GENETIC BASIS OF ANTIBODY DIVERSITY

OBJECTIVES: When you finish this section, you should be able to:

1. Define the following terms: allelic exclusion, isotype switching, affinity maturation, antibody repertoire, alternative RNA splicing, recombination signal sequence.

2. Describe the genes that encode Ig Heavy and Light chains.

3. Describe the sequence of Ig gene rearrangement that occurs during B cell differentiation.

4. Discuss how diversity in antibody specificity is achieved.

5. Discuss the mechanisms of heavy chain class switching.

5. Calculate the number of possible Igs which can be produced from a given number of V, J, D, and C genes.

GENETIC BASIS OF ANTIBODY DIVERSITY

- Estimates of antibody specificities in an individual range between 10^6 - 10^8 .
- If 1 gene encoded 1 immunoglobulin of a given specificity, approximately 10^6 - 10^8 genes would be required to JUST encode antibodies.
- This is impossible because the entire human genome is made up of 3.2×10^9 base pairs.
- Average protein = 107 amino acids = $107 \times 3 = 3.2 \times 10^2$ base pairs
- If 1 gene encodes 1 protein, the total # of genes will be equivalent to

$$\frac{3.2 \times 10^9}{3.2 \times 10^2} = 1 \times 10^7 \text{ genes}$$

- However, estimates are that only 31,000 to 35,000 genes in the human genome actually encode proteins!
- This number of genes is unable to encode ALL antibody specificities!

DILEMMA

- Since only 31-35 thousand genes in the human genome actually encode proteins,
- How then is antibody diversity (between 1 million to 100 million specificities) achieved with such a limited number of genes?

SOME TERMINOLOGY FIRST

GENOME: Total inherited DNA - consists of linear array of genes. All somatic (diploid) cells have the same complement of inherited genes. However, certain genes are expressed in some cells but not in others (e.g. only B lymphocytes express immunoglobulin genes)

EXON: Coding region of a gene (transcribed into mRNA and translated into protein).

INTRON: Non-coding region of a gene (transcribed into RNA but not translated into protein). Introns are removed from RNA transcript by **SPLICING**.

LEADER sequence: Exon preceding a gene encoding a protein to be secreted. Approximately 60-90 bp and encodes 20-30 amino acids. Hydrophobic and used to transport nascent protein into endoplasmic reticulum

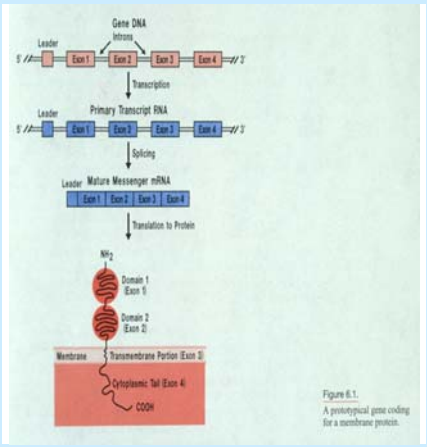


Figure 6.1.
A prototypical gene coding for a membrane protein.

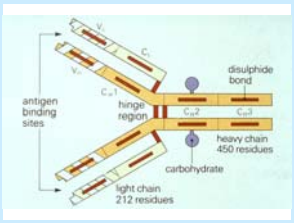
Organization of Immunoglobulin Genes

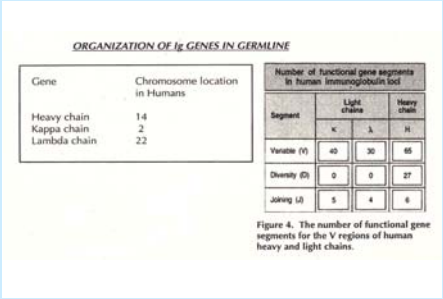
- Numerous V region genes are preceded by Leader or signal sequences (60-90 bp) exons interspersed with introns.

- Heavy chain contains **V** (Variable), **D** (Diversity), **J** (Joining) and **C** (Constant) region gene segments.
  - **V - D - J - C**

- Light chain contains **V**, **J**, and **C** region gene segments
  - **V - J - C**

- Constant region genes are sub-divided into exons encoding domains
  (CH_1,CH_2, CH_3, CH_4)





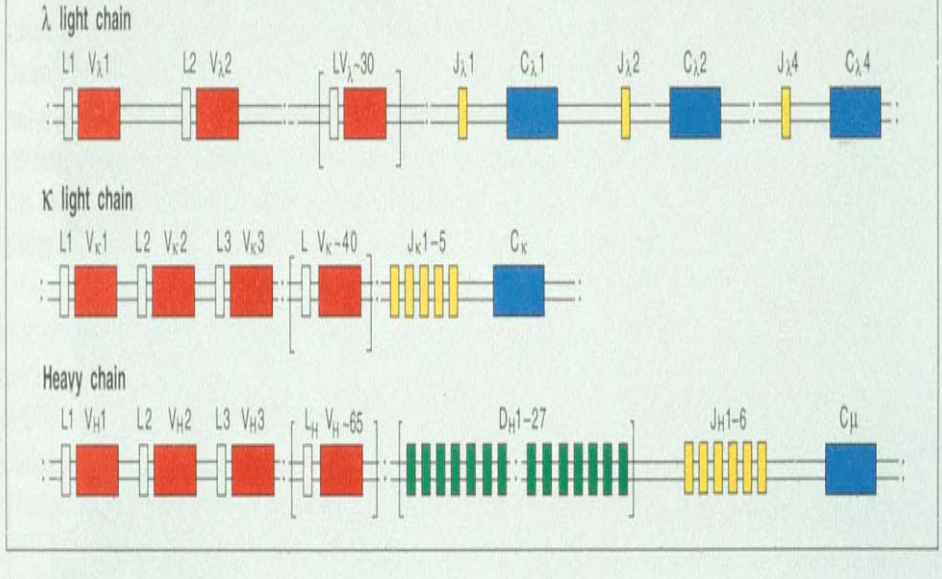
ORGANIZATION OF Ig GENES IN GERMLINE

Gene	Chromosome location in Humans
Heavy chain	14
Kappa chain	2
Lambda chain	22

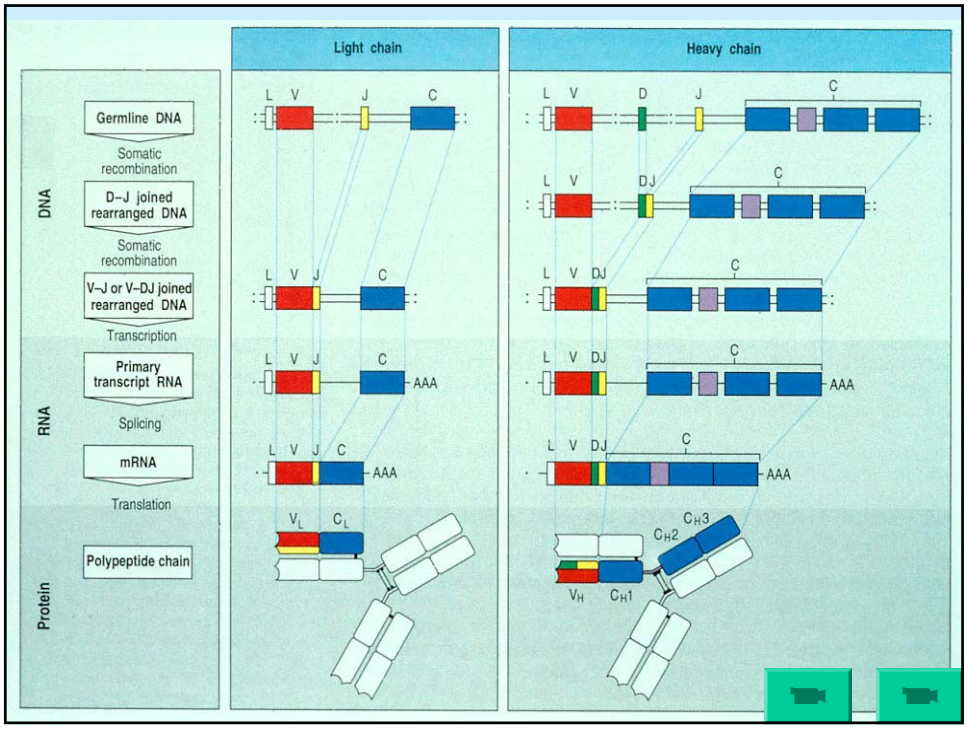
Number of functional gene segments in human immunoglobulin loci

Segment	Light chains κ	Light chains λ	Heavy chain H
Variable (V)	40	30	65
Diversity (D)	0	0	27
Joining (J)	5	4	6

Figure 4. The number of functional gene segments for the V regions of human heavy and light chains.



Genomic organization of the heavy and light chain gene segments in humans



CHARACTERISTICS OF IMMUNOGLOBULIN GENE RE-ARRANGEMENT

1. **Involves Allelic Exclusion.**
   - Only one of two parental alleles of Ig is expressed in a B cell.
   - Either kappa or lambda light chain is expressed by a B cell (light chain isotype exclusion).

2. **Ig rearrangement occurs prior to antigen exposure.**

• **A. Heavy chain re-arrangement**
   - Re-arrangement occurs in a precise order:
   - Heavy chain re-arranges before Light chain.
   - D-J joining occurs first to form DJ and is followed by V-DJ joining to form VDJ.
   - Production of μ heavy chain by re-arrangement of one allele inhibits re-arrangement on other allele.
     - If re-arrangement on first allele is non-productive (due to mutations, deletions or frame shifts that generate stop codons), then re-arrangement on the second allele is stimulated.
     - Therefore, in any antibody-producing B cell, one allele is productively re-arranged and the other is either not re-arranged (in germ line configuration) or is aberrantly re-arranged.

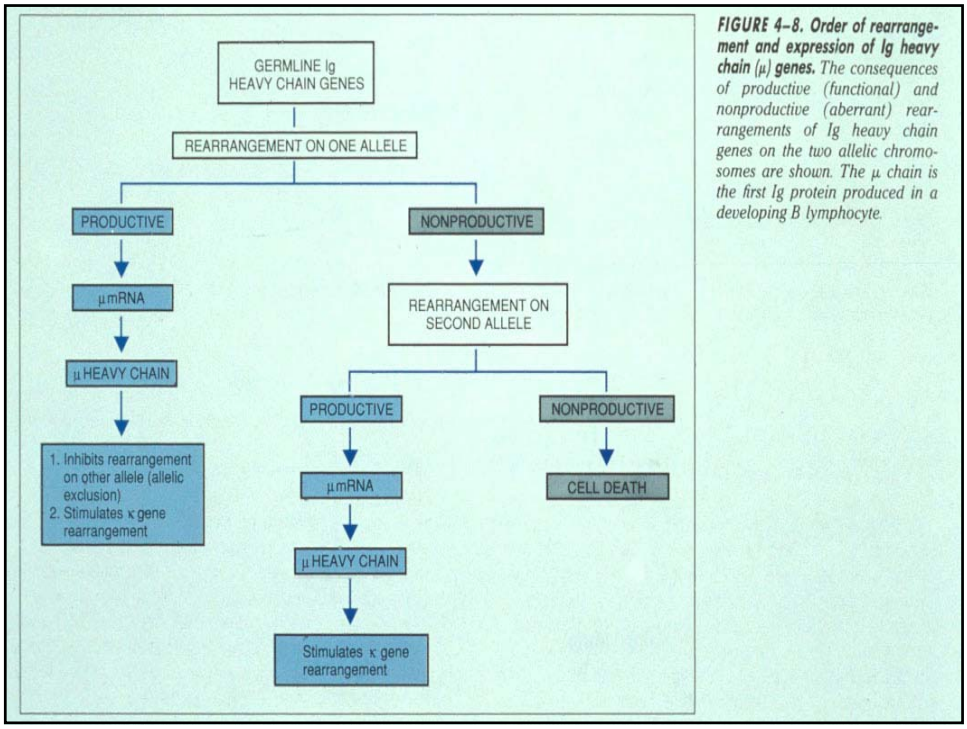


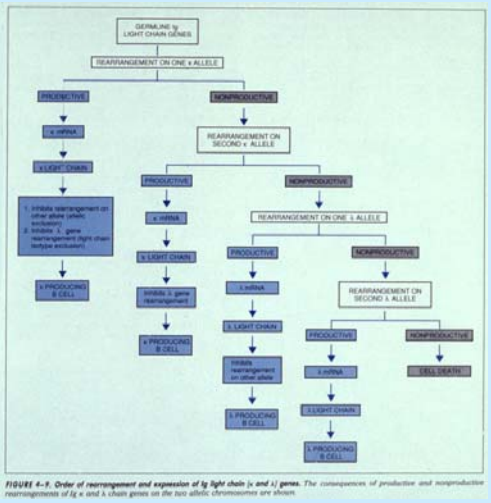
FIGURE 4–8. Order of rearrangement and expression of Ig heavy chain (μ) genes. *The consequences of productive (functional) and nonproductive (aberrant) rearrangements of Ig heavy chain genes on the two allelic chromosomes are shown. The μ chain is the first Ig protein produced in a developing B lymphocyte.*

B. Light chain re-arrangement

i.Kappa chain (κ) rearranges before lambda (λ) chain V-joining occurs.

ii. Productive arrangement on one allele blocks re-arrangement on other allele.

iii. If kappa protein is produced, re-arrangement of lambda chain is blocked. Otherwise lambda chain undergoes re-arrangement.



MECHANISM OF IMMUNOGLOBULIN RE-ARRANGEMENT

- Occurs principally via looping out (excision) of intervening gene sequences followed by ligation of Ig gene segments.
- Controlled by recombination signal sequences (RSS) located at joining sites.
- Consist of heptamer/nonamer (7/9) sequences interspersed by 12/23 base pair spacers.
- Recognized by Recombinases (enzymes with endonuclease and ligase activities). Consists of RAG1,2 proteins (lymphocyte-specific, and non-lymphocyte-specific DNA repair proteins (DNA ligase IV, DNA-dependent protein kinase (DNA-PK) and Ku, a protein that associates with DNA-PK
- Genes encoding recombinases are present in all cell types but are expressed only in lymphoid (B &T) cells.
- Recombination activating genes 1 and 2 (RAG-1, RAG-2) have been identified which stimulate Ig gene rearrangement. Have endonuclease activity.

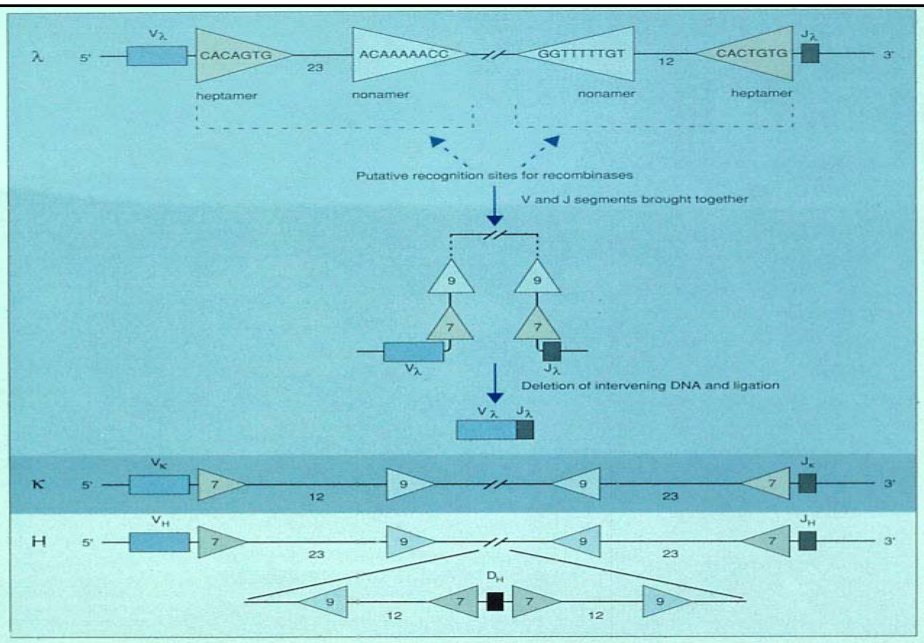
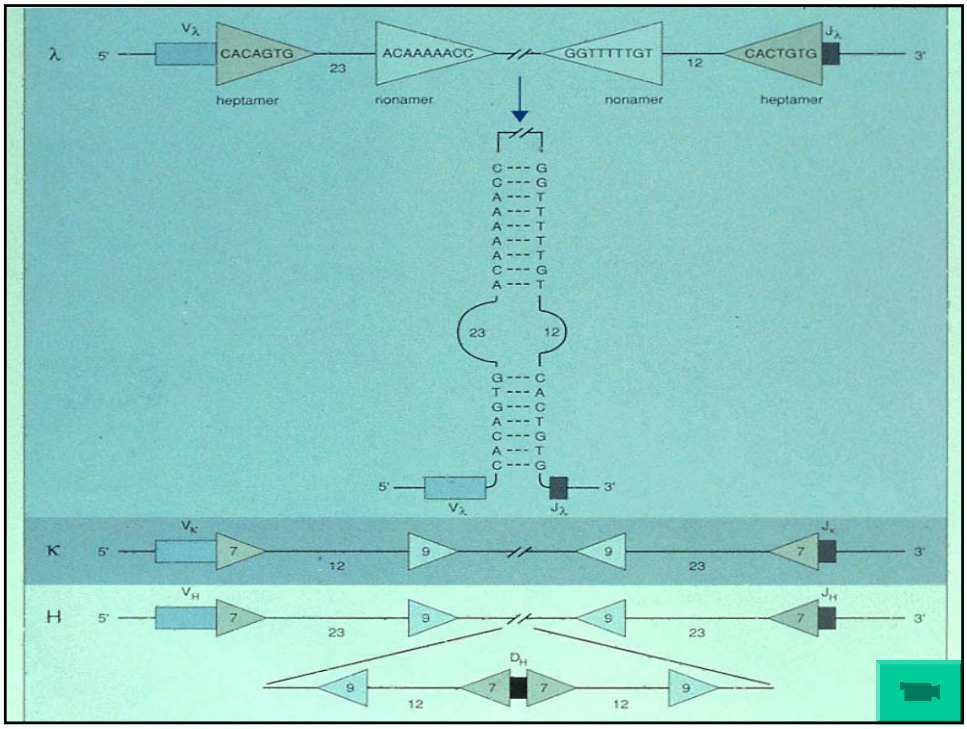


FIGURE 4–10. DNA recognition sequences for recombinases that mediate Ig gene rearrangement. *Conserved heptamer (7 bp) and nonamer (9 bp) sequences, separated by 12 or 23 base pair spacers, are located adjacent to V and J exons (for κ and λ loci), or V, D, and J exons (in the H chain locus). Recombinases presumably recognize these regions and bring the exons together, forming loops of non-coding intervening DNA that are excised and the exons are joined. Alternatively, the exons can be joined by a process of inversion, followed by excision and ligation. (Adapted with permission from Tonegawa, S. Somatic generation of antibody diversity. Nature 302:575–581, 1983. Copyright © 1983, Macmillan Magazines Ltd.)*



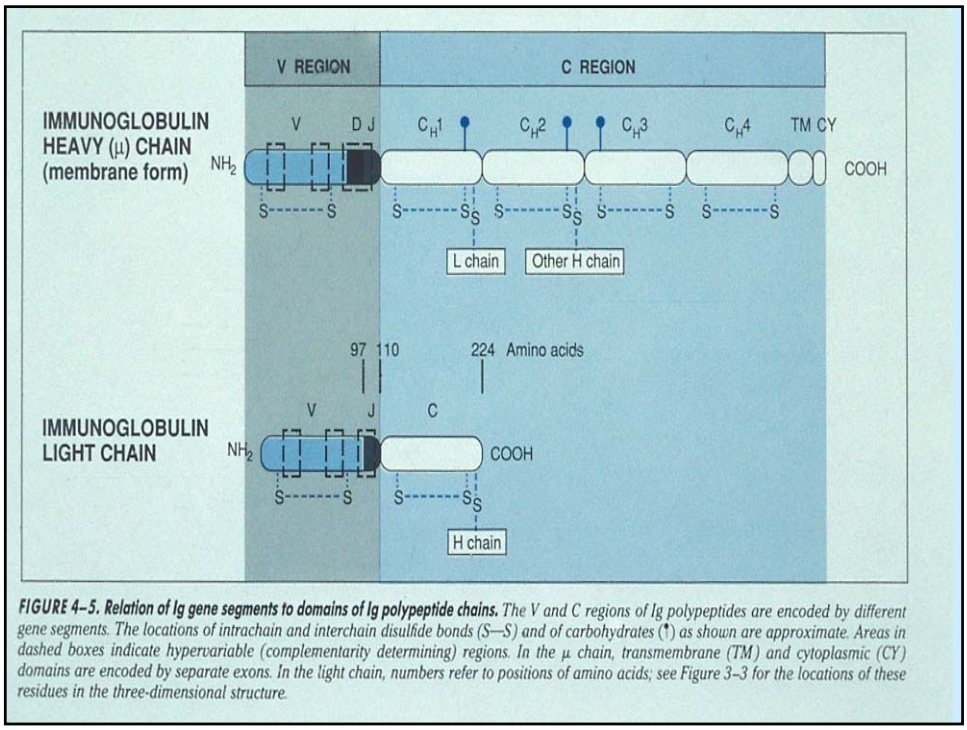


FIGURE 4–5. Relation of Ig gene segments to domains of Ig polypeptide chains. *The V and C regions of Ig polypeptides are encoded by different gene segments. The locations of intrachain and interchain disulfide bonds (S—S) and of carbohydrates (•) as shown are approximate. Areas in dashed boxes indicate hypervariable (complementarity determining) regions. In the μ chain, transmembrane (TM) and cytoplasmic (CY) domains are encoded by separate exons. In the light chain, numbers refer to positions of amino acids; see Figure 3–3 for the locations of these residues in the three-dimensional structure.*

MECHANISMS FOR GENERATING ANTIBODY DIVERSITY

- **Presence of multiple V genes in the germ line.**

- **Combinatorial Diversity - due to potentially different associations of different V, D and J gene segments**.

- **Junctional Diversity**
- i. Imprecise joining
- ii. N/P region (insertional) diversity occurs in VDJ joining (heavy chain) as well as VJ join of light chain. Arises from addition of up to 20 nucleotides by terminal deoxynucleotidyl transferase (TdT).

- **Somatic Hypermutation**
- i. Occurs randomly after antigenic stimulation and principally in CDR1, CDR2, CDR3 regions (more frequent in CDR3). Introduces point mutations at a higher rate than for normal mammalian genes. Mutation rate of V genes is 1 base pair change per 10^3 base pairs/cell division; it is 10^{-7} in other mammalian genes.
- ii. Can give rise to Ig with different (new) antigen specificities leading to high or low affinity Abs. High affinity B cell clones are selectively expanded (Affinity Maturation).
- iii. Affinity maturation is associated with isotype switching.

- **Random Assortment of H and L chains**.

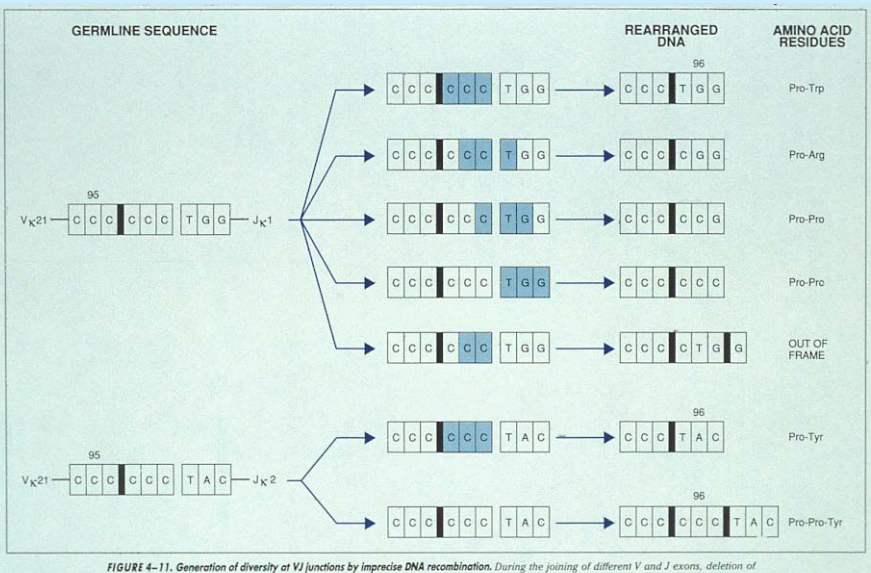


FIGURE 4–11. Generation of diversity at VJ junctions by imprecise DNA recombination. During the joining of different V and J exons, deletion of nucleotides (indicated as shaded boxes) may lead to the generation of novel nucleotide and amino acid sequences or to out-of-frame nonfunctional recombinations. All the indicated amino acid sequences have been detected in different V,21 containing mouse antibodies. Two examples are shown, one in which J,1 was used and another in which J,2 was used. (Adapted with permission from Weigert, M., R. Perry, D. Kelley, T. Hunkapiller, J. Schilling, and L. Hood. The joining of V and J gene segments creates antibody diversity. Nature 283:497–499, 1980. Copyright © 1980, Macmillan Magazines Ltd.)

Mechanisms Contributing to the Generation of Antibody Diversity in Humans

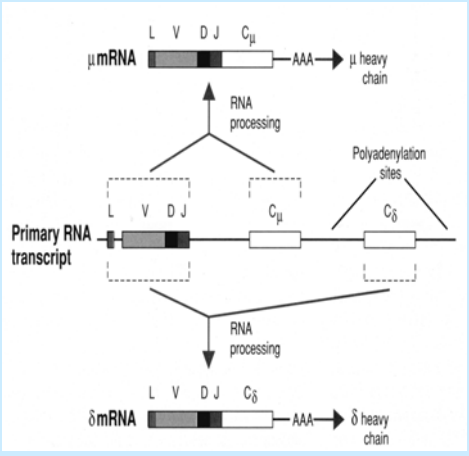
Germ line genes	**H**	κ	λ
V segments	65	40	30
J segments	6	5	4
D segments	27	0	0
Combinatorial Joining			
V x J (x D)	11,000	200	120
H-L chain associations			
H x κ	2.2×10^6	from 177	
H x λ	1.3×10^6	segments	

EXPRESSION OF DIFFERENT CLASSES AND TYPES OF IMMUNOGLOBULINS

•Co-expression of IgM and IgD

• Both IgM and IgD are co-expressed on the surface of a mature B lymphocyte.

• Occurs by alternative RNA splicing.



ISOTYPE SWITCHING

Is the conversion of an immunoglobulin from one isotype to another (e.g. IgG to IgE) while retaining the same antigen specificity.

- Switching is dependent on antigenic stimulation and is induced by cytokines released by helper T cells and requires engagement of CD40L [e.g. IL-4 triggers switching from IgM to IgE or IgG4 (humans); IFN-γ triggers switching from IgM to IgG2a (mice)]. Cyokines are thought to alter chromatin structure making switch sites more accessible to recombinases for gene transcription.
- Involves switch sites located in introns upstream of each C_H segment (except Cδ).
- Switch sites consist of multiple copies of conserved repeat sequences [(GAGCT)n GGGGGT)] where *n* can vary from 3-7.
- Class switching occurs usually in activated B cells (including memory cells) and not in naïve B cells and involves heavy chain genes. These cells (you will recall) already have rearranged VDJ genes at the DNA level and were producing IgM and IgD.

ISOTYPE SWITCHING (Contd.)

- If this B cell chooses to class switch to say IgG at the DNA level (by DNA deletion, which is irreversible), then the switch regions of IgM and IgD (recall that they are both controlled by one switch region) join with that of IgG to excise the IgM and IgD constant region gene segments.
- When this occurs what you have left is the recombined VDJ gene cluster still separated from the IgG constant region gene segments by intronic DNA.
- The switched gene will now be transcribed into primary RNA that undergoes splicing to now bring the VDJ and C regions together to form VDJC.
- The take home lesson is that VDJ (or VJ) segments do not join to C region genes at the DNA level whether it is during classical gene re-arrangement or class switching. The joining of VDJ to C to form VDJC occurs only at the RNA level!
- **Mechanism of switching is unclear. However, class specific recombinases are thought to recognize and bind to switch sites to facilitate recombination.**

Isotype switching can occur by:

A. Switch recombination (Deletion of DNA)

-primary mechanism of isotype switching

-is irreversible

B. Alternative splicing of primary RNA transcript (rare)

-Explains co-expression of multiple isotypes by a single B cell.