Antibody Diversity

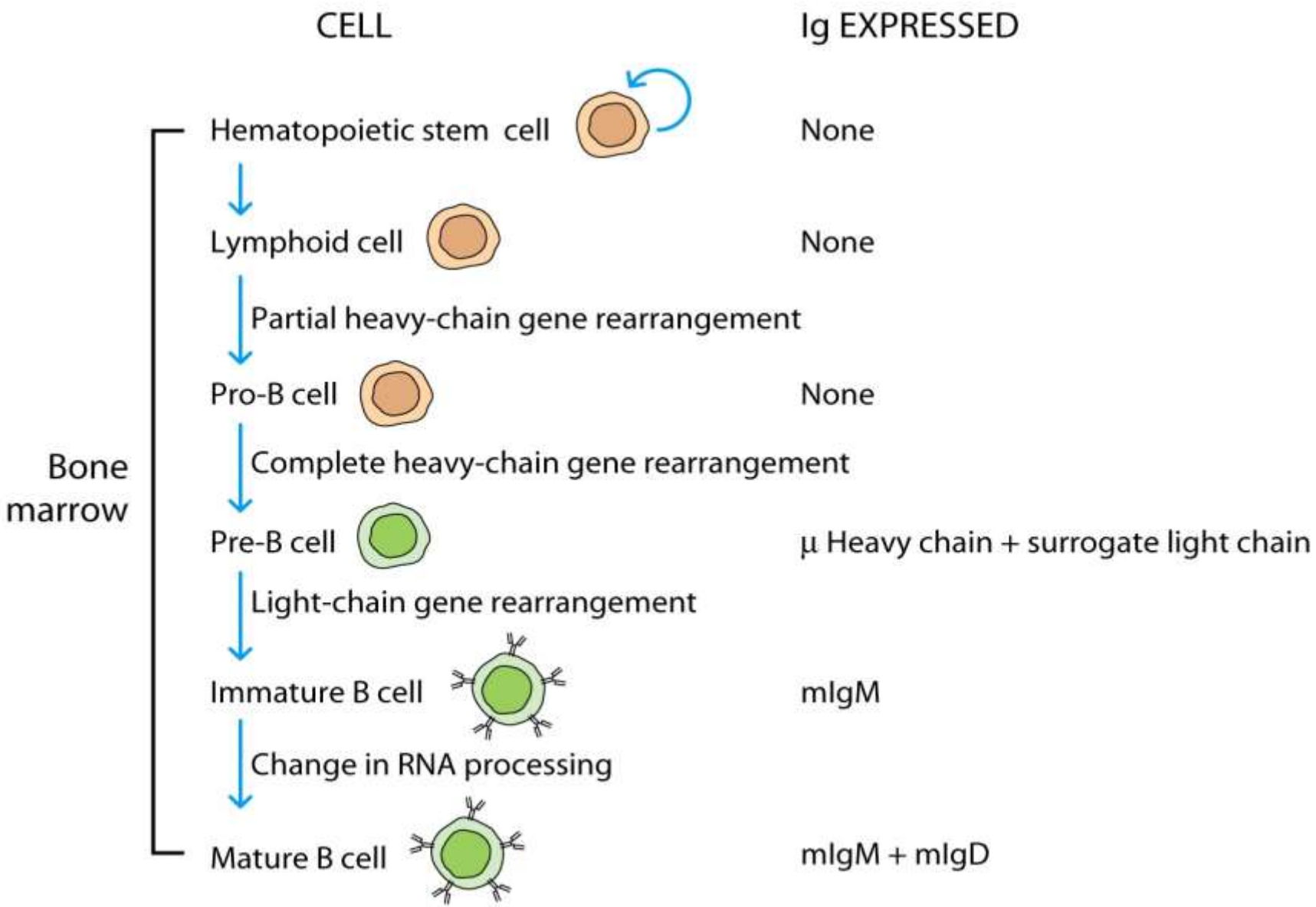
Problem…the immune system makes over one billion different antibody proteins

- In 1950's: central dogma stated DNA—to RNA—to protein
- One gene for one polypeptide hyphothesis
- Required millions of genes just for the immune system
- Does not seem possible, but most scientists thought it might be
- Today we know the human genome is less than 30,000 genes
- ***So, what is really going on???***

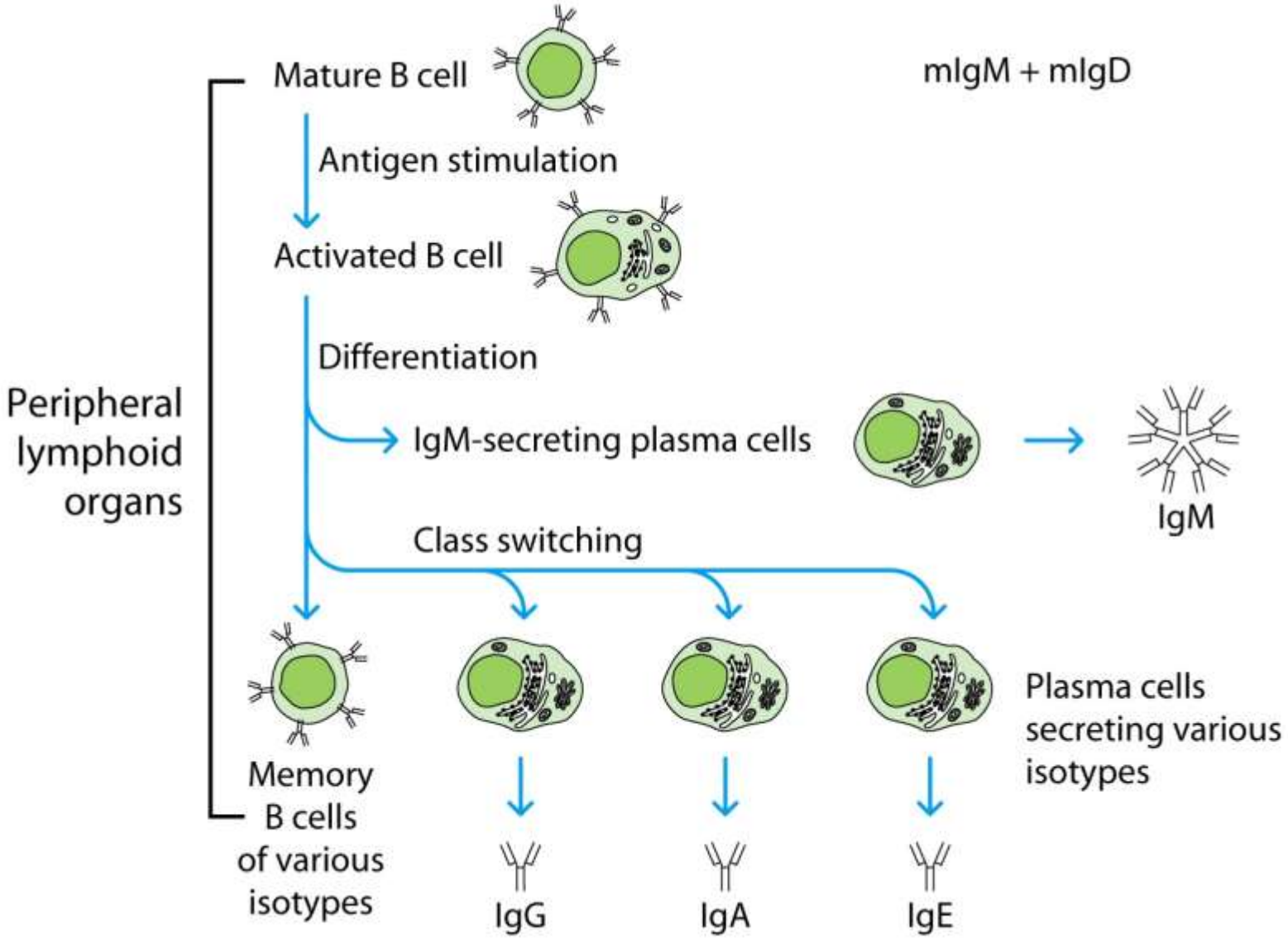
Current theory must account for the following known properties of antibodies

- The vast diversity of antibody specificities
- The presence in Ig heavy and light chains of a variable region at the amino-terminal end and a constant region at the carboxyl-terminal end
- The existence of isotypes with the same antigenic specificity, which result from the association of a given variable region with different heavy-chain constant regions

B lymphocyte development



B lymphocyte development (2)



Germ-line vs somatic-variation theories

- Germ-line: stated that each antibody had its own gene….nothing special, but required billions of genes to account for numbers of antibodies
- Somatic-variation: some mutation and recombination created vast number of genes for antibody formation
- This introduced a new concept: targeted mutation or recombination of DNA: is it possible??
- Paradox: how could stability be maintained in C region and diversity exist in V region?

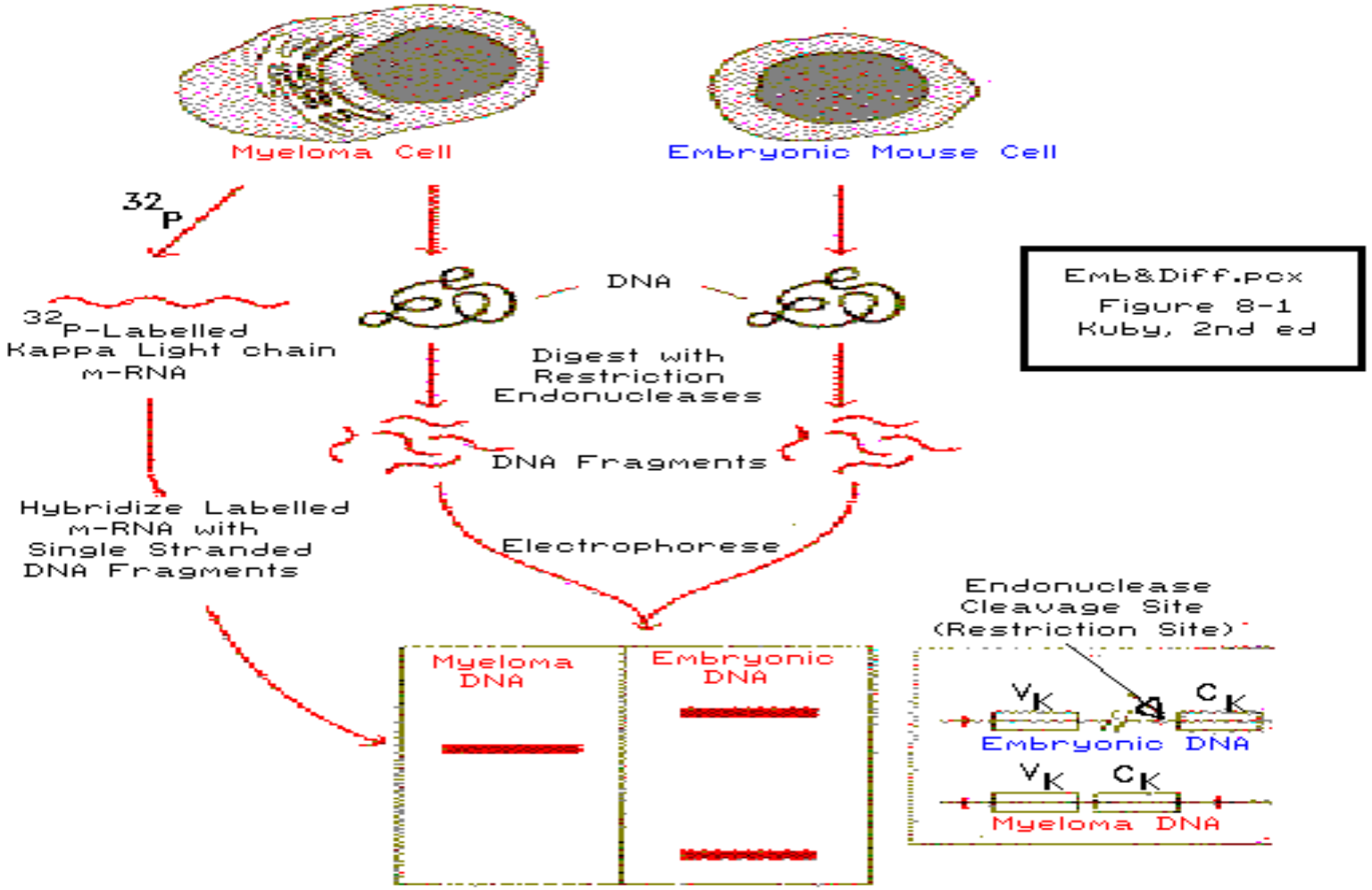
Dreyer and Bennett- likely to open pandora box.

- In 1965 proposed radical theory to account for diversity of antibodies
- Each antibody was coded for by two separate genes (One gene for one polypeptide hyphothesis DIFFER)
- One for the variable region
- One for the constant region
- Combined at the DNA level and expressed single mRNA
- Suggested 1000's of variable region genes and only one constant region gene
- Most scientists did not like this idea called it absurd and rejected it

BUT?

Tonegawa's demonstration-(OPEN the Pandora Box)

- 1976—used restriction enzymes and DNA probes to show that germ cell DNA contained several smaller DNA segments compared to DNA taken from developed lymphocytes (myeloma cells)

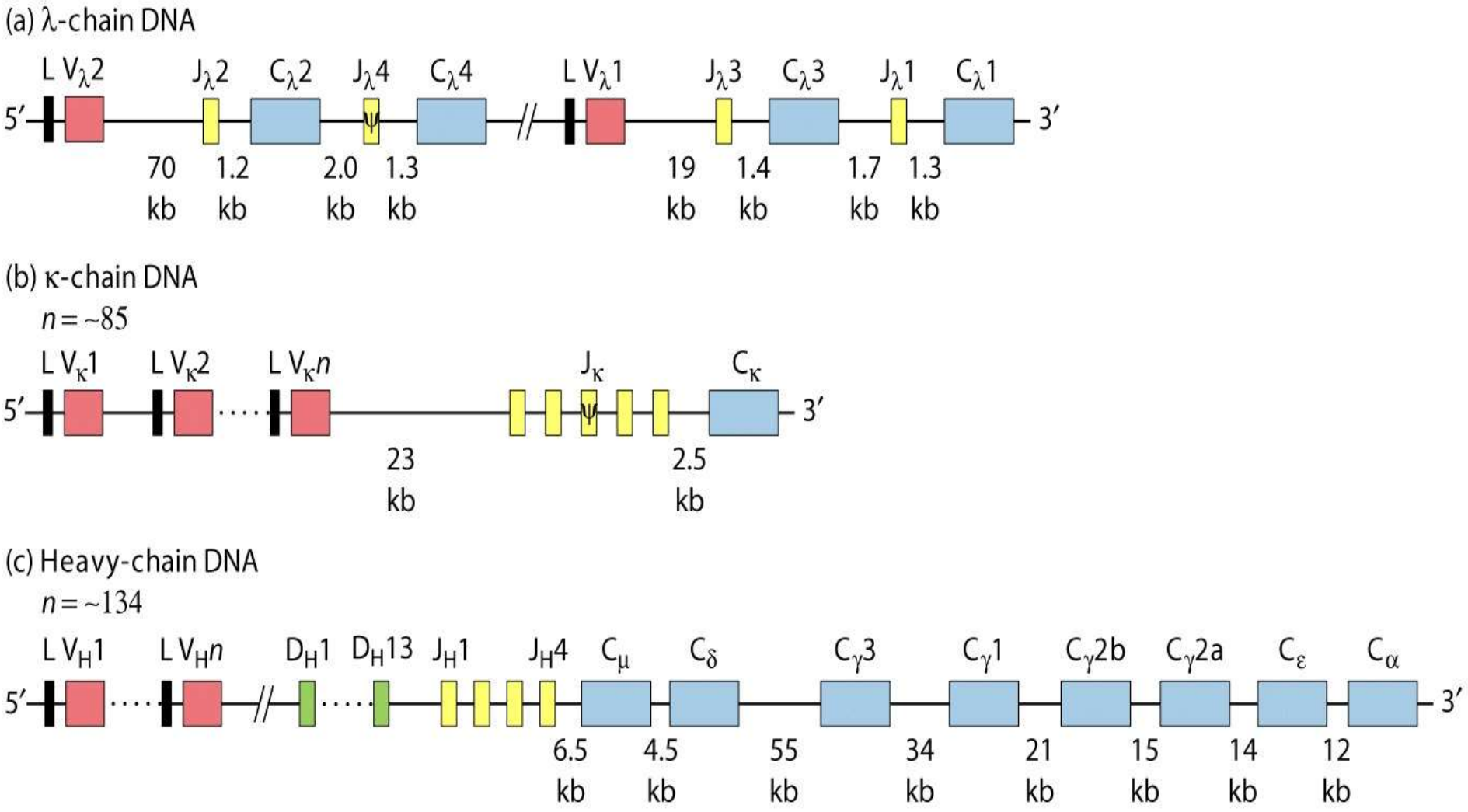


Genes for immunoglobulin proteins are found on different chromosomes

TABLE 5-1 Chromosomal locations of immunoglobulin genes in human and mouse

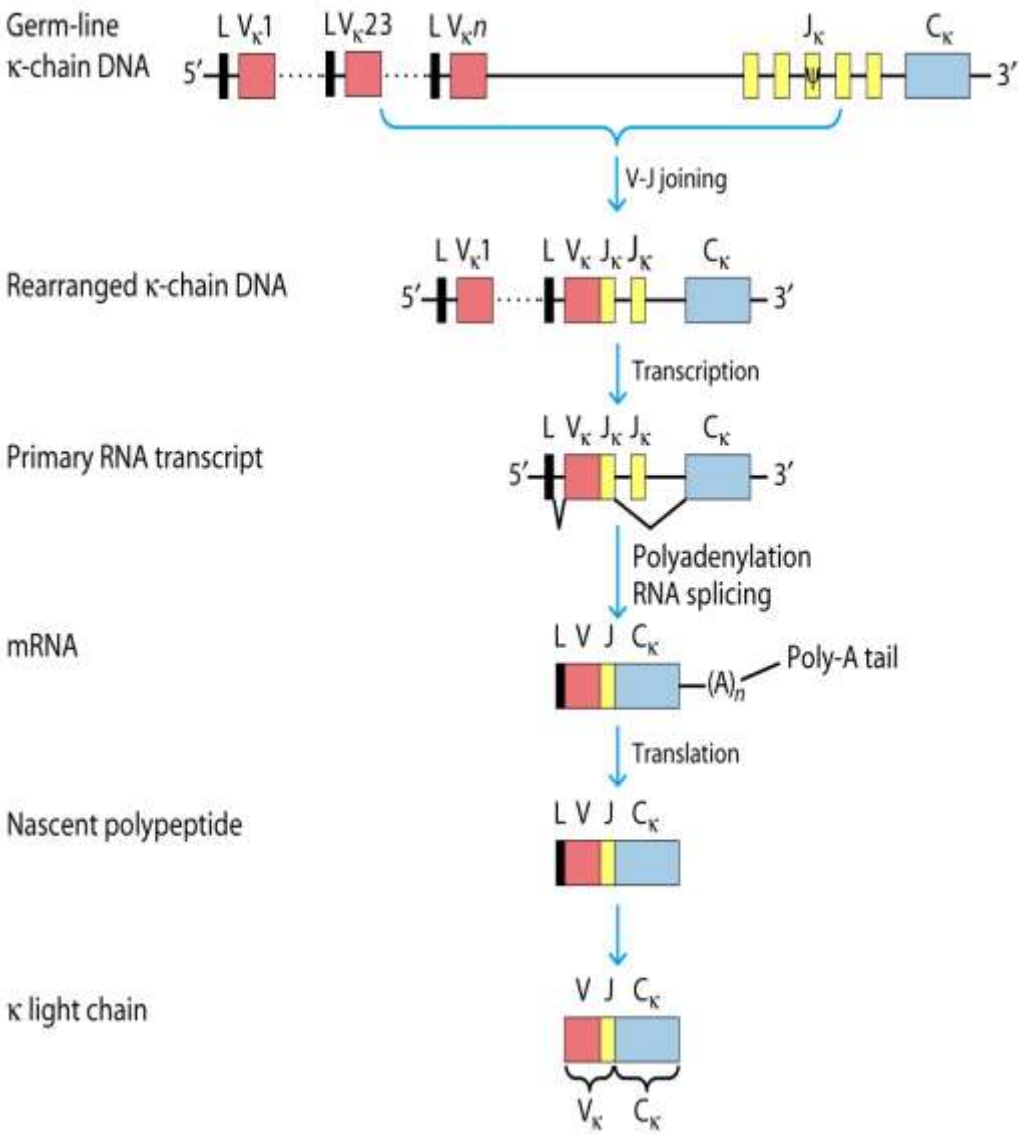
Gene	CHROMOSOME	
	Human	Mouse
λ Light chain	22	16
κ Light chain	2	6
Heavy chain	14	12

Multigene organization of Ig genes

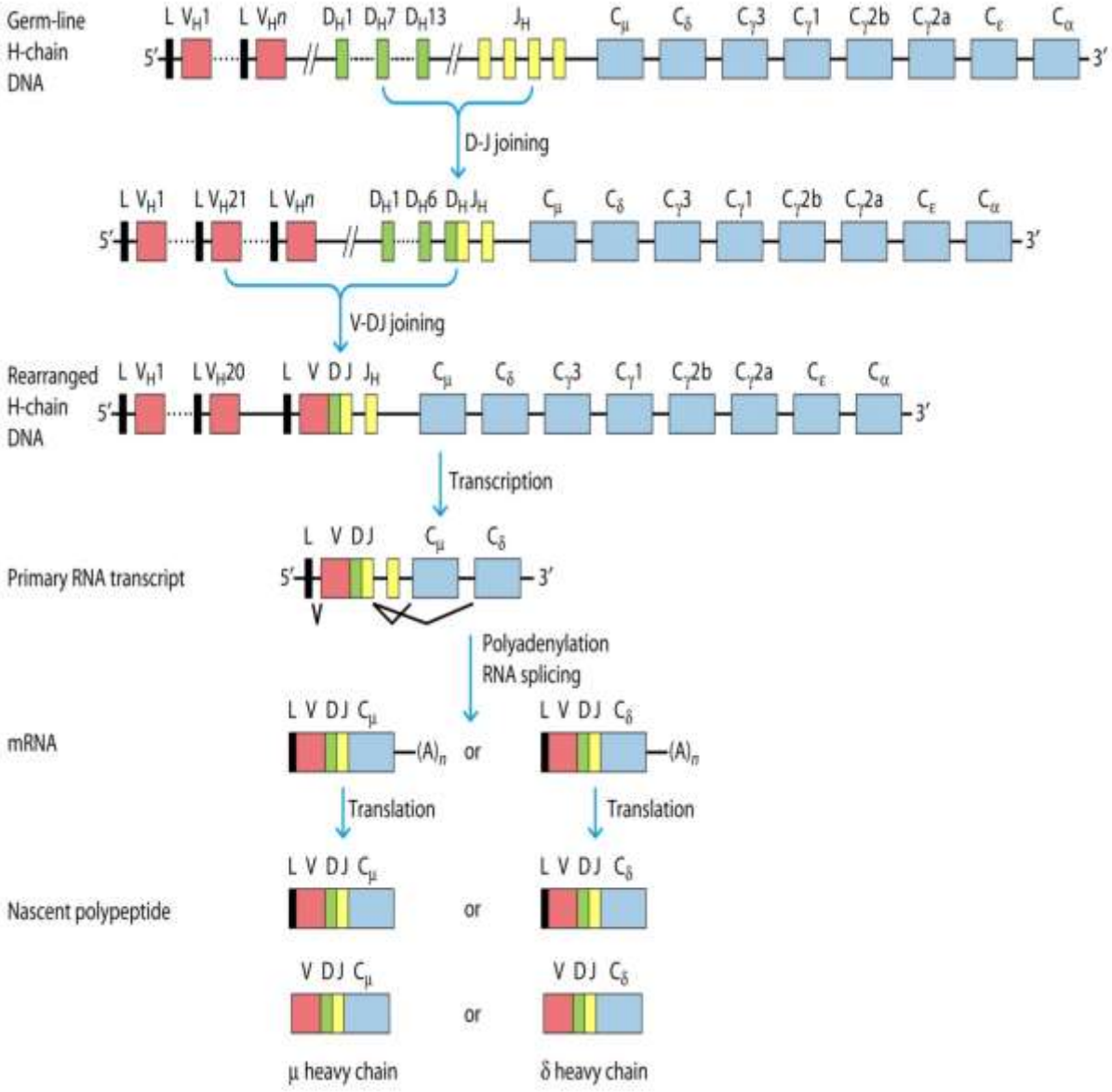


Kappa light chain rearrangement

When the nucleotide sequence was compared with the known amino acid sequence of the –lamda chain variable region, an unusual discrepancy was observed. First 97 amino acids of the lamda chain variable regioncorresponded to the nucleotide codon sequence, the remaining13carboxyl-terminal
amino acids of the protein's variable region did not match. It turned out that many base pairs away a separate, 39-bp gene segment, called J for *joining*, *encoded* the remaining 13 amino acids of the lamda-chain variable
region.



Heavy chain rearrangement



V_H gene- amino acids 1 to 94
J_H gene -amino acids 98 to 113.

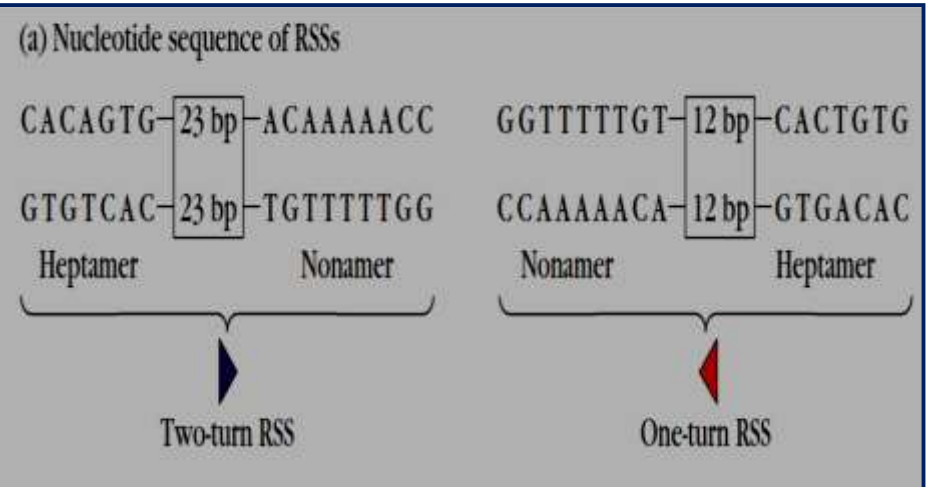
however, neither of these gene segments carried the information to encode amino acids 95 to 97.

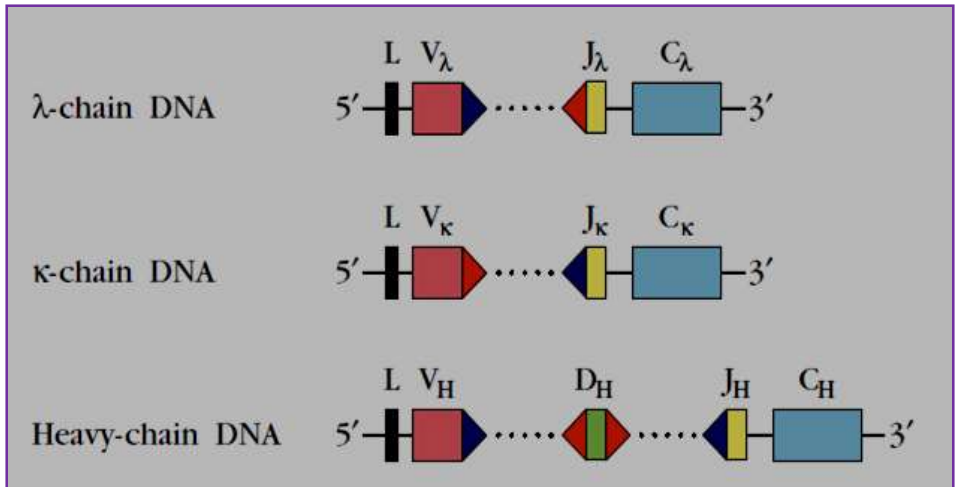
When the nucleotide sequence was determined for a rearranged myeloma DNA and compared with the germ-line DNA sequence, an additional nucleotide sequence was observed between the V_H and J_H gene segments. This nucleotide sequence corresponded to amino acids 95 to 97 of the heavy chain

Mechanism of variable region rearrangements

- Each V, D and J segments of DNA are flanked by special sequences (RSS—recombination signal sequences) of two sizes. One RSS is located 3 to each V gene segment, 5 to each J gene segment, and on both sides of eachD gene segment.
- RSS contain palindromic heptamer and a conserved AT-rich nonamer sequence separated by an intervening sequence of 12 or 23 base pairs.
- Only single turn can combine with a double turn sequence
- Joining rule ensures that V segment joins only with a J segment in the proper order

V-(D)-J recombination is catalyzed by enzymes collectively called **V(D)J recombinase (RAG 1 and RAG 2)** and the enzyme **terminal deoxynucleotidyl transferase (TdT)** are the only lymphoid-specific gene products

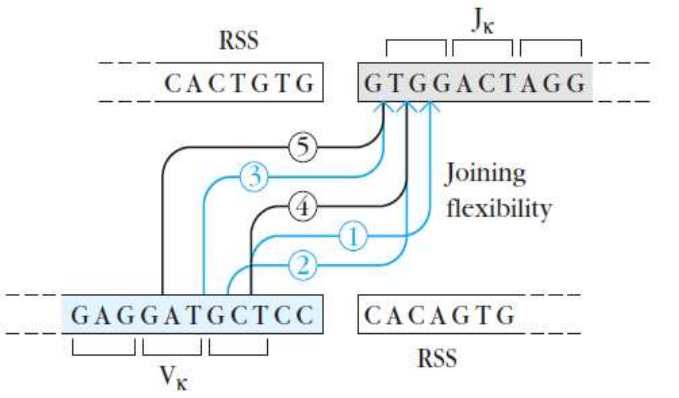




Generation of Antibody Diversity:

- Multiple germ-line gene segments
- Combinatorial V-(D)-J joining
- Junctional flexibility
- P-region nucleotide addition (P-addition)
- N-region nucleotide addition (N-addition)
- Somatic hypermutation
- Combinatorial association of light and heavy chains

Randomness in joining process – junctional felxibility



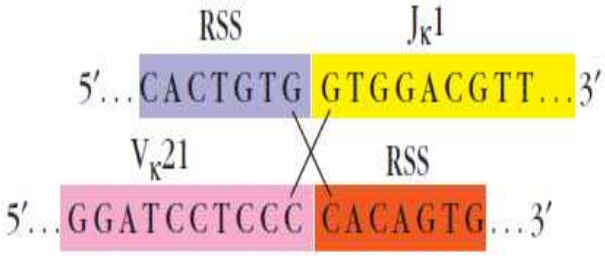
Productive rearrangements

①	Glu	Asp	Ala	Thr	Arg
	GAG	GAT	GCG	ACT	AGG
②	Glu	Asp	Gly	Thr	Arg
	GAG	GAT	GGG	ACT	AGG
③	Glu	Asp	Trp	Thr	Arg
	GAG	GAT	TGG	ACT	AGG

Nonproductive rearrangements

④	Glu	Asp	Ala	Asp	**Stop**
	GAG	GAT	GCG	GAC	**TAG** G
⑤	Glu	Val	Asp	**Stop**	
	GAG	GTG	GAC	**TAG**	G

Experimental Evidence



Pre-B cell lines	Coding joints ($V_\kappa 21 J_\kappa 1$)	Signal joints (RSS/RSS)
Cell line #1	5′-GGATCC GGACGTT-3′	5′-CACTGTG CACAGTG-3′
Cell line #2	5′-GGATC TGGACGTT-3′	5′-CACTGTG CACAGTG-3′
Cell line #3	5′-GGATCCTC GTGGACGTT-3′	5′-CACTGTG CACAGTG-3′
Cell line #4	5′-GGATCCT TGGACGTT-3′	5′-CACTGTG CACAGTG-3′

N nucleotide addition and P nucleotide addition

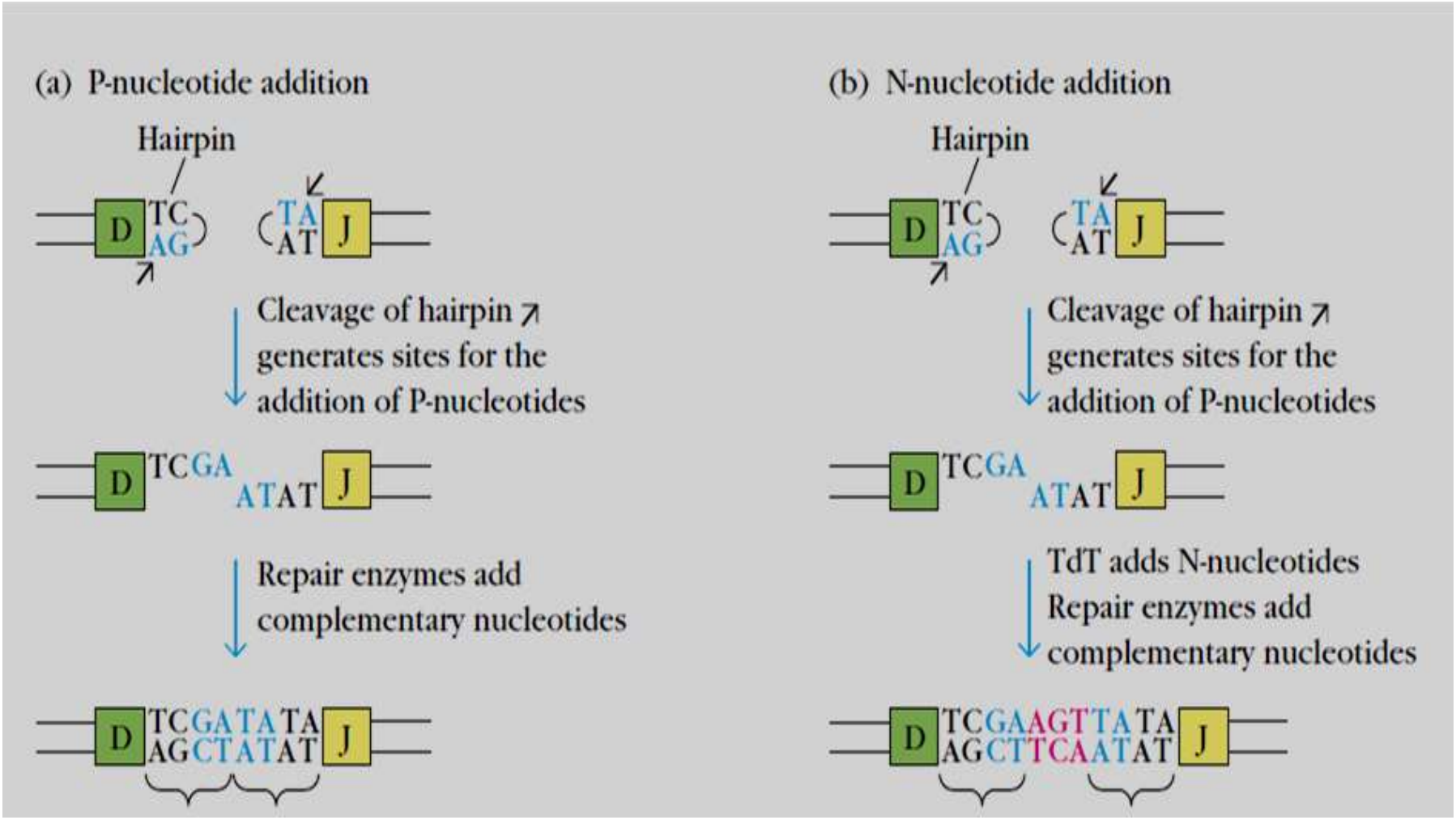
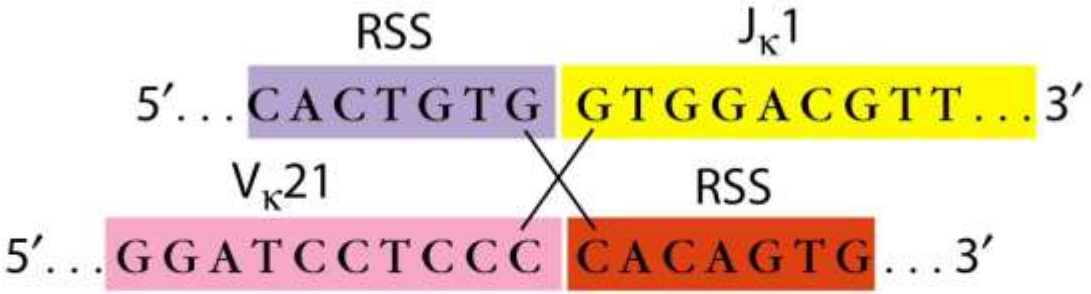


TABLE 5-3 Sources of sequence variation in complementarity-determining regions of immunoglobulin heavy- and light-chain genes

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	+	+	+

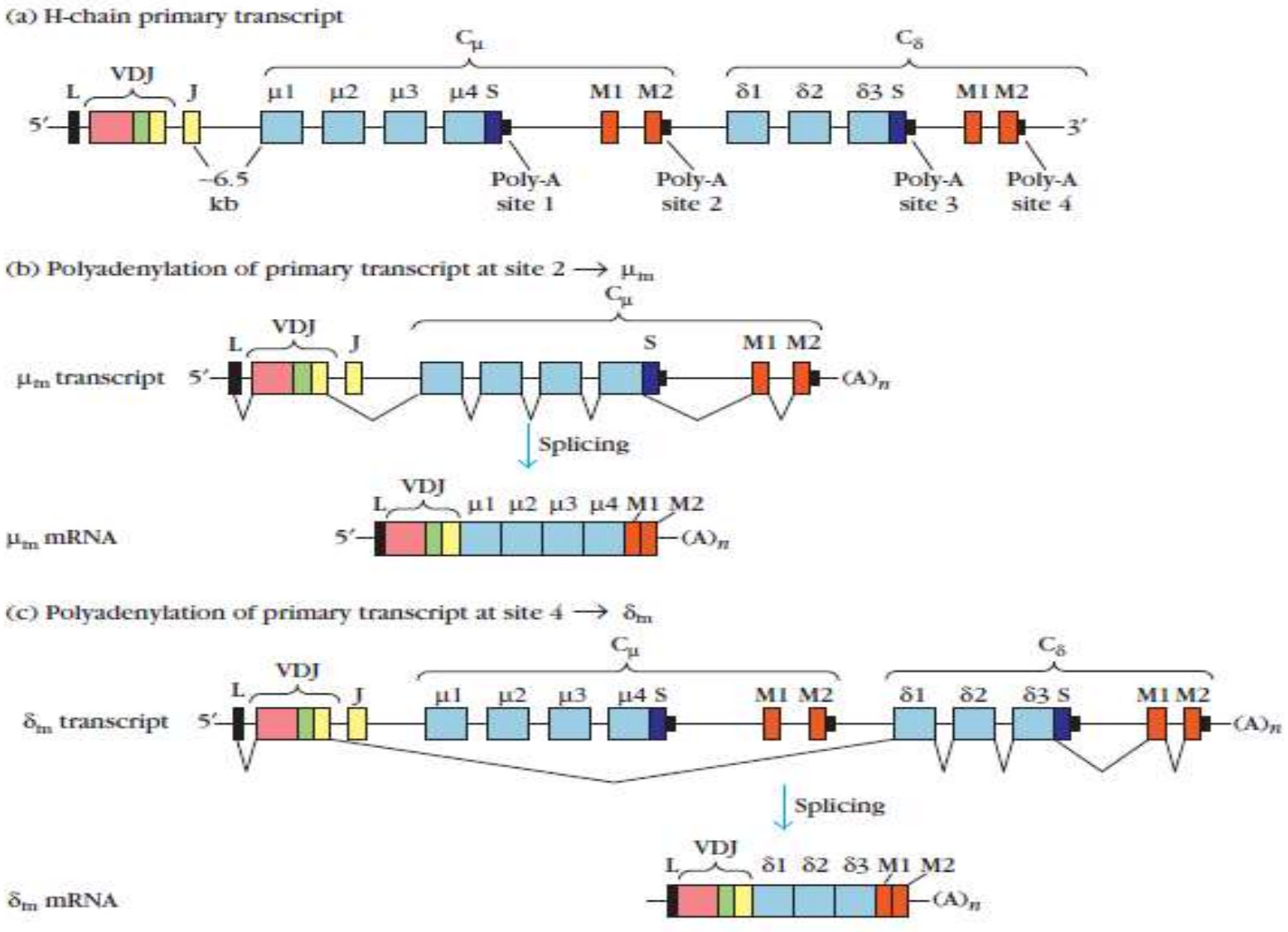
*N-nucleotide addition occurs only in heavy-chain DNA.

Imprecise joining generates diversity

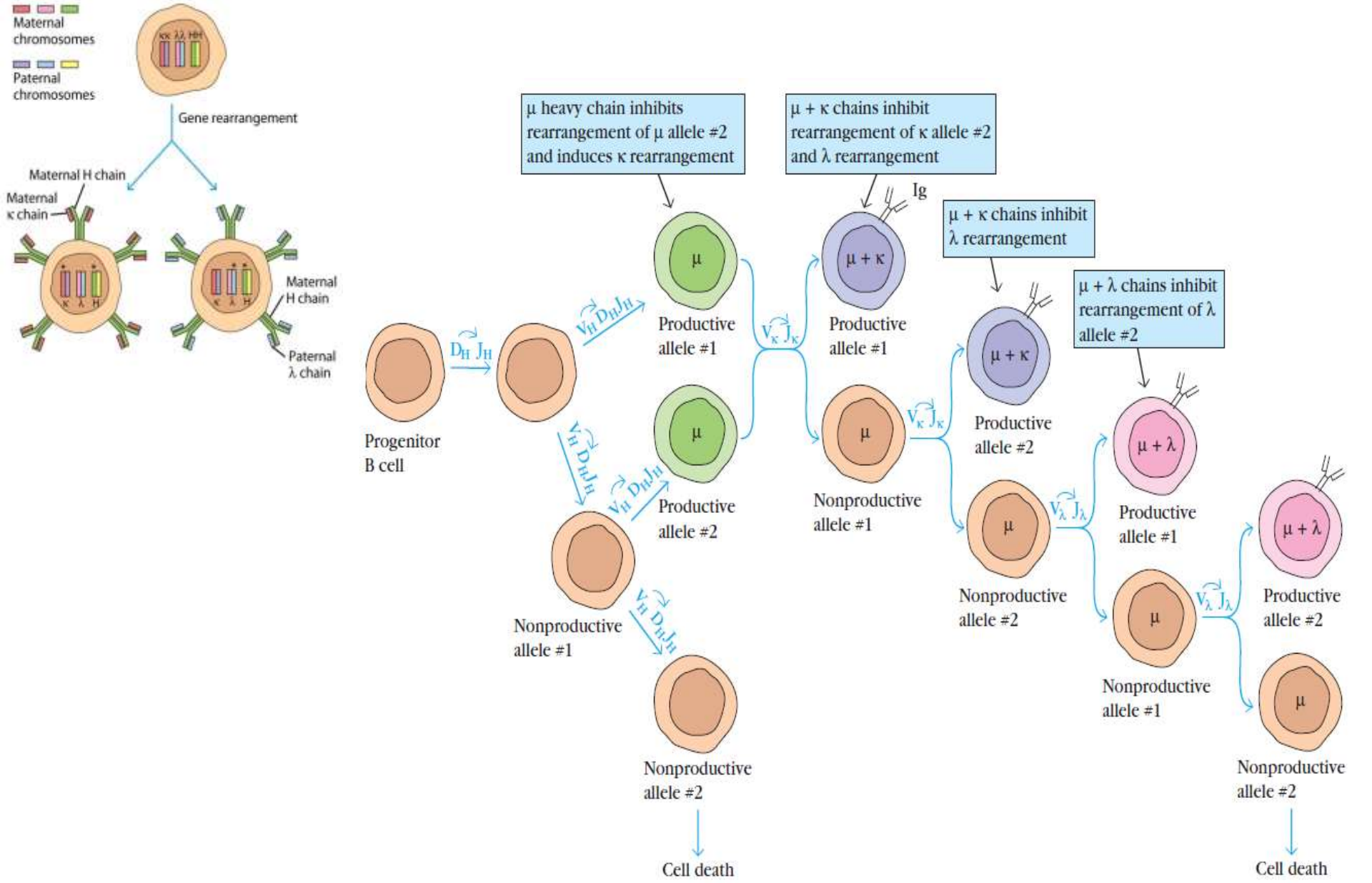


Pre-B cell lines	Coding joints ($V_\kappa 21$ $J_\kappa 1$)	Signal joints (RSS/RSS)
Cell line #1	5′-GGATCC GGACGTT-3′	5′-CACTGTG CACAGTG-3′
Cell line #2	5′-GGATC TGGACGTT-3′	5′-CACTGTG CACAGTG-3′
Cell line #3	5′-GGATCCTC GTGGACGTT-3′	5′-CACTGTG CACAGTG-3′
Cell line #4	5′-GGATCCT TGGACGTT-3′	5′-CACTGTG CACAGTG-3′

SIMULTANEOUS EXPRESSION OF IgM AND IgD



Allelic exclusion: only one chromosome is active in any one lymphocyte



Diversity calculations

TABLE 5-2 Combinatorial antibody diversity in humans and mice

Multiple germ-line segments	Heavy chain	LIGHT CHAINS: κ	LIGHT CHAINS: λ
ESTIMATED NUMBER OF SEGMENTS IN HUMANS*			
V	51	40	30
D	27	0	0
J	6	5	4
Combinatorial V-D-J and V-J joining (possible number of combinations)	$51 \times 27 \times 6 = 8262$	$40 \times 5 = 200$	$30 \times 4 = 120$
Possible combinatorial associations of heavy and light chains†	$8262 \times (200 + 120) = 2.64 \times 10^6$		
ESTIMATED NUMBER OF SEGMENTS IN MICE*			
V	134	85	2
D	13	0	0
J	4	4	3
Combinatorial V-D-J and V-J joining (possible number of combinations)	$134 \times 13 \times 4 = 6968$	$85 \times 4 = 340$	$2 \times 3 = 6$
Possible combinatorial associations of heavy and light chains†	$6968 \times (340 + 6) = 2.41 \times 10^6$		

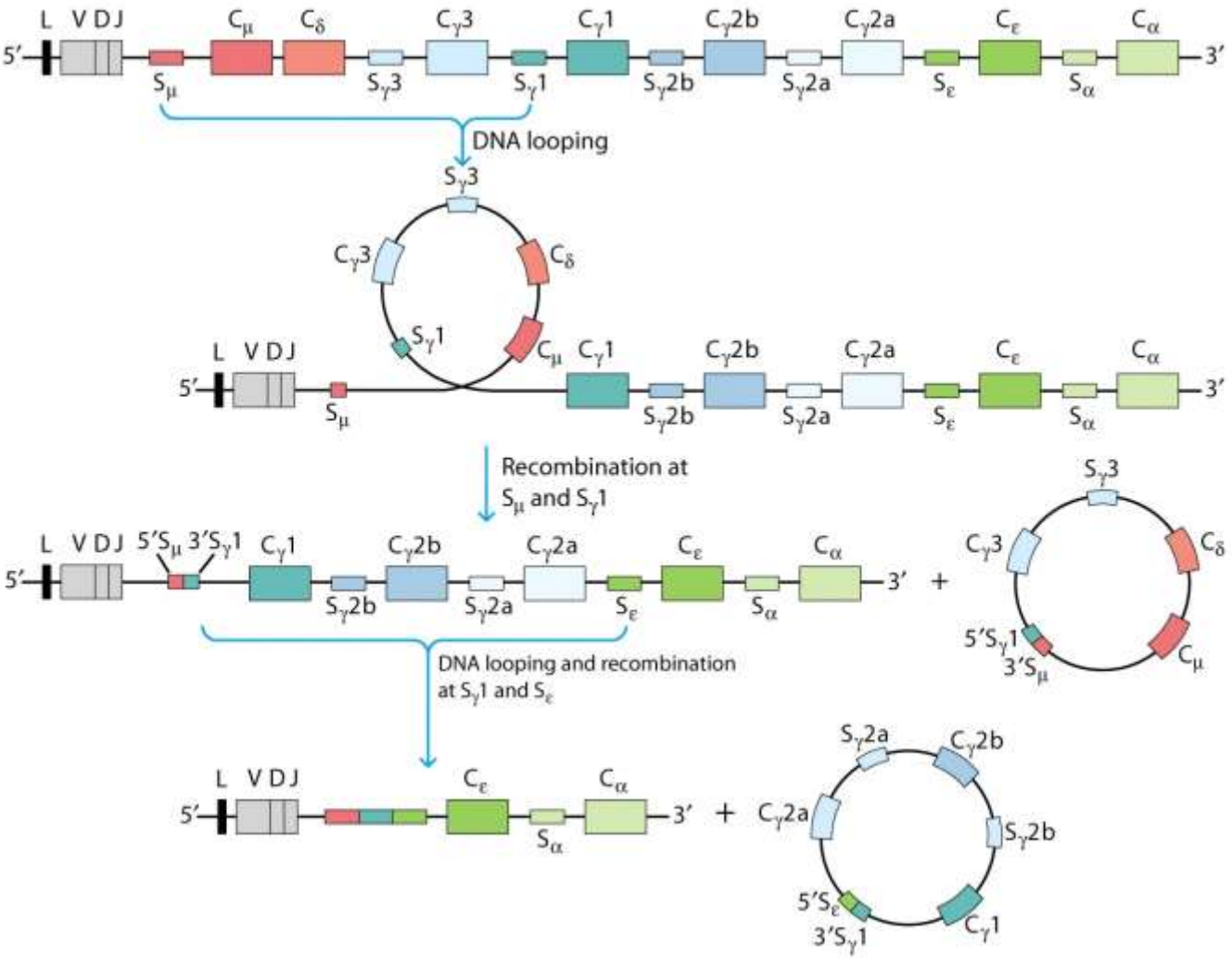
Somatic hypermutation adds even more variability

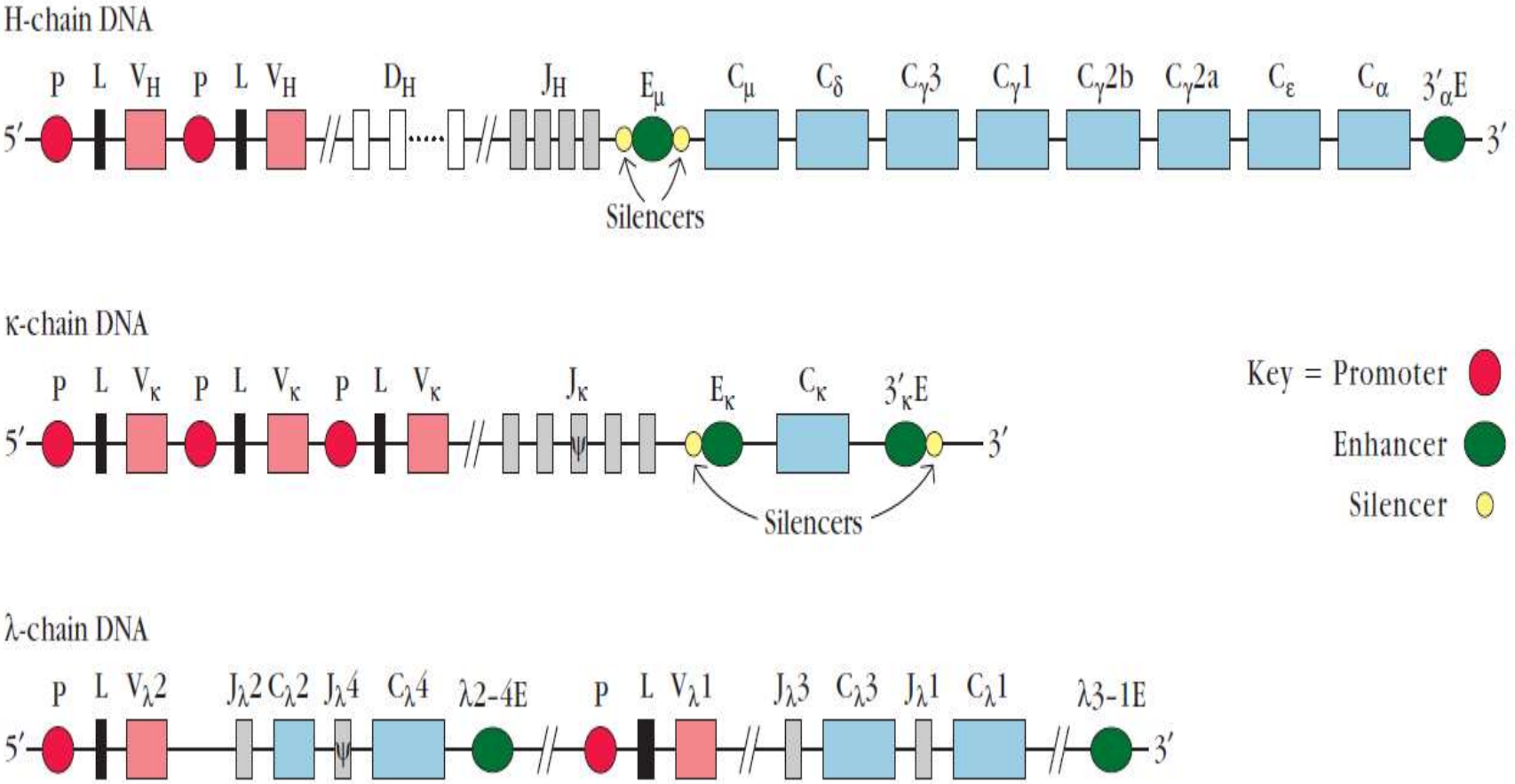
- Occur in Already-Rearranged Gene Segments in variable region.
- Target 1500 nucleotide in VJ and VDJ region.
- frequency approaching 10^{-3} / bp / generation. This rate is at least a hundred thousand-fold higher (hence the name *hypermutation)*than the spontaneous mutation rate, about10^{-8}/bp/generation

Combination of heavy and light chains adds final diversity of variable region

- 8262 possible heavy chain combinations
- 320 light chain combinations
- Over 2 million combinations
- P and N nucleotide additions, junctional flexibility,somatic hypermutation multiply this by 10^4
- Possible combinations over 10^{10}

Class switching among constant regions: generation of IgG, IgA and IgE with same antigenic determinants—idiotypes





Promoters: short nucleotide sequences**,** extending about 200 bp upstream from the transcription initiation site, that promote initiation of RNA transcription in a specific direction

Enhancers: nucleotide sequences situated some distance upstream or downstream from a gene that activate transcription from the promoter sequence in an orientation-independent manner

Silencers: nucleotide sequences that down-regulate transcription, operating in both directions over a distance.