

MUNIBE (Antropología y Arqueología)	Suplemento N.º6	149-167	SAN SEBASTIAN	1988	ISSN 0027 - 3414
-------------------------------------	-----------------	---------	---------------	------	------------------

Nouvelles perspectives dans l'étude des Polymorphismes génétiques sanguins.

Philippe LEFEVRE-WITIER *

RESUME

Depuis 10 à 15 ans, l'étude de la variabilité génétique humaine connaît des changements profonds. En effet, les systèmes génétiques polymorphiques mis en évidence par les travaux en immunologie (groupes sanguins) ou en biologie des protéines sont maintenant identifiés plus directement au niveau de l'ADN chromosomique. Les structures du gène et du génome humains se révèlent de plus en plus complexes et offrent à la biologie moléculaire, à l'immuno-génétique, à l'anthropologie des populations un champ de recherche d'une extrême richesse. L'auteur tente de résumer ces apports récents à la discipline anthropologique selon cinq axes principaux: les nouveaux marqueurs génomiques, les progrès des marqueurs génétiques «traditionnels», les liaisons observées entre l'immunologie, la génétique et la pathologie, la cartographie génétique et l'analyse statistique de la variabilité populationnelle. Chaque approche est illustrée d'exemples tirés de travaux de recherches sur les populations basco-pyrénéennes. En conclusion sont rappelées les différentes théories interprétatives du polymorphisme observé dans le groupe humain, ainsi que la remise en question de nombreuses théories et affirmations anthropologiques qu'impliquent les nouveaux acquis moléculaires et le développement des méthodes et outils d'analyse statistique.

Dans l'étude des polymorphismes sanguins, cinq approches scientifiques sont actuellement les plus riches d'avenir. Ce sont:

A—La lecture du génome, le déchiffrement de l'ADN, support des messages chromosomiques.

B—Les progrès continus dans la mise en évidence des produits du génome, autrement dit les marqueurs «traditionnels», comme les groupes sanguins et tissulaires et les variants enzymatiques et protéiques.

C—Le développement des connaissances sur les associations entre marqueurs génétiques et phénotypes pathologiques ou non.

D—La situation chromosomique des systèmes polymorphiques et leur participation à la cartographie générale du génome humain, cartographie qui procède du progrès dans les domaines précédentes, mais aussi.

E—De l'amélioration des outils d'analyse statistique de la variabilité génétique, dans les familles ou les populations humaines.

Ces approches sont ici séparées pour la commodité de l'exposé, mais bien évidemment, elles sont profondément liées, interactives, et chaque avancée des recherches sur l'une de ces voies retentit sur les autres, comme le prouve chacune des réunions scientifiques qui jalonnent ces progrès rapides et constants depuis 30 ans; tels sont récemment le

congrès international de génétique de population de BERLIN en 1986, le colloque «carte génétique humaine - HGM n.º 9» en septembre 1987 à PARIS, l'atelier international HLA n.º 10 en décembre 1987 à NEW YORK et bientôt les réunions d'anthropologie génétique du 12ème Congrès International des Sciences anthropologiques et ethnologiques à ZAGREB (juillet 1988 - Symposium sur «la génétique des isolats» et sur «génétique et structures des populations»).

A - EXPLORATION GENOMIQUE

Les progrès dans l'exploration directe du génome, des messages génétiques portés par l'ADN, sont liés depuis 1970 à la découverte d'un jeune suisse, WERNER ARBER; il apprend en effet aux biologistes à utiliser les véritables «bistouris à gènes» que sont les enzymes de restriction ou endonucléases qui coupent la double hélice en des sites très précis. Trois manipulations, qui donnent bien la mesure de l'astuce de l'ingénierie génétique, prendront le relais de cette chirurgie: la recombinaison, le clonage, le «blotting» ou buvardage.

Je ne ferai que rappeler rapidement ces techniques qui sont maintenant bien connues.

Le clonage du gène est son amplification, c'est à dire sa multiplication et sa purification par un double mécanisme:

La recombinaison permet le transport dans une bactérie des séquences nucléotidiques par un vecteur (plasmide ou phage) auxquels elles ont été fixées. (Figure 1).

* Centre de Recherches sur le Polymorphisme Génétique des Populations Humaines.
Centre National de la Recherche Scientifique
Hospital Purpan
31300 TOULOUSE. FRANCE.

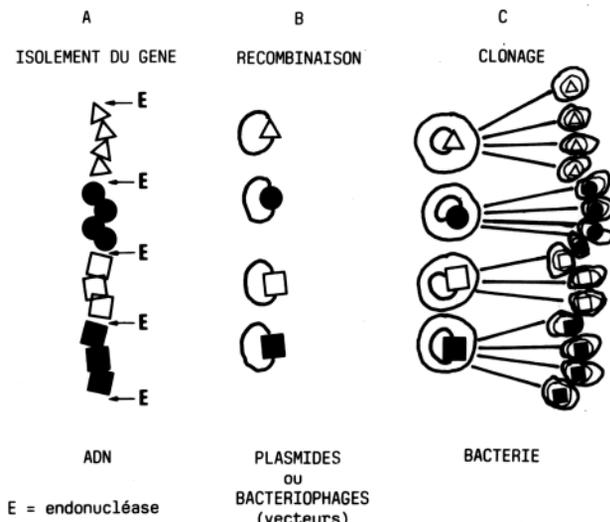


Figure 1: Recombinaison et clonage d'un gène permettant soit isolement et son amplification.

d'une part les vecteurs se fixent électivement le plus souvent à une séquence nucléotidique; d'autre part, la bactérie devient colonie et produit ainsi une grande quantité de la même séquence.

Le blotting de SOUTHERN permet de fixer les clones bactériens et les séquences d'ADN qu'ils renferment sur une membrane de nitrocellulose et d'identifier ces séquences par autoradiographie avec une sonde spécifique ADN complémentaire radioactive.

Il faut souligner la niveau de lecture du génome, que permettent ces techniques:

a) soit on coupe et on clône de l'ADN nucléaire «naturel» dit génomique, dont on veut identifier les structures fragmentaires. On peut alors avoir deux attitudes:

- établir les séquence nucléotidique de ces fragments, ce qui est devenu facile grâce aux travaux, en particulier de W. GILBERT et A. MAXAM à Harvard depuis 10 ans (1977); il s'agit là d'une lecture directe de la totalité des messages le long du chromosome.

- considérer aussi le polymorphisme de ces fragments produit par une ou plusieurs endonucléases chez la même individu ou des individus différents; en effet, outre les mutations entraînant des produits différents du gène ou des éléments du gène, lui-même inclu dans ces fragments, existent des mutations juxtaposées des sites des enzymes de restriction, créant un polymorphisme parallèle; ce polymorphisme des longueurs de fragments obtenus par les enzymes de restriction (Restriction Fragments Length Polymorphism) est relativement facile à mettre en évidence et souvent révélateur d'une mutation génique inconnue ou soupçonnée: UN POLYMORPHISME EN REVELE UN AUTRE. (Figure 2).

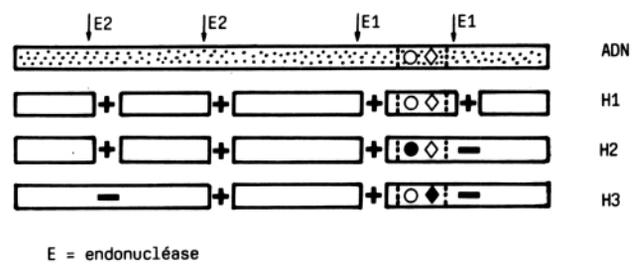


Figure 2: Mise en évidence par deux enzymes de restriction (endonucléases E1 et E2) d'un polymorphisme de fragments incluant d'autres mutations géniques (haplotypes H1, H2 et H3).

b) soit on sélectionne des ARN messagers dans une cellule spécialisée (par exemple réticulocytes du sang qui produisent beaucoup de chaînes d'hémoglobine alpha et beta) et on construit un ADN complémentaire (c ADN) par une transcriptase reverse. Cet ADN complémentaire, gène «artificiel», comme l'exprime J.C. KAPLAN (1987), sera cloné puis séquencé, ou bien fournira une sonde marquée par un produit radioactif pour une identification spécifique des messages de l'ADN étudié.

Dans l'étude des RFLP, comme dans celles des c ADN, il s'agit bien de lecture indirecte d'ADN.

Ces différents niveaux de lecture, ces «artifices» techniques sont à la base d'acquisitions majeures conceptuelles et pratiques dans le domaine des polymorphismes génétiques, et en particulier sanguins puisque les meilleures démonstrations nous sont venues, une fois encore, des systèmes des globines humaines (hémoglobine: chaînes alpha et bêta).

Nous le résumerons rapidement en fournissant quelques exemples et illustrations dans les deux domaines qui nous préoccupent ici:

- la structure fondamentale du gène et son évolution,

- l'hémotypologie ou l'anthropologie moléculaire populationnelle.

Cette séparation n'étant elle-même à nouveau qu'un artifice de présentation tant les apports dans un de ces domaines retentissent dans l'autre.

A-1) Au plan fondamental

L'anatomie et la physiologie du gène sont depuis quelques années en constant remaniement. La lecture indirecte par les c ADN a entraîné un véritable bouleversement des notions génétiques classiques. On a découvert que les c ADN ne traduisaient que certaines parties des gènes, les parties réellement codantes (Exons) pour la fabrication des protéines polymorphiques ou non; de nombreuses autres séquences non codantes se révèlent à l'examen de l'ADN, mais ces séquences non codantes ne sont pas retenues par les ARN cytoplasmiques des eucar-

yotes, donc n'apparaissent pas dans les c ADN: un «épissage» (Splicing) a nettoyé l'ARN nucléaire. Ainsi le gène n'apparaît plus simple et homogène, mais modelé, complexe, «éclaté». C'est l'anglais D. FLAVELL (1980) qui fit cette constatation pour le gène des alpha et beta globines de l'hémoglobine, puis l'ovalbumine de poule. Parmi les séquences non codantes (transcrites, non excisées, non traduites), des introns (transcrits, excisés, non traduits) dont la fonction reste inconnue et des pseudogènes à l'état de reliques évolutives sans fonction apparente, mais réservoirs potentiels d'informations. Cet éclatement spatial déborde même la séquence informative chromosomique puisque, si certains éléments isogéniques sont en tandem sur le même chromosome, d'autres sont dispersés sur ce même chromosome ou sur des chromosomes différents. (Figure n.° 3)

Ces découvertes, dues au génic génétique, nous contraignent à une nouvelle compréhension des systèmes polymorphes, essentiellement de leur complexité, très partiellement de leurs rôles.

A-1-1) Compréhension dans un premier temps de la complexité des «familles géniques» dont nous donnerons deux exemples:

a) famille de la bêta-globine de l'hémoglobine dont la structure comme celle de l'alpha-globine est maintenant assez bien élucidée. Sur le parcours du chromosome 11, sur un segment d'ADN de 50 kb, de son extrémité 5' vers son extrémité 3', nous trouvons un pseudogène bêta 2, un gène epsilon embryonnaire, deux gènes foetaux G-gamma et A-gamma, un pseudogène bêta 1 et enfin les deux gè-

nes des chaînes bêta de l'hémoglobine de l'adulte normal: delta pour Hb A2 et beta pour Hb A. (Figure 4).

En outre, la séquence complète de ces gènes et des parties juxtagéniques a été déterminée depuis 1980 par R.A. SPRITZ et B.G. FORGET, ainsi que R.M. LAWN, T. MANIATIS, F.E. BARALLE et leurs collaborateurs.

A partir de cette séquence, toutes les anomalies nucléotidiques des exons donnant naissance à des variants structuraux, pathologiques ou non, peuvent être définies ainsi que bien entendu des mutations intervenant dans les séquences non codantes, les introns, les pseudogènes.

On n'est plus devant une «série allélique» de protéines produites par des gènes «hypothétiques» en un locus, mais devant un objet dont on peut saisir toutes les facettes, les bizarreries anatomiques, les zones inactivées, les cicatrices évolutives, les commandes encore incomprises, etc.. C'est une véritable archéologie du gène qui se fait jour, avec un décryptage progressif des liens existant entre structure, temps et fonction. Certains de ces mécanismes d'apparition et d'évolution de ces isogènes ou allèles sont progressivement élucidés. Ainsi la redondance des messages le long du chromosome semble un des facteurs importants de la variabilité; toute répétition, recopiage, duplication, s'accompagnant visiblement d'erreurs correspondant aux mutations ponctuelles observées. Un «bégaiement», parfois un «bafouillage», provoque ces messages presque identiques et normaux, comme ceux des gènes G-gamma et A-gamma ou comme ceux des gènes delta et bêta des chaînes normales de la bêtaglobine, mais aussi les messages déviés des pseudogènes ou des gènes pathologiques. De la même façon

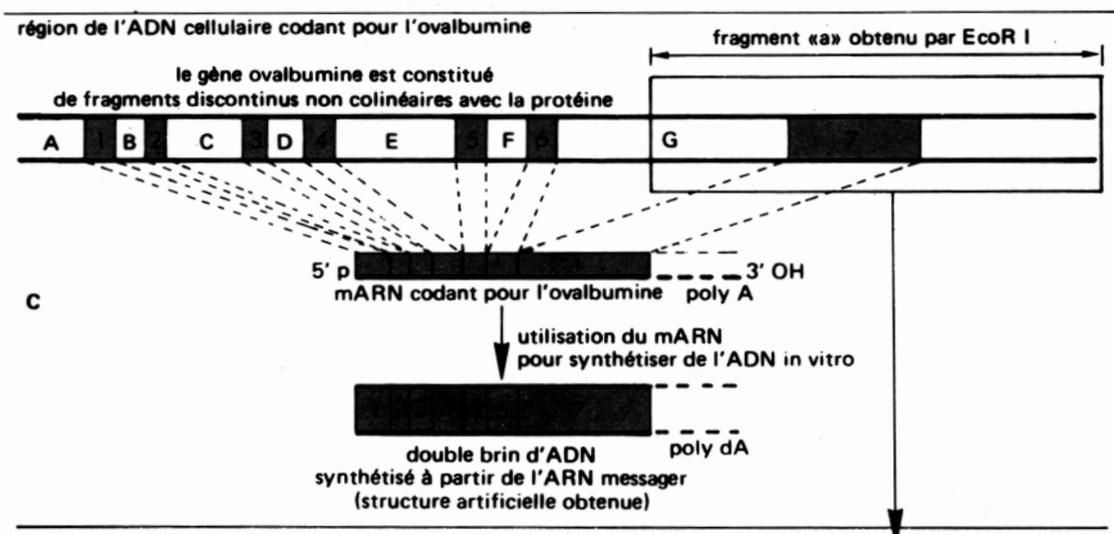


Figure 3: Production de l'ovalbumine de poule: schéma de l'épissage (splicing) survenant entre les messages de l'ADN chromosomique et l'ARN messager. Exemple de synthèse d'un ADN complémentaire «artificiel». In ENCYCLOPEDIA UNIVERSALIS, 1979, p. 147, fig. 11.

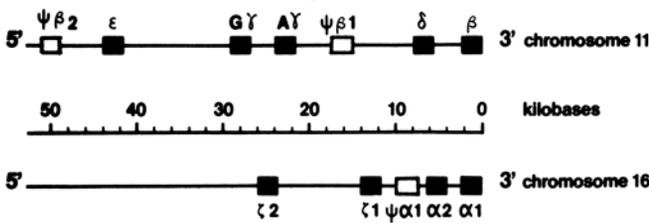


Figure 4: Comparaison entre les structures géniques des deux familles alpha (chromosome 16) et bêta (chromosome 11) de la globine humaine.

qu'on tentait de reconstituer l'apparition de ces mutations dans les espèces à partir de la comparaison des enchainements d'acides aminés des polypeptides, on peut maintenant comparer directement les séquences d'ADN. On utilise toujours la notion de temps minimum de fixation d'une divergence de message ou UTE (Unité de Temps d'Evolution) et de Distance Mutationnelle Minima (soit le nombre de mutations minima pour passer d'une séquence à une autre); mais on peut maintenant utiliser toutes les séquences chromosomiques au fur et à mesure que les détermine le génie génétique. Ainsi les deux gènes G-gamma et A-gamma de l'hémoglobine foetale permettent la reconstitution évolutive suivante, (SHEN, 1981): l'absence de cette duplication chez les singes du Nouveau Monde et sa présence régulière, chez les hommes, les grands singes anthropomorphes et les singes «catarhiniens» d'Afrique et d'Asie, conduit à situer son apparition dans une fourchette de temps postérieure à -40 millions d'années et antérieure à -25 millions d'années. (figure 5).

Comme G-gamma et A-gamma de la bêta-globine, les deux gènes alpha 1 et alpha 2 de la famille alpha-globine du chromosome 16, présentent une homologie presque parfaite. Ces phénomènes de duplication presque parfaite, situés en tandem sur les chromosomes 11 et 16 intriguent les généticiens qui ont fait de subtiles hypothèses de crossing-over inégaux et répétés permettant leur maintien dans une «évolution concertée» assortie de systèmes rectifiant les divergences dues à la dérive génique ou systèmes de «conversion génique». Ces crossing-over et mécanismes associés se produisent à certains points chauds (séquence de 5.000 kb) riches en nucléotides TG et GC, comme nous le verrons pour la formation des haplotypes dans certains systèmes génétiques complexes, à productions antigénique ou protéiques multiples.

En effet, malgré leur richesse apparente en messages, les familles alpha et bêta des globines ne traduisent, en temps donné, qu'un seul produit protéique et leurs gènes ne s'expriment que selon leur

ordre de 5' en 3', depuis la vie embryonnaire jusqu'à l'âge adulte.

b) Les familles géniques du complexe majeur d'histocompatibilité (CMH) offrent une bien plus grande complexité. Rappelons que la structure sur le chromosome 6 du CMH est la suivante: du centromère du chromosome 6 et sur le bras court se trouvent successivement, après le locus de glyoxylase 1, les familles de gènes de classe II réparties sur plusieurs loci, avec probablement un point chaud entre DP et D (en effet la distance de recombinaison de 1 à 2 centimorgan ne traduit que ce point chaud et il est probable que la distance physique ou nucléotidique soit beaucoup plus courte); puis les familles de gènes du complément (classe III) dont les quatre positions actuellement sont C2, Bf, C4A, C4B à l'intérieur d'un fragment de 120 kb (la plus grande distance est entre Bf et C4A, soit 30 kb). Enfin les familles A, C et B de classes I sont mieux connues car clonées récemment par B.R. KOLLER et collaborateurs (1986) ainsi que ZIEGLER et collaborateurs (1986); outre les 3 loci A, C et B, 14 autres positions géniques ont été découvertes dont certaines correspondent à des pseudogènes. (figure 6).

Comme il est maintenant bien connu, chaque locus comporte un nombre souvent très élevé d'allèles donnant naissance à des antigènes spécifiques. En outre ces familles présentent des associations préférentielles ou «déséquilibres de linkage» conduisant à des combinaisons haplotypiques très nombreuses et caractéristiques de certains individus ou de populations; nous en verrons quelques exemples infra.

A-1-2) Cet ensemble de familles de gène du CMH nous démontre un subtil mécanisme évolutif; en effet, la réponse immune est pilotée génétiquement par deux contrôles différents: les réactions à médiations cellulaires dépendent des familles antigéniques HLA A B C pour la marquage de l'identité cellulaire, de certaines familles D pour les opérations de différenciation et de coopération cellulaire concernant les actions des lymphocytes et des macrophages et en-

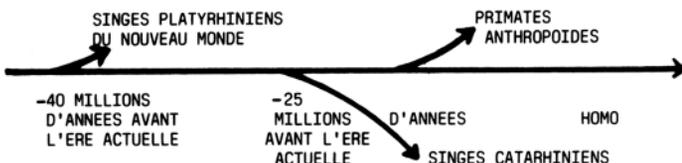


Figure 5: Evolution des messages génétiques de la chaîne bêta de la globine humaine: apparition de la duplication de gènes gamma (G-gamma et A-gamma) responsables de la production des chaînes protéiques bêta-foetales.

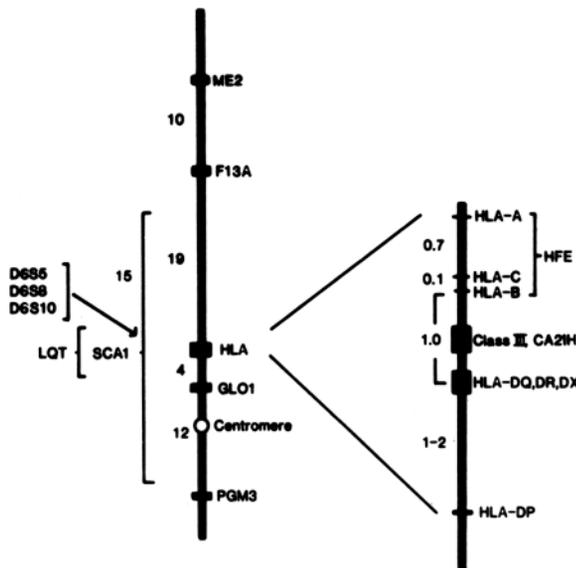


Figure 6: Structure schématique du Centre Majeur d'Histo-compatibilité (C.M.H.) sur le chromosome 6.

fin des familles de classe III (C2, C4, Bf) pour les réactions impliquant la présence du complément. La réaction humorale (la production d'anticorps) dépend d'autres systèmes génétiques dont peut-être les gènes DR correspondant à la série I de la souris. Les agents infectieux étant extrêmement divers, la réponse *polygénique* et à *double commande* offre un grand nombre de combinaisons phénotypiques dont la distribution «normale» garantit une réponse moyenne et/ou exceptionnelle: les combinaisons les plus défavorables disparaissent régulièrement à une des extrémités de la courbe. On serait devant une formule de polymorphisme équilibré et donc hétéro-entretenu.

A-2) Sur le plan de l'anthropologie génétique

L'examen systématique de fragments de restriction d'ADN par des sondes appropriées opère comme un formidable révélateur de polymorphismes au sein des populations humaines et donc complète et renouvelle le stock de marqueurs dont nous disposons avec un pouvoir discriminant et des interprétations phylogénétiques de grande valeur.

Rappelons que pour explorer ce nouveau polymorphisme, il faut impérativement une enzyme de restriction pour établir des fragments et une sonde pour situer les fragments par rapport à un locus génique particulier. Chaque sonde mise au point fait apparaître de nouvelles mutations, car une base sur 100 serait polymorphe (A.J. JEFFREYS, 1982). Il faut donc s'attendre à une avalanche de données dont le front nous parvient déjà. Une revue dans ce domaine serait une lourde tâche.

A-2-1) L'exemple des polymorphismes de restriction liés à la mutation S (NB.1) de l'hémoglobine humaine en Afrique et aux Etats Unis, maintenant classique, est repris par bien des auteurs. Nous en résumerons rapidement les éléments essentiels.

Dans un premier temps, l'étude de H.W. KAN et A.M. DOZY (1978) des fragments de restriction de la bêta-globine coupés par l'endonucléase HpaI a montré que la mutation S pouvait se trouver dans deux types de fragments: 7,6 kilobases (kb) ou 13 kilobases par perte d'un site HpaI du côté 3' du gène bêta.

La formation de fragments de restriction par 8 endonucléases a montré que 11 sites pouvaient être le lieu de mutations entraînant une anomalie de l'action enzymatique de ces endonucléases. Les combinaisons de ces mutations forment de haplotypes spécifiques de certaines zones géographiques et de leur population. Ainsi, au Nigéria et en Algérie, la mutation S drépanocytaire est à 100 % dans un haplotype caractéristique dit «BENIN» incluant en particulier le fragment 13 kb découvert par KAN et DOZY. Au Sénégal, la mutation S est à 82% dans un haplotype caractéristique à fragment 7,6 kb. De même en République Centre Africaine et au Zaïre où la liaison avec un autre haplotype caractéristique dit «BANTOU» atteint 86%.

Les auteurs responsables de ces travaux, G. MEARS, R.L. NAGER, J. PAGNIER et D. LABIE (1984) postulent l'apparition de la mutation bêta-S postérieurement à la formation de ces matrices géniques haplotypiques et affirment donc l'origine multicentrique de la mutation bêta S. Les travaux les plus récents prouvent en outre l'existence d'une matrice génique originale portant le trait drépanocytaire indien et moyen oriental (KULOSIK, 1986) (RAO, 1988). (Figure 7). Ces affirmations ruinent toutes les théories anthropologiques antérieures de diffusion de l'hémoglobine drépanocytaire HbS et ouvrent de nouveaux horizons, non seulement à l'étude des migrations intra-africaines, mais aussi aux descriptions des structures génétiques des populations caraïbes, méso et nord américaines, Les mêmes approches sont actuellement réalisées pour l'hémoglobine anormale asiatique E.

A-2-2) Les polymorphismes géniques des délétions des familles de gènes de l'alpha- et de la bêta-globine sont aussi intéressants pour les études de populations où ces délétions constituent de véritables «marqueurs» accompagnés d'anémies plus ou moins graves selon les génotypes et/ou phénotypes réalisés. Globalement ces anémies sont distinguées en bêta ou alpha-Thalassémies.

(NB. 1) Mutation ponctuelle du 6ème acide aminé de la chaîne bêta-globulinique entraînant l'apparition d'une grave anémie hémolytique appelée anémie falciforme au drépanocytose.

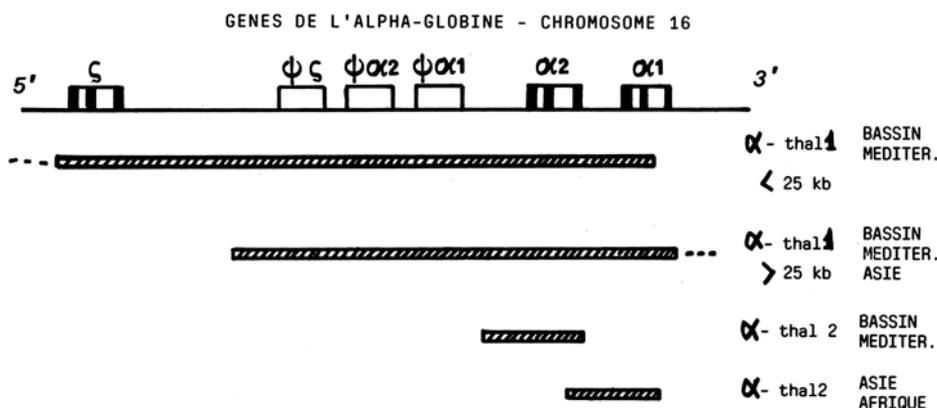


Figure 8: Représentation simplifiée des délétions géniques de l'alpha-globine humaine, responsables de différentes formes cliniques et géographiques de l'anémie Alpha-thalassémie. D'après SPRITZ et FORGET, 1983.

de chaînes embryonnaires sigma; un accident mortel (hydrops foetalis) entraîne le décès du fœtus à la naissance par production d'hémoglobine BART'S à 4 chaînes gamma (gamma-41, car aucune chaîne alpha normale n'apparaît (ORKIN, 1980 - PRESSLEY, 1980).

b) les délétions beaucoup plus courtes ne supprimant l'expression que d'un seul des gènes alpha 1 ou alpha 2 et maintenant donc une production minimale de chaînes globiniques alpha à la naissance. A l'état hétérozygote ces délétions retrouvées en Asie, en Méditerranée et en Afrique n'entraînent aucun trouble clinique. Cependant en Asie et dans le bassin méditerranéen, leurs combinaisons avec des délétions larges provoquent une anémie alpha-thalassémique dite «maladie hémoglobine bêta-4» par association préférentielle de chaînes globiniques bêta. En Afrique, les formes homozygotes n'existent que par association d'haplotypes à délétions courtes et donc les formes cliniques sévères, telles que la maladie hémoglobine bêta-4 ou l'hémoglobine BART'S (hydrops foetalis), sont pratiquement inconnues (EMBURY, 1980 - HIGGS, 1980).

A-2-3) Les travaux récents de J.S. WAINSCOAT (1986) ont étendu aux gènes normaux beta de la globine humaine dans plusieurs populations du monde les études de polymorphismes de restriction liés en haplotypes. Les huit populations ainsi examinées sont des anglais, des chypriotes, des italiens, des in-

diens de l'Inde, des thaïlandais, des mélanésiens de deux groupes A et B, des polynésiens et enfin deux groupes d'Afrique subsaharienne A et B, dont nous ne connaissons pas l'origine.

601 chromosomes étudiés ont révélé 14 haplotypes sur les 32 possibles. Certains des haplotypes sont tout à fait caractéristiques des groupes mélanésiens, polynésiens et africains noirs; d'autres sont communs à toutes les populations nonnégroïdes (et non pas non-africaines). Aucun des quatre chromosomes les plus communs ne peut être obtenu à partir des trois autres par un simple crossing-over ou une mutation ponctuelle. Ils semblent donc être des structures stables; ce sont eux qui devaient exister avant la séparation des rameaux humains que WAINSCOAT situe à 100.000 ans avant J.C. Les distances génétiques (de BALAKRISHNAN et SANGHVI) sont traduites par un dendrogramme assez réaliste des proximités et distances attendues et objectivant l'originalité du phylum africain pour ce locus de la bêta-globine (figure 9). Sur le plan évolutif, ces résultats vont dans le sens d'une séparation préalable entre les rameaux africains et eurasiens, telle que mise en évidence, depuis longtemps, par les recherches sur les types protéiniques (M. NEI, 1982), alors que les résultats des groupes sanguins et tissulaires ont toujours fait supposer une séparation des rameaux négroïde et caucasioïde postérieure à celle du rameau mongoloïde. Cet isolement ancien du rameau négroïde dans un milieu tropical et équatorial très particulier expliquerait peut-être sa grande divergence génétique confirmée dans de nombreux systèmes de

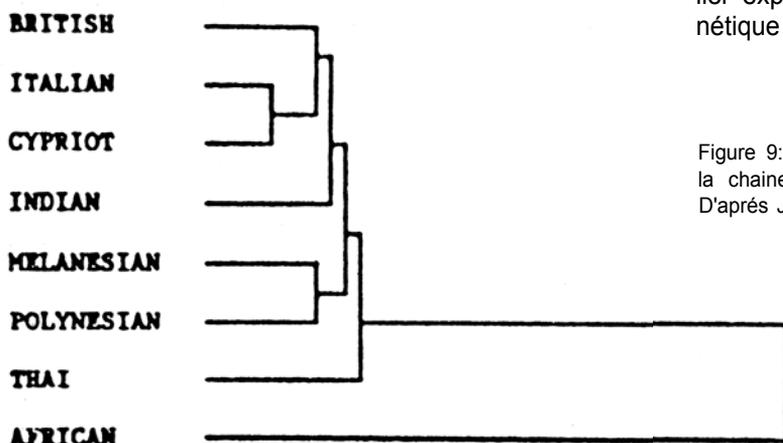


Figure 9: Classification hiérarchique selon la structure génique de la chaîne de la bêta-globine pour huit populations humaines. D'après J. S. WAINSCOAT, 1986.

marqueurs: cellulaires, protéiques, mais aussi génomiques, tels que les loci de l'hormone de croissance, de la dihydro-folate-réductase, de l'insuline et de l'albumine.

Si WAINSCOAT «suggère» que l'ancêtre commun de ces groupes explorés pour l'ADN de la beta-globine normale pourrait être africain et que la séparation des rameaux aurait pu prendre place sur ce continent, d'autres auteurs l'affirment beaucoup plus bruyamment dans un récent travail consacré à l'ADN mitochondrial humain. Ainsi, Rebecca L. CANN et collaborateurs (1987) tentent de tracer l'arbre phylogénétique de 133 types d'ADN mitochondriaux définis par 12 enzymes de restriction sur cinq populations de quelques individus:

- 20 noirs «africains», en fait américains,
- 34 asiatiques de Chine et du Sud-Est asiatique,
- 46 caucasoïdes incluant des africains du Maghreb et des Moyens Orientaux,
- 21 aborigènes australiens,
- 26 aborigènes mélanésiens de Nouvelle Guinée.

L'avantage de cette étude du polymorphisme de restriction est qu'elle utilise un matériel ADN mitochondrial à mutation plus rapide que l'ADN nucléaire, et qui donc accroît la diversité observable; ce matériel obéit en outre à une transmission maternelle, échappant ainsi à tous les jeux de recombinaison dus à l'appareil à chaque conception et à chaque génération. L'arbre obtenu relie chaque «haplotype» au voisin et détermine pas à pas la structure ancestrale possible pour les expliquer tous. (Figure 10)

Comme dans le travail de JIM WAINSCOAT, certains haplotypes africains, mais pas tous, semblent issus de la même racine très ancienne et ce résultat voudrait confirmer encore cette hypothèse d'une séparation primaire d'un bloc africain et d'un rameau à diversification géographique beaucoup plus large. Le lignage maternel ancêtre de ces deux rameaux était-il africain? REBECCA L. CANN en donne pour argument que situer cette séparation des filles a et b d'une Eve mitochondriale en Afrique «simplifierait le nombre de migrations intercontinentales à envisager pour rendre compte de la distribution géographique des haplotypes ADN mitochondriaux actuels». Cet argument paraît un peu faible et on peut se demander, comme devant le travail de WAINSCOAT, lequel des deux rameaux primaires est plus ancestral que l'autre? Au plan temporel, sur la base de 2% à 4% de différences dans la séquence d'ADN mitochondrial par million d'années, le maximum d'écart observé de 0,57% fait vivre notre Eve autour du 200ème millénaire avt J.C., ce qui pose certains problèmes de compatibilité avec les données archéologiques; mais plus encore que dans le travail de

WAINSCOAT, le déploiement migratoire à partir d'un point africain s'avérant indispensable pour expliquer la typologie haplotypaire, c'est une véritable recolonisation planétaire qui est évoquée par REBECCA L. CANN, car aucune hybridation avec des peuplements antérieurs issus des *Homo erectus* dispersés dans l'Eurasie n'est concevable. Peut-on admettre un monophylétisme de *Homo sapiens* aussi rigoureux? Peut-on admettre une démonstration basée sur aussi peu d'individus et de populations? Peut-on admettre une hypothèse basée autant sur le programme de représentation statistique des données que sur ces données elles-mêmes?

Pourtant des arguments complémentaires seraient apportés par l'étude génomique du chromosome Y dont les séquences sont transmises uniquement par le lignage mâle, ce qui donc théoriquement nous ramène à l'ancêtre ADAM. En effet, une famille de gènes de la sous-région Yq11 révèle, selon G. LUCOTTE (1988), un polymorphisme de restriction (par l'en-

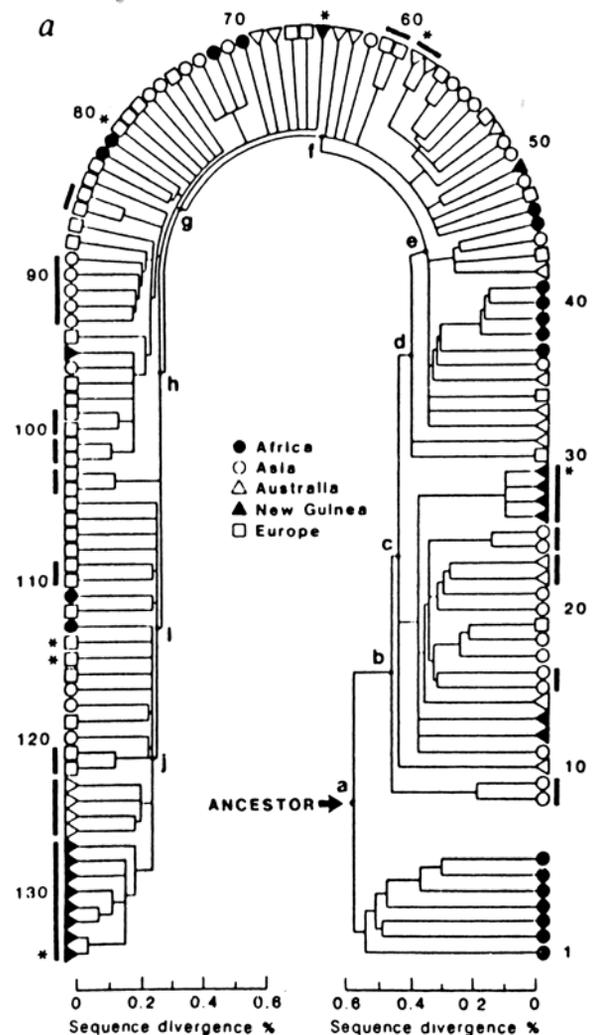


Figure 10: Classification hiérarchique selon la structure génétique d'ADN mitochondrial, définie par 12 enzymes de restriction, pour cinq «populations» humaines. D'après R. L. CANN et al., 1987.

zyme Taq 1) et une sonde adéquate identifie cinq fragments représentant 5 polymorphismes dont les allèles seraient apparus par mutations ponctuelles; sans entrer dans les détails de ce travail, disons qu'entre les haplotypes ainsi définis une filiation peut être établie des primates anthropoïdes jusqu'aux hommes actuels à travers le père ADAM. Et comme le dit J.S. WAINSCOAT avec beaucoup d'humour «si cette étude du chromosome Y prouve qu'ADAM était contemporain d'EVE, alors on pourra même imaginer qu'ils se sont rencontrés».

On a compris que ces études génomiques passionnantes et révolutionnaires nous conduisaient aux perspectives les plus vertigineuses et aux limites de l'anthropologie-fiction.

De nombreux anthropologues s'insurgent contre ces conclusions hâtives et contre l'utilisation un peu abusive de données archéologiques et paléontologiques à l'appui de résultats extrêmement parcellaires.

Comme l'ont exprimé nos collègues LUCA CAVALLI-SFORZA et ANDRE LANGANEY lors du 2ème congrès international de démographie historique à PARIS (juin 1987), «la plus grande prudence devrait être de miser dans un domaine où les données étant rares, il est plus facile d'être spectaculaire que rigoureux».

La génétique moléculaire et les polymorphismes de restriction ne peuvent donc effacer les nombreux travaux sur les polymorphismes antigéniques et protéiques, produits géniques, qui depuis trente ans ont permis une accumulation de données dans des systèmes génétiques nombreux, très informatifs, permettant donc des études statistiques très sophistiquées et de riches comparaisons avec les données archéologiques, démographiques, linguistiques et socio-culturelles dans de nombreuses populations du monde. Ces systèmes polymorphes font d'ailleurs chaque jour l'objet de nouvelles découvertes dont nous ne donnerons dans cet exposé que quelques exemples.

B - MARQUEURS TRADITIONNELS

Dans le choix de nos exemples interviendra évidemment l'intérêt qu'ils présentent pour la génétique des populations pyrénéennes.

B-1) Complexe majeur d'histocompatibilité: gènes de classes II et III.

Le premier de ces exemples, nous le prendrons dans le système du complexe majeur d'histocompatibilité (CMH) et en particulier dans les familles des marqueurs génétiques de HLAD, dits de classe

II, et des marqueurs génétiques du complément, dits de classe III, dont les nombreux allèles ne sont connus que depuis ces 6 ou 7 dernières années. La classe I, rappelons le, est constituée par les loci HLA A, C et B.

B-1-1) La carte du CMH nous a montré la localisation de ces deux familles II et III par rapport aux loci de la classe I. La famille de classe II est certainement très complexe et on lui suppose au moins 12 loci; 4 sont déjà connus et ont permis de reconnaître des séries d'allèles dont la revue a eu lieu en 1984 au 9ème atelier international d'histocompatibilité. Ce sont les loci DR avec 14 antigènes, DW avec 19 antigènes, DQ avec 3 antigènes, ainsi que DZ, DX et DP dont les spécificités sérologiques sont encore mal définies. DQ et DR ont un linkage assez fort et ce sont donc essentiellement pour le moment les variations de fréquences de spécificité de DR et DRW qui sont informatives pour les études de populations; cependant celles des spécificités DP le deviendront certainement à brève échéance.

La liste des antigènes du locus DR s'établit provisoirement ainsi: DR 1, 2, 3, 4, 5 (W11, W12), 6 (W13, W14), 7 (W8, W9, W10, W52 (MT2), WR3 (MT3). Les apports récents de cette famille de gènes sont de trois types:

a) L'information due à leur polymorphisme n'est pas très différente de celle apportée par les loci HLA A, C et B pour les grands rameaux humains comme le montre le tableau 1 tiré de W.R. MAYR (1986).

b) Cependant, à l'intérieur de populations à structures génétiques assez homogènes, telles les populations d'Europe et du Proche Orient, le polymorphisme du DR peut révéler des différences assez nettes; ainsi en est-il, par rapport à un pool de populations caucasoïdes, des différences génétiques au locus DR mesurables pour des échantillons de la Sardaigne, de l'Irlande et de Cardiff (Pays de Galles) ainsi que de l'Arabie (MAYR, 1986). Cette rupture est due principalement aux variations de certains allèles DR en Europe, comme les fréquences élevées des antigènes DR2, DR3 et DR4 qui vont décroissantes du Nord au Sud tandis qu'augmente le DR5; ces allèles DR2, DR3, DR4 ne sont certainement pas seuls en cause car leur liaison avec les B7, B8 et B15

loci HLA considérés	distance CAUC / NEG.	distance CAUC / MONG.	distance NEG / MONG.
HLA - DR	0,007	0,036	0,049
HLA - A+B+C	0,019	0,028	0,047
HLA - A+B+C+DR	0,017	0,029	0,046

Tableau 1: Distances génétiques (f) entre rameaux humains selon la répartition des marqueurs génétiques des principaux loci du système HLA. D'après W. R. MAYR, 1986.

est bien connue dans les haplotypes du CMH. Cependant les études réalisées dans la dernière décennie sur les habitants des Pyrénées et sur ceux de provinces françaises nous ont prouvé (B. BETUEL, 1986):

- que si le DR5 présente bien des valeurs élevées en Catalogne, ces valeurs étaient relativement les mêmes en Dauphiné, en Savoie et en Lorraine.
- que si le DR2 et DR4 présentaient bien les valeurs les plus faibles dans les Pyrénées et dans le Midi-pyrénéen, par contre le DR3, le DR1 et le DR7, maintenaient en Béarn et au Pays Basque les fréquences les plus élevées de France (en association avec les HLA-B18 et A3 pour le DR3, les HLA-B39 et B27 pour le DR1, les HLA-B29 et A23 pour le DR7).

Il semble donc que certains schémas de répartition doivent être en constante révision.

c) La région D du CMH étant plus ou moins similaire à la région I de la souris et les gènes DR en particulier intervenant dans les phénomènes de reconnaissance et de coopération entre les différentes cellules impliquées dans la réponse immunitaire, la présence de certains allèles DR en association avec d'autres marqueurs HLA sera observée chez des individus présentant des maladies autoimmunes et ces corrélations positives permettront de déterminer certains profils génétiques sensibles ou prédisposés à de telles affections. Nous envisagerons plus loin ces problèmes.

B-1-2) Les marqueurs de classe III sont les allotypes des gènes C4, C2 et Bf (facteur properdine). Les produits de ces gènes agissent comme activateurs du Complément C3. Les trois loci sont situés entre les HLA-B et les gènes de classe II dans l'ordre vraisemblable C2, C4, Bf encore qu'un doute subsiste quant à la position de Bf. La fonction du Complément comme médiateur dans tous les processus de défense immunitaire et la situation de sa commande dans le CMH ont provoqué des recherches intensives dont le résultat fut en quatre à cinq ans une explosion des marqueurs de cette famille (MAUFF, 1985). Nous les évoquerons pour les trois loci considérés:

— si C2 est très monomorphe avec le plus souvent l'expression C et très rarement les formes BC et AC, il semblerait que le locus soit par contre le lieu d'un riche polymorphisme de restriction. Les explorations par «blotting» devraient donc faire apparaître de nouvelles mutations.

— BF (facteur properdine): deux sous-types ont été mis en évidence au début: F rapide et S lent, puis deux variants plus rares, F, et SO₇; actuellement on observe 13 variants rapides et 5 variants lents, rares, ainsi qu'un variant très lent (S 1.1) découvert en

Auvergne par l'Unité INSERM 100 de TOULOUSE (HAUPTMANN, ABBAL, 1986) (MOENNARD, 1986).

— Le polymorphisme de C4 est extrêmement complexe. Deux loci C4A et C4B sont le lieu d'un grand nombre de mutations: en C4A, 6 mutations et un allèle AQ0 silencieux, en C4B. 28 mutations et un allèle silencieux BQ0.

Ces différents allèles donnent de nombreuses combinaisons génotypiques, y compris avec les allèles silencieux, ce qui ne facilite ni les calculs de fréquence en génétique de population, ni les interprétations de corrélations entre facteurs du complément et maladies.

Le tableau 2 rapporte quelques résultats des études actuelles sur les principaux marqueurs Bf (facteur properdine).

On remarque sur ce tableau 2, les fréquences tout à fait particulières aux Sardes, aux Basques, aux Touaregs et aux Tunisiens pour les quatre allèles identifiés de ce système.

Pour C4 A et B, 11 combinaisons géniques rendent compte de 75% des combinaisons retrouvées; 3 en représentent la majorité (60 à 74%), au moins chez les caucasoïdes; ce sont A3B1, A3BQ0, AQOB1. Citons quelques exemples de répartitions planétaires originales: la diminution de A3BQ0 chez les mongoloïdes, la diminution très forte de AQ0 chez les mongoloïdes, et l'absence totale de combinaison A4B2 chez les négroïdes.

Dans les études actuelles de populations, l'élément très informatif est constitué par certains déséquilibres de liaison intéressants, d'une part les 3 ou 4 loci de classe III formant ainsi des «complotypes», d'autre part les associations gamétiques de ces complotypes avec les autres gènes de classe II et I pour former des haplotypes particuliers.

Ainsi sont résumés dans les deux tableaux suivants N.° 3 et N.° 4 quelques résultats de l'enquête HLA sur les provinces françaises (P.F.) entreprise de-

FACTEUR B	BfF	BfS	BfF1	BfS07	Nb
ORIENTAUX	.126	.873	-	-	174
NEGROIDES	.519	.461	.010	.010	102
CAUCASOIDES	.204	.778	.012	.010	675
BASQUES FR	.296	.550	.139	.015	201
BASQUES ES	.270	.562	.145	.020	24
SARDES (a)	.219	.578	.198	.005	218
SARDES (b)	.295	.445	.216	.043	139
TUNISIENS	.287	.617	.019	.083	375
TOUAREGS	.345	.577	.022	.056	161

Tableau 2: Fréquences des variants Bf (Facteur properdine) dans quelques populations humaines. D'après HAUPTMANN, ABBAL, 1986, modifié.

puis 1980 dans nos laboratoires toulousains de l'INSERM et du CNRS (actuellement fédérés dans la formation LIGHT (Laboratoire d'ImmunoGénétique Humaine de Toulouse).

Le premier tableau 3 ne donne que les principaux haplotypes HLA A et B retrouvés dans l'enquête et d'une façon schématique leur présence ou non dans 3 régions qui intéressent particulièrement les études basques:

- le Béarn, très proche géographiquement et génétiquement du Pays Basque, (BE) (DEMUR, 1986),
- la Catalogne à l'autre extrémité pyrénéenne (CA) (PREVOST, 1984),
- enfin les Cévennes proches de la Méditerranée (CE) (CONSTANS, 1986).

Le second tableau 4 complique ce premier schéma en indiquant les haplotypes plus complexes correspondants aux associations avec les produits des gènes des classes II et III.

	HAPLOTYPES DES LOCI HLA A et B				
	Workshop 9	P. E.	BE	CA	CE
A1-B8	6.72	6.41	\	//	/
A3-B7	4.77	5.11	//	±	±
A29-B44	1.80	4.33	/	/	±
A11-B35	2.04	2	//	0	±
A1-B17	1.66	1.94	0	/	0
A25-B18	1.38	0.45	-	-	-
A23-B44			-	-	/
A23-B45			-	//	-
A30-B18			//	//	-
A24-B14			/	-	-
A3-B14			-	//	-
A2-B51			-	-	//

Tableau 3: Principaux haplotypes des marqueurs de classe I (loci A, C et B) du système HLA dans les Provinces françaises, et en particulier en Béarn, Catalogne et Cévennes.

HAPLOTYPES DU C.M.H.		
PAYS BASQUE ET BEARN	CATALOGNE	CEVENNES
	A1, Cw7, B8, A3, B1, S/F, DR3	
A30, Cw5, B18	A3, BQ0, F1, DR3	
A29, C, B44	A3, B1, F, DR7	S
A1, Cw6, B17	A6, B1, S, DR7	
	A23, Cw4, B45, A3B1, F, DR7/ B44	
A3, Cw7, B7	A3, B1, S, DR2	A2, Cw7, B51, A3B1, S, DR5

Tableau 4: Principaux haplotypes du Complexe Majeur d'Histocompatibilité (loci HLA A, C et B et loci des facteurs du complément C4A, C4B et Bf) dans quatre «Provinces» du Sud-ouest de la France: Pays Basque, Béarn, Catalogne et Cévennes.

De ces deux tableaux on peut tirer les commentaires suivants:

a) autour de la fréquence particulièrement élevée de BF F1 au Pays Basque et dans les Pyrénées, s'organise un complotype A3, BQ0, F1 puis un haplotype assez particulier; en effet, on observe un linkage fort, d'une part entre ce complotype et la combinaison A30, C5, B18 fréquente en Europe du sud et d'autre part avec DR3; l'haplotype complet est ainsi: A30, CW5, B18, A3, BQ0, F1, DR3. Mis en évidence dès 1978 par A. CAMBON-DEMOUZON et E. OHAYON (1982) au Pays Basque où il est en net déséquilibre de linkage, cet haplotype est retrouvé dans toute la chaîne pyrénéenne jusqu'en Catalogne.

b) Une autre association avec DR3 concerne le complotype AQ0, B1 et Bf S ou F, et dans ce cas, l'haplotype est complété par l'association HLA A1, C7, B8 dont la fréquence dans toute l'Europe du Nord, et donc en France est bien connue; cette fréquence est forte en Cévennes et Catalogne mais par contre faible en Béarn et en Pays Basque.

c) Le complotype le plus banal A3, B1, Bf F (ou S) semble aussi bien associé à DR7 qu'à DR2 ou DR5. Les facteurs HLA A-C-B associés constituent:

— soit un haplotype relativement commun aux populations pyrénéennes (et languedociennes): A29, C-, B44, A3, B1, F, DR7,

— soit des haplotypes bien caractéristiques de zones que nous avons étudiées dans le cadre des provinces françaises:

- BEARN: A3, CW7, B7, A3, B1, S, DR2
- CATALOGNE: A23, C4, B45, A3, B1, F, DR7
- CEVENNES: A23, C4, B44, A3, B1, F, DR7
ou
A2, C7, B51, A3, B1, S, DR5

d) Citons enfin un complotype plus original: A6, B1, S lié à DR7 et à HLA A1, C6, B17 dans un haplotype caractéristique du Pays Basque et d'autres populations pyrénéennes. Un déséquilibre de linkage partiel semble en prouver la présence même en Catalogne.

Ainsi voit-on apparaître grâce à cette association gamétique de marqueurs, à la fois une continuité, une homogénéité de la chaîne pyrénéenne pour certains haplotypes et parallèlement des contrastes et des gradients pour certains autres. Ces acquis recourent parfaitement ce que nous observions déjà dans les systèmes de groupes ABO, Rhésus et autres (BATAILLE, 1974) (MARCELLI-BARGE, 1988)

B-2) Groupes sériques Gm des immunoglobulines

Ces différences entre structures génétiques des populations réparties sur la chaîne pyrénéenne peu-

HAPLOTYPES DU SYSTEME Gm				
HAPL.	γ^1	γ^2	γ^3	
A	4	23	5, 10, 11, 13, 14	CAUC
B	1, 2, 17		21, 28	} CAUC
C	1, 17		21, 28	
D	1, 4	23	5, 10, 11, 13, 14	MONG.
E	1, 17	23	5, 10, 11, 13, 14	MONG.
F	1, 17		10, 11, 13, 15, 16	ORIENT MEDIT.
G	1, 17		10, 11, 13, 15, 16	6 QUEBEC
H	1, 17		10, 11, 13, 15, 28	ISSEQ.
I	1, 17		5, 10, 11, 13, 14	ORIENT AFRIQUE
J	1, 17		5, 10, 11, 13, 14, 28	AFRIQUE
K	1, 17		5, 10, 11, 13, 14	6 DJIBOUTI MEDIT.
L	1, 17		5, 10, 11, 14, 6	SAHARA
M	1, 17		5, 11, 14, 6	AHAGGAR
N	1, 17		5, 11, 6, 24	AFRIQUE NOIRE

Tableau 5: Répartition schématique des principaux haplotypes de système Gm des immunoglobulines selon les grands groupes humains.

vent être objectivées aussi par les haplotypes du système GM dont les marqueurs garnissent les chaînes gamma 1, gamma 2 et gamma 3 des immunoglobulines humaines, autrement dit des anticorps.

Le tableau 5 montre schématiquement cette répartition des marqueurs selon les sous-classes d'immunoglobulines et selon les rameaux humains.

Un tableau 6 construit avec les résultats de diverses enquêtes pyrénéennes et africaines depuis une dizaine d'années met en évidence, avec la même précision que les marqueurs du CMH, les éléments de similarité et de rupture dans le continuum des populations pyrénéennes. On y constate:

a) l'homogénéité des résultats pour les marqueurs caucasoïdes (haplotypes A, B et C) et le contraste des fréquences de ces marqueurs avec celles calculées dans les populations du nord du Sahara et du Sahara central (Touaregs),

b) la présence sur les groupes pyrénéens (basques, béarnais, catalans), cévenol et corse, des fréquences calculables d'un haplotype Gm commun aux populations proches orientales et africaines le Gm 1, 17 / - / 5 10 11 13 14 (I).

c) la présence uniquement dans les populations proches de la Méditerranée (Catalogne, Cévennes et Corse) d'haplotypes rares et communs à l'Afrique du Nord et au Sahara (LEFÈVRE-WITIER, 1982):

HAPLOTYPES Gm	BAS.	BEA	CAT.	CEV.	COR.	NORD SAHARA	TOUA REGS
A 4/-/5*	.683	.598	.654	.682	.710	.387	.189
B+C 1(2)17/-/21	.304	.370	.293	.294	.225	.161	.399
I 1,17/-/5*	.008	.017	.013	.010	.046	.310	.276
K 1,17/-/5*6	-	-	+	+	-	+	-
L 1,17/-/5,10,11,14,6	-	-	+	-	+	+	+
F 1,17/-/10,11,13,15,16	-	-	+	+	-	+	-
M 1,17/-/5,11,14,6	-	-	-	-	-	-	+

Tableau 6: Fréquences de quelques haplotypes du système Gm dans certaines provinces françaises méridionales et des populations du Sahara central algérien.

- haplotype «Djibouti» (K): les facteurs Gm 13 et GM 6 sont présents; cependant ici le Gm 6 n'est pas situé sur les chaînes immunoglobulines gamma 3, alors que dans les populations d'Afrique Noire il est constamment associé sur ces chaînes avec le Gm 24.
- haplotype «Saharien» (L): le facteur Gm 13 a ici disparu et, dans ce cas, on peut faire l'hypothèse que le Gm 6 est probablement sur les chaînes gamma 3.
- haplotype «Proche oriental» (F): les facteurs Gm15 et 16 sont ici présents alors qu'en Afrique Saharienne et Noire existe seulement le Gm 15.

d) L'absence dans ces mêmes populations de l'haplotype «Touareg du Hoggar» (M), propre à ces groupes de pasteurs montagnards.

Ces deux exemples voulaient montrer la finesse de description des populations par des systèmes très polymorphes, qui forment, grâce à la proximité des loci, des haplotypes relativement stables mais capables aussi de délétion et de recombinaison donnant de nouvelles associations parfois déconcertantes. Rappelons encore que ces examens de systèmes polymorphiques aussi complexes que la HLA, les complotypes ou le Gm, doivent évidemment faire l'objet d'enquêtes familiales pour en tester les transmissions de génération à génération.

En fait le choix de ces exemples de marqueurs traditionnels (dont la recherche est surtout le fait d'examen sérologiques) n'était pas neutre. Nous avons ainsi préparé la terrain aux quelques considérations que nous voudrions exprimer sur le problème des relations entre polymorphismes génétiques et pathologie.

Nous en écarterons évidemment les relations directes, c'est à dire les pathologies, souvent moléculaires, dues elles-mêmes à des mutations non-sens perturbant un métabolisme à différents niveaux d'intervention protéiques, enzymatiques ou non.

C - ASSOCIATIONS POLYMORPHISMES GENETIQUES ET MALADIES

Pour bien des mutations des polymorphismes génétiques antigéniques ou protéiques, un lien direct peut être établi entre la structure moléculaire et la pathologie observable. Nous l'avons entrevu pour les hémoglobines anormales, telles les mutations S ou C, et nous aurions pu citer de nombreuses protéines ou enzymes non-sens relevant de mutations ponctuelles ou de délétions comparables à celles des chaînes alpha ou bêta de la globine.

Pour certains systèmes génétiques comme ABO, Rhésus ou Kell, les valeurs sélectives des génotypes relèvent en partie des conflits immunologiques générés par les différents antigènes. Cependant de nombreux systèmes polymorphes paraissent neutres; cette situation de neutralité génait une pensée scientifique très néodarwinienne, et on a tenté pendant des décennies de lier les structures génétiques observées à certaines pathologies, soit endémo-épidémiques comme la tuberculose, la syphilis, le choléra (ou même la peste), ou plus individuelle comme l'ulcère ou le cancer des voies digestives (plus fréquents chez les sujets A) ou le cancer du col utérin pour les femmes O.

Ces associations ont souvent montré des résultats discordants selon les auteurs.

La découverte des polymorphismes du CMH (complexe majeur d'histocompatibilité) a donné un nouvel essor à ces études d'association entre structures génétiques et pathologies. Le tableau suivant 7 offre une image des principales associations reconnues actuellement et sur la compréhension desquelles travaillent de nombreuses équipes de recherche. C'est que la fonction même de différentes familles de gène HLA dans l'identification du soi et du non-soi et dans la défense immunitaire, fait espérer saisir certains des mécanismes physiopathologiques les plus intimes des affections liées aux variants génétiques du CMH.

En amont de ces espoirs, on peut mettre en évidence par l'observation des associations:

— d'une part, les valeurs sélectives de certains profils génétiques dont on peut calculer la fréquence et donc apprécier le poids dans l'adaptation des populations;

— d'autre part, estimer le risque relatif couru par les individus porteurs d'un ou plusieurs éléments de ces profils génétiques.

C'est ce risque relatif (c'est à dire la probabilité pour un individu porteur de contracter la maladie par rapport à un individu non porteur) qui figure dans la colonne de droite de tableau 7.

Les exemples les plus connus actuellement sont les liaisons:

MALADIES	HLA	RISQUE RELATIF
Hémochromatose idiopathique	A3/B14	8/5
Maladie de Hodgkin	A1	1,4
Psoriasis commun	CW6	13
Spondylarthrite ankylosante	B27	87
Syndrome de Reiter	B27	37
Uvéite antérieure aiguë	B27	10
Arthrite psoriasique	B27	11
Périarthrite de l'épaule	B27	6
Arthrite rhumatoïde juvénile	B27	4
Maladie de Behcet	B5	6,3
Thyroïdite subaiguë de Quervain	BW35	14
Dermatite herpétiforme	DR3/B8	15/9
Maladie coeliaque	DR3/B8	11/8
Syndrome de Sjögren	DR3/B8	10/3
Maladie d'Addison	DR3/B8	6/4
Maladie de Basedow	DR3/B8	4/2
Diabète juvénile	DR3/B8 : DR4	4/2 ; 6,4
Myasthénie grave	DR3/B8	13/6
Hépatite chronique active	DR3/B8	7/2
Déficit en C2	DR2/B18/A25	
Syndrome de Goodpasture	DR2	16
Sclérose en plaque	DR2/B5	4/2
Névrite optique	DR2	2
Pemphigus (Juifs)	DR4	14
Arthrite rhumatoïde	DR4	4
Maladie de Berger	DR4	4
L.E.D. (hydralazine)	DR4	5,6
Thyroïdite d'Hashimoto	DR5	3
Anémie pernicieuse	DR5	5

Tableau 7: Risque relatif de maladie selon la structure génétique HLA. D'après J. DAUSSET, HLA 1982. Flammarion, Paris.

— entre HLA A3 et l'hémochromatose idiopathique; nous avons vu que le gène contrôlant le métabolisme du fer dans l'hémoglobine est extrêmement proche du locus HLA-A.

— entre HLA I et maladie de Hodgkin, une hémopathie maligne. Cette association se précise actuellement et nous aurons certainement bientôt des données nouvelles.

— entre HLA B27 et la spondylarthrite ankylosante ainsi que d'autres syndromes à manifestations inflammatoires des articulations. Notons que cette liaison n'intéresse que l'hémisphère nord.

— entre HLA DR3 et diabète juvénile insulino-dépendant, maladie où la sécrétion insulinaire du pancréas est supprimée par un processus d'autodestruction, donc auto-immun.

— entre HLA DR2 et sclérose en plaque, maladie dans laquelle une inflammation de la moelle épinière provoque des foyers de destruction avec des conséquences très invalidantes et souvent mortelles par aggravation.

J'ai souhaité arriver à ces deux exemples parce que ces deux dernières maladies semblent d'une fréquence relativement importante à l'extrémité occidentale des Pyrénées (et donc en pays basque, béarnais et bigourdan) et que leurs associations avec des marqueurs génétiques du système HLA fait l'objet de travaux approfondis depuis une dizaine d'années dans les laboratoires d'ImmunoGénétique Humaine de Toulouse (Fédération LIGHT Unité INSERM 100 et CRPG-CNRS, ex Centre d'Hématologie).

En ce qui concerne le diabète insulino-dépendant (DID), nous avons vu qu'un haplotype du CMH relativement fréquent dans les populations pyrénéennes était porteur du variant DR3. Il se trouve qu'au Pays Basque le risque relatif pour le DID chez les porteurs de DR3, et souvent donc de cet haplotype, est pour le moment le plus élevé du monde. 90% des malades portent ce DR 3 et seulement 19% des témoins (CAMBON-DEMOUZON, 1982). Une observation complémentaire a révélé dans cette même population basque que les porteurs hétérozygotes DR3/DR4 présentaient un diabète plus précoce; ceci a été retrouvé dans d'autres populations européennes; mais chez les basques, il faut noter que le DR4 est très rare, ce qui est un élément favorable pour cette population; d'autant que DR2 et DR7 sont combinés dans d'autres haplotypes et paraissent porteurs d'un caractère de «non susceptibilité».

Le taux de mortalité de la Sclérose en plaques apparaît très élevé dans les Hautes Pyrénées, par rapport aux taux français ou toulousains. La prévalence (taux d'apparition annuelle) est en outre de 40, 2/100.000 habitants, soit une des plus élevée d'Europe. L'enquête actuelle du LIGHT tente d'évaluer la fréquence et la gravité de l'affection dans les Pyrénées Atlantiques, pour comparer les données épidémiologiques bigourdanes avec celle des basques français et des béarnais. La susceptibilité à contracter la Sclérose en plaques semble liée surtout à HLA DR2 en Europe du Nord, peut être à d'autres antigènes HLA dans les Pyrénées, l'Aquitaine, le Toulousain et le Languedoc-Roussillon. Les plus récents résultats obtenus sur 71 familles révèlent une association préférentielle avec DR2, peut être avec DR6 (72% de malades DR2 et/ou DR6 contre 37 % chez les témoins). Une étude du polymorphisme Gm montre que l'association avec l'haplotype Gm4/-/5 aggrave la susceptibilité à la Sclérose en plaques en particulier lorsque les individus sont homozygotes pour cet haplotype Gm (BLANC, 1986). Des associations faibles positives ou négatives ont été aussi constatées entre systèmes HLA et système Gm dans le diabète insulino-dépendant.

Ces observations introduisent ici la notion développée par J. DAUSSET et L. DEGOS d'association génique (index E) ou d'interaction génique ou de synergie génique, quelqu'en soit le mécanisme, qui peut se mesurer dans les familles ou les populations de malades comme un déséquilibre de linkage. Cette approche est importante dans les maladies polygéniques et les marqueurs utilisés ne seront pas seulement les traditionnels mais aussi les génomiques et bien sûr tous les polymorphismes de restriction découverts par les RFLP (WHITTINGHAM, 1984) (DAUSSET, 1985).

Outre leur intérêt en génétique des populations et en anthropologie d'adaptation, tous ces travaux

ouvrent le chapitre considérable de la médecine dite «prédictive», puisqu'on peut progressivement «déduire la présence d'un gène de celle d'un marqueur qui lui est étroitement lié, c'est à dire savoir si un enfant à naître est voué à une maladie ou si un sujet déjà né, enfant ou adulte, est prédisposé à telle ou telle affection» (FREZAL, 1987).

D - CARTE GENETIQUE

Ces travaux sur les associations polymorphismes génétiques et maladies ont enfin une incidence considérable sur la cartographie de l'ensemble du génome humain; depuis des années, c'est parallèlement à des fins scientifiques fondamentales, mais bien plus encore médicales, que vise cette entreprise mobilisant les équipes de pointe de la génétique moléculaire autant que celles de l'immunologie et de la génétique mathématique.

En effet, la découverte et la localisation de nouveaux polymorphismes permettent, comme nous l'avons vu, de lier un type de mutation pathologique à une mutation ou une combinaison de mutations transmises «en bloc» d'un parent à un enfant, c'est à dire souvent à un «haplotype» de ces polymorphismes. Outre le degré d'hétérozygotie que créent ces polymorphismes, ils représentent des balises sur les chromosomes et c'est par rapport à ces balises qu'on tentera de situer grossièrement le gène responsable d'une maladie. Tout le problème de la contribution des polymorphismes génétiques à la cartographie génomique tient dans ce problème de «définition», c'est à dire de pouvoir séparateur ou localisateur sur les chromosomes. Nous le résumerons rapidement:

Les méthodes de ségrégation ou de linkage sont fondées sur des comparaisons de probabilité («lord score») essayant de situer, à partir des transmissions familiales, les positions respectives des commandes géniques et leur ordre sur un même chromosome. Actuellement, grâce au Centre d'Etude de Polymorphismes Humains créé en 1983 par J. DAUSSET (Paris) et R. WHITE (Salt Lake City), 40 familles caucasoïdes comportant 3 générations et groupant 517 individus (avec une moyenne de fratries de 8,2 enfants) sont à l'étude pour 181 marqueurs connus, dont les polymorphismes «traditionnels» ainsi que de nombreuses sondes détectant des polymorphismes nouveaux. Cette opération devrait aboutir à une carte primaire du génome balisé par environ 200 marqueurs sur les 23 chromosomes, et permettre d'utiliser ensuite ces balises pour continuer à cartographier des gènes de maladies héréditaires. Mais cette approche est basée sur des taux de recombinaison, sur des crossing-over entre chromosomes parentaux, et son pouvoir séparateur est d'environ

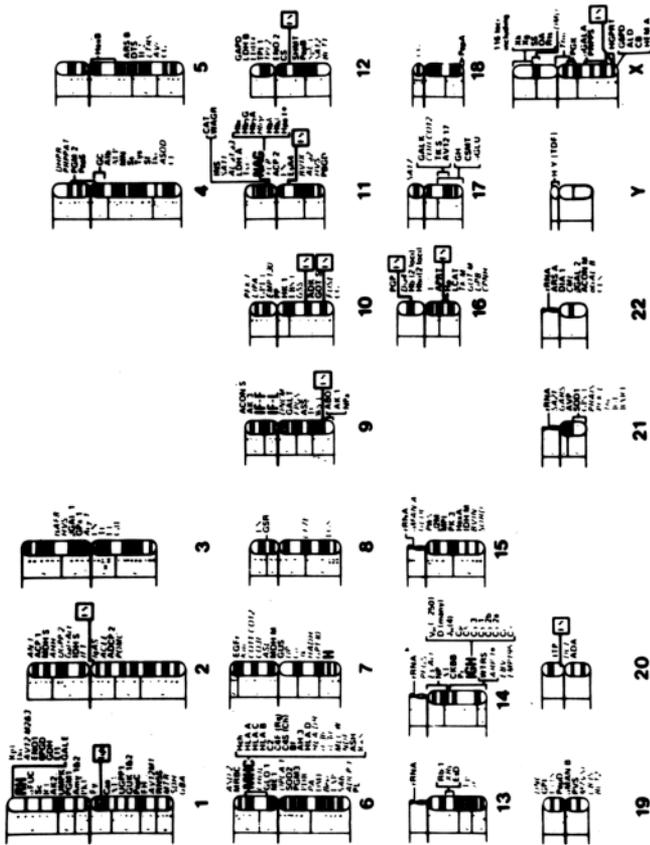


Figure 11: Carte du génome humain. D'après Mc KUSICK, 1981, in FELLOUS, HLA, 1982.

25 centimorgan par rapport à un repère, soit 25 millions de paires de base. Or rappelons-nous qu'un exon ne mesure que 1.000 à 5.000 paires de base. La technique d'étude génomique par sondes sépare, elle, des fragments de 50.000 paires de bases. Donc pour actuellement «marcher le long du génome»:

- soit on fait des sauts énormes (25 centimorgan),
- soit on marche presque «à cloche pied» à raison de 20 pas pour parcourir 1 seul centimorgan.

Les techniques récentes cherchent à trouver un outil, c'est à dire des polymorphismes de restriction, dont le pouvoir séparateur tournerait autour d'un centimorgan et ceci en attendant le séquençage total automatique; ce séquençage n'est pas impossible puisque les chercheurs japonais viennent de l'appliquer au génome de bactérie *Escherichia coli* (3.000 gènes seulement!).

Les 11 et 12 septembre 1987, à Paris, avait lieu la réunion de la Carte Génétique Humaine (HGM 9), après plusieurs jours d'un atelier très animé; cette réunion actualise la carte génomique; la figure 11 est celle de la carte génétique de 1981 du Professeur V.A. Mc. Kusick, coordinateur des acquis dans ce

domaine. La carte génétique humaine de 1987 est beaucoup plus chargée puisque le nombre de gènes cartographiés en 1985 a doublé et atteint maintenant 2 à 3.000. La génome humain étant estimé à 50.000 ou 100.000 gènes pour certains auteurs, le travail de cartographie est loim d'être terminé (Mc. KUSICK, 1987).

E - STATISTIQUES ET REPRESENTATIONS DES DONNEES

Je ne dirai rien des traitements statistiques et des représentations tirées de ces traitements. Nos maîtres et collègues LUCA CAVALLI-SFORZA, ALBERTO PIAZZA et DEREK H. ROBERTS, nous en ont montré suffisamment sur ce point, au cours de cette même réunion. Cependant j'aimerais vous présenter trois comparaisons de populations tirées des données accumulées sur 15 provinces françaises par nos laboratoires toulousains (LIGHT). Ces données génétiques concernent les systèmes polymorphiques antigéniques Rhésus, Gm et de quatre protéines sériques: Group-component, Transferrine, Alpha-antitrypsine, Haptoglobine. Les comparaisons de populations ont été réalisées par la distance du χ^2 (OHAYON, CAMBON-THOMSEN, 1987). (Figure 12).

Si nous nous contentons de suivre les positions génétiques de Béarn, parce que cette population évoque les structures génétiques du Pays Basque et de la Catalogne, nous constatons que:

— dans le premier dendrogramme Rhesus, le Béarn est proche des Flandres, du Poitou et du Finistère Sud. La Catalogne es par contre proche du Finistère Nord et du Limousin.

— dans le second dendrogramme Gm, le Béarn est isolé, la Catalogne reste proche du Finistère Nord mais aussi de la Corse.

— dans le troisième dendrogramme «Protéines sériques», le Béarn rejoint l'Auvergne, tandis que la Catalogne est proche génétiquement, nou plus du Finistère Nord mais du Finistère Sud, du Beaujolais et du Poitou.

A considérer ces représentations et ces comparaisons système génétique par système génétique, et ce pour les populations connaissant une culture et un environnement relativement homogènes, une certaine inquiétude nait quant aux synthèses tentées par les traitements globaux de ces données et les interprétations qu'il faut ensuite, très laborieusement, en tirer. Certes il apparait nécessaire d'accumuler toutes les données réunies sur les différents systèmes polymorphiques explorés au laboratoire car selon l'image citée par ALBERTO PIAZZA «une hirondelle ne fait pas le printemps»; mais à combien d'hirondelles est-il devinable??

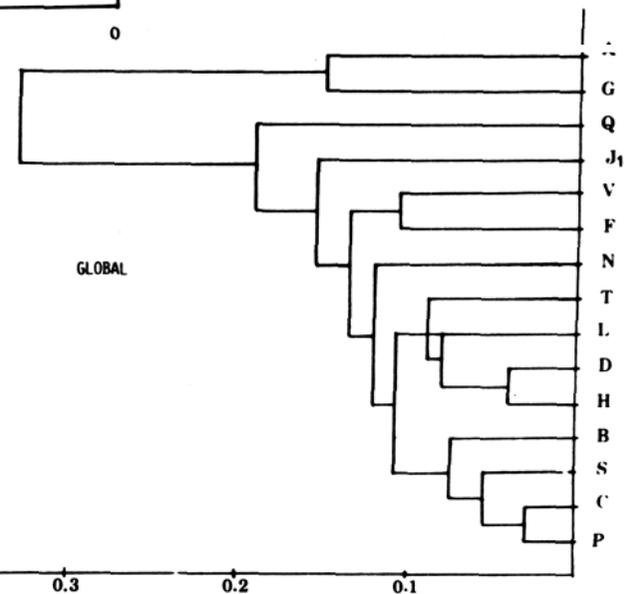
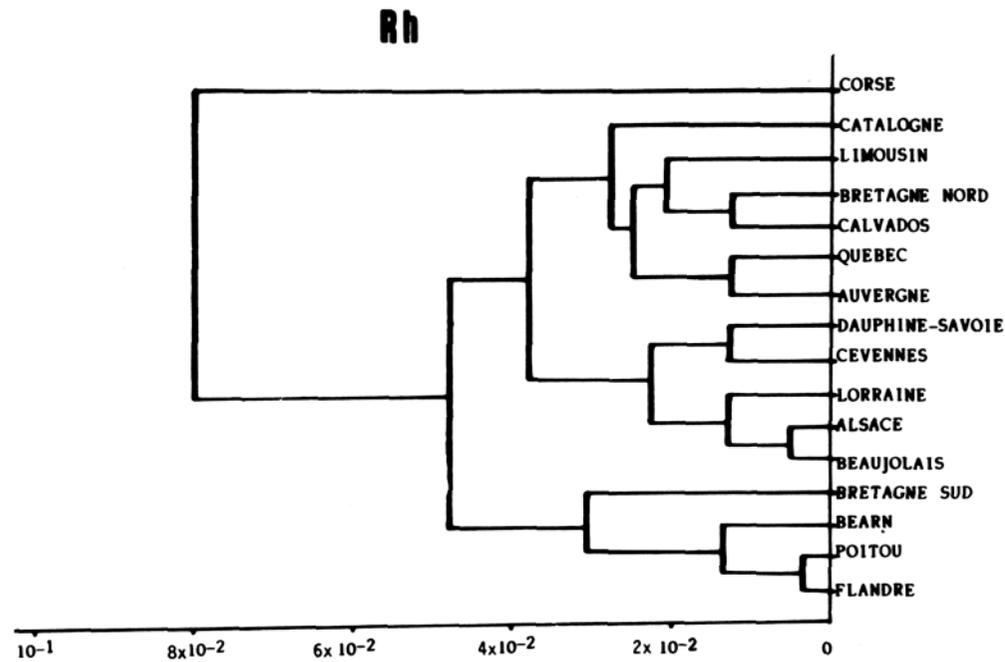
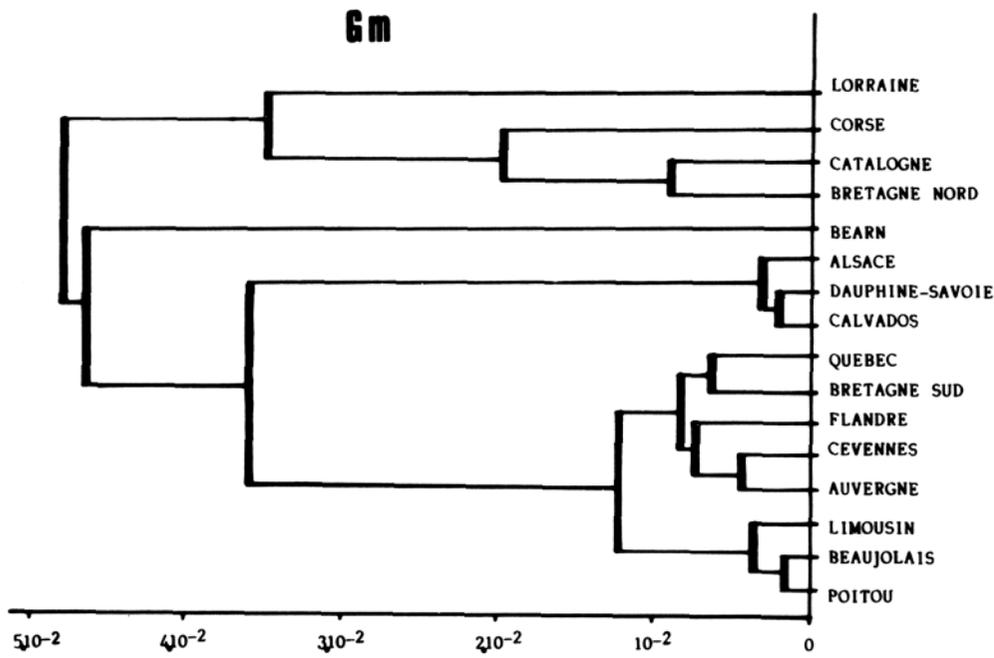


Figure 12: Comparaisons des classifications hiérarchiques entre 15 Provinces françaises et la québec obtenues selon les fréquences des marqueurs génétiques (distances du Chi2) du système Rhésus, du système Gm et de quatre protéines sériques (Group-component, Transferrine, Haptoglobine, Alpha-1-chymotrypsine). A: Auvergne, G: Béarn, Q Quebec, J1: Alsace, V: Corse, F Flandres, N: Bretagne nord, T: Basse Normandie, L Limousin, D: Dauphiné-Savoie, H: Cévennes, B: Beaujolais, Bourgogne, S: Bretagne su, C: Catalogne, P: Poitou. D'après E. CHAYON et A. CAMBON-THOMSEN, 1986.

CONCLUSIONS PROVISOIRES

A cet exposé, je vois actuellement deux types de conclusions: générales sur les systèmes polymorphiques, particulières sur leurs applications aux populations pyrénéennes et au peuplement de la chaîne.

a) Générales

Le terme de polymorphismes génétiques «sanguins» pratiquement ne signifie plus grand chose; les recherches sur l'ADN, sur les groupes HLA se font sur des cellules très différentes qu'on peut obtenir par biopsie d'organes, aspiration de moelle osseuse, ponction de liquide amniotique, etc... cependant le sang reste le tissu d'élection des généticiens par sa commodité de prélèvement, de manipulation, de conservation et de transport.

Les polymorphismes génétiques connaissent actuellement leur second souffle: souffle révolutionnaire conduisant à une «nouvelle génétique», donc à une nouvelle anthropologie génétique; ils restent l'outil fondamental de progrès dans la connaissance de la variabilité humaine. Mais ils sont un outil coûteux car il apparait de plus en plus nécessaire de mener de front la génétique «moléculaire» du génome et celle des produits de ce génome, c'est à dire des marqueurs «traditionnels» antigéniques et protéiques. Les laboratoires seront rares qui pourront s'offrir de telles études...!

Les progrès en termes de description seront certainement plus rapides que ceux touchant à la signification du polymorphisme et au vu des résultats, il faudra souvent louvoyer entre les théories interprétatives, rappelées encore par FRANCOIS GROS (1987) lors de la réunion sur la Carte génétique de l'Homme le 12 septembre 1987:

— théorie de la neutralité où les polymorphismes paraissent responsables de certains désavantages mais plutôt rares par rapport à leur invraisemblable richesse due au hasard,

— théorie physiologique impliquant que si tant de formes moléculaires existent c'est qu'elles doivent avoir des rôles spécifiques que nous ignorons encore,

— théorie régulatoire pour laquelle les polymorphismes ont une capacité à affronter des milieux et des situations métaboliques très diverses,

— théorie moléculaire où toutes les structures conservées dans le génome par l'évolution feraient partie d'un ensemble fonctionnel, d'une superstructure chromatinienne que nous ignorons encore totalement.

b) Particulières

Nous avons essayé de montrer que la «nouvelle génétique» et les nouveaux polymorphismes génétiques remettront vraisemblablement en question bien des acquis sur les caractéristiques des populations pyrénéennes actuelles, basques en particulier, comme sur l'origine de ces caractéristiques génétiques dont nous avons déjà exposé un résumé à ZAGREB (Yougoslavie) en 1981 (LEFEVRE-WITIER, 1981).

Pour le moment les laboratoires toulousains (LIGHT) continuent à collecter les matériels cellulaires nécessaires à ces études sur les populations basques et béarnaises ou sur les populations catalanes de plaine ou de montagne. Un problème demeure: celui des Pyrénées centrales dont le peuplement semble relativement original et pour lesquelles nous manquons de données sur les deux versants.

Il faut espérer que les recherches en cours apporteront des arguments complémentaires à la controverse sur la parenté des basques avec les différents groupes méditerranéens: euroméditerranéens ibéro-insulaires, afro-méditerranéens.

Une suggestion pourrait conclure cet exposé: Pourquoi au décours de ce Congrès Mondial Basque ne pourrait-on mettre en place une Fondation anthropologique dont les ressources permettraient un complément de coordination des recherches espagnoles et françaises ou autres sur le peuplement basco-pyrénéen?

BIBLIOGRAFIA

- BARALLE, F.E., SHOULDERS, C.C., PROUDFOOT, N.J.
1980. The primary structure of the human epsilon-globin gene, *Cell*, 21, 621-626.
- BATAILLE, C.
1974. *Mariages et marqueurs génétiques en CAPCIR (Pyrénées Orientales)*, Toulouse (Thèse Faculté Médecine): Université Paul Sabatier.
- BERNARDS, R. and FLAVELL, R.A.
1980. Physical mapping of the globin gene deletion in hereditary persistence of foetal haemoglobin (HPFH), *Nucleic Acid Res* 8, 1521-1534.
- BETUEL, H., GEBUHRER, L., LAMBERT, J.
1986. Marqueurs HLA de classe II dans les Provinces Françaises, *Human Population Genetics - Génétique des Populations Humaines*, Paris: INSERM, Colloque 142, 265-278.
- BLANC, M., CLANET, M., BERR, C., DUGOUJON, J.M., RUYDAVET, J.B., DUCOS, S.J., RASCOL, A. and ALPEROVITCH, A.
1986. Immunoglobulins allotypes and susceptibility to Multiple Sclerosis - an epidemiological and genetic study in the Hautes Pyrenees county of France, *Journal of the Neurological Sciences* 75, 1-5.

- CAMBON-DEMOUZON, A., OHAYON, E., HAUPTMANN, G., SEVIN, A., ABBAL, M., SOMMER, E., VERGNES, H. et DUCOS, J.
1982. HLA-A, B, C, Dr antigens, Bf, C4 and glyoxalase in diabetes mellitus (IDDM), *Tissue antigens*, 19, 336.
- CANN, R.L., STONEKING, M. and WILSON, A.C.
1987. Mitochondrial DNA and human evolution, *Nature* 325, 31-36.
- CAVALLI-SFORZA, L.L.
1987. Génétique et histoire du peuplement: remarques. Commentaires A2 - le peuplement du monde avant 1800, *Annales de Démographie historique* 49, 53-55.
- CONSTANS, J., SEIGNALET, J., MULLER, J.Y.
1986. Les Cévennes, *Human Population Genetics - Génétique des Populations Humaines*, Paris: INSERM, Colloque 142, 361.
- DAUSSET, J., PLA, M.
1985. *HLA*, Paris: Flammarion.
- DEMUR, C., CALOT, M.
1986. Béarn, *Human Population Genetics - Génétique des Populations Humaines*, Paris: INSERM, Colloque 142, 362.
- EMBURY, S.H., MILLER, J.A., DOZY, A.M., KAN, Y.W., CHAN, V., TOD, D.
1980. Two defferent molecular organizations account for the single alpha-globin gene of the alpha-thalassemia 2 genotype, *J. Clin. Invest.*, 66, 1319-1325.
- FLAVELL, R.A.
1980. The analysis of gene expression in hegher organisms, *Trends In Biochem. Sci.* 5, 313-315.
- FREZAL, J.
1987. La nouvelle génétique et la carte des gènes de l'Homme, Rapport *Human Gene Map N.°9*, Paris, 11/12-09-87.
- GROS, F.
1986. *Les secrets du gène*, Paris: Odile JACOB, Seuil.
- HAYPTMANN, G., ABBAL, M.
1986. Marqueurs du complément (Bf, C4A, C4B) dans les Provinces Françaises, *Human Population Genetics - Génétique des Populations Humaines*, Paris: INSERM, Colloque 142, 279-294.
- HIGGS, D.R. PRESSLEY, L., CLEGG, J.B., WEATHERALL, D.J., SERJEANT, D.R.
1980. Alpha-thalassemia in black populations, *Johns Hopkins Med. J.*, 146, 300-310.
- HUMPHRIES, R.K., LEY, T.J., GOLDSMITH, M.E., CLINE, A., KANTOR, A.J., NIENHUIS, A.W.
1982. Silent nucleotide substitution in bêta + thalassemia gene activates a crypto-splice site in bêta-globin RNA coding sequence, *Blood* 60 (supp), 54a.
- JEFREYS, A.J. and HARRIS, S.
1982. Processes of gene maturation, *Nature* 296, 9.
- KAN, Y.W. and DOZY, A.M.
1978. Polymorphism of DNA sequence adjacent to human be-taglobin structural gene: relationship to sickle-cell mutation, *Proc. Nat. Acad. Sci. USA*, 75, 5631.
- KAPLAN, J.C.
1981. ADN et Génétique médicale *Génétique médicale, acquisitions er perspectives*, Paris: Flammarion Médecine.
- KOLLER, B.H., GERAGHTY, T., DEMARS, R., ORR, H.T.
1986. Cloning and mapping of each member of the HLA class I gene family from a human LCL, *Human Gene Map n.° 8*, Munich: Kaarger.
- KULOSIK, A.E., WAINSCOAT, J.S., SERJEANT, G.R., KAR, B.C., AL AKAMY, B.B., ESSAN, G.J.F., FALUSI, A.G., HAQUE, S.K., HILALI, A.M., KATE, S., RANASINGHE, A.E.P. AND WEATHERALL, D.J.
1986. *Geographical survey of bêta-S-globin genes haplotypes: evidence for independant origin of the sickle-celle mutation*, *Am. J. Hum. Gen.*, 39, 239-244.
- LANGANEY A. et SANCHEZ-MAZAS A.
1987. Le réseau génétique humain et son histoire - génétique et histoire du peuplement. Rapport A2 - Le peuplement du monde avant 1800, *Annales de Démographie historique*, 49, 44-52.
- LAWN, T.M., ESTRATIADIS A., O'CONNEL, C. and MANIATIC, T.
1980. The nucleotid sequence of the beta-globin gene, *Cell*, 21, 647-651.
- LEFEVRE-WITIER, Ph.
1981. Hémotypologie des habitants du CAPCIR (Pyrénées Orientales), des basques français (Pyrénées Atlantiques) et de quelques autres communautés pyrénéennes, *7th School of Biological Anthropology -ZAGREB - 7/10 septembre, 1981*.
1982. *IDELES - Village du Hoggar - Structures et dynamique génétiques d'une communauté du Sahara algérien*, Toulouse (Thèse de doctorat d'état en biologie humaine): Université Paul Sabatier.
- LUCOTTE, G.
1988. Origin of modern humans: evidence from a Y chromosome specific sequence polymorphic DNA probe, *Origin and Dispersal of Modern Humans*, Cambridge: Cambridge University Press (sous presse).
- Mac KUSICK, V.A.
1987. Concluding remarks, *Human Gene Map N.°9*, Paris, 11/12-09-87.
- MARCELLI-BARGE, A., POIRIER, J.C., PREVOST, P., MENEAU, Ph., LEFEVRE-WITIER, Ph et HORS, J.
1988. Marqueurs HLA, complotypes, marqueurs érythrocytaires, sériques et enzymatiques en Catalogne Française, *Les Marqueurs Génétiques dans les Provinces Françaises* PARIS: INSERM-DOIN (Collection «Grandes enquêtes») (sous presse).
- MAUFF, G.
1985. The complement system: genetics and function, *Trends in Haematology*, I WARSAW Institute of Haematology.

- MAXAM, A.M. and GILBERT, W.
1977. A new method for sequencing DNA, *Proc. Nat. Acad. Sci. Usa* 74, 560-564.
- MAYR, W.R.
1986. Use of serologically defined HLA class II markers in population genetics, *Human Population Genetics - Génétique des Populations Humaines*, Paris: INSERM, Colloque 142, 133-142.
- MOENNARD, C.
1986. *Auvergne, Human Population Genetics - Génétique des Populations Humaines*, Paris: INSERM, Colloque 142, 358.
- NEĪ, M. and ROYCHAUDHOURY, A.K.
1982. *Evol. Biol.*, 14, 1-59.
- ORKIN, S.H., MICHELSON, A.
1980. *Partial deletion of the alphaglobin structural gene in human alpha-thalassemia*, *Nature* 286, 538-540.
- OTTOLENGHI, S., GIGLIONI, B., TARAMELLI, R. et al.
1982. Molecular comparison of delta-bêta-thalassemia and hereditary persistence of foetal haemoglobin DNAs; evidence of a regulatory area?, *Proc. Nat. Acad. Sci. USA*, 79, 2347-2351.
- PAGNIER, J., MEARS, G., DUNDA-BELKHODJA, O., SCHAEFFER-REGO, K.E., BELJORD, C., NAGEL, R.L. and LABIE D.
1984. Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa, *Proc. Nat. Acad. Sci. USA* 81, 1771-1773.
- PRESSLEY, L., HIGGS, D.R., CLEGG, J.B., WEATHERALL, D.J.
1980. Gene deletion in alpha-thalassemia prove that the 5' sigma-locus is functional, *Proc. Nat. Acad. Sci. USA* 77, 3586-3589.
- PREVOST, P., MENEAU, Ph., LEFEVRE-WITIER, Ph., MARCELLI-BARGE, A., POIRIER, J.C., SCHMID, M. et HORS, J.
1984. Etude préliminaire des marqueurs HLA en Catalogne Française, *Revue Française de Transfusion et Immuno-Hématologie*, XXVII, 2, 181-190.
- RAO, S.K.,
1988. *Variabilité génétique de la drépanocytose. Influence de l'expression des gènes non-bêta de globine*, Thèse de doctorat de biochimie de l'université Paris V. Paris.
- SHEN, S.H., SLIGHTOM, J. and SMITHIES, O.
1981. *A history of the human foetal globin gene duplication*, *Cell* 26, 191.
- SPRITZ, R.A., DERIEL, J.K., FORGET, B.G. and WEISSMAN, S.M.
1980. Complete nucleotide sequence of the human deltaglobin gene, *Cell* 21, 639-646.
- SPRITZ, R.A. and FORGET, B.G.
1983. The thalassemias: molecular mechanisms of human genetic disease, *Am. J. Hum. Genet.* 35, 333-361.
- SPRITZ, R.A., JAGADEESWARAN, P., CHOUDARY, P.V. et al.
1981. Base substitution in an intervening sequence of a bêta + thalassemia human globin gene, *Proc. Nat. Acad. Sci.*, 78, 2455-2459. USA
- TRECARTIN, R.F., LIEBHABER, S.A., CHANG, J.C., LEE, K.Y., KAN, Y.W.
1981. Bêta o thalassemia in Sardinia is caused by a nonsense mutation, *J. Clin. Invest.*, 68, 1012-1017.
- TREISMAN, T., PROUDFOOT, N.J., SHANDER, M., MANIATIS, T.
1982. A single-base change at a splice site in a bêta O thalassemia gene causes abnormal RNA splicing, *Cell*, 29, 903-911.
- WAINSCOAT, J.S., HILL, A.V.S., BOYCE, A.L., FLINT, J., HERNANDEZ, M., THEIN, S.L., OLD, J.M., LYNCH, J.R., FALUSI, A.G., WEATHERALL, D.L. and CLEGG, J.B.
1986. Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms, *Nature* 319, 491-493.
- WHITTINGHAM, S., MATHEWS, J.D., SCHANFIELD, M.S., TAIT, B.D., MACKAY, I.R.
1981. Effect of gene interaction on susceptibility to disease, *Tissue antigens* 17, 252-254.
- ZIEGLER, A., ZIMMER, F.J., FONATSCH, C., SO, A., TROWSDALE, J.
1986. Mapping of class I and II genes using HLA deletion mutant cell lines, *Human Gene Map n.° 8*, München: Kaarger.