

Triplets répétés, maladies neurodégénératives et psychiatriques : mécanismes et gènes candidats

**Christian Néri
Howard M. Cann
Jean Dausset**

Initialement détectées dans le syndrome de l'X fragile, les mutations dynamiques avec expansions variables de triplets répétés polymorphes (CGG, CAG, CTG ou GAA) constituent une nouvelle classe de mutations responsables de plusieurs maladies neurodégénératives héréditaires ; elles pourraient aussi être impliquées dans les formes familiales de maladies psychiatriques multifactorielles telles que la psychose maniaco-dépressive et la schizophrénie. De nouvelles méthodes de caractérisation de ces mutations et de clonage des gènes impliqués ont été développées. L'étude détaillée des effets intracellulaires des expansions déjà identifiées permettent déjà une meilleure compréhension des mécanismes biologiques et biochimiques en jeu. Le recensement exhaustif de nouveaux triplets répétés au niveau de l'ADN génomique et des ADNc, de répétitions de glutamine au niveau des protéines, permet d'identifier des candidats à l'expansion qui pourraient être en cause dans des affections aux mécanismes génétiques encore inconnus.

ADRESSE

C. Néri : docteur ès sciences. H.M. Cann : docteur en médecine, directeur de recherche à l'Inserm. J. Dausset : docteur en médecine, professeur honoraire au Collège de France. CEPH, 27, rue Juliette-Dodu, 75010 Paris, France.

TIRÉS À PART

C. Néri.

L'expansion de triplets répétés CGG, CAG, CTG, ou GAA est à l'origine de huit maladies héréditaires (*Tableau I*) : le retard mental lié à l'X fragile (FRAXA), l'atrophie musculaire spinobulbaire (SBMA), la dystrophie myotonique (DM), l'ataxie spinocérébelleuse de type 1 (SCA1), la maladie de Machado-Joseph (MJD) ou ataxie spinocérébelleuse de type 3 (SCA3), l'atrophie dentatorubro-pallidoluysienne (DRPLA), la maladie

de Huntington (HD) (*m/s* n° 4, vol. 10, p. 472), et l'ataxie de Friedreich (AF) [1-3]. DM, SCA1, SCA3, et HD sont des maladies transmises en dominance. D'une façon générale, la longueur des répétitions trinuécléotidiques impliquées dans les huit maladies ci-dessus est fortement polymorphique dans la population normale. Le passage d'un allèle normal à un allèle muté semble progressif, et il existe un seuil pathologique dans la taille des triplets répétés,

RÉFÉRENCES

1. Willems PJ. Dynamic mutations hit double figures. *Nature Genet* 1994; 8: 213-5.
2. Campuzano V, Montermini L, Dolores Moltò M, Pianese L, Cossée M, et al. Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* 1996; 271: 1423-7.
3. Koenig M, Campuzano V, Cossée M, Mandel JL. Ataxie de Friedreich : les expansions frappent encore. *médecine/sciences* 1996; 12: 431-5.
4. Trottier Y, Lutz Y, Stevanin G, Imbert G, Devys D, Cancel G, Saudou F, Weber C, David G, Tora L, Agid Y, Brice A, Mandel JL. Cellular localization of the Huntington's disease protein and discrimination of the normal and mutant form. *Nature Genet* 1995; 10: 104-10.
5. Servadio A, Koshy B, Armstrong D, Antalffy B, Orr HT, Zoghbi HY. Expression analysis of the ataxin-1 protein in tissues from normal and spinocerebellar ataxia type 1 individuals. *Nature Genet* 1995; 10: 94-8.
6. Wells RD. Molecular basis of genetic instability of triplet repeats. *J Biol Chem* 1996; 271: 2875-8.
7. Nelson DL. Allelic expansion underlies many genetic diseases. *Growth Genet Horm* 1996; 12: 1-4.
8. Gacy AM, Goellner G, Juranic N, Macura S, McMurray CT. Trinucleotide repeats that expand in human disease form hairpin structures *in vitro*. *Cell* 1995; 81: 533-40.
9. Eichler EE, Mac Pherson JN, Murray A, Jacobs PA, Chakravarti A, Nelson DL. Haplotype and interspersal analysis of the *FMRI* CCG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. *Hum Mol Genet* 1996; 5: 319-30.
10. Brown XT, Zhong N, Dodkin C. Positive fragile X microsatellite associations point to a common mechanism of dynamic mutation evolution. *Am J Hum Genet* 1996; 58: 641-3.
11. Sutherland GR, Richards RI. Simple tandem DNA repeats and human genetic diseases. *Proc Natl Acad Sci USA* 1995; 92: 3636-41.
12. Takiyama Y, Igarashi S, Rogeava EA, Endo K, Rogeav EI, et al. Evidence for intergenerational instability in the CAG repeat in the *MJD1* gene and for the conserved haplotypes at flanking markers amongst Japanese and Caucasian subjects with Machado-Joseph disease. *Hum Mol Genet* 1995; 4: 1137-46.

variable en fonction de la maladie considérée. Deux classes de mutations dynamiqués ont été mises en évidence (*Tableau 1*): d'une part, les expansions de grande taille (jusqu'à plusieurs milliers de triplets), présentes dans les régions non codantes des transcrits (*FRAXA*, *DM*) ou dans les introns humains (*AF*), et à l'origine de diminution ou d'absence de synthèse des protéines correspondantes; d'autre part, les expansions CAG (*SBMA*, *SCA1* et *SCA3*, *DRPLA*, et *HD*), plus modérées (jusqu'à 62 à 121 triplets selon la maladie), présentes dans les régions codantes des transcrits humains, et probablement à l'origine de gains de fonction. A l'exception des gènes *SBMA* (récepteur des androgènes) et *DM* (myotnine kinase), les fonctions des protéines codées par les six autres gènes de maladie ne sont pas connues [1-3]. Dans le cas des mutations dynamiqués codantes, les protéines mutées portent un domaine polyglutaminique ([Gln]*n*) de taille anormale codé par l'expansion CAG, et sont synthétisées en même quantité que les formes normales, comme cela a été montré pour *HD* [4] et *SCA1* [5]. Plusieurs questions restent sans réponse quant aux mécanismes d'apparition des mutations dynamiqués, à leurs conséquences intracellulaires, et à leur impact sur la santé humaine.

Mécanismes de l'expansion

Les expansions de triplets répétés ont été reproduites chez *E. coli*, ce qui a fourni un modèle très utile pour l'étude des mécanismes d'apparition des mutations dynamiqués [6]. Il n'existe à l'heure actuelle aucun modèle permettant d'expliquer complètement l'apparition de ces mutations. Il a été récemment proposé que les mécanismes déterminant l'apparition des mutations dynamiqués diffèrent selon la taille des expansions (grandes ou modérées) observées chez les malades [7]. L'expansion de triplets répétés semble liée à des phénomènes complexes d'arrêt et de glissement de l'ADN polymérase lorsque celle-ci rencontre un triplet répété lors de la réplication de l'ADN [8]. L'adoption de structures secondaires en épingle

à cheveux du brin néosynthétisé et les phénomènes de réparation de l'ADN favoriseraient l'incorporation d'un triplet répété plus long que celui du brin d'origine. Les triplets répétés constitueraient donc des « pièges à polymérase ». Les pertes d'interruption dans les triplets répétés observées chez les malades (*SCA1* et *FRAXA*) suggèrent que l'homogénéité du triplet répété favorise son instabilité [7]. Seules les répétitions CGG/GCC ou CAG/CTG contenant au moins 25 triplets pourraient être sujettes à expansion du fait de structures secondaires en épingle à cheveux plus stables dont la formation serait favorisée par la présence de deux nucléotides C/G dans le triplet [8]. Cependant, la découverte récente d'une expansion GAA de grande taille (jusqu'à 800 répétitions) responsable de l'ataxie de Friedreich remet ce concept en cause et, plus largement, les tentatives de modélisation de l'expansion [2, 3]. L'importance des déséquilibres de liaison entre mutations dynamiqués (*FRAXA*, *DM*, *HD*, *SCA3*, *DRPLA*) et structure interne de la répétition trinucleotidique (*FRAXA*) [9], ou marqueurs polymorphiques proches [10-12], suggère des effets fondateurs pour l'expansion. Par ailleurs, les mutations dynamiqués semblent spécifiques de l'homme parmi les mammifères. Ainsi, les analogues murins des gènes de *HD*, *SBMA*, *SCA1* et *SCA3* contiennent des répétitions [CAG]*n* courtes (*n* < 5) et invariables; chez le chimpanzé, les phénomènes de polymorphisme sont plus modérés que chez l'homme, du fait de nombreuses mutations ponctuelles au niveau des répétitions CAG qui limiteraient les risques d'expansion [13, 14].

Gain de glutamines: coopérativité dans le ciblage des neurodégénérescences

Quelles sont les conséquences intracellulaires des expansions de CAG codant pour des polyglutamines? Un gain de fonction (« gain de glutamines ») est très probablement en cause: la formation d'interactions de forte affinité entre les protéines de la maladie et des protéines « partenaires » serait à l'origine de la dégénérescence prématurée de certaines

Tableau I
MALADIES À TRIPLETS RÉPÉTÉS

Maladies	Mode de transmission	Localisation	Gène	Répétition trinuécléotidique		Nombre de répétitions			Biais de transmission
				Type	Localisation	Normal	Prémuté	Malade	
FRAXA	Lié à l'X	Xq27.3	FMR-1*	CGG	5' non traduit	6 à 54	50 à 200	200 à 4 000	Maternelle
DM	Dominant	19q13.3	Protéine myotonine kinase**	GTG	3' non traduit	5 à 30	nd	45 à 3 000	Maternelle Paternelle
AF	Récessif	9q13	X25*	GAA	Intron	7 à 22	nd	200 à 900	Maternelle
SBMA	Lié à l'X	Xq11.12	RA	CAG	Région codante	17 à 26	nd	40 à 62	Paternelle
SCA1	Dominant	6p22.23	Ataxine 1*	CAG	Région codante	6 à 39	nd	41 à 81	Paternelle
SCA3 MJD	Dominant	14q24.3-32	MJD 1*	CAG	Région codante	13 à 36	nd	68 à 79	Paternelle
DRPLA	Dominant	12p12.ter	atrophine 1*	CAG	Région codante	7 à 23	nd	49 à 75	Paternelle
HD	Dominant	4p16.3	IT 15*	CAG	Région codante	11 à 34	nd	37 à 121	Paternelle

Les trois premières maladies indiquées correspondent aux expansions de grande taille, les cinq suivantes aux expansions modérées. FRAXA : syndrome de l'X fragile, SCA1 : ataxie spinocérébelleuse dominante de type 1, SCA3 : maladie de Machado-Joseph ou ataxie spinocérébelleuse dominante de type 3, DRPLA : atrophie dentato-rubro-pallidoluisienne, HD : maladie de Huntington, SBMA : atrophie musculaire spino-bulbaire, DM : dystrophie myotonique de Steinert, HTZ : hétérozygotie, nd : non décrit. * : fonction inconnue. ** : fonction supposée. RA : récepteur des androgènes.

populations neuronales (*m/s* n° 4, vol. 12, p. 535; n° 6-7, vol. 12, p. 852) [15, 16]. Ces interactions auraient aussi lieu lorsque les polyglutamines sont de taille normale, mais de plus faible affinité et sans conséquence pour le neurone. Les gènes des cinq maladies concernées par ce mécanisme (SBMA, SCA1, SCA3, DRPLA, et HD) sont exprimés dans le système nerveux central et dans d'autres tissus. Les interactions spécifiques des protéines de maladies avec certains partenaires synthétisés en plus grande quantité, voire uniquement, dans certaines zones du cerveau expliqueraient donc la diversité des processus de neurodégénérescence. Ces spécificités d'interaction sont attribuées à une participation des acides aminés situés en dehors des polyglutamines. Ainsi, la huntingtine (la protéine HD) interagit avec un partenaire de fonction inconnue synthétisé uniquement dans le cerveau (HAP-1, pour *huntingtin-associated protein*), et avec une plus forte affinité si elle est mutée (*m/s* n° 2, vol. 12,

p. 258) [17]. HAP-1 ne se lie pas à l'atrophine-1 (la protéine DRPLA) normale et semble synthétisée en plus grande quantité dans les zones de dégénérescence caractéristiques de la maladie de Huntington. Des partenaires de l'ataxine-1 ont aussi été identifiés: ils sont différents de HAP-1 (H. Zoghbi, communication personnelle). Il y aurait donc coopérativité des protéines à polyglutamines et de leurs partenaires dans le ciblage des neurodégénérescences (figure 1) [16]. Une étude récente a montré que l'expression d'un fragment de l'ADNc SCA3 codant en majorité pour une polyglutamine de grande taille (79 résidus) induit l'apoptose des cellules COS-7 après transfection, ainsi que la dégénérescence précoce des cellules de Purkinje chez la souris hétérozygote après transgénèse [18]. En revanche, la protéine SCA3 entière ainsi qu'un fragment de cette protéine contenant un domaine (Gln)₃₅ restent sans effet. Les auteurs ont aussi observé un retard de migration sur gel du

fragment contenant le domaine (Gln)₇₉ et d'autres protéines dénaturées, ce qui suggère l'existence de liaisons covalentes avec d'autres protéines. Cette étude souligne donc les propriétés biochimiques propres aux expansions polyglutamines de grande taille; elle met possiblement en cause des anomalies de la protéolyse des protéines qui les contiennent comme facteur de neurotoxicité et comme dénominateur commun des différentes maladies à expansion CAG. Elle diffère sur ce point d'une étude précédente de transgénèse chez la souris, fondée sur l'utilisation d'un ADNc complet SCA1 (82 répétitions CAG), et ayant conduit à l'apparition tardive d'un phénotype ataxique (*m/s* n° 2, vol. 12, p. 258) [19]. Dans ce cas, on n'a observé aucun agrégat sur coupes de cerveau après incubation avec un anticorps contre l'ataxine-1. On n'observait pas non plus, avec ce même anticorps, de signaux sur *western blot* caractéristiques de la formation de complexes protéiques. Ces deux travaux [18, 19]

RÉFÉRENCES

13. Djian P, Hancock JM, Chana S. Codon repeats in genes associated with human diseases: fewer repeats in the genes of non-human primates and nucleotide substitutions concentrated at the sites of reiteration. *Proc Natl Acad Sci USA* 1996; 93: 417-21.
14. Limprasert P, Nouri N, Heyman RA, Nopparatana C, Kamonsilp M, Deininger PL, Keats BJB. Analysis of CAG repeat of the Machado-Joseph gene in human, chimpanzee and monkey populations: a variant nucleotide is associated with the number of CAG repeats. *Hum Mol Genet* 1996; 5: 207-13.
15. Housman D. Gain of glutamines, gain of function? *Nature Genet* 1995; 10: 3-4.
16. Ross CA. When more is less: pathogenesis of glutamine repeat neurodegenerative diseases. *Neuron* 1995; 15: 493-6.
17. Li XJ, Li SH, Sharp AH, Nucifora Jr FC, Schilling G, Lanahan A, Worley P, Snyder SH, Ross CA. A Huntingtin-associated protein enriched in brain with implications for pathology. *Nature* 1995; 378: 398-402.
18. Ikeda H, Yamaguchi M, Sugai S, Aze Y, Narumiya S, Kakizuka A. Expanded polyglutamine in the Machado-Joseph disease protein induces cell death *in vitro* and *in vivo*. *Nature Genet* 1996; 13: 196-202.
19. Burright EN, Clark HB, Servadio A, Matilla T, Feddersen RM, Yunis WS, Duvick LA, Zoghbi HY, Orr HT. SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. *Cell* 1995; 82: 937-48.
20. Stott K, Blackburn JM, Butler PJG, Perutz M. Incorporation of glutamine repeats makes protein oligomerize: implications for neurodegenerative diseases. *Proc Natl Acad Sci USA* 1995; 92: 6509-13.
21. Green H. Human genetic diseases due to codon reiteration: relationship to an evolutionary mechanism. *Cell* 1993; 74: 955-6.
22. Zeitlin S, Liu JP, Chapman DL, Papaioannou VE, Efstratiadis A. Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. *Nature Genet* 1995; 11: 155-63.
23. Burke JR, Enghild JJ, Martin ME, Jou YS, Myers RM, Roses AD, Vance JM, Strittmatter WJ. Huntingtin and DRPLA proteins selectively interact with the enzyme GAPDH. *Nature Med* 1996; 2: 347-50.
24. Bao J, Sharp AH, Wagster MV, Becher M, Schilling G, Ross CA, Dawson VL, Dawson TM. Expansion of polyglutamine repeat in Huntingtin leads to abnormal protein interactions involving calmodulin. *Proc Natl Acad Sci USA* 1996; 93: 5037-42.

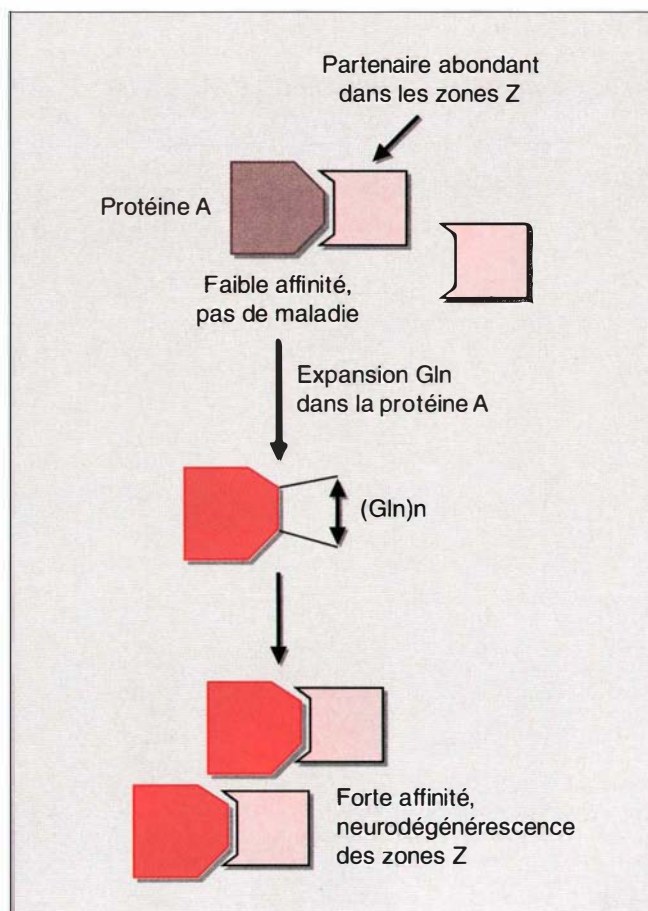


Figure 1. Schématisation du modèle du gain de fonction des expansions polyglutaminiques. Avant expansion, la protéine de maladie qui porte un domaine polyglutaminique (Gln)*n* de taille normale ainsi que sa protéine partenaire dans le système nerveux sont capables d'interagir à faible affinité et sont disponibles pour effectuer leurs fonctions, en particulier dans la cellule neuronale. A la suite de l'expansion du domaine (Gln)*n*, cette interaction s'opère à plus forte affinité, et la moindre disponibilité du partenaire, spécifiquement synthétisé dans le SNC contrairement à la

protéine, provoque la dégénérescence de populations neuronales dans les zones SNC où il est plus fortement présent (zones Z). L'hypothèse selon laquelle des fragments de protéolyse de la protéine pourraient participer à l'étiologie des maladies à expansion de polyglutamines n'est pas représentée sur ce schéma.

indiquent que la présence d'une expansion Gln dans une protéine lui confère des propriétés anormales d'interaction avec elle-même ou avec d'autres protéines, résultats à rapprocher de l'oligomérisation de l'inhibiteur de la chymotrypsine-2 provoquée par incorporation artificielle d'un domaine polyglutaminique (Gln)*n* [20]. L'acquisition de nouvelles propriétés par les polyglutamines à la suite de leur expansion pourrait donc provoquer la dégénérescence du neurone selon deux modes impliquant, soit la protéine de maladie entière, soit ses fragments de protéolyse. Dans ce cadre, l'hypothèse initiale de l'implication des transglutaminases et des protéases du cerveau dans l'accumulation anormale de résidus Gln-Lys provenant de la dégradation de complexes entre

les résidus de l'expansion Gln et les résidus Lys des protéines du cerveau reste aussi à clarifier [21]. Il n'existe pas de données publiées sur la structure tridimensionnelle des protéines portant des expansions de polyglutamines, si ce n'est les propriétés apolaires de peptides synthétiques contenant plus d'une dizaine de résidus Gln, probablement dues à des repliements en feuillet β [20].

L'étude de la huntingtine a permis d'obtenir d'autres données sur la biologie et la biochimie des protéines à polyglutamines. L'inactivation des deux allèles de l'analogue du gène *HD* chez la souris conduit à une dégénérescence cellulaire importante au niveau de l'ectoderme et à la mort de l'embryon au bout de dix jours [22]. Cela suggère que la huntingtine normale retarde les proces-

sus d'apoptose, mais, dans le cadre du modèle « gain de glutamines », les phénomènes de neurodégénérescence dans la maladie de Huntington resteraient en premier lieu liés au déficit du partenaire HAP-1 sous sa forme libre (*m/s n° 4, vol. 12, p. 535*). Il semble aussi que les peptides polyglutamines, la huntingtine et l'atrophine-1 interagissent avec la GAPDH (glyceraldéhyde 3-phosphate déshydrogénase) (*m/s n° 6-7, vol. 12, p. 852*), une protéine dont l'interaction avec des domaines riches en Gln est bien connue [23]. Cela est aussi vrai pour l'ataxine-1 (H. Zoghbi, communication personnelle). L'interaction de la GAPDH avec plusieurs protéines de maladies apporte peu d'éléments pour comprendre le ciblage des neurodégénérescences car la GAPDH est une enzyme ubiquitaire. Enfin, la mise en évidence d'un complexe protéique contenant la calmoduline et la huntingtine (à partir d'extraits de cerveau de rat) [24], ainsi que la rétention accrue de la huntingtine humaine par la calmoduline (sur colonne d'affinité) lorsque son domaine polyglutamine est de taille anormale [24], permettront peut-être de mieux comprendre l'étiologie de la maladie de Huntington.

Expansion CAG et familles de malades : détection et interprétations

Trois critères permettent de supposer l'existence de mutations dynamiques dans les familles de malades. Ces mutations se traduisent sur le plan clinique par un phénomène d'anticipation (critère 1), à savoir l'apparition des symptômes dans les familles de malades de plus en plus précoce au cours des générations, due à l'augmentation progressive de la taille du triplet répété. Il est possible de détecter des expansions CAG directement dans l'ADN des malades à l'aide de la technique RED (*repeat expansion detection*, critère 2), sans toutefois pouvoir les localiser ou cloner les gènes qui les portent [25]. Cette technique ne constitue pas un outil de diagnostic mais permet simplement de supposer l'implication d'une expansion, sous réserve qu'elle ségrege avec la maladie. Enfin, l'utili-

sation d'un anticorps monoclonal (1C2), capable de détecter une expansion Gln sur *western blot* à partir de lymphoblastes de malades (critère 3) [4], permet de préciser la nature codante d'expansions CAG détectées en RED. Récemment, une autre approche a été décrite : la détection cytogénétique d'expansions CAG directement au niveau des chromosomes de malades par FISH (*fluorescent in situ hybridization*) (*m/s n° 5, vol. 12, p. 653*) [26]. L'efficacité réelle de cette approche reste à définir.

L'utilité de l'anticorps 1C2 est en cours d'évaluation pour le clonage d'expression des gènes de SCA2 et SCA7, maladies pour lesquelles la ségrégation des expansions Gln détectées par 1C2 avec le phénotype malade indique clairement une mutation dynamique CAG codante [27, 28]. Existe-t-il des maladies autres que SCA2 et SCA7, causées par l'expansion de triplets répétés ? Des phénomènes d'anticipation ont été observés pour la démence familiale, la maladie de Parkinson, et le syndrome de Marner*, en général au niveau d'une seule famille, ce qui est insuffisant pour conclure quant au rôle d'une expansion. D'autres observations, fondées sur au moins deux des critères mentionnés ci-dessus (anticipation, expansion CAG dans l'ADN), laissent penser que les formes familiales de la psychose maniaco-dépressive [29, 30] et de la schizophrénie [30-32] pourraient être, en tout ou en partie, attribuées à une mutation dynamique CAG. Si les études de détection des expansions sont d'interprétation relativement aisée pour les maladies neurodégénératives monogéniques (ségrégation mendélienne avec le phénotype malade), il n'en est pas de même pour l'étude des familles atteintes de maladies psychiatriques. La variabilité des tableaux cliniques fait que la définition de ces maladies est plus difficile [33]. Schizophrénie et psychose maniaco-dépressive sont des maladies polyfactorielles et génétiquement hétérogènes, ce qui rend complexe l'analyse génétique (*m/s n° 8-9, vol. 12, p. 1001*). Les phéno-

mènes d'anticipation pourraient, en outre, procéder de phénomènes génétiques n'ayant lieu que dans des conditions sociologiques particulières comme les mariages entre malades [34]. Quelles que soient les conclusions des recherches accomplies sur les expansions CAG dans la schizophrénie ou la psychose maniaco-dépressive, la prudence est de rigueur : un assez grand nombre de malades a-t-il été étudié ? Les risques de biais dans les diagnostics effectués (comme les phénotopies) ont-ils été correctement évalués ? Plusieurs études de localisation génétique illustrent maintenant la nature génétique complexe de la schizophrénie [35] et de la psychose maniaco-dépressive [36]. Les faibles valeurs obtenues dans les études de liaison génétique, par exemple pour la psychose maniaco-dépressive [37, 38], ne sont pas compatibles avec la notion d'un gène majeur responsable de la maladie. Il est donc nécessaire d'effectuer les recherches d'expansions sur un grand nombre de familles, si possible larges et issues d'isolats comme la communauté Amish de Pennsylvanie [37], ou les familles de la Vallée Centrale au Costa Rica [38], et recrutées dans le cadre d'actions concertées utilisant des protocoles standardisés de définition clinique. Il apparaît aussi judicieux de rechercher des expansions autres que les répétitions CAG, voire d'autres types de mutations permettant de mieux comprendre pourquoi les tableaux cliniques sont aussi variés et se superposent parfois au sein d'une même famille (susceptibilité multiple).

Clonage des gènes : alternatives, ADNc humains et (CAG/CTG)_n

Clonage positionnel ou recherche de gènes candidats ? La recherche de gènes candidats semble bien adaptée au problème des maladies neurodégénératives monogéniques puisqu'un grand nombre d'entre elles seraient dues à des expansions de triplets répétés, ainsi qu'aux maladies psychiatriques du fait de leur nature polygénique. La recherche de gènes candidats est entreprise à partir d'échantillons de malades (ADN ou

* Décrit en 1949, il fait partie des cataractes héréditaires dont le tableau clinique est principalement une cataracte zonulaire. Sa transmission est autosomique dominante.

RÉFÉRENCES

25. Schalling M, Hudson TJ, Buetow KH, Housman DE. Direct detection of novel expanded trinucleotide repeats in the human genome. *Nature Genet* 1993; 4: 135-9.
 26. Haaf T, Sirugo G, Kidd KK, Ward DC. Chromosomal localization of long trinucleotide repeats in the human genome by fluorescence *in situ* hybridization. *Nature Genet* 1996; 12: 183-5.
 27. Gispert S, Lunkes A, Santos N, Orozco G, Ha-Hao D, *et al.* Localization of the candidate gene D-amino acid oxidase outside the refined 1-cM region of spinocerebellar ataxia 2. *Am J Hum Genet* 1995; 57: 972-5.
 28. Benomar A, Krols L, Stevanin G, Cancel G, Le Guern E, *et al.* The gene for autosomal dominant cerebellar with pigmentary macular dystrophy maps to chromosome 3p12-p21.1. *Nature Genet* 1995; 10: 84-8.
 29. Lindblad K, Nylander PO, De Bruyn A, Sourey D, Zander C, *et al.* Detection of expanded CAG repeats in bipolar affective disorder using the repeat expansion detection (RED) method. *Neurobiol Disease* 1995; 2: 55-62.
 30. O'Donovan MC, Guy C, Craddock N, Murphy KC, Cardno AG, Jones LA, Owen MJ, McGuffin P. Expanded CAG repeats in schizophrenia and bipolar disorders. *Nature Genet* 1995; 10: 380-1.
 31. Thibault F, Martinez M, Petit M, Jay M, Campion D. Further evidence for anticipation in schizophrenia. *Psychiatry Research* 1995; 59: 25-33.
 32. Morris AG, Gaitonde E, McKenna PJ, Mollon JD, Hunt DM. CAG repeat expansions and schizophrenia: association with disease in females and with early age-at-onset. *Hum Mol Genet* 1995; 4: 1957-61.
 33. Tsuang MT, Faraone SV, Lyons MJ. Identification of the phenotype in psychiatric genetics. *Eur Arch Psychiatry Clin Neurosci* 1993; 243: 131-42.
 34. Spence MA, Bishop DT, Boehnke M, Elston RC, Falk C, Hodge SE, Ott J, Rice J, Merikangas K, Kupfer D. Methodological issues in linkage analyses for psychiatric disorders: secular trends, assortative mating, bilineal pedigrees. Report of the McArthur Foundation Network I Task Force on Methodological Issues. *Hum Hered* 1995; 43: 166-72.
 35. Moises HW, Yang L, Kristbjarnarson H, Wiese C, Byerley W, *et al.* An international two-stage genome-wide search for schizophrenia susceptibility genes. *Nature Genet* 1995; 11: 321-4.
 36. Rish N, Botstein D. A manic depressive history. *Nature Genet* 1996; 12: 351-3.
- protéines), ou à partir d'échantillons d'individus non atteints. Dans le premier cas, citons: (1) la recherche d'expansion CAG par enrichissement en fragments de restriction contenant des répétitions [CAG]*n*, comparaison des profils de migration avec l'ADN normal, et séquençage des fragments candidats [39], technique prometteuse dans le principe et en cours de mise au point, (2) la détection par FISH d'expansion CAG dans les chromosomes (*m/s n° 5, vol. 12, p. 653*) [26], et (3), lorsqu'on met en évidence une expansion polyglutaminique, le criblage de banques d'expression ou l'immunopurification d'extraits cellulaires sur colonne d'affinité à l'aide de l'anticorps 1C2. Dans le deuxième cas, citons: (1) le séquençage à grande échelle de régions candidates, l'avantage étant de fournir directement des informations sur toute la région analysée (promoteurs, séquence complète du gène, organisation exon-intron); (2) la recherche à grande échelle de nouvelles répétitions [CAG]*n* polymorphiques dans l'ADN génomique, approche par définition exhaustive mais peu puissante dans les prédictions de l'expression des répétitions trinucleotidiques candidates [40]; et (3) la recherche à grande échelle de nouveaux ADNc portant des répétitions [CAG]*n* polymorphiques. Moyennement exhaustive du fait de la nature partielle des clones d'ADNc, mais précise dans les informations apportées, c'est l'approche qui a été choisie au CEPH. Nous avons procédé au criblage à grande échelle d'ADNc humains de référence (clones d'ADNc numérotés et accessibles à tout utilisateur) pour la présence de répétitions [CAG]*n* polymorphiques. Cribler des ADNc par hybridation avait déjà été entrepris [41, 42], mais de façon ponctuelle et non exhaustive. Nous avons fondé notre approche sur l'hybridation à stringence élevée de 100 128 clones d'ADNc de cerveau humain et sur l'analyse de plus de 200 000 séquences partielles EST (*expressed sequence tag*), disponibles dans la banque de données Genbank et représentant environ 50 000 gènes uniques exprimés dans différents tissus [43]. Le séquençage des ADNc sélectionnés par hybridation et représentant de nouveaux gènes, l'analyse de polymorphisme dans 40 ADN de référence du CEPH, la localisation des meilleurs candidats au niveau des hybrides somatiques et des YAC [44], et la recherche des cadres de lecture ouverts ont permis une caractérisation précise des candidats obtenus [45]. Actuellement, 23 longues répétitions [CAG]*n* ($9 < n < 29$) ont été identifiées: 9 sont polymorphiques, dont 3 fortement polymorphiques. Nos données montrent que les répétitions [CAG]*n* > 9 sont très rares dans les ARN messagers de cerveau (environ 1/2500) et, *a fortiori*, celles qui sont polymorphiques (1/5000) voire fortement polymorphiques (1/15000). Longueur et homogénéité du motif répété (au moins 10 répétitions CAG ininterrompues) semblent conditionner le polymorphisme. Cependant, il n'existe pas de corrélation absolue entre la longueur d'une répétition [CAG]*n* > 9 et son degré de polymorphisme. La liste actuelle de nouveaux ADNc portant des répétitions [CAG]*n* polymorphiques (*Tableau II*) est à la disposition de la communauté scientifique sur le serveur internet du CEPH*, et constitue une source non négligeable de gènes candidats étant donné leur rareté. En effet, la faible fréquence observée pour les répétitions [CAG]*n* fortement polymorphiques dans notre étude est comparable à celle obtenue par Gastier *et al.* [40], qui ont isolé, à partir de l'ADN génomique, 14 nouvelles répétitions (CAG)*n* potentiellement très polymorphiques, dont cinq seraient exprimées d'après l'analyse faite à l'aide du programme de prédiction d'exons GRAIL. Les ADNc sélectionnés dans notre étude ayant été obtenus à l'aide d'une amorce oligo-dT, ils représentent essentiellement les régions 3' des ARNm, et faiblement les régions 5'. Il est donc nécessaire de procéder au criblage de ces régions à l'aide d'ADNc obtenus par ancrage au niveau de l'extrémité 5' des ARNm, ou par ancrage au hasard à l'aide d'hexamères, cette dernière approche étant en cours dans d'autres laboratoires [46].

* Adresse URL: <http://www.cephb.fr>

Tableau II					
CARACTÉRISTIQUES DES CLONES D'ADNc CONTENANT UN (CAG) _n POLYMORPHIQUE (SOURCE CEPH)					
Nom du clone Mode de sélection Répétition 5'-3'	Banque d'origine	Homologie dans Genbank (EST, numéro d'accèsion)	Hétérozygotie	Nombre d'allèles (nombre de copies)	Localisation
2.116 Hybridation (CAG)20	Cerveau fœtal, Institut Max Plank, Berlin	aucune	0,88	9 (14-24)	3p14
2.81 Hybridation (CTG)19	Cerveau fœtal Institut Max Plank, Berlin	aucune	0,58	3 (13-19)	chromosome 19
2.119 Hybridation (CAG)28	Cerveau fœtal Institut Max Plank, Berlin	aucune	0,22	5 (23-28)	4p15 et 4q28.3
2.46 Hybridation, Genbank (CAG)10	Cerveau fœtal, Institut Max Plank, Berlin & Cerveau fœtal, Stratagene # 936206	T07007	0,68	9 (8-18)	12q13.3 et 3p21
2.70 Hybridation (AAG)7(CAG)8	Cerveau fœtal, Institut Max Plank, Berlin	aucune	0,05	2 (8, 9)	19q13.43
i.8 Hybridation, Genbank (CAG)13	Cerveau fœtal, normalisée, consortium IMAGE Livermore	R18580	0,90	15 (8-25)	13q13.1-q13.2
i.180 Genbank (CAG)12	Tissu mammaire 3NbHBst, normalisée Placenta Nb2HP, normalisée	R48249 H25944	0,32	2 (6, 9)	1q32-q41
i.181 Genbank (CAG)10	Foie et rein fœtal, normalisée	T85390	0,58	2 (7, 10)	chromosome 3, 11, 8
i.182 Genbank (CAG)12	Tissu mammaire 3NbHBst, normalisée	T49359	0,05	3 (9-12)	chromosome 5

* Adresse URL : <http://www.cephb.fr>

Conclusion : perspectives thérapeutiques, analyse du génome humain et gènes candidats

Malgré la complexité des mutations dynamiques, les efforts accomplis ces trois dernières années ont permis de prendre conscience des subtilités des mécanismes et des perturbations biochimiques en jeu. Les données obtenues fournissent de nouvelles bases de réflexion concernant le traitement des maladies à expansion CAG,

et l'efficacité de la recherche de gènes candidats à l'échelle du génome humain.

Si le modèle du gain de fonction des répétitions de glutamines se confirme, il est possible de rechercher des molécules (peptides synthétiques, composés organiques) capables de bloquer l'activité anormale des expansions de polyglutamines au niveau intracellulaire. A ce titre, les travaux visant à identifier les partenaires des protéines de maladies et à définir le rôle intrinsèque

des polyglutamines sont essentiels [17, 18, 47]. Ces études sont en cours dans plusieurs laboratoires, et sont fondées sur l'utilisation des techniques de double-hybride dans la levure [48, 49], de transfection de cellules non neuronales [18], et de transgénèse chez la souris [18, 19, 22]. Couplées au criblage de banques combinatoires de peptides ou de composés organiques [50], de telles recherches pourraient permettre la découverte de molécules leurres susceptibles de se fixer spécifiquement

RÉFÉRENCES

37. Ginns EI, Ott J, Egeland JA, Allen CR, Fann CSJ, *et al.* A genome-wide search for chromosomal loci linked to bipolar affective disorder in the Old Order Amish. *Nature Genet* 1996; 12: 431-5.
38. Freimer NB, Reus VI, Escamilla ME, McInnes LA, Spesny M, *et al.* Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22-q23. *Nature Genet* 1996; 12: 436-41.
39. Broude NE, Chandra A, Smith CL. Differential display of trinucleotide repeat-containing sequences in human genome. In: Bentley D, Green E, Waterson R, ed. *Genome Mapping and Sequencing*. New York: Cold Spring Harbor, 1995: 41.
40. Gastier JM, Brody T, Pulido JC, Businga T, Sunden S, *et al.* Development of a screening set for new (CAG/CTG)_n dynamic mutations. *Genomics* 1996; 32: 75-85.
41. Li SH, McInnis MG, Margolis RL, Antonarakis SE, Ross CA. Novel triplet repeat containing genes in human brain: cloning, expression, and length polymorphisms. *Genomics* 1993; 16: 572-9.
42. Riggins GJ, Lokey LK, Chastain JL, Leiner HA, Sherman SL, Wilkinson KD, Warren S. Human genes containing polymorphic trinucleotide repeats. *Nature Genet* 1992; 2: 186-91.
43. Jordan B. La valse des étiquettes. *médecine/sciences* 1995; 11: 273-6.
44. Chumakov I, Rigault P, Le Gall I, Belanne-Chantelot C, Billault A, *et al.* A YAC contig map of the human genome. *Nature* 1995; 377: 174-82.
45. Néri C, Albanese V, Lebre AS, Holbert S, Saada C, *et al.* Survey of CAG/CTG repeats in human cDNAs representing new genes: candidates for inherited neurological diseases. *Hum Mol Genet* 1996; 5: 1001-9.
46. Margolis RL, Breshel TS, Li SH, Kidwai AS, Antonarakis SE, McInnis MG, Ross CA. Characterization of cDNA clones containing cDNA CCA trinucleotide repeats derived from human brain. *Somat Cell Mol Genet* 1995; 21: 279-84.
47. Davidson JD, Servadio A, DuVick, Zoghbbi HY, Orr HT, Burchright EN. Characterization of ataxin-1 self-interactions and identification of genes encoding putative interacting proteins. *Proc Am Soc Hum Genet* 1995: abstract no 197.
48. Fields S, Song OK. A novel genetic system to detect protein-protein interactions. *Nature* 1989; 340: 245-6.
49. Plessis A, Camonis JH. Le système double-hybride, mode d'emploi. *médecine/sciences* 1994; 10: I-IX.
50. Chabala JC. Solid-phase combinatorial chemistry and novel tagging methods for identifying leads. *Curr Opin Biotechnol* 1995; 6: 632-9.

sur les expansions Gln, et de préserver ainsi la disponibilité des protéines-partenaires dans le neurone. La recherche de molécules actives de petite taille est indispensable si on veut apporter le principe actif directement au niveau du tissu cérébral; la spécificité et les caractéristiques physicochimiques et pharmacocinétiques des molécules identifiées resteront des paramètres déterminants pour la réussite d'une approche de ce type.

Plusieurs nouveaux gènes candidats pour les maladies à expansion CAG ont été identifiés à partir de l'ADN génomique ou des ADNc, et sont en cours d'analyse à la recherche d'une expansion dans les familles de malades chaque fois que celles-ci ont été localisées dans les régions candidates de maladies neurodégénératives monogéniques; ils sont plus systématiquement recherchés pour les maladies psychiatriques du fait de leur caractère polygénique et de la faiblesse des localisations génétiques. Outre les groupes de gènes candidats dont sont caractérisés la séquence, le caractère polymorphe, et la localisation

dans le génome humain, il existe des listes brutes de clones d'ADNc de référence sélectionnés par hybridation CAG ou CGG*. Les recherches de gènes candidats les plus récentes sont fondées sur des analyses à grande échelle, et laissent espérer un inventaire complet des répétitions trinucleotidiques potentiellement instables dans les gènes humains. La comparaison des données publiées par différents laboratoires montre qu'il existe peu d'éléments en commun entre les divers groupes de clones candidats obtenus. Ainsi, le groupe de candidats (ADN génomique) sélectionnés par Gastier *et al.* [40] ne contient pas l'ensemble des clones d'ADNc sélectionnés au CEPH [45]. Cela suggère qu'aucune approche ne garantit un inventaire complet des candidats au niveau des banques analysées, et qu'il est nécessaire de poursuivre la recherche de nouvelles répétitions trinucleotidiques candidates à l'expansion ■

* Notamment sur le serveur WWW du consortium IMAGE, Lawrence Livermore National Laboratory, CA, USA; contact: Dr G. Lennon.

Summary

Triplet repeats, neurodegenerative and neuropsychiatric diseases: mechanisms and candidate genes

Initially found to be the causative mutation for the fragile X syndrome, the expansion of polymorphic triplet repeats (CGG, CAG, CTG or GAA) in human transcripts or introns (dynamic mutations) is a new class of mutations for genetic diseases. CAG expansion in coding sequences is a generalized model for the molecular pathogenesis of several inherited neurodegenerative diseases. Dynamic mutations could also be involved in neuropsychiatric diseases such as bipolar affective disorder and schizophrenia. Over the past three years, new methodologies have been developed in order to characterize dynamic mutations in patients, as well as to rapidly clone the disease genes. The detailed analysis of known dynamic mutations and their consequences at the cellular level leads to a better understanding of the biological and biochemical mechanisms underlying these diseases. Specifically, several proteins interacting with polyglutamine stretches in proteins encoded by CAG

repeats in the genes of spinocerebellar ataxia 1, Huntington's disease, and dentatorubral-pallidolysian atrophy have been identified. Protein-protein interactions may account for the selective neurodegeneration observed in each disease. The survey of potentially expanded triplet repeats in genomic DNA, cDNAs, and proteins is providing the scientific community with an increasing number of candidate genes. These approaches should also permit the putative role of dynamic mutations in neuropsychiatric diseases to be further documented. These multifactorial diseases indeed show complex clinical and genetic traits, and are probably polygenic. Large-scale testing of patients and their families is the only way to detect all mutations involved. As for therapeutic approaches, the search for lead compounds with ability to block the gains of function subsequent to glutamine expansion in neurodegenerative disease proteins may be promising.