

Texte de la 440^e conférence de l'Université de tous les savoirs donnée le 19 juillet 2002

Génétique, Populations et Maladies

De la génétique des populations à la génétique épidémiologique et à la santé publique

Par Anne Cambon Thomsen

Gènes, Populations et Maladies : pour aborder ce thème, c'est la vision de la Génétique des populations humaines que j'ai choisie et qui va, à travers quelques exemples de projets de recherche, nous conduire vers la génétique épidémiologique et vers des problématiques de santé publique.

La Génétique des Populations, en quelques mots ?

Un regard sur l'histoire et la géographie des populations humaines à travers leurs gènes !

C'est notamment à travers le livre de Cavalli-Sforza, Menozzi et Piazza (1994) sur l'histoire et la géographie des gènes humains (1)* que cette discipline peut être illustrée. Elle a existé bien avant les apports de la génétique moléculaire. Le message à retenir ? Il est possible, à partir d'études de marqueurs génétiques, tels que groupes sanguins et gènes HLA, de retracer une géographie génétique qui est globalement cohérente avec l'histoire des populations.

Mais commençons par quelques définitions (2).

Tout d'abord, la Génétique des Populations, c'est l'étude des caractéristiques génétiques des populations, permettant de définir leur structure génétique et de déterminer l'influence de cette structure génétique sur leur évolution.

Deux stratégies, deux approches et deux objectifs sont possibles :

Etudier les populations et utiliser les marqueurs génétiques et leurs polymorphismes pour en savoir plus sur l'histoire d'une population et des migrations qui s'y rapportent ou étudier l'évolution du génome, en utilisant les données sur les populations pour révéler les spécificités d'évolution des marqueurs génétiques au sein du génome. Dans un cas la connaissance de la population est le but de la recherche, dans l'autre la population est le moyen de répondre à des questions de recherche sur le génome.

* Les chiffres entre parenthèse renvoient aux références en fin de texte

Quant à la génétique épidémiologique, c'est l'analyse de la contribution génétique aux causes et mécanismes des pathologies ; une nuance se profile lorsqu'on utilise le terme d'épidémiologie génétique mettant l'accent sur l'analyse des facteurs de risques génétiques de développer une pathologie. Ce qu'il faut savoir, c'est que beaucoup de maladies font entrer en jeu dans leur développement des facteurs génétiques et des facteurs environnementaux, et que l'analyse des interactions entre ces facteurs constitue aujourd'hui un axe essentiel de recherche.

Que dire maintenant de la rencontre entre génétique des populations et génétique épidémiologique ?

Lors d'une étude de génétique épidémiologique, intégrer des éléments de génétique des populations est essentiel, le choix des modalités d'étude dépendant des caractéristiques de la population. C'est ainsi qu'il sera intéressant, si l'étude de la pathologie nécessite de disposer de nombreux cas au sein d'une même famille, de réaliser une étude sur plusieurs générations dans une population présentant beaucoup de familles nombreuses. L'étude de la génétique d'une maladie récessive pourra avancer plus rapidement si l'on étudie des populations à fort degré de consanguinité.

En résumé, le fait de connaître les structures et les caractéristiques génétiques des populations permet d'optimiser les stratégies d'étude par le choix de la population d'étude.

Il faut savoir qu'en génétique, et notamment pour pouvoir appliquer des programmes statistiques d'analyses génétiques, il est essentiel de connaître la distribution ou fréquence des marqueurs génétiques dans les populations.

A travers quelques exemples, voici une démonstration de l'imbrication et de la complémentarité de ces deux disciplines et de leur nécessaire collaboration !

Prenons un exemple d'actualité dans un champ d'investigation en vogue : la pharmacogénétique et la pharmacogénomique.

Termes composés reposant sur la découverte récente des bases génétiques de la variabilité de réponse à certains traitements selon les individus. Un problème important qui concerne notamment l'apparition d'effets secondaires, parfois très indésirables ! De là, émerge l'idée d'une possibilité d'adaptation des traitements en fonction de la constitution génétique des malades pour les gènes pertinents vis à vis des molécules utilisées. D'où la nécessité de connaître la répartition de marqueurs génétiques d'intérêt dans les populations, par exemple pour un marqueur impliqué dans la survenue d'un effet secondaire... C'est ainsi que la génétique des populations, initialement orientée vers l'histoire du monde, le processus de

peuplement et la structure génétique des groupes humains, devient une dimension indispensable aujourd'hui de la recherche médicale et de la pharmacologie.

Après cette introduction, trois axes d'études seront évoqués, à travers des projets menés au cours de mes travaux de recherche :

- Un aspect de génétique des populations classique, à travers une étude franco-québécoise réalisée dans les années 1980-1990 et connue sous le nom "Les marqueurs génétiques dans les Provinces françaises";
- Deux applications en médecine de la génétique des populations, d'une part sur les études de génétique de pathologies, avec un exemple, le Diabète de type 1 et d'autre part en transplantation où la notion de diversité génétique et de compatibilité tissulaire est essentielle, avec un exemple portant sur le domaine des greffes de moelle osseuse.

Marqueurs génétiques et Provinces françaises

Années 1980-1985 : Premier exemple d'un projet de recherche prospectif de Génétique des Populations au niveau national, à travers une étude sur la répartition de marqueurs génétiques dans la population française (3 - 6).

Une diversité génétique se profile t-elle derrière la diversité culturelle au sein même d'un pays ?

La notion de diversité génétique n'est pas une notion récente: Ainsi chez l'homme, les groupes sanguins ont été décrits au début du XX^e siècle et très rapidement ont suivi les premières connaissances sur le typage de la diversité de ces systèmes génétiques dans des populations.

Dans les années 1980, existaient des données éparses, via les centres de transfusion et des informations sur les groupes sanguins et divers autres systèmes génétiques, notamment dans certaines régions, telles que Pyrénées, Bretagne, Alsace. Ces données témoignaient de l'existence d'une diversité génétique en France. Autre fait de cette période: l'intérêt grandissant pour les analyses d'association entre le système HLA et des maladies.

Pourquoi le système HLA ? C'est un système génétique crucial en immunologie, impliqué dans les greffes pour la compatibilité et la survie des greffons et dans la génétique de maladies, notamment maladies comportant des dérèglements immunitaires, comme les maladies auto-immunes.

Devant la multiplication des études décrivant des différences de distribution des marqueurs HLA entre groupes de malades et groupes de témoins (non malades) il devenait indispensable de connaître la distribution des groupes HLA en France, dont on n'avait pas d'idée globale.

Voilà le point de départ de l'étude "Marqueurs génétiques dans les Provinces françaises".

Cette étude a donc porté sur l'analyse de marqueurs HLA et non HLA, sur la base des connaissances de l'époque en terme de marqueurs et de procédés de typage, par des études au niveau des cellules, ou des protéines du sérum, et avant l'avènement des études menées directement sur l'ADN.

Quelles sont les caractéristiques de cette étude ?

Son originalité tout d'abord : c'est une étude prospective coordonnée et organisée, basée sur une stratégie et une méthodologie communes. La plupart des études de génétique de populations portant sur plusieurs échantillons de populations étaient habituellement basées sur une compilation des données recueillies dans la littérature scientifique, mais qui pouvaient porter sur des échantillons de taille différente, recueillis selon des modes d'échantillonnage variables, étudiés pour des panels de marqueurs génétiques hétérogènes, avec des méthodologies de typage non standardisées.

Sa réalisation concrète : l'étude s'est axée sur 15 régions rurales et peu brassées en France et au Québec et est basée sur un panel de 50 à 100 familles par zones, dont les membres, issus de cette région depuis au moins trois générations consentaient à répondre à un questionnaire et à donner un échantillon de sang pour cet usage scientifique.

Bilan de l'étude : 15 régions de France (Béarn, Catalogne, Corse, Cévennes, Savoie-Dauphiné, Beaujolais-Bourgogne, Lorraine, Alsace, Auvergne, Limousin, Poitou, Flandre, Basse Normandie, Bretagne Nord, Bretagne Sud) et le Québec, 1382 familles participantes avec en général 4 membres prélevés, 1360 familles analysées du point de vue des marqueurs génétiques, correspondant à environ 2700 individus non apparentés (car on calcule les fréquences sur les parents), 26 locus génétiques étudiés dont 8 locus dans la région HLA, grâce à la collaboration de 20 laboratoires et de 3 centres d'analyses.

Avant de dévoiler les résultats, voici un détour utile sur le système HLA (7, 8, 9).

Première caractéristique du système HLA, sa complexité et son grand nombre de gènes. Le complexe HLA est un système génétique situé sur le bras court du chromosome 6 et comprenant de nombreux gènes, notamment les locus HLA-A, HLA-B, HLA-DR. Pour représenter cette région du génome, on fait appel à des cartes génétiques, qui se densifient et se complexifient au fil des années et des études, avec aujourd'hui plus de 200 gènes ou fragments de gènes identifiés. Chaque personne a deux variants (ou allèles) pour chaque locus, l'un provenant de son père, l'autre de sa mère. La combinaison sur un même chromosome d'origine paternelle ou maternelle des variants à plusieurs gènes voisins (combinaison HLA-A, B, DR dans notre exemple) constitue un haplotype, dont on peut aussi

calculer la fréquence dans une population, certains étant plus fréquemment retrouvés que d'autres. Imaginez la complexité d'un haplotype ou d'une combinaison HLA à 200 locus ! En pratique, pour les applications courantes en médecine on se limite à une information HLA-A, B, DR, parfois quelques locus de plus, comme on le voit sur la carte des principaux gènes de la région HLA présentée en figure 1(9).

Deuxième caractéristique du système HLA, son très grand polymorphisme : 243 allèles HLA-A, 478 allèles HLA-B, 307 allèles HLA-DRB1 ; à comparer, par exemple à la variabilité existant pour le système sanguin A, B, O et ses 3 allèles principaux.

Ce polymorphisme immense génère une multiplicité de combinaisons d'haplotypes possibles, puisqu'en théorie, chaque allèle d'un locus peut être associé à chaque allèle d'un autre locus sur le même chromosome.

Troisième caractéristique, la distribution contrastée des combinaisons haplotypiques dans les populations, avec des combinaisons fréquentes, d'autres rares, voire absentes selon les populations.

Quatrième caractéristique : l'existence d'un phénomène appelé association gamétique ou déséquilibre de liaison. Il correspond à une association préférentielle d'allèles à des locus voisins et résulte en une fréquence de certaines combinaisons différente de celle attendue selon le hasard. Dans une population, s'il y a réarrangement aléatoire, les associations entre allèles se font au hasard et s'il y a déséquilibre de liaison, certaines combinaisons d'allèles seront favorisées. Il faut savoir que ce phénomène de déséquilibre de liaison évolue dans le temps, au fil des générations, selon le taux de recombinaison dans la région génomique étudiée, celui-ci traduit les possibilités d'échange entre deux chromosomes lors de la méiose : un déséquilibre de liaison au sein du génome est détruit par un fort taux de recombinaison et inversement perdure très longtemps en l'absence de recombinaison entre les locus concernés. Grâce à ce phénomène, il existe des haplotypes conservés; par exemple HLA-A1, B8, DR3, très largement étudié. Evolution du génome, sélection et migrations concourent à expliquer ce phénomène, dont l'analyse permet notamment de retracer les migrations de populations (10).

Etude de populations (individus non apparentés seulement) ou de familles, phénotypes et haplotypes : pourquoi et comment ?

Pour rappel, le phénotype est le résultat du typage d'un individu, composé de paires de données pour chaque locus typé et l'haplotype est une combinaison d'allèles à plusieurs locus portée par un chromosome et héritée du père ou de la mère** ; "Travailler en haplotype"

** Phénotype (3 paires): HLA-A (01,02), HLA-B (44, 44), Microsatellite (3, 4)

signifie savoir réorganiser les paires de données du phénotype en deux haplotypes constitué des deux combinaisons l'une d'origine paternelle et l'autre maternelle. Lors d'une étude de populations, on obtient des phénotypes HLA et par une étude de familles, on obtient directement des haplotypes en examinant leur transmission en bloc de parents à enfants. Pour passer du phénotype à l'haplotype en l'absence de données familiales, on procède à des estimations statistiques de fréquences des haplotypes dans les populations. Dans l'étude "Provinces françaises", nous avons choisi de travailler sur des familles pour avoir accès directement aux données haplotypiques et pour étudier des interactions éventuelles entre systèmes génétiques.

Carte synthétique de la diversité génétique en France

Et maintenant, résultats et synthèse du projet "Provinces françaises", avec une carte de la diversité génétique en France ! Soulignons que ce projet, mené il y a 20 ans, fut un projet d'envergure nationale, mobilisant beaucoup de chercheurs et de compétences. Notons qu'aujourd'hui les études génétiques de populations au niveau de pays entiers reviennent sur le devant de la scène avec des projets beaucoup plus ambitieux, grâce à l'outil moléculaire et avec des échantillonnages plus conséquents (50 000 personnes au Québec, près de 300 000 en Islande, 500 000 en Angleterre, 1 million en Estonie, pour des projets actuellement en cours).

Prenons un allèle ou une famille d'allèles HLA...

Par exemple HLA-DR4, un marqueur associé à plusieurs maladies auto-immunes, autrement dit plus fréquemment détecté dans une population de patients que dans une population témoin lors d'une étude "cas-contrôles".

Pour DR4, on observe un gradient décroissant Ouest/Est et Nord/Sud avec un maximum dans le Finistère et un minimum en Béarn, en Pays basque et en Corse. Ainsi, à partir de notre étude sérologique et cellulaire, on constatait une variabilité des fréquences alléliques selon les régions au niveau d'un même pays, variabilité qui fut retrouvée et confirmée par les études au niveau de l'ADN plus tard.

C'est une connaissance essentielle, car observer une variabilité et des différences entre régions permet d'attirer l'attention sur l'importance qu'il y a à vérifier l'origine géographique des individus inclus dans une étude génétique: par exemple lors d'une étude "cas-contrôle", il est primordial de connaître l'origine des malades et des témoins, afin d'éviter un biais au niveau

Haplotype (2 triplets): HLA-A, HLA-B, Microsatellite (1, 44, 3) (2, 44, 4)

des résultats et conclusions de l'étude. C'est important, par exemple, dans des études immunogénétiques de la polyarthrite rhumatoïde, maladie autoimmune associée justement avec certains allèles HLA-DR de la "famille" DR4. C'est une des premières applications pratiques de l'étude Provinces françaises.

Prenons maintenant quelques gènes et haplotypes HLA...

Premier exemple : HLA-DR2 et HLA-A2, B7, DR2.

Pour l'allèle DR2, on observe une fréquence importante dans la moitié Nord et notamment à l'Ouest de la France (Finistère) et une relative rareté dans le Sud (Pays basque, Béarn, Corse). Quand on analyse la distribution de fréquence de l'haplotype HLA-A2, B7, DR2 (un haplotype fréquent en France), il y a toujours une fréquence élevée dans la moitié Nord mais une zone de fréquence maximale apparaît dans le Sud-Est. On voit ici l'intérêt des analyses multi-locus (haplotypes) par rapport aux analyses d'un seul locus, par l'information supplémentaire qu'elles apportent. Cette combinaison HLA est plus fréquemment retrouvée chez les malades atteints d'une autre maladie autoimmune touchant le système neurologique, la sclérose en plaques.

Autre exemple : HLA-DR3 et HLA-A1, B8, DR3.

Notons que c'est l'haplotype le plus fréquent en France et en Europe. La différence de distribution de l'allèle DR3 et de l'haplotype indiqué est frappante : Pour DR3 seul, le Sud-Ouest, Béarn et Pays basque en particulier et le nord de la France sont les zones de fréquence maximale, et pour l'haplotype HLA-A1, B8, DR3, c'est le Nord-Ouest qui est la zone de prépondérance, avec une quasi-absence dans le Sud-Ouest... Explication : Dans le Sud-Ouest, HLA-DR3 est porté le plus souvent par un autre haplotype, un haplotype particulier rencontré fréquemment en Pays basque, Aquitaine, Sardaigne et Espagne. A travers cet exemple, on voit comment apparaissent des profils de populations en Europe.

Ou encore HLA-DR7 et HLA-A29, B44, DR7.

Un exemple qui montre un parallélisme entre les cartes de fréquence de l'allèle DR7 et de l'haplotype, ce qui signifie que DR7 est porté essentiellement sur cette combinaison.

Sur la base de ces exemples, on perçoit que les données de fréquences alléliques et de fréquences haplotypiques sont des informations complémentaires et que l'interprétation de ces données est informative en terme de génétique de populations et pour d'autres applications.

A travers ces quelques exemples, se révèle la complexité des analyses de données de fréquences de marqueurs génétiques, système par système. Une analyse synthétique nécessite le recours à des méthodes d'analyse globale, sur l'ensemble des données. En voici quelques exemples : distances génétiques globales, analyse multifactorielle, analyse en composantes principales ou construction de figures correspondant à des arbres phylogénétiques (approche utilisée pour l'étude des phylogénies entre espèces).

Voilà le résultat d'une analyse en composantes principales sur les données de la population française: Nord et Sud différenciés, avec la nette séparation des régions du Nord, Bretagne, Normandie et Flandre d'une part et des régions du Sud, Béarn, Catalogne d'autre part ou encore Cévennes et Auvergne qui s'individualisent par rapport à Dauphiné et Bourgogne et une nette différence en Alsace et Lorraine.

Grâce à des analyses globales, une géographie génétique est accessible.

C'est pour synthétiser l'information sur la variabilité génétique en France, que l'ensemble des données de l'étude Provinces françaises a été analysé par un généticien de Turin, Alberto Piazza.

Voici maintenant les résultats de l'analyse en composantes principales, sous forme de cartes de géographie génétique, permettant de distinguer les zones présentant des différences significatives pour les distributions de fréquence des marqueurs génétiques des systèmes étudiés. La variabilité commune à plusieurs gènes est décomposée et illustrée en cartes différentes. Une partie considérable de la variabilité génétique mise en évidence a pu être résumée sur 3 cartes, dont chacune synthétise une géographie génétique de la France différente et indépendante des autres. A Piazza et les auteurs de l'étude les ont interprétées comme la représentation de l'influence des migrations de populations qui ont peu à peu constitué la structure génétique de notre pays. Ces trois cartes sont publiées en version noir et blanc (11), la quatrième synthétisant les informations de ces trois cartes étant publiée ci-après (Figure 2).

Sur une première carte, on observe un gradient Nord/Sud (ou Sud/Nord), révélant de plus l'existence de différences entre l'Alsace et la Lorraine et montrant la spécificité du Pays basque et du Béarn. Il faut savoir que cette carte représente 20% de la variabilité génétique totale dans notre étude. On peut la considérer comme la carte de la composante méditerranéenne, comme si elle décrivait un centre de diffusion d'anciennes populations Ibero-Liguriennes au sud-est et un analogue centre de diffusion au Sud-Ouest, siège des Paleo-basques. Dès 500 avant J.C. plusieurs grandes régions se sont individualisées : outre l'aire méditerranéenne économique et culturelle et la région celtique couvrant le bassin

parisien et l'est, on voit s'établir une complémentarité agro-pastorale entre Pyrénées et Aquitaine au sud de la Garonne, tandis que le climat et les différences dans l'origine du peuplement se traduisent par un contraste croissant entre les zones nord-nord-est et sud-sud-ouest du massif central. La pénétration méditerranéenne vers le nord trouve son équilibre avec l'établissement des populations originaires de l'Europe centrale. Selon la deuxième carte, qui synthétise de façon indépendante, environ 15% de l'information présente dans les données, c'est un gradient inverse que l'on met en évidence. Si on essaye d'interpréter la signification de ces observations en terme de peuplement, ce sont les vagues successives de diffusion depuis l'Europe centrale qui se trouvent évoquées : premières installations de populations du Danube au 3^{ème} et 4^{ème} millénaire avant notre ère, établissement des celtes en Gaule au cours du demi-millénaire précédant l'ère chrétienne, enfin prospection d'un vaste territoire vers l'est sur lequel les Francs ont étendu leur royaume pendant le Haut Moyen-Age. Quant à la troisième carte qui correspond à 12% de la variabilité totale, c'est un gradient est/ouest qu'on y observe. On peut interpréter cette carte comme résultant des installations de populations de l'Aquitaine à la Bretagne : les populations qui habitaient entre la Loire et la Garonne ont été si différentes de celles de l'intérieur qu'elles ont gardé leurs caractéristiques sans être complètement assimilées même pendant la domination des gaulois.

A travers ces trois cartes, c'est environ 50% de la variabilité génétique de l'étude qui est résumée. Et d'ores et déjà, on retrouve, grâce à l'étude de quelques marqueurs sur ces populations d'origine locale, les grandes tendances de l'histoire du peuplement de la France.

Quant à la superposition informatique des trois cartes, elle aboutit à une carte complexe, qui se révèle comparable à la carte des dialectes et langues en France (Figure 2). Vérifié dans quelques pays par généticiens et linguistes, ce parallélisme entre carte génétique et carte linguistique montre qu'il y a une grande cohérence entre la géographie génétique, abordée via les marqueurs génétiques et la géographie linguistique. En terme de première conclusion, on peut considérer que les gènes qui sont transmis au fil des générations sont aussi des documents historiques dont la mémoire est très longue. Grâce à l'étude Provinces françaises, on a acquis une certaine idée de la géographie génétique de la France et des idées précises sur quelques régions en particulier.

C'est à partir de cette étude et de la démonstration de la possibilité de synthétiser de nombreuses informations à partir d'études génétiques, que nous allons aborder l'implication des études de populations pour les études de maladies.

Génétique et maladie: Exemple du Diabète insulino-dépendant ou de type 1

C'est à travers l'exemple du diabète insulino-dépendant ou diabète de type I que j'illustrerai ce thème "populations et maladies".

Tout d'abord, quelques mots sur le diabète.

Le diabète insulino-dépendant est une maladie multifactorielle, impliquant, dans son déterminisme et son développement, des facteurs génétiques et des facteurs d'environnement. Depuis les années 1970, on sait que ce diabète est une pathologie associée au système HLA, via un mécanisme impliquant un dérèglement auto-immunitaire sous la dépendance du système HLA. Parmi les facteurs génétiques, des allèles HLA de susceptibilité ou de protection et des haplotypes en déséquilibre de liaison ont été mis en évidence, au fil des études, des années et de l'évolution des techniques. Les facteurs génétiques de la région HLA représentent 40% de la composante génétique du diabète de type 1. A la fin des années 1980-1990, de nombreuses incompréhensions et interrogations persistaient: pourquoi observait-on une variabilité des associations selon les populations ? Un allèle ou un haplotype associé à la maladie dans une population ne l'était pas dans une autre. C'est alors que deux facteurs ont permis d'avancer : d'une part, le progrès technologique avec la connaissance moléculaire précise permettant de révéler des différences jusqu'alors invisibles et d'autre part, le recours à la génétique des populations.

Pour illustrer l'apport combiné de la génétique moléculaire et de l'étude de plusieurs populations à l'élucidation de telles interrogations on peut tout d'abord évoquer l'étude des séquences HLA-DR et DQ dans différentes populations. Celles-ci ont révélé 1) que l'association la plus forte de cette forme de diabète était avec des marqueurs des gènes HLA-DQ plutôt que HLA-DR alors que l'on avait jusque là surtout étudié DR et dans des populations d'origine européenne; 2) que les déséquilibres de liaison entre DR et DQ variaient selon les populations ce qui expliquait plusieurs des associations apparemment différentes entre populations lorsqu'on étudiait DR ; un même allèle HLA-DQ était associé à des allèles DR différents selon les populations et vice-versa; 3) que l'ensemble des allèles DQ très associés positivement au diabète insulino-dépendant portaient une caractéristique de séquence commune, localisée sur un site particulier (correspondant au codon de l'acide aminé en position 57 de la chaîne HLA-DQ β) ; 4) que c'étaient en fait des combinaisons HLA-DR, DQ qui étaient importantes, que celles-ci soient sur un même chromosome (cis) ou portées en partie sur l'un, en partie sur l'autre chromosome (trans). Ainsi finesse d'analyse moléculaire et analyse dans diverses populations aidaient à progresser de façon complémentaire (12).

Puisqu'on parle de populations et de géographie, parlons de l'incidence du diabète de type 1, en examinant une carte de son incidence en Europe (13). L'incidence est le nombre de nouveaux cas apparaissant par an, sur un territoire donné et dans une tranche d'âge donnée.

Le diabète de type 1 en Europe ? Un gradient nord/sud ! Au Nord, les pays où c'est une maladie fréquente et devenant un véritable problème de santé publique. Au Sud, les pays où il est relativement rare, comme au Portugal ou en Italie. Son incidence est également basse en France. Une anomalie au milieu de ce gradient : la Sardaigne, avec une très forte incidence comparable à celle des pays du Nord ! Pour essayer d'expliquer ce fait, quelques éléments de génétique des populations. On sait, grâce à des études cas-témoins que l'haplotype (HLA-A30, B18, DR3, DQA1*0501, DQB1*0201) est associé au diabète insulino-dépendant, en particulier dans le Sud de l'Europe. On sait aussi que les zones de fréquence maximale de cet haplotype sont la Sardaigne tout d'abord, et le Pays basque ensuite. Rappelons que dans l'étude Provinces françaises, on a montré que la distribution génétique en Béarn et en Pays basque, régions géographiquement proches, était similaire. Les distributions de fréquences géniques sont des phénomènes qui évoluent lentement par rapport aux évolutions culturelles et sociales. Prenons, l'exemple du Pays basque et du Béarn : la distribution génétique trouvée chez les Basques aujourd'hui recouvre la zone où on parlait basque au VI^e siècle et se révèle caractéristique, notamment par la fréquence de l'haplotype (HLA-A30, B18, DR3). Si on récapitule: le diabète insulino-dépendant est fréquent au Nord de l'Europe et en Sardaigne et est associé à l'haplotype HLA-A30, B18, DR3; A noter encore: alors que le diabète de type I est associé essentiellement aux marqueurs HLA de classe II, DR, DQA et DQB, d'autres haplotypes porteurs de la combinaison classe II (DR3, DQA1*0501, DQB1*0201) identique à celle portée par l'haplotype sus-cité porteur aussi de HLA-B18, sont moins fortement associés au diabète type I.

Face à ces données, beaucoup de questions auxquelles on a essayé de répondre, notamment par les travaux de mon équipe.

Le diabète est-il plus fréquent dans le pays basque qu'ailleurs en France ?

Quels sont les haplotypes HLA associés au diabète chez les basques et chez les sardes ?

Ces haplotypes sont-ils porteurs de facteurs de susceptibilité au diabète, absents sur les autres haplotypes porteurs de DR3 ?

Autant de questions difficiles à résoudre dans la région HLA ! A cause du phénomène de déséquilibre de liaison, il existe un "effet d'entraînement" au sein d'un haplotype HLA

résultant en une difficulté à séparer les effets des différents locus HLA-A, B, DR, DQ ou autre.

Par exemple, dans le cas de HLA-A1, B8, DR3, l'association observée avec la maladie est-elle due à DR3, A1 ou B8 ?

D'où notre problématique majeure : caractériser plus finement cet haplotype associé au diabète. Une problématique sous-tendue par les questions suivantes : Comment mettre en évidence d'autres facteurs génétiques sur cet haplotype ? Quels marqueurs utiliser pour définir cet haplotype ? Existe-il des différences au niveau de cet haplotype, entre les Basques, les Sardes, les patients, les témoins ?

Pour répondre à ces questions, une idée : utiliser, en supplément des marqueurs HLA, les marqueurs microsatellites.

Petite définition des microsatellites, peut-être connus des habitués des conférences de génétique et... sans aucun rapport avec l'espace !

Les microsatellites sont des motifs d'ADN sous forme de petites séquences, répétées dans le génome. Ils ont une distribution ubiquitaire, avec une fréquence d'environ 1 microsatellite par 10 kbases ; il en existe dans le système HLA comme ailleurs dans le génome.

Comme caractéristique essentielle de ces microsatellites, on notera leur polymorphisme, ou nombre de répétitions, variable d'un individu à un autre: au même endroit du génome, 10, 15 ou 20 répétitions... selon la personne. Il s'agit d'un polymorphisme de longueur, dû à un nombre variable de répétitions. Autre caractéristique : un taux de mutation élevé par rapport à celui des gènes, d'où l'intérêt *a priori* de microsatellites dans la région HLA, qui mutant plus rapidement, étaient censés en théorie pouvoir échapper au phénomène de déséquilibre de liaison.

Notre hypothèse était que grâce aux microsatellites, on arriverait à révéler des facteurs de susceptibilité aux maladies situés dans la région HLA, mais indétectables par typage des gènes HLA seuls à cause du déséquilibre de liaison entre eux !

C'est alors que fut entreprise, dans notre laboratoire et ailleurs, une série d'études de microsatellites chez les basques, chez les sardes, chez des malades et chez des témoins sains.

Résultat de ces études : quelques réponses... et encore des questions (14, 15, 16, 17) !

Première information : il n'y a pas plus de diabète dans le Pays basque qu'ailleurs en France.

Deuxième information: on observe une forte association entre la maladie et l'haplotype d'intérêt HLA-A30, B18, DR3, un haplotype dont la composition est identique chez Basques, Sardes, qu'il s'agisse de patients ou de témoins : on a donc acquis des informations sur

d'autres marqueurs de ce même haplotype, mais ces marqueurs même s'il s'agit de microsatellites sont également très fortement soumis au phénomène de déséquilibre de liaison, donc de peu d'utilité pour résoudre le problème posé.

D'où vient la fréquence de l'incidence du diabète en Sardaigne ? Elle ne peut être liée seulement à la forte fréquence de l'haplotype HLA-A30, B18, DR3 : sachant que la distribution des gènes et les fréquences géniques sont des phénomènes à évolution lente, que l'augmentation de l'incidence du diabète est récente et que l'haplotype est identique entre patients et témoins, un ou plusieurs autres facteurs entrent probablement en jeu.

Résultat pour nous, chercheurs : l'acquisition de quelques connaissances mais aussi une grande déception... Côté positif, on enregistrait la caractérisation très approfondie d'un haplotype d'intérêt, on disposait de marqueurs variés et nombreux sur les haplotypes, on avait utilisé des informations issues de la génétique des populations pour vérifier une association entre profil génétique et maladie, en utilisant les propriétés de populations particulières, présentant un haplotype globalement rare par ailleurs en Europe. Côté négatif, c'était la désillusion face aux espoirs déçus de l'utilisation de microsatellites pour disséquer les facteurs génétiques de la région HLA, en constatant que les microsatellites étaient soumis aussi au déséquilibre de liaison...

En résumé, des études prometteuses mais guère d'avancées par rapport aux maladies. Ceci a changé encore depuis puisqu'en construisant les études différemment pour arriver à "neutraliser" dans les analyses statistiques l'effet dû à HLA-DR, DQ et en utilisant une série plus étendue de microsatellites il semble bien qu'on soit maintenant arrivé à mettre en évidence grâce à l'utilisation de tels marqueurs de nouvelles zones de la région HLA qui interviennent dans la susceptibilité génétique au diabète de type 1 (18). Il y a souvent un long temps qui s'écoule entre la survenue d'une idée, l'émission d'une hypothèse et l'obtention du résultat permettant de la valider. La recherche procède par paliers, avec bien des détours. Un résultat négatif ou décevant peut être à l'origine de nouvelles orientations. C'est exactement ce qui s'est passé avec nos résultats sur les microsatellites et c'est un nouveau développement qui fera l'objet de ma troisième partie.

Génétique et transplantation : exemple greffe moelle osseuse

Genèse d'une idée

Pour aller plus loin, grâce à l'imagination fertile des chercheurs et devant l'intérêt potentiel des microsatellites, une autre idée s'est imposée : Pourquoi ne pas utiliser les microsatellites dans une optique inverse de ce que nous avons envisagé ? L'idée est simple : Typer des

microsatellites de la région HLA, en exploitant leur facilité de typage et leur propriétés de déséquilibre de liaison avec HLA afin d'évaluer dans quelle mesure il est possible de prévoir un groupage HLA à partir d'un typage de microsatellites.

Et pourquoi pas, envisager des applications pour les transplantations ?

Quel est l'intérêt de cette démarche ? Se servir de ces connaissances pour optimiser le choix des personnes à typer HLA dans certains domaines de la transplantation, générant des économies ou du moins une gestion des ressources optimisée. Nous allons voir comment.

A nouveau, un projet a démarré, sur une étude des microsatellites de la région HLA, avec comme objectifs la détermination des relations entre microsatellites et gènes HLA et la possibilité d'une prédiction HLA par un typage microsatellite.

Dans le cadre de cette étude menée au sein de mon laboratoire, un panel de six microsatellites a été choisi, couvrant la région HLA, autour de HLA-A, B et DR, trois gènes essentiels au niveau immunologique et bien connus de la région. Au niveau des échantillons, le travail s'est effectué sur des individus préalablement typés pour HLA et que nous avons typés pour ces microsatellites.

Synthèse des résultats : un schéma (19, 20, 21) qui résume l'information acquise par un typage de microsatellites de la région HLA sur les gènes et les haplotypes HLA (Figure 3). La quantification de cette information a mis en jeu un calcul d'entropie.

Commentaire et interprétation :

Sur la partie supérieure du schéma, premier exemple : par typage du microsatellite MIB seul, situé à proximité du gène HLA-B, on obtient 40 à 60% de l'information sur HLA-B.

Deuxième exemple : par typage d'un seul microsatellite, on acquiert une information importante (jusqu'à 40%) de l'information sur une combinaison HLA-A, B, DR.

Sur la partie inférieure du schéma, résultat encore plus intéressant : en prenant en compte à la fois les résultats des typages de 3 microsatellites, c'est plus de 80% d'information sur les gènes HLA qui est obtenue et plus de 90% sur les haplotypes HLA-A, B, DR.

A partir de la frustration et déception initiales de l'étude sur les maladies et grâce à des propriétés déterminées par des études de population, notre équipe, en retournant l'idée de départ, avait découvert des éléments très intéressants, et très prometteurs en applications et utilisations potentielles. Ces études sont à la base d'un projet européen qui est en cours actuellement (www.euromado.org)

D'un échec à un programme européen dynamique

Cette idée de prédiction d'haplotypes HLA à partir de typages de marqueurs microsatellites a trouvé un domaine d'application possible dans le contexte des greffes de moelle osseuse. C'est un aspect que j'aborderai par une description de l'organisation des greffes et par la notion de diversité HLA dans le registre des donneurs volontaires de moelle osseuse en France, qui nous amèneront ensuite à comprendre l'intérêt potentiel des histoires de microsatellites dans le contexte des greffes.

Commençons par le *contexte et le fonctionnement de la filière de la transplantation de cellules souches hématopoïétiques* correspondant le plus souvent à des greffes de moelle osseuse (22, 23).

Il faut noter tout d'abord, que la greffe de moelle osseuse est une issue thérapeutique pour de nombreuses pathologies hématologiques. Prenons l'exemple d'un malade atteint de leucémie, qui a besoin d'une greffe de moelle osseuse. Quel est le donneur idéal ? Son frère ou sa soeur HLA identique. Donneur idéal n'existant que rarement dans une famille. Il y a 1 chance sur 4 d'être HLA identique au sein d'une fratrie. La compatibilité HLA en transplantation et en particulier en greffe de moelle osseuse, constitue un élément essentiel de la survie de la greffe et des patients, notamment à cause de la "réaction du greffon contre l'hôte". Lors d'une transplantation de cellules souches hématopoïétiques, il existe deux possibilités de rejet, car on greffe des cellules immuno-compétentes, qui deviendront les cellules du sang du patient. D'une part le receveur peut rejeter la greffe ou celle-ci ne pas "prendre" et d'autre part, le greffon, formé de cellules ayant une capacité immunologique, peut reconnaître le receveur comme étranger et causer des phénomènes pathologiques graves en attaquant les tissus du receveur. C'est pourquoi la compatibilité HLA est primordiale en greffe de moelle osseuse et nécessite une recherche de donneur compatible, dans la famille ou ailleurs.

Ailleurs ? Dans les fichiers de donneurs volontaires de moelle osseuse (et aujourd'hui aussi dans les banques de sang placentaire qui constituent aussi de bonnes sources de cellules souches hématopoïétiques). Apparus en 1974 en Angleterre et généralisés depuis les années 1980 dans les pays développés pratiquant les greffes de moelle osseuse, les registres de donneurs constituent un panel d'individus, prêts à devenir donneurs, en faisant don de leurs cellules si besoin, pour un patient en attente de greffe. Plus concrètement, c'est un registre de données HLA, les donneurs potentiels étant soumis à un groupage HLA lors de leur inscription.

Phases d'une procédure de greffe. En l'absence de donneur familial, il y a interrogation des registres de donneurs sur la base du groupage HLA du patient, en France et dans le monde, pour rechercher un donneur compatible. Suivent alors, si un donneur est identifié, une convocation du donneur, une opération de vérification de typage, le recueil de son consentement, le prélèvement de moelle osseuse et enfin la transplantation.

Transplantation et génétique des populations : quel rapport ?

Statistiquement, posséder un groupage HLA fréquent offre une bonne chance de trouver un donneur, dans la famille ou au sein d'un registre, alors que pour un individu à groupe rare, cela sera plus difficile, voire aléatoire. De fait, s'instaure une inégalité d'accès à la thérapeutique de greffe de moelle osseuse, fondée sur la diversité génétique. La connaissance de cette diversité et de sa répartition peut-elle aider à mettre en place des stratégies de constitution des registres de donneurs mieux à même de répondre à la demande des patients ? C'est ainsi que l'idée d'étudier la diversité génétique au sein du registre de donneurs potentiels de cellules souches hématopoïétiques en France est née, inspirée de l'étude *Provinces françaises*, précédemment évoquée.

Qu'a révélé cette étude de la diversité génétique dans le fichier français de donneurs volontaires de moelle osseuse, menée en collaboration avec l'Association France Greffe de Moelle (<http://www.fgm.fr>) qui gère le fichier de donneurs en France (24) ?

La distribution des groupes HLA au sein du registre français de donneurs permet de mettre en évidence en France 5 groupes de régions, correspondant à des regroupements au sein des 20 régions administratives, sur la base de la structure en haplotypes HLA. Au sein de ces régions, il y a plus ou moins de diversité ou variabilité HLA.

Dès lors, apparaissent des possibilités de stratégies pour organiser les campagnes de sensibilisation de la population, selon les zones, en ciblant les régions où l'on observe une grande diversité HLA. Notons qu'il est intéressant de retrouver cette caractéristique d'une diversification extrême de la distribution HLA au sein de la France, à la fois dans une étude sur le registre des donneurs de moelle, menée sur des individus sans indication d'origine géographique, et dans une étude sur des familles d'origine géographique ciblée, telle que *Provinces Françaises*.

Et encore une fois, on constate que la génétique des populations débouche vers une application en santé publique...

Quelques mots maintenant sur le registre français de donneurs de moelle.

Créé en 1987, il répond au fait que, en France, 70% des indications de greffe de moelle osseuse nécessitent une recherche de donneurs hors de la famille, faute d'un donneur HLA compatible familial. Pour donner une idée des chiffres, voici quelques données concrètes.

Aujourd'hui, il y a plus de 110 000 donneurs dans le fichier français, et au niveau mondial, plus de 8 millions de donneurs sont enregistrés (<http://www.fgm.fr>). En 2001, en France, 856 patients étaient en attente de greffe, parmi lesquels 234 ont été greffés, dont 57 à partir de donneurs du registre français. Simple calcul et triste constatation : un trop grand écart entre les 234 patients greffés et les 856 nécessitant une greffe. A partir de ce constat, un nouveau projet a été élaboré, en réponse à une question : Comment augmenter la diversité génétique HLA dans les registres de donneurs de moelle osseuse ?

Première stratégie qui vient à l'esprit: accroître la taille des registres, en augmentant le nombre de donneurs. Mais il y a deux problèmes : une limitation financière due au coût des groupages HLA et une limitation génétique créant un plafonnement ou seuil en terme de variabilité à cause des différences de fréquence des combinaisons HLA dans la population. Au delà d'une certaine taille de registre il faut ajouter de très nombreux donneurs pour augmenter à peine la diversité génétique du registre... si l'on recrute les donneurs au hasard. En résumé, beaucoup d'argent et beaucoup de personnes pour à peine un petit peu plus de diversité.

Pour contourner ces problèmes, il faut rechercher un moyen de distordre le recrutement pour le registre, pour favoriser l'arrivée de nouveaux phénotypes HLA et réduire la répétition des groupages fréquents, tels que HLA-A1, B8, DR3 par exemple.

C'est ici qu'apparaît l'idée de filtre moléculaire en première étape de sélection de donneur pour le registre et c'est là que reviennent les microsatellites que nous paraissions avoir oublié. Selon nos premières expériences, le principe d'une prédiction de groupage HLA par un typage microsatellite est applicable. Pour optimiser la constitution du fichier, l'idée est d'exploiter cette propriété, présentant un avantage en terme de coût et permettant un choix de priorité de typage HLA au niveau des nouveaux donneurs potentiels selon qu'ils apportent ou non une nouvelle diversité au fichier. L'idée serait donc de repérer par un premier typage de marqueurs microsatellites si un nouveau donneur apportera ou non un phénotype HLA nécessaire pour optimiser le registre. Si cette stratégie a été étudiée dans un premier temps avec des microsatellites, d'autres méthodes moléculaires sont potentiellement intéressantes, avec d'autres marqueurs tels que les polymorphismes de nucléotide unique (abréviation SNP en anglais : *Single Nucleotide Polymorphism*), présents dans la région HLA. Dès aujourd'hui, des possibilités apparaissent pour imaginer des stratégies, basées sur des technologies moléculaires et incrémentées d'études médico-économiques, avec de très intéressantes

perspectives en santé publique. C'est aussi vers une adaptation de la politique de sensibilisation que s'orientent ces stratégies, en permettant de varier ces informations selon les populations et de se concentrer sur les populations susceptibles d'optimiser la diversité génétique du registre, lors des campagnes d'information et d'incitation au don, l'ensemble de la population n'étant pas forcément intéressée ou même informée.

Peu à peu, à partir de cette idée de filtre moléculaire comme moyen d'optimisation des programmes de typage et des stratégies de recrutement de donneurs, s'est construit un programme européen appelé MADO, pour *Marrow Donors* ou Donneurs de Moelle (www.euromado.org).

Alors, en quelques mots, voici l'histoire du projet MADO. Suite à un avant-projet réalisé dans mon laboratoire, s'est construit ce consortium européen dans le cadre du Cinquième Programme-Cadre européen de Recherche et développement (PCRD). C'est un projet que je coordonne, et qui regroupe 14 partenaires dans 5 pays, Angleterre, Hollande, France, Italie et Hongrie.

De par son sujet et l'approche envisagée, c'est un projet multidisciplinaire, nécessitant des compétences et des points de vues différents : développement technologique pour le choix des microsattellites ou SNP, gestion et organisation des registres, immunogénétique pour l'expertise sur les groupages HLA, économie de la santé, sociologie, droit et éthique. Autant d'aspects dans ce projet qui pose des questions éthiques et socio-culturelles : comment sensibiliser des minorités peu enclines au don de cellules ou d'organes, ou non informées ? Comment refuser le geste généreux d'un donneur sous prétexte qu'il n'a pas un phénotype rare ? Comment faire une information adaptée tout en respectant les personnes ?

Au-delà des questions scientifiques ou techniques, au-delà de la génétique des populations, ce programme interpelle sur des sujets sociétaux qui m'intéressent particulièrement.

Résumons ce projet en quelques mots: c'est un programme qui se propose d'optimiser les registres des donneurs de cellules souches hématopoïétiques, en augmentant l'adéquation immunogénétique entre les besoins des patients et les donneurs potentiels.

Pour ce projet et pour cet objectif, une stratégie a été élaborée : augmenter la proportion des donneurs ayant un phénotype rare ou peu fréquent dans les registres; un concept ou outil a été proposé : un filtre moléculaire, et une analyse sociologique vient compléter la stratégie du filtre moléculaire par l'étude de la population effectivement touchée par l'information et répondant en devenant "donneur potentiel". Cette dernière étude devrait fournir des

informations utiles pour organiser les campagnes de sensibilisation et de recrutement de donneurs.

Aujourd'hui, après information de la population par divers média quant aux possibilités de devenir donneur potentiel, on procède à un examen et à un typage HLA des volontaires qui se présentent signent un engagement volontaire. Cela aboutit à un registre interrogé lors d'une recherche de donneur pour les patients. Demain, peut-être, l'approche sera en deux temps, associée à l'utilisation du filtre moléculaire et à la notion de choix raisonné de typage : un fichier constitué de deux registres, un registre "Niveau I" basé sur une approche peut-être probabilistique de typage HLA, via le filtre moléculaire, et un registre optimisé basé sur un typage HLA haute définition, portant sur certains donneurs potentiels seulement selon la définition de priorités en terme de diversité génétique probable apportée et de besoins des patients.

Par ce projet, on perçoit en quoi le fait d'intégrer ce "modèle de filtre" pour instaurer des priorités de typage, implique à la fois une comparaison de l'organisation des registres en Europe, une évaluation des techniques, un éclairage sur les modèles sociologiques sous tendant les démarches de don et des enjeux légaux et éthiques.

Conclusion

Voilà un peu l'histoire de ce projet et l'histoire de l'articulation de ces recherches, depuis 1980 jusqu'à aujourd'hui. J'espère avoir illustré comment, à travers des projets divers, un ensemble de progrès dans la connaissance d'un système donné au niveau moléculaire et un appui réitéré sur la génétique des populations m'a permis d'aboutir à ces préoccupations de santé et de santé publique, qui constituent un retour vers ce à quoi me préparait ma formation de médecin.

Je conclurai mon propos par un hommage et une évocation du Professeur Mourant, un des pionniers de la génétique des populations des groupes sanguins et du Professeur Dausset, bien connu pour sa découverte des groupages HLA et prix Nobel 1980, qui s'étaient trouvés réunis à Toulouse lors d'un colloque sur l'étude *Provinces françaises* (3) et j'évoquerai aussi de grands noms français en génétique des populations humaines, en immunogénétique et en hématologie qu'il m'a été de rencontrer sur la route de mes propres recherches dès avant les années 1980 comme les professeurs Jacques Ruffié, Jean Bernard, Laurent Degos, André Langaney, Elie Ohayon et Jean Ducos.

REFERENCES:

1. Cavalli-Sforza LL, Menozzi P, Piazza A., *The history and geography of human genes*, Princeton, Princeton University Press, 1993, 518 p
2. Feingold J, Fellous M. Solignac M., *Principes de génétique humaine*, Paris, Herman Editeur des sciences et des arts, 1998, 586 p.
3. Ohayon E., Cambon-Thomsen A. Eds, « Génétique des populations humaines », *Colloque INSERM*, vol. 142, 1986, 407 p
4. Cambon-Thomsen A. « Diversité de la population française », *Pathologie Biologie*, 2001, 49, n°5, 407-409
5. Cambon-Thomsen A., Borot N., Neugebauer M., Sevin A., Ohayon E., and the participants to the genetic survey « Provinces Françaises ». Inter-regional variability between 15 French provinces and Quebec, *Collegium Antropologicum*, 1989, 13, 25-41
6. Cambon-Thomsen A., Sevin A., Sala V., Borot N., Ohayon E., avec la collaboration de l'ensemble des participants à l'étude « Provinces Françaises ». Caractéristiques de l'échantillonnage dans l'étude des marqueurs génétiques dans les Provinces Françaises (PF); référence particulière aux marqueurs HLA. In : *Approche multidisciplinaire des isolats humains*, Paris, INED, 1990 (Congrès et Colloques n° 3), pp. 184-219
7. Klein J., Sato A., « The HLA System— First of Two Parts *Advances in Immunology : N* », *Engl J Med*, 2000, 343:702-709, Sep 7, 2000.
8. Klein J., Sato A., « The HLA System— Second of Two Parts *Advances in Immunology : N* », *Engl J Med*, 2000, 343:782-786, Sep 14, 2000.
9. Mehra N K & Kaur G, « MHC-based vaccination approaches : progress and perspectives », *Exp. Rev. Mol. Med.*, Vol. 5, 24 February, <http://www.expertreviews.org/03005957h.htm>
10. MHC-based vaccination approaches: progress and perspectives. Expert reviews in molecular medicine. <http://www.expertreviews.org>
11. Feingold, J., « Le déséquilibre de liaison », *Médecine/Sciences* 7, 1991, pp.161-168
12. France. In Cavalli-Sforza LL, Menozzi P, Piazza A., « The History and Geography of Human Genes. Princeton », *Princeton University Press*, 1993, Chap. 5.8, pp 280-285
13. Caillat-Zucman S, Bach JF., « Genetic predisposition to IDDM », *Clin Rev Allergy Immunol.*, 2000 Dec;19(3):227-46.
14. Patterson CC, Dahlquist G, Soltesz G, Green A, « EURODIAB ACE Study Group. Europe and Diabetes », *Diabetologia*, 2001 Oct, 44 Suppl 3 : B9-16. Is childhood-onset type I diabetes a wealth-related disease? An ecological analysis of European incidence rates.
15. Levy-Marchal C., « Evolution de l'incidence du diabète insulno-dépendant chez l'enfant en France », *Rev Epidemiol Santé Publique*, 1998 Jun, 46(3):157-63.
16. Cambon-De Mouzon A., Ohayon E., Hauptmann G., Sevin A., Abbal M., Sommer E., Vergnes H., Ducos J., « HLA-A, B, C, DR antigens, Bf, C4 and glyoxalase I (GLO) polymorphisms in French Basques with insulin-dependent diabetes mellitus (IDDM) », *Tissue Antigens*, 1982, 19, 366-379

17. Briant L., Jongeneel C. V., Sommer E., Sevin A., Abbal M., Ohayon E., Cambon-Thomsen A., « HLA extended haplotypes including three TNF polymorphic microsatellites in normal caucasoids and in IDDM families », In *HLA 1991*, Eds. K. Tsuji, M. Aizawa, T. Sasazuki, Vol. II Oxford University Press, 1992, pp 492-495
18. Crouau-Roy B., Bouzekri N., Carcassi C., Clayton J., Contu C., Cambon-Thomsen A. « Strong association between microsatellites and an HLA-B, DR haplotype (B18, DR3): implication for the evolution of the microsatellites », *Immunogenetics*, 1996, 43, 255-260
19. Johansson S., Lie B. A., Todd J. A. , Pociot F., Nerup J., Cambon-Thomsen A., Kockum I., Akselsen H. E., Thorsby E., Undlien D. E., « Evidence of at least two type 1 diabetes susceptibility genes in the HLA complex distinct from HLA-DQB1, -DQA1 and -DRB1 », *Genes Immun*, 2003, 4(1):46-53
20. Foissac A. Analyse de microsatellites et génétique humaine. [Thèse de doctorat, génétique humaine], Université Paul Sabatier; Toulouse. 1999
21. Foissac A, Fort M, Clayton J, Abbal M, Raffoux C, Moine A, Bensa JC, Bignon JD, Mercier P Cambon-Thomsen A., « Microsatellites in the HLA region : HLA prediction and strategies for bone marrow donors registries », *Transplant. Proc.*, 2001, 33, 491-492
22. Cambon-Thomsen A, Foissac A, Fort M, Clayton J, Massonier L, Abbal M, Moine A, Bensa JC, Bignon JD, Mercier P, Raffoux C., « Microsatellites in the HLA region : Association with HLA and strategies for bone marrow donors Registries”, in *HLA 2002, Immunobiology of the human MHC*, Vol 2, Editors J Hansen & B Dupont, IHWC Press, Seattle, in press
23. Charron DJ. « HLA matching in unrelated donor bone marrow transplantation », *Curr Opin Hematol*, 1996 Nov, 3(6):416-22. Review
24. Charron D., « Immunogenomics of hematopoietic stem cell transplantation », *Transfus Clin Biol*, 2003 Jun, 10(3) :156-8.
25. Lonjou C., Clayton J., Cambon-Thomsen A., Raffoux C., « HLA-A, B, DR haplotype frequencies in France : implications for recruitment of potential bone marrow donors », *Transplantation*, 1995, 60, 375-383

LEGENDES DES FIGURES:

Figure 1: Carte simplifiée de la région HLA (d'après réf 9)

Figure 2: Carte synthétique de la variabilité génétique en France d'après l'étude des marqueurs génétiques dans les provinces françaises

Cette carte est produite par superposition de 3 cartes issues de l'analyse en composante principale menée par A Piazza (12) sur la base des données recueillies lors de l'étude "Marqueurs génétiques dans les provinces françaises" (3).

Figure 3: Information obtenue sur HLA par typage de microsatellites

01/03/2007 1:49

Schéma d'après la thèse d'A Foissac (20) et l'ouvrage (22).

En gris information obtenue sur les gènes HLA, en noir information obtenue sur les haplotypes HLA après une analyse d'entropie des données.