Chapter 5
Organization and Expression of Immunoglobulin Genes

Genetic Models
• How to account for:
  – 1) Vast diversity of antibody specificities
  – 2) Presence of Variable regions at the amino end of Heavy and Light chains, and a Constant region at the carboxyl end
  – 3) Existence of isotypes (different Heavy chains) with same antigenic specificity

Models to Explain Antibody Diversity
1) The Germ Line Theory: “genome posses a large repertoire of antibody genes to account for all the diversity”
2) The Somatic Variation Theory: “genome posses a relatively small number of antibody genes and diversity is generated by mutation and recombination of these genes during somatic development”

The two-gene model
• Developed by Dreyer and Bennet in 1965
• Two separate genes code for the Heavy and Light chains. One codes for the V region and the other for the C region
• These genes come together during at the DNA level to form a continuous message
• There are thousands of V genes in germ line but only one gene for the C region
Three genetic loci encode immunoglobulin molecules:

- Two loci encoding the light chains
  - kappa locus
  - lambda locus
- One locus encoding the heavy chain

These three loci are located on different chromosomes.

### Table 5-1: Chromosomal Locations of Immunoglobulin Genes in Human and Mouse

<table>
<thead>
<tr>
<th>Gene</th>
<th>Human</th>
<th>Mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa chain</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Lambda</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

**Multigene Families**

- **Light Chains**: V, J and C gene segments.
- **Lambda**: Humans (30V, 4J and 7C genes)
- **Kappa**: Humans (40V, 5J and 1C genes)
- **Heavy Chains**: V, D, J and C gene segments
- **Heavy Chains**: Humans (50V, 25D, 6J and 8 C genes)

**Number of functional gene segments in human immunoglobulin loci**

<table>
<thead>
<tr>
<th>Segment</th>
<th>Light chains</th>
<th>Heavy chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable (V)</td>
<td>40 30 40</td>
<td>40</td>
</tr>
<tr>
<td>Diversity (D)</td>
<td>0 0 25</td>
<td>25</td>
</tr>
<tr>
<td>Joining (J)</td>
<td>5 4 6</td>
<td>6</td>
</tr>
</tbody>
</table>

The loci encoding immunoglobulins have a unique structure.
- composed of “gene segments”
- The heavy chain locus has multiple V (variable) segments, multiple D (diversity) segments, multiple J (joining) segments and multiple C (constant) segments.

During maturation, one of each V, D and J segment is randomly “chosen” and used to encode the final antibody molecule.

Germline configuration of the heavy chain locus (mice)

The kappa locus has a similar structure - BUT - does not have D segments.

A kappa chain is encoded by one V segment, one J segment and one C segment.

**Kappa locus**

\[
\begin{align*}
V_k^1 & \quad V_k^2 \quad V_k^n \quad J_k \quad C_k
\end{align*}
\]

**Lambda locus**

\[
\begin{align*}
V_\lambda^2 & \quad J_\lambda^2 \quad C_\lambda^2 \quad J_\lambda^4 \quad C_\lambda^4 \quad V_\lambda^1 \quad J_\lambda^3 \quad C_\lambda^3 \quad J_\lambda^1 \quad C_\lambda^1
\end{align*}
\]

- In heavy chains, the V, D and J segments encode the variable domain while the C segment encodes the constant domain.
- In light chains, the V and J segments encode the variable domain while the C segment encodes the constant domain.
The kappa and lambda loci undergo similar rearrangement. Since there are no D segments, there is a single V→J rearrangement.

The final light chain mRNA contains one VJC unit.

**Recombinant Signal Sequences (RSS)**
- Nucleotide sequence of RSSs:
  - Kappa: CACAGTG
  - Lambda: GTGTCAC
  - Heptamer: CACAGTG
  - Nonamer: GTGTCAC

**RSS** – At 3’ of V genes, 5’ of J genes and at both sides in D genes

**Rule:** 12 (1 turn) or 23 (two turn) base pairs with conserved flanking heptamer and nonamer

* Only 1 to 2 turn
Rearrangement of gene segments is mediated by the RAG1/RAG2 enzyme complex (recombinases).
- The RAG1/RAG2 complex recognizes the heptamer/nonamer sequences and cuts one strand of the DNA.

A hairpin forms...

Step 2: The 5' end of the cut strand reassociates with the second strand resulting in a double-stranded break and synthesis in the 3' direction.

The hairpin is cut at a random site...

Step 3: The heptamer sequences are ligated. An endonuclease cleaves the hairpin at a random site.

Palindromic sequences may form...

Step 4: Endonuclease cleavage may result in short palindromes – additional nucleotides resulting from this are known as P-nucleotides.

Terminal deoxynucleotidyl transferase (Tdt)

An enzyme that randomly adds in nucleotides during joining of coding gene segments.

Step 5: Tdt adds P-nucleotides randomly to the single stranded end.
The join is repaired...

Step 6: The two single-stranded ends pair. Unpaired nucleotides are trimmed by an exonuclease and the coding joint is repaired.

Note: Keep in mind that this random rearrangement can lead to PRODUCTIVE and NON-PRODUCTIVE gene rearrangements.

**Generation of antibody diversity**

1. Multiple germine V, D and J gene segments
2. Combinatorial V-J and V-D-J joining
3. Somatic hypermutation
4. Junctional flexibility
5. P-nucleotide addition
6. N-nucleotide addition
7. Combinatorial association of heavy and light chains

**Combinatorial V-J and V-D-J joining**

- Humans:
  - Heavy Chain: V (51), D (27), J (6) = 8262
  - Light Chain: Kappa – V (40), J (5) = 200
    Lambda – V(30), J (4) = 120

\[ 8262 \times (200 \times 120) = 2.64 \times 10^6 \]

**Junctional flexibility**

- Generated through V, D and J combinations
- Joining of Recombination Signal Sequences = **Signal Joint**
- Joining of Coding Sequences = **Coding Joint**
- Signal Joints ALWAYS joined precisely, but joining of Coding Joints is IMPRECISE
- Good = Antibody diversity
- BAD = Non-productive rearrangements
### P-nucleotide addition
- Cleavage of the Hairpin at the end of the coding sequence by endonuclease (Artemis) is random
- Generates a short single strand of nucleotides at the end of the Coding sequence
- Addition of complementary nucleotides to this strand forms a palindrome sequence (P nucleotides)

### N-nucleotide addition
- Once complementary nucleotides to this strand have been added to form a palindrome sequence (P nucleotides)
- The enzyme TdT (terminal deoxynucleotidyl transferase) fills the gap with N nucleotides.
- This enzyme can add randomly up to 15 N nucleotides (non-genomic)
- N nucleotides can be added to the D-J and V-DJ in the H chain.
- This mechanisms does not happen in the Light chain

### Table 5-3

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>CDR1</th>
<th>CDR2</th>
<th>CDR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence encoded by:</td>
<td>V segment</td>
<td>V segment</td>
<td>( V_{\delta}, D_{\gamma/\delta} ) junctions</td>
</tr>
<tr>
<td>Junctional flexibility</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>P-nucleotide addition</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>N-nucleotide addition</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Somatic hypermutation</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*Note: N-nucleotide addition occurs only in heavy-chain DNA.
**Somatic Hypermutation**

- Generated **point mutations** in gene segments for variable regions (VDJ and VJ segments)
- Takes place in secondary lymphoid organs (~ 1 week after contact with antigen)
- In mature B cells mutations are clustered in CDRs regions
- **Affinity maturation**—selection process leading to survival of those B cells with high affinity for the antigen

---

**Generation of Diversity**

<table>
<thead>
<tr>
<th>B cell receptor (Immunoglobulin)</th>
<th>Heavy</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>V gene segments</td>
<td>1000</td>
<td>300</td>
</tr>
<tr>
<td>D gene segments</td>
<td>15</td>
<td>-</td>
</tr>
<tr>
<td>J gene segments</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>N region insertion</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>Junctional diversity</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Somatic mutation</td>
<td>??</td>
<td>??</td>
</tr>
</tbody>
</table>

- **V x D x J**: 1000 x 15 x 4
- **V x J**: 300 x 4

Total: $6 \times 10^4$  
Combinatorial association: $7.2 \times 10^7$

- **Ag independent process**
- **Clonal selection**

---

**ALLELIC EXCLUSION:**

- We have two copies (alleles) of each Ig gene - one inherited from our father and one from our mother.
- In most cases, both genes are expressed.
- But Antibody genes are different!!! Only one heavy chain allele and one light chain allele is expressed!!!
- This is termed **allelic exclusion** (one allele is excluded). Once a productive arrangement is made, the other allele is suppressed.
- Why? To ensure that each B cell makes antibody of a single specificity.
Class Switching

- Antigen stimulation of a B cells \(\rightarrow\) Antibodies with same variable Heavy (VDJ) with any \(C_H\) gene segment
- Process dependent on Switch Regions
- Switch Regions (2-3 kb) are located upstream from each \(C_H\) segment, except \(IgD (C_\delta)\)
- Process driven by cytokines:
  - IL-4 \(\rightarrow\) IgM to IgG1 or IgE
  - IFN-\(\gamma\) \(\rightarrow\) IgM to IgG2a
- Players in regulation: 1) switch regions, 2) switch recombinases, 3) cytokine signals

Expression of membrane or secreted Immunoglobulin

- In mature B cells \(\rightarrow\) membrane forms; in Plasma cells \(\rightarrow\) secreted forms
- Process depends on differential processing of primary transcript
- Remember: IgG, IgD, IgA (3 \(C_H\) domains), IgM and IgA (4 \(C_H\) domains).
- Domain 3/4 contains the Secretory (hydrophilic) nucleotide sequence (\(S\)) at its 3'.
- Two Exons at 3' encode the \(M1\) (trans-membrane) and \(M2\) (cytoplasmic) segments.
- Primary transcript contains two PolyA sites: If cleavage at Poly A site 1 = Secreted Form. If cleavage at PolyA site 2 = Membrane Form
In mature B cells: both IgM and IgD, then cleavage must occurs at SITE 4 and both processing pathways MUST occur simultaneously.
- RAG-1/RAG-2 cleave ONE strand of DNA
- This occurs at the border of the RSS heptamer and the coding gene
- The 3' OH group attacks a phosphodiester bond on the other DNA strand
- This results in hairpin DNA strand on the coding region.
- Other enzymes get involved and remove the "junk" and bring together the coding regions

**Junctional Diversity:**
- Terminal deoxynucleotidyl transferase (TdT) is important for creating junctional diversity
- What are SIGNAL JOINTS and CODING JOINTS?
- Hairpin must be opened → the enzyme Artemis
- Cleavage is random and can happened at any site in the hairpin
- Replication results in a short inverted repeat or palindrome = P nucleotides
- TdT can now introduce random nucleotides into the coding joints = N nucleotides
- Keep in mind that all this introduced variability may results in functional and non-functional Ig (or TCR) genes.