Lecture 5

Organization and Expression of Immunoglobulin Genes

Although we are always taught about astonishing specificity in biology, many mechanisms are only as specific as they need to be. - Gerhart, John, and Marc Kirschner, 1997

I. Uniqueness of the System

A. The Problem -
   1. The number of possible antibodies that an individual can produce is vast.
      a. Vast is a good term here. Technically there can't be an infinite number.
      b. Estimates range from $10^8$ to $10^{11}$, but frankly no one really knows.
   2. Any one antibody-producing plasma cell synthesizes antibodies with one and only one CDR (complementarity determining region or antigen binding region).
   3. Any one antibody-producing plasma cell can make more than one class of antibodies, each with the same CDR.

B. The Solution
   1. Mixing instructions for chains (heavy and light peptides).

C. THIS IS A VERY BIG DEAL.
   1. Other Systems with Changes to the Germline DNA.
      a. Chromosome Diminution. This is a loss that accompanies the decision to become a somatic as opposed to a germ cell, and does not seem to be involved in the regulation of genes involving differentiated state.
      b. Gene amplification. There are a number of systems in which cells make extra copies of particular genes (rRNA for example).
      c. McClintock's jumping genes. She thought they were important in developmental regulation, but they turned out just to be examples of DNA
parasitism.

d. Neuron development. In mice, at least, as neurons in certain regions of the brain develop, they relax their controls on transposable elements. Such elements then move around randomly in the nuclei.

2. What Specifically Makes the Adaptive Immune System Unique

a. Differentiation of B (and T) cells involves clipping DNA out of specific regions of the immunoglobulin genes.

b. The clipping is not precise within those regions.

c. The clipping takes place at different parts of the regions in different cells.

d. The clipping does not take place in regions other than those for immune proteins.

e. The clipping results in different cells (and their progeny) ultimately having different versions of the immunoglobulin genes.

f. These site-specific DNA rearrangements are unique to the vertebrate adaptive immune system.

II. Gene to RNA to Protein

A. Orientation: I’m going to show the human version. Many texts show versions from the mouse, which differ from the human. *(figure 1)*

1. There are two different genes for the light chain, and one for the heavy, all on different chromosomes.

2. Since we are diploid, there is a pair of each chromosome and therefore each of the following arrangements may occur twice in the genome.
3. Many of the elements of these genes are repeated variations of interchangeable parts.

4. When DNA elements occurs as tandem repeats, individual often very in the number of these repeats.

5. In this lecture, I refer to the Ig genes as genes and the subparts of the genes are gene regions.

6. Do not confuse the terms "region" with "exon."

B. Lambda (λ) Gene Expression (figure 2)


2. One leader-variable and one J-C pair come together at random.

3. This activates the promoter of the selected leader only, enabling transcription from the beginning of that leader only.

4. RNA polymerase transcribes the VL and JC into a primary transcript. (figure 3)

5. Message processing removes the introns from between the V-L and the J-C, adds a poly A tail and 5’ cap (not shown).

6. The ribosome attaches to the message, begins translating, attaches to the RER, and pushes the nascent peptide into the ER lumen.

7. Enzymes clip off the leader, leaving a light chain with a variable domain and a constant domain.

C. Kappa (κ) Gene Expression (figure 4)

1. Gene family in humans includes a series of about 40 Vκ (leader – variable) regions, 5 functional Jκ (joining) regions, and Cκ one constant region.

2. Gene rearrangement places 1 VL next to 1 J gene region, again activating only the promoter of the selected VL.

3. RNA polymerase transcribes a message precursor with one VL, the selected J, the constant region and any remaining Js and introns between them.

4. During processing, introns and extra Js get clipped out, leaving a message with the same structure as that of the lambda.

5. Translation proceeds as above, producing a light chain with the same overall structure.
D. Heavy Gene Expression  (*figure 5*)

1. As with the kappa light, family begins with sequence of about 40 V-Ls.

2. Next is a series of about 20 short D (diversity) segments, each coding for 3 amino acids.

3. In humans, 5 or 6 J regions follow.

4. Finally there is a series of constant regions, 1μ, 1δ, 4 different γs, 1 ε, and 2 αs. The order is μ, δ, γ3, γ1, α1, γ2, γ4, ε, and α2.

5. First the a D regions joins with a J, cutting out all the extra downstream Ds and upstream Js between them, but leaving any downstream J's. (*figure 6*)

6. Then one of the VLs joins with a D, removing all the extra downstream VLs and upstream Ds. (*figure 7*)

7. The initial primary transcript starts with this LVDJ regions, continues through any remaining J's and introns, and then copies through the μ and δ constant regions and stops. (*figure 8*)

8. The transcript now undergoes alternative message splicing to include either the μ or the δ constant exons, but not both. (*figure 9*)

9. Translation proceeds as with the other messages.

III. Rearrangement in Developmental Context – in the bone marrow.

A. Variable Region Rearrangements - (*figure 10*)

1. First, the developing B cell rearranges a heavy chain gene. If you get a functioning gene, fine, you express it and shut down the other heavy chain gene.

2. If not your try the other heavy chain gene, if that works, you proceed to the light gene. If not the cell undergoes apoptosis.

3. First you rearrange one kappa gene and then, if that does work, the other, again turning off the unused genes.

4. If neither works, you proceed to the lambda, first one then the other.

5. Only if all four genes prove to be duds does the cell apoptose.
B. Comments on the Final Peptides

1. Whether you start with a lambda or kappa gene, you wind up coding for peptides that superficially look the same.

2. The peptides for these genes NEVER have membrane-spanning regions.

3. Once you have decided on a heavy chain gene the RNA from this gene can be alternatively spliced into four **different** peptides.

<table>
<thead>
<tr>
<th></th>
<th>μ (M class) with M1 and M2</th>
<th>δ (D class) with M1 and M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ (M class)</td>
<td>soluble</td>
<td>δ (D class) soluble</td>
</tr>
</tbody>
</table>

a. Initially, a developing cell only makes the μ version with membrane-spanning exons.

b. Upon maturity, the cell now makes both μ and δ peptides in the same cell.

c. Once stimulated to produce antibody, the cell makes primarily μ peptides, but without the membrane spanning region. Thus they will make soluble M-class antibodies.

d. A cell rarely makes δ peptides without the membrane-spanning regions.

IV. Mechanism Details: Breaking and Joining (*figure 11*)

A. Signal Structure

1. Recombination signal sequences (RSSs) flank the V, D, and J segments
   a. 3’ end of V
   b. both sides of the D
   c. 5’ end of the J

2. Each RSS has (*figure 12*)
   a. 7 nucleotide palindrome (reads the same forward and back). This region will participate in the ligation of adjacent V(D)J segments.
   b. spacer, seems to be important in lining up the adjacent V(D)J regions so that they join properly:
      i. κ 12 base (1 turn of helix - V) or 23 base (2 turns of helix - J)
      ii. λ 23 base (2 turns of helix - V) or 12 base (1 turn of helix - J)
      iii. heavy chain, V and J, both are two turn, D is one turn
c. 9 nucleotides rich in AT – aid in attachment to the multienzyme complex that performs the whole complex function. (figure 13)

3. Signal sequences with one-turn spacers can only join with those with two-turn spacers, so this protects against misjoining.

4. The enzyme complex responsible for joining is called the V(D)J recombinase.

B. Mechanism of Rearrangement

1. The genes to be arranged moved to and specific location in the nucleus and open up, detaching form the nucleosomes.

2. First the processing complex grabs one of the RSS signals, than it grabs the complement.

3. One-turn signal juxtaposes to two-turn signal.

4. One strand of the DNA between the coding and signal sequence cleaves.

5. The 5'OH end of the cut strand attacks the opposite side of the uncut strand of the same DNA molecule.

6. This produces a hairpin loop at the downstream end of the V and the upstream end of the J (light chains).

   a. The nucleotides within the loop were originally part of the palindromic sequences.

   b. The one- and two-turn sequences, along with the AT region signal end in a flush cut with a 5’ phosphorylated end.

7. Enzymes clip the hairpin loop. Now the gene regions end in a double strand of DNA with an open end, the clip site used to join the V(D)J and add variability at the junction.

   a. A few nucleotides get trimmed off

   b. The ends of two of the single stands are ligated.

   c. Nucleotides are added to fill in the unmatched singled stranded regions (P-nucleotide addition), matching the ends generated by the cut.

   d. For heavy chains only, up to 15 additional nucleotides can be added at random in the junction
8. Repair and ligation of the coding joint, with release and digestion of the signal sequences (which are retained if there is an inversional joining.)

9. At a later stage of development, the parts of the genes coding for the hypervariable region can mutate, introducing further variation (much more later).

C. Commitment to Producing a Single Functional CDR (Antigen Binding Region)

1. Because the joining process between V\textsubscript{L} and J\textsubscript{L} and V\textsubscript{H}, D\textsubscript{H}, and J\textsubscript{H} regions contains so many random elements, about 2/3 of the time, any one junction will produce a frame shift.

2. If any one junction in any part of the gene gets out of phase, this produces a non-productive rearrangement.

3. If the message remains in phase, this usually produces a productive rearrangement and the message for a working peptide. N and P addition can also throw in random stop codons, even if the message as a whole remains in frame.

4. Recall that these cells are diploid and that these rearrangements will take place on both homologous chromosomes.

5. However you must get both a good heavy and a good light of some kind or the cell dies by apoptosis. Only about 8% make it through this.

6. Allelic Exclusion - you only get one good productively rearranged gene.
   a. The heavy rearranges first. Once there's a good µ heavy chain, this shuts off rearrangement of the other heavy gene.
   b. The light rearranges next. First it tries a κ, and if one works, the cell shuts down rearrangement.
   c. If the κ doesn't work, it tries a λ. Once one of them works, the cell shuts down rearrangement.
   d. If nothing works, the cell apoptoses.

1. Once a cell has a working heavy gene and a working light, then it shuts off its recombination enzyme genes.

D. Dangers and Mechanisms (figure 14)

V. Reviewing the Generation of Antibody Diversity.

A. Mixing and Matching Germ Line Genes
1. Recall that both mice and humans have multiple $V_H$, $D_H$, $J_H$, $V_K$, $J_K$, $V_\lambda$, and $J_\lambda$ regions on homologous chromosomes.

2. You can put any heavy chain with any light.

3. So just from recombinatorial options in the germ line, there's a lot of potential variability.

B. Fooling Around at the Junction (figure 15)

1. Coding junctions do not always join precisely.

2. P-nucleotide addition. Generates a palindrome when it matches the nucleotides opened at the clipping of the hairpin loop.

3. N-nucleotides- those optionally inserted in the middle of this.

4. The region of the protein coded by the junction winds up in the CDR loops of the variable region.

C. Somatic Hypermutation (figures 16-17)

1. The pre-B cell leaves the bone marrow able to make one and only one kind of CDR (although it can add this to the different classes of antibody heavy C chains).

2. However, it can still change or mutate the region coding for the hypervariable loops.

3. This occurs later in the process of antibody stimulation, maturation, and selection outside the bone marrow.

VI. Class Switching

A. Where and When

1. Class switching typically in the lymph nodes after exposure to antigen.

2. Class switching takes place after the system has been producing antibodies for a week or more.

3. Thus the first antibodies produced in response to an infection are Class M and you don’t start to see G (or other classes) until later in the infection.

B. How

1. Class switching involves further rearrangements to the DNA, but these are brought about by a separate set of enzymes.

2. The enzymes are induced by signals from the T_H cells.

3. Recombination sites are called switch regions, and are designated by 2-3 kb sequences of DNA with multiple copies of short consensus sequences.

4. Before switching, the cell expresses the $\mu$ and $\delta$ constant regions.
5. Switching involves removing these and whatever other constant gene region(s) stand(s) between the VDJ recombined region and the constant to be expressed.

These may be removed in sequence as a cell class-switches down its options.

VII. Synthesis and Secretion

A. A Farewell Look at the Complexity of the Heavy Chain Gene (figure 18)

1. Our first look showed the different gene regions, but not the internal structure of the constant region or most of the signals.

2. If we enlarge a portion of the rearranged gene showing just one upstream VL and the downstream region into the first gamma exon, we can show more detail.

3. At a first pass you can see that the turquoise lines indicate splice signals, which will allow introns to be processed from the message.

4. Focus on the part of the gene that codes for the variable domain, which is flanked by an unused VL and an unused J, both of which still have the complete RSS.

5. Locate the functional parts of the instructions for the variable domain:
   a. The activated promoter
   b. The VDJ in use
   c. The rearrangement joints with the location of the N-nucleotide addition

6. Now focus on the constant chain instructions

7. The C\(_\mu\) and C\(_\delta\) domains are coded for by separate exons for each constant domain with intervening introns.
   a. The sequence is preceded by a switch signal.
   b. Thus there are four C\(_\mu\) exons and three C\(_\delta\), the D class having the hinge regions and therefore greater flexibility.
   c. Both complex regions have two more exons (M1 and 2) downstream from the constant exons.
   d. Both C\(_\mu\) and C\(_\delta\) have a polyA splicing option in both the last Ig exon and the second membrane spanning exon.
   e. This explains the four possible option (M or D antibody, M or D Ig receptor) that a message from this region could specify.

8. Beware of confusing terms! Class M refers to the whole antibody, IgM, whether soluble of membrane bound. This class has the \(\mu\) constant region. If it is membrane-
bound then the μ has M 1 and 2 at the end.

9. This drawing does not show most of the rest of the constant instruction, but you can see the switch site and first exon for the γ3 gene region.

10. All other gene regions have the two membrane-spanning exons, expressed only in memory cells.
B. Processing in the GERL (Golgi-ER-Lysosome)

1. L and H transcripts exit the nucleus separately, attach to ribosomes, and begin peptide synthesis.

2. The initial signal sequence (leader) causes the ribosome to attach to the RER and the insert the L and H peptide separately into the lumen.

3. Only if the message for the H chain ends on a membrane-spanning region, the peptide will remain anchored in the membrane.

4. In the RER lumen, enzymes clip off the leader and begin to add oligosaccharides to the peptides.

5. Assembly and processing occurs in the primarily in the lumen of the RER and the product is then sent to the Golgi.

6. First the heavy chains are put together, then the Ls are added (for the G class first one heavy and one light associate).

7. After assembly, enzymes oxidize the disulfide bonds, nails the structure into position, and adjust the oligosaccharide into the version characteristic of the antibody.

Resources:

http://bcs.whfreeman.com/immunology6e/content/cat_070/Stanford%20VIBE/index.html