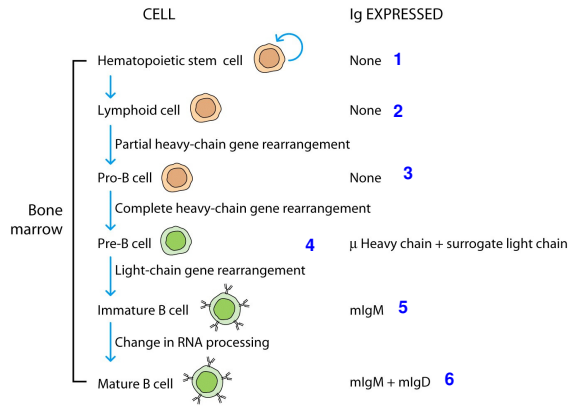
Chapter 5

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**

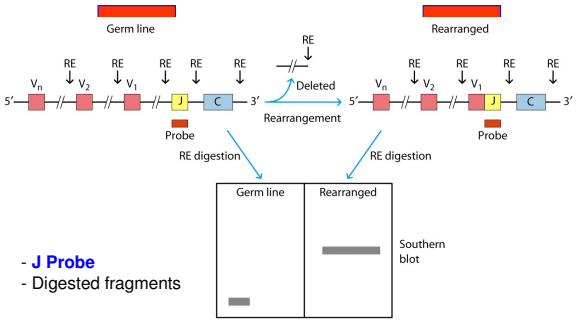
Models to Explain Antibody Diversity

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

Tonegawa (1976): Immunoglobulin gene rearrangement



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

Gene	Chromosome	
	Human	Mouse
λ Light chain	22	16
κ Light chain	2	6
Heavy chain	14	12

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

Segment	Light chains		Heavy chain
	κ	λ	H
Variable (V)	40	30	40
Diversity (D)	0	0	25
Joining (J)	5	4	6

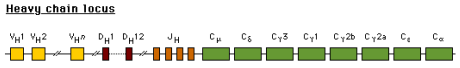
Figure 4-3 Immunobiology, 6/e, (© Garland Science 2005)

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.

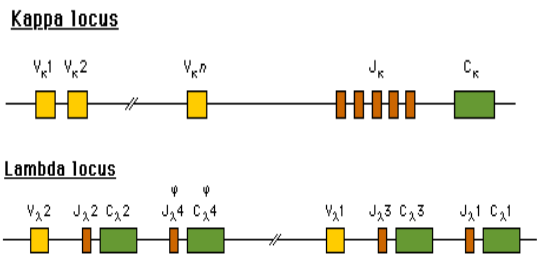
During maturation, one of each V, D and J segment is randomly "chosen" and used to encode the final antibody molecule.

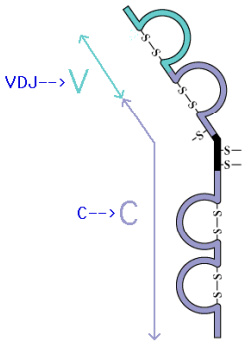
Germline configuration of the heavy chain locus (mice)

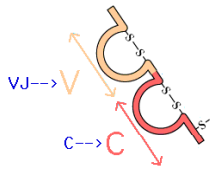


The kappa locus has a similar structure - BUT - does not have D segments.

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

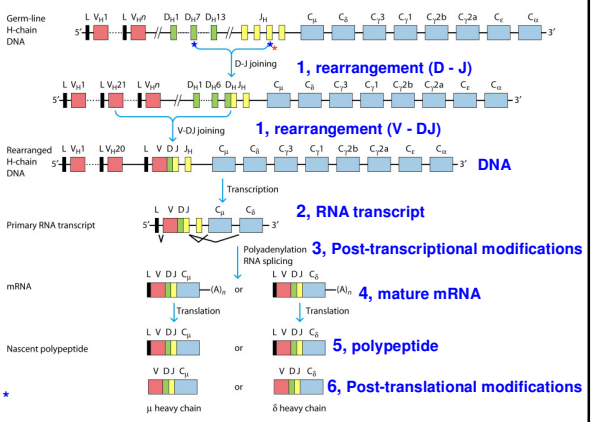






- In heavy chains, the V, D and J segments encode the variable domain while the C segment encodes the constant domain.

- In light chains, the V and J segments encode the variable domain while the C segment encodes the constant domain.

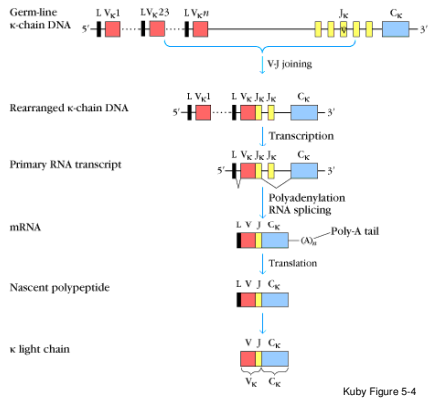


The kappa and lambda loci undergo similar rearrangement.

Since there are no D segments, there is a single V-->J rearrangement.

The final light chain mRNA contains one VJC unit.

Gene rearrangement Kappa light Chain



Kuby Figure 5-4

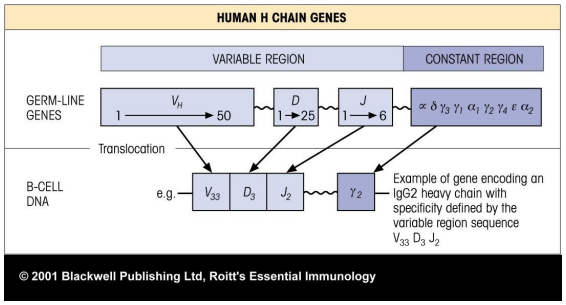


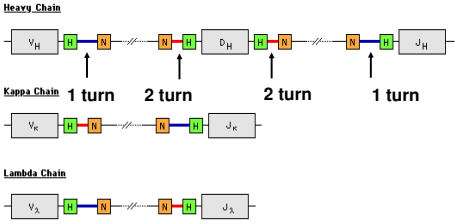
Figure 3.7

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 and nonamer with a 12 or 23 bp spacer.

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



Recombinant Signal Sequences (RSS)

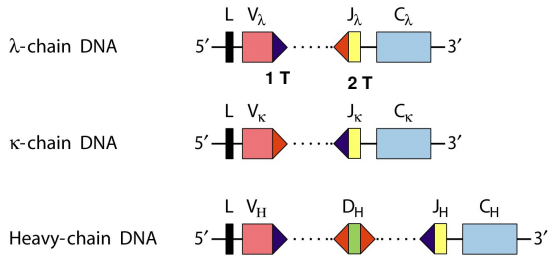
(a) Nucleotide sequence of RSSs

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**

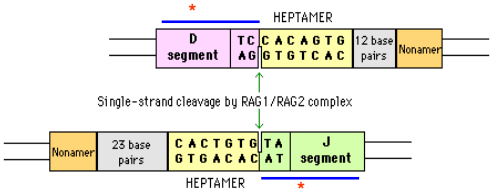
(b) Location of RSSs in germ-line immunoglobulin DNA



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

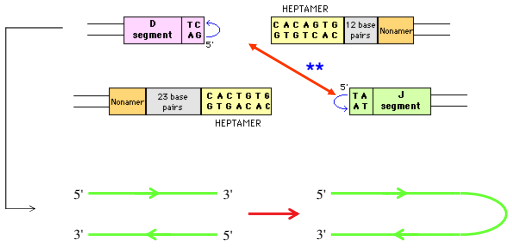
-The RAG1/RAG2 complex recognizes the heptamer/nonamer sequences and cuts one strand of the DNA.

Step 1: The RAG1/RAG2 complex recognizes the RSS's and mediates single-strand DNA cleavage



A hairpin forms…

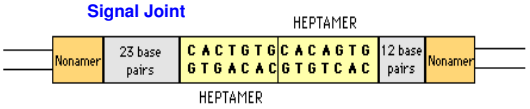
Step 2: The 5' cut end of the cut strand reacts with the uncut strand resulting in a double-stranded break and hairpin formation



The hairpin is cut at a random site…

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

Signal Joint



Coding Joint

Endonuclease (Artemis)

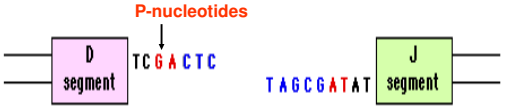
Palindromic sequences may form…

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

Terminal deoxynucleotidyl transferase (Tdt)

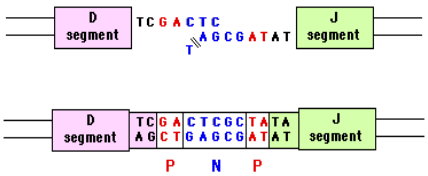
An enzyme that randomly adds in nucleotides during joining of coding gene segments.

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



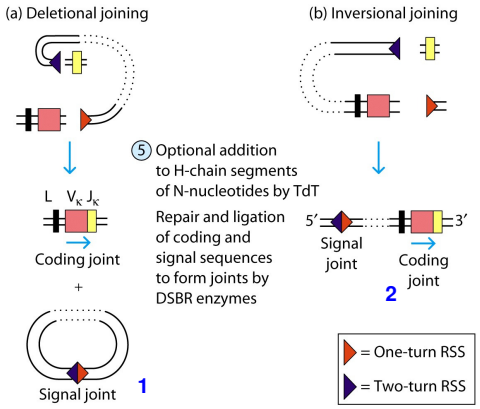
The join is repaired...

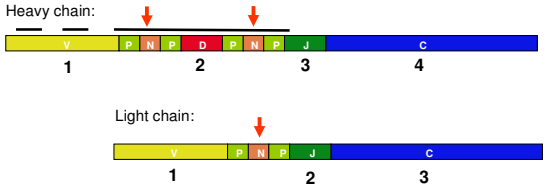
Step 6: The two single-stranded ends pair. Unpaired nucleotides are trimmed by an exonuclease and the coding joint is repaired



Coding Joint

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





The final "gene" encoding the antibody produced by a B cell (and T cells) consists of a number of different segments.

This process of recombination of different gene segments and addition of **P** and **N nucleotides** ensures that an enormous number of different antigen specificities are possible.

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**

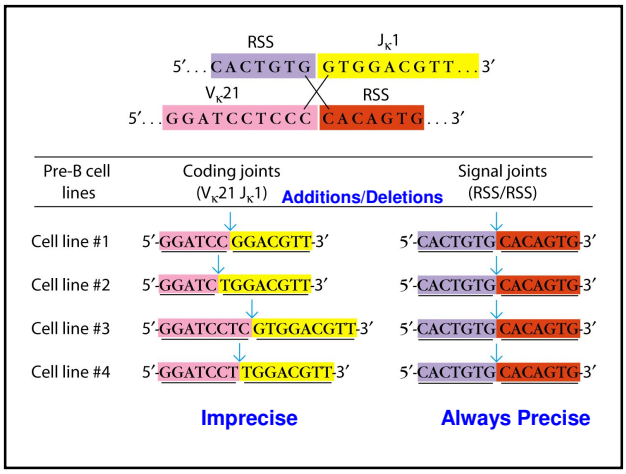
Combinatorial V-J and V-D-J joining

- 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 x (200 x 120) = **2.64×10^6**

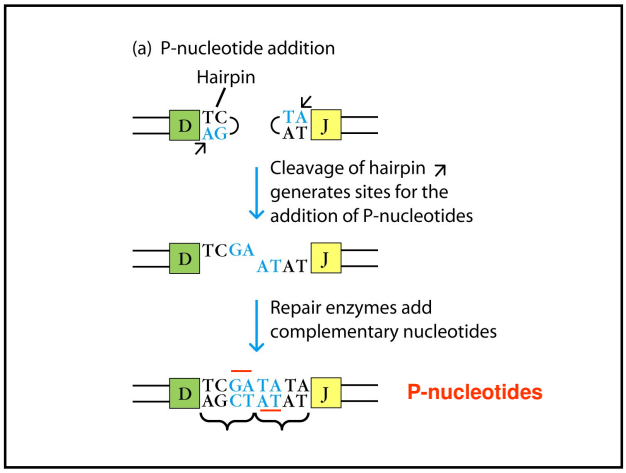
Junctional flexibility

- 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 nuclotides can be added to the **D-J** and **V-DJ** in the H chain.
- This mechanisms does not happen in the Light chain

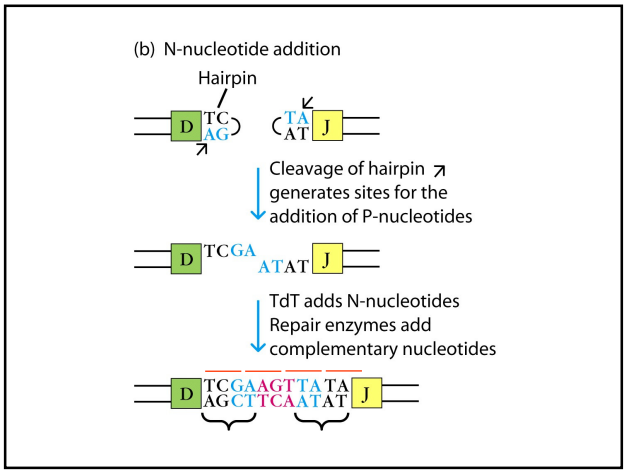
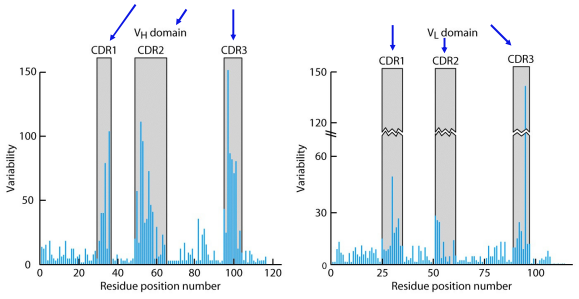


TABLE 5-3 Sources of sequence variation in complementarity-determining regions of immunoglobulin heavy- and light-chain genes

Source of variation	CDR1	CDR2	CDR3 ****
Sequence encoded by:	V segment	V segment	V_L-J_L junction; V_H-D_H-J_H junctions
Junctional flexibility	–	–	+
P-nucleotide addition	–	–	+
N-nucleotide addition*	–	–	+
Somatic hypermutation	+	+	+

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

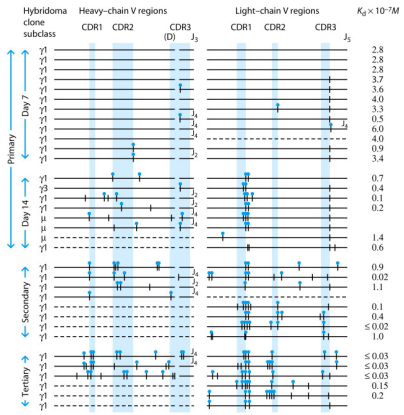
Complementarity-Determining Regions, or CDRs.



Heavy Chain **Light Chain**

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

	B cell receptor (Immunoglobulin)	
	Heavy	Light
V gene segments	1000	300
D gene segments	15	-
J gene segments	4	4
N region insertion	++	-
Junctional diversity	+++	+
Somatic mutation	+	+
	V x D x J 1000 x 15 x 4	V x J 300 x 4
Total	6×10^4	1.2×10^3
Combinatorial association	7.2×10^7	

2.64×10^6

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

ALLELIC EXCLUSION:

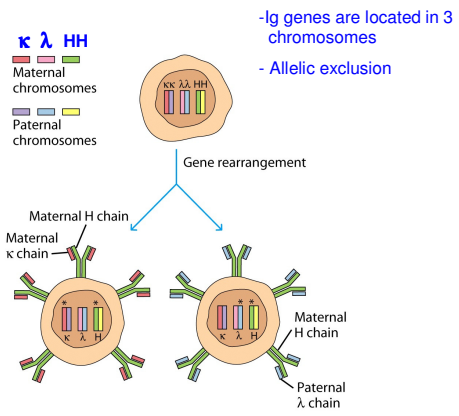
-We have two copies (alleles) of each Ig gene - one inherited from our father and one from our mother.

- In most cases, both genes are expressed.

- But Antibody genes are different! …. Only one heavy chain allele and one light chain allele is expressed!!!

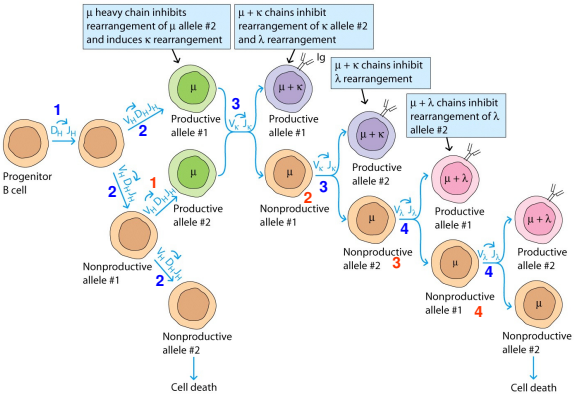
- This is termed **allelic exclusion** (one allele is excluded). Once a productive arrangement is made, the other allele is suppressed

- Why? To ensure that each B cell makes antibody of a single specificity.



-Ig genes are located in 3 chromosomes

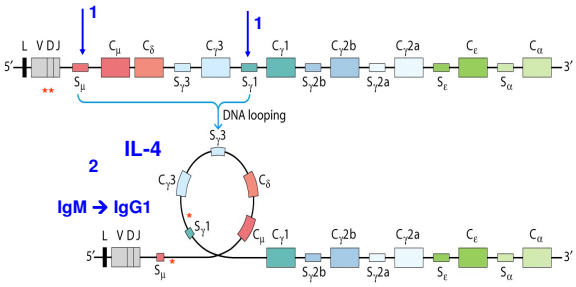
- Allelic exclusion

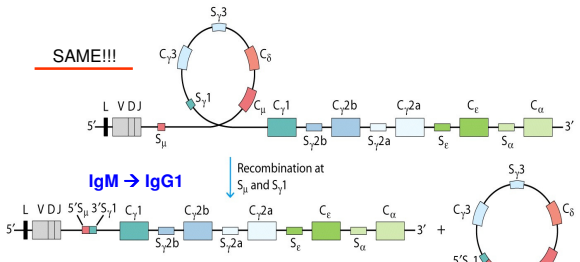


Class Switching

- 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 (Cδ)
- 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

Class Switching





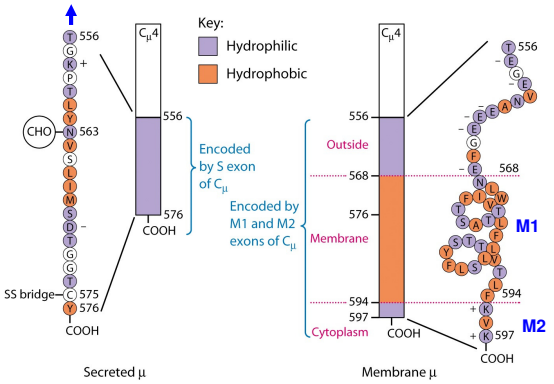
-**Recombination between Switch Regions**

-**Switching only proceeds downstream**

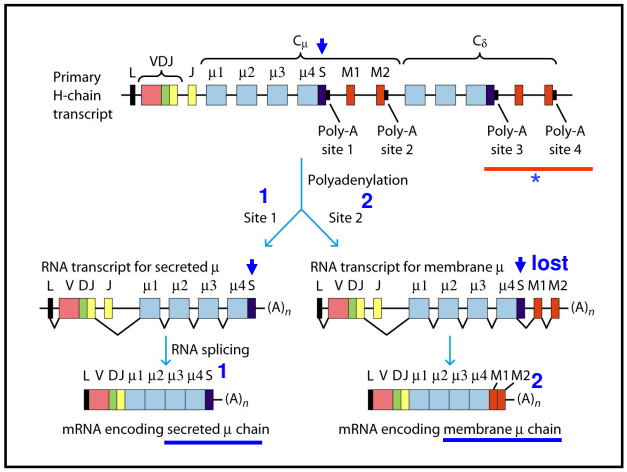
Note: Same specificity but different H chain

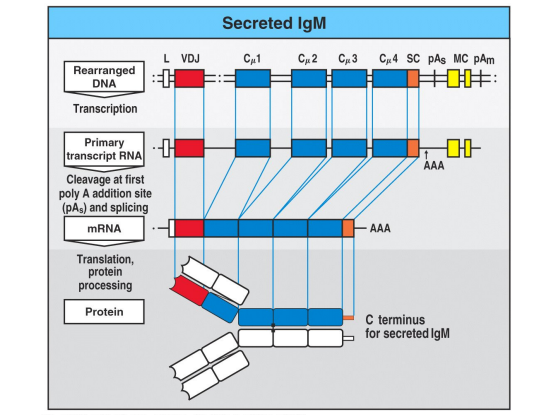
Expression of membrane or secreted Immunoglobulin

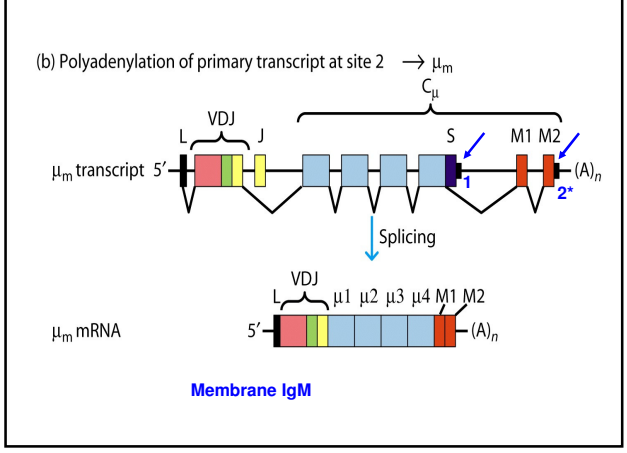
- In mature B cells → membrane forms; in Plasma cells→ secreted forms
- Process depends on **differential processing** of primary transcript
- Remember: IgG, IgD, IgA (3 C_H domains), IgM and IgA (4 C_H domains).
- Domain 3/4 contains the Secretory (hydrophilic) nucleotide sequence (S) at its 3'.
- Two Exons at 3' encode the M1 (trans-membrane) and M2 (cytoplasmic) segments.
- Primary transcript contains two PolyA sites: If cleavage at Poly A site I = Secreted Form. If cleavage at PolyA site 2 = Membrane Form

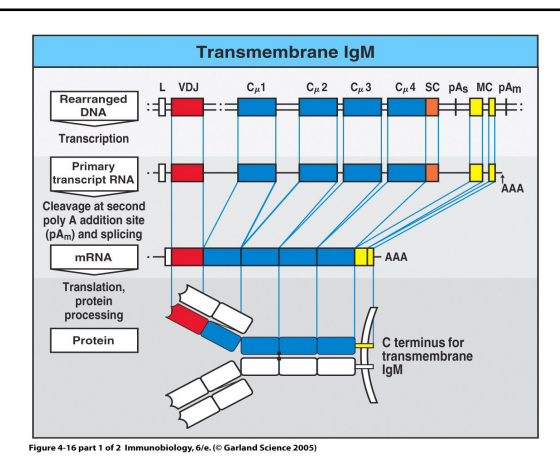


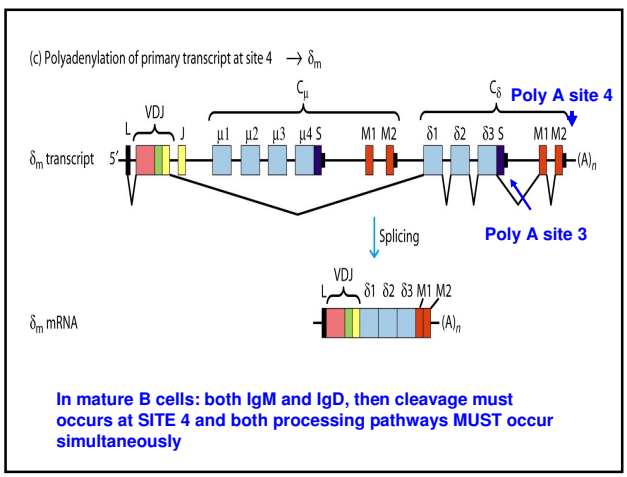
S component **M1 and M2 components (exons)**







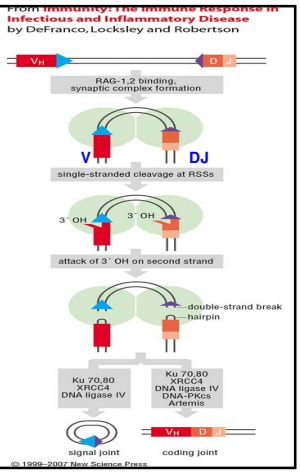


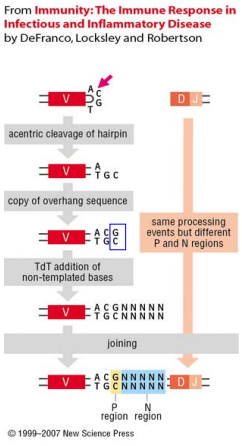
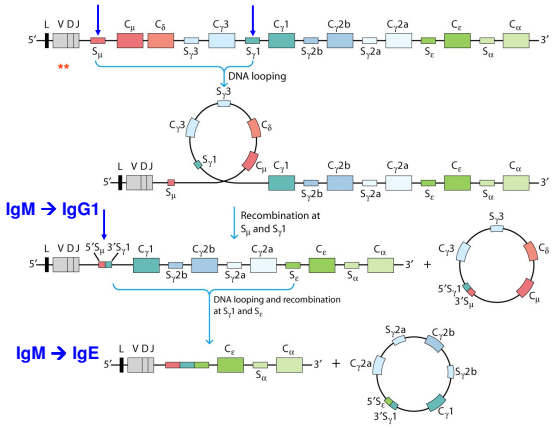


In mature B cells: both IgM and IgD, then cleavage must occurs at SITE 4 and both processing pathways MUST occur simultaneously

The End

- **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.