

Prédisposition génétique à la bilharziose chez l'Homme : Méthodes de recherche et application à l'étude de l'infection par *Schistosoma Mansoni*

par Laurent Abel¹, Sandrine Marquet², Christophe Chevillard², Nasr Eldin MA ElWali^{2,3},
Dominique Hillaire², Alain Dessein²

¹ INSERM Unité 436, « Modélisation mathématique et statistique en biologie et en médecine »,
CHU Pitié-Salpêtrière, 91 Bd de l'Hôpital, 75013 Paris.

Tél: 01 40 77 96 15. Fax: 01 45 85 15 29. Email : abel@biomath.jussieu.fr

² INSERM Unité 399, « Immunologie et génétique des maladies parasitaires », Faculté de Médecine,
Marseille, France.

³ Institut de Médecine nucléaire et de biologie moléculaire, Université de Gezira, Sudan.

Reçu le 26 novembre 1999

RÉSUMÉ

Le développement des méthodes de génétique épidémiologique qui exploite l'information considérable apportée par les récentes cartes génétiques et la disponibilité croissante de gènes candidats a conduit à des avancées majeures dans l'identification des gènes humains de prédisposition à la bilharziose. Deux phénotypes ont jusqu'à présent été étudiés dans l'infection par *Schistosoma mansoni* : les niveaux d'infection par le parasite mesurés par le nombre d'œufs dans les selles, et la fibrose hépatique sévère causée par *S. mansoni* évaluée par échographie. La première étude réalisée sur des familles brésiliennes a démontré l'existence d'un gène majeur contrôlant les niveaux d'infection par *S. mansoni*, appelé SM1 qui a été localisé dans la région chromosomique 5q31-q33. Cette région contient plusieurs gènes candidats impliqués dans la régulation de la réponse Th1/Th2,

et le rôle direct de polymorphismes situés dans ces gènes est en cours d'exploration. La seconde étude conduite au Soudan a également montré la présence d'un gène majeur influençant le développement d'une fibrose hépatique sévère lors de l'infection à *S. mansoni* appelé SM2. Ce gène n'est pas situé dans la région 5q31-q33, mais a été localisé en 6q22-q23 très proche du gène IFNGR1 codant pour le récepteur de l'Interferon- γ , une cytokine au rôle fortement anti-fibrogénique. Ces résultats indiquent que deux loci distincts contrôlent la prédisposition humaine à la bilharziose, SM1 situé en 5q31-q33 qui joue vraisemblablement un rôle dans la différenciation Th1/Th2, et SM2 en 6q22-q23 influençant la progression vers la maladie hépatique avec une implication possible dans la régulation de l'IFN- γ .

SUMMARY Genetic predisposition to human schistosomiasis: methods of investigation and application to the study of infection by *Schistosoma mansoni*

The development of genetic epidemiology methods using recent human genetic mapping information together with the growing availability of candidate genes has led to major advances in the identification of host genes in human schistosomiasis. Two phenotypes have been studied so far in the infection by *Schistosoma mansoni* : infection levels by the parasite as measured by the faecal egg counts, and the severe hepatic fibrosis caused by *S. mansoni* assessed by ultrasound examination. The first study was performed on Brazilian pedigrees and provided strong evidence for a major gene controlling infection levels by *S. mansoni* denoted as SM1 which was mapped to chromosome 5q31-q33. This region contains several candidate genes involved in the regulation of the Th1/Th2 response, and the direct role of polymor-

phisms located within these genes is under investigation. The second study conducted in Sudan also showed the presence of a major gene influencing the development of severe hepatic fibrosis due to *S. mansoni* infection denoted as SM2. This gene is not located in the 5q31-q33 region, but maps to chromosome 6q22-q23 and is closely linked to the IFN- γ RI gene encoding the receptor of the strongly anti-fibrogenic cytokine Interferon- γ . These findings indicate that two distinct genetic loci control human predisposition to schistosomiasis, SM1 located in the 5q31-q33 region which is likely to play a role in the Th1/Th2 differentiation, and SM2 in 6q22-q23 influencing disease progression with a possible involvement in the regulation of IFN- γ .

Le rôle majeur des facteurs génétiques dans la capacité d'un organisme à combattre une infection a été démontré dans les modèles expérimentaux (McLeod R. *et al.*, 1995), où il est possible de contrôler la maladie, le milieu, et les croisements entre animaux. Chez l'Homme, un rôle possible de facteurs génétiques dans l'expression des maladies infectieuses est suggéré par l'existence d'une grande variabilité individuelle de réponse à l'agent infectieux, contrastant souvent avec de fortes corrélations familiales. Les méthodes expérimentales sont bien évidemment inapplicables chez l'Homme, et c'est à partir de l'observation de ces concentrations familiales et de l'information apportée par les marqueurs génétiques que l'épidémiologie génétique cherche à déterminer la part respective des facteurs génétiques et de milieu (en particulier, les facteurs de contamination par l'agent infectieux) dans l'étiologie des ressemblances familiales observées, et à identifier la nature des facteurs génétiques en cause (Abel L. & Dessein A.J. 1997; Abel L. & Dessein A.J. 1998). Le contrôle génétique de la réponse aux maladies infectieuses comme la bilharziose peut être étudiée par l'intermédiaire de différents traits ou phénotypes. Il peut s'agir de phénotypes cliniques (atteint/sain) comme, par exemple, les atteintes hépatiques graves de bilharziose. Il peut s'agir aussi de phénotypes biologiques (quantitatifs) reflétant soit les niveaux d'infection par l'agent infectieux (nombre d'œufs de *Schistosoma mansoni* dans les selles), soit la réponse immunitaire (taux d'anticorps ou de cytokines spécifiques ou non d'un certain antigène de schistosomes). Les analyses de ces différents phénotypes sont tout à fait complémentaires permettant de déterminer par exemple si le contrôle génétique de telle réponse immunitaire permet d'expliquer le contrôle de l'expression de tel phénotype clinique.

Les méthodes d'épidémiologie génétique sont de nature statistique et permettent d'analyser ces différents phénotypes dans le but de faire la distinction entre ce qui est expliqué par des facteurs de milieu et ce qui peut être dû à des facteurs génétiques (Khoury M.J. *et al.*, 1993). Elles vont donc combiner des informations de nature épidémiologique et génétique. Les données épidémiologiques concernent le recueil de l'ensemble des facteurs de risque qui peuvent influencer la pathologie étudiée (facteurs d'exposition à l'agent infectieux, âge...). Les informations génétiques sont représentées par la connaissance des liens familiaux (recueil de données familiales) et par le typage de marqueurs génétiques. Les développements récents de la génétique moléculaire, comme l'établissement d'une carte génétique dense du génome humain basée sur des marqueurs microsatellites très polymorphes (Dib C. *et al.*, 1996) et la disponibilité croissante de polymorphismes mononucléotidiques (Single Nucleotide Polymorphism ou SNP) (Kruglyak L., 1999) ont créé des outils fondamentaux pour ces études génétiques. La méthodologie générale en génétique épidémiologique est de nature statistique, et comporte deux grands types d'études, les analyses de liaison génétique (ou analyses de linkage) et les études d'association (Abel L. & Dessein A.J. 1998). En géné-

tique humaine des maladies infectieuses, les analyses de liaison sont utilisées pour localiser une région chromosomique contenant un ou quelques gène(s) d'intérêt. Un de leurs avantages est de pouvoir explorer l'ensemble du génome par criblage complet (genome screen) permettant de détecter l'effet de gènes dont le rôle était *a priori* inconnu. Bien que ces analyses de liaison permettent d'affirmer le rôle d'une région chromosomique, elles ne permettent pas une localisation très précise des gènes influençant des phénotypes aussi complexes. L'étape suivante est alors de tester directement, par des études d'association, le rôle de polymorphismes de gènes candidats situés dans les régions identifiées. Ces études d'association peuvent également être réalisées en première intention en testant le rôle de gènes candidats par hypothèse (par exemple des gènes impliqués dans la réponse immunitaire). Dans tous les cas, le rôle d'un polymorphisme ne peut être validé que par des études fonctionnelles, soulignant la complémentarité indispensable des études de génétique épidémiologique et de génétique moléculaire.

La bilharziose, ou schistosomose, est la deuxième grande maladie parasitaire après le paludisme, affectant plus de 200 millions de personnes dans le monde. *Schistosoma mansoni* est le parasite responsable de la bilharziose hépato-intestinale, la forme la plus fréquente chez l'Homme. Le risque de contamination est d'autant plus grand que le contact avec l'eau infestée par la forme jeune du parasite est plus prolongé, et le facteur de milieu prépondérant est dans ce cas l'importance de ce contact. En collaboration avec le groupe d'Alain Dessein (Unité INSERM 399), nous recherchons depuis plusieurs années les facteurs génétiques impliqués dans la réponse à cette affection parasitaire. Nous nous sommes intéressés à l'étude de deux phénotypes, tout d'abord les niveaux d'infection par *S. mansoni*, puis le phénotype clinique correspondant au développement d'une fibrose hépatique sévère. Dans les deux cas, nous avons utilisé une stratégie d'étude dite modèle-dépendante (ou paramétrique) basée sur la succession analyse de ségrégation – analyse de liaison génétique (Abel L. & Dessein A. J. 1998) afin de localiser les régions chromosomiques d'intérêt pour ces deux phénotypes. Cette approche modèle-dépendante nécessite de spécifier le modèle décrivant la relation entre le phénotype observé (par exemple les niveaux d'infection) et les facteurs influençant le phénotype comme l'existence d'un gène majeur, les facteurs d'exposition à l'agent infectieux (niveaux de contact avec l'eau infestée), l'âge... Ce modèle est en général estimé par l'analyse de ségrégation et ensuite utilisé pour l'analyse de liaison génétique par lod-score classique. Il s'agit de l'approche la plus puissante en analyse de liaison génétique lorsque le modèle peut être correctement estimé.

La première étude sur les niveaux d'infection par *S. mansoni* s'est déroulée dans un village du nord-est brésilien où ce parasite est hyperendémique (Abel L. *et al.*, 1991). L'ensemble des familles de ce village ont été recueillies pour l'analyse. Le phénotype reflétant l'intensité de l'infection par le parasite était le nombre d'œufs de schistosomes dans les selles mesuré à plu-

sieurs reprises et avant tout traitement. Nous avons tout d'abord montré l'influence importante des niveaux d'exposition à l'eau infestée et de l'âge sur les niveaux d'infection qui expliquait environ 30 % de la variance des niveaux d'infection (Abel L. *et al.*, 1991). Néanmoins, après prise en compte de ces facteurs, persistait une grande variabilité individuelle de niveaux d'infection avec une forte agrégation familiale. L'analyse de ségrégation a alors montré que ces distributions familiales étaient expliquées par l'existence d'un gène majeur contrôlant l'intensité de l'infection par *S. mansoni* avec environ 3 % de la population prédisposée à de fortes infections (Abel L. *et al.*, 1991). Ce gène, appelé SM1, avait un effet très important puisqu'il rendait compte d'environ 40 % de la variance des niveaux d'infection. La deuxième étape de cette étude a été alors de localiser SM1 par analyse de linkage. Le typage des marqueurs génétiques couvrant l'ensemble du génome (un total de 250 marqueurs) a été effectué en collaboration avec le Généthron par Sandrine Marquet du groupe d'Alain Dessein. L'analyse de linkage a été effectuée par la méthode des lod scores en utilisant pour les paramètres du gène majeur les valeurs obtenues par l'analyse de ségrégation. Les résultats observés sur les 250 marqueurs ont montré qu'une seule région chromosomique en 5q31-q33 donnait, pour deux marqueurs adjacents, des lod-scores supérieurs à + 3 (Marquet S. *et al.*, 1999; Marquet S. *et al.*, 1996). Un typage complémentaire de marqueurs génétiques dans cette région a alors été réalisé et a permis de localiser SM1 avec des valeurs de lod-score très significatives (supérieures à 4.5) pour deux marqueurs situés à 3 centimorgans l'un de l'autre : D55636 et un marqueur situé dans le gène du récepteur de CSF1 (Colony Stimulating Factor 1) qui est un facteur stimulé par les infections intervenant dans la production des phagocytes mononucléaires (Marquet S. *et al.*, 1996). Outre le gène du récepteur de CSF1, la région 5q31-q33 identifiée par notre analyse de linkage contient plusieurs autres gènes candidats potentiels pour SM1 tels ceux codant pour l'interleukine 4 (IL-4), l'IL-5, l'IL-13, et la sous-unité p40 de l'IL-12. Il est à noter que la liaison des niveaux d'infection par *S. mansoni* dans la région 5q31-q33 a été confirmée dans une population du Sénégal (Muller-Myhsok B. *et al.*, 1997). De plus, parallèlement à cette étude génétique, une étude immunologique évaluant la réponse spécifique T auxiliaire chez les sujets de l'étude a montré que les sujets homozygotes prédisposés aux fortes infections produisent 10 à 100 fois plus d'IFN- γ que d'IL-5 et d'IL-4 alors que les sujets homozygotes résistants présentent un profil de sécrétion inverse (Rodrigues Jr V. *et al.*, 1999). Ces résultats indiquent que SM1 pourrait avoir un rôle important dans la différenciation des lymphocytes T auxiliaires et dans la balance Th1/Th2, et des études sont en cours pour tester le rôle de polymorphismes situés dans les gènes candidats de la région 5q31-q33.

L'autre phénotype d'intérêt dans la bilharziose à *S. mansoni* est le phénotype maladie correspondant au développement d'une fibrose hépatique sévère avec les

conséquences graves de l'hypertension portale. La raison pour laquelle seulement une proportion relativement faible (2 à 10 %) des sujets infectés par le parasite développe une maladie grave n'est pas connue, mais plusieurs observations suggèrent le rôle de facteurs génétiques (Mohamed-Ali Q. *et al.*, 1999). Nous avons récemment conduit une large étude sur l'ensemble des familles d'un village soudanais où le phénotype a été évalué par échographie hépatique. La prévalence des fibroses hépatiques graves avec hypertension portale était de 6 %. L'analyse de ségrégation réalisée sur les familles de ce village a mis en évidence un gène majeur contrôlant ces formes graves de fibrose hépatique (Dessein A.J. *et al.*, 1999). La fréquence de l'allèle prédisposant à la fibrose hépatique sévère était estimée à 0,16. Les homozygotes pour cet allèle ont une pénétrance estimée à 50 % après 9 (pour les hommes) ou 14 ans pour les femmes) de résidence dans la zone d'exposition; cette pénétrance de 50 % est atteinte après 19 ans de résidence pour les hétérozygotes hommes, alors que, pour tous les autres sujets, la pénétrance reste inférieure à 2 % après 20 ans d'exposition. Une analyse de liaison génétique utilisant ce modèle a été réalisée sur quatre régions génétiques (incluant la région 5q31-q33) et a montré que ce gène était localisé dans la région 6q22-q23 où est situé le gène IFNGR1 codant l'une des chaînes du récepteur de l'IFN- γ qui est une cytokine ayant un rôle anti-fibrogénique connu (Dessein A.J. *et al.*, 1999). Ces résultats suggèrent donc que des polymorphismes situés dans le gène IFNGR1, qui sont en cours d'investigation, pourraient influencer le développement d'une fibrose hépatique sévère dans l'infection à *S. mansoni*. De plus, ils indiquent que ce gène IFNGR1 est un excellent gène candidat pour l'étude du contrôle de la survenue pathologique de fibrose dans d'autres maladies.

En conclusion, nos résultats indiquent que les niveaux d'infection par *S. mansoni* et le développement ultérieur d'une fibrose hépatique sévère sont sous le contrôle de facteurs génétiques différents. L'identification précise de ces facteurs génétiques permettra une meilleure compréhension des mécanismes biologiques intervenant dans le contrôle des niveaux d'infection (rôle de la balance Th1/Th2) et du développement de la maladie hépatique (rôle éventuel de l'IFN- γ). De plus, cette identification devrait avoir des applications importantes pour la stratégie de contrôle de la bilharziose. Elle pourra conduire à l'identification de marqueurs permettant une détection précoce des individus à risque pour les formes graves de la maladie. Cette distinction entre sujets susceptibles et résistants sera très importante pour l'évaluation des programmes de vaccination et de chimiothérapie, et devrait conduire à la définition de nouvelles stratégies d'intervention en fonction du statut susceptible/résistant du sujet. La détermination de la fonction des gènes de prédisposition devrait également fournir de nouvelles possibilités thérapeutiques plutôt orientées vers la restauration d'une réponse immunitaire partiellement déficiente.

BIBLIOGRAPHIE

- Abel L., Demenais F., Prata A., Souza A. E., Dessein A., Evidence for the segregation of a major gene in human susceptibility/resistance to infection by *Schistosoma mansoni*. *Am. J. Hum. Genet.*, 1991, 48, 959-970.
- Abel L., Dessein A. J., The impact of host genetics on susceptibility to human infectious diseases. *Curr. Opin. Immunol.*, 1997, 9, 509-516.
- Abel L., Dessein A. J., Genetic epidemiology of infectious diseases in humans: design of population-based studies. *Emerg. Infect. Dis.*, 1998, 4, 593-603.
- Dessein A.J., Hillaire D., Elwali N. E., Marquet S., Mohamed-Ali Q., Mirghani A., Henri S., Abdelhameed A. A., Saeed O. K., Magzoub M. M., Abel L., Severe hepatic fibrosis in *Schistosoma mansoni* infection is controlled by a major locus that is closely linked to the interferon-gamma receptor gene. *Am. J. Hum. Genet.*, 1999, 65, 709-721.
- Dib C., Faure S., Fizames C., Samson D., Drouot N., Vignal A., Millasseau P., Marc S., Hazan J., Seboun E., Lathrop M., Gyapay G., Morissette J., Weissenbach J., A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*, 1996, 380, 152-154.
- Khoury M. J., Beaty T. H., Cohen B. H., Fundamentals of genetic epidemiology; Oxford University Press, New York, 1993.
- Kruglyak L., Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.*, 1999, 22, 139-144.
- Marquet S., Abel L., Hillaire D., Dessein A., Full results of the genome-wide scan which localises a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. *Eur. J. Hum. Genet.*, 1999, 7, 88-97.
- Marquet S., Abel L., Hillaire D., Dessein H., Kalil J., Feingold J., Weissenbach J., Dessein A. J., Genetic localization of a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. *Nat. Genet.*, 1996, 14, 181-184.
- McLeod R., Buschman E., Arbuckle L. D., Skamene E., Immunogenetics in the analysis of resistance to intracellular pathogens. *Curr. Opin. Immunol.*, 1995, 7, 539-552.
- Mohamed-Ali Q., Elwali N. E., Abdelhameed A. A., Mergani A., Rahoud S., Elagib K. E., Saeed O. K., Abel L., Magzoub M. M., Dessein A. J., Susceptibility to Periportal (Symmers) Fibrosis in Human *Schistosoma mansoni* Infections: Evidence That Intensity and Duration of Infection, Gender, and Inherited Factors Are Critical in Disease Progression. *J. Infect Dis.*, 1999, 180, 1298-1306.
- Muller-Myhsok B., Stelma F. F., Guisse-Sow F., Muntau B., Thyé T., Burchard G. D., Gryseels B., Horstmann R. D., Further evidence suggesting the presence of a locus, on human chromosome 5q31-q33, influencing the intensity of infection with *Schistosoma mansoni*. *Am. J. Hum. Genet.*, 1997, 61, 452-454.
- Rodrigues Jr V., Piper K., Couissinier-Paris P., Bacelar O., Dessein H., Dessein A. J., Genetic control of schistosome infections by the SM1 locus of the 5q31-q33 region is linked to differentiation of type 2 helper T lymphocytes. *Infect. Immun.*, 1999, 67, 4689-4692.