**Lec(3) Immunotechnology MSc Biotechnology**

**Prof. Dr. Ekhlass Noori Ali**

**IMMUNOGLOBULIN GENES**

Immunoglobulin Genes Are Formed Through DNA Rearrangement in B Cells

To contend with the almost unlimited variety of antigens that it may encounter, the human immune system is able to produce antibody molecules with an estimated 108 different unique specificities for antigen. How can so many different antibody proteins be encoded in the genes of every human being? The antigen specificity of an antibody is determined by amino acid sequences within its paired heavy- and light-chain variable domains, which together form the antigen-binding site. To produce antibodies with many different specificities, the immune system must   
have the genetic capability to produce a very large number of different variable domain sequences. The sequence of the constant region, on the other hand, is generally the same for all heavy or light chains of a given immunoglobulin class and has no effect on antigen specificity. In fact, the entire family of immunoglobulin proteins consists of a relatively small number of different constant region domains linked in various combinations with an almost unlimited assortment of variable region sequences. In 1965, Dreyer and Bennett first recognized that these interchangeable combinations of protein domains must be the result of an active reshuffling of gene fragments that took place within the B-cell chromosomes. This was a revolutionary insight, because it implied that a cell could efficiently manipulate its chromosomes to change the structure of genes that it had inherited. And yet this proved to be only a part of the story: Nearly a decade later, Tonegawa made the astonishing discovery that the inherited chromosomes contain no immunoglobulin genes at all, but only the building blocks from which these genes can be assembled. Since that time, studies by many investigators have revealed in detail the extraordinary process through which a B-cell precursor assembles an immunoglobulin gene.

As with most human genes, the information that codes for an immunoglobulin protein is dispersed along the DNA strand in multiple coding segments (exons) that are separated by regions of noncoding DNA (introns); after the gene is transcribed into RNA, introns are removed from the transcript and the exons are joined together by RNA splicing. Unlike nearly all other genes, however, the immunoglobulin DNA sequences that are found in germ cells or other nonlymphoid cell types do not exist as intact, functional genes. This is because the exons that code for variable domains are normally broken up along the chromosome into still smaller gene segments; these segments each lack some of the features needed for proper RNA splicing and so cannot function individually as exons. Before a developing B cell can begin to synthesize immunoglobulin, it must first fuse two or three of these gene segments together to assemble a complete variable region exon. This fusion of gene segments is achieved through a highly specialized process that requires cutting, rearrangement, and rejoining of the chromosomal DNA strands. Only developing lymphocytes possess the enzymatic machinery needed to carry out this process of immunoglobulin gene rearrangement.

**Light-Chain Genes**

The kappa light-chain genes are simplest and are therefore considered first. All of the genetic information needed to produce kappa chains lies within a single locus on chromosome 2 (Figure 7-8). The constant domain of the protein (amino acid residues 109â€“214) is encoded by an exon called CÎº, and only one copy of this exon is found on the chromosome. The sequence encoding any given variable domain, however, is contained in two separate gene segments called the variable (VÎº) and joining (JÎº) segments. The VÎº segment encodes approximately the first 95 amino acids of the variable domain; the shorter JÎº segment codes for the remaining 13 (amino acids 96â€“108). In contrast to the single CÎº exon, multiple VÎº and JÎº segments are present, each with a somewhat different DNA sequence. The five JÎº segments are clustered together near the CÎº exon, whereas approximately 30â€“35 different VÎº segments lie scattered over a region that spans roughly 1 million base pairs (bp) of DNA (less than 1% of the length of chromosome 2). This wide separation between VÎº and JÎº segments is found in the DNA of all nonlymphoid cells. When an immature hematopoietic cell becomes committed to the B-lymphocyte lineage, however, it selects one VÎº and one JÎº segment and fuses these together. This process of V/J joining is accomplished by highly precise enzymatic manipulation of specific sites in the chromosomal DNA. In some instances, this involves precise deletion of all the DNA that normally separates the VÎº and JÎº segments; in others, the two segments are brought together by inverting a portion of the chromosomal strand with no overall loss of DNA (Figure 7-9). The result in either case is that the VÎº and JÎº segments become permanently and covalently joined to one another, side by side on the rearranged chromosome, to form a single continuous exon. Transcription can then begin at one end of the VÎº segment, and pass through both the fused VÎº/JÎº exon and the nearby CÎº

exon. When transcribed together, these two exons contain all of the information needed to synthesize a particular kappa protein. The organization of the kappa genes thus accounts for the unusual properties of this light-chain protein family. As there is only one CÎº exon, all kappa proteins must have identical constant region sequences. On the other hand, because the cell can choose from among many alternative VÎº and JÎº segments, and can join these together in various combinations, a large number of different variable domain sequences can result. For example, 30 VÎº and 5 JÎº segments could in theory give rise to (30 Ã— 5 =) 150 different variable domains. This reshuffling process, known as combinatorial joining, is the most important source of light-chain protein diversity.

Lambda light chains arise from a similar gene complex on chromosome 22. Joining of VÎ» and JÎ» segments occurs in a manner identical to that of the kappa segments. A given chromosome 22, however, may contain up to six slightly different copies of the CÎ» exon (corresponding to various subtypes of lambda protein), each with a nearby JÎ» segment. A VÎ» segment (of which there are approximately 100) may fuse to any of these alternative JÎ» segments, and the resulting VÎ»/JÎ» exon can then be transcribed together with the adjacent CÎ» exon. The B cell selects only one of the available JÎ» segments for V/J joining, and in so doing determines which CÎ» subtype will be expressed (Figure 7-10).

**Heavy-Chain Genes**

All immunoglobulin heavy chains are derived from a single region spanning 685,000 bp on chromosome 14 (Figure 7-11). Each heavy-chain constant region is encoded by a cluster of several short exons. The Âµ constant region, for example, is divided among five exons known collectively as the CÂµ sequence. Constant region (CH) sequences for each of the nine heavy-chain isotypes are arrayed in tandem along the chromosome in the following order: CÂµ, CÎ´, CÎ³3, CÎ³1, CÎ±1, CÎ³2, CÎ³4, CÎµ, CÎ±2; only a single copy of each is present. The six JH segments and approximately 65 VH segments are arranged in a manner analogous to those of the kappa gene. In contrast to the light-chain genes, however, a third type of gene segment, called the diversity (DH) segment, must also be used in forming a heavy-chain variable region. Several of these DH segments (the exact number is unknown), each coding for two or three amino acids, lie between the JH and VH segments on the unrearranged chromosome. In assembling the heavy-chain gene, a B cell must complete two DNA rearrangement events, first bringing together one DH and one JH segment and subsequently linking these to a VH segmentâ€”a sequence termed V/D/J joining.

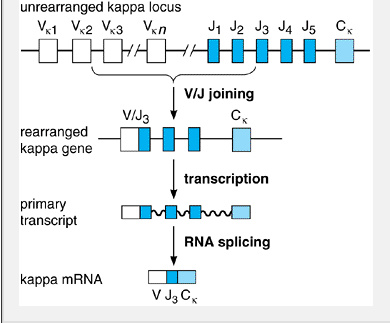
****

Figure 7-8. Assembly and expression of the Îº light-chain locus. A DNA rearrangement event fuses one V segment (in this example, VÎº2) to one J segment (JÎº3) to form a single exon. The V/J exon is then transcribed together with the unique CÎº exon, and the transcript is spliced to form mature Îº mRNA. Note that any unrearranged J segments on the primary transcript are removed as part of the intron during RNA splicing.

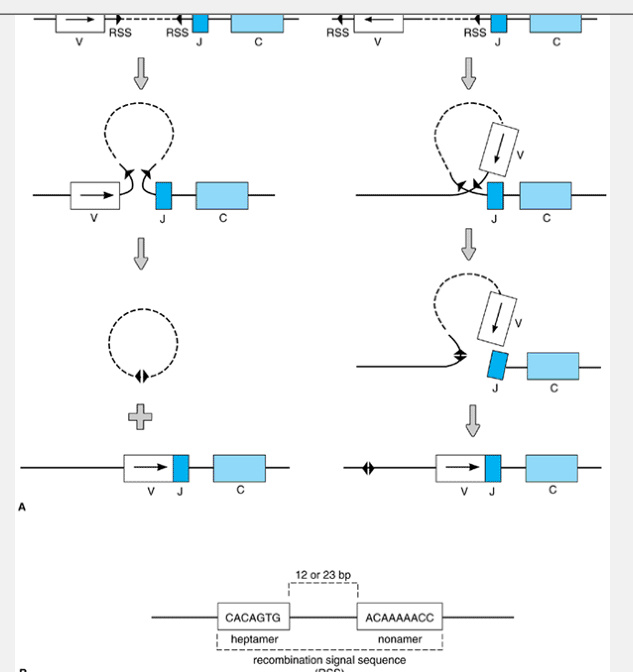


Figure 7-9. Mechanism of immunoglobulin Îº gene rearrangement. A: Site-specific cleavage and religation of the chromosomal DNA is guided by a pair of recombination signal sequences (RSS) flanking V and J gene segments. The chromosome segment undergoes either deletion (left) or inversion (right), depending on the original orientation of the V segment with respect to the J and C segments. If deletion occurs, the DNA that originally separated V and J is released as a covalently closed circle and subsequently is degraded. B: The RSS, consisting of a pair of short DNA sequences (heptamer and nonamer) separated by either 12 or 23 bp of DNA, marks all sites of recombinase action in both heavy- and light-chain genes.

Use of the DH segment greatly increases the amount of heavy-chain diversity that can be produced. For example, 65 VH, 10 DH, and 6 JH segments could give rise to (65 Ã— 10 Ã— 6 =) 3900 different heavy-chain variable domains, and these, when combined with 150 kappa-chain variable domains, could form (150 Ã— 3900 =) nearly 600,000 different antigen-binding sites! Even using a relatively small number of gene segments, then, the immune system can generate enormous antibody diversity through combinational joining.

Additional diversity of immunoglobulin variable regions arises because the V/(D)/J (ie, either V/J or V/D/J) rearrangement process is somewhat imprecise, so that the site at which one segment fuses with another can vary by a few nucleotides. As a result, the DNA coding sequence that remains at the junction between any two segments can also vary. Moreover, during assembly of a heavy-chain gene (but not of light-chain genes), a few nucleotides of random sequence (called N regions) are often inserted at the points of joining between the V, D, and J segments; these insertions are carried out by terminal deoxynucleotidyl transferase (TdT), a nuclear enzyme that is expressed in immature lymphocytes. The variations in gene sequence that result from imprecise joining or from the insertion of N regions contribute substantially to overall antibody diversity. Moreover, these processes affect the sequences within each V/(D)/J exon that code for the third hypervariable region (CDR3) of the heavy- or light-chain

variable domain; hence, the diversity they engender has a disproportionately strong effect on antigen specificity. At the same time, however, these processes greatly increase the risk that two segments may be joined in an improper translational reading frame, resulting in a nonfunctional gene. In practice, such unsuccessful rearrangements occur frequently and generally cannot be reversed or repaired; they represent a cost paid by the immune system in exchange for greater potential gene diversity.

In general, only the CH region located immediately downstream of the V/D/J exon can be expressed. Because the V/D/J exon is originally assembled at a site adjacent to the CÂµ locus, the gene always produces Âµ heavy chains when it is first rearranged. For this reason, naive B lymphocytes always express IgM on their surfaces. Expression of one of the other CH regions can occur only after a cell becomes activated in the periphery, as is described in Chapter 8. One important exception to this rule is the CÎ´ sequence, which lies very near the CÂµ region and is often transcribed along with the V/D/J and CÂµ exons. This produces RNA that can be spliced to yield either Âµ or Î´ mRNA (see Figure 7-11) and so enables the cell simultaneously to express IgM and IgD antibodies that have identical variable domain sequences. Such coexpression of IgM and IgD on the surface membrane is a common phenotype of mature B lymphocytes.

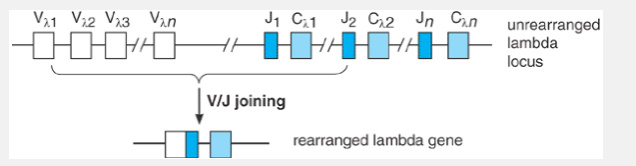


Figure 7-10. Assembly of a Î» light-chain gene. An individual Î» locus contains up to six alternative CÎ» exons, each with a nearby JÎ» segment. In this example, DNA rearrangement fuses VÎ»1 with JÎ»2; the resulting gene produces light chains that contain CÎ»2.

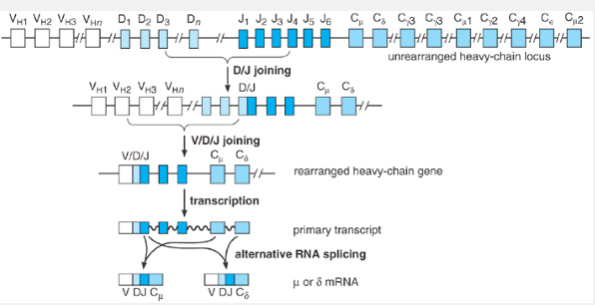


Figure 7-11. Rearrangement and expression of the heavy-chain locus. Unlike the light-chain genes, assembly of a heavy-chain V region exon requires two sequential DNA rearrangement events involving three different types of gene segments. The DH and JH segments are joined first and are then fused to a VH segment. Nine alternative C-region sequences are present; of these, however, only CÂµ and CÎ´ are initially transcribed. The primary transcript can be spliced in either of two ways to generate mRNAs that encode Âµ or Î´ heavy chains with identical V domains. This diagram is highly schematic: Each CH sequence is actually composed of multiple exons whose aggregate length is more than three times longer than that of the V/D/J exon.

**The Molecular Basis of V/(D)/J Rearrangement**

Active gene rearrangements of the type that produce V/(D)/J joining were first thought to be a unique property of the immunoglobulin genes. Subsequently, however, it was found that the genes encoding T-cell antigen receptors (TCRs) also are assembled from germline V, D, J, and C segments through a virtually identical series of DNA rearrangements (see Chapter 9). Among the similarities, for example, the rearrangement sites in both immunoglobulin and TCR genes always coincide with so-called recombination signal sequences (see Figure 7-9B)â€”a pair of short DNA sequences (7 and 9 bp long, respectively) that are located immediately adjacent to each unrearranged V, D, or J segment. It is now thought that rearrangement of both the immunoglobulin and TCR gene families is carried out by the same molecular machinery: a system of enzymes and other proteins known collectively as the V/(D)/J recombinase. The most important components of the recombinase are two nuclear proteins called RAG-1 and RAG-2 (the products of recombination activating genes 1 and 2, respectively), which are expressed in immature B- and T-lineage cells. Acting together, RAG-1 and RAG-2 have the ability to recognize and cleave DNA specifically at a recombination signal sequence, making them the critical ingredients in these early steps

of recombination. By contrast, later steps of recombination, such as religating the various gene segments together, are carried out by cellular enzymes that are also involved in more common forms of DNA repair that occur in all cell types.

Because RAG-1 and RAG-2 (and possibly unknown accessory factors) are expressed only in lymphoid cells, V/(D)/J recombination appears to be absolutely confined to this lineage: no nonlymphoid cell type has yet been proven to manipulate its chromosomes in this way. Until recently, it was also thought that recombinase was expressed only in the early phases of lymphoid development that take place within the lymphopoietic organs. It is true that, by the time a naive B cell emerges from the marrow, it has rearranged both its heavy- and light-chain genes, and has ceased to express recombinase activity, so that it is unable to perform further V/(D)/J rearrangements. We now know, however, that under some circumstances recombinase can become active again in mature lymphocytes, which use it to modify their rearranged genes further in a process called receptor editing (see Chapter 8). The orderly manner in which recombinase is expressed and carries out its tasks thus defines specific stages of lymphocyte ontogeny, as will be discussed in the following chapters.

**Natural Antibiotics**

In both humans and animals, a high proportion of circulating antibodies normally are able to bind pathogen-specific molecules such as bacterial lipopolysaccharide (LPS), phosphatidylcholine, or mannans. Because these antibodies are continually expressed even without specific immunization, can activate complement, and presumably have a role in innate (ie, natural) immunity against pathogens, they are called natural antibodies. Other natural antibodies react with endogenous antigens that are expressed by injured or distressed host cells. For example, when a tissue is deprived of oxygen, preexisting antibodies may bind the injured cells, activate complement, and immediately trigger an inflammatory response. The origins and function of natural antibodies are controversial. In mice, most are IgM antibodies produced by a subset of B lymphocytes called B-1 B cells (see Chapter 8), using a very limited number of VH and VL gene segments. This may indicate that part of the immunoglobulin gene repertoire has evolved to provide rapid, innate recognition of pathogens and distressed tissues.

**Immunoglobulin Gene Rearrangements & B-Cell Malignancy**

Apart from their role in generating antibody diversity, immunoglobulin gene rearrangements are gaining increasing importance in clinical diagnosis and research. Rearrangement of these genes can be detected in biopsies or blood specimens by using a technique known as the Southern blot, and their presence provides a highly sensitive and specific means of diagnosing lymphoid cancers (see Chapter 18). Perhaps more importantly, errors in immunoglobulin gene rearrangement are now thought to contribute to the genesis of several major types of leukemia and lymphoma. For example, the cells of Burkitt's lymphoma, a B-lymphocytic malignancy, usually contain a specific chromosomal abnormality called t(8,14), in which a portion of chromosome 8 has been translocated onto chromosome 14 (Figure 7-12). In this translocation, breakage of chromosome 14 occurs within the immunoglobulin heavy-chain locus, whereas the breakpoint on chromosome 8 coincides with a cellular protooncogene known as c-myc, which encodes the transcription factor c-Myc (see Chapter 1). As a result, the c-myc gene is moved to a position directly adjacent to the heavy-chain gene. It is thought that this proximity to the transcriptionally active heavy-chain locus alters the expression of the protooncogene, and that this, along with other damage to c-myc that can occur during translocation, contributes to malignant transformation. Less commonly, Burkitt's lymphoma may lack t(8,14) and instead exhibit a closely related anomaly in which the c-myc locus is translocated into the kappa or lambda light-chain gene on chromosomes 2 or 22, producing the same effects.

A similar type of chromosomal anomaly, designated t(14,18), is observed in at least 90% of cases of follicular lymphomaâ€” the most common human B-cell

malignancy. In this translocation, the gene on chromosome 18 that encodes the cytoplasmic membrane protein Bcl-2 is moved to a position immediately adjacent to the heavy-chain locus on chromosome 14. B-cells carrying the t(14,18) anomaly express unusually high levels of structurally normal Bcl-2 protein and hence are resistant to being killed by many of the physiologic processes that normally induce apoptosis (Chapter 1). As a result, they tend to accumulate in great numbers and evolve into a malignancy (Chapter 43). In both t(14,18) and the Burkitt's anomalies, the chromosomal breakpoint in affected immunoglobulin loci occurs directly beside a J segment, which strongly implies that each of these translocations results in part from an error in immunoglobulin gene rearrangement.

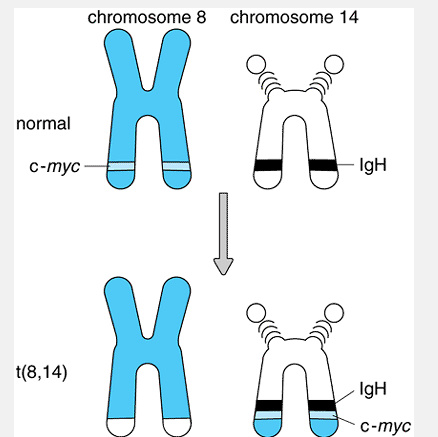


Figure 7-12. The t(8,14) chromosomal anomaly of Burkitt's lymphoma. A reciprocal translocation of genetic material exchanges the distal ends of the long arms of chromosomes 8 and 14. This transposes the c-myc protooncogene from chromosome 8 into the active immunoglobulin heavy-chain locus on chromosome 14 and contributes to the development of a malignancy.