

## Le répertoire B humain : expression normale et pathologique

Sylvie Huck  
Moncef Zouali

L'extraordinaire diversité du répertoire des immunoglobulines découle de mécanismes génétiques originaux : la recombinaison VDJ, l'ajout de nucléotides au hasard dans les zones de jonction, des réarrangements secondaires et des hypermutations somatiques aux stades tardifs de développement des cellules B. L'étude récente des gènes *V* a permis de les classer en familles d'après l'analogie de leurs séquences et de découvrir de nombreux pseudogènes et des gènes orphelins, situés sur des chromosomes différents. L'assemblage combinatoire par hasard du gène *V* peut conduire à la production d'autoanticorps mais, normalement, les cellules B qui les hébergent subissent une délétion clonale, sont rendues silencieuses, ou modifient leur spécificité par un mécanisme d'« édition » du récepteur. L'incapacité de cellules B à modifier ainsi leur récepteur pourrait être à l'origine de certaines maladies auto-immunes. Enfin, l'infection par le VIH modifie aussi l'expression du répertoire des cellules B : l'interaction de la protéine d'enveloppe gp120 avec les immunoglobulines contenant une partie variable  $V_{H3}$  stimule la sécrétion d'immunoglobulines par les cellules B  $V_{H3}^+$  dont une large partie du répertoire est mobilisée, ce qui rappelle l'action des superantigènes.

### ADRESSE

S. Huck : chargée de recherche à l'Inserm. Laboratoire d'immunogénétique moléculaire, UMR9942, IGMM, 1919, route de Mende, BP 5051, 34033 Montpellier Cedex 01, France. M. Zouali : directeur de recherche à l'Inserm. Institut Pasteur, département d'immunologie, 28, rue du Docteur-Roux, 75015 Paris, France.

La réponse immunitaire des vertébrés supérieurs comprend un ensemble complexe d'événements cellulaires et moléculaires destinés à neutraliser spécifiquement l'agent infectieux, la greffe incompatible ou la cellule tumorale. Plastique et dynamique, le système immunitaire fascine par sa spécificité, sa diversité et son adaptabilité. Il y a environ un

siècle, déjà, les dessins de Paul Ehrlich montraient des entités en forme de courte clé à molette, appelées immunoglobuline (Ig) ou anticorps (Ac), avec une complémentarité structurale pour leur cible, l'antigène (Ag), leur permettant de s'y fixer. On sait désormais que l'Ig est un dimère d'une chaîne lourde (H) et d'une chaîne légère (L), chacune étant organisée en plusieurs domaines, ou

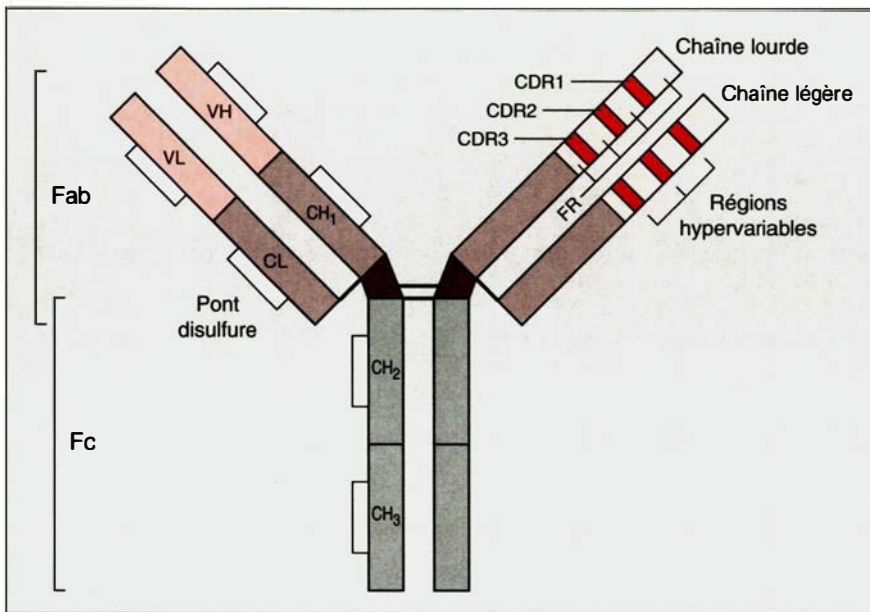


Figure 1. **Schéma d'une Ig d'isotype IgG.** L'Ig est un dimère d'une chaîne légère (L) – qui peut être d'isotype  $\kappa$  ou  $\lambda$  – associée à une chaîne lourde (H) – qui peut être d'isotype  $\mu$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$ ,  $\alpha$ . C'est la chaîne lourde qui confère sa classe à l'Ig (IgM, IgD, IgE, IgG, IgA), ainsi que ses caractéristiques fonctionnelles. Toutes les chaînes polypeptidiques sont associées entre elles par des liaisons covalentes de type pont disulfure (S-S). La chaîne légère comprend un domaine variable et un domaine constant; la chaîne lourde comprend un domaine variable et trois ou quatre domaines constants selon l'isotype. La structure en domaines est favorisée par des liaisons disulfures intrachaîne qui forment un pont entre le début et la fin d'un domaine. L'Ac contient deux sites identiques de reconnaissance, ou paratopes, pour l'antigène. Chaque paratope, formé par les domaines variables VH et VL, est plus spécifiquement constitué par les régions hypervariables CDR1, CDR2 et CDR3 (cette dernière correspondant à la jonction V-J ou V-D-J). Les CDR (en rouge) sont localisés entre les régions charpentes (FR, en rose).

boucles, de structure tridimensionnelle bien définie et comprenant un pont disulfure reliant le début et la fin d'un même domaine (figure 1). Le site de reconnaissance de l'Ag est localisé dans les domaines variables situés en position amino-terminale de la protéine, alors que les domaines constants assurent les fonctions effectrices. Les Ac d'un individu donné sont dotés d'un potentiel de reconnaissance considérable; la taille de notre génome étant insuffisante pour la synthèse de la quasi-infinité d'Ac possibles, ce système utilise, pour la production de la diversité, des mécanismes génétiques tout à fait originaux: le domaine variable (V) et le domaine constant d'une même chaîne polypeptidique d'Ig

sont codés par un ensemble de segments d'ADN, que des réarrangements vont assembler [1, 2]. Ainsi, le domaine V polypeptidique est codé par un gène V (choisi parmi plusieurs dizaines de gènes) réarrangé sur l'un des petits segments d'ADN appelés de jonction (J), et aussi sur l'un des segments dits de diversité (D) dans le cas de la chaîne H (figure 2). Divers gènes V réarrangés peuvent être combinés à la même région constante, permettant ainsi la synthèse d'Ac différents, reconnaissant des Ag différents, mais dotés de fonctions biologiques identiques. La diversité est renforcée par le fait que n'importe quelle chaîne H peut être associée à n'importe quelle chaîne L. L'imprécision des jonctions D-J, V-DJ ou

V-J amplifie cette variabilité. Deux autres facteurs contribuent aussi à augmenter la diversité. D'abord, il y a délétion ou addition de nucléotides au hasard dans les zones de jonction, créant ainsi une diversité de type N. Il en résulte que chaque lymphocyte exprimera une diversité unique de sa troisième région hypervariable, ou CDR3, la séquence nucléotidique de cette région servant d'ailleurs d'empreinte moléculaire pour suivre l'évolution clonale des cellules B. Ensuite, des hypermutations somatiques, dont le mécanisme reste inconnu, apparaissent aux stades tardifs du développement des cellules B et pendant la réponse secondaire [3]. C'est surtout au niveau des centres germinatifs du follicule lymphoïde que les lymphocytes portant une Ig de forte affinité pour l'Ag sont positivement sélectionnés et subissent une expansion clonale qui aboutira à la sécrétion de l'Ac spécifique, et aussi à une mémoire immunologique qui caractérise l'immunité des vertébrés supérieurs, et dont les mécanismes restent à l'étude.

Au cours des dernières années, nos connaissances des mécanismes moléculaires de la génération de la diversité des Ac humains ont progressé de façon significative. Nous discuterons ici les données les plus récentes concernant le répertoire B humain à l'état normal, mais aussi dans des situations pathologiques, en mettant l'accent sur les maladies auto-immunes systémiques et l'infection par le VIH.

### Chaînes lourdes : localisation télomérique du locus VH humain sur le chromosome 14

Plusieurs laboratoires se sont intéressés à l'étude de l'organisation et du contenu des loci VH et DH. L'étude détaillée de YAC et de cosmides chevauchants a permis d'établir une première carte du locus VH, situé sur la région télomérique du chromosome 14 au niveau de 14q32.39, directement en amont des gènes JH, couvrant 0,8 Mb et comprenant 64 gènes VH d'un répertoire total estimé alors à environ 120 éléments VH [4]. Il avait été montré auparavant que cette région est le siège de translocations chromosomiques associées à

des transformations néoplasiques du lymphocyte B. Ces translocations, qui semblent être la conséquence de l'instabilité chromosomique associée aux réarrangements fréquents des gènes V des Ac sont, par exemple, bien connues dans les lignées cellulaires dérivées du lymphome de Burkitt où un allèle a subi une translocation chromosomique t(8;14) (q24;q32), ce qui donne naissance à un chromosome contenant la région télomérique du bras long du chromosome 14 (*m/s n° 2, vol. 3, p. 112*). Ce type de translocation peut être utilisé pour identifier de nouveaux gènes grâce à des « marches » le long des points de cassure chromosomique. C'est ainsi que la translocation t(8;14) (q24;q32) a été récemment mise à profit pour cloner la partie distale du locus VH et, ainsi, étendre la carte de 150 kb vers l'extrémité télomérique de ce locus [5] (*figure 3*). Il en résulte que la distance entre les gènes JH et le gène VH le plus distal est de seulement environ 1100 kb, et le nombre total moyen de gènes VH par haplotype est d'environ quatre-vingt-dix [6], nombre plus faible qu'initialement présumé.

## RÉFÉRENCES

1. Tonegawa S. Somatic generation of antibody diversity. *Nature* 1983; 302: 575-81.
2. Sigaux F. Physiologie et pathologie de la recombinaison V(D)J. *médecine/sciences* 1994; 10: 995-1005.
3. Berek C, Milstein C. Mutation drift and repertoire shift in the maturation of the immune response. *Immunol Rev* 1987; 96: 23-41.
4. Matsuda F, Shin EK, Nagaoka H, Matsumura R, Haino M, Fukita Y, Taka-ishi S, Imai T, Riley JH, Anand R, Soeda E, Honjo T. Structure and physical map of 64 variable segments in the 3' 0.8-megabase region of the human immunoglobulin heavy-chain locus. *Nature Genet* 1993; 3: 88-94.
5. Cook GP, Tomlinson IM, Walter G, Riethman H, Carter NP, Buluwela L, Winter G, Rabbitts TH. A map of the human immunoglobulin VH locus completed by analysis of the telomeric region of chromosome 14q. *Nature Genet* 1994; 7: 162-8.
6. Cook GP, Tomlinson IM. The human immunoglobulin VH repertoire. *Immunol Today* 1995; 16: 237-42.
7. Zachau HG. The immunoglobulin kappa locus-or-what has been learned from looking closely at one-tenth of a percent of the human genome. *Gene* 1993; 135: 167-73.
8. Fripiat JP, Williams SC, Tomlinson IM, Cook GP, Cherif D, Le Paslier D, Collins JE, Dunham I, Winter G, Lefranc MP. Organization of the human immunoglobulin lambda light-chain locus on chromosome 22q11.2. *Hum Mol Genet* 1995; 4: 983-91.
9. Sasso EH, Willems vDK, Bull AP, Milner EC. A fetally expressed immunoglobulin VH1 gene belongs to a complex set of alleles. *J Clin Invest* 1993; 91: 2358-67.

répartis sur une région de 800 kb [8]. Les gènes Vλ ont été classés dans dix familles, elles-mêmes réparties en trois groupes dans le locus. Le groupe le plus distal par rapport au locus Cλ comprend les familles VλIV, λVI, λVIII et λX, le groupe central comprend VλI, λV, λVII et λIX, et enfin le groupe proximal comprend VλII et λIII [8]. Globalement, les gènes d'une même famille ne sont donc pas éparpillés tout au long du locus, contrairement à la situation des gènes VH et Vκ humains.

## Polymorphisme du locus VH humain

L'hétérogénéité génomique est une propriété importante des gènes des Ig humaines. Les études de RFLP et de séquençage ont initialement suggéré l'existence d'un petit nombre d'allèles de gènes VH. Dans d'autres études, des polymorphismes d'insertion impliquant plusieurs gènes VH ont été détectés dans une partie de la population [6]. C'est ainsi que le gène télomérique 1-69 est l'un des gènes VH les plus polymorphes, avec 13 allèles connus. Certains individus possèdent une copie unique du gène; chez d'autres le gène 1-69 est dupliqué [9]. Il est remarquable que, alors que des pseudogènes ont été identifiés sur le chromosome 14, on ignore si le rapport du nombre des gènes fonctionnels à celui des pseudogènes est identique pour tous les haplotypes. Au total, bien que l'étendue du polymorphisme des gènes des Ac humains reste inconnue, l'hétérogénéité du locus VH semble très élevée à la fois pour les gènes fonctionnels et pour les pseudogènes, avec des polymorphismes de délétion et d'insertion et des allélismes.

Non seulement ce polymorphisme diversifie considérablement le répertoire entre individus, mais il pourrait avoir une signification fonctionnelle. Il est par exemple possible qu'il joue un rôle dans l'affinité des Ac pour leurs ligands. En effet, même les mutations dans les régions charpentes (FR) de l'Ig (*figure 1*) peuvent affecter l'affinité de la liaison Ag-Ac [10]. L'expression de variants alléliques particuliers pourrait aussi avoir des conséquences sur l'association entre chaînes H et L, ou sur l'interaction des régions V des Ac humains avec les

## Chaînes légères : duplication du locus Vκ et complexité du locus Vλ humains

Le locus Vκ comprend environ 1600 kb et son organisation est maintenant bien élucidée, avec deux portions de 800 kb chacune, l'une distale, l'autre proximale par rapport au locus Jκ, séparées par une région non codante d'environ 800 kb (*figure 4*). Les 76 gènes germinaux de ce locus peuvent être classés en quatre familles (Vκ1 à Vκ4) qui comprennent, respectivement, 29, 27, 14 et un gènes. Les cinq gènes des trois autres familles Vκ5, Vκ6, Vκ7 ne sont pas retrouvés dans les réarrangements productifs. Parmi les gènes Vκ connus, 32 seulement sont fonctionnels, 25 sont des pseudogènes et 16 contiennent des déficits mineurs. Ces derniers sont susceptibles d'être exprimés chez certains individus [7]. Localisé en amont de l'ensemble des régions constantes dont le nombre varie de 7 à 10 gènes Cλ, chacun précédé par un segment de jonction Jλ, le locus Vλ contient plus de 50 gènes

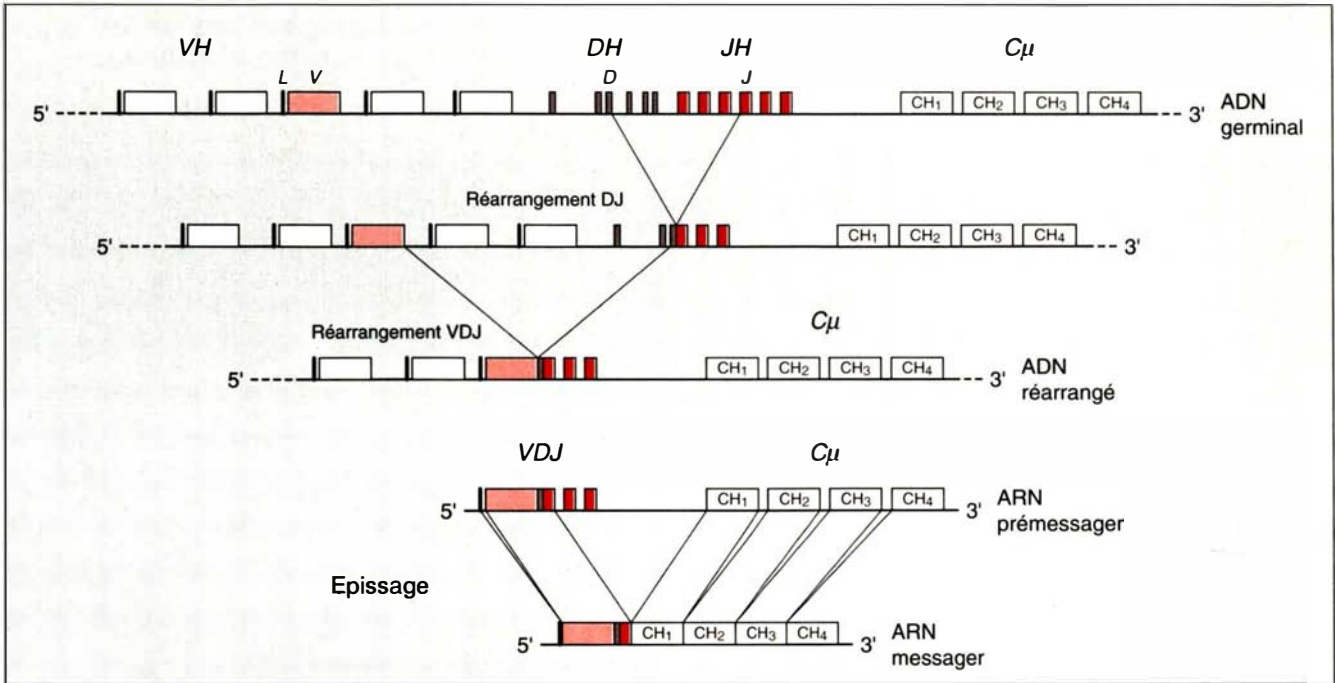


Figure 2. **Événements moléculaires conduisant à la synthèse d'une chaîne H d'immunoglobuline humaine.** L'extrémité télomérique du chromosome 14 comprend, de 5' en 3', un ensemble de gènes variables VH, un groupe de segments de diversité DH, puis un ensemble de segments de jonction JH suivis de gènes constants C. Au cours du développement du lymphocyte B, l'ADN, au départ en configuration germinale, subit les réarrangements au niveau des loci des IG. La première étape du réarrangement s'effectue entre l'un des segments DH, parmi la trentaine supposée, et l'un des segments JH, parmi les 6 gènes JH fonctionnels. Le réarrangement entraîne une délétion de l'ADN intermédiaire situé entre les deux segments choisis. La seconde étape consiste à réarranger l'un des gènes VH (et sa séquence signal L) sur le produit DH-JH. L'ADN se trouve alors en configuration réarrangée, spécifique du lymphocyte B. L'ADN est transcrit en ARN pré-messager. Celui-ci est épissé, et l'intron entre le peptide signal et le gène VH, ainsi que l'intron entre le produit VDJ et le gène C, sont délévés. La chaîne lourde est enfin traduite, exportée vers la lumière du réticulum endoplasmique où elle formera une Ig par association à la chaîne légère et dimérisation.

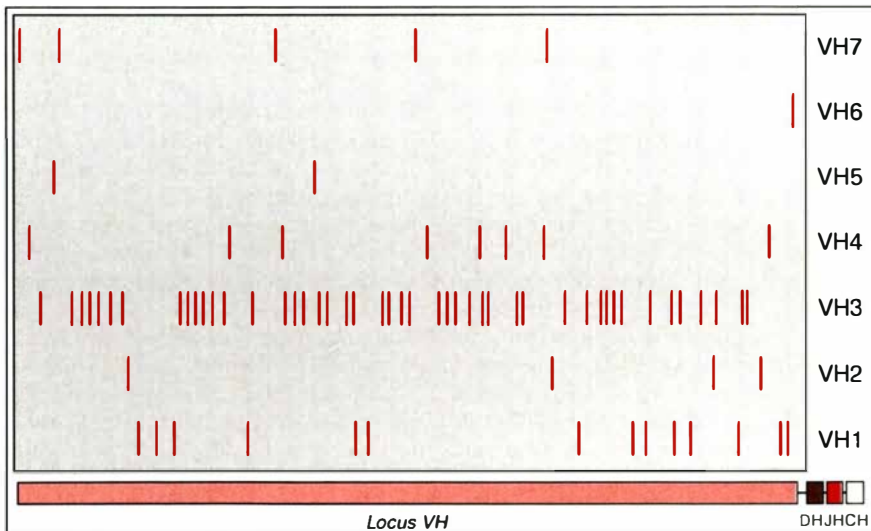


Figure 3. **Organisation des sept familles de gènes VH humains sur le chromosome 14.** Cette figure, dans laquelle les distances sont approximatives, représente les gènes fonctionnels et les pseudogènes présents sur le chromosome 14. Certains gènes VH peuvent être dupliqués ou absents selon les haplotypes (non représenté). La dispersion des membres des différentes familles contraste avec le regroupement des gènes de familles VH chez la souris. Il est à noter que le gène VH6, le plus proximal du locus constant, n'est pas le plus fréquemment exprimé. (D'après [4, 6].)

## RÉFÉRENCES

10. Foote J, Winter G. Antibody framework residues affecting the conformation of the hypervariable loops. *J Mol Biol* 1992; 224: 487-99.
11. Zouali M. B-cell superantigens: implications for selection of the human antibody repertoire. *Immunol Today* 1995; 16 : 399-405.
12. Yang PM, Olsen NJ, Siminovitch KA, Olee T, Kozin F, Carson DA, Chen PP. Possible deletion of a developmentally regulated heavy-chain variable region gene in autoimmune diseases. *Proc Natl Acad Sci USA* 1990; 87 : 7907-11.
13. Kodaira M, Kinashi T, Umemura I, Matsuda F, Noma T, Ono Y, Honjo T. Organization and evolution of variable region genes of the human immunoglobulin heavy chain. *J Mol Biol* 1986; 190 : 529-41.
14. Weill JC, Reynaud A. The chicken B cell compartment. *Science* 1987; 238 : 1094-8.
15. Sanz I. Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions. *J Immunol* 1991; 147 : 1720-9.
16. Childs G, Maxson R, Cohn RH, Kedes L. Orphans : dispersed genetic elements derived from tandem repetitive genes of eucaryote. *Cell* 1981; 23 : 651-63.
17. Robinson MA, Mitchell MP, Wei S, Day CE, Zhao TM, Concannon P. Organization of human T-cell receptor beta-chain genes: clusters of V beta genes are present on chromosomes 7 and 9. *Proc Natl Acad Sci USA* 1993; 90 : 2433-7.

superantigènes B [11], qui seront discutées plus loin. Enfin, puisque les polymorphismes alléliques des gènes *VH* peuvent conduire à un nombre élevé d'haplotypes qui codent pour différentes protéines et qui varient d'un individu à l'autre, il est possible que de tels polymorphismes puissent être associés à des prédispositions génétiques à développer certaines maladies [12]. Il convient donc de mettre au point des stratégies pour étudier ces polymorphismes et pour évaluer leur emplacement, leur expression et la ségrégation des différents allèles. De tels allèles de gènes *V* pourraient devenir des cibles d'immuno-interventions spécifiques dans les maladies auto-immunes et dans certaines maladies allergiques.

## Pseudogènes *V* humains

L'organisation des *loci* des Ig est propre à chaque espèce, tant dans le nombre de familles de gènes *V*, ou dans le nombre de gènes constants, que dans l'agencement des structures en tandem : *Vn-Jn-Cn*, *Vn-J-C (Vn-J-C)n*, *(Vn-Jn-C)n*, etc. Les gènes des Ig subissent donc de fréquentes duplications qui favorisent l'émergence et l'accumulation de pseudogènes. Pour le *locus VH*, il y a environ 30 pseudogènes *VH* sur le chromosome 14 [4, 5]. En ce qui concerne le *locus κ*, qui comprend 76 gènes germinaux, seuls 32 gènes *Vκ* sont potentiellement fonctionnels [7]. Le nombre de pseudogènes est donc élevé, avec environ 45 % de l'ensemble des gènes *VH* [6] et *Vκ* [7]; on estime que les pseudogènes *Vλ* sont également abondants. Alors que certains des pseudogènes *VH* ont largement divergé par rapport aux gènes fonctionnels, d'autres sont assez conservés et la distribution de mutations entre les régions hypervariables et la charpente y est analogue à celle observée pour les gènes fonctionnels [13], ce qui peut suggérer que des forces sélectives identiques opèrent sur les gènes fonctionnels et les pseudogènes, vraisemblablement au travers d'un mécanisme de conversion génique. Ce mécanisme est d'ailleurs utilisé pour la création de la diversité des Ig des oiseaux [14] et pourrait être impliqué dans la génération de la grande diversité des CDR3 des Ac humains [15].

## L'énigme des gènes *V* humains « orphelins »

Il avait été rapporté que certains membres de familles multigéniques, dits gènes « orphelins », sont isolés de leur famille originelle [16]. Au moins 25 gènes *Vκ* ont été localisés eux aussi en dehors du *locus κ* fonctionnel (qui se trouve sur le bras court du chromosome 2), à la fois sur le bras long du chromosome 2 et sur d'autres chromosomes [7] (Tableau I). De même, à côté des 51 gènes *Vβ* fonctionnels du récepteur T, qui sont groupés en 24 familles le long du bras long du chromosome 7 (7q32-35), environ 10% des gènes *Vβ* connus sont situés sur le chromosome 9 [17]. Deux groupes indépendants ont récemment démontré que des gènes *VH* humains peuvent aussi avoir d'autres emplacements chromosomiques [6]. Grâce à l'utilisation de cellules hybrides d'origine humaine et animale, et de l'hybridation *in situ* à fluorescence, plusieurs gènes *VH* et deux groupes de gènes *DH* ont été mis en évidence sur les chromosomes 15 (15q11-12) et 16 (16p11) (Tableau I). Il semble que même les gènes *VH* orphelins sont polymorphes. Ainsi, la duplication des gènes *VH* orphelins du chromosome 16 semble absente chez certains individus. De même, chez 11 donneurs non apparentés, le groupe de gènes orphelins *DH (D5-b)* n'a été détectable que chez deux individus. Cette fréquence des gènes orphelins soulève la question du caractère fonctionnel de ces gènes isolés de leur contexte chromosomique. Certains gènes orphelins *VH* et *DH* des chromosomes 15 et 16 ont des phases de lecture ouvertes et des signaux de recombinaison fonctionnels. Des transcrits germinaux de deux gènes orphelins *VH*, *HC16-15* et *HC16-16*, ont par ailleurs été isolés du foie fœtal humain [18]; ces gènes pourraient donc être accessibles aux recombinaisons qui catalysent les réarrangements VDJ. Par ailleurs, un réarrangement inhabituel *VH-DH* a été trouvé dans une lignée cellulaire B humaine [19], alors que normalement le réarrangement *DH-JH* précède l'assemblage *VH-DH/JH*. Selon cette voie de réarrangement, le produit d'un assemblage *VH-DH* se produisant sur le chromosome 15 pourrait

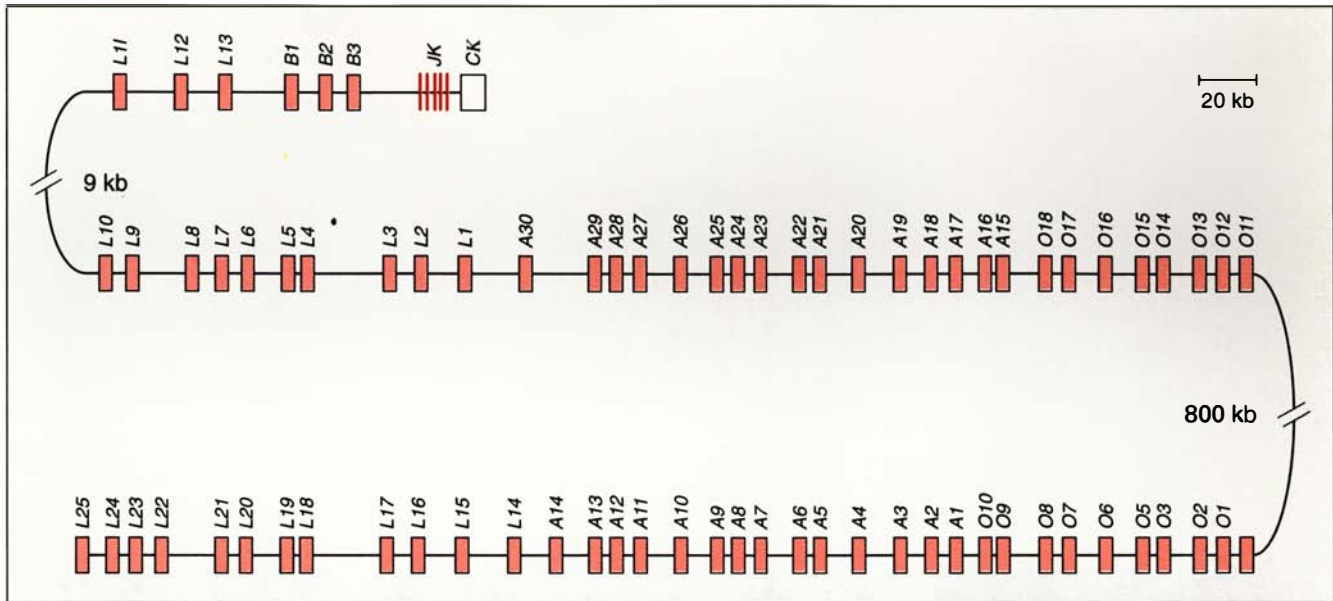


Figure 4. **Organisation dupliquée du locus Vκ humain.** Le locus κ est formé de deux régions dupliquées identiques à 98%. Un premier ensemble d'environ 850kb contenant 40 gènes Vκ est situé directement en amont du gène unique constant Cκ, lui-même précédé par 5 segments Jκ. Le second ensemble, d'environ 850 kb, est séparé du précédent par environ 800 kb, et contient 36 gènes Vκ en orientation opposée par rapport à Cκ. (Adapté à partir de [7].) Dans la nomenclature des gènes Vκ, chaque gène est d'abord désigné par une lettre (A, O, L, B) qui correspond à une région bien définie du locus.

Localisation chromosomique	Locus VH		Locus DH		Locus Vκ	
	Gènes fonctionnels	Gènes orphelins	Gènes fonctionnels	Gènes orphelins	Gènes fonctionnels	Gènes orphelins
Chromosome 14	Télomère		Télomère			
Chromosome 15		15q11-12		15q11-12		
Chromosome 16		16p11				
Chromosome 2					Bras court	
Chromosome 2						Bras long
Chromosome 1						Plusieurs
Chromosome 22						Plusieurs

Par isolement de clones génomiques et par les techniques d'hybridation de fluorescence in situ, de nombreux gènes variables ont été localisés en dehors de leur contexte chromosomique habituel. Ces gènes, dits orphelins, se sont révélés être majoritairement des pseudogènes. Douze gènes orphelins κ sont localisés sur le bras long du chromosome 2, le bras court portant le locus κ fonctionnel, et proviendraient d'une inversion péricentrique du chromosome.

Les chromosomes contenant les loci fonctionnels sont en gras.

devenir la cible de réarrangements ultérieurs sur des segments JH ou DHJH du chromosome 14, d'où une fonction potentielle pour ces gènes orphelins. La mise en évidence de ces gènes orphelins et leur dispersion chromosomique illustre la plasticité génomique des loci V humains.

### Mécanismes génétiques particuliers aux anticorps humains

Historiquement, ce sont les mécanismes de la génération de la diversité des Ac murins qui ont été élucidés les premiers. Par la suite, des

séquences nucléotidiques spécifiques de réarrangements des gènes des Ig ont été identifiées. Pour les chaînes lourdes, celles en aval d'un VH et en amont d'un JH sont similaires et contiennent un espaceur de 23 nucléotides (type 23), alors que celles bordant un segment D contiennent

ment un espaceur de 12 nucléotides (type 12). Le réarrangement se fait par reconnaissance, appariement, puis jonction entre une séquence de type 12 et une autre de type 23. Outre les mécanismes mentionnés dans l'introduction de cet article, d'autres événements participant à la diversité du CDR3 ont été décrits plus récemment. En effet, des adjonctions de nucléotides palindromiques (P) au niveau des jonctions, et des fusions entre deux gènes *DH* ont été observées. Les Ac humains disposent de mécanismes supplémentaires. D'abord des segments DIR (longs segments D avec d'irréguliers signaux de recombinaisons), flanqués par des séquences de réarrangements de type 12 et/ou 23, sont aussi utilisés. Ces segments, qui peuvent s'assembler avec un segment *DH*, représentent environ 7% du répertoire des régions CDR3 [15]. On note aussi l'expression des gènes *DH* dans les trois phases de lecture, alors que chez la souris une seule phase est privilégiée et l'utilisation très fréquente de fusions directes ou inversées. Il faut enfin souligner que la formation d'environ 40% des régions CDR3 ne peut pas être expliquée par des mécanismes conventionnels de réarrangements. Cette diversité pourrait trouver son origine dans un processus de conversion génique [15]. Il en résulte un répertoire très large d'Ac, avec plus de 10<sup>11</sup> spécificités, ce qui dépasse le nombre de cellules B disponibles à un moment donné du développement de l'individu.

## RÉFÉRENCES

18. Cuisinier AM, Gauthier L, Boubli L, Fougereau M, Tonnelle C. Mechanisms that generate human immunoglobulin diversity operate from the 8th week of gestation in fetal liver. *Eur J Immunol* 1993; 23: 110-8.
19. Shin EK, Matsuda F, Fujikura J, Akamizu T, Sugawa H, Mori T, Honjo T. Cloning of a human immunoglobulin gene fragment containing both VH-D and D-JH rearrangements: implication for VH-D as an intermediate to VH-D-JH formation. *Eur J Immunol* 1993; 23: 2365-7.
20. Hillson JL, Oppliger IR, Sasso EH, Milner ECB, Wener MH. Emerging human B cell repertoire. Influence of developmental stage and interindividual variation. *J Immunol* 1992; 149: 3741-52.
21. Zouali M. Development of human antibody variable genes in systemic autoimmunity. *Immunol Rev* 1992; 128: 73-99.
22. Goodnow CC, Crosbie JS, Adelstein S, Lavoie TB, Smith-Gill SJ, Brink RA, Pritchard-Briscoe H, Wotherspoon JS, Loblay RH, Raphael K, Trent RJ, Basten A. Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature* 1988; 334: 676-82.
23. Nemazee DA, Burki K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* 1989; 337: 562-6.
24. Hartley SB, Cooke MP, Fulcher DA, Harris AW, Cory S, Basten A, Goodnow CC. Elimination of self-reactive B lymphocytes proceeds in two stages: arrested development and cell death. *Cell* 1993; 72: 325-35.
25. Cyster JG, Hartley SB, Goodnow CC. Competition for follicular niches excludes self-reactive cells from the recirculating B-cell repertoire. *Nature* 1994; 371: 389-95.

vée au cours des étapes précoces du développement pourrait refléter la sélection d'un certain nombre de gènes *V* par des déterminants du soi, présents au cours de la vie fœtale. Les études moléculaires du répertoire VH primaire au cours du deuxième trimestre de la vie fœtale sont en accord avec cette hypothèse [18, 20]. Ces analyses montrent qu'un petit nombre de gènes sont exprimés par les cellules pré-B du fœtus humain: l'élément *DHQ52* (figure 5), le gène *JH1*, certains gènes *VH*, et des régions CDR3 beaucoup plus courtes que chez l'adulte caractérisent ce répertoire. Il est intéressant de noter que les gènes *VH* exprimés précocement ne sont pas ceux qui se trouvent en position proximale par rapport au locus *JH*. Il semble donc que plusieurs mécanismes de régulation régissent leur expression, avec notamment des éléments régulateurs en *cis* situés au niveau des régions non codantes des gènes *VH*.

A l'âge néonatal, on n'observe plus la surabondance des gènes *DHQ52* (figure 5) et *JH1*, et on note l'apparition d'autres membres du répertoire VH. On note aussi une augmentation de la taille des régions CDR3 avec ajout de séquences N, dont le nombre peut atteindre sept nucléotides. Puisque, par le jeu des mutations somatiques et d'ajout de séquences N, chaque codon supplémentaire peut augmenter la diversité d'un facteur 20, cet enrichissement nucléotidique s'accompagne d'une augmentation du répertoire de plusieurs ordres de grandeur.

### Expression temporellement programmée du répertoire variable

De longue date, on sait que les spécificités Ac sont acquises de façon ordonnée au cours du développement. Ainsi, la réponse Ac contre des antigènes polysaccharidiques apparaît plus tardivement que celle dirigée contre les antigènes protéiques, ce qui, d'ailleurs, pose des problèmes de vaccination contre des agents pathogènes dotés d'une capsule polysaccharidique, comme *Haemophilus influenzae*. Un certain nombre d'Ac d'apparition précoce sont polyréactifs; cette forte polyréactivité obser-

### Sélection de l'expression des gènes V

Même à l'âge adulte, de nombreuses études indiquent que l'utilisation des différents éléments génétiques *VH*, *DH*, *JH* et *VK* ne se fait pas au hasard. Après l'expression fréquente de certains gènes au cours des étapes précoces de l'ontogenèse, il y a normalisation du répertoire qui reste cependant remarquable par l'absence de corrélation entre la complexité génomique des loci et l'expression des différents gènes. Ainsi, pour les gènes *VH*, le gène 1-69 est utilisé avec une fréquence élevée par les cellules B fœtales, par celles du sang périphérique de l'adulte sain, mais aussi par

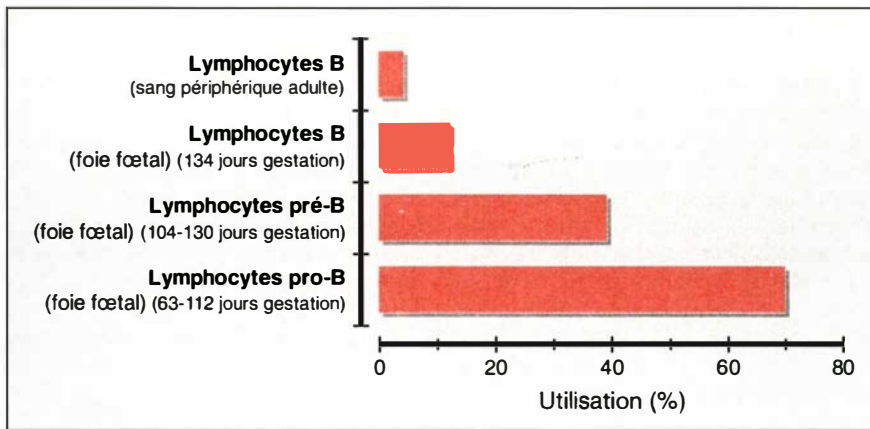


Figure 5. **Évolution de l'expression du segment génique DHQ52 au cours de l'ontogenèse.** La fréquence de l'utilisation du gène DHQ52 dans les réarrangements du locus VH humain, (exprimée en pourcentage), a été étudiée au cours de l'évolution du lymphocyte B, du stade pro-B, dans le foie fœtal, au stade B mûr, dans le sang périphérique [20]. On note qu'aux stades immatures le gène DHQ52 est exprimé avec une grande fréquence, alors que le lymphocyte mûr exprime principalement d'autres segments géniques DH.

les cellules B leucémiques, par les autoAc et par les Ac dirigés contre les agents infectieux [6, 21]. Initialement, on pensait que la proximité des différents éléments génétiques VH par rapport au locus JH pouvait à elle seule déterminer la probabilité d'expression des gènes VH. Cependant, la localisation du gène 1-69 à plus de 900 kb des gènes JH [5] indique que la fréquence de l'assemblage d'un élément génétique doit être déterminée par les facteurs qui régissent l'assemblage des gènes V, tels que les distances entre les éléments qui se recombinent et les gènes VH, les différences au niveau des séquences de recombinaison, l'accessibilité et la topologie de l'ADN génomique, les gènes qui codent pour les recombinases et ceux qui règlent leur expression, et des mécanismes post-transcriptionnels qui détermineraient les sélections cellulaires. Il est aussi possible que, tôt au cours de l'ontogenèse, les Ag du soi exercent une sélection positive ou négative sur l'expression de certains gènes VH. L'étude de ces différentes questions pourrait avoir un impact important sur la compréhension des mécanismes de sélection du répertoire VH fonctionnel.

### **La tolérance des cellules B au soi et la genèse des autoanticorps pathogènes**

Chez l'individu sain, la combinaison de certains gènes V du répertoire B peut conduire à la formation d'Ac capables de reconnaître les antigènes

du soi. Très tôt au cours du développement, le répertoire préimmunitaire se caractérise par l'expression préférentielle de certains gènes V en configuration germinale et par une autoréactivité. La sélection négative, quant à elle, élimine les cellules productrices d'autoAc de forte affinité, tandis que les autoAc de faible affinité et polyréactifs feront partie des Ig circulantes. Ces autoAc naturels pourraient servir de première ligne de défense contre des agents infectieux. Comment et pourquoi des autoAc sont-ils produits en grande quantité, et deviennent-ils agressifs et pathogènes dans les maladies auto-immunes ? Des expériences récentes de transgénèse chez la souris ont permis de mettre en évidence différents mécanismes de tolérance au soi des cellules B : l'anergie, la délétion clonale, l'exclusion par compétition et l'édition des récepteurs des cellules B. D'abord, l'anergie clonale (ou inactivation fonctionnelle) a été mise en évidence chez des souris doublement transgéniques, qui expriment du lysozyme « soluble » et des récepteurs B antilysozyme. Dans ce contexte de reconnaissance, les lymphocytes B circulent normalement, mais ne produisent pas d'Ac spécifiques du lysozyme [22]. Ils sont dits anergiques. Quant à la délétion clonale, son rôle dans la tolérance immunitaire B fut longtemps soupçonné. Avec l'utilisation de souris transgéniques H-2<sup>l/k</sup> exprimant des récepteurs B anti-H-2κ, il a été montré que les souris n'ont pas de cellules B porteuses de l'Ac anti-H-2κ, ni d'Ac sériques de cette spécificité [23]. Cette situation

se retrouve chez des souris doublement transgéniques pour le lysozyme « membranaire » et les Ac correspondants, où l'élimination des cellules B résulte d'un arrêt de leur développement au stade B immature, suivi de la mort par apoptose dans la moelle osseuse [24]. Troisièmement, l'absence de clones autoréactifs B circulants peut aussi être due au phénomène d'exclusion par compétition, où la majorité des cellules B autoréactives restent bloquées, par compétition, à l'extérieur des follicules lymphoïdes qui laissent pénétrer les cellules B normales. Tout se passe comme s'il y avait une compétition pour l'entrée dans le follicule lymphoïde, entre les cellules normales et les cellules autoréactives qui, en présence de l'autoantigène, ne tarderont pas à mourir [25]. Enfin, d'autres études ont montré que des réarrangements successifs des chaînes L endogènes des Ig permettaient aux cellules d'échapper à l'autorecognition en changeant leur spécificité [26]. Ce dernier mécanisme, appelé « édition des récepteurs », permet une diversification du répertoire B. Dans ce modèle, la cellule n'est pas limitée à un choix de vie ou de mort ; elle peut, par l'acquisition de nouveaux récepteurs, se doter de capacités de reconnaissance différentes.

Dans le lupus érythémateux disséminé (LED) humain, maladie auto-immune systémique d'origine inconnue, des autoAc de forte affinité sont produits en grande quantité. L'étude des régions VH des Ac anti-ADN indique de fréquentes mutations avec



remplacements dans les régions hypervariables par rapport aux gènes germinaux du même patient [27], ce qui suggère l'existence de pressions de sélection. Pour les gènes  $V\kappa$ , de nombreux gènes utilisés sont issus de la région 3' du *locus*, et associés à des segments  $J\kappa$  localisés en amont [28]. Ces profils de réarrangements observés chez les malades permettent de formuler l'hypothèse selon laquelle le déficit réside dans les mécanismes moléculaires correcteurs qui, normalement, permettent aux lymphocytes B d'échapper à l'autoréactivité par l'édition de nouveaux récepteurs.

### Le répertoire B dans l'infection par le VIH

La variabilité du génome du virus de l'immuno-déficience humaine (VIH), son polymorphisme biologique et son affinité pour les cellules du système immunitaire soulèvent nombre de questions quant aux stratégies adoptées par le système immunitaire pour répondre à l'agression virale. A la diversité des manifestations cliniques associées à l'infection par le VIH correspond un large spectre de signes biologiques. Au niveau de la réponse immunitaire humorale, alors que les individus contaminés produisent généralement une forte réponse Ac, une proportion seulement des sujets développe des Ac neutralisants à titre élevé. Détectables surtout aux phases initiales de l'infection, ces derniers sont capables de conférer une protection, comme le montrent les résultats de protection du chimpanzé par injection expérimentale d'Ac. En revanche, aux étapes tardives de l'infection humaine, le titre des Ac protecteurs diminue, ce qui, en conjonction avec d'autres facteurs, contribue à l'aggravation de la maladie.

Des études récentes ont révélé un nouveau mécanisme par lequel le virus induirait des anomalies importantes au niveau du recrutement des cellules B de l'individu infecté. En effet, on pensait initialement que l'interaction des cellules B avec l'Ag se produisait exclusivement par un mécanisme d'activation clonale, où seulement moins de 0,01 % des cellules B sont recrutées. Un autre mode d'interaction vient cependant d'être décrit par lequel des molécules, appelées superantigènes B, reconnaissent

des cellules B humaines exprimant des régions V d'une famille donnée, comme par exemple la protéine A du staphylocoque ou la glycoprotéine gp120 du VIH-1 [11, 29, 30]. Ce type d'interaction peut conduire au recrutement d'une large proportion du répertoire B, comme les cellules de la famille  $VH3$  qui domine normalement le répertoire humain (> 50 %), à la fois au niveau génomique et somatique [31]. Ce sont justement les cellules B correspondant aux membres de cette famille qui peuvent être recrutées par gp120 [30] avec comme conséquences possibles une activation puis une délétion clonale. A la lumière de l'effet de l'injection expérimentale de superantigènes T à la souris [32, 33], on peut proposer l'hypothèse suivante: au cours des étapes précoces de l'infection, la «superantigénicité» de gp120 induirait une activation et une expansion du répertoire B exprimant des gènes de la famille  $VH3$ . Dans une seconde phase, il y aurait une déplétion en cellules B  $VH3^+$ . L'individu contaminé disposerait alors de moins de la moitié de son répertoire B pour produire des Ac anti-gp120 protecteurs, ce qui expliquerait la difficulté de production d'une réponse neutralisante efficace.

Ces acquis récents ont ajouté une nouvelle dimension à notre compréhension de la sélection du répertoire B humain par le VIH. En outre, ils pourraient offrir de nouvelles perspectives d'immuno-intervention pour le SIDA. En effet, l'étude des bases moléculaires des propriétés superantigéniques B de gp120 et l'identification du site responsable de cette activité permettraient de synthétiser une molécule gp120 dépourvue d'activité superantigénique. Il est permis de penser qu'une telle molécule n'induirait pas la délétion clonale d'une partie importante du répertoire B et pourrait donc s'avérer plus efficace dans la protection ■

#### Remerciements

Ce travail a été rendu possible grâce au soutien de l'Institut Pasteur, de l'ARC, et de SIDACTION.

#### TIRÉS À PART

M. Zouali.

#### RÉFÉRENCES

26. Prak EL, Trounstine M, Huszar D, Weigert M. Light chain editing in kappa-deficient animals: a potential mechanism of B cell tolerance. *J Exp Med* 1994; 180: 1805-15.
27. Demaison C, Chastagner P, Thèze J, Zouali M. Somatic diversification in the heavy chain variable region genes expressed by human autoantibodies bearing a lupus-associated nephritogenic anti-DNA idiotype. *Proc Natl Acad Sci USA* 1994; 91: 514-8.
28. Bensimon C, Chastagner P, Zouali M. Human lupus anti-DNA autoantibodies undergo essentially primary V kappa gene rearrangements. *EMBO J* 1994; 13: 2951-62.
29. Hillson JL, Karr NS, Oppliger IR, Mannik M, Sasso EH. The structural basis of germline-encoded VH3 immunoglobulin binding to staphylococcal protein A. *J Exp Med* 1993; 178: 331-6.
30. Berberian L, Goodglick L, Kipps TJ, Braun J. Immunoglobulin VH3 gene products: natural ligands for HIV gp120. *Science* 1993; 261: 1588-91.
31. Zouali M. Unravelling antibody genes. *Nature Genet* 1994; 7: 118-20.
32. White J, Herman A, Pullen AM, Kubo R, Kappler JW, Marrack P. The Vβ-specific superantigen staphylococcal enterotoxin B stimulation of mature T cells and clonal deletion in neonatal mice. *Cell* 1989; 56: 27-35.
33. Lafon M, Scott-Algara D, Marche PN, Cazenave PA, Jouvin-Marche E. Neonatal deletion and selective expansion of mouse T cells by exposure to rabies virus nucleocapsid superantigen. *J Exp Med* 1994; 180: 1207-15.

## Summary

### The human B-cell repertoire: expression in normal and pathological conditions

The timing and mechanisms that govern B-cell development are unique in several respects. Immunoglobulin (Ig) genes are exquisitely assembled in B cells and this process is site-specific. In order to express Ig on their surface, B-cell progenitors must rearrange variable (*V*) gene segments on the Ig heavy (*H*), and  $\kappa$  and  $\lambda$  light (*L*) chain loci in a process known as VDJ recombination. During human ontogeny, B-cell progenitors can choose from the libraries of variable gene segments. This primary repertoire is further expanded by junctional flexibility, secondary rearrangements, *i.e.* of VL to VL and somatic hypermutation. In human, the organization and the content of the Ig loci have been the focus of considerable attention. The recent elucidation of the structure of the *VH* and *V $\kappa$*  loci revealed that human Ig *V* genes exhibit unique features. For both loci, the *V* genes have been categorized into families of sequence-related members which are intermingled among each other and include approximately 45% of pseudogenes. Even more intriguing are the observations that some *VH* and *V $\kappa$*  genes are dispersed not only outside the functional loci, but also on other chromosomes. The molecular events that gave rise to these « orphon » genes underlie the extensive degree of plasticity of human Ig genes. While in normal conditions, they could play a role in amplifying the repertoire, in pathological conditions they may in turn

become the targets for chromosomal translocations in B-cell neoplasia. Other remarkable features of the human *VH* locus include its dominance by members of the *VH3* gene-family and its high degree of polymorphism, with insertion deletion/polymorphisms and allelisms, which may lead to a high number of haplotypes with different coding potentials. The possibility that certain polymorphisms are associated with disease susceptibility has not been fully addressed. The elucidation of the content of the human Ig *V* loci is providing insight into the mechanisms of *V* gene expression in pathological conditions. A number of studies focused on *V* gene usage in relation to systemic autoimmunity. In principle, random combinatorial assembly of Ig *V* gene elements may lead to generation of a proportion of antibodies with high-affinity to self. In normal conditions, the corresponding B cells will be clonally deleted, rendered silent, or switched to a different specificity, a mechanism that has been termed receptor editing. In systemic autoimmunity, however, anti-self high-affinity antibodies are overexpressed. Molecular cloning, nucleotide sequencing and comparative genomic PCR revealed that human pathogenic autoantibody *V* genes have incurred extensive somatic selection events. There is also evidence that the corresponding B cells underwent essentially primary L-chain rearrangements. This observation raises the possibility that,

in systemic autoimmunity, a subset of B cells may be unable to revise their receptors and to extinguish their high-affinity for self. This blockade could be genetically determined or somatically acquired. Finally, infection with the human immunodeficiency virus (HIV) results in several B-cell abnormalities. The first evidence that HIV antigens may alter repertoire expression came from the demonstration that the major envelope protein of HIV-1, gp120, binds to *V $H$ 3*<sup>+</sup> B cells and to serum *V $H$ 3*<sup>+</sup> Ig from normal individuals, and this interaction is independent of the L-chain isotype. The gp120 interaction with *V $H$ 3* gene products also seems to be functional because gp120 selectively induces Ig secretion by *V $H$ 3* B cells. These characteristics are reminiscent of the properties of B-cell superantigens which are able to trigger a large proportion of the B-cell repertoire. More recently, studies of peripheral blood lymphocytes from HIV seropositive subjects and AIDS patients show marked changes in *V $H$ 3*<sup>+</sup> B cells during different clinical stages of HIV infection. It is possible that the sequential expansion and reduction of *V $H$ 3* B cells are related to the superantigenic properties of gp120. Understanding the molecular basis of this activity and the cellular mechanisms responsible for the clonal fate of B cells during HIV infection are relevant to designing novel strategies for immunointervention.