



HAL
open science

Développement d'une stratégie de criblage du gène MECP2 chez des patientes atteintes d'un syndrome de Rett : mise en évidence d'un spectre hétérogène de mutations

Violaine Bourdon

► To cite this version:

Violaine Bourdon. Développement d'une stratégie de criblage du gène MECP2 chez des patientes atteintes d'un syndrome de Rett : mise en évidence d'un spectre hétérogène de mutations. Médecine humaine et pathologie. Université Henri Poincaré - Nancy 1, 2002. Français. NNT : 2002NAN10292 . tel-01748672

HAL Id: tel-01748672

<https://hal.univ-lorraine.fr/tel-01748672v1>

Submitted on 29 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

UNIVERSITE NANCY I – HENRI POINCARÉ
FACULTE DE MÉDECINE

N°

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE NANCY I

Spécialité Génie Biologique et Médical

présentée et soutenue publiquement
par

Violaine BOURDON

le 6 juin 2002

**Développement d'une stratégie de criblage du gène *MECP2* chez des
patientes atteintes d'un syndrome de Rett :
Mise en évidence d'un spectre hétérogène de mutations**

Directeur de thèse

Professeur Philippe JONVEAUX

JURY

Professeur Anne MONCLA
Professeur Michel KOENIG
Docteur Jamel CHELLY
Professeur Simone GILGENKRANTZ
Professeur Philippe JONVEAUX

Rapporteur
Rapporteur
Examineur
Examineur
Examineur



Remerciements

Merci aux professeurs Anne Moncla, Michel Koenig, Simone Gilgenkrantz et au docteur Jamel Chelly de me faire le grand honneur de juger ce travail.

Je voudrais remercier toutes les personnes qui ont participé de loin comme de près à ce travail, et en particulier :

Un grand merci à vous, Philippe Jonveaux, pour la confiance que vous avez toujours su me témoigner, et ce depuis un coup de téléphone un certain matin de novembre 1997, merci.

Un grand merci à toi, Christophe, pour ta disponibilité, ta générosité. Tu m'as appris que chaque problème a une solution, même si la solution entraîne d'autres problèmes plus compliqués encore ! Les voies de la Science sont inépuisables ! Merci.

Un grand merci à vous, Marie-José, votre esprit cartésien et votre logique sont imparables, et ô combien appréciables ! Merci.

Un grand merci à Michelle, Madame Southern,
Brigitte, Madame ARN,
Claudine et Florence, Mesdames Extraction d'ADN et
Lignées,

Vos savoir-faire sont une richesse inestimable.

Merci à toutes les personnes du laboratoire qui se sont intéressées à mon travail et qui m'ont soutenue par un mot, un geste. Merci « aux petites mains » qui m'ont aidée par-ci par-là.

Et puis, on dit que les gens du Nord sont froids et réservés, et que les gens du Sud sont accueillants et superficiels. Originnaire du Sud, je dis que les Lorrains sont chaleureux, je n'ai jamais trouvé de porte fermée...

Je remercie toute ma famille pour son soutien et son réconfort inépuisables

Un immense merci à ma maman, tu vois j'y suis presque ! et pardon pour le souci que je te cause !

Merci, Merci, Merci à ma grande sœur adorée !!! (mon juge préféré !)

Merci à mon petit frère chéri ! (I will think about you lost in London city !)

MERCI à toi, David, pour ta présence, ta patience et ton amour ! (Allez, plus que 38 ans avant la retraite dans le Sud !)

Merci à mon père.

Enfin, j'aurais presque envie de remercier *France Telecom* qui m'a quasiment fait croire que Nancy était à Velaux !

Je dédie ce travail à « toutes les petites filles Rett »

Je remercie vivement l'association française du syndrome de Rett pour son soutien.

Table des matières

Liste des figures	9
Listes des tableaux	10
Abréviations	11
<u>PREAMBULE</u>	<u>14</u>
<u>INTRODUCTION</u>	<u>15</u>
A. Les retards mentaux	16
I. Définition	16
II. Les retards mentaux liés au chromosome X (RMLX)	16
1. Définitions	16
2. Données génétiques des RMLX	17
3. Mécanismes physiopathologiques impliqués dans les RMLX	18
B. Le syndrome de Rett	22
I. Historique et clinique	22
1. Historique	22
2. Clinique	22
2.1. Les critères de diagnostic	22
2.2. L'histoire naturelle	24
2.3. Les variants	26
II. Etiologie du syndrome de Rett	27
1. Syndrome de Rett, maladie génétique	27
1.1. Quelques cas familiaux	27
1.2. Les anomalies cytogénétiques associées au syndrome de Rett	27
1.3. Quel mode de transmission ?	29
1.3.1. Mise en évidence de mutations au sein de l'ADN mitochondrial	29
1.3.2. Recherche d'expansions de triplets	29
1.3.3. Hypothèse d'une transmission dominante liée au chromosome X d'origine maternelle	30
1.3.4. Hypothèse d'une transmission dominante liée au chromosome X d'origine paternelle	31
1.3.5. Hypothèse d'une maladie digénique impliquant un locus autosomique et un locus porté par le chromosome X	31
1.4. Inactivation du chromosome X et syndrome de Rett	33
1.4.1. Quelques notions sur l'inactivation du chromosome X	33
1.4.2. Etudes de l'inactivation du chromosome X dans le syndrome de Rett	36
2. Syndrome de Rett, maladie génétique liée au chromosome X	37

2.1. Deux translocations t(X;A) associées au syndrome de Rett	37
2.2. Etudes de liaison génétique	38
2.2.1. Principe	38
2.2.2. Carte d'exclusion du syndrome de Rett	40
2.2.3. Recherche du gène responsable du syndrome de Rett	42
2.2.3.1. Etudes de gènes candidats	42
2.2.3.2. Stratégie de clonage par une approche combinée	42
2.2.3.3. Criblage systématique de la région Xq28	43
C. Régulation de la transcription	44
I. Quelques rappels sur la transcription	44
II. Régulation de la transcription	46
1. Modifications de la chromatine	47
1.1. Acétylation et désacétylation	48
1.1.1. Les HAT	49
1.1.2. Les HDAC	50
2. Méthylation de l'ADN	53
2.1. Définition et fonctions de la méthylation	53
2.1.1. Les acteurs de la méthylation	55
2.2. Méthylation et répression transcriptionnelle	58
2.2.1. Les protéines de liant aux cytosines méthylées ou Methyl CpG-binding proteins (MeCP)	59
2.2.2. La protéine MeCP2	63
2.2.2.1. Organisation et expression du gène <i>MECP2</i>	63
2.2.2.2. Structure et fonction de la protéine MeCP2	65
2.2.2.3. Fonction des autres MeCP	68
D. Objet de notre étude	69
RESULTATS	70
I. Présentation des patients	71
II. Recherche de mutations dans le gène <i>MECP2</i> à l'échelle moléculaire	72
1. Première stratégie de criblage : en trois étapes	72
1.1. La technique de CSGE	73
1.2. Séquençage et interprétation des variations nucléotidiques	74
1.3. Les résultats obtenus par CSGE et séquençage	76
1.3.1. Les mutations faux-sens	78
1.3.2. Les mutations non-sens et perturbant le cadre de lecture	79
1.3.3. Une mutation à un site d'épissage	81
1.3.4. Les mutations silencieuses ou polymorphismes	83
2. Seconde stratégie de criblage du gène <i>MECP2</i> : en deux étapes	83
2.1. La technique de DHPLC	83
2.2. Application de la DHPLC au gène <i>MECP2</i>	85
2.3. Les résultats obtenus par DHPLC	88
3. Deux cas particuliers	89

3.1. Mutations somatiques en mosaïque	89
3.1.1. La mutation 1165del26 : une découverte fortuite	89
3.1.2. Criblage systématique	94
3.2. La substitution 1126C>T, P376S	94
4. Les résultats obtenus par Southern blot	97
5. Analyse de la région 3' non traduite du gène <i>MECP2</i>	103
III. Recherche de mutations dans le gène <i>MECP2</i> à l'échelle chromosomique	104
DISCUSSION	106
<hr/>	
A. Syndrome de Rett, et gène <i>MECP2</i>	107
I. Le spectre mutationnel du gène <i>MECP2</i>	107
1. Quelques points chauds de mutations	107
2. Mutations en mosaïque somatique	114
3. Analyse des régions régulatrices et non codantes du gène <i>MECP2</i>	115
II. Origine parentale des mutations dans le gène <i>MECP2</i>	115
III. Gène <i>MECP2</i>, syndrome de Rett et retard mental chez le garçon	117
IV. Corrélation génotype-phénotype et rôle de l'inactivation du chromosome X	122
1. Corrélation génotype-phénotype	122
2. Influences de l'inactivation du chromosome X sur le phénotype	123
2.1. Dans les cas sporadiques	123
2.2. Dans les cas familiaux	124
V. Hypothèse d'un second locus impliqué dans le syndrome de Rett	125
B. Neuropathologie du syndrome de Rett et fonction de la protéine MeCP2	129
I. Quelques définitions anatomiques et fonctionnelles	129
II. Données anatomopathologiques et neuropathologiques du syndrome de Rett	130
III. Conséquences de mutations dans le gène <i>MECP2</i>	132
1. Etudes fonctionnelles de protéines MeCP2 anormales	132
2. Conséquences des mutations sur la fonction du complexe de répression transcriptionnelle	135
3. Les modèles animaux	136
3.1. Définitions	136
3.2. Modèles murins du syndrome de Rett	137

4. Recherche des gènes cibles de la protéine MeCP2	141
CONCLUSION	142
MATERIEL ET METHODES	146
A. Patients	147
B. Méthodes	147
I. Techniques de biologie moléculaire	147
1. Extraction d'ADN génomique total	147
1.1. A partir de sang	147
1.2. A partir de cellules buccales	148
1.3. A partir de bulbes de cheveux	148
2. Amplification <i>in vitro</i> de séquences d'ADN par PCR	148
2.1. Les conditions d'amplification de l'ADN pour la CSGE	149
2.2. Les conditions d'amplification de l'ADN pour la DHPLC	150
3. Technique de CGSE	152
4. Technique de DHPLC	152
5. Le séquençage	153
6. Analyse de fragments par PCR fluorescente semi-quantitative	154
7. La technique de Southern blot	154
8. Obtention de l'ADNc du gène <i>MECP2</i>	156
8.1. Extraction d'ARN totaux et transcription inverse (RT)	156
8.2. PCR	156
9. Inactivation du chromosome X	156
II. Technique de cytogénétique	157
1. Préparation des chromosomes	157
2. Hybridation <i>in situ</i> en fluorescence	157
2.1. Prétraitement des chromosomes	157
2.2. Obtention de la sonde	158
2.3. Dénaturation des chromosomes et de la sonde	159
2.4. Hybridation	159
2.5. Lavages	159
2.6. Détection immunocytochimique	159
2.7. Contre-coloration et observation de la lame	160
BIBLIOGRAPHIE	161
PUBLICATIONS	183
Liste des publications	184
Résumé en anglais	185

Liste des figures

Figure 1 :	Cartographie des gènes de RMLX.	18
Figure 2 :	Cartographie du locus morbide impliqué dans le syndrome de Rett.	41
Figure 3 :	Complexe d'initiation de la transcription.	46
Figure 4 :	Facteurs et co-répresseurs associés aux HDAC1 et 2.	51
Figure 5 :	Les deux modèles possibles (A ou B) de l'effet de la désacétylation sur la chromatine.	52
Figure 6 :	Les ADN méthyltransférases.	56
Figure 7 :	Les différents exons sexe-spécifiques du gène <i>Dnmt1</i> .	56
Figure 8 :	La famille des protéines MBD.	61
Figure 9 :	Localisation et structure du gène <i>MECP2</i> sur le chromosome X.	64
Figure 10 :	Complexe de répression transcriptionnelle via MeCP2.	66
Figure 11 :	Position des 8 couples d'amorces utilisés pour la CSGE et le séquençage systématique, sur le gène <i>MECP2</i> .	72
Figure 12 :	Principe de la CSGE.	74
Figure 13 :	Mutation 1454-1457del.	80
Figure 14 :	Séquences consensus des jonctions exon/intron et du site de branchement.	81
Figure 15 :	Caractérisation de l'effet de la mutation IVS3-2A>C sur la transcription du gène <i>MECP2</i> .	82
Figure 16 :	Interactions entre l'ADN, le TEAA et la matrice.	84
Figure 17 :	Chromatogramme théorique de DHPLC.	85
Figure 18 :	Position des 6 couples d'amorces utilisés pour la DHPLC sur le gène <i>MECP2</i> .	86
Figure 19 :	Mutations détectées par DHPLC.	87
Figure 20 :	Détection de mutations en mosaïque somatique par les techniques d'hétéroduplex.	91
Figure 21 :	Evaluation semi-quantitative du taux de mosaïcisme par PCR fluorescente.	93
Figure 22 :	Résultats de l'étude de l'inactivation du chromosome X pour la famille B.	95
Figure 23 :	Ségrégation de la substitution 1126C>T et étude du profil d'inactivation.	96
Figure 24 :	Sonde utilisée pour le Southern blot.	97
Figure 25 :	Réarrangements au sein du gène <i>MECP2</i> détectés par Southern blot.	99
Figure 26 :	Caractérisation des différents remaniements détectés par Southern blot.	101
Figure 27 :	Localisation génomique des différents remaniements identifiés par Southern blot.	102
Figure 28 :	Séquence analysée dans la région 3'UTR.	104
Figure 29 :	Recherche de remaniements du gène <i>MECP2</i> par FISH.	105
Figure 30 :	Les différentes mutations identifiées dans le gène <i>MECP2</i> au cours de ce travail.	108
Figure 31 :	Localisation sur le gène <i>MECP2</i> des mutations identifiées chez des garçons avec un retard mental.	120
Figure 32 :	Modèle de transmission familiale d'un second gène potentiel dans le syndrome de Rett.	127
Figure 33 :	Le système Cre-Lox.	137

Liste des tableaux

Tableau I :	Les principaux RMXS pour lesquels les gènes responsables sont identifiés.	20
Tableau II :	Les gènes des RMX.	21
Tableau III :	Syndrome de Rett : les critères nécessaires.	23
Tableau IV :	Syndrome de Rett : les critères d'appoint.	23
Tableau V :	Syndrome de Rett : les critères d'exclusion.	24
Tableau VI :	Syndrome de Rett : les quatre stades évolutifs.	25
Tableau VII :	Quelques anomalies cytogénétiques décrites associées à un syndrome de Rett.	28
Tableau VIII :	Quelques syndromes dus à des mutations instables.	30
Tableau IX :	Les différents génotypes et phénotypes associés selon le modèle digénique de Bühler.	32
Tableau X :	Les trois classes d'ARN polymérase eucaryotes.	44
Tableau XI :	Quelques éléments régulateurs en cis et leurs facteurs de transcription associés.	45
Tableau XII :	Les familles HAT.	49
Tableau XIII :	Mutations (A) et polymorphismes (B) identifiés dans le gène <i>MECP2</i> par CSGE et séquençage.	77-78
Tableau XIV :	Mutations et polymorphismes identifiés dans le gène <i>MECP2</i> par DHPLC.	88
Tableau XV :	Les huit réarrangements identifiés par Southern blot.	100
Tableau XVI :	Mutations identifiées dans le gène <i>MECP2</i> chez des patientes atteintes d'un syndrome de Rett et publiées dans la littérature.	109-112
Tableau XVII :	Les différents modèles murins du syndrome de Rett.	139
Tableau XVIII :	Séquences des neuf couples d'amorces et paramètres d'amplification permettant le criblage du gène <i>MECP2</i> par CSGE	150
Tableau XIX :	Les six couples d'amorces utilisés pour le criblage du gène <i>MECP2</i> par DHPLC.	151
Tableau XX :	Autres amorces utilisées	151
Tableau XXI :	Conditions de criblage du gène <i>MECP2</i> par DHPLC.	153

Abréviations

3'UTR : 3' untranslated region
Ac : groupement acétyl
AdML : adenovirus major late promotor
ADN : acide désoxyribonucléique
ADNc : ADN complémentaire
AMPc : adénosine monophosphate cyclique
ARN : acide ribonucléique
ARNm : ARN messenger
ARNr : ARN ribosomique
ARNt : ARN de transfert
ATP : adénosine triphosphate
ATR-X : alpha-thalassémie avec retard mental lié au chromosome X
°C : degré celsius
Cellules ES : cellules souches embryonnaires
CEP X : sonde centromérique du chromosome X
cpm : coup par minute
CRE : cyclic AMP response element
CREB : cyclic AMP response element binding
CSGE : conformation sensitive gel electrophoresis
CTP : cytidine triphosphate
DAPI : 4', 6-diamidino-2-phenylindole
ddNTP : didésoxynucléoside triphosphate
del : délétion
derX : dérivé du chromosome X
DHPLC : denaturing high performance liquid chromatography
DNMT : DNA methyltransferase
dNTP : désoxynucléoside triphosphate
DO : densité optique
dup : duplication
EDTA : éthylènediaminetétraacétate
EST : expressed sequence tag
FISH : fluorescence *in situ* hybridization
GFP : green fluorescent protein
GTP : guanosine triphosphate
HAT : histone-acétyl transférase
HDAC : histone désacétylase
Hz : Hertz
ICF syndrome : immunodeficiency, centromere instability and facial anomalies syndrome
K : lysine
kb : kilobase
KCl : potassium chlorure
KH₂PO₄ : potassium dihydrogenophosphate
MAR/SAR : matrix associated regions / scaffold associated regions
Mb : mégabase
MBD : methyl binding domain
MeCP2 : methyl CpG-binding protein 2
mg : milligramme
MgCl₂ : magnesium chlorure
min : minute
ml : millilitre
mM : millimolaire
Na₂HPO₄ : di-sodium hydrogenophosphate
NaCl : sodium chloride
NLS : nuclear localisation signal

NMD : nonsense mediated RNA decay
p : bras court d'un chromosome
PAC : P1 artificial chromosome
pb : paire de bases
PBS : phosphate buffered saline
PCR : polymerase chain reaction
PCR-ASO : PCR-allele specific oligonucleotides
q : bras long d'un chromosome
RMLX : retard mental lié au chromosome X
RMX : retard mental lié au chromosome X non spécifique
RMXS : retard mental lié au chromosome X spécifique ou syndromique
RT-PCR : reverse transcriptase polymerase chain reaction
SDS : sodium dodecyl sulfate
sec : seconde
snARN : small nuclear ARN
SSC : solution saline citrate
t(X;A) : translocation entre un chromosome X et un autosome
TAE : tampon Tris-acétate EDTA
TE : tampon Tris-EDTA
TEMED : N, N, N', N'-tetramethylethylenediamine
TF : transcription factor
Tm : melting temperature
TPA : tetradecanoyl-phorbol-acetate
TRD : transcriptional repressor domain
TRE : élément de réponse au TPA
Tris-HCl : tris(hydroxyméthyl)-aminométhane chlorhydrate
TSA : trichostatine
U : unité d'enzyme
U.V. : ultra-violets
UTP : uridine triphosphate
V : volt
Xce : X-controlling element
Xic : centre d'inactivation du chromosome X
Xtel : télomère du chromosome X
YAC : yeast artificial chromosome
µg : microgramme
µl : microlitre
µM : micromolaire
θ : fréquence de recombinaison

PREAMBULE

La méthylation des cytosines et l'acétylation/désacétylation des histones constituent deux modifications épigénétiques importantes du génome eucaryote. Elles sont, toutes les deux, impliquées dans la répression transcriptionnelle. La découverte de la famille protéique MeCP (Methyl-CpG-binding protein) et notamment de la protéine MeCP2, archétype de cette famille, a permis d'établir un lien entre ces deux modifications épigénétiques. En effet, la protéine MeCP2, une fois liée aux cytosines méthylées des dinucléotides 5'CpG3' réprime indirectement la transcription par l'intermédiaire d'un complexe multiprotéique comprenant en plus du co-répresseur Sin-3A, les histones désacétylases 1 et 2 (HDAC1 et 2). Les HDAC1 et 2, en désacétylant les histones du nucléosome entraînent la condensation de la chromatine. Récemment, le gène *MECP2* codant la protéine MeCP2 a été impliqué dans un retard mental syndromique lié au chromosome X : le syndrome de Rett. Le syndrome de Rett, pathologie rare (1 fille sur 10 000 à 15 000) se caractérise par une période périnatale normale (jusqu'à 6-18 mois) avant l'apparition des premiers symptômes correspondant à une régression des acquisitions psychomotrices.

Nous avons recherché les mutations dans le gène *MECP2* chez 255 patientes atteintes d'un syndrome de Rett et chez 15 garçons retardés mentaux. Nous avons mis en évidence un spectre mutationnel hétérogène du gène *MECP2* chez les patientes atteintes d'un syndrome de Rett. En effet, en plus de mutations ponctuelles, le gène *MECP2* peut être remanié sur plusieurs centaines de paires de bases. D'autre part, nous rapportons, pour la première fois chez des filles, l'existence de mutations en mosaïque somatique.

INTRODUCTION

A. Les retards mentaux

I. Définition

Le retard mental est défini par l'Association Américaine de Psychiatrie (American Psychiatric Association, 1994) comme "un fonctionnement intellectuel général significativement inférieur à la moyenne, qui s'accompagne de limitations manifestes du fonctionnement adaptatif dans les secteurs d'aptitude, tels la communication, l'apprentissage scolaire, la vie sociale, la responsabilité individuelle, le travail, le loisir, la santé et la sécurité. Le tout doit survenir avant l'âge de 18 ans". Le diagnostic du retard mental et sa sévérité sont évalués par la mesure du niveau des performances intellectuelles grâce à des tests psychométriques standardisés. Le résultat de ces tests est exprimé en quotient intellectuel (QI) chez l'adulte et en quotient de développement (QD) chez l'enfant. Cent est la moyenne générale du QI, et un écart de 15 correspond à une déviation standard (DS). Une personne est considérée comme atteinte d'un retard mental à partir d'un QI inférieur à 70, c'est-à-dire à partir de -2 DS. On distingue les retards mentaux légers qui sont définis par un QI entre 50 et 70 et qui affectent entre 2 à 3% de la population, des retards mentaux sévères caractérisés par un QI < 50 qui touchent 0,3% des individus (Chelly 1999). La cause (souffrance fœtale, embryopathie infectieuse, syndrome malformatif, aberration chromosomique) est identifiée dans la moitié des retards mentaux légers et dans trois-quart des retards mentaux sévères (Chelly 2000).

II. Les retard mentaux liés au chromosome X (RMLX)

1. Définitions

Diverses études épidémiologiques ont montré que la fréquence des retards mentaux inexplicables est plus élevée chez les garçons avec un sex-ratio variant de 1,3 à 1,9 selon les séries (Lehrke 1972). Ce phénomène est actuellement interprété comme étant dû à des mutations récessives touchant des gènes localisés sur le chromosome X dont l'effet est plus apparent chez les garçons du fait de leur hémizygotie. Herbst et Miller (Herbst et al. 1980) ont

estimé la fréquence des RMLX à 1,8 pour 1000 naissances masculines, soit environ 16% de l'ensemble des retards mentaux (Chiurazzi et al. 2001). Bien que les retards mentaux liés au chromosome X représentent un ensemble de maladies très hétérogène, on peut les répartir en deux grands groupes :

-Les retards mentaux spécifiques ou syndromiques (RMS) qui associent un retard mental à des anomalies neurologiques, morphologiques, viscérales, radiologiques ou biochimiques caractéristiques. Ces retards mentaux syndromiques peuvent être subdivisés arbitrairement en quatre classes : 1) les retards mentaux associés à des anomalies congénitales (syndrome de l'X fragile impliqué dans 15 à 20% des RMLX, par exemple), 2) les maladies dans lesquelles le symptôme majeur peut être l'atteinte cérébrale et/ou musculaire (myopathie de Duchenne), 3) les retards mentaux dus à des anomalies du métabolisme (déficit en ornithine transcarbamylase), et 4) les syndromes supposés dominants létaux chez les hémizygotés et affectant les filles (syndrome d'Aicardi).

-Les retards mentaux non spécifiques (RMX) définis par un retard mental non progressif isolé sans autre anomalie clinique.

2. Données génétiques des RMLX

Dans une compilation récente de la littérature, Chiurazzi (Chiurazzi et al. 2001) a répertorié quelques 202 retards mentaux liés au chromosome X dont 136 retards mentaux syndromiques et 66 retards mentaux non spécifiques. L'identification des gènes impliqués dans les RMS est facilitée par les caractéristiques cliniques, radiologiques ou biochimiques spécifiques des patients qui permettent de regrouper plusieurs familles et de définir ainsi des intervalles génétiques réduits par des analyses de liaisons génétiques. Sur les 136 RMS potentiels dénombrés, 26 gènes responsables sont aujourd'hui identifiés (figure 1), 58 autres locus sont cartographiés. Dans le cas des RMX, les résultats des études de liaisons génétiques ne peuvent être additionnés. Ainsi, un intervalle génétique, le plus souvent de grande taille, est défini pour chacune des 66 familles présentant un RMX. Les 66 intervalles de liaison délimités sont partiellement chevauchants, ils peuvent être regroupés en 10-12 régions non chevauchantes (Chiurazzi et al. 2001 ; Chelly 1999). Ces données préliminaires indiquent qu'il existe au minimum 10 à 12 gènes différents impliqués dans les RMX dont 8 ont été identifiés actuellement (figure 1)(Castellvi-Bel et al. 2001).

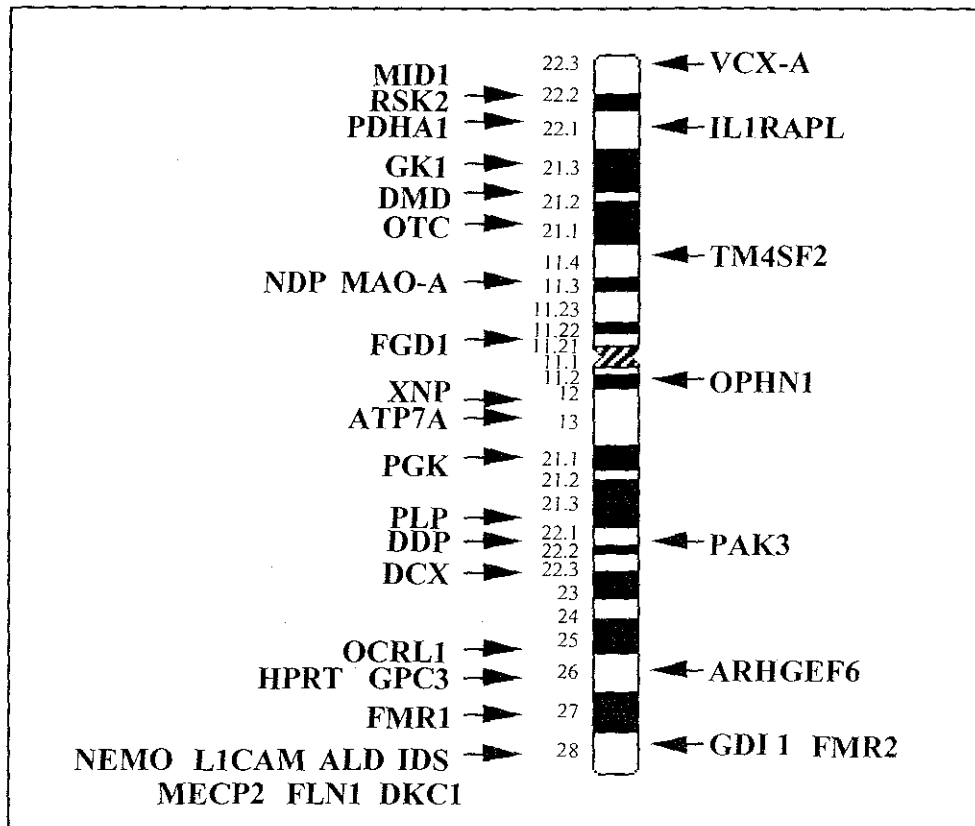


Figure 1 : Cartographie des gènes de RMLX

A gauche de l'idéogramme du chromosome X (bandes G) sont indiqués les 26 gènes clonés des RMSX. A droite, les 8 gènes des RMX.

3. Mécanismes physiopathologiques impliqués dans les RMLX

Etant donné l'extrême complexité de l'ontogenèse et du fonctionnement du système nerveux central, la diversité des gènes potentiellement impliqués dans les retards mentaux n'est pas une surprise, et dans l'immense majorité des cas, on est loin de comprendre le lien entre la présence d'un gène défectueux et la déficience cognitive observée. Cependant, on distingue deux grands types de mécanismes physiopathologiques selon le type de retard mental. Il semble, en effet, que dans les retards mentaux syndromiques, les gènes mis en cause ont le plus souvent de nombreux effets pléiotropes sur des mécanismes généraux, comme la régulation de la transcription, pouvant ainsi expliquer la présence de nombreux signes extraneurologiques. Par exemple, la protéine XH2 ou XNP (ADN-hélicase et ATPase ADN dépendante) est un facteur de transcription qui régule l'expression de plusieurs gènes par modification de la conformation chromatinienne et qui est responsable, lorsqu'elle est défectueuse, du syndrome ATR-X (Alpha-thalassémie avec retard mental lié au chromosome X) (Gibbons et al. 1995 ; Cardoso et al. 1998) (tableau I). A l'inverse, les produits des gènes

récemment impliqués dans les retards mentaux non spécifiques sont associés à des processus cellulaires ciblant *a priori* essentiellement la morphogenèse et la plasticité neuronale. Ces processus incluent l'organisation du cytosquelette (rhoGAP, PAK3, ARHGEF6, TMSF4), le recyclage des vésicules synaptiques (GDI1) ou la régulation de la synthèse d'un messager secondaire impliqué dans la plasticité synaptique ou neuronale (IL1RAPL) (Chelly 2000) (tableau II).

Tableau I : Les *principaux RMXS* pour lesquels les gènes responsables sont identifiés (d'après Chelly 2000).

Syndrome	Signes cardinaux	Gène	Fonction protéique	Pathogénie
Aarskog-Scott dysplasie faciogénitale	hypertélorisme, brachydactylie, scrotum en châte, petite taille	<i>FGD1</i>	Rho-GEF (règle cdc42) active la voie Ras-MAPkinase	troubles de la croissance cellulaire
ATR-X	a-thalassémie, hypogénitalisme	<i>XH2 ou XNP</i>	ADN hélicase, facteur de transcription modifie la structure chromatinienne	régulation de nombreux gènes
Coffin-Lowry	hypertélorisme, faciès grossier, phalanges en baguettes, petite taille	<i>RSK-2</i>	voie RAS-MAPkinase	activateur des facteurs de transcription c-fos, c-jun
Opitz G	fente labio-palatine, hypertélorisme, anomalies génito-urinaires	<i>MID-1</i>	doigt de zinc, régulateur transcriptionnel ?	troubles de fusion de la ligne médiane
X fragile (FRAXA)	faciès allongé, grandes oreilles, macro-orchidie	<i>FMRI</i>	protéine se liant aux ARNm, impliquée dans leur transport	anomalies de structure des épines dendritiques ?
CRASH	hypoplasie du Corps calleux, Retard mental, Adduction des pouces, paraplégie Spastique, Hydrocéphalie	<i>LICAM</i>	molécule d'adhérence	troubles de la migration
X-LIS/HLSC	hétérotopies laminaires, lissencéphalie, épilepsie	<i>Doublecortine</i>	liaison au cytosquelette	trouble de la migration des neurones
Lowe, syndrome oculo-cérébro-rénal	cataracte, hypotonie sévère, rachitisme vitamino résistant, tubulopathie	<i>OCRL-1</i>	transport vésiculaire dans le Golgi, assemblage du cytosquelette	troubles de la migration neuronale, et de la différenciation épithéliale

Tableau II : Les gènes des RMX (d'après Chelly 2000 ; Chiurazzi et al. 2001 ; Fukami et al. 2000).

Gène	Fonction protéique	Pathogénie
<i>VCX-A</i>	inconnue	inconnue
<i>ILIRAPL</i>	protéine homologue des protéines accessoires des récepteurs de l'IL-1	plasticité synaptique
<i>TM4SF2</i>	protéine à 4 domaines transmembranaires, forme un complexe avec les intégrines	organisation du cytosquelette, signalisation intracellulaire
<i>OPHN-1</i>	activateur des Rho-GTPases Rho-GAP (GTPase activating protein)	organisation du cytosquelette, signalisation intracellulaire
<i>PAK3</i>	effecteur des Rho-GTPases kinase (p21-activated protein)	organisation du cytosquelette, signalisation intracellulaire
<i>ARHGEF6</i>	protéine Rho GEF, activateur des protéines Rho en échangeant le GDP lié par du GTP	organisation du cytosquelette, signalisation intracellulaire
<i>FMR2</i>	facteur de transcription	inconnue
<i>RabGDI</i>	régulateur des Rab3-GTPase (Rab GDP-dissociation inhibitor)	transport vésiculaire, croissance neuritique

B. Le syndrome de Rett

I. Historique et clinique

1. Historique

C'est en 1966 qu'Andréas Rett, neuropédiatre autrichien de l'université de Vienne, décrit pour la première fois en langue allemande un syndrome caractérisé par une atrophie cérébrale et une hyperammoniémie (Rett, 1966). Une traduction anglaise de cette publication dans le "Handbook of Clinical Neurology" en 1977 n'empêcha pas que ce syndrome reste longtemps méconnu voire même oublié, avant d'être de nouveau découvert dans différents pays (Suède, Portugal, Japon, France...) au cours des années 70-80. En 1983, Hagberg publie en collaboration avec Aicardi, Dias et Ramos, une série de 35 patientes atteintes du syndrome dit "**syndrome de Rett**", car toutes présentent des caractéristiques initialement décrites par le neuropédiatre autrichien (Hagberg et al. 1983).

2. Clinique

Ce syndrome défini comme le résultat d'un trouble global du développement cérébral, est caractérisé par un développement initial normal suivi d'une perte progressive des fonctions intellectuelles et des différentes acquisitions motrices ; cette perte est accompagnée d'une diminution du périmètre crânien, du développement de mouvements stéréotypés des mains et d'une apraxie de la marche.

2.1. Les critères de diagnostic

Jusqu'en octobre 1999, aucun marqueur biologique n'était disponible, de sorte que le diagnostic restait purement clinique. Des critères précis ont été établis initialement lors de la IIème Conférence Internationale sur le Syndrome de Rett de 1984 (Hagberg et al. 1985), puis ceux-ci ont été modifiés et complétés en 1988 lors de la Vème Conférence Internationale sur le syndrome de Rett (Trevathen et al.1988).

Ainsi le syndrome de Rett est défini selon 9 critères nécessaires (tableau III), 8 critères

d'appoint (tableau IV) et 7 critères d'exclusion qui permettent d'éliminer les maladies métaboliques, les maladies progressives de surcharge, avec notamment atteinte de la rétine et les maladies à début anténatal ou périnatal (tableau V) (Echenne et al. 1989).

Tableau III : Syndrome de Rett : les critères nécessaires

- Période prénatale et périnatale normale
- Développement psychomoteur normal dans les 6 premiers mois de vie
- Périmètre crânien normal à la naissance
- Ralentissement de la vitesse de croissance du périmètre crânien entre 5 mois et 4 ans
- Pertes de praxies manuelles entre 6 et 30 mois, temporairement associées à des troubles de la communication et des conduites sociales
- Troubles graves du développement du langage, dans son versant réceptif, comme dans son aspect expressif, avec retard psychomoteur sévère
- Mouvements stéréotypés et très polymorphes des mains, apparaissant une fois l'apraxie manuelle installée
- Apraxie et ataxie du tronc et de la marche apparues entre 1 et 4 ans
- Hésitations diagnostiques entre 2 et 5 ans

Tableau IV : Syndrome de Rett : les critères d'appoint

- Troubles respiratoires :
apnées périodiques au cours de la veille,
hyperventilation intermittente,
expulsion forcée d'air et de salive.
- Anomalies électroencéphalographiques :
ralentissement de l'activité de fond à la veille,
ondes lentes intermittentes rythmiques (3-5 Hz),
décharges épileptiques, avec ou sans crise clinique.
- Epilepsie
- Spasticité, associée souvent à une amyotrophie et à des dystonies
- Troubles vasomoteurs périphériques
- Scoliose
- Retard statural
- Hypotrophie des pieds

Tableau V : Syndrome de Rett : les critères d'exclusion

<ul style="list-style-type: none">-Présence d'un retard de croissance intra-utérin-Hypertrophie d'organe ou signe de surcharge viscérale-Rétinopathie ou atrophie optique-Microcéphalie à la naissance-Signes de souffrance cérébrale néonatale-Maladie métabolique ou autre affection neurologique progressive identifiable-Troubles neurologiques acquis d'origine infectieuse ou traumatique

2.2. L'histoire naturelle

Le syndrome de Rett évolue selon une séquence de 4 stades cliniques définis par Hagberg et Witt-Engerstrom (Hagberg et al. 1986), qui apparaissent après un **intervalle libre normal** de 6 à 18 mois (tableau VI).

Le syndrome de Rett débute par **une phase de stagnation (Stade I)** détectable par un ralentissement de nouvelles acquisitions et la manifestation de traits autistiques.

Au cours du **Stade II**, est notée **une régression rapide** des fonctions cognitives et motrices, la marche devient apraxique avec une ataxie du tronc, se développent des comportements peu usuels, telles les stéréotypies manuelles ou oro-manuelles orientées selon la ligne médiane du corps devant la poitrine. Ces stéréotypies se manifestent par des frottements, des claquements, des battements et des mouvements répétitifs de flexion/extension des doigts sur la paume.

Suit **une phase pseudo-stationnaire (Stade III)**, au cours de laquelle on peut noter une atténuation des traits autistiques et une amélioration du contact principalement visuel de l'enfant avec son environnement malgré l'absence de développement du langage. L'enfant est capable de quelques acquisitions. Des accès d'angoisse et de tachypnée sont fréquents, entrecoupés de périodes d'apnée toujours diurnes. Des crises d'épilepsie de tous types apparaissent dans 94% des cas (Steffenburg et al. 2001).

La maladie atteint un plateau, voire diminue au stade IV dans la plupart des cas en ce qui concerne les crises convulsives et les irrégularités respiratoires. Le **Stade IV** correspond à **une phase de détérioration motrice** pouvant entraîner une perte complète de l'autonomie motrice. Il n'y a pas d'aggravation du retard mental. La scoliose se développe chez au moins la moitié des sujets. Plus tard, le tableau neurologique peut évoquer un syndrome parkinsonien. La croissance staturo-pondérale est ralentie, la puberté survient à l'âge normal.

Tableau VI : Syndrome de Rett : les quatre stades évolutifs

I. Phase de stagnation
Début : 6 à 18 mois
-Stagnation des performances motrices, -Diminution de l'intérêt pour les jeux, -Hypotonie, -Ralentissement de la croissance céphalique.
Durée : plusieurs mois.
II. Phase de régression rapide
Début : 1 à 3 ans
-Régression rapide du comportement, -Perte de l'usage des mains, -Crises convulsives, -Stéréotypies manuelles, -Manifestations autistiques, -Troubles ventilatoires, -Perte du langage, -Motricité maladroite, -Insomnie, -Automutilations.
Durée : quelques semaines à quelques mois.
III. Phase de stabilisation apparente
Début : 2 à 10 ans
-Retard mental sévère, -Régression des traits autistiques, -Amélioration du contact, -Crises convulsives, -Stéréotypies manuelles caractéristiques, -Ataxie-apraxie et spasticité, -Dysfonctionnement respiratoire, -Perte de poids avec conservation de l'appétit, -Bruxisme.
Durée : quelques mois à des années.
IV. Phase de détérioration motrice tardive
Début : après 10 ans
-Perte de la motricité, -Scoliose, atrophie musculaire, hypertonie, -Syndromes pyramidal et extra-pyramidal marqués, -Retard de croissance, -Absence de langage, -Amélioration du contact visuel, -Epilepsie moins sévère, -Troubles trophiques des extrémités (pieds).
Durée : plusieurs années.

La durée de vie des patientes peut varier entre 2 ans et l'âge adulte. La survie est réduite de 75% à l'âge de 35 ans par rapport à la population générale (Ellaway et al. 1999). Un tiers des décès se produit au cours des 20 premières années de vie, principalement pendant le sommeil, sans relation apparente avec les crises convulsives, la cause du décès reste inexpliquée. Un dysfonctionnement cardiaque correspondant à une bradycardie avec intervalle QT corrigé long pourrait être à l'origine de ces morts subites (Guideri et al. 1999).

2.3. Les variants

Toute petite fille présentant **tous** les critères nécessaires et **aucun** des critères d'exclusion, avec ou sans critères d'appoint est atteinte d'un **syndrome de Rett dit classique ou typique**. L'absence d'un seul critère nécessaire, même en présence de tous les critères d'appoint, exclut le diagnostic.

Toutefois, avec le recul et l'expérience, il est clairement apparu que certaines petites filles présentaient des symptômes du syndrome de Rett dont l'apparition et/ou la sévérité ne correspondaient pas tout à fait à celles établies initialement. Entre 1993-1995, Hagberg et al. (Hagberg et al. 1994) définirent 5 variants possibles du syndrome de Rett permettant ainsi de circonscrire la symptomatologie extrêmement variable de cette pathologie.

Le premier variant défini comme la **forme atypique** du syndrome de Rett, correspond à un phénotype assez voisin de la forme classique, cependant il est dominé par des crises convulsives sévères et précoces et la maladie débute avant les 6 premiers mois de vie de l'enfant.

Le deuxième variant appelé forme **congénitale ou précoce** doit son nom à une absence d'intervalle libre avant l'apparition de la phase de régression.

A l'opposé, le troisième variant correspond à des patientes dont la phase de régression est plus tardive.

Le quatrième variant étiqueté **forme frustré** présente les 4 stades évolutifs aux mêmes périodes de la vie mais avec une sévérité moins marquée. Ainsi, l'usage des mains par exemple est préservé et les stéréotypies sont minimales.

Enfin, le cinquième variant défini par Hagberg correspond à des patientes ayant conservé le langage.

Il aura fallu quasiment 30 ans, depuis la description princeps d'Andréas Rett en 1966, jusqu'à la définition de variants en 1995 par Hagberg, en passant par l'établissement de critères internationaux de diagnostic stricts pour définir cliniquement le syndrome de Rett.

Cette maladie est reconnue maladie pan-éthnique avec une prévalence exclusivement féminine de 1 fillette sur 10000-15000 naissances.

II. Etiologie du syndrome de Rett

1. Syndrome de Rett, maladie génétique

1.1. Quelques cas familiaux

L'étiologie du syndrome de Rett est restée jusqu'à récemment (octobre 1999) inconnue, mais l'hypothèse d'une origine génétique est apparue assez rapidement comme hautement probable et a fait l'objet de nombreuses discussions.

Bien que la majorité des cas, c'est-à-dire plus de 99%, soit sporadique, la description de premiers cas familiaux (Murphy et al. 1986), de jumeaux monozygotes concordants (Tariverdian et al. 1987 ; Partington 1988), de gémellité discordante (Migeon et al. 1995) et d'un cas de transmission d'une mère atteinte à sa fille (Witt-Engerström et al. 1992) plaident dès 1986 en faveur d'une origine génétique.

1.2. Les anomalies cytogénétiques associées au syndrome de Rett

Une dizaine d'anomalies cytogénétiques associées à une forme plus ou moins typique de syndrome de Rett a été décrite dans la littérature. Quelques unes de ces anomalies référencées dans le tableau VII, impliquent des autosomes différents ne permettant donc pas de cibler un chromosome en particulier. Deux translocations entre un chromosome X et un autosome associées au syndrome de Rett ont été par ailleurs rapportées : $t(X;3)(p21.3;q13.31)$ (Zoghbi et al. 1990-a), et $t(X;22)(p11.2;p11)$ (Journel et al. 1990).

Tableau VII : Quelques anomalies cytogénétiques décrites associées à un syndrome de Rett

	<i>de novo</i>		familiale
	constitutionnelle	mosaïque	constitutionnelle
délétions	del(3)(p25.1p25.2) Wahlström et al. 1999	del(18)(q22.3) Gordon et al. 1993	del(18)(q21.1q22.3) Gustavsson et al. 1999
duplications	dup(6)(p11.2) Wahlström et al. 1986 dup(11)(q23.3) Delobel et al. 1998		
translocations	t(X;3)(p21.3;q13.31) (Zoghbi et al. 1990-a)		t(X;22)(p11.2;p11) (Journel et al. 1990)

Plusieurs sites chromosomiques fragiles ont été mis en évidence sur certains chromosomes humains. Ces sites correspondent à des cassures des chromosomes visualisables dans des conditions particulières de culture (concentrations faibles en acide folique et en thymine, ou culture en présence de méthotrexate ou de fluorodésoxyuridine). Les bases moléculaires et la signification de ces sites restent une énigme. On pense qu'ils représenteraient une visualisation cytogénétique d'une activité génique (Austin et al. 1992 ; Hecht et al. 1991). Ainsi, une tendance à la cassure chromosomique chez les patientes atteintes du syndrome de Rett constituerait d'une part un critère de diagnostic biologique et d'autre part une piste en faveur de l'origine génétique de la maladie. Plusieurs recherches de ces sites ont été effectuées. Certaines ont mis en évidence une relation entre le syndrome de Rett et le site fragile en Xp22 (Wahlström 1985, 1986, 1990 ; Gillberg et al. 1984, 1985). Cependant, d'autres études (Moore et al. 1986 ; Romeo et al. 1986) ont contredit l'existence de cette relation. Une dernière étude en 1994 (Telvi et al. 1994) a recherché tous les sites potentiels de cassure et a montré que l'ensemble des chromosomes des patientes était plus fragile par rapport à ceux d'une population témoin. Les auteurs concluent quant à la potentialité d'un facteur génétique qui fragiliserait les chromosomes des patientes atteintes du syndrome de Rett.

1.3. Quel mode de transmission ?

Avant la découverte du gène coupable, si l'origine génétique semblait quasiment acquise, le mode de transmission de la maladie ne paraissait pas aussi évident. Plusieurs modèles ont été proposés. Beaucoup sont basés sur un mode de transmission dominant. De plus, les modèles diffèrent sur l'origine paternelle ou maternelle de la mutation, ainsi que sur le mode de transmission autosomique et/ou gonosomique voire mitochondrial.

1.3.1. Mise en évidence de mutations au sein de l'ADN mitochondrial

A partir de 1989, plusieurs équipes ont mis en évidence, chez certaines patientes présentant un syndrome de Rett, des mitochondries déficientes aussi bien au niveau de leur structure que de leur fonction (Ruch et al. 1989 ; Dotti et al. 1993). Plusieurs recherches de mutations dans le génome mitochondrial ont révélé l'existence de différentes mutations par substitution de bases. Alors que la transition T10463C localisée dans la région codant l'ARN de transfert de l'arginine (ARNt^{Arg}) n'a été retrouvée qu'une seule fois (Lewis et al. 1995, 1998), deux autres mutations putatives ont été retrouvées de façon récurrente et localisées dans la même région 2650-3000 codant l'ARN 16S. Ainsi la mutation A2706G a été retrouvée chez 14 patientes et chez 12 de leur mère, et la mutation C2835T 6 fois et chez 5 des mères (Tang et al. 1997). Toutefois, deux études plus récentes, une en Italie (Cardaioli et al. 1999), et une en Espagne (Armstrong et al. 2000) n'ont pas retrouvé la mutation C2835T ni aucune autre sur une cohorte de 53 patientes atteintes d'un syndrome de Rett.

1.3.2. Recherche d'expansion de triplets

Au début des années 90, un nouveau type de mutations a été mis en évidence dans la maladie de l'X-fragile (Oberlé et al. 1991 ; Verkerk et al. 1991), puis a été retrouvé dans une quinzaine de maladies. Ces nouvelles mutations qualifiées de mutations instables correspondent à une expansion de triplets nucléotidiques localisée soit dans les régions 5' ou 3' non traduites des gènes, soit dans la région codante, soit dans un intron (tableau VIII). L'effet des ces amplifications de triplets sur l'expression des gènes est variable. Elles peuvent entraîner une inactivation du gène. C'est le cas du gène *FMRI*, responsable du syndrome de l'X-fragile, qui dans les cas de mutations complètes (à partir de 200-230 CGG) est inactivé du fait de la

méthylation des triplets localisés dans la région 5' du gène. A l'inverse, ces expansions peuvent créer un gain de fonction, comme pour la maladie de Huntington dans laquelle l'amplification de CAG confère à la protéine un effet toxique de part la présence d'une polyglutamine. Il existe des maladies qui sont dues non pas à des répétitions de triplets mais de pentanucléotides dans le cas de l'ataxie spino-cérébelleuse de type 10 (Matsuura et al. 2000) ou de dodécanucléotides pour l'épilepsie myoclonique progressive (Lalioi et al. 1997).

Tableau VIII : Quelques syndromes dus à des mutations instables

Maladie	Mode de transmission	Locus	Gène	Trinucléotide	Localisation
X fragile	dominant lié à l'X	Xq27.3 (FRAXA)	<i>FMRI</i>	CGG	exon en 5' non traduit
Chorée de Huntington	dominant autosomique	4p16.3	<i>IT15 (HD)</i>	CAG	région codante
Dystrophie myotonique	dominant autosomique	19q13.3	<i>DMPK</i>	CTG	exon en 3' non traduit
Ataxie de Friedreich	récessif autosomique	9q13	<i>X25</i>	GAA	intron 1

Pour ces mutations instables, le nombre de répétitions présente la particularité d'augmenter au cours des générations successives, aggravant ainsi les risques de développer la maladie, la sévérité et la précocité des manifestations cliniques (phénomène d'anticipation).

Des expansions de triplets ont été recherchées dans des cas familiaux de syndrome de Rett présentant une aggravation du phénotype au cours des générations pouvant faire penser au phénomène d'anticipation avec des individus prémutés atteints de formes frustrées. C'est l'équipe de Uta Francke qui réalisa cette analyse en 1997 en recherchant de manière systématique dans tout le génome, les répétitions de dix triplets possibles par la technique de RED (repeat expansion detection) et exclut le rôle d'une expansion importante de triplets dans la pathogenèse du syndrome de Rett (Hofferbert et al. 1997).

1.3.3. Hypothèse d'une transmission dominante liée au chromosome X d'origine maternelle

Le syndrome de Rett est une affection touchant quasi exclusivement les filles. L'hypothèse d'une transmission dominante liée au chromosome X d'origine maternelle a été émise afin

d'expliquer l'atteinte exclusive des filles par un effet létal chez les garçons. Ce mode de transmission expliquerait la majorité des cas sporadiques par une néomutation, et les quelques cas familiaux par l'occurrence nécessaire de deux événements : la mutation, et une inactivation préférentielle du chromosome X muté ou un mosaïcisme germlinal permettant d'expliquer l'absence de phénotype chez la mère vectrice et donc la transmission de la maladie.

Toutefois, l'absence de biais dans le sex-ratio en faveur des filles que ce soit dans les fratries des cas sporadiques (Killian 1986), ou au cours des générations dans les cas familiaux (Wahlstrom et al. 1986) va à l'encontre de ce modèle.

1.3.4. Hypothèse d'une transmission dominante liée au chromosome X d'origine paternelle

Afin de pallier le problème dû à l'absence de biais dans le sex-ratio, Thomas a proposé en 1996 un modèle de transmission dominante liée au chromosome X de néomutations d'origine paternelle. Il s'est appuyé sur les exemples de plusieurs maladies dont l'origine de la mutation est paternelle (syndrome d'Apert, déficience en ornithine transcarbamylase, achondroplasie). Ce modèle expliquerait l'atteinte exclusive des filles car le père transmet son chromosome X à ses filles uniquement. Toutefois, la fréquence des néomutations ponctuelles semble augmenter avec l'âge chez les hommes et aucune élévation de l'âge de la procréation n'a été constatée chez les pères des petites filles atteintes d'un syndrome de Rett.

1.3.5. Hypothèse d'une maladie digénique impliquant un locus autosomique et un locus porté par le chromosome X

Pour ce modèle, Bühler (Bühler et al. 1990, 1999) s'est inspiré du modèle d'interférence métabolique non allélique décrit par Johnson au début des années 80 (Johnson, 1980). A partir de quatre observations sur le syndrome de Rett : i) consanguinité plus fréquente chez les parents de patientes, ii) l'hypothèse d'un mosaïcisme germlinal ne peut expliquer tous les cas familiaux décrits, iii) peu de renseignements sur les cas de fausses-couches dans les familles atteintes, iv) chez les familles présentant un (ou plusieurs) cas de syndrome de Rett, est noté un accroissement du nombre d'individus montrant des anomalies (comme un retard mental non spécifique et/ou des handicaps moteurs), Bühler propose l'implication de deux locus l'un lié au chromosome X (allèle normal X, allèle muté X*) l'autre à un autosome (A /A*), d'où le

nom du modèle d'interférence non allélique. L'interaction entre les produits de l'allèle sain A et l'allèle muté X* engendrerait le développement de la maladie chez les filles, les hommes AAX*Y seraient eux vecteurs de la maladie. L'association entre les produits des deux allèles mutés A* et X* causerait la létalité chez les hommes et les filles AA*XX* seraient vectrices (tableau IX).

Tableau IX : Les différents génotypes et phénotypes associés selon le modèle digénique de Bühler (Bühler et al. 1990 ; Bühler et al. 1999).

Fille		Garçons	
Génotypes	Phénotypes	Génotypes	Phénotypes
AAX*X	Syndrome de Rett	AAX*Y	Vecteur
AA*X*X	Vecteur (variant ?)	AA*X*Y	Létal
AA*XX	Normal	AA*XY	Normal
AAXX	Normal	AAXY	Normal
AA*X*X*	Létal	A*A*X*Y	Létal?
AAX*X*	Vecteur (variant ?)	A*A*XY	Normal
A*A*XX	Normal		
A*A*X*X	Vecteur?		

Ce modèle expliciterait les cas sporadiques et les cas familiaux. La majorité des mutations serait d'origine paternelle, ce qui expliquerait l'atteinte préférentielle des filles et la prépondérance des cas sporadiques car les filles atteintes ne se reproduisent que très rarement (un seul cas décrit dans la littérature par Witt-Engerstrom et al. 1992). Par ailleurs, les quelques cas familiaux seraient dus à une transmission d'origine maternelle beaucoup moins fréquente.

L'intervention du second locus autosomique modulerait le résultat d'une transmission dominante liée au chromosome X qui n'est pas applicable telle quelle au syndrome de Rett. L'association entre les produits des allèles mutés X* et A* sera létale chez le garçon et confèrera le statut de vectrice à un fille. Lorsque le produit de l'allèle X* interagit avec celui de l'allèle A, la fille sera atteinte du syndrome de Rett, le garçon sera vecteur. Ainsi une fille vectrice est responsable des cas familiaux. Par ailleurs, toutes les combinaisons ne sont pas létales chez un garçon (il peut aussi être vecteur) expliquant ainsi l'absence de biais dans le sex-ratio. De plus, les garçons vecteurs peuvent être responsables des quelques cas familiaux

à transmission paternelle décrits (Budden et al. 1986). Enfin, les différentes possibilités entre les deux locus donnant les individus vecteurs expliqueraient la variabilité phénotypique du syndrome de Rett.

1.4. Inactivation du chromosome X et syndrome de Rett

1.4.1. Quelques notions sur l'inactivation du chromosome X

La présence de deux chromosome X chez la femme et d'un seul chez l'homme nécessite l'existence d'un mécanisme de compensation de dosage. Le chromosome Y, propriété exclusive des hommes, ne contient qu'un nombre limité de gènes impliqués dans la détermination sexuelle, la spermatogenèse, ou correspondant à des homologues de gènes présent sur le chromosome X, tout ceci ne permettant pas de compenser l'absence d'un chromosome X. Cette compensation de dosage des gènes liés au chromosome X est réalisée précocément au cours de l'embryogenèse féminine par une inactivation transcriptionnelle aléatoire d'un des deux chromosomes X (d'origine maternelle ou paternelle) (Lyon 1961). Le mécanisme d'inactivation du chromosome X est supposé se dérouler en plusieurs étapes : i) comptage des chromosomes X présents dans la cellule, ii) choix du chromosome actif (Xa), iii) initiation de l'inactivation sur le futur chromosome inactif (Xi), expansion de l'inactivation, et iv) maintien de l'état inactif. Une fois qu'un chromosome X est inactivé dans une cellule, toutes les cellules filles auront le même chromosome X inactif. C'est un processus clonal. L'état inactif est transmis de manière épigénétique. Ainsi, toutes les filles sont porteuses d'une mosaïque constituée de deux populations cellulaires dont l'une exprime les allèles à partir d'un chromosome X, et l'autre population cellulaire à partir de l'autre chromosome X. Le chromosome X inactif présente toutes les caractéristiques de l'hétérochromatine constitutive. Il se réplique de façon tardive au cours du cycle cellulaire, il est hyperméthylé, sa chromatine est condensée, ses histones H4 hypoacétylées, et il y a une accumulation de macrohistone H2A1.2, enfin, il peut être visualisé à l'interphase sous la forme du corpuscule de Barr (revue Brown et al. 2000).

Une région particulière, nommée centre d'inactivation du chromosome X (Xic) localisée en Xq13.2 (Brown et al. 1991) est impliquée dans le mécanisme d'inactivation du chromosome X. Chez l'Homme, un gène localisé dans cette région : le gène *XIST* (X-Inactivation-Specific Transcript) présente la particularité de coder un ARN nucléaire non traduit *XIST* impliqué dans l'initiation de l'inactivation du chromosome X (Brown et al. 1992). Au moment de la

formation du zygote femelle, les deux chromosomes X sont actifs (Epstein et al. 1978) et expriment à faible quantité *XIST*. L'inactivation du chromosome X se déclenche juste avant la gastrulation (Kay et al. 1993). En effet, il est observé à ce moment-là, une augmentation de la quantité d'ARNm *XIST*, certainement par une stabilisation de l'ARNm plutôt que par une transcription accrue (Panning et al. 1997 ; Sheardown et al. 1997). L'ARNm *XIST* ainsi stabilisé recouvre le futur chromosome inactif et se propage en *cis* tout le long du chromosome entraînant ainsi son inactivation (Brown et al. 1992). Récemment chez la souris, un deuxième gène situé dans le Xic et impliqué dans les premières étapes de l'inactivation du chromosome X, a été découvert : le gène *Tsix* codant un ARN antisens du gène *Xist* (Lee et al. 1999). Ce gène semble agir en amont de *Xist* dans la cascade des événements conduisant à l'inactivation du chromosome X. En effet, il interviendrait dans le choix du chromosome X à inactiver (Mlynarczyk et al. 2000). Durant la différenciation, une diminution de la transcription du gène *Tsix* par l'action d'un répresseur putatif du gène (rep*Tsix*) serait corrélée à la stabilisation de l'ARNm *Xist* (Pereira et al. 2000). Ainsi, le niveau de l'ARNm *Xist* avant la différenciation cellulaire apparaît être régulé par l'ARN *Tsix*. La répression du gène *Tsix* serait alors la première étape du processus d'inactivation du chromosome X chez la souris (Lee et al. 1999 ; Debrand et al. 1999).

L'inactivation du chromosome X chez la fille est associée à une méthylation différentielle des îlots CpG entre le chromosome X actif et le chromosome X inactif. A l'exception de quelques locus (par exemple le locus DXS255), le chromosome X actif est hypométhylé et le chromosome X inactif hyperméthylé. Un petit groupe de gènes (15% de quelques 200 gènes du chromosome X étudiés (Carell et al. 1999)) échappent, toutefois, à cette inactivation et sont exprimés à partir des deux chromosomes X, c'est-à-dire à partir des chromosomes X actif et inactif. Ces gènes sont localisés soit dans une des deux régions pseudoautosomiques (PAR1 en Xp22.3 et PAR2 en Xq28) qui possèdent chacune leur région homologue sur le chromosome Y, soit en dehors de ces régions mais ces gènes ont alors un homologue, fonctionnel ou non, sur le chromosome Y (Brown et al. 1997).

L'inactivation du chromosome X est un phénomène aléatoire, cependant trois mécanismes peuvent expliquer une inactivation biaisée ou sélective.

Lors d'une inactivation aléatoire du chromosome X, les contributions maternelle et paternelle sont totalement équivalentes. Comme tout mécanisme aléatoire, celui-ci peut être représenté selon une distribution binomiale dans laquelle la grande majorité des femmes (80 à 90%) ont dans leurs cellules un chromosome X actif soit d'origine maternelle soit d'origine paternelle. Les extrémités de la courbe "gaussienne" correspondent aux femmes (10 à 20% de la

population générale) présentant par pur hasard une inactivation entièrement biaisée (95-5% ou 5-95%) en faveur de l'un ou l'autre chromosome X. Ainsi, une telle inactivation biaisée peut être à l'origine d'un "phénotype malade" chez une femme si le chromosome X porteur d'une mutation est préférentiellement conservé actif (Willard 1996).

Un deuxième mécanisme implique une contre-sélection des cellules qui montreraient un cycle cellulaire anormal si le chromosome X porteur d'une mutation était gardé actif. Une telle contre-sélection a été mise en évidence chez des femmes conductrices par exemple de l'ATR-X (Gibbons et al. 1992). Le produit du gène *XNP* responsable de cette maladie serait impliqué dans le remodelage de la chromatine et/ou l'activation/répression de locus spécifiques (McDowell et al. 1999). Ce mécanisme est rencontré également lors de translocations équilibrées entre un chromosome X et un autosome t(X;A). En effet, dans 77% des cas, le chromosome X normal est préférentiellement inactivé, ce qui confère aux cellules un patrimoine génétique proche de la normale à l'exception des deux points de cassure (Mattéi et al. 1982). Toutefois, il a été montré que dans les 23% restants impliquant souvent un point de cassure proche d'une des deux régions télomériques, c'est le dérivé du chromosome X qui est préférentiellement inactivé avec potentiellement une propagation de l'inactivation sur une partie du fragment de l'autosome adjacent. Dans ce cas-là, l'inactivation est à l'origine d'une disomie fonctionnelle partielle du chromosome X et d'une monosomie partielle de l'autosome entraînant de multiples anomalies congénitales et un retard mental (Schmidt et al. 1992).

Une inactivation biaisée peut également être due à une mutation dans un gène impliqué dans le processus de l'inactivation du chromosome X. A titre d'exemple, une mutation ponctuelle dans le promoteur du gène *XIST* a été associée à une inactivation biaisée dans deux familles différentes (Plenge et al. 1997).

Enfin, chez la souris, 20 à 30% des inactivations biaisées sont associés au locus *Xce* (X-controlling element) (Cattanach 1975). Les études cartographiques suggèrent que le locus est situé dans le *Xic* mais en aval du gène *Xist* (Simmler et al. 1993). Les différents allèles du gène *Xce* sont corrélés avec des niveaux variables d'expression du gène *Xist* (Brockdorff et al. 1991) et de méthylation de l'îlot CpG du locus *Tsix* (Avner et al. 1998). Il interviendrait ainsi dans le choix du chromosome à inactiver. Dans les cellules femelles hétérozygotes au locus *Xce*, le chromosome X qui porte un allèle fort *Xce* restera préférentiellement actif (Avner et al. 2001). L'homologue humain du gène *Xce* n'est pas connu, bien que des familles présentant un profil d'inactivation biaisée hérité, apparemment non porteuses d'une maladie liée au chromosome X, soient décrites (Busque et al. 1996 ; Naumova et al. 1998).

1.4.2. Etudes de l'inactivation du chromosome X dans le syndrome de Rett

Plusieurs études sur l'inactivation du chromosome X chez des patientes atteintes d'un syndrome de Rett ont été réalisées en utilisant soit des techniques de cytogénétique qui permettent de suivre et de différencier les cinétiques de réplication des deux chromosomes X (Kormann-Bortolotto et al. 1992 ; Panasiuk et al. 1997), soit des techniques de biologie moléculaire par étude du degré de méthylation des allèles paternel et maternel à certains locus (Carrel et al. 1996 ; Allen et al. 1992).

Bien que la plupart des approches par biologie moléculaire (Zoghbi et al. 1990-b ; Webb et al. 1993 ; Anvret et al. 1994) n'ait pas mis en évidence de biais dans l'inactivation, certaines (Camus et al. 1996 ; Kreprischi et al. 1998) ont montré l'existence d'une inactivation modérément biaisée du chromosome X d'origine paternelle (entre 65-35% et 80-20%) dans une partie des patientes (18/30 et 15/24, respectivement).

Les approches cytogénétiques ont montré que seule la bande Xp21.2 présentait une altération de sa cinétique de réplication dans quelques cellules de patientes. En effet, cette bande se réplique plus précocément chez les patientes que chez des témoins (Kormann-Bortolotto et al. 1992 ; Panasiuk et al. 1997). Cette chronologie incorrecte pourrait entraîner que le(s) gène(s) localisé(s) dans cette bande soit (soient) anormalement actif(s) pendant une période du cycle cellulaire, déséquilibrant ainsi le dosage génique et interférant possiblement avec le développement normal du cerveau. Ces résultats pouvaient constituer un point de départ quant à la recherche du gène responsable du syndrome de Rett. Toutefois, les analyses de liaison (développées ci-après) n'ont, par la suite, pas confirmé le lien de cette région avec le syndrome de Rett.

2. Syndrome de Rett, maladie génétique liée au chromosome X

Les nombreuses théories sur le mode de transmission possible du syndrome de Rett ont longtemps reflété la difficulté de compréhension de l'étiologie de cette maladie et ont ainsi freiné la recherche du gène responsable du syndrome de Rett.

Le clonage des gènes responsables de maladies est fondé depuis une quinzaine d'années, non plus sur l'étude des produits des gènes (clonage fonctionnel), mais sur l'étude du génome (clonage positionnel) du fait du développement de nouvelles approches moléculaires grâce à la mise en évidence d'un nombre presque illimité de séquences polymorphes transmises de façon mendélienne, à l'amélioration grandissante des techniques de biologie moléculaire (techniques de clonage de grands fragments, techniques de détection de mutations...).

Le clonage positionnel est une approche visant à identifier des gènes responsables de maladies génétiques dont la fonction et le produit sont inconnus, en se basant sur leur localisation chromosomique et en effectuant une analyse moléculaire de la région correspondante (Collins 1992). Il s'agit d'un processus composé de plusieurs étapes qui débute par l'assignation à une région chromosomique donnée du gène impliqué dans la maladie étudiée. Cette localisation préliminaire est suivie d'une analyse moléculaire de la région, consistant à établir le plus finement possible la localisation génétique, à développer la carte physique locale et à cloner l'ADN de la région. Il devient alors possible d'entreprendre une recherche des séquences exprimées dans l'intervalle critique aboutissant à l'identification d'un ou plusieurs gènes candidats. L'implication d'un de ces gènes dans la pathologie étudiée est définitivement démontrée par la mise en évidence de mutations délétères chez les patients.

La première étape du clonage positionnel, c'est-à-dire la localisation chromosomique du gène morbide, est réalisée par l'observation d'anomalies chromosomiques cytogénétiquement détectables associées à la maladie, telles que des délétions, des duplications, des translocations, des inversions ou des sites fragiles, et par des études de liaisons génétiques dans des familles où co-ségrègent le locus morbide et des allèles de marqueurs génétiques.

2.1. Deux translocations t(X;A) associées au syndrome de Rett

En 1990, deux translocations équilibrées entre un chromosome X et un autosome (Zoghbi et al. 1990-a ; Journel et al. 1990) ont été publiées chez des patientes présentant un syndrome de Rett. Malheureusement, les points de cassure de ces deux translocations étaient trop éloignés

pour délimiter un intervalle génique possible. En effet, la translocation t(X;3) *de novo* (Zoghbi et al. 1990-a) désigne un point de cassure en Xp21.3 et la translocation t(X;22) familiale (Journel et al. 1990) en Xp11.22 ; la distance physique entre les deux points de cassure sur le chromosome X est d'environ 20 Mb. Par ailleurs, comme nous le verrons ci-après, ces deux régions (en Xp21.3 et Xp11.22) seront exclues dès les premières analyses de liaison à l'aide des quelques cas familiaux de syndrome de Rett (Ellison et al. 1992).

2.2. Etudes de liaison génétique

2.2.1. Principe

En l'absence d'anomalies chromosomiques décelables, les gènes morbides peuvent être localisés par des études de liaison génétique dans les familles où existe une ségrégation du trait phénotypique avec un marqueur génétique. L'efficacité et la puissance de cette approche ont été significativement améliorées par le développement d'une carte génétique de très haute résolution, fondée sur le polymorphisme d'un grand nombre de locus microsatellites hautement informatifs et couvrant l'ensemble du génome humain (Weissenbach et al. 1992 ; Dib et al. 1996).

Les analyses de liaison génétique permettent de déterminer un intervalle de quelques centiMorgans (cM) en fonction du nombre de méïoses, c'est-à-dire de quelques millions de paires de bases, dans lequel il est hautement probable de retrouver le gène morbide. Il s'agit d'une distance génétique calculée à partir de méïoses informatives où l'on a constaté une co-ségrégation d'un allèle marqueur et de l'allèle morbide dans plusieurs familles atteintes de la même maladie. Le traitement statistique des analyses de liaison s'appelle le lod score (lod : logarithm of odds) décrit initialement par Morton (Morton 1955), puis adapté et amélioré par Lathrop et al. (Lathrop et al. 1984). La distance génétique entre le locus morbide et un marqueur donné est fonction de la fréquence de recombinaison θ qui varie de 0 (marqueurs liés) à 0,5 (marqueurs indépendants). Un procédé de calcul statistique permet d'évaluer le rapport de vraisemblance de deux hypothèses opposées : celle d'une liaison ($\theta < 0,5$) et celle d'une absence de liaison ($\theta = 0,5$), symbolisé respectivement par $L(\theta)$ et par $L(\theta_{0,5})$. Le rapport des vraisemblances pour une série de valeurs de θ permet de calculer la meilleure probabilité. Pour faciliter le calcul, on utilise le logarithme décimal du rapport des vraisemblances ou lod score symbolisé par $Z(\theta) = \log(L(\theta)/L(\theta_{0,5}))$. Le calcul de Z est effectué

pour toutes les valeurs de θ comprises entre 0 et 0,5. On estime qu'il faut un lod score $Z(\theta) \geq 3$ pour pouvoir affirmer que la liaison est réelle ($Z(\theta)=3$ signifie que l'hypothèse d'une liaison entre le gène morbide et le marqueur est 1000 fois plus probable qu'une absence de liaison). Un lod score inférieur à -2 permet au contraire d'exclure toute liaison génétique entre les deux locus considérés.

Cette analyse de liaison peut se révéler très efficace mais à plusieurs conditions : i) étudier un nombre élevé de méioses, ce qui implique des familles nombreuses étendues sur plusieurs générations ; ii) toutes les familles recrutées doivent présenter strictement la même pathologie et le même mode de transmission, autrement dit, le matériel réuni doit offrir le maximum de garanties d'homogénéité génétique pour permettre le cumul des résultats qu'impose le traitement statistique et donc l'obtention d'un lod score significatif. Pour les maladies héréditaires fréquentes, la réunion de ces deux critères ne pose pas de problèmes. Les conditions sont nettement moins favorables dans les cas de maladies monogéniques rares comme le syndrome de Rett. En effet, les difficultés telles que l'incertitude du diagnostic clinique, le petit nombre ($<1\%$) ainsi que la petite taille des cas familiaux, une possible létalité précoce ont constitué des obstacles majeurs à une localisation précise du locus responsable du syndrome de Rett.

Une alternative de cette technique est la cartographie d'exclusion qui repose sur une analyse de concordance/discordance entre les allèles de différents marqueurs d'ADN et le locus morbide potentiel. Cette approche ne peut s'appliquer qu'à des familles dans lesquelles l'origine parentale de la mutation est strictement déterminée afin de pouvoir associer le phénotype malade à un chromosome (maternel ou paternel). Ainsi, l'analyse des familles dans lesquelles des demi-sœurs ont la même mère et présentent un syndrome de Rett a été privilégiée. Dans ces familles, il a été supposé qu'il y a concordance lorsque les deux demi-sœurs possèdent les mêmes allèles, hérités de la mère pour la région potentiellement associée au locus de la maladie. Et de la même façon, les régions discordantes montrent des haplotypes différents chez les deux demi-sœurs et peuvent donc être exclues du chromosome X. Un recoupement des résultats obtenus à partir de plusieurs familles permet de cerner un ou quelques locus potentiels.

2.2.2. Carte d'exclusion du syndrome de Rett

Trois familles ont principalement permis d'élaborer la carte d'exclusion du chromosome X et d'orienter les recherches vers une région candidate (figure 2).

Dans la famille 1 (figure 2A), la mère a eu trois enfants, deux filles atteintes et un garçon, avec trois maris différents. Le pedigree particulier de cette famille indique avec quasi certitude que l'origine de la mutation est maternelle. L'étude de cette famille avec des marqueurs de plus en plus nombreux a fait l'objet de quatre publications (Archidiacono et al. 1991 ; Ellison et al. 1992 ; Curtis et al. 1993 et Schanen et al. 1997). Les résultats obtenus ont conduit à l'exclusion de la région péricentromérique du chromosome X entre les marqueurs DXS704 en Xp21.3 et DXS456 en Xq22.2 (figure 2B).

En 1997, l'étude d'une famille supplémentaire (famille 2, figure 2A, 2B) (Schanen et al. 1997) chez laquelle une tante et une nièce sont atteintes et la mère certainement vectrice présentant une atteinte légère a étendu la carte d'exclusion jusqu'en Xp22.2 au niveau du marqueur DXS1053. Sur le bras long du chromosome, seul un marqueur a pu être éliminé (DXS456).

En 1998, trois publications (Xiang et al. 1998 ; Sirianni et al. 1998 ; Webb et al. 1998) ciblerent la région candidate en Xq28 grâce au cumul des données de nouvelles familles dont la plus remarquable : la famille 3 (figure 2A). Cette famille, rapportée par Sirianni (Sirianni et al. 1998) est d'origine brésilienne comportant une fratrie de 9 enfants dont 3 fillettes présentent un syndrome de Rett. C'est la première fois qu'un cas familial avec plusieurs (en l'occurrence ici 3) enfants atteints d'un syndrome de Rett est décrit. Ceci a permis aux auteurs de réaliser une analyse de liaison multipoint, en supposant que la mutation était d'origine maternelle. Les auteurs se sont appuyés sur 3 arguments i) présence dans la fratrie de filles atteintes et de filles saines, ii) déséquilibre apparent du sex-ratio (un seul garçon dans toute la fratrie, mort au bout de 3 jours), iii) une inactivation biaisée du chromosome X de la mère (95%-5%). L'analyse de la famille a donné un lod score positif ($Z=1,2$) seulement pour la région en Xq28 à partir du marqueur DXS998 (figure 2B). Toutefois, la valeur significative argumentant en faveur de l'existence d'une liaison entre un marqueur et le locus morbide devant être supérieure ou égale à 3, les auteurs ont alors cumulé les données de cette famille avec les données précédemment publiées et ont déterminé un "lod score cumulatif" de 2,9 uniquement pour la région Xq28. Ces résultats apportaient un argument supplémentaire quant à la localisation probable du gène responsable du syndrome de Rett en position Xq28,

localisation d'abord évoquée par Xiang (Xiang et al. 1998). Un mois plus tard, Webb (Webb et al. 1998) publiait une analyse de liaison entre 6 nouvelles familles confirmant ainsi l'implication vraisemblable de la région Xq28.

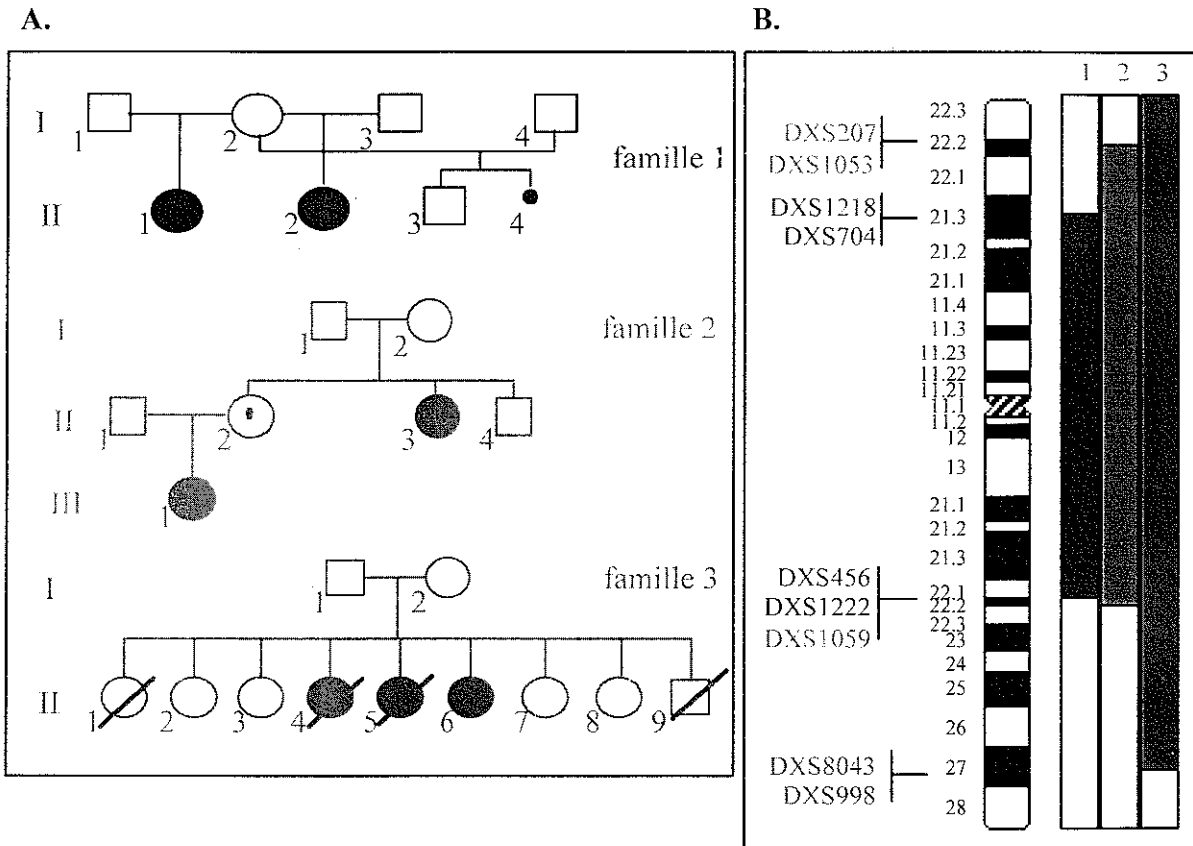


Figure 2 : Cartographie du locus morbide impliqué dans le syndrome de Rett.

A. Pedigrees des familles ayant permis la cartographie du locus morbide dans la région Xq28.
B. Carte d'exclusion du chromosome X. A gauche de l'idéogramme du chromosome X, sont indiqués les marqueurs délimitant les zones exclues (en rouge, en vert et en bleu) par les études de cartographie des familles 1 (rouge), 2 (vert) et 3 (bleu) respectivement.

La région impliquée dans le syndrome de Rett étant ainsi déterminée, "il ne restait plus qu'à" identifier le gène coupable.

2.2.3. Recherche du gène responsable du syndrome de Rett

2.2.3.1. Etudes de gènes candidats

En parallèle de l'approche par cartographie d'exclusion, de nombreux laboratoires ont réalisé une recherche de mutations dans des gènes candidats. Un des tous premiers gènes candidats écarté a été le gène codant l'ornithine transcarbamylase (OTC) localisé en Xp21.1 dont la déficience entraîne l'hyperammoniémie de type II, symptôme pensé initialement comme caractéristique du syndrome de Rett (Rett 1966). Ensuite, les chercheurs se sont focalisés sur des gènes exprimés soit majoritairement, soit exclusivement dans le système nerveux central, étant donné la physiopathologie du syndrome de Rett. Un dizaine de gènes candidats (revue Schanen 1999) ont été criblés, aucun n'a présenté de mutations chez les patientes atteintes de syndrome de Rett.

2.2.3.2. Stratégie de clonage par une approche combinée

Après assignation d'une maladie génétique à une région chromosomique donnée, il est désormais possible d'interroger les bases de données pour avoir accès à la liste des gènes ou des transcrits localisés dans cette région. Les données disponibles sur ces gènes (séquence, expression, fonction, synténie, etc) sont comparées aux caractéristiques de la maladie afin d'identifier le meilleur gène candidat. La recherche de mutations chez des individus atteints permettra de confirmer ou d'infirmer que le gène candidat désigné est réellement responsable de la maladie génétique étudiée.

La région Xq28, bien que donnant une coloration homogène en technique de bandes chromosomiques G, est définie comme étant assez hétérogène au niveau génomique. En effet, elle est composée de régions qui sont répétées sur le chromosome X, de la région pseudo-autosomique 2 (PAR2), des séquences classiques sub-télomériques et télomériques, et enfin d'un fragment de 350 kb dupliqué dans les régions péri-centromériques de quelques autres autosomes (Eichler et al. 1996, 1997). De plus, cette région est considérée, de part notamment sa forte concentration en bases azotées guanines et cytosines, comme l'une des régions les plus riches en gènes du génome humain (17 gènes impliqués dans les RMLX sont déjà décrits dans cette région (Chiurazzi et al. 2001)).

En dépit d'un élargissement des gènes candidats possibles de cette région, en tenant compte d'autres symptômes de la maladie, tels les problèmes de croissance malgré une nutrition adéquate (Motil et al. 1994 ; Motil et al. 1998), ou l'ostéopénie malgré une homéostasie du calcium, du phosphate et du système endocrinien normale (Haas et al. 1997 ; Leonard et al. 1999), aucun gène coupable n'a été caractérisé.

2.2.3.3. Criblage systématique des gènes de la région Xq28

C'est un criblage mutationnel systématique des gènes localisés dans la région Xq28 qui eut raison de l'identification du gène responsable du syndrome de Rett. En octobre 1999, Amir et al. (Amir et al. 1999) décrivent des mutations dans le gène *MECP2* (Methyl CpG-binding protein 2) chez 25% des patientes étudiées.

C. Régulation de la transcription

I. Quelques rappels sur la transcription

La transcription, étape fondamentale du cycle de la cellule, correspond à la synthèse d'ARN (ARNm, ARNt et ARNr). Cette étape se déroule principalement dans le noyau, et de façon plus limitée dans les mitochondries ou autre organelle (chloroplastes) de la cellule. La synthèse d'ARN nécessite une ARN polymérase ADN dépendante qui, à partir du brin matrice de l'ADN et des précurseurs (ATP, CTP, UTP et GTP) fabrique dans le sens 5'-3', la chaîne d'ARN par complémentarité au brin matrice.

On distingue trois molécules d'ARN polymérase différentes permettant la synthèse des différentes classes d'ARN (tableau X).

Tableau X : Les trois classes d'ARN polymérase eucaryotes

classes	gènes transcrits
I	ARNr 28S, ARNr 18S, ARNr 5,8S
II	gènes codant les protéines et la plupart des sn ARN
III	ARN 5S, ARNt, sn ARN U6, ARN 7SL, ARN 7SK, ARN 7SM

La transcription des gènes eucaryotes implique des interactions entre des séquences régulatrices actives en *cis* et des facteurs de transcription protéiques actifs en *trans*. En effet, l'ARN polymérase ne peut initier seule la transcription.

Le promoteur, ensemble de courtes séquences regroupées en amont de la séquence codante du gène, constitue le signal de reconnaissance pour les facteurs de transcription qui se lient à l'ADN afin de positionner la polymérase au site unique d'initiation de la transcription.

Il existe différents types de séquences promotrices, chacune fonctionnant avec des facteurs particuliers qui caractérisent certains gènes.

Ainsi, les gènes ayant une expression spatio-temporelle bien précise (comme la β -globine) incluent toujours dans leurs séquences promotrices une séquence consensus TATA(A/T)A(A/T) appelée boîte TATA située 25 pb en amont du site d'initiation de la transcription. Cette séquence est responsable de la sélection précise de la base qui définit le

premier nucléotide transcrit, une mutation dans cet élément n'empêche pas l'initiation de la transcription mais entraîne une hétérogénéité dans le premier nucléotide transcrit.

D'autres gènes, tels les gènes domestiques n'ont pas de boîte TATA mais souvent une boîte appelée GC, correspondant à la séquence consensus GGGCGG. En plus de ces boîtes TATA ou GC, les séquences promotrices classiques incluent souvent une boîte CAAT, située 80 pb en amont du site d'initiation, déterminant important de l'efficacité du promoteur.

De la même façon, il existe différents facteurs de transcription dont certains définissent avec l'ARN polymérase le complexe minimal de la transcription. Le tableau XI en présente quelques uns.

Tableau XI : Quelques éléments régulateurs en *cis* et leurs facteurs de transcription associés

Eléments <i>cis</i>	Séquence consensus	Facteurs associés en <i>trans</i>
Boîte GC	GGGCGG	Sp1
Boîte TATA	TATA(A/T)A(A/T)	TFIID
Boîte CAAT	CCAAT	Nombreux, C/EBP, CTF/NFI
TRE	GTGAGTA/CA	Famille AP-1, JUN/FOS
CRE	GTGACGTA/CAA/G	Famille CREB/ATF

Une séquence des événements de formation du complexe basal de la transcription et d'initiation de la transcription des gènes codant des polypeptides a été proposée (Roeder et al. 1991). Après reconnaissance d'une séquence promotrice particulière par ses facteurs de transcription, le facteur de transcription TFII B se fixerait à l'ADN et serait impliqué dans la détermination de la distance qui sépare le promoteur du site d'initiation de la transcription. Le facteur TFII F recruterait alors l'ARN polymérase II sur l'ADN. L'élément TFII E interagirait avec le complexe TFII F/ARN polymérase II et permettrait la fixation de TFII H responsable de la phosphorylation de la grande sous-unité de la polymérase entraînant ainsi son activation. A ce promoteur minimal déterminant la position précise du démarrage de la transcription et le niveau basal d'expression, viennent s'ajouter des activateurs (enhancers) actifs en *cis*, augmentant la production d'ARNm de gènes spécifiques ou des inhibiteurs (silencers) diminuant l'activité transcriptionnelle. A la différence des séquences promotrices dont la position est relativement constante, ces enhancers ou silencers sont localisés à des distances variables et parfois considérables du site d'initiation de la transcription (par exemple un élément régulateur du gène *DFN3* est situé à 40-50 kb du promoteur (Cremers et al. 1995)).

Comme dans le promoteur minimal, ces séquences fixent des protéines qui régulent l'expression des gènes à distance en introduisant des boucles au niveau de l'ADN leur permettant ainsi d'interagir positivement ou négativement avec le complexe transcriptionnel (figure 3).

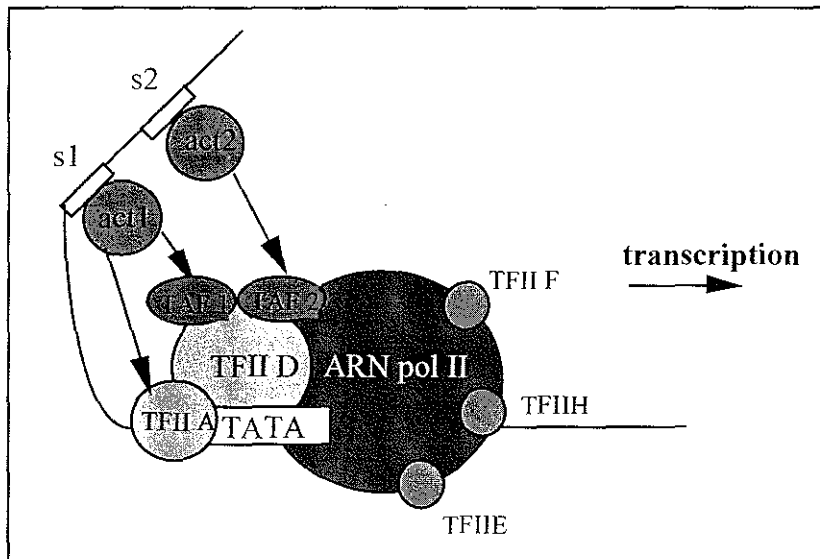


Figure 3 : *Complexe d'initiation de la transcription (d'après Roeder 1991)*

TAF : TFII D associated factor
act : activateur
s1, s2 : séquences régulatrices distales

II. Régulation de la transcription

Le contenu en ADN d'un type spécifique de cellules est exactement identique à celui de tout autre type cellulaire du même organisme. C'est la variation du profil d'expression des gènes au sein de la cellule qui est à l'origine de sa spécificité. Les fonctions essentielles (métabolisme, croissance et division cellulaire) sont déterminées par les gènes dits de ménage ou domestiques exprimés dans toutes les cellules. L'expression de nombreux autres gènes est étroitement restreinte au type cellulaire (expression génique spécifique de tissu). Cependant même les gènes à expression spécifique de tissu sont transcrits à des niveaux très bas dans quasiment tous les types cellulaires par la transcription illégitime (Kaplan et al. 1992).

L'expression génique spatio-temporelle spécifique à chaque type cellulaire est due à un **contrôle épigénétique** de l'expression des gènes. Ceci implique un "marquage" de l'ADN, sans altérer sa séquence nucléotidique. A l'heure actuelle, au moins deux mécanismes à l'origine de ce marquage sont connus, il s'agit de la **méthylation** de l'ADN et de l'**acétylation/désacétylation** des histones. Ces deux mécanismes agissent de concert dans la **régulation transcriptionnelle** des gènes.

1. Modifications de la chromatine

Dans les cellules eucaryotes, le matériel génétique est organisé en chromatine, structure complexe composée d'ADN et de protéines. L'unité fondamentale de la chromatine est le nucléosome composé de 146 pb d'ADN enroulées autour d'un octamère de protéines histones comprenant un tétramère (H3-H4)₂ et deux dimères H2A-H2B (positionnés chacun sur une face du tétramère) et d'une région d'ADN dite de liaison dont la longueur varie selon les espèces et le type cellulaire. Un 5^{ème} type d'histone, l'histone H1 ou histone de liaison, se lie à l'ADN de liaison et aux histones de l'octamère. Elle stabilise les niveaux supérieurs de compaction de la chromatine.

Il existe deux types de chromatine : l'hétérochromatine et l'euchromatine. L'hétérochromatine correspond à un état très condensé de la chromatine, et est généralement associée à une répression transcriptionnelle. On distingue l'hétérochromatine constitutive localisée au niveau des régions centromériques et subtélomériques, de l'hétérochromatine facultative qui adopte cette conformation condensée dans certaines régions du génome de manière cellule-spécifique. L'euchromatine représente un état moins condensé de la chromatine, et est généralement associée à une activité transcriptionnelle.

La chromatine, par sa structure dynamique, est déterminante pour toutes les fonctions du génome : transcription, réplication, réparation et recombinaison. Cette nature dynamique est rendue possible par des modifications post-traductionnelles réversibles des histones : acétylation, phosphorylation, méthylation, ubiquitinylation, glycosylation et ADP-ribosylation. Les quatre histones nucléosomiques H3, H4, H2A et H2B sont composées chacune de deux domaines : un domaine carboxy-terminal enchassé dans la structure du nucléosome et une extrémité amino-terminale basique pointant à l'extérieur du nucléosome. Cette extrémité, encore appelée "queue" de l'histone, de longueur variable (de 16 à 44 acides aminés, Spencer et al. 1999) est riche en lysines. C'est au niveau de ce domaine N-terminal qu'ont lieu la plupart des modifications, excepté l'ubiquitinylation qui se produit dans le

domaine C-terminal. Ces modifications constitueraient un véritable code épigénétique des histones (Jenuwein et al. 2001). Le rôle de ces différentes modifications n'est pas encore clairement élucidé. Il semble que l'ubiquitinylation des histones affecte la ségrégation des chromosomes et la transcription. L'histone H2B ubiquitinylée serait associée à de l'ADN actif transcriptionnellement (Spencer et al. 1999). H1 et H3 seraient phosphorylées sur des résidus sérines et thréonines particuliers de manière dépendante du cycle cellulaire. En effet, une phosphorylation plus importante de ces histones est observée au cours des mitoses pendant la condensation des chromosomes. La découverte récente chez l'homme de l'histone H3 méthyltransférase (SUV39H1) a mis en évidence que la méthylation de l'histone H3 interviendrait dans l'organisation de la chromatine en euchromatine et hétérochromatine. Parmi toutes les modifications post-traductionnelles des histones, le rôle de l'acétylation/désacétylation des histones dans la régulation transcriptionnelle des gènes est aujourd'hui le mieux connu.

1.1. Acétylation et désacétylation

Une corrélation entre l'acétylation des histones et l'activité transcriptionnelle a été suspectée dès le milieu des années 60 (Allfrey et al. 1964). Les régions hyperacétylées de la chromatine contiennent des unités actives de transcription alors qu'une chromatine hypoacétylée est inactive transcriptionnellement. L'état d'acétylation des histones résulte d'un équilibre entre deux activités enzymatiques antagonistes : l'activité histone-acétyl transférase (HAT) qui permet de greffer un groupement acétyl, et l'activité histone désacétylase (HDAC) qui permet de l'éliminer. Les HAT ou HDAC peuvent agir sur des régions chromatiniennes variables allant d'un nucléosome (par exemple le gène *his3* de *S. Cerevisiae*) jusqu'à la quasi-totalité d'un chromosome (le chromosome X inactivé chez les femelles des mammifères).

Les premières HAT ou HDAC ont été identifiées chez la levure, ensuite c'est par une recherche d'homologies dans les banques de données que l'on a pu identifier les homologues murins et humains.

1.1.1. Les HAT

Les protéines nucléaires possédant une activité histone-acétyl transférase sont réparties en six familles selon les variations observées dans le domaine conservé responsable de l'activité enzymatique.

Ces différentes protéines appartiennent à des complexes multiprotéiques incluant facteurs de transcription et autres corégulateurs, elles sont impliquées dans une grande variété de voies de transduction de signal (tableau XII).

Tableau XII : Les familles HAT

HAT	Organismes	Fonctions
Famille GCN5/PCAF Gcn5 PCAF	Levure→Homme Homme	Co-activateur Co-activateur
Famille MYST Sas2 Sas3 Esa1 MOF Tip60 MOZ	Levure Levure Levure Drosophile Homme Homme	Inactivation Inactivation Cycle cellulaire Compensation de dosage Interaction avec HIV Tat Leucémogénèse
Famille TAFII250	Levure→Homme	Facteur associé à TBP
Famille CBP/p300	Ver→Homme	Co-activateur global
Famille SRC ATF2	Souris et Homme Levure→Homme	Co-activateur des récepteurs stéroïdiens Activateur séquence-spécifique
Famille GNAT TFIIIC	Homme	Initiation de la transcription

d'après Marmorstein et al. 2001

L'acétylation ciblée d'histones nucléosomiques par le recrutement d'une HAT elle-même associée à un complexe protéique particulier, agirait de concert avec des facteurs dits de remodelage des nucléosomes pour activer la transcription par modification de la structure chromatinienne. En effet, des facteurs de remodelage des deux grandes familles SWI/SNF et ISWI (définies par un domaine hélicase appelée swi2/snf2 et ISWI respectivement, qui est couplé à une sous-unité catalytique ATPase) ont la capacité de rendre l'ADN plus accessible. L'acétylation neutraliserait les charges positives portées par les résidus lysine, ce qui

affaiblirait les interactions ADN/histones. Ces interactions diminuées, les facteurs de remodelage (des deux familles SWI/SNF et ISWI) modifieraient la structure nucléosomique en une forme dite altérée, par glissement des histones sur l'ADN libre rendant ainsi accessibles les séquences promotrices cibles (Taddei et al. 2000).

1.1.2. Les HDAC

Il existe trois classes d'enzymes présentant une activité histone désacétylase chez les eucaryotes supérieurs, définies en fonction de leurs similarités avec les séquences des HDAC de levure. La classe I comprend les HDAC1, 2, 3 et 8 qui sont homologues à l'HDAC Rpd3 de la levure. La classe II correspond aux HDAC4, 5, 6 et 7 homologues à l'HDAC Hda1 de levure. Et enfin la classe III, récemment déterminée (Khochbin et al. 2001), regroupe les protéines SIRT1 à 7 homologues à la protéine Sir2 de levure. Toutes ces protéines possèdent un domaine catalytique conservé capable de désacétyler les histones ; on distingue les trois classes par les séquences extérieures au domaine catalytique (Khochbin et al. 2001). Par ailleurs, une étude récente des données de la littérature et des bases de données suggère que le nombre de HDAC chez l'homme pourrait atteindre 17 (Gray et al. 2001).

A la différence des HAT, les HDAC appartiennent à des complexes multiprotéiques qui sont impliqués dans la répression de la transcription. En effet, dans ces complexes multiprotéiques, les HDAC sont couplées à des co-répresseurs. Tous les complexes ne sont pas encore élucidés à l'heure actuelle, deux seulement fonctionnant avec HDAC1 et 2 commencent à être bien caractérisés. Ces deux complexes, Sin3/HDAC et NuRD/Mi2/NRD ont en commun en plus de HDAC1 et 2, les protéines RbAp48 et RbAp46 (Rb associated protein 48/46 kb), et diffèrent par les co-répresseurs associés.

Le premier complexe, conservé depuis la levure jusqu'aux mammifères, comprend le co-répresseur Sin3p chez la levure, ou mSin3A ou mSin3B chez les mammifères.

Le second complexe est caractérisé par le co-répresseur Mi2 ou NuRD (Nucleosome remodeling histone deacetylase complex) identifié chez les mammifères comme facteur de remodelage du nucléosome (Knoepfler et al. 1999).

De nombreux autres co-répresseurs associés aux HDAC1 et 2 ont été identifiés ou sont en cours d'identification, laissant supposer que les deux complexes distincts formés soit à partir de Sin3 soit à partir de Mi2 ne représentent que la partie immergée de l'iceberg correspondant aux différentes voies de répression transcriptionnelle de la cellule.

La figure 4 illustre les différents co-répresseurs connus associés aux HDAC1 et 2.

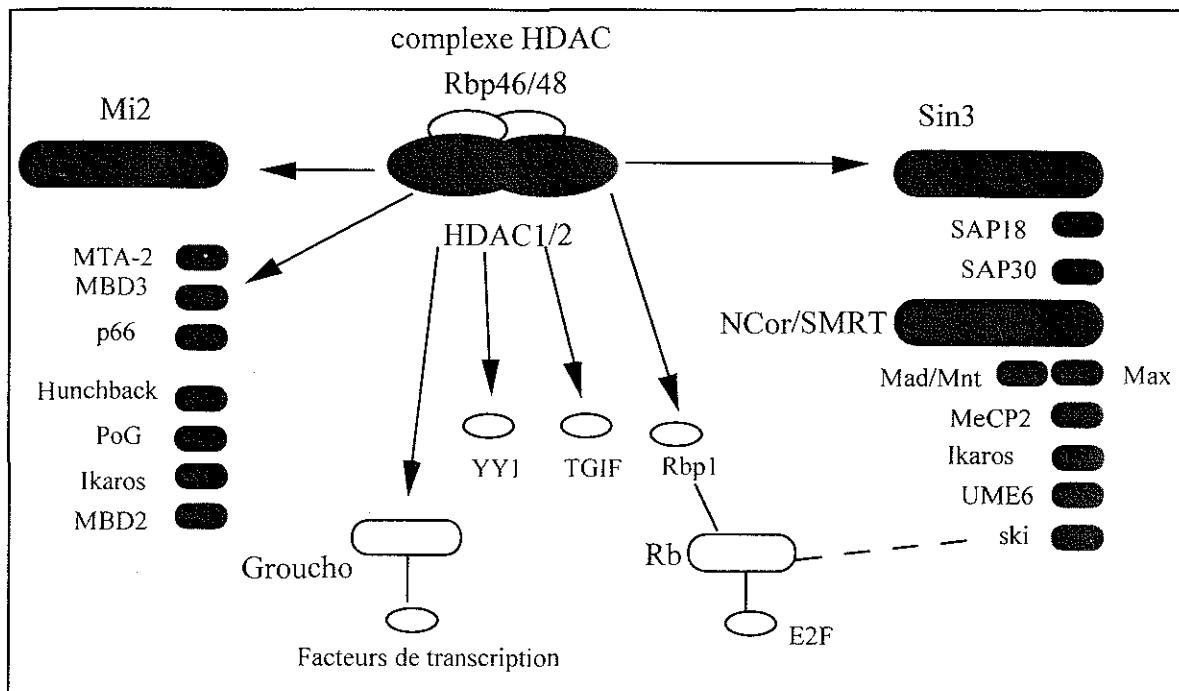


Figure 4 : Facteurs et co-répresseurs associés aux HDCA 1 et 2 (d'après Knoepfler et al. 1999)

Sont représentés les différents co-répresseurs interagissant avec HDAC1 et HDAC2 (en vert).
 En bleu sont indiqués les facteurs associés au co-répresseur Sin3
 En rouge ceux associés au co-répresseur Mi2,
 Les flèches montrent les interactions directes entre HDAC1, 2 et les co-répresseurs ou facteurs de transcription.

La troisième HDAC de la classe I, HDAC3 fait partie d'un autre complexe nucléaire contenant SMRT et N-CoR qui sont des co-répresseurs de récepteurs nucléaires (Khochbin et al. 2001).

Les complexes comprenant les HDAC des classes II et III n'ont pas été isolés. Les membres de la classe II ne semblent pas participer à l'un et/ou l'autre complexe Sin3/HDAC, NuRD/Mi2/NRD. Ils sont caractérisés par leur capacité de circulation entre le noyau et le cytoplasme. Cette propriété serait dépendante du type cellulaire. Ainsi leur activité nucléaire serait hautement tributaire de leur localisation cellulaire (Khochbin et al. 2001).

De nouveaux co-répresseurs seront peut-être identifiés ultérieurement associés spécifiquement à ces deux classes de HDAC, ou bien ces HDAC pourraient être impliquées dans des voies de répression transcriptionnelle indépendantes de co-répresseurs en désacétylant directement par exemple des protéines du complexe transcriptionnel.

Comment la désacétylation des histones entraîne-t-elle une répression transcriptionnelle par modification de la structure chromatinienne ? Actuellement, deux modèles sont proposés. Dans le premier, la désacétylation des histones H3 et H4 libérerait les charges positives du résidu lysine qui s'allieraient par des interactions ioniques avec les groupements phosphates de l'ADN chargés négativement. Cette interaction empêcherait les facteurs de transcription de se fixer au niveau de leur séquence cible. Dans le second modèle, la désacétylation conduirait à une compaction de la chromatine par interaction entre nucléosomes adjacents (figure 5).

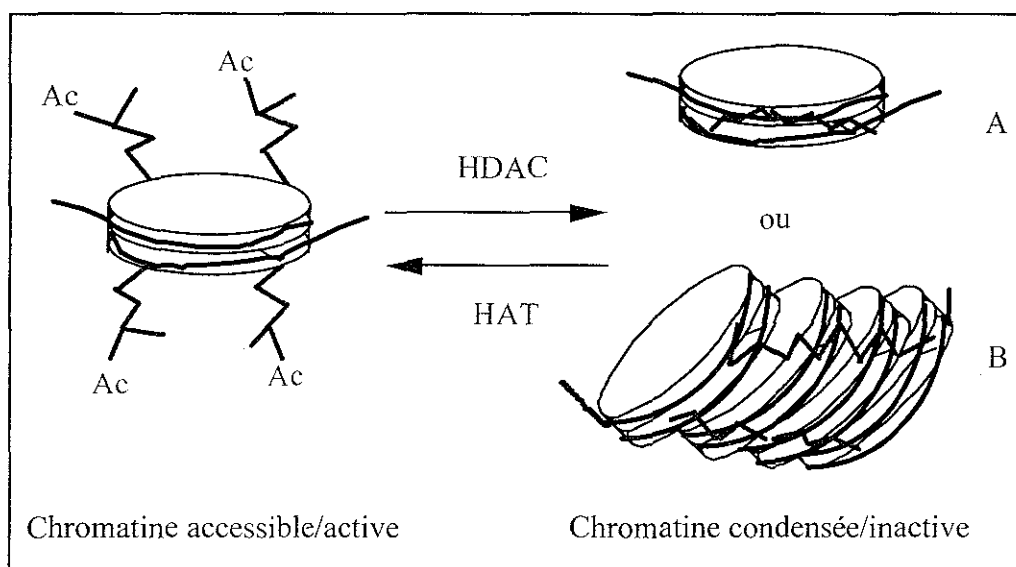


Figure 5 : Les deux modèles possibles (A ou B) de l'effet de la désacétylation sur la chromatine

La haute affinité des queues des histones désacétylées pour l'ADN favorise le premier mécanisme, mais des études de diffraction aux rayons X de cristaux de nucléosome semblent en faveur du second (Bestor 1998).

Ces deux modèles sont peut-être l'illustration de l'action différente des multiples complexes pour réprimer la transcription dans la cellule. Des gènes pourraient nécessiter un remodelage plus complexe de la chromatine grâce à l'intervention du facteur Mi2 pour être efficacement inactivés. Ces mêmes gènes seraient alors résistants à la répression par le complexe comprenant le co-répresseur Sin3.

En conclusion, les activités histone-acétyl-transférase et histone désacétylase incorporées via des corégulateurs au sein de complexes protéiques incluant des facteurs de transcription

spécifiques, seraient ciblées vers des promoteurs particuliers. Par exemple, l'hétérodimère Mad-Max en se fixant à son site de reconnaissance recruterait au niveau de ses gènes cibles, HDAC2 via le corepresseur Sin3. Dans le cas des gènes activés en réponse à l'AMPc, le facteur CREB se lie aux séquences CRE (cyclic AMP response element) en amont des promoteurs des gènes concernés, et recrute l'histone-acétyl-transférase CBP/p300 qui interagit avec une autre HAT (PCAF). Ici, CBP/p300 est à la fois une HAT et un corégulateur. Dans le cas des récepteurs des hormones thyroïdiennes, le complexe RXR/TR en absence de son ligand recrute une HDAC via mSin/NCor/SMRT alors que, en présence du ligand, il recrute une HAT au même site.

Il est à noter enfin, que les HAT et les HDAC agiraient sur d'autres protéines que les histones nucléosomiques comme par exemple des facteurs de transcription ou des protéines chromatinienne non-histones (HMG (High mobility group)) pour réguler la transcription de la cellule.

2. Méthylation de l'ADN

2.1. Définition et fonctions de la méthylation

La méthylation de l'ADN constitue une modification épigénétique majeure des génomes eucaryotes. A la différence des procaryotes où adénines et cytosines peuvent être méthylées, chez les eucaryotes, la méthylation n'affecte que les cytosines, elle se produit sur le carbone 5 du noyau pyrimidique. De plus, seules les cytosines appartenant à un doublet 5'CpG3' sont méthylables.

Ces dinucléotides sont relativement sous-représentés dans le génome humain, leur fréquence est cinq fois inférieure à celle attendue par un simple calcul statistique. Chez l'homme, 60 à 90% des cytosines de ces dinucléotides 5'CpG3' sont méthylées (Bird 1986). La 5-méthylcytosine peut être spontanément désaminée et donner une thymine (Frederico et al. 1990 ; Shen et al. 1994). La thymine étant alors mésappariée avec une guanine, devrait être reconnue par le système de réparation spécifique de l'ADN, excisée et remplacée par une cytosine. Mais la rareté des séquences 5'CpG3' indique probablement que le système de réparation est relativement inefficace dans ce cas précis. La conséquence est que ces cytosines méthylées constituent des points chauds de mutation (C→T ou G→A).

Certaines régions du génome riches en 5'CpG3' (correspondant à environ 1% de la totalité des dinucléotides 5'CpG3') ont été caractérisées car elles donnaient de nombreux petits

fragments après digestion par une enzyme de restriction sensible à la méthylation c'est-à-dire contenant dans son site de coupure un dinucléotide 5'CpG3' (comme l'enzyme HpaII dont la séquence cible est CCGG). Cette hydrolyse a mis en évidence deux propriétés de ces régions : une concentration élevée en doublets 5'CpG3' et un état non méthylé. Ces régions d'ADN riches en 5'CpG3' sont appelées " îlots CpG " ou îlots HTF (HpaII Tiny Fragments). Ils présentent donc la particularité d'être regroupés sur 1 à 2 kb dans des régions précises du génome (Bird 1986). Ces îlots sont localisés au niveau des promoteurs d'approximativement 60% des gènes transcrits par l'ARN polymérase II, dont la quasi-totalité des gènes domestiques (Antequera et al. 1994). Le gène codant la glucose-6-phosphate déshydrogénase (G6PD) présente la particularité de posséder une paire d'îlots CpG non pas dans sa région promotrice mais à proximité de son extrémité 3' (Bird 1986).

La méthylation de l'ADN est à l'origine de plusieurs mécanismes impliqués dans la répression transcriptionnelle. Comme nous l'avons vu précédemment, la méthylation joue un rôle essentiel dans l'inactivation des gènes situés sur un des deux chromosomes X chez les femelles de mammifères.

Par ailleurs, la méthylation est impliquée dans la répression de gènes soumis à l'empreinte parentale. Chez les mammifères, quelques gènes ne sont exprimés qu'à partir d'un seul allèle en fonction de son origine parentale (on parle d'hémizygotie fonctionnelle), l'allèle réprimé étant hyperméthylé. Les syndromes d'Angelman et de Prader-Willi constituent une illustration de référence de l'empreinte parentale. En effet, une absence de contribution paternelle par délétion ou disomie maternelle de la région critique (en 15q11-q13) entraîne un syndrome de Prader-willi. A l'opposé, une absence de contribution maternelle est à l'origine d'un syndrome d'Angelman. Ainsi une même région du génome engendre deux pathologies distinctes.

De plus, la méthylation des cytosines contribuerait à l'immobilisation de transposons ou de rétrovirus endogènes, agents potentiellement hautement mutagènes par insertion ou recombinaison illégitime. Les gènes soumis à inactivation sur le chromosome X chez la femme ou à empreinte parentale rendent compte de moins de 10% des cytosines méthylées du génome ; la grande majorité se situe au niveau des transposons qui sont assez abondants (les éléments transposables au nombre de 10^6 représentent environ 40% du génome humain). La méthylation de l'ADN constituerait une défense de la cellule hôte contre ces éléments (Bestor 2000).

Une perturbation des profils généraux de la méthylation entraîne de nombreux dysfonctionnements cellulaires. Il a été montré que la méthylation était essentielle pour un

développement murin correct. En effet, des souris privées du gène codant l'ADN méthyltransférase 1 ne sont pas viables (Li et al. 1992). De la même façon, une déméthylation des séquences satellites classiques de l'ADN chez l'homme entraîne une instabilité chromosomique avec une immunodéficience caractéristiques du syndrome ICF (Immunodeficiency, Centromere instability and Facial anomalies) (Xu et al. 1999).

Enfin, à l'opposé, une hyperméthylation aberrante a des conséquences importantes sur la cancérogenèse ou sur les tentatives d'intervention thérapeutiques. La méthylation de promoteurs de gènes suppresseurs de tumeurs contribue directement à la progression de certains cancers (Baylin et al. 2000), alors que la méthylation de l'ADN exogène introduit dans des cellules compromet les efforts de thérapie génique (Bird et al. 1999).

2.1.1. Les acteurs de la méthylation

Actuellement, trois familles d'ADN cytosine-5-méthyltransférase (DNMT) sont connues chez les mammifères : la famille DNMT1, la famille DNMT2 et la famille DNMT3. Ces trois familles regroupent cinq DNMTs présentant toutes un domaine catalytique caractérisé par dix motifs de séquences dont six sont fortement conservés (excepté pour DNMT3L) (figure 6). On retrouve la quasi-totalité de ces motifs dans toutes les DNMTs depuis la bactérie jusqu'à l'homme en passant par les champignons et les plantes (Colot et al. 1999).

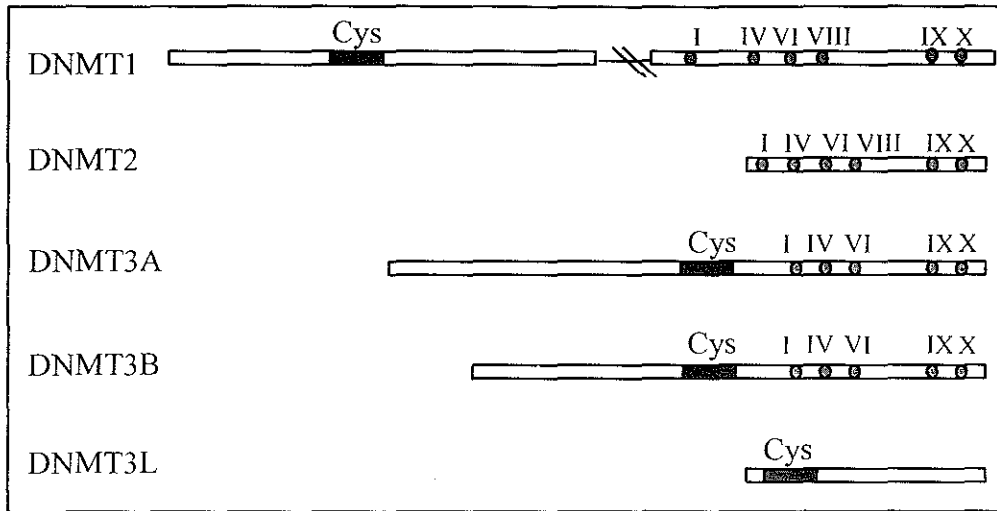


Figure 6 : Les ADN méthyltransférases (d'après Bestor, 2000)

Les motifs I-X définissent le domaine catalytique des ADN méthyltransférases.
Cys : région riche en cystéines

-La famille DNMT1

Le gène *Dnmt1* présente la particularité unique d'être exprimé à partir de promoteurs différents qui sont spécifiques du sexe et d'un exon (figure 7).

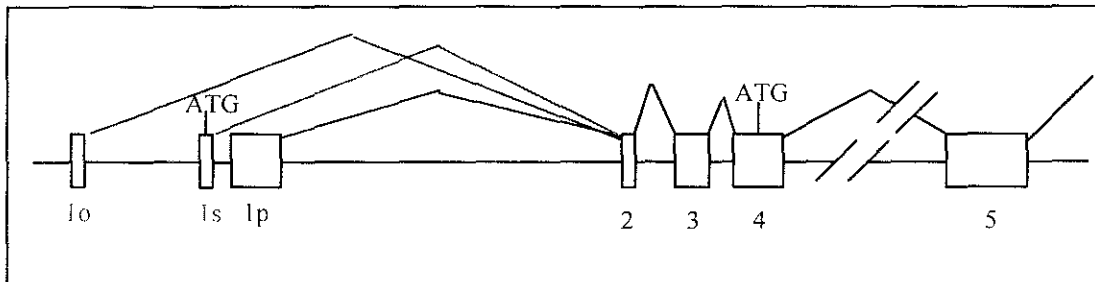


Figure 7 : Les différents exons sexe-spécifiques du gène *Dnmt1* (d'après Bestor 2000)

1o = exon spécifique de l'expression du gène dans les ovocytes
1s = exon spécifique de l'expression du gène dans le soma (des deux sexes)
1p = exon spécifique de l'expression du gène dans les spermatozoïdes au stade pachytène
 En bleu, est indiqué l'épissage se produisant dans les ovocytes,
 En vert, l'épissage spécifique des cellules somatiques,
 En rouge, l'épissage spécifique des spermatozoïdes.

Ainsi, il existe une protéine de 1620 acides aminés, exprimée seulement au niveau du soma, et une protéine 118 acides aminés (plus courte du côté N-terminal) exprimée uniquement dans les ovocytes. Les spermatozoïdes possèdent aussi leur propre (exon et) promoteur actif

entraînant un épissage alternatif spécifique du stade pachytène de la méiose. Bien que ce transcrit spécifique soit synthétisé, aucune protéine n'est détectée dans les spermatocytes.

Yoder et al. (Yoder et al. 1997) ont montré que chez la souris, la protéine DNMT1 avait une préférence 5 à 30 fois supérieure pour l'ADN hémiméthylé par rapport à de l'ADN non méthylé, d'où la conclusion qu'elle serait responsable de la méthylation dite de maintenance et agirait notamment après le passage de la fourche de réplication. Récemment, Fuks et al. (Fuks et al. 2000) ont localisé dans DNMT1 un domaine de répression transcriptionnelle pouvant interagir avec HDAC1 et ainsi réprimer la transcription.

-La famille DNMT2

Dnmt2 a été identifié par recherches d'homologies avec *Dnmt1* dans les bases de données regroupant l'ensemble des EST (Expressed Sequence Tag). Fait surprenant, elle est homologue à *pmt1p* de *S. pombe*, organisme non connu pour méthyler son ADN. A la différence de *Dnmt1*, une interruption du gène *Dnmt2* dans les cellules souches embryonnaires (cellules ES) n'a pas de conséquence sur les profils de méthylation, et n'affecte pas la capacité de ces cellules à méthyler un ADN viral nouvellement intégré. Se pose la question du rôle biologique de cette famille d'ADN méthyltransférases. Un début d'explication du rôle physiologique proviendrait de l'observation que seuls les organismes possédant des homologues de *Dnmt2* présentent une fonction et une structure du centromère identiques ; ce qui n'est pas le cas par exemple de *S. cerevisiae* ou *C. elegans* qui n'ont pas d'homologues de *Dnmt2* et qui présentent une structure centromérique différente (Bestor 2000).

-La famille DNMT3

Cette famille comprend trois méthyl-transférases : DNMT3A, DNMT3B et DNMT3L. DNMT3L a été cartographiée dans la partie distale de la région critique du syndrome de Down en 21q22.2-22.3. On ne sait pas si ce gène est impliqué ou non dans l'étiologie de ce syndrome. DNMT3A et DNMT3B seraient responsables de la méthylation *de novo* de régions particulières du génome. Un double mutant murin *Dnmt3A^{-/-}/Dnmt3B^{-/-}* est incapable de méthyler des séquences virales nouvellement intégrées, alors que chaque simple mutant hétérozygote peut le faire. L'interruption de *Dnmt3A* comme de *Dnmt3B* chez la souris est létale (Okano et al. 1999).

DNMT3B est la seule méthyl-transférase, à ce jour, dont une mutation à l'état homozygote dans le domaine C-terminal est impliquée dans une pathologie humaine : le syndrome ICF.

L'instabilité centromérique caractéristique de ce syndrome s'explique par une déméthylation des séquences satellites classiques de l'ADN visible cytogénétiquement au niveau des régions péri-centromériques des chromosomes 1, 9 et 16. Bien que ces régions soient totalement méthylées, aucun des patients n'est homozygote pour une mutation nulle, suggérant que la perte totale de fonction de DNMT3 serait létale chez l'homme comme c'est le cas chez la souris (Bestor 2000).

2.2. Méthylation et répression transcriptionnelle

Deux modèles, l'un direct et l'autre indirect, sont envisagés pour expliquer la façon dont la méthylation des 5'CpG3' réprime la transcription.

La méthylation peut interférer directement avec la transcription en empêchant la liaison de la machinerie basale ou de facteurs plus spécifiques avec les séquences promotrices cibles. La méthylation de ces séquences serait perçue comme une "mutation" dans le site de liaison pour les facteurs de transcription et empêcherait donc leur fixation. Il a été montré, en effet, que certains facteurs de transcription comme les facteurs de transcription AP-2 et CREB/ATF impliqués respectivement dans l'expression du gène proenkephaline (Comb et al. 1990) et des gènes répondant à l'AMPc (Iguchi-Arigo et al. 1989) étaient sensibles à la méthylation de leur promoteur respectif. Toutefois, ce n'est pas le cas de tous les facteurs de transcription, par exemple le facteur de transcription Sp1 dont la séquence cible est la boîte CG (GGGCGG) est indifférent à son état de méthylation (Holler et al. 1988).

Ceci constitue un premier argument en faveur de l'existence d'un mécanisme d'action indirecte de la méthylation sur la transcription. D'autres observations sont venues conforter cette hypothèse. Une de ces observations montre que l'injection dans des noyaux du gène méthylé de la thymidine kinase du virus simplexe de l'herpès (HSV tk) n'entraînait sa répression transcriptionnelle qu'au bout de huit heures (Buschhausen et al. 1987). Ce résultat est incompatible avec un mécanisme instantané de répression. La cinétique de mise en place de la répression pourrait témoigner d'un temps nécessaire à l'établissement d'une interaction entre l'ADN méthylé et des composés nucléaires. Ainsi, ce mécanisme indirect suppose l'intervention de protéines nucléaires particulières qui reconnaîtraient spécifiquement l'ADN méthylé et empêcheraient secondairement l'interaction entre le complexe transcriptionnel et le promoteur. Ce modèle a été validé par la découverte de ces protéines nucléaires qui se lient à l'ADN méthylé et répriment la transcription (Boyes et al. 1991).

2.2.1. Les protéines se liant aux cytosines méthylées ou Methyl CpG binding proteins (MeCP)

Au début des années 90, deux protéines ayant une affinité pour l'ADN méthylé ont été caractérisées : MeCP1 (Meehan et al. 1989) et MeCP2 (Lewis et al. 1992).

Aujourd'hui, ces deux protéines appartiennent à une famille constituée de trois membres supplémentaires : MBD2, MBD3 et MBD4 (pour Methyl Binding Domain) identifiés par Hendrich et al. (Hendrich et al. 1998) grâce à une recherche, dans les bases de données mammaliennes des EST, de séquences homologues au domaine de liaison aux cytosines méthylées (MBD) de MeCP2.

Seule MeCP1 est en réalité un complexe multiprotéique (poids moléculaire natif par gel filtration entre 400-800 kD (Meehan et al. 1989)) dont la composition n'est pas totalement élucidée. MBD2 constitue le composant de liaison à l'ADN méthylé du complexe MeCP1 (Ng et al. 1999). MeCP1 étant un complexe multiprotéique, il n'apparaît pas directement dans la liste des membres de la famille des MeCP, mais apparaît par son MBD. Ainsi, les cinq membres de la famille des MeCP sont : MeCP2, MBD1, MBD2, MBD3 et MBD4.

Toutes ces protéines ont en commun le MBD composé d'environ 70 acides aminés présentant un haut degré de similarité (50-70%), et dont la séquence ne correspond à aucun des cinq motifs d'interaction protéine/ADN précédemment connus [motif hélice-coude-hélice (ou helix-turn-helix), motif en doigt de zinc (ou zinc finger), motif glissière à leucines (ou leucine zipper), motif hélice-boucle-hélice (ou helix-loop-helix) et le motif feuillet β antiparallèles (ou antiparallel β -sheet)]. Ceci suggère fortement que le MBD constitue un nouveau motif d'interaction protéine-ADN. Les homologies de séquences au sein du MBD montrent que les protéines pourraient être regroupées en deux sous-classes : d'un côté MeCP2 et MBD4, et de l'autre côté MBD1, 2 et 3. Dans les cinq gènes, le MBD est interrompu par un intron toujours localisé à la même position témoignant d'un probable gène ancestral commun (Hendrich et al. 1998). Enfin, les similarités de séquences entre ces cinq protéines sont uniquement restreintes à leur MBD, excepté pour MBD2 et 3 qui présentent une similarité de 71,1% sur la totalité de leur séquence (Hendrich et al. 1998).

Des expériences de Northern blot ou de RT-PCR (reverse transcriptase-PCR) sur ces cinq gènes ont mis en évidence une expression ubiquitaire chez la souris. Dans les cellules ES, seules MBD3, MBD4 et MBD2 en plus faible quantité, sont détectées. Cette diminution, voire pour MBD1 cette absence, de transcrits peut être corrélée avec le fait que la méthylation est connue pour être peu importante dans la viabilité des cellules ES (Lei et al. 1996). Une

quantité normale de transcrits *MBD3* et *MBD4* ne suppose pas forcément une synthèse des protéines correspondantes.

Du fait de leur expression ubiquitaire, ces cinq gènes sont considérés comme des gènes de ménage et comme attendu pour de tels gènes, des îlots CpG ont été caractérisés dans leur extrémité 5' (Hendrich et al. 1999). Les localisations chromosomiques ont été déterminées pour les quatre gènes *MBD1*, 2, 3 et 4 (Hendrich et al. 1999). A la différence du locus *MECP2* (Vilain et al. 1996), ces quatre gènes sont autosomiques. Les gènes *MBD1* et 2 sont cartographiés tous les deux sur le chromosome 18 au niveau de la bande q21, le locus *MBD3* est positionné en 19q13 et *MBD4* en 3q21.

La figure 8 résume l'organisation et la structure protéique des gènes *MBD1*, 2, 3, 4 et *MECP2*.

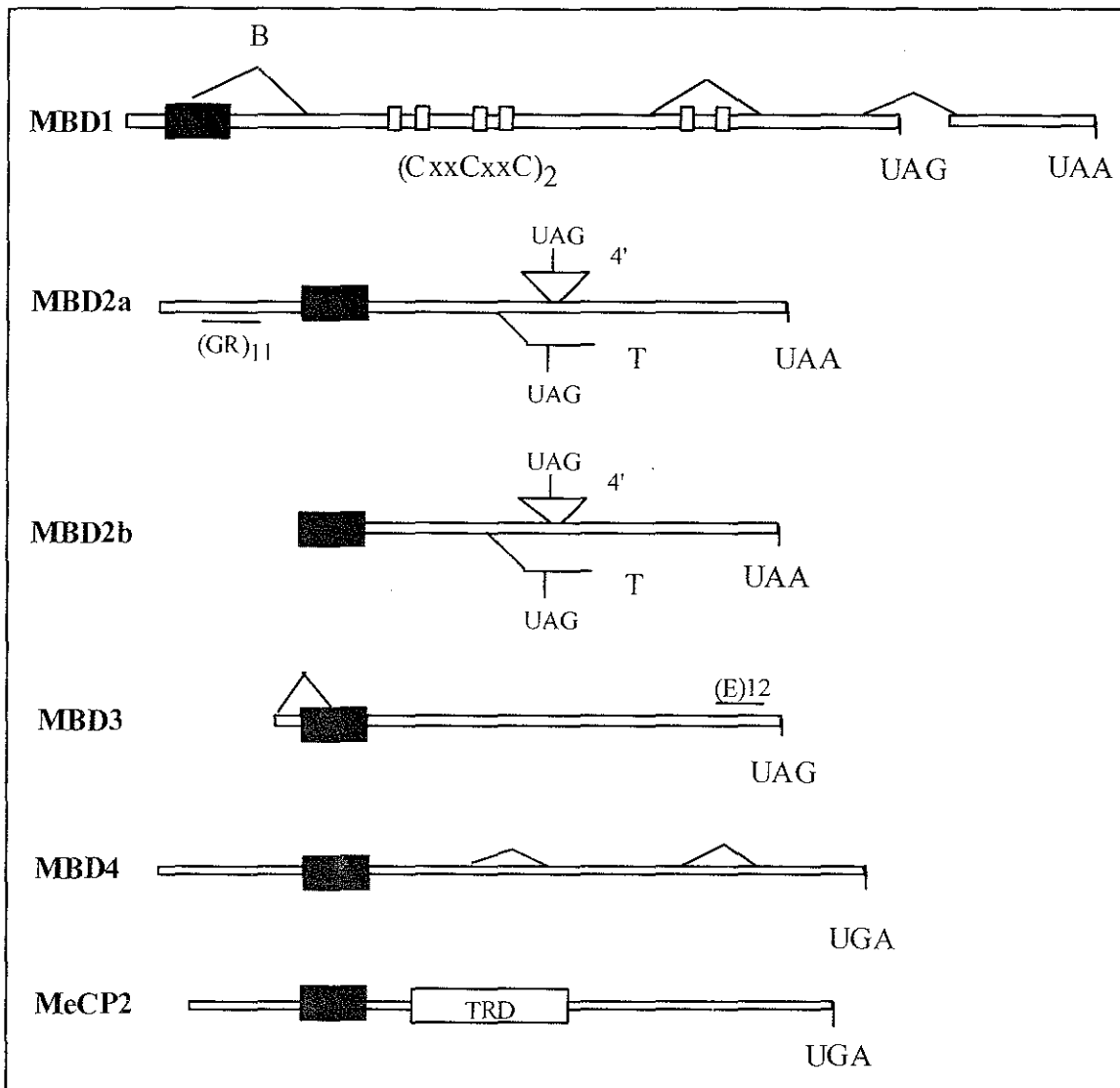


Figure 8 : La famille des protéines MBD

■ représente le MBD (Methyl binding domain)

∧ symbolise les différentes possibilités d'épissage

Les deux formes de la protéine MBD2 sont indiquées, elles correspondent à l'initiation de la transcription soit au premier (MBD2a), soit au second (MBD2b) codon méthionine.

Le gène MBD2 peut donner des transcrits plus courts si l'exon 4' ou l'exon 3 ne sont pas épissés. T représente le transcrit plus court, par insertion de l'exon3, spécifiquement synthétisé dans les testicules.

B représente le site d'épissage spécifiquement utilisé dans le cerveau pour le gène MBD1

TRD: domaine de répression transcriptionnelle

(GR)₁₁: motif de 11 glycine-arginine

(E)₁₂: motif de 12 acides glutamiques

Hendrich et al. (Hendrich et al. 1998) ont montré par des expériences de retard sur gel complétées par des études de transfection avec la GFP (green fluorescent protein) que MBD1, MBD2 et MBD4 peuvent former des complexes spécifiques avec l'ADN méthylé *in vitro*. Les protéines MBD ne nécessitent qu'une seule paire de dinucléotides 5'CpG3' méthylés symétriquement pour se lier à l'ADN. Ce n'est pas le cas du complexe MeCP1 qui a besoin de douze paires de cytosines méthylées pour interagir avec l'ADN (Meehan et al. 1989). Ces protéines se lient à l'ADN sous forme de monomères (Ballestar et al. 2001).

Malgré sa forte similarité de séquence protéique avec MBD2 (71,1%), seule MBD3 ne semble pas posséder la capacité de liaison avec l'ADN méthylé et montre un profil de localisation *in vivo* tout à fait différent. Le très haut degré de conservation entre les homologues murin et humain de ces deux protéines (97,6% d'identité en acides aminés pour MBD2, et 93,8% pour MBD3) laisse supposer que ces deux protéines sont fonctionnelles. Une hypothèse serait que le gène *MBD3* proviendrait d'une duplication du gène *MBD2* et qu'il aurait acquis une nouvelle spécificité de liaison à l'ADN méthylé non détectée par les approches utilisées (Hendrich et al. 1998). Une analyse de la séquence du motif MBD suggère une seconde hypothèse. Ohki et al. (Ohki et al. 1999) ont montré que des mutations au niveau d'acides aminés hautement conservés depuis le Xénope jusqu'à l'homme et entre les cinq MBD, pouvaient fortement gêner la capacité d'interaction de la protéine avec l'ADN méthylé. Deux résidus (en position 30, en position 34) différents dans le MBD du gène *MBD3* des mammifères par rapport aux résidus des autres membres de la famille MBD pourraient expliquer l'absence de liaison de la protéine MBD3 à l'ADN méthylé (Ballestar et al. 2001). Une mutagenèse ciblée au niveau de ces deux acides aminés (positions 30 et 34) permettrait de vérifier cette dernière hypothèse.

Une fois la capacité d'interaction de ces protéines avec l'ADN méthylé prouvée, il restait à comprendre comment ces protéines intervenaient pour réprimer la transcription. Mais avant cela, arrêtons nous un peu plus longuement sur la protéine MeCP2, archétype de la famille MBD et récemment impliquée dans le syndrome de Rett (Amir et al. 1999).

2.2.2. La protéine MeCP2

2.2.2.1. Organisation et expression du gène *MECP2*

Le gène *MECP2*, couvrant quelques 80 kb dans la région Xq28, est localisé entre les locus IRAK (interleukin I receptor-associated kinase) et RCP (red opsin), région synténique retrouvée sur le chromosome X de la souris (Dragich et al. 2000). Le gène est composé de quatre exons qui sont transcrits depuis le télomère vers le centromère du chromosome X. La région codante comprend 1461 nucléotides et se répartit entre les exons 2 et 4 (figure 9). La région 3' non traduite (3'-UTR, 3'-untranslated region) dans l'exon 4 est inhabituellement longue : 8,5 kb, et présente des parties bien conservées entre l'homme et la souris (Coy et al. 1999 ; Reichwald et al. 2000). Comme les autres gènes *MBD*, le gène *MECP2* est considéré comme un gène domestique du fait de son expression ubiquitaire et de son îlot 5'CpG3' en 5' (Reichwald et al. 2000). Il serait notamment régulé par le facteur de transcription Sp1. En effet, l'inactivation des deux allèles chez la souris ($Sp1^{-/-}$) abaisse le taux d'expression du gène *MECP2* (Marin et al. 1997), et plusieurs sites potentiels de fixation pour Sp1 ont été identifiés au sein de l'îlot 5'CpG3' du gène *MECP2* (Reichwald et al. 2000). Ce gène est soumis au processus d'inactivation chez la souris comme chez l'homme (Adler et al. 1995 ; D'Esposito et al. 1996).

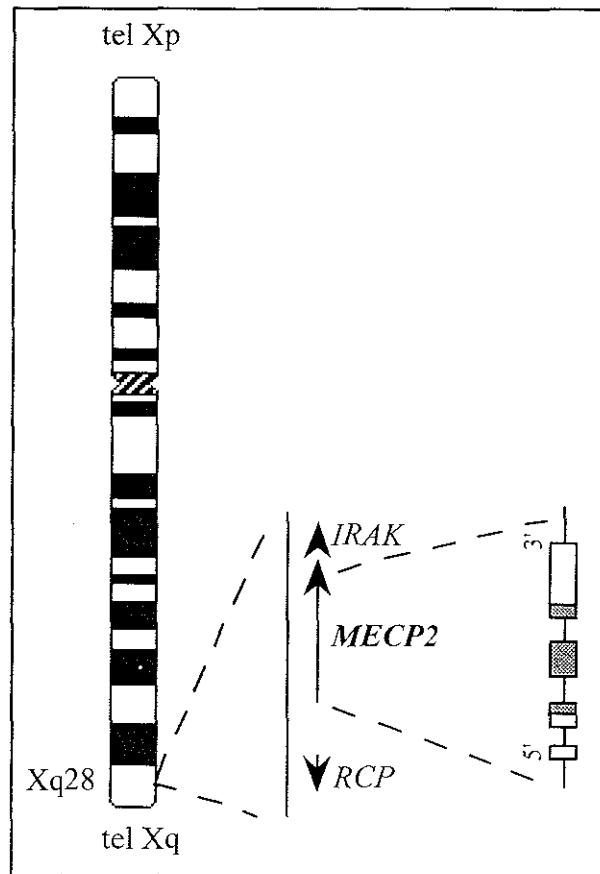


Figure 9 : Localisation et structure du gène *MECP2* sur le chromosome X
 Les flèches indiquent le sens de transcription des différents gènes *IRAK*, *MECP2* et *RCP*.
 Le gène *MECP2* est composé de quatre exons; la région codante se répartit entre les exons 2 et 4 (partie rouge).

Chez les mammifères, le gène *MECP2* est largement exprimé dans les tissus embryonnaires et adultes, avec toutefois, une expression à un niveau faible au début du développement (Meehan et al. 1992). La région 3'UTR contient cinq sites potentiels de polyadénylation dont l'utilisation fait l'objet de controverses (Coy et al. 1999 ; Reichwald et al. 2000). En effet, Coy et al ont détecté par Northern blot principalement trois transcrits d'environ 1,8 kb, 5 kb et 10,1 kb, alors que Reichwald et al. ont détecté deux transcrits majeurs de 1,9 kb et 10 kb et un plus faiblement exprimé supérieur à 7,5 kb. Tous ces transcrits ont été retrouvés dans la plupart des tissus étudiés bien qu'une variation d'expression tissu-spécifique ait été mise en évidence. Ils sont exprimés dans des quantités équivalentes dans le poumon, les reins, le tube digestif, les glandes thyroïdiennes et surrénales, alors que le plus petit transcrit est plus abondant au niveau des muscles cardiaque et squelettiques, du tissu lymphoïde, du foie et du placenta (D'Esposito et al. 1996 ; Coy et al. 1999 ; Reichwald et al. 2000). De plus, Coy et al ont mis en évidence une variation du profil d'expression des différents transcrits selon les

stades du développement. Alors que le transcrit de 10,1 kb est plus abondant dans le cerveau fœtal, le transcrit de 5 kb est plus abondant dans le cerveau adulte. Le rôle biologique de ces différents transcrits n'est pas encore élucidé, Reichwald a montré que les transcrits de 1,9 kb et 10 kb avaient la même durée de demi-vie (3-4 h), supposant alors que l'efficacité de leur traduction respective pouvait différer. Le haut degré de conservation entre l'homme et la souris de la région 3' non traduite suggère une fonction importante pour le plus long transcrit de 10 kb. L'expression de plusieurs autres gènes est connue pour être régulée au cours de l'étape de polyadénylation, mais les mécanismes moléculaires à l'origine de cette régulation ne sont pas élucidés (Edwards-Gilbert et al. 1997).

2.2.2.2. Structure et fonction de la protéine MeCP2

Le polypeptide MeCP2 contient un domaine de liaison aux cytosines méthylées (MBD) (Nan et al. 1993), un domaine de répression transcriptionnelle (TRD) (Nan et al. 1998 ; Jones et al. 1998) et deux signaux potentiels de localisation nucléaires (NLS) (Nan et al. 1996 ; Kudo et al. 1998). Le domaine MBD comprend 85 acides aminés du côté N-terminal, il est nécessaire et suffisant pour l'interaction avec l'ADN méthylé en présence ou en absence de chromatine assemblée. Par ailleurs, des expériences d'immunofluorescence *in situ* chez la souris ont montré que la protéine MeCP2 se localisait tout le long des chromosomes mais était plus particulièrement concentrée au niveau de l'hétérochromatine péricentromérique connue pour être hautement méthylée (Nan et al. 1993).

La résolution du domaine MBD par résonance magnétique nucléaire a révélé une nouvelle structure en forme de "coin" (wedge-shaped structure). Cette structure consiste en quatre feuillets β antiparallèles et en une boucle avec une poche hydrophobe prédite pour interagir avec la paire de cytosines méthylées dans le sillon majeur de l'ADN (Wakefield et al. 1999).

Le domaine TRD a été délimité par Nan et al. (Nan et al. 1997) entre les acides aminés 207 et 310 en remplaçant le MBD par le domaine de liaison à l'ADN de GAL4 fusionné à différentes protéines MeCP2 délétées et en testant la fonction répressive de la protéine de fusion ainsi formée sur un gène rapporteur présentant la séquence de liaison de GAL4.

Une fois liée aux cytosines méthylées, la protéine MeCP2 réprime la transcription par l'intermédiaire d'un complexe multiprotéique comprenant le co-répresseur Sin3A et les HDAC1 et 2 qui modifient la structure de la chromatine en une forme non accessible à la transcription. En d'autres termes, la méthylation des cytosines induit la désacétylation des

histones, le remodelage de la chromatine et donc la répression transcriptionnelle (Nan et al. 1998 ; Jones et al. 1998) (figure 10).

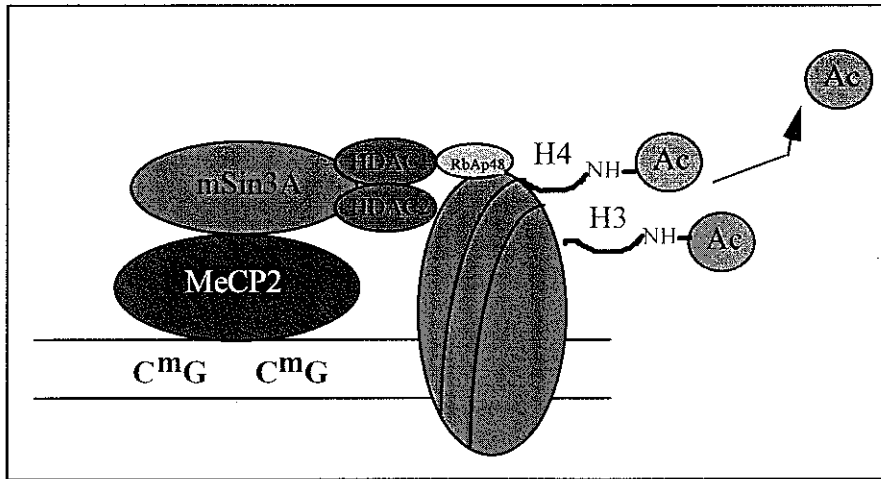


Figure 10 : Complexe de répression transcriptionnelle via MeCP2 (d'après Razin, 1998)

NHAc: groupes amino-acétyl liés aux résidus lysines H3 et H4.

C^mG : cytosine méthylée d'un dinucléotide 5'CpG3'

MeCP2 établit ainsi un lien fondamental entre la méthylation de l'ADN et la désacétylation des histones, ce qui constitue une étape importante dans notre compréhension des mécanismes de répression épigénétique des gènes chez les vertébrés.

Récemment, deux articles (Yu et al. 2000 ; Kaludov et al. 2000) ont montré que la protéine MeCP2 pouvait réprimer la transcription indépendamment des HDAC. Cette (ou ces) voie(s) alternative(s) a (ont) été supposée(s) suite à quelques unes des observations suivantes : i) MeCP2 peut réprimer la transcription à partir d'ADN nu dans un système *in vitro*, c'est-à-dire en absence d'histones (Nan et al. 1997) ; ii) la répression transcriptionnelle par MeCP2 du gène de la b-actine n'est que partiellement éliminée par la trichostatine (TSA, inhibiteur de toutes les HDAC connues) (Jones et al. 1998 ; Nan et al. 1998 ; Cameron et al. 1999) ; et iii) une déméthylation du promoteur inactif du gène *FMR1* dans les cellules des patients atteints du syndrome de l'X-fragile entraîne une activation du gène, alors que l'acétylation des histones H3 et H4 par l'action seule de la TSA ne permet pas la ré-expression du gène.

La première publication (Yu et al. 2000) démontre par des expériences de transfection que la protéine MeCP2 peut réprimer la transcription de manière HDAC-dépendante ou HDAC-indépendante en fonction du promoteur étudié. En effet, MeCP2 exerce une forte répression

sur le promoteur du virus SV40 qui ne peut être éliminée par la TSA, alors qu'une plus faible répression exercée sur le promoteur AdML (Adenovirus major late promotor), en utilisant le même plasmide rapporteur et dans le même contexte cellulaire, est totalement levée par la TSA. De plus, par une étude mutationnelle du domaine TRD de la protéine MeCP2, les auteurs ont mis en évidence une région impliquée spécifiquement dans la voie de répression HDAC-indépendante, et la participation du co-répresseur mSin3A dans cette voie. Dans cette région, une mutation inactive totalement l'effet répresseur sur le promoteur SV40 du fait de l'inhibition de l'interaction entre MeCP2 et mSin3A, cette même mutation ne lève pas l'effet répresseur sur le promoteur AdML.

Dans la seconde publication (Kaludov et al. 2000), les auteurs ont mis au point et utilisé un système de transcription simplifié dans lequel facteurs de transcription purifiés, ADN nu et protéine MeCP2 sont ajoutés "manuellement". Ils ont validé leur système en retrouvant la répression ciblée et à distance (jusqu'à 500 pb) par MeCP2 au niveau d'un ADN méthylé. Leur système a permis de montrer que MeCP2 pouvait interférer avec la transcription en empêchant la formation du complexe d'initiation par une interaction directe de MeCP2 avec TFIIB, composant essentiel de la machinerie basale de transcription, et, ceci, indépendamment des HDAC, puisqu'elles sont absentes du système.

Enfin, MeCP2 agirait non seulement comme un répresseur transcriptionnel dynamique mais aussi jouerait un rôle dans l'architecture de la chromatine inactivée en se liant aux sites d'attachements de la matrice nucléaire (MAR/SAR). En effet, l'homologue de MeCP2 chez le poulet est la protéine ARBP (Attachment region binding protein) qui présente une haute affinité pour ces régions (MAR/SAR) (Weitzel et al. 1997).

En conclusion, la protéine MeCP2 joue certainement un rôle central dans plusieurs voies de répression transcriptionnelle. On peut supposer que le choix de la voie est influencé par le type cellulaire, la séquence d'ADN et par l'architecture locale de la chromatine (Wade, 2001). MeCP2 est la première protéine qui a permis d'établir un lien entre deux grandes modifications épigénétiques de la chromatine (désacétylation et méthylation) et la répression transcriptionnelle. La découverte, par la suite, de rôles similaires pour les autres membres de la famille MBD témoigne, sans doute, de la très grande complexité de la régulation des gènes dans la cellule eucaryote à l'origine de la spécificité cellulaire.

2.2.2.3. Fonction des autres MeCP

Des approches similaires ont été utilisées pour rechercher l'implication des autres membres de la famille MBD dans la répression transcriptionnelle. Les protéines MBD1, 2 et 3 sont aussi associées à des HDAC.

MBD1 réprime la transcription grâce à une coopération entre son domaine MBD, ses motifs riches en cystéines (motifs CxxC analogues à ceux décrits dans DNMT1 et dans la protéine ALL-1 ou HRX du groupe trithorax de la Drosophile), son domaine TRD (localisé dans la partie C-terminale de la protéine et riche en résidus hydrophobes à la différence du TRD hautement basique de MeCP2 (Fujita et al. 2000)) ainsi que des HDAC. Toutefois, ni HDAC1, ni Sin3 ne sont impliqués dans le mécanisme de répression par MBD1. Ce mécanisme pourrait constituer une autre voie d'inactivation transcriptionnelle via les HDAC mais indépendante de celle de MeCP2 (Ng et al. 2000). Enfin, MBD1 possède la particularité de réprimer la transcription indépendamment de la méthylation. En effet, un des trois motifs CxxC peut se lier à l'ADN non méthylé (Fujita et al. 2000).

Le complexe MeCP1 agirait par le biais d'un complexe contenant les HDAC1, 2 et les protéines de liaison aux histones RbAp48 et RbAp46 ainsi que probablement MBD2 comme protéine de liaison à l'ADN (Ng et al. 1999). Toutefois, il existe une controverse quant au rôle de MBD2. En effet, Bhattacharya et al. (Bhattacharya et al. 1999) ont identifié une activité de déméthylase pour la forme tronquée MBD2b de la souris. Ng et al. n'ont pas retrouvé cette activité, et montrent par ailleurs son rôle de répresseur dans le complexe MeCP1.

MBD3 a été aussi identifiée comme composant d'un complexe HDAC avec en plus une activité de remodelage des nucléosomes (Mi2/NuRD) (Wade et al. 1999).

Seule MBD4 n'a pas été retrouvée associée avec des HDAC. Dès sa découverte en 1998, Hendrich et al. (Hendrich et al. 1998) ont montré que la protéine MBD4 possédait des similarités de séquences avec les enzymes bactériennes de réparation de l'ADN. En 1999, Hendrich a mis en évidence que MBD4 (rebaptisée thymine glycosylase MBD4) était capable d'éliminer efficacement une thymine ou un uracile appartenant à un mésappariement 5'TpG3'/3'Gp^mC5' *in vitro*. Le MBD de MBD4 se lie préférentiellement au mésappariement 5'TpG3'/3'Gp^mC5' résultant de la désamination de la cytosine méthylée en thymine. Cette protéine MBD4 serait bien impliquée dans la réparation de l'ADN et diminuerait donc la fréquence des mutations au niveau des cytosines méthylées.

D. Objet de notre étude

La recherche de l'origine génétique du syndrome de Rett pendant une quinzaine d'années, a abouti à l'implication du gène *MECP2* apportant ainsi un marqueur biologique indispensable au diagnostic moléculaire de la maladie et donc au conseil génétique. Au cours de ce travail, nous avons recherché les mutations au sein du gène *MECP2* chez 270 patients non apparentés. Nous avons, pour cela, développé et utilisé un panel d'approches complémentaires qui ont confirmé l'implication du gène *MECP2* dans le syndrome de Rett et mis en évidence différents mécanismes mutationnels de ce gène. En effet, nous montrons que des mutations dans le gène *MECP2* peuvent être en mosaïque somatique, et que des réarrangements sur plusieurs centaines de paires de bases existent dans ce gène. De plus, nous avons recherché des remaniements au niveau chromosomique du locus *MECP2*.

RESULTATS

I. Présentation des patients

Nous avons étudié 270 patients (255 filles et 15 garçons) dont les prélèvements sanguins nous ont été adressés par le biais de l'association française du syndrome de Rett, ou par des neuropédiatres depuis toute la France et la Belgique. Le laboratoire fait partie du consortium national de diagnostic moléculaire du syndrome de Rett regroupant au total quatre laboratoires : le laboratoire de Biochimie et Génétique Moléculaire au CHU Cochin, Paris (Dr T. Bienvenu), le laboratoire d'Hormonologie et de Biologie Moléculaire au CHU Bicêtre, Paris (Dr N. de Roux), INSERM U491-Génétique Médicale et Développement au CHU La Timone, Marseille (Dr L. Villard) et le laboratoire de Génétique au CHU Brabois, Nancy (Pr P. Jonveaux).

Parmi les 255 patientes étudiées, nous pouvons distinguer deux catégories principales : une regroupant les patientes avec un syndrome de Rett typique (92) et l'autre celles présentant un phénotype atypique (140). Il est à noter que le terme "atypique" désigne ici toutes les patientes n'ayant pas un syndrome de Rett classique (et pas seulement les patientes présentant le variant dénommé syndrome de Rett atypique défini par Hagberg et al. (Hagberg et al. 1994)). Il est parfois difficile de poser précisément le diagnostic de syndrome de Rett, étant donné l'aspect clinique hétérogène et évolutif de cette pathologie. Pour 23 patientes, nous n'avons eu aucun renseignement clinique.

Par ailleurs, 15 garçons ont été analysés. Les phénotypes de ces garçons sont très variables. Trois nous ont été adressés dans le contexte d'un retard mental familial pour lequel les données généalogiques sont compatibles avec une transmission liée au chromosome X. Les autres patients montrent un retard global des acquisitions motrices associé à un retard mental et une absence de langage. Sont notés parfois une microcéphalie (5 cas), des convulsions (4 cas), des stéréotypies (3 cas) et un retard de croissance intra-utérin (2 cas).

II. Recherche de mutations dans le gène *MECP2* à l'échelle moléculaire

1. Première stratégie de criblage du gène *MECP2* : en trois étapes

Dans un premier temps, nous avons choisi la technique de CSGE (Conformation Sensitive Gel Electrophoresis). Pour réaliser cette technique, nous avons eu recours aux couples d'amorces précédemment publiés par Amir et collaborateurs (Amir et al. 1999). La séquence du gène *MECP2* est divisée en huit fragments chevauchants permettant l'amplification de toute la région codante du gène ainsi que les zones de jonction exon/intron (figure 11). Etant donné la petite taille de la région codante du gène *MECP2* (1461 nucléotides), et sachant que la technique de CSGE, comme toute technique de précriblage, ne détecte pas l'ensemble des mutations, nous avons décidé de séquencer systématiquement la séquence codante du gène *MECP2* pour les patients qui ne présentaient aucune mutation délétère détectée par CSGE. Parallèlement, nous avons mis au point une approche par Southern blot afin de mettre en évidence des remaniements de taille plus importante non détectés par la CSGE et le séquençage.

Ainsi, la recherche de mutations dans le gène *MECP2* s'est articulée, en premier lieu, autour de trois approches : i) précriblage par la technique de CSGE, ii) séquençage, iii) technique de Southern blot.

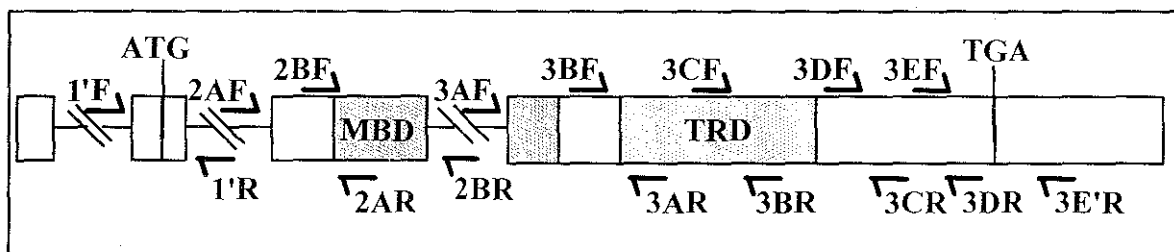


Figure 11 : Position des 8 couples d'amorces utilisés pour la CSGE et le séquençage systématique, sur le gène *MECP2*.

1.1. La technique de CSGE

Initialement décrite par Ganguly et al. (Ganguly et al. 1993), la CSGE détecte les variations de séquence hétérozygotes présentes dans une région préalablement amplifiée par PCR. En effet, des modifications de séquence à l'état hétérozygote engendrent la formation d'hétéroduplex par l'appariement entre un brin du produit de PCR correspondant à l'allèle normal et le brin complémentaire de l'amplicon dérivé de l'allèle muté. Le mésappariement présent au sein des hétéroduplex modifie la structure spatiale de ces molécules d'ADN double-brins et affecte donc leur migration dans un gel d'acrylamide semi-dénaturant. Une diminution de la mobilité des hétéroduplex est observée par rapport à celle des homoduplex. Théoriquement, quatre bandes correspondant aux deux hétéroduplex et aux deux homoduplex sont obtenues (figure 12). Dans la pratique, la séparation des deux hétéroduplex (et des deux homoduplex) est rarement visualisable sur le gel d'acrylamide. Ainsi, deux bandes sont généralement révélées, une caractérisant les deux hétéroduplex, et l'autre les homoduplex. Chez les garçons, du fait de l'hémizygotie pour le chromosome X, un mélange du produit de PCR du patient avec celui d'un témoin garçon est nécessaire. Ce mélange subit une dénaturation à 98°C suivie d'une réassociation à 68°C permettant la formation d'hétéroduplex "artificiels" s'il y a une différence de séquence entre les deux individus.

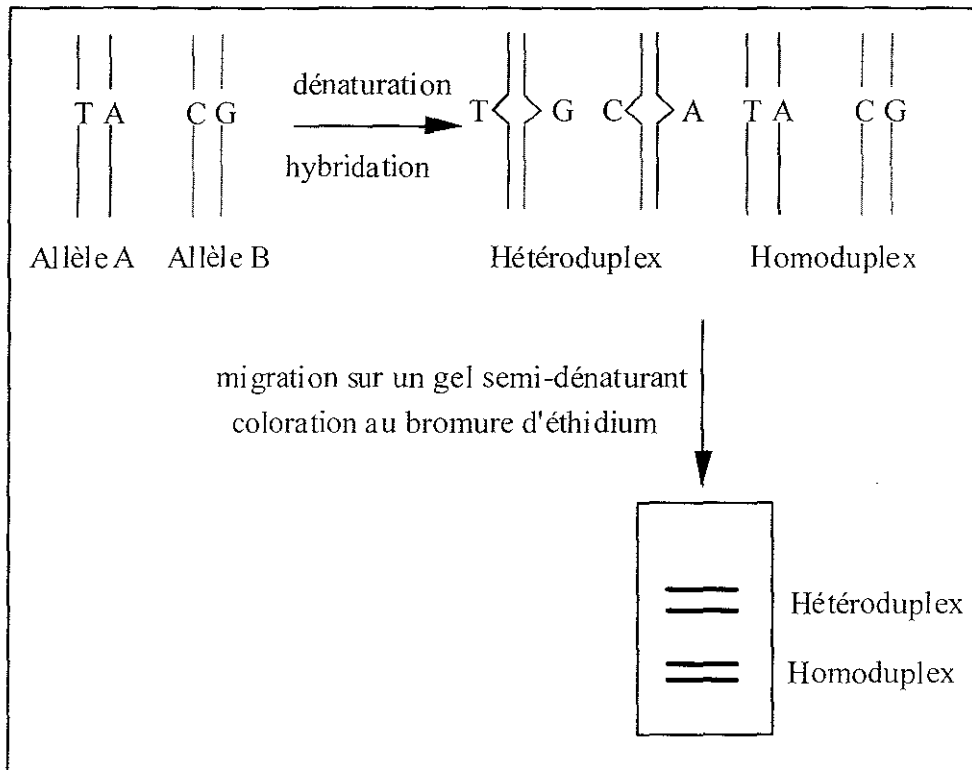


Figure 12 : *Principe de la CSGE*

1.2. Séquençage et interprétation des variations nucléotidiques

La nature (mutation délétère ou polymorphisme neutre) des modifications de séquence détectées par CSGE (et par toute technique de précriblage) ainsi que leur position sont déterminées par le séquençage du produit de PCR.

L'interprétation des résultats du séquençage peut être plus ou moins évidente. Nous pouvons regrouper les résultats en deux catégories : celle pour laquelle l'interprétation en mutation délétère ou en mutation silencieuse, est évidente (catégorie 1), et celle pour laquelle il est difficile de conclure (catégorie 2).

Dans la catégorie 1, nous rassemblons les mutations non-sens, les mutations entraînant un décalage du cadre de lecture et les mutations silencieuses. Les mutations non-sens ou les délétions/insertions décalant le cadre de lecture entraînent l'apparition d'un codon stop précoce. La protéine potentiellement synthétisée sera tronquée et pour les délétions/insertions, ne contiendra pas les bons acides aminés en aval du codon stop prématuré. Elle sera donc certainement non fonctionnelle. Le caractère délétère est hautement probable dans ce cas-là, bien qu'un doute subsiste si le codon stop apparaît à proximité du codon stop naturel

(Mazoyer et al. 1996). Les mutations silencieuses qui ne causent aucune modification de la séquence protéique, résultent de la dégénérescence du code génétique au niveau de la troisième position du codon de certains acides aminés. Elles peuvent, dans de rares cas, entraîner une modification de l'épissage.

La catégorie 2 comprend les mutations faux-sens et les délétions/insertions dont le nombre de nucléotides impliqués est un multiple de trois. Dans le cas des mutations faux sens, plusieurs approches complémentaires sont possibles afin de conclure à un polymorphisme ou à une mutation délétère. La première consiste à se référer aux données de la littérature afin de vérifier si l'anomalie a déjà été décrite en tant que mutation ou polymorphisme. En parallèle, il faut rechercher cette anomalie chez les parents. Si aucun des deux parents ne porte ce variant, c'est en faveur d'une mutation *de novo* délétère. Si un des deux parents le porte, l'interprétation dépend du mode de transmission de la maladie (dominant, récessif, lié au chromosome X). Si le changement d'acide aminé n'est pas décrit dans la littérature, et en plus de la recherche familiale, un séquençage de la région incriminée chez une centaine d'individus de la population générale peut permettre de répondre à la question. En effet, si l'anomalie est retrouvée ne serait-ce qu'une seule fois dans la population générale, on conclut à un polymorphisme rare. En revanche, l'absence de la modification chez des témoins n'exclut ni un polymorphisme rare ni une mutation. Si le gène est conservé entre les espèces, et/ou s'il appartient à une famille de gènes, on peut comparer les séquences protéiques et vérifier si l'acide aminé affecté par la mutation est conservé ou non. Si c'est le cas, c'est un argument supplémentaire en faveur d'une mutation.

Les délétions (ou insertions) comprenant un nombre de nucléotides multiple de trois n'entraînent pas de décalage du cadre de lecture. Toutefois, elles peuvent ou non se produire en phase avec le cadre de lecture du gène. Si elles se produisent en phase, la protéine potentiellement synthétisée présentera une délétion (ou une insertion) proportionnelle au nombre de nucléotides délétés (ou insérés) sans qu'il y ait de modification de la séquence en aval de la délétion (ou de l'insertion). Si la délétion/insertion comprenant un nombre de nucléotides multiple de trois, ne se produit pas en phase avec le cadre de lecture, en plus de la délétion (ou de l'insertion), l'acide aminé faisant la jonction risque d'être modifié. Dans les deux cas, la conséquence sera plus ou moins grave en fonction du nombre de nucléotides délétés ou insérés et/ou en fonction de la région atteinte (domaine fonctionnel ou catalytique par exemple).

Pour toutes les mutations de la catégorie 2, seules des études fonctionnelles de la protéine anormale donneront une réponse plus proche de la réalité quant à la conséquence de la

mutation. De telles études fonctionnelles sont envisagées en dernier recours, dans le cas de maladies fréquentes (comme la mucoviscidose) pour lesquelles il est indispensable de savoir si la mutation a un effet délétère ou non afin de confirmer ou d'infirmier le diagnostic et surtout pour la réalisation d'un diagnostic prénatal.

1.3. Les résultats obtenus par CSGE et séquençage

Par CSGE et séquençage systématique, nous avons détecté 27 mutations et 12 polymorphismes différents chez 84 patients. 75% des 27 mutations et des 12 polymorphismes ont été détectés par CSGE.

Ces 27 mutations se répartissent aussi bien en **mutations faux-sens** (7) et **non-sens** (6) dues à des substitutions, en **mutation à un site d'épissage** (1) qu'en **mutations décalant le cadre de lecture (mutations frameshift)** (13) dues à des délétions/insertions. Dans notre série, la CSGE détecte toutes les altérations dues à une insertion ou à une délétion. Par contre, cette technique n'a mis en évidence que 62,5% (15/24) des substitutions (mutations et polymorphismes confondus). Ce taux faible de détection peut s'expliquer par le fait que la CSGE est sensible à la position de l'anomalie dans le fragment (si elle est proche d'une extrémité, elle peut ne pas être détectée), et à la séquence intrinsèque du produit de PCR. En effet, une variation nucléotidique dans une séquence riche en bases G et C sera plus difficile à mettre en évidence car la déstabilisation engendrée par le mésappariement modifiera moins facilement la conformation spatiale des hétéroduplex et donc leur migration par rapport à celle des homoduplex. Le tableau XIII résume l'ensemble des mutations et des polymorphismes identifiés au sein du gène *MECP2* chez des patientes atteintes d'un syndrome de Rett ainsi que leur récurrence.

Tableau XIII : Mutations (A) et polymorphismes (B) identifiés dans le gène *MECP2* par CSGE et séquençage

L'astérisque désigne les mutations ou polymorphismes détectés par CSGE. Les mutations ou polymorphismes sont écrits selon les recommandations de Antonarakis et al. (Antonarakis et al. 1998)

A.

Mutations		Réurrence
faux-sens	316C>T, R106W*	4
	438C>G, D156E	1
	473C>T, T158M*	5
	481G>T, G161W*	1
	905C>G, P302R*	1
	916C>T, R306C	5
	917G>A, R306H	1
non-sens	146C>A, S49X	2
	502C>T, R168X	7
	508C>T, Q170X	1
	763C>T, R255X*	7
	808C>T, R270X	4
	880C>T, R294X*	8
à un site d'épissage	IVS3-2A>C*	1
délétion/insertion	117-118insA*	1
	439delG*	1
	480delT*	1
	752-753insCGCCC*	1
	806delG*	2
	808delC*	1
	1150-1193del*	1
	1155-1157insT*	1
	1155-1184del*	1
	1157-1188del*	1
	1158-1201del*	2
	1164-1207del*	2
1454-1457del	1	

B.

Polymorphismes	Réurrence
210C>T, S70	1
IVS3+22G>C*	1
IVS3-18insG*	3
IVS3-61G>C*	1
582C>T, S194*	1
897C>T, T299*	1
1070C>T, S357	1
1189G>A, E397K*	1
1197C>T, P399	1
1330G>A, A444T*	1
1470G>A*	1
1560-1561insT*	1

Nous avons démontré, lorsque nous disposions de l'ADN des parents (70% des cas) qu'il s'agissait toujours d'une néo-mutation, confirmant le caractère sporadique du syndrome de Rett.

1.3.1. Les mutations faux-sens

Toutes les mutations faux-sens excepté la mutation G161W, ont été décrites de manière récurrente et par différents auteurs (Amir et al. 1999 ; Wan et al. 1999 ; Amano et al. 2000 ; Amir et al. 2000 ; Bienvenu et al. 2000 ; Buyse et al. 2000 ; Cheadle et al. 2000 ; De Bona et al. 2000 ; Hampson et al. 2000 ; Hupkke et al. 2000 ; Xiang et al. 2000 ; Auranen et al. 2001 ; Giunti et al. 2001 ; Hoffbuhr et al. 2001 ; Inui et al. 2001 ; Laccone et al. 2001 ; Sung Lae Lee et al. 2001 ; Monros et al. 2001 ; Nicolao et al. 2001 ; Nielsen et al. 2001-a ; Yamada et al. 2001 ; Yamashita et al. 2001 ; Vacca et al. 2001 ; Zappella et al. 2001 ; Bourdon et al. 2001-a). Elles affectent toutes un domaine fonctionnel (MBD ou TRD) de la protéine MeCP2, ce qui argumente quant à leur implication dans la pathogenèse du syndrome de Rett.

La mutation G161W *de novo*, décrite pour la première fois par notre étude, entraîne le changement du résidu hydrogène de la glycine en un hétérocycle (noyau indole) du tryptophane. On peut supposer que l'encombrement stérique résultant de cette substitution est à l'origine d'une modification conformationnelle de la protéine MeCP2. Par ailleurs, la glycine appartient à la séquence consensus du MBD caractérisant la famille des protéines se

liant aux cytosines méthylées (Wakefield et al. 1999). Enfin, cet acide aminé est conservé depuis le xénope jusqu'aux mammifères, et un séquençage de 100 chromosomes d'individus témoins n'a jamais révélé une telle substitution à la position 161 de la protéine. Il semble donc licite de penser, devant ce faisceau d'arguments, que ce changement d'acide aminé G161W est bien une mutation délétère.

1.3.2. Les mutations non-sens et les mutations perturbant le cadre de lecture

Les mutations non-sens et les mutations à l'origine d'un décalage du cadre de lecture provoquent directement ou indirectement l'apparition d'un codon stop prématuré. Selon la position de ce codon stop prématuré, l'ARNm synthétisé sera reconnu et dégradé par le système de dégradation des ARNm (NMD : nonsense-mediated mRNA decay, Maquat 1996 ; Culbertson 1999) empêchant ainsi la synthèse de protéines trop incomplètes.

Nous avons identifié six mutations non-sens différentes retrouvées chez 29 patientes ; toutes (R168X, Q170X, R255X, R270X et R294X) excepté la mutation S49X, ont été maintes fois décrites dans la littérature (Amir et al. 1999 ; Wan et al. 1999 ; Amano et al. 2000 ; Bienvenu et al. 2000 ; Buyse et al. 2000 ; Cheadle et al. 2000 ; De Bona et al. 2000 ; Hampson et al. 2000 ; Hupkke et al. 2000 ; Xiang et al. 2000 ; Auranen et al. 2001 ; Laccone et al. 2001 ; Nielsen et al. 2001-a ; Vacca et al. 2001 ; Bourdon et al. 2001-a).

Nous pouvons classer les mutations perturbant le cadre de lecture en deux catégories différentes : d'un côté les mutations qui n'impliquent que quelques nucléotides (<5 pb), et de l'autre, celles qui impliquent quelques dizaines de nucléotides. A l'exception de la mutation 1454-1457del, toutes les mutations de la première catégorie touche un domaine fonctionnel de la protéine. La mutation 806delG a été décrite une dizaine de fois dans la littérature (Wan et al. 1999 ; Cheadle et al. 2000 ; Giunti et al. 2001 ; Nicolao et al. 2001 ; Vacca et al. 2001 ; Bourdon et al. 2001-a). La mutation 1454-1457del présente la particularité de se situer à moins de 10 pb du codon stop du gène *MECP2*. Nous n'avons pu déterminer les bornes exactes de cette délétion étant donné que les quatre nucléotides potentiellement délétés appartiennent à une répétition AG. Quelle que soit la position de la délétion, elle provoque l'insertion de 25 acides aminés au sein de la protéine MeCP2 (figure 13).

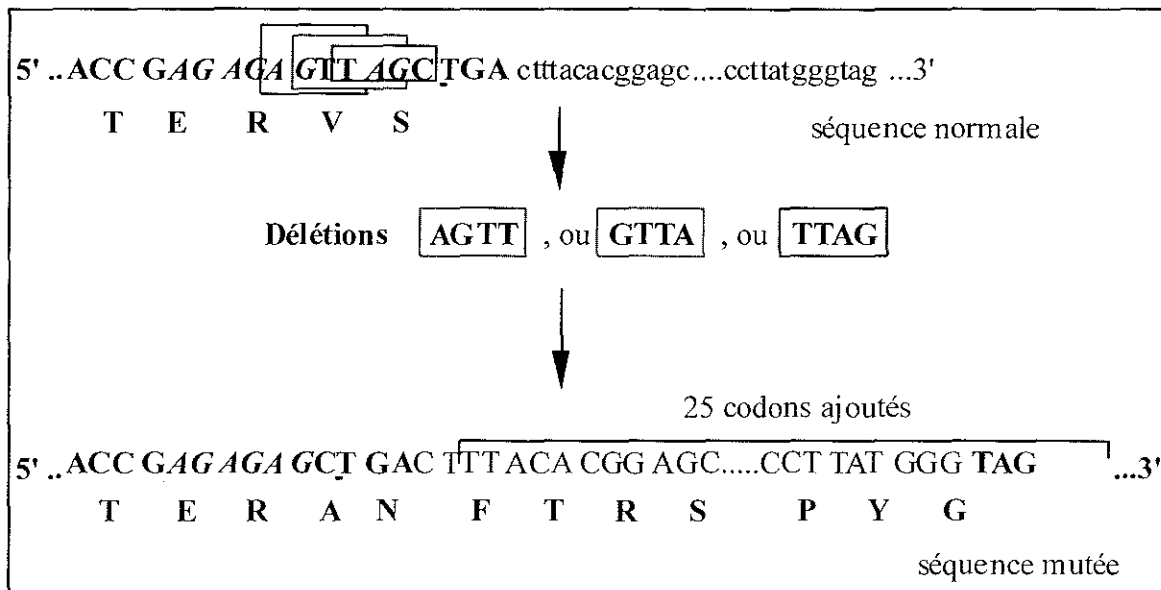


Figure 13 : Mutation 1454-1457 del

Le codon stop TGA de la séquence normale est souligné marquant la fin de la séquence traduite (en majuscule). Du fait de la répétition de quatre AG (notée en italique), les bornes de la délétion ne peuvent être déterminées exactement. La séquence mutée se termine par le codon stop TAG (en gras) situé 25 codons en aval du codon stop de la séquence normale (souligné). Est indiquée pour chaque séquence nucléotidique codante la séquence protéique correspondante.

La cytosine en position 808 peut être impliquée soit dans la mutation non-sens R270X du fait de sa substitution par une thymine, soit dans la mutation perturbant le cadre de lecture 808delC par sa délétion.

La deuxième catégorie de mutations à l'origine de décalages du cadre de lecture constitue une caractéristique mutationnelle du gène *MECP2*. En effet, cette deuxième catégorie est composée dans notre série, de cinq délétions de 26 à 43 pb qui affectent une région précise du gène (entre les positions 1150 et 1164) située après les deux domaines fonctionnels. Des délétions similaires ont été publiées dans la quasi totalité des articles relatant une analyse mutationnelle du gène *MECP2* (Amano et al. 2000 ; Bienvenu et al. 2000 ; Buyse et al. 2000 ; Cheadle et al. 2000 ; De Bona et al. 2000 ; Hampson et al. 2000 ; Hupkke et al. 2000 ; Xiang et al. 2000 ; Auranen et al. 2001 ; Giunti et al. 2001 ; Hoffbuhr et al. 2001 ; Inui et al. 2001 ; Laccone et al. 2001 ; Sung Jae Lee et al. 2001 ; Monros et al. 2001 ; Nicolao et al. 2001 ; Nielsen et al. 2001-a ; Yamada et al. 2001 ; Yamashita et al. 2001 ; Vacca et al. 2001 ; Zappella et al. 2001 ; Bourdon et al. 2001-a).

1.3.3. Une mutation à un site d'épissage

De part et d'autre d'un intron, au niveau des jonctions exon/intron se trouvent des séquences consensus de quelques paires de bases : le site donneur d'épissage en 5' de l'intron (**GUAAGU**), le site accepteur d'épissage en 3' (**PyNCAG**). En plus de ces deux séquences, intervient dans la formation de la boucle au cours de la réaction d'épissage, le site de branchement (**PyNPyPyPuAPy**) situé une vingtaine de bases en amont de l'extrémité 3' de l'intron. A l'exception des cinq bases **GU, A, AG** qui sont conservées dans 100% des ARNm, et dont une modification est délétère, une altération au niveau d'une des bases plus ou moins conservées peut également engendrer un mauvais épissage de l'intron responsable d'un décalage du cadre de lecture et donc de l'introduction d'un codon stop précoce (figure 14).

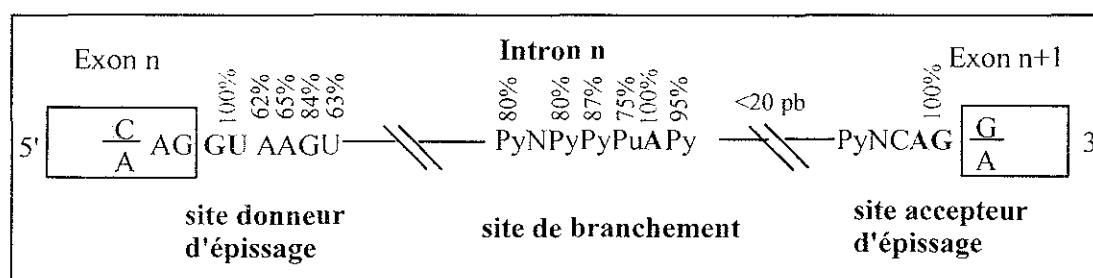


Figure 14 : Séquences consensus des jonctions exon/intron et du site de branchement. Les pourcentages indiquent la fréquence à laquelle la base est retrouvée dans différents ARNm.

Nous avons identifié une mutation au site accepteur d'épissage de l'intron 3 du gène *MECP2*, il s'agit de la transversion IVS3-2A>C. Afin de déterminer la conséquence de cette mutation, nous avons amplifié par RT-PCR, à partir de l'ARNm extrait d'une lignée lymphoblastique de la patiente, un fragment de 526 pb (par le couple d'amorces 2BF/3AR, figure 11) comprenant la jonction entre les exons 3 et 4. Le séquençage du produit d'amplification a révélé l'existence de deux populations d'ADNc chez cette patiente. En effet, nous avons retrouvé la séquence normale, et une séquence comportant une délétion de 7 pb. L'analyse de cette séquence délétée a montré que la mutation engendrait l'utilisation d'un site accepteur d'épissage cryptique localisé 7 pb en aval dans l'exon 4. Ainsi, pour l'allèle délété, le spliceosome ne peut utiliser le site normal du fait de la mutation, mais utilise un site cryptique 7 pb en aval entraînant un décalage du cadre de lecture sur dix codons avant l'introduction d'un codon stop précoce (figure 15). Si l'ARNm n'est pas détruit par "le système de

dégradation" des ARNm (nonsense-mediated mRNA decay, Maquat 1996 ; Culbertson 1999), ce qui est probablement le cas étant donné que nous amplifions les deux populations d'ARNm par RT-PCR, la protéine résultante se termine par un domaine MBD tronqué.

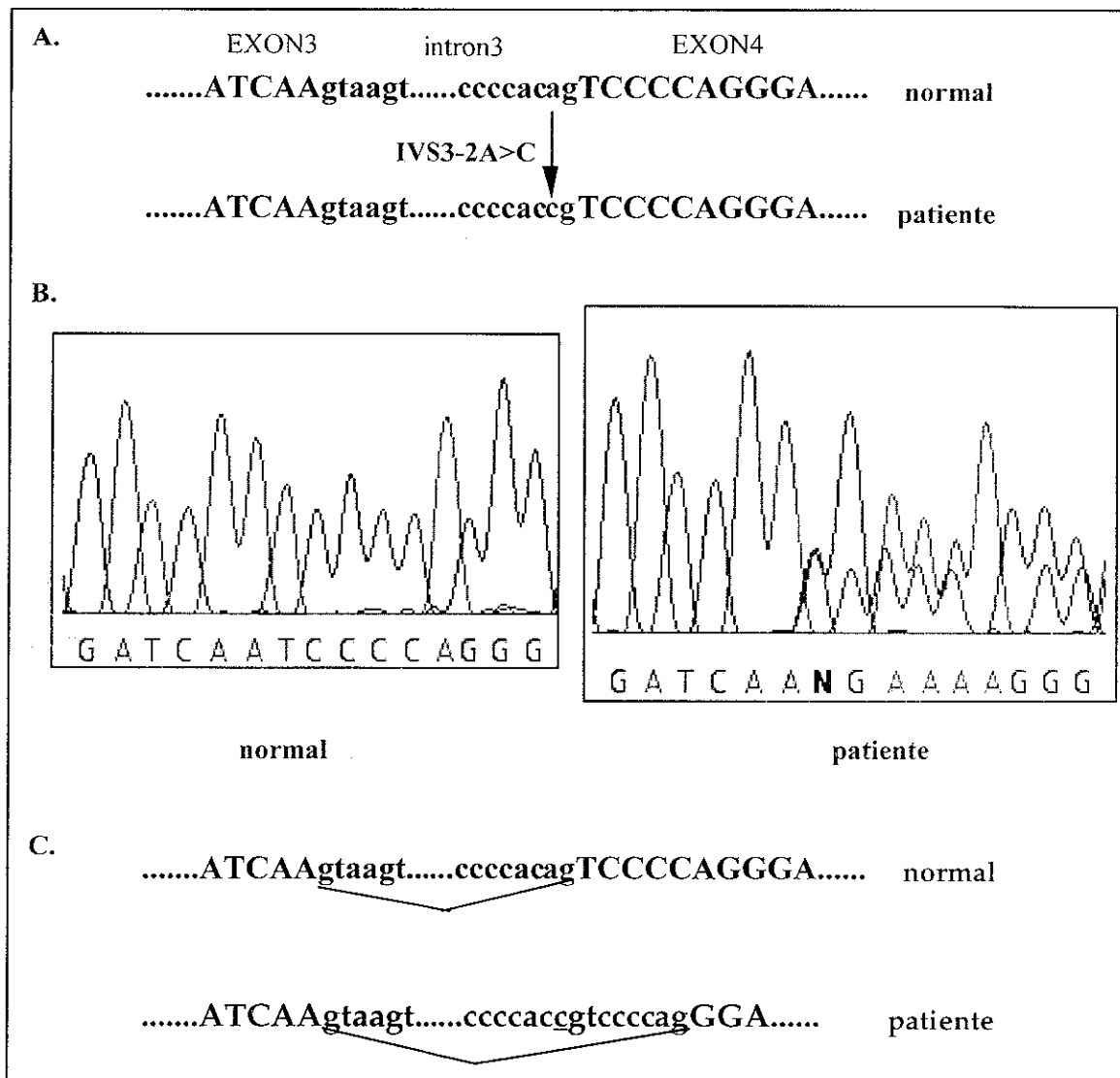


Figure 15 : Caractérisation de l'effet de la mutation IVS3-2A>C sur la transcription du gène MECP2

A. Identification de la transversion IVS3-2A>C au site accepteur d'épissage de l'intron 3 (séquence en rouge).

B. Le séquençage du produit de RT-PCR de 526 pb met en évidence deux populations d'ADNc chez la patiente.

C. L'interprétation de la double séquence montre une délétion de 7 pb chez la patiente. Ceci s'explique par l'utilisation d'un site accepteur d'épissage cryptique localisé 7 pb plus loin (le nucléotide souligné indique la base mutée).

1.3.4. Les mutations silencieuses ou polymorphismes

Nous avons identifié cinq mutations silencieuses (S70, S194, T299, S357 et P399) dont deux avaient déjà été rapportées dans la littérature (S194 par Amir et al. 2000 et T299 par Cheadle et al. 2000).

Les deux substitutions à l'origine des changements d'acides aminés E397K et A444T ont été classées parmi les polymorphismes car elles ont été identifiées chez des patientes présentant par ailleurs une mutation classique (T158M pour E397K et R168X pour A444T). De plus, la substitution E397K a été aussi publiée comme un polymorphisme par l'équipe de Zoghbi (Wan et al. 1999). Tous les autres polymorphismes (IVS3+22G>C, IVS3-18insG, IVS3-61G>C, 1470G>A, 1560insT) touchent des régions introniques ou non codantes du gène *MECP2* et ont été détectés pour trois cas chez des patientes porteuses par ailleurs de mutations délétères.

Aucune variation de séquence (mutation délétère ou polymorphisme) n'a été détectée chez les huit garçons étudiés par CSGE et séquençage.

2. Seconde stratégie de criblage du gène *MECP2* : en deux étapes

L'acquisition d'un système d'analyse d'hétéroduplex par DHPLC (Denaturing High Performance Liquid Chromatography) couplé à un collecteur de fragments nous a permis d'éliminer l'étape de séquençage systématique et de baser ainsi notre recherche de mutations dans le gène *MECP2* sur i) précriblage par DHPLC et ii) technique de Southern blot pour déceler de plus grands remaniements.

2.1. La technique de DHPLC

La DHPLC, technique mise au point récemment (Oefner et al. 1995, 1998 ; Liu et al. 1998 ; O'Donovan et al. 1998) repère les hétéroduplex formés spontanément au cours des cycles de la PCR. La séparation des produits de PCR se fait dans une colonne composée d'une phase stationnaire (ou matrice) et d'une phase mobile. La phase stationnaire est formée de billes non poreuses électriquement neutres et la phase mobile est un tampon contenant du TEAA (Tri-Ethyl-Ammonium-Acetate) à 0,1M. Le TEAA lie les molécules d'ADN aux billes en établissant d'une part des interactions hydrophobes avec les billes, et d'autre part des interactions ioniques avec les molécules d'ADN (hétéroduplex et homoduplex) (figure 16).

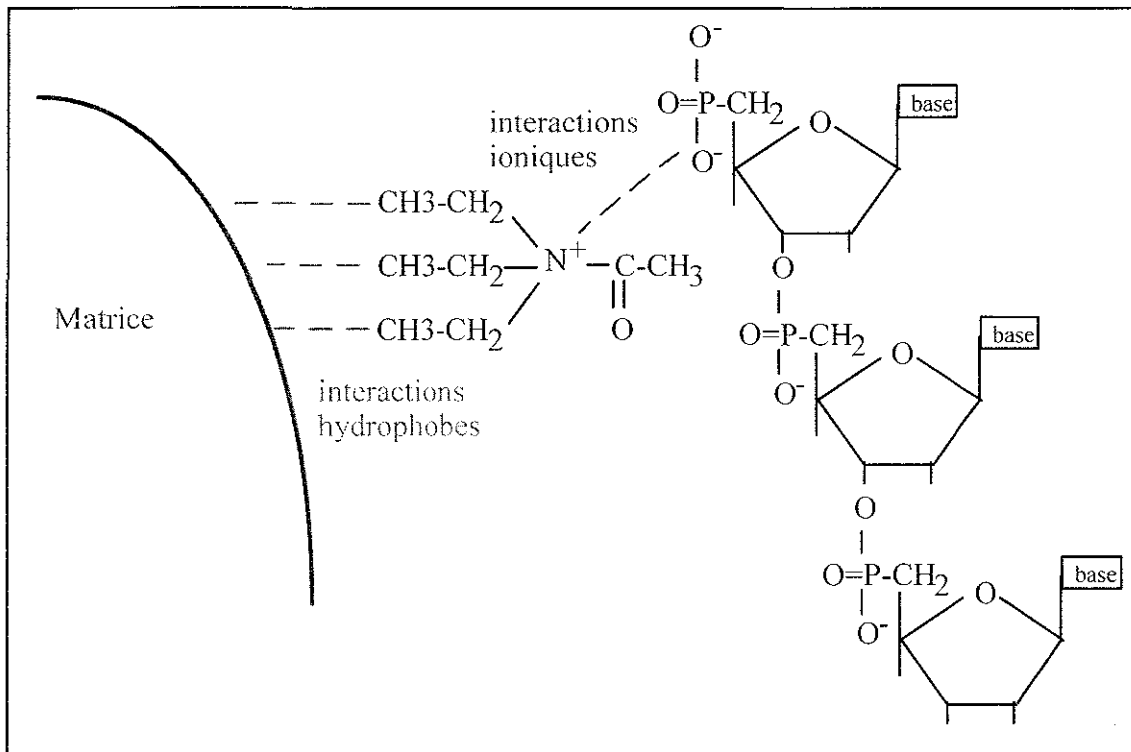


Figure 16 : Interactions entre l'ADN, le TEAA et la matrice
 La molécule de TEAA est indiquée en bleu.
 Seul un brin de la molécule d'ADN est schématisé en noir.

Les différentes molécules d'ADN sont éluées par une solution de concentration croissante en acétonitrile, en fonction du nombre d'interactions qui les lient à la matrice via le TEAA. Du fait du mésappariement, la conformation des hétéroduplex n'est plus celle de l'ADN double brin de type B, les groupements phosphates ne sont plus localisés à l'extérieur de la molécule, il y a donc moins d'interactions entre les hétéroduplex et la matrice qu'entre les homoduplex et la matrice. Les hétéroduplex sont ainsi élués avant les homoduplex. En jouant sur la température, on favorise la déstabilisation des hétéroduplex, ce qui permet de mieux les différencier des homoduplex. Les produits élués sont détectés par mesure de l'absorbance à 260 nm, les résultats sont obtenus sous forme d'un chromatogramme. En théorie, le chromatogramme doit contenir quatre pics, un pour chaque molécule (figure 17). Dans la pratique c'est plus souvent deux (un pour les hétéroduplex, un pour les homoduplex) voire trois pics (un pour chaque homoduplex et un pour les deux hétéroduplex) qui sont obtenus. Cette technique est à l'heure actuelle la plus performante, son taux de détection de variations de séquences est supérieur à 96% (Xiao et al. 2001).

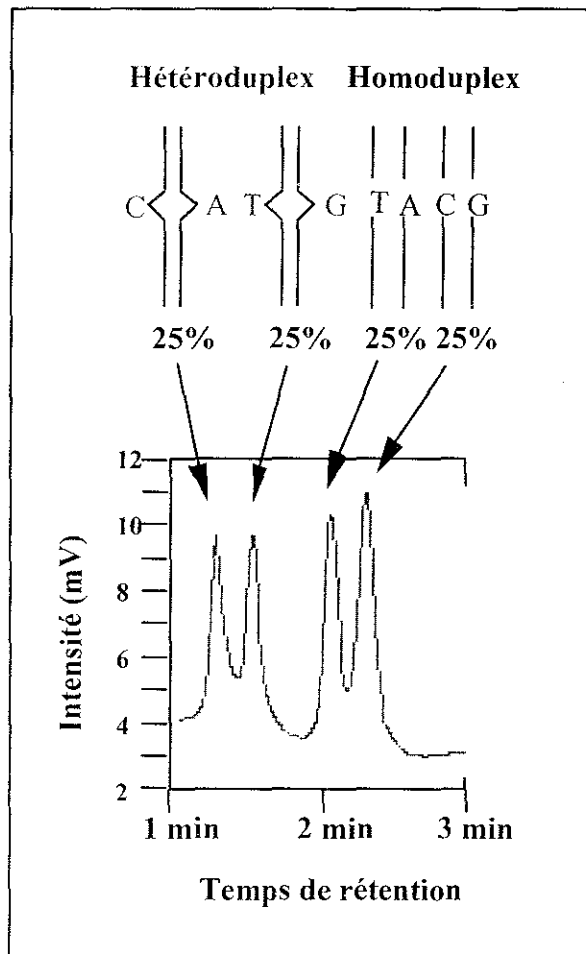


Figure 17 : Chromatogramme théorique de DHPLC (d'après Xiao et al. 2001)

2.2. Application de la DHPLC au gène *MECP2*

La technique de DHPLC nécessite une étape de mise au point permettant la détermination des températures optimales d'analyse des produits de PCR pour détecter des mutations. Nous avons réduit le nombre de couples d'amorces utilisés pour la CSGE de huit à six et nous avons modifié certaines amorces d'après la publication de Buyse et al. (Buyse et al. 2000) (figure 18).

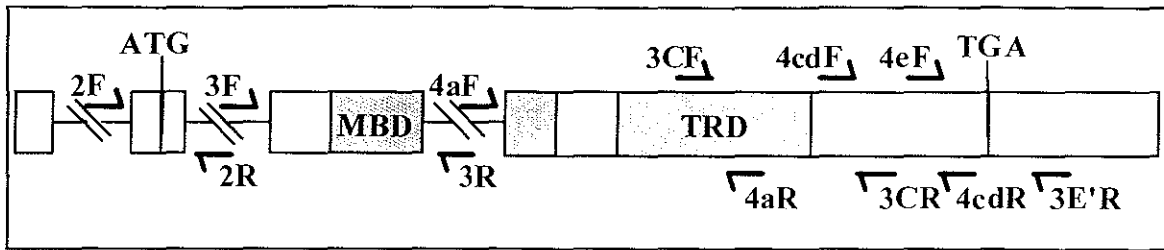


Figure 18 : Position des 6 couples d'amorces utilisés pour la DHPLC sur le gène MECP2

Afin de définir les températures d'analyse, nous avons utilisé comme témoins les 27 mutations précédemment identifiées par CSGE et séquençage systématique. Aucune mutation n'ayant été rapportée dans l'exon 2 jusqu'à présent, nous nous sommes appuyés sur les conditions définies par Buyse et al. (Buyse et al. 2000). Les produits de PCR 3 et 4a, étant les plus grands (625 pb et 583 pb), contiennent plusieurs domaines de fusion nécessitant plusieurs températures d'analyse : deux (59°C et 65°C) pour le fragment 3 et trois (61°C, 63°C et 66°C) pour le fragment 4a. La technique de DHPLC détectant les mutations à l'état hétérozygote, nous avons, comme pour la CSGE, mélangé les produits de PCR des garçons avec ceux d'un témoin masculin.

Les 27 mutations ont été identifiées dans une ou plusieurs conditions d'analyse par des profils chromatographiques caractéristiques. Des exemples de chromatogrammes sont montrés dans la figure 19.

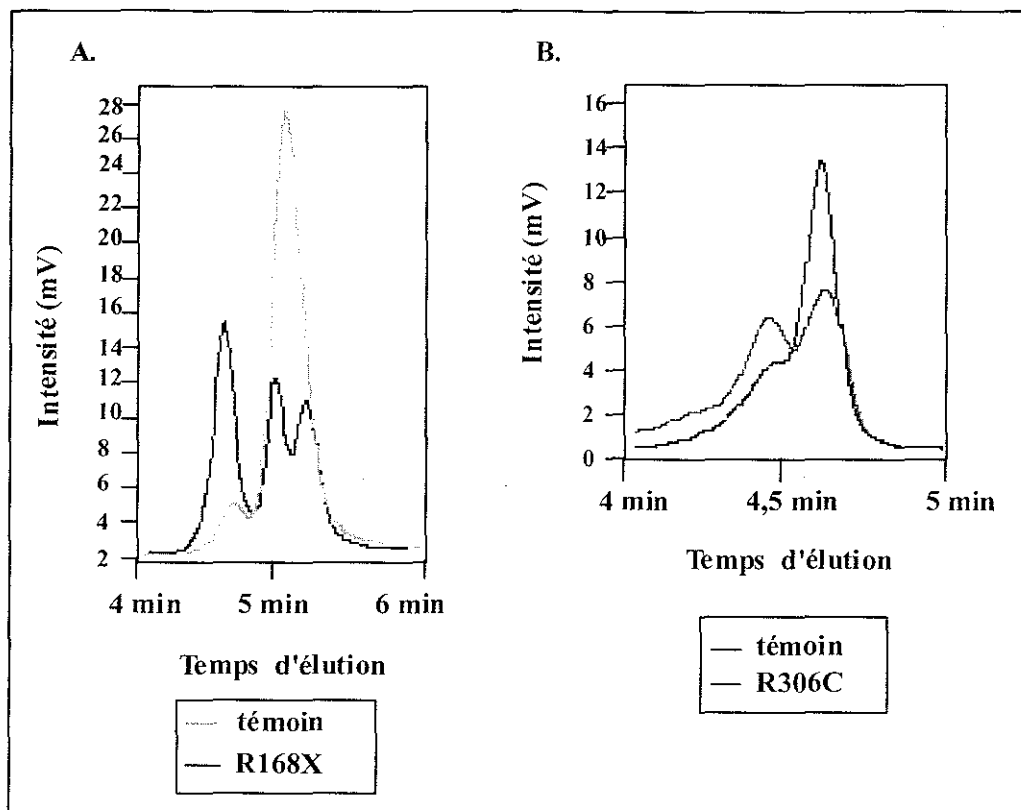


Figure 19 : Mutations détectées par DHPLC

*A. Le chromatogramme de la mutation R168X présente trois pics comparé à celui du témoin.
 B. Le chromatogramme de la mutation R306C montre deux pics comparé à celui du témoin.
 C'est un séquençage du produit de PCR correspondant au chromatogramme anormal qui a permis de déterminer la nature de la mutation.*

Une fois que les conditions de détection des 27 mutations par DHPLC ont été établies, nous les avons validées en deux étapes. Une première série de neuf patientes a été étudiée en parallèle par CSGE/séquençage systématique et DHPLC. Nous avons obtenu 100% de concordance entre les deux approches. Une seconde série de dix patientes a été testée à l'aveugle par DHPLC. Les produits de PCR des patientes ne présentant pas de chromatogrammes anormaux ont été séquençés de façon systématique, aucune mutation ou polymorphisme n'a été décelée, confirmant un taux de détection de 100% des mutations connues au laboratoire. Nous avons ainsi décidé de ne plus séquençer les produits de PCR des patients sans anomalie détectée par DHPLC, tout en gardant à l'esprit que nous avons déterminé des conditions d'analyse à partir de mutations connues, sous-entendant la possibilité de ne pas détecter d'autres mutations.

2.3. Les résultats obtenus par DHPLC

Nous avons étudié 104 patients par DHPLC (97 filles et 7 garçons). Douze nouvelles variations de séquence correspondant à dix substitutions et deux délétions, ont été décelées, montrant que les conditions d'analyse définies n'empêchent pas la détection de variations de séquence inconnues jusqu'alors au laboratoire. Le tableau XIV résume tous les mutations et polymorphismes détectés par DHPLC dans le gène *MECP2*.

Tableau XIV: *Mutations et polymorphismes identifiés dans le gène MECP2 par DHPLC.*

L'astérisque désigne les nouvelles variations de séquence détectées par DHPLC.

Mutations		Récurrence
faux-sens	455C>G, P152R*	1
	397C>T, R133C*	2
	473C>T, T158M	1
	916C>T, R306C	1
non-sens	28G>T, E10X*	1
	502C>T, R168X	3
	763C>T, R255X	2
	808C>T, R270X	4
délétions	106-112del*	1
	696delC*	1
	806delG	1
variant	982C>G, L328V*	1
polymorphismes	135C>T, P45*	1
	IVS2-85G>A*	1
	IVS3+94G>A*	1
	897C>T, T299	1
	1233C>T, S411*	1
	1240A>C, K414Q*	1
	1474C>T*	1

Parmi les douze nouvelles modifications de séquence, il y a cinq mutations qui n'étaient pas connues au laboratoire (E10X, 106-112del, P152R, R133C et 696delC) dont trois sont publiées : P152R, R133C et 696delC (Amir et al. 2000 ; Bienvenu et al. 2000 ; Buyse et al.

2000 ; Cheadle et al. 2000 ; Hupkke et al. 2000 ; Yamashita et al. 2001 ; Xiang et al. 2000 ; Giunti et al. 2001 ; Vacca et al. 2001 ; Nicolao et al. 2001 ; Hoffbuhr et al. 2001). La mutation E10X est proche de l'extrémité N-terminale de la protéine, elle se situe en amont du premier domaine fonctionnel (MBD) de la protéine. Très peu de mutations aussi précoces dans la séquence codante du gène *MECP2* ont été décrites dans la littérature (mutation Q19X (Kim et al. 2000), K22X (Giunti et al. 2001), 42ins8 (Vacca et al. 2001), 46delC (Monros et al. 2001) et 76delC (Nielsen et al. 2001-a)). Il est probable que les ARNm synthétisés à partir de ces allèles mutés soient dégradés avant toute traduction. La faible récurrence de ces mutations précoces suggère qu'une diminution de moitié de la quantité de protéine MeCP2 produite (dans le cas d'une inactivation du chromosome X aléatoire) a un effet plus délétère que les autres mutations. Le rôle potentiel de l'inactivation du chromosome X dans l'expression de ces mutations sera discuté dans le paragraphe A IV-2.1 de la discussion.

Parmi les dix nouvelles substitutions, six sont des polymorphismes (P45, IVS2-85G>A, IVS3+94G>A, S411, K414Q et C1474T). Deux sont des polymorphismes neutres : P45 et S411. La substitution K414Q a été retrouvée chez le père de la patiente. Les trois polymorphismes restant touchent des régions introniques (IVS2-85G>A et IVS3+94G>A) ou non codante du gène *MECP2* (C1474T).

Deux de ces polymorphismes (P45 et IVS3+94G>A) ont été identifiés chez des garçons, ce sont les seules anomalies de séquences détectées chez les 15 garçons étudiés au total dans cette étude par CSGE/séquençage systématique et DHPLC.

Enfin, la dernière substitution (L328V) identifiée chez une fille, n'a jamais été décrite dans la littérature. Nous attendons les prélèvements des parents afin de poursuivre l'analyse.

3. Deux cas particuliers

3.1. Mutations somatiques en mosaïque (Bourdon et al. 2001-b)

3.1.1. La mutation 1165del26 : une découverte fortuite

Une analyse par DGGE (Denaturing gradient gel electrophoresis) réalisée au laboratoire de Biochimie et Génétique Moléculaire du CHU Cochin (Paris) sur l'ADN extrait de lymphoblastes d'une patiente (patiente 1) présentant un syndrome de Rett typique, avait mis en évidence la mutation 1165del26. Afin de déterminer le caractère *de novo* ou familial de cette mutation, un prélèvement de sang frais de cette patiente ainsi que ceux de ses parents

nous ont été envoyés par la suite. Connaissant exactement la mutation, nous avons directement séquencé la région incriminée, nous n'avons pas retrouvé la mutation sur l'ADN extrait des leucocytes. Surpris par ce résultat, nous avons envoyé un aliquot de notre ADN extrait des leucocytes à Paris afin qu'ils réalisent la technique de DGGE. En parallèle, nous avons réalisé la technique de CSGE, afin de confirmer ou d'infirmer l'absence de la mutation sur l'ADN leucocytaire.

La méthode DGGE est fondée sur une électrophorèse dans un gel de polyacrylamide contenant une concentration linéairement croissante d'agents dénaturants de l'ADN (urée et formamide). Le fragment analysé migre sous l'influence du champ électrique, jusqu'au moment où il atteint la concentration critique d'agents dénaturants qui dissocie ses brins d'ADN au niveau de la région la moins stable (appelée domaine de fusion). La molécule d'ADN partiellement dénaturée présente une structure branchée qui interagit avec la réticulation du gel de polyacrylamide, ralentissant considérablement sa vitesse de migration. La séquence d'un domaine de fusion détermine le degré de stabilité local. Le changement d'une seule base altère la température de fusion d'un domaine donné, ce qui explique que deux produits de PCR différant par une seule base dans un domaine de fusion, cessent leur migration à des positions distinctes du gel. De la même façon que la technique de CSGE, lorsqu'un fragment d'ADN contient une mutation à l'état hétérozygote, quatre bandes peuvent être observées (deux correspondant aux homoduplex et deux aux hétéroduplex).

L'analyse à partir de l'ADN leucocytaire par DGGE et par CSGE, a mis en évidence trois bandes : les deux hétéroduplex et un seul homoduplex, l'homoduplex correspondant au fragment délété est absent (figure 20. A-ligne 2, B-ligne 1). Nous étions face à un cas de mutation en mosaïque somatique chez une fille, jamais décrit jusqu'alors dans la littérature pour le gène *MECP2*.

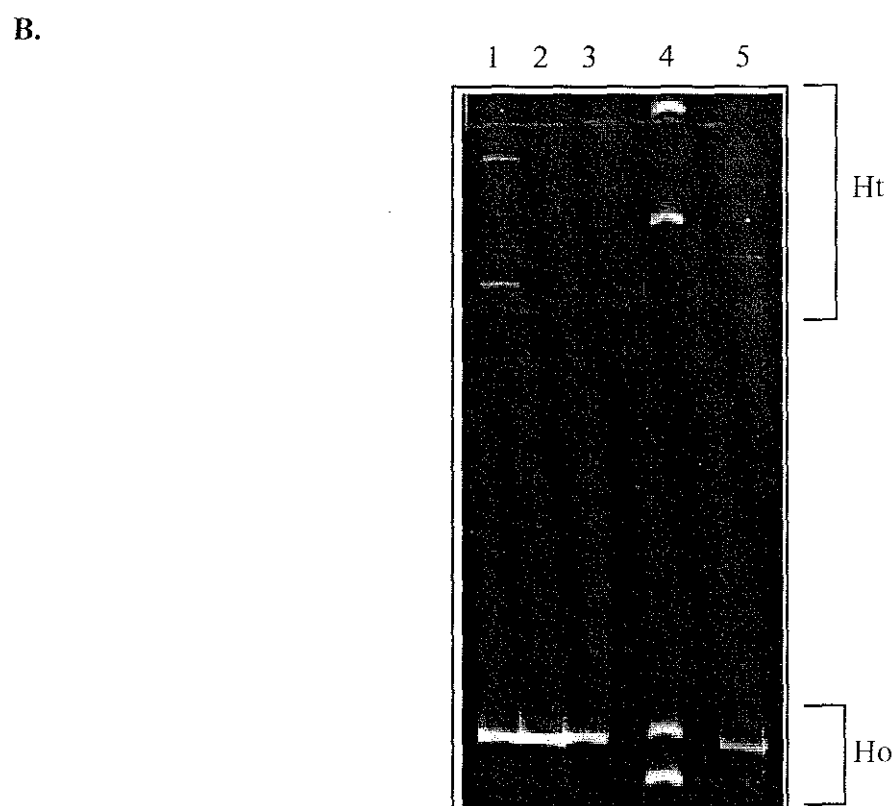
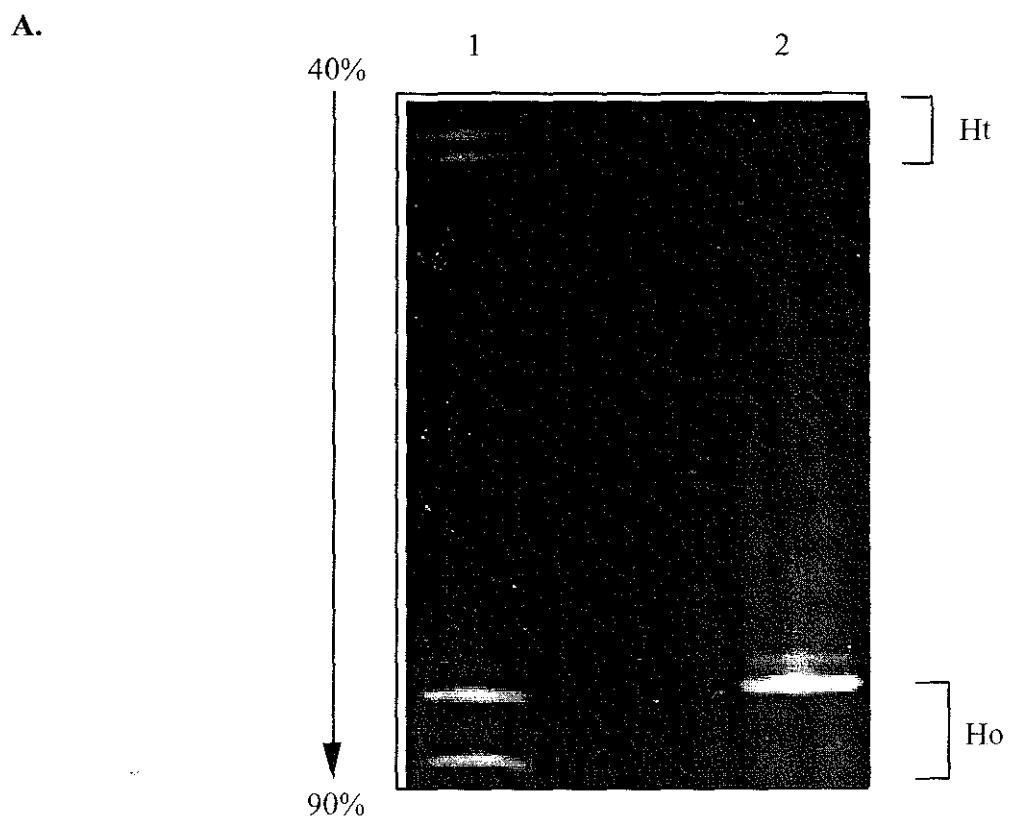


Figure 20 : Détection de mutations en mosaïque somatique par les techniques d'hétéroduplex

A. Résultats obtenus par DGGE pour la patiente 1 (photo transmise par T. Bienvenu).

Ligne 1 : ADN extrait à partir de lymphoblastes. Toutes les bandes attendues sont présentes (2 hétéroduplex, 2 homoduplex).

Ligne 2 : ADN extrait à partir de leucocytes. Il manque l'homoduplex délété.

Ht : Hétéroduplex, Ho : Homoduplex; 40%-90% : gradient de formamide utilisé, la flèche indique le sens de migration.

B. Résultats obtenus par CSGE pour la patiente 1 (lignes 1-3), et pour la patiente 2 (ligne 5).

Patiente 1 : Ligne 1 : l'absence de l'homoduplex délété signe la mosaïque somatique.

Lignes 2 et 3 : le père et la mère de la patiente 1 ne sont pas porteurs de cette mutation.

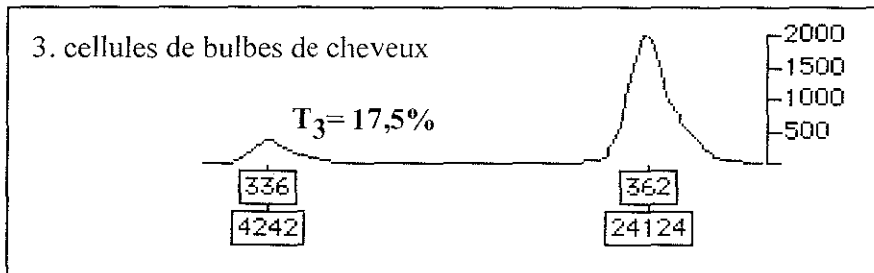
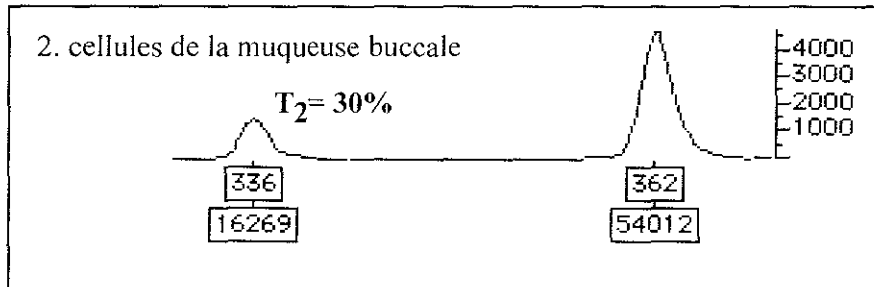
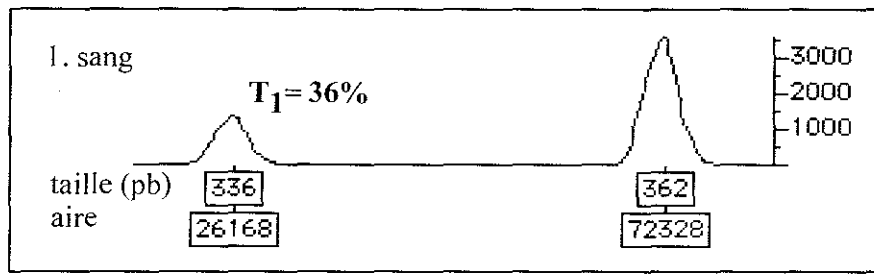
Patiente 2 : Ligne 5 : mutation en mosaïque somatique détectée chez la patiente 2 à partir de leucocytes. La variation du profil électrophorétique par rapport à celui de la patiente 1 suggère que les délétions somatiques en mosaïque sont différentes chez les patientes 1 et 2.

Ligne 4 : exemple de profil obtenu par CSGE pour une délétion constitutionnelle de 31 pb localisée dans la même région du gène MECP2. Les bandes des deux homoduplex ont la même intensité.

Afin de quantifier le degré de mosaïcisme, nous avons développé une approche semi-quantitative par PCR fluorescente. Nous avons utilisé le couple d'amorces 3D (figure 11) dont l'amorce 3DR a été marquée en 5' par le 6-FAM, amplifiant ainsi la région comprenant la délétion. En calculant le rapport : aire du pic de l'allèle délété sur la somme des aires des pics de l'allèle normal et de l'allèle muté, nous avons estimé que 36% des leucocytes de la patiente 1 portent la mutation (figure 21.A-1).

Nous avons ici une représentation du mosaïcisme au niveau du sang, mais qu'en est-il des autres tissus de la patiente et notamment du cerveau, organe principalement atteint par le syndrome de Rett ? Pour tenter de répondre à cette question, nous avons réalisé cette approche semi-quantitative sur deux autres tissus d'origine embryonnaire différente que le sang : des cellules buccales (origine mésoblastique), et des cheveux (origine ectoblastique, la plus proche du cerveau). Nous avons retrouvé un taux de mosaïcisme du même ordre dans les cellules buccales (30%) et un taux plus faible dans les bulbes de cheveux (17,5%) (figure 21. A.2-3). Ainsi la mutation en mosaïque a dû se produire très tôt au cours de l'embryogenèse car elle est retrouvée dans les différents feuilletts embryonnaires. Par ailleurs, ce cas confirme que les lignées lymphoblastoïdes ne représentent pas toujours la réalité, puisque dans la lignée issue des lymphocytes B de la patiente, toutes les cellules paraissent porter la mutation. Ceci est dû soit à une sélection positive des cellules mutées transformées par le virus EBV, soit plus vraisemblablement à un événement aléatoire.

A.



B.

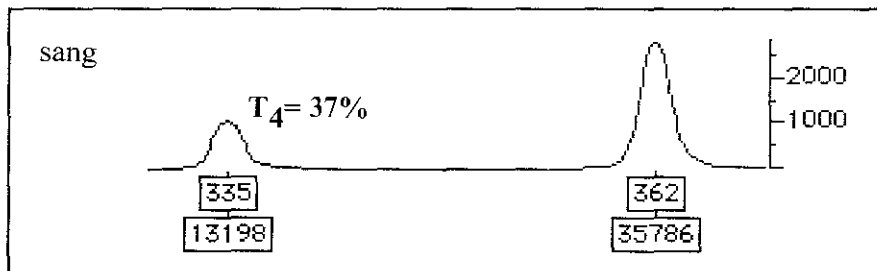


Figure 21: Evaluation semi-quantitative du taux de mosaïcisme par PCR fluorescente

A. Patiente 1 : Electrophorégrammes analysés par le logiciel Genotyper® pour trois tissus différents: sang (1), cellules de la muqueuse buccale (2) et cellules de bulbes de cheveux (3). Chaque profil est présenté avec le ratio (T_1 , T_2 et T_3) correspondant au rapport entre l'aire du pic délété et la somme des aires des 2 pics.

B. Patiente 2 : Profil obtenu par le logiciel Genotyper® à partir de sang, accompagné du ratio du pic délété (T_4).

3.1.2. Criblage systématique

Forts de cette découverte, nous nous sommes demandés si d'autres cas de mutations somatiques en mosaïque ne nous avaient pas échappé. Etant donné que la région mutée mise en cause chez la patiente 1 est une région fréquemment impliquée dans des petites délétions dans le syndrome de Rett, nous avons criblé une cohorte de 29 patientes présentant un syndrome de Rett typique et pour lesquelles nous n'avons pas mis de mutation en évidence. Sur les 29 cas étudiés, nous avons retrouvé un cas supplémentaire de mutation en mosaïque somatique (patiente 2). Cette mutation correspond à une délétion de 26 paires de bases avec un taux de mosaïcisme de 37% équivalent à celui rencontré chez la patiente 1 (figure 21.B). L'approche par CSGE a confirmé la mosaïque somatique par la présence des deux hétéroduplex et d'un seul homoduplex (l'homoduplex correspondant à l'allèle délété étant absent) (figure 20.B-ligne 5). Par contre, le séquençage ne détectant pas l'allèle délété, nous n'avons pu dans un premier temps, localiser précisément les bornes de la délétion. L'utilisation du collecteur de fragment couplé au système d'analyse d'hétéroduplex par DHPLC nous a permis de purifier l'allèle délété (sans le cloner) et de le séquencer. Il s'agit de la délétion : 1164-1189del.

Ainsi, par ces deux cas, nous montrons les limites de détection du séquençage, notamment pour le fragment du gène *MECP2* étudié, puisque nous ne détectons pas la mutation qui est présente dans environ 35% des cellules, soit 17,5% de chromosomes X. Des données récentes de la littérature (Costa et al. 2000 ; Dunger et al. 2001) confirment la faiblesse de détection du séquençage. Dunger et al. montrent même qu'il faut au moins 50% d'allèle mutant du gène *c-fos* pour le détecter, ce qui correspond aux conditions de détection des mutations à l'état hétérozygote.

3.2. La substitution 1126C>T, P376S

Parmi les modifications faux-sens identifiées, nous avons mis en évidence le changement P376S. Il présente la particularité de se situer en aval des deux domaines fonctionnels, ce qui n'est le cas d'aucune mutation faux-sens décrite chez des filles atteintes d'un syndrome de Rett jusqu'à aujourd'hui. Est-ce un polymorphisme ou une mutation délétère ?

Le séquençage de cette région chez 100 témoins féminins n'a jamais révélé cette modification. Nous l'avons également recherchée chez les parents, et l'avons retrouvée chez la mère de la patiente. Se posait alors la question des profils d'inactivation des chromosomes

X chez la mère et la fille, puisque la mère n'est pas atteinte. En effet, si le changement faux-sens P376S est responsable de la maladie de l'enfant, seules une inactivation biaisée d'un des chromosomes X chez la mère et une inactivation aléatoire chez l'enfant peuvent expliquer les phénotypes différents entre l'enfant et la mère. L'étude de l'inactivation des chromosomes X au locus HUMARA (Allen et al. 1992) a mis en évidence une inactivation biaisée (87,5-12,5%) chez la mère et une inactivation aléatoire chez la fille (60-40%) (figure 22). Ces résultats préliminaires sont en faveur du caractère délétère de la variation faux-sens P376S.

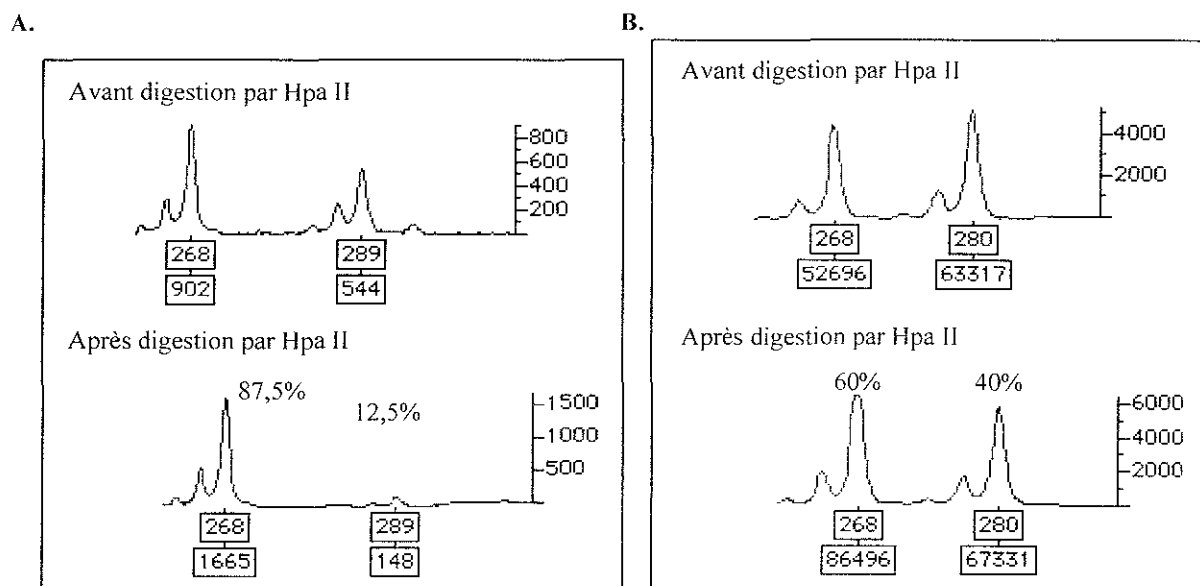


Figure 22 : Résultats de l'étude de l'inactivation du chromosome X pour la famille B.

A. La mère est informative au locus HUMARA, elle présente deux allèles: 268 pb et 289 pb. **La digestion par Hpa II met en évidence un biais d'inactivation.** L'allèle 268 est porté par le chromosome X préférentiellement inactivé dans 87,5% de ses cellules, l'allèle 289 correspond au chromosome X actif dans seulement 12,5% des cellules. Les pourcentages sont calculés à partir de l'aire des pics indiquée sous les pics.

B. La fille présente l'allèle 268 préférentiellement inactivé chez sa mère et un autre allèle (de son père) présents dans les proportions équivalentes après la digestion par Hpa II. **Elle a une inactivation aléatoire de ses chromosomes X.**

Une étude plus approfondie des membres de la famille de la patiente (figure 23) a mis en évidence que la modification P376S est présente chez la grand-mère maternelle (I.1) et fait surprenant chez l'oncle maternel (II.4). Elle n'a été retrouvée ni chez la tante maternelle (II.3), ni chez le grand-père maternel (I.2) (figure 23). La présence de la substitution chez l'oncle maternel exclut l'implication de cette substitution dans la pathologie puisque celui-ci ne présente aucun symptôme, et permet de conclure à un polymorphisme. De plus, l'analyse du

profil d'inactivation des chromosomes X montre qu'il est biaisé chez la grand-mère (I.1, 94-6%) porteuse de la modification ainsi que chez la tante (II.3, 88,5-11,5%) non porteuse de la substitution (figure 23). Il semble donc exister dans cette famille une inactivation biaisée familiale coségrégeant fortuitement avec la pathologie, puisque seule l'enfant présente une inactivation aléatoire associée à un phénotype.

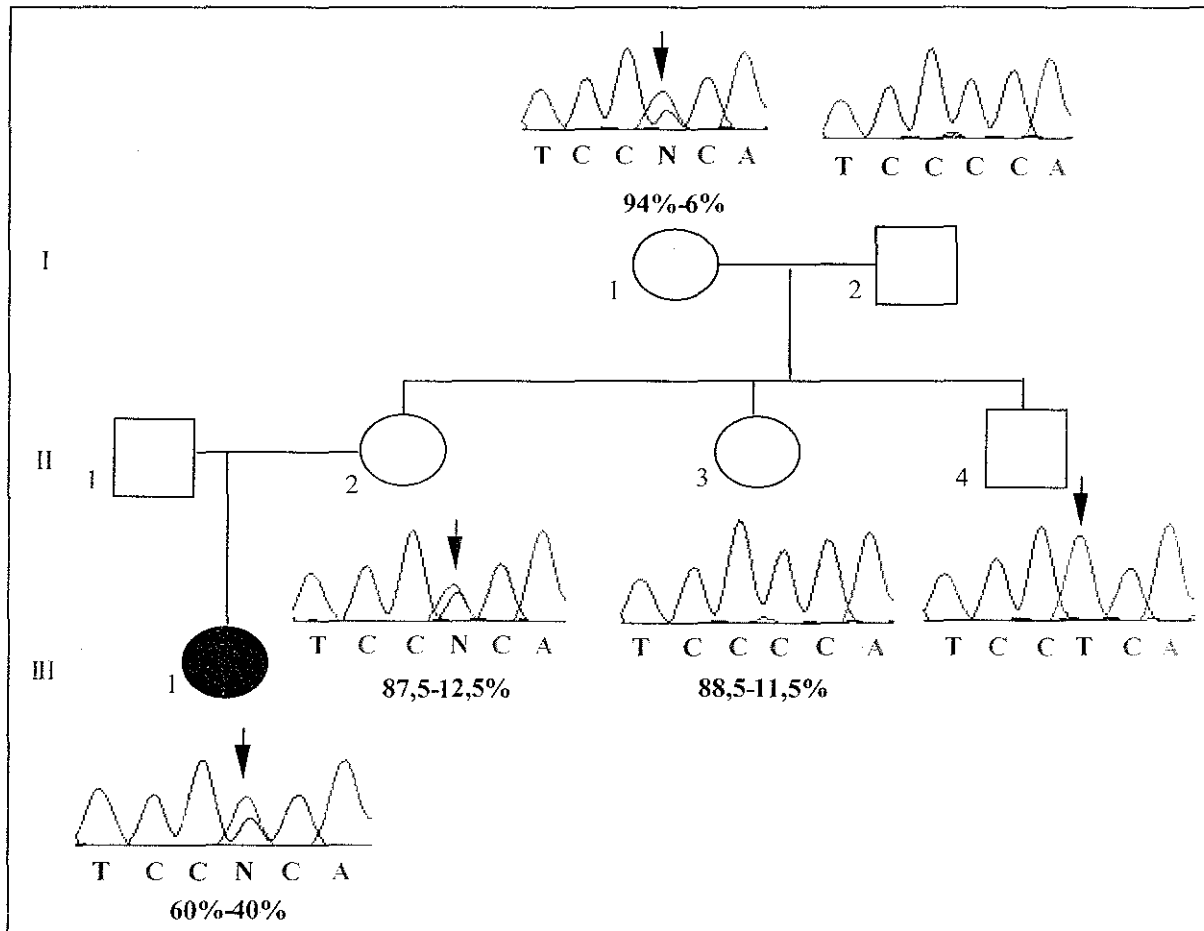


Figure 23 : Ségrégation de la substitution 1126C>T et étude du profil d'inactivation

Sont présentées les séquences de chaque individu étudié montrant la ségrégation de la substitution 1126C>T (identifiée par la flèche). Elle est présente chez I-1, II-2, II-4, III-1, et absente chez II-3.

Les pourcentages indiquent les résultats obtenus lors de l'étude de l'inactivation du chromosome X. Celle-ci est biaisée chez I-1, II-2, II-3, et est aléatoire chez III-1.

La conséquence de cette substitution sur la protéine MeCP2 peut être discutée, bien que seules des expériences au niveau protéique permettraient d'en savoir plus. En effet, les deux domaines fonctionnels de la protéine restent intacts, la protéine n'est modifiée qu'au niveau d'un seul acide aminé, par ailleurs non conservé entre le xénope, le poulet, la souris et

l'homme. Cependant, l'acide aminé modifié est une proline dont l'hétérocycle intervient généralement dans le repliement tridimensionnel des protéines. Il semble donc que la proline en position 376 ne soit pas essentielle dans la structure tridimensionnelle de la protéine MeCP2, étant donné que son changement par une sérine n'a pas de conséquence pathologique apparente.

4. Les résultats obtenus par Southern blot

Toutes les recherches de mutations dans le gène *MECP2* publiées sont basées soit sur une technique de séquençage direct (Xiang et al. 2000 ; Auranen et al. 2001 ; Cheadle et al. 2001) soit sur une double approche : précriblage par une analyse d'hétéroduplex par DGGE, ou SSCP, ou DHPLC, ou CSGE (Amir et al. 1999 ; Bienvenu et al. 2000 ; Buyse et al. 2000 ; De Bona et al. 2000 ; Obata et al. 2000) et séquençage. Certaines mutations créant ou abolissant un site de restriction, la digestion du produit de PCR peut être utilisée soit en précriblage soit pour confirmer la présence de la mutation (Wan et al. 1999 ; Hanefeld et al. 2000). Nous avons développé, en plus des techniques classiques de précriblage par CSGE puis par DHPLC, une approche de recherche de grands réarrangements du gène *MECP2* par Southern blot. La sonde utilisée correspond à l'ADNc du gène *MECP2* qui est amplifié en deux fragments chevauchants à l'aide de deux couples d'amorces (MEC1/3BR et 3CF/3E'R) (figure 24).

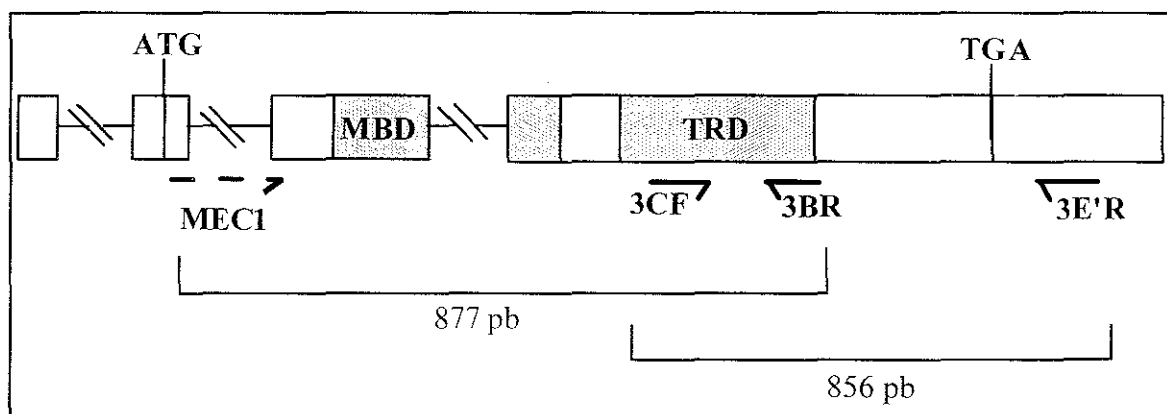


Figure 24 : Sonde utilisée pour le Southern blot

La sonde du Southern blot correspond à l'ADNc partiel du gène *MECP2*. Elle est obtenue par deux RT-PCR à partir d'ARNm, amplifiant deux fragments chevauchants (couples MEC1/3BR et 3CF/3E'R). L'amorce MEC1 est à cheval sur les exons 2 et 3.

Afin de choisir les enzymes de restriction permettant d'analyser la région génomique du gène *MECP2*, nous avons réalisé une première technique sur deux ADN témoins avec cinq enzymes de restriction différentes : BamHI, BglII, EcoRI, HindIII et TaqI, choisies en fonction de la fréquence théorique de leur site de restriction. Ces études préliminaires ont montré que les enzymes de restriction BamHI, BglII et HindIII étaient les plus adéquates pour l'analyse par Southern blot du gène *MECP2*. En effet, les fragments résultant de la digestion par l'enzyme BglII, reconnus par la sonde ADNc du gène *MECP2* ont une taille supérieure à 6,5 kb, les fragments provenant de la digestion par l'enzyme BamHI ou HindIII ont une taille plus petite, inférieure à 6 kb. Nous avons toujours réalisé cette technique avec les trois enzymes BglII, BamHI et HindIII, en parallèle afin de reconnaître les polymorphismes de sites de restriction. Le Southern blot est effectué une fois le précriblage par CSGE ou DHPLC opéré.

Sur les 270 patients étudiés, en plus des 34 mutations identifiées par CSGE/séquençage systématique et DHPLC, nous avons mis en évidence 11 profils de digestion anormaux, tous sont retrouvés avec au moins deux enzymes de restriction (figure 25).

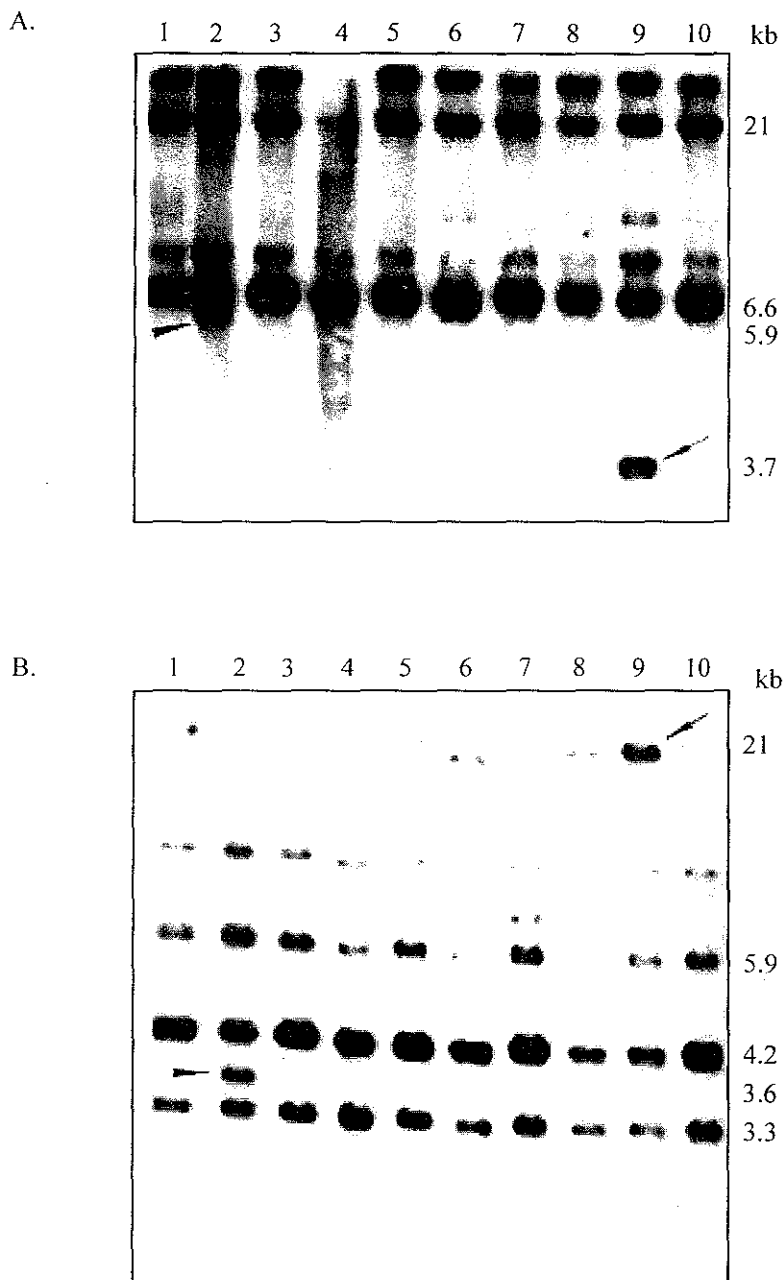


Figure 25 : Réarrangements au sein du gène *MECP2* détectés par Southern blot.

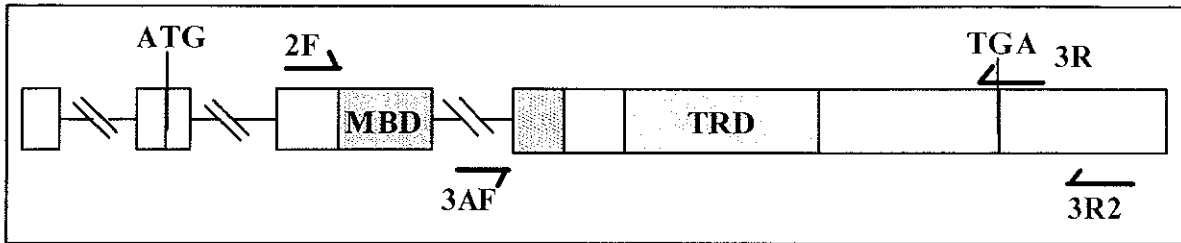
Sont montrés les résultats obtenus avec deux enzymes de restriction : *Bgl*III (A) et *Bam*HI (B) pour 10 patientes. La patiente 2 présente un fragment de jonction avec les deux enzymes (tête de flèche). La patiente 9 montre un fragment de jonction avec *Bgl*III et un fragment plus intense avec *Bam*HI (flèche).

Parmi ces 11 remaniements, 8 ont été totalement caractérisés. Etant donné que l'intron 3 du gène *MECP2* est assez petit (756 pb), nous avons étudié, dans un premier temps, ces remaniements au niveau génomique. Pour tous les remaniements détectés par Southern blot, nous avons amplifié par PCR toute la région codante (depuis l'exon 3 jusqu'au codon stop dans l'exon 4, couple d'amorces 2F/3R figure 26.A) à l'exception des 26 premières paires de bases codantes localisées dans l'exon 2 du gène *MECP2* (l'intron 2 s'étend sur 60 kb). Sur les

11 cas analysés, 5 ont montré deux bandes (figure 26.B-lignes 1 à 5), une correspondant à l'allèle normal de 2197 pb et une autre bande soit plus petite (figure 26.B-lignes 1 à 4), soit plus grande (figure 26.B-ligne 5). Ainsi dans quatre cas, nous avons une délétion et dans un cas, une insertion. Les bandes anormales étant trop grandes pour être séquencées directement (toutes sont supérieures à 1,5 kb), des PCR successives sont réalisées afin de cerner le plus finement possible les bornes des remaniements. Le produit de PCR est ensuite séquencé. Il a été alors possible de caractériser au nucléotide près les 5 remaniements : IVS3-107/937del, 382-1186del, 716-1267del, 730-1166del et 1159-1160ins300.

Lorsque la PCR avec les amorces 2F/3R ne met pas en évidence une deuxième bande, une autre PCR avec les amorces 3AF/3R2 centrée sur le dernier exon, est réalisée (figure 26.A). Trois autres remaniements ont pu être confirmés par cette PCR : 488-1190del, 1129-2029del+ins6 et 1157-2406del (figure 26.B-lignes 6 à 8).

A.



B.

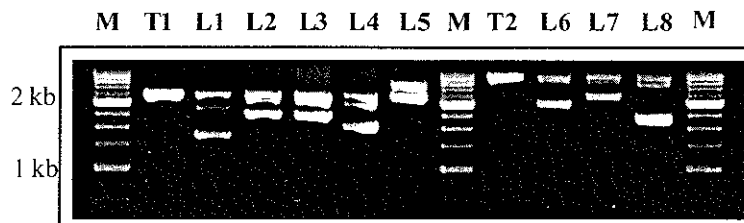


Figure 26 : Caractérisation des différents remaniements détectés par Southern blot

A. Position des deux couples d'amorces (2F/3R et 3AF/3R2).

B. Résultats des deux PCR chez huit patientes :

M : marqueur 250 pb, **T1 :** témoin de l'amplification 2F/3R (2197 pb), **T2 :** témoin de l'amplification 3AF/3R2(2911 pb).

Lignes 1-5 : PCR réalisée avec le couple 2F/3R : pour cinq patientes, présence d'une seconde bande.

Lignes 6-8 : PCR réalisée avec le couple 3AF/3R2 : pour trois patientes, présence d'une seconde bande.

Ainsi, les huit remaniements caractérisés correspondent dans sept cas à des délétions et dans un cas à une insertion au niveau de l'exon 4 du gène *MECP2*. Le tableau XV et la figure 27 résument la position et la taille de ces huit réarrangements identifiés par Southern blot.

Tableau XV : Les huit réarrangements identifiés par Southern blot.

Patient	Remaniements
1	IVS3-107del666
2	382del1804
3	488del1702
4	716del1551
5	730del1436
6	1129del900+ins6
7	1157del1249
8	1159-1160ins300

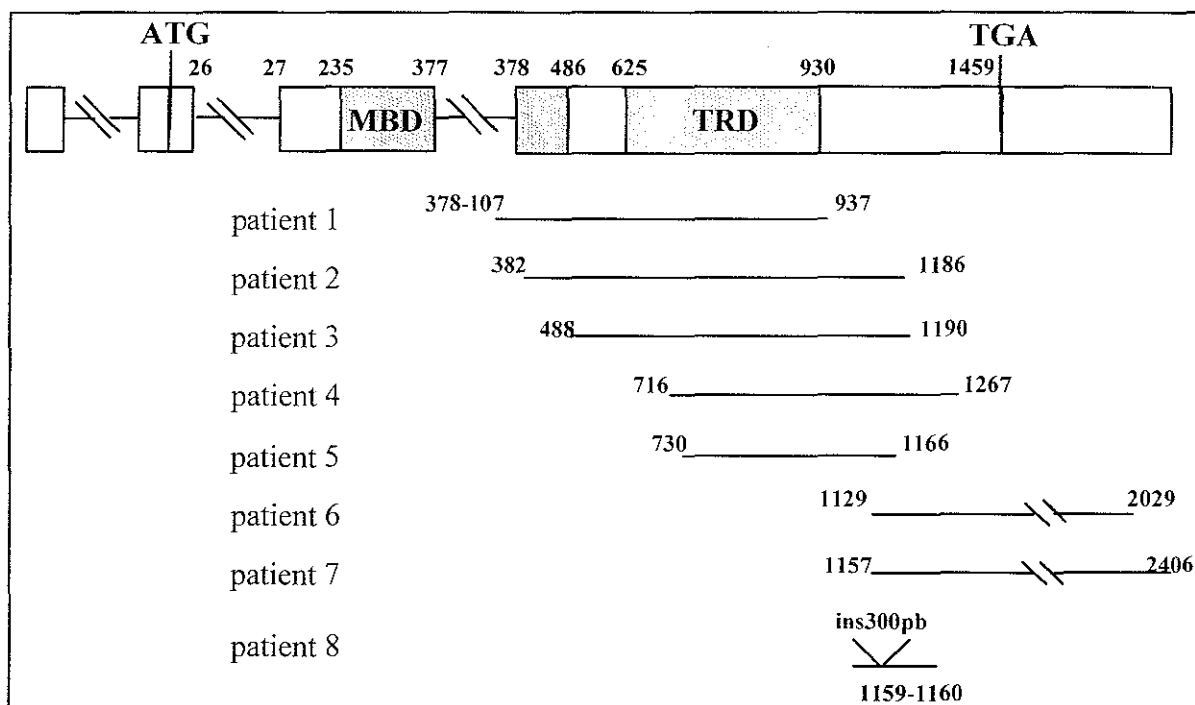


Figure 27 : Localisation génomique des différents remaniements identifiés par Southern blot

Se pose la question des trois cas non élucidés. L'hypothèse d'un polymorphisme de site de restriction peut être écartée car ces remaniements sont mis en évidence par au moins deux enzymes différentes. Dans un cas, une étude de l'ARNm (extrait de lymphoblastes) a été faite par RT-PCR successives, aucune n'a mis en évidence de bande de taille différente par rapport à l'allèle normal. Une hypothèse pourrait être l'existence d'une insertion (ou duplication) de trop grande taille qui empêcherait une amplification par PCR.

Ainsi, la technique de Southern appliquée au gène *MECP2* nous a permis de détecter 11 remaniements supplémentaires, ce qui représente 11,5% de toutes les mutations identifiées dans notre série. Certes, la technique de Southern blot est assez lourde (en temps, et nécessite la manipulation de radioactivité), et certains de ces remaniements auraient pu être détectés par la stratégie de long-PCR utilisée par Cheadle et al. (Cheadle et al. 2000) ou Bienvenu et al. (Bienvenu et al. 2001). Toutefois, trois de ces remaniements : les délétions IVS3-107del666, 1129del900+ins6 et 1157del1249, n'auraient pas été décelés car la borne 5' de la délétion IVS3-107del66, et les bornes 3' des deux autres, sont soit en aval soit en amont (respectivement) des amorces utilisées par Cheadle et al. (Cheadle et al. 2000), ou par Bienvenu et al. (Bienvenu et al. 2001).

5. Analyse de la région 3' non traduite du gène *MECP2*

Comme nous l'avons vu dans l'introduction, le gène *MECP2* est caractérisé par une région 3' non traduite étonnamment longue (8,5 kb) contenant plusieurs sites potentiels de polyadénylation (Coy et al. 1999 ; Reichwald et al. 2000). La comparaison de la séquence de cette région entre les homologues humains et murins a révélé plusieurs blocs de séquences hautement conservés, séparés par des régions plus divergentes. Coy (Coy et al. 1999) et Reichwald (Reichwald et al. 2000) ont mis en évidence que le transcrit de 10 kb était le transcrit le plus abondant dans le cerveau et que son site de polyadénylation appartenait à une des régions hautement conservées entre la souris et l'homme.

Etant donné que l'atteinte neurologique est majeure dans le syndrome de Rett, nous avons analysé chez 34 patientes et par CSGE uniquement, une région de 300 pb encadrant le site de polyadénylation correspondant au transcrit de 10 kb. Trois variants ont été obtenus et le séquençage a montré qu'ils correspondaient tous les trois à une insertion d'une cytosine dans un motif poly-cytosines (figure 28). L'analyse chez 100 femmes témoins a mis en évidence cette insertion dans sept cas, montrant qu'il s'agissait d'un polymorphisme neutre assez fréquent (3,5%) dans la population générale. De plus, deux des trois patientes sont porteuses par ailleurs, de mutations classiques. Il s'agit donc bien d'un polymorphisme.

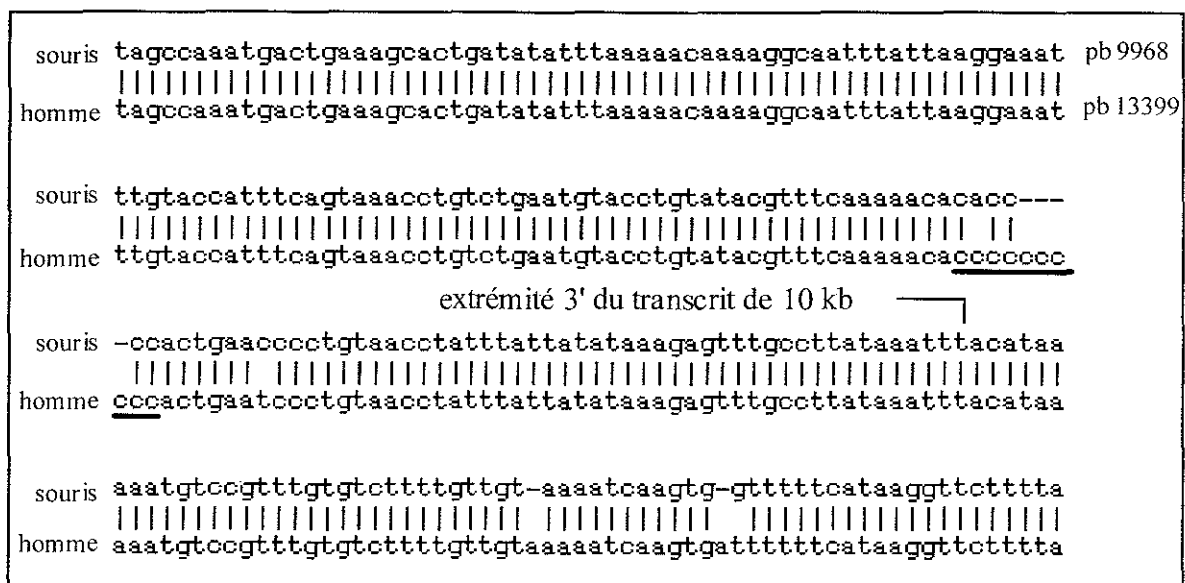


Figure 28 : Séquence analysée dans la région 3' UTR

La région soulignée correspond au motif poly-cytosines dans lequel a été trouvé l'insertion d'une cytosine chez 3 patientes.

Les numéros de paires de bases correspondent aux positions déterminées à partir des séquences AJ132922 (souris), et AF030876 (homme).

III. Recherche de réarrangements dans le gène *MECP2* à l'échelle chromosomique (Bourdon et al. 2001-c)

La recherche de mutations dans le gène *MECP2* par des techniques de biologie moléculaire (CSGE, séquençage, DHPLC, Southern blot et PCR fluorescente semi-quantitative) nous a permis de détecter 79% de mutations dans les cas typiques de syndrome de Rett, et 13% dans les cas atypiques. Par contre, les grands remaniements génomiques échappent à ces techniques de biologie moléculaire. Nous avons posé l'hypothèse que chez certaines patientes, l'ensemble de la région du locus *MECP2* pouvait être délété. Nous avons appliqué une technique d'hybridation fluorescente *in situ* (ou FISH) sur chromosomes en utilisant le PAC 671D9 isolé et séquencé par Kathrin Reichwald de l'équipe du Dr Matthias Platzer de l'Institut de Biotechnologie Moléculaire de Iena (Allemagne). Le PAC 671D9 comprend dans un insert d'environ 100 kb la totalité du locus *MECP2* (Reichwald et al. 2000). Ce PAC nous a permis d'explorer la région chromosomique du gène *MECP2* chez 26 patientes pour lesquelles nous avons une lignée lymphoblastique et qui présentent un syndrome de Rett typique, mais sans mutation dans la région codante du gène *MECP2*. Aucun remaniement n'a été mis en évidence (le caryotype des patientes est, par ailleurs, normal) (figure 29). Une approche similaire réalisée chez trois patientes n'a décelé aucun remaniement (Nielsen et al.

2001). Pour certains gènes, la présence de séquences répétées flanquantes favorisent une recombinaison illégitime entraînant soit une délétion soit une duplication du locus à l'origine de pathologies génétiques. Par exemple, le gène *STS* en Xp22 est responsable de l'ichtyose liée au chromosome X par une délétion chez 90% des hommes atteints (Hernandez-Martin et al. 1999). D'après notre étude, il semble donc que ce mécanisme mutationnel n'affecte pas le gène *MECP2*, dans le cadre du syndrome de Rett.

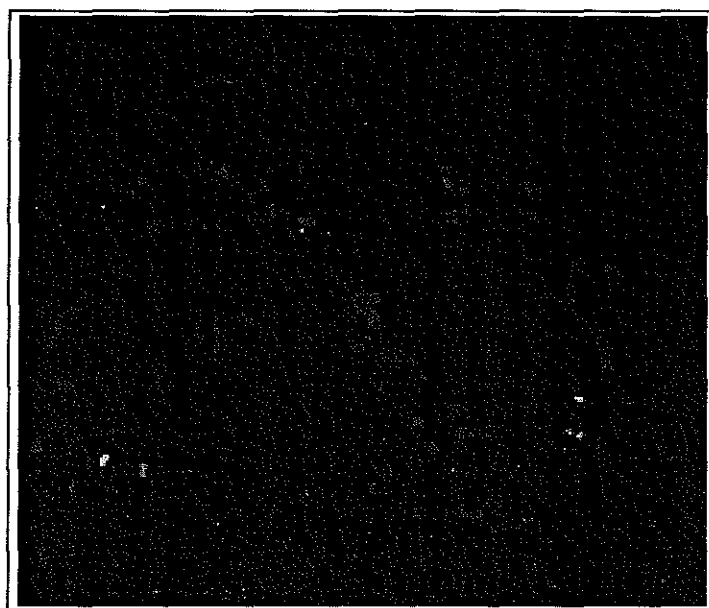


Figure 29 : Recherche de remaniements du gène *MECP2* par FISH

*La métaphase montre les deux signaux contrôles turquoise localisés aux centromères des chromosomes X ainsi que les deux signaux verts correspondant à l'hybridation de la sonde (PAC 671D9) sur les deux allèles du gène *MECP2* dans la région Xq28. Aucun remaniement n'est détecté.*

DISCUSSION

A. Syndrome de Rett et gène *MECP2*

I. Le spectre mutationnel du gène *MECP2*

1. Quelques points chauds de mutations

En 1999, Amir et al. (Amir et al. 1999) mirent en évidence 5 mutations dans le gène *MECP2* chez 29 patientes étudiées, ce qui représentait un taux de couverture du syndrome de Rett d'environ 20%. Lors de la deuxième publication (Wan et al. 1999), le gène *MECP2* fut incriminé dans 50% des cas. Aujourd'hui, les analyses le mettent en cause dans 80 à 85%, voire même 100% des cas (Auranen et al. 2001). Cette augmentation est à corrélérer fortement avec une sélection de plus en plus stricte des patientes présentant un syndrome de Rett typique.

Nous avons étudié 270 patients parmi lesquels 92 filles présentent un syndrome de Rett classique. Au total, 95 mutations ont été identifiées, soit environ 35% de mutations dans le gène *MECP2* pour l'ensemble des patients testés (filles et garçons confondus). Plus précisément, une mutation dans le gène *MECP2* est impliquée chez 79% des 92 patientes atteintes d'un syndrome de Rett typique, et chez 13% des patientes porteuses un syndrome atypique. Ces 95 mutations correspondent en réalité à 45 mutations différentes réparties tout le long du gène avec plusieurs positions récurrentes (figure 30).

Les données de la littérature rapportent l'étude d'environ 1000 patientes ; le tableau XVI illustre la quasi-totalité des différentes mutations publiées ainsi que leur fréquence.

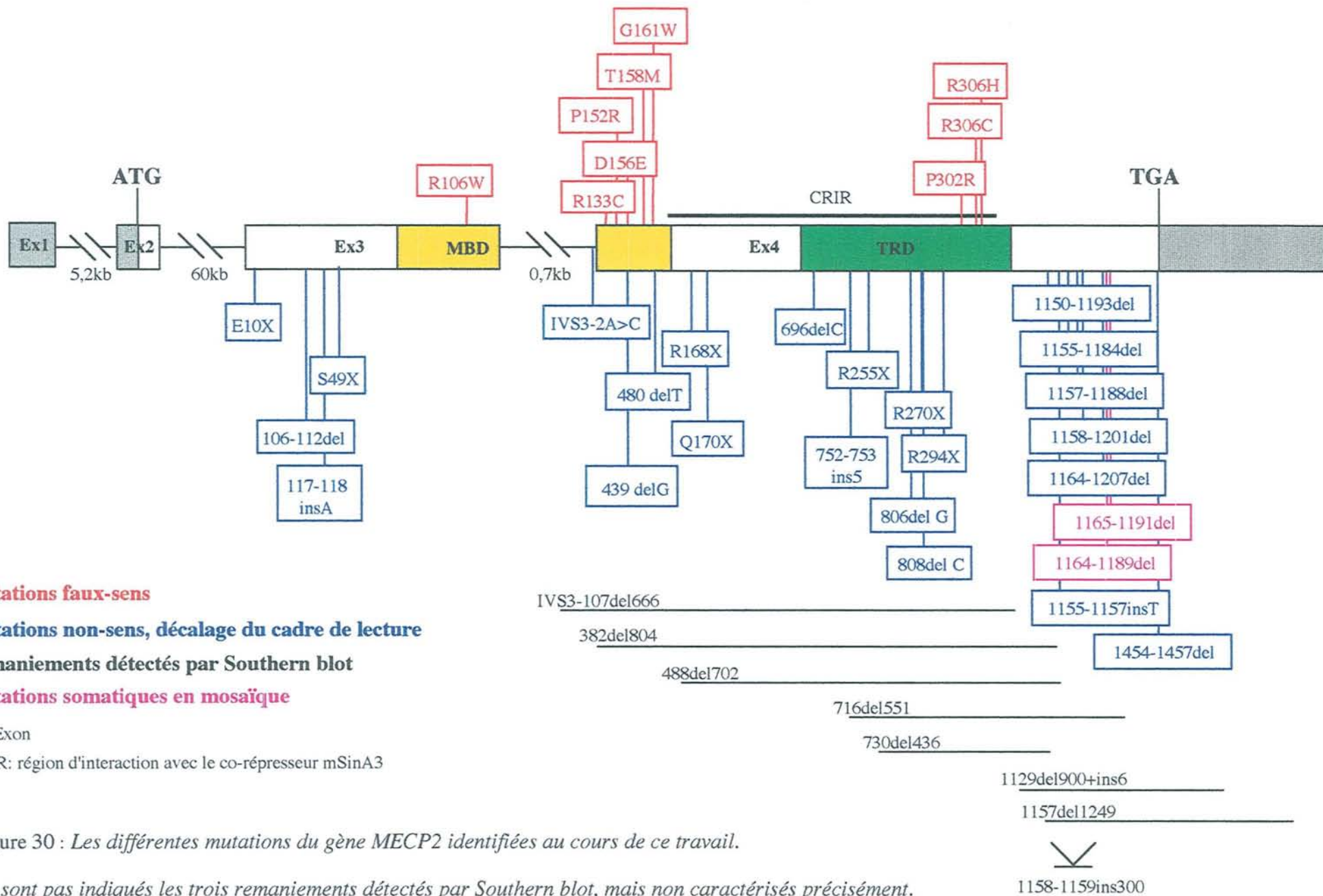


Figure 30 : Les différentes mutations du gène MECP2 identifiées au cours de ce travail.

Ne sont pas indiqués les trois remaniements détectés par Southern blot, mais non caractérisés précisément.

Mutations	Localisation	Domaine (a)	Nucleotide substitué (b)	Acide aminé modifié	Récurrence (c)	Références (d)
faux-sens	exon3	MBD	289C>T	D97Y	1	23)
	exon3	MBD	291C>A	D97E	1	7)
	exon3	MBD	301C>A	P101T	1	4)
	exon3	MBD	302C>A	P101H	1	4)
	exon3	MBD	302 C>T	P101L	1	4)
	exon3	MBD	301C>T	P101S	1	17)
	exon3	MBD	317G>A	R106Q	1	3)
	<u>exon3</u>	<u>MBD</u>	<u>316C>T</u>	<u>R106W</u>	<u>25+4</u>	<u>1),2),4),5),6),7),14),15),16),17),20),23),</u> <u>notre étude</u>
	exon3	MBD	ND	Y120D	1	12)
	exon3	MBD	372G>C	L124F	1	14)
	<u>exon4</u>	<u>MBD</u>	<u>397C>T</u>	<u>R133C</u>	<u>27+2</u>	<u>1),2),4),5),7),11),12),14),15),17),19),20),21),23)</u> <u>notre étude</u>
	exon4	MBD	398G>T	R133L	1	9)
	exon4	MBD	398G>A	R133H	1	23)
	exon4	MBD	401C>G	S134C	6	4),5),14),16),17)
	<u>exon4</u>	<u>MBD</u>	<u>455C>G</u>	<u>P152R</u>	<u>11+1</u>	<u>2),4),5),14),16),17),21),23),</u> <u>notre étude</u>
	exon4	MBD	463T>A	F155I	1	7)
	exon4	MBD	464T>C	F155S	2	1),2)
	<u>exon4</u>	<u>MBD</u>	<u>468C>G</u>	<u>D156E</u>	<u>2+1</u>	<u>2),11),notre étude</u>
	exon4	MBD	472A>G	T158A	1	17)
	<u>exon4</u>	<u>MBD</u>	<u>473>T</u>	<u>T158M</u>	<u>75+6</u>	<u>1),2),3),4),5),7),11),12),14),15),16),17),19),20),21),23),</u> <u>notre étude</u>
	<u>exon4</u>	<u>MBD</u>	<u>481G>T</u>	<u>G161W</u>	<u>1</u>	<u>notre étude</u>
	exon4	TRD	674C>G	P225R	6	2),4),9), 11),20)
	exon4	TRD	898G>A	V300I	1	21)
exon4	TRD	904C>G	P302A	2	5),16)	
<u>exon4</u>	<u>TRD</u>	<u>905C>G</u>	<u>P302R</u>	<u>1+1</u>	<u>3),notre étude</u>	

Mutations	Localisation	Domaine (a)	Nucleotide substitué (b)	Acide aminé modifié	Réurrence (c)	Références (d)
faux-sens	exon4	TRD	905C>T	P302L	3	4),15),20)
	exon4	TRD	914A>G	K305R	2	20),23)
	<u>exon4</u>	<u>TRD</u>	<u>916C>T</u>	<u>R306C</u>	<u>39+6</u>	<u>2),3),4),5),6),7),11),12),14),15),16),17),19),20),21),23),</u> <u>notre étude</u>
	<u>exon4</u>	<u>TRD</u>	<u>917G>A</u>	<u>R306H</u>	<u>1+1</u>	4), notre étude
	exon4	TRD	964C>G	P322A	3	3),16),21)
	exon4	TRD	965C>T	P322L	2	5),23)
	exon4	CTS	1461A>C	X487C	1	3)
non-sens	<u>exon3</u>		<u>28G>T</u>	<u>E10X</u>	<u>1</u>	<u>notre étude</u>
	exon3		55C>T	Q19X	1	8)
	exon3		64A>T	K22X	3	16),17),21)
	<u>exon3</u>		<u>146C>A</u>	<u>S49X</u>	<u>2</u>	<u>notre étude</u>
	exon4	MDB	423C>G	Y141X	7	2),9),11),15),16),17)
	<u>exon4</u>	<u>CRIR</u>	<u>502C>T</u>	<u>R168X</u>	<u>73+10</u>	<u>2),3),4),5),6),7),9),12),14),15),16),17),20),21),23),</u> <u>notre étude</u>
	<u>exon4</u>	<u>CRIR</u>	<u>508C>T</u>	<u>Q170X</u>	<u>2+1</u>	<u>2),14),</u> notre étude
	exon4	CRIR	592A>T	R198X	1	3)
	exon4	CRIR	611C>T	S204X	1	22)
	exon4	CRIR	622C>T	Q208X	1	17)
	exon4	TRD	730C>T	Q244X	1	23)
	<u>exon4</u>	<u>TRD</u>	<u>763C>T</u>	<u>R255X</u>	<u>63+9</u>	<u>1),2),3),4),5),6),7),9),11),15),16),17),20),21),23),</u> <u>notre étude</u>
	exon4	TRD	766A>T	K256X	1	2)
	<u>exon4</u>	<u>TRD</u>	<u>808C>T</u>	<u>R270X</u>	<u>60+8</u>	<u>2),3),4),5),9),11),12),14),15),16),17),20),21),23),</u> <u>notre étude</u>
	<u>exon4</u>	<u>TRD</u>	<u>880C>T</u>	<u>R294X</u>	<u>43+8</u>	<u>2),3),4),5),7),9),11),14),15),16),17),20),21),23),</u> <u>notre étude</u>
site d'épissage	<u>intron3</u>	<u>MBD</u>	<u>IVS3-2A>C</u>		<u>1</u>	<u>notre étude</u>
	intron3	MBD	IVS3-2A>G		3	2),5),23)

Mutations	Localisation	Domaine (a)	Nucléotide modifié (b)	Réurrence (c)	Références (d)
délétions/ insertions	exon3		42ins8	1	17)
	exon3		46delC	1	20)
	exon3		76delC	1	15)
	exon3		100del4	1	11)
	<u>exon3</u>		<u>106del7</u>	<u>1</u>	<u>notre étude</u>
	<u>exon3</u>		<u>117-118insA</u>	<u>1</u>	<u>notre étude</u>
	exon3	MBD	258delCA	1	5)
	exon3	MBD	345delC	1	10)
	exon4	MBD	407del507+ ins8	1	4)
	exon4	MBD	411delG	1	6)
	exon4	MBD	431delA	1	9)
	<u>exon4</u>	<u>MBD</u>	<u>439delG</u>	<u>1</u>	<u>notre étude</u>
	exon4	MDB	451delG	1	11)
	<u>exon4</u>	<u>MDB</u>	<u>480delT</u>	<u>1</u>	<u>notre étude</u>
	exon4	CRIR	543del2	1	13)
	exon4	CRIR	563delG	1	17)
	exon4	CRIR	620insT	2	1),2)
	exon4	TRD	654del4bp	1	4)
	exon4	TRD	677insA	1	3)
	<u>exon4</u>	<u>TRD</u>	<u>696delC</u>	<u>1+1</u>	<u>21), notre étude</u>
	exon4	TRD	704insG	2	16),21),23)
	exon4	TRD	706delG	1	2)
	exon4	TRD	710insT	1	22)
	exon4	TRD	726del4	1	12)
	exon4	TRD	736insAT	2	16),21)
	exon4	TRD	747insC	1	23)
	<u>exon4</u>	<u>TRD</u>	<u>752-753ins5</u>	<u>1</u>	<u>notre étude</u>
	exon4	TRD	766dup14	1	15)
	exon4	TRD	770delG	1	14)
	<u>exon4</u>	<u>TRD</u>	<u>806delG</u>	<u>8+3</u>	<u>2),4),6),16), 17),21), notre étude</u>
<u>exon4</u>	<u>TRD</u>	<u>808delC</u>	<u>1+1</u>	<u>23), notre étude</u>	
exon4	TRD	852del4	1	13)	

Mutations	Localisation	Domaine (a)	Nucléotide modifié (b)	Réurrence (c)	Références (d)
délétions/ insertions	exon4	TRD	855del4	1	23)
	exon4	TRD	863del18	1	15)
	exon4	TRD	880delC	1	14)
	exon4	TRD	881del289+ 1187 del8	1	11)
	exon4	TRD	895del202	1	10)
	exon4	CTS	989ins18del8	1	10)
	exon4	CTS	996insAG	1	10)
	exon4	CTS	1193insC	1	16)
	exon4	CTS	1124del53	2	10),11)
	exon4	CTS	1129del4ins4+ 1149del46	1	13)
	exon4	CTS	1038del82+ 1169del171+ 1169dup137	1	23)
	exon4	CTS	1127ins10bp	1	7)
	exon4	CTS	1042del14+ 1060del14+ 1100del3+ 1137del26+ 1177del8	1	18)
	exon4	CTS	1050del15+ 1087del28+ 1137del70	1	18)
	exon4	CTS	1189insA	1	11)
	exon4	CTS	1194insT	1	3)
	exon4	CTS	1195insT	1	2)
	exon4	CTS	1198insC	1	21)
	exon4	CTS	1269insC	1	14)
	exon4	CTS	1364insC	1	5)
<u>exon4</u>	<u>CTS</u>	<u>1454del4</u>	<u>1+1</u>	<u>23), notre étude</u>	
<u>exon4</u>	<u>CTS</u>	<u>1011-1169 del 7à 191pb</u>	<u>61+8</u>	<u>2),3),4),5),7), 9),13),15),16),17), 18),19),20),21),22), notre étude</u>	

Tableau XVI : Mutations identifiées dans le gène MECP2 chez des patientes atteintes d'un syndrome de Rett, et publiées dans la littérature.

Les mutations soulignées correspondent aux mutations que nous avons identifiées au cours de ce travail. La couleur bleue indique les mutations du gène *MECP2* les plus fréquentes.

(a) *MDB* : domaine de liaison aux cytosines méthylées, *TRD* : domaine de répression transcriptionnelle, *CRIR* : région d'interaction avec le co-répresseur *Sin3A* (Nan et al. 1998), *CTS* : segment C-terminal, *ND* : non déterminé

(b) Les nucléotides sont numérotés à partir de la première base du codon initiateur selon Wan et al. (Wan et al 1999).

(c) Par exemple pour la mutation 316C>T, 25+4 : signifie que la mutation a été rapportée 25 fois dans la littérature et que nous l'avons identifiée quatre fois dans notre série.

(d) 1) Amir et al. 1999, 2) Amir et al. 2000, 3) Bienvenu et al. 2000, 4) Cheadle et al. 2000, 5) Huppke et al. 2000, 6) Wan et al. 1999, 7) Xiang et al. 2000, 8) Kim et al. 2000, 9) De Bona et al. 2000, 10) Bienvenu et al. 2001, 11) Yamada et al. 2001, 12) Inui et al. 2001, 13) Sung Jae Lee et al. 2001, 14) Yamashita et al. 2001, 15) Nielsen et al. 2001-b, 16) Giunti et al. 2001, 17) Vacca et al. 2001, 18) Lebo et al. 2001, 19) Zappella et al. 2001, 20) Monros et al. 2001, 21) Nicolao et al. 2001, 22) Dragich et al. 2001, 23) Hoffbuhr et al. 2001.

La récurrence importante de certaines mutations (R106W, R133C, T158M, R168X, R255X, R270X, R294X, R306C, et des petites délétions dans la région 3' codante du gène) met en évidence des points chauds de mutations au sein du gène *MECP2*.

Dans notre étude, les mutations (R106W, R133C, T158M, R168X, R255X, R270X, R294X, R306C) représentent à elles seules environ 45% de la totalité des mutations identifiées. Ces mutations sont donc à rechercher en priorité dans le cadre du diagnostic moléculaire du syndrome de Rett.

Wan et al. (Wan et al. 1999) ont attiré l'attention dès le début sur le fait que la majorité des substitutions correspondaient à des transitions C>T localisées au niveau de sites 5'CpG3'. Ceci est le cas des huit mutations ponctuelles les plus fréquentes, précédemment citées. Les dinucléotides 5'CpG3' sont sous-représentés dans le génome humain du fait de leur hypermutabilité. En effet, 60 à 90% de ces dinucléotides ont une cytosine méthylée chez l'Homme et cette cytosine méthylée peut être désaminée spontanément pour donner une thymine. Il a été montré que le système de réparation de l'ADN ne reconnaissait pas efficacement le mésappariement T-G, contribuant ainsi au maintien de la substitution par transition C>T au cours du cycle suivant de réplication de l'ADN (Bird 1986). Ce mécanisme de mutagenèse du gène *MECP2* ne constitue pas une particularité, il est retrouvé dans de nombreux autres gènes (Krawczak et al. 1996).

D'autre part, la région codante 3' de l'exon 4 du gène *MECP2* (entre les nucléotides 1011 et 1169) est encline à de petites délétions (de 7 à 191 pb). L'analyse de la séquence de cette région et notamment à proximité des bornes des délétions, met en évidence des répétitions de pentanucléotides : CCACC qui seraient responsables de mésappariements à l'origine des délétions (Bienvenu et al. 2000).

2. Mutations en mosaïque somatique

La découverte, un peu par hasard il est vrai, d'une délétion en mosaïque somatique nous a alertés quant à ce mode possible de mutation du gène *MECP2*. L'existence de mutations ou de remaniements en mosaïque a été rapportée dans plusieurs cas d'anomalies chromosomiques, de mutations mitochondriales, de répétitions de triplets (Bernards et al. 1994), ainsi que dans un nombre croissant de maladies dominantes ou récessives liées au chromosome X [myopathie de Duchenne (Bunyan et al. 1994), hémophilies A et B (Leuer et al. 2001, Costa et al. 2000)]. Le dépistage de mutation en mosaïque dépend principalement de la méthode utilisée pour identifier les mutations constitutionnelles. Bien que de nos jours, la méthode de prédilection soit le séquençage direct, nous avons montré la faiblesse de cette approche pour la détection de mutations en mosaïque. L'utilisation d'une PCR fluorescente semi-quantitative nous a permis de confirmer l'existence de telles mutations dans le gène *MECP2* et de quantifier de manière relative, le nombre de cellules mutées. L'application systématique de cette technique chez des patientes atteintes d'un syndrome de Rett typique ne présentant pas jusqu'alors de mutations reconnues, a mis en évidence un cas supplémentaire de délétion en mosaïque somatique, soulignant la récurrence de ce type de mutation dans le gène *MECP2*. Toutefois, notre approche par PCR fluorescente semi-quantitative se limite à la détection des délétions (ou insertion) en mosaïque localisées dans la partie 3' codante de l'exon 4 (ces petites délétions représentent, dans notre série, 8,5% des mutations identifiées). Récemment, une approche par DHPLC a identifié une mutation ponctuelle (E365X) en mosaïque (Bowles et al. 2001). Les auteurs n'ont pu déterminer par séquençage la nature de la mutation initialement détectée par DHPLC, confirmant le peu de fiabilité du séquençage pour la détection de mutations en mosaïque. Par ce travail, les auteurs renforcent l'hypothèse selon laquelle des mutations ponctuelles ou des délétions/insertions en mosaïque du gène *MECP2* dans le syndrome de Rett constituent un événement non négligeable. De plus, la DHPLC semble être une alternative intéressante à l'utilisation de PCR allèle-spécifiques (PCR-ASO) notamment pour la détection de mutations ponctuelles en mosaïque; en effet, l'approche par

PCR-ASO nécessite la connaissance préalable de la mutation et donc une amorce particulière pour chaque mutation (soit une quarantaine d'amorces dans notre série). Toutefois, le seuil de détection de telles mutations est probablement plus bas par PCR-ASO que par DHPLC.

3. Analyse des régions régulatrices et non codante du gène *MECP2*

Au cours de l'année 2000, Kathrin Reichwald (Reichwald et al. 2000) a mis en évidence une nouvelle région transcrite en 5' du gène *MECP2*. En effet, en comparant la séquence homologue de la souris à celle de l'homme, elle a détecté le premier exon du gène *MECP2*, il se situe 5,2 kb en amont de l'exon 2 (initialement considéré comme le premier) et comporte 69 pb. N'ayant pas déterminé les conditions idéales d'analyse de cette portion du gène *MECP2* (la région est très riche en CG), nous n'avons pu réaliser le criblage. Une fois au point, l'analyse de cet exon révélera, peut-être, de nouvelles mutations chez les patientes présentant un syndrome de Rett typique sans mutation détectée dans la séquence codante du gène *MECP2* ; un tel cas a été identifié par T. Bienvenu (communication personnelle). De plus, une analyse plus approfondie de la région 3' non traduite hautement conservée du gène identifiera probablement de nouvelles mutations.

II. Origine parentale des mutations dans le gène *MECP2*

Afin d'expliquer l'atteinte féminine exclusive du syndrome de Rett (sans une augmentation du nombre de fausses-couches), Thomas a proposé en 1996 (Thomas 1996) que les mutations *de novo* se produisaient principalement dans la lignée germinale masculine, touchant ainsi seulement la descendance féminine, puisque les garçons reçoivent de leur père le chromosome Y. Thomas (Thomas 1996) s'appuie sur l'exemple de plusieurs maladies (l'hyperammoniémie de type II, les hémophilies A et B, le syndrome de Lesch-Nyhan) pour lesquelles des études de biologie moléculaire ont montré l'existence d'un biais dans l'origine parentale des mutations.

Dans le cas du syndrome de Rett, cette hypothèse est étayée par le fait que la majorité des mutations dans le gène *MECP2* a pour origine une transition C>T à un site 5'CpG3', cette transition étant due à une probable désamination spontanée de la cytosine méthylée, accroissant la mutabilité de ces sites. D'une façon générale, ces dinucléotides 5'CpG3' sont

beaucoup plus méthylés dans la lignée germinale masculine que dans la lignée germinale féminine, et de plus, le chromosome X est totalement méthylé, autrement dit inactif, chez l'homme pendant la spermatogenèse (Driscoll et al. 1990). Il est donc probable que le gène *MECP2*, localisé sur le chromosome X, soit également plus méthylé et donc plus sensible, dans la lignée germinale masculine, à une désamination.

Quatre études sur l'origine des mutations dans le gène *MECP2* ont été publiées. Dans la première analyse, à l'aide d'hybrides somatiques, Amir (Amir et al. 2000) a étudié quatre patientes présentant chacune une mutation correspondant à une transition C>T. Il a mis en évidence une origine paternelle dans deux cas sur quatre. Par ailleurs, il a montré qu'une même mutation (R168X) peut avoir aussi bien une origine paternelle que maternelle. Bien que cette étude soit trop petite pour tirer des conclusions générales sur l'origine des mutations dans le syndrome de Rett, il est intéressant de noter que des transitions C>T à des sites 5'CpG3' peuvent avoir une origine maternelle.

A l'aide de polymorphismes associés à la mutation, Girard et al. (Girard et al. 2001) ont rapporté dans 5 cas sur 7, une origine paternelle. Dans cette étude, les deux mutations d'origine maternelle ne correspondent pas à une transition C>T.

Dans les deux autres publications (Kondo et al. 2000 et Trappe et al. 2001), les auteurs ont obtenu des résultats similaires : une origine paternelle de la mutation dans la majorité des cas et ceci indépendamment du type et de la localisation de la mutation. Trappe et al (Trappe et al. 2001) ont décelé des délétions ayant une origine paternelle. Dans les syndromes d'Apert et de Crouzon, l'exclusivité de l'origine paternelle des mutations est corrélée à une augmentation significative de l'âge paternel (Moloney et al. 1996 et Glaser et al. 2000). Ce mécanisme ne semble pas être responsable de l'origine majoritairement paternelle des mutations *de novo* du syndrome de Rett, étant donné qu'aucune augmentation significative de l'âge des pères des patientes, au moment de la conception, n'a été mise en évidence dans la série de Trappe et al. (Trappe et al. 2001).

Ainsi, bien que les séries soient réduites, ces études convergent pour donner une explication moléculaire de l'atteinte sporadique exclusive des filles dans le syndrome de Rett en accord avec le modèle de Thomas (Thomas 1996). L'origine paternelle des mutations avait, par ailleurs été pressentie lors des études d'inactivation du chromosome X chez des patientes atteintes d'un syndrome de Rett, avant même que le gène *MECP2* ne soit identifié. En effet, Camus et al (Camus et al. 1994, 1996) et Krepischi et al. (Krepischi et al. 1998) avaient montré dans des cas d'inactivation modérément biaisée (65-35%), une inactivation préférentielle du chromosome X paternel.

L'origine paternelle des mutations explique l'atteinte exclusivement féminine du syndrome de Rett sans biais dans le sex-ratio de la descendance. En effet, les garçons sont naturellement protégés des formes sporadiques de la maladie puisqu'ils ne reçoivent pas leur chromosome X de leur père, il en découle donc que le syndrome de Rett ne peut plus être considéré comme létal chez les garçons. De plus, la faible proportion de mutations d'origine maternelle dans les cas sporadiques, peut être interprétée comme étant proportionnelle au taux plus faible de méthylation des dinucléotides 5'CpG3' dans la lignée germinale féminine. Par ailleurs, plusieurs mutations d'origine maternelle ont été retrouvées dans les cas familiaux (Amir et al. 1999 ; Amir et al. 2000 ; Wan et al. 1999 ; Cheadle et al. 2000). La recherche de la mutation chez les parents de la mère vectrice révélerait certainement une origine paternelle. Enfin, la récurrence modérée des cas familiaux dans le syndrome de Rett, pourrait résulter de la probabilité faible de l'occurrence de deux événements : mutation dans le gène *MECP2* et contre-sélection des cellules mutées par une inactivation biaisée.

III. Gène *MECP2*, syndrome de Rett, et retard mental chez le garçon

A deux reprises, Rett (Rett 1966) et Hagberg (Hagberg et al. 1983) décrivirent uniquement des filles atteintes du syndrome de Rett. Ces premières descriptions firent penser que le mode de transmission était dominant lié au chromosome X avec un effet létal chez les garçons. L'absence de biais dans le sex-ratio remit en cause cette hypothèse. En 1996, le modèle de Thomas (Thomas 1996) proposant une origine paternelle des néomutations a permis de résoudre le problème : les garçons sont protégés de la maladie puisqu'ils reçoivent le chromosome Y de leur père. Ainsi quelque soit l'hypothèse, les garçons ne semblaient pas concernés par le syndrome de Rett. Toutefois, dès le début des années 90, des hommes présentant certains critères cliniques nécessaires et aucun des critères d'exclusion du syndrome de Rett furent rapportés (pour revue Jan et al. 1999). Aussi en 1994, Hagberg (Hagberg et al. 1994) proposa d'inclure parmi les différentes catégories de variants du syndrome de Rett, une catégorie correspondant aux garçons. Malgré cette définition, et étant donné la rareté de ces cas, très peu de garçons ont été diagnostiqués : Jan et al (Jan et al. 1999) ont répertorié un dizaine de cas.

Seuls les garçons appartenant à des cas familiaux de syndrome de Rett sont véritablement considérés comme atteints de syndrome de Rett. Avant la découverte de l'implication du gène

MECP2 dans le syndrome de Rett, trois familles avec des garçons atteints avaient été décrites (Schanen et al. 1997 ; Sirianni et al. 1998). Avec la découverte de l'implication du gène *MECP2*, des mutations ont été détectées dans deux des trois familles confirmant le diagnostic, ainsi que dans une famille supplémentaire rapportée par Villard (Villard et al. 2000). Dans tous ces cas, les garçons ont présenté une encéphalopathie congénitale sévère suivie d'un décès soit dans la période néonatale (Sirianni et al. 1998), soit au cours de la première année (Schanen et al. 1998 ; Villard et al. 2000). Plus récemment, une délétion de 32 pb en position 1154 a été caractérisée dans une famille (Hoffbuhr et al. 2001). La mère porteuse de la délétion est asymptomatique grâce à une inactivation biaisée. Elle a transmis la délétion à au moins un de ses fils qui est mort à 21 mois à la suite de complications respiratoires. Les mutations identifiées chez ces garçons (R168X, 806delG, T158M, et 1154del32, Wan et al. 1999; Villard et al. 2000 ; Hoffbuhr et al. 2001) sont des mutations retrouvées classiquement chez des filles atteintes d'un syndrome de Rett.

Par ailleurs, un garçon présentant une encéphalopathie non létale proche du syndrome de Rett a été décrit avec une mutation somatique en mosaïque dans le gène *MECP2* (241del2) expliquant probablement la survie de cet enfant (Clayton-Smith et al. 2000) et confirmant la grande variabilité phénotypique du syndrome de Rett observée dans les cas sporadiques et les cas familiaux chez les garçons. Enfin, Monros et al. (Monros et al. 2001) rapporte le cas d'un garçon atteint d'un syndrome de Rett classique (à 7 ans, il remplit huit des neuf critères nécessaires et sept des huit critères d'appoint), présentant la mutation R133H dans le gène *MECP2*. Toutefois, les auteurs signalent que les parents n'ont pu être testés car ils sont décédés : seul un séquençage de 156 chromosomes les conduit à conclure que cette mutation est responsable de la pathologie. Récemment, cette mutation a été identifiée chez une femme atteinte d'un syndrome de Rett (Hoffbuhr et al. 2001).

Ainsi des patients de sexe masculin peuvent être atteints du syndrome de Rett.

Etant donné la grande variabilité phénotypique observée aussi bien chez les filles que chez les garçons, plusieurs auteurs ont recherché des mutations dans le gène *MECP2* chez des individus des deux sexes atteints de retards mentaux divers (Orrico et al. 2000 ; Meloni et al. 2000 ; Imessaoudene et al. 2001 ; Couvert et al. 2001 ; Yntema et al. 2001 ; Yntema et al. 2002).

Orrico et al (Orrico et al. 2000) ont analysé 75 patientes et une famille, présentant différents types de retard mentaux. Des mutations connues ont été retrouvées chez les patientes avec un syndrome de Rett de forme classique ou atypique. Une nouvelle mutation (T203M) a été identifiée chez une femme présentant un autisme classique avec quelques signes du syndrome

de Rett. La mutation faux-sens A140V a été caractérisée dans la famille étudiée chez laquelle la mère est porteuse de cette mutation ainsi que sa fille et quatre de ses fils. La mère et la fille, âgées respectivement de 63 et 41 ans, présentent un retard mental léger non syndromique alors que les quatre garçons sont beaucoup plus gravement atteints (retard mental sévère, langage limité, aucune phase de régression après un début de développement normal n'a été noté chez ces garçons). L'inactivation du chromosome X chez la mère et la fille est aléatoire. Dans cette famille, la mutation A140V est responsable d'une atteinte légère chez les femmes et d'une atteinte grave chez les hommes, sans biais d'inactivation du chromosome X, le gène *MECP2* semble donc se comporter comme un gène impliqué dans un retard mental non spécifique familial dominant lié au chromosome X avec une expressivité variable entre les deux sexes.

L'équipe de Renieri (Meloni et al. 2000) a détecté une mutation non-sens (Q406X) dans une famille chez laquelle d'une part les femmes porteuses de la mutation ne sont pas atteintes et ont une inactivation du chromosome X aléatoire, et d'autre part les garçons porteurs de l'allèle muté présentent un retard mental sévère accompagné de quelques symptômes observés au cours du syndrome de Rett (absence de langage, ataxie de la marche, crises convulsives, bruxisme) et de différences importantes (pas de pertes des acquisitions, et pas de microcéphalie acquise).

Enfin, Yntema et al. (Yntema et al. 2002) ont détecté une délétion en phase de 240 pb (1161-1400del) dans une famille de RMX. Seuls les garçons portant la délétion présente un retard mental non spécifique léger. La femme vectrice n'est pas atteinte, probablement du fait d'une inactivation extrêmement biaisée (entre 15-85% et 0-100%).

Des études récentes confirment et complètent ces données. En effet, deux mutations (E137G et R167W) dans le gène *MECP2* ont été identifiées dans des familles de retard mental non syndromique avec des garçons atteints et des femmes vectrices normales sans biais d'inactivation (Couvert et al. 2001). Par ailleurs, Couvert et al. ont analysé quelques 185 patients retardés mentaux pour lesquels la recherche de triplets CGG dans le gène *FMRI*, responsable du syndrome de l'X-fragile, s'est révélée négative. Ils ont identifié les modifications de séquence suivantes: A140V, P399L, et R453Q ; la mutation A140V identifiée deux fois dans cette étude, avait déjà été rapportée par Orrico et al. (Orrico et al. 2000). Dans cette série, la fréquence de mutations dans le gène *MECP2* chez les garçons retardés mentaux est d'environ 2%, proche de celle du gène *FMRI* estimée à 3-4% dans les retards mentaux. Une étude similaire à partir de 500 patients retardés mentaux non atteints du syndrome de l'X-fragile confirme la fréquence proche de 2% de mutations au sein du gène

MECP2 dans la population hollandaise (Yntema et al. 2001). De plus, la mutation A140V a été aussi rapportée dans une famille atteinte du syndrome PPM-X (psychose, signes pyramidaux et macro-orchidisme, lié au chromosome X). C'est la première fois qu'elle est décrite dans un retard mental lié au chromosome X syndromique. Cette mutation semble être un point chaud de mutation dans le cas de retards mentaux liés au chromosome X chez les garçons (Klauck et al. 2002). Imessaoudene et al. (Imessaoudene et al. 2001) ont analysé 78 enfants (40 filles, 38 garçons) diagnostiqués cliniquement comme atteints d'un syndrome d'Angelman (syndrome pouvant constituer un diagnostic différentiel du syndrome de Rett surtout avant l'âge de 5 ans) mais pour lesquels le profil de méthylation de la région 15q11q13 est normal. Au total, les auteurs ont identifié 6 mutations (dont une récurrente) : 5 parmi les filles et 1 parmi les garçons. Les 5 mutations détectées chez les filles (R106W, R255X (deux fois), 803delG et R270X) sont des mutations classiques du syndrome de Rett. Une ré-évaluation clinique a montré que quatre des cinq patientes présentaient les critères du syndrome de Rett ; la cinquième patiente a une encéphalopathie progressive mais qui a débuté dans la période néonatale. La sixième mutation (G428S) a été caractérisée chez un garçon atteint d'une encéphalopathie non progressive à début néonatal, elle est héritée de la mère asymptomatique qui a une inactivation aléatoire. La figure 31 rassemble les différentes mutations décelées chez des garçons retardés mentaux.

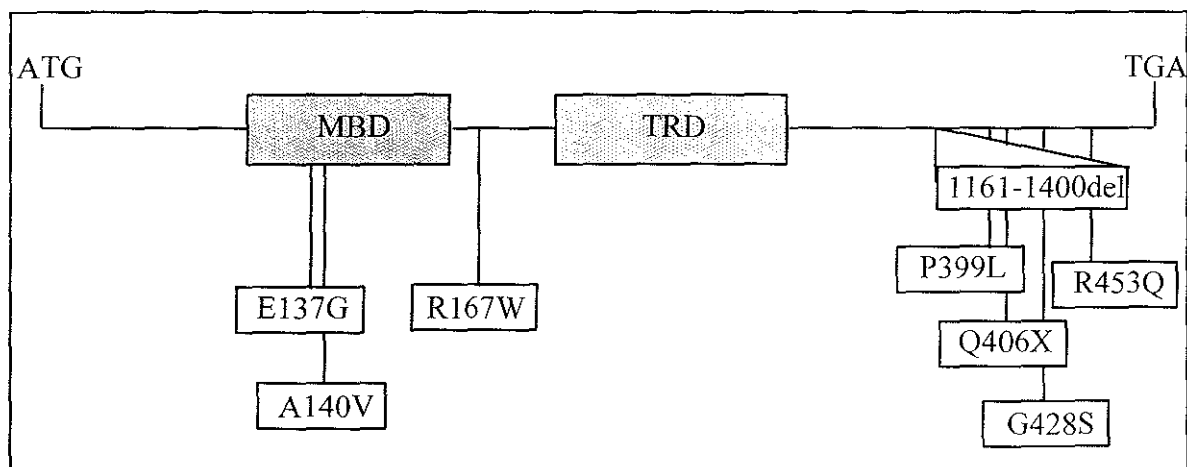


Figure 31 : Localisation sur le gène *MECP2* des mutations identifiées chez des garçons avec un retard mental.

L'ensemble de ces résultats semble mettre en évidence une implication grandissante du gène *MECP2* dans des retards mentaux autres que le syndrome de Rett. En effet, jusqu'à

aujourd'hui, toutes les mutations identifiées chez les garçons présentant un retard mental sont différentes de celles détectées chez les patientes atteintes d'un syndrome de Rett et de plus, les femmes vectrices ne présentent pas de biais d'inactivation, excepté dans la famille rapportée par Yntema et al. (Yntema et al. 2002). Ceci montre que les mutations dans le gène *MECP2* sont responsables d'une grande hétérogénéité clinique (syndrome de Rett, retards mentaux non spécifiques, syndrome PPM-X). Par ailleurs, les mutations à l'état hémizygote identifiées dans le cadre de retards mentaux, compatibles avec la vie suggèrent que les acides aminés modifiés auraient un rôle moins crucial dans le fonctionnement normal de la protéine. On peut noter que plus de la moitié des mutations décrites chez les garçons se situent dans la partie C-terminale de la protéine, en aval des deux domaines fonctionnels (figure 31). L'ensemble de ces données posent donc la question d'un recrutement élargi des patients dans le cadre de la recherche de mutations dans le gène *MECP2*. Cependant, des éléments récents viennent pondérer ces premiers résultats. En effet, Moncla et al. (Moncla et al. 2002) ont identifié la substitution S359P chez un garçon retardé mental dont la description clinique correspondait à celle des cas précédemment publiés; une étude familiale sur trois générations a montré que cette substitution était en réalité un polymorphisme car elle est portée par un cousin sain. Ce résultat exclut toute implication de cette substitution dans la pathologie du patient. De la même façon, nous avons analysé au laboratoire une cohorte de 250 patients présentant un retard mental sans expansion de triplets CGG dans le gène *FMRI*. Les premiers résultats ont révélé qu'un patient portait le changement faux-sens P376S que nous avons identifié comme un polymorphisme dans la famille B (partie Résultats paragraphe 3.2). Cette modification faux-sens a été par la suite publiée comme étant un polymorphisme (Sung Jae Lee et al. 2001). Ces deux résultats soulignent l'importance des études familiales (en plus de la recherche dans la population générale), et attirent l'attention sur les précautions à prendre lors d'un conseil génétique avant de conclure à une mutation responsable de la pathologie.

IV. Corrélation génotype-phénotype et rôle de l'inactivation du chromosome X

1. Corrélation génotype-phénotype

Etant donné la très grande variabilité phénotypique du syndrome de Rett chez les filles, dès les premières publications de mutations dans le gène *MECP2*, les auteurs ont tenté de dégager une corrélation entre le type de la mutation et le phénotype de la patiente.

Amir et al. (Amir et al. 2000) ont établi une corrélation positive entre la présence de mutations tronquantes et deux paramètres : les troubles respiratoires et les quantités faibles d'acide homovanillique dans le liquide cérébro-spinal. La scoliose leur semblait plus fréquente chez les patientes présentant une mutation faux-sens. Toutefois, les autres traits cliniques caractéristiques étudiés (âge de l'apparition de la maladie, du décès, crises convulsives et retard de croissance) étaient équivalents quelque soit la mutation. Une autre étude (Cheadle et al. 2000) a évalué l'existence d'une corrélation génotype-phénotype par établissement de scores selon les symptômes présentés. Les patientes porteuse d'une mutation faux-sens semblaient moins atteintes que celles avec une mutation tronquante. Cheadle et al. ont, d'autre part, trouvé que les mutations localisées dans le domaine TRD ou en aval de celui-ci engendraient des atteintes moins graves que celles localisées en amont de ce domaine. Cependant, les auteurs ne pouvaient conclure statistiquement quant à une véritable relation entre le type de mutation et le phénotype. La mutation R255X, par exemple, est à l'origine de phénotypes tout à fait différents. Une patiente porteuse de cette mutation peut saisir voire tenir brièvement un objet dans ses mains, comprendre, dire quelques mots et marcher seul. Une autre patiente avec cette même mutation, est incapable d'utiliser ses mains, ne parle pas et ne marche pas (Cheadle et al. 2000). De la même manière, Huppke et al. (Huppke et al. 2000) et Nielsen et al. (Nielsen et al. 2001-a) n'ont pu définir une corrélation nette entre le génotype et le phénotype.

Ces résultats montrent que très certainement la variabilité clinique du syndrome de Rett est influencée par des facteurs autres que le type de la mutation au sein du gène *MECP2*.

Un de ces facteurs est l'inactivation du chromosome X et plus précisément du gène *MECP2*.

2. Influences de l'inactivation du chromosome X sur le phénotype

2.1. Dans les cas sporadiques

Avant la découverte de l'implication du gène *MECP2* dans le syndrome de Rett, l'étude de l'inactivation du chromosome X chez les patientes avait montré que celle-ci était aléatoire à modérément biaisée (65-35%) (Zoghbi et al. 1990-b ; Webb et al. 1993, Anvret et al. 1994 ; Camus et al. 1994, 1996).

Amir et al. (Amir et al. 2000) ont réalisé une étude du profil d'inactivation du chromosome X chez 41 patientes mutées dans le gène *MECP2* et trouvé 91% de cas avec une inactivation aléatoire dans les leucocytes. Chez deux patientes, ils ont pu réaliser l'étude à partir de tissu cérébral qui a révélé, là encore, une inactivation aléatoire.

Parmi les patientes ayant une inactivation biaisée, deux ont une mutation introduisant un codon stop précoce : Y141X et 411delG. Suite à la description de ces mutations entraînant la fabrication d'une protéine tronquée, l'hypothèse selon laquelle une inactivation préférentielle du chromosome X muté protégerait la patiente d'effets trop délétères de ces mutations a été proposée. Ces mutations conduiraient à une perte totale de fonction de la protéine MeCP2 dans certaines cellules, du fait de l'action du "système de surveillance des ARNm" (the nonsense-mediated mRNA decay, Maquat 1996 ; Culbertson 1999), empêchant ainsi la synthèse de protéine altérée. Cette hypothèse est corroborée par la fréquence très faible de mutations localisées dans la région 5' du gène et par les premiers résultats obtenus chez la souris (Tate et al. 1996). En effet, Tate et al. ont montré que des embryons chimères mâles *Mecp2*^{-y} ne sont pas viables après 12,5 jours de gestation.

Toutefois, le cas d'une mutation non-sens (Q19X) tronquant très précocement la protéine a été décrite chez une patiente qui ne présente qu'une inactivation modérément biaisée (30-70%) dans le sang (Kim et al. 2000). De plus, Nielsen et al. (Nielsen et al. 2001-a) ont rapporté une délétion d'une cytosine (76delC) engendrant une protéine putative de 31 acides aminés. Cette patiente présente d'une part un profil d'inactivation du chromosome X de 65-35%, et d'autre part un phénotype classique du syndrome de Rett non aggravé comme l'aurait laissé supposer l'hypothèse d'un effet très délétère de ce type de mutation associé à une inactivation quasiment aléatoire.

Enfin, nous avons réalisé l'étude de l'inactivation du chromosome X à partir des leucocytes, au locus du gène codant le récepteur aux androgènes (HUMARA, Allen et al. 1992) chez 59

de nos patientes présentant une mutation dans le gène *MECP2*. 86% des patientes (51/59) montrent une inactivation aléatoire, une seule patiente possède une inactivation fortement biaisée (>95-5%) et les sept cas restant ne sont pas informatifs à ce locus. La patiente montrant une inactivation biaisée est porteuse d'une mutation non-sens particulière qui engendre la survenue d'un codon stop précoce, après les 48 premiers acides aminés de la protéine (mutation S49X). Cette patiente a les caractères typiques du syndrome de Rett, excepté l'épilepsie. Ce cas-là confirmerait l'hypothèse d'une protection contre les mutations trop délétères. Cependant, nous avons identifié cette même mutation une deuxième fois ainsi que deux autres mutations précoces (117insA et E10X) associées à des profils d'inactivation du chromosome X aléatoires. Aucune des patientes ne montre de phénotype aggravé.

2.2. Dans les cas familiaux

Le rôle de l'inactivation du chromosome X, sur l'expression de la maladie dans les cas familiaux de syndrome de Rett, est beaucoup plus clairement établi. En effet, dans plusieurs cas familiaux décrits (Wan et al. 1999), la mère vectrice obligatoire porte bien la mutation mais n'a pas de phénotype (ou extrêmement léger) grâce à une inactivation complètement biaisée. Ces cas sont extrêmement importants à dépister dans le cadre d'un conseil génétique éclairé puisque le risque de récurrence de la maladie est alors de 50%.

Toutefois, la recherche de mutations chez les parents peut se révéler infructueuse dans le cas de mutations germinales en mosaïque. Un tel cas de mosaïque germinale a été rapporté (Wan et al. 1999). Cette observation, certes rare, peut conduire à proposer un diagnostic prénatal pour tout couple ayant déjà une enfant atteinte de syndrome de Rett confirmé sur le plan moléculaire.

En conclusion, l'extrême variabilité clinique du syndrome de Rett ne semble donc pas définie uniquement par le type de mutation et le profil d'inactivation du gène *MECP2*, puisqu'un profil d'inactivation aléatoire est retrouvé dans la quasi-totalité des cas sporadiques, et qu'une même mutation peut donner des phénotypes divers. L'analyse des conséquences des différentes mutations sur la fonction de la protéine MeCP2 et potentiellement sur ses gènes cibles aidera peut-être à mieux comprendre cette variabilité.

V. Hypothèse d'un second locus impliqué dans le syndrome de Rett

Chez 15 à 20% des patientes présentant un syndrome de Rett typique, aucune mutation dans la séquence codante (et pour notre série, aucun remaniement chromosomique) du gène *MECP2* n'a été mise en évidence. L'analyse des régions régulatrices et non codantes permettra peut-être de mettre en évidence de nouvelles mutations. Toutes ? La question se pose, soulevant l'hypothèse d'un autre locus potentiel. La possibilité de plusieurs locus dans l'étiologie du syndrome de Rett, avait été suggérée par Bühler et al. au début des années 90 (Bühler et al. 1990). En effet, les auteurs proposaient un modèle d'interférence interallélique entre un locus lié au chromosome X et un locus autosomique. Plus récemment, Xiang et al. (Xiang et al. 2000) ont testé le gène *MECP2* chez cinquante-neuf cas sporadiques et chez neuf familles. Des mutations dans le gène *MECP2*, ont été détectées dans seulement 35% des patientes, incitant les auteurs à analyser six autres locus candidats autosomiques (*UBE21*, *CDR2*) et gonosomiques (*UBE1*, *GdX*, *SOX3* et *GABRA3*) choisis en fonction de leurs rôles connus soit dans l'ubiquitinylation des protéines (*UBE1*, *UBE21*, *GdX*), soit dans des fonctions neuronales (*SOX3*, *GABRA3*, *CDR2*). Aucune mutation dans ces gènes candidats n'a été décelée.

Deux articles ont, par la suite, étayé l'hypothèse d'un second locus. Dans le premier, Rosenberg et al. (Rosenberg et al. 2001) rapporte le cas d'une enfant atteinte d'un syndrome de Rett associé au caryotype suivant : 46,X,r(X). Des expériences par hybridation *in situ* ont montré que le point de cassure, pour le bras court, était proximal de la bande Xp11.3, et que pour le bras long, il était en (ou distal de) Xq13. L'anneau ne contient pas le gène *MECP2*. L'inactivation du chromosome X chez cette patiente, est totalement biaisée. Comme attendu dans les cas de réarrangements du chromosome X, c'est l'anneau qui est exclusivement inactivé (il contient le gène *XIST*). Le gène *MECP2*, porté par le chromosome X actif, n'est quant à lui pas muté. Les auteurs proposent que le second locus putatif aurait une pénétrance et une expressivité plus faibles que le gène *MECP2*, puisqu'il serait responsable de seulement 20% des cas de syndrome de Rett. De plus, ce second locus rendrait compte de la proportion importante de cas familiaux de syndrome de Rett non mutés dans le gène *MECP2*, supérieure à 70% d'après ces auteurs. Leur modèle prédit que les femmes hétérozygotes seraient asymptomatiques du fait de la faible pénétrance de l'allèle muté du second locus et d'une inactivation aléatoire. Par contre, l'apparition du syndrome de Rett chez les patientes serait due à l'association de la mutation avec une inactivation préférentielle du chromosome X normal. A partir de ce modèle, les auteurs expliquent que chez leur patiente, l'expression de la

pathologie est due à la présence de l'anneau qui est préférentiellement inactivé, maintenant ainsi actif l'autre chromosome X portant une mutation dans le second locus. Cette situation est différente de celle rencontrée dans les pathologies dites du point de cassure. Ces pathologies du point de cassure ont été mises en évidence entre autres, chez des femmes porteuses d'une t(X;A) et présentant par exemple, un RMLX de transmission récessive. L'expression du retard mental est due chez ces femmes à l'interruption du gène sur le chromosome X transloqué associée à une inactivation préférentielle du chromosome X normal. D'après le modèle de Rosenberg et al. (Rosenberg et al. 2001), les femmes porteuses d'un réarrangement du chromosome X et atteintes d'un syndrome de Rett exprimeraient donc dans toutes leurs cellules l'allèle muté, comme si elles étaient hémizygotes. Si ce modèle est vrai, devraient être décrits des garçons présentant un syndrome de Rett classique. Afin d'expliquer l'absence de garçons par un effet létal, les auteurs avancent deux hypothèses : soit une expression à partir des deux chromosomes X, avant la mise en place de l'inactivation chez les filles, soit une expression de gènes à partir du chromosome X inactif. La première hypothèse suppose l'absence de garçons par un effet létal de la mutation au deuxième locus, car les garçons ne portent que la copie mutée du gène, donc l'allèle normal ne pourra jamais s'exprimer comme chez les filles. Il est à noter toutefois, que l'inactivation du chromosome X se met en place relativement tôt au cours de l'embryogenèse (lors de la formation de la cavité de segmentation à 4-5 jours (Pereira et al. 2000)), laissant peu de temps d'expression à l'allèle normal avant son extinction. Dans la seconde hypothèse, ce sont des gènes échappant à l'inactivation qui compensent l'effet létal de la mutation au second locus, ces gènes seraient alors localisés dans la région péricentromérique du chromosome X car c'est la région présente dans l'anneau. Une objection peut être faite à l'encontre du modèle proposé par Rosenberg et al. (Rosenberg et al. 2001), il nécessite la présence indispensable d'une inactivation biaisée chez les patientes. Nous avons étudié le profil d'inactivation du chromosome X chez une vingtaine de patientes ayant un syndrome de Rett typique et pas de mutation dans la séquence codante du gène *MECP2*, aucune des patientes ne présente une inactivation biaisée du chromosome X, c'est-à-dire un rapport d'inactivation entre les deux chromosomes X supérieur à 9/1.

Le second article (Villard et al. 2001) s'appuie sur l'étude de quatre familles chez lesquelles les filles (deux par famille) présentent un syndrome de Rett typique mais pas de mutation dans le gène *MECP2*. L'étude de l'inactivation du chromosome X montre que les quatre mères ont une inactivation biaisée du chromosome X ainsi que six des huit filles testées. Les auteurs proposent qu'une mutation dans le second gène permettrait sa réactivation à partir du

chromosome X inactif. Chez les mères, pour expliquer l'absence de la pathologie, ce serait le chromosome portant la mutation qui serait actif, alors que le chromosome normal serait inactif. Chez les filles atteintes, le chromosome muté est inactif, mais l'allèle muté échappant à l'inactivation reste actif, ces filles expriment donc deux doses du gène, comme dans le cas d'une disomie fonctionnelle. Ainsi, c'est un effet de dose qui est responsable dans ce modèle, du syndrome de Rett : l'inactivation biaisée n'a qu'un rôle annexe dans l'expression de la maladie, comparé au modèle de l'article précédent, permettant la ségrégation dans la descendance de l'allèle muté. Les garçons ne sont pas atteints, car qu'ils portent le chromosome muté ou le chromosome normal, dans tous les cas, ils n'expriment qu'une seule dose du gène (figure 32). Ces deux articles semblent témoigner directement de l'existence d'un second locus impliqué dans le syndrome de Rett.

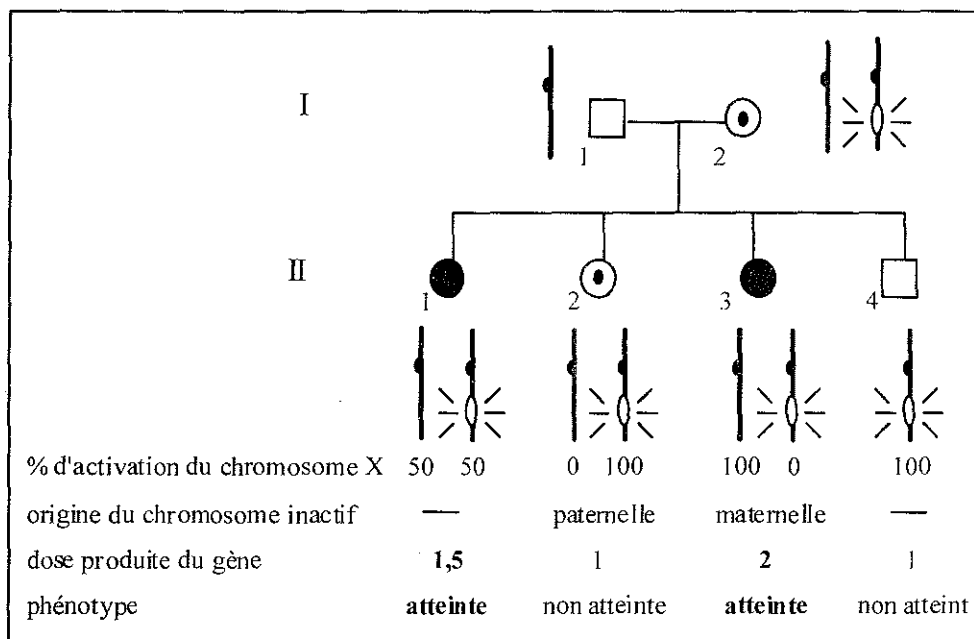


Figure 32 : *Modèle de transmission familiale d'un second gène potentiel dans le syndrome de Rett (d'après Villard et al. 2001).*

Pédigree hypothétique de deux parents et de leurs quatre enfants ayant reçu chacun le chromosome muté de leur mère. Le gène muté échappe à l'inactivation et est responsable du syndrome de Rett. Il est symbolisé par un cercle vide, les traits illustrant l'expression du gène indépendamment de l'inactivation du chromosome. Les chromosomes en rouge sont inactivés exclusivement.

Avant la publication de la responsabilité du gène *MECP2* dans le syndrome de Rett par Amir et collaborateurs (Amir et al. 1999), deux translocations associées à un syndrome de Rett avaient été décrites en 1990, impliquant toutes les deux un chromosome X mais des points de cassure distincts et un autosome différent : t(X;3) (p21.3;q13.31) (Zoghbi et al. 1990-a), et t(X;22) (p11.2;p11) (Journel et al. 1990). Ces deux translocations pourraient constituer aujourd'hui une piste pour le clonage d'un second locus impliqué dans le syndrome de Rett. Seule la description d'un YAC chevauchant le point de cassure sur le chromosome X a été publiée pour la translocation *de novo* (X;3) (Ellison et al. 1993). C'est dans l'hypothèse d'un second gène que nous poursuivons au laboratoire l'étude de la translocation (X;22) familiale, après nous être assurés que les patientes porteuses de la translocation (la mère et ses deux filles) n'étaient pas mutées dans le gène *MECP2*. Nous avons isolé, par FISH, un BAC chevauchant le point de cassure sur le chromosome X. Grâce au projet de séquençage du génome humain, le clone chevauchant est totalement séquencé. Une fois le point de cassure localisé précisément dans le clone, nous envisageons une stratégie de clonage *in silico*. La localisation du point de cassure est actuellement en cours. La première approche de cartographie physique du point de cassure dans le BAC par PCR sur l'ADN extrait d'hybrides somatiques interspécifiques (contenant uniquement le derX) de la mère et d'une des deux filles a donné des résultats discordants, empêchant de déterminer la région précise du point de cassure. Nous réalisons actuellement cette localisation par Southern blot en utilisant comme sondes des fragments de quelques kb du BAC amplifiés par PCR, hybridées sur l'ADN génomique des trois patientes.

La translocation (X;22), et probablement la translocation (X;3), sont associées à une inactivation préférentielle du chromosome X normal (Mattei et al. 1982). On peut leur appliquer soit le modèle de Rosenberg et al. (Rosenberg et al. 2001) soit celui de Villard et al. (Villard et al. 2000). D'après les deux modèles, la pathologie ne proviendrait pas du point de cassure du chromosome X transloqué. Si cette hypothèse est vraie, le second gène potentiellement impliqué dans le syndrome de Rett peut être situé n'importe où sur le chromosome X, expliquant ainsi l'existence de deux points de cassure distincts sur le chromosome X entre les deux translocations décrites. Seule la connaissance exacte du point de cassure, pour chacune des deux translocations, permettra de savoir si la pathologie est due au point de cassure, ou non, d'où l'intérêt de poursuivre l'étude du point de cassure en Xp11.2 pour notre cas.

B. Neuropathologie du syndrome de Rett et fonction de la protéine MeCP2

I. Quelques définitions anatomiques et fonctionnelles

Le système nerveux central (SNC) ou névraxe se compose de trois parties : la moelle épinière, le tronc cérébral et le cerveau. Le **cortex cérébral** ou **pallium** est une fine couche de substance grise (entre deux et quatre millimètres) recouvrant les deux hémisphères. Il est composé de six couches cellulaires superposées (couches I à VI). Le cortex présente une série de sillons dont les plus profonds (scissures) délimitent les **lobes cérébraux** et les plus superficiels les circonvolutions. Les scissures divisent la face externe de chaque hémisphère en quatre lobes : le lobe frontal, le lobe pariétal, le lobe occipital et le lobe temporal. Ces quatre lobes regroupent les différentes aires fonctionnelles du cerveau. Ainsi, le **lobe frontal** contient l'**aire motrice**, le **lobe temporal** l'**aire auditive**, le **lobe pariétal** l'**aire somesthésique** (sensibilités cutanées à l'environnement, et sensibilités propres au corps) et le **lobe occipital** l'**aire visuelle**.

Les **corps striés** ou **striatum** correspondent aux **noyaux caudé** et **lenticulaire**. Ils sont constitués d'amas de corps cellulaires de neurones d'où le nom de noyau ; ils interviendraient dans la régulation des mouvements lents inconscients accompagnant les mouvements volontaires.

L'**hypophyse** ou **glande pituitaire**, organe glandulaire et nerveux est le lieu de synthèse de plusieurs hormones responsables de phénomènes physiologiques variés (croissance tissulaire, action sur la glande thyroïde, les glandes génitales...).

Le **cervelet** est composé de trois parties : le **vermis** médian et deux **hémisphères cérébelleux** latéraux. Il est responsable du maintien de l'équilibre et de la posture, du contrôle du tonus musculaire et de la coordination de l'activité motrice volontaire.

II. Données anatomopathologiques et neuropathologiques du syndrome de Rett

Bien que le signe clinique caractéristique du syndrome de Rett soit l'atteinte cérébrale, le cerveau de patientes ne montre guère d'altérations morphologiques marquées, si ce n'est une taille réduite. Une étude comparée entre 13 patientes atteintes d'un syndrome de Rett et 10 témoins par une approche IRM quantitative (Imagerie par Résonance Magnétique) a été réalisée (Murakami et al. 1992). L'approche par IRM quantitative mesure les dimensions et les aires de différentes structures cérébrales ; Murakami et al. (Murakami et al. 1992) ont étudié les structures suivantes : lobes cérébraux, corps striés, hypophyse, corps calleux, lobules et hémisphères cérébelleux, 4ème ventricule, tronc cérébral. L'ensemble des structures mesurées a une taille diminuée de 12 à 34% chez les patientes, et ceci sans qu'il y ait une preuve d'une diminution du nombre de cellules, d'inflammation, de glioses ou de défaillance de la migration neuronale. Cette réduction globale de la taille cérébrale est plus marquée pour certaines structures cérébrales comme le cortex préfrontal, frontal, temporal externe, le noyau caudé et le corps calleux. Cette hypotrophie ne progresse pas avec l'âge des patientes, témoignant plutôt d'un arrêt de développement que d'un processus dégénératif.

Une des caractéristiques anatomopathologiques observée dans les retards mentaux est la diminution de l'arborisation dendritique. Afin de mettre en évidence une diminution de l'arborisation spécifique du syndrome de Rett, une expérience a été réalisée en comparant l'arborisation dendritique de différentes aires du cortex cérébral (aires pré-motrice, motrice, visuelle...) entre des patientes atteintes d'un syndrome de Rett, des patientes trisomiques 21 (caractérisées par une diminution de la masse du cerveau et une raréfaction de certaines arborisations dendritiques) et des témoins (Armstrong et al. 1998). Les résultats de cette expérience ont montré que l'arborisation dendritique des neurones des couches III et V des cortex frontal, moteur et sensoriel est significativement réduite chez les patientes atteintes d'un syndrome de Rett. Est associée à cette arborisation réduite, une diminution de l'expression de certaines protéines : la cyclooxygénase-2 et MAP2 (Microtubule-Associated Protein 2) qui sont impliquées dans la formation du cytosquelette neuronal, dans le développement dendritique et dans l'établissement de la polarité neuronale. Il est à noter aussi une réduction des épines dendritiques des neurones des couches III et V. La réduction de la taille de l'arborisation du neuropile axono-dendritique ainsi que la conservation du nombre de neurones entraînent une augmentation de la densité cellulaire et le ralentissement de la

croissance céphalique. L'arbre dendritique de l'aire cérébrale visuelle n'est pas atteint.

De plus, Kitt et al. (Kitt et al. 1995) ont observé une hypopigmentation de la substance noire due à une quantité décroissante de neuromélanine dans les neurones.

Enfin, l'atteinte du cervelet est bien particulière chez les patientes présentant un syndrome de Rett. En effet, comme les autres structures cérébrales, le cervelet est hypoplasique, mais à la différence des autres structures, c'est la seule qui évolue avec l'âge des patientes vers une atrophie. Les données autopsiques confirment l'existence de l'atrophie. Les causes de cette atrophie du cervelet ne sont pas identifiées avec certitude (carences nutritionnelles, crises épileptiques, traitements anti-épileptiques, ou atteinte liée directement à la maladie).

En conclusion, la symptomatologie observée dans le syndrome de Rett peut être expliquée par les différentes aires cérébrales atteintes :

- Le handicap moteur est dû à l'atteinte des aires pré-motrices et motrices du cortex cérébral,
- Le déficit des fonctions de mentalisation, d'abstraction et d'anticipation est dû à l'atteinte des aires préfrontales,
- Les émotions (comportement autistique, crises d'angoisse) à l'atteinte du système limbique,
- L'hypotonie, l'ataxie de la marche à l'atteinte du cervelet.

Les aires indemnes comme les aires visuelles, correspondent à des fonctions effectivement préservées dans le syndrome de Rett.

Avec la découverte du gène *MECP2* comme responsable du syndrome de Rett, sans doute a-t-il été pensé que la pathogenèse de cette maladie serait comprise et ceci d'autant plus que la fonction de la protéine codée par le gène *MECP2* était déjà connue au moment de son implication dans le syndrome de Rett. Il n'en est rien. Bien au contraire, le rôle de la protéine MeCP2 dans le syndrome de Rett soulève beaucoup plus de questions qu'il ne donne de réponses. Ainsi, parmi les questions essentielles : i) comment des mutations dans le gène *MECP2*, acteur important de la répression transcriptionnelle perturbent-elles le développement et le maintien de **certaines** arborisations dendritiques ? ii) quelle est la cascade de gènes impliqués ? Des analyses fonctionnelles des différentes mutations dans le gène *MECP2*, l'étude de modèles animaux ainsi que l'analyse du profil d'expression de gènes du cerveau chez des patientes atteintes d'un syndrome de Rett clarifieront probablement le mécanisme pathogénique du syndrome de Rett, en établissant le lien entre la perturbation de processus dépendants de modifications épigénétiques du génome (méthylation de l'ADN, acétylation des histones) et la réduction de taille générale du cerveau.

III. Conséquences de mutations dans le gène *MECP2*

1. Etudes fonctionnelles de protéines MeCP2 anormales

Les mutations non-sens ou celles perturbant le cadre de lecture se situant dans la partie 5' de la séquence codante d'un gène entraînent l'apparition d'un codon stop précoce. Ces mutations conduisent à la synthèse d'ARNm dits non-sens qui sont reconnus par le système de surveillance des ARN et sont dans la plupart des cas rapidement dégradés par le mécanisme de dégradation spécifique des ARN (NMD : nonsense mediated RNA decay, Maquat 1996 ; Culbertson 1999). Si ces mutations sont localisées dans la région 3' de la séquence codante d'un gène, les ARNm synthétisés échapperont au NMD et donneront des protéines tronquées. Excepté les mutations (E10X, 42ins8, 46delC, Q19X, K22X, 76delC, 100del4, 106del7, S49X et 117insA) localisées au début de la région codante du gène *MECP2* et entraînant certainement une dégradation des ARNm synthétisés par le NMD, la majorité des mutations tronquantes du gène *MECP2* sont localisées dans la partie 3' de la séquence codante (dans le dernier tiers) et conduisent probablement à la synthèse de protéines incomplètes. Si ces protéines sont effectivement fabriquées chez les patientes atteintes d'un syndrome de Rett, quel est leur niveau de fonctionnalité ?

Si on se réfère aux expériences qui ont permis de caractériser les deux domaines fonctionnels de la protéine MeCP2 (Meehan et al. 1992 ; Nan et al. 1993, 1997 ; Chandler et al. 1999), il semble que la protéine privée d'un domaine garde la fonctionnalité du domaine intact.

Une première étude *in vitro* sur les conséquences fonctionnelles de mutations faux-sens (R106W, R133C, F155S et T158M) localisées dans le domaine MBD de la protéine MeCP2 de xénope a montré que les protéines anormales dues aux mutations R106W, R133C et F155S ont une affinité réduite d'un facteur 100 pour l'ADN méthylé ; la mutation T158M réduit l'affinité de la protéine seulement d'un facteur 2 (Ballestar et al. 2000).

La même équipe a réalisé une étude similaire mais sur la protéine humaine (Yusufzai et al. 2000). Ils ont étudié l'effet de cinq mutations faux-sens (R106W, R133C, F155S, T158M et R306C) et sept mutations non-sens (L138X, R168X, E235X, R255X, R270X, V288X et R294X) sur l'interaction de la protéine avec l'ADN méthylé, sur les capacités de réprimer la transcription et sur la stabilité de la protéine aberrante. Toutes les mutations faux-sens

(R106W, R133C, F155S et T158M) ainsi que la mutation non-sens (L138X) localisées dans le domaine MBD de la protéine MeCP2 abolissent l'interaction spécifique de la protéine avec l'ADN méthylé. Autrement dit, les protéines anormales perdent leur capacité de reconnaître l'ADN méthylé mais pas leur capacité d'interaction avec l'ADN non méthylé, sauf la protéine tronquée par la mutation L138X qui est incapable de se lier à l'ADN. Les autres mutations situées dans le domaine TRD conservent leurs propriétés de liaison avec l'ADN méthylé.

Les auteurs supposent que les gènes cibles de la protéine MeCP2 ne seraient plus réprimés faute d'interaction de la protéine MeCP2 avec l'ADN méthylé, et que les protéines mutées pourraient peut-être se lier à des séquences non méthylées et altérer ainsi l'expression de certains gènes. Les protéines contenant les mutations non-sens (R168X, E235X, R255X, R270X, V288X et R294X) localisées dans le domaine TRD sont incapables de réprimer la transcription d'un gène rapporteur. Par contre, la mutation R306C (très fréquemment retrouvée chez les patientes) ne semble pas perturber la capacité de répression. En effet, la protéine mutée réprime la transcription avec la même intensité que la protéine normale. Les études sur la stabilité des différentes protéines mutées *in vivo* dans des ovocytes de xénope mettent en évidence que les protéines les plus tronquées sont les moins stables (8 heures après l'arrêt de la traduction, il ne reste que 10% de protéines mutées par la mutation L138X alors qu'il reste 60% de la protéine mutée R294X).

Une troisième étude (Kudo et al. 2001) a porté sur quatre mutations faux-sens (R106W, R133C, F155S, et T158M). Afin d'analyser la liaison des protéines mutées avec l'ADN méthylé, les auteurs ont utilisé des cellules de souris qui ont les régions d'hétérochromatine enrichies en cytosines méthylées. Il a été montré que la protéine MeCP2 normal s'accumule, chez la souris, dans ces régions et est facilement visualisable en fluorescence sous forme de grains (Nan et al. 1993). Les auteurs ont transfecté les cellules de souris, par les protéines MeCP2 mutées fusionnées à la GFP. Les auteurs ont mis en évidence une coloration diffuse dans les noyaux et non granuleuse, témoignant d'une mauvaise liaison des protéines portant les mutations R106W ou F155S ou T158M avec les cytosines méthylées. Ils ont étudié, par ailleurs, ces mêmes mutations dans des cellules de drosophiles (la drosophile est dépourvue de méthylation) dans le but d'analyser l'effet de ces mutations sur la répression transcriptionnelle conduite par MeCP2. Ils ont obtenu des résultats superposables aux précédents : les protéines mutées ne répriment pas le gène rapporteur. Seule la mutation R133C ne semble réduire ni l'affinité de la protéine pour l'ADN méthylé (les auteurs obtiennent une coloration granuleuse), ni l'action de répression de la protéine, à la différence des résultats précédemment publiés (Ballestar et al. 2000, Yusufzai et al. 2000). Les auteurs invoquent

l'implication du contexte cellulaire qui aiderait, dans certains cas, la protéine anormale à retrouver une affinité pour l'ADN méthylé, ou bien supplérait par une protéine de l'hôte. D'autres expériences sont nécessaires pour déterminer l'origine de ces différences.

Enfin, Kudo et al. (Kudo et al. 2002) ont utilisé les mêmes systèmes d'études de la protéine MeCP2 (c'est-à-dire les cellules de souris et de drosophile) pour analyser l'effet de deux mutations faux-sens (A140V et E137G) mises en évidence chez des garçons. Ils ont étudié en parallèle deux mutations faux-sens fréquentes dans le syndrome de Rett : R106W et T158M. A la différence de ces deux mutations, les mutations A140V et E137G n'abolissent pas la capacité de la protéine MeCP2 d'interagir avec l'ADN méthylé. Les auteurs ont obtenu une coloration granuleuse dans les cellules de souris identique à celle des cellules transfectées par la protéine MeCP2 normale. Les résultats obtenus dans les cellules de drosophile montre, en accord avec les résultats de fluorescence, que la mutation A140V ne réduit aucunement l'activité de répression transcriptionnelle de la protéine mutée ; elle aurait plutôt tendance à l'augmenter. La protéine portant la mutation E137G diminue de façon modérée la répression transcriptionnelle par rapport au témoin. Ainsi, les auteurs concluent que ces mutations ne semblent modifier que très peu l'activité normale de la protéine MeCP2, ce qui expliquerait la survie des garçons à l'âge adulte. Ces derniers résultats confirmeraient l'hypothèse selon laquelle les acides aminés modifiés chez les garçons auraient un rôle moins crucial dans le fonctionnement de la protéine MeCP2.

Les résultats de ces études soulignent deux effets possibles des mutations sur la protéine MeCP2 : i) soit les mutations dans le domaine MBD entraînent une perte totale de fonction de la protéine puisque celle-ci ne peut plus se lier à l'ADN méthylé, quelque soit la position de la mutation testée dans le domaine MBD. Dans ces cas-là, le syndrome de Rett serait dû à une haplo-insuffisance de la protéine MeCP2. La moitié de l'activité biologique normale est insuffisante pour empêcher l'apparition de la maladie. L'haplo-insuffisance peut s'expliquer par le fait que la protéine MeCP2 normale se lie à l'ADN méthylé via un dinucléotide 5' CpG 3' symétriquement méthylé ; de telles séquences sont présentes environ toutes les 150 pb dans le génome humain (Chandler et al. 1999), la quantité de protéine normale ne suffirait donc pas pour compenser le manque de protéines dû à une mutation. ii) soit la protéine mutée dans son domaine MBD garde une fonction résiduelle d'interaction avec l'ADN non méthylé suffisante pour acquérir une nouvelle fonction en agissant sur de nouveaux gènes cibles. Dans ce cas-là, les mutations dans le gène *MECP2* auraient un effet dominant négatif.

2. Conséquences des mutations sur la fonction du complexe de répression transcriptionnelle

La protéine MeCP2 interagit avec le corépresseur Sin3A pour recruter les HDAC1 et 2 qui en désacétylant les résidus lysine (K) des histones H3 et H4 condensent la chromatine et la rendent ainsi transcriptionnellement inactive.

Wan et al (Wan et al. 2001) ont testé l'effet de deux mutations tronquant le domaine TRD (R168X et 803delG) sur l'état acétylé (ou non) des histones dans des cellules de patients atteints d'un syndrome de Rett. Ils ont examiné le niveau d'expression des protéines MeCP2 mutées ainsi que le niveau d'acétylation des histones à partir de cellules cultivées provenant d'une femme pour la mutation R168X, et d'un homme pour la mutation 803delG. Les résultats témoignent d'une hyperacétylation spécifique de l'histone H4 par rapport à l'histone H3. L'histone H4 peut être acétylée au niveau de quatre résidus lysine particuliers : K5, K8, K12 et K16 (l'acétylation de H3 se produit sur les résidus K9, K14, K18 et K23). L'utilisation d'un anticorps dirigé spécifiquement contre chaque résidu acétylé indique que l'hyperacétylation a lieu sur le résidu lysine 16 de l'histone H4. La présence de l'histone H4 hyperacétylée en K16 en présence de la protéine MeCP2 déficiente peut être interprétée de deux façons : i) soit l'absence de la protéine MeCP2 fonctionnelle entraîne la formation d'un complexe represseur qui ne peut recruter les HDAC et notamment une HDAC spécifique de la désacétylation du résidu 16 de H4, et donc H4 reste acétylée. ii) soit un gène cible de la répression par la protéine MeCP2 serait une HAT spécifique du résidu 16 de H4, et ainsi en absence de la protéine MeCP2 fonctionnelle, la HAT hyperacétylerait K16 de H4. Ces deux cas sont possibles, étant donné que des HAT spécifiques du résidu 16 de H4 ont été identifiées chez la drosophile : HAT MOF (Akhtar et al. 2000), et chez *S. cerevisiae* : GCN5 (Kuo et al. 1996 ; Zhang et al. 1998) et que des HDAC spécifiques ont aussi été rapportées. Par exemple, les protéines SIR2 de la levure et des mammifères ont une activité HDAC qui cible préférentiellement les résidus lysine 16 de H4 mais aussi les résidus 9 et 14 de H3 (Imai et al. 2000).

Toutes ces données suggèrent un rôle fonctionnel important de l'état d'acétylation du résidu 16 de H4 dans la régulation des gènes. Ces résultats constituent une étape supplémentaire dans la compréhension du fonctionnement du complexe represseur associé à la protéine MeCP2.

3. Les modèles animaux

3.1. Définitions

Une autre approche pour comprendre les conséquences de la déficience d'une protéine consiste à utiliser des modèles animaux (les principaux organismes modèles étant le nématode, la drosophile ou la souris). Le modèle animal le plus utilisé aujourd'hui dans la compréhension de la physiopathologie de l'espèce humaine est la souris. En effet, la souris présente un développement similaire à celui de l'Homme. De plus, l'inactivation de gènes homologues (à ceux de l'Homme) chez la souris donne des phénotypes proches entre les deux espèces, et enfin, la souris est facile à élever. Les progrès récents de la biologie moléculaire, en facilitant les manipulations du génome de souris, fournissent de nouveaux moyens d'explorer les fonctions de différents gènes. Pour cela, un gène d'intérêt est invalidé et les effets sont analysés dans la descendance. L'utilisation de cellules souches embryonnaires (cellules ES) de souris offre la possibilité de réaliser une mutagenèse ciblée sur le gène d'intérêt par recombinaison homologue. Après sélection des cellules souches ayant acquis le gène muté, celles-ci sont transfectées dans un blastocyste qui est réimplanté dans une souris pseudo-gestante. La souris développée à partir du blastocyste est une mosaïque pour le gène muté (seules les parties de la souris dérivant des cellules ES mutées sont porteuses du transgène). Si les cellules ES mutées sont à l'origine de la lignée germinale de la souris, la descendance de celle-ci obtenue par un croisement avec une souris normale, portera la mutation dans toutes les cellules à l'état hétérozygote. Un croisement entre deux souris hétérozygotes permettra d'obtenir des souris homozygotes mutées pour le gène d'intérêt, c'est le **knock-out**, à la condition que l'invalidation des deux copies du gène soit viable. Dans le cas des mutations dominantes, on peut étudier des souris hétérozygotes pour la mutation, c'est le **knock-in**.

Il est possible d'invalider un gène de façon conditionnelle par le système **Cre-loxP**. Le gène *cre* code une recombinase du phage P1 qui catalyse la recombinaison entre deux courtes séquences de 34 pb appelées loxP (figure 33). Ainsi, après avoir obtenu une lignée contenant la séquence à muter "floxée", c'est-à-dire encadrée par deux séquences loxP, c'est en fonction de la présence de la recombinase que la recombinaison homologue et donc la mutagenèse a lieu. Le gène *cre* peut être mis sous le contrôle de promoteurs actifs de manière tissu-spécifique.

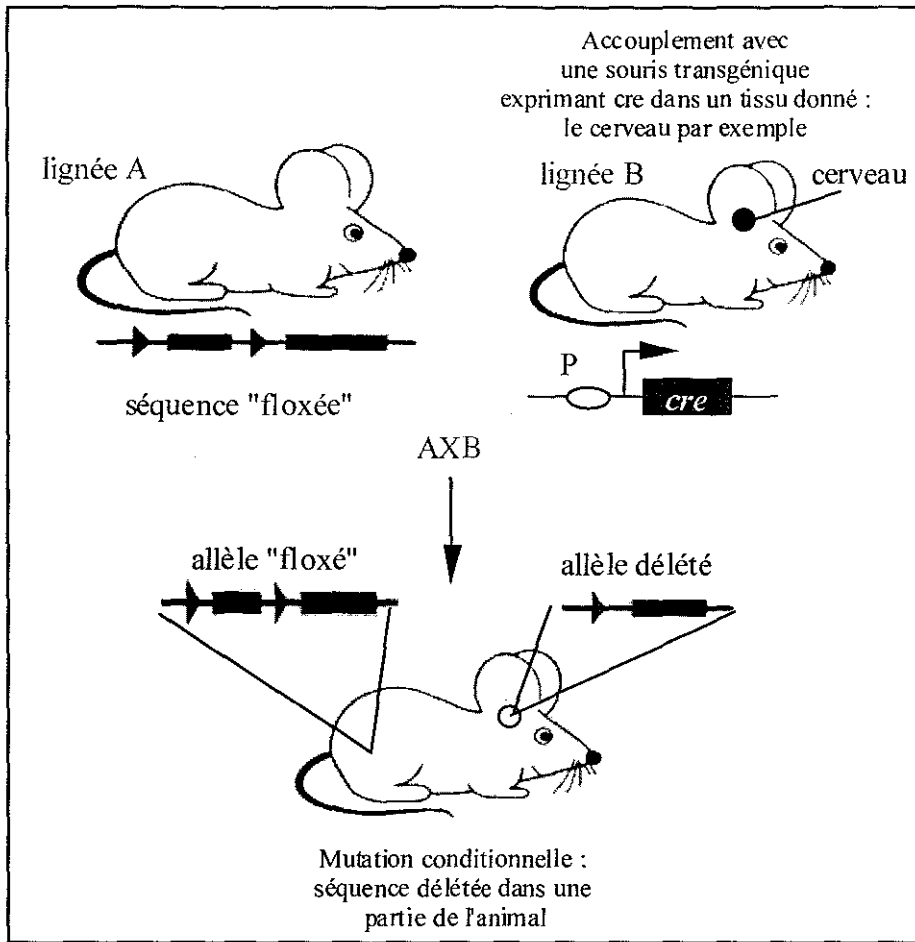


Figure 33 : Le système Cre-LoxP (d'après Babinet et al. 2000)

3.2. Modèles murins du syndrome de Rett

Une première étude à partir d'un modèle murin (Tate et al. 1996) avait montré que si le gène *MECP2* n'était pas indispensable à la multiplication de cellules ES mâles, il l'était en revanche, pour le développement des embryons mâles dérivés des cellules ES mutées. Récemment, de nouveaux modèles murins ont été établis en utilisant le système de recombinaison Cre-loxP (Chen et al. 2001 ; Guy et al. 2001). L'équipe américaine (Chen et al. 2001) a obtenu quatre modèles murins : des souris mâles déficientes en protéine MeCP2 (*Mecp2*^{-y}, modèle 1), et deux types de souris avec une mutation conditionnelle spécifique du cerveau. Une des mutations transitoires du cerveau a été obtenue grâce au contrôle du gène *cre* par le promoteur *Nestin* qui entraîne plus de 90% de recombinaison spécifiquement dans le cerveau dès le douzième jour de gestation (modèle 2). La seconde mutation conditionnelle est obtenue par le transgène appelé CamK-cre93 (la recombinaise cre codée par le transgène

est sous le contrôle du promoteur de la kinase Cam) ; l'utilisation de ce transgène permet d'activer, dans le cerveau, la recombinaison entre les séquences loxP tardivement dans la période post-natale (modèle 3).

Les modèles 1 et 2 ont le même phénotype avec la même évolution. Les souris mâles sont normales jusqu'à la cinquième semaine, puis présentent des tremblements, des troubles respiratoires et comportementaux. Après une période d'obésité et d'hypoactivité, leur état général se dégrade et ces souris meurent environ vers la dixième semaine. Les femelles hétérozygotes (modèle 4) sont normales durant les quatre premiers mois de vie, puis elles prennent du poids et présentent des troubles de la coordination motrice. L'autopsie révèle une réduction sensible de la taille du cerveau et de la taille des neurones, mais pas de signe de dégénérescence. Le modèle 3 montre une histoire naturelle analogue mais d'apparition plus tardive.

L'équipe anglaise de Bird (Guy et al. 2001) a aussi obtenu des souris filles homozygotes *Mecp2*^{-/-} ou mâles hémizygotés *Mecp2*^{-y} (modèle 1') et des souris avec une mutation conditionnelle dans le cerveau en utilisant le même promoteur *Nestin* que Chen et collaborateurs (Chen et al. 2001) (modèle 2'). Les souris homozygotes *Mecp2*^{-/-} ou hémizygotés *Mecp2*^{-y} sont viables et fertiles. Si leur développement, en particulier le poids corporel, varie en fonction des lignées utilisées, il n'existe pas de différence sur le plan neurologique. Les souris *Mecp2*^{nul} apparaissent normales à la naissance et dans les semaines qui suivent. Entre les troisième et huitième semaines, sont observées une réduction des mouvements avec une incoordination motrice, une contracture des pattes arrière, une respiration irrégulière et une dentition anormale. Les souris mâles *Mecp2*^{-y} ont des testicules internes. L'étude histologique de différents tissus dont le cerveau n'a montré aucune anomalie. La mort survient en général vers le cinquante-quatrième jour. Comme dans l'article précédent, le phénotype des souris mutées soit uniquement dans le cerveau, soit dans toutes leurs cellules, est identique. Guy et al. ont, par ailleurs, observé le développement des souris femelles hétérozygotes *Mecp2*^{+/-} (modèle 3'). Celles-ci ne présentent aucun symptôme pendant plusieurs mois et donnent des portées saines. A partir de trois mois, certaines montrent une inertie et une contracture des pattes arrière. A l'âge de 9 mois, près de la moitié des souris ont des symptômes évidents, incluant souvent une respiration irrégulière. Mais certaines souris peuvent rester asymptomatiques pendant une année.

Le tableau XVII résume les modèles mutins obtenus (génotypes et phénotypes) par les deux équipes.

Tableau XVII : Les différents modèles murins du syndrome de Rett

		génotype	phénotype
Chen et al. 2001	Modèle 1	<i>Mecp2</i> ^{-/y}	Normal pendant les 1ères semaines Fertile
	Modèle 2	<i>Mecp2</i> ^{-/y} - <i>Nestin-cre</i>	Comportement anormal : -nervosité -tremblement -troubles respiratoires -extrémités froides -petite taille du cerveau Mort à 10 semaines
	Modèle 3	<i>Mecp2</i> ^{-/y} - <i>CamK-cre93</i>	Normal jusqu'à 3 mois Symptômes : -gain de poids -marche ataxique -activité réduite
	Modèle 4	<i>Mecp2</i> ^{-/+}	Normal jusqu'à 4 mois Symptômes : -gain de poids -activité réduite -marche ataxique
Guy et al. 2001	Modèle 1'	<i>Mecp2</i> ^{-/y} , <i>Mecp2</i> ^{-/-}	Normal jusqu'à 3-8 semaines Symptômes : -marche raide, non coordonnée -mouvements réduits
	Modèle 2'	<i>Mecp2</i> ^{-/y} - <i>Nestin-cre</i>	-troubles respiratoires -mauvaise dentition -contracture des membres arrière -pas de défaut sensoriel -pour les mâles : testicules internes Mort à 54 jours
	Modèle 3'	<i>Mecp2</i> ^{-/+}	Développement initial normal possible jusqu'à 12 mois Portées saines A 3 mois : inertie, contracture des membres arrière A 9 mois : troubles respiratoires

Plusieurs conclusions peuvent être tirées de ces deux articles. Ces deux études, similaires dans leurs résultats, démontrent que contrairement à ce qui avait été précédemment publié le gène *MECP2* n'est pas indispensable au développement embryonnaire murin. De plus, les phénotypes murins sont relativement superposables au phénotype du syndrome de Rett,

attestant de la validité de ces modèles murins pour étudier la pathogenèse de la maladie humaine.

Par ailleurs, les symptômes observés chez les souris homozygotes mutées sont probablement dus à l'absence de la protéine MeCP2 dans le cerveau, étant donné que les mêmes symptômes sont retrouvés, et ce, que la mutation ait lieu dans tous les tissus ou spécifiquement dans le cerveau. Cette atteinte caractéristique du cerveau atteste peut-être d'une compensation de la protéine MeCP2 mutée par un autre répresseur de la famille MBD dans les autres tissus. Guy et al. ont testé l'hypothèse d'un rôle de la protéine MBD2, composant du complexe de répression transcriptionnelle MeCP1. Les double-mutants *Mecp2*^{-/-} et *Mbd2*^{-/-} ont exactement le même phénotype que les souris *Mecp2*^{-/-}. Il ne semble donc pas exister d'interaction fonctionnelle entre ces deux protéines.

Enfin, l'équipe de Jaenisch (Chen et al. 2001) a montré que l'absence d'expression du gène *MECP2* n'a pas d'effet sur le **développement cérébral** puisque les souris dont la mutation n'est exprimée qu'après la formation du cerveau présentent les mêmes symptômes que les autres souris. Ceci est confirmé par l'observation suivante : l'apparition des premiers symptômes est indépendante du programme développemental de la souris et donc certainement de l'Homme. En effet, l'apparition des symptômes a lieu à environ 6 mois chez la souris hétérozygote comme chez l'Homme. L'échelle de temps étant tout à fait différente entre les deux espèces, cette correspondance dans le temps réel et non dans le temps intrinsèque de chaque espèce, semble indiquer que la déficience en protéine MeCP2 agit après le développement cérébral au sein des neurones postmitotiques. L'atteinte est cellule-autonome car l'autopsie des souris mutées uniquement au niveau des neurones postmitotiques (modèle 3, Chen et al. 2001) a révélé que seules les cellules portant la mutation étaient touchées ; les cellules du cervelet non porteuses de la mutation sont indemnes. Ce trouble cellule-autonome suggère un possible effet de seuil au-dessous duquel la maladie ne se déclenche pas. Autrement dit, chez les filles atteintes d'un syndrome de Rett, les cellules exprimant la mutation étant (en absence de tout biais d'inactivation) deux fois moins nombreuses que chez les garçons mutés, il faudra plus de temps pour atteindre le niveau de dysfonctionnement cérébral entraînant l'apparition des premiers symptômes. Si ce raisonnement est correct, des stratégies thérapeutiques pourraient être envisagées dans le but d'empêcher après la naissance chez les filles, le développement des neurones déficients en MeCP2 plutôt que de corriger l'insuffisance neuronale inscrite dans le génome.

En conclusion, l'atteinte cerveau-spécifique due à une déficience de la protéine MeCP2 mise en évidence aussi dans les modèles murins est paradoxale étant donné l'expression ubiquitaire du gène *MECP2* et son rôle de répresseur global de la transcription. L'identification des gènes cibles de la protéine MeCP2 constituera une étape déterminante dans la compréhension de la pathogenèse du syndrome de Rett.

4. Recherche des gènes cibles de la protéine MeCP2

Colantuoni et al. (Colantuoni et al. 2001) ont étudié l'expression de nombreux gènes du cerveau chez des patientes atteintes d'un syndrome de Rett. Pour cela, ils ont extrait les ARN totaux et les protéines de la partie du cortex cérébral correspondant aux aires 1 à 5 de Brodmann de six patientes atteintes. A partir de l'ARN, les auteurs ont utilisé quatre technologies indépendantes de puces à ADNc ainsi que l'hybridation soustractive et la RT-PCR. De plus, l'expression protéique de quelques gènes a été testée par immunoblot. Par ces approches combinées, les auteurs ont relevé une expression différentielle de plusieurs gènes chez les patientes par rapport aux contrôles (un contrôle de même âge pour chaque patiente). En effet, le taux de transcription de plusieurs gènes impliqués dans la structure et/ou la maturation des neurones est diminué (par exemple: neuronal pentraxine 1, calcineurine, CAM kinase IIb). En plus de cette diminution générale, les gènes codant des marqueurs présynaptiques (comme l'annexine VI, la syntaxine 1A) sont aussi fortement sous-exprimés dans le cerveau des patientes Rett, expliquant probablement le nombre réduit de synapses identifiables. Ceci est en accord avec les descriptions neuropathologiques soulignant une diminution du volume neuronal, de l'arborisation dendritique et du nombre de synapses et d'épines dendritiques dans le cerveau des patientes. En parallèle à cette réduction d'expression de gènes neuronaux, les auteurs ont mis en évidence une augmentation d'expression de plusieurs gènes codant des protéines spécifiques de la glie.

Bien que ces résultats obtenus témoignent de l'état d'expression de certains gènes à un moment donné, différent de celui de l'apparition de la maladie, ils constituent une caractérisation moléculaire de la neuropathologie du syndrome de Rett, et initient des pistes quant à l'identification de gènes cibles du gène *MECP2*.

CONCLUSION

La découverte de la responsabilité du gène *MECP2* dans le syndrome de Rett mit fin à une longue et laborieuse quête génétique, et a confirmé l'émergence de maladies à composante épigénétique, dans lesquelles des systèmes de régulation de la transcription sont défectueux. En plus du syndrome de Rett, entrent également dans cette catégorie le syndrome ICF avec l'implication du gène *DNMT3B* codant une ADN méthyltransférase responsable de la méthylation *de novo* du génome, et le syndrome *ATRX* dans lequel est en cause la protéine XNP impliquée dans le remodelage de la chromatine.

Par ce travail, nous avons contribué à l'établissement du spectre mutationnel du gène *MECP2* dont la connaissance est essentielle pour confirmer d'un point de vue moléculaire le diagnostic clinique de cette maladie. Nous avons, ainsi, confirmé l'existence de points chauds de mutations (représentés par huit mutations ponctuelles et par de petites délétions dans la région 3' de la séquence codante du gène *MECP2*) et découvert grâce à des approches complémentaires des mécanismes mutationnels non décrits jusqu'alors dans le gène *MECP2*. En effet, la technique de Southern blot a identifié des remaniements sur plusieurs centaines de paires de bases (jusqu'à 1,2 kb) et la PCR fluorescente semi-quantitative, des mutations en mosaïque somatique. L'utilisation du Southern blot étant assez lourde, une approche par PCR multiplex comparative de courts fragments fluorescents est en cours de mise au point. Cette méthode est basée sur l'amplification simultanée de plusieurs courtes séquences exoniques (<300 pb) réparties tout le long du gène. Une telle approche est déjà utilisée pour plusieurs autres gènes comme le gène *BRCA1* dans les formes héréditaires de cancer du sein et de l'ovaire (Robinson et al. 2000), ou le gène *MSH2* dans le cancer du colon non polyposique (Wang et al. 2002).

Si l'ensemble des mutations dans le gène *MECP2* paraît quasiment connu, notamment pour les substitutions, ceci ne correspond qu'à une étape supplémentaire dans la compréhension de la pathogenèse du syndrome de Rett. Plusieurs questions restent en suspens : i) comment expliquer les 15 à 20% de patientes ayant un syndrome de Rett typique, mais ne montrant pas de mutations dans la séquence codante du gène *MECP2* ? ii) existe-t-il un second gène impliqué dans le syndrome de Rett ? iii) comment la protéine MeCP2, répresseur transcriptionnel ubiquitaire, agit-elle pour entraîner une atteinte aussi spécifique du système nerveux central ? iv) quels sont ses gènes cibles ?

Afin d'apporter des éléments de réponse aux deux premières questions, nous poursuivons l'étude d'une translocation familiale t(X;22) associée à un syndrome de Rett. Le point de cassure sur le chromosome X correspond peut-être au second gène impliqué dans le syndrome

de Rett.

Quant à l'action de la protéine MeCP2 dans le fonctionnement cérébral, les modèles murins récemment établis apportent un premier élément de réponse. Il semble que la perte d'expression du gène *MECP2* retentit essentiellement sur la physiologie des neurones postmitotiques, donc après le développement du cerveau, et que le trouble touche les cellules indépendamment les unes des autres, ce qui expliquerait le ralentissement de la croissance céphalique chez les filles par atteinte sélective des neurones dont le chromosome X muté est actif. Par ailleurs, les mutants murins (*Mecp2*^{-ly} et *Mecp2*^{-ly} *Nestin-cre*) montrent que les conséquences de l'inactivation du gène *MECP2* dans toutes les cellules de la souris ou uniquement dans le cerveau sont les mêmes : la protéine MeCP2 aurait une activité spécifique non compensée dans le cerveau par d'autres protéines comme cela serait le cas dans les autres organes du corps. Tout ceci reste à démontrer. Un deuxième élément de réponse provient de deux articles récemment publiés (LaSalle et al. 2001 ; et Shabazian et al. 2002). Dans ces deux articles, les auteurs ont étudié le profil d'expression de la protéine MeCP2 *in situ* dans le cerveau et dans d'autres organes. Ces deux articles montrent que, malgré une transcription ubiquitaire du gène *MECP2*, la quantité de protéine synthétisée n'est pas équivalente dans tous les tissus : la protéine MeCP2 est beaucoup plus produite dans le cerveau (2 à 19 fois plus que dans les autres organes). Toutefois, son expression dans les autres tissus et notamment les poumons, suggère que l'atteinte neuronale spécifique du syndrome de Rett ne dépend pas uniquement de la quantité de la protéine synthétisée. A partir de cerveaux de patientes atteintes d'un syndrome de Rett (portant la mutation R168X), LaSalle et al. (LaSalle et al. 2001) ont analysé l'expression de la protéine MeCP2 tronquée. La protéine anormale est exprimée dans quasiment la moitié des cellules du cerveau et du cervelet et ce en relation avec une inactivation aléatoire. Ces résultats indiquent que la mutation dans le gène *MECP2* n'affecterait pas la migration normale des cellules au cours du développement du système nerveux central, confirmant l'hypothèse que l'effet des mutations dans le gène *MECP2* se manifesterait en période postnatale, après le développement du cerveau. Shabazian et al. (Shabazian et al. 2002) ont suivi l'expression de la protéine MeCP2 tout au long de l'embryogenèse de la souris et de l'homme. Ces auteurs montrent que la protéine MeCP2 est plus abondante dans les neurones matures et que son expression est corrélée avec l'ontogenèse du système nerveux central. La corde spinale et le tronc cérébral se développe en premier, avant les hémisphères cérébraux. La protéine MeCP2 est exprimée en premier lieu dans la corde spinale et dans le tronc cérébral, puis plus tardivement dans le cortex cérébral et l'hippocampe. Dans le cortex cérébral, l'expression de la protéine MeCP2 suit la maturation

des neurones. Le pourcentage de neurones exprimant la protéine MeCP2 ne cesse d'augmenter depuis la naissance jusqu'à l'âge de 10 ans. Ainsi, il se pourrait qu'en absence de protéine MeCP2 fonctionnelle, les neurones des patientes atteintes d'un syndrome de Rett, ne puissent atteindre leur état de maturation complet.

Enfin, un dernier élément de réponse quant au rôle de la protéine MeCP2 dans le fonctionnement cérébral, émane de l'utilisation des puces d'expression. Une première étude met en évidence une dérégulation importante de plusieurs gènes et notamment de ceux qui codent des protéines associées à des fonctions synaptiques (Colantuoni et al. 2001).

Tout ceci constitue les prémices des prochaines étapes dans notre compréhension de la pathogenèse du syndrome de Rett avec en vue le développement d'approches thérapeutiques.

**MATERIEL
et METHODES**

A. Patients

Le plus souvent, nous avons travaillé à partir de deux types de prélèvements sanguins, un sur EDTA et l'autre sur héparine-lithium. Chaque prélèvement est accompagné d'un consentement libre et éclairé signé par les parents.

L'ADN a été extrait à partir du prélèvement sanguin sur EDTA avec le kit Nucleon BACC3 (Amersham). Les lignées lymphoblastoïdes continues sont établies à partir du prélèvement sanguin sur héparine-lithium. Les lymphocytes B sont transformés en cellules lymphoblastoïdes sous l'action du virus Epstein-Barr (EBV).

Toutes les extractions d'ADN génomique et les établissements de lignées continues ont été réalisés par Claudine Hainaut, Michaël Brice et Florence Rousselet (Banque de Génomes du CHU de Nancy).

B. Méthodes

I. Techniques de biologie moléculaire

1. Extraction d'ADN génomique total

1.1. A partir de sang

L'extraction d'ADN à partir de leucocytes est réalisée selon le protocole du kit Nucleon BACC 3 (Amersham). Brièvement, une résine cationique colorée permet de fixer l'ADN de façon non covalente après une lyse des cellules et une déprotéinisation. Après une précipitation par l'éthanol et une resuspension dans du tampon TE 10:1 pH 8 (Tris-HCl 10 mM pH 8, EDTA 1 mM pH 8), l'ADN est dosé par spectrophotométrie en mesurant sa densité optique (DO) à 260 nm. La pureté des ADN est vérifiée par le dosage des protéines à 280 nm, le rapport $DO_{260\text{ nm}}/DO_{280\text{ nm}}$ doit être compris entre 1,7 et 2. S'il est inférieur à 1,7, il existe une contamination protéique importante nécessitant une seconde extraction. Un rapport supérieur à 2 témoigne de la présence d'une quantité importante d'ARN qui peuvent être éliminées par un traitement à la ribonucléase A (0,1 mg/ml, 15 min à 37°C). Les ADN sont conservés à court terme à 4°C et stockés à long terme à -20 °C.

1.2. A partir de cellules buccales

Les cellules buccales sont récupérées en frottant l'intérieur des joues à l'aide d'une dizaine de cotons-tiges. Les cotons-tiges sont transportés dans 5 ml de tampon stérilisé (NaCl 100 mM, Tris-HCl 10 mM pH 8, EDTA 25 mM pH 8, SDS 0,5%) à température ambiante. Après un traitement par la protéinase K (10 mg/ml) pendant 2-3 heures à 65°C (ou une nuit à 37°C), l'ADN est extrait par 1 ml de guanidine hydrochloride (6M) et 0,5 ml d'acétate d'ammonium (7,5M). Afin d'éliminer les cotons-tiges sans perdre de l'ADN, ceux-ci sont centrifugés (1000 rpm, 5 min) dans une seringue placée dans un tube de 15 ml. L'ADN peut être purifié, après l'ajout d'un volume de chloroforme froid, à l'aide de la résine du kit Nucleon (300 µl, centrifugation à 2500 rpm pendant 5 min). Le surnageant est récupéré, l'ADN précipité par deux volumes d'éthanol absolu froid, centrifugé, lavé par de l'éthanol 70% et enfin resuspendu dans du tampon TE 10:1.

1.3. A partir de cellules de bulbes de cheveux (Walsh et al. 1991)

Un cm de cheveu **avec son bulbe** sont mélangés avec 200 µl de solution de Chelex 100 (résine chélatante à haute affinité pour les ions polyvalents. BioRad). L'ensemble est mis à bouillir pendant 8 min. 40 µl de cette solution sont utilisés directement pour faire une PCR dans un volume final de 100 µl.

2. Amplification *in vitro* de séquences d'ADN par PCR (Mullis et al. 1986)

Nous avons utilisé les conditions de PCR suivantes, dans un volume final de 50 µl :

- 100 ng d'ADN génomique total,
- tampon de réaction 1X (spécifique de l'enzyme utilisée),
- une amorce sens et une amorce anti-sens (1 µM chacune),
- dNTP (250 µM),
- MgCl₂ (1 ou 1,5 mM),
- et 2 U d'enzyme Taq polymérase.

NB : Toutes les concentrations données sont les concentrations finales.

La température de fusion (ou T_m) spécifique de l'hybridation d'une amorce avec la région complémentaire de l'ADN génomique peut être calculée selon la formule simplifiée : $T_m = 2(A+T) + 4(G+C) - 5$ où A, T, G et C symbolisent les sommes respectives des bases

Adénine, Thymine, Guanine et Cytosine présentes dans la séquence de l'amorce. Pour chacune des deux amorces, la T_m est calculée et une température moyenne est ainsi définie théoriquement pour la réaction d'amplification. Plusieurs essais à différentes températures sont parfois nécessaires avant de déterminer la bonne température.

La qualité et la quantité du produit de PCR sont vérifiées par électrophorèse sur un gel d'agarose à 0,8% dissous dans du tampon TBE 1X (10X : Tris borate 0,9 mM, EDTA 20 mM, pH 8,2) et mélangé à 0,5 mg/ml de bromure d'éthidium (Br Et : agent intercalant de l'ADN double brin, visualisant celui-ci par les U.V). La taille du produit de PCR est estimée par comparaison avec un marqueur de taille 100 pb (Marqueur XIV, Roche).

2.1. Les conditions d'amplification de l'ADN pour la CSGE

Après une dénaturation de l'ADN génomique total pendant 5 min à 95°C, 35 cycles comportant chacun, une étape de dénaturation à 95°C, une étape d'hybridation des amorces selon une température spécifique (T_m), et une étape d'élongation à 72°C, sont réalisés. La durée de chacune de ces étapes a été fixée à 1 min. Suit une extension finale à 72°C pendant 10 min pour terminer la synthèse des brins néosynthétisés.

Tableau XVIII : Séquences des neuf couples d'amorces et paramètres d'amplification permettant le criblage du gène *MECP2* par CSGE

séquence (5'-3')	température d'hybridation	Concentration en MgCl ₂	taille (pb)
1' R : caatgggggctttcaactta 1' F : aaaacagatggccaaaccag	55	1 mM	218
2A R : ggggtcatcatacatgggtc 2A F : cctgcctctgctcactgtt	59	1 mM	339
2B R : gtcccccgacccccacct 2B F : agcccgtgcagccatcagcc	70	1 mM	349
3A R : ctcccaggacttttctcca 3A F : tttgtcagagcgttgtcacc	59	1 mM	379
3B R : ctgcacagatcggatagaagac 3B F : aaccacctaagaagcccaaa	59	1 mM	379
3C R : tgagtgggtgatgggtgg 3C F : ggcaggaagcgaagctgag	66	1 mM	365
3D R : ctccctcccctcggtgtttg 3D F : tggtaagcccctgctggt	59	1 mM	413
3E' R : ccaactactcccacctgaagc 3E F : ggagaagatgccagaggag	66	1 mM	385
polyA10R* : tagccaaatgactgaaagcac polyA10 F* : gctgagtattcttgccttcac	53	1 mM	321

2.2. Les conditions d'amplification de l'ADN pour la DHPLC

Les fragments 2, 3, 4a, et 4cd sont amplifiés par la technique de touch-down PCR. Quelques cycles sont réalisés à une température d'hybridation des amorces haute afin d'éviter des amplifications aspécifiques, puis celle-ci est diminuée progressivement afin de permettre l'amplification.

Programme utilisé : 95°C pendant 5 min

10 cycles: 95°C pendant 30 sec

65°C pendant 30 sec, avec une diminution de 1,5°C par cycle

72°C pendant 1 min 45 sec

28 cycles: 95°C pendant 30 sec

51°C pendant 30 sec

72°C pendant 1 min 30 sec

72°C pendant 10 min

Tableau XIX : Les six couples d'amorces utilisés pour le criblage du gène *MECP2* par DHPLC

séquence (5'-3')	Concentration en MgCl ₂	taille (pb)
2-For: taagctgggaaatagcctagtagtac 2-Rev: ttatatggcacagtttggcacag	1 mM	427
3-For: aggacatcaagatctgagtgtat 3-Rev: ggtcatttcaagcacacctg	1,5 mM	625
4a-For: cgctctgccctatctctga 4a-Rev: acagatcggatagaagactcctt	1,5 mM	583
3C R: tgagtgggtggtgatgggtgg 3C F: ggcaggaagcgaaaagctgag	1 mM	365
4cd-For: ggaaaggactgaagacctgtaag 4cd-Rev: ctccctcccctcggtgtttg	1 mM	371
3E F: ggagaagatgccagaggag 3E' R: ccaactactcccaccctgaagc	1 mM	385

L'ensemble des amorces (tableaux XVIII et XIX) a été décrite par Amir et al. (1999), Wan et al. (1999) pour l'approche par CSGE et par Buyse et al. (2000) pour la DHPLC ; excepté les amorces marquées d'un astérisque.

Le tableau XX indique les noms et séquences des amorces nécessaires à la synthèse de la sonde utilisée pour la recherche de grands réarrangements par Southern blot et pour déterminer les bornes des remaniements identifiés par Southern blot.

Tableau XX :

nom	séquence (5'-3')
2F	gaagaaaagtcagaagaccag
3R	gtgtaaagtcagctaactctc
MEC1	tgtagggctcaggaagaaaag

3. Technique de CSGE

Nous avons utilisé le protocole de Philippe et al. (Philippe et al.1997) adapté de Ganguly et al. (Ganguly et al. 1993).

Après une dénaturation de 5 min à 98°C et une réassociation à 68°C pendant une heure favorisant la formation d'hétéroduplex, 3 à 10 µl de produits de PCR de 300-400 pb sont déposés sur un gel semi-dénaturant (polyacrylamide 10%, formamide 15%, éthylène glycol 10%, tampon TTE 0,5X (20X: Tris 89 mM, Taurine 7 mM, EDTA 0,1 mM, pH 8,8-9,0), persulfate d'ammonium 1% et TEMED 0,055%), 0,8 mm d'épaisseur, et 40 cm de hauteur sur 21cm de largeur.

L'électrophorèse est réalisée avec du tampon TTE 0,25X dans le réservoir supérieur et du tampon TTE 1X dans le réservoir inférieur. Une pré-électrophorèse d'une heure à 500 V est effectuée avant le dépôt des échantillons. Puis, une fois déposés, ceux-ci migrent au moins 16 heures à 500 V. La migration terminée, le gel est en partie démoulé (seule la plaque de verre liée au réservoir est enlevée), puis coloré par du bromure d'éthidium (0,5 mg/ml) pendant 20 min, et rincé à l'eau. La région d'intérêt contenant les fragments d'ADN est repérée à l'aide d'une lampe U.V. à mains, découpée et transférée grâce à du papier Wattman 3M sur une table U.V. reliée à une caméra permettant une analyse plus fine du gel.

4. Technique de DHPLC

Afin de déterminer les températures et les gradients d'acétonitrile d'élution pour analyser le gène *MECP2* par DHPLC, nous avons utilisé les 27 mutations déjà connues au laboratoire ainsi que les conditions publiées par Buyse et al. (Buyse et al. 2000). Les conditions mises au point sont résumées dans le tableau XXI.

Tableau XXI : Conditions de criblage du gène *MECP2* par DHPLC.

Fragment	Température	Gradient d'acétonitrile
2	59°C	54%
3	59°C	61%
	65°C	53%
4a	61°C	60%
	63°C	56%
	66°C	52%
3C	63,5°C	57%
4cd	64,5°C	57%
3E	65°C	54%

5. Le séquençage

Les produits de PCR sont purifiés soit directement, soit après extraction de la bande d'intérêt depuis un gel d'agarose, à l'aide de kits (QIAquick Gel extraction de Qiagen ou Nucleospin de Macherey-Nagel) afin d'éliminer principalement les amorces, les dNTP, et les produits aspécifiques lorsque la bande est extraite du gel. Les produits ainsi purifiés sont dosés sur un gel d'agarose 0,8% par comparaison de l'intensité de la bande avec celle d'un marqueur de poids moléculaire 100 pb (Marqueur XIV, Roche) dont on connaît pour chaque bande la quantité d'ADN.

La méthode de séquençage utilisée est la suivante. Le fragment d'ADN est amplifié dans un sens par PCR avec une incorporation des quatre didésoxynucléotides triphosphates (ddNTP) marqués chacun par un fluorochrome différent, ainsi la réaction de séquençage est effectuée dans un seul tube. La réaction de séquençage est réalisée selon le protocole du kit Big Dye Terminator contenant les 4 ddNTP, le tampon de réaction, les dNTP et l'enzyme (Applied Biosystems).

Dans un volume réactionnel de 20 µl, sont nécessaires :

- 20 à 30 ng de produit de PCR purifié (quantité pour des fragments de 300-400 pb),
- 4 µl du mélange du kit Big Dye Terminator,
- 2 µl de tampon (5X : 400 mM Tris-HCl, 10 mM MgCl₂, pH 9,0),
- 1 µl d'amorce à 25 ng / µl.

Les paramètres du programme d'amplification sont les suivants : 10 sec à 96°C, 5 sec à 50°C, et 4 min à 60°C et ceci pendant 25 cycles.

Les produits de PCR sont alors purifiés sur des colonnes DyEx (Qiagen) afin d'éliminer les ddNTP fluorescents non incorporés. Les éluats séchés sont resuspendus dans 40 µl de tampon TSR (Template Suppression Reagent. Applied Biosystems). Les échantillons sont analysés sur un séquenceur automatique ABI 310 (Applied Biosystems).

Chaque échantillon est séquencé dans les deux sens, et analysé visuellement par comparaison avec une séquence contrôle normale (séquence du gène *MECP2*, n° d'accension : AF030876).

6. Analyse de fragments par PCR fluorescente semi-quantitative

La PCR fluorescente est réalisée dans les mêmes conditions que la PCR décrite dans le paragraphe 2.1., excepté qu'une des deux amorces, en l'occurrence ici l'amorce anti-sens est marquée par un fluorochrome, le 6-FAM (6-carboxy-fluorescein), à son extrémité 5'.

Le produit de PCR est contrôlé sur gel d'agarose 0,8%. Les échantillons sont dilués au 1/5 ou au 1/10 selon l'intensité de la bande visualisée sur le gel d'agarose, puis 1 µl de la dilution est mélangé avec 12 µl d'une solution de formamide/dextran sulfate et 0,5 µl d'un standard interne de taille Rox HD400 (Applied biosystems) et dénaturés 2 min à 95°C avant d'être chargés sur le séquenceur automatique ABI 310. L'analyse des tailles et des aires des pics sont générées par les logiciels Genescan et Genotyper (Perkin-Elmer).

7. La technique de Southern blot (Southern, 1975)

10 µg d'ADN génomique total sont digérés par 20 U d'enzymes de restriction (les enzymes de restriction les plus utilisées ont été : BamHI, BglII, et HindIII) dans un volume final de 50 µl en présence de 5 µl de tampon 10X spécifique de l'enzyme et de 0,5 µl de BSA 100X (albumine de sérum bovin) nécessaire à certaines enzymes (BamHI).

Le mélange est incubé à une température spécifique (37°C, pour les 3 enzymes pré-citées) pendant une nuit.

5 µl de l'ADN ainsi digéré sont d'abord déposés sur un gel d'agarose 0,8% afin de vérifier la bonne digestion de l'ADN. Ensuite, si la digestion a été efficace, le reste du produit de digestion est déposé avec un marqueur de taille (phage lambda digéré par HindIII) sur un gel 0,8%/TAE 1X (50X : Tris-acétate 2M, EDTA 50 mM, pH 8,0)/Br Et. L'électrophorèse est réalisée sur la nuit sous une tension de 40 V dans du tampon TAE 1X.

Le lendemain, la séparation des fragments d'ADN digéré génomique est analysée par observation du gel sur une table U.V.. Les distances de migration des différents fragments du marqueur de taille sont mesurées (en cm) par rapport à la position des puits. La partie du gel contenant le marqueur de taille est éliminée afin d'éviter des hybridations non spécifiques. L'ADN est ensuite transféré et fixé sous forme simple-brin par action de la soude 0,4 N sur une membrane de Nylon Hybond N+ (Amersham) chargée positivement par un phénomène de capillarité. Le transfert s'effectue pendant une nuit (4 heures au minimum).

Après un rinçage rapide dans du SSC 2X (solution saline citrate, SSC 20X : NaCl 3M, Citrate de sodium 0,3 M, pH7), la membrane de nylon est placée dans un tube à hybridation en présence de 15 ml de tampon d'hybridation (NaCl 750 mM, héparine 50 µg/ml, SDS 1%, dextran sulfate 5%, ADN de sperme de hareng 50 µg/ml) et laissée pré-hybrider pendant au moins 2 heures à 65°C sous agitation afin d'éviter des hybridations non spécifiques par saturation de la membrane par l'ADN de sperme de hareng.

La sonde est marquée par la technique de random priming (Random Primed DNA labeling kit, Roche) en incorporant de l' $\alpha^{32}\text{P}$ dCTP (2.10^6 cpm/ml). Après purification sur une colonne de Séphadex® G50 (Nick™ column, Pharmacia Biotech), la sonde est dénaturée 10 min dans de l'eau bouillante et transférée dans le tube à hybridation contenant la membrane. L'hybridation est réalisée pendant une nuit à 65°C.

Le lendemain, la membrane est lavée plusieurs fois dans des conditions de plus en plus stringentes à 65 °C (SSC 2X pendant 15 min, SSC 2X/SDS 0,1% pendant 30 min, et SSC 0,1X/SDS 0,1% pendant 10 min). Tous ces lavages limitent le bruit de fond dû à des hybridations aspécifiques.

Un film d'autoradiographie (Biomax, Kodak) est exposé à la membrane pendant 2 à 4 jours dans une cassette à -80°C avant d'être développé par une développeuse automatique.

8. Obtention de l'ADNc du gène *MECP2*

8.1. Extraction d'ARN totaux et transcription inverse (RT)

Les ARN totaux sont extraits à l'aide de la solution TRIZOL® (TRIZOL® Reagent, Gibco-Brl). La transcription inverse est réalisée en deux étapes à partir d'1 µg d'ARN totaux extrait de leucocytes par le kit Reverse-transcriptase Superscript II (Gibco-Brl) en utilisant 500 ng d'hexanucléotides selon le protocole du fabricant.

8.2. PCR

A partir des ADNc totaux générés, l'ADNc complet du gène *MECP2* a été obtenu par deux amplifications spécifiques avec les couples d'amorces suivants : MEC1/3BR, 3CF/3ER (tableaux XVII et XIX). Dans un volume final de 50 µl, 3 µl d'ADNc, 1 µM d'amorces, 1,5 µM de MgCl₂ à 50 mM, 250 µM de dNTP, tampon 1X et 2 U de Taq polymérase sont utilisés. Après une dénaturation de 5 min à 95°C, 35 cycles sont réalisés comprenant chacun : 1 min à 95°C, 1 min au T_m et 1 min à 72°C, suivis de 10 min à 72°C. La T_m est de 62°C pour le couple MEC1/3BR, et de 56°C pour le couple 3CF/3E'R.

9. Inactivation du chromosome X

L'inactivation du chromosome X a été étudiée au locus HUMARA en Xq13 (Allen et al. 1992). 1 µg d'ADN génomique total est digéré par 2 µl d'enzyme HpaII (10U/ µl) pendant 2 heures dans un volume réactionnel de 50 µl. Au bout de 2 heures, 2 µl d'enzyme sont additionnés pour une durée au moins équivalente. L'enzyme HpaII est sensible à la méthylation, c'est-à-dire qu'elle ne reconnaît son site de restriction et ne digère l'ADN que si la séquence de son site n'est pas méthylée. Une fois la digestion terminée, l'enzyme est inactivée 5 min à 95°C et le produit est purifié sur colonne (kit Wizard® DNA Clean-Up system, Amersham).

2 µl d'ADN génomique digéré purifié et 0,5 µl d'ADN génomique non digéré sont amplifiés en parallèle par PCR fluorescente à l'aide d'amorces spécifiques du locus HUMARA, dans les conditions suivantes: tampon de réaction 1X, MgCl₂ 1,5 mM, dNTP 200 µM, amorces 0,2 µM (chacune) et 2,5 U de Taq polymérase. Après une dénaturation de 5 min à 95°C, 35 cycles

sont effectués comprenant chacun : 30 sec à 95°C, 30 sec à 59°C et 45 sec à 72°C, suivis d'une élongation à 72°C. Les qualité et quantité des produits d'amplification sont vérifiées sur un gel d'agarose de 2%. Des bandes de 250-300 pb sont attendues. Selon leur intensité, les produits de PCR sont plus ou moins dilués avant d'être analysés sur le séquenceur ABI 310 par les logiciels Genescan et Genotyper.

II. Technique de cytogénétique

1. Préparation des chromosomes

Les chromosomes métaphasiques sont obtenus à partir des lignées lymphoblastoïdes continues. Les cellules sont bloquées en métaphase pendant 50 min à 37°C par l'action de la colchicine (1 µg/ml, agent anti-mitotique bloquant la formation du fuseau achromatique). La membrane cellulaire est distendue par turgescence sous l'effet du KCl (75 mM) pendant 15 min à 37°C, et les chromosomes dispersés de façon optimum. Après une préfixation et plusieurs lavages par du fixateur (3 volumes d'éthanol / 1 volume d'acide acétique) dans le but d'éliminer au maximum le cytoplasme, les cellules sont étalées sur des lames dégraissées Superfrost.

2. Hybridation *in situ* en fluorescence

2.1. Prétraitement des chromosomes

Dans un premier temps, les lames sont observées en contraste de phase afin de vérifier la qualité des métaphases obtenues. Ensuite, elles sont traitées par une protéase (pepsine à 100 µg/ml) pendant 5 min à 37°C, dans le but de débarrasser les chromosomes des protéines nucléaires. Après deux lavages de 5 min à température ambiante dans une solution de PBS 1X (NaCl 170 mM, KCl 3 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1,8 mM, pH 7,3), les chromosomes sont traités par une solution de formaldéhyde 1% pendant 2 min. Suit de nouveau un lavage dans du PBS 1X, avant une déshydratation dans trois bains successifs d'éthanol de concentrations croissantes (70%, 85% et 100%), 1 min chacun. Les lames sont alors prêtes à l'emploi, si elles ne sont pas utilisées immédiatement, elles peuvent être conservées à -20°C.

2.2. Obtention de la sonde

La sonde utilisée est le PAC 671D9 (Reichwald et al. 2000) qui contient un insert d'ADN génomique humain d'environ 100 kb comprenant le gène *MECP2* dans sa totalité. Ce PAC nous a été fourni par l'équipe du Dr Matthias Platzer de l'Institut de Biotechnologie moléculaire de Iena (Allemagne).

1 µg de ce PAC est marqué par incorporation de 16dUTP-biotine par Nick translation selon le protocole du kit Bio-Nick (Gibco-brl).

Une fois la qualité du marquage vérifiée sur un gel d'agarose 0,8% par obtention d'un "smear" compris entre 500 et 2000 pb, une première précipitation à l'éthanol est réalisée afin d'éliminer les nucléotides marqués non incorporés. Le culot est resuspendu dans du tampon TE 10:1, puis mélangé avec une quantité 50 fois supérieure d'ADN cot 1 (Gibco-brl). Cet ADN cot 1 correspond à un condensé de séquences répétées du génome humain. Lorsqu'il est mélangé à une sonde, il bloque les séquences répétées contenues dans cette sonde, empêchant ainsi la sonde de se fixer aux séquences répétées lorsqu'elle est hybridée sur les chromosomes cibles.

L'ensemble (sonde et ADN cot 1) est précipité de nouveau à l'éthanol afin de le resuspendre dans un petit volume (13 µl) de tampon d'hybridation (formamide désionisée 50%, sulfate de dextran 10%, SSC 2X, Tween 20 1%, pH 7).

1 µl de sonde témoin (dans notre cas, sonde complémentaire des séquences α-satellites du centromère du chromosome X) directement marquée par un fluorochrome émettant dans le bleu, peut-être ajoutée (CEP X spectrum aqua VYSIS).

La sonde ainsi préparée est prête à être utilisée, dans le cas contraire, elle peut être stockée à -20°C.

2.3. Dénaturation des chromosomes et de la sonde

Les chromosomes sont dénaturés 2 min à 74°C dans une solution de formamide 70%/SSC 2X, pH 7, et la sonde 5 min à 74°C.

2.4. Hybridation

Les 13 µl de sonde ainsi dénaturée sont déposés sur les chromosomes eux-mêmes dénaturés, recouverts d'une lamelle scellée par de la colle à rustines. L'ensemble est incubé une nuit à 37°C dans une chambre humide.

2.5. Lavages

Les lavages éliminent d'une part l'excès de sonde et d'autre part les hybridations non spécifiques. La lame est lavée 2 min dans une solution stringente (SSC 0,4X/ Nonidet P40 0,3%, pH 7-7,5) à 73°C, puis 15 sec dans une solution moins stringente (SSC 2X/ Nonidet P40 0,1%, pH 7-7,5) à température ambiante.

2.6. Détection immunocytochimique

La détection immunocytochimique est seulement nécessaire lorsque la sonde est marquée de façon indirecte par un haptène tel la biotine. Dans notre cas, nous avons utilisé un kit de détection à la fluorescéine (FITC : fluorochrome émettant dans le vert) (kit Biotine-FITC detection, Oncor-Appligène).

La lame est incubée 10 min dans une solution BN block (Bicarbonate de sodium 0,1 mM, Nonidet P40 0,025%, pH 8,0, poudre de lait 5%) afin de saturer la lame et d'éviter ainsi une fixation non spécifique des anticorps. 60 µl d'une solution d'avidine marquée à la fluorescéine (avidine-FITC) sont déposés sur la lame, recouverts d'une lamelle et l'ensemble est incubé 10 min à 37°C dans une chambre humide.

Après 3 lavages de 2 min chacun, à température ambiante dans une solution de BN (Bicarbonate de sodium 0,1 mM, Nonidet P40 0,025%, pH 8,0), une amplification du signal est effectuée par une première incubation de la lame à 37°C en présence de 60 µl d'anticorps anti-avidine biotinylés. Puis après 3 lavages de 2 min à température ambiante dans la solution

de BN, une seconde incubation de la lame à 37°C avec 60 µl d'avidine-FITC est réalisée. Finalement, la lame est lavée trois fois dans la solution de BN.

2.7. Contre-coloration et observation de la lame

Une fois séchée à l'abri de la lumière, 20 µl de DAPI (1 µg/ml, VYSIS), agent intercalant fluorescent émettant dans le bleu et mélangé avec une solution "antifade" évitant une extinction trop rapide du signal fluorescent, sont déposés sur la lame. La pénétration du DAPI n'étant pas tout à fait instantanée, une attente de 3-4 heures est nécessaire afin d'obtenir une meilleure définition des chromosomes.

La lame est observée à l'aide d'un microscope à fluorescence équipé d'un système de filtres d'émission et d'excitation spécifiques des trois fluorochromes utilisés (DAPI, fluorescéine et aqua). De plus, le microscope est relié à une caméra afin de photographier les métaphases intéressantes. L'analyse de ces images numérisées, obtenues lors de trois captures successives (une pour chaque fluorochrome) est réalisée à l'aide du logiciel d'analyse d'images Smart Capture (VYSIS).

BIBLIOGRAPHIE

279 références bibliographiques :

Adler, D., Quaderi, N., Brown, S., Chapman, V., Moore, J., Tate, P. and Disteché, C. (1995) The X-linked methylated DNA binding protein, Mecp2, is subject to X inactivation in the mouse. *Mamm Genome*, **6**, 491-492.

Akhtar, A. and Becker, P. (2000) Activation of transcription histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol Cell*, **5**, 367-375.

Allen, R., Zoghbi, H., Moseley, A., Rosenblatt, H. and Belmont, J. (1992) Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet*, **51**, 1229-1239.

Allfrey, V., Faulkner, R. and Mirsky, A. (1964) *Proc Natl Acad Sci USA*, **51**, 786-794.

Amano, K., Nomura, Y., Segawa, M. and Yamakawa, K. (2000) Mutational analysis of the MECP2 gene in Japanese patients with Rett syndrome. *J Hum Genet*, **45**, 213-236.

American, P. A. (1994) DSM-IV, Diagnostic and statistical manual of mental disorders. 4th ed. Washington DC: The American Psychiatric Association, 1994.

Amir, R., Van den Veyver, I., Wan, M., Tran, C., Francke, U. and Zoghbi, H. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*, **23**, 185-188.

Amir, R., Van den Veyver, I., Schultz, R., Malicki, D., Tran, C., Dahle, E., Philippi, A., Timar, L., Percy, A., Motil, K., Lichtarge, O., Smith, E., Glaze, D. and Zoghbi, H. (2000) Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann Neurol*, **47**, 670-679.

Antequera, F., MacLeod, D. and Bird, A. (1991) Specific protection of methylated CpGs in mammalian nuclei. *Cell*, **58**, 509-517.

Antequera, F. and Bird, A. (1994) Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci USA*, **90**, 11995-11999.

Antonarakis, S. and Group, t. N. W. (1998) Recommendations for a nomenclature system for human gene mutations. *Hum Mutat*, **11**, 1-3.

Anvret, M. and Wahlström, J. (1994) Rett syndrome: random X chromosome inactivation. *Clin Genet*, **45**, 274-275.

Archidiacono, N., Lerone, M., Rocchi, M., Anvret, M., Ozcelik, T., Francke, U. and Romeo, G. (1991) Rett syndrome: exclusion mapping following the hypothesis of germinal mosaicism for new X-linked mutations. *Hum Genet*, **86**, 604-606.

Armstrong, D., Dunn, K. and Antalffy, B. (1998) Decreased dendritic branching in frontal, motor and limbic cortex in Rett syndrome compared with trisomy 21. *J Neuropathol Exp Neurol*, **57**, 1013-1017.

Armstrong, J., Pineda, M. and Monros, E. (2000) Mutation analysis of 16S rRNA in patients with Rett syndrome. *Pediatr Neurol*, **23**, 85-87.

Auranen, M., Vanhala, R., Vosman, M., Levander, M., Varilo, T., Hietala, M., Riikonen, R., Peltonen, L. and Jarvela, I. (2001) MECP2 gene analysis in classical Rett syndrome and in patients with Rett-like features. *Neurology*, **56**, 611-617.

Austin, M., Collins, J., Corey, L., Nance, W., Neale, M., Schieken, R. and Brown, J. (1992) Aphidicolin-inductible common fragile-site expression: Results from a population survey in twins. *Am J Hum Genet*, **50**, 76-83.

Avner, P., Prissette, M., Arnaud, D., Courtier, B., Cecchi, C. and Heard, E. (1998) Molecular correlates of the murine Xce locus. *Genet Res*, **72**, 217-224.

Avner, P. and Heard, E. (2001) X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet*, **2**, 59-67.

Babinet, C. and Cohen-Tannoudji, M. (2000) Vingt ans d'interventions délibérées sur le génome de la souris. Une révolution dans l'approche génétique de la biologie des mammifères. *Med Sci*, **1**, 31-42.

Ballestar, E., Yusufzai, T. and Wolffe, A. (2000) Effects of Rett syndrome mutations of the Methyl-CpG-binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. *Biochemistry*, **39**, 7100-7106.

Ballestar, E. and Wolffe, A. P. (2001) Methyl-CpG-binding proteins. Targeting specific gene repression. *Eur J Biochem*, **268**, 1-6.

Baylin, S. and Herman, J. (2000) DNA hypermethylation in tumorigenesis. *TIG*, **16**, 168-174.

Bernards, A. and Gusella, J. (1994) The importance of genetic mosaicism in human disease. *N Engl J Med*, **331**, 1447-1449.

Bestor, T. (1998) Methylation meets acetylation. *Nature*, **393**, 311-312.

Bestor, T. (2000) The DNA methyltransferases of mammals. *Hum Mol Genet*, **9**, 2395-2402.

Bhattacharya, S., Ramchandani, S., Cervoni, N. and Szyf, M. (1999) A mammalian protein with specific demethylase activity for mCpG DNA. *Nature*, **397**, 579-583.

Bienvenu, T., Carrie, A., de Roux, N., Vinet, M., Jonveaux, P., Couvert, P., Villard, L., Arzimanoglou, A., Beldjord, C., Fontes, M., Tardieu, M. and Chelly, J. (2000) MECP2 mutations account for most cases of typical forms of Rett syndrome. *Hum Mol Genet*, **9**, 1377-1384.

Bienvenu, T., Souville, I., Poirier, K., Aquaviva, C., Burglen, L., Amiel, J., Heron, B., Kaminska, A., Couvert, P., Beldjord, C. and Chelly, J. (2001) Five novel frameshift mutations in exon 3 and 4 of the *MECP2* gene identified in Rett patients: consequences for molecular diagnosis strategy. *Hum Mutat*, **18**, 251-252.

Bird, A. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, **321**, 209-213.

Bird, A. and Wolffe, A. (1999) Methylation-induced repression-belts, braces, and chromatin. *Cell*, **99**, 451-454.

Bowles, K., Buyse, I., Honn, K., Bashford, M., Arn, P. and Roa, B. (2001) Identification of a putative mosaic *MECP2* mutation in an atypical Rett syndrome patient by denaturing high performance liquid chromatography. *Am J Hum Genet*, **69-Suppl**, 640.

a-Bourdon, V., Philippe, C., Labrune, O., Amsellem, D., Arnould, C. and Jonveaux, P. (2001) A detailed analysis of the *MECP2* gene : prevalence of recurrent mutations and gross DNA rearrangements in Rett syndrome patients. *Hum Genet*, **108**, 43-50.

b-Bourdon, V., Philippe, C., Bienvenu, T., Koenig, B., Tardieu, M., Chelly, J. and Jonveaux, P. (2001) Evidence of somatic mosaicism for a *MECP2* mutation in females with Rett syndrome: diagnostic implications. *J Med Genet*, **38**, 867-870.

c-Bourdon, V., Philippe, C., Grandemenge, A., Reichwald, K. and Jonveaux, P. (2001) Deletion screening by fluorescence in situ hybridization in Rett syndrome patients. *Ann Genet*, **44**, 191-194.

Boyes, J. and Bird, A. (1991) DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*, **64**, 1123-1134.

Brockdorff, N., Ashworth, A., Kay, G., Cooper, P., Smith, S., McCabe, V., Norris, D., Penny, G., Patel, D. and Rastan, S. (1991) Conservation of position and exclusive expression of mouse *Xist* from the X inactive chromosome. *Nature*, **351**, 329-331.

Brown, C., Ballabio, A., Rupert, J., Lafreniere, R., Grompe, M., Tonlorenzi, R. and Willard, H. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, **349**, 38-44.

Brown, C., Hendrich, B., Rupert, J., Lafreniere, R., Xing, Y., Lawrence, J. and Willard, H. (1992) The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, **71**, 527-542.

Brown, C., Carrel, L. and Willard, H. (1997) Expression of genes from the human active and inactive X chromosomes. *Am J Hum Genet*, **60**, 1333-1343.

Brown, C. and Robinson, W. (2000) The causes and consequences of random and non random X chromosome inactivation in humans. *Clin Genet*, **58**, 353-363.

Budden, S. (1986) Rett syndrome: studies of 13 affected girls. *Am J Med Genet*, **1-Suppl**, 99-109.

- Bunyan, D., Robinson, D., Collins, A., Cockwell, A., Bullman, H. and Whittaker, P. (1994) Germline and somatic mosaicism in a female carrier of Duchenne muscular dystrophy. *Hum Genet*, **93**, 541-544.
- Buschhausen, G., Wittig, B., Graessmann, M. and Graessmann, A. (1987) Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc Natl Acad Sci USA*, **84**, 1177-1181.
- Busque, L., Mio, R., Mattioli, J., Brais, E., Blais, N., Lalonde, Y., Maragh, M. and Gilliland, D. (1996) Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood*, **88**, 59-65.
- Buyse, I., Fang, P., Hoon, K., Amir, R., Zoghbi, H. and Roa, B. (2000) Diagnostic testing for Rett syndrome by DHPLC and direct sequence analysis of the MECP2 gene: identification of several novel mutations and polymorphisms. *Am J Hum Genet*, **67**, 1428-1436.
- Bühler, E., Malik, N. and Alkan, M. (1990) Another model for inheritance of Rett syndrome. *Am J Med Genet*, **36**, 126-131.
- Bühler, E., Malik, N. and Alkan, M. (1999) Rett syndrome and genetic drift. *Brain dev*, **21**, 175-178.
- Cameron, E., Bachman, K., Myöhänen, S., Herman, J. and Baylin, S. (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat Genet*, **21**, 103-107.
- Camus, P., Abbadi, N. and Gilgenkrantz, S. (1994) X inactivation in Rett syndrome: a preliminary study showing partial preferential inactivation of paternal X with the M27□ probe. *Am J Med Genet*, **50**, 307-308.
- Camus, P., Abbadi, N., Perrier, M., Chery, M. and Gilgenkrantz, S. (1996) X chromosome inactivation in 30 girls with Rett syndrome: analysis using the probe. *Hum Genet*, **97**, 247-250.
- Cardaioli, E., Dotti, M., Hayek, G., Zappella, M. and Federico, A. (1999) Studies in mitochondrial pathogenesis of Rett syndrome: Ultrastructural data from skin and muscle biopsies and mutational analysis at mtDNA nucleotides 10463 and 2835. *J Submicrosc Cytol Pathol*, **31**, 301-304.
- Cardoso, C., Timsit, S., Villard, L., Khrestchatisky, M., Fontes, M. and Colleaux, L. (1998) Specific interaction between the XNP/ATR-X gene product and SET domain of the human EZH2 protein. *Hum Mol Genet*, **7**, 679-684.
- Carrel, L. and HF., W. (1996) An assay for X inactivation based on differential methylation at the fragile X locus, FMR1. *Am J Med Genet*, **64**, 27-30.
- Carrel, L., Cottle, A., Goglin, K. and HF., W. (1999) A first generation X-inactivation profile of the human X chromosome. *Proc Natl Acad Sci USA*, **96**, 14440-14444.

- Castellvi-Bel, S. and Mila, M. (2001) Genes responsible for nonspecific mental retardation. *Mol Genet Metab*, **72**, 104-108.
- Cattanach, B. (1975) Control of chromosome inactivation. *Ann Rev Genet*, **9**, 1-18.
- Chandler, S., Guschin, D., Landsberger, N. and Wolffe, A. (1999) The methyl-CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. *Biochemistry*, **38**, 7008-7018.
- Cheadle, J., Gill, H., Fleming, N., Maynard, J., Kerr, A., Leonard, H., Krawczak, M., Cooper, D. N., Lynch, S., Thomas, N., Hughes, H., Hulten, M., Ravine, D., Sampson, J. and Clarke, A. (2000) Long-read sequence analysis of the MECP2 gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Hum Mol Genet*, **9**, 1119-1129.
- Chelly, J. (1999) Breakthroughs in molecular and cellular mechanisms underlying X-linked mental retardation. *Hum Mol Genet*, **8**, 1833-1838.
- Chelly, L. (2000) Retards mentaux liés au chromosome X. *Med sciences*, **16**, 363-372.
- Chen, R., Akbarian, S., Tudor, M. and Jaenisch, R. (2001) Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat Genet*, **27**, 327-331.
- Chiurazzi, P., Hamel, B. C. and Neri, G. (2001) XLMR genes: update 2000. *Eur J Hum Genet*, **9**, 71-81.
- Clayton-Smith, J., Watson, P., Ramsden, S. and Black, G. (2000) Somatic mutation in MECP2 as a non-fatal neurodevelopmental disorder in males. *Lancet*, **356**, 830-832.
- Colantuoni, C., Jeon, O., Hyder, K., Chenchik, A., khimani, A., Narayanan, V., Hoffman, E., Kaufmann, W., Naidu, S. and Pevsner, J. (2001) Gene expression profiling in postmortem Rett syndrome brain: differential gene expression and patient classification. *Neurobiol Dis*, **8**, 847-865.
- Collins, F. (1992) Positional cloning: let's not call it reverse anymore. *Nat Genet*, **1**, 3-6.
- Colot, V. and Rossignol, J. (1999) Eukaryotic DNA methylation as an evolutionary device. *Bioessays*, **21**, 402-411.
- Comb, M. and Goodman, H. (1990) CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Res*, **18**, 3975-3982.
- Costa, J., Vidaud, D., Laurendreau, I., Vidaud, M., Fressinaud, E., Moisan, J., David, A., Meyer, D. and Lavergne, J. (2000) Somatic mosaicism and compound heterozygosity in female hemophilia B. *Blood*, **96**, 1585-1587.
- Couvert, P., Bienvenu, T., Aquaviva, C., Poirier, K., Moraine, C., Gendrot, C., Verloes, A., Andres, C., Le Fevre, A., Souville, I., Steffann, J., des Portes, V., Ropers, H., Yntema, H., Fryns, J., Briault, S., Chelly, J. and Beldjord, C. (2001) MECP2 is highly mutated in X-linked mental retardation. *Hum Mol Genet*, **10**, 941-946.

- Coy, J., Sedlacek, Z., Bachner, D., Delius, H. and Poustka, A. (1999) A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3' untranslated region of the methyl-CPG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. *Hum Mol Genet*, **8**(7), 1253-1262.
- Cremers, F., Bitner-Glindziez, M., Pembrey, M. and Roppers, H. (1995) Mapping and cloning hereditary deafness genes. *Curr Opin Genet Dev*, **5**, 371-375.
- Culbertson, M. (1999) RNA surveillance unforeseen consequences for gene expression, inherited genetic disorders and cancer. *Trends Genet*, **15**, 74-80.
- Curtis, A., Headland, S., Lindsay, S., Thomas, N., Boye, E., Kamakari, S., Roustan, P., Anvret, M., Wahlstrom, J., McCarthy, G., Clarke, A. and Bhattacharya, S. (1993) X chromosome linkage studies in familial Rett syndrome. *Hum Genet*, **90**, 551-555.
- D'Esposito, M., Quaderi, N., Ciccodicola, A., Bruni, P., Esposito, T., D'Urso, M. and Brown, S. (1996) Isolation, physical mapping, and northern analysis of the X-linked human gene encoding methyl CpG-binding protein, MECP2. *Mamm Genome*, **7**, 533-535.
- De Bona, C., Zappella, M., Hayek, G., Meloni, I., Vitelli, F., Bruttini, M., Cusano, R., Loffredo, P., Longo, I. and Renieri, A. (2000) Preserved speech variant is allelic of classic Rett syndrome. *Eur J Hum Genet*, **8**, 325-330.
- Debrand, E., Chureau, C., Arnaud, D., Avner, P. and Heard, E. (1999) Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. *Mol Cell Biol*, **19**, 8513-8525.
- Delobel, B., Delannoy, V., Pini, G., Zapella, M., Tardieu, M., Vallée, L. and Croquette, M. (1998) Identification and molecular characterization of a small 11q23.3 de novo duplication in a patient with Rett syndrome manifestations. *Am J Med Genet*, **80**, 273-280.
- 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. and Weissenbach, J. (1996) A comprehensive genetic map of the human genome based on 5264 microsatellites. *Nature*, **380**, 152-154.
- Dotti, M., Manneschi, L., Malandrini, A., De Stefano, N., Carnerale, F. and Federico, A. (1993) Mitochondrial dysfunction in Rett syndrome. *Brain Dev*, **15**, 103-106.
- Dragich, J., Houwink-Manville, I. and Schanen, C. (2000) Rett syndrome: a surprising result of mutation in MECP2. *Hum Mol Genet*, **9**, 2365-2375.
- Driscoll, D. and Migeon, B. (1990) Sex difference in methylation of single-copy genes in human meiotic germ cells: implications for X chromosome inactivation, parental imprinting, and origin of CpG mutations. *Somat Cell Mol Genet*, **16**, 267-282.
- Dunger, S., Neumann, S., Zell, R., Birch-Hirschfeld, E., Stelzner, A., Paschke, R., Kinne, R. and Sickinger, S. (2001) Mutation detection in mosaic situations: RNA mismatch assay and denaturing gradient gel electrophoresis are more sensitive than conventional cycle sequencing. *Anal Biochem*, **294**, 89-93.

Echenne, B., Bressot, N., Cheminal, R., Van Leuwen, F., Compagnon, P. and Azaïs, M. (1989) Le syndrome de Rett. *Annales de Pédiatrie*, **36**, 661-665.

Edwalds-Gilbert, G., Veraldi, K. and Milcarek, C. (1997) Alternative poly(A) site selection in complex transcription units: means to end? *Nucleic Acids Res*, **25**, 2547-2561.

Eichler, E., Lu, F., Shen, Y., Antonacci, R., Jurecic, V., Doggett, N., Moyzis, R., Baldini, A., Gibbs, R. and Nelson, D. (1996) Duplication of a gene-rich cluster between 16p11.1 and Xq28: A novel pericentromeric-directed mechanism for paralogous genome evolution. *Hum Mol Genet*, **5**, 899-912.

Eichler, E., Budarf, M., Rocchi, M., Deaven, L., Doggett, N., Baldini, A., Nelson, D. and Mohrenweiser, H. (1997) Interchromosomal duplications of the adrenoleukodystrophy locus: A phenomenon of pericentromeric plasticity. *Hum Mol Genet*, **6**, 991-1002.

Ellaway, C. and Christodoulou, J. (1999) Rett syndrome: clinical update and review of recent genetic advances. *J Paediat Child Health*, **35**, 419-426.

Ellison, K., Fill, C., Terwilliger, J., DeGennaro, L., Martin-Gallardo, A., Anvret, M., Percy, A., Ott, J. and Zoghbi, H. (1992) Examination of X chromosome markers in Rett syndrome: exclusion mapping with a novel variation on multilocus linkage analysis. *Am J Hum Genet*, **50**, 278-287.

Ellison, K., Roth, E., McCabe, E., Chinault, A. and Zoghbi, H. (1993) Isolation of a yeast artificial chromosome contig spanning the X chromosomal translocation breakpoint in a patient with Rett syndrome. *Am J Med Genet*, **47**, 1124-1134.

Epstein, C., Smith, S., Travis, B. and Tucker, G. (1978) Both X chromosomes function before visible X-chromosome inactivation in female mouse embryos. *Nature*, **274**, 500-503.

Frederico, L., Kunkel, T. and Shaw, B. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry*, **29**, 2532-2537.

Fujita, N., Shimotake, N., Ohki, I., Chiba, T., Saya, H., Shirakawa, M. and Nakao, M. (2000) Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1. *Mol Cell Biol*, **20**, 5107-5118.

Fukami, M., Kirsch, S., Schiller, S., Richter, A., Benes, V., Franco, B., Mutoya, K., Rao, E., Merker, S., Niesler, B., Ballabio, A., Ansorge, W., Ogata, T. and Rappold, A. (2000) A member of a gene family on Xp22.3, VCX-A, is deleted in patients with X-linked nonspecific mental retardation. *Am J Hum Genet*, **67**, 563-573.

Fuks, F., Burgers, W., Brehm, A., Hughes-Davies, L. and Kouzarides, T. (2000) DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat genet*, **24**, 88-91.

- Ganguly, A., Rock, M. and Prockop, D. (1993) Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. *Proc Natl Acad Sci US A*, **90**, 10325-10329.
- Gibbons, R., Suthers, S., Wilkie, A., Buckle, V. and Higgs, D. (1992) X-linked alpha thalassemia/mental retardation syndrome: localization to Xq12-21.31 by X inactivation and linkage analysis. *Am J Hum Genet*, **51**, 1136-1149.
- Gibbons, R., Picketts, D., Villard, L. and Higgs, D. (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α -thalassemia (ATR-X syndrome). *Cell*, **80**, 837-845.
- Gillberg, C., Wahlstrom, J. and Hagberg, B. (1984) Infantile autism and Rett's syndrome: common chromosomal denominator. *Lancet*, **2**, 1094-1095.
- Gillberg, C., Wahlstrom, J. and Hagberg, B. (1985) A "new" chromosome marker common to the Rett syndrome and infantile autism? The frequency of fragile sites at X p22 in 81 children with infantile autism, childhood psychosis and the Rett syndrome. *Brain Dev*, **7**, 365-367.
- Girard, M., Couvert, P., Carrie, A., Tardieu, M., Chelly, J., Beldjord, C. and Bienvenu, T. (2001) Parental origin of de novo MECP2 mutations in Rett syndrome. *Eur J Hum Genet*, **9**, 231-236.
- Giunti, L., Pelagatti, S., Lazzerini, V., Guardicci, S., Lapi, E., Coviello, S., A., C., Ombroni, L., Andreucci, E., Sani, I., Brusaferrri, A., Lasagni, A., Ricotti, G., Giometto, B., Nicolao, P., Gasparini, P., Granatiero, M. and Giovannucci-Uzielli, M. (2001) Spectrum and distribution of MECP2 mutations in 64 Italian Rett syndrome girls: tentative genotype/phenotype correlation. *Brain Dev*, **23-Suppl**, 242-245.
- Glaser, R., Jiang, W., Boyadjiev, S., Tran, A., Zachary, A., Van Maldergem, L., Johnson, D., Walsh, S., Oldridge, M., Wall, S., Wilkie, A. and Jabs, E. (2000) Paternal origin of FGFR2 mutations in sporadic cases of Crouzon syndrome and Pfeiffer syndrome. *Am J Hum Genet*, **66**, 768-777.
- Gordon, K., Mok Siu, V., Sergovich, F. and Jung, J. (1993) 18q-mosaicism associated with Rett syndrome phenotype. *Am J Med Genet*, **46**, 142-144.
- Gray, S. and Ekstrom, T. (2001) The human histone deacetylase family. *Exp Cell Res*, **262**, 75-83.
- Guideri, F., Acampa, M., Hayek, G., Zapella, M. and Di Perri, T. (1999) Reduced heart rate variability in patients affected with Rett syndrome. A possible explanation for sudden death. *Neuropediatrics*, **30**, 146-148.
- Gustavsson, P., Kimber, E., Wahlström, J. and Anneren, G. (1999) Monosomy 18q syndrome and atypical Rett syndrome in a girl with an interstitial deletion (18) (q21.1q22.3). *Am J Hum Genet*, **82**, 348-351.

- Guy, J., Hendrich, B., Holmes, M., Martin, J. E. and Bird, A. (2001) A mouse *Mecp2*-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet*, **27**, 322-326.
- Haas, R., Dixon, S., Sartoris, D. and Hennessy, M. (1997) Osteopenia in Rett syndrome. *J Pediatr*, **131**, 771-774.
- Hagberg, B., Aicardi, J., Dias, K. and Ramos, O. (1983) A progressive syndrome of autism, dementia, ataxia and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol*, **14**, 471-479.
- Hagberg, B. (1985) Rett syndrome: Swedish approach to analysis of prevalence and cause. *Brain Dev*, **7**, 276-80.
- Hagberg, B. and Witt Engerstrom, I. (1986) Rett syndrome: A suggested staging system for describing impairment profile with increasing age towards adolescence. *Am J Med Genet*, **1**, 47-59.
- Hagberg, B. and Skjeldal, O. (1994) Rett variants: a suggested model for inclusion criteria. *Pediatr Neurol*, **11**, 5-11.
- Hampson, K., Woods, C., Latif, F. and Webb, T. (2000) Mutations in the *MECP2* gene in a cohort of girls with Rett syndrome. *J Med Genet*, **37**, 610-612.
- Hecht, F. and Hecht, B. (1991) Genetic activity and fragile sites. *Cancer Genet Cytogenet*, **55**, 113.
- Hendrich, B. and Bird, A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol*, **18**, 6538-6547.
- Hendrich, B., Abbott, C., McQueen, H., Chambers, D., Cross, S. and Bird, A. (1999) Genomic structure and chromosomal mapping of the murine and human *Mbd1*, *Mbd2*, *Mbd3*, and *Mbd4* genes. *Mamm Genome*, **10**, 906-912.
- Herbst, D. and Miller, J. (1980) X-linked mental retardation II: the frequency in British Columbia. *Am J Med Genet*, **7**, 461-469.
- Hernandez-Martin, A., Gonzalez-Sarmiento, R. and De Unamuno, P. (1999) X-linked ichthyosis: an update. *Br J Dermatol*, **141**, 617-627.
- Hoffbuhr, K., Devaney, J., LaFleur, B., Sirianni, N., Scacheri, C., Giron, J., Schuette, J., Innis, J., Marino, M., Philippart, M., Narayanan, V., Umansky, R., Kronn, D., Hoffman, E. and Naidu, S. (2001) *MeCP2* mutations in children with the phenotype of Rett syndrome. *Neurology*, **56**, 1486-1495.
- Hofferbert, S., Schanen, N., Budden, S. and Francke, U. (1997) Is Rett syndrome caused by a triplet repeat expansion? *Neuropediatrics*, **28**, 179-183.
- Holler, M., Westin, G., Jiricny, J. and Schaffner, W. (1988) Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev*, **2**, 1127-1135.

- Huppke, P., Laccone, F., Kramer, N., Engel, W. and Hanefeld, F. (2000) Rett syndrome: analysis of MECP2 and clinical characterization of 31 patients. *Hum Mol Genet*, **9**, 1369-1375.
- Igushi-Ariga, S. and Schaffner, W. (1989) CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev*, **3**, 612-619.
- Imai, S., Armstrong, C., Kaeberlein, M. and Guarente, L. (2000) Transcriptional silencing and longevity protein Sir2 is an NAD-dependant histone deacetylase. *Nature*, **403**, 795-800.
- Imessaoudene, B., Bonnefont, J., Royer, G., V., C.-D., Lyonnet, S., Lyon, G., Munnich, A. and Amiel, J. (2001) MECP2 mutation in non-fatal, non-progressive encephalopathy in a male. *J Med Genet*, **38**, 171-174.
- Inui, K., Akagi, M., Ono, J., Tsukamoto, H., Shimono, K., Mano, T., Imai, K., Yamada, M., Muramatsu, T., Sakai, N. and Okafa, S. (2001) Mutational analysis of MECP2 in Japanese patients with atypical Rett syndrome. *Brain Dev*, **23-Suppl**, 212-215.
- Jan, M., Dooley, J. and Gordon, K. (1999) Male Rett syndrome variant: application of diagnostic criteria. *Pediatr Neurol*, **20**, 238-240.
- Jenuwein, T. and Allis, C. (2001) Translating the histone code. *Science*, **293**, 1074-1079.
- Johnson, W. (1980) Metabolic interference and the +/- heterozygote. A hypothetical form of simple inheritance which is neither dominant nor recessive. *Am J Hum Genet*, **32**, 374-386.
- Jones, P., Veenstra, G., Wade, P., Vermaak, D., Kass, S., Landsberger, N., Strouboulis, J. and Wolffe, A. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet*, **19**, 187-191.
- Journel, H., Melki, J., Turleau, C., Munnich, A. and de Grouchy, J. (1990) Rett phenotype with X/autosome translocation: possible mapping to the short arm of chromosome. *Am J Med Genet*, **35**, 142-147.
- Kaludov, N. and Wolffe, A. (2000) MeCP2 driven transcriptional repression *in vitro*: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery. *Nucleic Acids Res*, **28**, 1921-1928.
- Kaplan, J., Kahn, A. and Chelly, J. (1992) Illegitimate transcription: its use in the study of inherited disease. *Hum Mutat*, **1**, 357-360.
- Kay, G., Penny, G., Patel, D., Ashworth, A., Brockdorff, N. and Rastan, S. (1993) Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell*, **72**, 171-182.
- Khochbin, S., verdel, A., lemercier, C. and Seigneurin-Berny, D. (2001) Functional significance of histone deacetylase diversity. *Curr Opin Genet dev*, **11**, 162-166.

- Killian, W. (1986) On the genetics of Rett syndrome: analysis of family and pedigree data. *Am J Med Genet*, **1-suppl**, 369-376.
- Kim, S. and Cook, J. (2000) Novel de novo nonsense mutation of MECP2 in a patient with Rett syndrome. *Hum Mutat*, **15**, 382-383.
- Kitt, C. and Wilcox, B. (1995) Preliminary evidence for neurodegenerative changes in the substantia nigra of Rett syndrome. *Neuropediatrics*, **26**, 114-118.
- Klauck, S., Lindsay, S., Beyer, K., Splitt, M., Burn, J. and Poustka, A. (2002) A mutation hot spot for nonspecific X-linked mental retardation in the MECP2 gene causes the PPM-X syndrome. *Am J Hum Genet*, **70**, 1034-1037.
- Knoepfler, P. and Eisenman, R. (1999) Sin meets NuRD and other tails of repression. *Cell*, **99**, 447-450.
- Kondo, I., Morishita, R., Fukuda, T., Obata, K., Matsuishi, T., Yamashita, Y., Kuwajima, K., Horiuchi, I. and Omori, I. (2000) The spectrum and parental origin of *de novo* mutations of methyl-CpG-binding protein 2 gene (MECP2) in Rett syndrome. *Am J Hum Genet*, **67-Suppl2**, 386.
- Kormann-Bortolotto, M., Woods, C., Green, S. and Webb, T. (1992) X-inactivation in girls with Rett syndrome. *Clin Genet*, **42**, 296-301.
- Krawczak, M. and Cooper, D. (1996) Single base-pair substitutions in pathology and evolution: two sides to the same coin. *Hum Mutat*, **8**, 23-31.
- Krepischi, A., Kok, F. and Otto, P. (1998) X chromosome-inactivation patterns in patients with Rett syndrome. *Hum Genet*, **102**, 319-321.
- Kudo, S. (1998) Methyl-CpG-binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated. *Mol Cell Biol*, **18**, 5492-5499.
- Kudo, S., Nomura, Y., Segawa, M., Fujita, N., Nakao, M., Dragich, J., Schanen, C. and Tamura, M. (2001) Functional analyses of MeCP2 mutations associated with Rett syndrome using transient expression systems. *Brain Dev*, **23-Suppl**, 165-173.
- Kudo, S., Nomura, Y., Segawa, M., Fujita, N., Nakao, M., Hammer, S., Schanen, C., Terai, I. and Tamura, M. (2002) Functional characterisation of MeCP2 mutations found in male patients with X linked mental retardation. *J Med Genet*, **39**, 132-136.
- Kuo, M., Brownell, J., Sobel, R., Ranalli, T., Cook, R., Edmondson, D., Roth, S. and Allis, C. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*, **383**, 269-272.
- Laccone, F., Huppke, P., Hanefeld, F. and Meins, M. (2001) Mutation spectrum in patients with Rett syndrome in the German population: Evidence of hot spot regions. *Hum Mutat*, **17**, 183-190.

- Laloti, M., Scott, H. and Antonarakis, S. (1997) What is expanded in progressive myoclonus epilepsy? *Nat Genet*, **17**, 17.
- LaSalle, J., Goldstine, J., Balmer, D. and Greco, C. (2001) Quantitative localization of heterogeneous methyl-CpG-binding protein 2 (MeCP2) expression phenotypes in normal and Rett syndrome brain by laser scanning cytometry. *Hum Mol Genet*, **10**, 1729-1740.
- Lathrop, G., Lalouel, J., Julier, C. and Ott, J. (1984) Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci USA*, **81**, 3443-3446.
- Lebo, R., Ikuta, T., Milunsky, J. and Milunsky, A. (2001) Rett syndrome from quintuple and triple deletions within the *MECP2* deletion hotspot region. *Clin Genet*, **59**, 406-417.
- Lee, J., Davidow, L. and Warshawsky, D. (1999) *Tsix*, a gene antisense to *Xist* at the X-inactivation center. *Nat Genet*, **21**, 400-404.
- Lehrke, R. (1972) A theory of X-linkage of major intellectual traits. Response to Dr. Anastasi and to the Drs. Nance and Engel. *Am J Ment Defic*, **76**, 626-631.
- Lei, H., Oh, S., Okano, M., Jüttermann, R., Goss, K., Jaenisch, R. and Li, E. (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development*, **122**, 3195-3205.
- Leonard, H., Thomson, M., Glasson, E., Fyfe, S., Leonard, S., Bower, C., Christodoulou, J. and Ellaway, C. (1999) A population-based approach to the investigation of osteopenia in Rett syndrome. *Dev Med Child Neurol*, **41**, 323-328.
- Leuer, M., Oldenburg, J., Lavergne, J., Ludwig, M., Fregin, A., Eigel, A., Ljung, R., Goodeve, A., Peake, I. and Olek, K. (2001) Somatic mosaicism in hemophilia A: a fairly common event. *Am J Hum Genet*, **69**, 75-87.
- Lewis, J., Meehan, R., Henzel, W., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905-914.
- Lewis, D., Erichson, C. and Castora, F. (1995) Mutation analysis of mitochondrial transfer RNA genes in Rett syndrome. *Pediatr Neurol*, **11**, 143-144.
- Lewis, D., Burgess, C., Naidu, S., Castora, F. and Norfolk, V. (1998) Single-stranded conformational polymorphism analysis of mtDNA in Rett syndrome. *Ann Child Neurol*, **38**, 352.
- Li, E., Bestor, T. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915-926.
- Liu, W., Smith, D., Rechtzigel, K., Thibodeau, S. and James, C. (1998) Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. *Nucleic Acids Res.*, **26**, 1396-1400.

- Lyon, M. (1961) Gene action in the X-chromosome of the mouse. *Nature*, **190**, 372-373.
- Maquat, L. (1996) Defects in RNA splicing and the consequence of shortened translational reading frame. *Am J Hum Genet*, **59**, 279-286.
- Marin, M., Karis, A., Visser, P., Grosveld, F. and Philipson, S. (1997) Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. *Cell*, **89**, 619-628.
- Marmorstein, R. and Roth, S. (2001) Histone acetyltransferases: function, structure, and catalysis. *Curr Opin Genet Dev*, **11**, 155-161.
- Mattei, M., Mattei, J., Ayme, S. and Giraud, F. (1982) X-autosome translocations: cytogenetic characteristics and their consequences. *Hum Genet*, **61**, 295-309.
- Matsuura, T., Yamagata, T., Burgess, D., Rasmussen, A., Grewal, R., Watase, K., Khajavi, M., McCall, A., Davis, C., Zu, L., Achari, M., Pulst, S., Alonso, E., JL., N., Nelson, D., Zoghbi, H. and Ashizawa, T. (2000) Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat Genet*, **26**, 191-194.
- Mazoyer, S., Dunning, A., Serova, O., Dearden, J., Puget, N., Healey, C., Gayther, S., Mangion, J., Stratton, M., Lynch, H., Goldgar, D., Ponder, B. and Lenoir, G. (1996) A polymorphic stop codon in *BRCA2*. *Nat Genet*, **14**, 253-254.
- McDowell, T., Gibbons, R., Sutherland, H., O'Rourke, D., Bickmore, W., Pombo, A., Turley, H., Gatter, K., Picketts, D., Buckle, V., Chapman, L., Rhodes, D. and Higgs, D. (1999) Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc Natl Acad Sci*, **96**, 13983-13988.
- Meehan, R., Lewis, J., McKay, S., Kleiner, E. and Bird, A. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, **58**, 499-507.
- Meehan, R., Lewis, J. and Bird, A. (1992) Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res*, **20**, 5085-5092.
- Meloni, I., Bruttini, M., Longo, I., Mari, F., Rizzolio, F., D'Adamo, P., Denvriendt, K., Fryns, J., Toniolo, D. and Renieri, A. (2000) A mutation in the Rett syndrome gene, MECP2, causes X-linked mental retardation and progressive spasticity in males. *Am J Hum Genet*, **67**, 982-985.
- Migeon, B., Dunn, M., Thomas, G., Schmeckpeper, B. and Naidu, S. (1995) Studies of X inactivation and isodisomy in twins provide further evidence that the X chromosome is not involved in Rett syndrome. *Am J Hum Genet*, **56**, 647-653.
- Mlynarczyk, S. and Panning, B. (2000) X inactivation: *Tsix* and *Xist* as yin and yang. *Current Biology*, **10**, R899-R903.

- Moloney, D., Slaney, S., Oldrige, M., Wall, S., Sahlin, P., Stenman, G. and Wildie, A. (1996) Exclusive paternal origin of new mutations in Apert syndrome. *Nat Genet*, **13**, 9-10.
- Moncla, A., Kpebe, A., Missirian, C., Mancini, J. and Villard, L. (2002) Polymorphisms in the C-terminal domain of MECP2 in mentally handicapped boys: implications for genetics counselling. *Congrès: Jères Assises de Génétique Humaine et Médicale. 18-19-20 Janvier 2002*.
- Monros, E., Armstong, J., Aibar, E., Poo, P., Canos, I. and Pineda, M. (2001) Rett syndrome in Spain: mutation analysis and children correlations. *Brain Dev*, **23-Suppl**, 251-253.
- Moore, J., Tuck-Muller, C., Murphy, M., Naidu, S. and Thomas, G. (1986) Chromosome studies in 10 patients with the Rett syndrome. *Am J Med Genet*, **24**, 345-354.
- Morton, N. (1955) Sequential tests for the detection of linkage. *Am J Hum Genet*, **7**, 277-318.
- Motil, K., Schultz, R., Brown, B., Glaze, D. and Percy, A. (1994) Altered energy balance may account for growth failure in Rett syndrome. *J Child Neurol*, **9**, 315-319.
- Motil, K., Schultz, R., Wong, W. and Glaze, D. (1998) Increased energy expenditure associated with repetitive involuntary movement does not contribute to growth failure in girls with Rett syndrome. *J Pediatr*, **132**, 228-233.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol*, **51**, 263-273.
- Murakami, J., Courchesne, E., Haas, R., Press, G. and Yeung-Courchesne, R. (1992) Cerebellar and cerebral abnormalities in Rett syndrome: a quantitative MR analysis. *AJR*, **159**, 177-183.
- Murphy, M., Naidu, S. and Moser, H. (1986) Rett syndrome : observational study in 33 families. *Am J Med Genet*, **24**, 73-76.
- Nan, X., Meehan, R. and Bird, A. (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res*, **21**, 4886-4892.
- Nan, X., Tate, P., Li, E. and Bird, A. (1996) DNA methylation specifies chromosomal localization of MeCP2. *Mol Cell Biol*, **16**, 414-421.
- Nan, X., Campoy, F. and Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, **88**, 471-481.
- Nan, X., Ng, H., Johnson, C., Laherty, C., Turner, B., Eisenman, R. and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386-369.
- Naumova, A., Olien, L., Bird, L., Smith, M., Verner, A., Leppert, M., Morgan, K. and Sapienza, C. (1998) Genetic mapping of X-linked loci involved in skewing of X chromosome inactivation in the human. *Eur J Hum Genet*, **6**, 552-562.

Ng, H., Zhang, Y., Hendrich, B., Johnson, C., Turner, B., Erdjument, B., Tempst, P., Reinberg, D. and Bird, A. (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet*, **23**, 58-61.

Ng, H., Jeppesen, P. and Bird, A. (2000) Active repression of methylated genes by the chromosomal protein MBD1. *Mol Cell Biol*, **20**, 1394-1406.

Nicolao, P., Carella, M., Giometto, B., Tavolato, B., Cattin, R., Giovannucci-Uzielli, M., Vacca, M., Della Regione, F., Piva, S., Bortoluzzi, S. and Gasparini, P. (2001) DHPLC analysis of the MECP2 gene in Italian Rett patients. *Hum Mutat*, **18**, 132-140.

a-Nielsen, J., Henriksen, K., Hansen, C., Silaharoglu, A., Schwartz, M. and Tommerup, N. (2001) MECP2 mutations in Danish patients with Rett syndrome: high frequency of mutations but no consistent correlations with clinical severity or with the X chromosome inactivation pattern. *Eur J Hum Genet*, **9**, 178-184.

b-Nielsen, J., Ravn, K. and Schwartz, M. (2001) A 77-year-old woman and a preserved speech variant among the Danish Rett patients with mutations in MECP2. *Brain Dev*, **23-Suppl**, 230-232.

O'Donovan, M., Oefner, P., Roberts, S., Austin, J., Hoogendoorn, B., Guy, C., Speight, G., Upadhyaya, M., Sommer, S. and McGuffin, P. (1998) Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. *Genomics*, **52**, 44-49.

Obata, K., Matsuishi, T., Yamashita, Y., Fukuda, T., Kuwajima, K., Horiuchi, I., Nagamitsu, S., Iwanaga, R., Kimura, A., Omori, I., Endo, S., Mori, K. and Kondo, I. (2000) Mutation analysis of the methyl-CpG binding protein 2 gene (MECP2) in patients with Rett syndrome. *J Med Genet*, **37**, 608-610.

Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M. and Mandel, J. (1991) Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science*, **252**, 1097-1102.

Oefner, P. and Underhill, P. (1995) Comparative DNA sequencing by denaturing high-performance liquid chromatography (DHPLC). *Am J Hum Genet*, **57**, A266.

Oefner, P. and Underhill, P. (1998) Detection of nucleic acid heteroduplex molecules by denaturing high-performance liquid chromatography and methods for comparative sequencing. *US Patent*, **5**, 976.

Okano, M., Bell, D., Haber, D. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247-257.

Ohki, I., Shimotake, N., Fujita, N., Nakao, M. and Shirakawa, M. (1999) Solution structure of the methyl-CpG-binding domain of the methylation-dependent transcriptional repressor MBD1. *Embo J*, **18**, 6653-61.

Orrico, A., Lam, C.-W., Galli, L., Dotti, M., Hayek, G., Tong, S., Poon, P., Zappella, M., Federico, A. and Sorrentino, V. (2000) MECP2 mutation in male patients with non-specific X-linked mental retardation. *FEBS Letters*, **481**, 285-288.

Panasiuk, B., Midro, A. and Zadrozna-Tolwinska, B. (1997) Earlier finishing of Xp21.2 subband replication of the inactive X chromosome in Rett syndrome girl but not in her 47,XXX mother. *Clin Genet*, **52**, 120-125.

Panning, B., Dausman, J. and Jaenisch, R. (1997) X chromosome inactivation is mediated by RNA stabilisation. *Cell*, **90**, 907-916.

Partington, M. (1988) Rett syndrome in monozygotic twins. *Am J Hum Genet*, **29**, 633-637.

Pereira, L. and Vasques, L. (2000) X-chromosome inactivation: lessons from transgenic mice. *Gene*, **255**, 363-371.

Philippe, C., Porter, D., Emerton, M., Wells, D., Simpson, A. and Monaco, A. (1997) Mutation screening of the EXT1 and EXT2 genes in patients with hereditary multiple exostoses. *Am J Genet*, **61**, 520-528.

Plenge, R., Hendrich, B., Schwartz, C., Arena, J., Naumova, A., Sapienza, C., Winter, R. and Willard, H. (1997) A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. *Nat Genet*, **17**, 353-356.

Razin, A. (1998) CpG methylation, chromatin structure and gene silencing—a three-way connection. *Embo J*, **17**, 4905-4908.

Reichwald, K., Thiesen, J., Wiehe, T., Weitzel, J., Stätling, W., Kioschis, P., Poustka, A., Rosenthal, A. and Platzer, M. (2000) Comparative sequence analysis of the MECP2-locus in human and mouse reveals new transcribed regions. *Mamm Genome*, **11**, 182-190.

Rett, A. (1966) Über ein eigenartiges himatrophisches Syndrom bei Hyperammonämie im Kindesalter. *Wien Med Wochenschr*, **116**, 723-738.

Robinson, M., Chu, C., Turnier, G., Bishop, D. and Taylor, G. (2000) Exon deletions and duplications in BRCA1 detected by semiquantitative PCR. *Genet Testing*, **4**, 49-54.

Roeder, R. (1991) The complexity of eukaryotic transcription initiation: regulation of preinitiation complex assembly. *Trends Biochem Sci*, **16**, 402-408.

Romeo, G., Archidiacono, N., Ferlini, A. and Roccho, M. (1986) Rett syndrome: Lack of association with fragile site Xp22 and strategy for genetic mapping of X-linked new mutations. *Am J Med Genet*, **24**, 355-359.

Rosenberg, C., Wouters, C., Szuhai, K., Dorland, R., Pearson, P., Poll-The, B., Colombijn, R., Breuning, M. and Lindhout, D. (2001) A Rett syndrome patient with a ring X chromosome: further evidence for skewing of X inactivation and heterogeneity in the aetiology of the disease. *Eur J Hum Genet*, **9**, 171-177.

Ruch, A., Kurczynski, T. and Velasco, M. (1989) Mitochondrial alterations in Rett syndrome. *Pediatr. Neurol*, **5**, 320-323.

Schanen, N., Roth Dahle, E., Capozzoli, F., Holm, V., Zoghbi, H. and Francke, U. (1997) A new Rett syndrome family consistent with X-linked inheritance expands the X chromosome exclusion map. *Am J Hum Genet*, **61**, 634-641.

Schanen, N. (1999) Molecular approaches to the Rett syndrome gene. *J Child Neurol*, **14**, 806-814.

Schmidt, M. and Du Sart, D. (1992) Functional disomies of X chromosome influence the cell selection and hence the X inactivation pattern in females with balanced X;autosome translocations: a review of 122 cases. *Am J Med Genet*, **42**, 161-169.

Shahbazian, M., Antalffy, B., Armstrong, D. and Zoghbi, H. (2002) Insight into Rett syndrome : MeCP2 levels display tissue-and cell-specific differences and correlate with neuronal maturation. *Hum Mol Genet*, **11**, 115-124.

Sheardown, S., Duthie, S., Johnston, C., Newall, A., Formstone, E., Arkell, R., Nesterova, T., Alghisis, G., Rastan, S. and Brockdorff, N. (1997) Stabilization of *Xist* RNA mediates initiation of X-chromosome. *Cell*, **91**, 99-107.

Shen, J., Rideout, W. and Jones, P. (1994) The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res*, **22**, 972-976.

Simmler, M., Cattanach, B., Rasberry, C., Rouguelle, C. and Avner, P. (1993) Mapping the murine *Xce* locus with (CA)_n repeats. *Mamm Genome*, **4**, 523-530.

Sirianni, N., Naidu, S., Pereira, J., Pillotto, R. and Hoffman, E. (1998) Rett syndrome: confirmation of X-linked dominant inheritance, and localization of the gene to Xq28. *Am J Hum Genet*, **63**, 1552-1558.

Solage, A. and Cedar, H. (1978) Organization of 5-methylcytosine in chromosomal DNA. *Biochemistry*, **17**, 2934-2938.

Southern, E. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol*, **88**, 503-517.

Spencer, V. and Davie, J. (1999) Role of covalent modifications of histones in regulating gene expression. *Gene*, **240**, 1-12.

Steffenburg, U., Hagberg, G. and Hagberg, B. (2001) Epilepsy in a representative series of Rett syndrome. *Acta Paediatr*, **90**, 34-39.

Sung Jae Lee, S., Wan, M. and Francke, U. (2001) Spectrum of MECP2 mutations in Rett syndrome. *Brain Dev*, **23-Suppl**, 138-143.

Taddei, A., Ray-Gallet, D. and Almouzni, G. (2000) Assemblage et remodelage: le nucléosome sous influence. *Med sciences*, **16**, 603-610.

- Tang, J., Qi, Y., Bao, X. and Wu, X. (1997) Mutational analysis of mitochondrial DNA of children with Rett syndrome. *Pediatr Neurol*, **17**, 327-330.
- Tariverdian, G., Kantner, G. and Vogel, F. (1987) A monozygotic twin pair with Rett syndrome. *Hum Genet*, **75**, 88-90.
- Tate, P., Skarnes, W. and Bird, A. (1996) The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nat Genet*, **12**, 205-208.
- Telvi, L., Leboyer, M., Chiron, C., Feingold, J. and Ponsot, G. (1994) Is Rett syndrome a chromosome breakage syndrome? *Am J Med Genet*, **51**, 602-605.
- Thomas, G. (1996) High male:female ratio of germ-line mutations: An alternative explanation for postulated gestational lethality in males in X-linked dominant disorders. *Am J Hum Genet*, **58**, 1364-1368.
- Trappe, R., Laccone, F., Cobilanschi, J., Meins, M., Huppke, P., Hanefeld, F. and Engel, W. (2001) Mecp2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am J Hum Genet*, **68**, 1093-1101.
- Trevarthen, E., Moser, H. and Group., t. D. C. W. (1988) Diagnostic criteria for Rett syndrome. The Rett Syndrome Diagnostic Criteria Work Group. *Ann Neurol*, **23**, 425-428.
- Vacca, M., Filippini, F., Budillon, A., Rossi, V., Della Ragione, F., De Bonis, M., Mercadante, G., Manzati, E., Gualandi, F., Bigoni, S., Trabanelli, C., Pini, G., Calzolari, E., Ferlini, A., Meloni, I., Hayek, G., Zappella, M., Renieri, A., D'Urso, M., D'Esposito, M., Macdonald, F., Kerr, A., Dhanjal, S. and Hulten, M. (2001) MECP2 gene mutation analysis in the British and Italian Rett syndrome patients: hot spot map of the most recurrent mutations and bioinformatic analysis of a new MECP2 conserved region. *Brain Dev*, **23-Suppl**, 246-250.
- Verkerk, A., Pieretti, M., Sutcliffe, J., Fu, Y., Kuhl, D., Pizzuti, A., Reiner, O., Richards, S., Victoria, M., Zhang, F. and al., e. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, **65**, 905-914.
- Vilain, A., Apiou, F., Vogt, N., Dutrillaux, B. and Malfoy, B. (1996) Assignment of the gene for methyl-CpG-binding protein 2 (MECP2) to human chromosome band Xq28 by in situ hybridization. *Cytogenet Cell Genet*, **74**, 293-294.
- Villard, L., Kpebe, A., Cardoso, C., Chelly, J., Tardieu, M. and Fontes, M. (2000) Two affected boys in a Rett syndrome family. *Neurology*, **55**, 1188-1193.
- Villard, L., Levy, N., Xiang, F., Kpebe, A., Labelle, V., Chevillard, C., Zhang, Z., Schwartz, C., Tardieu, M., Chelly, J., Anvret, M. and Fontes, M. (2001) Segregation of totally skewed pattern of X chromosome inactivation in four familial cases of Rett syndrome without MECP2 mutation. *J Med Genet*, **38**, 435-442.

- Wade, P., Geggion, A., Jones, P., Ballestar, E., Aubry, F. and Wolffe, A. (1999) Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat Genet*, **23**, 62-66.
- Wade, P. (2001) Methyl CpG-binding proteins and transcriptional repression. *BioEssays*, **23**, 1131-1137.
- Wahlström, J. (1985) Genetic implications of Rett syndrome. *Brain Dev*, **7**, 573-574.
- Wahlström, J. and Anvret, M. (1986) Chromosome findings in the Rett syndrome and a test of a two-step mutation theory. *Am J Med Genet*, **24**, 361-368.
- Wahlström, J., Witt-Engerstom, I., Mellquist, L., Anvret, M. and Oden, A. (1990) The Rett syndrome related to fragile Xp22 in caffeine-induced lymphocyte culture. *Brain Dev*, **12**, 128-130.
- Wahlström, J., Uller, A., Johannesson, T., Holmqvist, D., Darnfors, C., Vujic, M., Tonny, B., Hagberg, B. and Martinsson, T. (1999) Congenital variant Rett syndrome in a girl with terminal deletion of chromosome 3p. *J Med Genet*, **36**, 343-345.
- Wakefield, R., Smith, B., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A. and Barlow, P. (1999) The solution structure of the domain from MeCP2 that binds to methylated DNA. *J Mol Biol*, **291**, 1055-1065.
- Walsh, P., Metzger, D. and Higuchi, R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-Based typing from forensic material. *Biotechniques*, **10**, 506-513.
- Wan, M., Lee, S. S., Zhang, X., Houwink, M. I., Song, H. R., Amir, R. E., Budden, S., Naidu, S., Pereira, J. L., Lo, I. F., Zoghbi, H. Y., Schanen, N. C. and Francke, U. (1999) Rett syndrome and beyond: recurrent spontaneous and familial MECP2 mutations at CpG hotspots. *Am J Hum Genet*, **65**, 1520-1529.
- Wan, M., Zhao, K., Lee, S. S. and Francke, U. (2001) MECP2 truncating mutations cause histone H4 hyperacetylation in Rett syndrome. *Hum Mol Genet*, **10**, 1085-1092.
- Wang, Y., Friedl, W., Sengteller, M., Jungck, M., Filges, I., Propping, P. and Mangold, E. (2002) A modified multiplex PCR assay for detection of large deletions in *MSH2* and *MLH1*. *Hum Mutat*, **19**, 279-286.
- Watson, P., Black, G., Ramsden, S., Barrow, M., Super, M., Kerr, B. and Clayton-Smith, J. (2001) Angelman syndrome phenotype associated with mutations in *MECP2*, a gene encoding a methyl CpG binding protein. *J Med Genet*, **38**, 224-228.
- Webb, T., Watkiss, E. and Woods, C. (1993) Neither uniparental disomy nor skewed X-inactivation explains Rett syndrome. *Clin Genet*, **44**, 236-240.
- Webb, T., Clarke, A., Hanefeld, F., Pereira, J., Rosenbloom, L. and Woods, C. (1998) Linkage analysis in Rett syndrome families suggests that there may be a critical region at Xq28. *J Med Genet*, **35**, 997-1003.

- Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, A., Millasseau, P., Vaysseix, G. and Lathrop, M. (1992) A second-generation linkage map of the human genome. *Nature*, **359**, 794-801.
- Weitzel, J., Buhrmester, H. and Stratling, W. (1997) Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG binding protein MeCP2. *Mol Cell Biol*, **17**, 5656-5666.
- Willard, H. (1996) X chromosome inactivation and X-linked mental retardation. *Am J Hum Genet*, **64**, 21-26.
- Witt-Engerstom, I. and Forslund, M. (1992) Mother and daughter with Rett syndrome. *Dev Med Child Neurol*, **12**, 131-135.
- Xiang, F., Zhang, Z., Clarke, A., Joseluiz, P., Sakkubai, N., Sarojini, B., Delozier-Blanchet, C., Hansmann, I. and Edström, L. (1998) Chromosome mapping of Rett syndrome: a likely candidate region on the telomere of Xq. *J Med Genet*, **35**, 297-300.
- Xiang, F., Buervenich, S., Nicolao, P., Bailey, M., Zhang, Z. and Anvret, M. (2000) Mutation screening in Rett syndrome patients. *J Med Genet*, **37**, 250-255.
- Xiao, W. and Oefner, P. (2001) Denaturing high-performance liquid chromatography: a review. *Hum Mutat*, **17**, 439-474.
- Xu, G., Bestor, T., Bourc'his, D., Hsieh, C., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. and Viegas-Pequignot, E. (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*, **402**, 187-191.
- Yamada, Y., Miura, K., Kumagai, T., Hayakawa, C., Miyazaki, S., Matsumoto, A., Kurosawa, K., Nomura, N., Taniguchi, H., Sonta, S., Yamanaka, T. and Wakamatsu, N. (2001) Molecular analysis of Japanese patients with Rett syndrome: identification of five novel mutations and genotype-phenotype correlation. *Hum Mutat*, **18**, 253-259.
- Yamashita, Y., Kondo, I., Fukuda, T., Morishima, R., Kusaga, A., Iwanaga, R. and Matsuishi, T. (2001) Mutation analysis of the methyl-CpG-binding protein 2 gene (MECP2) in Rett patients with preserved speech. *Brain Dev*, **23-Suppl**, 157-160.
- Yntema, H., Kleefstra, T., Oudakker, A., Jacobs, A., de Vries, B., Sistermans, E., Brunner, H., Hamel, B. and van Bokhoven, H. (2001) High frequency of MECP2 mutations MR males: implications for routine DNA diagnostics. *Am J Hum Genet*, **69-Suppl**, 632.
- Yntema, H., Oudakker, A., Kleefstra, T., Hamel, B., van Bokhoven, H., Chelly, J., Kalscheuer, V., Fryns, J., Raynaud, M., Moizard, M. and Moraine, C. (2002) In-frame deletion in MECP2 causes mild nonspecific mental retardation. *Am J Med Genet*, **107**, 81-83.
- Yoder, J., Soman, N., Verdine, G. and Bestor, T. (1997) DNA methyltransferases in mouse tissues and cells. Studies with a mechanism-based probe. *J Mol Biol*, **270**, 385-395.
- Yu, F., Thiesen, J. and Stratling, W. (2000) Histone deacetylase-independent transcriptional repression by methyl-CpG-binding protein 2. *Nucleic Acids Res*, **28**, 2201-2206.

Yusufzai, T. and Wolffe, A. (2000) Functional consequences of Rett syndrome mutations on human MeCP2. *Nucleic Acids Res*, **28**, 4172-4179.

Zappella, M., Meloni, I., Longo, I., Hayek, G. and Renieri, A. (2001) Preserved speech variants of the Rett syndrome. *Am J Med Genet*, **104**, 14-22.

Zhang, W., Bone, J., Edmondson, D., Turner, B. and Roth, S. (1998) Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *EMBO J*, **17**, 3155-3167.

a-Zoghbi, H., Ledbetter, D., Schultz, R., Percy, A. and Glaze, D. (1990) A de novo X;3 translocation in Rett syndrome. *Am J Med Genet*, **35**, 148-151.

b-Zoghbi, H., Peroy, A., Schultz, R. and Fill, C. (1990) Patterns of X chromosome inactivation in the Rett syndrome. *Brain Dev*, **12**, 131-135.

PUBLICATIONS

Liste des publications

Bourdon, V., Philippe, C., Labrune, O., Amsallem, D., Arnould, C. and Jonveaux, P. (2001) A detailed analysis of the *MECP2* gene : prevalence of recurrent mutations and gross DNA rearrangements in Rett syndrome patients. *Hum Genet*, **108**, 43-50.

Bourdon, V., Philippe, C., Grandemenge, A., Reichwald, K. and Jonveaux, P. (2001) Deletion screening by fluorescence in situ hybridization in Rett syndrome patients. *Ann Genet*, **44**, 191-194.

Bourdon, V., Philippe, C., Bienvenu, T., Koenig, B., Tardieu, M., Chelly, J. and Jonveaux, P. (2001) Evidence of somatic mosaicism for a MECP2 mutation in females with Rett syndrome: diagnostic implications. *J Med Genet*, **38**, 867-870.

Gilgenkrantz, S. and Bourdon, V. (2001) *MECP2* et retard mental. *Med sciences*, **17**, 810-812.

Philippe, C., Arnould, C., Bourdon, V., Grégoire, M., Berger, R. and Jonveaux, P. (2000) Un nouveau gène de fusion dans les leucémies aiguës promyélocyaires. *Med sciences*, **16**, 575-577.

Arnould, C., Philippe, C., Bourdon, V., Grégoire, M. J., Berger, R. and Jonveaux, P. (1999) The signal transducer and activator of transcription STAT5b gene is a new partner of retinoic acid receptor α in acute promyelocytic-like leukemia. *Hum Mol Genet*, **8**, 1741-1749.

ORIGINAL INVESTIGATION

Violaine Bourdon · Christophe Philippe
Orienne Labrune · Daniel Amsallem · Cécile Arnould
Philippe Jonveaux

A detailed analysis of the *MECP2* gene: prevalence of recurrent mutations and gross DNA rearrangements in Rett syndrome patients

Received: 25 July 2000 / Accepted: 25 September 2000 / Published online: 13 December 2000

© Springer-Verlag 2000

Abstract Mutations in the X-linked methyl-CpG-binding protein 2 gene (*MECP2*) have been found to be a cause of Rett syndrome (RTT). In order to provide further insights into the distribution and the spectrum of mutations, we investigated, in addition to the whole coding sequence, a phylogenetically conserved sequence within the 3' untranslated region (3' UTR) of the *MECP2* gene for 55 sporadic RTT, including 47 typical and 8 nonclassical cases. We have developed an approach based on conformation-sensitive gel electrophoresis, sequence analysis and, for the first time, Southern blot analysis. Mutation detection, including unreported gross DNA rearrangements, was achieved in 79% of classical RTT and 25% of nonclassical RTT patients. The high prevalence of recurrent mutations allows us to propose a molecular diagnosis strategy for RTT.

Introduction

Rett syndrome (RTT, MIM 312750) is a childhood neurodevelopmental disorder that occurs almost exclusively in female children, with an estimated prevalence of 1 in 10,000–15,000 (Rett 1966; Hagberg et al. 1983). Affected girls are considered to have normal development for the first 6–18 months, followed by a period of regression, marked in particular by loss of purposeful hand use and speech. Seizures, autistic features, ataxia, growth retardation with deceleration of head circumference growth and breathing abnormalities often accompany a profound mental handicap. The genetic basis of RTT has been difficult to establish, because most cases are sporadic. How-

ever, mapping studies with rare familial RTT cases have focused attention on Xq28 (Webb et al. 1998), and recently mutations in the X-linked *MECP2* gene were identified in RTT patients (Amir et al. 1999). *MECP2* encodes the methyl-CpG-binding protein 2 (MeCP2), an abundant chromosomal protein that binds specifically to methylated CpG dinucleotides and seems to act as a global transcriptional repressor (Nan et al. 1997). The protein consists of an 85-aminoacid methyl-CpG binding domain (MBD), which is essential for its binding to 5-methylcytosine, and an 104-aminoacid transcriptional repression domain (TRD) which interacts with the co-repressor Sin3A and histone deacetylases, the resulting co-repressor complex inducing core histones deacetylation and transcriptional silencing (Nan et al. 1998).

Wan et al. (1999) reported initially that approximately 50% of a small group of sporadic and familial RTT have missense or truncating mutations in the *MECP2* gene. More recently, studies reported data showing that *MECP2* mutations account for most cases of typical of RTT (up to 80%) (Amir et al. 2000; Bienvenu et al. 2000; Cheadle et al. 2000; De Bona et al. 2000; Huppke et al. 2000; Xiang et al. 2000), suggesting that another major RTT locus is unlikely. So far, different screening techniques have been used to detect *MECP2* mutations, including denaturing gradient gel electrophoresis, single-strand conformation polymorphism analysis, heteroduplex analysis and DNA sequencing.

In the present study, to provide further insights into the distribution and the spectrum of mutations at the *MECP2* locus, we analyzed 47 typical sporadic cases of Rett syndrome and 8 of nonclassical RTT. We used three approaches to scan the *MECP2* gene for mutations: conformation-sensitive gel electrophoresis, sequence analysis and, for the first time, Southern blot analysis in order to detect gross gene rearrangements. We found the disease-causing mutation in 79% of classical RTT and 25% of nonclassical RTT patients, which provides additional support for *MECP2* as the major locus in the pathogenesis of Rett syndrome.

V. Bourdon · C. Philippe · C. Arnould · P. Jonveaux (✉)
Laboratoire de Génétique Médicale, UPRES 952,
e-mail: p.jonveaux@chu-nancy.fr.
Tel.: +33-3-83153771, Fax: +33-3-83153772

O. Labrune · D. Amsallem
Service de pédiatrie,
Centre Hospitalier Universitaire Besançon, France

Materials and methods

Patients and DNA samples

From the 55 sporadic RTT of our series, 43 patients were referred to our laboratory through the French Rett Syndrome Association and 12 were ascertained from the Neuropediatrics unit of the Centre Hospitalier Universitaire de Besançon, France (Dr. Amsellem). We studied 47 patients with classical sporadic RTT, including a pair of monozygotic twins. All these patients were diagnosed according to the international criteria based on the Rett Syndrome Diagnostic Criteria Work Group (Trevvarthen et al. 1988). We also investigated 8 female patients who had some features of sporadic RTT but who did not fulfill all the diagnostic criteria and were considered as nonclassical cases. Blood samples were obtained from patients and available parents after informed consent. Total genomic DNA was extracted using the Nucleon BACC3 kit (Amersham Life Technologies) according to the manufacturer's protocol.

Mutation detection

Conformation-sensitive gel electrophoresis and sequence analysis

Primer pairs described by Amir et al. (1999) and modified by Wan et al. (1999) were used to amplify, from genomic DNA, the three MECP2-coding exons and flanking intronic sequences in eight overlapping fragments of 300–400 bp. Standard PCR was carried out in a 50- μ l reaction volume containing 100 ng genomic DNA, 1 μ M primers, 250 μ M dNTP, 5 μ l reaction buffer (Eurobio; 670 mM TrisHCl pH 8.8; 160 mM $(\text{NH}_4)_2\text{SO}_4$; 0.1% Tween 20), 1 mM MgCl_2 (Eurobio) and 2 U EurobioTaq DNA polymerase. Cycling parameters were 95°C for 5 min, followed by 35 cycles of 95°C for 1 min, (T_m) for 1 min, 72°C for 1 min, and a final step of extension at 72°C for 10 min. The (T_m) varies from 55°C to 70°C for the three exons.

The amplified products were analyzed by conformation-sensitive gel electrophoresis (CSGE). Briefly, after denaturation at 98°C for 5 min, products were allowed to reanneal for 60 min at 68°C and electrophoresed at 450–500 V overnight, on a conformation-sensitive semi-denaturant polyacrylamide gel to resolve heteroduplexes (adapted from Ganguly et al. 1994; Philippe et al. 1997).

When CSGE variants were detected, products were purified with the QIAGEN QIAquick gel extraction kit and cycle sequenced using the ABI PRISM big dye terminator cycle sequencing ready reaction kit (Perkin Elmer). Cycle sequencing products were purified on DyeEx spin kit (QIAGEN), sequenced on both strands on an ABI 310 automatic sequencer and analyzed visually by comparison between the mutated sequence and a normal control sequence.

For patients without any CSGE variants, we sequenced the whole coding sequence and flanking intronic sequences of the MECP2 gene.

For the missense mutations never reported so far in the literature, and when nucleotide substitution generated or abolished a restriction enzyme site, PCR-RFLP was performed.

Southern blot analysis

Genomic DNA 10 μ g was digested with 20 U of several restriction enzymes (BamHI, BglII, EcoRI, HindIII, NcoI, PstI, SacI, New England Biolabs), electrophoresed overnight on a 0.8% agarose gel, and transferred to a nylon membrane (Hybond N⁺, Amersham) by upward capillary transfer. After a prehybridization step at 65°C for at least 2 h in hybridization buffer (NaCl 750 mM, heparin 50 μ g/ml, SDS 1%, dextran sulfate 5%, salmon sperm DNA 50 mg/l), membranes were hybridized at 65°C for 16 h with the DNA probe labeled with [α -³²P]dCTP by a random primed reaction, washed several times at variable stringency (SSC 2X at room

Table 1 Additional primers used in our study

Primers	Sequences
PolyA10F	5'-GCTGAGTATTCTTTGCTTTTCAC-3'
PolyA10R	5'-TAGCCAAATGACTGAAAGCAC-3'
MEC1	5'-TGTTAGGGCTCAGGGAAGAAAAG-3'
3BR	5'-CTGCACAGATCGGATAGAAGAC-3'
3CF	5'-GGCAGGAAGCGAAAAGCTGAG-3'
3ER	5'-CGGTAAGAAAACATCCCCAA-3'
3BF	5'-AACCACCTAAGAAGCCCCAAA-3'
3DR	5'-CTCCCTCCCCTCGGTGTTTG-3'
2BF	5'-AGCCCGTGCAGCCATCAGCC-3'
3AR	5'-CTTCCCAGGACTTTTCTCCA-3'

temperature, SSC 2X+SDS 0.1% at 65°C, and SSC 0.1X at 65°C), and exposed to X-ray films at -80°C for 2–4 days.

The probe was obtained by RT-PCR with the primer pairs MEC1/3BR and 3CF/3ER (Table 1), which allow amplification of the whole coding region of the MECP2 gene. The reverse transcription was realized on 1 μ g of total RNA isolated from blood using trizol solution (Gibco-Brl), with the Superscript II reverse transcriptase kit (Gibco-Brl) using 500 ng of random oligohexamer primers according to the manufacturer's recommendations.

X-Chromosome inactivation assay

To determine the pattern of X-chromosome inactivation for the patient with R49X mutation, we performed a gene methylation assay at the fragile X locus, as previously described (Carrel and Willard 1996).

Results

We identified a mutation into the MECP2 gene in 37 of 47 classical RTT patients (79%), and in 2 (25%) of 8 cases with suggestive but nondiagnostic features of RTT (atypical RTT). Therefore, this represents a total of 39 patients with mutations in the MECP2 gene. The 39 genomic alterations reported in our series correspond to 18 different mutations, which can be divided into four categories: 5 nonsense mutations, 5 missense mutations, 7 frameshift mutations and 1 splice site mutation.

All missense mutations were shown to have arisen de novo, which is consistent with their pathogenicity. Four missense mutations (R106W, T158M, R306C and P302R) have already been described elsewhere (Amir et al. 1999, 2000; Wan et al. 1999; Bienvenu et al. 2000; Cheadle et al. 2000; Huppke et al. 2000; Xiang et al. 2000), and the first three were recurrent in our study (R106W in 4 cases, T158M in 3 cases and R306C in 2 cases). All missense mutations are located in a functional domain (MBD or TRD), arguing for their pathogenicity. The G161W mutation has not been described previously. As the hydrogen residue of the glycine is changed into a double cycle, one can speculate that this missense mutation creates a sterical bulk, which disturbs the conformation of the MeCP2 protein. The G residue belongs to the MBD consensus sequence defined by a multiple sequence alignment encompassing the MBDs from the five members of the MBD

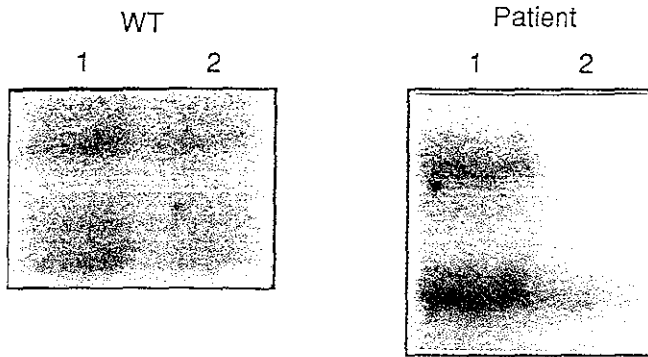


Fig. 1 X-chromosome inactivation study for the patient with the R49X mutation, showing a skewed pattern (WT wild type, 1 PCR on undigested DNA, 2 PCR after Hpa II digest)

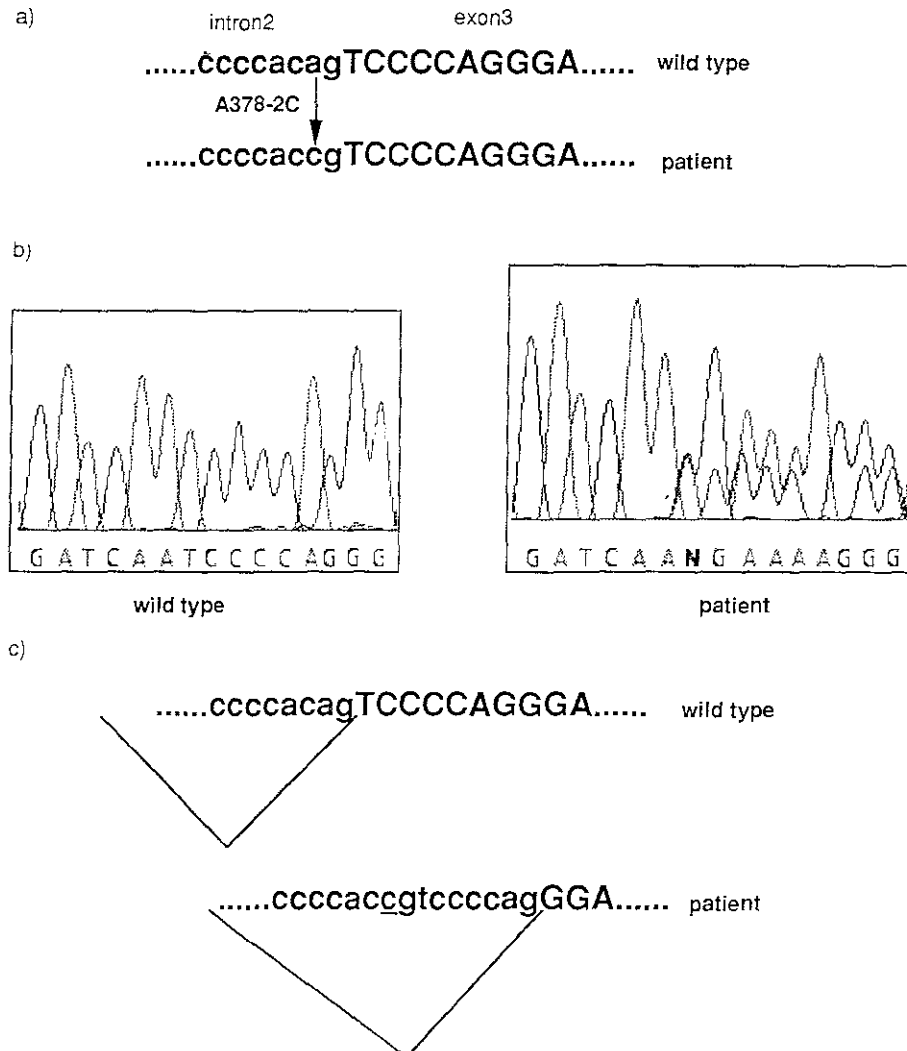
family (Wakefield et al. 1999). Also, the sequence analysis of both parents confirmed the *de novo* occurrence of this mutation. Moreover, the substituted glycine is conserved in the *MECP2* gene from mammals to *Xenopus laevis*, and direct sequencing of 100 control chromosomes

from unrelated unaffected individuals has never revealed this aminoacid substitution at position 161. This mutation can be detected by PCR-RFLP, since it abolishes a BsrI restriction site.

Two silent polymorphisms C582T (S194) and C897T (T299) that have been described previously (Bienvenu et al. 2000; Cheadle et al. 2000) were also detected; the C897T substitution was also present in the patient's mother. We also characterized an unreported variant outside the coding region, G1470A located 9 bp after the stop codon. All these variants have been found in patients with a nonsense mutation.

In our study, two mutational mechanisms, nucleotide substitution and duplication / deletion led to the generation of truncated proteins. Nucleotide substitutions create a premature stop codon in 17 unrelated cases divided into five different nonsense mutations. Among these five mutations, four are recurrent (R168X, R255X, R270X and R294X) and have been already reported (Amir et al. 1999, 2000; Wan et al. 1999; Bienvenu et al. 2000; Cheadle et al. 2000; De Bona et al. 2000; Huppke et al. 2000; Xiang et al. 2000). We identified an early-truncating mutation,

Fig. 2a–c Characterization of the splice site mutation (A378–2C) at the RNA level. **a** The A→C transversion (arrow), at the acceptor splice site of intron 2 was identified by sequence analysis on genomic DNA. **b** Sequencing of the RT-PCR-product with the primers 2BF/3AR used to amplify a fragment of 526 bp surrounding the junction between exon 2 and exon 3. **c** The use of a cryptic acceptor splice site 7 bp within exon 3 leads to a 7-bp deletion in the resulting mRNA. The mutated nucleotide is *underlined*



R49X, which is located upstream from the MBD. The pattern of X-chromosome inactivation (XCI) in peripheral blood leukocyte DNA in this patient is skewed (Fig. 1).

We identified a splicing mutation (A378-2C transversion) that affects the acceptor splice site of exon 3. We studied the patient's *MECP2* mRNA to determine the consequence of this mutation. To do so, we used the primer pair 2BF/3AR (Table 1) to amplify, by RT-PCR, a fragment of 526 bp including the junction between exon 2 and exon 3. The sequence analysis of this RT-PCR product revealed two populations of cDNA, the normal population and a population with a 7-bp deletion. Therefore, the spliceosome uses the next cryptic acceptor splice site 7 bp further on in exon 3, leading to a 7-bp deletion, which causes a frameshift-derived premature stop codon (Figure 2). Splicing mutations at the same acceptor site were reported with another nucleotide change (A378-2G) twice in two different studies (Huppke et al. 2000; Amir et al. 2000), but without being entirely characterized at the mRNA level.

Frameshift mutation due to single nucleotide deletion was identified in one case. The 439 delG creates a TAA stop codon after 13 missense aminoacids in the MBD. A 5-bp duplication (747dupCGCCC) introduces a premature stop codon 40 aminoacids further within the TRD. Interestingly, four different but overlapping deletions were identified in a portion of 57 bp located at the 3' end of the *MECP2* gene. These deletions consist of a 31-bp deletion and of three different 43-bp deletions. All these deletions led to the same premature stop codon (P403X), and one-fifth of the protein (83 aa) is missing in the resulting proteins.

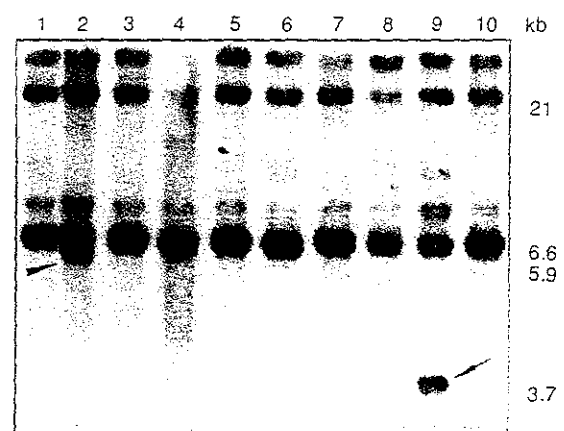
The third exon of the *MECP2* gene contains a large (>8.5 kb) 3' UTR, with several polyadenylation sites that enable the generation of multiple transcripts of different lengths (Coy et al. 1999). Sequence comparison between the human and mouse homologues revealed several blocks of very high homology separated by less conserved sequences. Interestingly, *MECP2* is highly expressed in fetal brain, where the 10.1-kb transcript containing the longest 3' UTR is the predominant isoform. We analyzed 300 bp around the 10.1-kb transcript polyadenylation signal in 34 patients by CSGE and found three variants (data not shown). The sequence analysis revealed for the three variants an insertion of a cytosine in the poly-C tract within this region. We performed the same study on 100 control women (200 X chromosomes), and found the same insertion in 7 cases, suggesting a neutral and relatively frequent (3.5%) polymorphic variant. In addition, comparison of the 3' UTR of the human and mouse *MECP2* genes revealed that this variant falls in a small region of the fragment surrounding the polyadenylation signal, which is not conserved between the two species.

Southern blot analysis of the *MECP2* gene

Using Southern blot analysis, two additional rearrangements were identified (patients 2 and 9) (Fig. 3). Thanks

to a PCR approach with the primer pair 3BF/3DR (Table 1) on the patient from whom the DNA was taken, we could delimit the first rearrangement. It corresponds to a 436-bp deletion (736del436 bp), which eliminates two-thirds of the TRD and creates a premature stop codon 3 missense aminoacids further. The 5' and the 3' ends of this deletion are flanked by a 6-bp direct repeat that may have been involved in the deletion process (Fig. 4). For the second one, we confirmed the rearrangement using seven different restriction enzymes. At the present time, we have not yet found direct evidence that this rearrangement results in a truncated protein, but this seems the most likely scenario.

a) Bgl II digest



b) Bam HI digest

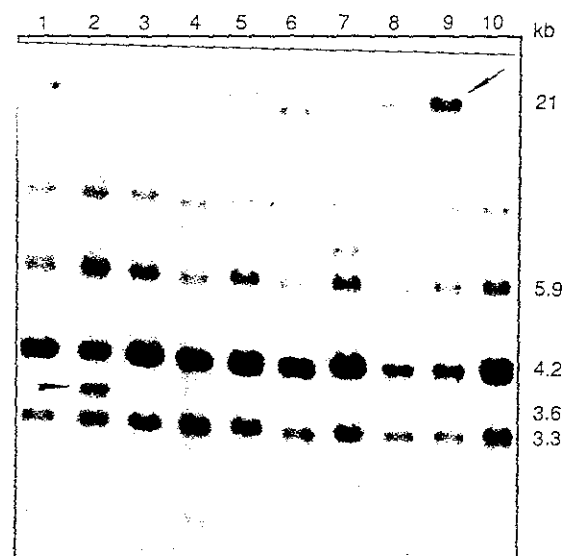
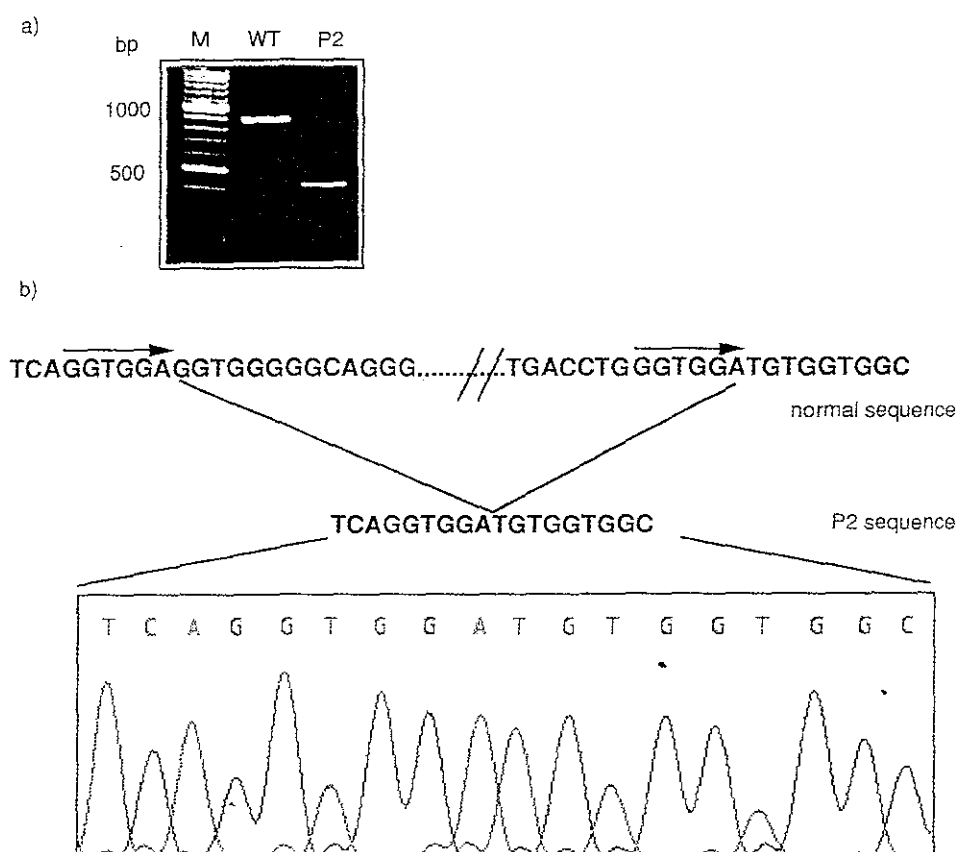


Fig. 3a,b *MECP2* rearrangements detected by Southern blotting. At least three different restriction enzymes (BamHI, BglII and EcoRI) were used to screen RTT patients for gross DNA rearrangement. The results obtained with two restriction enzymes **a** BglII and **b** BamHI for ten patients are shown. Patient 2 presents a junction fragment with BglII and BamHI (arrowhead). Patient 9 shows a junction fragment with BglII, and a more intense fragment with BamHI (arrow)

Fig. 4a, b Characterization of the rearrangement in patient 2 at DNA level. **a** Fragment amplified with the primer pair 3BF/3DR surrounding the deletion after purification with the QIAquick gel extraction kit (*M* 100 bp DNA molecular weight marker, *WT* wild type, *P2* patient 2). **b** Sequencing of the deleted fragment. The *arrows* show the 6-bp direct repeats most probably involved in the deletion process. The two *slashes* and the *dotted line* indicate a portion of the deleted sequence that is not shown



Discussion

In this paper, we report a detailed analysis of the *MECP2* gene in a large number of Rett syndrome patients. Besides conventional mutation screening procedures with CSGE and direct sequencing of the whole coding sequence, Southern blot analysis allowed us to detect unreported larger DNA rearrangements within the *MECP2* gene. In addition, we extended the search for mutations within an evolutionarily conserved region of the *MECP2* 3' UTR.

Altogether, we identified an heterogeneous spectrum of disease-causing *MECP2* mutations in a high proportion (79%) of classical RTT patients. As previously described, we detected mutations in nonclassical RTT (25%) broadening the range of phenotypes related to RTT. The mutations reported in our study include two types of DNA modifications: nucleotide substitutions (72%) and deletions / duplication (28%). These alterations are scattered all along the *MECP2* gene. During this study, we identified 18 different mutations with two novel mutations (G161W and R49X). So far, early truncation in the *MECP2* gene associated with RTT has been described once in the literature (Kim et al. 2000). Clinical evaluation of the girl with the R49X early truncation showed all criteria of classical RTT except for epilepsy. She is unable to walk unsupported. These mutations are likely to lead to a complete loss of function of the MeCP2 protein. More-

over, the nonsense-mediated mRNA decay reduce mRNA abundance decreasing the synthesis of truncated protein (Maquat 1996; Culbertson 1999). The low frequency of early truncations in the *MECP2* gene should suggest a deleterious effect. The nonrandom XCI found in the patient may have contributed to protect the patient against the deleterious effect of the R49X mutation.

In addition, and for the first time, using the *MECP2* cDNA probe we detected two gross DNA rearrangements extending the spectrum of *MECP2* mutations in RTT. Such rearrangements have commonly been observed in patients with other X-linked diseases as Duchenne muscular dystrophy. Smaller recurrent deletions of 43 bp and 31 bp within the same region between bp 1156 and 1164 were identified; these deletions do not affect the two functional domains but result in the loss of one fifth of the protein. Other small deletions (from 7 to 170 bp) have recently been described in the same region (Amir et al. 2000; Bienvenu et al. 2000; Cheadle et al. 2000; De Bona et al. 2000; Huppke et al. 2000; Xiang et al. 2000), supporting the idea that this region is certainly a hot spot for deletions, notably around the nucleotides 1156-1164, where 56% of reported deletions occur. Moreover, as it has already been reported (Bienvenu et al. 2000; De Bona et al. 2000), we found CCACC direct repeats at or near the joining site in the four deletions located at the 3' end of the *MECP2* gene reported in this paper. Interestingly, it has been shown that the deletion of the carboxy-terminal

Table 2 Summary of all MECP2 mutations already associated with Rett syndrome. This table includes mutations published so far together with mutations reported in our study, which are *underlined* [CRIR CoRepressor Interacting Region (Nan et al 1998), CTS C-terminal segment]

Mutations	Location	Domain	Nucleotide change ^a	Aminoacid change	Restriction site ^b	Recurrence ^c	References ^d
Missense	Exon 2	MBD	C291A	D97E	None	1	(7)
	Exon 2	MBD	C301A	P101T	None	1	(4)
	Exon 2	MBD	C302A	P101H	Nla III(+)	1	(4)
	Exon 2	MBD	C302A	P101L	None	1	(4)
	Exon 2	MBD	G317A	R106Q	None	1	(3)
	<u>Exon 2</u>	<u>MBD</u>	<u>C316T</u>	<u>R106W</u>	<u>Nla III(+)</u>	(12+4)	<u>Our study</u> , (1), (2), (4), (5), (6), (7)
	Exon 3	MBD	C397T	R133C	None	8	(1), (2), (4), (5), (7)
	Exon 3	MBD	G398T	R133L	None	1	(9)
	Exon 3	MBD	C401G	S134C	None	2	(4), (5)
	Exon 3	MBD	C455G	P152R	Nla IV (-)	5	(2), (4), (5)
	Exon 3	MBD	T463A	F155I	None	1	(7)
	Exon 3	MBD	T464C	F155S	Tfi(+)	2	(1), (2)
	Exon 3	MBD	C468G	D156E	None	1	(2)
	<u>Exon 3</u>	<u>MBD</u>	<u>C473T</u>	<u>T158M</u>	<u>Nla III(+)</u>	(15+3)	<u>Our study</u> , (1), (2), (3), (4), (5), (7)
	<u>Exon 3</u>	<u>MBD</u>	<u>G481T</u>	<u>G161 W</u>	<u>Bsr I(-)</u>	1	<u>Our study</u>
	Exon 3	TRD	C674G	P225R	None	4	(2), (4), (9)
	Exon 3	TRD	C904G	P302A	None	1	(5)
	<u>Exon 3</u>	<u>TRD</u>	<u>C905G</u>	<u>P302R</u>	<u>Bse RI(+)</u>	2	<u>Our study</u> , (3)
	Exon 3	TRD	C905T	P302L	None	1	(4)
	<u>Exon 3</u>	<u>TRD</u>	<u>C916T</u>	<u>R306C</u>	<u>Hha I(-)</u>	(16+2)	<u>Our study</u> , (2), (3), (4), (5), (6), (7)
	Exon 3	TRD	G917A	R306H	Hha I(-)	1	(4)
	Exon 3	TRD	C964G	P322A	Dra II(+)	1	(3)
	Exon 3	TRD	C965T	P322L	Alu I(+)	1	(5)
	Exon 3	CTS	A1461C	X487C	Fnu4HI (+)	1	(3)
Nonsense	Exon 2		C129T	Q19X	Bpm I	1	(8)
	Exon 2		<u>C146A</u>	<u>S49X</u>	<u>None</u>	1	<u>Our study</u>
	Exon 3	MDB	C423G	Y141X	None	2	(2), (9)
	<u>Exon 3</u>	<u>CRIR</u>	<u>C502T</u>	<u>R168X</u>	<u>Hph I(+)</u>	(38+3)	<u>Our study</u> , (2), (3), (4), (5), (6), (7), (9)
	Exon 3	CRIR	C508T	Q170X	None	1	(2)
	Exon 3	CRIR	A592T	R198X	Mae II(+)	1	(3)
	<u>Exon 3</u>	<u>TRD</u>	<u>C763T</u>	<u>R255X</u>	<u>None</u>	(24+6)	<u>Our study</u> , (1), (2), (3), (4), (5), (6), (7), (9)
	Exon 3	TRD	A766T	K256X	None	1	(2)
	<u>Exon 3</u>	<u>TRD</u>	<u>C808T</u>	<u>R270X</u>	<u>Nla IV(-)</u>	(18+3)	<u>Our study</u> , (2), (3), (4), (5), (9)
	<u>Exon 3</u>	<u>TRD</u>	<u>C880T</u>	<u>R294X</u>	<u>None</u>	(17+4)	<u>Our study</u> , (2), (3), (4), (5), (7), (9)
Splicing site	<u>Intron 2</u>	<u>MBD</u>	<u>A378-2C</u>	<u>L138X</u>	<u>Bsl I(+)</u>	1	<u>Our study</u>
	Intron 2	MBD	A378-2G		Nla IV (+)	2	(2), (5)
Deletions/ insertions	Exon 3	MBD	258delCA		None	1	(5)
	Exon 3	MBD	407del507+ins8			1	(4)
	Exon 3	MBD	411delG	L138X	None	1	(6)
	Exon 3	MBD	431delA		None	1	(9)
	<u>Exon 3</u>	<u>MBD</u>	<u>439delG</u>	<u>V159X</u>	<u>None</u>	1	<u>Our study</u>
	Exon 3	TRD	654del4bp		None	1	(4)
	Exon 3	TRD	706delG		None	1	(2)
	<u>Exon 3</u>	<u>TRD</u>	<u>730del436bp</u>	<u>P391X</u>		1	<u>Our study</u>
	Exon 3	TRD	803delG		None	1	(2)
	Exon 3	TRD	806delG	V288X	Nla IV(-)	4	(4), (6)
	Exon 3	CRIR	620insT	E235X	Bsg I(-)	2	(1), (2)
	Exon 3	TRD	677insA		None	1	(3)
	<u>Exon 3</u>	<u>TRD</u>	<u>747dup5bp</u>	<u>V288X</u>	<u>Fau I(+)</u>	1	<u>Our study</u>
	Exon 3	CTS	1127ins10bp			1	(7)
Exon 3	CTS	1194insT		Bspw I(+)	1	(3)	
Exon 3	CTS	1195insT		None	1	(2)	

Table 2 (continued)

Mutations	Location	Domain	Nucleotide change ^a	Aminoacid change	Restriction site ^b	Recurrence ^c	References ^d
	Exon 3	CTS	1364insC		None	1	(5)
	Exon 3	CTS	1150del43bp	P403X		1	Our study
	Exon 3	CTS	1157del31bp	P403X		1	Our study
	Exon 3	CTS	1158del43bp	P403X		1	Our study
	Exon 3	CTS	1164del43bp	P403X		2	Our study
	Exon 3	CTS	1096-1165del 7 to 170bp			21	(2), (3), (4), (5), (7), (9)

^aNucleotides were numbered from the first base of the translation initiation ATG codon (GenBank accession No AF030876 for the complete sequence and accession No 4826829 for the mRNA sequence) according to Wan et al. (1999)

^b(+) generated, (-) abolished

^cex: (12+4) this mutation has been reported 12 times in the literature and was found again 4 times in our study

^d(1) Amir et al. (1999); (2) Amir et al. (2000); (3) Bienvenu et al. (2000); (4) Cheadle et al. (2000); (5) Huppke et al. (2000); (6) Wan et al. (1999); (7) Xiang et al. (2000); (8) Kim et al. (2000); (9) De Bona et al. (2000)

63 aminoacids of MeCP2 protein impairs its binding with the nucleosomal DNA during the transcription regulation process (Chandler et al. 2000).

So far, the search for mutations has been limited to the coding sequence of MECP2. The 3' UTRs may exert post-transcriptional regulation enabling rapid modulation of protein levels in response to different developmental or physiological changes (Decker and Parker 1995). The strong conservation of several blocks within the 3' UTRs of MECP2 genes in different species indicates that these sequences are under selection pressure and therefore functionally important. Interestingly, MECP2 is highly expressed in fetal brain, where the 10.1-kb transcript containing the longest 3' UTR is the predominant isoform. Using CSGE analysis of a phylogenetically conserved 300-bp sequence around the 10.1-kb transcript-specific polyadenylation site, we detected only polymorphic variants, which is an argument suggesting that disease-causing mutations in this region will be a rare event in RTT patients.

Including the data reported so far in the literature (Amir et al. 1999, 2000; Wan et al. 1999; Bienvenu et al. 2000; Cheadle et al. 2000; De Bona et al. 2000; Huppke et al. 2000; Xiang et al. 2000), the spectrum of mutations into the MECP2 gene associated with RTT now encompasses missense (34%), nonsense (46%), and frameshift (20%) mutations. However, the occurrence of mutations mainly in exon 3 (89%) and the multiple recurrence of specific mutations (Table 2) point to true mutational hot spots that could propose molecular diagnosis strategies of RTT. Indeed, three missense mutations (R106W, T158M, and R306C) represent about a quarter (23%) of all mutations in our study, and between 16% and 32% in studies reported in the literature. In the same way, four nonsense mutations (R168X, R255X, R270X and R294X) are very frequent in reported studies. These seven mutations represent 64% of all mutations in our study, 72% in Amir and Wan's studies, 64% in Xiang's study, 60% in Cheadle's study, 58% in Huppke's study, 62% in Amir et al.'s (2000), and 63% in Bienvenu's study. Among these seven

mutations, five (R106W, T158M, R306C, R168X and R270X) were detectable by PCR-RFLP, which represents a more rapid alternative to direct sequencing. Besides these substitutions, the remaining most frequent diseases causing mutations are deletions. The small deletions clustered in a 200-bp region could be identified by PCR, and the deleted alleles are easily resolved from the normal allele by electrophoresis on a 2% agarose gel. For the larger rearrangements, only Southern blot analysis can readily detect them, because even with long-distance PCR (Cheadle et al. 2000), the likelihood that a primer will be located in the deletion is not negligible. So, to detect at least 70% of Rett syndrome-causing mutations, we propose beginning with PCR-RFLP for the more frequent missense or nonsense mutations and a simple PCR for small deletions. In a second step, CSGE and Southern blot analysis can be realized, allowing screening of all the MECP2 gene coding sequence. To complete the mutation detection, sequencing of the whole coding sequence and flanking intronic sequences of the three exons of the MECP2 gene will be achieved.

We failed to identify any mutation in 21% of cases with typical RTT. The data reported initially by Amir et al. (1999) found MECP2 mutations in 25% of RTT patients using only one mutation detection technique (CSGE). We now report on a mutation detection rate of 79%, and a rate of 80% has recently been published (Cheadle et al. 2000). Mutation screening in the promoter region and the use of techniques to detect the loss of one allele (as FISH analysis) might reveal additional mutations. Altogether, these results suggest that MECP2 is the major locus in classical RTT, lowering the likelihood of the possible involvement of a minor second X-linked locus.

Acknowledgements We would like to thank all the patients and their families for their participation, all doctors who provided us with blood samples. C. Hainaut, B. Leotard, M. Franck and A. Leuret for their technical contribution and l'Association Française du Syndrome de Rett, l'Association Française contre les Myopathies and the Ministère de l'Education Nationale, de la Recherche et de la Technologie (VB) for their financial support.

References

- Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet* 23:185-188
- Amir RE, Van den Veyver IB, Schultz R, Malicki DM, Tran CQ, Dahle EJ, Philippi A, Timar L, Percy AK, Motil KJ, Lichtarge O, Smith EO, Glaze DG, Zoghbi HY (2000) Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann Neurol* 47:670-679
- Bienvenu T, Carrie A, de Roux N, Vinet MC, Jonveaux P, Couvert P, Villard L, Arzimanoglou A, Beldjord C, Fontes M, Tardieu M, Chelly J (2000) *MECP2* mutations account for most cases of typical forms of Rett syndrome. *Hum Mol Genet* 9:1377-1384
- Carrel L, Willard HF (1996) An assay for X inactivation based on differential methylation at the fragile X locus, *FMR1*. *Am J Med Genet* 64:27-30
- Chandler SP, Guschin D, Landsberger N, Wolffe AP (1999) The methyl-CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. *Biochemistry* 38:7008-7018
- Cheadle JP, Gill H, Fleming N, Maynard J, Kerr A, Leonard H, Krawczak M, Cooper DN, Lynch S, Thomas N, Hughes H, Hulten M, Ravine D, Sampson JR, Clarke A (2000) Long-read sequence analysis of the *MECP2* gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Hum Mol Genet* 9:1119-1129
- Coy JF, Sedlacek Z, Bachner D, Delius H, Poustka A (1999) A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3' untranslated region of the methyl-CpG-binding protein 2 gene (*MeCP2*) suggests a regulatory role in gene expression. *Hum Mol Genet* 8:1253-1262
- Culbertson MR (1999) RNA surveillance unforeseen consequences for gene expression, inherited genetic disorders and cancer. *Trends Genet* 15:74-80
- De Bona C, Zappella M, Hayek G, Meloni I, Vitelli F, Bruttini M, Cusano R, Loffredo P, Longo I, Renieri A (2000) Preserved speech variant is allelic of classic Rett syndrome. *Eur J Hum Genet* 8:325-330
- Decker CJ, Parker R (1995) Diversity of cytoplasmic functions for the 3' untranslated region of eukaryotic transcripts. *Curr Opin Cell Biol* 7:386-392
- Ganguly A, Rock MJ, Prockop DJ (1993) Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. *Proc Natl Acad Sci USA* 90:10325-10329
- Hagberg B, Aicardi J, Dias K, Ramos O (1983) A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* 14:471-479
- Huppke P, Laccone F, Kramer N, Engel W, Hanefeld F (2000) Rett syndrome: analysis of *MECP2* and clinical characterization of 31 patients. *Hum Mol Genet* 9:1369-1375
- Kim SJ, Cook EJ (2000) Novel *de novo* nonsense mutation of *MECP2* in a patient with Rett syndrome. *Hum Mutat* 15:382-383
- Maquat LE (1996) Defects in RNA splicing and the consequence of shortened translational reading frame. *Am J Hum Genet* 59:279-286
- Nan X, Campoy FJ, Bird A (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 88:471-481
- Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393:386-389
- Philippe C, Porter D, Emerton ME, Wells DE, Simpson AHRW, Monaco AP (1997) Mutation screening of the *EXT1* and *EXT2* genes in patients with hereditary multiple exostoses. *Am J Hum Genet* 61:520-528
- Rett A (1966) Über ein eigenartiges himatrophisches Syndrom im Kindesalter. *Wien Med Wochenschr* 11:723-726
- Trevarthen E, Moser HW, Diagnostic Criteria Working Group (1988) Diagnostic criteria for Rett syndrome. *Ann Neurol* 23:425-428
- Wakefield RI, Smith BO, Nan X, Free A, Soteriou A, Uhrin D, Bird AP, Barlow PN (1999) The solution structure of the domain from MeCP2 that binds to methylated DNA. *J Mol Biol* 291:1055-1065
- Wan M, Lee SS, Zhang X, Houwink MI, Song HR, Amir RE, Budden S, Naidu S, Pereira JL, Lo IF, Zoghbi HY, Schanen NC, Francke U (1999) Rett syndrome and beyond: recurrent spontaneous and familial *MECP2* mutations at CpG hotspots. *Am J Hum Genet* 65:1520-1529
- Webb T, Clarke A, Hanefeld F, Pereira JL, Rosenbloom L, Woods CG (1998) Linkage analysis in Rett syndrome families suggests that there may be a critical region at Xq28. *J Med Genet* 35:997-1003
- Xiang F, Buervenich S, Nicolao P, Bailey ME, Zhang Z, Anvret M (2000) Mutation screening in Rett syndrome patients. *J Med Genet* 37:250-255

Letters to the Editor

J Med Genet
2001;38:867-870

Laboratoire de
Génétique Médicale,
CHU Nancy-Brabois,
Avenue du Morvan,
54511 Vandoeuvre les
Nancy, France
V Bourdon
C Philippe
P Jonveaux

Laboratoire de
Génétique et
Physiopathologie des
Retards
Mentaux-ICGM,
Faculté de Médecine
Cochin, Paris, France
T Bienvenu
J Chelly

Site du Neuhof, Centre
de Ressources en
Matière de Déficiences
Sensorielles et de
Polyhandicaps Louis
Braille - Raoul
Clainchard - Auguste
Jacoutôt, Strasbourg,
France
B Koenig

Département de
Pédiatrie, Service de
Neurologie, CHU
Bicêtre, Le Kremlin
Bicêtre, France
M Tardieu

Correspondence to:
Dr Jonveaux,
p.jonveaux@chu-nancy.fr

Evidence of somatic mosaicism for a *MECP2* mutation in females with Rett syndrome: diagnostic implications

Violaine Bourdon, Christophe Philippe, Thierry Bienvenu, Bernadette Koenig, Marc Tardieu, Jamel Chelly, Philippe Jonveaux

EDITOR—Rett syndrome (RTT) (MIM 312750) is an X linked dominant neurodevelopmental disorder that occurs almost exclusively in females. Affected girls are considered to have a normal perinatal period followed by a

period of regression, loss of acquired purposeful manual and speech skills, hand wringing, gait disturbance, and growth retardation.^{1,2}

A gene for RTT has been identified in the Xq28 region which encodes the methyl-CpG

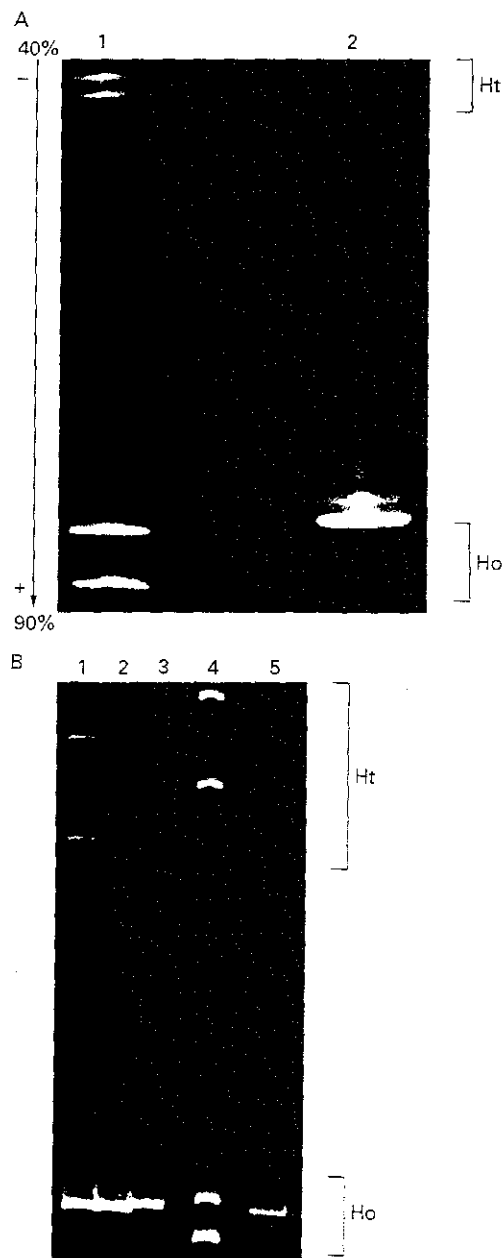


Figure 1 Detection of the two somatic mutations by heteroduplex analysis. (A) DGGE results for case 1. DNA extracted from a lymphoblastoid cell line (lane 1) and a fresh blood sample (lane 2). The homoduplex band corresponding to the deleted allele is missing in DNA extracted from lymphocytes. Ht, heteroduplex; Ho, homoduplex; 40%-90% formamide gradient. (B) CSGE results for case 1 (lanes 1-3) and case 2 (lane 5). Case 1: the somatic mutation is shown on DNA extracted from a fresh blood sample (lane 1); the father (lane 2) and the mother (lane 3) do not carry the 26 bp deletion. Case 2: a somatic deletion was detected on DNA extracted from lymphocytes as indicated by the absence of the deleted homoduplex band (lane 5). Lane 4 depicts a CSGE pattern of a 31 bp deletion localised in the deletion prone region of the *MECP2* gene; the two homoduplex bands are of equal intensity.

binding protein 2 (MeCP2) involved in transcriptional silencing.^{3,4} This disorder most frequently occurs sporadically and results from a de novo mutation, although a few familial cases have been reported. Many studies⁵⁻¹⁶ have shown that the *MECP2* gene is mutated in approximately 80% of patients with classical RTT and the *MECP2* mutation spectrum

includes missense, nonsense, and frameshift mutations, as well as larger rearrangements like deletions encompassing a few hundred bp.¹⁰ The failure to detect *MECP2* mutations in the remaining 20% may indicate the presence of mutations in unexplored regions of the *MECP2* gene, such as regulatory elements or non-coding regions, notably in the new first exon¹⁷ or in an additional RTT locus.

Here, we report for the first time mosaicism for a somatic *MECP2* mutation found in two unrelated females affected with RTT. These two girls were diagnosed according to the international criteria of the Rett Syndrome Diagnostic Criteria Work Group.¹⁸

Case reports

The first patient (case 1) is 13 years old. She suffers from classical Rett syndrome with 7/9 of the necessary criteria, 4/8 of the supportive criteria, and none of the exclusion criteria.¹⁸ More specifically, she had a normal neonatal period and head circumference at birth and a phase of social withdrawal at the age of 12 months when she lost purposeful hand skills and developed stereotypic hand movements, ataxia, and apraxia. She suffered from breathing dysfunction and peripheral vasomotor disturbances. She had severely impaired development but acquired independent walking at the age of 24 months. However, she did not acquire microcephaly or develop epilepsy.

The second patient (case 2) was reported as an atypical case of RTT without any period of regression. Both mental and motor development were very slow. At the age of 4 years, she had acquired microcephaly (-2 SD) and had very limited ambulation, but her hand use was correct without hand wringing movements. She developed epilepsy and progressive scoliosis. She is a placid girl without useful speech but she communicates well by eye movements.

Methods and results

For case 1, an initial study on DNA extracted from a lymphoblastoid cell line by denaturing gradient gel electrophoresis (DGGE) and sequencing showed that she carried a 26 bp deletion starting at position 1165. To confirm this mutation, DNA was extracted from a fresh blood sample and the deletion was assessed by direct sequencing. Surprisingly and despite a careful examination of the sequence, we did not find the 26 bp deletion with DNA extracted from leucocytes. This sample was reanalysed by DGGE and heteroduplexes were detected while the homoduplex corresponding to the deleted band was absent (fig 1A). We confirmed this result by conformation sensitive gel electrophoresis (CSGE) analysis, which showed the heteroduplexes but not the mutant homoduplex (fig 1B). The results obtained from peripheral blood lymphocytes suggested mosaicism for a somatic mutation.

In order to determine the level of mosaicism, we used a semiquantitative approach based on fluorescent PCR. The *MECP2* gene exon 3 portion containing the deletion was PCR amplified, the reverse primer being conjugated

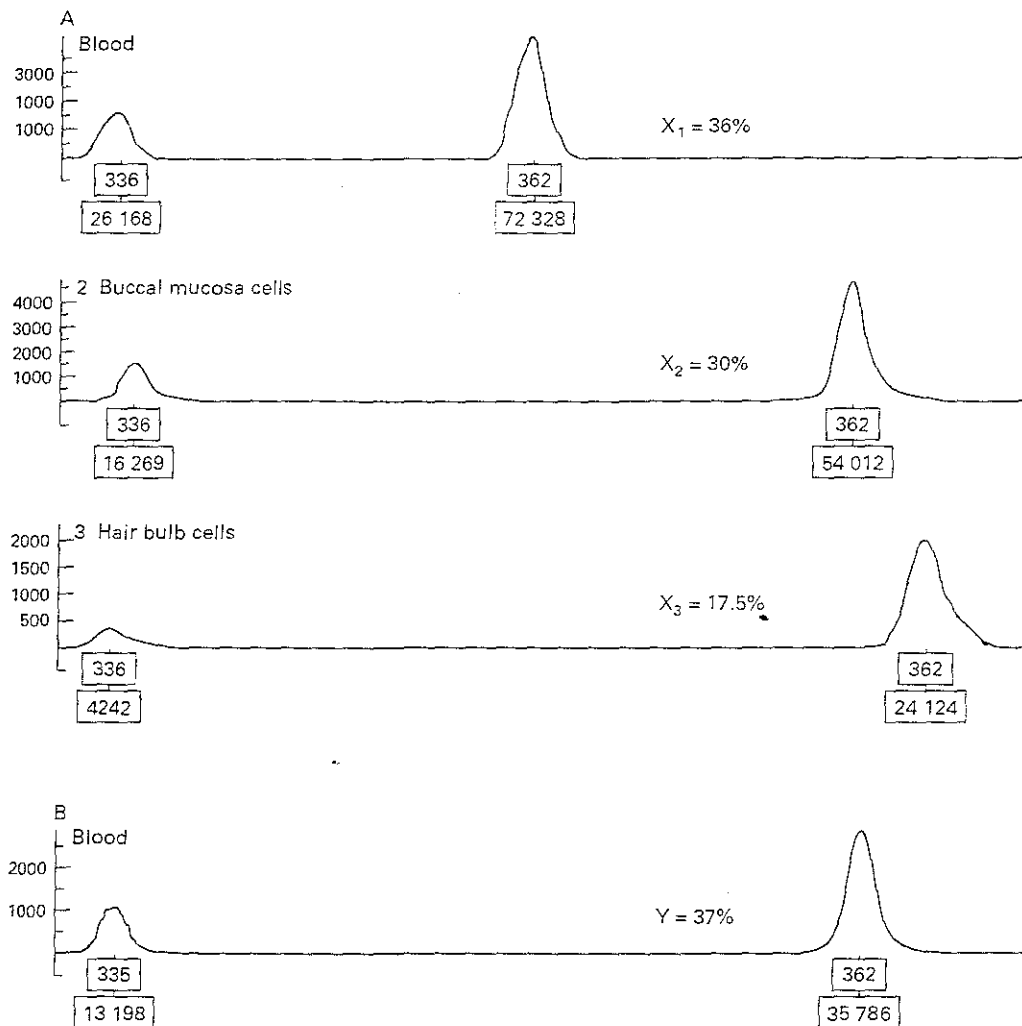


Figure 2 Semiquantitative fluorescent PCR of the somatic mosaicism rate. (A) Case 1. Genotyper traces of the fluorescent PCR products obtained with three different tissues, blood (1), buccal mucosa cells (2), and hair bulb cells (3), shown with the three respective ratios of peak areas (X_1 , X_2 , X_3). For each peak, the fragment size in bp and the peak area calculated by Genescan is indicated. We assumed that the mosaicism rate could be estimated by calculating the ratio between the deleted and the normal peak areas. (B) Case 2. Genotyper trace of a fluorescent PCR product obtained from blood with the ratio of peak areas (Y).

to 6-FAM (6-carboxy-fluorescein). PCR products were analysed on an ABI 310 sequencer and peak areas were generated by ABI Genescan and Genotyper software. The ratio between the deleted and normal peak areas showed that only 36% of lymphocytes harboured the deletion, that is, 18% of X chromosomes bore the 26 bp deletion (fig 2A). This semiquantitative approach confirms that case 1 does have somatic mosaicism for the *MECP2* deletion. The relatively low level of somatic mosaicism could explain the normal sequencing result. Thus, mosaicism was quantified in different tissues. DNA was extracted from buccal mucosa cells¹⁹ and hair bulb cells.²⁰ The level of mosaicism was about the same in buccal mucosa cells (30%) as in lymphocytes, but lower in hair bulbs cells (17.5%) (fig 2A).

Discussion

On the basis of these results, we hypothesised that some patients with RTT may in fact carry a somatic mutation. Small deletions (from 7 to 170 bp) within the region between bp 1096 and 1165 of the *MECP2* gene have been recurrently identified.^{5,7,9,10,12,13,19} They do not affect

the two functional domains but result in the loss of one fifth of the protein. Interestingly, it has been shown that the deletion of the carboxy-terminal 63 amino acids of the MeCP2 protein impairs binding with the nucleosomal DNA during the transcription regulation process.²¹ These recurrent deletions may be the result of palindromic and quasipalindromic sequences within this region, which are believed to form secondary structures that render the region vulnerable to deletions. Therefore, using our fluorescent PCR approach, we reanalysed the 3' region of the *MECP2* gene, between bp 1096 and 1165, in a cohort of 29 patients diagnosed as typical or atypical RTT; for these patients, we failed to detect any mutation using a bidirectional sequencing strategy of the entire *MECP2* coding region. A second somatic mosaicism for a 27 bp deletion was identified in peripheral blood lymphocytes from case 2 with atypical RTT; the mosaicism rate was quantified with our fluorescent approach to be about 37% (fig 2B). We confirmed this result by CSGE analysis (fig 1B).

In both cases, numerical aberrations of the X chromosomes as a cause for the uncommon fluorescent PCR patterns were excluded by the presence of a normal 46,XX karyotype.

These two patients show a similar deletion with an equivalent mosaicism rate in blood, but a distinct clinical presentation. X inactivation study on proband 1 with typical Rett syndrome showed a random pattern of inactivation in the peripheral blood. Although the results have to be extrapolated from the peripheral blood cells, it would suggest that in the brain the majority of mutated X chromosomes may remain active in the girl with classical Rett syndrome. Our results illustrate clearly once again the difficulty in establishing a correlation between genotype and phenotype in RTT.

Recently, a boy with a mosaic mutation has been described.²² To our knowledge, we show for the first time that somatic mosaicism for *MECP2* mutation in girls is not infrequent (two somatic mutations on 102 putative RTT cases studied) and may cause different phenotypes. These clinical and molecular findings suggest that multiple forms of mosaicism (X inactivation mosaicism and somatic mosaicism) may be present in a single patient with RTT. Mosaicism has been documented for chromosomal abnormalities, mitochondrial mutations, triplet repeats,²³ and in a growing number of dominant and recessive X linked gene disorders, such as Duchenne muscular dystrophy,²⁴ haemophilia B,²⁵ Conradi-Hünemann-Happle syndrome,²⁶ and double cortex/lissencephaly syndrome.²⁷ Because a proportion of cells carry the mutation not only in blood but also in tissues deriving from other cell lineages, it must be assumed that the mutation occurred very early during embryogenesis.

Finally, the detection of mosaic mutation depends mainly on the method used for the identification of mutations within the *MECP2* gene. Nowadays, the method of choice for identifying deleterious mutations relies on direct DNA sequencing. The ability of this method to detect mosaic mutations is poor, which is particularly true when the mosaicism rate is low. Our findings underline the need for at least two complementary approaches, such as methods based on heteroduplex analysis and sequencing, for an efficient screening of the *MECP2* gene.

We thank Dr Deblay for critical advice, Dr Florence Rousselet for her technical contribution, and l'Association Française du Syndrome de Rett, l'Association Française contre les myopathies, and the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie for their financial support.

- 1 Rett A. On an unusual brain atrophy syndrome in hyperamonemia in childhood. *Wien Med Wochenschr* 1966;11:723-6.
- 2 Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* 1983;14:471-9.
- 3 Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999;23:185-8.

- 4 Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* 1997;21:471-81.
- 5 Amir RE, Van den Veyver IB, Schultz R, Malicki DM, Tran CQ, Dahle EJ, Philippi A, Timar L, Percy AK, Motil KJ, Lichtarge O, Smith EO, Glaze DG, Zoghbi HY. Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann Neurol* 2000;47:670-9.
- 6 Amano K, Nomura Y, Segawa M, Yamakawa K. Mutational analysis of the *MECP2* gene in Japanese patients with Rett syndrome. *J Hum Genet* 2000;45:213-36.
- 7 Bienvenu T, Carrie A, de Roux N, Vinet MC, Jonveaux P, Couvert P, Villard L, Arzamanoglou A, Beldjord C, Fontes M, Tardieu M, Chelly J. *MECP2* mutations account for most cases of typical forms of Rett syndrome. *Hum Mol Genet* 2000;9:1377-84.
- 8 Buyse JM, Fang P, Hoon KT, Amir RE, Zoghbi HY, Roa BB. Diagnostic testing for Rett syndrome by DHPLC and direct sequence analysis of the *MECP2* gene: identification of several novel mutations and polymorphisms. *Am J Hum Genet* 2000;67:1428-36.
- 9 Cheadle JP, Gill H, Fleming N, Maynard J, Kerr A, Leonard H, Krawczak M, Cooper DN, Lynch S, Thomas N, Hughes H, Hulten M, Ravine D, Sampson JR, Clarke A. Long-read sequence analysis of the *MECP2* gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Hum Mol Genet* 2000;9:1119-29.
- 10 De Bona C, Zappella M, Hayek G, Meloni I, Vitelli F, Brutini M, Cusano R, Loffredo P, Longo I, Renieri A. Preserved speech variant is allelic of classic Rett syndrome. *Eur J Hum Genet* 2000;8:325-30.
- 11 Hampson K, Woods CG, Latif F, Webb T. Mutations in the *MECP2* gene in a cohort of girls with Rett syndrome. *J Med Genet* 2000;37:610-12.
- 12 Huppke P, Laccone F, Kramer N, Engel W, Hanefeld F. Rett syndrome: analysis of *MECP2* and clinical characterization of 31 patients. *Hum Mol Genet* 2000;9:1369-75.
- 13 Obata K, Matsuishi T, Yamashita Y, Fukuda T, Kuwajima K, Horiuchi I, Nagamitsu S, Iwanaga R, Kimura A, Omori I, Endo S, Mori K, Kondo I. Mutation analysis of the methyl-CpG binding protein 2 gene (*MECP2*) in patients with Rett syndrome. *J Med Genet* 2000;37:608-10.
- 14 Wan M, Lee SS, Zhang X, Houwink MI, Song HR, Amir RE, Budden S, Naidu S, Pereira JL, Lo IF, Zoghbi HY, Schanen NC, Francke U. Rett syndrome and beyond: recurrent spontaneous and familial *MECP2* mutations at CpG hotspots. *Am J Hum Genet* 1999;65:1520-9.
- 15 Xiang F, Buervenich S, Nicolao P, Bailey ME, Zhang Z, Anvret M. Mutation screening in Rett syndrome patients. *J Med Genet* 2000;37:250-5.
- 16 Bourdon V, Philippe C, Labrune O, Amsellem D, Arnould C, Jonveaux P. A detailed analysis of the *MECP2* gene: prevalence of recurrent mutations and gross DNA rearrangements in Rett syndrome patients. *Hum Genet* 2001;108:43-50.
- 17 Reichwald K, Thiesen J, Wiehe T, Weitzel J, Stätling WH, Kioschis P, Poustka A, Rosenthal A, Platzer M. Comparative sequence analysis of the *MECP2*-locus in human and mouse reveals new transcribed regions. *Manus Genome* 2000;11:182-90.
- 18 Trevarthen E, Moser HW, and the Diagnostic Criteria Working Group. Diagnostic criteria for Rett syndrome. The Rett Syndrome Diagnostic Criteria Work Group. *Ann Neurol* 1988;23:425-8.
- 19 Philippe C, Porter DE, Emerton ME, Wells DE, Simpson AHRW, Monaco AP. Mutation screening of the *EXT1* and *EXT2* genes in patients with hereditary multiple exostoses. *Am J Hum Genet* 1997;61:520-8.
- 20 Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 1991;10:506-13.
- 21 Chandler SP, Guschin D, Landsberger N, Wolfe AP. The methyl-CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. *Biochemistry* 1999;38:7008-18.
- 22 Clayton-Smith J, Watson P, Ramsden S, Black GCM. Somatic mutation in *MECP2* as a non-fatal neurodevelopmental disorder in males. *Lancet* 2000;356:830-2.
- 23 Bernards A, Gusella JF. The importance of genetic mosaicism in human disease. *N Engl J Med* 1994;331:1447-9.
- 24 Bunyan DJ, Robinson DO, Collins AL, Cockwell AE, Bullman HMS, Whitaker PA. Germline and somatic mosaicism in a female carrier of Duchenne muscular dystrophy. *Hum Genet* 1994;93:541-4.
- 25 Costa J-M, Vidaud D, Laurendreau I, Vidaud M, Fressinaud E, Moisan JP, David A, Meyer D, Lavergne JM. Somatic mosaicism and compound heterozygosity in female hemophilia B. *Blood* 2000;96:1585-7.
- 26 Has C, Bruckner-Tuderman L, Müller D, Floeth M, Folkers E, Donnai D, Traupe H. The Conradi-Hünemann-Happle syndrome (*CDPX2*) and emopamil binding protein: novel mutations, and somatic and gonadal mosaicism. *Hum Mol Genet* 2000;9:1051-5.
- 27 Gleeson JG, Minnerath S, Kuzmicki RJ, Dobyns WB, Young ID, Ross ME, Walsh CA. Somatic and germline mosaic mutations in the doublecortin gene are associated with variable phenotypes. *Am J Hum Genet* 2000;67:574-81.

Deletion screening by fluorescence in situ hybridization in Rett syndrome patients

Violaine Bourdon^a, Christophe Philippe^a, Agnès Grandemenge^a,
Kathrin Reichwald^b, Philippe Jonveaux^{a*}

^a Laboratoire de génétique médicale, EA 3441, CHU Brabois, rue du Morvan, 54511 Vandoeuvre-les-Nancy cedex, France

^b Department of Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr.11, 07745 Jena, Deutschland

Received 11 June 2001; accepted 3 October 2001

Abstract – Mutations in the X-linked methyl-CpG-binding protein 2 (MECP2) gene have been found to be a cause of Rett syndrome (RTT). Mutation screening was based on various techniques including denaturing gradient gel electrophoresis, single-strand conformation polymorphism analysis, heteroduplex analysis, DNA sequencing and recently Southern Blot analysis. Mutation detection was achieved in 80% of typical RTT with a high prevalence of recurrent mutations. In order to provide further insights into the spectrum of MECP2 rearrangements in patients without any point mutation or small deletion/insertion in the coding region MECP2 gene, we screened 25 classical RTT females using fluorescence *in situ* hybridization analysis. No deletion were found in our group, suggesting that MECP2 gross rearrangements are a rare cause of Rett syndrome. © 2001 Éditions scientifiques et médicales Elsevier SAS

Rett syndrome / MECP2 gene / deletion / FISH

1. Introduction

Rett syndrome (RTT, MIM 312750) is a childhood neurodevelopmental disorder which almost affects girls, with an estimated prevalence of 1 in 10 000–15 000 [21, 11]. Patients with classic RTT appear to develop normally until age 6–18 months, at which time they enter a period of developmental regression. The affected girls suffer from gradual loss of speech and purposeful hand use, development of microcephaly, seizures, ataxia, autistic features, intermittent hyperventilation and stereotypic hand movements. Although most cases are sporadic, linkage studies in the rare familial RTT cases pointed to Xq28 as the most likely candidate region [25]. By systematic mutation analysis of genes in Xq28, Amir et al. [2] identified the RTT gene as MECP2, the methyl-CpG-binding protein 2, that is normally involved in transcriptional silencing [17]. Numerous studies have since found various mutations (missense, nonsense,

frameshift) in the coding region of the MECP2 gene, in as many as 80% of patients [1–9, 12, 14–16, 18, 19, 23–26]. Most mutations occur within the two functional domains, the methyl-CpG-binding domain (MBD) and the transcriptional repression domain (TRD) of MeCP2. Moreover, small deletions (about 20–80 bp) focused on a limited region localised after the second functional domain (TRD).

The failure to detect MECP2 mutations in the remaining 20% patients may indicate the presence of mutations in unexplored regions of the MECP2 gene, such as regulatory elements or noncoding regions. Most MECP2 mutations reported so far lead to inactive MeCP2 proteins (premature stop codons or amino-acid changes in the functional domains) arguing for an haplo-insufficiency of the MECP2 gene in RTT. Therefore, the complete loss of one allele might be involved in RTT. In this study, we looked for MECP2 deletion by fluorescence in situ hybridization (FISH) on a panel of 25 girls with classic RTT.

* Correspondence and reprints.

E-mail address: p.jonveaux@chu-nancy.fr (P. Jonveaux).

2. Materials and methods

2.1. Patients

We studied 25 female patients with classical sporadic RTT without any mutations identified in the MECP2 coding region after complete sequencing and Southern blot analysis [6]. They were referred through the French Rett Syndrome Association and were diagnosed according to the Rett Syndrome Diagnostic Criteria Work Group [22]. Blood samples were obtained after informed consent.

2.2. Fluorescence *in situ* hybridization

Metaphase spreads from control females and RTT patient lymphoblastoid cell lines were prepared according to standard procedures. A slides pretreatment by pepsin digest during 10 min (10 µg/mL) was made to eliminate cytoplasm. Chromosomal DNA denaturation was obtained in 70% formamide, 2X SSC, pH 7 at 73°C during 2 minutes. The probe used was the PAC clone 671D9 containing only the MECP2 gene [accession number: AF030876, 20]. 1 µg was labelled with biotin-16-dUTP by nick translation using Bio-Nick kit (Gibco-BRL, Germany). For each experiment, the labelled PAC DNA was used together with a CEP X Spectrum Aqua probe (Vysis, USA) specific to the X chromosome α -satellite region. The probe mixture was denatured for 10 minutes at 73°C and directly added to the slide. Slides were incubated overnight at 37°C, and washed for 2 min with 0,4XSSC/0,3% Nonidet P40, pH 7, and for 15 s with 2XSSC/0,1% Nonidet P40, pH 7. Biotin-labelled probe was detected according to the protocol of the fluorescein detection kit (Oncor-Appligene, France). Finally, slides were counterstained with DAPI I (4'6 diamino-2 phenylindole) (Vysis, USA) and analysed on a Zeiss epifluorescence microscope (Jena, Germany) with a Sensys CCD camera (Photometrics, USA) and IPLab Spectrum Imaging Software (Vysis, USA).

3. Results and discussion

Preliminary studies of the MECP2 coding region using a mutation screening strategy based on conformation-sensitive gel electrophoresis, sequence analysis and Southern blot analysis did not reveal any mutation in 25 classic RTT patients. Therefore, further analysis by FISH was performed in order to exclude large deletions of the RTT region in Xq28 in this



Figure 1. Metaphase chromosome spread shows the 2 control aqua spots on the X centromeres, and the 2 green spots due to the PAC 671d9 probe hybridisation on the 2 MECP2 alleles in Xq28 region. No deletion was detected in any of the 25 RTT patients studied.

cohort of 25 patients. All patients have a normal female karyotype. The 671d9 PAC clone which contains the entire MECP2 gene, gave signals on the distal Xq-region on both X chromosomes in more than 80% of the metaphases from both normal control females, and the 25 MECP2 mutation free patients (*figure 1*). In addition, no duplication of the MECP2 locus was identified in any of the 25 RTT patients analysed. Recently, Nielsen et al. [18] did not find any deletion of the entire MECP2 gene in 3 classical RTT patients, supporting our findings.

Repeated sequences are very common in deletion prone regions [27]. As an illustration the steroid sulfatase gene (STS) is completely deleted in up to 90% of patients with X-linked-ichthyosis [13], or conversely, the phenylalanine hydroxylase gene, poor in long highly homologous segments, is deleted in only 5% of patients with phenylalanine hydroxylase deficiency [10]. Based on the data presented here, the MECP2 gene seems to belong to the second gene category; indeed complete loss (or duplication) of the MECP2 gene seems to be a rare event in typical RTT cases. However, more severe phenotype associated with congenital malformations could be caused by contiguous gene syndrome due to microdeletion in Xq28 region encompassing the MECP2 gene.

In conclusion, gross rearrangements seem not to be a MECP2 gene mutational mechanism in RTT; another way to illustrate a possible haplo-insufficiency

of the MECP2 gene in RTT could be an expression study by RNA quantification. A decreased expression of the MECP2 gene would reveal mutations in regulatory elements. Finally, if any additional mutations in MECP2 non-coding regions are identified in RTT patients, a second disease locus might be involved in this neurodevelopmental disorder.

Acknowledgments

We thank all the patients and their families for their participation, all medical doctors who provided us with blood samples, and l'Association Française du Syndrome de Rett, l'Association Française contre les Myopathies and the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie for their financial support. Moreover, we are grateful towards C. Hainaut, B. Petitjean, C. Jacquot, K. Brochet for their technical contribution.

References

- [1] Amano K., Nomura Y., Segawa M., Yamakawa K., Mutational analysis of the MECP2 gene in Japanese patients with Rett syndrome, *J. Hum. Genet.* 45 (2000) 213–236.
- [2] Amir R., Van den Veyver I., Wan M., Tran C., Francke U., Zoghbi H., Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2, *Nat. Genet.* 23 (1999) 185–188.
- [3] Amir R.E., Van den Veyver I., Schultz R., Malicki D.M., Tran C.Q., Dahle E.J., Philippi A., Timar L., Percy A.K., Motil K.J., Lichtarge O., Smith E.O., Glaze D.G., Zoghbi H.Y., Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes, *Ann. Neurol.* 47 (2000) 670–679.
- [4] Auranen M., Vanhala R., Vosman M., Levander M., Varilo T., Hietala M., Riikonen R., Peltonen L., Jarvela I., MECP2 gene analysis in classical Rett syndrome and in patients with Rett-like features, *Neurology* 56 (2001) 611–617.
- [5] Bienvenu T., Carrie A., de Roux N., Vinet M., Jonveaux P., Couven P., Villard L., Arzimanoglou A., Beldjord C., Fontes M., Tardieu M., Chelly J., MECP2 mutations account for most cases of typical forms of Rett syndrome, *Hum. Mol. Genet.* 9 (2000) 1377–1384.
- [6] Bourdon V., Philippe C., Labrune O., Amsellem D., Arnould C., Jonveaux P., A detailed analysis of the MECP2 gene: prevalence of recurrent mutations and gross DNA rearrangements in Rett syndrome patients, *Hum. Genet.* 108 (2001) 43–50.
- [7] Buysse I., Fang P., Hoon K., Amir R., Zoghbi H., Roa B., Diagnostic testing for Rett syndrome by DHPLC and direct sequence analysis of the MECP2 gene: identification of several novel mutations and polymorphisms, *Am. J. Hum. Genet.* 67 (2000) 1428–1436.
- [8] Cheadle J., Gill H., Fleming N., Maynard J., Kerr A., Leonard H., Krawczak M., Cooper D.N., Lynch S., Thomas N., Hughes H., Hulten M., Ravine D., Sampson J., Clarke A., Long-read sequence analysis of the MECP2 gene in Rett syndrome patients: correlation of disease severity with mutation type and location, *Hum. Mol. Genet.* 9 (2000) 1119–1129.
- [9] De Bona C., Zappella M., Hayek G., Meloni I., Vitelli F., Brutini M., Cusano R., Loffredo P., Longo I., Renieri A., Preserved speech variant is allelic of classic Rett syndrome, *Eur. J. Hum. Genet.* 8 (2000) 325–330.
- [10] Eisensmith R.C., Woo S.L.C., Molecular basis of phenylketonuria and related phenylalaninemias: mutations and polymorphisms in the human phenylalanine hydroxylase gene, *Hum. Mut.* 1 (1992) 13–23.
- [11] Hagberg B., Aicardi J., Dias K., Ramos O., A progressive syndrome of autism, dementia, ataxia and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases, *Ann. Neurol.* 14 (1983) 471–479.
- [12] Hampson K., Woods C., Latif F., Webb T., Mutations in the MECP2 gene in a cohort of girls with Rett syndrome, *J. Med. Genet.* 37 (2000) 610–612.
- [13] Hernandez-Martin A., Gonzalez-Sarmiento R., De Unamuno P., X-linked ichthyosis: an update, *Br. J. Dermatol.* 141 (1999) 617–627.
- [14] Huppke P., Laccione F., Kramer N., Engel W., Hanefeld F., Rett syndrome: analysis of MECP2 and clinical characterization of 31 patients, *Hum. Mol. Genet.* 9 (2000) 1369–1375.
- [15] Kim S., Cook J., Novel *de novo* nonsense mutation of MECP2 in a patient with Rett syndrome, *Hum. Mutat.* 15 (2000) 382–383.
- [16] Laccione F., Huppke P., Hanefeld F., Meins M., Mutation spectrum in patients with Rett syndrome in the German population: Evidence of hot spot regions, *Hum. Mutat.* 17 (2001) 183–190.
- [17] Nan X., Campoy F.J., Bird A., MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin, *Cell.* 88 (1997) 471–481.
- [18] Nielsen J., Henriksen K., Hansen C., Silaharoglu A., Schwartz M., Tommerup N., MECP2 mutations in Danish patients with Rett syndrome: high frequency of mutations but no consistent correlations with clinical severity or with the X chromosome inactivation pattern, *Eur. J. Hum. Genet.* 9 (2001) 178–184.
- [19] Obata K., Matsuishi T., Yamashita Y., Fukuda T., Kuwajima K., Horiuchi I., Nagamitsu S., Iwanaga R., Kimura A., Omori I., Endo S., Mori K., Kondo I., Mutation analysis of the methyl-CpG binding protein 2 gene (MECP2) in patients with Rett syndrome, *J. Med. Genet.* 37 (2000) 608–610.
- [20] Reichwald K., Thiesen J., Wiehe T., Weitzel J., Stätling W., Kioschis P., Poustka A., Rosenthal A., Platzer M., Comparative sequence analysis of the MECP2-locus in human and mouse reveals new transcribed regions, *Mamm. Genome* 11 (2000) 182–190.
- [21] Rett A., On a unusual brain atrophy syndrome in hyperammonemia in childhood, *Wien. Med. Wochenschr.* 116 (1966) 723–726.
- [22] Trevarthen E., Moser H., Diagnostic Criteria Work Group. Diagnostic criteria for Rett syndrome, *Ann. Neurol.* 23 (1988) 425–428.
- [23] Vacca M., Filippini F., Budillon A., Rossi V., Mercadante G., Manzati E., Gualandi F., Bigoni S., Trabonelli C., Pini G., Calzolari E., Ferlini A., Meloni I., Hayek G., Zappella M., Renieri A., D'Urso M., D'Esposito M., MacDonald F., Kerr A., Dhanjal S., Hulten M., Mutation analysis of the MECP2 gene in British and Italian Rett syndrome females, *J. Mol. Med.* 78 (2001) 648–655.

[24] Wan M., Lee S.S., Zhang X., Houwink M.I., Song H.R., Amir R.E., Budden S., Naidu S., Pereira J.L., Lo I.F., Zoghbi H.Y., Schanen N.C., Francke U., Rett syndrome and beyond: recurrent spontaneous and familial MECP2 mutations at CpG hotspots, *Am. J. Hum. Genet.* 65 (1999) 1520–1529.

[25] Webb T., Clarke A., Hanefeld F., Pereira J., Rosenbloom L., Woods C., Linkage analysis in Rett syndrome families suggests that there may be a critical region at Xq28, *J. Med. Genet.* 35 (1998) 997–1003.

[26] Xiang F., Buervenich S., Nicolao P., Bailey M.E., Zhang Z., Anvret M., Mutation screening in Rett syndrome patients, *J. Med. Genet.* 37 (2000) 250–255.

[27] Yen P.H., Li X.M., Tsai S.P., Johnson C., Mohandas T., Shapiro L.J., Frequent deletions of the human X- chromosome distal short arm result from recombination between low copy repetitive elements, *Cell* 61 (1990) 603–610.

MECP2 et retard mental

Il aura fallu beaucoup de temps et d'efforts pour découvrir la cause moléculaire du syndrome de Rett, cette encéphalopathie extrêmement sévère qui n'atteint que les petites filles. Mais depuis la découverte de l'implication dans cette maladie du gène *MECP2* (pour *methyl-CpG-binding protein-2*), les recherches entreprises, nombreuses désormais, commencent à porter leurs fruits.

Le gène *Mecp2* a été découvert chez le rat en 1993 [1]. Il appartient à la famille des gènes codant pour des protéines possédant un domaine de fixation aux dinucléotides CpG méthylés (MBD). Il fut retrouvé ensuite chez la souris, localisé sur le chromosome X [2], puis chez l'homme où il est soumis à l'inactivation [3]. Par la suite, il fut démontré que s'il n'était pas indispensable à la multiplication des cellules souches embryonnaires (ES) mâles, il l'était en revanche pour le développement des embryons de souris mâles [4].

MECP2 possède, en plus de son domaine MBD, un domaine de répression transcriptionnelle (TRD) pouvant fonctionner à distance *in vitro* et *in vivo* [5] (figure 1) et interagissant avec un complexe contenant un répresseur de transcription, mSin3A, et une histone désacétylase (*m/s* 1997, n° 10, p. 1205). L'utilisation de trichostatine A, inhibitrice de la désacétylation, avait révélé que cette réaction est essentielle dans le mécanisme de répression. Avec ses domaines MBD et TRD, agissant sur la méthylation de l'ADN et la désacétylation des histones, la protéine *MECP2*, concentrée dans l'hétérochromatine péricentromérique, agit sur le modelage de la chromatine et l'extinction à distance d'autres gènes qui restent encore à trouver. *MECP2* était le candidat idéal pour le syndrome de Rett [6]: létalié chez les garçons (comme chez les souris

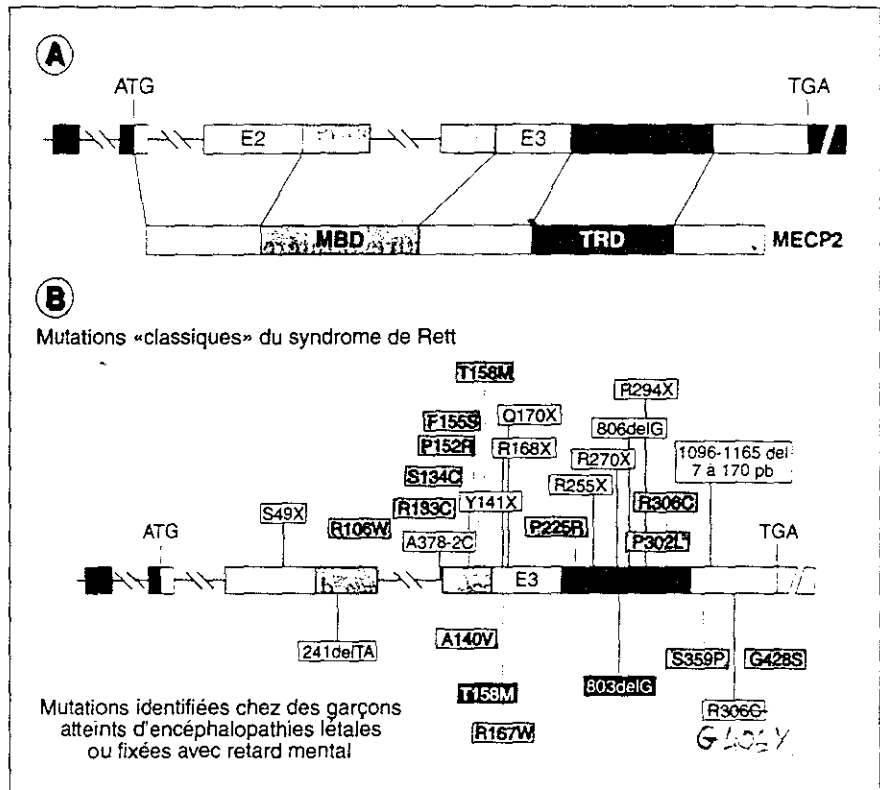


Figure 1. **Gène *MECP2* et retards mentaux.** **A.** Représentation du gène *MECP2* et de la protéine qui contient deux domaines fonctionnels MBD (methyl-CpG binding domain) et TRD (domaine de répression transcriptionnelle). **B.** Localisation des mutations de *MECP2* dans le syndrome de Rett classique et dans les deux formes d'atteinte chez le garçon [8]. Les mutations faux sens sont en rose, et non sens en gris clair. Les mutations responsables d'encéphalopathies létales chez les garçons sont représentées en rouge (faux sens) et en gris foncé.

mâles) et encéphalopathie sévère et progressive chez les filles, d'autant plus que le gène se situe en Xq28, dans la région où le locus du syndrome de Rett avait fini - à grand peine en raison de la rareté des cas familiaux - par être localisé. Encore fallait-il le démontrer, ce qui fut fait en 1999 par une équipe américaine ([7] et *m/s* 1999, n° 11, p. 1334). Depuis, l'étude de grandes cohortes de filles atteintes de syndrome de

Rett a confirmé le rôle de *MECP2* et a permis de répertorier une dizaine de mutations [8]. Le pourcentage de détection des mutations se situe aux alentours de 80%. Celles-ci surviennent *de novo* dans la plupart des cas et l'X muté est le plus souvent d'origine paternelle. Chez les filles atteintes, l'inactivation de l'X est aléatoire (90% des cas environ). Toutefois, dans certains cas familiaux, une origine maternelle est

retrouvée, et s'explique par une inactivation préférentielle de l'X porteur d'une mutation de *MECP2* chez des femmes qui ne présentent alors peu ou pas de symptômes. Mais lorsqu'il est transmis à leurs filles – qui ont une inactivation au hasard de leur X – le chromosome X porteur de la mutation entraînera les manifestations caractéristiques du syndrome de Rett.

Tout récemment, deux types de modèles murins ont été étudiés. Ils donnent un aperçu du retentissement des mutations de *MECP2* sur le fonctionnement cérébral.

Une équipe américaine a obtenu deux modèles murins : des souris avec absence de *Mecp2* (*Mecp2*^{-/-}) et des souris avec une mutation conditionnelle dans le cerveau (par le système Cre-loxP) [9]. Les deux modèles murins ont le même phénotype, avec la même évolution : les souris mâles sont normales jusqu'à la cinquième semaine, puis présentent des tremblements, des troubles respiratoires et comportementaux. Après une période d'obésité et d'hypoactivité, leur état général se dégrade et elles meurent environ vers la dixième semaine. Les femelles hétérozygotes sont normales durant les quatre premiers mois de vie, puis elles prennent du poids et présentent des troubles de la coordination motrice. L'autopsie révèle une réduction sensible du poids du cerveau et de la taille des neurones, mais pas de signes de neurodégénérescence. Si la délétion de *Mecp2* est reportée à la période post-natale, on observe le même phénotype, mais d'apparition un peu plus tardive.

Une équipe anglaise, par des techniques similaires, a obtenu des souris homo-*Mecp2*^{-/-} ou hémizygotés *Mecp2*^{-y} [10]. Celles-ci sont viables et fertiles. Si leur développement, en particulier le poids corporel, varie selon les lignées de souris utilisées, il n'existe en revanche pas de différence sur le plan neurologique. Les souris *Mecp2*^{nul} apparaissent normales à la naissance et dans les semaines qui suivent puis, entre 3 et 8 semaines, on observe une réduction des mouvements, avec incoordination motrice, contracture des pattes de derrière et respiration irrégulière. L'étude histo-

logique des différents tissus, le cerveau en particulier, n'a montré aucune anomalie. La mort survient en général vers le 54^e jour. Les femelles hétérozygotes, qui semblent normales pendant plusieurs mois, finissent elles aussi par présenter des troubles du comportement.

Bien que les résultats ne soient pas absolument superposables, ces deux études démontrent que, contrairement à ce qui avait été publié précédemment [4], le gène *Mecp2* n'est pas indispensable au développement embryonnaire et que l'absence d'expression de *Mecp2* ne retient pas sur le développement cérébral. Le phénotype qui apparaît chez les souris est moins sévère que chez l'homme, puisque les mutants mâles survivent jusqu'à l'âge adulte. Il ressemble cependant beaucoup aux manifestations cliniques du syndrome de Rett qui, on le sait, n'apparaissent que vers l'âge de six mois. Il semble donc que la perte d'expression de *MECP2* retient essentiellement sur la physiologie des neurones post-mitotiques et que le trouble est cellule-autonome, ce qui expliquerait le ralentissement de la croissance céphalique chez les filles par atteinte sélective des neurones dont l'X muté est actif.

Le gène *MECP2* étant ubiquitaire et devant agir comme répresseur transcriptionnel dans toutes les cellules, il est surprenant de constater que son inactivation a pour seul retentissement une atteinte cérébrale spécifique.

Mais il y a plus surprenant encore. Contrairement à ce que l'on avait cru jusqu'à présent, on découvre des mutations du gène *MECP2* chez des garçons. Elles se manifestent par deux formes syndromiques différentes : 1) l'une correspond à une encéphalopathie convulsivante avec décès précoce, pressentie dès 1998, avant même la découverte du gène grâce à quelques cas familiaux [11] ; 2) la seconde est beaucoup plus inattendue. Il s'agit d'un retard mental compatible avec une survie à l'âge adulte. Dans les cas familiaux, les femmes qui transmettent la maladie ont une activation aléatoire de l'X, mais ne présentent que peu ou pas de symptômes. On retrouve aussi des formes sporadiques

qui sont jusqu'à présent sous-évaluées en raison de l'absence de critères spécifiques [12, 13].

Il convient donc de vérifier l'intégrité de *MECP2* chez les garçons porteurs de retards mentaux pour lesquels les recherches d'autres étiologies (X fragile, syndrome d'Angelman, entre autres) se sont avérées négatives, afin de pouvoir établir éventuellement une corrélation génotype-phénotype. Dès à présent, il semble que les mutations dans cette forme non spécifique chez le garçon soient différentes des mutations des formes de syndrome de Rett classique. Elles n'altéreraient que partiellement la fonction de *MECP2* (en dehors des domaines MBD et TRD) (figure 1). Mais ceci mérite encore d'autres études.

Le gène *MECP2*, qui, comme *RSK2* impliqué dans une autre encéphalopathie progressive, le syndrome de Coffin Lowry (*m/s* 1997, n° 1, p. 107), intervient dans le remodelage chromatinien, a donc encore beaucoup à nous apprendre sur les mécanismes d'activation ou d'extinction des gènes.

1. Lewis JD, Meehan RR, Henzel WJ, et al. Purification, sequence, and cellular localisation of a novel chromosomal protein that binds to methylated DNA. *Cell* 1992; 69: 905-14.
2. Quaderi NA, Meehan RR, Tate PH, et al. Genetic and physical mapping of a gene encoding a methyl CpG binding protein, *Mecp2*, to the mouse X chromosome. *Genomics* 1994; 22: 648-51.
3. D'Esposito M, Quaderi NA, Ciccociolla A, et al. Isolation, physical mapping, and Northern analysis of the X-linked human gene encoding methylCpG-binding protein, *MECP2*. *Mammalian Genome* 1996; 7: 533-5.
4. Tate P, Skarnes W, Bird A. The methyl-CpG binding protein *MeCP2* is essential for embryonic development in the mouse. *Nat Genet* 1996; 12: 205.
5. Nan X, Campoy FJ, Bird A. *MeCP2* is a transcriptional repressor with abundant binding sites in genomics chromatin. *Cell* 1997; 88: 471-81.
6. Willard HF, Hendrich BD. Breaking the silence in Rett syndrome. *Nat Genet* 1999; 23: 127-8.
7. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999; 23: 185-8.
8. Bourdon V, Philippe C, Labrune O, Amsellem D, Amould C, Jonveaux P. A detailed analysis of the *MECP2* gene: prevalence of recurrent mutations and gross DNA rearrangements in Rett syndrome patients. *Hum Genet* 2001; 108: 43-50.
9. Chen RZ, Akbarian S, Tudor M, Jaenisch R. Deficiency of methyl-CpG binding protein-2 in

CNS neurons results in a Rett-like phenotype in mice. *Nat Genet* 2001; 27: 327-31.

10. Guy J, Hendrich B, Holmes M, Martin JE, Bird A. A mouse *MeCP2*-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet* 2001; 27: 322-6.

11. Schanen NC, Kurczynski TW, Brunelle D, Woodcock MM, Dure LS, Percy ALK. Neonatal encephalopathy in two boys in families with recurrent Rett syndrome. *J Child Neurol* 1998; 13: 229-31.

12. Meloni I, Bruttini M, Longo I, et al. *MECP2* mutations in Rett syndrome gene, *MECP2*, causes X-linked mental retardation and progressive spasticity in males. *Am J Hum Genet* 2000; 67: 982-5.

13. Orrico A, Lam C-W, Galli L, et al. *MECP2* mutation in male patients with non-specific X-linked mental retardation *FEBS* 2000; 24106: 1-4.

Simone Gilgenkrantz

9, rue Basse, 54330 Clérey-sur-Brenon, France.

Violaine Bourdon

Laboratoire de génétique médicale, UPRES-INRA 952, CHRU, rue du Morvan, 54511 Vandœuvre-les-Nancy, France.

Association Française de Cytométrie (AFC)

7^e Congrès annuel de l'AFC
9 au 12 octobre 2001
le Nouveau centre
des congrès de Saint-Étienne

En dehors des sujets classiques, de nouveaux thèmes tels que génome, transcriptome, protéome et biopuces seront abordés. Tous les résumés feront l'objet de communications courtes (2 min), moyennes (10 min) ou longues (20 min). Par ailleurs, de nombreuses firmes présenteront leurs nouveaux produits dans tous les domaines de la cytométrie.

Pour tout renseignement
écrire à: Alpha Visa Congrès/
AFC Cytométrie 2001,
624, rue des Grèzes,
34070 Montpellier;
ou faites parvenir un e-mail à:
afc2001@alphavisa.com

■■■ BRÈVES ■■■

■■■ **Rétrovirus et schizophrénie, quand *pol* s'exprime.** C'est à l'aboutissement d'une longue quête, puisqu'elle a commencé il y a plus de 15 ans [1], qu'une équipe américano-allemande menée par Hakan Karlsson (Johns Hopkins, Baltimore USA) nous invite en démontrant la présence de séquences proches du gène rétroviral *pol* dans le liquide céphalo-rachidien de patients atteints de schizophrénie. Près de 30% des 35 patients testés, atteints d'une schizophrénie d'apparition récente, présentaient en effet ces séquences, alors que ce n'était le cas d'aucun des 52 témoins (et d'un seul patient schizophrène chronique sur 20). Ces séquences correspondaient, plus particulièrement, à celles rencontrées dans la famille des rétrovirus humains endogènes HERV-W. Une étude *post-mortem* parallèle a permis aux auteurs de confirmer une augmentation de la transcription de rétrovirus de la même famille dans la région frontale du cerveau d'autres patients schizophrènes. D'une façon intéressante, les auteurs ont observé dans les séquences *pol* retrouvées chez les patients des codons *stop* qui semblaient devoir interdire la réplication des virus, ce qui suggère une synthèse endogène plutôt qu'une source exogène. Il est clair – et les auteurs ne le prétendent sûrement pas – que la schizophrénie n'est pas une maladie mono-factorielle. Les études de familles ont depuis longtemps démontré, notamment, le rôle de facteurs génétiques (toutefois encore mal déterminés). L'existence d'un élément rétroviral, qui agirait essentiellement lors du déclenchement de la maladie, semble donc mis à l'ordre du jour. En dehors de l'intérêt général que présentent ces résultats – dont on attend néanmoins la réplication, l'histoire des recherches étiopathogéniques des maladies psychiatriques étant riche en fausses bonnes nouvelles – ils pourraient déboucher sur des applications thérapeutiques, le traitement de la schizophrénie passant alors, dans un curieux détour, par la pharmacopée antivirale.

[1. Crow TJ. *Br J Psychiatry* 1984; 145: 243-53.]

[2. Karlsson H, et al. *Proc Natl Acad Sci USA* 2001; 98: 4634-9.]

■■■ **Des cellules progénitrices amplifiées à partir du cerveau adulte... chez l'homme aussi.** Dans une courte lettre adressée à *Nature*, Fred Gage et son équipe (*Salk Institute*, CA, USA) concluent une quête de près de 10 ans en démontrant que le cerveau adulte de l'homme contient des cellules capables de proliférer *in vitro* et de donner naissance à des neurones [1]. De telles cellules avaient été observées chez la souris par l'équipe de Samuel Weiss (Université de Calgary, Canada) en 1992 [2], et l'équipe de Fred Gage avait, il y a trois ans, rapporté l'existence d'une neurogenèse dans l'hippocampe adulte chez l'homme [3]. La boucle est bouclée, ce qui vaut pour la souris vaut pour l'homme: des cellules souches adultes peuvent être amplifiées *in vitro*, dans des conditions de culture particulières nécessitant de fortes concentrations de FGF-2 et un cofacteur, la forme glycosylée de la cystatine C. Incidemment, ces travaux montrent que les limites apparentes des capacités de ces cellules souches identifiées chez la souris adulte valent aussi pour l'homme: seulement 1% à 1‰ des cellules différenciées sont des neurones, et plus l'âge du cerveau dans lequel les cellules souches sont prélevées est avancé, plus petit est le nombre de passages possible avant l'entrée en sénescence. On attend à présent les mises au point qui permettront de contourner ces problèmes et de créer, à partir de ce matériel, les banques de cellules implantables dont tout le monde rêve.

[1. Palmer TD, et al. *Nature* 2001; 411: 42-3.]

[2. Reynolds BA, Weiss S. *Science* 1992; 255: 1707-10.]

[3. Eriksson PS, et al. *Nat Med* 1998; 4: 1313-7.]

The signal transducer and activator of transcription STAT5b gene is a new partner of retinoic acid receptor α in acute promyelocytic-like leukaemia

Écile Arnould⁺, Christophe Philippe⁺, Violaine Bourdon, Marie José Grégoire, Roland Berger¹ and Philippe Jonveaux[§]

⁺Laboratoire de Génétique, UPRES-INRA 952, CHRU, Rue du Morvan, 54511 Vandoeuvre les Nancy, France and [§]INSERM U434, SD 401 No. 434 CNRS, CEPH, 27 Rue Juliette Dodu, 75010 Paris, France

Received April 19, 1999; Revised and Accepted June 9, 1999

Acute promyelocytic leukaemia (APL) exhibits a characteristic t(15;17) translocation that fuses the promyelocytic leukaemia (PML) gene on 15q22 to the retinoic acid receptor α (RARA) gene on 17q12–q21.1. In a small subset of acute promyelocytic-like leukaemias (APL-L), RARA is fused to a different partner: the promyelocytic leukaemia zinc finger (PLZF) gene on 11q23, a nucleophosmin (NPM) gene on 5q35 or the nuclear mitotic apparatus (NuMA) gene on 11q13. We report on a molecular characterization of a RARA gene rearrangement in a patient with APL-L and demonstrate that the STAT5b gene is fused with RARA. STAT5b belongs to a janus kinase (JAK)–STAT signalling pathway. Remarkably, the STAT5b component of the chimeric protein is delocalized from the cytoplasm to the nucleus, where it displays a microspeckled pattern. Therefore, the usual features of this APL-L might result from dysregulation of the JAK/STAT5 signal transducing pathways in the patient leukaemic cells. In this study, we identified STAT5b as a new gene fused to RARA in leukaemia; this is the first human tumour bearing a structurally abnormal STAT gene.

INTRODUCTION

Chromosomal aberrations leading to gene fusion play an important role in a great number of haematopoietic malignancies. Acute promyelocytic leukaemia (APL) accounts for 10–15% of all acute myeloid leukaemias. The vast majority of APL exhibits a characteristic t(15;17) translocation that fuses the promyelocytic leukaemia (PML) gene to the retinoic acid receptor α (RARA) gene (1), resulting in an aberrant PML–RARA protein associated with leukaemogenesis. The fusion protein inhibits differentiation and promotes survival of myeloid precursor cells. In APL-L patients (1–2%), the RARA gene is fused to a different partner as a result of variant rearrangements: apart from PML, promyelocytic leukaemia zinc finger

(PLZF) (2), nucleophosmin (NPM) (3) and nuclear mitotic apparatus (NuMA) (4) have been found fused to RARA.

We have previously reported an APL-L case associated with a derivative chromosome 17 (5). A RARA gene rearrangement was found by Southern blot analysis. The PML–RARA gene fusion present in almost 100% of APL was ruled out (5).

In the present study, we demonstrate that this APL-L patient harbours a STAT5b–RARA gene fusion. STAT5b belongs to a family of latent cytosolic transcription factors activated by janus kinase (JAK) tyrosine kinases. The STAT family is composed of seven members in mammals (STAT1–STAT4, STAT5a, STAT5b and STAT6). After cytokine interaction with their receptor, JAKs are auto-transphosphorylated. Subsequently, the intracellular receptor docking site is phosphorylated by the activated JAK. STAT5b interacts with the activated receptor through its SH2 domain and in turn becomes phosphorylated (on Tyr699) by JAK. Phosphorylated STAT5b dimerizes in a head-to-tail configuration and STAT5b homodimers and STAT5a/b heterodimers proceed to the nucleus and regulate gene transcription through interaction between the DNA-binding domain and IFN- γ activation sequence (GAS) (for reviews on the JAK/STAT pathway see refs 6–8). STAT5 proteins were initially characterized in sheep as prolactin-regulated mammary gland factors (9). The human STAT5b protein is highly homologous to STAT5a except for the C-terminal *trans*-activation domain (10,11). A great diversity of extracellular signalling polypeptides, including cytokines, growth factors and hormones, have been shown to trigger STAT5 protein activation (8). STAT5 proteins are mainly implicated in mammopoiesis, lactogenesis and body growth regulation. However, these transcription factors are also involved in haematopoiesis (12–17) and several lines of evidence implicate the Stat5 proteins in proliferation induced by mitogenic cytokines such as interleukin-2 (IL-2), IL-3, IL-7 (18–22), granulocyte colony-stimulating factor (G-CSF) (23) and erythropoietin (24).

The STAT5b gene gives rise to both long (94 kDa) and short (80 kDa) functionally distinct isoforms generated by protein processing relying on a nuclear serine protease (25). This natural short variant acts as a dominant negative STAT5b transcription factor (11,26) and is preferentially activated in multipotent haematopoietic progenitors (25).

These authors contributed equally to this work.

To whom correspondence should be addressed. Tel: +33 3 83 15 37 69; Fax: +33 3 83 15 37 72; Email: p.jonveaux@chu-nancy.fr

STAT5a^{-/-} and/or STAT5b^{-/-} knockout mice have phenotypes mainly attributable to deficiencies in prolactin and growth hormone functions (14). STAT5a^{-/-}b^{-/-} double mutant mice have a profound deficiency in peripheral T cell proliferation together with a reduced number of cytokine colony-forming unit responses of bone marrow cells (14). Moreover, STAT5a^{-/-} and STAT5b^{-/-} mice exhibit a defect in natural killer cell proliferation and function (21). STAT5s regulate expression of many genes (8,21,27) coding for milk proteins (β -casein and β -lactoglobulin) and for proteins involved in haematopoiesis and the immune response [oncostatin M, *c-myc*, cytokine-inducible SH2 domain containing protein (CIS), IL-2R α , IL-2R β , perforin and pim-1]. Under growth hormone stimulation, STAT5b controls liver gene expression in a sex-specific manner (28).

We report on the molecular characterization of the STAT5b-RARA gene fusion at both the cDNA and the genomic levels together with immunolocalization studies of the chimeric protein in the patient leukaemic cells. We discuss the potential implication of the STAT5b component of the fusion protein in leukaemogenesis in the present APL-L.

RESULTS

Fusion of STAT5b to RARA

In order to isolate the gene fused to RARA, a 5' rapid amplification of cDNA ends (5'-RACE) was performed on total RNAs extracted from the patient bone marrow cells (Fig. 1a). A Blastn analysis of a 5'-RACE product-derived sequence showed an in-frame fusion between STAT5b and RARA genes (Fig. 1a). The resulting protein is composed of 1038 amino acids; a schematic representation of the anticipated structure of the STAT5b-RARA protein is shown in Figure 1b. The putative STAT5b-RARA protein contains all known functional domains of RARA except for the N-terminal *trans*-activating function. Apart from all RARA functional domains, the fusion protein contains the STAT5b coiled-coil region (29), DNA-binding domain and a truncated SH2 domain. Tyr699 essential for STAT5b activation and the C-terminal transactivation domain with Ser731 involved in regulation of STAT5b transcriptional activity are absent from the chimeric STAT5b-RARA protein. The reciprocal RARA-STAT5b mRNA was not detected after RT-PCR on patient bone marrow total RNAs (data not shown). The STAT5b rearrangement was confirmed at the genomic level by Southern blot analysis with STAT5b cDNA (Fig. 2).

Characterization of the fusion at the genomic level was done with a PAC which contains all the STAT5b coding region and with a breakpoint spanning phage clone (Fig. 3). Sequence analysis showed a 3 bp insertion at the joining site; the 2.4 kb chimeric intron is mainly composed of STAT5b-derived intronic sequence. No homologies were detected with highly repeated long (LINES) and short interspersed elements (SINES). The breakpoint in RARA gene intron 2 is located in the so-called 'extremely restricted region' (ERR) of 50 bp which has been identified as an illegitimate recombination hotspot involved in PML-RARA rearrangements (30,31). Moreover, we found short identical stretches of DNA between STAT5b and RARA at the joining points in the present case (Fig. 3), resembling those observed in PML-RARA fusions (32). The model of the t(15;17) translocation proposed by

Yoshida *et al.* (32) is therefore likely to be relevant to this STAT5b-RARA gene fusion. The reciprocal chimeric RARA-STAT5b intron, estimated to be slightly larger than 17 kb long [the normal RARA intron 2 is 17 kb long (33)], was not detected by PCR with two different oligonucleotide pairs on patient genomic DNA (data not shown). This suggests a small interstitial deletion ranging from STAT5b intron *n* to RARA intron 2 as the most probable mechanism for formation of the STAT5b-RARA gene fusion. Duplication of the 17q21.3-q23 region observed after comparative genomic hybridization (CGH) analysis (see Materials and Methods) is probably a secondary event in leukaemogenesis.

The STAT5b gene maps to 17q21.1-q21.2 close to the RARA locus

STAT5a and STAT5b were both mapped at 17q11.2 (10) by FISH; however, the metaphase chromosome resolution did not allow precise localization of STAT5b in the 17q proximal region. Our FISH analysis with PAC 196p17 on normal prometaphases seems to indicate that the STAT5b and RARA genes are very close to each other in 17q21.1-q21.2 and that the physical distance between the two genes is most likely <3 Mb (34; Fig. 4).

Localization of the STAT5b-RARA protein in leukaemic cells

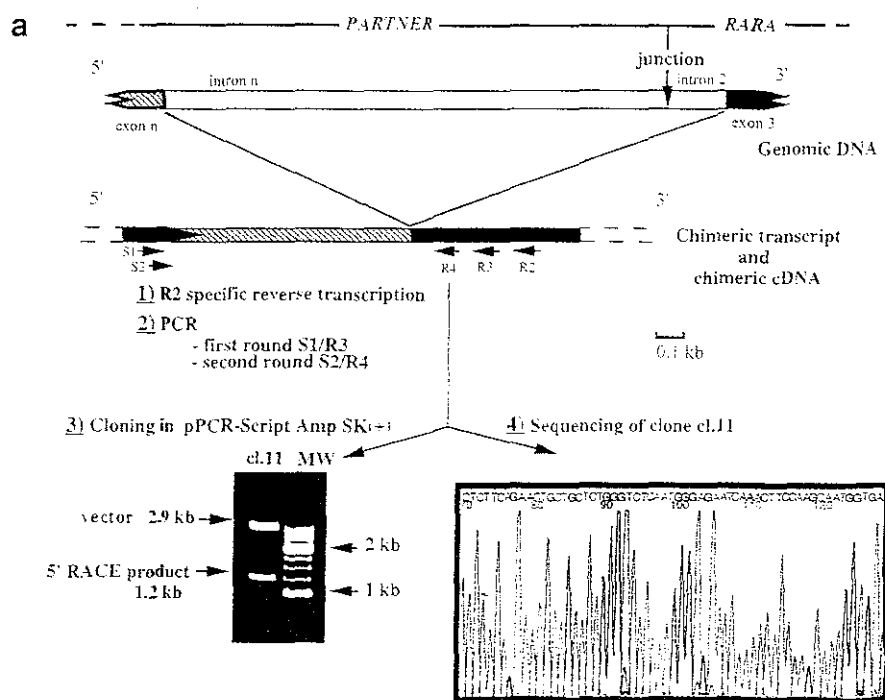
Immunolocalization studies on patient bone marrow cells show that the STAT5b-RARA protein is mainly localized in the nucleus but also to a lesser extent in the cytoplasm and displays a microspeckled pattern, whereas the normal STAT5b protein shows a normal diffuse cytoplasmic localization (Fig. 5).

DISCUSSION

In this study, we identified STAT5b as the fifth gene fused to RARA in APL and APL-L leukaemias. PML, PLZF, NPM, NuMA and STAT5b (in its phosphorylated form) are nuclear proteins. However, for the first time, the protein derived from the RARA gene partner can be localized in the cytosol (in its latent form). To date, all RARA partners contain an N-terminal protein-protein interaction domain; the STAT5b coiled-coil domain is located between amino acids 232 and 321 (29).

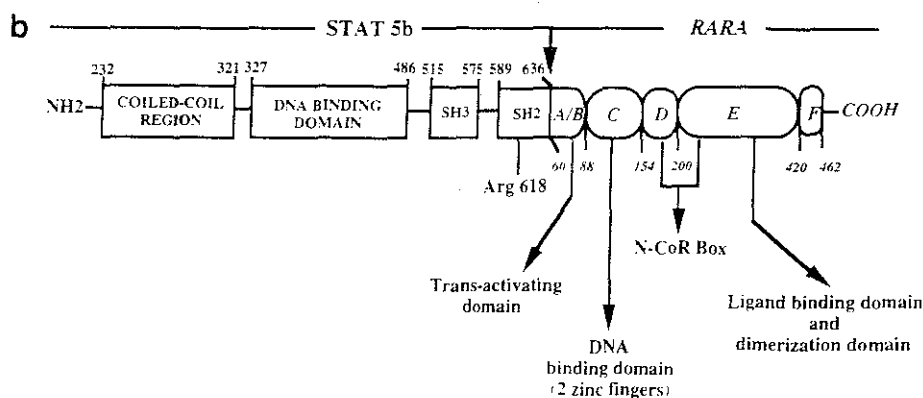
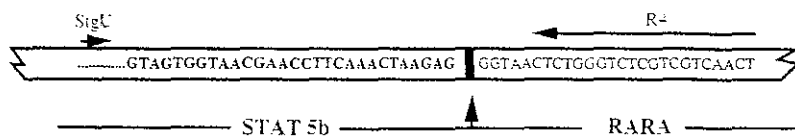
RARA is essential for proper myeloid cell differentiation in response to retinoic acid (RA). As for other APLs and APL-Ls, the chromosome 17 breakpoint in the present case is located in intron 2 of the RARA gene. This suggests that an abnormal receptor with almost all functional domains of the native RARA is the hallmark for inhibition of myeloid cell differentiation. The arrest of maturation of myeloid cells in APL-L reported in this study is possibly due to an abnormal RARA protein.

The well-characterized PML-RARA and PLZF-RARA oncogenic proteins mediate leukaemogenesis through recruitment of histone deacetylase leading to aberrant chromatin acetylation and transcriptional repression of RA-regulated genes (35-38). An N-CoR box has been implicated in the ability of RARA to interact with co-repressors (35). Moreover, in PML-RARA-associated APLs, delocalization of the normal PML protein in blast cells promotes survival of myeloid precursors, as PML has recently been shown to be involved in apoptotic pathways (39,40). Patients with the PML-RARA, NPM-RARA and NuMA-RARA gene fusions



MW : molecular weight marker

5) Sequence comparison (blastn)



1. Isolation and characterization of a STAT5b-RARA fusion. (a) A RARA-specific cDNA library was obtained after reverse transcription with oligonucleotide R2 (1) RNAs from patient bone marrow cells. We chose the avian myeloblastoid virus (AMV) reverse transcriptase because of its high thermostability in order to avoid problems to RNA secondary structures in the 5'-portion of the chimeric transcript. After two rounds of PCR (2), a smear was obtained and the 5'-RACE products were cloned into pPCR-Script Amp SK(+) (3). Sequencing with primer T3 (4) of clone no. 11, which contains a 1.2 kb insert, and database searches using the Blastn program (<http://bi.nlm.nih.gov/BLAST/>), allowed the characterization of a fusion between STAT5b and the third exon of RARA (5). 100% homology was detected with human STAT5b or RARA (accession no. M61111) from nt 574 corresponding to the 5'-end of exon 3 up to nt 616, the 5'-end of R4 (the oligonucleotide used for the nested PCR). In our chimeric cDNA sequence, exon 3 from RARA was fused to the STAT5b gene, as revealed by a perfect homology between our chimeric cDNA sequence and the STAT5b sequence (up to nt 2052). The presence of a chimeric cDNA was confirmed by PCR on the patient cDNA library with oligonucleotides StgU and R4 (data not shown). The 234 bp PCR product was sequenced, confirming the STAT5b-RARA gene fusion found by 5'-RACE (data not shown). The STAT5b intron fused to RARA intron 2 (intron n) because the STAT5b genomic structure is unknown. (b) The fusion between STAT5b exon 3 and the third exon of RARA preserved the reading frame in frame. The predicted chimeric transcript, from the STAT5b initiation codon to the end of the 3'-untranslated region of RARA, is 4532 nt long and encodes a fusion protein of 1038 amino acid residues. The fusion protein contains a proline which replaces Threonine and Glutamine in the wild-type RARA and STAT5b proteins, respectively.

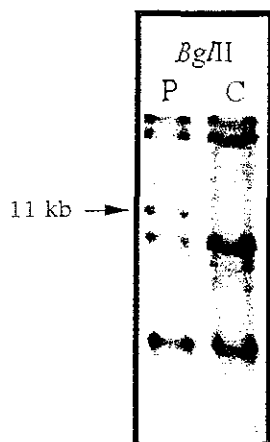


Figure 2. Detection of the STAT5b gene rearrangement at the genomic level by Southern blot analysis. Southern blot analysis was done on DNA from patient bone marrow cells (P) and from a normal control (C) with the STAT5b cDNA probe. A junction fragment is detected with *Bgl*II (11 kb). The presence of a rearranged band confirms, at the genomic level, the result obtained with total RNAs after 5'-RACE-PCR.

show complete remission after differentiation therapy with all-*trans* RA (ATRA) (41). Addition of ATRA mediates replacement of the repressor complex by an activator complex with acetyltransferase activity (42,38). The transcriptional activator complex is composed of CREB binding protein (CBP)/p300 and histone acetyltransferase. APL-Ls with the PLZF-RARA gene fusion are not sensitive to ATRA because the PLZF component of the chimeric protein is also able to recruit a repressor complex with histone deacetylase activity but RA has not effect on the PLZF-repressor complex (36–38). It is noteworthy that our case did not respond to ATRA. The STAT5b protein-protein interaction domain has been shown to associate with N-Myc interactor (Nmi), which in turn enhances association of CBP/p300 with STAT5 (29). The chimeric STAT5b-RARA protein could sequester CBP/p300 cofactors and therefore prevent ATRA to release the repressor complex from the RA target genes. Alternatively, an ATRA-insensitive association of N-CoR co-repressor with the N-terminal coiled-coil domain derived from the STAT5b component of the fusion protein could explain the unresponsiveness of the present APL-L to RA.

Apart from its role in prolactin and growth hormone-induced functions, STAT5b is also involved in differentiation and proliferation of myeloid progenitors. Notably, G-CSF, which plays a critical role in granulopoiesis, stimulates STAT5 transcription factor activity (23). The STAT5b component of the fusion protein could therefore participate in leukaemogenesis in the present APL-L and explain the atypical phenotype as compared with classical APL. Thus far, several reports link aberrations in the JAK/STAT pathways to malignant phenotypes. In *Drosophila*, a dominant mutant Jak kinase (*hopTum-1*) causes leukaemia-like abnormalities (43). Constitutive STAT5 activation has been found in T cell leukaemia/lymphoma (44,45) and in malignant T lymphocytes derived from cutaneous anaplastic large T cell lymphoma and Sezary syndrome (46). The STAT5 transcription factors were also found constitutively activated mainly in acute lymphoblastic leukaemias (ALL) (47,48) and in haematopoietic cell lines transformed by BCR/ABL tyrosine kinase (49,50). It was also found constitutively activated in acute myeloid leukaemias

(47,48,51,52) and in familial erythrocytosis (53). Before this report, TEL-JAK2 fusions in a T cell childhood acute lymphoblastic leukaemia (54) and in myeloid leukaemias (55) were the unique gene rearrangements in the JAK/STAT pathways described in human cancers. All TEL-JAK2 variants strongly activate STAT5 (56). The finding of constitutively activated STAT5 transcription factors in numerous malignant haematopoietic disorders strongly supports the idea of STAT5-dependent oncogene activation. Very recently, Nieborowska-Skorska *et al.* (57) have demonstrated a causal involvement of STAT5 activation in BCR/ABL-mediated leukaemogenesis both *in vitro* and *in vivo*.

The present APL-L with a gene fusion between STAT5b and RARA is the first malignancy that harbours a rearranged member of the STAT gene family. The STAT5b-RARA chimeric protein contains 636 amino acids from the N-terminus of STAT5b: the fusion site occurs within the C-terminal SH2 domain of STAT5b, before Tyr699. The SH2 domain coordinates interaction of Stat5b with the phosphotyrosine docking site on the cytoplasmic part of the cytokine receptor. The phosphotyrosine docking site recruits the STAT5b protein by interacting with a critical arginine residue (Arg618) in the highly conserved core of SH2. When recruited, Tyr699 on STAT5b is phosphorylated [Tyr(P)], the activated form of STAT5b is then able to dimerize by reciprocal SH2-Tyr(P) interactions. The highly conserved core of the SH2 domain with Arg618 is present in the STAT5b-RARA protein: it could bind to the cytokine receptor. The cytoplasmic truncated STAT5b protein could therefore act in a dominant negative manner by competing with the normal STAT5b transcription factor for the interaction with the receptor docking site. Studies on homo- and heterodimerization between the STAT1 and STAT2 proteins showed that an unphosphorylated STAT can dimerize with a Tyr(P)-STAT: a single Tyr(P)-SH2 interaction therefore seems sufficient for dimerization (58,59). If the chimeric STAT5b-RARA protein could 'dimerize' by a single Tyr(P)-SH2 interaction with the wild-type STAT5a and STAT5b proteins, it will sequester normal STAT proteins and perturbate the JAK/STAT5 pathways. Moreover, the fusion protein is mainly nuclear and contains a functional DNA-binding region. One could anticipate that stable homo- or heterodimers with normal STAT5a/b proteins are generated and bind to GAS elements. STAT5b-RARA-STAT5a/b heterodimers could either constitutively activate gene transcription or act in a dominant negative manner on activated wild-type STAT5a/b transcription factors. If the STAT5b-RARA gene fusion leads to a constitutively activated STAT5b transcription factor, cytokine independent growth may participate in leukaemogenesis in this APL-L. Lastly, and as mentioned above, the STAT5 transcription factors have been shown to interact with other proteins such as CBP/p300 to form enhanceosomes (29). In the present APL-L case, abnormal STAT5b proteins could sequester co-activators of other signalling pathways, including the retinoid pathway.

As patient material is no longer available, ongoing transfection studies with the chimeric cDNA in the Ba/F3 cell line, which is dependent on IL-3 for growth, will allow testing of the effect of the fusion protein on the STAT5 pathways. Moreover, co-transfection experiments with an RA-inducible reporter plasmid will also permit the study of the effect of STAT5b-RARA protein on RA-regulated genes.

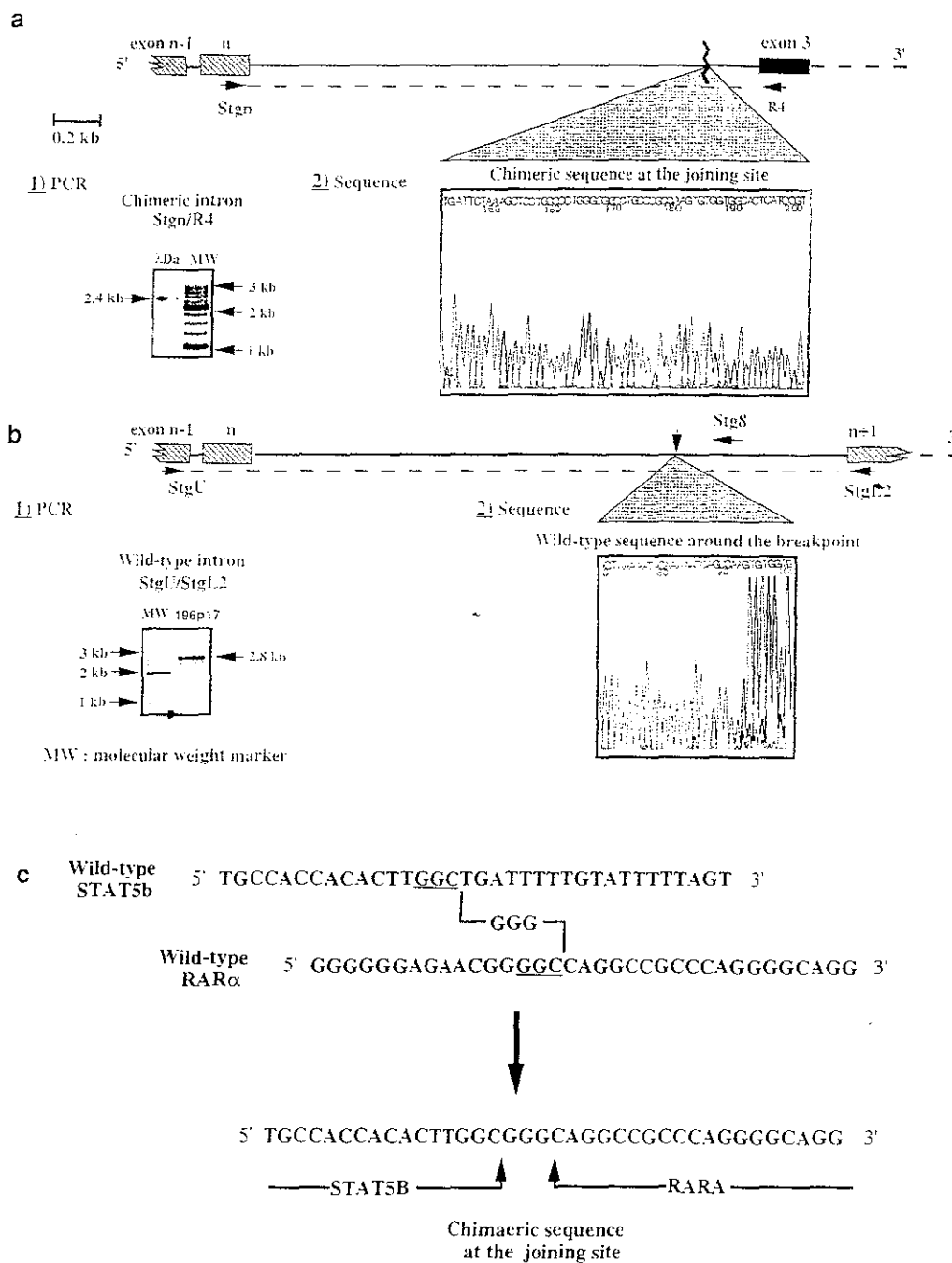


Fig. 3. Characterization of the STAT5b-RARA rearrangement at the genomic level. The characterization of the rearrangement at the genomic level was done with PAC 196p17 and a breakpoint spanning phage clone. PAC 196p17 was obtained by PCR screening of the RPCII PAC library (60); it contains all the STAT5b coding region (data not shown). The phage clone Da encompasses the STAT5b-RARA joining sequence. Southern blot analysis of *EcoRI*-digested phage Da DNA with a 5.5 kb *EcoRI* RARA cDNA probe revealed a 6.5 kb junction fragment (data not shown). (a) The chimeric intron was PCR amplified on phage Da DNA which contains the joining sequence. The PCR reaction was performed with R4 and Stgn, an oligonucleotide localized at the 3'-end of STAT5b exon *n*. The Stgn 3' nucleotide corresponds to the last STAT5b-1 nucleotide in the chimeric STAT5b-RARA cDNA. The chimeric intron is ~2.4 kb long. (2) Sequence analysis starting from RARA exon 3 showed a normal RARA cDNA sequence for 344 bp, the chimeric intron is therefore essentially composed of STAT5b-derived intronic sequence. (b) The wild-type STAT5b intronic sequence around the breakpoint was obtained on a 2.8 kb StgU/StgL2 PCR product. The StgL2 3'-end nucleotide corresponds to the 5'-end of STAT5b exon *n* + 1; StgU is localized upstream from Stgn at the cDNA level. (1) The PCR reaction was performed on PAC 196p17 DNA. The PAC 196p17 insert contains an internal *NcoI* site; its size estimated by field inversion gel electrophoresis (FIGE Mapper; Bio-Rad) to be ~100 kb. Southern blot analysis with STAT5b cDNA allowed us to estimate the STAT5b genomic size to be <50 kb (data not shown). (2) The Stg8 oligonucleotide was designed thanks to an StgL2-derived intronic sequence and allowed us to approach the wild-type intronic sequence around the breakpoint. Sequence analysis from StgU gave us partial information on the intron-exon structure of STAT5b. Exon *n* is 131 bp long (n 1922 to 2052; accession no. U47686) and is split off from exon *n* - 1 by a 101 bp intron *n* - 1. The donor and acceptor splice sites in introns *n* - 1 and *n* follow the AG rule. (c) Comparison between the germline sequences (accession no. S57794 for the 3'-end of RARA intron 2) and the chimeric intronic sequence revealed a 3 bp identity at the joining site between STAT5b and RARA genes. Moreover, short 3 bp identical stretches of DNA between STAT5b and RARA are present upstream from splice sites in wild-type STAT5b and RARA introns.

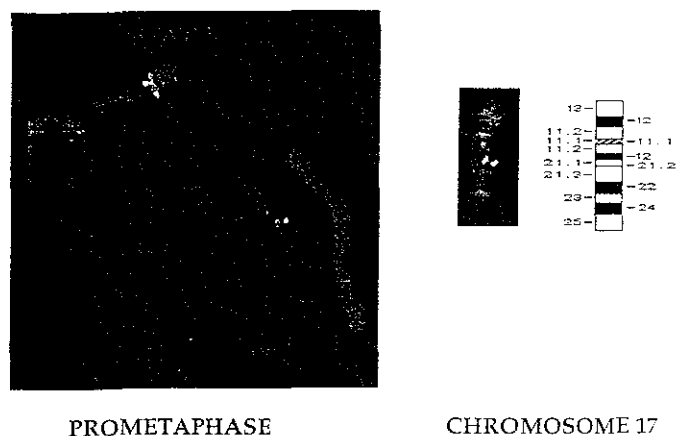


Figure 4. Localization of the STAT5b gene in relation to the RARA locus by FISH on normal prometaphases. PAC 196p17 was biotin labeled and co-hybridized with digoxigenin-labeled Smith-Magenis/D17S258 (17p11.2) and RARA (17q21.1) probes (Oncor) on normal prometaphases and nuclei. On prometaphase chromosome 17, the STAT5b green signal, observed after FITC-avidin detection, co-localizes with the RARA red signal visualized after rhodamine-anti-digoxigenin detection, as demonstrated by yellow signals on 17q21.1–q21.2. No background was noted at any other chromosome location using the PAC clone 196p17 as a probe for FISH analysis.

MATERIALS AND METHODS

Patient

Details of the patient, a 67-year-old man, have been previously published (5). He had acute myeloid leukaemia (AML M1 in the French/American/British classification). However, in the bone marrow a minority of blast cells showed morphological features suggestive of the microgranular variant of APL, M3v. Blast cells failed to respond to ATRA *in vitro*. Cytogenetic analysis of bone marrow cells revealed a derivative chromosome 17 larger than a normal chromosome 17. Chromosome microdissection (5) and CGH analysis allowed us to demonstrate that extra material on der(17) was the result of a duplication of the 17q21.3–q23 region (data not shown).

5'-RACE PCR

Total RNAs from patient bone marrow cells were extracted according to the TRIzol protocol (Sigma). Amplification of the unknown 5'-end of the chimeric mRNA was performed with a 5'/3'-RACE kit (Boehringer Mannheim). The incubation step for first strand cDNA synthesis was performed at 55°C. Gene-specific first strand cDNA synthesis was performed using primer R2 (for oligonucleotides used in this study see Table 1) localized in exon 4 of the RARA gene (accession no. X06614). After dA tailing of the purified chimeric cDNA, a first PCR with an oligo(dT) anchor primer and R3 followed by a second semi-nested PCR using an anchor primer and R4 allowed us to obtain a 1.2 kb PCR product. Both PCRs were carried out with the Expand Long Template system (Boehringer Mannheim) at an annealing temperature of 63°C. Secondary PCR products were cloned using the PCR-Script AMP SK(+) cloning kit (Stratagene). After transformation of Epicurian Coli XL1-Blue MRF Kan supercompetent cells, plasmid DNA was extracted from white colonies with the Qiagen Plasmid Minikit and subsequently sequenced using primers T3 and T7.

Table 1. Oligonucleotides used in this study

Name	Location	Sequence (5'→3')
RARA		
R2	Exon 5	GCACCTCCTTCTTCTTCTG
R3	Exon 3	CAGCCTCACAGGCGCTGAC
R4	Exon 3	CTGGGCACTATCTTTCAGAACT
STAT5b		
Stgn	Exon <i>n</i>	TGCTTGGGAAGTTTGATTCTC
StgU	Exon <i>n</i> - 1	GTTTGACGGTGTGATGGAAGTG
StgL2	Exon <i>n</i> + 1	AAGTCTCTGGTGGTAAAAGG
Stex5'	First exon	CTCAGCAGCTCCAAGGAGAAGC
St165U	Last exon	AGTCTTTCATCTTCACCA
Stg8	Intron <i>n</i>	CCAGCACTTTGGGAGGCC
St165L	3'-UTR	AAACACATACTCGCACTCG

Construction of the patient genomic library

A phage library was constructed from patient genomic DNA extracted from bone marrow leukaemic cells by inserting partially *Mbo*I-digested DNA into the *Xho*I site of the λ FixII vector (Stratagene) and the library was packaged with Gigapack II Plus Packaging extracts according to the manufacturers' protocols. The total phage library (10^6 phages) was screened with a 5.5 kb *Eco*RI genomic probe used to confirm an RARA gene rearrangement in the patient (1). A total of five bacteriophages were isolated after three rounds of screening. Bacteriophage DNAs were prepared with the Qiagen Lambda Maxi kit. Hybridization of the RARA 5.5 kb *Eco*RI genomic probe on *Eco*RI-digested phage DNA identified two phages (Ea and Da) spanning the junction fragment. Hybridization of chemiluminescence-labeled R4 and Stgn oligoprobes confirmed the presence of exonic sequences derived from both STAT5b and RARA in the 6.5 kb junction fragment.

Sequencing of the junction fragment was performed on the R4/Stgn PCR product of DNA from phage clones Ea and Da. PCR products were analysed by electrophoresis on a 0.8% agarose gel. The 2.4 kb band was purified with a Gel Extraction kit (Qiagen). Initially, sequencing was performed with R4 and Stgn; subsequently nested primers were designed to reach the joining site.

Sequencing reaction

Purified PCR products (30–100 ng) or cloned DNA (800 ng) were cycle sequenced using dye terminator chemistry with the AmpliTaq FS enzyme (Applied Biosystems) and were run on an ABI 373 DNA sequencer.

Isolation of a genomic clone containing the STAT5b gene

PAC 196p17 containing the entire coding region of the STAT5b gene was obtained after PCR screening of the RPCII PAC library constructed by Ioannou *et al.* (60). Primary and secondary pools as well as individual microtitre plates were purchased from the UK HGMP Resource Centre (Cambridge, UK). Initial PCR screening was performed with primers St165U and St165L flanking the STAT5b stop codon. DNA from PAC 196p17 was prepared with the Qiagen Maxi-prep kit. The presence of the 5'-end of STAT5b was shown by Southern blot analysis with the Stex5' oligoprobe. Labelling by terminal transferase with digoxigenin-

cell pellets were then resuspended in the isotonic solution at 5×10^6 cells/ml. Aliquots of 250 000 cells were spread out on SuperFrost slides with a Cytospin at 400 r.p.m. for 5 min. Cells were subsequently air dried and fixed in cold 70% alcohol for 5 min at 4°C and washed twice for 5 min in cold phosphate-buffered saline (pH 7.2). For double-labelling, slides were then co-incubated for 2 h at room temperature with a mouse monoclonal antibody to RARA [Ab9a(F)(9a-9A6); 1/100] (Prof. P. Chambon, Strasbourg, France) and with rabbit polyclonal IgG STAT5b N-20 or STAT5b C-17 (2 mg/ml; Santa Cruz Biotechnology). After a second incubation with rhodamine-conjugated sheep anti-rabbit antibody (1/40; Oncor) for 30 min at room temperature, a final incubation with FITC-conjugated rabbit anti-mouse antibody (1/40; Dako) for 30 min at room temperature was performed. The slides were washed in phosphate-buffered saline and mounted with antifade (Oncor).

Southern blotting

Total genomic DNA from lymphocytes or from bone marrow cells was extracted with the Nucleon BACC3 DNA isolation kit (Amersham). Southern blot analysis was performed following standard methods. Rearrangement of the STAT5b locus was investigated with the full-length STAT5b cDNA.

FISH

FISH analysis was performed with PAC 196p17 on cytogenetic preparations from a normal male. Prometaphase chromosomes were prepared following the standard procedures. The D17S258 Smith-Magenis probe (Oncor) were used as a marker for chromosome 17p11.2. The D17S258 probe also contains a RARA-specific probe. PAC 196p17 DNA (17q12) was labelled by nick-translation (Biotin-Nick Translation Mix; Boehringer Mannheim) with 16-dUTP-biotin. Suppression of repetitive sequences with Cot-1 DNA followed by co-hybridization of D17S258, RARA and PAC 196p17 were done following standard methods. Washes, signal detection and amplification were performed according to the chromosome *in situ* hybridization protocol (Oncor). The chromosomes were counterstained with DAPI antifade (Oncor) and visualized with a Zeiss Axioplan fluorescent microscope equipped with a PSI Power Gene FISH System analyser.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. C.C.A. Bernard (Melbourne, Australia), Dr O. Bernard (INSERM U434), Prof. P. Leblond, F. Wuilque, Drs M.T. Daniel and J. Buisine for helpful discussions. They are also indebted to Profs P. Chambon (IGBMC-LGME-U184-UPL) and W.J. Leonard (Bethesda, MD) for the RARA antibody and the STAT5b cDNA, respectively. We are grateful to L. Maury for help with the patient genomic library screening and to F. Poly for PAC library screening. We also thank Profs G. Faure and M.C. Bene for help with immunocytochemistry and M. Franck and B. Petitjean for technical assistance. This work was supported by Association pour la Recherche contre le Cancer, Ligue Nationale contre le Cancer, la Fondation contre la Leucémie and Association Française de lutte contre les Myopathies. C.A. was supported by a fellowship from the Ministère de la Recherche et de l'Espace (no. 94951338083).

REFERENCES

- de Thé, H., Chomienne, C., Lanotte, M., Degos, L. and Dejean, A. (1990) The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor α gene to a novel transcribed locus. *Nature*, **347**, 558–561.
- Chen, Z., Brand, N.J., Chen, A., Chen, S.J., Tong, J.H., Wang, Z.Y., Waxman, S. and Zelen, A. (1993) Fusion between a novel Kruppel-like zinc finger gene and the retinoic acid receptor α locus due to a variant t(11;17) associated with acute promyelocytic leukaemia. *EMBO J.*, **12**, 1161–1167.
- Redner, R.L., Rush, E.A., Faas, S., Rudert, W.A. and Corey, S.J. (1996) The t(5;17) variant of acute promyelocytic leukemia expresses a nucleophosmin-retinoic acid receptor fusion. *Blood*, **87**, 882–886.
- Wells, R.A., Catzavelos, C. and Kamel-Reid, S. (1997) Fusion of retinoic acid receptor α to NuMA, the nuclear mitotic apparatus protein, by a variant translocation in acute promyelocytic leukaemia. *Nature Genet.*, **17**, 109–113.
- Jonveaux, P., Le Coniat, M., Derre, J., Flexor, M.A., Daniel, M.T. and Berger, R. (1996) Chromosome microdissection in leukemia: a power tool for the analysis of complex chromosomal rearrangements. *Genes Chromosomes Cancer*, **15**, 26–33.
- Darnell, J.E.Jr (1997) STATs and gene regulation. *Science*, **277**, 1630–1635.
- Pellegrini, S. and Dusanter-Fourt, I. (1997) The structure, regulation and function of the Janus kinases (JAKs) and the signal transducers and activators of transcription (STATs). *Eur. J. Biochem.*, **248**, 615–633.
- Heim, M.E. (1999) The Jak-STAT pathway: cytokine signalling from the receptor to the nucleus. *J. Recept. Signal Transduct. Res.*, **19**, 75–120.
- Wakao, H., Gouilleux, F. and Groner, B. (1994) Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *EMBO J.*, **13**, 182–191.
- Lin, J.X., Mietz, J., Modi, W.S., John, S. and Leonard, W.J. (1996) Cloning of human Stat5b. *J. Biol. Chem.*, **271**, 10738–10744.
- Moriggl, R., Gouilleux-Gruart, V., Jähne, R., Berchtold, S., Gartmann, C., Liu, X., Hennighausen, L., Sotiropoulos, A., Groner, B. and Gouilleux, F. (1996) Deletion of the carboxyl-terminal transactivation domain of MGF-Stat5 results in sustained DNA binding and a dominant negative phenotype. *Mol. Cell. Biol.*, **16**, 5691–5700.
- Gouilleux, F., Pallard, C., Dusanter-Fourt, I., Wakao, H., Haldosen, L.-A., Norstedt, G., Levy, D. and Groner, B. (1995) Prolactin, growth hormone, erythropoietin and granulocyte-macrophage colony stimulating factor induce MGF-Stat5 DNA binding activity. *EMBO J.*, **14**, 2005–2013.
- Welte, T., Koch, F., Schuler, G., Lechner, J., Doppler, W. and Heufler, C. (1997) Granulocyte-macrophage colony-stimulating factor induces a unique set of STAT factors in murine dendritic cells. *Eur. J. Immunol.*, **27**, 2737–2740.
- Teglund, S., McKay, C., Schuetz, E., van Deursen, J.M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G. and Ihle, J.N. (1998) Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. *Cell*, **93**, 841–850.
- al-Shami, A., Mahanna, W. and Naccache, P.H. (1998) Granulocyte-macrophage colony-stimulating factor-activated signaling pathways in human neutrophils. *J. Biol. Chem.*, **273**, 1058–1063.
- Caldenhoven, E., van Dijk, T.B., Tjimmen, A., Raaijmakers, J.A.M., Lammers, J.W.J., Koenderman, L. and de Groot, R.P. (1998) Differential activation of functionally distinct STAT5 proteins by IL-5 and GM-CSF during eosinophil and neutrophil differentiation from human CD34⁺ hematopoietic stem cells. *Stem Cells*, **16**, 397–403.
- Welte, T., Leitenberg, D., Dittel, B.N., al-Ramadi, B.K., Xie, B., Chin, Y.E., Janeway, C.A.Jr, Bothwell, A.L.M., Bottomly, K. and Fu, X.Y. (1999) STAT5 interaction with the cell receptor complex and stimulation of T-cell proliferation. *Science*, **283**, 222–225.
- Wilks, A.F. and Oates, A.C. (1996) The JAK/STAT pathways. *Cancer Surv.*, **27**, 139–163.
- Mui, A.L., Wakao, H., Kinoshita, T., Kitamura, T. and Miyajima, A. (1996) Suppression of interleukin-3-induced gene expression by a C-terminal truncated Stat5: role of Stat5 in proliferation. *EMBO J.*, **15**, 2425–2433.
- van der Plas, D.C., Smiers, F., Pouwels, K., Hoefsloot, L.H., Lowenberg, B. and Touw, I.P. (1996) Interleukin-7 in human B cell precursor acute lymphoblastic leukemia cells and murine BAF3 cells involves activation of STAT1 and STAT5 mediated via the interleukin-7 receptor alpha chain. *Leukemia*, **10**, 1317–1325.
- Imada, K., Bloom, E.T., Nakajima, H., Horvath-Arcidiacono, J.A., Udy, G.B., Davey, H.W. and Leonard, W.J. (1998) Stat5b is essential for natural killer cell-mediated proliferation and cytolytic activity. *J. Exp. Med.*, **7**, 2067–2074.

- Matikainen, S., Sarenva, T., Ronni, T., Lehtonen, A., Koskinen, P.J. and Juikunen, I. (1999) Interferon- α activates multiple STAT proteins and upregulates proliferation-associated *IL-2Ra*, *c-myc* and *pim-1* genes in human T cells. *Blood*, **93**, 1980–1991.
- Dong, F., Xiuwen, L., de Koning, J.P., Touw, J.P., Henninghausen, L., Lamer, A. and Grimley, P.M. (1998) Stimulation of STAT5 by granulocyte colony-stimulating factor (G-CSF) is modulated by two distinct cytoplasmic regions of the G-CSF receptor. *J. Immunol.*, **161**, 6503–6509.
- Chrétien, S., Varlet, P., Verdier, F., Gobert, S., Cartron, J.P., Gisselbrecht, S., Mayeux, P. and Lacombe, C. (1996) Erythropoietin-induced erythroid differentiation of the human erythroleukemia cell line TF-1 correlates with impaired STAT5 activation. *EMBO J.*, **15**, 4174–4181.
- Meyer, J., Jücker, M., Ostertag, W. and Stocking, C. (1998) Carboxyl-truncated STAT5b is generated by a nucleus-associated serine protease in early hematopoietic progenitors. *Blood*, **91**, 1901–1908.
- Wang, D., Stravopodis, D., Teglund, S., Kitazawa, J. and Ihle, J.N. (1996) Naturally occurring dominant negative variants of STAT5. *Mol. Cell. Biol.*, **16**, 6141–6148.
- Yu, C.R., Ortaldo, J.R., Curiel, R.E., Young, H.A., Anderson, S.K. and Gosselin, P. (1999) Role of a STAT binding site in the regulation of the human perforin promoter. *J. Immunol.*, **162**, 2785–2790.
- Gerbert, G.A., Park, S.H. and Waxman, D.J. (1999) Termination of growth hormone pulse-induced STAT5b signaling. *Mol. Endocrinol.*, **13**, 38–56.
- Zhu, M., John, S., Berg, M. and Leonard, W.J. (1999) Functional association of Nmi with Stat5 and Stat1 in IL-2 and INF- γ -mediated signaling. *Cell*, **96**, 121–130.
- Fashiro, S., Koromura, N., Tanaka, K., Suzuki, K., Kyo, T., Dohy, H., Niwa, J. and Kamada, N. (1994) Identification of illegitimate recombination hot spot of the retinoic acid receptor alpha gene involved in 15;17 chromosomal translocation of acute promyelocytic leukemia. *Oncogene*, **9**, 1939–1945.
- Fashiro, S., Wang, Z., Kotomura, N., Niwa, O., Ueda, K. and Kamada, N. (1997) Nuclear proteins binding to the recombination hotspot region of the retinoic acid receptor alpha gene. *Leukemia*, **11** (suppl. 3), 285–286.
- Foshida, H., Naoe, T., Fukutani, H., Kiyoi, H., Kubo, K. and Ohno, R. (1995) Analysis of the joining sequences of t(15;17) translocation in human acute promyelocytic leukemia: sequence non-specific recombination between the PML and RARA genes within identical short stretches. *Genes Chromosomes Cancer*, **12**, 37–44.
- Fiverio, D., Lo Coco, F., D'Adamo, F., Biondi, A., Fagioli, M. and Grignani, F. (1992) Identification of DNA rearrangements at the retinoic acid receptor- α (RAR- α) locus in all patients with promyelocytic leukemia (APL) and mapping of APL breakpoints within the RAR- α second intron. *Leukemia*, **7**, 3331–3336.
- Fontana, J., Lawrence, J., Carter, K.C. and Gerdes, M.J. (1992) Extending the capabilities of interphase chromatin mapping. *Nature Genet.*, **2**, 171–172.
- Fong, S.H., David, G., Wong, C.W., Dejean, A. and Privalsky, M.L. (1997) SMRT corepressor interacts with PLZF and with the PML-retinoic acid receptor α (RAR α) and PLZF-RAR α oncoproteins associated with acute promyelocytic leukemia. *Proc. Natl Acad. Sci. USA*, **94**, 9028–9033.
- Fong, R.J., Nagy, L., Inoue, S., Shao, W., Miller, W.H.Jr and Evans, R.M. (1998) Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature*, **391**, 811–813.
- Fornace, A.J., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Croce, M., Zamir, I., Seiser, C., Grignani, F., Lazar, M.A., Minucci, S. and Ricci, P.G. (1998) Fusion proteins of the retinoic acid receptor- α recruit histone deacetylase in promyelocytic leukaemia. *Nature*, **383**, 815–818.
- Fornace, A.J., Inoue, S., Zhu, J., Soderstrom, M., Waxman, S. and Zelent, A. (1998) Reduced retinoic acid-sensitivities of nuclear receptor corepressor binding to PML- and PLZF-RAR α underlie molecular pathogenesis of treatment of acute promyelocytic leukemia. *Blood*, **91**, 2634–2642.
- Fornace, A.J., De Bels, F., Koken, M., Feunteun, J., Ameisen, J.C. and de Thé, H. (1998) PML induces a novel caspase-independent death process. *Nature Genet.*, **20**, 259–265.
- Fornace, A.J., Ruggero, D., Ronchetti, S., Zhong, S., Gaboli, M., Rivi, R. and Pandolfi, P.P. (1998) Pml is essential for multiple apoptotic pathways. *Nature Genet.*, **20**, 266–272.
- Fornace, A.J., Nervi, C., Avvisati, G. and Mandelli, F. (1998) Acute promyelocytic leukemia: a curable disease. *Leukemia*, **12**, 1866–1880.
- Fornace, A.J., Lin, R.J., Schiltz, R.L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M.L., Nakatani, Y. and Evans, R.M. (1997) Nuclear receptor activator ACTR is a novel histone acetyltransferase and forms a multiprotein activation complex with PCAF and CBP/p300. *Cell*, **90**, 569–580.
43. Luo, H., Hanratty, W.P. and Dearolf, C.R. (1995) An amino acid substitution in *Drosophila* hopTum-1 Jak kinase causes leukemia-like hematopoietic defects. *EMBO J.*, **14**, 1412–1420.
44. Migone, T.S., Lin, J.X., Cereseto, A., Mulloy, J.C., O'Shea, J.J., Franchini, G. and Leonard, W.J. (1995) Constitutively activated Jak-Stat pathway in T cells transformed with HTLV-I. *Science*, **269**, 79–81.
45. Takemoto, S., Mulloy, J.C.C., Cereseto, A., Migone, T.S., Patel, B.K.R., Marsooka, M., Yamaguchi, K., Takatsuki, K., Kamihira, S., White, J.D., Leonard, W.J., Waldmann, T. and Franchini, G. (1997) Proliferation of adult T cell leukemia/lymphoma cells is associated with the constitutive activation of JAK/STAT proteins. *Proc. Natl Acad. Sci. USA*, **94**, 13897–13902.
46. Zhang, Q., Nowak, I., Vonderheid, E.C., Rook, A.H., Kadin, M.E., Nowel, P.C., Shaw, L.E. and Wasik, M.A. (1996) Activation of Jak/STAT proteins involved in signal transduction pathway mediated by receptor for interleukin-2 in malignant T cell lymphocytes derived from cutaneous anaplastic large T-cell lymphoma and Sezary syndrome. *Proc. Natl Acad. Sci. USA*, **93**, 9148–9153.
47. Gouilleux-Gruart, V., Debierre-Grockiego, F., Gouilleux, F., Capiod, J.C., Claisse, J.F., Delobel, J. and Prin, L. (1997) Activated Stat related transcription factors in acute leukemia. *Leukemia Lymphoma*, **28**, 83–88.
48. Weber-Nordt, R., Egen, C., Wehinger, J., Ludwig, W., Gouilleux-Gruart, V., Mertelsman, R. and Finke, J. (1996) Constitutive activation of STAT proteins in primary lymphoid and myeloid leukemia cells and in Epstein-Barr virus (EBV)-related lymphoma cell lines. *Blood*, **88**, 809–816.
49. Carlesso, N., Frank, D.A. and Griffin, J.D. (1996) Tyrosyl phosphorylation and DNA binding activity of signal transducers and activators of transcription (STAT) proteins in hematopoietic cell lines transformed by Bcr/Abl. *J. Exp. Med.*, **183**, 811–820.
50. Shuai, K., Halpern, J., ten Hoeve, J., Rao, X. and Sawyers, C.L. (1996) Constitutive activation of STAT5 by the BCR-ABL oncogene in chronic myelogenous leukemia. *Oncogene*, **13**, 247–254.
51. Hayakawa, F., Towatari, M., Iida, H., Wakao, H., Kiyoi, H., Naoe, T. and Saito, H. (1998) Differential constitutive activation between STAT-related proteins and MAP kinase in primary acute myelogenous leukemia. *Br. J. Haematol.*, **101**, 521–528.
52. Xia, Z., Baer, M.R., Block, A.W., Baumann, H. and Wetzler, M. (1998) Expression of signal transducers and activators of transcription proteins in acute myeloid leukemia blasts. *Cancer Res.*, **58**, 3173–3180.
53. Arcasoy, M.O., Harris, K.W. and Borger, B.G. (1999) A human erythropoietin receptor gene mutant causing familial erythrocytosis is associated with deregulation of the rates of Jak2 and Stat5 inactivation. *Exp. Hematol.*, **27**, 63–74.
54. Lacroix, V., Boureux, A., Delle Valle, V., Poiré, H., Quang, C.T., Mauchauffe, M., Berthou, C., Lessard, M., Berger, R., Ghysdael, J. and Bernard, O.A. (1997) A TEL-JAK2 fusion protein with constitutive kinase activity in human leukemia. *Science*, **278**, 1309–1312.
55. Peeters, P., Raynaud, S.D., Cools, J., Wlodarska, I., Grosgeorge, J., Philip, P., Monpoux, F., Van Rompaey, L., Baens, M., Van den Berghe, H. and Marynen, P. (1997) Fusion of TEL, the ETS-variant gene 6 (ETV6), to the receptor-associated kinase JAK2 as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia. *Blood*, **90**, 2535–2540.
56. Schwaller, J., Frantsve, J., Aster, J., Williams, I.R., Tomasson, M.H., Ross, T.S., Peeters, P., Van Rompaey, L., Van Eten, R.A., Ilaria, R.J., Marynen, P. and Gilliland, D.G. (1998) Transformation of hematopoietic cell lines to growth-factor independence and induction of a fatal myeloid and lymphoproliferative disease in mice by retrovirally transduced TEL/JAK2 fusion genes. *EMBO J.*, **17**, 5321–5333.
57. Nieborowska-Skorska, M., Wasik, M.A., Stupianek, A., Salomoni, P., Kitamura, T., Calabretta, B. and Skorski, T. (1999) Signal transducer and activator of transcription (STAT)5 activation by BCR/ABL is dependent on intact Src homology (SH)3 and SH2 domain of BCR/ABL and is required for leukemogenesis. *J. Exp. Med.*, **189**, 1229–1242.
58. Zhang, X., Blenis, J., Li, H.C., Schindler, C. and Chen-Kiang, S. (1995) Requirement of serine phosphorylation for formation of STAT-promoter complexes. *Science*, **267**, 1990–1994.
59. Gupta, S., Yan, H., Wong, L., Ralph, S., Krolewski, J. and Schindler, C. (1996) The SH2 domains of STAT1 and STAT2 mediate multiple interactions in the transduction of IFN- α signals. *EMBO J.*, **15**, 1075–1084.
60. Ioannou, P.A., Amemiya, C.T., Games, J., Kroisel, P.M., Shizuya, H., Batzer, M.A. and de Jong, P.J. (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet.*, **6**, 84–89.

Un nouveau gène de fusion dans les leucémies aiguës promyélocyaires

La grande majorité des patients atteints de leucémies aiguës promyélocyaires (LAP), hémopathies qui représentent 10% à 15% des leucémies non lymphoïdes, présentent une translocation acquise, t(15;17), véritable signature chromosomique de la maladie. Celle-ci entraîne une fusion entre les gènes *PML* (*promyelocytic leukemia*) et *RARA* (récepteur à l'acide rétinoïque) [1]. Dans 1% à 2% des LAP, dont certaines ont des caractères cytologiques atypiques, le gène *RARA* est fusionné à un autre partenaire: *PLZF* (*promyelocytic leukemia zinc finger*) dans la translocation t(11;17); *NPM* (*nucleophosmin*); *NuMA* (*nuclear mitotic apparatus protein*). Ces quatre partenaires de *RARA* n'ont pas de structure commune, mais possèdent tous un domaine amino-terminal d'interaction protéine-protéine et un domaine d'homodimérisation [2]. Le point de cassure est dans l'intron 2 (ou exceptionnellement dans l'exon 3) du gène *RARA*, et la protéine hybride X-*RARA* est toujours fortement exprimée dans les LAP, qu'elles soient typiques ou non, alors que l'expression du transcrite réciproque *RARA-X* est variable.

Les protéines oncogéniques X-*RARA* (surtout *PML-RARA* et *PLZF-RARA*) forment des dimères avec les récepteurs *RARA* et *RXR* (*retinoid X receptor*) normaux et se comportent comme des récepteurs dominants négatifs. Ainsi, dans les LAP associées à la fusion *PML-RARA*, *PML* est séquestrée par les hétérodimères *ML/PML-RARA*, et les corps nucléaires, contenant normalement la protéine *PML*, sont désorganisés, ce qui peut contribuer à accroître la survie des blastes bloqués au stade

promyélocyte, *PML* étant impliquée dans l'apoptose [3, 4]. *RARA* fait partie de la superfamille des récepteurs nucléaires qui s'hétérodimérisent avec *RXR* et se fixent sur des éléments de réponse aux rétinoïdes (*RARE*), formant ainsi une unité transcriptionnelle inductible par les rétinoïdes (*AR*) dont l'activité est modulée par des co-facteurs transcriptionnels [5, 6]. En l'absence de ligand, l'hétérodimère *RARA-RXR* recrute des co-répresseurs (*SMRT* ou *NCoR*, *mSin3* et une histone désacétylase), l'interaction entre le dimère *RARA-RXR* et le complexe co-répresseur s'effectuant au niveau de la boîte *N-CoR* située dans la partie centrale des récepteurs nucléaires (figure 1) [7]. La désacétylation localisée des histones entraîne la compaction de la chromatine et donc la répression transcriptionnelle des gènes contrôlés par les acides rétinoïques (*AR*). La présence d'*AR* à des concentrations physiologiques (10^{-9} M) induit la dissociation du complexe et favorise l'association de *RARA-RXR* avec des co-activateurs (tels que *CBP*, *p300*) possédant une activité acétyltransférase. Il en résulte une décompaction de la chromatine favorable à l'expression des gènes impliqués dans la différenciation de la lignée granulocytaire.

Des études récentes ont démontré que les oncoprotéines *PML-RARA* et *PLZF-RARA* interagissent de façon anormale avec les co-répresseurs de la transcription [2]. Cette interaction est stable à des concentrations physiologiques d'*AR* et il en résulte un blocage de la différenciation de la lignée myéloïde. Le complexe entre les co-répresseurs et la protéine *PML-RARA* est déstabilisé à des doses

pharmacologiques d'*AR* (10^{-6} M), ce qui explique la différenciation induite par l'*ATRA* (acide rétinoïque tout-*trans*) dans la vaste majorité des LAP. En effet, un aspect remarquable des LAP associées à *PML-RARA*, *NPM-RARA* et *NuMA-RARA* est la rémission complète qu'entraîne l'administration de concentrations supra-physiologiques (10^{-6} M) d'*ATRA*. A l'inverse, l'interaction entre *PLZF-RARA* et les co-répresseurs est insensible à l'*ATRA*. En fait, *PLZF* possède un domaine de liaison avec les co-répresseurs (domaine *BTB/POZ*) situé dans sa partie amino-terminale, le complexe formé entre *BTB/POZ* et les co-répresseurs n'est pas déstabilisé en présence de rétinoïdes, et ce même à des concentrations supra-physiologiques [2]. *PLZF-RARA* entraîne donc une répression constitutive des gènes contrôlés par les *AR*.

STAT5b est fusionné à *RARA* dans un cas de LAP atypique

Nous avons récemment identifié, chez un patient atteint de LAP atypique, insensible à l'action de l'acide rétinoïque tout-*trans*, une nouvelle fusion de *RARA* avec un gène partenaire original codant pour le facteur de transcription *STAT5b* (*signal transducer and activator of transcription*) [8]. D'un point de vue chromosomique, le gène de fusion *STAT5b-RARA* dérive d'une délétion interstitielle de la région 17q21.1-q21.2 survenue à la suite d'une recombinaison illégitime intramoléculaire (les gènes *STAT5b* et *RARA* ne sont distants que de quelques Mb sur le chromosome 17). *STAT5b* est un facteur de transcription présent dans le cytoplasme sous

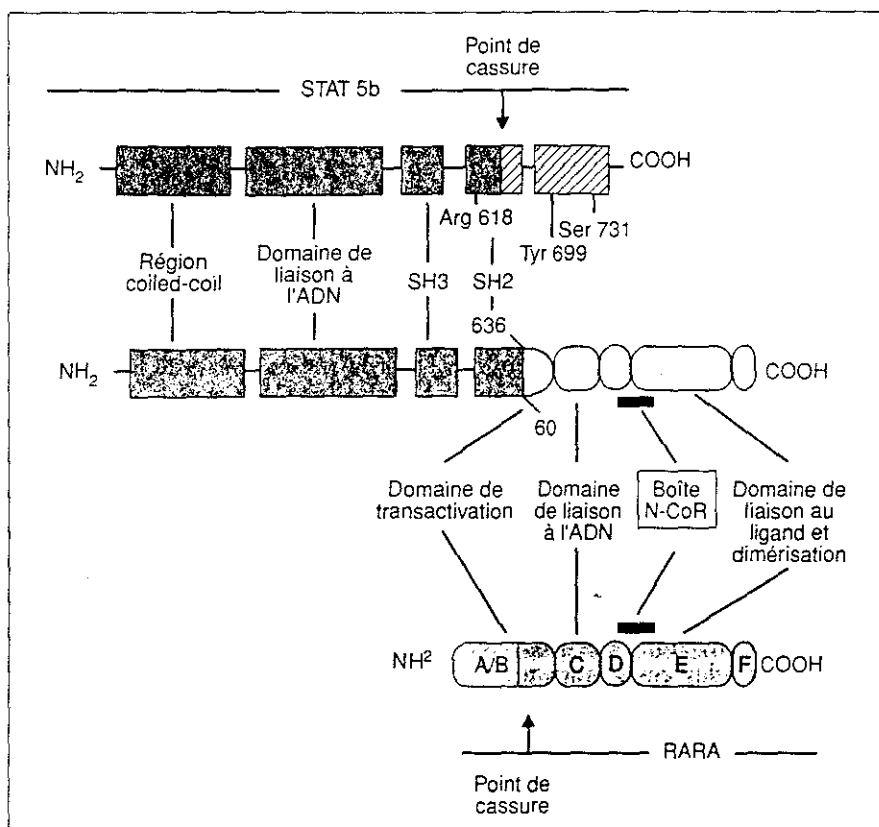


Figure 1. Anatomie de la protéine hybride STAT5b/RARA. L'oncoprotéine STAT5b/RARA est longue de 1038 acides aminés, elle contient la quasi-totalité des domaines fonctionnels de RARA à l'exception du domaine de transactivation A/B qui est tronqué; le point de cassure est classiquement situé dans l'intron 2 du gène codant pour le récepteur aux rétinoïdes. Comme les autres partenaires de RARA au cours de LAP typiques ou atypiques, STAT5b contient un domaine d'interaction protéine-protéine (coiled-coil, entre les acides aminés 232 et 321) et un domaine de dimérisation, tous deux présents au sein de la protéine de fusion. Le domaine de liaison à l'ADN ainsi que le domaine SH3 (Src homology 3) sont également présents au sein de la protéine hybride. Le point de fusion se situe dans le domaine SH2, la région carboxy-terminale, contenant les résidus tyrosine 699 (essentielle au processus d'activation) et sérine 731 (intervenant dans la modulation de l'activité du facteur de transcription), est donc absente. À l'inverse, l'arginine en position 618, intervenant dans le processus de dimérisation, est présente; il est donc envisageable, comme cela a déjà été décrit pour STAT1 et STAT2, que la protéine de fusion s'hétérodimérise avec des facteurs STAT normaux par le biais d'une seule interaction SH2-Tyr phosphorylée.

forme de monomère inactif, et qui, en réponse à la liaison d'un facteur de croissance exogène sur son récepteur, est phosphorylé par les Janus kinases (JAK) sur une tyrosine située dans la partie carboxy-terminale des STAT [9]. L'activation induit la dimérisation de deux monomères STAT, et leur migration dans le noyau où ils modulent la transcription de gènes cibles. STAT5b a été

initialement identifié dans le tissu mammaire [10] et, depuis, la réalisation de l'inactivation des gènes *STAT5A* et *STAT5B* a permis de préciser leur fonction dans l'hématopoïèse [11]. Des études d'immunocytochimie sur les blastes de la moelle osseuse du patient nous ont permis de localiser la protéine hybride majoritairement dans le noyau où elle pourrait être

en mesure d'interférer avec les voies des rétinoïdes mais également avec les voies JAK-STAT (STAT5 et STAT3). De plus, la protéine STAT5b-RARA est présente en faible quantité dans le cytoplasme, où elle pourrait affecter les premières étapes de la transduction du signal par la voie JAK-STAT.

L'oncoprotéine STAT5b-RARA (1038 acides aminés), produit du gène hybride, fusionne un récepteur RARA délété d'un de ses domaines d'activation, et un facteur STAT5b qui a perdu son extrémité carboxy-terminale, critique pour l'activation par les JAK (figure 1). Cette fusion STAT5b-RARA est doublement intéressante: (1) elle identifie un nouveau partenaire de RARA, et (2) elle représente le premier remaniement de structure connu d'un gène de la famille STAT dans un processus néoplasique. Reste à comprendre son implication dans le processus leucémique. A ce jour, la responsabilité des membres de la famille STAT dans certains processus malins se fonde sur des arguments indirects, soit une dérégulation de leur activité, soit une activité constitutive associée à l'expression de certains oncogènes [12]. Par exemple, la protéine de fusion Tel/JAK2, identifiée dans les blastes d'un enfant atteint de leucémie lymphoblastique T avec une translocation t(9;12) (p24;p13) phosphoryle et active STAT 1, 3 et 5 [13].

STAT5b et PLZF: un modèle commun de résistance à la thérapie par différenciation dans les LAP atypiques

Notre patient était résistant à l'action de l'ATRA comme c'est le cas pour les patients porteurs de la translocation t(11;17) fusionnant PLZF et RARA. Le domaine d'interaction protéine-protéine dans la partie amino-terminale des protéines STAT est conservé dans la protéine hybride (figure 1), et pourrait, comme c'est le cas du domaine BTB/POZ de PLZF, s'associer avec un complexe répresseur, et ce même en présence de fortes concentrations d'ATRA. STAT5 est aussi capable de s'associer avec CBP/p300, un co-activateur de la transcription qui est recruté par

RARA-RXR en présence d'AR. Il est donc possible que la protéine hybride, si elle se lie à CBP/p300, puisse interférer avec l'action de ces co-activateurs, il est aussi possible que, dans la protéine hybride, la partie STAT5b crée un encombrement stérique gênant l'accessibilité de l'ATRA.

Des études de transfection de la protéine hybride STAT5b/RARA *in vitro* mais également *in vivo* permettront d'appréhender les propriétés biochimiques et biologiques de cette oncoprotéine et d'évaluer les perturbations qu'elle entraîne tant sur les voies JAK/STAT que sur les voies de signalisation faisant intervenir les récepteurs nucléaires.

1. Lavau C, Jansen J, Weis K, Lamond A, Dejean A. Leucémie aiguë promyélocytaire et acide rétinolique: le paradoxe. *Med Sci* 1994; 10: 817-24.
2. Lin RJ, Egan DA, Evans RM. Molecular genetics of acute promyelocytic leukemia. *Trends Genet* 1999; 15: 179-84.
3. Quignon F, De Bels F, Koken M, Feunteun J,

Ameisen JC, de Thé H. PML induces a novel caspase-independent death process. *Nat Genet* 1998; 20: 259-65.

4. Wang ZC, Ruggero D, Ronchetti S, et al. Pml is essential for multiple apoptotic pathways. *Nat Genet* 1998; 20: 266-72.

5. Gelman L, Staels B, Auwerx J. Rôle des co-facteurs transcriptionnels dans la transduction des signaux hormonaux par les récepteurs nucléaires. *Med Sci* 1997; 13: 961-70.

6. Taddei A, Almouzni G. Les acétyl-transférases et les désacétylases des histones: des co-régulateurs de la transcription. *Med Sci* 1997; 13: 1205-9.

7. Dhordain P, Albagli O. SMRT est un co-répresseur commun aux récepteurs nucléaires et aux protéines oncogéniques à domaine POZ, LAZ3 et PLZF. *Med Sci* 1998; 14: 219-22.

8. Arnould C, Philippe C, Bourdon V, Grégoire M, Berger R, Jonveaux P. The signal transducer and activator of transcription STAT5b gene is a new partner of retinoic acid receptor α in acute promyelocytic-like leukemia. *Hum Mol Genet* 1999; 8: 1741-9.

9. Vignais ML. Protéines JAK et STAT dans la transmission du signal cellulaire. *Med Sci* 1997; 13: 1277-84.

10. Wakao H, Gouilleux F, Groner B. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *EMBO J* 1994; 13: 182-91.

11. Morrighi R, Topham D, Teglund S, et al. Stat5 is required for IL-2-induced cell cycle progression of peripheral T cells. *Immunity* 1999; 10: 249-59.

12. Frank D. STAT signaling in the pathogenesis and treatment of cancer. *Mol Med* 1999; 5: 432-56.

13. Lacronique V, Boureux A, Valle VD, et al. A TEL-JAK2 fusion protein with constitutive kinase activity in human leukemia. *Science* 1997; 278: 1309-12.

Christophe Philippe
Cécile Arnould
Violaine Bourdon
Marie-José Grégoire
Philippe Jonveaux

Laboratoire de génétique, UPRES UA 952, CHRU, rue du Morvan, 54511 Vandœuvre-lès-Nancy, France.

Roland Berger

Inserm U. 434, SD 401 N° 434 Cnrs, CEPH, 27, rue Juliette-Dodu, 75010 Paris, France.



26^e SYMPOSIUM EUROPÉEN DES PEPTIDES

Montpellier, France
10-15 septembre 2000

- Le 26^e Symposium Européen des Peptides (26th EPS) aura lieu à Montpellier, France du 10 au 15 septembre 2000. C'est un événement biennal, qui regroupe plus d'un millier de personnes et qui est le congrès de référence dans le monde du Peptide (le dernier symposium, qui s'est déroulé en France, a été organisé par le Professeur Bricas en 1968). Il est organisé sous les auspices de la Société Européenne des Peptides (EPS) et, cette année, du Groupe Français des Peptides et Protéines (GFPP). L'organisateur, le Professeur Jean Martinez, vous attend à Montpellier.
- Un présymposium sur le suivi analytique des réactions organiques sur support solide aura lieu le samedi 9 septembre 2000 et est organisé par le Professeur Jean-Louis Aubagnac.
- Consultez notre site web pour toute information et inscription.

Site web : <http://ww2.pharma.univ.montp1.fr/26-EPS>

Date limite d'inscription : 1^{er} mars 2000

Development of a mutation screening strategy of the *MECP2* gene in Rett syndrome patients : Identification of an heterogeneous spectrum of mutations.

RESUME en anglais

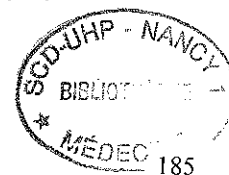
Rett syndrome (RTT) is a progressive neurodevelopmental disorder affecting almost exclusively females with an incidence of 1 in 15,000. Defined by accurate clinical criteria, RTT is characterized by a period of early normal growth and development followed by regression with loss of speech and acquired motor skills, stereotypical hand movements, and seizures. Although 99,5% of cases are sporadic, exclusion mapping using rare familial cases allowed the assignment of the gene responsible for RTT to Xq28. Systematic mutation analysis of candidate genes in the critical region resulted in the identification of mutations in the *MECP2* gene.

During this work, we carried out a mutation analysis of *MECP2* gene in 255 females with RTT and 15 mentally retarded boys. To identify *MECP2* mutations, we developed several complementary approaches (FISH, Southern blot, semi-quantitative PCR, CSGE-sequencing, then DHPLC-sequencing) and provided insights into an heterogeneous spectrum of mutations at the *MECP2* locus in RTT patients. Accordingly, we showed that in addition of point mutations, rearrangements involving several hundreds base pairs can be detected. Moreover and for the first time in RTT girls, we reported cases of somatic mosaicism. Finally, a study at the chromosome level of the *MECP2* gene did not reveal any deletion in RTT patients. We found 79% of mutations in RTT patients, our mutation detection rate is in agreement with the literature data.

Our mutation screening strategy of the *MECP2* gene is an essential molecular diagnosis tool to confirm the clinical diagnosis. The link between *MECP2* function (global repressor of methylated genes transcription) and disease expression is still unknown. Many studies are necessary to understand the physiological mechanism leading to RTT and eventually to develop therapeutic strategies.

MOTS CLES :

Rett syndrome, *MECP2* gene, mutations, somatic mosaicism, gross rearrangements



Nom : Mademoiselle BOURDON
Prénom : Violaine

DOCTORAT DE L'UNIVERSITÉ HENRI POINCARÉ, NANCY 1

en BIOLOGIE SANTÉ ENVIRONNEMENT

Spécialité : GÉNIE BIOLOGIQUE ET MÉDICAL

VU, APPROUVÉ ET PERMIS D'IMPRIMER

Nancy, le 24 juin 2002

Le Président de l'Université

