

Une protéase, la calpaïne 3, est responsable d'une forme de dystrophie des ceintures

Les dystrophies des ceintures (LGMD, pour *limb girdle muscular dystrophy*) sont des myopathies transmises sur un mode autosomique dominant ou récessif, caractérisées principalement par une atrophie progressive et une faiblesse des muscles proximaux des membres. Les symptômes apparaissent durant les vingt premières années de la vie et évoluent progressivement, la perte de la marche survenant le plus souvent dix à vingt ans plus tard. La maladie s'exprime avec une grande hétérogénéité clinique. En outre, l'absence de définition nosologique spécifique rend difficile la distinction entre les malades LGMD et ceux présentant d'autres maladies neuromusculaires (telles que les dystrophies facio-scapulo-humérales [1], les myopathies de Becker, et surtout les amyotrophies spinales [2]), conditions qui ont conduit Michel Fardeau à qualifier le diagnostic LGMD de diagnostic « fourre-tout ».

C'est dans ce contexte particulier qu'il faut comprendre l'importance de la découverte par Michel Fardeau et Dominique Hillaire au sud de l'île de La Réunion d'une importante concentration de familles comprenant une proportion élevée de malades, dont les données cliniques correspondent aux critères de la forme récessive de la LGMD, tels qu'ils furent décrits il y a plus d'un siècle par W. Erg.

C'est sur cet ensemble de familles, caractérisées par un fort taux de consanguinité, qu'un gène responsable de la forme récessive *LGMD2A* (MIM n° 253600*) a pu être localisé sur le bras long du chromosome 15, grâce à la mise en évidence d'une liaison avec une sonde RFLP (*restric-*

tion fragment length polymorphism) au locus *D15S25* [3]. La localisation génétique a ensuite été confirmée sur un ensemble de familles Amish [4]. Des études portant sur des familles brésiliennes [5] et métropolitaines ont montré l'hétérogénéité génétique de la forme récessive (*m/s* n° 9, vol. 8, p. 1005) puisque, pour certaines familles, le locus morbide a pu être exclu de la région du chromosome 15, un tiers des familles seulement semblant liées à ce chromosome. Depuis, une équipe anglaise a localisé un deuxième locus, *LGMD2B*, sur le chromosome 2 [6].

Nous avons récemment, première surprise de ce travail, exclu le locus *LGMD2A* dans six familles Amish du sud de l'État de l'Indiana. Ces familles sont reliées par de multiples liens de consanguinité à celles du nord de cet État pour lesquelles l'implication du locus *LGMD2A* avait été démontrée précédemment. Cette étude a également permis de révéler l'existence d'un troisième locus *LGMD* puisque dans cette population le locus *LGMD2B* ainsi que les loci impliqués dans d'autres maladies neuromusculaires ont pu être exclus [7]. En raison de l'absence d'indices cytogénétiques, biochimiques ou autres, l'identification du gène responsable de la forme *LGMD2A* a fait appel aux stratégies de clonage positionnel. Après sa localisation primaire, le gène a pu être encadré par deux marqueurs microsatellites : *D15S129* dérivé du locus *D15S25* et l'un des premiers microsatellites développé à Généthon *D15S143*, par analyse des recombinants présents dans les familles réunionnaises. Le criblage de la banque de chromosomes artificiels de levure (YAC) du CEPH avec les sondes correspondantes a permis l'isolement de clones qui ont été utilisés en hybridation *in*

situ pour définir l'intervalle cytogénétique *LGMD2A* en 15q15.1-q21.1. L'établissement de la carte physique de cette région sous la forme d'un continuum ininterrompu de YAC a été réalisé à l'aide de STS (*sequence-tagged sites*) de différentes origines. Ce continuum comprenait 72 STS dont 25 polymorphes, ce qui correspond à une résolution génétique et physique de respectivement 140 et 400 kb. Un minimum de 11 clones couvrent la région 15q15.1 à 15q21.1, estimée à 10-12 mégabases et 7 cM [8].

Dans le but de réduire cet intervalle, l'analyse d'autres méioses ainsi que d'autres marqueurs de la région a été nécessaire. Seules les familles pour lesquelles l'appartenance au chromosome 15 a été certifiée par analyse génétique ont pu être étudiées. Le développement ciblé de nouveaux marqueurs de type microsatellite, produits à partir de YAC de la région, a permis la reconnaissance de nouveaux recombinants proximaux et distaux et, par conséquent, la réduction de l'intervalle *LGMD2A* entre deux marqueurs distants de 1 cM [9]. En outre, l'examen des haplotypes réunionnais a également mis en évidence, deuxième surprise, 6 haplotypes porteurs différents [9]. Cette hétérogénéité allélique est incompatible avec l'hypothèse d'un effet fondateur dans cette population consanguine, effet qui avait été présumé au début de cette étude.

Le nouvel intervalle correspond à une distance physique d'environ 3-4 Mb. La prévalence relativement faible et l'hétérogénéité génétique limitent le nombre de familles pouvant être utilisées dans l'analyse génétique, et par conséquent la restriction de cet intervalle par examen des recombinants. L'analyse de l'« homozygotie par descendance »* des familles consanguines n'a pas

* Numéro de la classification Mendelian inheritance in man.

permis de mieux cerner les bornes. L'étude des déséquilibres de liaison a, cependant, suggéré un emplacement dans la moitié proximale de l'intervalle 15q15.1-15q15.3 défini par les recombinants.

Afin d'établir un catalogue des transcrits de cette région, la méthode de sélection d'ADNc a été choisie. L'hybridation sélective de banques d'ADNc musculaires avec l'ADN de YAC de l'intervalle a permis l'isolement de 3 gènes connus et de 2 séquences codantes précédemment localisés dans l'intervalle, ainsi que l'identification de 10 nouvelles séquences exprimées [10].

Considérant les manifestations cliniques de la maladie, nous avons choisi d'examiner de façon prioritaire, parmi les ADNc isolés par hybridation sélective, ceux correspondant à des messagers spécifiques du tissu musculaire. Un de ces ADNc, précédemment cloné par Sorimachi *et al.* [11], nous est ainsi apparu comme candidat fonctionnel, d'autant qu'il était localisé dans la région de plus fort déséquilibre de liaison [9]. Il code pour une protéase intracellulaire, la calpaïne (protéase à cystéine dépendante du calcium) spécifique du muscle squelettique.

Les calpaïnes sont des protéases non lysosomiales soupçonnées d'avoir un rôle régulateur. Elles comprennent deux isoenzymes ubiquitaires, CANP1 et CANP2 (présentes dans toutes les cellules de vertébrés examinées), deux protéines spécifiques de l'estomac et CANP3, propre au muscle squelettique. Les formes ubiquitaires sont constituées chacune d'une large sous-unité de 80 kDa associée à la même petite sous-unité de 30 kDa, toutes codées par des gènes différents. Les larges sous-unités sont subdivisées en 4 domaines [12]. Les domaines I et III ne présentent pas d'analogie avec des protéines connues. Le domaine II est semblable aux autres protéases à cystéine avec un site actif composé d'une cystéine, d'une histidine et d'une asparagine [11]. Le domaine IV comprend 4 boucles EF, sites potentiels

de liaison du calcium. Trois régions particulières, sans analogie connue, sont également présentes dans la protéine spécifique du muscle, NS, IS1 et IS2, cette dernière contenant un motif de translocation nucléaire. Cette protéine est aussi caractérisée par une autolyse très rapide.

Après avoir isolé des cosmides correspondant à ce gène, nous en avons déterminé la séquence ainsi que l'organisation génomique. Le gène couvre 40 kb et est composé de 24 exons, codant pour une protéine de 821 acides aminés. CANP3 présente une forte analogie avec les autres calpaïnes, en particulier avec celles spécifiques du muscle d'autres espèces comme le rat et le porc. L'analyse en *Northern blot* montre un transcrite de 3,4 à 3,6 kb exprimé spécifiquement dans les muscles sans indication d'épissage alternatif.

L'analyse systématique de ce gène par la technique de détection électrophorétique des doubles brins hétérologues, par séquençage direct des exons ainsi que par analyse du produit de transcription illégitime (*m/s n° 1, vol. 6, p. 55*) dans les familles certifiées LGMD2A a conduit à l'identification de 7 mutations différentes [13]. Elles comprennent 1 mutation stop, 1 d'épissage, 2 insertions-délétions et 3 faux sens. Il faut remarquer que le parti-pris de s'intéresser à des gènes spécifiques aurait pu s'avérer un mauvais choix. En effet, on sait aujourd'hui que de nombreuses affections lésant un tissu particulier sont, en fait, liées à l'altération de gènes exprimés dans tous les tissus. Le gène SMA, responsable de l'amyotrophie spinale récemment isolé, en est un éclatant exemple [2].

Compte tenu du taux élevé de consanguinité dans la communauté Amish, la mutation transformant l'arginine 769 en glutamine se retrouve à l'état homozygote chez tous les malades Amish du nord de l'Indiana (tout en étant absente chez les patients du sud de cet État). A l'inverse, dans la population réunionnaise, nous avons, à ce jour, identifié 5 mutations différentes, confirmant l'hétérogénéité allélique suggérée par l'analyse des haplotypes [9]. Ces observations sont en contradiction avec

l'hypothèse d'un effet fondateur suggéré par les études généalogiques des familles de l'île de la Réunion. En effet, tous les malades appartiennent à un petit isolat génétique qui semble issu d'un ancêtre ayant immigré dans l'île vers 1670. La présence de ces nombreuses mutations est surprenante, compte tenu de la prévalence de la maladie estimée à 10^{-5} , incompatibilité que nous avons qualifiée de « paradoxe réunionnais ».

Les hypothèses d'erreur sur la prévalence, d'un avantage sélectif ou d'une exposition à des effets mutagènes semblent peu probables. Nous pensons que le mode de transmission de la maladie, considéré jusqu'à présent comme monogénique, est en fait plus complexe et que l'expression des mutations de la calpaïne pourrait dépendre d'un fond génétique particulier (impliquant le génome mitochondrial ou nucléaire). Ce modèle de transmission di- ou même oligo-génique [13] pourrait également dépasser le cadre des dystrophies des ceintures et permettre d'expliquer les observations faites dans d'autres affections comme la maladie de Hurler (*m/s n° 11, vol. 9, p. 1289*) ou la leucodystrophie métrachromatique, ainsi que la difficulté de reproduire certains phénotypes dans les souris transgéniques (*m/s n° 5, vol. 10, p. 565*).

Après avoir confirmé le rôle de CANP3 dans l'étiologie de LGMD2A, nous avons étendu l'analyse à d'autres familles. Dix mutations (5 faux sens, 1 non sens et 4 décalages du cadre de lecture) dont deux déjà identifiées ont pu être mises en évidence, permettant la reconnaissance de 8 nouvelles familles LGMD2A. Il découle de l'étude rapportée dans *Cell* [13] que les mutations de la calpaïne sont nombreuses, sans mutation prédominante, et réparties sur l'ensemble du gène.

Le gène de la calpaïne 3 serait donc, troisième surprise, le premier gène connu impliqué dans l'étiologie d'une dystrophie ne faisant pas intervenir une protéine de structure, mais plutôt une enzyme protéolytique qui serait altérée ou absente chez les malades. Peu de progrès dans la compréhension de la fonction de cette calpaïne ont été faits depuis son

* Excès d'homozygotie des marqueurs à proximité du locus morbide chez les patients atteints d'une affection récessive et issus de parents consanguins.

1. Kaplan J. Deux percées sur le front des maladies neuro-musculaires : la myopathie facio-scapulo-humérale et la myopathie autosomique récessive maghrébine. *médecine/sciences* 1992 ; 8 : 985-8.
2. Bürglen L, Lefebvre S, Reboullet S, Clermont O, Burllet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, Le Paslier D, Frézal J, Cohen D, Weissenbach J, Munnich A, Melki J. Identification et caractérisation d'un gène déterminant dans les amyotrophies spinales. *médecine/sciences* 1995 ; 11 : 149-51.
3. Beckmann JS, Richard I, Hillaire D, Broux O, Antignac C, Bois E, Cann H, Cottingham Jr RW, Feingold N, Feingold J. A gene for limb-girdle muscular dystrophy maps to chromosome 15 by linkage. *CR Acad Sci III* 1991 ; 312 : 141-8.
4. Young K, Foroud T, Williams P, Jackson CE, Beckmann JS, Cohen D, Conneally PM, Tischfield J, Hodes ME. Confirmation of linkage of limb-girdle muscular dystrophy, type-2, to chromosome-15. *Genomics* 1992 ; 13 : 1370-1.
5. Passos-Bueno MR, Richard I, Vainzof M, Fougereuse F, Weissenbach J, Broux O, Cohen D, Akiyama J, Marie SKN, Carvalho AA, Guilherme L, Kalil J, Tsanaclis AM, Zatz M, Beckmann JS. Evidence of genetic heterogeneity in the autosomal recessive adult forms of limb-girdle muscular dystrophy following linkage analysis with 15q probes in Brazilian families. *J Med Genet* 1993 ; 30 : 385-7.
6. Bashir R, Strachan T, Keers S, Stephenson A, Mahjneh I, Marconi G, Nashef L, Bushby KMD. A gene for autosomal recessive limb-girdle muscular dystrophy maps to chromosome 2p. *Hum Mol Genet* 1994 ; 3 : 455-7.
7. Allamand V, Broux O, Bourg N, Richard I, Tischfield J, Hodes ME, Conneally PM, Fardeau M, Jackson CE, Beckmann JS. Genetic heterogeneity of autosomal recessive limb-girdle muscular dystrophy in a genetic isolate (Amish) and evidence for a new locus. *Hum Mol Genet* 1995 ; 4 : 459-64.
8. Fougereuse F, Broux O, Richard I, Allamand V, Pereira de Souza A, Bourg N, Brenguier L, Devaud C, Pasturaud P, Roudaut C, Chiannikulchai N, Hillaire D, Bui H, Chumakov I, Weissenbach J, Cherif D, Cohen D, Beckmann JS. Mapping of a chromosome 15 region involved in limb-girdle muscular dystrophy. *Hum Mol Genet* 1994 ; 3 : 285-93.
9. Allamand V, Broux O, Richard I, Fougereuse F, Chiannikulchai N, Bourg N, Brenguier L, Devaud C, Pasturaud P, Pereira de Souza A, Roudaut C, Tischfield JA, Conneally PM, Fardeau M, Cohen D, Jackson CE, Beckmann JS. Preferential localization of the limb-girdle muscular dystrophy type 2A gene in the proximal part of a 1cM 15q15.1-q15.3 interval. *Am J Hum Genet* 1995 (sous presse).
10. Chiannikulchai N, Pasturaud P, Richard I, Auffray C, Beckmann JS. A primary expression map of chromosome 15q15 region containing the recessive form of limb-girdle muscular dystrophy (*LGMD2A*) gene. *Hum Mol Genet* 1995 ; 4 (sous presse).
11. Sorimachi H, Imajoh-Ohmi S, Emori Y, Kawasaki H, Ohno S, Minami Y, Suzuki K. Molecular cloning of a novel mammalian calcium-dependent protease distinct from both m- and mu- type. Specific expression of the mRNA in skeletal muscle. *J Biol Chem* 1989 ; 264 : 20106-11.
12. Ohno S, Emory Y, Imajoh S, Kawasaki H, Kishigami M, Suzuki K. Evolutionary origin of a calcium dependent protease by fusing of genes for a thiol protease and a calcium-binding protein? *Nature* 1984 ; 312 : 566-70.
13. Richard I, Broux O, Allamand V, Fougereuse F, Chiannikulchai N, Bourg N, Brenguier L, Devaud C, Pasturaud P, Roudaut C, Hillaire D, Passos-Bueno MR, Zatz M, Tischfield JA, Fardeau M, Cohen D, Jackson CE, Beckmann JS. Mutations in the proteolytic enzyme, calpain 3, cause limb-girdle muscular dystrophy type 2A. *Cell* 1995 ; 81 : (sous presse).
14. Sorimachi H, Toyama-Sorimachi N, Saido TC, Kawasaki H, Sugita H, Miyasaka M, Arahata K, Ishiura S, Suzuki K. Muscle-specific calpain, p94, is degraded by autolysis immediately after translation, resulting in disappearance from muscle. *J Biol Chem* 1993 ; 268 : 10593-605.

Remerciements

Nous remercions l'apport de C. Auffray, B. Bakker, P. Bucher, I. Cadjee, H. Cann, D. Caterina, P.M. Conneally, N. Fonknechten, D. Fugman, et J. Weissenbach. Mais avant tout, nous remercions tout particulièrement les patients, leurs familles et cliniciens et Bernard Barataud sans la participation et les encouragements duquel ce travail n'aurait jamais vu le jour. Ces recherches ont bénéficié du soutien de l'Association Française contre les Myopathies.