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 the large repertoire of antibody genes to account for all the antibody 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, one codes for the V region and the other for the C region
- These genes come together at the DNA level to form a continuous message
- There must be 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>
<thead>
<tr>
<th>TABLE 5-1</th>
<th>Chromosomal locations of immunoglobulin genes in human and mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
<td>Human</td>
</tr>
<tr>
<td>gamma Light chain</td>
<td>22</td>
</tr>
<tr>
<td>kappa Light chain</td>
<td>2</td>
</tr>
<tr>
<td>heavy chain</td>
<td>14</td>
</tr>
</tbody>
</table>

Multigene Families
- 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)

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): (IgG1 – 4)

Gene rearrangement Kappa light Chain
- 1, rearrangement – D\text{\(H\)}7/J\text{\(H\)}3
- 2, rearrangement – D\text{\(H\)}7/J\text{\(H\)}3 and V\text{\(H\)}21
- 3, Post-transcriptional modifications
- 4, mature mRNA
- 5, polypeptide
- 6, Post-transcriptional modifications
- 7, light chain
- 8, Post-transcriptional modifications

Post-transcriptional modifications
- mRNA
- Transcription
- Polyadenylation
- Poly-A tail

Primary RNA transcript
- 1, V\text{\(J\)}3 joining
- 2, Rearranged DNA
- 3, RNA transcript

Post-transcriptional modifications
- 4, mature mRNA
- 6, polypeptide
- 7, light chain

Gene rearrangement Kappa light Chain
- 1, rearrangement – D\text{\(H\)}7/J\text{\(H\)}3
- 2, rearrangement – D\text{\(H\)}7/J\text{\(H\)}3 and V\text{\(H\)}21

Multigene Families
- 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)

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): (IgG1 – 4)

Gene rearrangement Kappa light Chain
- 1, rearrangement – D\text{\(H\)}7/J\text{\(H\)}3
- 2, rearrangement – D\text{\(H\)}7/J\text{\(H\)}3 and V\text{\(H\)}21

Multigene Families
- 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)
What mechanism ensures correct joining of gene segments during rearrangement of the heavy and light chain loci?

- Recombination signal sequences (RSS) - conserved sequences in regions just upstream or downstream of gene segments.

- Consist of a conserved heptamer (green) and nonamer (orange) with a 12 or 23 bp spacer.

- The one-turn (red)/two-turn rule (blue) - (12/23 rule) - recombination occurs only between a segment with a 12 bp spacer and a 23 bp spacer.

Recombinant Signal Sequences (RSS)

(a) Nucleotide sequence of RSS

<table>
<thead>
<tr>
<th>RSS Segment</th>
<th>Heptamer</th>
<th>Nonamer</th>
</tr>
</thead>
<tbody>
<tr>
<td>V genes</td>
<td>CACACTG</td>
<td>ACAAAAAAC</td>
</tr>
<tr>
<td>D genes</td>
<td>CGTTCAC</td>
<td>TGTTTTTGG</td>
</tr>
<tr>
<td>J genes</td>
<td>CCAAAACA</td>
<td>GTGACAC</td>
</tr>
</tbody>
</table>

- 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

A hairpin forms...

Step 1: The RAG1/RAG2 complex recognizes the heptamer/nonamer sequences in RSS and cuts one strand of the DNA.

Step 2: The 5' end of the cut strand reacts with the second strand resulting in a double-stranded break and hairpin formation

- Rearrangement of gene segments is mediated by the RAG1/RAG2 enzyme complex (V(D)J recombinases).

- The RAG1/RAG2 complex recognizes the heptamer/nonamer sequences in RSS and cuts one strand of the DNA.
The hairpin is cut at a random site by Artemis...

Signal Joint

Coding Joint

Endonuclease (Artemis)

Signal Joint

Coding Joint

Endonuclease (Artemis)

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. Only in H chain.

Non-genomic nucleotides!!

Step 5: Tdt adds N-nucleotides randomly to the single stranded ends.

P-nucleotides

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

P nucleotides

Generation of antibody diversity

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

- **Humans**:
  - Heavy Chain: V (51), D (27), J (6) = 8262
  - Light Chain: Kappa – V (40), J (5) = 200
    Lambda – V(30), J (4) = 120
  
  \[8262 \times (200 \times 120) = 2.64 \times 10^6\]

Where does other diversity comes from?

3. Junctional flexibility (diversity)

- Generated through V, D and J combinations
- Joining of Recombination Signal Sequences (RSS) = Signal Joint (heptamer + heptamer)
- 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

4. 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)

5. 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 (and to the V-J in the L chain)
- Only in H chain
6. Somatic Hypermutation

- Generated **point mutations** (substitutions) in gene segments for variable regions (VDJ and VJ segments)
- Takes place in **secondary lymphoid organs** (~ 1 week after contact with antigen) – Germinal centers (~ 1 per 1000 bp per cell division)
- In mature B cells mutations are clustered in CDRs regions
- **Somatic hypermutation leads to Affinity maturation**- selection process leading to survival of those B cells with high affinity for the antigen

**TABLE 5-3** Sources of sequence variation in complementary-determining regions of immunoglobulin heavy- and light-chain genes

<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_{	ext{J, JJ}} junction; V_{	ext{D, DJ}} 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>

*N-nucleotide addition occurs only in heavy-chain DNA.*

**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.

**Generation of antibody diversity**

1. Multiple germline 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

\[
2.64 \times 10^6 \rightarrow 7.2 \times 10^6 \text{ variabilities}!!!
\]
**Class Switching**

- Antigen stimulation of a B cells → 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_δ)
- Process driven by cytokines:
  - IL-4 → IgM to IgG1 or IgE
  - IFN-γ → IgM to IgG2a
- Players in regulation: 1) switch regions, 2) switch recombinases, 3) cytokine signals

**AID Enzyme**

- Activation induced cytidine deaminase
- RNA editing enzyme
- Deamination of cytosine → uracyl → repair induces base modifications!!!
- Mediates SOMATIC HYPERMUTATION, GENE CONVERSION, and CLASS switching recombination
Expression of membrane or secreted Immunoglobulin

- In mature B cells → membrane forms; in Plasma cells → secreted forms
- Process depends on differential processing of primary transcript
- Remember: IgG, IgD, IgA (3 C_H domains), IgM and IgE (4 C_H domains).
- Domain 3/4 contains the Secretory (hydrophilic) nucleotide sequence (S) at its carboxyl end.
- 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 occur at SITE 4 and both processing pathways MUST occur simultaneously
Synthesis, Assembly and Secretion

- Plasma cells produce and secrete 1000 Ab/sec!
- Assembly in the RER

WHERE TO GO?
- Membrane-bound → hydrophobic sequence into the membrane
- Secreted → hydrophilic sequence, no trans-membrane component

-H and L chains are made in different polyribosomes in RER
- No fusion of Ab molecule with secretory molecules

Quality Control in RER

- BiP – Immunoglobulin Heavy Chain Binding Protein
- Destroy incomplete Abs
The End