in V\_HB usage (Fig. 3). Other similarities suggest that IGHV9P to IGHV14 and IGHV3P to IGHV8 (Fig. 1) may be recent duplicons. Similar duplicons have been reported in the human V\_H genome (52). Thus, whatever factors control V\_H usage would be likely to affect all of the genes in each of these duplicon cassettes. Although overall V\_H usage is not positional, proportional usage for V\_H genes within each cassette may be the case (e.g., IGHV3P to IGHV8 in the first cassette and IGHV9P to IGHV14 in the second cassette).

Swine genetics, especially the fine genetics of lymphocyte receptors, are still in infancy compared with mouse genetics. Most of the nearly 30 V\_H genes described for swine were mostly recovered from outbred farm pigs so that the number very likely contains numerous alleles (9). The BAC library used in this study was made from the Large White breed that was interbred at Institut Nationale Recherche Agronomique, whereas the data presented in Fig. 3 were derived from United States animals, largely Landrace, Yorkshires, and their crosses that are preferred over Landrace, Yorkshire, and their crosses that are preferred over inbred animals for fecundity, health, and absence of viral disease, all of which are also important criteria for their use in fetal and neonatal isolator studies. We know from studies on Ig genetics that even randomly inbred swine can be homozygous in the IGHC locus (53). This background might therefore explain one of two troubling observations made in this study. The first concerns our failure to localize V\_C within the region of the V\_H locus that was mapped, despite its prominent expression (Fig. 3) (9, 39, 40). IGHV14 (V\_HY) (Fig. 1, Table II) shares CDR1 with VpH1 and CDR2 with IGHV4 (V\_HA) and IGHV10 (V\_HA*); all porcine V\_H genes have essentially identical framework sequences. It is altogether possible that IGHV14 (V\_HY) and VpH1 are alleles and the BAC library came...
FIGURE 2. A, Phylogenetic tree of porcine IGHV genes detected in this study. Human $\beta_2$-microglobulin (huB2M: NM_004048.2) was used as an unrelated reference. Exon 2 sequences (or sequences corresponding to exon 2 of functional IGHV segments) were used for the alignment. Because all porcine $V_{H}$ genes belong to the $V_{H}3$ family of clan III, we designated the various subgroups as $V_{H}3-1$, -2, and -3, respectively. Bootstrap values for 1000 replicates are indicated beside the branches. B, Genomic structure in the vicinity of the region coding IGHV. Repetitive sequences such as LINE/L1 and PRE1 are indicated by black boxes. IGHV genes are indicated by white boxes. Areas surrounded by dashed lines are highly conserved regions. C, Phylogenetic tree estimating the divergence of the two major duplicons within the mapped locus. Line, long interspersed nucleotide element; Pre, porcine repeat element family belonging to short interspersed nucleotide elements.
from a swine homozygous for IGHV14. The second observation concerns IGHV5 (VH14) that is expressed in low frequency in outbred North American swine (9). However, based on its position in the first duplication (Fig. 1), high expression might be expected. IGHV5 (VH14) shares CDR1 with IGHV8 (VH4) and CDR2 with IGHV11 (VH4), respectively. Because all three (IGHV5, IGHV8, and IGHV11) map to separate loci, IGHV5 cannot be an allele of either IGHV8 or IGHV11. Recent progress with the porcine genome project might resolve these issues.

Although the gene mapping studies we present can explain the expression of only two DH segments and one JH, they fail to explain the basis of selection for VH gene usage and expression. Clearly, it is not position in the locus. At DG30 in liver, there is very little transcription and no BCR expression as determined by flow cytometry and immunohistochemistry (54). Because this is not true at DG95, the changes in VH usage in fetal BCRs at DG95 could be due to selection by intrinsic ligands for B cells that use certain VH genes to form their BCRs.

Theories to explain positive selection of B cells expressing certain surface receptors date to the pre-B selection hypothesis of Melchers (55) and the work of Pozisil et al. (56). The latter studies used Alicia rabbits that carry a deficient VH1. These rabbits show increased expression of a VH1-like gene that was generated by gene conversion following gut colonization. As indicated above, the authentic VH1 gene is normally used in >90% of rearrangements in newborn rabbits (18). These observations by Melchers (55) and the Pozisil team (56) can be interpreted to mean that B cells at certain stages of development display BCRs that are recognized by stromal receptors that promote their survival and proliferation.

The dominant role of CDR3 in determining Ab specificity is well accepted and was experimentally shown in studies by Padian (57) and Xu and Davis (58). The latter group showed that mice with a single VH gene could make Abs to nearly all of the same Ags as control littermates with a full germline VH repertoire as long as the DH and JH region was intact. Because of limited VH, DH, and JH usage in swine, we estimated that >95% of the repertoire is determined by CDR3 (59). In the case of TCR, CDR1 and CDR2 are purported to be mainly responsible for MHC recognition, whereas CDR3 is responsible for peptide (Ag) recognition. Because TCR and BCR are structural and genetic homologs, what is true for the goose may also be true for the swan. Perhaps VH usage by surviving B cells in swine is a result of their recognition by certain stromal receptors that recognize epitopes expressed in FPRs due to the conformation imposed by the VH genes they use. Perhaps this also explains the data of Pozisil et al. (56) in rabbits with a mutated VH11 gene.

Finally, our data reveal that porcine IgG3, regarded as the ancestral IgG of swine (41) and which comprise 60% of all IgG expression in the fetal IPP (25), is most 3′ proximal to C6. Because speciation preceded subclass diversification, IgG3 in swine, human, and mouse are not homologous genes (1, 41). Coincidentally, the Cγ gene called IgG3 in all three species is the most 5′ Cγ gene and appears to be involved in early natural Ab responses to bacteria such as those expressed by marginal zone B cells. In swine, the dominant expression of IgG3 in the fetal IPP of piglets would be consistent with this concept because the IPP is considered the site of Ag-independent repertoire development (29).

Our data argue that there is a practical reason to understand how the IGH locus is organized and how the Ab repertoire is developed, and, in this case, to expose the Achilles’ heel of a species that supplies 60% of the world’s meat supply. Although a mutation of VH1 in the Alicia still allows rabbits a second chance, probably by gene conversion, a mutation of JH5 in pigs destroys the entire Ab system of this species, as has been experimentally demonstrated (M. Mendicino et al., personal communication). Thus, JH5 of swine represents a deleterious genetic target of spontaneous or designed mutation.

Disclosures

The authors have no financial conflicts of interest.

References


Corrections


The fourth author’s name was published incorrectly. The correct spelling is Francois Puimi.

www.jimmunol.org/cgi/doi/10.4049/jimmunol.1290085