**Ig diversity generation**

- Gene rearrangement: A limited number of **gene segments** recombine to generate a great repertoire.

Nomenclature

Protein: a functional unit that can be one or more polypeptides

Chain: one polypeptide within a protein (e.g. heavy chain)

Domain: a part of a polypeptide chain that folds into a separable structural area (e.g. Ig domain)

Region: in this course, this term is used define areas of Ig and TCR that are either constant or variable; these include domains from both heavy and light chain (or TCR and TCR chain)

Exon: region of genomic DNA containing protein-coding information or untranslated regions of mRNA

Gene segment: this term applies specifically to V(D)J rearrangement; see below

# Antibody Diversity Generation - variable region

\*\* Preview: Ab diversity is generated by four main processes.

1. Multiple copies of the different variable region gene segments recombine (VDJ gene rearrangement)

2. Additional junctional diversity during recombination process

3. Pairing of different heavy and light chain variable regions

4. Somatic hypermutation

**1. V(D)J gene rearrangement**

Variable regions of both light and heavy chains are encoded by different gene segments that join together during DNA recombination. (Fig. 4.16) Unrearranged Ig loci are in **germline configuration**

\* Light chain variable region is encoded by two gene segments that come together; V and J

\* Heavy chain variable region is encoded by three gene segments that come together: V, D, and J

Nomenclature

**V** Makes up most of the variable region sequence (95-101 aa; each variable region, which is composed of one Ig domain, is appx. 110 aa). Each V segment preceded by a leader sequence (L) for ER translocation.

**D** For "diversity"; only in heavy chain; adds extra degree of diversity. A few aa.

**J** For "joining"; next to the C region (up to 13 aa).

Light chain variable segments: V and J

Heavy chain variable segments: VH, DH, and JH

Light chain constant region gene: encoded by C

Heavy chain constant region gene: encoded by C

- There are multiple copies of the three variable region gene segments. (Figs. 4.16, 4.18)

- locus, locus and H chain locus are on different chromosomes.

- They are organized slightly differently.  locus is different because of four paired J-C segments.

- One recombination event for light chain, two for heavy chain due to D segment. (Fig. 4.17)

- V, (D), and J gene segments join by DNA recombination.

- Heavy chain: DJ joining first then V and DJ joining.

- Light chain: VJ joining only (no D segment).

- Different V, (D) and J segments are selected at random during development of individual B cell

- DNA sequences encoding variable region and constant region are separated by great distances in genome (represented by dots in lines of figure); only in B cells are the genes rearranged so that variable and constant region sequences come close together.

- The constant region sequences are not joined to the variable region sequences by DNA recombination; instead they join at the RNA splicing level. Only after RNA splicing are all gene sequences joined as one continuous chain before translation.

Sequence of individual Ig chain production (Fig. 4.34)

Germline (unrearranged) DNA

DJ joining (if it is a heavy chain)

V and (D)J joining

Transcription of a primary transcript RNA

Splicing out of introns and intervening sequences - generation of mRNA

Peptide processing (leader sequence deletion) and glycosylation

Mechanism of DNA rearrangement

- Recombination signal sequence (**RSS**) : Heptamer – 12 or 23 bases of spacer – nonamer (Fig. 4.19). Sequence of heptamer and nonamer is conserved; spacer length but not sequence is conserved

- 12 or 23 base pairs = one or two helical turns in DNA. Brings heptamer and nonamer sequences to one side of the helix; can now be recognized by recombination machinery

\* VDJ recombination occurs on the genes on the same chromosome (therefore not between  and  or between light and heavy chains).

\* VDJ recombination occurs between a gene segments with a 12mer RSS and another gene segment with a 23mer RSS (The 12/23 rule). (Fig. 4.20)

- DH flanked by12mer RSS and VH and JH flanked by 23mer RSS: VH and JH therefore do not join together.

Machinery of DNA recombination

- V(D)J recombination occurs only in B and T cells (T cell receptors undergo VDJ recombination, but with different sets of V, D and J gene segments for TCRand chains).

Definitions: endonuclease vs. exonuclease🡪 endonuclease can cleave intact double-stranded DNA; exonuclease chews away from the ends.

**V(D)J recombinase**: - Complex of several enzymes that act in concert to carry out variable region gene recombination. (Fig. 4.20)

- Composed of:

*\** **RAG-1 and RAG-2** gene products (recombination-activating gene): RAG proteins form a heterodimeric endonuclease complex, directly recognize RSS (one 12 and one 23)

- the two RSS are brought together and then cleavage and DNA repair proceeds

\* RAG1/2 are expressed only in developing lymphocytes (bone marrow pro-B and pre-B cells; and thymocytes). This explains cell type specificity.

\* Normal cellular DNA cleavage and repair machinery involved in DNA repair (found in many other cell types): **DNA-PK**, Ku70/Ku80, DNA ligase IV, XRCC4, Artemis.

Humans lacking DNA-PK have severe combined immunodeficiency (Scid) because of no mature B cells or T cells; also more sensitive to DNA damage in other cells

-products of reaction are **signal joint** and **coding joint**

-Joining of the coding joint is not precise – additional level of diversity

**2. Junctional diversity**

- CDR1 and 2: encoded by V segment. Note that there is some variability even in the framework regions of different V segments (Fig. 4.8 from last lecture)but that the greatest difference between V segments is in the nucleotides encoding the CDR1 and CDR2 residues. When V gene segments duplicated and diverged, natural selection preserved sequence of framework regions more than CDR loops.

- CDR3 is encoded by the sequences at the junction between V and J in light chain/ partially by D in heavy chain.

- CDR3 diversity is aided by addition and/or deletion of nucleotides at the junction of gene segments.

- Added nucleotides are called P-nucleotides and N-nucleotides. (Fig. 4.21)

- Because of inprecise nature of joining, 2/3 of recombination events will not produce an exon that preserves the reading frame and no functional protein will be produced: **non-functional rearrangement**.

- **Allelic exclusion**: once a given Ig chain is successfully rearranged and functional protein produced, no more gene rearrangement occurs on that chain (other chromosome or second light chain locus). Ensures single Ig specificity per B cell.

**3. Pairing of different heavy and light chain variable regions**

- Different pairing can generate different specificity: a given light chain can pair with any different heavy chain and generate different specificity.

**\*** How many different specificities can the above mechanisms generate?

1. Multiple copies of the different variable region gene segments

: 35 V x 5 J = 175

: 30 V x 4 J = 120

295 possible light chains

40 VH x 23 DH x 6 JH = 5,520 possible heavy chains

2. Pairing of different H and L chain variable regions

295 x 5,520 = 1.6x106 different Ab specificities

1 and 2 = Combinatorial Diversity

Further increased diversity due to:

3. Additional diversity during recombination process itself (junctional diversity)

4. Somatic hypermutation

An individual is thought to have 109 - 1011 Ig specificities

Co-expression of IgM and IgD on naïve mature B cells

- All B cells start out by producing a transcript in which the VDJ exon is linked to the µ constant exons that are immediately 3’ in the locus (Fig. 4.22): B cells in late maturation stage in the bone marrow have IgM on cell surface.

- Naive mature B cells that have left the bone marrow but have not seen Ag express both IgM and IgD on cell surface. This is not from gene rearrangement; rather, a primary RNA transcript actually has both C and C, which are encoded by exons that are adjacent in the heavy chain locus (Fig. 4.22). **Alternative splicing** allows the cells to generate two species of mRNA from the same transcriptional unit; and therefore to express both IgD and IgM. (Fig. 4.23) The function of IgD is not clear; very low amounts of secreted IgD in plasma and lymph. IgD has flexible hinge region (not in IgM) so might increase spectrum of antigen recognition.

- surface expression of IgM and IgD requires association with two proteins called Ig and Ig (Fig. 4.25). These transmembrane proteins have longer cytoplasmic tails than Ig proteins and provide signaling function to the BCR.

Transmembrane vs. secreted immunoglobulins (Fig. 4.26)

An antibody can be either secreted by B cells or remain expressed on the cell surface.

- IgM initially on the surface of B cells.

- After an Ag encounter, some B cells differentiate into cells that secrete immunoglobulins (called plasma cells).

- Example: C is actually a cluster of exons that encode different Ig domains of a heavy chain (Fig. 4.26). Each C gene has secretion-coding (SC) and membrane-coding (MC) exons. After the primary transcript is made, alternative splicing occurs and either SC or MC is lost. The remaining exon determines whether Ig is membrane-bound or secreted.