

2m11.3094.9

**Université de Montréal**

Variabilité génétique au niveau des gènes de réparation de l'ADN

Par  
Caroline Meloche

Faculté des études supérieures

Mémoire présenté à la Faculté des études supérieures  
En vue de l'obtention du grade de Maître en science (M.Sc.)  
En Biologie Moléculaire

Avril 2003

© Caroline Meloche, 2003



QH

506

U54

2003

V.010

**Direction des bibliothèques**

**AVIS**

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

**NOTICE**

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé :  
Variabilité génétique au niveau des gènes de réparation de l'ADN

Présenté par :  
Caroline Meloche

A été évalué par un jury composé des personnes suivantes :

Dr Abdellah Belmaaza

.....  
président-rapporteur

Dr Daniel Sinnett

.....  
directeur de recherche

Dr Francine Durocher

.....  
membre du jury

## RÉSUMÉ

La réparation de l'ADN joue un rôle fondamental dans le maintien de l'intégrité génétique et la protection contre le cancer. Nous proposons que les polymorphismes fonctionnels modifiant soit l'activité des protéines ou les niveaux de transcription des gènes codant pour des composantes de réparation de l'ADN pourraient influencer la susceptibilité inter-individuelle au cancer. Initialement, nous avons évalué, à l'aide d'études d'association, l'importance de polymorphismes présents dans les gènes XRCC1, XRCC3 et hOGG1 dans la susceptibilité à la leucémie lymphoblastique aiguë (LLA). Ensuite, nous avons procédé à l'identification de variants d'ADN dans la région régulatrice de 15 gènes de réparation des cassures doubles-brins chez 40 individus représentant diverses ethnies. L'impact fonctionnel des polymorphismes a été évalué *in silico* à l'aide de prédictions de gain/perte de sites de liaison pour des facteurs de transcription et *in vitro* par expériences EMSA. Finalement, nous avons mesuré l'impact de certains de ces polymorphismes dans la susceptibilité à la LLA. Nous avons montré que l'allèle Cys-326 du gène hOGG1 est associé avec une diminution de 42% du risque de développer la LLA. L'analyse des régions régulatrices a permis l'identification de 144 polymorphismes dont 73% modifient potentiellement la liaison à des facteurs de transcription. L'approche EMSA a permis de confirmer que des polymorphismes dans les gènes Ku70, Ku80, MRE11A et RAD50 mènent à une modification de la liaison à des facteurs de transcription. L'étude d'association de 5 variants du gène Ku70 nous a permis d'identifier 9 haplotypes dont 3 se retrouvent uniquement chez les patients LLA. Cette étude illustre l'importance de caractériser la variabilité génétique au niveau des gènes de réparation de l'ADN afin d'améliorer notre compréhension de l'étiologie des maladies complexes.

Mots clés : polymorphisme, région régulatrice, facteurs de transcription, leucémie lymphoblastique aiguë, cancer, génétique, facteurs de susceptibilité.

## SUMMARY

DNA repair plays a fundamental role in genetic integrity and cancer protection. We propose that functional polymorphisms modifying either the level of transcription or protein activity of genes part of the DNA repair machinery could influence the differences in inter-individual susceptibility to cancer. We have used association studies to evaluate the importance of polymorphisms present in genes XRCC1, XRCC3 and hOGG1 in acute lymphoblastic leukemia (ALL) susceptibility. We identified variants in the regulatory region of 15 genes part of the double-strand break repair pathway for 40 individuals picked from different ethnic groups. The functional impact of these polymorphisms has been evaluated *in silico* using transcription factor binding site gain and/or loss, and *in vitro* by EMSA. Finally, we measured the impact of some of these polymorphisms in ALL susceptibility. We have shown that the Cys-326 allele of the hOGG1 gene is associated with a 42% reduction in the risk of developing ALL. The analysis of regulatory regions allowed us to identify 144 polymorphisms, of which 73% potentially modify transcription factor binding sites. These modifications were confirmed by EMSA for the Ku70, Ku80, MRE11A and RAD50 genes. Association study done with 5 variants of the Ku70 gene identified 9 haplotypes, of which 3 were found only in ALL patients. This study demonstrates the importance of the characterization of genetic variants in DNA repair genes to increase our knowledge of complex disease aetiology.

**Key words:** polymorphism, regulatory region, transcription factors, acute lymphoblastic leukemia, cancer, genetic, susceptibility factors.

<b>TABLE DES MATIÈRES</b>	
<b>RÉSUMÉ</b>	<b>ii</b>
<b>SUMMARY</b>	<b>iii</b>
<b>TABLE DES MATIÈRES</b>	<b>iv</b>
<b>LISTE DES TABLEAUX</b>	<b>vii</b>
<b>LISTE DES FIGURES</b>	<b>ix</b>
<b>LISTE DES ABRÉVIATIONS</b>	<b>xi</b>
<b>REMERCIEMENTS</b>	<b>xv</b>
<b>CHAPITRE I</b>	<b>1</b>
<b>INTRODUCTION</b>	<b>1</b>
<b>1 CANCER</b>	<b>2</b>
1.1 INCIDENCE	2
1.2 GÉNÉTIQUE DU CANCER	2
1.2.1 Processus multi-étapes	2
1.2.2 Activation de proto-oncogènes	3
1.2.3 Inactivation de gènes suppresseurs de tumeur	7
1.3 LE CANCER : MALADIE COMPLEXE	10
<b>2. LA BIOTRANSFORMATION DES CARCINOGENES</b>	<b>13</b>
<b>3. RÉPARATION DE L'ADN</b>	<b>14</b>
3.1 RÉPARATION PAR EXCISION DE BASES (BER)	14
3.2 RÉPARATION PAR EXCISION DE NUCLÉOTIDES (NER)	15
3.3 RÉPARATION COUPLÉE À LA TRANSCRIPTION (TCR)	20
3.4 RÉPARATION DES MÉSAPPARIEMENTS (MMR)	20
3.5 RÉPARATION DES CASSURES DOUBLES-BRINS (DSBs REPAIR)	23
3.5.1 Réparation par recombinaison homologue (HRR)	24
3.5.2 Réparation non-homologue (NHEJ)	28
<b>4 NATURE D'UN POLYMORPHISME</b>	<b>30</b>
4.1 TYPE	30
4.2 EFFET D'UN POLYMORPHISME	31
4.2 DÉTECTION DE POLYMORPHISMES	32
4.2.1 Séquençage	32
4.2.2 SSCP	35
4.3 DHPLC	37

<b>5 INTER-RELATION ENTRE LE MÉTABOLISME DES XÉNOBIOTIQUES ET LA RÉPARATION DE L'ADN</b>	<b>41</b>
<b>6 LEUCÉMIE DE L'ENFANT</b>	<b>44</b>
6.1 INCIDENCE	44
6.2 CARACTÉRISTIQUES DE LA LLA	44
6.3 RÉPARATION DE L'ADN ET LEUCÉMOGÈNE	45
<b>7 RÉGULATION DE L'EXPRESSION GÉNÉRIQUE</b>	<b>46</b>
7.1 STRUCTURE GÉNÉRALE D'UN PROMOTEUR	46
7.2 LA GÉNÉTIQUE RÉGULATRICE	52
<b>8. HYPOTHÈSE</b>	<b>53</b>
<b>9 OBJECTIFS DU PROJET</b>	<b>54</b>
<b>CHAPITRE II</b>	<b>55</b>
<b>MATÉRIEL ET MÉTHODES</b>	<b>55</b>
<b>1 POPULATIONS ÉTUDIÉES</b>	<b>56</b>
1.1 POPULATION DE REPRÉSENTATION MONDIALE	56
1.2 LE GROUPE « CAS »	56
1.3 LE GROUPE « TÉMOIN »	58
<b>2 ÉTUDES D'ASSOCIATION</b>	<b>58</b>
<b>3 EXTRACTION D'ADN</b>	<b>59</b>
<b>4 DÉTECTION DE SNPs PAR HPLC DÉNATURANT</b>	<b>59</b>
4.1 DESSIN D'AMORCES	59
4.2 AMPLIFICATION PCR DE L'ADN GÉNOMIQUE	60
4.3 DHPLC	60
4.4 SÉQUENÇAGE	61
4.4 Identification de séquences conservées entre l'homme et la souris	61
4.5 PRÉDICTION <i>IN SILICO</i>	62
<b>5 HYBRIDATION SPÉCIFIQUE D'UN ALLÈLE ( ASO)</b>	<b>62</b>
5.1 PRINCIPE	62
5.2 TRANSFERT DES PRODUITS DE PCR SUR MEMBRANE (DOT BLOT)	64
5.3 MARQUAGE RADIOACTIF DES SONDAS ASO	64
5.4 HYBRIDATION	66
<b>6 GEL DE RETARDATION (EMSA)</b>	<b>67</b>
6.1 PRINCIPE	67
6.2 EXTRAITS NUCLÉAIRES	69
6.3 APPARIEMENT DES SONDAS	69
6.4 MARQUAGE DES SONDAS	72
6.5 RÉACTION DE LIAISON	72
6.6 ÉLECTROPHORÈSE	73
<b>CHAPITRE III</b>	<b>74</b>
<b>RÉSULTATS</b>	<b>74</b>



<b>1 OUTIL DE GENOTYPAGE</b>	<b>75</b>
<b>2 POLYMORPHISMES DANS LES GÈNES XRCC1 (ARG399GLN), XRCC3 (THR241MET) ET HOGG1 (SER326CYS)</b>	<b>75</b>
2.1 DÉTERMINATION DE LA FRÉQUENCE DE CES VARIANTS CHEZ LES PATIENTS LLA ET CHEZ LES TÉMOINS	75
2.2 DÉTERMINATION DES ALLÈLES ANCESTRAUX	76
<b>3 DÉTECTION DE POLYMORPHISMES DANS LES GÈNES DE RÉPARATION DES CASSURES DOUBLES-BRINS.</b>	<b>82</b>
3.1 DESCRIPTION DES POLYMORPHISMES IDENTIFIÉS	82
<b>4 ÉTUDE <i>IN VITRO</i> DE L'IMPACT DE RSNPs SUR LA LIAISON DE FACTEURS DE TRANSCRIPTION</b>	<b>90</b>
<b>5 ÉTUDE D'ASSOCIATION POUR LE GÈNE KU70</b>	<b>94</b>
5.1 DÉTERMINATION DE LA FRÉQUENCE DES VARIANTS DU GÈNE KU70	94
5.2 DÉTERMINATION DES ALLÈLES ANCESTRAUX	95
5.3 ANALYSE D'HAPLOTYPES	95
<b>CHAPITRE IV</b>	<b>103</b>
<b>DISCUSSION</b>	<b>103</b>
<b>CONCLUSION</b>	<b>118</b>
<b>BIBLIOGRAPHIE</b>	<b>119</b>
<b>ANNEXE I.</b>	<b>xvi</b>
<b>ANNEXE II .</b>	<b>xxi</b>
<b>ANNEXE III.</b>	<b>xxii</b>
<b>ANNEXE IV.</b>	<b>xxiv</b>
<b>ANNEXE V.</b>	<b>xxvi</b>
<b>CURRICULUM VITAE</b>	<b>lxv</b>

## LISTE DES TABLEAUX

Tableau I. Liste partielle d'oncogènes connus.	5
Tableau II. Liste partielle de gènes suppresseurs de tumeur.	9
Tableau III. Défaut de réparation dans la voie NER et TCR.	19
Tableau IV. Syndromes associés à un défaut de HRR.	27
Tableau V. Liste partielle d'étude d'association montrant un effet entre la présence de polymorphismes et le risque de développer un cancer.	43
Tableau VI: Composition des groupes continentaux étudiés	57
Tableau VII. Pouvoir de détection des variants.	57
Tableau VIII. Liste des sondes ASO et conditions d'hybridation.	65
Tableau IX. Liste des sondes utilisées pour le gel de retardement.	70
Tableau X. Fréquence allélique des gènes XRCC1, XRCC3 et hOGG1 chez les patients LLA et les témoins.	79
Tableau XI. Distribution des génotypes du gène XRCC1, XRCC3 et hOGG1 chez les patients LLA et chez les témoins.	80
Tableau XII. Allèles ancestraux.	81
Tableau XIII. Compilation des résultats obtenus pour la détection de rSNPs.	84
Tableau XIV. Description détaillée des polymorphismes identifiés dans le gène Ku70.	87
Tableau XV. Sites testés par gel de retardement et prédiction <i>in silico</i> de la présence de sites de liaison de facteurs de transcription.	91
Tableau XVI. Résultats des expériences de gel de retardement. + : présence d'un résultat concluant, - : absence de résultat concluant.	93
Tableau XVII. Distribution des génotypes chez les cas et les témoins.	96
Tableau XVIII. Allèles ancestraux.	96
Tableau XIX. Fréquence des haplotypes chez les patients et les témoins.	101
Tableau XX. Fréquence des haplotypes chez les individus d'origine différente.	101
Tableau XXI. Distribution des génotypes chez les cas et les témoins.	102
Tableau XXII : Liste des amorces utilisées pour l'amplification d'ADN génomique ainsi que les conditions de PCR.	xvi
Tableau XXIII : Conditions de PCR.	xxi
Tableau XXIV : Protocole pour l'amplification PCR.	xxii

	viii
Tableau XXV : Température d'analyse au dHPLC.	xxiv
Tableau XXVI. Description détaillée des polymorphismes identifiés dans le gène ATM.	xxvii
Tableau XXVII. Description détaillée des polymorphismes identifiés dans le gène BRCA1.	xxix
Tableau XXVIII. Description détaillée des polymorphismes identifiés dans le gène BRCA2.	xxxi
Tableau XXIX. Description détaillée des polymorphismes identifiés dans le gène DCLER1C.	xxxiv
Tableau XXX. Description détaillée des polymorphismes identifiés dans le gène Ku80.	xxxvii
Tableau XXXI. Description détaillée des polymorphismes identifiés dans le gène Lig4.	xli
Tableau XXXII. Description détaillée des polymorphismes identifiés dans le gène MRE11A.	xliv
Tableau XXXIII. Description détaillée des polymorphismes identifiés dans le gène NBS1.	xlix
Tableau XXXIV. Description détaillée des polymorphismes identifiés dans le gène PRKDC.	lii
Tableau XXXV. Description détaillée des polymorphismes identifiés dans le gène RAD50.	liv
Tableau XXXVI. Description détaillée des polymorphismes identifiés dans le gène RAD51.	lvi
Tableau XXXVII. Description détaillée des polymorphismes identifiés dans le gène XRCC2.	lix
Tableau XXXVIII. Description détaillée des polymorphismes identifiés dans le gène XRCC3.	lxi
Tableau XXXIX. Description détaillée des polymorphismes identifiés dans le gène XRCC4.	lxiii

## LISTE DES FIGURES

Figure 1. Processus multi-étapes du cancer du côlon.	4
Figure 2. Rôle des mutations somatiques spontanées dans le rétinoblastome une maladie chez l'enfant marquée par des tumeurs rétiniennes.	8
Figure 3. Carcinogénèse environnementale et facteurs de susceptibilité de l'hôte.	11
Figure 4. Mécanisme de cancérogenèse et mécanismes de protection.	12
Figure 5. Mécanisme de réparation par excision de bases.	16
Figure 6. Mécanisme de réparation par excision de nucléotides.	18
Figure 7. Mécanisme de réparation des mauvais appariements (MMR).	21
Figure 8. Représentation des voies de réparation des cassures doubles-brins.	29
Figure 9. Principe de séquençage.	34
Figure 10. Représentation schématique de la technique SSCP.	36
Figure 11. Représentation graphique de la formation des hétéroduplexes.	38
Figure 12. Résolution des hétéroduplexes.	38
Figure 13. Profil de dénaturation à 3 températures d'un fragment de 185 pb tel qu'évalué par WAVEMAKER.	38
Figure 14. Représentation d'une courbe de titration de températures.	40
Figure 15. Représentation hypothétique de l'inter-relation du métabolisme des xénobiotiques et de la réparation de l'ADN.	42
Figure 16. A. Modèle typique d'un gène avec les composantes jouant un rôle dans l'activation et l'inactivation de ce gène B. Activation d'un gène et assemblage du complexe de pré-initiation de la Pol II .	47
Figure 17. Éléments de séquence d'un cœur de promoteur.	49
Figure 18. Modèle d'interaction entre différents activateurs.	51
Figure 19. Méthode de génotypage par l'approche PCR-ASO	63
Figure 20. Gel de retardation.	68
Figure 21. Résultats d'hybridation par la méthode PCR-ASO.	77
Figure 22. Représentation graphique des allèles étudiés.	78
Figure 23. Effets des polymorphismes sur les sites putatifs de liaison de facteurs de transcription.	84
Figure 24. Légende pour les figures des représentations graphiques des régions promotrices des gènes étudiés.	85
Figure 25. Représentation graphique de la région régulatrice du gène Ku70.	86
Figure 26. Résultats de gel de retardation pour le polymorphisme à la position -1687 du gène Ku70.	92

Figure 27. Représentation des 9 haplotypes reconstruit par Phase.	97
Figure 28. Réseau des haplotypes chez les témoins	98
Figure 29. Réseau des haplotypes chez les cas.	99
Figure 30. Recombinaisons pouvant expliquer l'apparition des haplotypes H6 (A), H7 (B) et H8 (C).	100
Figure 31. Représentation graphique de la région régulatrice du gène ATM.	xxvi
Figure 32. Représentation graphique de la région régulatrice du gène BRCA1.	xxviii
Figure 33. Représentation graphique de la région régulatrice du gène BRCA2.	xxx
Figure 34. Représentation graphique de la région régulatrice du gène DCLER1C.	xxxiii
Figure 35. Représentation graphique de la région régulatrice du gène Ku80.	xxxvi
Figure 36. Représentation graphique de la région régulatrice du gène Lig4.	xl
Figure 37. Représentation graphique de la région régulatrice du gène MRE11A.	xliii
Figure 38. Représentation graphique de la région régulatrice du gène NBS1.	xlviii
Figure 39. Représentation graphique de la région régulatrice du gène PRKDC.	li
Figure 40. Représentation graphique de la région régulatrice du gène RAD50.	liii
Figure 41. Représentation graphique de la région régulatrice du gène RAD51.	lv
Figure 42. Représentation graphique de la région régulatrice du gène XRCC2.	lviii
Figure 43. Représentation graphique de la région régulatrice du gène XRCC3.	lx
Figure 44. Représentation graphique de la région régulatrice du gène XRCC4.	lxii

## LISTE DES ABRÉVIATIONS

°C : degré Celcius

μL : microlitre

μM : micromolaire

A : adénine

ADN : acide désoxyribonucléique

Ala : alanine

Arg: arginine

ASO : allele specific oligonucleotide

ATM : ataxia telangiectasia mutated

ATP : adénosine triphosphate

BER : base excision repair

BRCA1 : breast cancer 1

BRCA2 : breast cancer 2

BRE : élément de reconnaissance TFIIIB

BS : Bloom syndrome

C : cytosine

CS : Cockayne syndrome

cSNP : coding single nucleotide polymorphism

Cys: cytéine

DCLER1C : DNA cross-link repair 1C

ddATP : didésoxy adénosine triphosphate

ddNTP : didésoxynucléotide triphosphate

DHPLC : denaturing high pressure liquid chromatography

dNTP : désoxynucléotide triphosphate

DPE : downstream core promoter

DSB : double-strand break

DTT : Dithiothreitol

EDTA : ethylenediamine-tetraacetic acid

EGTA : ethylene glycol tetra acetic acid

EMSA : electromobility shift assay

ERCC1 : excision repair cross-complementing repair deficiency, complementation group 1

F-SSCP : fluorescent single stranded conformation polymorphism

G : guanine

Gln : glutamine

GNMT : glycine N-méthyltransférase

gSNP : genomic single nucleotide polymorphism

HCC : carcinome hépatocellulaire

HEPES : N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid

HNPCC : hereditary nonpolyposis colorectal cancer

hOGG1 : human 8-oxoguanine DNA glycosylase

HRR : homologous recombination repair

Inr : élément initiateur

iSNP : intronic single nucleotide polymorphism

kb : kilobase

KCl : chlorure de potassium

Ku70 : thyroid autoantigen 70kDa (Ku antigen)

Ku80 : thyroid autoantigen 80kDa (Ku antigen)

Lig4 : ligase 4

LLA : leucémie lymphoblastique aiguë

M : molaire

mA : milliampère

Mer : oligomer

Met: méthionine

MgCl<sub>2</sub> : chlorure de magnésium

min : minute

mL : millilitre

mM : millimolaire

MMR : mismatch repair

MRE11A : MRE11 meiotic recombination 11 homolog A (*S. cerevisiae*)

NaCl : chlorure de sodium  
NaH<sub>2</sub>PO<sub>4</sub> : sodium phosphate  
NaOH : hydroxyl de sodium  
NBS1 : Nijmegen breakage syndrome 1  
NER : nucleotide excision repair  
NHEJ : non-homologous end joining  
nM : nanomolaire  
OR : Odds ratio  
p: probabilité  
pb : paire de base  
PCNA : proliferating-cell nuclear antigen  
PCR : polymerase chain reaction  
PMSF : Phenylmethylsulfonylfluoride  
Pol : polymérase  
PRKDC : protein kinase catalytic subunit  
RAD : radiation  
RFC : DNA replication factor C  
RPA : Replication protein A  
rSNP : regulatory single nucleotide polymorphism  
s : seconde  
SDS : sodium dodecyl sulfate  
Ser : sérine  
SNP : single nucleotide polymorphism  
SSCP : single stranded conformation polymorphism  
T : thymine  
TBS : tris buffered saline  
TCR : transcription coupled repair  
TEAA : acétate triéthylammonium  
TF : transcription factor  
Thr : thréonine  
TTD : trichothiodystrophie



U : unité enzymatique

UV : ultraviolet

VNTR : variable number of tandem repeats

XP : Xeroderma pigmentosum

XRCC : X-ray repair complementing defective repair

## REMERCIEMENTS

J'aimerais remercier mon directeur de recherches, Daniel Sinnett, pour m'avoir accueilli dans son laboratoire ainsi que pour son aide tout au long de ma maîtrise. J'aimerais également remercier tout le monde à l'étage B du pavillon vidéotron que ce soit pour les discussions constructives durant ces 2 années autant que pour l'ambiance très agréable au labo et très divertissante pendant les partys! J'aimerais remercier plus particulièrement toutes les personnes qui m'ont aidée pendant mon projet : Patrick, Simon, Josiane, Rocio, Hélène, Sarah, Marie-Christine, Sylvie, Sandra, Dominik, Vania, Stéphane, Dominika, Nathalie et Chantal. Merci également à Gino, Zeina et Alexandre qui m'ont supportée moralement.

Finalement, j'aimerais remercier ma mère et mon père pour m'avoir encouragée à poursuivre mes études et qui continue de le faire. Sans vous, je ne serais pas rendue jusqu'ici!

Merci au CRSNG pour le support financier.

**CHAPITRE I**  
**INTRODUCTION**

## **1 Cancer**

### **1.1 Incidence**

Dans les pays industrialisés, le cancer est une des maladies les plus répandues. On estime que 136 900 nouveaux cas de cancer et 66 200 décès causés par cette maladie surviennent annuellement au Canada (Statistiques canadiennes sur le cancer, 2002). L'incidence du cancer est toujours en progression et de façon plus marquée chez la femme. Toutefois, le taux de mortalité excepté pour le cancer du poumon chez la femme est en baisse. De plus, depuis le début des années 50, les taux de mortalité par les cancers de l'enfant ont reculé de plus de 50%, les baisses les plus importantes étant survenues après 1970.

### **1.2 Génétique du cancer**

#### **1.2.1 Processus multi-étapes**

Les analyses génétiques nous ont permis d'apprendre beaucoup de choses sur la biologie normale et sur les maladies en étudiant les propriétés des mutations qui mènent à des perturbations de processus normaux. C'est particulièrement vrai dans le cas du cancer. Il est maintenant clair que le cancer est dû à une série de mutations qui s'accumulent dans la cellule. La tumorigenèse humaine est un processus multi-étapes complexe impliquant des altérations dans les proto-oncogènes et les gènes suppresseurs de tumeur. L'effet de ces mutations s'exprime de différentes façons: augmentation de l'habileté des cellules à proliférer, diminution de la susceptibilité des cellules à l'apoptose ou augmentation du taux général de mutations de la cellule qui favorise l'apparition de mutations menant à une dysfonction de la prolifération ou de l'apoptose. Des études statistiques sur l'âge de l'incidence des tumeurs solides humaines indiquent que 5 à 6 événements mutationnels indépendants contribuent à la formation d'une tumeur (Peto, R. 1977). Dans les leucémies, seulement 3 ou 4 mutations dans différents gènes peuvent être nécessaires. L'étude de la carcinogénèse chimique chez la souris a permis de comprendre la nature multi-étapes du cancer

(Weinberg, RA.1989). Depuis, les études d'un bon nombre de cancers ont clairement confirmé la nature multi-étapes du cancer (figure 1) .

### 1.2.2 Activation de proto-oncogènes

Les proto-oncogènes codent pour des protéines qui sont impliquées dans le contrôle de la croissance cellulaire. L'altération de la structure et/ou de l'expression des proto-oncogènes peut mener à leur activation et devenir des oncogènes capables d'induire la transformation cellulaire. Les oncogènes peuvent être classés en 5 groupes selon leurs propriétés fonctionnelles et biochimiques (tableau I) : (1) les facteurs de croissance, (2) les récepteurs de facteurs de croissance, (3) les transducteurs de signaux, (4) les facteurs de transcriptions et (5) les phosphorylateurs de protéines, incluant les régulateurs de la mort programmée cellulaire. L'activation des oncogènes peut s'effectuer selon trois mécanismes (Bishop, JM. 1991) qui résultent en l'altération de la structure des proto-oncogènes ou en l'augmentation d'expression de ces proto-oncogènes :

- 1- Mutations : substitutions de bases, des délétions et des insertions sont capables d'activer des proto-oncogènes. Les mutations ponctuelles sont fréquemment détectées dans la famille ras des proto-oncogènes (Rodenhuis, S. 1992).
- 2- L'amplification génique : expansion du nombre de copies d'un gène à l'intérieur du génome d'une cellule. Des études ont montré que 3 familles de proto-oncogènes, myc, erb B et ras, sont amplifiées dans un nombre important de tumeurs (<http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowSection&rid=cmed.section.d1e41591>).

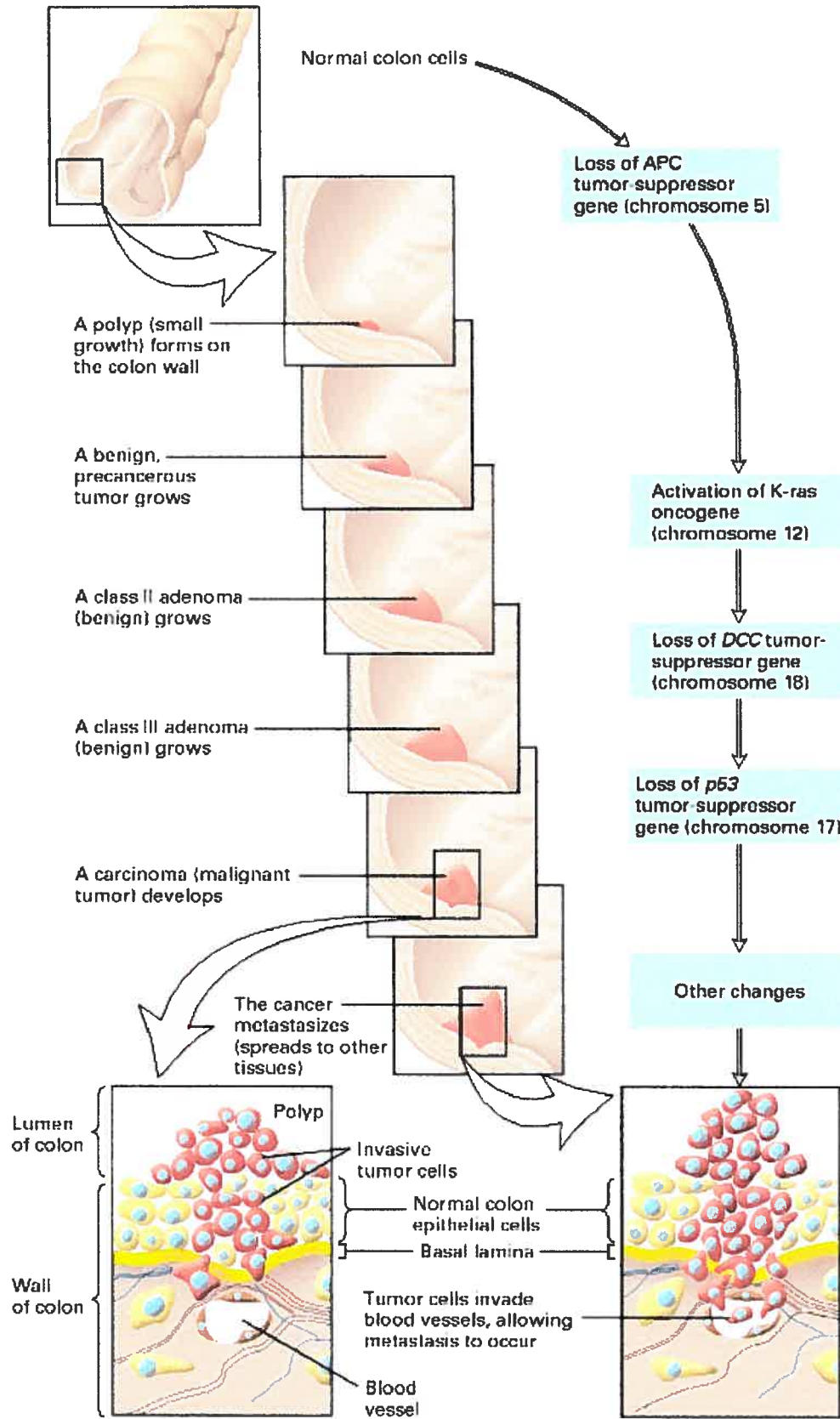


Figure 1. Processus multi-étapes du cancer du côlon.

Tableau I. Liste partielle d'oncogènes connus

Oncogène	Chromosome	Néoplasme	Mécanisme d'activation	Fonction protéique
<i>Facteurs de croissance</i>				
v-sis	22q12.313.1	Gliome/fibrosarcome	Production constitutive	Chaîne B de FGF
int2	11q13	Carcinome mammaire	Production constitutive	Membre de la famille FGF
KS3	11q13.3	Sarcome de Kaposi	Production constitutive	Membre de la famille FGF
HST	11q13.3	Carcinome de l'estomac	Production constitutive	Membre de la famille FGF
<i>Récepteurs de facteurs de croissance</i>				
Tyrosine kinases: protéines membranaires				
ErbB	7p1.11.3	Carcinome des cellules squameuses	Amplification génique	Récepteur EGF
Fms	5q33-34 (FMS)	Sarcome	Activation constitutive	Récepteur CSF-1
<i>Transducteurs de signaux</i>				
Tyrosines kinases cytoplasmiques				
SRC	20p1213	Carcinome du colon	Activation constitutive	Protéine tyrosine kinase
ABL	9q34.1	CML	Réarrangement de l'ADN (activation constitutive/protéine de fusion)	Protéine tyrosine kinase
Protéines G associées à la membrane				
H-RAS	11p15.5	Carcinomes du colon, poumon, pancréas	Mutation ponctuelle	GTPase
K-RAS	12p11.112.1	AML, carcinome de la thyroïde, mélanome	Mutation ponctuelle	GTPase
N-RAS	1p1113	Carcinome, mélanome	Mutation ponctuelle	GTPase
Sérine/thréonine kinases: cytoplasmique				
v-mos	8q11 (MOS)	Sarcome	Activation constitutive	Protéine kinase (ser/thr)
v-raf	3p25 (RAF-1)	Sarcome	Activation constitutive	Protéine kinase (ser/thr)
<i>Facteurs de transcription</i>				
v-myc	8q24.1 (MYC)	Carcinome myélocytomateuse	Dérèglement de l'activité	Facteur de transcription
N-MYC	2p24	Neuroblastome	Dérèglement de l'activité	Facteur de transcription
L-MYC	1p32	Carcinome du poumon	Dérèglement de l'activité	Facteur de transcription
v-fos	14q2122	Ostéosarcome	Dérèglement de l'activité	Facteur de transcription API
v-jun	p3132	Sarcome	Dérèglement de l'activité	Facteur de transcription API
v-ski	1q2224	Carcinome	Dérèglement de l'activité	Facteur de transcription
v-rel	2p1214	Leucémie lymphatique	Dérèglement de l'activité	Mutant NFkB
v-ets-1	11p23q24	Érythroblastose	Dérèglement de l'activité	Facteur de transcription
v-ets-2	21q24.3	Érythroblastose	Dérèglement de l'activité	Facteur de transcription
v-erbA1	17p1121	Érythroblastose	Dérèglement de l'activité	Facteur de transcription T3
v-erbA2	3p2224.1	Érythroblastose	Dérèglement de l'activité	Facteur de transcription T3
<i>Autres</i>				
BCL2	18q21.3	Lymphomes à cellules B	Activation constitutive	Protéine antiapoptotique
MDM2	12q14	Sarcome	Amplification génique/augmentation de la protéine	Complexe avec p53

- 3- Les réarrangements chromosomiques : détectés dans les malignités hématologiques ainsi que dans quelques tumeurs solides (Croce, CM. 1987; Solomon, E. *et al.* 1991; Rabbits, TH. 1994).

Ces réarrangements consistent principalement en translocations et moins fréquemment en inversions chromosomiques. Ces réarrangements peuvent causer des malignités par deux mécanismes différents : (1) l'activation de la transcription de proto-oncogènes ou (2) la création de protéines de fusion. L'activation de la transcription d'un proto-oncogène se produit lorsque le réarrangement chromosomique place un proto-oncogène près d'un récepteur des immunoglobulines ou des récepteurs des cellules T. La transcription du proto-oncogène tombe sous le contrôle du promoteur de ces gènes. La translocation t(8;14)(q24;q32), trouvée dans environ 85% des cas de lymphomes de Burkitt est un exemple bien caractérisé d'activation de la transcription d'un proto-oncogène. Le réarrangement place le gène c-myc, situé sur le chromosome 8q24 sous le contrôle des éléments régulateurs du locus de la chaîne lourde des immunoglobulines situés à 14q32 (Dalla, R. *et al.* 1982). La fusion de gènes mène à la production de chimère. En général, les deux gènes contribuent au pouvoir transformant de la protéine chimérique produite. Le premier exemple de gène de fusion a été découvert en clonant le point de cassure du chromosome de Philadelphie dans la leucémie myélocytaire chronique (LMC) (Groffen, J. *et al.* 1984). La translocation t(9;22)(q34;q11) mène à la fusion de c-abl et bcr. La protéine de fusion a une activité tyrosine kinase augmentée et a une localisation cellulaire anormale. Puisque la néoplasie est un processus multi-étapes, plus d'un de ces mécanismes contribuent à la genèse de tumeurs en altérant un nombre de gènes associés au cancer. L'expression complète du phénotype néoplasique, incluant la capacité à métastaser, implique une combinaison d'activation de proto-oncogènes et l'inactivation ou la perte de gènes suppresseurs de tumeur.



### 1.2.3 Inactivation de gènes suppresseurs de tumeur

La notion de gène suppresseur de tumeur découle de l'hypothèse de Knudson (1971) qui proposait que la transformation d'une cellule normale en une cellule cancéreuse nécessitait au moins deux mutations, soit l'hypothèse des « two-hits ». Dans le cas de la forme héréditaire du rétinoblastome, une copie du gène défectueux, *Rb-1*, est héritée des parents et est présente dans toutes les cellules. Environ 80% des porteurs développent une tumeur dans les deux yeux avant l'âge de 5 ans, 15% dans un seul œil tandis que 5% sont des porteurs asymptomatiques. Les analyses subséquentes ont démontré que le deuxième allèle de *Rb-1* était muté dans les cas de rétinoblastome (Yandell, DW. *et al.* 1989). Étant donné que ces mutations mènent à l'inactivation complète de *Rb-1* le terme anti-oncogène ou gène suppresseur de tumeur a été utilisé pour décrire ce type de gène impliqué dans le cancer (figure 2). Le tableau II contient une liste partielle de gènes suppresseurs de tumeur. On peut remarquer que les gènes suppresseurs de tumeurs jouent des rôles dans la régulation du cycle cellulaire, la régulation de la transcription, l'apoptose, la réparation de l'ADN, etc. Les gènes impliqués dans la réparation de l'ADN forment un sous-groupe de gènes suppresseurs de tumeur. Comme les autres, ils sont inactivés dans certains types de cancer. Contrairement aux gènes suppresseurs de tumeurs classiques qui sont directement impliqués dans l'inhibition de la croissance ou la différenciation, l'inactivation de plusieurs des protéines de réparation dans les cellules tumorales résulte en une augmentation du taux de mutations (phénotype mutateur) dans d'autres gènes cellulaires, incluant les proto-oncogènes et les autres gènes suppresseurs de tumeurs (<http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowSection&rid=cmed.chapter.d1e48373>).

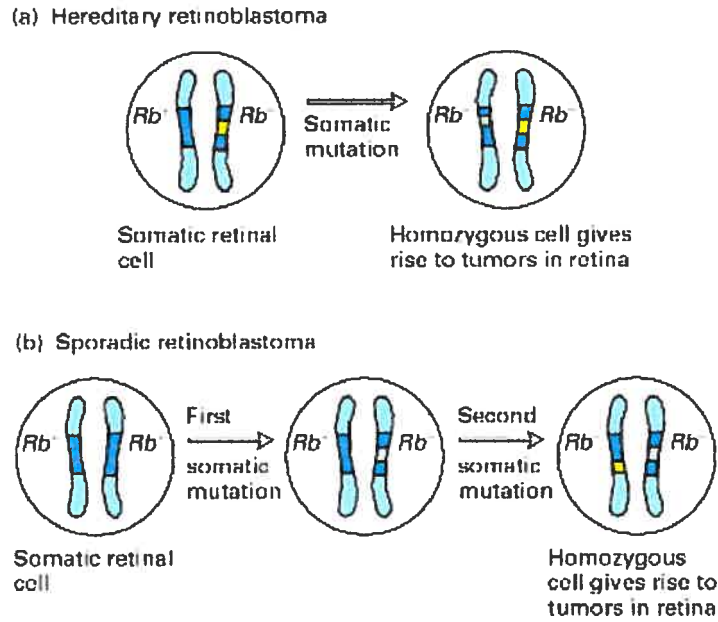


Figure 2. Rôle des mutations somatiques spontanées dans le rétinoblastome, une maladie chez l'enfant marquée par des tumeurs rétinienne. Les tumeurs surviennent dans les cellules de la rétine qui portent deux allèles du gène *Rb* muté. (a) Dans le rétinoblastome héréditaire, l'enfant reçoit un allèle normal  $Rb^+$  d'un parent et un allèle muté  $Rb^-$  de l'autre parent. Une mutation dans une cellule somatique hétérozygote qui inactive l'allèle normal va produire une cellule homozygote pour deux allèles mutés  $Rb^-$ . (b) Dans le rétinoblastome sporadique, l'enfant reçoit deux allèles normaux  $Rb^+$ . Deux mutations somatiques différentes, inactivant les deux allèles dans une cellule particulière, sont requises pour produire une cellule homozygote  $Rb^-/Rb^-$ .

Tableau II. Liste partielle de gènes suppresseurs de tumeur

Gène	Syndrome cancéreux héréditaire associé	Cancers avec mutations somatiques	Fonction de la protéine
<i>RB1</i>	Rétinoblastome familial	Rétinoblastome, ostéosarcome, SCLC, sein, prostate, vessie, pancréas, oesophage, autres	Régulateur de la transcription; liaison avec E2F
<i>TP53</i>	Syndrome Li-Fraumeni	Approx. 50% de tous les cancers	Facteur de transcription; régule le cycle cellulaire et l'apoptose
<i>p16</i>	Mélanome familial, Carcinome pancréatique familial	Approx. 25-30% de plusieurs types de cancer (e.g., sein, poumon, pancréas, vessie)	Inhibiteur des cyclines-dependantes kinases (i.e., Cdk4 and Cdk6)
<i>p19<sup>ARF</sup></i>	?	Approx. 15% de plusieurs types de cancer	Régule la stabilité de Mdm-2 et augmente la stabilité de p53
<i>APC</i>	FAP, syndrome Gardner, syndrome de Turcot	Colorectal, gastrique, endometrial	Régule les niveaux de b-caténine dans le cytosol; liaison aux microtubules
<i>BRCA1</i>	Cancer héréditaire du sein et ovarien	Ovarien (~10%), rare dans le cancer du sein	Réparation de l'ADN; forme un complexe avec Rad 51 et BRCA2; régulation de la transcription
<i>BRCA2</i>	Cancer héréditaire du sein, cancer pancréatique	Mutation rare dans le cancer pancréatique	Réparation de l'ADN; forme un complexe avec Rad 51 et BRCA1
<i>WT-1</i>	WAGR, syndrome Denys-Drash	Tumeur de Wilm	Facteur de transcription
<i>NF-1</i>	Neurofibromatose type 2	Mélanome, neuroblastome	p21ras-GTPase
<i>PTEN/SMAC1</i>	syndrome Cowden; cas sporadique de cancer du syndrome juvénile polypose	Gliome, sein, prostate, carcinome folliculaire de la thyroïde, carcinome de la tête et du cou	Phosphoinositide 3-phosphatase; protéine tyrosine phosphatase
<i>MSH2, MLH1, PMS1, PMS2, MSH6</i>	Hereditary non-polyposis colorectal cancer (HNCC)	Colorectal, gastrique, endometrial	Réparation des mauvais appariements de base

### 1.3 Le cancer : maladie complexe

Des études épidémiologiques ont montré que le risque de cancer diffère non seulement entre les populations, mais aussi sur une base individuelle. Cette variabilité inter-individuelle dans la susceptibilité au cancer peut être due à plusieurs facteurs : race, genre, état de santé, âge, diète et la variabilité génétique qui est à la base de mon projet (figure 3). En effet, la réponse de l'hôte face à son environnement dépend notamment de polymorphismes fonctionnels dans les gènes impliqués dans le métabolisme des carcinogènes ainsi que les enzymes impliquées dans la réparation de l'ADN. Certaines études ont mis en évidence l'effet de l'exposition à des carcinogènes chimiques comme le benzène, la peinture, les pesticides et certains médicaments anticancéreux sur le taux de cancers pédiatriques (McBride, ML. 1998). Le lien entre la variation inter-individuelle dans le métabolisme des carcinogènes et la formation d'adduits au niveau des macromolécules a été établi, il y a de ça plusieurs années (Harris, CC. 1990). Par conséquent, si l'efficacité du métabolisme des carcinogènes ou l'efficacité du système de réparation est diminuée, le fardeau mutationnel augmentera (phénotype mutateur) et aura donc un impact sur la probabilité d'activer les proto-oncogènes ou inactiver les gènes suppresseurs de tumeur d'où une élévation du risque d'initier la cancérogenèse (Figure 4).

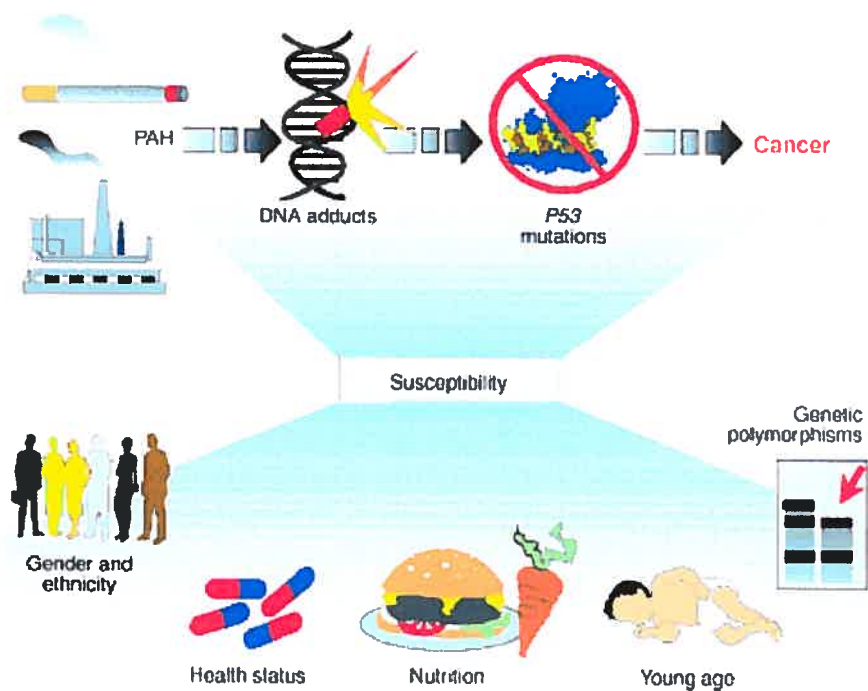


Figure 3. Carcinogénèse environnementale et facteurs de susceptibilité de l'hôte.

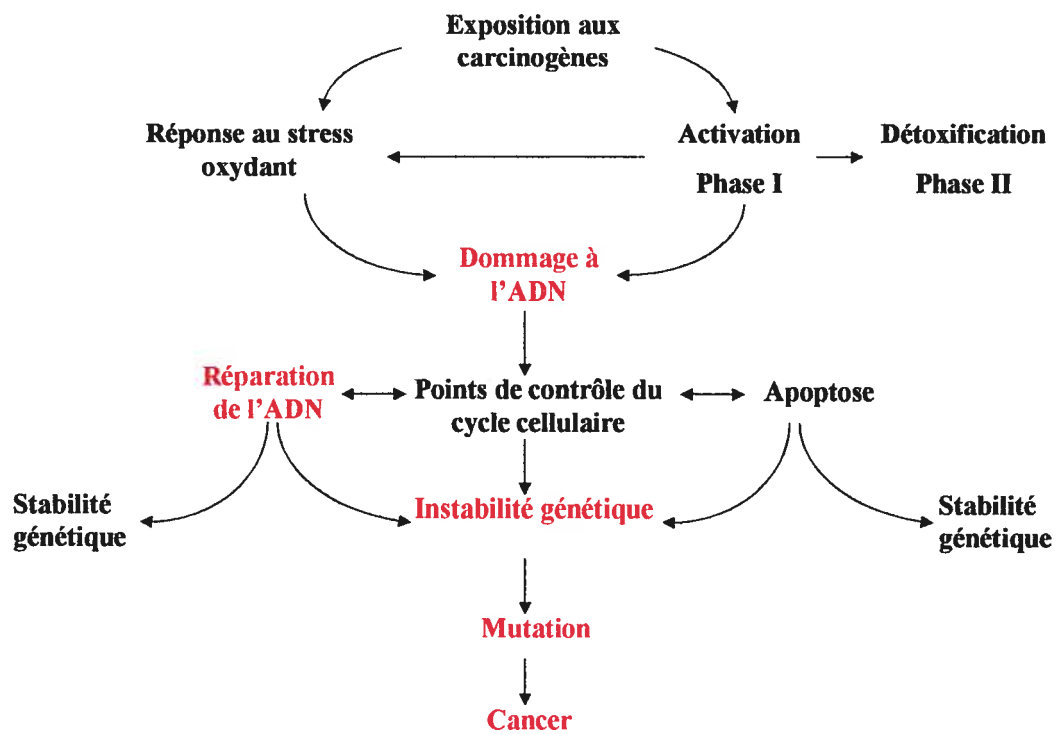


Figure 4. Mécanisme de cancérogenèse et mécanismes de protection.

## 2. La biotransformation des carcinogènes

Les facteurs environnementaux reconnus pour causer le cancer chez l'humain incluent :

- 1- les substances chimiques comme les organochlorés (pesticides), les solvants organiques (benzène) et les hydrocarbures aromatiques polycycliques produits lors de la combustion de matières organiques (fumée de cigarette, cuisson des aliments, etc)
- 2- les facteurs physiques comme les UV ou les radiations ionisantes

L'exposition aux carcinogènes requiert généralement une activation métabolique pour interagir avec des macromolécules comme les protéines et l'ADN. Les processus d'activation et de conjugaison sont gouvernés par les enzymes de la phase I et de la phase II respectivement. L'activation métabolique des carcinogènes est gouvernée par les enzymes de la phase I (par exemple cytochromes p450) qui catalysent des réactions d'hydrolyse, de réduction et d'oxydation. Les enzymes de la phase II (par exemple glutathione-S-transférases) contribuent à la détoxification des substances activées pendant la phase I (figure 4). Les enzymes de la phase II catalysent des réactions de glucuronidation, sulfatation, acétylation, méthylation et conjugaison avec le glutathion ou les acides aminés. Ces réactions transforment les molécules activées en produits moins réactifs et solubles facilitant ainsi leur élimination de l'organisme par l'urine ou la bile (Lai, C et Shields, P.G. 1999). Certains polymorphismes modifiant l'activité biochimique de ces enzymes du métabolisme des carcinogènes ont fait l'objet de nombreuses études (Krajnovick, M. *et al.* 2001).

### 3. Réparation de l'ADN

Le maintien de l'intégrité du matériel génétique est d'une importance capitale pour la survie d'un organisme. Il est démontré que dans une cellule de mammifère métaboliquement active, il survient jusqu'à 10 000 lésions à l'ADN chaque jour ce qui est incompatible avec la vie si elles ne sont pas éliminées (Lindahl, T. 1993). Par conséquent, les cellules eucaryotes possèdent des mécanismes pour repérer, signaler et réparer ces dommages et ainsi protéger le génome contre l'apparition de mutations pouvant mener à des maladies comme le cancer. Plusieurs agents endommagent l'ADN, incluant les rayons UV, les radiations ionisantes, les produits réactifs de l'oxygène et une variété de composés chimiques endogènes et exogènes. Les lésions induites par ces agents sont différentes incluant des bris simple-brin et doubles-brins de l'ADN, l'altération de bases, la formation d'adduits, des liaisons inter- et intra-brin et autres dommages (Friedberg, EC. *et al.* 1995). De plus, lors de la réplication de l'ADN des erreurs peuvent survenir et causer des mauvais appariements et des délétions. Il y a trois principales voies pour éliminer ces lésions : (1) la réparation par excision de bases (BER), (2) par excision de nucléotides (NER) et (3) la réparation des mauvais appariements. De plus, il existe deux autres voies de réparation pour les cassures doubles-brins de l'ADN soit la réparation par recombinaison homologue (HRR) et la réparation non-homologue (NHEJ).

#### 3.1 Réparation par excision de bases (BER)

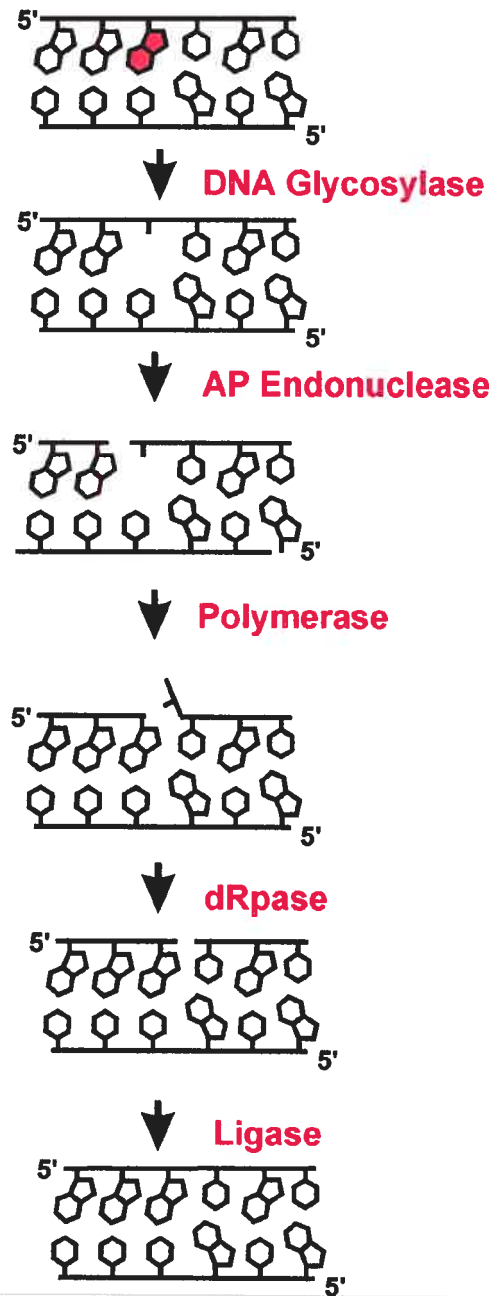
La réparation par excision de bases (figure 5) élimine principalement les adduits qui ne déforment pas l'ADN comme ceux produits par les agents méthylants ou les bases oxydées, réduites ou fragmentées causées par les radiations ionisantes et les dommages oxydants. Il y a deux principaux événements qui surviennent dans ce type de réparation. Premièrement, il y a hydrolyse du lien N-glycosylique qui lie la base endommagée à la chaîne désoxyribose-phosphate créant ainsi un site AP (apurinique ou apyrimidique). La classe d'enzyme responsable de cette étape est appelée



glycosylase. Il existe au moins 5 glycosylases, dont la glycosylase 8-hydroxyguanine (hOGG1) qui contribue à l'élimination des lésions 7,8-dihydro-8-oxoguanine (8-oxoG) qui sont les lésions les plus abondantes induites par les radiations ionisantes et le stress oxydant (Sancar, A. 1995). La seconde étape est le clivage du sucre abasique résultant par une AP endonucléase en 5' du site AP engendrant ainsi des résidus 3'OH et 5'désoxyribose phosphate. Jusqu'à maintenant, une seule endonucléase HAP-1 a été rapportée chez l'humain (Wood, R. D. 1996). L'achèvement de la réparation nécessite l'élimination du résidu désoxyribose-phosphatidique en 5' généré par l'AP endonucléase. Cette étape est réalisée par l'activité phosphodiesterase de l'ADN polymérase  $\beta$ . La brèche est ensuite remplie par l'ADN polymérase  $\beta$  et ensuite scellée par l'ADN ligase III. XRCC1 joue également un rôle dans la BER. Cette protéine à multidomaines interagit avec l'ADN et avec trois enzymes, la polymérase poly ADP-ribose (PARP), l'ADN ligase III et l'ADN polymérase  $\beta$  toutes trois impliquées dans cette voie de réparation (Kim, S-U. *et al.* 2002). XRCC1 jouerait le rôle de protéine adaptatrice et il a été démontré que XRCC1 maintient la stabilité et le niveau de l'ADN ligase III *in vivo* (Caldecott, K. *et al.* 1994).

### 3.2 Réparation par excision de nucléotides (NER)

La réparation par excision de nucléotides (figure 6) intervient sur les lésions encombrantes ou qui provoquent une distorsion importante de l'hélice d'ADN. Elle élimine les adduits formés par l'action des carcinogènes chimiques comme le cisplatine et des photoproduits de l'ADN engendrés par les rayons ultraviolets de type C (UV-C) (Wood, RD. 1996; Lindhal, T. *et al.* 1997). Chez l'homme, le système NER est le mécanisme principal impliqué dans la réparation des dommages produits par les UV. Cette voie de réparation implique la participation des produits d'au moins 30 gènes. Il a été établi que l'étape initiale de reconnaissance du dommage est la liaison de XP-C qui recrute les autres composantes de cette voie de réparation (Sugasawa, H. *et al.* 1998). XP-E reconnaît les lésions cyclobutane pyrimidine puisqu'il possède une haute affinité pour les dommages causés par les UV (Keeney, S. *et al.* 1993). La protéine XP-A est importante pour la vérification des dommages et



par

Figure 5. Mécanisme de réparation excision de bases. ([web.mit.edu/bevin/www/alkback.shtml](http://web.mit.edu/bevin/www/alkback.shtml))

pour l'organisation de l'appareillage de réparation autour de la lésion (Eker, AP. *et al* 1992). XP-A interagit avec RPA, ERCC1 et TFIIH qui participent également à la NER. Plusieurs hélicases dont XP-B et XP-D participent à la reconnaissance et à l'élimination des lésions. XP-B et XP-D font parties des sous-unités de TFIIH (Schaeffer, L. *et al.* 1993; Schaeffer, L. *et al.* 1994; Lehmann, AR. 1998). La fonction de ces hélicases est d'ouvrir l'hélice d'ADN autour du dommage pour permettre aux nucléases d'inciser l'ADN endommagé. Ensuite, un oligonucléotide de 30 nucléotides est excisé. Deux incisions sont produites sur le brin endommagé, une de chaque côté de la lésion. Deux nucléases différentes créent ces incisions. ERCC1-XPF et XP-G sont deux nucléases ayant une spécificité différente. XP-G coupe en 3' de la région simple-brin de l'ADN et ERCC1-XPF coupe en 5' (O'Donovan, A. *et al.* 1994). Le résultat est l'élimination de la portion d'ADN endommagé. La fonction de RPA dans ce type de réparation est de stabiliser l'ADN autour de la lésion en coopération avec d'autres enzymes. La dernière étape est le remplissage de la brèche. Des études *in vitro* ont démontré que les facteurs de réplication RPA, RF-C ainsi que PCNA et les ADN polymérase  $\delta$  et  $\epsilon$  (Hubscher, U. et Thommes, P. 1992; Shivji, MK. *et al.* 1995) sont requis pour une synthèse efficace. Ensuite, la ligase I scelle la brèche réparée (Lindahl, T. et Barnes, DE. 1992). Trois désordres génétiques peu fréquents à transmission autosomale récessive sont associés à une déficience dans la voie de réparation NER : Xeroderma pigmentosum (XP), le syndrome de Cockayne (CS) et la forme photosensible de trichothiodystrophie (TTD). Ces désordres montrent des phénotypes cliniques distincts (Tableau III).

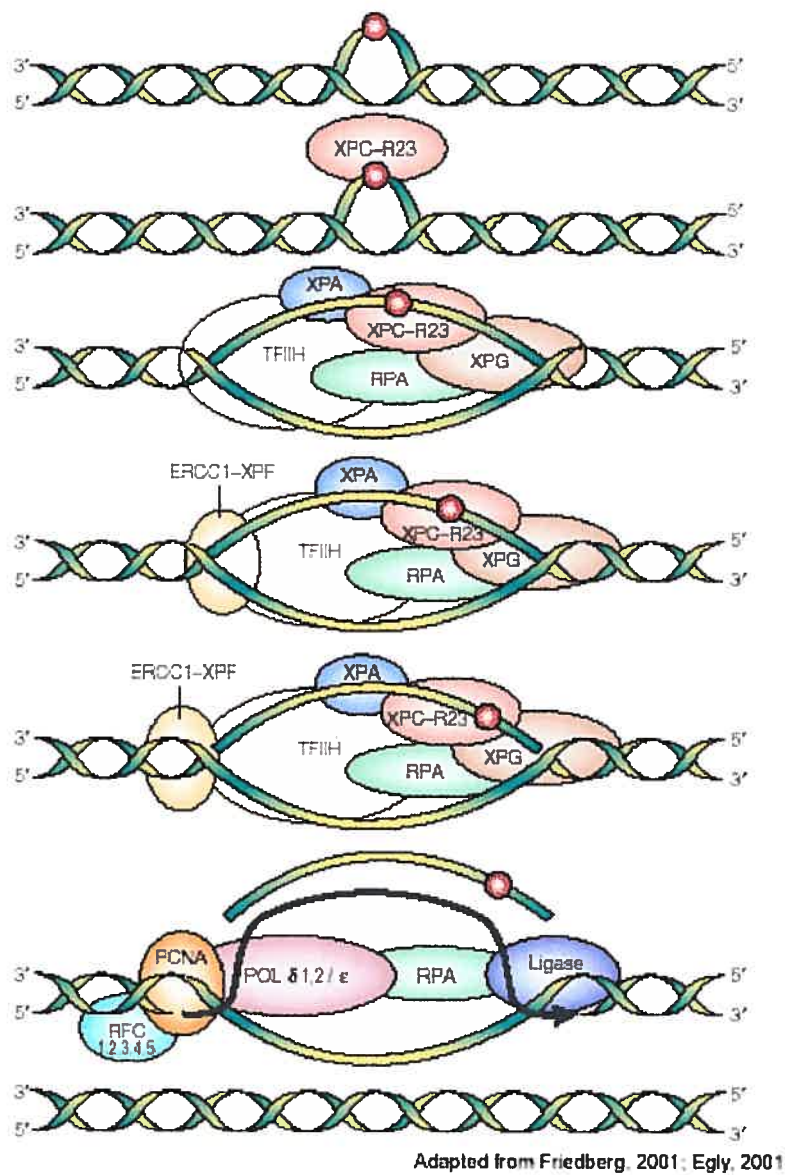


Figure 6. Mécanisme de réparation par excision de nucléotides.

Tableau III. Défaut de réparation dans la voie NER et TCR

<b>Syndrome</b>	<b>Gène défectueux</b>	<b>Voie touchée</b>	<b>Phénotype</b>	<b>Références</b>
Xeroderma pigmentosum (XP)	XP-A, -B, -D, -F, -G	NER + TCR (NER)	extrême photosensibilité pigmentation anormale de la peau incidence du cancer de la peau très élevée	Kramer <i>et al.</i> 1987
Syndrome Cockayne (CS)	CS-A, CS-B (90%) XP-B, -D, -G	TCR (NER +BER) TCR (NER +BER)	Retard mental Retard de croissance neurodégénération dégénération rétinienne	Nance et Berry 1992
Forme photosensible de trichothiodystrophie (TTD)	XP-B XP-D autre sous-unité de TFIIH*	TCR (NER +BER) TCR (NER +BER) Transcription	Photosensibilité Retard mental Retard de croissance Cheveux cassants et déficients en sulfure	Itin et Pittelkow 1990

\*pas de photosensibilité

### 3.3 Réparation couplée à la transcription (TCR)

Une des causes les plus communes à l'arrêt de l'ARN polymérase est la présence de lésions sur le brin transcrit de gènes actifs (Tornaletti, S. et Hanawalt, PC. 1999). Le blocage d'une ARN polymérase II au niveau d'un gène essentiel est suffisant pour mener à la mort de la cellule affectée. Pour contourner ce problème, la réparation des dommages dans les gènes activement transcrits se produit plus rapidement que dans les régions non transcrites du génome. Cette réparation plus rapide n'est pas seulement due à la plus grande accessibilité de l'ADN dans les régions activement transcrites puisqu'il a été démontré que c'est spécifiquement le brin transcrit qui est réparé préférentiellement (Mellon, I. *et al.* 1987). Le brin non-transcrit du même gène est réparé à un taux similaire aux régions non transcrites. Initialement, il a été proposé que la TCR était un sous-mécanisme de la NER, mais maintenant il est connu que plusieurs lésions éliminées par BER sont réparées via la TCR (Svejstrup, J. 2002). La première étape est la reconnaissance de l'arrêt de l'ARN polymérase II. Cette étape est effectuée par les protéines CSA et CSB. Ensuite, l'ARN polymérase II est enlevée et TFIIH est recruté au niveau de la lésion. Par la suite, la lésion est réparée soit par la voie NER ou BER selon le type de lésion.

### 3.4 Réparation des mésappariements (MMR)

Les cellules eucaryotes disposent d'un système de réparation similaire à celui d'*E. coli* pour corriger les mésappariements d'un seul nucléotide ainsi que les petites boucles (jusqu'à 4 nucléotides mésappariés) qui surviennent lors de la réplication, de la recombinaison et de la désamination des cytosines (Modrich, P. et Lahue, R. 1996; Kolodner, R. 1996). L'hétérodimère MSH2/MSH6 (MutS $\alpha$ ) reconnaît un mésappariement d'un nucléotide ou une boucle de un à deux nucléotides et l'hétérodimère MSH2/MSH3 (MutS $\beta$ ) reconnaît une boucle de plus de deux

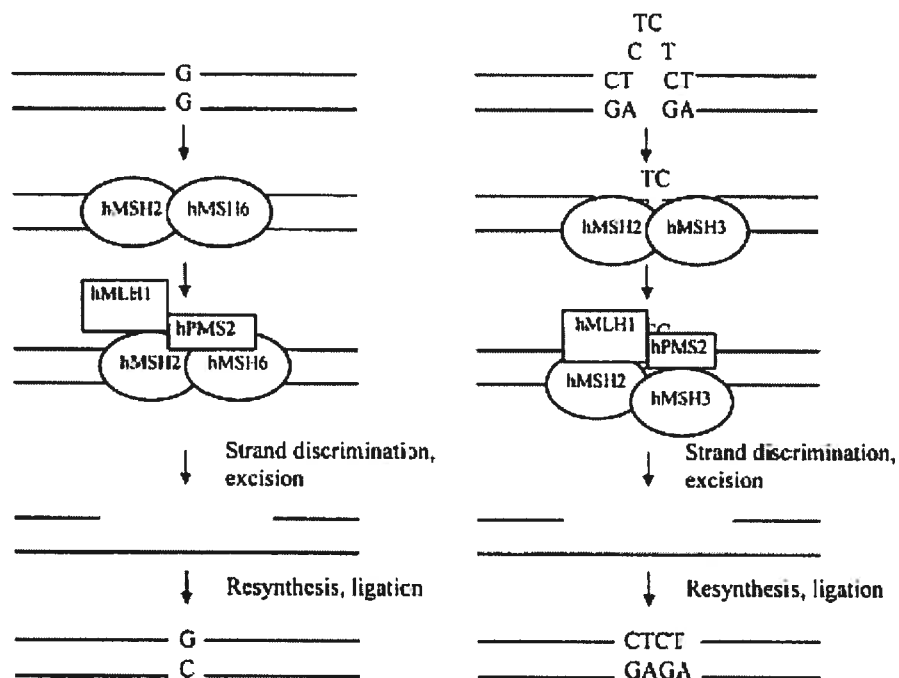


Figure 7. Mécanisme de réparation des mauvais appariements (MMR). La partie de gauche représente la réparation d'un mésappariement d'un seul nucléotide. La partie de droite illustre la réparation d'une petite boucle. (<http://www.ub.rug.nl/eldoc/dis/medicine/y.wu/c2.pdf>).

nucléotides (figure 7). Le complexe ADN-MutS est reconnu à son tour par l'hétérodimère MutL $\alpha$  composé de protéines homologues de MutL de *E. coli*, MLH1/PMS2 (Aquilina, G. et Bignami, M. 2001). L'étape de resynthèse/ligation est similaire à celle de la réparation par excision de nucléotides. Des études ont démontré un rôle important de PCNA, le facteur accessoire des ADN polymérases  $\delta$  et  $\epsilon$ , dans le système de réparation des mésappariements (Umar, A. *et al.* 1996).

Le cancer du côlon est d'origine héréditaire dans environ 10 à 20% des cas. Un défaut au niveau du gène APC explique en partie ces cas héréditaires. Par contre, les cas qui ne sont pas causés par un défaut dans ce gène sont associés au désordre *Hereditary nonpolyposis colorectal cancer* (HNPCC) qui est responsable d'environ 2 à 5% de tous les cas de cancer du côlon (Boland, CR. 1998; Vasen, HFA. *et al.* 1991). Une mutation dans un des gènes impliqués dans la MMR peut être montrée dans la majorité des cas. Parmi les gènes de la MMR, des mutations dans le gène MLH1 et MSH2 comptent pour plus de 90% des mutations détectées dans les cas de HNPCC (Peltomäki, P. et Vasen, HFA. 1997). Des mutations dans le gène MSH6 ont été trouvées chez quelques familles, et de rares mutations dans PMS1 et PMS2 ont été décrites. Les cellules affectées de ces individus portent un allèle non fonctionnel d'un gène de la voie MMR; si le second allèle est perdu ou inactivé, la cellule ne peut plus réparer les erreurs survenant lors de la réplication de l'ADN. Les cellules accumulent des centaines d'erreurs/mutations à chaque division cellulaire. Les cellules tumorales ont donc un phénotype mutateur. Puisque plusieurs mutations surviennent dans les répétitions de mononucléotides, dinucléotides ou trinuécléotides (i.e les microsatellites) retrouvés un peu partout dans le génome, le phénotype mutateur est également appelé le phénotype d'instabilité des microsatellites (MI). Environ 10 à 15% de l'ensemble des cas de cancers du côlon montre un tel phénotype mutateur (Aaltonen, LA. 1993 ; Ionov, Y. *et al.* 1993 ; Esteller, M. *et al.* 1998). Récemment, l'inactivation de MLH1 dans ces cas a été démontrée et est due à des changements épigéniques comme la méthylation de séquences régulatrices de la transcription (Leung, LY. *et al.* 1999; Cunningham, JM. *et al.* 1998; Herman, JG. *et al.* 1998).



Presque tous les cancers du côlon avec instabilité des microsatellites ont des mutations au niveau des séquences répétitives présentes dans les régions codantes des gènes TGF  $\beta$  IIR et BAX. Les mutations du gène TGF  $\beta$  IIR permettent d'échapper à l'inhibition de la croissance induite par TGF  $\beta$  tandis que les mutations dans BAX interfèrent avec son effet anti-apoptotique (<http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowSection&rid=cmed.section.d1e120808>). Donc, la carcinogénèse dans les cancers du côlon avec instabilité des microsatellites peut impliquer des mutations dans des gènes critiques différents de ceux impliqués dans les autres cas de cancer du côlon (APC, K-ras, p53, etc).

### 3.5 Réparation des cassures doubles-brins (DSBs repair)

Les lésions doubles-brins de l'ADN (DSBs) peuvent être générées par une variété d'agents génotoxiques, incluant les radiations ionisantes. Elles peuvent également survenir lors de recombinaison somatique comme la recombinaison V(D)J qui est le processus par lequel les domaines variables des immunoglobulines et des récepteurs des cellules T sont assemblés à partir des éléments V, D et J dans les cellules pré-B dans la moelle osseuse et dans les cellules pré-T dans le thymus (Fugmann, SD. *et al.* 2000). Il est crucial que les cellules reconnaissent ces bris et les réparent puisqu'un seul bris double-brin peut mener à la mort cellulaire, et un défaut de réparation peut mener à l'apparition de mutations ou la perte de matériel chromosomique. Ces lésions sont de potentiels déclencheurs de carcinogénèse via l'activation de proto-oncogènes ou l'inactivation de gènes suppresseurs de tumeur.

Il existe deux voies de réparation des cassures doubles-brins : la réparation par recombinaison homologue (HRR) et la réparation non-homologue (NHEJ). Dans les cellules de mammifères, la situation est différente de celle observée chez la levure. Jusqu'à récemment, la voie NHEJ était perçue comme la principale voie de réparation des lésions doubles-brins. De récentes études ont démontré que la contribution de la recombinaison homologue avait été sous-estimée chez les mammifères. Il semble que les deux mécanismes de réparation sont d'importance égale mais qu'ils seraient

utilisés à différents moments du cycle cellulaire. La voie NHEJ serait surtout employée pendant la phase G1 jusqu'en phase S tandis que la voie par HRR serait utilisée pendant la phase S tardive jusqu'en G2 (West, SC. *et al.* 2000).

La réparation des lésions doubles-brins est déclenchée suite à la détection du bris. Depuis plusieurs années, il a été démontré qu'une cascade de phosphorylation de protéines mène à l'arrêt de la transcription, la réplication et la réparation. Il a été suggéré que la protéine mutée dans l'ataxie telangiectasie (ATM) puisse être une des premières kinases activées dans la réponse cellulaire aux DSBs (Burna, S. *et al.* 2001). Dans une cellule non-perturbée, il a été démontré que ATM est séquestrée sous forme de dimères. Après une lésion de l'ADN, le domaine kinase d'une molécule d'ATM phosphoryle l'autre molécule ATM avec laquelle elle interagit. ATM est ensuite libérée du complexe et peut phosphoryler ses cibles dont l'histone H2AX, p53, BRCA1 et peut-être BRCA2 (Bakkenist, CJ. et Kastan, MB. 2003). Le mécanisme exact par lequel ATM est activée est pour le moment inconnu. Chez les mammifères, la phosphorylation de H2AX joue un rôle crucial dans le recrutement de la machinerie de réparation ou des facteurs qui signalent la présence d'une lésion. Cette phosphorylation permet la décondensation de la chromatine autour de la lésion et ainsi facilite l'accès aux facteurs de réparation. L'activation d'ATM serait causée par la lésion double-brin qui provoque un changement dans la chromatine (Barket, J. et Lukas, J. 2003).

### **3.5.1 Réparation par recombinaison homologue (HRR)**

Lors de la recombinaison homologue, la machinerie de réparation utilise une molécule d'ADN homologue, c'est-à-dire la chromatide sœur ou le chromosome homologue, comme gabarit pour une réparation efficace des bris d'ADN double-brin se produisant lors de la méiose ou de la mitose (De Jager, M. 2001). Des analyses génétiques chez la levure *S. cerevisiae* ont permis d'identifier un groupe de gènes dont les produits jouent un rôle important dans la HRR: RAD50, RAD51, RAD52, RAD54, RAD55, RAD57, RAD59, MRE11 et XRS2. On connaît les homologues de ces gènes chez les mammifères (Jackson, SP. 2001) à l'exception des gènes RAD55

et RAD57. Il existe au moins 5 gènes qui ont été classés comme membres de la famille RAD51 : XRCC2 et XRCC3 (Cartwright, R. *et al.* 1998; Liu, N. *et al.* 1998) ainsi que RAD51B, RAD51C et RAD51D (Albala, JS. *et al.* 1997; Rice, MC. *et al.* 1997; Cartwright, R. *et al.* 1998; Dosanjh, MK. *et al.* 1998; Pittman, DL. *et al.* 1998). L'alignement des séquences des membres de la famille RAD51 suggèrent que RAD51D et XRCC3 ressemblent plus à RAD57 tandis que XRCC2 est plus similaire à RAD55 (Thacker, J. 1999; Tsutsui, Y *et al.* 2000). Il a été démontré que BRCA1 et BRCA2 sont impliquées dans la recombinaison et la réparation homologue (Zhang, N. *et al.* 1998; Venkitaraman, AR. 2000). Rad51 interagit avec BRCA1 et BRCA2 (Scully, R. *et al.* 1997). Par contre, la fonction de la liaison de BRCA1 et Rad51 n'est pas claire car ce dimère est retrouvé en quantité minime dans les cellules représentant environ 2 à 5% de la quantité totale de ces protéines. Pour ce qui est de la liaison de BRCA2 et de RAD51, il a été suggéré que BRCA2 séquestre RAD51 prévenant ainsi l'activation de la réparation pendant le métabolisme normal de la cellule. Le relargage de RAD51 surviendrait suite à la phosphorylation de BRCA2 possiblement par ATM après la détection d'une lésion double-brin (Venkitaraman, A.R. 2002).

La première étape de la HRR (figure 8) est le clivage de l'ADN de chaque côté de la lésion dans la direction 5'-3' par une nucléase dont l'activité est associée au complexe MRE11, RAD50 et XRS2 (NBS1 est l'homologue chez les mammifères) (Haber, JE. 1998; Paull, TT et Gellert, M. 1998). BRCA1 qui interagit avec ce complexe (Zhong, Q. *et al.* 1999; Wang, Y. *et al.* 2000) pourrait inhiber l'activité nucléase de MRE11, permettant ainsi de réguler la longueur et la persistance de la dégradation de l'ADN au site de cassure. La seconde étape débute par la liaison de RAD52 aux extrémités 3' simple-brin créées précédemment. RAD52 permet la liaison de deux brins d'ADN complémentaires. De plus, RAD52 faciliterait la liaison de RAD51 à l'ADN avec l'aide de RAD54 dont la fonction serait d'ouvrir la chromatine. RAD51 forme un filament nucléoprotéique qui permet la recherche de la séquence homologue non-endommagée. Ensuite, RAD51 catalyse d'une façon ATP-dépendante l'échange des brins d'ADN c'est-à-dire que la molécule endommagée s'insère au niveau du duplexe d'ADN intact. Tous les membres de la famille RAD51

participeraient à cette étape puisque des études ont démontré l'interaction directe entre RAD51, XRCC3 et RAD51C ainsi qu'une interaction entre RAD51B, RAD51C, RAD51D et XRCC2 (Masson, JY. *et al.* 2001; Liu, N. *et al.* 2002). Il est donc possible que ces protéines interagissent ensemble, formant un plus gros complexe protéique ou « recombinosome » qui s'assemblerait au niveau des lésions double-brins d'ADN pour permettre leur réparation. Par contre, une étude a montré que les deux complexes protéiques, RAD51-XRCC3-RAD51C et RAD51B-RAD51C-RAD51D-XRCC2 ne s'associeraient pas ensemble (Liu, N. *et al.* 2002). La troisième étape consiste à la synthèse de l'ADN par l'ADN polymérase qui copie l'information à partir du brin non-endommagé. Cette étape fait également intervenir PCNA qui facilite la liaison de l'ADN polymérase  $\delta$  ou  $\epsilon$ . Finalement, il y a migration et résolution de la jonction de Holliday et ligation de l'ADN. Les enzymes participant à cette étape sont toujours inconnues.

Plusieurs désordres sont associés à des défauts dans les gènes impliqués dans la HRR (Tableau IV). Comme on peut le constater, tous ces syndromes causent une augmentation du risque de développer un cancer. De plus, parmi les gènes impliqués dans la HRR se trouvent les gènes BRCA1 et BRCA2. Des récentes études ont clairement démontré que ces deux gènes étaient impliqués dans le développement du cancer du sein et ovarien (Durocher, F. *et al.* 1996 ; Simard, J. *et al.* 1994). Les cas familiaux de cancers du sein compte pour environ 5 à 10% de tous les cas et bon nombre d'entre eux peuvent être liés à la présence de mutations germinales dans ces gènes (Ford, D. *et al.* 1998; Couch, FJ. *et al.* 1997). Les cas familiaux de cancers ovariens quant à eux comptent pour 8% de tous les cas et de récentes études suggèrent que la plupart sinon la totalité de ceux-ci peuvent être corrélés avec la présence de mutations dans les gènes BRCA1 et BRCA2 (Gayther, SA. *et al.* 1999). La majorité des mutations identifiées dans ces gènes mènent à la formation d'une protéine tronquée. Le cancer se développerait quand la seconde copie du gène est perdue. Par conséquent, BRCA1 et BRCA2 agissent comme des gènes suppresseurs de tumeur classiques dont la perte d'un seul allèle prédispose au développement d'un cancer.

Tableau IV. Quelques syndromes associés à un défaut de HRR

<b>Syndrome</b>	<b>Gène défectueux</b>	<b>Phénotype</b>	<b>Références</b>
Ataxie-telangiectasie (A-T)	ATM	neurodégénération immunodéficience dilatation des petits vaisseaux sanguins perte de fonction cérébelleuse radiosensibilité prédisposition aux cancers (60 à 180 fois dont 70% sont des lymphomes et des leucémies de type T)	LH Thompson 2002
Syndrome Nijmegen breakage (NBS)	NBS1	immunodéficience radiosensibilité prédisposition aux cancers	LH Thompson 2002
Ataxie-telangiectasie-like (ATLD)	MRE11	radiosensibilité prédisposition aux cancers (60 à 180 fois dont 70% sont des lymphomes et des leucémies de type T)	LH Thompson 2002
Syndrome de Bloom (BS)	BLM (interagit avec Rad51)	neurodégénération dilatation des petits vaisseaux sanguins perte de fonction cérébelleuse immunodéficience retard de croissance sensibilité au soleil augmentation de la fréquence du diabète augmentation du risque de cancers	J German 1993

### 3.5.2 Réparation non-homologue (NHEJ)

Chez les eucaryotes, une composante importante du système de réparation NHEJ est l'ADN ligase IV (figure 8). Elle permet la ligation directe de l'ADN et fonctionne en association avec la protéine adaptatrice XRCC4 (Jackson, SP.2001). Par contre, ce complexe ne permet pas à lui seul la reconnaissance du bris. L'hétérodimère Ku formé de Ku80 et Ku70 se lie de chaque côté de la lésion aux extrémités de l'ADN. Cet hétérodimère recrute ensuite PRKDC (DNA-PKcs), une protéine kinase de la famille des phosphoinositol kinases (Calson, P. 2003). Une fois liée à l'ADN, l'activité kinase de la protéine est activée. À son tour Artemis (DCLER1C) est recrutée et phosphorylée par DNA-PKcs ce qui mène à une modification de son activité (Ma, Y. *et al.* 2002). Ce complexe possède une activité endonucléase capable de cliver les structures en épingles à cheveux formées lors de la recombinaison V(D)J et de cliver les extensions 3' tandis qu'Artemis, seule, possède une activité exonucléase 5'-3' sur l'ADN simple-brin (Ma, Y. *et al.* 2002). Ce complexe permettrait donc de modifier les extrémités de la lésion nécessaire à sa réparation par XRCC4 et Lig IV. Jusqu'à la découverte d'Artemis, on pensait que le complexe Mre11-Rad50-Xrs2 effectuait cette étape comme c'est le cas chez *S. cerevisiae* (Boulton, SJ. et Jackson, SP. 1998).

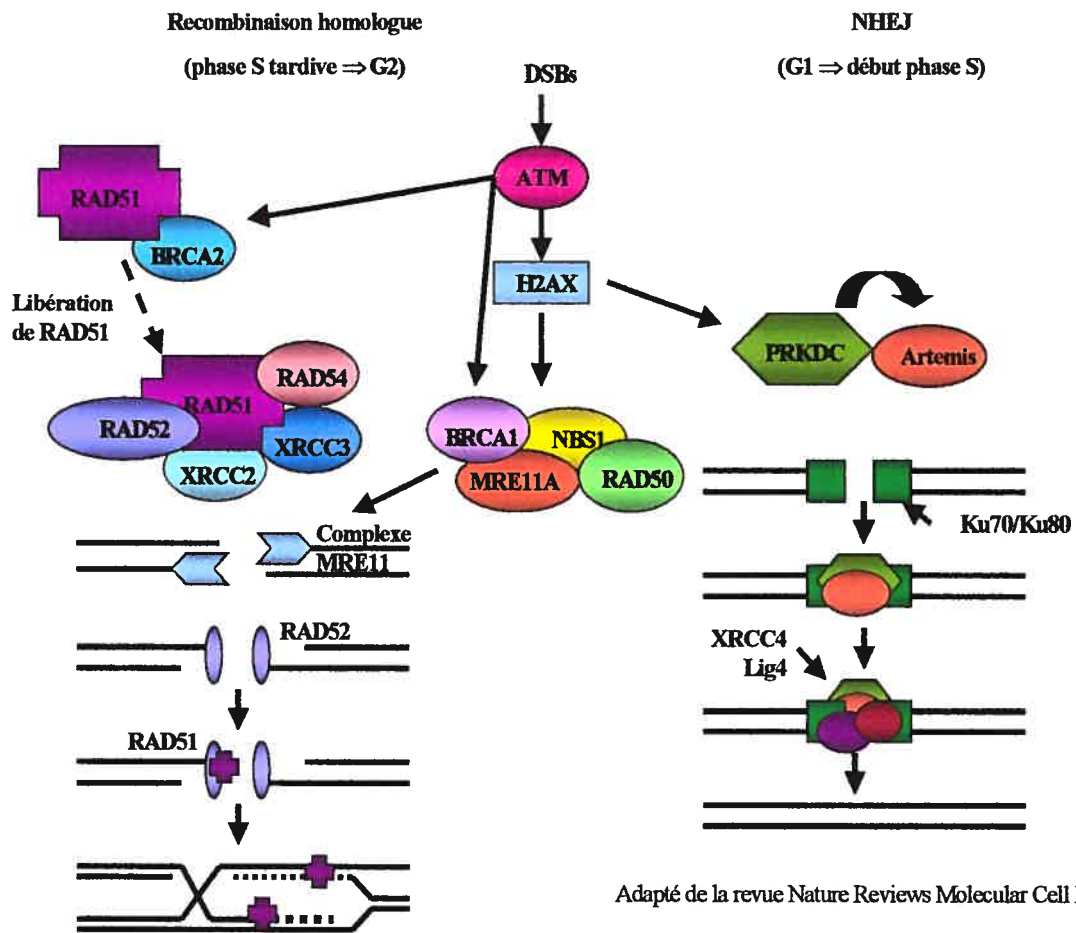


Figure 8. Représentation des voies de réparation des cassures doubles-brins.

## 4 Nature d'un polymorphisme

### 4.1 Type

Un polymorphisme est une différence au niveau de la séquence d'ADN observée entre les individus. Ce terme peut être appliqué à plusieurs situations allant du trait génétique ou de maladies dans une population jusqu'à une variation dans la séquence d'ADN ou de protéines. Les variations génétiques présentes à plus de 1% dans la population sont considérées polymorphiques.

Il existe différents types de polymorphismes dont les polymorphismes de longueur et les polymorphismes simples. Dans tout le génome, il y a une variété de différentes séquences d'ADN dispersées, chacune d'elles étant répétées plusieurs fois à la suite l'une de l'autre. Le nombre de répétitions peut varier d'une douzaine à plus d'une centaine chez différents individus. Ces séquences sont connues sous le nom de « nombre variable de répétition en tandem » (VNTRs). Un autre type de polymorphisme est les microsatellites. Il existe plus de 50 000 copies de séquences répétées dC-dA (répétées de 10 à 60 fois en tandem) dispersées à travers le génome. Puisque les séquences plus longues répétées en tandem (VNTR) ont été également nommées minisatellites, les séquences plus courtes répétées dC-dA ont été nommées microsatellites. Le nombre de répétition d'un locus particulier varie de façon polymorphique entre les individus et puisque ces séquences sont héritées de façon stable, elles peuvent servir de marqueurs polymorphiques. La différence dans le nombre de répétition entre deux microsatellites peut être de seulement quelques nucléotides. Habituellement, le nombre de répétitions d'un marqueur microsatellite est stable, mais dans quelques cancers, comme le cancer colorectal, le nombre de répétitions de microsatellite dans les tumeurs diffèrent du nombre retrouvé dans le tissu colorectal normal de ces mêmes patients.



Le type de polymorphisme le plus simple et le plus commun est le changement simple d'un nucléotide (SNP). Il se produit à une position au niveau de la séquence d'ADN où il peut y avoir plus d'une base possible. Il est retrouvé à un taux de 1 à chaque 300 bp ce qui représente environ  $10^6$  SNPs communs chez l'humain. Différents types de SNP existent : les insertions et les délétions d'une paire de nucléotides et les substitutions soit par transition entre deux purines ou deux pyrimidines ou par transversion entre une purine et une pyrimidine.

#### **4.2 Effet d'un polymorphisme**

L'effet d'un polymorphisme simple dépend de sa localisation. Un SNP qui se trouve au niveau d'une séquence génomique non-codante à l'extérieur d'un gène (gSNP) n'aura aucun effet. Toutefois, si un SNP se situe au niveau de la séquence codante d'un gène (cSNP), il peut avoir un impact fonctionnel. Comme les séquences codantes sont traduites en protéine, une substitution de base peut mener à un changement d'acide aminé. Elle peut altérer le repliement de la protéine, affecter un site important pour des modifications post-traductionnelles comme un site de clivage protéolytique, un site de phosphorylation ou de glycosylation ou modifier la traduction d'une autre façon. Par contre, comme le code génétique est dégénéré, une substitution de base se traduit souvent par le même acide aminé et ne cause pas d'effet sur la protéine. Dans le cas d'insertion ou de délétion au niveau de la séquence codante, il se produit un changement de cadre de lecture car elles causent un décalage d'un nucléotide et changera la lecture de tous les acides aminés suivants. Les cSNPs peuvent également causer l'apparition d'un codon stop prématuré menant à la production d'une protéine incomplète. Les cSNPs peuvent donc produire un changement de la forme ou de la taille de la protéine codée. Un tel changement peut altérer la fonction biologique de la protéine soit en éliminant son activité, la réduisant ou l'augmentant.

Les SNPs qui se trouvent à l'intérieur d'un intron sont nommés iSNPs. Si le iSNP se situe près d'une jonction intron-exon, la production du transcrit primaire peut être

affectée et un intron peut être conservé ou un exon peut être sauté. Un SNP dans la région non traduite en 3' peut avoir un effet sur la stabilité de l'ARNm.

Un polymorphisme au niveau de la séquence régulatrice (rSNP) peut affecter le niveau de transcription du gène. Il peut affecter la liaison de l'ARN polymérase, faire apparaître des nouveaux sites de liaison de facteurs de transcription ou en faire disparaître modifiant ainsi le niveau de transcription. Cette classe de polymorphisme est à la base de mon projet car peu d'attention jusqu'à maintenant a été attribuée aux rSNPs puisque leur effet fonctionnel est plus difficile à évaluer étant donné qu'ils ne provoquent pas un changement direct de la séquence de la protéine codée.

## 4.2 Détection de polymorphismes

Pour détecter de nouveaux SNPs, il existe différentes méthodes. Il y a le séquençage, le SSCP (single-stranded conformational polymorphism), les analyses d'hétéroduplexes entre autres par dHPLC et la détection enzymatique des mauvais appariements. Dans cette section, une description des méthodes les plus utilisées en plus des avantages et inconvénients est donnée.

### 4.2.1 Séquençage

Il y a deux méthodes de séquençage, la méthode de Maxam et Gilbert (Maxam, AM. et Gilbert, M. 1977) et la méthode de Sanger (Sanger, F. *et al.* 1977). La méthode la plus utilisée est la seconde. Le séquençage est basé sur l'observation que lorsque l'ADN polymérase ajoute un nucléotide synthétique modifié à un brin en processus de synthèse, la polymérisation arrête. Les nucléotides synthétiques utilisés pour arrêter la réaction sont des didésoxyribonucléotides (ddNTP) qui ne possèdent pas de OH sur le carbone en 3' de leur groupe désoxyribose. Par exemple, en présence de ddATP, la terminaison de la chaîne se produira lorsqu'un A apparaît sur le brin synthétisé (un T sur la séquence de référence). Si la bonne proportion de dATP

normaux par rapport aux ddATP est utilisée, seulement quelques molécules se termineront à chaque T de la séquence de référence. Ceci générera une série de nouveaux brins d'ADN, chacun partant d'une extrémité de la chaîne jusqu'à la position d'un T. Si on utilise des ddNTPs radioactifs, les produits sont séparés par électrophorèse sur gel de polyacrylamide. Quatre réactions différentes sont préparées, une pour chaque ddNTP différent. On peut également utiliser des ddNTPs marqués avec 4 fluorescents de couleurs différentes. La migration se fait par capillaire, chacun des quatre ddNTPs fluorescent d'une couleur différente lorsqu'ils sont illuminés par un laser et un scanner automatique donne la lecture de la séquence (figure 9).

Le séquençage manuel est une technique très efficace pour la détection de polymorphismes. Par contre, le débit n'est pas très rapide. Pour ce qui est du séquençage automatique, le débit est très élevé mais les coûts le sont également même si ceux-ci ont beaucoup diminué depuis quelques années (Ellis, LA. 2000). De plus, comme chaque ddNTP fluoresce à une intensité différente la détection des hétérozygotes est parfois difficile puisqu'on peut confondre avec le bruit de fond.

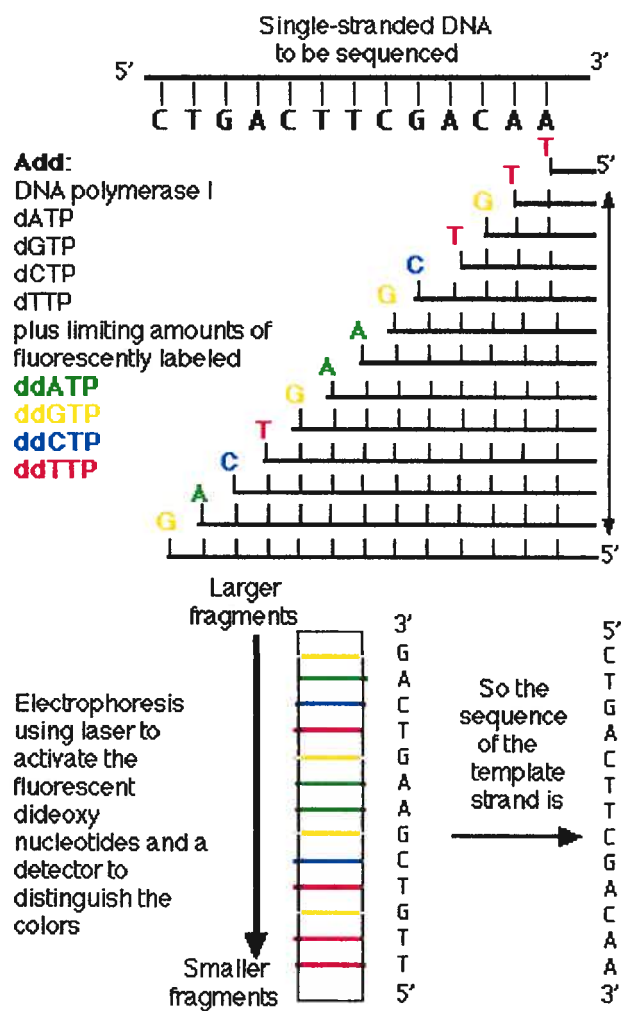


Figure 9. Principe de séquençage.

(<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/D/DNAsequencing.html>)

### 4.2.2 SSCP

Une technique simple de détection de polymorphismes est « single-stranded conformation polymorphism » (SSCP). Cette technique se base sur le principe qu'il y a une différence de mobilité sur un gel non-dénaturant entre deux molécules simple-brin d'ADN à cause de la structure secondaire qu'ils forment. Les molécules d'ADN simple-brin peuvent se refermer sur elles-mêmes grâce à des interactions intra-brin entre les bases formant une structure secondaire. L'altération d'une simple base dans une molécule d'ADN courte modifiera l'interaction intra-brin entre les bases. Un fragment d'ADN de même longueur, mais ayant une structure secondaire différente migrera différemment lors d'électrophorèse sur gel non-dénaturant. Il suffit donc d'amplifier par PCR en utilisant des dNTP radioactifs la région d'intérêt à tester, de dénaturer les produits de PCR suivi d'une renaturation pour permettre la formation de structures secondaires de chaque brin d'ADN. Ensuite, les produits sont séparés par électrophorèse sur gel non-dénaturant. Les fragments contenant un SNP peuvent être identifiés par leur profil de migration différent (figure 10). Une nouvelle alternative au SSCP traditionnel est le F-SSCP. Les produits de PCR sont marqués à chaque extrémité avec de la fluorescence. L'électrophorèse s'effectue par capillaire et la lecture et l'analyse se fait automatiquement par un scanner à fluorescence.

Le SSCP classique comporte quelques limitations. D'un point de vue théorique, il est pratiquement impossible de prédire la structure secondaire. En pratique, le SSCP est limité par la taille du fragment et nécessite l'essai de plusieurs conditions pour obtenir le maximum de sensibilité. Il est connu que l'analyse d'un fragment plus long que 300 pb diminue grandement la sensibilité de la technique (Hayashi, K. 1991). La modification du SSCP en F-SSCP a permis l'amélioration de la technique. Elle est maintenant beaucoup plus adaptée à une détection à haut débit, plus efficace pour détecter les polymorphismes et plus facile d'utilisation. L'efficacité de détection se situe autour de 95% (Gross, E. 1999; Ellis, LA. *et al.* 2000; Dobson-Stone, C. *et al.* 2000).

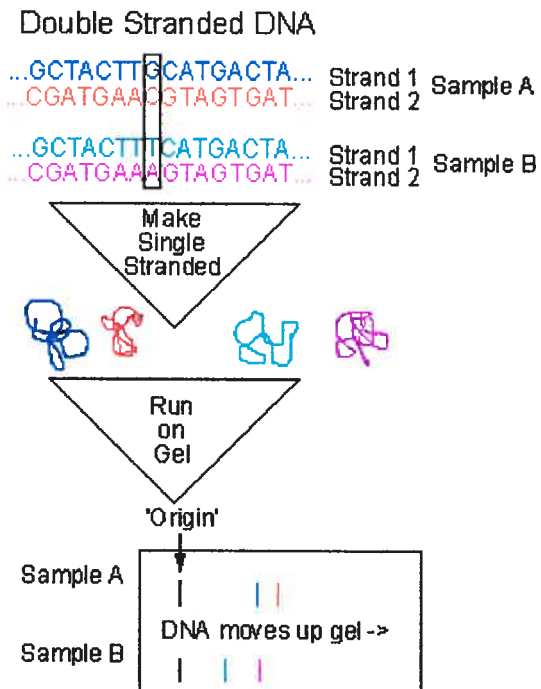


Figure 10. Représentation schématique de la technique SSCP.  
 (www.amonline.net.au/evolutionary\_biology/ images/sscp.gif)

### 4.3 dHPLC

La technique du HPLC dénaturant (denaturing high performance liquid chromatography, dHPLC) a été développée récemment comme méthode de détection de polymorphismes pouvant compétitionner quant à la simplicité et la sensibilité avec le SSCP. Le dHPLC est une technique qui permet, entre autres, la détection rapide de la présence de polymorphisme en identifiant les échantillons hétérozygotes. Les produits de PCR, préalablement dénaturés, sont refroidis lentement pour permettre aux échantillons hétérozygotes de former des hétéroduplexes et des homoduplexes tel qu'illustré à la figure 11. Par la suite, la séparation de ces différents duplexes permet la détection de polymorphismes. Comme les hétéroduplexes commencent à se dénaturer plus rapidement que les homoduplexes à cause du mauvais appariement au niveau du polymorphisme, ils seront élués en premier de la colonne (figure 12). Pour détecter les individus hétérozygotes, le fragment de PCR doit être partiellement dénaturé (50 à 99,9%) autour de la région contenant le polymorphisme. Comme la position des polymorphismes est inconnue on sélectionne les températures nécessaires pour pouvoir analyser en totalité le fragment de PCR (figure 13).

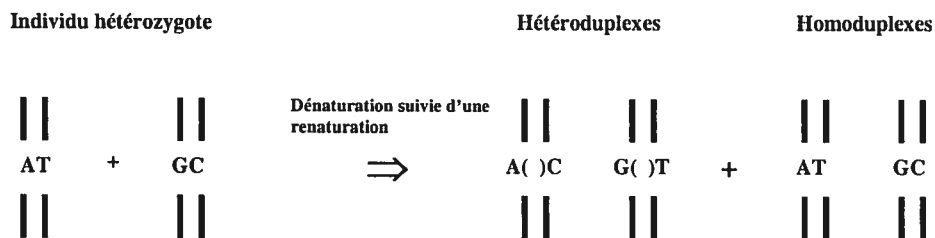


Figure 11. Représentation graphique de la formation des hétéroduplexes.

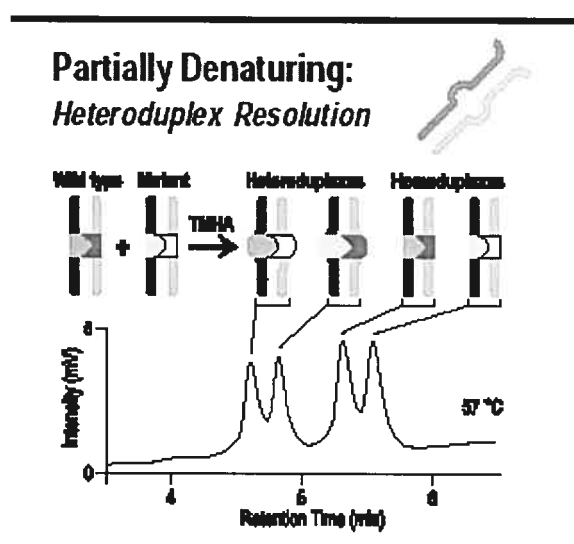


Figure 12. Résolution des hétéroduplexes.

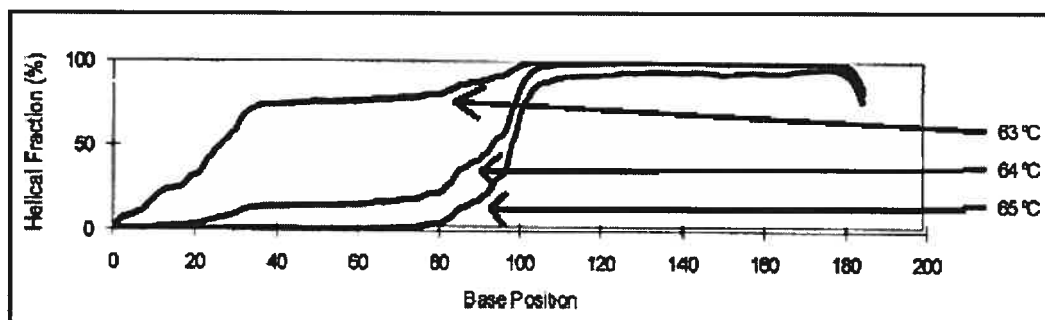


Figure 13. Profil de dénaturation à 3 températures d'un fragment de 185 pb tel qu'évalué par WAVEMAKER. La recherche de polymorphismes inconnus sur la totalité du fragment nécessitera l'analyse à 2 températures différentes : les SNPs du côté 5' seraient détectés à une plus basse température que celle du côté 3'.



Quand on observe que le pic se déplace vers un temps de rétention plus court (vers la gauche, figure 14), cela signifie que cette température permet le début de la dénaturation du fragment. À ce point, la présence d'un polymorphisme pourra être détectée par l'apparition d'un ou deux pics supplémentaires qui seront élués juste avant les homoduplexes.

Cette technique est facile d'utilisation, possède un débit élevé et a un pouvoir de détection des polymorphismes très élevé. Des études ont rapporté un taux d'efficacité allant de 95 à 100% (Gross, E. 2000; Premstaller, A. 2002; Dobston-Stone, C. *et al.* 2000). Contrairement au SSCP, le dHPLC peut détecter la présence de polymorphismes dans des fragments allant jusqu'à 1,5 kb. Par contre, un désavantage du dHPLC est la difficulté de distinguer les homoduplexes entre eux. En d'autres mots, le génotypage est effectué en parallèle avec cette approche.

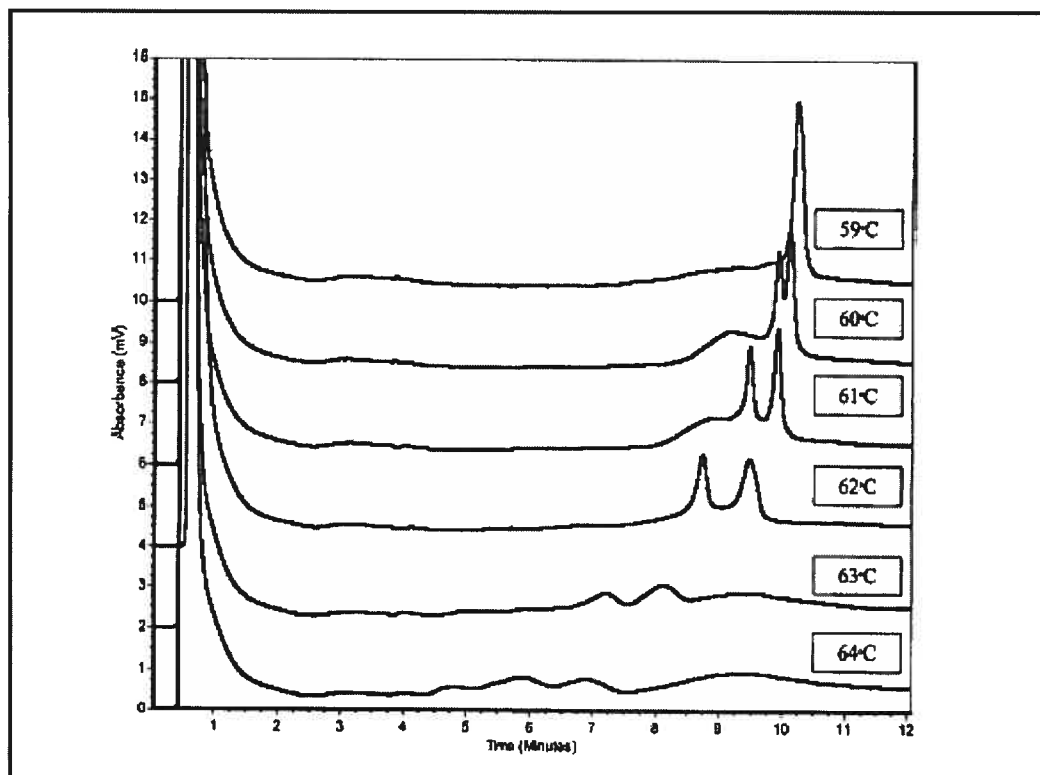


Figure 14. Représentation d'une courbe de titration de températures. Dans cet exemple, la température idéale se situe entre 61 et 62°C.

## **5 Inter-relation entre le métabolisme des xénobiotiques et la réparation de l'ADN**

Il existe une relation importante entre le métabolisme des xénobiotiques, la réparation de l'ADN et la variation génétique. En effet, si on reprend le modèle décrit précédemment, on peut comprendre que l'effet d'un polymorphisme fonctionnel au niveau de n'importe quelle composante de ces deux voies peut avoir un impact sur la susceptibilité à développer un cancer (figure 15). Par exemple, un polymorphisme causant une augmentation de l'activité d'une enzyme de la phase I sera associé à une plus grande concentration d'intermédiaires réactifs capables d'interagir avec les macromolécules, d'où une augmentation du risque de dommages causés à l'ADN. Par conséquent, si un autre polymorphisme situé dans une enzyme de réparation cause une diminution de son activité, ces dommages seront réparés moins efficacement et par conséquent le risque de développer un cancer est accru. Plusieurs études ont déjà montré une association avec la présence de polymorphismes dans les gènes de réparation et un risque plus élevé de développer différents types de cancer (tableau V). Par exemple, on peut remarquer que la présence de polymorphismes dans certains gènes de la voie BER, comme hOGG1, est souvent associée avec une susceptibilité élevée au cancer du poumon. Il est logique de penser que la présence d'un polymorphisme diminuant l'activité d'une de ces enzymes place les cellules à une plus grande fragilité face au stress oxydant comme la fumée de cigarettes.

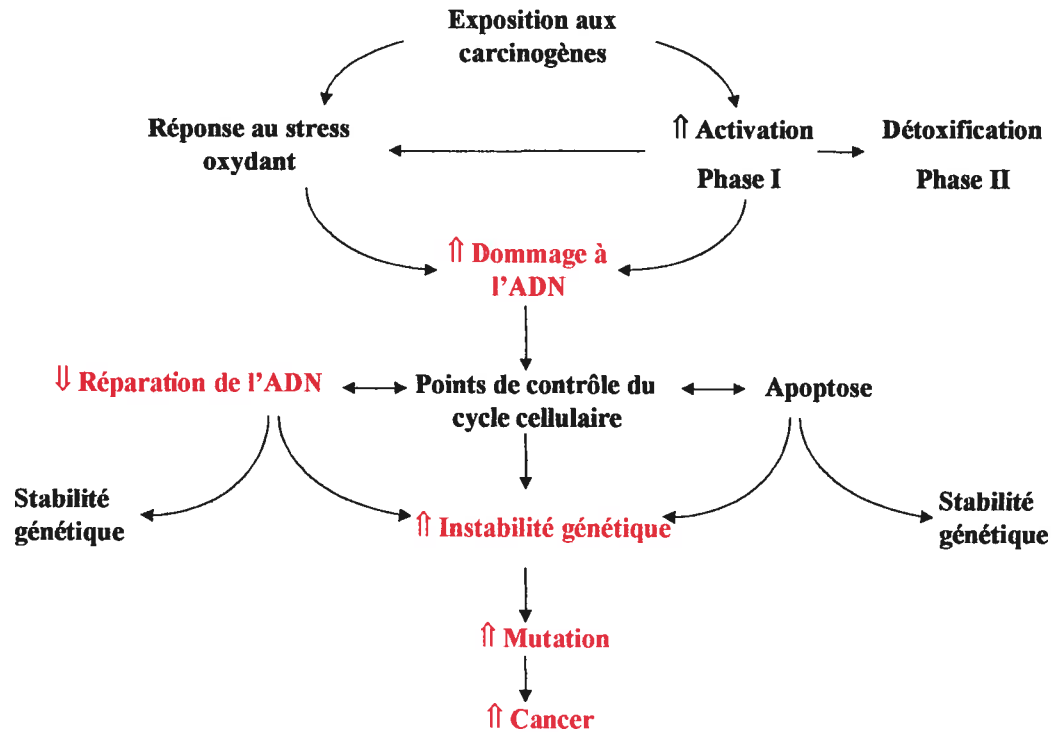


Figure 15. Représentation hypothétique de l'inter-relation du métabolisme des xénobiotiques et de la réparation de l'ADN.

Tableau V. Liste partielle d'étude d'association montrant un effet entre la présence de polymorphismes et le risque de développer un cancer

Gène	Voie de réparation	Position	Changement	Association avec le cancer	Références
XRCC1	BER	194	Arg-Trp	poumon, sein, colorectal	W Ladiges 2002
		399	Arg-Gln	poumon, sein, colorectal, peau	W Ladiges 2002
XRCC3	HRR	241	Thr-Met	mélanome, vessie	SL Winsey <i>et al.</i> 2000
hOGG1	BER	326	Ser-Cys	poumon	L Le Marchand 2002
				cancer orolaryngial prostate	A Elahi 2002 J Xu 2002
PRKDC	NHEJ	2140	Arg-Cys	sein	Y Yu <i>et al.</i> 2001

## **6 Leucémie de l'enfant**

### **6.1 Incidence**

En Amérique du Nord, un individu sur trois sera atteint d'un cancer au cours de sa vie. Cependant, les cancers pédiatriques ne représentent que 1% de l'ensemble des cas (Statistiques canadiennes du cancer, 2002). Par contre, ils constituent la première cause de décès chez les enfants de 0 à 19 ans malgré une amélioration de plus de 50% du taux de survie depuis 50 ans. Les cancers les plus fréquents sont les cancers du système hématopoïétique représentant 26% des nouveaux cas et 32% des décès. Il existe plusieurs types de leucémie selon le type de cellules affectées, soit la leucémie myéloïde aiguë ou chronique et la leucémie lymphoblastique aiguë ou chronique. Dans le cadre de mon projet, la leucémie lymphoblastique aiguë (LLA) a retenu notre attention. La LLA constitue plus de 75% des cas de leucémies. Le taux de survie sur 5 ans pour les enfants atteints de LLA était seulement de 5% au début des années 60, pour atteindre 80% aujourd'hui.

### **6.2 Caractéristiques de la LLA**

Le développement de la LLA est causé par la prolifération maligne des lymphoblastes, les cellules progénitrices des lymphocytes T et B, qui se trouvent dans la moelle osseuse, le sang périphérique et d'autres organes. La moelle osseuse produit tellement de ces cellules qu'elle ne peut plus produire suffisamment de globules rouges, de leucocytes et de plaquettes résultant en une anémie et une susceptibilité aux infections sévères. La LLA peut se développer à partir de n'importe quelle cellule lymphoïde, incluant les cellules primitives. Près de 85% des cas de LLA sont le résultat de l'expansion clonale d'une cellule précurseur d'un lymphocyte B (LLA pré-B). L'expansion d'un précurseur des cellules T représente environ 15% des cas (LLA pré-T) tandis que la leucémie à cellules B matures ne constitue qu'environ 1% des cas. La forme pré-B est considérée comme étant la forme la plus facilement curable. Des anomalies génétiques (gains ou pertes chromosomiques, translocations

menant à la formation de gènes de fusion, délétions ou inactivation de gènes suppresseurs de tumeur) sont retrouvées dans les cellules de plus de la moitié des patients atteints de LLA. La découverte de ces anomalies a contribué grandement à la compréhension de la pathogénèse et du pronostic de la maladie, par contre peu est connu sur la susceptibilité génétique et les facteurs environnementaux. Comme d'autres cancers sporadiques, la LLA peut être considérée comme une maladie génétique complexe où une série de gènes moduleraient l'effet de facteurs extérieurs, modifiant ainsi le risque individuel de développer cette maladie (Krajinovic, M. *et al.* 2001, Krajinovic, M. *et al.* 2002a, Krajinovic, M. *et al.* 2002b; Sinnott, D. *et al.* 2000; Infante-Rivard, C. *et al.* 2002).

### 6.3 Réparation de l'ADN et leucémogénèse

L'instabilité génétique est la marque de commerce des cancers. Deux principaux types d'instabilité ont été décrits, soit l'instabilité des microsatellites discutée précédemment retrouvée principalement dans les cancers du côlon et l'instabilité résultant de réarrangements chromosomiques. Des évidences suggèrent que cette instabilité résulterait d'une perturbation du système de réparation des cassures double-brin (Khanna, KK. et Jackson, SP. 2001; Schar, P. 2001). Comme la leucémie est souvent associée avec la présence de translocations et de délétions, elle est un bon modèle pour étudier ce phénomène d'instabilité. Un groupe a récemment démontré *in vitro* qu'une augmentation de l'activité de la voie NHEJ et plus spécifiquement une activité dérégulée de l'hétérodimère Ku70/Ku80 contribuerait à l'instabilité génétique observée dans la leucémie myéloïde (Gaymes, TJ. *et al.* 2002). Plusieurs désordres génétiques mentionnés précédemment comme le syndrome de Bloom ou l'ataxie télangiectasie sont associés avec un risque élevé de développer des malignités hématologiques. De plus, des études ont démontré que ATM était relié avec la pathogénèse de la leucémie lymphocytaire chronique à cellules B. La perte d'un allèle suite à une délétion et la présence de mutations dans le second allèle ont été retrouvés chez ces patients suggérant un rôle pour ATM dans la leucémogénèse (Schaffner, C. *et al.* 1999). Dans notre laboratoire, nous avons également montré une association entre des variants présents dans le gène de réparation MLH1 (Mathonnet, G. *et al.*

sous presse). De plus, le variant A676G du gène MLH1 est associé avec un risque plus élevé de rechutes ou de décès suggérant ainsi un rôle important de cet allèle dans la réponse au traitement de la LLA (Krajinovic, M. *et al.* 2002).

## **7 Régulation de l'expression génique**

### **7.1 Structure générale d'un promoteur**

Dans un gène typique, une région appelée le cœur du promoteur est localisée immédiatement en amont du gène. Le cœur du promoteur lie l'ARN polymérase II (Pol II) et ses facteurs accessoires (« la machinerie générale de transcription ») et permet à la Pol II d'initier la transcription au bon endroit. Immédiatement en amont du cœur du promoteur se situe le promoteur régulateur et plus loin soit en amont ou en aval se trouvent les séquences enhancers (figure 16). Les promoteurs régulateurs ainsi que les enhancers lient des protéines appelées facteurs de transcription qui contrôlent la transcription du gène. L'activation se produit suite au recrutement de la machinerie générale au cœur du promoteur via l'interaction entre les activateurs liés au promoteur. Quelques activateurs sont exprimés de façon ubiquitaire tandis que d'autres ne sont exprimés que dans certains types cellulaires régulant les gènes nécessaires à une fonction cellulaire particulière. Pour qu'un gène soit transcrit activement, la chromatine doit être remodelée. L'ADN doit être décondensé au niveau des régions régulatrices de façon à être accessible pour la liaison de facteurs de transcription (Carey, M. et Smale, ST. 2000) .



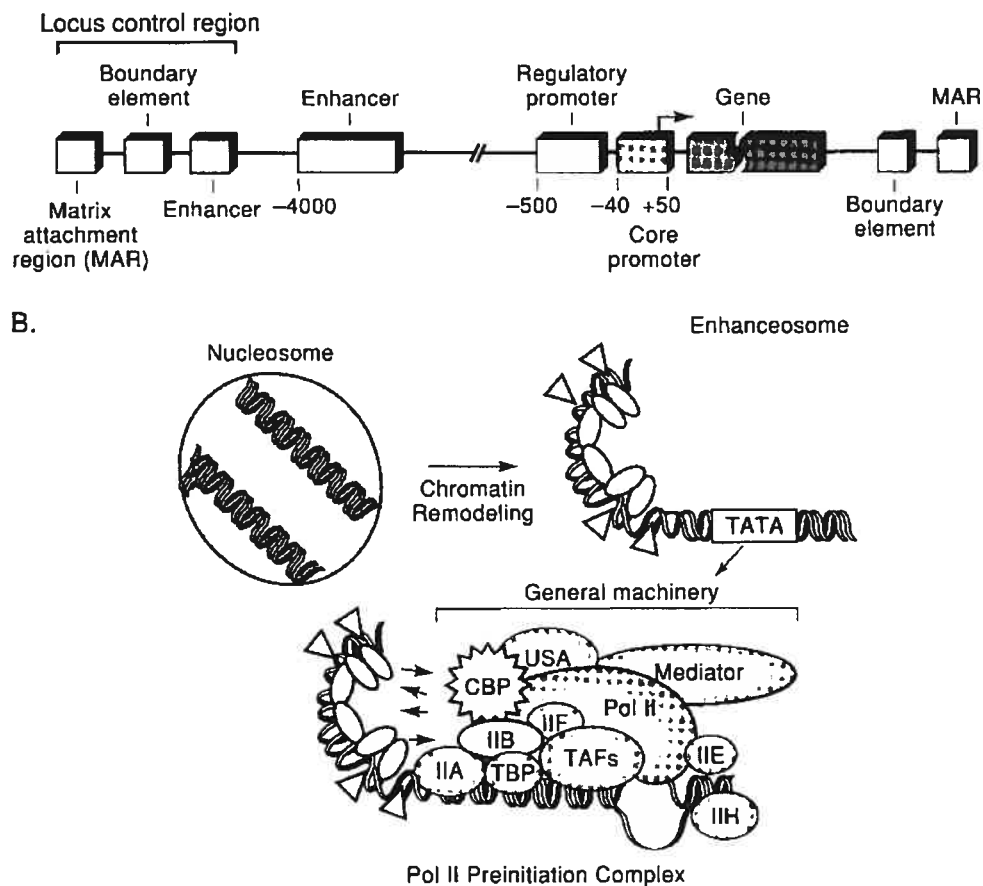


Figure 16. A. Modèle typique d'un gène avec les composantes jouant un rôle dans l'activation et l'inactivation de ce gène B. Activation d'un gène et assemblage du complexe de pré-initiation de la Pol II (Carey, M. et Smale, ST. 2000).

Un cœur de promoteur typique comprend des séquences d'ADN entre -40 et +50 approximativement par rapport au site d'initiation de la transcription (Smale, ST. 1994). Les éléments du cœur du promoteur (1) lient et contrôlent l'assemblage du complexe de pré-initiation contenant la Pol II, les facteurs de la transcription générale et les coactivateurs ; (2) positionnent le site de l'initiation de la transcription et contrôlent la directionnalité de la transcription ; et (3) répondent aux activateurs et répresseurs proximaux ou distaux dans une cellule. Dans la majorité des cas, les éléments du cœur du promoteur ne jouent pas un rôle direct dans la régulation de la transcription.

Un cœur de promoteur typique contient les éléments suivants (figure 17):

1. Le motif TATA. Cet élément de séquence, dont le consensus est TATAAA, est localisé 25-30 pb en amont du site d'initiation de la transcription. La boîte TATA est suffisante pour activer la transcription quand un activateur se lie à un élément régulateur qui se trouve près. La sous-unité TBP de TFIID se lie directement au motif TATA. La liaison de TFIID à la boîte TATA permet le recrutement du reste de la machinerie de transcription générale.
2. L'élément initiateur (Inr). Bien que le Inr possède les mêmes fonctions que la boîte TATA en permettant la formation du complexe de pré-initiation, en déterminant la localisation du site de départ et en médiant l'action d'activateurs, il chevauche directement le site d'initiation de la transcription. L'activité fonctionnelle de Inr dépend de la séquence PyPyA+1NT/ApyPy qui est reconnue par les protéines TAF<sub>II</sub> et Pol II.
3. L'élément en aval du cœur du promoteur (downstream core promoter element ; DPE). Le DPE est une séquence de 7 nucléotides (séquence consensus RGA/TCGTG) qui est située 30 pb en aval du site d'initiation. Il est trouvé dans beaucoup de gènes de mammifères.

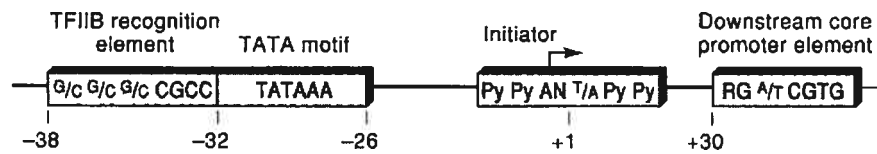


Figure 17. Éléments de séquence d'un cœur de promoteur.

4. L'élément de reconnaissance TFIIB (BRE). Le BRE lie TFIIB et se situe à la position -32 à -38, juste en amont de la boîte TATA et à la séquence G/C, G/C, G/A, CGCC.

La structure spécifique du cœur du promoteur d'un gène particulier joue un rôle de régulation de la transcription non pas en interagissant avec des protéines spécifiques au type cellulaire mais en fournissant un contexte approprié pour une activation efficace.

La régulation de la transcription est contrôlée par la liaison de protéines à des séquences spécifiques d'ADN au niveau du promoteur régulateur et des enhanceurs. Le promoteur régulateur est défini comme étant la région environnant le cœur du promoteur sur une distance de quelques centaines de paires de bases du site d'initiation de la transcription. Un enhanceur est défini comme une région contrôlée trouvée à une plus grande distance du site d'initiation de la transcription, soit en amont ou en aval d'un gène ou dans un intron. Les enhanceurs lient des activateurs et d'autres protéines qui sont impliquées dans le remodelage de la chromatine. Une fois les activateurs liés, l'ADN forme une boucle permettant ainsi à l'ADN d'interagir avec les protéines liées au cœur du promoteur et au promoteur régulateur (figure 18). Ces interactions stabilisent l'assemblage du complexe de transcription.

Les activateurs sont des protéines modulaires formées de domaines distincts pour lier l'ADN et activer la transcription. Le domaine de liaison à l'ADN dirige l'activateur vers des sites spécifiques tandis que le domaine d'activation interagit avec la machinerie générale pour la recruter au niveau du promoteur. Les protéines régulatrices sont regroupées en différentes familles par la séquence et la structure de leur domaine de liaison à l'ADN. Le domaine d'activation de la transcription réfère à une grande variété de domaines protéiques qui interagissent directement avec les composantes de la machinerie transcriptionnelle ou avec des coactivateurs. La définition d'un domaine d'activation est une région d'une protéine qui stimule la

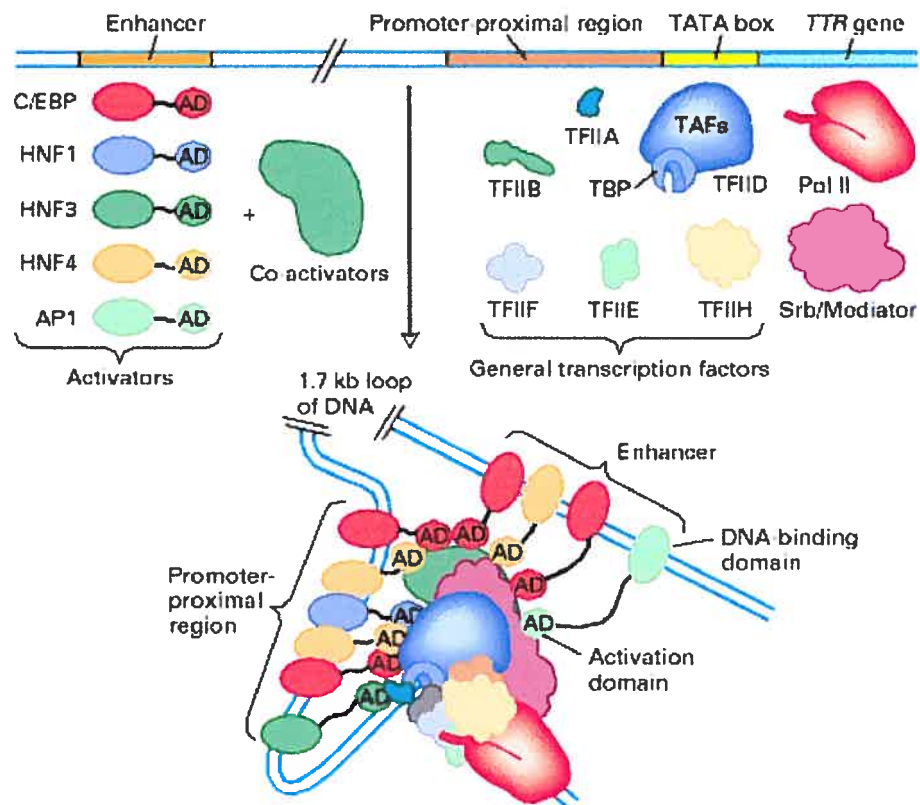


Figure 18. Modèle d'interaction entre différents activateurs. (Lodish, H *et al.* 2000)

transcription quand elle est liée à un domaine de liaison à l'ADN. Plusieurs activateurs possèdent plus d'un domaine d'activation.

## 7.2 La génétique régulatrice

Un polymorphisme au niveau des régions régulatrices ne change pas la séquence protéique, mais peut affecter le niveau de l'expression du gène. Il est donc plus difficile de prédire l'impact du polymorphisme sur la seule base de la séquence d'ADN. Un tel polymorphisme au niveau des régions régulatrices peut mener à l'apparition ou à la disparition de sites de liaison de facteurs de transcription causant ainsi une perturbation au niveau de l'expression du gène. Il existe quelques exemples dans la littérature de polymorphismes dans les régions régulatrices ayant un impact fonctionnel. Le gène glycine N-méthyltransférase (GNMT) code pour une protéine impliquée dans le maintien de la stabilité génétique et est souvent moins exprimée dans les cas de carcinomes hépatocellulaires (HCC) (Tseng, T-L. 2003). Dans cette étude, l'impact fonctionnel de polymorphismes en 5' a été caractérisé à l'aide d'études de gène rapporteur et d'EMSA (electromobility shift assay). Le résultat était que la diminution d'expression de ce gène pourrait être causée par la disparition d'un site de liaison d'un facteur de transcription. Une analyse fonctionnelle d'un variant au niveau du promoteur du gène codant pour la chaîne I du récepteur de l'interféron-gamma impliqué dans la susceptibilité à la malaria a été publiée (Jülicher, S. *et al.* 2003). Dans cette étude, le polymorphisme coïncide avec un site de liaison d'un facteur de transcription abolissant la liaison de ce facteur au niveau du promoteur. Par conséquent, la baisse du niveau d'expression de ce gène observée pourrait être expliquée par ce polymorphisme. Un article de revue regroupant l'étude de 151 variants fonctionnels retrouvés dans 107 gènes différents montre que 133 de ces variants se situent à l'intérieur de 2 kb en amont du site d'initiation de la transcription, 5 dans la région 3' UTR, 7 dans les introns et 6 à une distance plus grande que 2 kb.

## 8. Hypothèse

L'hérédité joue un rôle important dans le développement de maladies complexes comme l'hypertension, l'asthme et le cancer. La variabilité génétique pourrait expliquer la différence de susceptibilité observée entre les individus. Des polymorphismes au niveau des régions codantes peuvent avoir un impact fonctionnel en modifiant la structure ou l'activité de la protéine. Également, puisque le niveau protéique régule plusieurs voies métaboliques, des polymorphismes régulateurs peuvent influencer des processus biologiques et modifier le risque de développer certaines maladies en modifiant le niveau de transcription du gène correspondant. Grâce aux nombreux efforts portés au projet du génome humain, le développement de stratégies à haut débit pour identifier des séquences d'ADN qui sont importantes pour la régulation des gènes chez les mammifères devient de plus en plus important. Par contre, jusqu'à tout récemment la caractérisation du promoteur d'un seul gène demandait beaucoup de temps et d'efforts. Par conséquent, des données sur la structure détaillée des promoteurs de chacun des gènes du génome humain ne sont pas disponibles. Pour mener une étude sur la variabilité observée au niveau des régions régulatrices de certains gènes, il faut par conséquent délimiter une région suffisamment grande pour éviter de manquer des éléments de régulation. Nous avons défini arbitrairement la région régulatrice comme un domaine de 2kb situé en amont du premier exon. Nous proposons que la variabilité d'expression ou d'activité causée par la présence des rSNPs dans les gènes codant pour des composantes de réparation de l'ADN est associée à une susceptibilité inter-individuelle face aux maladies complexes comme le cancer.

## 9 Objectifs du projet

Les objectifs spécifiques de mon projet étaient :

1. Évaluer à l'aide d'études d'association l'importance de cSNPs présents dans les gènes de réparation XRCC1, XRCC3 et hOGG1 potentiellement impliqués dans la susceptibilité à la leucémie lymphoblastique aiguë
2. Identifier des polymorphismes dans les régions régulatrices (rSNPs) en 5' de 15 gènes impliqués dans la réparation des cassures doubles-brins
3. Évaluer l'impact fonctionnel des rSNPs, à l'aide d'analyses *in silico* : prédiction de gain/perte de sites de liaison pour des facteurs de transcription et *in vitro* : vérification des prédictions par expériences de gel de retardement (EMSA)
4. Mesurer l'impact des rSNPs « fonctionnels » dans la susceptibilité à la leucémie de l'enfant en effectuant une étude d'association.



**CHAPITRE II**

**MATÉRIEL ET MÉTHODES**

## **1 Populations étudiées**

### **1.1 Population de représentation mondiale**

Nous avons analysé par dHPLC 40 échantillons d'ADN (80 chromosomes) provenant d'individus appartenant à 5 groupes distincts représentant la population mondiale. La liste et l'origine des échantillons utilisés sont représentées dans le tableau VI. Les échantillons d'ADN utilisés dans le groupe d'Africains ainsi que dans le groupe d'Amérindiens proviennent de la collection Coriell. Les autres échantillons proviennent d'individus qui ont accepté de donner un échantillon de sang et certaines données démographiques (sexe, âge et origine ethnique) sur une base anonyme (non-nominatif). Ce projet a été accepté par le comité d'éthique institutionnel et un formulaire de consentement éclairé a été signé par chacun des participants. Nous avons utilisé ce nombre d'échantillons pour être en mesure de détecter les variants fréquents. Le tableau VII représente le pouvoir de détection en fonction de la fréquence du variant et de la taille de l'échantillon.

### **1.2 Le groupe « cas »**

Le groupe de patients génotypés dans les études d'association est constitué de 275 enfants atteints de LLA diagnostiqués à la clinique d'hémo-oncologie de l'hôpital Ste-Justine de Montréal. Les critères d'inclusion pour participer à l'étude sont : (1) être d'origine canadienne-française se basant sur le lieu de résidence, le nom, la langue parlée et le lieu de naissance; (2) la disponibilité du matériel biologique; (3) histoire médicale complète. Ces données sont documentées dans notre base de données. Le matériel biologique de ces patients provient du sang périphérique, de la moelle osseuse en rémission et/ou de frottis bucaux. Les patients et/ou leurs parents ont signé un formulaire de consentement dans lequel ils acceptent de participer à cette étude.

Tableau VI : Composition des groupes continentaux étudiés

<b>Groupe continental</b>	<b>Origine</b>	<b>Nombre d'individus</b>
Moyen-Orient	Égypte	1
	Maroc	1
	Liban	2
	Ashkenaze	2
	Sépharade	2
Africains	Pygmy Biaka	3
	Pygmy Mbuti	1
	Africains-américains	4
Asiatiques	Chine	2
	Vietnam	3
	Cambodge	2
	Laos	1
Européens	Québec	4
	Pologne	1
	Allemagne	1
	Bulgarie	1
	Serbie	1
Amérindiens	Amérique du Sud	8

Tableau VII. Pouvoir de détection des variants

<b>Nombre d'échantillons</b>	<b>Fréquence du polymorphisme</b>	<b>p</b>
8	0.05	0.56
	0.20	0.97
40	0.05	0.98
	0.20	0.99

p= pouvoir de détection

### **1.3 Le groupe « témoin »**

Le groupe témoin utilisé pour nos études d'association est composé de 330 individus non reliés. Les échantillons ont été obtenus sur une base volontaire et anonyme. Les critères d'inclusion, à l'exception de l'histoire médicale complète, sont identiques à ceux du groupe cas. La distribution géographique du groupe témoin est similaire à celle des patients afin de former un groupe se rapprochant le plus des cas. Ce groupe est composé de trois sous-groupes :

- 1- Sang du cordon ombilical prélevé dans la salle d'accouchement de l'hôpital (groupe contrôle A)
- 2- Individus venus au centre de prélèvement de l'hôpital pour des conditions non-cancéreuses (groupe contrôle CP)
- 3- Employés de l'hôpital (groupe contrôle MQF).

## **2 Études d'association**

Une étude d'association permet de mesurer l'impact de l'exposition ou d'une variabilité génétique, comme c'est le cas dans mon projet, sur la susceptibilité de développer la LLA. Dans une étude d'association, les sujets sont choisis sur la base de la présence (cas) ou l'absence (témoins) de la maladie spécifique étudiée, ici la LLA. L'approche la plus commune des études d'association est que les cas et les témoins sont sélectionnés dans la même population. L'association entre l'exposition ou une variation génétique et une maladie est le plus souvent exprimée par le risque relatif dans une étude de cohortes ou par le rapport de cotes (OR) dans les études d'association. Les risques relatifs et de cotes caractérisent l'association entre l'exposition ou la variation génétique et la maladie de façon relative. Dans le cas présent, le rapport de cote représente l'impact d'avoir un allèle plutôt qu'un autre en comparant la fréquence des allèles variants chez les cas à la fréquence de ceux-ci chez les témoins. La force des associations causales est mesurée avec des intervalles

de confiance de 95%. Le seuil de signification est fixé à 5%. Un OR supérieur à 1 indique une augmentation du risque de leucémie alors qu'un OR inférieur à 1 indique une protection. Le test du Chi carré est utilisé pour vérifier si la différence de distribution des génotypes entre les cas et les témoins est significative.

### **3 Extraction d'ADN**

L'ADN a été extrait des cellules mononucléées à l'aide de protocoles standards (Baccichet, A. *et al.* 1997, Sinnett, D. *et al.* 1988). La trousse d'extraction *DNA Purification Kit* de PUREGENE a été utilisée dans le cas d'extraction d'ADN à partir du sang périphérique. Le protocole d'extraction pour 10 ml de sang a été suivi selon les directives du fabricant. La concentration de l'ADN a été mesurée à l'aide d'un spectrophotomètre (Ultrospec 2000 de *Pharmacia Biotech*).

### **4 Détection de SNPs par HPLC dénaturant**

#### **4.1 Dessin d'amorces**

Les amorces de PCR ont été choisies à l'aide du programme Primer3 ([http://www-genome.wi.mit.edu/cgi-bin/primer/primer3\\_www.cgi](http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi)) afin de générer des produits de PCR de 150-450 pb de façon à couvrir une région de 2 kb en amont de l'exon 1 des gènes étudiés. Les caractéristiques de ces amorces sont présentées en annexe I. Pour chaque amorce, un BLAST (<http://www.ncbi.nlm.nih.gov/>) a été effectué pour s'assurer que celle-ci était spécifique à la séquence d'intérêt. Lorsque les fragments générés avaient un profil de dénaturation non-homogène, prédit par le programme WAVEMAKER, des clamps GC ont été ajoutées à une ou l'autre des amorces pour homogénéiser la température de dénaturation du fragment.

## 4.2 Amplification PCR de l'ADN génomique

Les amplifications PCR ont été effectuées dans un volume de 20 à 60  $\mu\text{l}$  de réactions contenant de 5 à 20 ng d'ADN génomique, de 0,1 à 0,42  $\mu\text{M}$  de chaque amorce, de 0,1 à 0,2 mM de dNTPs (Amersham Pharmacia Biotech), de 1,2 à 2 mM de  $\text{MgCl}_2$ , de 0,5 à 1,8 U d'ADN polymérase (Platinum Taq DNA polymerase (GibcoBRL) ou FastStart Taq DNA polymerase (Roche)). Les détails pour chaque réaction sont donnés en annexe II. Les PCRs se font dans un appareil GeneAmp PCR System 9700 (Applied Biosystems) en plaque de 96 puits. La dénaturation initiale de l'ADN et l'activation de la polymérase s'effectuent à 95°C pendant 5 min suivi de cycles de dénaturation à 95°C pendant 30s, d'appariement à une température variant de 51 à 63°C selon les cas, d'élongation à 72°C pendant 30s et d'une élongation finale à 72°C pendant 7 min. Pour les détails, voir l'annexe III. Le succès de chaque amplification a été vérifié par électrophorèse sur gel d'agarose 1%.

## 4.3 dHPLC

Le principe de la technique est décrit dans le chapitre 1 section 4.3. Les produits de PCR, préalablement dénaturés à 95°C pendant 4 min, sont refroidis pendant 45 min à 25°C pour permettre aux échantillons hétérozygotes de former des hétéroduplexes et des homoduplexes. La plaque de PCR 96-puits est ensuite placée dans l'échantillonneur automatique du dHPLC (WAVE DNA Fragment Analysis System, Transgenomic). À l'aide du programme WAVEMAKER, on sélectionne les températures à tester sur des échantillons contrôles. La liste des températures choisies pour chaque fragment est énumérée en annexe IV. Pour chaque échantillon, 5  $\mu\text{l}$  est injecté, les fragments sont élués à un débit de 1,5 ml/min. Entre chaque injection, la colonne est lavée avec 50% d'acétonitrile utilisant l'accélérateur du dHPLC pendant 0,5 min et ré-équilibrée. Les tampons utilisés sont préparés avec de l'acétate triéthylammonium (TEAA). La concentration du tampon A est de 0,1 M de TEAA et la concentration du tampon B est de 0,1 M de TEAA avec 25% d'acétonitrile. Le débit est de 16 échantillons par température par heure par dHPLC.

#### 4.4 Séquençage

Afin de déterminer la nature des rSNPs retrouvés dans les hétérozygotes, un échantillonnage des individus montrant un profil de dHPLC variable est séquencé. Comme contrôle, on séquence un échantillon provenant d'un profil de dHPLC non-variable. Initialement, les échantillons étaient séquencés de façon manuelle. Les produits de PCR à séquencer étaient préalablement purifiés en utilisant la trousse Rapid PCR Purification System (GibcoBRL). Une fois purifiés, les produits de PCR étaient séquencés à l'aide de la trousse Thermo Sequenase Radiolabeled Terminator Cycle Sequencing (USB). Les réactions ont été préparées suivant les instructions du fabricant. En résumé, 4 réactions sont préparées chacune contenant 0,5 µl d'un [ $\alpha$ -<sup>33</sup>P]ddNTP et 2 µl de mélange de nucléotides. À chaque tube fut ajouté 4,5 µl d'un mélange contenant 2 µl de tampon de réaction, 5 µl de produit de PCR, 1 µl d'amorce 25 µM et 2 µl d'enzyme. Les réactions furent ensuite placées dans un appareil thermocycler « Touchdown » de Hybaid pendant 45 cycles de 95°C, 30s; 55°C, 30s et 72°C, 30s. Ensuite, 4 µl de solution stop ont été ajoutés. Avant de déposer les échantillons sur gel, ils sont placés 10 min à 70°C. Un gel dénaturant d'acrylamide 6% (19 :1 acrylamide/bis acrylamide) fut utilisé et migré à 50 mA. Le gel est ensuite démoulé, transféré sur un papier Whatman, séché et placé dans une cassette avec un film BioMax de Kodak. Par contre, la plus grande partie du séquençage a été effectuée au Centre d'innovation Génome Québec et Université McGill à l'aide d'un appareil 3700 DNA analyser d'ABI (applied biosystems).

#### 4.4 Identification de séquences conservées entre l'homme et la souris

Pour affirmer qu'une région est hautement conservée entre l'homme et la souris, on doit trouver plus de 50 pb consécutives dont l'identité est supérieure à 80%. Une région moins conservée correspond à plus de 50 pb consécutives dont l'identité est inférieure à 80%. Les critères mentionnés précédemment corrélaient bien avec ce qui a été mentionné dans les articles de Hardison, RC. *et al.* 2003 et Elnitski, L. *et al.* 2003.

#### **4.5 Prédiction *in silico***

Ces prédictions proviennent du programme Matinspector (<http://www.genomatix.de/cgi-bin/matinspector.pl>). Ce programme utilise la base de données TRANSFAC. De plus, un score est attribué à chaque site prédit pour estimer la probabilité de liaison du facteur en se basant sur la comparaison de notre séquence avec la séquence consensus du site de liaison du facteur. Les résultats, dont le score est supérieur à 0,7, sont incorporés dans notre base de données.

### **5 Hybridation spécifique d'un allèle ( ASO)**

#### **5.1 Principe**

L'hybridation ASO (allele-specific-oligonucleotide) est une technique d'hybridation permettant le génotypage de mutations préalablement identifiées. La stabilité de l'association d'une sonde oligonucléotidique avec un brin d'ADN dépend non seulement de la température, mais également du degré d'appariement. Un mésappariement induit une instabilité de l'association entre la sonde et l'ADN, en comparaison à un appariement parfait. Par conséquent, on peut génotyper un échantillon d'ADN pour des variations connues en effectuant une série d'hybridations avec des oligonucléotides rendant compte de ces variations dans leur séquence (figure 19).



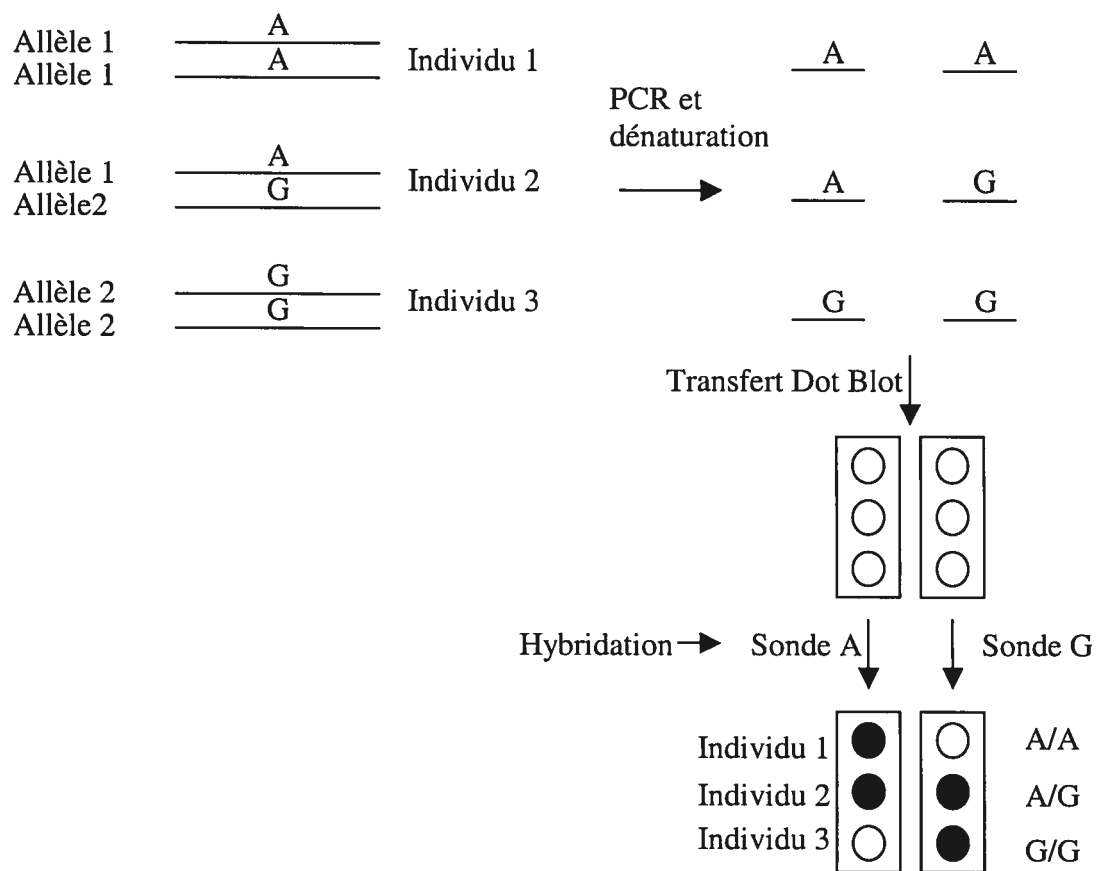


Figure 19. Méthode de génotypage par l'approche PCR-ASO.

## 5.2 Transfert des produits de PCR sur membrane (Dot blot)

La région génomique contenant le site polymorphique d'intérêt est amplifiée par la méthode PCR. Les amorces utilisées et les conditions d'amplification sont indiquées en annexe I. Les fragments de PCR sont dénaturés avant d'être transférés sur une membrane de nylon pour y être fixés. Pour chaque échantillon, 180  $\mu$ l de solution de dénaturation (0,4 N NaOH ; 20 mM EDTA ; 0,02% orange II) sont ajoutés directement dans le puits de la plaque de PCR. Des aliquots de 100  $\mu$ l sont ensuite transférés en duplicata sur des membranes de nylon (Hybond N+, Amersham) par aspiration (Minifold 1, Schleicher et Schuell). Les deux membranes sont alors brièvement rincées dans un peu d'eau distillée avant d'être disposées dans un four à UV (*Stratagene UV stratalinker* 1800). Les membranes sont alors exposées à 120 mJ de rayonnement UV afin de lier de façon covalente l'ADN à la membrane.

## 5.3 Marquage radioactif des sondes ASO

La liste complète des sondes oligonucléotidiques utilisées dans cette étude est présentée dans le tableau VIII. Des oligonucléotides de 15 mer complémentaires à la séquence correspondant à chaque variant, ont été synthétisés chez *Invitrogen*. Les sondes ont été marquées en 5' avec du [ $\gamma^{32}$ ]ATP (5000Ci/mmol) et la T4 kinase (Gibco BRL). Pour chaque sonde, 1,67  $\mu$ M est incubé pendant 30 min à 37°C en présence d'eau, de 0,325  $\mu$ M de [ $\gamma^{32}$ ]ATP, 1 X de tampon kinase F et 10 u de T4 kinase. Ensuite, 15  $\mu$ l de EDTA sont ajoutés pour stopper la réaction.

Tableau VIII. Liste des sondes ASO et conditions d'hybridation

Polymorphismes	Séquence des sondes	Tm	T d'hybridation	T de lavage
KU70-731A	GACGTAGAGGGCGTT	40	54 à 30*	40
KU70-731G	GACGTAGGGGGCGTT	42	54 à 30	40
KU70-533C	CGTTAGCCTTAAGTG,	34	54 à 30	30
KU70-533G	CGTTAGCGTTAAGTG	34	54 à 30	30
KU70-489G	AGACCACGCTCCTTC	40	54 à 30	30
KU70-489A	AGACCACACTCCTTC	37	54 à 30	30
KU70-1599G	GAGGCAGGAGTGCAC	42	54 à 30	30
KU70-1599A	GAGGCAGAAGTGCAC	40	54 à 30	30
KU70-1469C	GCATGACTCCTCCAT	37	54 à 30	30
KU70-1469T	GCATGATTCCTCCAT	34	54 à 30	30
KU70-1296C	TCGGCCAGCCGCCA	44	54 à 30	45
KU70-1296G	TCGGCCAGGCGCCA	44	54 à 30	45
XRCC1-R399	TCCC GGAGGTAAGGC	55	46	46
XRCC1-Q399	TCCCAGAGGTAAGGC	49	46	46
XRCC3-T241	GGCCA CGCTGCGTGA	63	55	55
XRCC3-M241	GGCCATGCTGCGTGA	60	55	55
hOGG1-S326	GCCAATCCCGCCATG	60	55	55
hOGG1-C326	GCCAATGCCGCCATG	62	55	55
MnSOD-V-9	TGGCTCCGGTTTTGG	57	50	50
MnSOD-A-9	TGGCTCCGGCTTTGG	60	50	50
MnSOD-I58	CCCAGATAGCTCTTC	42	37	37
MnSOD-T58	CCCAGACAGCTCTTC	46	37	37

\*Toutes les sondes dont la température d'hybridation indiquée est de 54 à 30 ont été hybridées simultanément dans le même four.

## 5.4 Hybridation

Les paires de membranes de nylon sur lesquelles furent fixés les échantillons d'ADN dénaturés sont placées dans deux tubes de verre différents. Un volume suffisant de solution d'hybridation contenant 1X SSPE (150 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, 1 mM EDTA, pH 7,4), 0,75 M NaCl, 70 mM Tris, 1% SDS et 100 mg d'héparine, préalablement chauffée, est ajouté dans les tubes puis ils sont placés dans un four rotatif (Modèle 400, Robbins Scientific) à la température d'hybridation spécifique pour une pré-hybridation (voir tableau VIII). Après 30 minutes, les sondes sont ajoutées, à raison de 2,5 µl de sondes marquées pour 12,5 ml de solution d'hybridation. Afin de s'assurer d'une plus grande spécificité des hybridations, il fut décidé d'ajouter en excès un compétiteur. Ce compétiteur est la sonde froide spécifique pour l'autre allèle étudié. Ici, 5 µl de sonde froide à 25µM furent ajoutés ce qui correspond à un ratio 15 :1 entre la sonde compétitrice et la sonde radioactive. Ces sondes froides permettent de saturer les sites auxquels les sondes marquées pourraient se lier de façon non spécifique. Les tubes sont ensuite remis dans le four pour une durée d'environ 30 minutes, temps amplement suffisant pour permettre l'hybridation dans le cas des hybridations simples ou bien jusqu'à ce que la température redescende à la température minimum de l'hybridation lorsque plusieurs hybridations sont effectuées en même temps dans un même four. Les membranes sont ensuite rincées 2 fois à température de la pièce avec une solution de SSPE 2X contenant 0,1% de SDS et ensuite 1 fois pendant 15 minutes à la température de lavage indiquée dans le tableau VIII. Les membranes sont alors transférées dans des sacs en plastique qui seront scellés et placés dans une cassette d'exposition avec un film (Kodak, X-omat) en présence d'écrans intensificateurs pendant une nuit à -80°C. Les membranes peuvent être réutilisées si nécessaire après décrochage des sondes dans une solution bouillante de SDS 0,5%.

## 6 Gel de retardation (EMSA)

### 6.1 Principe

La technique de gel de retardation permet de détecter les interactions ADN-protéines. Cette méthode a été largement utilisée dans l'étude de séquence d'ADN spécifique liant des protéines comme les facteurs de transcription. La technique est basée sur l'observation que des complexes de protéines et d'ADN migrent plus lentement à travers un gel non-dénaturant de polyacrylamide que des fragments libres d'ADN ou des oligonucléotides double-brin. Cette technique consiste à incuber une protéine purifiée ou un mélange de protéines (comme les extraits cellulaires ou nucléaires) avec un fragment d'ADN marqué au  $P^{32}$  contenant le site putatif de liaison de la protéine. La réaction est par la suite analysée sur un gel non-dénaturant de polyacrylamide (figure 20). La spécificité de la liaison de la protéine à l'ADN pour le site de liaison est établie par des réactions de compétition en utilisant des fragments d'ADN ou des oligonucléotides contenant le site de liaison de la protéine d'intérêt ou une autre séquence d'ADN non reliée.

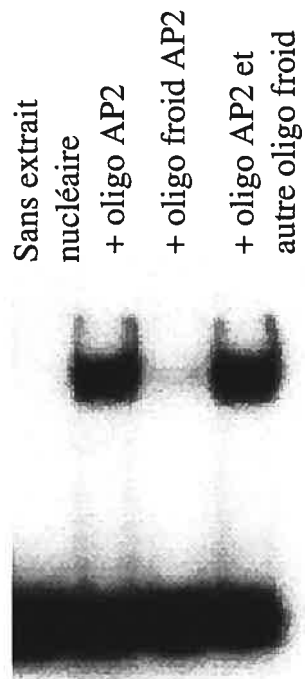


Figure 20. Gel de retardation. Ligne 1 : contrôle négatif; ligne 2 : contrôle positif (extrait nucléaire + oligo AP2 marqué au  $P^{32}$ ; ligne 3 : oligo AP2 marqué au  $P^{32}$  + oligo froid AP2 (compétiteur spécifique); ligne 4 : oligo AP2 marqué au  $P^{32}$  + oligo TFIIID froid (compétiteur non-spécifique).

## 6.2 Extraits nucléaires

Nous avons utilisé des cellules HeLa pour obtenir des extraits nucléaires nécessaires aux expériences de gel de retardation. Les cellules ont été récoltées et lavées avec 10 ml de TBS (Tris buffered saline). Ensuite les cellules ont été centrifugées pendant 5 min à 1500 x g. Le culot est resuspendu dans 1 ml de TBS, transféré dans un tube eppendorf et centrifugé dans une microfuge à 13 000 rpm pendant 15 sec. Le TBS est ensuite enlevé et le culot de cellules resuspendu dans 400 µl de tampon A froid (10 mM HEPES pH 7,9; 10 mM KCl; 0,1 mM EDTA; 0,1 mM EGTA; 1 mM DTT; 0,5 mM PMSF) en pipettant doucement. Les cellules sont placées dans la glace pendant 15 min, ensuite 25 µl de 10% Nonidet NP-40 est ajouté et le tube est vortexé vigoureusement pendant 10 sec. L'homogénat est centrifugé pendant 30 sec dans une microfuge à 13 000 rpm. Le culot de noyaux est resuspendu dans 50 µl de tampon C froid (20 mM HEPES pH 7,9; 0,4 M NaCl; 1 mM EDTA; 1 mM EGTA; 1 mM DTT; 1 mM PMSF, 1 mM Aprotinin) et le tube est vigoureusement mélangé à 4°C pendant 15 min en le plaçant sur une plate-forme agitatrice. L'extrait nucléaire est centrifugé pendant 5 min dans une microfuge à 13 000 rpm à 4°C. Le surnageant est ensuite récupéré et les protéines sont dosées en utilisant le réactif Bradford (Biorad). L'extrait nucléaire est congelé à -80°C jusqu'à utilisation (Schreiber, E. *et al.* 1989).

## 6.3 Appariement des sondes

Les expériences de gel de retardation ont été effectuées dans le but de confirmer des prédictions informatiques. Les rSNPs découverts par dHPLC qui potentiellement font apparaître un site de liaison de facteurs de transcription ont été testés. La liste des sondes testées est énumérée dans le tableau IX.

Tableau IX. Liste des sondes utilisées pour le gel de retardement.

Gène	rSNP	Séquences oligonucléotidiques
DCLRE1C	-1571C	5'-CCAGTTCACTCAGCCTGGGCAAC-3' 3'-GGTCAAGTGAGGTCGGACCCGTTG-5'
	-1571G	5'-CCAGTTCACTGCAGCCTGGGCAAC-3' 3'-GGTCAAGTGACGTCGGACCCGTTG-5'
DCLRE1C	-1892G	5'-TGTCCGAAGAATCTCACGG-3' 3'-ACAGGCTCTTAGAGTGCC-5'
	-1892C	5'-TGTCCGAACAATCTCACGG-3' 3'-ACAGGCTGTAGAGTGCC-5'
KU70	-489G	5'-ACTCGAGACCACGCTCCTTC-3' 3'-TGAGCTCTGGTGCGAGGAAGG-5'
	-489A	5'-ACTCGAGACCACACTCCTTCC-3' 3'-TGAGCTCTGGGTGAGGAAGG-5'
KU70	-1687C	5'-CTCCGAGAGCCACCCGACCT-3' 3'-GAGGCTCTCGGTGGGCTGGA-5'
	-1687T	5'-CTCCGAGAGTCACCCGACCT-3' 3'-GAGGCTCTCAGTGGGCTGGA-5'
KU80	-205A	5'-TGCTCGGAGAGAATCTGCGC-3' 3'-ACGAGCCTCTCTTAGACGCG-5'
	-205G	5'-TGCTCGGAGGGAATCTGCGC-3' 3'-ACGAGCCTCCCTTAGACGCG-5'
KU80	-207A	5'-GCATGCTCGGAGAGAATCTGC-3' 3'-CGTACGAGCCTCTCTTAGACG-5'
	-207C	5'-GCATGCTCGCGGAGAATCTGC-3' 3'-CGTACGAGCCGCTCTTAGACG-5'
KU80	-297C	5'-GTCGTA TAGCGTTTCAGGC-3' 3'-CAGCATGATCGCAAAGTCCG-5'
	-297T	5'-GTCGTA TAGTGTTCAGGC-3' 3'-CAGCATGATCAAAAGTCCG-5'
KU80	-1669T	5'-ACTAATCAGATTTGGCTTCT-3' 3'-TGATTAGTCTAAACCGAAGA-5'
	-1669A	5'-ACTAATCAGATATGGCTTCT-3' 3'-TGATTAGTCTATACCGAAGA-5'
MRE11A	-1226A	5'-ATGTAAGAATTTTACTAAAGTA-3' 3'-TACATTCTTAAAAATGATTTTCAT-5'
	-1226C	5'-ATGTAAGAATTTTCCTAAAGTA-3' 3'-TACATTCTTAAAAAGGATTTTCAT-5'
NBS1	-692C	5'-TCCTGATCACATACATTTTG-3' 3'-AGGACTAGTGTATGTAAAAC-5'
	-692T	5'-TCCTGATCATATACATTTTG-3' 3'-AGGACTAGTATATGTAAAAC-5'
RAD50	-52A	5'-TTGTGGACTCCAGGTCCCTG-3' 3'-AACACCTGAGGTCCAGGGAC-5'
	-52G	5'-TTGTGGGCTCCAGGTCCCTG-3' 3'-AACACCCGAGGTCCAGGGAC-5'



RAD50	-1347 -1347INST	5'-TTCCGGTTTTTTTAACTGG-3' 3'-AAGGCCAAAAAATTGGACC-5' 5'-TTCCGGTTTTTTTAACTGG-3' 3'-AAGGCCAAAAAATTGGACC-5'
XRCC2	-1164G -1164C	5'-GAAAAAAGAAAAGTAAGTG-3' 3'-CTTTTTTCTTTTCATTCAC-5' 5'-GAAAAAACAAAAGTAAGTG-3' 3'-CTTTTTTGTTCATTCAC-5'
XRCC4	-1864T -1864C	5'-CCTCCCGAGTAGCTGGGACTACAGG-3' 3'-GGAGGGCTCATCGACCCTGATGTCC-5' 5'-CCTCCCGAGCAGCTGGGACTACAGG-3' 3'-GGAGGGCTCGTCGACCCTGATGTCC-5'

Les oligonucléotides ont été synthétisés chez *Invitrogen*. Chaque oligonucléotide était simple brin. Pour obtenir des sondes doubles-brins, les oligonucléotides 5'-3' (1,75  $\mu\text{M}$ ) ont été incubés avec les oligonucléotides 3'-5' correspondants (1,75  $\mu\text{M}$ ) dans un tampon d'appariement (10x= 100 mM Tris-HCl pH 7,5; 0,5 M NaCl; 10 mM EDTA) pendant 3 min à 95°C et ensuite 1 h à la température de la pièce.

#### 6.4 Marquage des sondes

Chaque sonde double-brin est ensuite marquée radioactivement. Le protocole de marquage consiste à incuber pendant 10 min une solution contenant 2  $\mu\text{l}$  de la sonde à marquer (1,75 pmol/ $\mu\text{l}$ ), 1  $\mu\text{l}$  du tampon d'échange 10x, 1  $\mu\text{l}$  de la kinase T4, 5  $\mu\text{l}$  d'eau et 1  $\mu\text{l}$  de  $\gamma\text{ATP}$  radioactif (3000 ci/mmol, Amersham). Après l'incubation, 1  $\mu\text{l}$  d'EDTA à 0,5 M est ajouté à la solution afin d'arrêter la réaction. Ensuite, 89  $\mu\text{l}$  de tampon TE ( 10 mM Tris-HCl pH 8,0 ; 1 mM EDTA) sont ajoutés et les sondes sont passées sur colonne G-25 (Amersham) pour éliminer la radioactivité non-incorporée.

#### 6.5 Réaction de liaison

Par la suite, 4 réactions contrôles sont préparées à partir de la trousse de gel de retardation (*Promega*) : un contrôle négatif contenant 7  $\mu\text{l}$  d'eau, 2  $\mu\text{l}$  de tampon de liaison 5 x ; un contrôle positif contenant 5  $\mu\text{l}$  d'eau, 2  $\mu\text{l}$  de tampon, 2  $\mu\text{l}$  d'extrait nucléaire HeLa fourni ; un essai avec un compétiteur spécifique contenant 4  $\mu\text{l}$  d'eau, 2  $\mu\text{l}$  de tampon, 2  $\mu\text{l}$  d'extrait nucléaire, 1  $\mu\text{l}$  de sonde compétitrice non-marquée (SP1)(1,75 pmol) et un essai avec un compétiteur non- spécifique contenant 4  $\mu\text{l}$  d'eau, 2  $\mu\text{l}$  de tampon, 2  $\mu\text{l}$  d'extrait nucléaire et 1  $\mu\text{l}$  de sonde non-compétitrice non-marquée (AP2) (1,75 pmol). Pour chaque sonde à tester, 2 réactions sont préparées : une réaction positive contenant 4,5  $\mu\text{l}$  d'eau, 2  $\mu\text{l}$  de tampon, 2,5  $\mu\text{l}$  d'extrait nucléaire HeLa préparée précédemment et un essai avec un compétiteur spécifique contenant 3,5  $\mu\text{l}$  d'eau, 2  $\mu\text{l}$  de tampon, 2,5  $\mu\text{l}$  d'extrait nucléaire et 1  $\mu\text{l}$  de la sonde testée non-marquée (1,75 pmol). Toutes les réactions sont incubées à 22°C pendant 10 min. Ensuite, 1  $\mu\text{l}$  de sonde marquée radioactivement est ajouté à chaque réaction

correspondante. Les tubes sont incubés 20 min à 22°C. Par la suite, 1  $\mu$ l de tampon de chargement 10X (250 mM Tris-HCl pH 7,5 ; 0,2% bleu de bromophénol ; 40% glycérol) est ajouté.

## **6.6 Électrophorèse**

Pour l'électrophorèse, on utilisait un gel non-dénaturant d'acrylamide 4% (60 :1 acrylamide/bis acrylamide) ou 6% (29 :1 acrylamide/bis acrylamide). Après avoir déposé les échantillons, l'électrophorèse se déroule à la température de la pièce à 350V, pour un gel 4% ou à 190V pour un gel 6% jusqu'à ce que le bleu soit rendu aux trois-quarts du gel. Après démoulage, le gel est séché (sur un papier Whatman), puis exposé en présence d'un film X-Omat de Kodak pendant la nuit à -80°C avec des écrans intensificateurs.

**CHAPITRE III**  
**RÉSULTATS**

## **1 Outil de géotypage**

Nous avons utilisé l'approche PCR-ASO afin de géotyper un grand nombre d'individus. Le principe de la méthode est décrit à la section 6 de la partie Matériel et Méthodes. Des résultats représentatifs sont montrés dans la figure 21. Cette approche a été utilisée pour les 2 études d'association de mon projet.

## **2 Polymorphismes dans les gènes XRCC1 (Arg399Gln), XRCC3 (Thr241Met) et hOGG1 (Ser326Cys)**

### **2.1 Détermination de la fréquence de ces variants chez les patients LLA et chez les témoins**

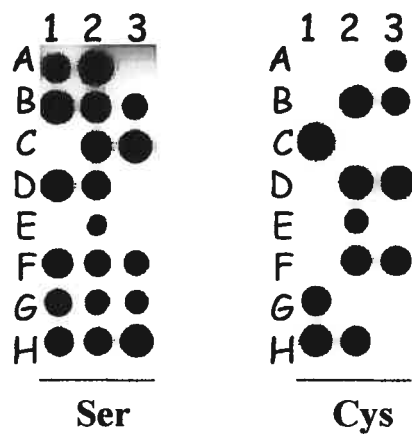
Nous avons déterminé la fréquence des 3 polymorphismes illustrés à la figure 22 chez un groupe d'enfants atteints de la LLA (n=275) et chez un groupe d'individus sains (n=330), les deux groupes sont formés de Canadiens français résidant dans la province de Québec. La fréquence allélique et la distribution des géotypes chez les patients et chez les témoins sont donnés dans les tableaux X et XI respectivement. Pour quelques individus, les géotypes n'ont pu être déterminés, ce qui explique la variation du nombre d'échantillons listés dans les tableaux. La distribution des géotypes du gène XRCC3 et hOGG1 suit la loi de Hardy-Weinberg. Par contre la distribution des géotypes du gène XRCC1 ne la suit pas (patients  $p=0,004$  et contrôles  $p=0,003$ ). Ce résultat pourrait être expliqué par la présence d'un autre polymorphisme qui serait en déséquilibre de liaison avec celui étudié contre lequel serait exercé une pression sélective. La fréquence allélique entre le groupe de patients et le groupe de témoins diffère pour le gène hOGG1 : l'allèle Cys326 est sous-représenté dans le groupe des patients (tableau X). Par contre, aucune différence de distribution de la fréquence allélique n'est observée pour les gènes XRCC1 et XRCC3. Lorsque les géotypes ont été déterminés, nous avons observé que les hétérozygotes hOGG1Ser326/Cys326 et les homozygotes Cys326 étaient présents à 30,88% chez les patients comparativement à 43,47% chez les témoins (tableau XI). Ces résultats suggèrent que le fait d'être porteur d'au moins un allèle Cys326 est associé à une diminution du risque de développer la LLA (OR = 0,58 ; IC95%=0,41-

0,81 ;  $p=0,002$ ). Une légère différence entre la fréquence des homozygotes Met241 du gène XRCC3 est observée entre les patients et les témoins. Ces résultats pourraient expliquer la diminution du risque, qui est non-significative, observée chez les individus homozygotes comparativement aux individus hétérozygotes et homozygotes Thr241 ( $OR=0,64$  ;  $IC95\%=0,39-1,05$  ;  $p=0,084$ ). La fréquence des génotypes du gène XRCC1 ne diffère pas entre les patients et les témoins ce qui suggère que ce polymorphisme n'est pas relié au risque de développer la LLA.

## **2.2 Détermination des allèles ancestraux**

Pour déterminer les allèles ancestraux pour chacun des polymorphismes génotypés, de l'ADN de primates non-humains a été analysé. Pour XRCC1, l'allèle ancestral a été déterminé par PCR/ASO et l'allèle trouvé est l'allèle majeur chez les humains. Pour XRCC3, l'allèle ancestral n'a pu être déterminé puisque le PCR ne fonctionnait pas avec l'ADN de chimpanzés. Pour hOGG1, l'allèle ancestral est l'allèle majeur chez l'humain. Par contre, puisque l'hybridation ne donnait aucun résultat, les échantillons de chimpanzés ont été séquencés. Une différence au niveau de la séquence a été observée près de la position du polymorphisme ce qui explique pourquoi l'hybridation ne fonctionnait pas (tableau XII).

**Ser326: GCC AAT CCC GCC ATG**  
**Cys326: GCC AAT GCC GCC ATG**



**A2= homozygotes Ser326**

**B2= hétérozygotes**

**D3= homozygotes Cys326**

Figure 21. Résultats d'hybridation par la méthode PCR-ASO.

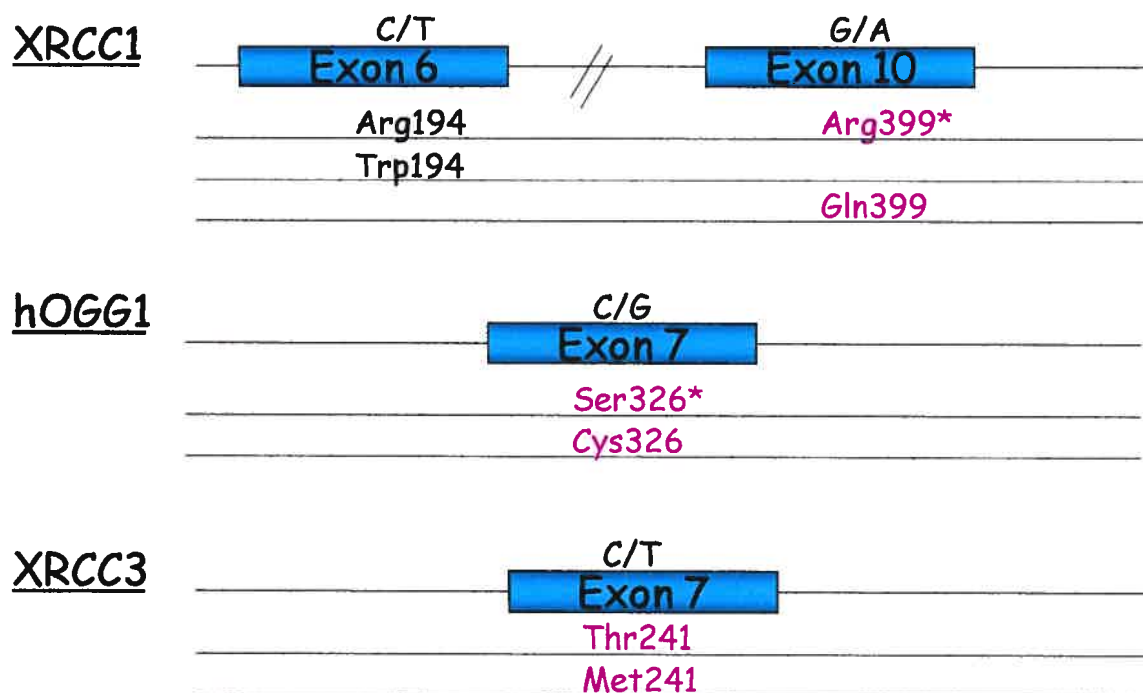


Figure 22. Représentation graphique des allèles étudiés. \*allèle ancestral (tableau XII).



Tableau X. Fréquence allélique des gènes XRCC1, XRCC3 et hOGG1 chez les patients LLA et les témoins

Allèles	patients LLA		Témoins		OR*	IC 95%	p
	n	%	n	%			
<b>XRCC1</b>							
Arg399	347	62.86	408	61.63			
Gln399	205	37.14	254	38.37	0,95	0,75-1,20	0,678
	552		662				
<b>XRCC3</b>							
Thr241	333	64.53	382	60.44			
Met241	183	35.47	250	39.56	0,84	0,66-1,07	0,160
	516		632				
<b>hOGG1</b>							
Ser326	447	82.17	491	74.62			
Cys326	97	17.83	167	25.38	<b>0,64</b>	<b>0,48-0,85</b>	<b>0,002</b>
<b>Total</b>	<b>544</b>		<b>658</b>				

\*OR calculés à partir du ratio du nombre d'allèle mineur versus l'allèle majeur chez les patients comparé avec le ratio chez les individus témoins.

Tableau XI. Distribution des génotypes du gène XRCC1, XRCC3 et hOGG1 chez les patients LLA et chez les témoins

Génotypes	patients LLA		Témoins		OR*	IC 95%	p
	n	%	n	%			
<b>XRCC1</b>							
Arg399/Arg399	98	35.51	113	34.14			
Arg399/Gln399	151	54.71	182	54.98	0.96	0.68-1.35	0.86
Gln399/Gln399	27	9.78	36	10.88	0.86	0.49-1.53	0.667
Arg399/Gln399 + Gln399/Gln399	178	64.49	218	65.86	0.94	0.67-1.31	0.733
	276		331				
<b>XRCC3</b>							
Thr241/Thr241	102	39.53	115	36.39			
Thr241/Met241	129	50.00	152	48.10	0.96	0.67-1.37	0.856
Met241/Met241	27	10.47	49	15.51	0.62	0.36-1.07	0.107
Met241/Met241**	27	10.47	49	15.51	0.64	0.39-1.05	0.084
	258		316				
<b>hOGG1</b>							
Ser326/Ser326	188	69.12	186	56.53			
Ser326/Cys326	71	26.10	119	36.17	<b>0.59</b>	<b>0.41-0.84</b>	<b>0.004</b>
Cys326/Cys326	13	4.78	24	7.29	0.54	0.33-0.93	0.087
Ser326/Cys326 + Cys326/Cys326	84	30.88	143	43.47	<b>0.58</b>	<b>0.41-0.81</b>	<b>0.002</b>
	272		329				

\*OR calculé à partir du ratio du nombre d'hétérozygotes, d'homozygotes des allèles mineurs ou de la somme des hétérozygotes et des homozygotes de l'allèle mineur versus le nombre d'homozygotes de l'allèle majeur chez les patients comparé avec le ratio chez les témoins

\*\*OR calculé à partir du ratio d'homozygotes de l'allèle mineur versus la somme des hétérozygotes et des homozygotes de l'allèle majeur chez les patients comparé avec le ratio chez les témoins

Tableau XII. Allèles ancestraux

Gène	Allèle majeur chez l'humain	Allèle ancestral
XRCC1	Arg399	Arg399
XRCC3*	Thr241	?
hOGG1**	Ser326	Ser326

\*L'amplification n'a pas fonctionné

\*\*Séquence chez l'humain GCC AAT CCC GCC ATG; Séquence chez le chimpanzé GCC AAT CCC GCT GTG; en italique c'est la position du polymorphisme et en gras c'est la différence de séquence

### **3 Détection de polymorphismes dans les gènes de réparation des cassures doubles-brins.**

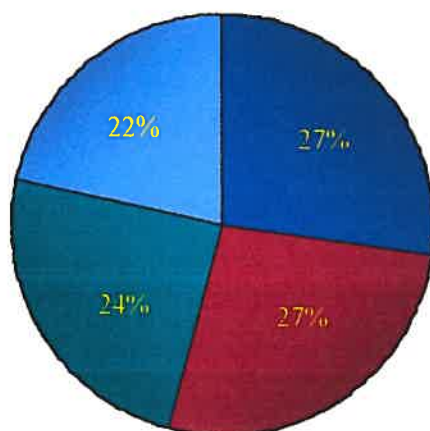
#### **3.1 Description des polymorphismes identifiés**

La détection de polymorphismes a été effectuée à l'aide de la technique dHPLC décrite à la section 4 du chapitre Matériel et Méthodes. Nous avons analysé les variations de séquence pour 15 gènes codant pour des composantes du système de réparation des cassures doubles-brins : ATM, BRCA1, BRCA2, DCLER1C, KU70, KU80, LIGIV, MRE11A, NBS1, PRKDC, RAD50, RAD51, XRCC2, XRCC3 et XRCC4. La distribution globale des polymorphismes au niveau des 2 kb analysés est illustrée dans le tableau XIII. La distribution est relativement uniforme lorsqu'on subdivise les 2 kb par fenêtre de 250 pb. La densité de polymorphismes par kilobase est de 4,5. Cette valeur se rapproche des valeurs déjà publiées par d'autres groupes (Stephens, M. *et al.* 2001). Au total, 144 rSNPs distincts ont été identifiés : 53 transversions, 75 transitions, 16 délétions/insertions. En comparant le nombre de profils variables identifiés et le nombre de profils variables dont la présence de polymorphismes a été confirmée par séquençage, on obtient une efficacité de détection par le dHPLC de 79,6 % (Tableau XIII). Des 144 polymorphismes identifiés, 39 (27%) rSNPs ne sont pas associés avec un site de liaison de facteurs de transcription. Par contre, 105 (73%) SNPs font apparaître ou disparaître des sites putatifs de liaison de facteurs de transcription. Parmi ceux-ci, 35 (24%) font seulement apparaître un ou plusieurs sites, 31 (22%) font seulement disparaître un ou plusieurs sites et 39 (27%) font apparaître et disparaître un ou plusieurs sites (figure 23). L'ensemble de ces résultats est ensuite intégré dans notre base de données (voir tableau XIV).

La représentation graphique de la région régulatrice du gène Ku70 est illustrée à la figure 25 où les 17 rSNPs identifiés par dHPLC et confirmés par séquençage sont représentés. Parmi ceux-ci, 4 sont retrouvés dans une région conservée entre l'homme et la souris. Aucun rSNP n'est situé dans une région répétitive. Le tableau XIV donne une description détaillée des rSNPs ainsi que l'effet de chaque rSNP sur le gain/perte de sites de liaison de facteurs de transcription. 76% des rSNPs identifiés dans le gène Ku70 ont un effet sur le gain ou la perte de sites de liaison : 3 (18%) font gagner et perdre des sites, 6 (35%) font seulement gagner des sites et 4 (23%) font seulement perdre des sites de liaison de facteurs de transcription. La représentation graphique de chaque région régulatrice de tous les autres gènes et les tableaux plus détaillés sont en annexe V.

Tableau XIII. Compilation des résultats obtenus pour la détection de rSNPs

# gènes analysés par dHPLC	# fragments	# kb par individu	# profils avec séquence	# profils avec SNPs	# SNPs distincts	Efficacité (%)
15	92	32,2	235	187	144	79.6



■ Sans effet ■ Sites apparaissent/disparaissent ■ Sites apparaissent ■ Sites disparaissent

Figure 23. Effets des polymorphismes sur les sites putatifs de liaison de facteurs de transcription.

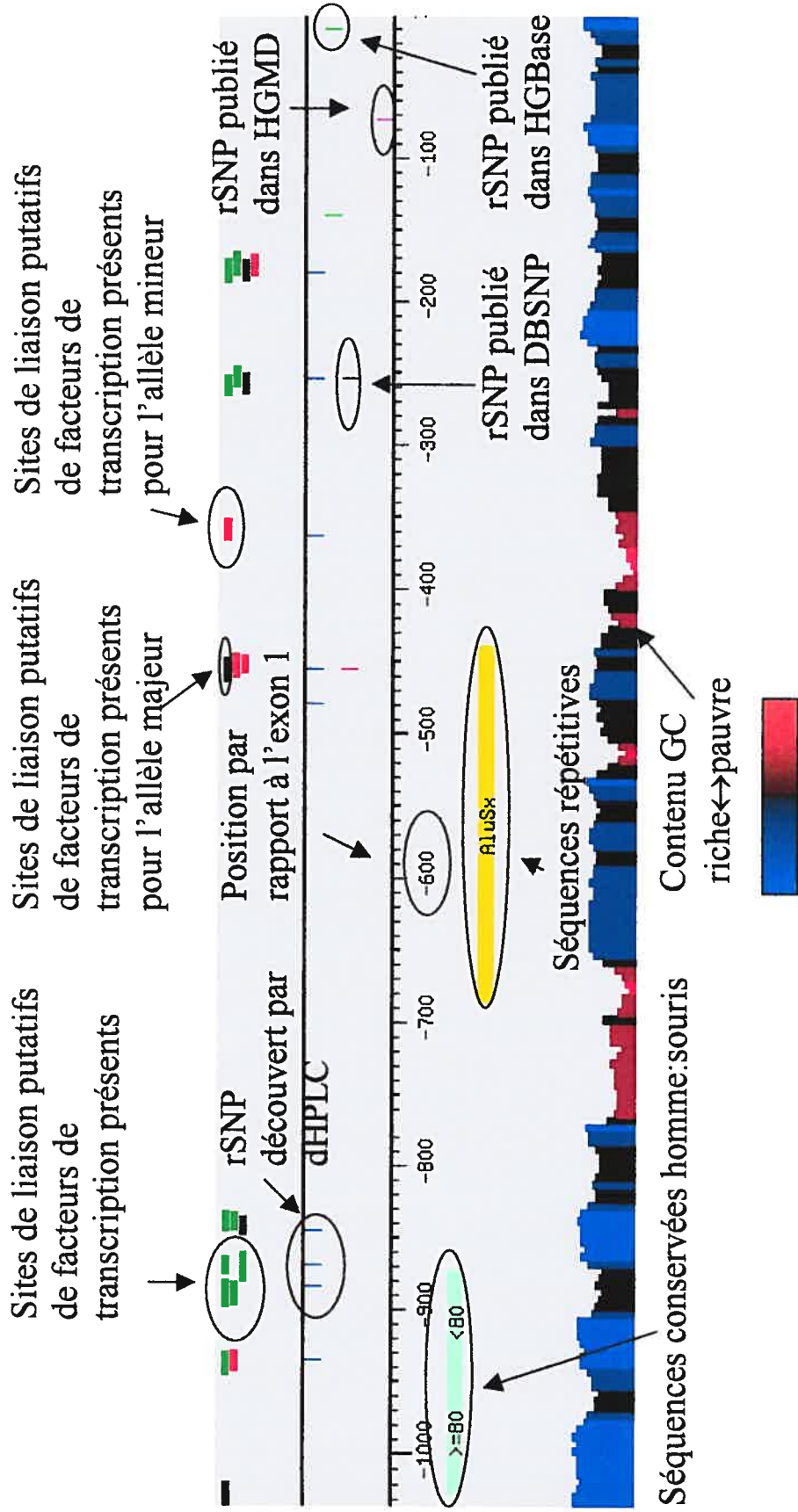


Figure 22. Légende pour les figures des représentations graphiques des régions promotrices des gènes étudiés. DBSNP, HGMD et HGBase sont des bases de données publiques.

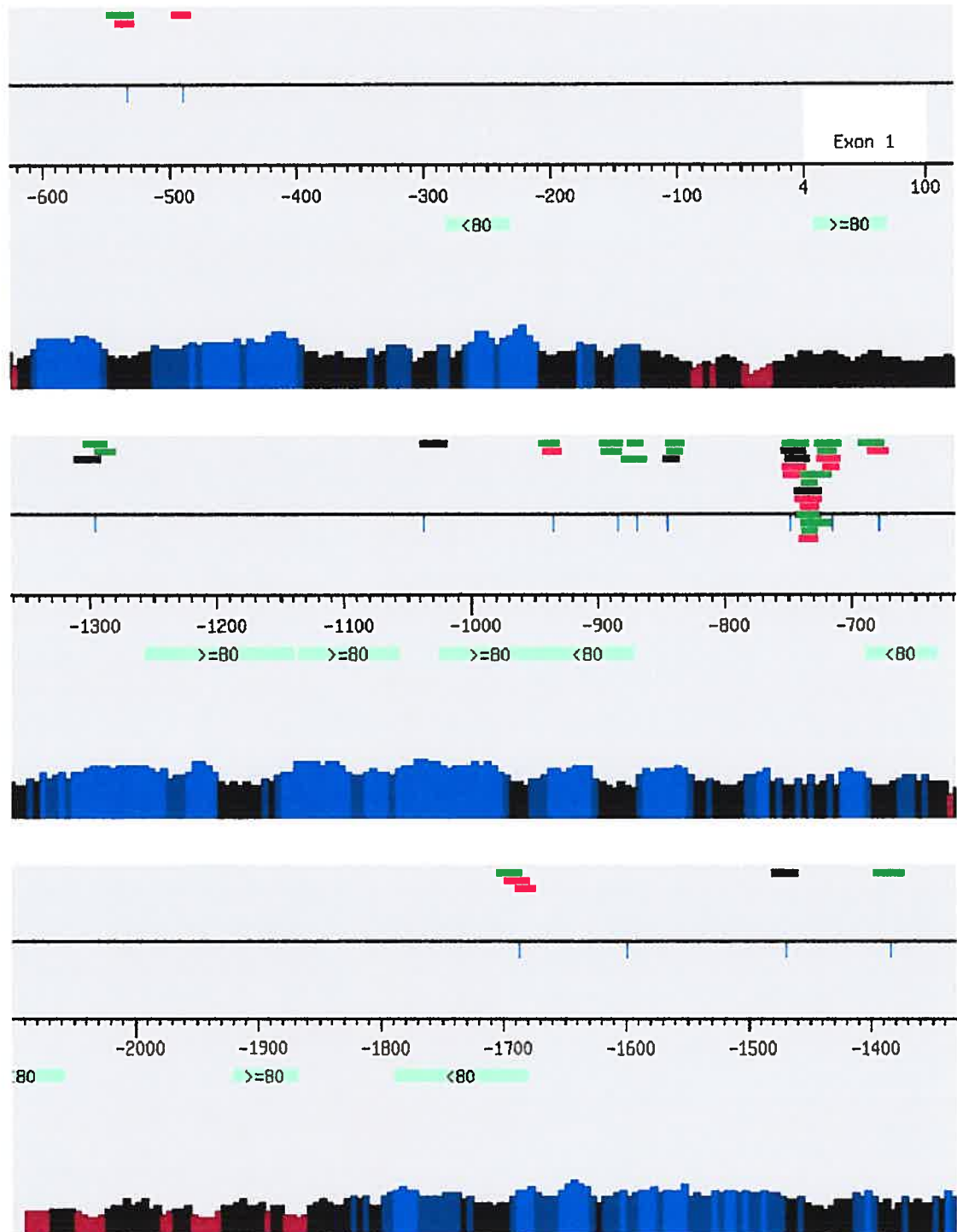


Figure 25. Représentation graphique de la région régulatrice du gène Ku70.



Tableau XIV. Description détaillée des polymorphismes identifiés dans le gène Ku70

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
KU70	Ex1.-748	C->T  transition	C	59	Muscle Initiator Sequence->0.877  Pax-6 paired domain protein->0.771  Tax/CREB complex->0.729	octamer-binding factor 1, POU-specific domain->0.88  serum response factor->0.663  Tax/CREB complex->0.775		
KU70	Ex1.-733	A->G  transition	A	55	cAMP-responsive element binding protein->0.98  GLI-Krüppel-related transcription factor, regulator of adenovirus E4 promoter->0.847  Mammalian C-type LTR CCAAT box->0.852	Activating transcription factor 6, member of b-zip family, induced by ER stress->0.987  GLI-Krüppel-related transcription factor, regulator of adenovirus E4 promoter->0.836  Mammalian C-type LTR CCAAT box->0.863  X-box-binding protein 1->0.911		
KU70	Ex1.-731	A->G  transition	A	57	cAMP-responsive element binding protein->0.98  GLI-Krüppel-related transcription factor, regulator of adenovirus E4 promoter->0.847  Mammalian C-type LTR CCAAT box->0.852	cAMP-responsive element binding protein->0.98  GLI-Krüppel-related transcription factor, regulator of adenovirus E4 promoter->0.824  Hypoxia induced factor-1 (HIF-1)->0.88  Mammalian C-type LTR CCAAT box->0.846		
KU70	Ex1.-715	C->A  transversion	C	63	COMP1, cooperates with myogenic proteins in multicomponent complex->0.777  nuclear factor Y (Y-box binding factor)->0.917	COMP1, cooperates with myogenic proteins in multicomponent complex->0.814  Ikaros 2, potential regulator of lymphocyte differentiation->0.988		

						nuclear factor Y (Y-box binding factor)->0.917 X-box binding protein RFX1->0.899	
KU70	Ex1.-678	T->C  transition	T	61	COMP1, cooperates with myogenic proteins in multicomponent complex->0.768	COMP1, cooperates with myogenic proteins in multicomponent complex->0.779  Ecotropic viral integration site 1 encoded factor->0.775	2953-3009
KU70	Ex1.-533	C->G  transversion	C	55	Tax/CREB complex->0.721	early growth response gene 3 product->0.811  Tax/CREB complex->0.721	
KU70	Ex1.-489	G->A transition	G	67		runt-factor AML-1->1	
KU70	Ex1.-1599	G->A transition	G	67			
KU70	Ex1.-1469	C->T  transition	C	59	BTB/POZ-bZIP transcription factor BACH1 forms heterodimers with the small Maf protein family->0.87		
KU70	Ex1.-1384	C->T  transition	C	53	Mammalian C-type LTR CCAAT box->0.851	Mammalian C-type LTR CCAAT box->0.858	
KU70	Ex1.-1296	C->G  transversion	C	76	Member of b-zip family, induced by ER damage/stress, binds to the ERSE in association with NF-Y->0.947  neural-restrictive-silencer-element->0.684  Nuclear factor 1 (CTF1)->0.815	Member of b-zip family, induced by ER damage/stress, binds to the ERSE in association with NF-Y->0.935  Nuclear factor 1 (CTF1)->0.821	
KU70	Ex1.-1687	C->T  transition	C	61	Ribonucleoprotein associated zinc finger protein MOK-2 (human)->0.982	Activator protein 1->0.963  Ribonucleoprotein associated zinc finger protein MOK-2 (human)->0.982  Sterol regulatory element binding protein->0.851	1853-1962

KU70	Ex1.-1037	DEL CCCGG nd	CCC GG	82	Epstein-Barr virus transcription factor R->0.843			
KU70	Ex1.-935	G->C transversion	G	73	VDR/RXR Vitamin D receptor RXR heterodimer site- >0.889	NF-kappaB->0.899 VDR/RXR Vitamin D receptor RXR heterodimer site- >0.86	2699-2770	
KU70	Ex1.-884	T->C transition	T	57	c-Myc/Max heterodimer->0.93 MIT (microphthalmia transcription factor) and TFE3->0.85	c-Myc/Max heterodimer->0.93 MIT (microphthalmia transcription factor) and TFE3->0.834	2699-2770	
KU70	Ex1.-869	DEL G nd	G	63	cAMP-responsive element binding protein->0.962  GLI-Krueppel-related transcription factor, regulator of adenovirus E4 promoter->0.836	cAMP-responsive element binding protein->0.962  GLI-Krueppel-related transcription factor, regulator of adenovirus E4 promoter->0.853		
KU70	Ex1.-845	G->C transversion	G	71	MYC-associated zinc finger protein related transcription factor- >0.922  stimulating protein 1 SP1, ubiquitous zinc finger transcription factor->0.994  Zinc finger transcription factor ZBP-89->0.961	MYC-associated zinc finger protein related transcription factor- >0.91  stimulating protein 1 SP1, ubiquitous zinc finger transcription factor->0.965		

#### 4 Étude *in vitro* de l'impact de rSNPs sur la liaison de facteurs de transcription

La présence de polymorphismes causant le gain et/ou la perte de sites de liaison de facteurs de transcription est un indice de la présence de rSNPs fonctionnels. Nous avons donc décidé de vérifier les prédictions *in silico* en utilisant la méthode de gel de retardation. Nous avons choisi de tester *in vitro* les sites où un rSNP menait à un gain de sites de liaison de facteurs de transcription (tableau XV). Pour chaque site testé, deux sondes d'ADN double-brin d'environ 20 nucléotides ont été utilisées chacune correspondant soit à la séquence de l'allèle majeur ou à l'allèle mineur. Nous avons examiné la capacité de chaque sonde à lier des facteurs nucléaires isolés à partir de cellules HeLa. L'analyse des complexes ADN-protéines par électrophorèse sur gel de polyacrylamide non-dénaturant a révélé, pour le rSNP Ku70-1687, l'apparition d'une bande proéminente et spécifique pour la sonde correspondant à l'allèle mineur (figure 26). Ce complexe est aboli totalement par l'ajout de la même sonde non -marquée. On observe peu ou pas de liaison de facteurs nucléaires à la sonde correspondant à l'allèle majeur. Ces résultats confirment la prédiction *in silico*. Des résultats similaires ont été observés pour les rSNPs Ku70-489 et Ku80-1669 (tableau XVI). L'expérience de gel de retardation avec la sonde de l'allèle majeur du gène Rad50-1347 a montré une bande spécifique de complexes ADN-protéines, tandis que la sonde de l'allèle mineur ne montre pas de liaison à des facteurs nucléaires. Par contre, ces résultats sont contraires à la prédiction *in silico* (tableau XVI). Finalement, les expériences pour le gène MRE11A-1226 ont montré la présence de bandes spécifiques avec les deux sondes testées. Toutefois, l'intensité de la bande avec la sonde de l'allèle majeur est supérieure à l'intensité de la bande avec la sonde de l'allèle mineur. Ces résultats ne concordent pas avec la prédiction. Les autres sites testés n'ont montré aucun résultat concluant (tableau XVI).

Tableau XV. Sites testés par gel de retardement et prédiction *in silico* de la présence de sites de liaison de facteurs de transcription

Gène	rSNP	Sites de liaison mutants
DCLRE1C	-1571C/G	Se-Cys tRNA gen transcription activating factor=0.803 X-box binding protein RFX1=0.936
DCLRE1C	-1892G/C	Sox-5=0.979
KU70	-489G/A	Runt-factor AML-1=1.000
KU70	-1687C/T	Activator protein 1=0.963 Ribonucleoprotein associated zinc finger protein MOK-2(human)=0.982 Sterol regulatory element binding protein=0.851
KU80	-205A/G	NF-kappa B=0.827
KU80	-207A/C	E2F, involved in cell cycle regulation, interacts with Rb p107 protein=0.804
KU80	-297C/T	AREB6 (Atp 1a1 regulatory element binding factor 6)=0.981
KU80	-1669T/A	Complex of Lmo2 bound to Tal-1, E2A proteins, and GATA-1, half-site 2=0.989 Serum response factor=0.687
MRE11A	-1226A/C	Heat shock factor 1=0.966 HMGI(Y) high-mobility-group protein 1(Y), architectural transcription factor organizing the framework of a nuclear protein-DNA transcriptional complex=0.932 Nuclear factor of activated T-cells=0.984 Octamer-binding factor 1=0.823
NBS1	-692C/T	Octamer-binding factor 1, POU-specific domain=0.864
RAD50	-52A/G	Zinc finger with interaction domain=0.884
RAD50	-1347/insT	myocyte enhancer factor=0.967
XRCC2	-1164G/C	Barbiturate-inducible element=0.892 HNF-3/Fkh homolog 2=0.938
XRCC4	-1864T/C	Activator protein 4=0.976

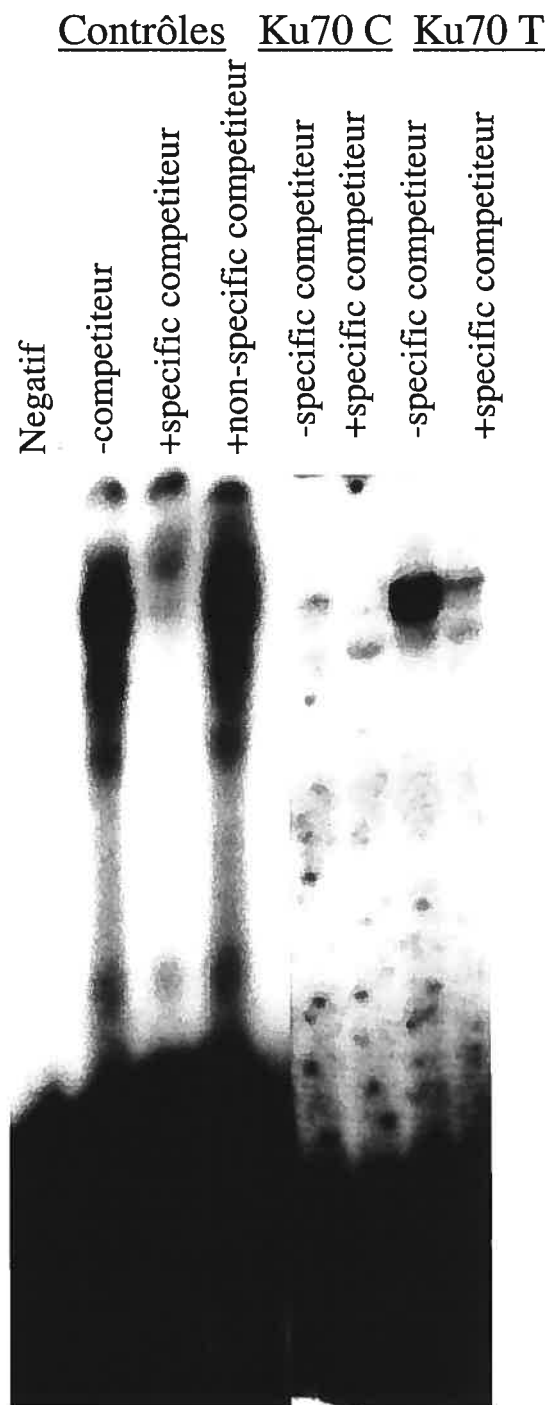


Figure 26. Résultats de gel de retardation pour le polymorphisme à la position -1687 du gène Ku70. Ligne 1=contrôle négatif, ligne 2=contrôle positif, ligne 3= contrôle avec compétiteur spécifique, ligne 4=contrôle avec compétiteur non-spécifique, ligne 5=sonde allèle majeur, ligne 6=sonde allèle majeur plus compétiteur spécifique, ligne 7=sonde allèle mineur, ligne 8=sonde allèle mineur plus compétiteur spécifique.

Tableau XVI. Résultats des expériences de gel de retardement

Gène	rSNP	Résultat	Variation entre l'allèle majeur et l'allèle mineur	Comparaison avec la prédiction
DCLRE1C	-1571C/G	-	non	-
DCLRE1C	-1892G/C	-	non	-
KU70	-489G/A	+	oui	pareil
KU70	-1687C/T	+	oui	pareil
KU80	-205A/G	-	non	-
KU80	-207A/C	-	non	-
KU80	-297C/T	-	non	-
KU80	-1669T/A	+	oui	pareil
MRE11A	-1226A/C	+	oui	différent
NBS1	-692C/T	-	non	-
RAD50	-52A/G	-	non	-
RAD50	-1347/insT	+	oui	contraire
XRCC2	-1164G/C	-	non	-
XRCC4	-1864T/C	-	non	-

+ : présence d'un résultat concluant, - : absence de résultat concluant.

## 5 Étude d'association pour le gène Ku70

Nous avons décidé d'étudier l'impact de certains polymorphismes du gène Ku70 sur la susceptibilité à développer la LLA à l'aide d'une étude d'association. Par les expériences d'EMSA, nous avons identifié des rSNPs ayant un impact sur la liaison de facteurs de transcription. Nous avons trouvé important de poursuivre l'analyse de ce gène, puisque certains polymorphismes semblent avoir un impact fonctionnel.

### 5.1 Détermination de la fréquence des variants du gène Ku70

Nous avons déterminé la fréquence de 5 polymorphismes présents dans le gène Ku70 : -1469C/T, -1296C/G, -731G/A, -533C/G et -489G/A chez un groupe d'enfants atteints de la LLA (n=250) et chez un groupe d'individus sains (n=279), les deux groupes sont formés de Canadiens français résidant dans la province de Québec ainsi que chez un groupe d'individus de représentation mondiale (tableau VI). Nous avons choisi d'étudier tous les polymorphismes identifiés par dHPLC présents chez les Canadiens français. La distribution des génotypes chez les patients et chez les témoins est donnée dans le tableau XVII. Pour quelques individus, les génotypes n'ont pu être déterminés, ce qui explique la variation du nombre d'échantillons listés dans le tableau. La distribution des génotypes des positions -1296, -731, -533 et -489 suit la loi de Hardy-Weinberg. Par contre, la distribution des génotypes du polymorphisme -1469 ne la suit pas dans le groupe témoin ( $p=0.00002$ ). Toutefois, on ne peut accorder une grande importance à ce résultat étant donné que le nombre d'individus dans chacune des classes n'est pas supérieur à 5 par conséquent la valeur statistique de ce test est grandement diminué. Aucune différence de distribution des génotypes entre les deux groupes n'est observée pour tous les polymorphismes étudiés.



## 5.2 Détermination des allèles ancestraux

Nous avons génotypé un chimpanzé, un orang-outang et un gorille pour déterminer les allèles ancestraux. Dans tous les cas, l'allèle majeur présent chez l'humain correspond à l'allèle ancestral qui est monomorphique chez les primates (tableau XVIII).

## 5.3 Analyse d'haplotypes

Par la suite, à l'aide des données de génotypage, nous avons construit les haplotypes en utilisant le programme Phase (Stephens, M. *et al.* 2001). Ce programme se base sur la fréquence allélique de chaque polymorphisme pour construire les haplotypes les plus probables. Le programme a permis la reconstruction de 9 haplotypes différents (voir figure 27). Les haplotypes 6, 7 et 8 sont retrouvés uniquement chez les patients, l'haplotype 9 est retrouvé seulement chez le groupe témoin et l'haplotype 5 est retrouvé uniquement chez les individus d'origine canadienne-française. Ensuite, nous avons déterminé la fréquence de ces 9 haplotypes chez les différents groupes (tableau XIX et XX). Une fois les fréquences calculées pour chaque haplotype dans le groupe cas et témoin, nous avons pu reconstruire le réseau selon lequel les différents haplotypes ont divergé de l'haplotype ancestral (figure 28 et 29) à l'aide du programme Network 3.1.1.1 ([http://www.fluxus\\_engineering.com](http://www.fluxus_engineering.com)). On peut remarquer que le réseau pour le groupe cas est plus complexe que le réseau pour le groupe témoin puisqu'il y a un plus grand nombre d'haplotypes. L'origine des haplotypes 6, 7 et 8 pourrait être expliquée par des événements de recombinaison entre les haplotypes présents dans le groupe témoin (figure 30). Les différentes possibilités de recombinaison ont été prédites en utilisant le programme Recomb (El-Mabrouk, N. et Labuda, D. 2002). La distribution des génotypes chez les patients et chez les témoins est donnée dans le tableau XXI. Aucune différence de distribution n'est observée entre les deux groupes.

Tableau XVII. Distribution des génotypes chez les cas et les témoins

Génotypes	patients LLA		Témoins		Chi carré p
	n	%	n	%	
<b>-1469</b>					
CC	197	0.89	245	0.91	0.056
CT	25	0.11	19	0.07	0.059
TT	0	0.00	4	0.01	0.191
<b>Total</b>	<b>222</b>		<b>268</b>		
<b>-1296</b>					
CC	80	0.37	95	0.36	0.973
CG	103	0.47	127	0.48	0.979
GG	35	0.16	41	0.16	0.992
<b>Total</b>	<b>218</b>		<b>263</b>		
<b>-731</b>					
GG	132	0.64	163	0.61	0.461
AG	67	0.32	89	0.33	0.506
AA	8	0.04	17	0.06	0.511
<b>Total</b>	<b>207</b>		<b>269</b>		
<b>-533</b>					
CC	215	1.00	268	0.99	0.206
CG	0	0.00	2	0.01	0.207
<b>Total</b>	<b>215</b>		<b>270</b>		
<b>-489</b>					
GG	199	0.98	265	0.99	0.447
GA	5	0.02	4	0.01	0.452
<b>Total</b>	<b>204</b>		<b>269</b>		

Tableau XVIII. Allèles ancestraux

Position	Allèle majeur	Allèle ancestral
-1469	C	C
-1296	C	C
-731	G	G
-533	C	C
-489	G	G

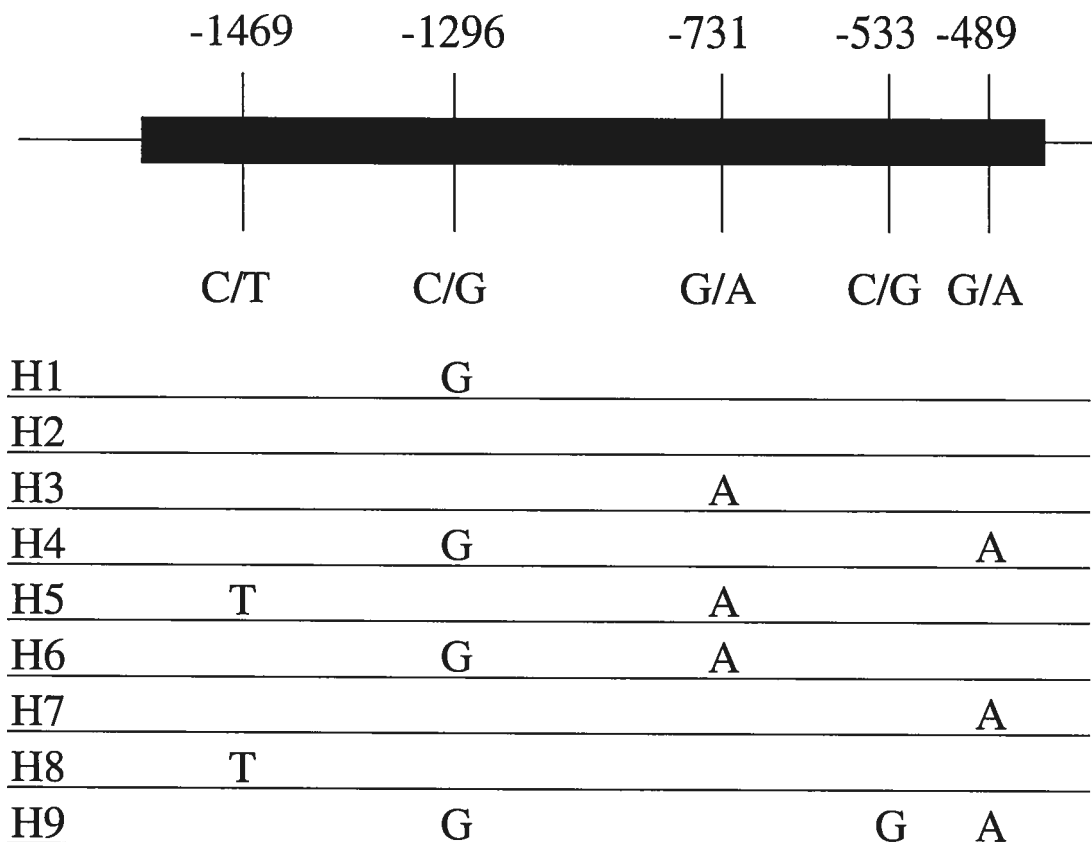


Figure 27. Représentation des 9 haplotypes reconstruit par Phase. L'haplotype 2 est l'haplotype ancestral.

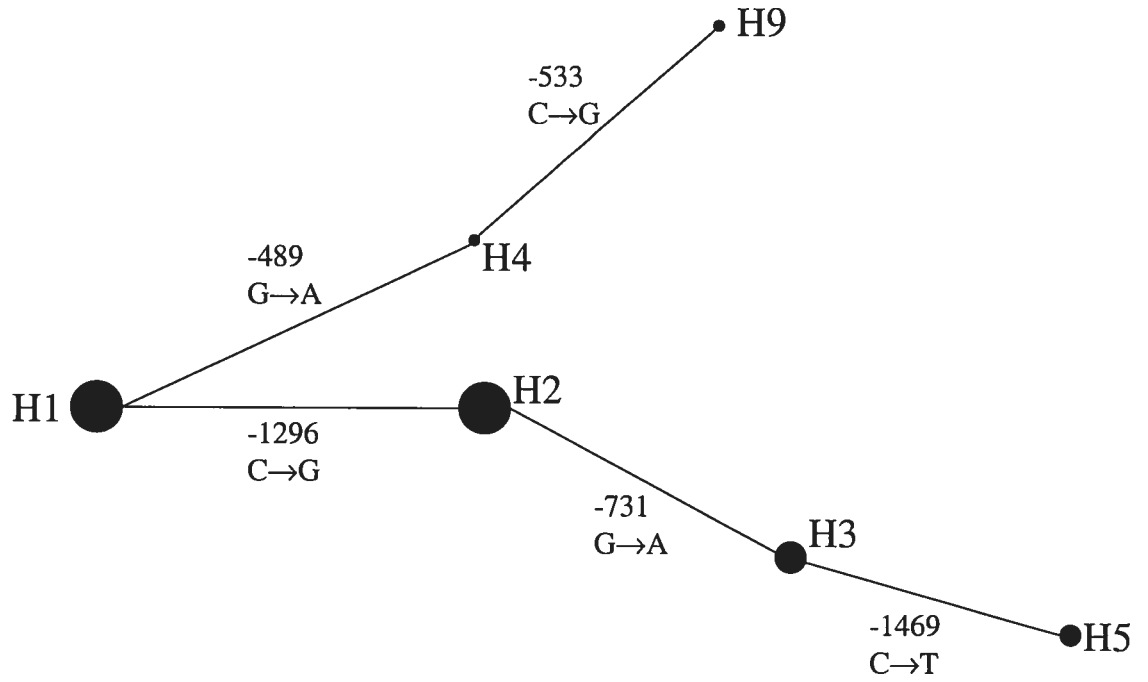


Figure 28. Réseau des haplotypes chez les témoins

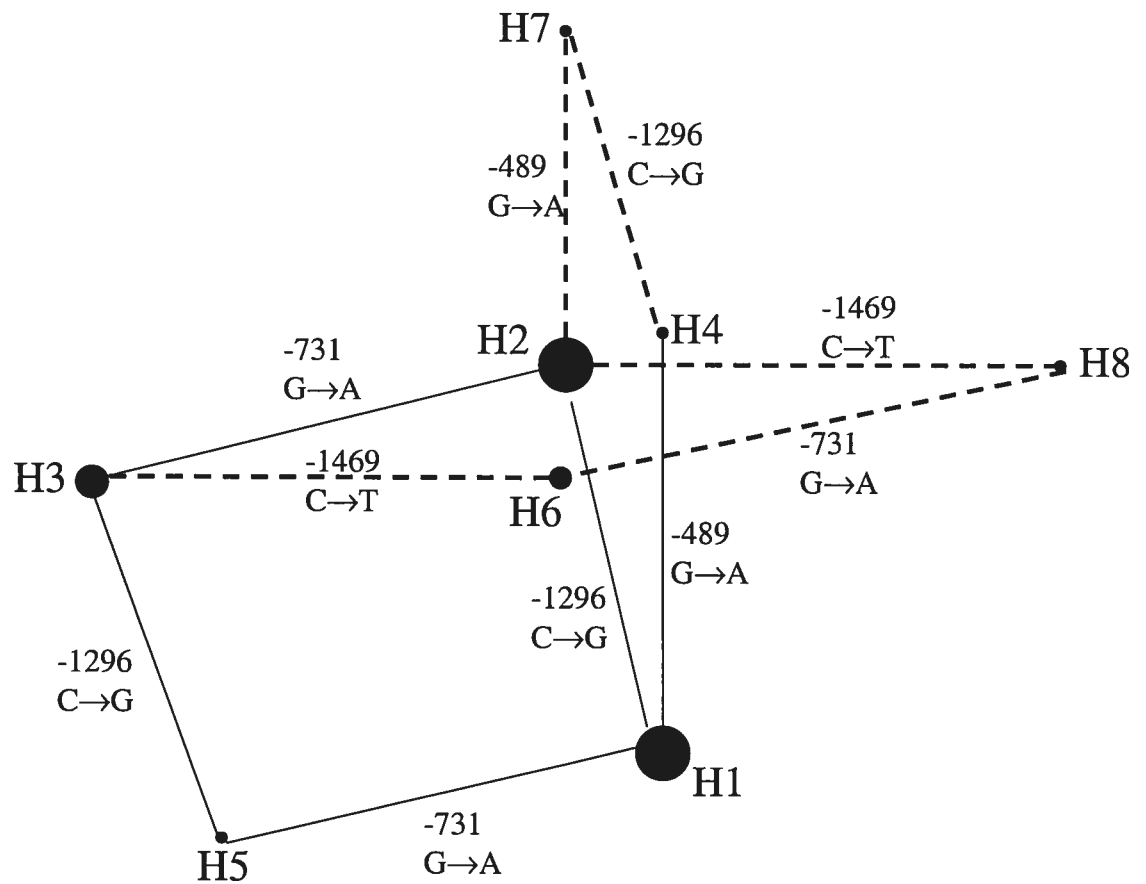


Figure 29. Réseau des haplotypes chez les cas. Les lignes pointillées indiquent que l'origine des haplotypes H6, H7 et H8 ne serait probablement pas expliquée par ces événements, mais plutôt par des événements de recombinaison (voir figure 30).

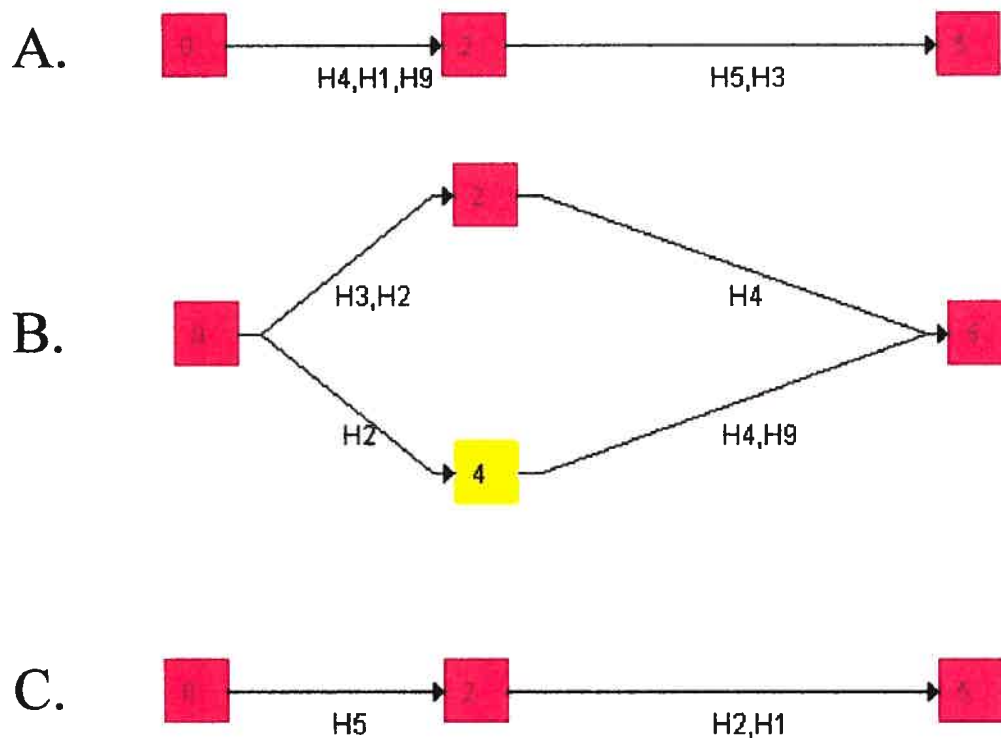


Figure 30. Recombinaisons pouvant expliquer l'apparition des haplotypes H6 (A), H7 (B) et H8 (C). En A, les haplotypes H4, H1 ou H9 peuvent recombiner avec H5 ou H3 pour former l'haplotype H6. La recombinaison se situe entre le 2<sup>e</sup> polymorphisme (-1296) et le 3<sup>e</sup> (-731). En B, les haplotypes H2 ou H3 peuvent recombiner avec H4 entre le 2<sup>e</sup> (-1296) et le 3<sup>e</sup> (-731) polymorphisme pour former l'haplotype H7. Aussi, l'haplotype H2 pourrait recombiner avec H4 ou H9 entre la position 4 (-533) et 5 (-489). En C, l'haplotype H5 peut recombiner avec H2 ou H1 pour former H8.

Tableau XIX. Fréquence des haplotypes chez les patients et les témoins.

Haplotypes	patients LLA (n=194)		Témoins (n=281)	
	n	%	n	%
H1	92	23.71	152	27.05
H2	166	42.78	245	43.59
H3	114	29.38	143	25.44
H4	4	1.03	1	0.18
H5	7	1.80	19	3.38
H6	3	0.77	0	0.00
H7	1	0.26	0	0.00
H8	1	0.26	0	0.00
H9	0	0.00	2	0.36
<b>Total</b>	<b>388</b>		<b>562</b>	

Tableau XX. Fréquence des haplotypes chez les individus d'origines différentes

Haplotypes	Populations mondiales	
	n	%
H1	15	30
H2	27	54
H3	7	14
H4	1	2
H5	0	0
H6	0	0
H7	0	0
H8	0	0
H9	0	0
<b>Total</b>	<b>50</b>	

Tableau XXI. Distribution des génotypes chez les cas et les témoins

Génotypes	patients LLA (n=194)		Témoins (n=281)		OR**	IC 95%	p			
	n	%	n	%						
H1;H1	26	8.84	40	14.87	0.89	0.52-1.51	0.76			
H1;H2	56	19.05	80	29.74						
H1;H3	24	8.16	39	14.50						
H1;H4	3	1.02	1	0.37						
H1;H5	12	4.08	7	2.60						
H1;H6	2	0.68	0	0.00						
Hétérozygotes H1	97	32.99	127	47.21	1.12	0.77-1.62	0.62			
Somme de H1;*	123	41.84	167	62.08	1.06	0.72-1.55	0.85			
H2;H2	35	11.90	37	13.75	1.36	0.82-2.24	0.29			
H2;H3	19	6.46	33	12.27						
H2;H4	1	0.34	3	1.12						
H2;H5	6	2.04	10	3.72						
H2;H7	1	0.34	0	0.00						
H2;H8	1	0.34	0	0.00						
H2;H9	0	0.00	1	0.37						
Hétérozygotes H2	84	28.57	127	47.21				0.85	0.89-1.24	0.46
Somme de H2;*	119	40.48	164	60.97				1.02	0.69-1.48	0.93
H3;H3	4	1.36	9	3.35	0.61	0.18-2.00	0.59			
H3;H5	4	1.36	4	1.49						
H3;H9	0	0.00	1	0.37						
Hétérozygotes H3	47	15.99	77	28.62	0.8	0.52-1.21	0.3			
Somme H3;*	51	17.35	86	31.97	0.76	0.50-1.14	0.22			
Hétérozygotes H4	4	1.36	4	1.49	1.39	0.34-5.65	0.91			
H5;H5	0	0.00	4	1.49	0.15	0.001-2.84	0.23			
Hétérozygotes H5	22	7.48	21	7.81	1.25	0.68-2.29	0.57			
Somme H5;*	22	7.48	25	9.29	1.49	0.79-2.79	0.28			
<b>Total</b>	<b>194</b>	<b>100</b>	<b>269</b>	<b>100</b>						

\*Tous les autres haplotypes

\*\*OR=calculé à partir du ratio du nombre de génotypes d'intérêt versus tous les autres génotypes chez les patients LLA et du ratio chez les individus contrôles



**CHAPITRE IV**

**DISCUSSION**

Dans ce projet, nous avons voulu évaluer l'impact de polymorphismes dans les gènes de réparation de l'ADN. L'altération de ces systèmes de réparation pourrait avoir un impact important au niveau du fardeau mutationnel du génome d'un individu qui expliquerait, du moins en partie, la susceptibilité inter-individuelle à certaines maladies comme le cancer.

À cause de la vocation pédiatrique de l'institution hôte, nous avons privilégié l'étude de la leucémie lymphoblastique aiguë (LLA) de l'enfant. Cette maladie offre une opportunité unique d'étudier l'effet de la variation génétique autant au niveau des gènes du métabolisme des carcinogènes qu'au niveau des gènes de réparation de l'ADN sur le risque de développer des cancers pédiatriques (Krajinovic, M. *et al.* 1999). Le jeune âge des patients et une courte période de latence entre l'apparition des mutations initiales et la détection de cellules tumorales facilitent l'identification de facteurs de risque comparativement au cancer chez l'adulte où plusieurs facteurs entrent en jeu à cause de la longue période de latence. Nous avons concentré nos efforts sur les Canadiens français qui sont caractérisés par un effet fondateur. Cette population fut fondée par quelques milliers d'immigrants de France au 17<sup>ème</sup> siècle (Labuda, D. *et al.* 1997 ; Bouchard, G. et Braekeleer, M. 1990). Après la conquête anglaise au milieu du 18<sup>ème</sup> siècle, la vague d'immigration française a pratiquement cessé. Cependant, grâce à une grande croissance démographique, cette population a pris de l'expansion jusqu'à atteindre 10 millions d'habitants au Québec et ailleurs en Amérique du Nord. Malgré l'effet fondateur, les fréquences alléliques testées chez les témoins concordent avec celles rapportées pour d'autres populations d'origine européenne (Winsey, S.L. *et al.* 2000 ; Matullo, G. *et al.* 2001), ce qui suggère que l'implication de nos résultats peut être étendue à ces groupes également. Un des avantages d'étudier les Canadiens français est la faible hétérogénéité génétique qui les caractérisent. Ce point est souvent problématique dans les études d'association. De plus, on observe une similarité socioculturelle qui diminue certains effets confondants comme le style de vie, la diète, etc.

Nous avons déterminé la fréquence des variants alléliques de XRCC1(Arg399Gln), XRCC3(Thr241Met) et hOGG1(Ser326Cys). Nous avons choisi d'étudier ces polymorphismes puisqu'ils se situent au niveau d'acides aminés conservés à travers l'évolution et qu'ils mènent à un changement de la séquence protéique indiquant potentiellement un impact fonctionnel. Nous avons démontré que les allèles majeurs chez l'humain correspondent aux allèles ancestraux, c'est-à-dire les allèles retrouvés chez les primates.

### **XRCC1(Arg399Gln)**

La fréquence allélique du variant XRCC1-399Gln, 0,38 chez les témoins, est similaire à celle rapportée dans d'autres études (entre 0,34 et 0,39) (Matullo, G. *et al.* 2001 ; Winsey, SL. *et al.* 2000 ; Nelson, HH. *et al.* 2002). Nous n'avons pas observé de différences entre la fréquence du variant chez les cas et chez les témoins ni de différences au niveau de la distribution des génotypes. Ces résultats suggèrent que ce polymorphisme ne joue pas de rôle dans la susceptibilité à la LLA. Cependant, de récentes études ont démontré que XRCC1-399Gln était associé avec une diminution de la capacité de réparation (Lunn, R. *et al.* 1999 ; Duell, E. *et al.* 2000), à une élévation d'échange entre les chromatides sœurs (Duell, E. *et al.* 2000 ; Abdel-Rahman, S. *et al.* 2000) et à une augmentation de la durée du cycle cellulaire (Hu, J. *et al.* 2001). De plus, ce même polymorphisme a été associé avec le risque de développer certaines tumeurs solides : cancer du cou et de la tête (Sturgis, E. *et al.* 1999 ; Lee, J. *et al.* 2001), cancer du sein (Duell, E. *et al.* 2001), cancer du poumon (Divine, K. *et al.* 2001), cancer de la vessie (Stern, H. *et al.* 2000), cancer de l'estomac (Shen, H. *et al.* 2000) et le cancer colorectal (Abdel-Rahman, S. *et al.* 2000). Par contre, d'autres études montrent des résultats différents : augmentation du risque associée à l'autre allèle (XRCC1-399Arg) (Sturgis, E. *et al.* 1999 ; Lee, J. *et al.* 2001 ; Duell, E. *et al.* 2001 ; Divine, K. *et al.* 2001 ; Shen, H. *et al.* 2000 ; Abdel-Rahman, S. *et al.* 2000), une diminution du risque (Stern, H. *et al.* 2000) ou aucun effet. (Matullo, G. *et al.* 2001 ; Ratnasinghe, D. *et al.* 2001 ; Butkiewicz, D. *et al.* 2001). Ces résultats peuvent s'expliquer par une étiologie différente entre les cancers,

des différences au niveau des designs expérimentaux ou tout simplement à la liaison avec un gène avoisinant.

### **XRCC3(Thr241Met)**

Le variant XRCC3-Met241 est représenté à une fréquence de 0,40 dans le groupe témoin ce qui est comparable aux valeurs publiées antérieurement qui se situent entre 0,40 et 0,43 (Matullo, G. *et al.* 2001 ; Winsey, SL. *et al.* 2000). Nous avons observé une différence dans la distribution des homozygotes Met241 entre les cas (10,47%) et les témoins (15,51%). Cette sous-représentation dans les patients est associée avec une diminution non significative du risque de développer la LLA (OR=0,64 ; IC95%=0,39-1,05 ; p=0,084). Jusqu'à présent, les efforts visant à démontrer le caractère fonctionnel de ce variant se sont avérés vains. (Araujo, FD. *et al.* 2002 ; Rafii, S. *et al.* 2003). D'autres études d'association ont montré que l'allèle Met241 augmentait le risque de cancer de la vessie, du sein et du cancer de la peau (Kuschel, B. *et al.* 2002 ; Winsey, SL. *et al.* 2000 ; Matullo, G. *et al.* 2001). Par conséquent, l'étude d'une plus grande cohorte permettrait de voir si l'effet qui semble être observé se confirme ou bien disparaît étant donné que nos résultats ne sont pas statistiquement significatifs.

### **hOGG1(Ser326Cys)**

Nous avons montré que le variant hOGG1-Cys326 est associé avec une diminution du risque de développer la LLA. Les génotypes Ser326/Cys326 et Cys326/Cys326 sont sous-représentés dans les cas suggérant un effet protecteur chez les porteurs de l'allèle Cys326 (OR = 0,58 ; IC95%=0,41-0,81 ; p=0,002). Ce gène code pour l'enzyme de réparation impliquée dans l'élimination des 8-hydroxyguanines (figure 5). Ces lésions sont très abondantes au niveau de l'ADN suite à une exposition endogène ou exogène au stress oxydant, aux radiations ionisantes ou à des produits carcinogènes (Dizdaroglu, M. 1991 ; Cadet, J. 1994 ; Urios, A. et Blanco, M. 1994). Ces lésions sont très mutagéniques, non pas parce qu'elles bloquent l'élongation du

brin synthétisé lors de la réplication de l'ADN, mais plutôt parce qu'elles sont préférentiellement appariées à une adénine ce qui mène à une accumulation de transversions G : C → T : A. Des mutations somatiques dans le gène hOGG1 ont été identifiées dans plusieurs types de tumeurs dont les tumeurs trouvées dans les poumons, reins, cou et tête (Chevallard, S. *et al.* 1998 ; Blons, H. *et al.* 1999). Ces mutations altèrent partiellement l'activité de hOGG1. Dans le cas du cancer du cou et de la tête, la région chromosomique contenant le gène hOGG1 est souvent délétée. En plus, des études de mutations du gène p53 dans les tumeurs ont montré la présence élevée de transversions G : C → T : A ce qui suggère que la fonction de hOGG1 est abolie par la délétion chromosomique observée (Blons, H. *et al.* 1999). Suite à ces résultats, il est raisonnable de penser qu'une diminution de l'activité du gène hOGG1 peut mener à une augmentation du risque de développer un cancer puisqu'on retrouve une augmentation du fardeau mutationnel qui peut mener à l'inactivation de gènes suppresseurs de tumeurs ou à l'activation d'oncogènes. D'un point de vue fonctionnel, Kohno, T. *et al.* en 1998 ont démontré que l'allèle Cys326 est environ 7 fois moins efficace pour compléter la déficience en réparation d'une souche de *E.coli* lors d'essais fonctionnels *in vitro*. Par contre, Dherin, C. *et al.* en 1999 n'ont pas observé de différence significative dans l'activité de OGG1 lors d'études *in vitro* utilisant la protéine de fusion OGG1-glutathione S-transférase (GST). Une étude récente par Janssen, K. *et al.* en 2001 a montré que l'activité de réparation de OGG1 dans des lymphocytes humains n'est pas dépendante du variant Ser326Cys. Parallèlement aux études fonctionnelles, les études d'association entre ce variant et le risque de cancer sont plus ou moins conséquentes. Sugimura, H. *et al.* en 1999 ont montré que les homozygotes Cys326, dans une population japonaise, ont un risque plus élevé de développer un cancer des poumons. Xing, D. *et al.* en 2003 ont trouvé que les homozygotes Cys326, dans une population chinoise, ont un risque plus élevé de développer le cancer de l'œsophage. Cependant, deux autres études n'ont pas confirmé ces associations (Chevallard, S. *et al.* 1998 ; Wikman, H. *et al.* 2000). Dans la population allemande, Wikman, H. *et al.* en 2000 ont trouvé une proportion plus importante d'homozygotes Ser326 chez un groupe d'individus atteints de cancer du poumon (64,8%) comparativement au groupe témoin (57,1%). Xu, J. *et al.* en 2002

ont montré, dans une population caucasienne, que les homozygotes Ser326 avaient entre 2 et 3 fois plus de risque de développer un cancer de la prostate. De plus, il y a une différence très significative dans la distribution des génotypes entre les différentes races et ethnies. La proportion d'homozygotes Ser326 est la plus élevée chez les Mélanésiens (74,5%), Hongrois (63,7) et Allemands (57,1), plus faible chez les Australiens caucasiens (39,9%), les Japonais (27,7%) et les Micronésiens (25,8%) et encore plus faible chez les Chinois (12% ; Wikman, H. *et al.* 2000 ; Sugimura, H. *et al.* 1999). Dans notre étude, nous avons observé une fréquence de 56,53% dans le groupe des témoins, ce qui est comparable aux résultats publiés chez les Caucasiens. En compilant l'ensemble de ces résultats, il semble que l'allèle Ser326 confère un risque de cancer chez les populations caucasiennes et l'allèle Cys326 confère un risque chez les populations asiatiques. D'autres études d'association et d'autres études fonctionnelles seront nécessaires pour comprendre la complexité des résultats en lien avec ce polymorphisme. La recherche d'autres polymorphismes serait également une chose à envisager afin d'éliminer la possibilité d'un polymorphisme fonctionnel situé ailleurs sur le gène.

### **Gènes de réparation des cassures doubles-brins**

Nous avons également voulu étudier la variation au niveau des gènes de réparation des cassures doubles-brins, puisque ces lésions sont souvent causées par les carcinogènes et mènent à des bris et des réarrangements chromosomiques. Dans des modèles de souris, l'absence d'une des voies de réparation des lésions doubles-brins, mènent à une instabilité génomique causant entre autres des translocations oncogéniques (Ferguson, DO. *et al.* 2001). La majorité des polymorphismes étudiés affectent les régions codantes ainsi que les régions introniques. Une raison justifiant ce biais est la facilité de prédire l'impact de ces polymorphismes en analysant la nature du changement d'acides aminés ou la position des sites d'épissages. Dans le cadre de ce projet, nous avons tourné notre attention sur les régions de régulation situées en 5' des gènes. La présence de polymorphismes dans les régions régulatrices

peut affecter le niveau de transcription du gène soit par le gain ou la perte de sites de liaison de facteurs de transcription.

### **Plate-forme de détection**

Nous avons utilisé le dHPLC comme plate-forme de détection. Nous l'avons choisi parce que cette technique a un pouvoir de détection se situant entre 95 et 100% (Gross, E. 2000 ; Premstaller, A. 2002 ; Dobston-Stonce, C. *et al.* 2000). Pour optimiser nos chances de détecter la présence de SNPs, les produits de PCR ont été analysés à plus d'une température. Dans la même perspective, nous avons choisi de diminuer nos faux-négatifs en étant très conservateur dans l'analyse des chromatogrammes. Nous avons séquencé tous les profils variables identifiés, même quand ceux-ci n'indiquaient pas clairement la présence d'hétérozygotes. Nous avons choisi de diminuer notre efficacité, qui est de 79,6% (tableau XIII), afin d'augmenter nos chances d'identifier la présence de polymorphismes.

Nous avons criblé les régions régulatrices définies arbitrairement comme 2 kb en amont de l'exon 1 de 15 gènes de réparation des lésions doubles-brins. Nous avons ainsi identifié 144 rSNPs dont seulement 10% étaient présents dans les bases de données publiques. Ces résultats correspondent à une densité de 4,6 rSNPs par kb. La densité observée est comparable aux valeurs publiées dans d'autres études ainsi qu'à la densité de cSNPs publiée par Stephens, J.C. *et al.* en 2001 qui est de 4,2. Ces données sont intéressantes puisqu'on s'attendrait à voir une densité de polymorphismes plus élevée dans les régions régulatrices que dans les régions codantes à cause de la présence accrue d'îlots CpG. En effet, la méthylation des îlots CpG est associée avec l'apparition fréquente de mutations qui implique les transitions G→A ou C→T. Par conséquent, comme la densité dans les régions régulatrices n'est pas supérieure à celle dans les régions codantes, une pression sélective s'exerce pour éviter l'apparition de variations dans ces régions. De plus, nous n'avons pas observé de différence entre le nombre de polymorphismes dans les régions GC riche

(69 rSNPs) et le nombre dans les régions AT riche (75 rSNPs). Nous avons identifié 75 transitions, 53 transversions et 16 insertions/délétions. Il est normal de trouver un plus grand nombre de transitions que de transversions. La pression évolutive favorise la conversion des G et C vers des A ou T afin de relaxer la force d'appariement des bases (Stephens, JC. *et al.* 2001).

L'identification de nouveaux polymorphismes de régulation permet d'identifier de nouveaux candidats pour des études génétiques et pharmacogénétiques. Ces rSNPs pourront être utilisés pour des études d'association afin de quantifier leur impact sur le développement de différents cancers dans différentes populations. Par contre, tous ces rSNPs n'auront pas nécessairement un effet fonctionnel. Pour accorder une importance plus grande à un rSNP plutôt qu'un autre, plusieurs approches furent utilisées : les prédictions *in silico* sur le gain/perte de sites de liaison de facteurs de transcription, la génomique comparative et la confirmation *in vitro* de certaines des prédictions *in silico*. De plus, des études avec des gènes rapporteurs permettraient d'évaluer la différence d'expression d'un promoteur comparativement à un autre. Le but ultime serait la validation *in vivo*, comme un modèle de souris.

### **Prédictions *in silico***

La création ou la perte d'un site de liaison de facteurs de transcription pourrait avoir des répercussions sur la transcription d'un gène. Nous avons donc utilisé une méthode *in silico* pour prédire l'effet des polymorphismes sur le gain et/ou la perte de sites de liaison de facteurs de transcription en se basant sur les séquences consensus (Jegga, AG. *et al.* 2002). Nous avons déterminé que parmi les 144 rSNPs, 35 font apparaître des sites de liaison de facteurs de transcription, 31 font disparaître des sites et 39 font apparaître et disparaître un ou plusieurs sites. Un score, basé sur la comparaison de la séquence analysée par rapport aux séquences consensus des facteurs de transcription, est attribué afin d'évaluer la prédiction. Par contre, la confirmation expérimentale de ces prédictions est nécessaire. Les bases de données



de sites consensus de liaison de facteurs de transcription sont limitées par le nombre de facteurs de transcription connus. Par contre, elles permettent de faire une sélection parmi les nombreux polymorphismes identifiés afin de poursuivre leur caractérisation fonctionnelle.

### **Génomique comparative**

La génomique comparative permet d'identifier des régions conservées au cours de l'évolution. Les séquences ayant un rôle dans la régulation de l'expression de gènes tendent à être conservées d'une espèce à l'autre (Pennacchio, LA. et Rubin, EM. 2001). Les polymorphismes qui se trouvent dans ces régions conservées ont par conséquent plus de chance d'avoir un impact fonctionnel. À cause de la disponibilité des séquences, la comparaison s'est limitée à l'homme et la souris. Par contre, la comparaison entre plusieurs espèces pourrait diminuer le nombre de régions conservées identifiées et donc augmenter les chances de trouver des régions plus importantes qui sont conservées à travers l'évolution de plusieurs espèces. Nous avons identifié 15% de régions conservées entre l'humain et la souris au niveau des 15 gènes analysés. Par contre, ce pourcentage varie de 0 à 44% selon le gène. Nous avons identifié 21 rSNPs dans des régions conservées ce qui représente 15% des rSNPs. Parmi ceux-ci, 6 sont associés à un gain de sites de liaison de facteurs de transcription, 2 sont associés à une perte de sites, 5 à un gain et une perte et 8 n'ont aucun effet de prédit. La fréquence de rSNPs dans les régions conservées est la même que pour les régions non conservées. Par contre, 62% de ces rSNPs modifient des sites de liaison de facteurs de transcription comparativement à 75% des rSNPs situés à l'extérieur de ces régions. Ces résultats suggèrent que ces régions se trouvent sous une pression sélective pour éviter la variation génétique qui pourrait avoir un impact fonctionnel au niveau de la régulation génique.

### **Analyses *in vitro***

Comme beaucoup de rSNPs modifient potentiellement des sites de liaison, nous avons testé *in vitro* 14 polymorphismes menant à un gain potentiel d'un tel site. Nous avons utilisé la technique du gel de retardement. Cette technique est couramment utilisée pour montrer la liaison de facteurs de transcription et également elle sert de point de départ pour l'identification de ces facteurs (Della Ragione, F. *et al.* 2003). Nous avons ainsi confirmé que l'allèle mineur Ku70-1687 est capable de lier des protéines dans un extrait de cellules HeLa, alors que l'allèle majeur en est incapable. Cette position est conservée entre l'humain et la souris, ce qui suggère que l'apparition de sites dans une région importante pourrait avoir un impact sur la transcription du gène. Pour connaître la nature des protéines qui se lient à l'allèle mineur, on pourrait faire des expériences de gel de retardement en utilisant des sondes compétitrices contenant le site consensus des différents facteurs prédits. Il serait aussi possible d'effectuer des expériences de supershift avec des anticorps spécifiques à chaque facteur de transcription prédit dont la liaison à l'ADN peut être compétitionnée par une sonde consensus. Il serait tentant de dire que l'apparition de sites mène à une augmentation de la transcription, mais des études de gène rapporteur sont nécessaires pour déterminer si le polymorphisme a un impact sur la transcription. Les résultats obtenus pour le polymorphisme à la position -489 du gène Ku70 confirme la prédiction *in silico*. Par contre, pour confirmer que c'est bien le facteur AML-1 qui se lie à la sonde, des expériences de supershift avec un anticorps sont nécessaires. Pour le gène Ku80-1669, les mêmes conclusions peuvent être tirées. La présence d'un complexe ADN-protéines détecté avec la sonde de l'allèle majeur du gène Rad50-1347 n'était pas attendue. La liaison d'aucun facteur n'était prédite pour cette sonde. Au contraire, on s'attendait à voir une bande avec la sonde de l'allèle mineur et aucune bande avec la sonde de l'allèle majeur. Il est certain que les séquences consensus ne sont pas connues pour tous les facteurs de transcription, ni l'existence de l'ensemble des facteurs de transcription. Si on observe une différence du niveau de la transcription entre les deux allèles, il serait intéressant de déterminer quelle protéine se lie à l'allèle majeur. Pour confirmer la nature de cette protéine, des

expériences de supershift sont impossibles puisqu'on ne connaît pas la protéine. Finalement, pour le gène MRE11A-1226, les résultats obtenus diffèrent des prédictions. Nous observons des bandes spécifiques avec les deux sondes. Toutefois l'intensité n'est pas la même. Des études de gène rapporteur seraient nécessaires pour évaluer l'impact de cette différence sur l'expression et la transcription. Pour les 9 autres sites testés, aucun résultat concluant n'a été obtenu. Au moins deux possibilités peuvent expliquer ces résultats. Tout d'abord, nos expériences se basent sur des prédictions informatiques, donc celles-ci peuvent être erronées. Deuxièmement, plusieurs facteurs entrent en jeu pour la réussite d'une expérience de gel de retardement. Les expériences devront être refaites en modifiant différents paramètres. Tout d'abord, différentes concentrations de sel lors de la préparation d'extraits nucléaires peuvent être utilisées. De plus, la concentration de sel dans le tampon de liaison, la nature du compétiteur, les tampons et la température d'électrophorèse peuvent varier.

### **Études d'association**

Nos résultats suggèrent que des polymorphismes au niveau des gènes Ku70 et Ku80 moduleraient la transcription de ces gènes par l'apparition de sites de liaison de facteurs de transcription. Une étude a montré par des expériences *in vitro* que des cellules de leucémie myéloïde ont une activité de réparation de la voie NHEJ supérieure à des cellules normales (Gaymes, TJ. *et al.* 2002). Par contre, on retrouve plus d'erreurs de réparation se traduisant par une présence accrue de délétions. Ces résultats suggèrent qu'un système de réparation NHEJ trop actif, en particulier lors d'une activité dérégulée de ku70 ou ku80, pourrait expliquer l'instabilité chromosomique observée dans les leucémies. Nous avons donc mené une étude d'association pour les polymorphismes présents dans la région régulatrice du gène Ku70 qui sont détectables dans la population européenne, incluant les Canadiens français. Nous avons ainsi sélectionné 5 polymorphismes : -1469C/T, -1296C/G, -731G/A, -533C/G et -489G/A. Nous n'avons observé aucune différence au niveau de la fréquence de ces variants ni au niveau de la distribution des génotypes entre les cas

et les témoins, tous d'origine canadienne-française. Il a été suggéré que l'utilisation d'haplotypes, plutôt que de SNP seul, serait plus avantageuse pour des études d'association dans le cas de maladies complexes (Lee, S-G. *et al.* 2002; Bonnen, PE. *et al.* 2002). Notre étude de reconstruction des haplotypes a mené à l'identification de neuf haplotypes différents (figure 27). L'haplotype ancestral est l'haplotype 2, qui d'ailleurs est le plus fréquent. En assumant un seul haplotype ancestral et que l'apparition de nouvelles mutations menant à la création de nouveaux haplotypes est le seul fruit des forces de l'évolution, un nombre maximal de  $n+1$  haplotypes est attendu, où  $n$  est le nombre de SNPs. La présence de 9 haplotypes est supérieure au nombre attendu de 6 haplotypes suggérant la présence de recombinaisons, de mutations récurrentes ou de conversions géniques (Bonnen, PE. *et al.* 2002). Les 3 haplotypes excédentaires peuvent être expliqués par des événements de recombinaison (figure 30). En se basant sur la fréquence des haplotypes, les recombinaisons les plus probables sont : entre l'haplotype 1 et 3 pour former l'haplotype 6, entre l'haplotype 2 et 4 pour former l'haplotype 7 et entre l'haplotype 5 et 2 pour former l'haplotype 8. La zone de recombinaison commune à ces trois événements se situe entre la position  $-1296$  et  $-731$ . Il est intéressant de souligner que les haplotypes recombinants H6, H7 et H8 sont présents uniquement chez les patients. Donc, la présence des 3 haplotypes spécifiques au groupe de cas pourrait être due à des événements de recombinaison entre les haplotypes les plus fréquents qui sont également présents dans le groupe témoin. Nous avons ensuite observé la distribution des génotypes des haplotypes. Aucune différence significative n'est observée entre le groupe des cas et le groupe des témoins. Par contre, la présence d'haplotypes spécifiques au groupe de cas est intéressante. Comme leur fréquence est très faible, on ne peut affirmer avec certitude qu'il pourrait y avoir une association avec la leucémie lymphoblastique aiguë et la présence de ces haplotypes. L'analyse d'une plus grande population de cas et de témoins serait nécessaires.

Nous avons également observé la distribution de ces haplotypes chez des individus de différentes origines ethniques (8 asiatiques, 8 moyen-orientaux, 8 amérindiens et 4 européens). Il est intéressant de noter que le polymorphisme  $-1469C/T$  qui mène à

l'apparition des haplotypes 5 et 8 n'est pas retrouvé chez les individus d'origine non canadienne-française. Par contre, on ne peut affirmer que ces haplotypes sont spécifiques aux Canadiens français puisque aucun africain n'a été analysé par génotypage. De plus, un individu hétérozygote -1469C/T d'origine africaine a été identifié par dHPLC et confirmé par séquençage. Pour être certain que ces haplotypes sont seulement présents chez les Africains et chez les Canadiens français, l'étude d'un échantillonnage plus important serait nécessaire.

Suite à cette étude, d'autres études d'association seront menées pour évaluer l'impact des rSNPs que nous avons identifiés dans les gènes de réparation des cassures doubles-brins sur le développement de la LLA. Plusieurs études ont démontré l'importance de l'intégrité de ces gènes. Par exemple, une modification de l'expression du gène ATM peut avoir des répercussions importantes. ATM est responsable de la reconnaissance des lésions doubles-brins et du recrutement des autres facteurs de réparation. Les souris homozygotes nulles pour ATM développent des lymphomes (Barlow, C. *et al.* 1996). Également, on observe un risque très élevé de développer des lymphomes ou des leucémies à cellules T chez les patients atteints d'ataxie télangiectasie (Meyn, MS. 1997). ATM est exprimée dans plusieurs tissus, mais plus spécialement dans le thymus, la rate et le système nerveux central ce qui explique la corrélation des systèmes touchés par une altération du gène ATM. On observe des défauts similaires lors d'une déficience du gène NBS1. Le type de tumeurs le plus fréquent chez les patients atteints de NBS est le lymphome non-hodgkinien, mais les leucémies et plusieurs autres tumeurs solides surviennent également. Les caractéristiques principales des défauts qui surviennent lors d'une déficience de ces gènes incluent l'instabilité chromosomique, l'hypersensibilité aux radiations ionisantes, l'immunodéficience, la susceptibilité aux infections et une augmentation de l'incidence des néoplasies, plus particulièrement les lymphomes. Par conséquent, la présence de rSNPs qui diminueraient l'expression de ces protéines pourrait causer une augmentation du risque de développer des lymphomes ou des leucémies.

La présence de mutations germinales dans les gènes suppresseurs de tumeurs BRCA1 et BRCA2 a été démontrée dans le cancer du sein, des ovaires et quelques cas de cancer de la prostate et du côlon. Ces mutations sont relativement peu fréquentes, mais très pénétrantes. Dans le cas du cancer du sein, seulement 2% des cas sont causés par ce type de mutation dans les gènes BRCA1 et BRCA2 et moins de 20% des cas familiaux sont expliqués par la présence de mutations dans ces gènes (Kuschel, B. *et al.* 2002). Par conséquent, le développement de maladies complexes s'explique par la contribution additive de plusieurs variants alléliques qui sont relativement fréquents, mais moins pénétrants. Donc, la présence de polymorphismes causant une diminution de leur expression pourrait être impliquée dans la susceptibilité à d'autres types de cancer puisqu'une augmentation de l'instabilité chromosomique, une diminution de la résistance à divers agents qui endommagent l'ADN, une mauvaise ségrégation des chromosomes, une diminution de la recombinaison homologue et une diminution des foyers de réparation après traitements aux radiations ionisantes sont observées dans des cellules en culture déficientes pour l'un de ces gènes (Thompson, LH. et Schild, D. 2002).

Tous les autres gènes impliqués dans la HRR, Rad51, Rad52, Rad54, Rad54B, XRCC2, XRCC3, Rad51B, Rad51C et Rad51D sont également impliqués dans la susceptibilité au cancer. Des mutations dans plusieurs de ces gènes ont été retrouvées dans plusieurs lignées cellulaires tumorales (Thomson, LH. et Schild, D. 2002). De plus, des translocations impliquant le gène RAD51B ont été trouvées dans les léiomyomes utérins (Schoenmakers, EF. *et al.* 1999 ; Takahashi, T. *et al.* 2001). Des polymorphismes dans les régions codantes du gène XRCC3 ont été associés avec une augmentation du risque de mélanome, de cancer de la vessie et récemment le cancer du sein (Winsey, SL. *et al.* 2000 ; Matullo, G. *et al.* 2001 ; Kuschel, B. *et al.* 2002). Également 2 groupes ont montré un lien avec la présence de polymorphismes dans le gène XRCC2 et une légère augmentation du risque de cancer du sein (Kuschel, B. *et al.* 2002 ; Rafii, S. *et al.* 2003).

L'implication de la voie de HRR dans le maintien de la stabilité chromosomique a été très bien démontrée. Par contre, la voie NHEJ semble jouer un rôle tout aussi important. Des défauts dans cette voie peuvent mener à une augmentation de transformations néoplasiques. Les lymphomes à cellules T et une augmentation de la transformation des fibroblastes ont été rapportés dans le contexte d'une déficience du gène Ku70 (Ferguson, DO. *et al.* 2000). En plus, une mutation dans le gène Lig4 a été impliquée dans la leucémie chez un patient décédé suite à une sensibilité aux radiations (Riballo, E. *et al.* 1999). L'introduction d'une déficience du gène PRKDC ou XRCC4 en plus d'une déficience en p53 mènent à l'apparition de lymphomes à cellules pro-B (Vanasse, GJ. *et al.* 1999 ; Gao, Y. *et al.* 2000). Ces lymphomes montrent la présence de translocations impliquant le locus de la chaîne lourde des Ig et le locus c-myc. Une autre étude a montré qu'une déficience pour le gène Lig4 ou Ku70 mène à une instabilité chromosomique, même en absence d'agents exogènes endommageant l'ADN, résultant en l'accumulation de translocations (Ferguson, DA. *et al.* 2000 ).

Il a été clairement démontré que les translocations sont impliquées dans plusieurs malignités, les translocations réciproques prédominent dans les tumeurs lymphoïdes et les translocations non-réciproques prédominent dans les tumeurs solides (Hilgenfeld, E. *et al.* 1999 ; Atkin, NB. 1986). Par conséquent, la présence de polymorphismes régulateurs dans les gènes impliqués autant dans la réparation HRR que dans la NHEJ pourrait être impliquée dans l'augmentation ou la diminution de l'instabilité génomique qui mène au développement du cancer.

## CONCLUSION

En conclusion, nous avons montré que la présence de l'allèle Cys-326 du gène hOGG1 est associée avec une diminution de 42% du risque de développer la LLA. Par contre, les polymorphismes XRCC1Arg399Gln et XRCC3Thr241Met ne sont pas associés avec la susceptibilité à la LLA. L'analyse des régions régulatrices des gènes de réparation des cassures doubles-brins de l'ADN a permis l'identification de 144 rSNPs. Nous avons montré que parmi ceux-ci 73% modifient potentiellement le nombre de sites de liaison de facteurs de transcription. Nous avons analysé à l'aide d'expériences de gel de retardation, 14 rSNPs qui menaient potentiellement à l'apparition de nouveaux sites de liaison. Cette analyse a montré une différence entre les allèles majeurs et les allèles mineurs dans 5 des cas : Ku70-489G/A, Ku70-1687C/T, Ku80-1669T/A, MRE11A-1226A/C et Rad50-1347/insT. Par la suite, l'étude d'association entre 5 variants du gène Ku70 (-1469C/T; -1296C/G; -731G/A; -533C/G et -489G/A) n'a montré aucune association avec la susceptibilité à la leucémie lymphoblastique aiguë. Par contre, nous avons identifié 3 haplotypes qui se retrouvent uniquement chez les patients.

Ce projet a permis d'identifier une liste de nouveaux rSNPs qui pourront être utilisés dans de nombreuses études. L'impact fonctionnel reste à être déterminé dans la majorité des cas, mais un point de départ est offert pour permettre d'expliquer l'étiologie de nombreux cancers.



**BIBLIOGRAPHIE**

- 1- Aaltonen LA, Peltomaki P et Sistonen P *et al.* Clues to the pathogenesis of familial colorectal cancer. *Science*. 1993; 260: 812-816.
- 2- Abdel-Rahman S, El-Zein RA. The 399Gln polymorphism in the DNA repair gene *XRCC1* modulates the genotoxic response induced in human lymphocytes by the tobacco-specific nitrosamine NNK. *Cancer Lett*. 2000; 159: 63-71.
- 3- Abdel-Rahman S, Soliman AS, Bondy ML, Omar S, El-Badawy SA., Khaled HM., Seifeldin IA., Levin B. Inheritance of the 194Trp and the 399Gln variant alleles of the DNA repair gene *XRCC1* are associated with increased risk of early-onset colorectal carcinoma in Egypt. *Cancer Lett*. 2000; 159: 79-86.
- 4- Albala JS, Thelen MP, Prange C, Fan W, Christensen M, Thompson LH, Lennon GG. Identification of a novel human RAD51 homolog, RAD51B. *Genomics*. 1997; 46 (3): 476-479.
- 5- Aquilina G, Bignami M. Mismatch repair in correction of replication errors and processing of DNA damage. *J. of cellular physiology*. 2001; 187: 145-154.
- 6- Araujo FD, Pierce AJ, Stark JM. et Jasin M. Variant XRCC3 implicated in cancer is functional in homology-directed repair of double-strand breaks. *Oncogene*. 2002; 21: 4176-4180.
- 7- Atkin NB. Lack of reciprocal translocations in carcinomas. *Cancer Genet Cytogenet*. 1986; 21(3): 275-278.
- 8- Baccichet A, Qualman SK, Sinnett D. Allelic loss in childhood acute lymphoblastic leukemia. *Leuk Res*. 1997; 21 (9): 817-823.
- 9- Bakkenist CJ, Kastan MB. DNA damage activates ATM through intermolecular autophosphorylation and dimmer dissociation. *Nature*. 2003; 421: 499-506.

- 10- Barlow C, Hirotsume S, Paylor R, Liyanage M, Eckhaus M, Collins F, Shiloh Y, Crawlet JN, Ried T, Tagle D, Wynshaw-Boris A. ATM-deficient mice: a paradigm of ataxia telangiectasia. *Cell*. 1996; 86: 159-171.
- 11- Bartek J, Lukas J. Damage alert. *Nature*. 2003; 421: 486-488.
- 12- Bartsch H, Nair U, Risch A, Rojas M, Wilkman H, Alexandrov K. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Epidemiol Biomarkers Prev*. 2000; 9: 3-28.
- 13- Barzilai A, Rotman G, Shiloh Y. ATM deficiency and oxidative stress: a new dimension of defective response to DNA damage. *DNA repair*. 2002; 1 (1): 3-25.
- 14- Berwick M, Vinesis P. Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. *J. of the NCI* 2000; 92 (11): 874-897.
- 15- Bishop JM. Molecular themes in oncogenesis. *Cell*. 1991; 64: 235-248.
- 16- Blons H, Radicella JP, Laccourreye O, Brasnu D, Beaune P, Boiteux S, Laurent-Puig P. Frequent allelic loss at chromosome 3p distinct from genetic alterations of the 8-oxoguanine DNA glycosylase 1 gene in head and neck cancer. *Mol Carcinog*. 1999; 26(4):254-260.
- 17- Boland CR. Hereditary nonpolyposis colorectal cancer. The Genetic Basis of Human Cancer. Edited by B Vogelstein, KW Kinzler. New York: McGraw-Hill, 1998: 333-346.
- 18- Bonnen PE, Wang PJ, Kimmel M, Chakraborty R, Nelson DL. Haplotype and linkage disequilibrium Architecture for Human cancer-associated genes. *Genome Research*. 2002; 12: 1846-1853.
- 19- Bos JL, Toksoz D, Marshall CJ *et al*. Prevalence of *ras* gene mutations in human colorectal cancers. *Nature*. 1987 ; 327: 293-297.
- 20- Bouchard G, De Braekeleer M. eds. Histoire d'un génome, *Presses de l'Université du Québec*. 1990.
- 21- Boulton SJ et Jackson SP. Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO J*. 1998; 17: 1819-1828.

- 22- Burna S, Chen BP, Murphy M, Kurimasa A, Chen DJ. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J. Bio. Mol.* 2001; 276 (45): 42462-42467.
- 23- Butkiewicz D, Rusin M, Enewold L, Shields PG, Chorazy M, Harris CC. Genetic polymorphisms in DNA repair genes and risk of lung cancer. *Carcinogenesis.* 2000; 22: 593-597.
- 24- Cadet J. DNA damage caused by oxidation, deamination, ultraviolet radiation and photoexcited psoralens. Dans: Hemminki K, Dipple A, Shuker DEG, Kadlubar F, Segerback, Bartsch H (eds.), DNA Adducts: Identification and Biological Significance, IARC Scientific Publications. 1994; 125: 245-276.
- 25- Calson P, Delteil C, Frit P, Drouet J, Salles B. Coordinated assembly of Ku and p460 subunits of the DNA-dependent protein kinase on DNA ends is necessary for XRCC4-ligase IV recruitment. *J. Mol. Biol.* 2003; 326: 93-103.
- 26- Carey M et Smale ST. Transcriptional regulation in eukaryotes, cold spring harbor laboratory press, 2000.
- 27- Cartwright R, Dunn AM, Simpson PJ, Tambini CE, Thacker J. Isolation of novel human and mouse genes of the recA/RAD51 recombination-repair gene family. *Nucleic Acids Res.* 1998; 26 (7): 1653-1659.
- 28- Cartwright R, Tambini CE, Simpson PJ, Thacker J. The XRCC2 DNA repair gene from human and mouse encodes a novel member of the recA/RAD51 family. *Nucleic Acids Res.* 1998; 26 (13): 3084-3089.
- 29- Chevillard S, Radicella JP, Levalois C, Lebeau J, Poupon MF, Oudard S, Dutrillaux B, Boiteux S. Mutations in OGG1 a gene involved in the repair of oxidative DNA damage, are found in human lung and kidney tumours. *Oncogene.* 1998; 16: 3083-3086.
- 30- Couch FJ, DeShano ML, Blackwood MA. *BRCA1* mutations in women attending clinics that evaluate the risk of breast cancer. *N Engl J Med.* 1997; 336: 1409-1415.
- 31- Croce CM. Role of chromosome translocations in human neoplasia. *Cell.* 1987; 49: 155-156.

- 32-Cunningham JM, Christensen ER, Tester DJ. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res* 1998; 58: 3455-3460.
- 33-D'amours D, Jackson SP. The MRE11 complex: at the crossroads of DNA repair and checkpoint signaling. *Nature reviews molecular cell biology*. 2002; 3: 317-327.
- 34-Dalla Favera R, Bregni M, Erikson J. Assignment of the *c-myc* oncogene to the region of chromosome 8 which is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci U S A*. 1982; 79: 7824-7827.
- 35-De Jager M, Dronkert MLG, Modesti M, Beerens CEMT, Kanaar R, Van Gent DC. DNA-binding and strand-annealing activities of human Mre11: implications for its roles in DNA double-strand break repair pathways. *Nucleic Acids Research*. 2001; 29 (6): 1317-1325.
- 36-Della Ragione F, Cucciolla V, Criniti V, Indaco S, Borriello A, Zappia V. p21Cip1 gene expression is modulated by Egr1: a novel regulatory mechanism involved in the resveratrol antiproliferative effect. *J Biol Chem* 2003.
- 37-Dherin C, Radicella JP, Dizdaroglu M, Boiteux S. Excision of oxidatively damaged DNA bases by the human hOgg1 protein and the polymorphic  $\alpha$ -hOgg1(Ser326Cys) protein which is frequently found in human populations. *Nucleic Acids Res*. 1999; 27: 4001-4007.
- 38-Divine K, Gilliland FD, Crowell RE, Stidley CA, Bocklage TJ, Cook DL, Belinsky SA. The XRCC1 glutamine allele is a risk factor for adenocarcinoma of the lung. *Mutat. Res*. 2001; 461: 273-278.
- 39-Dizdaroglu M. Chemical determination of free radical-induced damage to DNA. *Free Radic. Biol. Med*. 1991; 10: 225-242.
- 40-Dobson-Stone C, Cox RD, Lonie L, Southam L, Fraser M, Wise C, Bernier F, Hodgson S, Porter DE, Hamish A, Simpson RW et Monaca AP. Comparison of fluorescent single-strand conformation polymorphism analysis and denaturing high-performance liquid chromatography for detection of EXT1

- and EXT2 mutations in hereditary multiple exostose. *European Journal of Human Genetics*. 2000; 8: 24-32.
- 41-Dosanjh MK, Collins DW, Fan W, Lennon GG, Albala JS, Shen Z, Schild D. Isolation and characterization of RAD51C, a new human member of the RAD51 family of related genes. *Nucleic Acids Res*. 1998; 26 (5): 1179-1184.
- 42-Duell E, Millikan RC, Pittman GS, Winkel S, Lunn RM., Tse CJ, Eaton A, Mohrenweiser HW, Newman B, Bell DA. Polymorphisms in the DNA repair gene *XRCC1* and breast cancer. *Cancer Epidemiol. Biomarkers Prev*. 2001; 10: 217-222.
- 43-Duell E, Wiencke JK, Cheng TJ, Varkonyi A, Zuo ZF, Ashok TD, Mark EJ, Wain JC, Christiani DC, Kelsey KT. Polymorphisms in the DNA repair genes *XRCC1* and *ERCC2* and biomarkers of DNA damage in human blood mononuclear cells [Published erratum appears in *Carcinogenesis* (Lond.), 2001; 21: 1457]. *Carcinogenesis* (Lond.), 2000; 21: 965-971.
- 44-Duker NJ. Chromosome breakage syndromes and cancer. *American journal of medical genetics*. 2002; 115: 125-129.
- 45-Dulic A, Bates PA, Zhang X, Martin SR, Freemont PS, Lindahl T, Barnes DE. BRCT domain interactions in the heterodimeric DNA repair protein XRCC1-DNA ligase III. *Biochemistry*. 2001; 40: 5906-5913.
- 46-Durocher F, Tonin P, Shattuck-Eidens D, Skolnick M, Narod SA, Simard J. Mutation analysis of the BRCA1 gene in 23 families with cases of cancer of the breast, ovary, and multiple other sites. *J Med Genet* 1996; 3(10): 814-819.
- 47-Eker AP, Vermeulen W, Miura N, Tanaka K, Jaspers NG, Hoeijmakers JH, Bootsma D. Xeroderma pigmentosum group A correcting protein from calf thymus. *Mutat Res*. 1992; 274(3):211-224.
- 48-Elahi A, Zheng Z, Park J, Eyring K, McCaffrey T, Lazarus P. The human OGG1 DNA repair enzyme and its association with orolaryngeal cancer risk. *Carcinogenesis*. 2002 Jul; 23(7):1229-1234.
- 49-Ellis LA, Taylor CF, Taylor GR. A comparison of fluorescent SSCP and denaturing HPLC for high throughput mutation scanning. *Methods*. 2000; 15: 556-564.

- 50-El-Mabrouk N, Labuda D. Haplotypes histories as networks of recombinations. Soumis à *Genetics* 2002.
- 51-Elnitski L, Hardison RC, Li J, Yang S, Kolbe D, Eswara P, O'Connor MJ, Schwartz S, Miller W, Chiaromonte F. Distinguishing Regulatory DNA From Neutral Sites. *Genome Res.* 2003; 13: 64-72.
- 52-Esteller M, Levine R, Baylin SB. MLH1 promoter hypermethylation is associated with the microsatellite instability phenotype in sporadic endometrial carcinomas. *Oncogene.* 1998; 17: 2413-2417.
- 53-Feinstein E, Cimino G, Gale RP. p53 in chronic myelogenous leukemia in acute phase. *Proc Natl Acad Sci U S A.* 1991; 88: 6293-6297.
- 54-Ferguson DO, Alt FW. DNA double-strand break repair and chromosomal translocations: lessons from animal models. *Oncogene.* 2001; 20: 5572-5579.
- 55-Ford BN, Ruttan CC, Kyle VL, Brackley ME, Glickman BW. Identification of single nucleotide polymorphisms in human DNA repair genes. *Carcinogenesis.* 2000; 21(11): 1977-1981.
- 56-Ford D, Easton DF, Stratton M. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *Am J Hum Genet.* 1998; 62: 676-689.
- 57-Friedberg EC, Walker GC, Siede W. DNA repair and mutagenesis, ASM press, Washington, D C, 1995.
- 58-Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annu Rev Immunol.* 2000; 18: 495-527.
- 59-Gao Y, Ferguson DO, Xie W, Manis JP, Sekiguchi J, Frank KM, Chaudhuri J, Horner J, DePinho RA, Alt FW. Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature.* 2000; 404 (6780): 897-900.
- 60-Gaymes TJ, Mufti GJ, Rassool FV. Myeloid leukemias have increased activity of the nonhomologous end-joining pathway and concomitant DNA misrepair that is dependent on the Ku70/86 heterodimer. *Cancer Res.* 2002; 62 (10): 2791-2797.

- 61-Gayther SA, Russell P, Harrington P. The contribution of germline BRCA1 and BRCA2 mutations to familial ovarian cancer: no evidence for other ovarian cancer susceptibility genes. *Am J Hum Genet.* 1999; 65: 1021-1029.
- 62-Groffen J, Stephenson JR, Heistercamp N. Philadelphia chromosomal breakpoints are clustered within a limited region, *bcr*, on chromosome 22. *Cell.* 1984; 36: 93-98
- 63-Gross E, Norbert A, Goette J, Schwarz-Boeger U, Kiechle M. A comparison of BRCA1 mutation analysis by direct sequencing, SSCP and dHPCL. *Hum Genet.* 1999; 105: 72-78.
- 64-Haber JE. The many interfaces of Mre11. *Cell.* 1998; 95 (5): 583-586.
- 65-Hardison RC, Roskin KM, Yang S, Diekhans M, Kent WJ, Weber R, Elnitski L, Li J, O'Connor M, Kolbe D, Schwartz S, Furey TS, Whelan S, Goldman N, Smit A, Miller W, Chiaromonte F, Haussler D. Covariation in Frequencies of Substitution, Deletion, Transposition, and Recombination During Eutherian Evolution. *Genome Res.* 2003; 13: 13-26.
- 66-Harris CC. Interindividual variation in human chemical carcinogenesis: implications for risk assessment. Moolgavkar SH, editor. Scientific issues in quantitative risk assessment. Boston, MA: Birkhauser; 1990: 235-251.
- 67-Hayashi K. PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR methods appl.* 1991; 1: 34-38.
- 68-Herman JG, Umar A, Polyak. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A.* 1998; 95: 6870-6875.
- 69-Hilgenfeld E, Padilla-Nash H, Schrock E, Ried T. Analysis of B-cell neoplasias by spectral karyotyping (SKY). *Curr Top Microbiol Immunol.* 1999; 246: 169-174.
- 70-Hu J, Smith TR, Miller MS, Mohrenweiser HW, Golden A, Case LD. Amino acid substitution variants of *APE1* and *XRCC1* genes associated with ionizing radiation sensitivity. *Carcinogenesis (Lond.).* 2001; 22: 917-922.
- 71-Hubscher U, Thommes P. DNA polymerase epsilon: in search of a function. *Trends Biochem Sci.* 1992; 17 (2): 55-58.

- 72-Infante-Rivard C, Krajnovic M, Labuda D, Sinnett D. Childhood acute lymphoblastic leukemia associated with parental alcohol consumption and polymorphisms of carcinogen-metabolizing genes. *Epidemiology*. 2002; 13 (3): 277-281.
- 73-Ionov Y, Peinado MA, Malkhosyan S. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*. 1993; 363: 558-561.
- 74-Jackson SP. DNA damage signaling and apoptosis. *Biochemical society transactions*. 2001; 29 (6): 655-661.
- 75-Janssen K, Schlink K, Gotte W, Hippler B, Kaina B, Oesch F. DNA repair activity of 8-oxoguanine DNA glycosylase 1 (*OGGI*) in human lymphocytes is not dependent on genetic polymorphism Ser326/Cys326. *Mutat. Res*. 2001; 486: 207-216.
- 76-Jegga AG, Sherwood SP, Carman JW, Pinski AT, Phillips JL, Pestian JP, Aronow BJ. Detection and visualization of compositionally similar cis-regulatory element clusters in orthologous and co-ordinately controlled genes. *Genome Research*. 2002; 12: 1408-1417.
- 77-Jülicher S, Bongartz M, Luty AJF, Kremsner PG, Kun JFJ. Functional analysis of a promoter variant of the gene encoding the interferon-gamma receptor chain I. *Immunogenetics*. 2003; 54: 675-680.
- 78-Keeney S, Chang GJ, Linn S. Characterization of a human DNA damage binding protein implicated in xeroderma pigmentosum E. *J Biol Chem*. 1993; 268 (28): 21293-21300.
- 79-Khanna KK, Jackson SP. DNA double strand breaks: signaling, repair and the cancer connection. *Nat. Genet*. 2001; 27: 247-254.
- 80-Kim S-U, Park SK, Yoo K-Y, Yoon K-S, Choi JY, Seo J-S, Park W-Y, Kim J-H, Noh D-Y, Ahn S-H, Choe K-J, Strickland PT, Hirvonen A et Kang D. XRCC1 genetic polymorphism and breast cancer risk. *Pharmacogenetics* 2002; 12: 335-338.
- 81-Kohno T, Shinmura K, Tosaka M, Tani M, Kim SR, Sugimura H, Nohmi T, Kasai H, Yokota J. Genetic polymorphisms and alternative splicing of the



- hOGG1* gene, that is involved in the repair of 8-hydroxyguanine in damaged DNA. *Oncogene*. 1998; 16: 3219-3225.
- 82- Kolodner R. Biochemistry and genetics of eukaryotic mismatch repair. *Genes Dev*. 1996; 10 (12): 1433-1442.
- 83- Krajinovic M, Labuda D, Mathonnet G, Labuda M, Moghrabi A, Champagne J, Sinnett D. Polymorphisms in genes encoding drugs and xenobiotic metabolizing enzymes, DNA repair enzymes, and response to treatment of childhood acute lymphoblastic leukemia. *Clin Cancer Res*. 2002; 8(3):802-10.
- 84- Krajinovic M, Labuda D, Sinnett D. Childhood acute lymphoblastic leukemia: genetic determinants of susceptibility and disease outcome. *Rev Environ Health* 2001; 16(4): 263-279.
- 85- Krajinovic M, Labuda D, Sinnett D. Glutathione S-transferase P1 genetic polymorphisms and susceptibility to childhood acute lymphoblastic leukaemia. *Pharmacogenetics*. 2002; 12(8): 655-658.
- 86- Krajinovic M, Labuda D, Richer C, Karimi S, Sinnett D. Susceptibility to childhood acute lymphoblastic leukemia: influence of *CYP1A1*, *CYP2D6*, *GSTM1* and *GSTT1* genetic polymorphisms. *Blood*. 1999; 93: 1496-1501.
- 87- Kuschel B, Auranen A, McBride S, Novik KL, Antoniou A, Lipscombe JM, Day NE, Easton DF, Ponder BAJ, Pharoah PDP. Variants in DNA double-strand break repair genes and breast cancer susceptibility. *Hum. Mol. Genet*. 2002; 11: 1399–1407.
- 88- Labuda D, Zietkiewicz E, Labuda M. The genetic clock and the age of the founder effect in growing populations: a lesson from French Canadians and Ashkenazim. *Am. J. Hum. Genet*. 1997; 61: 768.
- 89- Ladiges W, Wiley J, MacAuley A. Polymorphisms in the DNA repair gene XRCC1 and age-related disease. *Mechanisms of ageing and development*. 2003; 124 (1): 27-32 .
- 90- Lai C, Shields PG. The role of interindividual variation in human carcinogenesis. *J. Nutr*. 1999; 129: 552S-555S.

- 91-Le Marchand L, Donlon T, Lum-Jones A, Seifried A, Wilkens LR. Association of the hOGG1 Ser326Cys polymorphism with lung cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2002; 11 (4): 409-412.
- 92-Lee J, Lee YC, Yang SY, Yang PW, Luh SP, Lee CJ, Chen CJ, Wu MT. Genetic polymorphism of XRCC1 and risk of the esophageal cancer. *Int. J. Cancer.* 2001; 95: 240-246.
- 93-Lee S-G, Kim B, Choi J, Kim C, Lee I, Song K. Genetic polymorphisms of XRCC1 and risk of gastric cancer. *Cancer letters.* 2002; 187: 53-60.
- 94-Lehmann AR. Dual functions of DNA repair genes: molecular, cellular, and clinical implications. *Bioessays.* 1998; 20 (2):146-155.
- 95-Leung LY, ST Yuen, LP Chung. hMLH1 promoter methylation and lack of hMLH1 expression in sporadic gastric carcinomas with high-frequency microsatellite instability. *Cancer Res.* 1999; 59: 159-164.
- 96-Lindahl T, Barnes DE. Mammalian DNA ligases. *Annu Rev Biochem.* 1992; 61: 251-281.
- 97-Lindahl T. Instability and decay of the primary structure of DNA. *Nature (London).* 1993; 362: 709-715.
- 98-Lindahl T, Karran P, Wood RD. DNA excision repair pathways. *Curr Opin Genet Dev.* 1997; 2: 158-69.
- 99-Liu N, Lamerdin JE, Tebbs RS, Schild D, Tucker JD, Shen MR, Brookman KW, Siciliano MJ, Walter CA, Fan W, Narayana LS, Zhou ZQ, Adamson AW, Sorensen KJ, Chen DJ, Jones NJ, Thompson LH. XRCC2 and XRCC3, new human Rad51-family members, promote chromosome stability and protect against DNA cross-links and other damages. *Mol Cell.* 1998; 1 (6): 783-793.
- 100- Liu N, Schild D, Thelen MP, Thompson LH. Involvement of Rad51C in two distinct protein complexes of Rad51 paralogs in human cells. *Nucleic Acids Res.* 2002; 30 (4): 1009-1015.
- 101- Lodish H, Berk A, Zipursky LS, Matsudaira P, Baltimore D, Darnell J. *Molecular Cell Biology* (10). Regulation of transcription initiation. 2000.

- 102- Lundin C, Erixon K, Arnaudeau C, Schultz N, Jenssen D, Meuth M, Helleday T. *Molecular and cellular biology*. 2002; 22 (16): 5869-5878.
- 103- Lunn R, Langlois RG, Hsieh LL, Thompson CL, Bell DA. XRCC1 polymorphisms: effects on aflatoxin B1-DNA adducts and glycophorin A variant frequency. *Cancer Res*. 1999; 59: 2557-2561.
- 104- Ma Y, Pannicke Y, Schwarz, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell*. 2002; 108: 781-794.
- 105- Masson JY, Stasiak AZ, Stasiak A, Benson FE, West SC. Complex formation by the human RAD51C and XRCC3 recombination repair proteins. *Proc Natl Acad Sci U S A*. 2001; 98 (15): 8440-8446.
- 106- Matullo G, Palli D, Peluso M, Guarrera S, Carturan S, Celentano E, Krogh V, Munnia A, Tumino R, Polidoro S, Piazza A, Vineis P. XRCC1, XRCC3, XPD gene polymorphisms, smoking and 32P-DNA adducts in a sample of healthy subjects. *Carcinogenesis*. 2001; 22 (9): 1437-1445.
- 107- Matullo G, Guarrera S, Carturan S, Peluso M, Malaveille C, Davico L, Piazza A, Vineis P. DNA repair gene polymorphisms, bulky DNA adducts in white blood cells and bladder cancer in a case-control study. *Int. J. Cancer*. 2001; 92: 562-567.
- 108- Maxam AM, Gilbert W A new method for sequencing DNA *Proc Natl Acad Sci USA*. 1977. 74: 560-564.
- 109- McBride ML. Childhood cancer and environmental contaminants. *Can J Public Health* 89 (Suppl 1).1998; S53-62, S58-98.
- 110- Mellon I, Spivak G, Hanawalt PC. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell*. 1987; 51 (2): 241-249.
- 111- Meyn MS. Chromosome instability syndromes : lessons for carcinogenesis. *Curr Top. Microbiol. Immunol*. 1997; 221: 71-148.
- 112- Modrich P, Lahue R. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem*. 1996; 65: 101-133.

- 113- Moshous D, Callebaut I, de Chasseval R, Corneo B, Cavazzana-Calvo M, Le Deist F, Tezca I, Sanal O, Bertrand Y, Philippe N, Fisher A, Villartay J-P. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell*. 2001; 105: 177-186.
- 114- Moshous D, Li L, de Chasseval R, Philippe N, Jabado N, Cowan MJ, Fischer A, de Villartay J-P. A new gene involved in DNA double-strand break repair and V(D)J recombination is located on human chromosome 10p. *Human Molecular Genetics*. 2000; 9 (4): 583-588.
- 115- Nelson HH, Kelsey KT, Mott LA, Karagas MR. The *XRCC1* Arg399Gln Polymorphism, Sunburn, and Non-melanoma Skin Cancer. *Cancer Research*. 2002; 62: 152-155.
- 116- Noordzij JG, Verkaik NS, van der Burg M, van Veelen LR, de Bruin-Versteeg S, Wiegant W, Vossen JMJJ, Weemaes CMR, de Groot R, Zdzienicka MZ, van Gent DC, van Dongen JJM. Radiosensitive SCID patients with Artemis gene mutations show a complete B-cell differentiation arrest at the pre-BCR checkpoint in bone marrow. *Blood*. 2002; DOI 10.1182.
- 117- O'Donovan A, Davies AA, Moggs JG, West SC, Wood RD. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature*. 1994; 371 (6496): 432-435.
- 118- Paull TT, Gellert M. The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. *Mol Cell*. 1998; 1 (7): 969-979.
- 119- Peltomäki P, Vasen HFA. International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. *Gastroenterology*. 1997; 113: 1146-1158.
- 120- Pennacchio LA, Rubin EM. Genomic strategies to identify mammalian regulatory sequences. *Nature Reviews Genetics*. 2001; 2: 100-109.
- 121- Peto R Epidemiology, multistage models, and short-term mutagenicity models. *Cold Spring Harbor Conf Cell Prolif*. 1977; 4: 1404-1428.

- 122- Pittman DL, Weinberg LR, Schimenti JC. Identification, characterization, and genetic mapping of Rad51d, a new mouse and human RAD51/RecA-related gene. *Genomics*. 1998; 49 (1): 103-111.
- 123- Premstaller A, Oefner PJ, Oberacher H, Huber CG. Capillary array high-performance liquid chromatography of nucleic acids and proteins. *Anal Chem*. 2002; 74 (18): 4688-4693.
- 124- Preston R. Interindividual variations in susceptibility and sensitivity: linking risk assessment and risk management. *Toxicology*. 1996; 111: 331-341.
- 125- Rabbits TH. Chromosomal translocations in human cancer. *Nature*. 1994; 372: 143-149.
- 126- Rafii S, Lindblom A, Reed M, Meuth M, Cox A. A naturally occurring mutation in an ATP-binding domain of the recombination repair gene XRCC3 ablates its function without causing cancer susceptibility. *Hum Mol Genet*. 2003; 12 (8): 915-923.
- 127- Rafii S, O'Regan P, Xinarianos G, Azmy I, Stephenson T, Reed M, Meuth M, Thacker J, Cox A. XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer. *Hum. Mol. Genet*. 2002; 11(12): 1433-1438.
- 128- Ratnasinghe D, Yao SX, Tangrea JA, Qiao YL, Andersen MR, Barrett MG, Giffen CA, Erozan Y, Tockman MS, Taylor PR. Polymorphisms of the DNA repair gene *XRCC1* and lung cancer risk. *Cancer Epidemiol. Biomarkers Prev*. 2001; 10: 119-123.
- 129- Riballo E, Critchlow SE, Teo SH, Doherty AJ, Priestley A, Broughton B, Kysela B, Beamish H, Plowman N, Arlett CF, Lehmann AR, Jackson SP, Jeggo PA. Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. *Curr Biol*. 1999; 9 (13): 699-702.
- 130- Rice MC, Smith ST, Bullrich F, Havre P, Kmiec EB. Isolation of human and mouse genes based on homology to REC2, a recombinational repair gene from the fungus *Ustilago maydis*. *Proc Natl Acad Sci U S A*. 1997; 94 (14): 7417-7422.
- 131- Roddam PL, Rollinson S, O'Driscoll M, Jeggo PA, Jack A, Morgan GJ. Genetic variants of NHEJ DNA ligase IV can affect the risk of developing

- multiple myeloma, a tumour characterised by aberrant class switch recombination. *J. Med. Genet.* 2002; 39 (12): 900-905.
- 132- Rodenhuis S. *Ras* and human tumors. *Semin Cancer Biol.* 1992; 3: 241-247.
- 133- Rowley JD. Chromosome abnormalities in human cancer. Dans: DeVita VT, Hellman S, Rosenberg SA, editors. Principles and practice of oncology. Philadelphia: Lippincott; 1989: 81-87.
- 134- Sancar A. DNA repair in humans. *Annu. Rev. Genetics.* 1995; 29: 69-105.
- 135- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA.* 1977; 74: 5463-5467.
- 136- Schaeffer L, Roy R, Humbert S, Moncollin V, Vermeulen W, Hoeijmakers JH, Chambon P, Egly JM. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science.* 1993; 260 (5104): 58-63.
- 137- Schaeffer L, Moncollin V, Roy R, Staub A, Mezzina M, Sarasin A, Weeda G, Hoeijmakers JH, Egly JM. The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.* 1994; 13 (10): 2388-2392.
- 138- Schaffner C, Stilgenbauer S, Rappold GA, Dohner H, Lichter P. Somatic ATM mutations indicate a pathogenic role of ATM in B-cell chronic lymphocytic leukemia. *Blood.* 1999; 94: 748-753,
- 139- Schar P. Spontaneous DNA damage, genomic instability, and cancer—when DNA replication escapes control. *Cell.* 2001; 104: 329-332.
- 140- Schoenmakers EF, Huysmans c, Van de Ven WJ. Allelic knockout of novel splice variants of human recombination repair gene RAD51B in t(12;14) uterine leiomyomas. *Cancer Res.* 1999; 59: 19-23.
- 141- Schreiber E, Matthias P, Müller M, Schffner W. Rapid detection of octamer binding proteins with “mini-extracts”, prepared from a small number of cell. *Nucleic Acids Research.* 1989; 17(15): 6419.
- 142- Scully R, Chen J, Plug A, Xiao Y, Weaver D, Feunteun J, Ashley T, Livingston DM. Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell.* 1997; 88 (2): 265-275.

- 143- Seedhouse C, Bainton R, Lewis M, Harding A, Russell N and Das-Gupta E. The genotype distribution of the XRCC1 gene indicates a role for base excision repair in the development of therapy-related acute myeloblastic leukemia. *Blood*. 2002; 100 (10): 3761-3766.
- 144- Shen H, Xu Y, Qian Y, Yu R, Qin Y, Wang X, Spitz M, Wei Q. Polymorphisms of the DNA repair gene XRCC1 and risk of cancer in a Chinese population. *Int. J. Cancer*. 2000; 88: 601-606.
- 145- Shiloh Y. ATM and related protein kinases: safeguarding genome integrity. *Nat Rev Cancer*. 2003; 3 (3): 155-168.
- 146- Shitvelman E, B Lifshitz, RP Gale, and E Canaani Fused transcript of *abl* and *bcr* genes in chronic myelogenous leukemia. *Nature*. 1985; 315: 550-552.
- 147- Shivji MK, Podust VN, Hubscher U, Wood RD. Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry*. 1995; 34 (15): 5011-5017.
- 148- Simard J, Tonin P, Durocher F, Morgan K, Rommens J, Gingras S, Samson C, Leblanc JF, Belanger C, Dion F, et al. Common origins of BRCA1 mutations in Canadian breast and ovarian cancer families. *Nat Genet*. 1994; (4): 392-398.
- 149- Sinnett D, Kajinovic M, Labuda D. Genetic susceptibility to childhood acute lymphoblastic leukemia. *Leukemia and lymphoma*. 2000; 38 (5-6): 447-462.
- 150- Sinnett D, Lavergne L, Melancon SB, Dallaire L, Potier M, Labuda D. Lesch-Nyhan syndrome: molecular investigation of three French Canadian families using a hypoxanthine-guanine phosphoribosyltransferase cDNA probe. *Hum Genet*. 1988; 81 (1): 4-8.
- 151- Slamon DJ, Godolphin W, Jones LA. Studies of *HER-2/neu* proto-oncogene in human breast and ovarian cancer. *Science*. 1989; 244: 707-712.
- 152- Smale ST. Core promoter architecture for eukaryotic protein-coding genes. *Transcription: Mechanisms and regulation* (Ed. Conaway, RC et Conaway, JW) Raven Press, New York. 1994: 63-81.

- 153- Solomon E, Borrow J, Goddard AD. Chromosome aberrations and cancer. *Science*. 1991; 254: 1153-1160.
- 154- Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han J-H, Duan J, Carr JL, Lee MS, Koshy B, Kumar MA, Zhang GE, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaner SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Rúaño G, Vovis GF. Haplotype Variation and Linkage Disequilibrium in 313 Human Genes. *Science*. 2001; 293 (5529): 489-493.
- 155- Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*. 2001; 68 (4): 978-989.
- 156- Stern M, Umbach DM, van Gils CH., Lunn RM, Taylor JA. DNA repair gene *XRCC1* polymorphisms, smoking and bladder cancer risk. *Cancer Epidemiol. Biomarkers Prev*. 2001; 10: 125-131.
- 157- Sturgis E, Castillo EJ, Li L, Zheng R, Eicher SA, Clayman GL, Strom SS, Spitz MR, Wei Q. Polymorphisms of DNA repair gene *XRCC1* in squamous cell carcinoma of the head and neck. *Carcinogenesis (Lond.)*. 1999; 20: 2125-2129.
- 158- Sugimura H, Kohno T, Wakai K, Nagura K, Genka K, Igarashi H, Morris BJ, Baba S, Ohno Y, Gao CM, Li ZY, Wang JD, Takezaki T, Tajima K, Varga T, Sawaguchi T, Lum JK, Martinson JJ, Tsugane S, Iwamasa T, Shinmura K, Yokota J. *Cancer epidemiology, biomarkers & prevention*. 1999; 8: 669-674.
- 159- Svejstrup JQ. Mechanisms of transcription-coupled DNA repair. *Nature reviews Molecular Cell Biology*. 2002; 3: 21-29.
- 160- Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics*. 2002; 3: 1-7.
- 161- Takahashi T, Nagai N, Oda H, Ohama K, Kamada K, Miyagawa K. Evidence for *RAD51L1/HMGIC* fusion in the pathogenesis of uterine leiomyoma. *Gene Chromosomes Cancer*. 2001; 30: 196-201.



- 162- Thacker J. The role of homologous recombination processes in the repair of severe forms of DNA damage in mammalian cells. *Biochimie*. 1999; 81 (1-2): 77-85.
- 163- Thompson LH, Schild D. Recombinational DNA repair and human disease. *Mutation research*. 2002; 509: 49-78.
- 164- Tornaletti S, Hanawalt PC. Effect of DNA lesions on transcription elongation. *Biochimie*. 1999; 81 (1-2): 139-146.
- 165- Tseng T-L, Shih Y-P, Huang Y-C, Wang C-K, Chen P-H, Chang J-G, Yeh K-T, Arthur Chen Y-M, Buetow KH. Genotypic and phenotypic characterization of a putative tumor susceptibility gene, GNMT, in liver cancer. *Cancer research*. 2003; 63: 647-654.
- 166- Tsutsui Y, Morishita T, Iwasaki H, Toh H, Shinagawa H. A recombination repair gene of *Schizosaccharomyces pombe*, *rhp57*, is a functional homolog of the *Saccharomyces cerevisiae* RAD57 gene and is phylogenetically related to the human XRCC3 gene. *Genetics*. 2000; 154 (4): 1451-1461.
- 167- Tuteja N, Tuteja R. Unraveling DNA repair in human: Molecular mechanisms and consequences of repair defect. *Critical Reviews in Biochemistry and Molecular Biology*. 2001; 36 (3): 261-290.
- 168- Umar A, Buermeyer AB, Simon JA, Thomas DC, Clark AB, Liskay RM, Kunkel TA. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell*. 1996; 87 (1): 65-73.
- 169- Urios A, Blanco M. Induction of SOS-independent mutations by benzo[a]pyrene treatment in *Escherichia Coli* cells deficient in MutY and MutM DNA glycosylases: possible role of oxidative lesions. *Mutat. Res*. 1994; 54: 1890-1894.
- 170- Vanasse GJ, Halbrook J, Thomas S, Burgess A, Hoekstra MF, Disteché CM, Willerford DM. Genetic pathway to recurrent chromosome translocations in murine lymphoma involves V(D)J recombinase. *J Clin Invest*. 1999; 103 (12): 1669-1675.
- 171- Vasen HFA, Mecklin JP, Meera Kahn P, Lynch HT. Hereditary non-polyposis colorectal cancer. *Lancet*. 1991; 17710: 26478-19244.

- 172- Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*. 2002; 108: 171-182.
- 173- Venkitaraman AR. Chromosomal instability and cancer predisposition: insights from studies on the breast cancer susceptibility gene, BRCA2. *Cold Spring Harb Symp Quant Biol*. 2000; 65: 567-572.
- 174- Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 2001; 412 (9): 607-614.
- 175- Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev*. 2000; 14 (8): 927-939.
- 176- Weinberg RA, editor. *Oncogenes and the molecular origins of cancer*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1989: 307-326.
- 177- West SC, Chappell C, Hanakahi LA, Masson J-Y, McIlwraith MJ, Van Dick E. Double-strand break repair in human cells. *Cold spring harbor symposia on quantitative biology*. 2000; LXV: 315-320.
- 178- Wikman H, Risch A, Klimek F, Schmezer P, Spiegelhalter B, Dienemann H, Kayser K, Schulz V, Drings P, Bartsch H. *hOGG1* polymorphism and loss of heterozygosity (LOH): significance for lung cancer susceptibility in a caucasian population. *Int. J. Cancer*. 2000; 88: 932-937.
- 179- Winsey SL, Haldar NA, Marsh HP, Bunce M, Marshall SE, Harris AL, Wojnarowska F, Welsh KI. A variant within the DNA repair gene XRCC3 is associated with the development of melanoma skin cancer. *Cancer research*. 2000; 60: 5612-5616.
- 180- Wood RD. DNA repair in eukaryotes. *Annu.Rev.Biochem*. 1996; 65: 135-167.
- 181- Xing D, Tan W, Lin D. Genetic polymorphisms and susceptibility to esophageal cancer among Chinese population (Review). *Oncol Rep*. 2003; 10 (5): 1615-1623.
- 182- Xu J, Zheng SL, Turner A, Isaacs SD, Wiley KE, Hawkins GA, Chang BL, Bleecker ER, Walsh PC, Meyers DA, Isaacs WB. Associations between

- hOGG1 sequence variants and prostate cancer susceptibility. *Cancer Res.* 2002 Apr 15; 62(8): 2253-2257.
- 183- Yandell DW, Campbell TA, Dayton SH. Oncogenic point mutations in the human retinoblastoma gene: their application to genetic counseling. *N Engl J Med.* 1989; 321: 1689-1695.
- 184- Yu Y, Okayasu R, Well MM, Silver A, McCarthy M, Zabriskie R, Long S, Cox R, Ullrich RL. Elevated breast cancer risk in irradiated BALB/c mice associates with unique functional polymorphism of the PRKDC gene. *Cancer research.* 2001; 61 (5): 1820-1824.
- 185- Zhang H, Tomblin G, Weber BL. BRCA1, BRCA2, and DNA damage response: collision or collusion? *Cell.* 1998; 92 (4): 433-436.
- 186- Zhong Q, Chen CF, Li S, Chen Y, Wang CC, Xiao J, Chen PL, Sharp ZD, Lee WH. Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science.* 1999; 285 (5428): 747-750.

## ANNEXE I.

Tableau XXII Liste des amorces utilisées pour l'amplification d'ADN génomique ainsi que les conditions de PCR.

Gène	Fragment	Nom des amorces	Séquences (5'-3')	Clamp (5'-3')	*	**
ATM	2	ATM.2F.EX1.-489	AGTGGCGGCACTGAACTC		2	A
		ATM.2R.EX1.-143	ATGCGCATCCAGTATCACG			
ATM	3	ATM.3F.EX1.-822	AGCCTCTGAAGAGAGGAGCAT	GCCCCGCCG	2	A
		ATM.3R.EX1.-451	AGAGGCGGGACAAATTGC			
ATM	4	ATM.4F.EX1.-1135	GTATTGCGTGGAGGATGGAG	GCCCCGCCG	2	A
		ATM.4R.EX1.-728	GCTTACGAGTTGAGACTGGC			
ATM	5	ATM.5F.EX1.-1486	GGAGGTTTCATCACAGATGC		2	A
		ATM.5R.EX1.-1083	ACTACGAATTTGGGAGTAGGG			
ATM	6a	ATM.6aF.EX1.-1788	GGCTGAGAAACCATGTTAAGA		2	A
		ATM.6aR.EX1.-1436	CGTATGCGACGTCCATAAAA	CGGGCGGGGG		
ATM	7	ATM.7F.EX1.-2016	TTAGGGCTTTCTTAAGTAGGGG		2	A
		ATM.7R.EX1.-1735	TTGGTAATCTAGTTGAGTTCCG	CGGGCGGGGG		
BRCA1	1	BRCA1.1F.EX1.-199	CTTCTCTTCCGTCTCTTTCC		1	A
		BRCA1.1R.EX1.47	CCCGTCCAGGAAGTCTCAG	CGGGCGGGGG		
BRCA1	2	BRCA1.2F.EX1.-365	TCCATTAGGGCGGAAAGAGT		2	A
		BRCA1.2R.EX1.-58	AAGCAGCAGCCTCTCAGAAT			
BRCA1	3a	BRCA1.3aF.EX1.-802	GAACTACGAGTGCGCAGACA		2	A
		BRCA1.3aR.EX1.-563	GCGCGCTTGTAATTGTCAAC			
BRCA1	3b	BRCA1.3bF.EX1.-611	ACGAAAACATAAACTCCAGTCC		2	A
		BRCA1.3bR.EX1.-165	CGGATGACGTAAGGAAAGAG			
BRCA1	4	BRCA1.4F.EX1.-1071	TGCTTCTTACCAGCTTCC		1	A
		BRCA1.4R.EX1.-665	ACTACAATTTCCATCCAGCC			
BRCA1	5	BRCA1.5F.EX1.-1345	CCAGAATGGACGCCAAGG		1	A
		BRCA1.5R.EX1.-986	CACCCACAGAGATAGCG			
BRCA1	6	BRCA1.6F.EX1.-1728	GGAGCTCACAGGGTTGG		3	A
		BRCA1.6R.EX1.-1328	CCTTGGCGTCAATTCTGG			
BRCA1	7	BRCA1.7F.EX1.-2020	ACACAGAAGTTCTCAAGTGC		3	E
		BRCA1.7R.EX1.-1628	CCCGCTTGAATTCTCACC			
BRCA2	2	BRCA2.2F.EX1.-611	GATACTACCAAGCCCTGCG	GCCCCGCCG	1	A
		BRCA2.2R.EX1.-230	TATGCTACTCCAAGTCCCC			
BRCA2	3	BRCA2.3F.EX1.-976	GTCAGTCTTCCAGTTCAGC	GCCCCGCCG	2	A
		BRCA2.3R.EX1.-564	TCGCTCATGAAGTGTGAGG			
BRCA2	5	BRCA2.5F.EX1.-1578	TCCTAGTGAGACACCCCTCC		4	B
		BRCA2.5R.EX1.-1258	AAGTCTAATTCGAAAAGAAATACG			
BRCA2	6	BRCA2.6F.EX1.-1823	ATGCCAGCTGGAAAATAATG		1	A
		BRCA2.6R.EX1.-1444	CCTGGTAGTAAGGCCAGGT	CGGGCGGGGG		
BRCA2	7	BRCA2.7F.EX1.-2259	TCACCGCACAAAAAGAAAG		1	A
		BRCA2.7R.EX1.-1737	TTCTAGTGCAGCTGCTCAATT	CGGGCGGGGG		

DCLREIC	1b	DCLREIC.1bF.EX1.55	AAAACCGCAGCTGAAGCC		2	A
		DCLREIC.1bR.EX1.-336	GGAACGAAGAATGATTCTAAGC	CGGGCGGGGG		
DCLREIC	2	DCLREIC.2F.EX1.-254	AAATCGAAACTCCAGAGCA		2	A
		DCLREIC.2R.EX1.-542	CCAAGATCACGCAGAGATGA	CGGGCGGGGG		
DCLREIC	3	DCLREIC.3F.EX1.-525	ATCTCTGCGTGATCTTGGG		2	A
		DCLREIC.3R.EX1.-880	TGGGCGATAGAACTAGACCC			
DCLREIC	4	DCLREIC.4F.EX1.-814	AAACAACCCAGAACACTTTCG		2	A
		DCLREIC.4R.EX1.-1186	CGAAGGTTTGTACAGTGC			
DCLREIC	5	DCLREIC.5F.EX1.-1104	TGGGTGAGCCATTGTACC		2	A
		DCLREIC.5R.EX1.-1494	AGTAGCTGGAGATGTGTGCC	CGGGCGGGGG		
DCLREIC	6	DCLREIC.6F.EX1.-1478	ACACATCTCCAGTACTCGG		2	A
		DCLREIC.6R.EX1.-1808	AACCATCAAAGGTGTTGCC	CGGGCGGGGG		
DCLREIC	7	DCLREIC.7F.EX1.-1694	TCCATGAAATCTTGAGGAGG	GCCCCGCGG	2	A
		DCLREIC.7R.EX1.-2087	TTCACTAGGTTTGGTCTGGG			
hOGG1	S326C	F	GGTGGCCCTAAAGGACTCTC		12	F
		R	AAGGTGCTTGGGGAATTCT			
KU70	1b	KU70.1bF.EX1.-218	CTGAAACGTGAGGGATAGCTG		2	A
		KU70.1bR.EX1.25	CACCCTGACATGTTGGCTACT			
KU70	1c	KU70.1cF.EX1.-501	ACTCGAGACCACGCTCCTT		2	A
		KU70.1cR.EX1.-176	CATCCCAGTAGCTGCCTCAT	CGGGCGGGGG		
KU70	2	KU70.2F.EX1.-803	TAGGTAGAAGCTGGTTGGGG		2	A
		KU70.2R.EX1.-379	AGATCTCAAACGCAGACAGG			
KU70	3b	KU70.3bF.EX1.-1320	CCCAAGTCTCCACCTC		3	A
		KU70.3bR.EX1.-751	CGCACTATATCGCGTCAGG	CGGGCGGGGGCGGGGGCCG		
KU70	4	KU70.4F.EX1.-1631	CTCACCTTTCCTACACCGC		3	B
		KU70.4R.EX1.-1229	CATCATGCTGGGTGAGGG			
KU70	5	KU70.5F.EX1.-1746	CTATCCCTGGGACTGCAAAG	GCCCCGCGG	3	A
		KU70.5R.EX1.-1533	GGTGTGGGGTAAACCAGC			
KU70	6	KU70.6F.EX1.-1975	TTAATGGAGCACAACCGAAA	GCCCCGCGG	2	A
		KU70.6R.EX1.-1720	AAGCTCCCTTTCAGTCC			
KU70	7	KU70.7F.EX1.-2105	CTGCTCATTGTTACCCGAAA	GCCCCGCGG	2	A
		KU70.7R.EX1.-1802	CAAGTCAGGATAAGAGGAGCC			
KU80	1	KU80.1F.EX1.-310	TCTGTCGTAAGCTGTTTCAGG		2	A
		KU80.1R.EX1.+23	CCAAGTCCATGGCTTCTTT			
KU80	2	KU80.2F.EX1.-563	CAAGAAAATCTAACTCCACAATAACC		2	A
		KU80.2R.EX1.-238	CGCACATTCTCTCCATTCC			
KU80	3	KU80.3F.EX1.-907	GGGCTCGTGATCAAGTAAAGC		2	A
		KU80.3R.EX1.-522	TCTACTTTTATTTTGGTTATTGTGG	CGGGCGGGGG		
KU80	5	KU80.5F.EX1.-1523	TGAAAATTTATGCTGTCTGTC		2	A
		KU80.5R.EX1.-1142	AGATGTGGAGAATCTATGAGCC			
KU80	6	KU80.6F.EX1.-1734	TCTGGATCGTACAAGAAGGG		2	A
		KU80.6R.EX1.-1453	GTCAGCTAAATCTCTCCAGC			
KU80	7	KU80.7F.EX1.-2049	CCAAGTGCAGTTGCTCACAC		2	A
		KU80.7R.EX1.-1707	CCTTGTCTCCCTTCTTGTACG			
LIG4	4	LIG4.4F.EX1.-746	CCAGAAGATGCCGGTTCC		3	C
		LIG4.4R.EX1.-1010	GGTCTCCCTCGAGACAAAATC			
LIG4	5	LIG4.5F.EX1.-964	GCAAGCAGTTGATCCACAC		3	C

		LIG4.5R.EX1.-1322	CATAGTTGATTCTGGTCATGC			
LIG4	6	LIG4.6F.EX1.-1297	CGTTGCATGACCAGAAATC	GCCCCGCGG	1	A
		LIG4.6R.EX1.-1669	ACATTACACAACCTCTAAGGGC			
LIG4	7	LIG4.7F.EX1.-1611	CAGTGGGCCCTGATAAAACA		1	C
		LIG4.7R.EX1.-1857	TGAAGAGCAGATACTCACAAAACA	CGGGCGGGGG		
LIG4	8	LIG4.8F.EX1.-1730	TCCGTCTAGCCCAAGAACAT	GCCCCGCGG	1	A
		LIG4.8R.EX1.-2060	TTTGGAGTTGCCTGGTTTC			
MnSOD	V-9A	F	GCACTAGCAGCATGTTGAGC		11	F
		R	CCGTAGTCGTAGGGCAGGT			
MnSOD	I58T	F	TCCACCTCTGCTAGGTCCTT		13	G
		R	TAGGGCTGAGGTTGTCCAG			
MRE11A	10	MRE11A.10F.EX1.-1492	ATGCCTTCTATAGACAAGTAGGAGA		5	A
		MRE11A.10R.EX1.-1855	GTTTATTCTGCTAAGCTGCCAT			
MRE11A	11	MRE11A.11F.EX1.-1722	AGGCTGGGAATTCGTATTG		5	A
		MRE11A.11R.EX1.-1978	TGTCTGGCAAAGACAACATGA			
MRE11A	8	MRE11A.8F.EX1.-813	CTGTAAGTCGGGTAGATGCG		5	A
		MRE11A.8R.EX1.-1182	CAGAGACACCCTTGAATCA			
MRE11A	9	MRE11A.9F.EX1.-1179	TCTGATTGGGTGACTGAAAAG		5	A
		MRE11A.9R.EX1.-1553	AAGCTGAAAAGATAAGGACCACA			
MRE11A	MRE11A-1	MRE11A-1F	CCAACATGCTTTTCCTTTCAA		6	A
		MRE11A-1R	CCATTGAAGGGTCCTTGATCT			
MRE11A	MRE11A-2	MRE11A-2F	AGGCGGGGTATAAACCTTA		6	A
		MRE11A-2R	CTTCGTGGTCCCCAGGTA			
MRE11A	MRE11A-3	MRE11A-3F	ACAGCCAAGAGCGGAGAAG		6	A
		MRE11A-3R	TCGCGACACTTCATGGATAA			
MRE11A	MRE11A-4	MRE11A-4F	AGAGCCGAAGTGGACTTGAA		7	A
		MRE11A-4R	CCAGTTAACTCTGGGCTTCC			
MRE11A	MRE11A-5	MRE11A-5F	CTCCTGTTTCTGCAGGTTCC		6	A
		MRE11A-5R	CACCAGCAGGATGACTGAAA			
MRE11A	MRE11A-6	MRE11A-6F	GATGTGGTGGGGTTGTCT		6	A
		MRE11A-6R	CCTGTTTCTAGTTTCACCAAGA			
MRE11A	MRE11A-7	MRE11A-7F	AGAGCACGGGAAAGGAAAAT		6	A
		MRE11A-7R	GGGATTCCAGAAGTCAGGTG			
NBS1	NBS-6	NSBI-F6	TGATGGGTTGGGAAGCTACT		6	A
		NSBI-R6	CTCCAAAGCAACATAAGGAAAA			
NBS1	NBS1-1	NSRI-F1	CCTTGTACCAAGCAAACACTGA		6	A
		NSBI-R1	AATGGCACCAAAAATAATCGAG			
NBS1	NBS1-2	NSRI-F2	TTTGATCGATGTATTACCAAGCA		6	A
		NSRI-R2	ATTGAACATTTCCCATGACA			
NBS1	NBS1-3	NSBI-F3	GAGCGTGTACATCCTTGTTC		6	A
		NSBI-R3	AGATGCAAACCTCCTAAGCACAAT			
NBS1	NBS1-4	NSBI-F4	AACCCACCCTTTTGTCTG		2	A
		NSBI-R4	CATGGCTCGCTCCTTAATC			
NBS1	NBS1-5	NBSI-5F	CTGGGAATACAGGAGCATGG		9	F
		NBSI-5R	CCACCTGTCTCAAGCTGGTT			
NBS1	NBS1-7	NSBI-F7	CTGCAAACGCACGAAACTAC		8	A
		NSBI-R7	GGGCAGCAGTTCCACAT			
PRKDC	PRKDC-1	PRKDC-F1	GTGTGCCCTAACACCACTT		6	A

		PRKDC-R1	GGCATAACGCCTAAAGCATC		
PRKDC	PRKDC-2	PRKDC-F2	AAAAAGCACCAGAGAAGCTG	8	A
		PRKDC-R2	TGACACAAGCTCGGAGTGAG		
PRKDC	PRKDC-3	PRKDC-F3	TCCTCGCCTCTACGTCTCTG	8	A
		PRKDC-R3	GCACAGACACCCACAGCA		
PRKDC	PRKDC-4	PRKDC4F	GGCAAAATGGCAATCGAATTA	8	F
		PRKDC4R	CGCCTTTCACGGTAACC		
PRKDC	PRKDC-5	PRKDC5F	CGCGACAAGGACAAGCTC	8	A
		PRKDC5R	TGGTGTGTAGGACCCAGTT		
PRKDC	PRKDC-6	PRKDC-6F	CCTTGGTCCAGATGGTCTGA	2	B
		PRKDC-6R	TTGGGTGGCTACTTGGTGTT		
PRKDC	PRKDC-7	PRKDC-7F	CTAGGCCCTCGCTTGTT	9	F
		PRKDC-7R	GCAGCGACAAGGTCTCC		
RAD50	1	RAD50.1F.EX1.-339	GCAGGAAGCTGTGAGTGC	2	A
		RAD50.1R.EX1.41	CCTCTATTCAAAACTCCGC	GCCCCGCGC	
RAD50	2	RAD50.2F.EX1.-682	TTTTGACATGAACTGACCCC	2	A
		RAD50.2R.EX1.-322	GCACTCACAGTTCCTGC		
RAD50	3	RAD50.3F.EX1.-1044	CTGAGACAGGTAACTTCTTTCC	2	A
		RAD50.3R.EX1.-619	CTGGAATGGGGTAGTGG		
RAD50	4	RAD50.4F.EX1.-1456	TGTTTTATTACATTTGCTTTGCC	2	A
		RAD50.4R.EX1.-972	TAATCCATTCCGATCAGTGC		
RAD51	1	RAD51.1F.EX1.-326	GTGGTTAGCCTCGAACTCC	2	A
		RAD51.1R.EX1.92	CTTCTTACTCGTTGCC		
RAD51	2	RAD51.2F.EX1.-687	CTTCTCGAGTTCCTCAGC	2	A
		RAD51.2R.EX1.-285	GGCGAGAGTATCGTCTGAGC		
RAD51	3	RAD51.3F.EX1.-1193	CCCAAGAAACCCTACTCAGC	2	A
		RAD51.3R.EX1.-615	TCCTACTCGATTCTGGAG		
RAD51	4	RAD51.4F.EX1.-1359	CCGTGCAGGCCTTATATGAT	2	A
		RAD51.4R.EX1.-1151	TGTCAAAACAACACCACAAGC		
RAD51	5	RAD51.5F.EX1.-1653	TGTTGTGCGTTGTCTTTGA	2	A
		RAD51.5R.EX1.-1299	GTGCCGAAACTGGAAGGTTA		
RAD51	6	RAD51.6F.EX1.-1986	CCTAGAGACCAAAGTCTCTTACC	2	A
		RAD51.6R.EX1.-1578	GTGAGCTATGATGGGCCACT		
XRCC1	R399Q	F	CTTTGCCCTCAGATCACAC	11	F
		R	CAGGATAAGGAGCAGGGTTG		
XRCC2	XRCC2-1	XRCC2-1F	AACCATGTTGCACTCACACG	10	A
		XRCC2-1R	CGTCCAATTCACACACTTC		
XRCC2	XRCC2-2	XRCC2-2F	GGAACAGTGTGGAGACTGAAAA	10	A
		XRCC2-2R	CCTGTGTGCCCACTTCTAT		
XRCC2	XRCC2-3	XRCC2-3F	GGCCTGTGTCTTTGAAAAA	10	A
		XRCC2-3R	GAATCGTGTGCCCACTTAC		
XRCC2	XRCC2-4	XRCC2-4F	GGGCGAGAGTGCAATGAG	5	A
		XRCC2-4R	TTCAAAAACGCTAGGAAAGAGC		
XRCC2	XRCC2-5	XRCC2-5F	ACTTGGAACCAAACCATTCC	5	A
		XRCC2-5R	GATCACTGGAAGTCGGGAGT		
XRCC2	XRCC2-6a	XRCC2-6Fa	CTGCCGCTCAGGTGTGAT	3	D
		XRCC2-6Ra	CTTTGTCCGTTCTCGGGTTA		
XRCC2	XRCC2-6b	XRCC2-6Fb	AACCATGAATCGCACACACA	3	D

		XRCC2-6Rb	GGACAGCTGGTGCTTTTCTT		
XRCC2	XRCC2-7	XRCC2-7F	TAACCCGAGAACGGACAAAG	10	A
		XRCC2-7R	GCCCTATGGAAGGCACTACA		
XRCC3	T241M	F	AAGAAGGTCCCGTACTGCT	11	F
		R	CTCACCTGGTTGATGCACAG		
XRCC3	13	XRCC3.13F.EX1.-1613	GCCTCCAACCAAGCTTCAAT	1	A
		XRCC3.13R.EX1.-1375	GCATTTAATCCCCTGGCTGT		
XRCC3	14	XRCC3.14F.EX1.-1786	ACTCCTCATGGTGGCGATAC	1	A
		XRCC3.14R.EX1.-1587	TGGGTTTCATTGAAGCTTGGT		
XRCC3	15	XRCC3.15F.EX1.-1955	CCTCATCTGGCCACACTTT	5	A
		XRCC3.15R.EX1.-1731	TGGAGAGTTGAGACCAGTGC		
XRCC3	16	XRCC3.16F.EX1.-2122	CACAAAGAAAGGCAGGGAAA	5	A
		XRCC3.16R.EX1.-1916	AACCGTAAGGAGCACGTGAG		
XRCC3	8	XRCC3.8F.EX1.-213	CCTCTCCAATCCGTGTAAGG	1	A
		XRCC3.8R.EX1.29	GGGTTCCGCACTCCTCTT		
XRCC4	1	XRCC4.1F.EX1.-411	TCCTAAACAAAGTAGGAAAGCG	1	B
		XRCC4.1R.EX1.16	CGTGATCTAAATCCCGCC		
XRCC4	2	XRCC4.2F.EX1.-774	CTGAGGCTAGACACCACTCC	1	A
		XRCC4.2R.EX1.-389	CCGCTTTCCTACTTTGTTTAGG		
XRCC4	3	XRCC4.3F.EX1.-1018	TCTTCCTAGTCTCAACTGCCATC	1	A
		XRCC4.3R.EX1.-718	AGGTGGATTGCCCTTCTAGG		
XRCC4	4	XRCC4.4F.EX1.-1368	ATCTACTAGGTGTGATCCGAGG	1	A
		XRCC4.4R.EX1.-967	CTCTGGACGAATTACCACC		
XRCC4	5	XRCC4.5F.EX1.-1566	CATCCAAGCTCACAGGAAT	1	A
		XRCC4.5R.EX1.-1278	TCAAGTTCACAGCAACTTCTGA		
XRCC4	6	XRCC4.6F.EX1.-1668	GATGCGAACTCAAAGATACTGAGA	2	B
		XRCC4.6R.EX1.-1526	ACAACTGGCAGGTGGTTTTT		
XRCC4	7	XRCC4.7F.EX1.-2123	AAGTGCTACAATTACTGTTTATGC	1	A
		XRCC4.7R.EX1.-1694	CCCAATGGTAATGACTTAATAAACC		

\*Référer au tableau XXIII. \*\*Référer au tableau XXIV.



## ANNEXE II .

Tableau XXIII : Conditions de PCR

	Kit DNA polymerase*	Amorces	MgCl <sub>2</sub>	dNTP	Enzyme	Solution	Conc. ADN	Volume final
		(mM)	(mM)	(mM)	(U)	G-C riche**	(ng)	(ml)
1	Platinum Taq	0,42	1,5	0,1	1,8		15	60
2	Platinum Taq	0,2	1,5	0,1	1,8		15	60
3	FastStart Taq	0,42	1,5	0,1	1,8	1X	15	60
4	Platinum Taq	0,42	2	0,2	1,8		15	60
5	Platinum Taq	0,2	1,5	0,1	1		10	40
6	Platinum Taq	0,2	1,5	0,1	0,5		20	25
7	Platinum Taq	0,2	1,8	0,1	0,5		20	25
8	FastStart Taq	0,2	1,5	0,1	1	1X	5	20
9	FastStart Taq	0,2	1,8	0,1	1	1X	5	20
10	Platinum Taq	0,2	1,2	0,1	1		10	40
11	Platinum Taq	0,1	1,5	0,1	0,5		10	40
12	Platinum Taq	0,14	1,5	0,1	0,5		10	40
13	Platinum Taq	0,1	1,2	0,1	0,5		10	40

\*Utilisés selon les directives du fabricant. \*\*Fourni dans le kit FastStart.

## ANNEXE III.

Tableau XXIV : Protocole pour l'amplification PCR

	Début	Touchdown	Amplification principale	Final
<b>Protocole A</b>				
Dénaturation	95°C, 5min	95°C, 30s	95°C, 30s	
Appariement		62°C-0,5°C par cycle jusqu'à 55°C, 30s	55°C, 30s	
Élongation		72°C, 30s	72°C, 30s	72°C, 7min
Cycles	1	14	24	1
<b>Protocole B</b>				
Dénaturation	95°C, 5min	95°C, 30s	95°C, 30s	
Appariement		60°C-0,5°C par cycle jusqu'à 53°C, 30s	53°C, 30s	
Élongation		72°C, 30s	72°C, 30s	72°C, 7min
Cycles	1	14	24	1
<b>Protocole C</b>				
Dénaturation	95°C, 5min	95°C, 30s	95°C, 30s	
Appariement		58°C-0,5°C par cycle jusqu'à 51°C, 30s	51°C, 30s	
Élongation		72°C, 30s	72°C, 30s	72°C, 7min
Cycles	1	14	24	1
<b>Protocole D</b>				
Dénaturation	95°C, 5min	95°C, 30s	95°C, 30s	
Appariement		61°C-0,5°C par cycle jusqu'à 54°C, 30s	54°C, 30s	
Élongation		72°C, 30s	72°C, 30s	72°C, 7min
Cycles	1	14	24	1
<b>Protocole E</b>				
Dénaturation	95°C, 5min	95°C, 30s	95°C, 30s	
Appariement		63°C-0,5°C par cycle jusqu'à 56°C, 30s	56°C, 30s	
Élongation		72°C, 30s	72°C, 30s	72°C, 7min
Cycles	1	14	24	1

**Protocole F**

Dénaturation 95°C, 5min  
Appariement

95°C, 30s  
60°C, 30s

Élongation  
Cycles 1

72°C, 30s  
35

72°C, 7min  
1

**Protocole G**

Dénaturation 95°C, 5min  
Appariement

95°C, 30s  
62°C, 30s

Élongation  
Cycles 1

72°C, 30s  
35

72°C, 7min  
1

## ANNEXE IV.

Tableau XXV. Température d'analyse au dHPLC

Gène	Fragment	Température		
ATM	2	64	65,5	
ATM	3	61	62	
ATM	4	60	61	
ATM	5	57,8	58,6	
ATM	6a	55	57,5	
ATM	7	58		
BRCA1	1	62		
BRCA1	2	60	60,5	
BRCA1	3a	62	64	
BRCA1	3b	58,8		
BRCA1	4	62	62,5	
BRCA1	5	62	62,5	
BRCA1	6	66	68	
BRCA1	7	59	66	
BRCA2	2	59,5	60	
BRCA2	3	59,5	62,3	
BRCA2	5	56,5		
BRCA2	6	55	57	
BRCA2	7	54	60	
DCLRE1C	1b	63,5	64,5	
DCLRE1C	2	57,5	62,5	
DCLRE1C	3	54,5	56,5	
DCLRE1C	4	60	63,5	
DCLRE1C	5	59	60,5	
DCLRE1C	6	57,5	61	
DCLRE1C	7	56,5	58,8	
KU70	1b	59	62	
KU70	2	63,5	65	
KU70	3b	64	65	
KU70	4	62,5	63,5	
KU70	5	63	65,5	
KU70	6	58,5	60,5	
KU70	7	58		
KU80	1	64,5		
KU80	2	52	56,5	
KU80	3	58	59,5	
KU80	5	55	62	63
KU80	6	58	59,5	
KU80	7	58	62	
LIG4	4	62,5	68	
LIG4	5	54	58	
LIG4	6	53	54,5	57,5
LIG4	7	55,7	57	

LIG4	8	54	55	
MRE11A	10	53	55,5	
MRE11A	11	53,5	54,1	
MRE11A	8	53,5	54	
MRE11A	9	54,5	55	
MRE11A	MRE11A-1	52,8	54,5	
MRE11A	MRE11A-2	58	58,4	
MRE11A	MRE11A-3	63	63,4	
MRE11A	MRE11A-4	61,5	52	
MRE11A	MRE11A-5	63	64	
MRE11A	MRE11A-7	56		
NBS1	NBS1-1	54,5	55	
NBS1	NBS1-3	52,8	53,2	
NBS1	NBS1-4	54	59,5	
NBS1	NBS1-5	59,2	59,5	
PRKDC	PRKDC-3	61	61,5	
PRKDC	PRKDC-5	66	66,5	
PRKDC	PRKDC-6	55,5	58	61,5
RAD50	1	61	64	
RAD50	2	62	65	66,5
RAD50	3	55	56,5	58
RAD50	4	52	55	56
RAD51	1	62	65,5	
RAD51	2	61,5	63	
RAD51	3	59	60	
RAD51	4	57	58	
RAD51	5	56	61	62,5
RAD51	6	60,5	61,5	63
XRCC2	XRCC2-1	55,3	55,8	
XRCC2	XRCC2-2	55,1	55,6	
XRCC2	XRCC2-4	56,5	57	
XRCC2	XRCC2-5	54,5	52	
XRCC2	XRCC2-6a	55,5	59	
XRCC2	XRCC2-6b	55,5	61	
XRCC2	XRCC2-7	57		
XRCC3	13	57		
XRCC3	14	63	63,8	
XRCC3	15	57,5	60	
XRCC3	16	59	60	
XRCC3	8	63,5		
XRCC4	1	63,5	65	
XRCC4	2	56	58	
XRCC4	3	57,5	58	
XRCC4	4	55,6	55,9	
XRCC4	5	55,5	57	
XRCC4	6	58,5		
XRCC4	7	54,5	60,5	

## ANNEXE V.

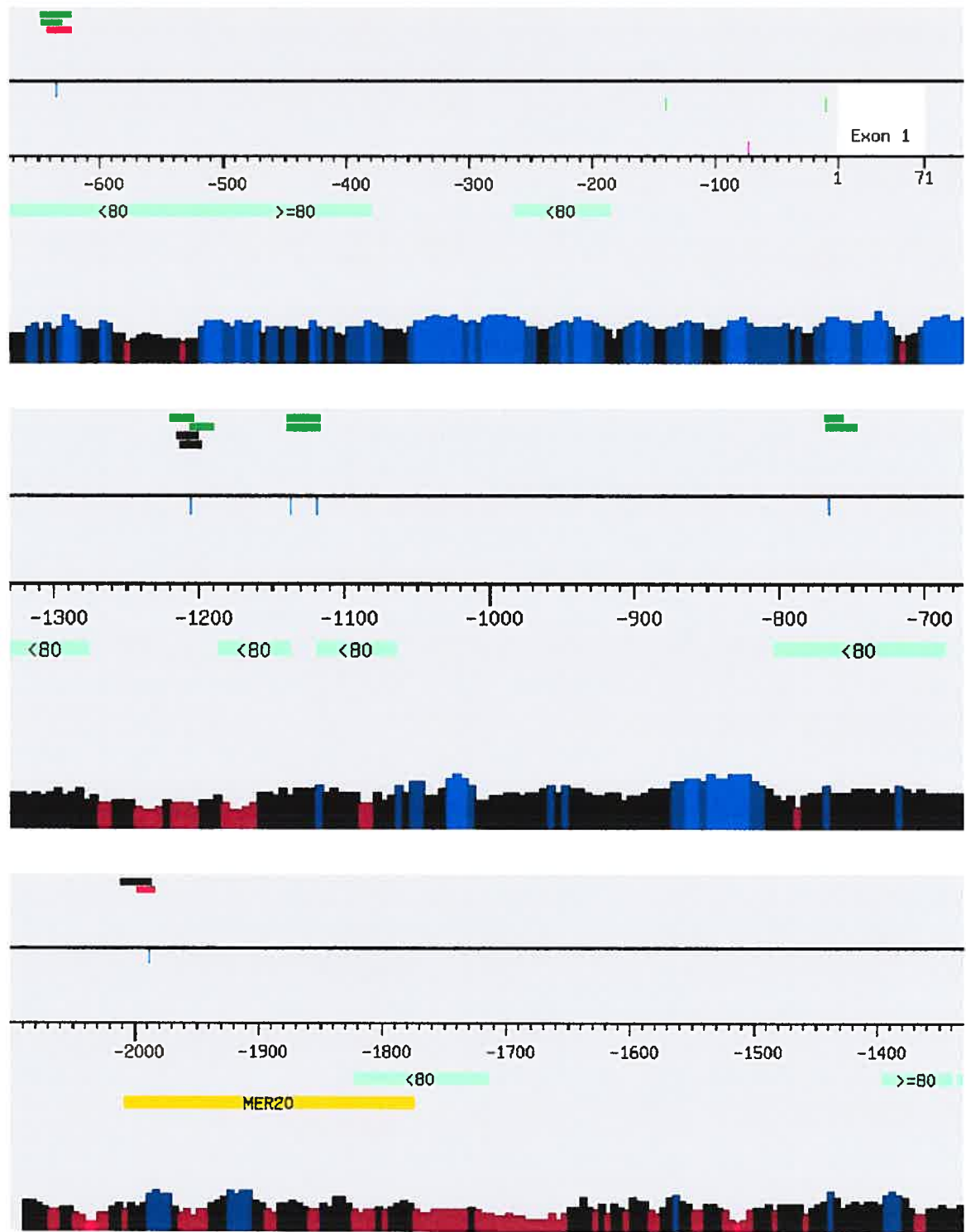


Figure 31. Représentation graphique de la région régulatrice du gène ATM.

Tableau XXVI. Description détaillée des polymorphismes identifiés dans le gène ATM.

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
ATM	Ex1.-766	C->T  transition	C	53	cAMP-response element-binding protein->0.866  Pax-3 paired domain protein, expressed in embryogenesis, mutations correlate to Waardenburg Syndrome->0.78	cAMP-response element-binding protein->0.866  Pax-3 paired domain protein, expressed in embryogenesis, mutations correlate to Waardenburg Syndrome->0.786	10013-10131	
ATM	Ex1.-635	T->A  transversion	T	61	cut-like homeodomain protein->0.797  Mammalian C-type LTR CCAAT box->0.877	cut-like homeodomain protein->0.797  Mammalian C-type LTR CCAAT box->0.864  Muscle Initiator Sequence->0.861	10143-10314	
ATM	Ex1.-1988	T->C  transition	T	51	Brachyury->0.693	c-Rel->0.947		MER20
ATM	Ex1.-1206	G->T  transversion	G	41	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.866  E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.741  HMGI(Y) high-mobility-group protein I (Y), architectural transcription factor organizing the framework of a nuclear protein-DNA transcriptional complex->0.92  TCF/LEF-1, involved in the Wnt signal transduction pathway->0.916	HMGI(Y) high-mobility-group protein I (Y), architectural transcription factor organizing the framework of a nuclear protein-DNA transcriptional complex->0.92  TCF/LEF-1, involved in the Wnt signal transduction pathway->0.876		
ATM	Ex1.-1137	A->G  transition	A	53	Aryl hydrocarbon receptor / Arnt heterodimers->0.937	Aryl hydrocarbon receptor / Arnt heterodimers->0.937	9630-9680	
ATM	Ex1.-1119	G->A  transition	G	57	Aryl hydrocarbon receptor / Arnt heterodimers->0.937	Aryl hydrocarbon receptor / Arnt heterodimers->0.937	9698-9753	

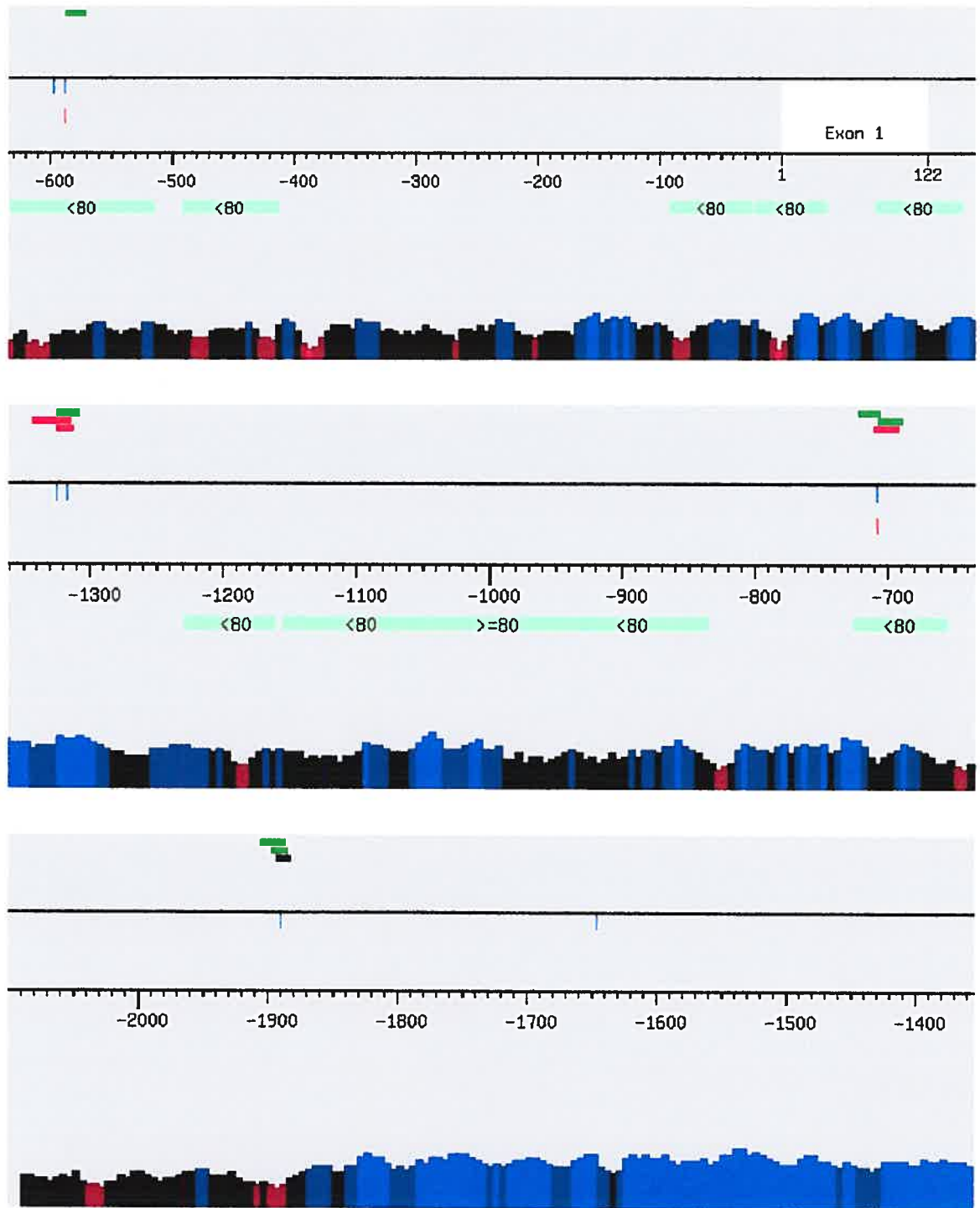


Figure 32. Représentation graphique de la région régulatrice du gène BRCA1.



Tableau XXVII. Description détaillée des polymorphismes identifiés dans le gène BRCA1

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
BRCA1	Ex1.-1317	C->G transversion	C	71				LTR12C
BRCA1	Ex1.-1325	C->T  transition	C	73	Pituitary Homeobox 1 (Ptx1)->0.794	B-cell-specific activating protein->0.804  Pituitary Homeobox 1 (Ptx1)->0.804  SF1 steroidogenic factor 1->0.988		LTR12C
BRCA1	Ex1.-708	A->G  transition	A	63	C/EBP binding site->0.855  Monomers of the nur subfamily of nuclear receptors (nur77, nurr1, nor-1)->0.947	C/EBP binding site->0.855  Glucocorticoid receptor, C2C2 zinc finger protein binds glucocorticoid dependent to GREs->0.82  Monomers of the nur subfamily of nuclear receptors (nur77, nurr1, nor-1)->0.947	2618-2689	
BRCA1	Ex1.-598	INS ACA nd	ACA	35			2711-2829	
BRCA1	Ex1.-588	A->G transition	A	43	Xenopus fork head domain factor 3->0.826	Xenopus fork head domain factor 3->0.826	2711-2829	
BRCA1	Ex1.-1890	T->C  transition	T	45	Cdx-2 mammalian caudal related intestinal transcr. factor->0.883  MyT1 zinc finger transcription factor involved in primary neurogenesis->0.996  Nuclear factor of activated T-cells->0.976	Cdx-2 mammalian caudal related intestinal transcr. factor->0.847  MyT1 zinc finger transcription factor involved in primary neurogenesis->0.899		LTR12C
BRCA1	Ex1.-1646	C->T transition	C	71				LTR12C

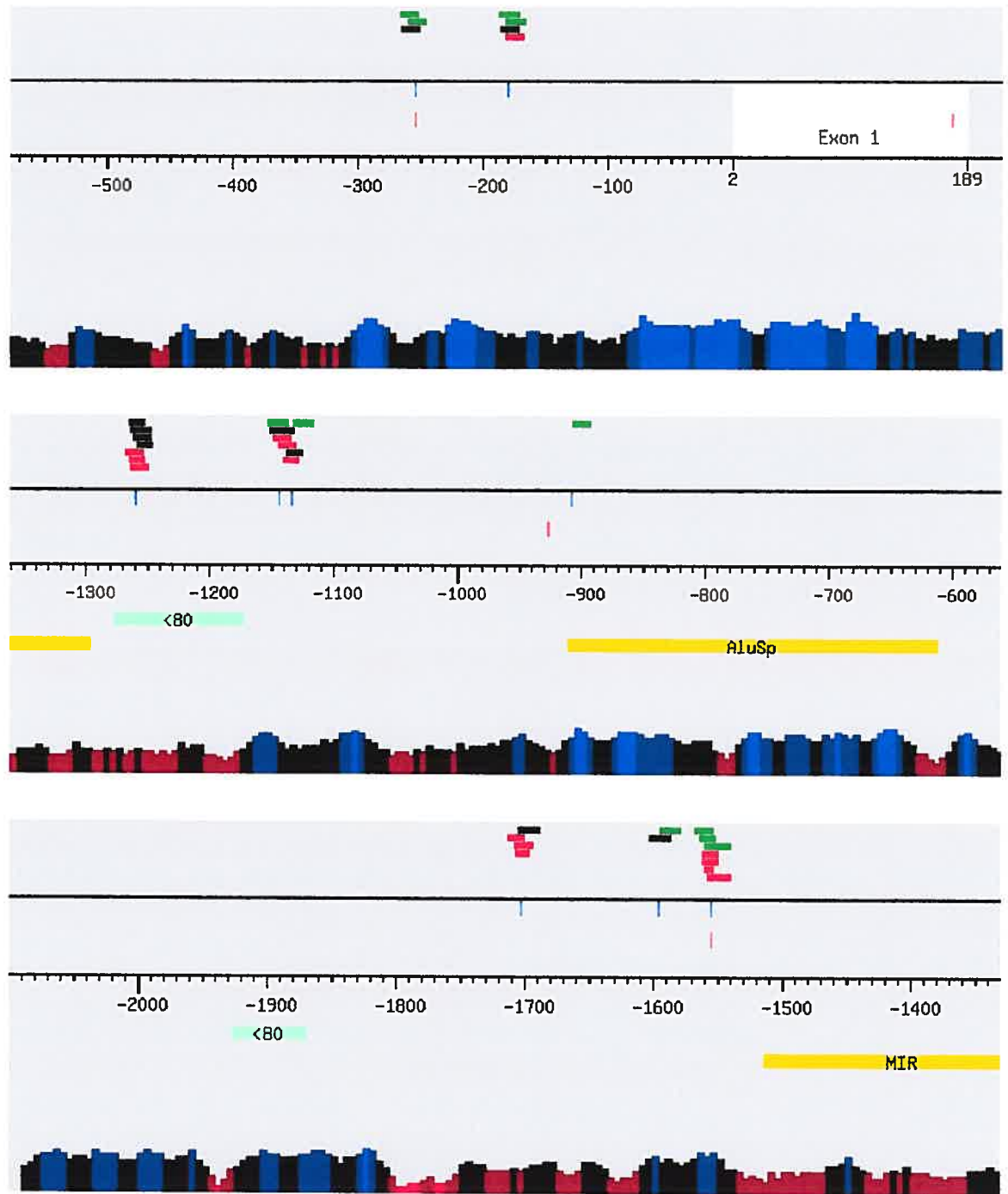


Figure 33. Représentation graphique de la région régulatrice du gène BRCA2.

Tableau XXVIII. Description détaillée des polymorphismes identifiés dans le gène BRCA2

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
BRCA2	Ex1.-254	G->A  transition	G	49	Gut-enriched Krueppel-like factor->0.954  Gut-enriched Krueppel-like factor->0.924  interferon regulatory factor 2->0.872	Gut-enriched Krueppel-like factor->0.924  interferon regulatory factor 2->0.866		
BRCA2		G->A  transition	G	59	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.765  Ecotropic viral integration site 1 encoded factor->0.88  Ecotropic viral integration site 1 encoded factor->0.753	Ecotropic viral integration site 1 encoded factor->0.88  Ecotropic viral integration site 1 encoded factor->0.75  interferon regulatory factor 2->0.822		
BRCA2	Ex1.-1144	G->A  transition	G	51	CCAAT/enhancer binding protein beta->0.953  Cone-rod homeobox-containing transcription factor / otx-like homeobox gene->0.968	Cone-rod homeobox-containing transcription factor / otx-like homeobox gene->0.974  Gut-enriched Krueppel-like factor->0.911  Gut-enriched Krueppel-like factor->0.915		
BRCA2	Ex1.-1260	DEL GTCTAA  nd	GTCTAA	33	Homeodomain proteins MSX-1 and MSX-2->0.981  Pituitary Homeobox 1 (Ptx1)->0.825  POU-factor Tst-1/Oct-6->0.947  prostate-specific homeodomain protein NKX3.1->0.863	Homeodomain proteins MSX-1 and MSX-2->0.981  octamer-binding factor 1->0.819  POU-factor Tst-1/Oct-6->0.894	82391-82495	
BRCA2	Ex1.-1134	C->T	C	49	Alpha (1)-fetoprotein transcription factor (FTF), liver receptor homologue-1 (LHR-1)->0.972	Cone-rod homeobox-containing transcription factor / otx-like homeobox gene->0.968		

		transition			Cone-rod homeobox-containing transcription factor / otx-like homeobox gene->0.968	MyT1 zinc finger transcription factor involved in primary neurogenesis->0.881		
<u>BRCA2</u>	<u>Ex1.-1555</u>	A->C  transversion	A	49	Epstein-Barr virus transcription factor R->0.844  stimulating protein 1 SP1, ubiquitous zinc finger transcription factor->0.924  Zinc finger transcription factor ZBP-89->0.986	Epstein-Barr virus transcription factor R->0.817  MYC-associated zinc finger protein related transcription factor->0.882  MZF1->0.99  serum response factor->0.885  stimulating protein 1 SP1, ubiquitous zinc finger transcription factor->0.928  Zinc finger transcription factor ZBP-89->0.996  Zinc finger transcription factor ZBP-89->0.947		
<u>BRCA2</u>	<u>Ex1.-1703</u>	T->C  transition	T	41	Hepatic nuclear factor 1->0.831	cellular and viral CCAAT box->0.957  nuclear factor Y (Y-box binding factor)->0.951  zinc finger with interaction domain->0.906		
<u>BRCA2</u>	<u>Ex1.-1596</u>	G->A  transition	G	49	Hox-1.3, vertebrate homeobox protein->0.847  upstream stimulating factor->0.927	Hox-1.3, vertebrate homeobox protein->0.847		
<u>BRCA2</u>	<u>Ex1.-908</u>	C->T transition	C	49	runt-factor AML-1->1	runt-factor AML-1->1		AluSp

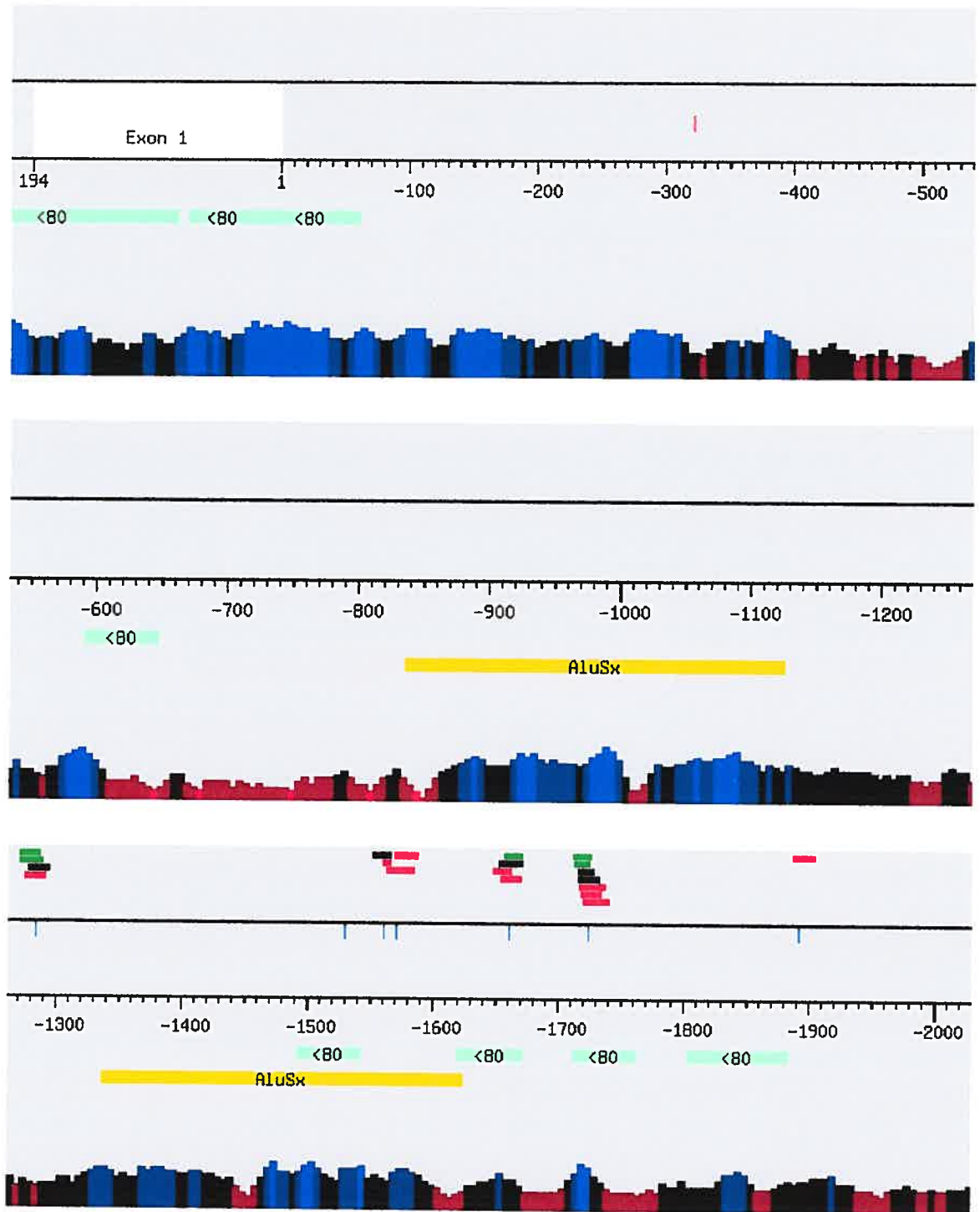


Figure 34. Représentation graphique de la région régulatrice du gène DCLER1C.

Tableau XXIX. Description détaillée des polymorphismes identifiés dans le gène DCLRE1C

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
DCLRE1C	Ex1.-1284	C->G  transversion	C	43	POZ/zinc finger protein, transcriptional repressor, translocations observed in diffuse large cell lymphoma->0.76  signal transducer and activator of transcription 1->0.783  TCF/LEF-1, involved in the Wnt signal transduction pathway->0.896	c-Ets-2 binding site->0.877  POZ/zinc finger protein, transcriptional repressor, translocations observed in diffuse large cell lymphoma->0.76  signal transducer and activator of transcription 1->0.787		
DCLRE1C	Ex1.-1530	G->C transversion	G	57			55959-56010	AluSx
DCLRE1C	Ex1.-1561	C->T  transition	C	57	Albumin D-box binding protein->0.864	TG-interacting factor belonging to TALE class of homeodomain factors->1		AluSx
DCLRE1C	Ex1.-1661	C->T  transition	C	41	Growth factor independence 1 zinc finger protein Gfi-1B->0.844  MIBP-1 / RFX1 complex->0.76	Ecotropic viral integration site 1 encoded factor->0.833  Growth factor independence 1 zinc finger protein Gfi-1B->0.84  Promyelocytic leukemia zinc finger (TF with nine Krueppel-like zinc fingers)->0.875	56087-56139	
DCLRE1C	Ex1.-1724	G->T  transversion	G	49	GC box elements->0.885  Myc associated zinc finger protein (MAZ)->0.929	Binding site for a Pbx1/Meis1 heterodimer->0.765  COMP1, cooperates with myogenic proteins in multicomponent complex->0.768	56179-56230	

					Myc associated zinc finger protein (MAZ)->0.919 Sterol regulatory element binding protein->0.867	GC box elements->0.884 Myc associated zinc finger protein (MAZ)->0.929 Tax/CREB complex->0.814		
<u>DCLRE1C</u>	<u>Ex1.-1571</u>	C->G  transversion	C	55		Se-Cys tRNA gene transcription activating factor->0.803  X-box binding protein RFX1->0.936		AluSx
<u>DCLRE1C</u>	<u>Ex1.-1892</u>	G->C transversion	G	49		Sox-5->0.979		

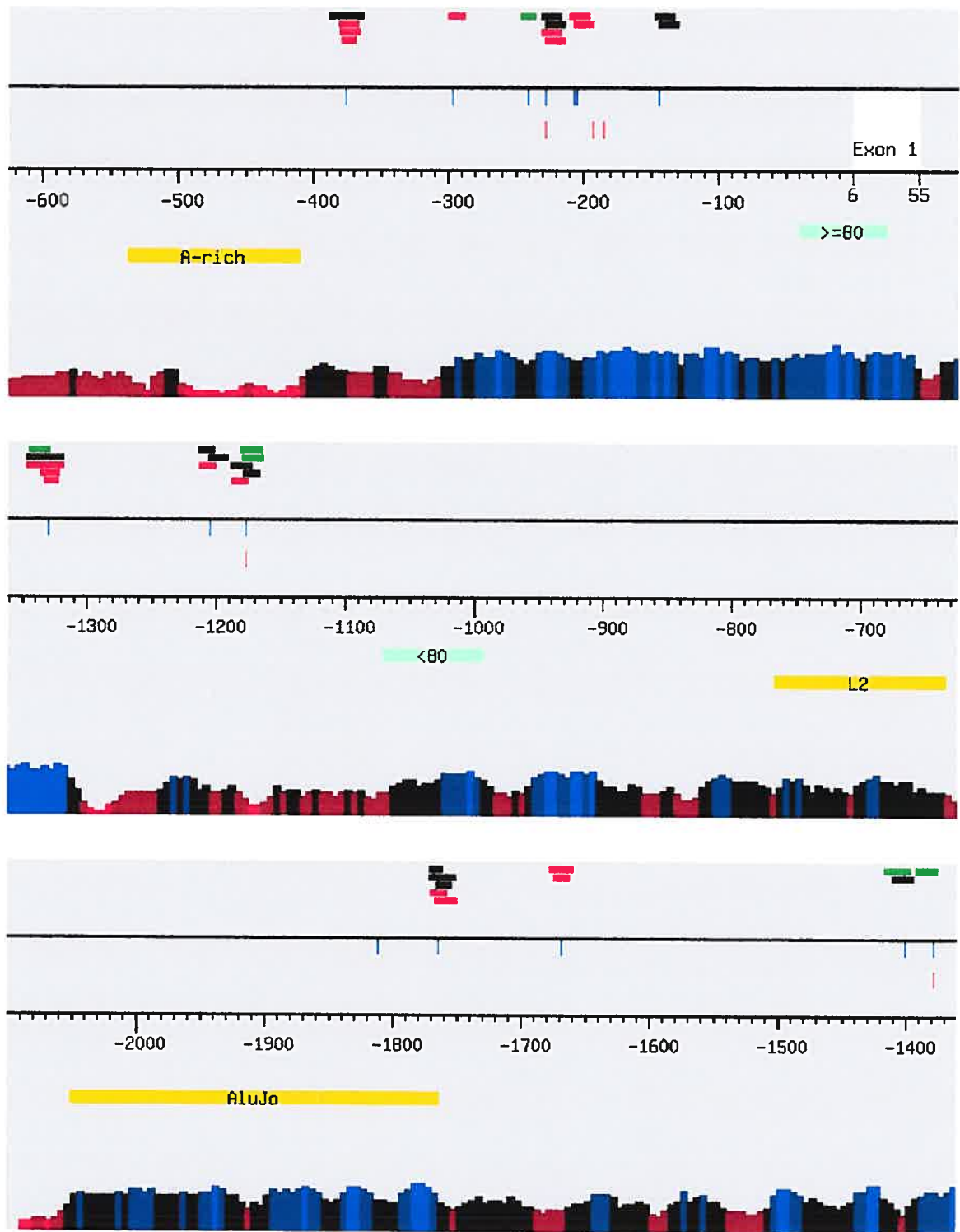


Figure 35. Représentation graphique de la région régulatrice du gène Ku80.



Tableau XXX. Description détaillée des polymorphismes identifiés dans le gène Ku80

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
KU80	Ex1.-376	G->C  transversion	G	41	T-Box factor 5 site (TBX5), mutations related to Holt-Oram syndrome->0.998	Hypoxia inducible factor, bHLH / PAS protein family->0.94 PAR-type chicken vitellogenin promoter-binding protein->0.875 X-box-binding protein 1->0.893		
KU80	Ex1.-297	C->T  transition	C	45		AREB6 (Atp1a1 regulatory element binding factor 6)->0.981		
KU80	Ex1.-1401	A->C  transversion	A	53	POZ/zinc finger protein, transcriptional repressor, translocations observed in diffuse large cell lymphoma->0.788 PPAR/RXR heterodimers->0.704	PPAR/RXR heterodimers->0.71		
KU80	Ex1.-1379	T->G transversion	T	63	GABP: GA binding protein->0.851	GABP: GA binding protein->0.859		
KU80	Ex1.-1330	C->T  transition	C	65	Hepatic nuclear factor 4->0.821  zebrafish PAX9 binding sites->0.808	B-cell-specific activating protein->0.83  cellular and viral CCAAT box->0.926 Hepatic nuclear factor 4->0.821 nuclear factor Y (Y-box binding factor)->0.818		
KU80	Ex1.-1205	T->C  transition	T	45	octamer-binding factor 1, POU-specific domain->0.904 Promyelocytic leukemia zink finger (TF with nine Krueppel-like zink fingers)->0.863	octamer-binding factor 1, POU-specific domain->0.935		

KU80	Ex1.-1177	A->G transition	A	29	Avian C-type LTR TATA box->0.835  cellular and viral TATA box elements->0.939  prostate-specific homeodomain protein NKX3.1- >0.864 Xenopus fork head domain factor 2- >0.928	cellular and viral TATA box elements- >0.95 PAX 2/5/8 binding site- >0.91  Xenopus fork head domain factor 2->0.902		
KU80	Ex1.-241	DEL TGCGCATGCTC  nd	TGCGCATGCTCGGCG	63	heat shock factor 1->0.94	heat shock factor 1->0.94		
KU80	Ex1.-228	DEL CGGGAATCTGC  nd	CGGGAATCTGCGCAT	59	E2F, involved in cell cycle regulation, interacts with Rb p107 protein- >0.817 NF-kappaB->0.825			
KU80	Ex1.-144	DEL CGGGAATCTGC  nd	CGGGAATCTGCGCAT	63	E2F, involved in cell cycle regulation, interacts with Rb p107 protein- >0.817 NF-kappaB->0.825			
KU80	Ex1.-207	A->C transversion	A	65		E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.804		
KU80	Ex1.-208	INS CGGGAATCTGC  nd	CGGGAATCTGCGCAT	59		E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.817 NF-kappaB- >0.843		
KU80	Ex1.-205	A->G transition	A	63		NF-kappaB- >0.827		
KU80	Ex1.-1669	T->A transversion	T	39		complex of Lmo2 bound to Tal-1, E2A proteins, and GATA-1, half- site 2->0.989 serum response factor->0.687		
KU80	Ex1.-1812	C->G transversion	C	59				AluJo
KU80	Ex1.-1765	DEL CTGT	CTGT	55	AREB6 (Atp1a1 regulatory element binding factor 6)- >0.997	AREB6 (Atp1a1 regulatory element binding factor 6)->0.997		AluJo

		nd		Erythroid krueppel like factor (EKLF)->0.955	Ecotropic viral integration site 1 encoded factor->0.789		
				Ribonucleoprotein associated zinc finger protein MOK-2 (mouse)->0.74			

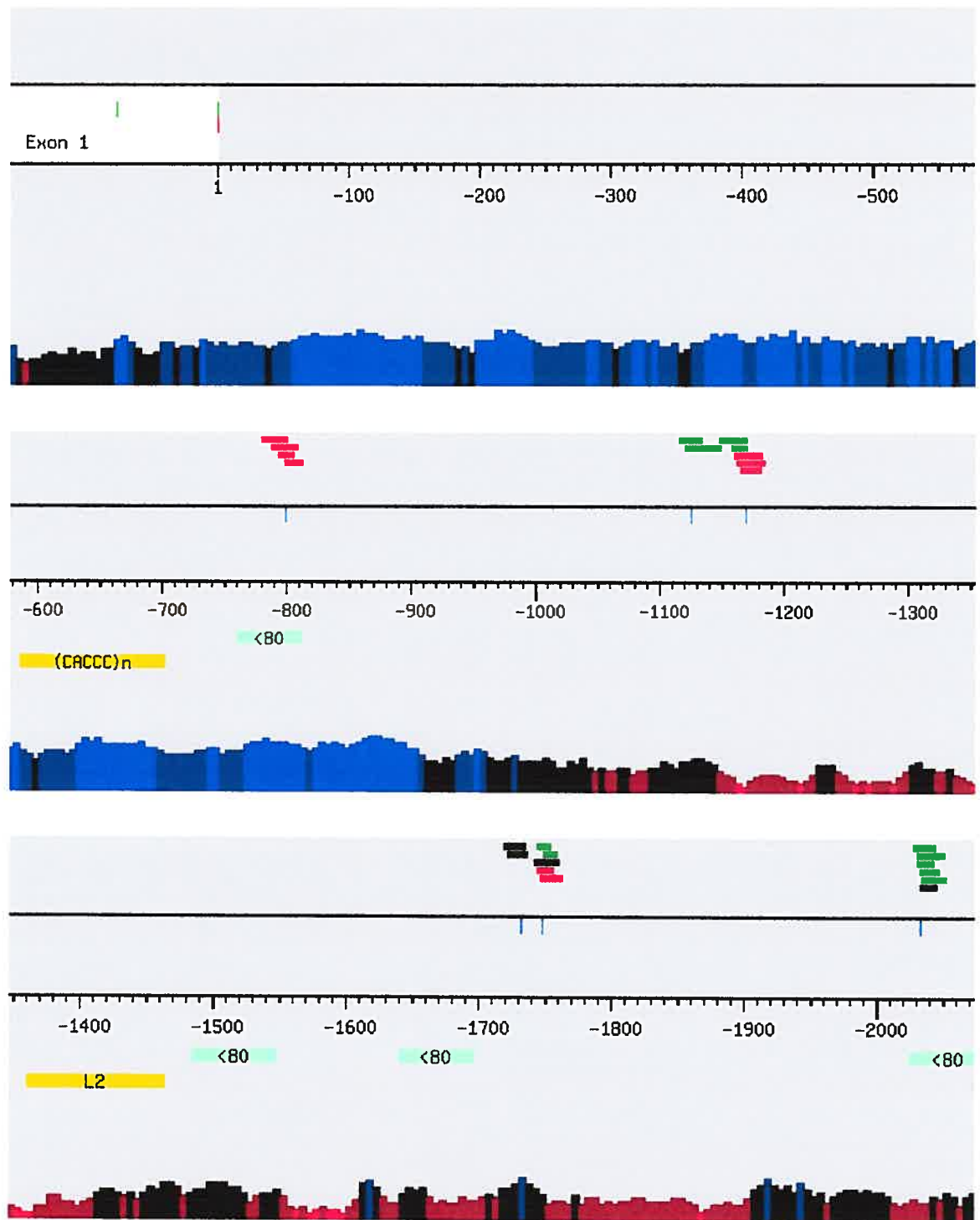


Figure 36. Représentation graphique de la région régulatrice du gène Lig4.

Tableau XXXI. Description détaillée des polymorphismes identifiés dans le gène Lig4

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
LIG4	Ex1.-1732	C->T transition	C	47	c-Rel->0.928 Elk-1->0.82			
LIG4	Ex1.-1748	A->G transition	A	43	heat shock factor 1->0.97 Progesterone receptor binding site->0.947 v-Myb->0.948	heat shock factor 1->0.98 Hepatic nuclear factor 1->0.779 MyT1 zinc finger transcription factor involved in primary neurogenesis->0.756 v-Myb->0.948		
LIG4	Ex1.-2033	T->C transition	T	35	Albumin D-box binding protein->0.872 CCAAT/enhancer binding protein beta->0.944 hepatic leukemia factor->0.885 homeo domain factor Nkx-2.5/Csx, tinman homolog low affinity sites->0.95 Homeobox protein MEIS1 binding site->0.88 POU transcription factor Brn-3->0.912	Albumin D-box binding protein->0.872 CCAAT/enhancer binding protein beta->0.944 hepatic leukemia factor->0.885 homeo domain factor Nkx-2.5/Csx, tinman homolog low affinity sites->0.95 POU transcription factor Brn-3->0.804	14865-14924	
LIG4	Ex1.-799	INS C nd	C	76		Activator protein 1->0.973 Activator protein 2->0.89 NF-kappaB (p50)->0.904 Ribonucleoprotein associated zinc finger protein MOK-2 (human)->0.996	13600-13652	
LIG4	Ex1.-1125	A->G transition	A	47	Myogenin / nuclear factor 1 or related factors->0.733 Progesterone receptor binding site->0.922	Myogenin / nuclear factor 1 or related factors->0.791 Progesterone receptor binding site->0.918		
LIG4	Ex1.-1169	A->T	A	25	myocyte enhancer factor->0.801	MEF2->0.99		

		transversion		MyT1 zinc finger transcription factor involved in primary neurogenesis->0.989	Muscle TATA box->0.858  myocyte enhancer factor->0.801  myocyte enhancer factor->0.947  MyT1 zinc finger transcription factor involved in primary neurogenesis->0.991		
--	--	--------------	--	---	---	--	--

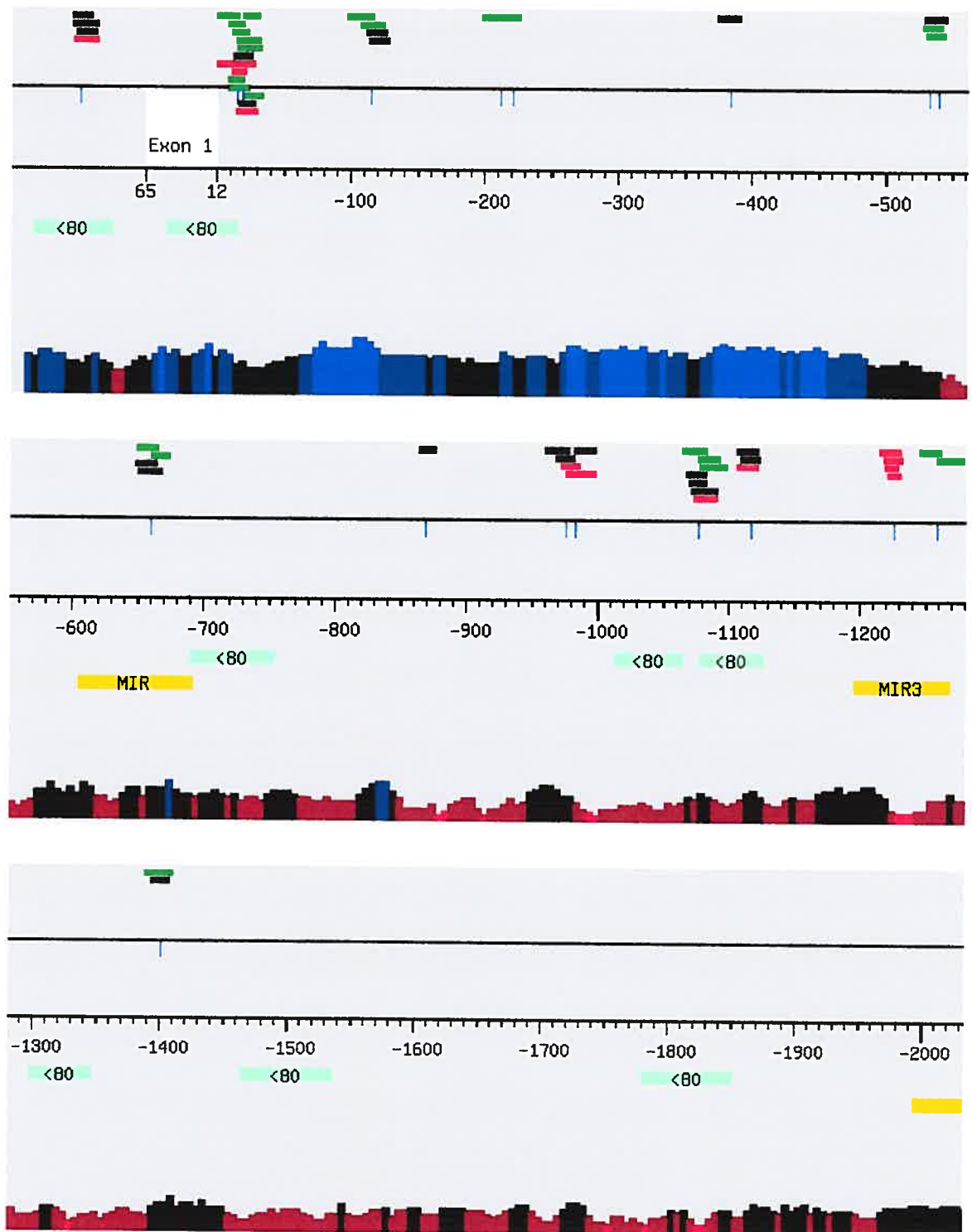


Figure 37. Représentation graphique de la région régulatrice du gène MRE11A.

Tableau XXXII. Description détaillée des polymorphismes identifiés dans le gène MRE11A

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
MRE11A	Ex1.-660	T->C  transition	T	47	Cdx-2 mammalian caudal related intestinal transcr. factor->0.853  cut-like homeodomain protein->0.813  GC box elements->0.895 Muscle TATA box->0.912	GC box elements->0.906  Muscle TATA box->0.844		MIR
MRE11A	Ex1.-540	T->A  transversion	T	35	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.866  E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.779	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.866  E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.779		
MRE11A	Ex1.-533	T->C transition	T	37	cut-like homeodomain protein->0.856			
MRE11A	Ex1.-384	T->C transition	T	71	upstream stimulating factor->0.938			
MRE11A	Ex1.-869	A->C transversion	A	18	MyT1 zinc finger transcription factor involved in primary neurogenesis->0.798			
MRE11A	Ex1.-19	G->A  transition	G	51	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.794  GATA-binding factor 1->0.992 PAX 2/5/8 binding site->0.883 Pax-3 paired domain protein, expressed in embryogenesis, mutations correlate to Waardenburg Syndrome->0.798  transcriptional repressor CDP->0.859	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.794  GATA-binding factor 1->0.992 Muscle TATA box->0.841 Pax-3 paired domain protein, expressed in embryogenesis, mutations correlate to Waardenburg Syndrome->0.793  transcriptional repressor CDP->0.872		



<u>MRE11A</u>	<u>Ex1.+49</u>	C->T  transition	C	53	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.741  E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.758  signal transducers and activators of transcription->0.944	signal transducers and activators of transcription->0.886	76823-76882
<u>MRE11A</u>	<u>Ex1.-15</u>	T->C  transition	T	51	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.794  PAX 2/5/8 binding site->0.883  Pax-3 paired domain protein, expressed in embryogenesis, mutations correlate to Waardenburg Syndrome->0.798  Pax1 paired domain protein, expressed in the developing vertebral column of mouse embryos->0.675  transcriptional repressor CDP->0.859  X-box binding protein RFX1->0.936	B-cell-specific activating protein->0.773  PAX 2/5/8 binding site->0.883  Pax-3 paired domain protein, expressed in embryogenesis, mutations correlate to Waardenburg Syndrome->0.809  Pax1 paired domain protein, expressed in the developing vertebral column of mouse embryos->0.616  transcriptional repressor CDP->0.859  winged helix protein, involved in hair keratinization and thymus epithelium differentiation->0.984  X-box binding protein RFX1->0.936	76923-76975
<u>MRE11A</u>	<u>Ex1.+315</u>	G->T transversion	G	65	B-cell-specific activating protein->0.775	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.909	
<u>MRE11A</u>	<u>Ex2.-508</u>	T->C transition	T	51	Myf5 myogenic bHLH protein->0.905		
<u>MRE11A</u>	<u>Ex2.-336</u>	A->G  transition	A	35	Avian C-type LTR TATA box->0.816  myocyte enhancer factor->0.919	Avian C-type LTR TATA box->0.816  myocyte enhancer factor->0.907	
<u>MRE11A</u>	<u>Ex2.-277</u>	C->T	C	37	POU factor Brn-2 (N-Oct 3)->0.924	HNF-3/Fkh Homolog 3 (= Freac-6)->0.971	

		transition			Ribonucleoprotein associated zinc finger protein MOK-2 (mouse)->0.762	Pax1 paired domain protein, expressed in the developing vertebral column of mouse embryos->0.651  POU factor Brn-2 (N-Oct 3)->0.924		
MRE11A	Ex1.-221	C->A transversion	C	55	Myogenin / nuclear factor 1 or related factors->0.776	Myogenin / nuclear factor 1 or related factors->0.793		
MRE11A	Ex1.-212	G->A transition	G	57				
MRE11A	Ex1.-115	G->T  transversion	G	67	COMP1, cooperates with myogenic proteins in multicomponent complex->0.773  Muscle Initiator Sequence->0.969 Wilms Tumor Suppressor->0.92 Wilms Tumor Suppressor->0.886	COMP1, cooperates with myogenic proteins in multicomponent complex->0.773  Muscle Initiator Sequence->0.952		
MRE11A	Ex1.-1401	A->C  transversion	A	41	Albumin D-box binding protein->0.87  Lentiviral TATA upstream element->0.852	Lentiviral TATA upstream element->0.769		
MRE11A	Ex1.-1259	T->A  transversion	T	31	Ecotropic viral integration site 1 encoded factor->0.812  hepatic leukemia factor->0.865	Ecotropic viral integration site 1 encoded factor->0.812  hepatic leukemia factor->0.865		MIR3
MRE11A	Ex1.-1226	A->C  transversion	A	29		heat shock factor 1->0.966 HMGI(Y) high-mobility-group protein I (Y), architectural transcription factor organizing the framework of a nuclear protein-DNA transcriptional complex->0.932  Nuclear factor of activated T-cells->0.984 octamer-binding factor 1->0.823		MIR3
MRE11A	Ex1.-976	T->C	T	35	runt-factor AML-1->1	Albumin D-box binding protein->0.842		

		transition			SEF1 binding site->0.757	Lentiviral TATA upstream element->0.729		
<u>MRE11A</u>	<u>Ex1.-1077</u>	G->T	G	33	Alpha (1)-fetoprotein transcription factor (FTF), liver receptor homologue-1 (LHR-1)->0.976	c-Ets-2 binding site->0.9		
		transversion			c-Ets-2 binding site->0.9	C/EBP binding site->0.853		
					C/EBP binding site->0.864	Pancreatic and intestinal lim-homeodomain factor->0.829		
					COUP antagonizes HNF-4 by binding site competition or synergizes by direct protein - protein interaction with HNF-4->0.856	signal transducers and activators of transcription->0.911		
					Pancreatic and intestinal lim-homeodomain factor->0.829			
					Ribonucleoprotein associated zinc finger protein MOK-2 (human)->0.982			
<u>MRE11A</u>	<u>Ex1.-1117</u>	T->G	T	35	c-Rel->0.924	GABP: GA binding protein->0.872	78038-78087	
		transversion			Elk-1->0.851			
<u>MRE11A</u>	<u>Ex1.-983</u>	T->C	T	29	hepatic nuclear factor 1->0.791			
		transition						

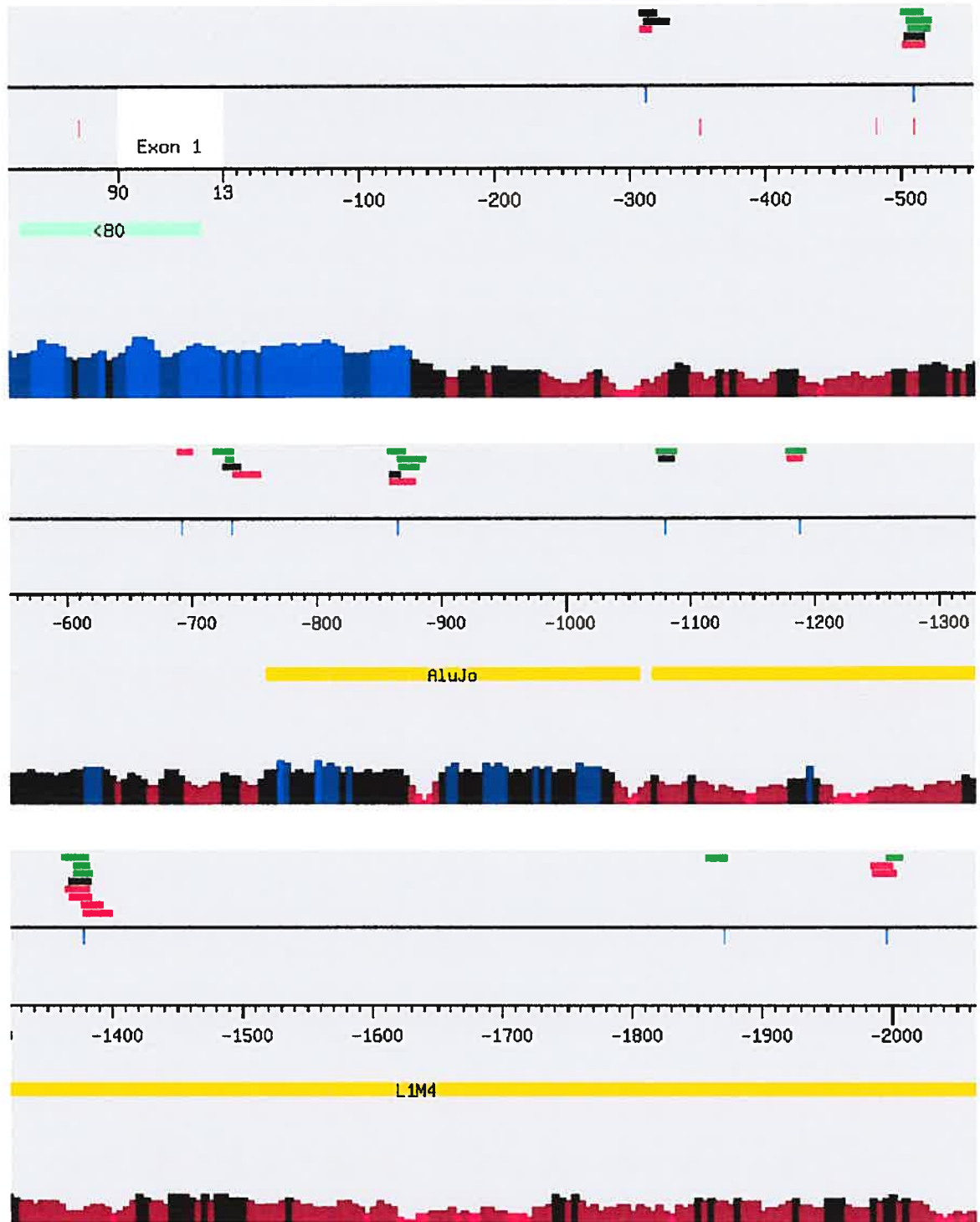


Figure 38. Représentation graphique de la région régulatrice du gène NBS1.

Tableau XXXIII. Description détaillée des polymorphismes identifiés dans le gène NBS1

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
NBS1	Ex1.-1871	A->G transition	A	35	Sox-5->0.986	Sox-5->0.983		L1M4
NBS1	Ex1.-1996	A->G  transition	A	35	SF1 steroidogenic factor 1->0.95	POZ/zinc finger protein, transcriptional repressor, translocations observed in diffuse large cell lymphoma->0.771  serum responsive factor->0.812 SF1 steroidogenic factor 1->0.95		L1M4
NBS1	Ex1.-510	G->A transition	G	43	c-Myc/Max heterodimer->0.924  MIT (microphthalmia transcription factor) and TFE3->0.839  octamer-binding factor 1->0.813  POU factor Brn-2 (N-Oct 3)->0.945	Avian C-type LTR TATA box->0.816  c-Myc/Max heterodimer->0.939  MIT (microphthalmia transcription factor) and TFE3->0.859  POU factor Brn-2 (N-Oct 3)->0.939		
NBS1	Ex1.-312	DEL AGTA  nd	AGTA	25	Cdx-2 mammalian caudal related intestinal transcr. factor->0.85  prostate-specific homeodomain protein NKX3.1->0.928	muscle-specific Mt binding site->0.969		
NBS1	Ex1.-1378	G->A transition	G	24	cAMP-responsive element binding protein 1->0.919  Fork head Related ACTivator-4->0.785  MyT1 zinc finger transcription factor involved in primary neurogenesis->0.798  Promyelocytic leukemia zink finger (TF with nine Krueppel-like zink fingers)->0.926	C/EBP binding site->0.862  cAMP-responsive element binding protein 1->0.919  cut-like homeodomain protein->0.85  HNF-3/Fkh Homolog 1->0.857		L1M4

						myogenic enhancer factor 2->0.74 MyT1 zinc finger transcription factor involved in primary neurogenesis->0.798  Promyelocytic leukemia zink finger (TF with nine Krueppel-like zink fingers)->0.87	
<u>NBS1</u>	<u>Ex1.-1188</u>	C->T  transition	C	33	Xenopus fork head domain factor 3->0.856	MyT1 zinc finger transcription factor involved in primary neurogenesis->0.776  Xenopus fork head domain factor 3->0.872	LIM4
<u>NBS1</u>	<u>Ex1.-692</u>	C->T  transition	C	39		octamer-binding factor 1, POU-specific domain->0.864	
<u>NBS1</u>	<u>Ex1.-732</u>	DEL A  nd	A	35	E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.805  Hox-1.3, vertebrate homeobox protein->0.837  MZF1->0.99	Hox-1.3, vertebrate homeobox protein->0.837  MEF2->0.983  MZF1->0.99	
<u>NBS1</u>	<u>Ex1.-865</u>	A->G  transition	A	35	cellular and viral TATA box elements->0.932  MEF2->0.963 NF-kappaB->0.884 Smad4 transcription factor involved in TGF-beta signaling->0.947	cellular and viral TATA box elements->0.932  MEF2->0.963 NF-kappaB->0.884 Ribonucleoprotein associated zinc finger protein MOK-2 (human)->0.996	AluJo
<u>NBS1</u>	<u>Ex1.-1080</u>	A->G  transition	A	37	Ecotropic viral integration site 1 encoded factor->0.854  GATA-binding factor 1->0.993	Ecotropic viral integration site 1 encoded factor->0.854	LIM4

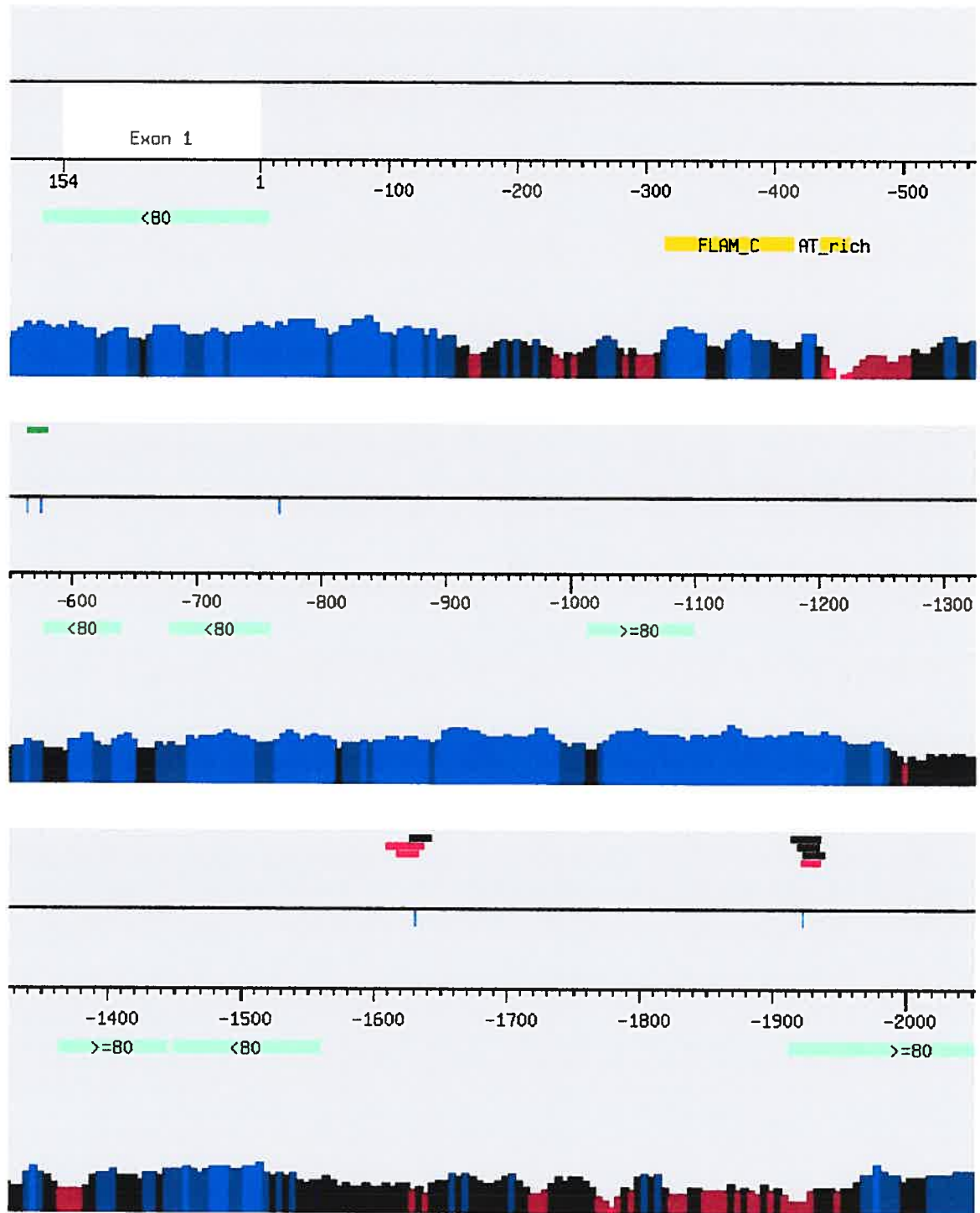


Figure 39. Représentation graphique de la région régulatrice du gène PRKDC.

Tableau XXXIV. Description détaillée des polymorphismes identifiés dans le gène PRKDC

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
PRKDC	Ex1.-1923	T->C transition	T	31	Mammalian C-type LTR TATA box->0.991 MEF2->0.983 Xenopus homeodomain factor Xvent-2; early BMP signaling response->0.852	Interferon regulatory factor 3 (IRF-3)->0.95	83086-83272	
PRKDC	Ex1.-1631	T->C transition	T	41	cellular and viral TATA box elements->0.915	Myogenin / nuclear factor 1 or related factors->0.725 Tal-1 beta/ITF-2 heterodimer->0.85		
PRKDC	Ex1.-766	G->A transition	G	73				
PRKDC	Ex1.-575	T->G transversion	T	59	RAR-related orphan receptor alpha2->0.826	RAR-related orphan receptor alpha2->0.829		
PRKDC	Ex1.-564	C->T transition	C	57				



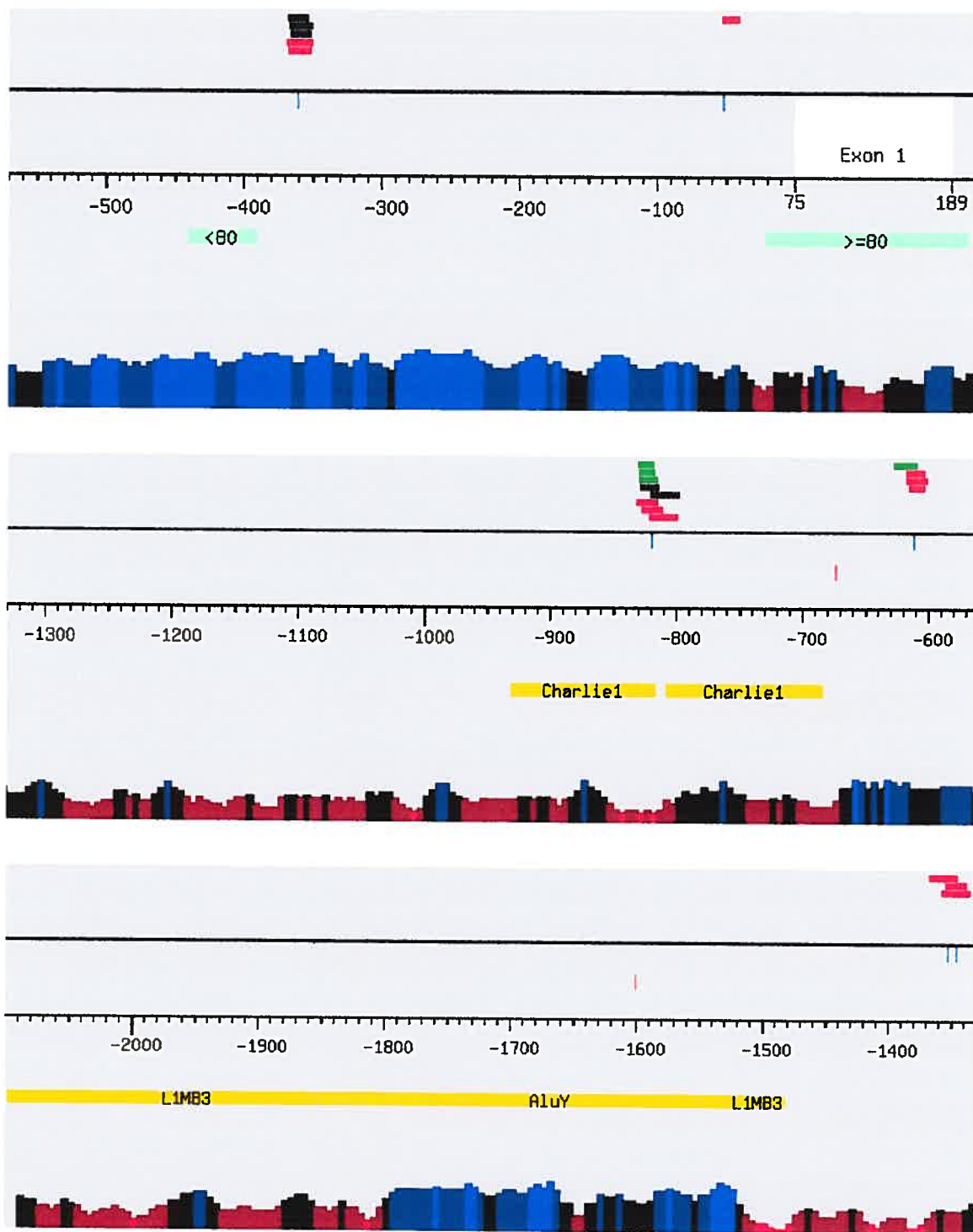


Figure 40. Représentation graphique de la région régulatrice du gène RAD50.

Tableau XXXV. Description détaillée des polymorphismes identifiés dans le gène RAD50

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
RAD50	Ex1.-52	A->G transition	A	51		zinc finger with interaction domain->0.884		
RAD50	Ex1.-820	DEL A nd	A	25	Bright, B cell regulator of IgH transcription->0.947 Homeodomain proteins MSX-1 and MSX-2->0.989 myocyte enhancer factor->0.832 octamer-binding factor 1->0.811 POU-factor Tst-1/Oct-6->0.929	Bright, B cell regulator of IgH transcription->0.949 Homeobox protein engrailed (en-1)->0.822 Homeobox protein engrailed (en-1)->0.887 Homeodomain proteins MSX-1 and MSX-2->0.989 myocyte enhancer factor->0.832 POU-factor Tst-1/Oct-6->0.905		Charlie1
RAD50	Ex1.-612	C->A transversion	C	57	signal transducer and activator of transcription 1->0.804	Alpha (1)-fetoprotein transcription factor (FTF), liver receptor homologue-1 (LHR-1)->0.944 COUP antagonizes HNF-4 by binding site competition or synergizes by direct protein - protein interaction with HNF-4->0.87 Hepatic nuclear factor 4->0.773 signal transducer and activator of transcription 1->0.804		
RAD50	Ex1.-361	C->T transition	C	76	Hypoxia inducible factor, bHLH / PAS protein family->0.964 upstream stimulating factor->0.997 X-box-binding protein 1->0.917	c-Myc/Max heterodimer->0.93 MIT (microphthalmia transcription factor) and TFE3->0.888		
RAD50	Ex1.-1347	INS T nd	T	37		myocyte enhancer factor->0.967		
RAD50	Ex1.-1353	INS A nd	A	35		Elk-1->0.954 myocyte enhancer factor->0.967		

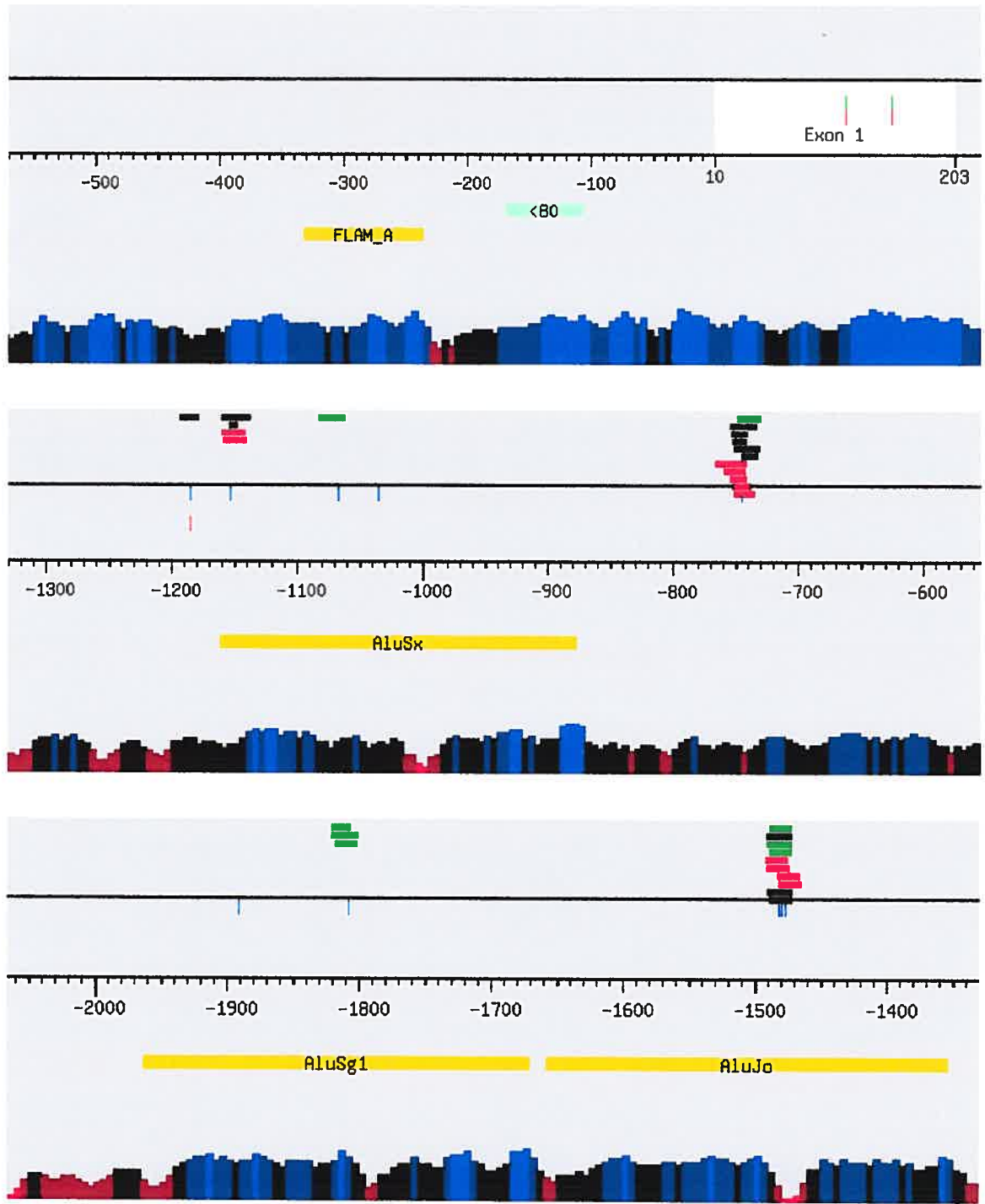


Figure 41. Représentation graphique de la région régulatrice du gène RAD51.

Tableau XXXVI. Description détaillée des polymorphismes identifiés dans le gène RAD51

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
RAD51	Ex1.-745	C->G  transversion	C	49	COMP1, cooperates with myogenic proteins in multicomponent complex->0.762  heat shock factor 1->0.97  octamer-binding factor 1, POU-specific domain->0.931  Pax-3 paired domain protein, expressed in embryogenesis, mutations correlate to Waardenburg Syndrome->0.796  SEF1 binding site->0.748  Tax/CREB complex->0.758	cut-like homeodomain protein->0.806  Mammalian C-type LTR CCAAT box->0.795  nuclear factor Y (Y-box binding factor)->0.927  octamer-binding factor 1, POU-specific domain->0.904  RAR-related orphan receptor alpha2->0.902  SEF1 binding site->0.731		
RAD51	Ex1.-1035	C->G  transversion	C	47				AluSx
RAD51	Ex1.-1067	T->C  transition	T	53	NF-E2 p45->0.884	NF-E2 p45->0.884		AluSx
RAD51	Ex1.-1153	T->G  transversion	T	55	myogenic enhancer factor 2->0.747  TG-interacting factor belonging to TALE class of homeodomain factors->1	C/EBP binding site->0.863  MIBP-1 / RFX1 complex->0.81		AluSx
RAD51	Ex1.-1185	A->T  transversion	A	47	Ras-responsive element binding protein 1->0.806			
RAD51	Ex1.-1482	A->T  transversion	A	37	Cdx-2 mammalian caudal related intestinal transcr. factor->0.86  Ecotropic viral integration site 1 encoded factor->0.834	Ecotropic viral integration site 1 encoded factor->0.819		AluJo
RAD51	Ex1.-1480	T->A	T	37	Cdx-2 mammalian caudal related intestinal transcr. factor->0.86	Cdx-2 mammalian caudal related intestinal transcr. factor->0.852		AluJo

		transversion			Ecotropic viral integration site 1 encoded factor->0.834	cellular and viral TATA box elements->0.949  Ecotropic viral integration site 1 encoded factor->0.806  Hepatocyte Nuclear Factor 3beta->0.965  POU factor Brn-2 (N-Oct 3)->0.926  Special AT-rich sequence-binding protein 1, predominantly expressed in thymocytes, binds to matrix attachment regions (MARs)->0.957	
<u>RAD51</u>	<u>Ex1.-1808</u>	C->T  transition	C	45	early growth response gene 3 product->0.783  Epstein-Barr virus transcription factor R->0.827  Member of b-zip family, induced by ER damage/stress, binds to the ERSE in association with NF-Y->0.945	early growth response gene 3 product->0.784  Epstein-Barr virus transcription factor R->0.82  Member of b-zip family, induced by ER damage/stress, binds to the ERSE in association with NF-Y->0.938	AluSg1
<u>RAD51</u>	<u>Ex1.-1477</u>	A->{T,DEL_A}  transversion	A	37	Cdx-2 mammalian caudal related intestinal transcr. factor->0.86  Ecotropic viral integration site 1 encoded factor->0.834		AluJo
<u>RAD51</u>	<u>Ex1.-1891</u>	A->C  transversion	A	61			AluSg1

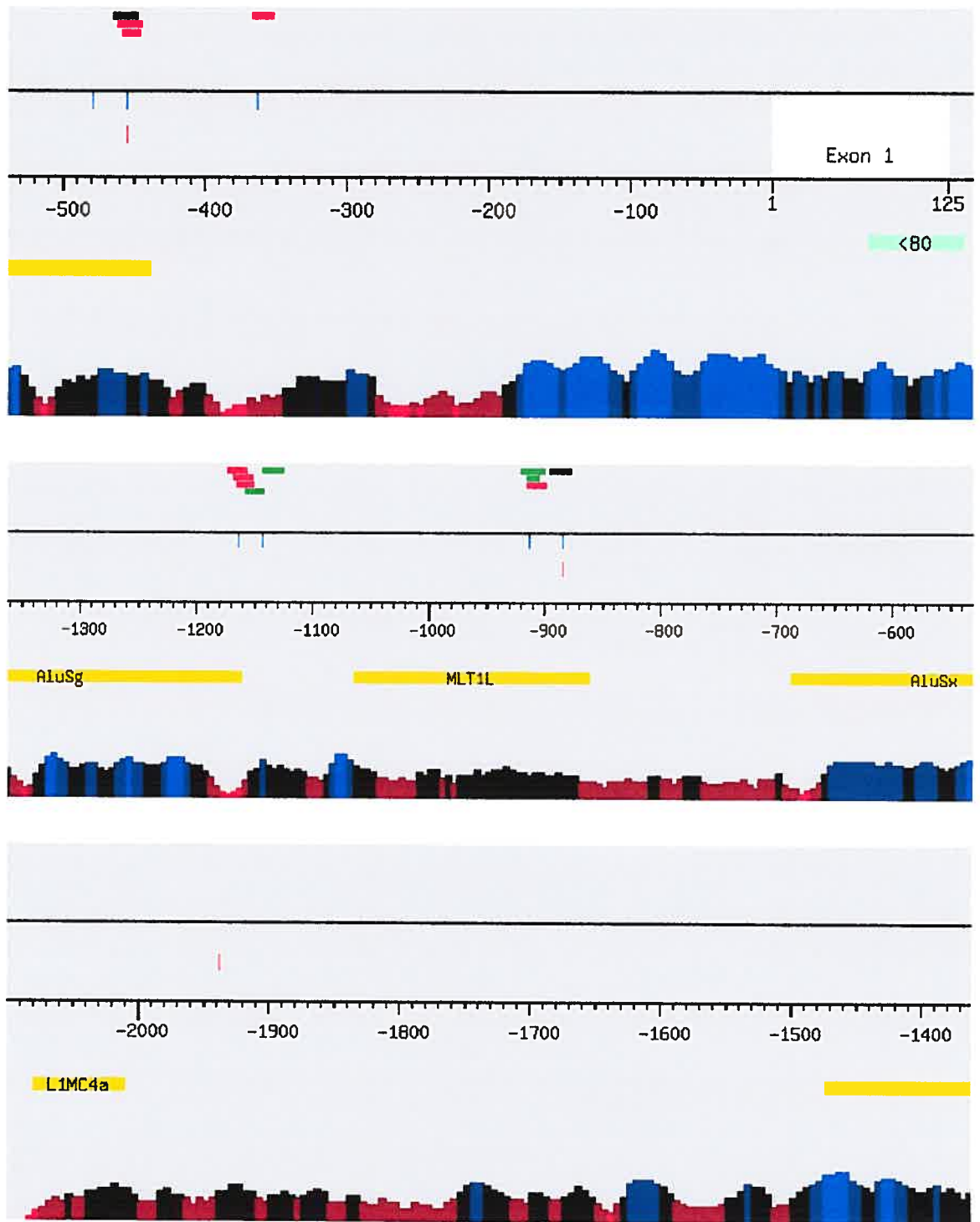


Figure 42. Représentation graphique de la région régulatrice du gène XRCC2.

Tableau XXXVII. Description détaillée des polymorphismes identifiés dans le gène XRCC2

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
XRCC2	Ex1.-1164	G->C transversion	G	25		barbiturate-inducible element->0.892 HNF-3/Fkh Homolog 2->0.938 sex-determining region Y gene product->0.953		AluSg
XRCC2	Ex1.-1143	A->G transition	A	45	Pax-6 paired domain protein->0.794 RAR-related orphan receptor alpha2->0.822	Pax-6 paired domain protein->0.795 RAR-related orphan receptor alpha2->0.822		
XRCC2	Ex1.-913	G->T transversion	G	47	cAMP-responsive element binding protein 1->0.811 PAR-type chicken vitellogenin promoter-binding protein->0.886	Avian C-type LTR TATA box->0.858 cAMP-responsive element binding protein 1->0.879 PAR-type chicken vitellogenin promoter-binding protein->0.899		MLTIL
XRCC2	Ex1.-884	G->C transversion	G	35	MIBP-1 / RFX1 complex->0.779			MLTIL
XRCC2	Ex1.-479	G->A transition	G	57				AluSx
XRCC2	Ex1.-455	C->A transversion	C	55	c-Ets-1(p54)->0.947	homeo domain factor Nkx-2.5/Csx, tinman homolog, high affinity sites->1 Homeobox protein engrailed (en-1)->0.827		AluSx
XRCC2	Ex1.-364	INS TTCT nd	TTCT	24		interferon regulatory factor 1->0.872		

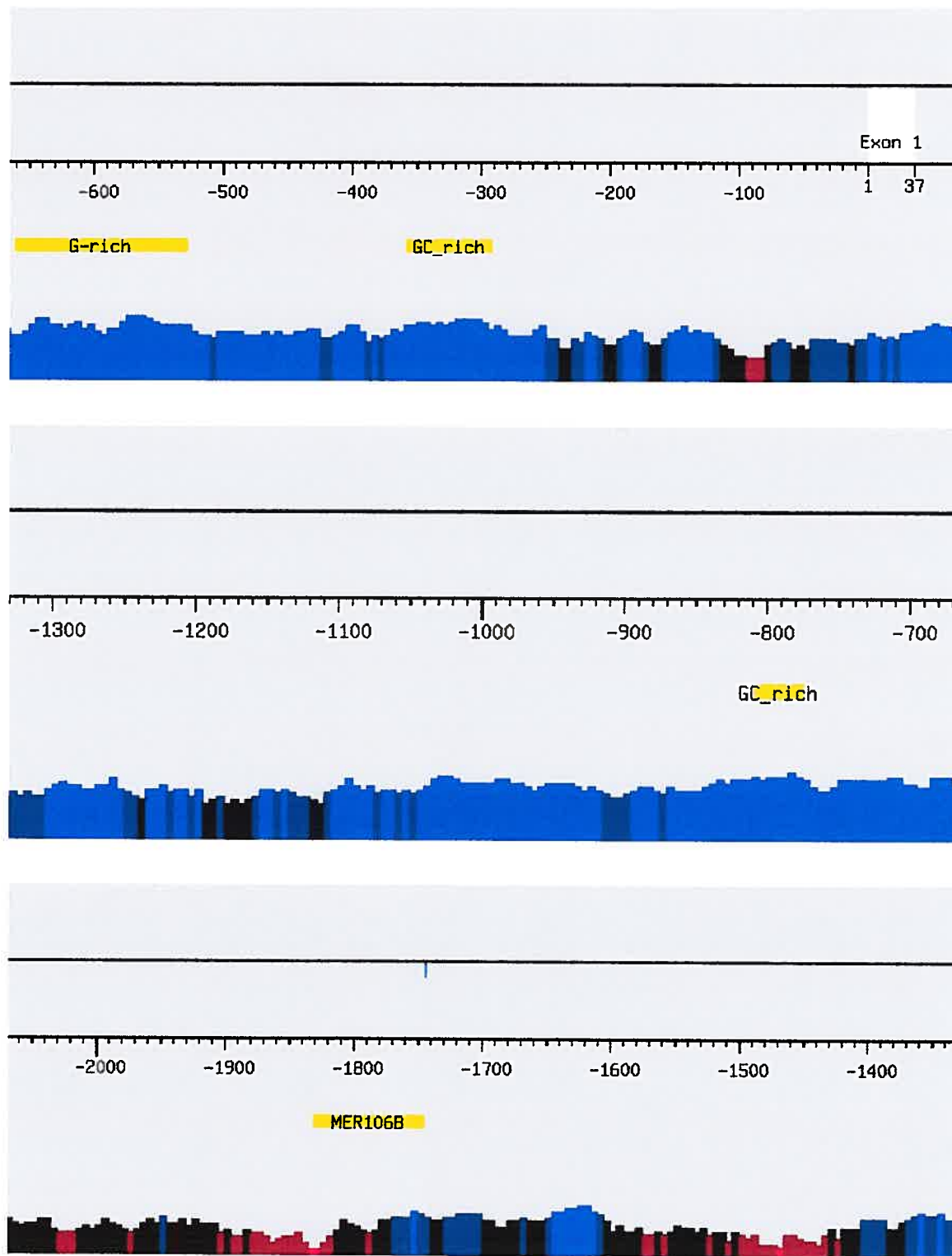


Figure 43. Représentation graphique de la région régulatrice du gène XRCC3.



Tableau XXXVIII. Description détaillée des polymorphismes identifiés dans le gène XRCC3

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
XRCC3	Ex3.-381	G->A transition	G	73	cut-like homeodomain protein->0.958  Farnesoid X - activated receptor (RXR/FXR dimer)->0.832  zebrafish PAX9 binding sites->0.787	cut-like homeodomain protein->0.77  cut-like homeodomain protein->0.958  Farnesoid X - activated receptor (RXR/FXR dimer)->0.832  Muscle Initiator Sequence->0.87		
XRCC3	Ex3.-281	T->C transition	T	61	zebrafish PAX9 binding sites->0.782	zebrafish PAX9 binding sites->0.801		
XRCC3	Ex1.-1744	G->C transversion	G	61				

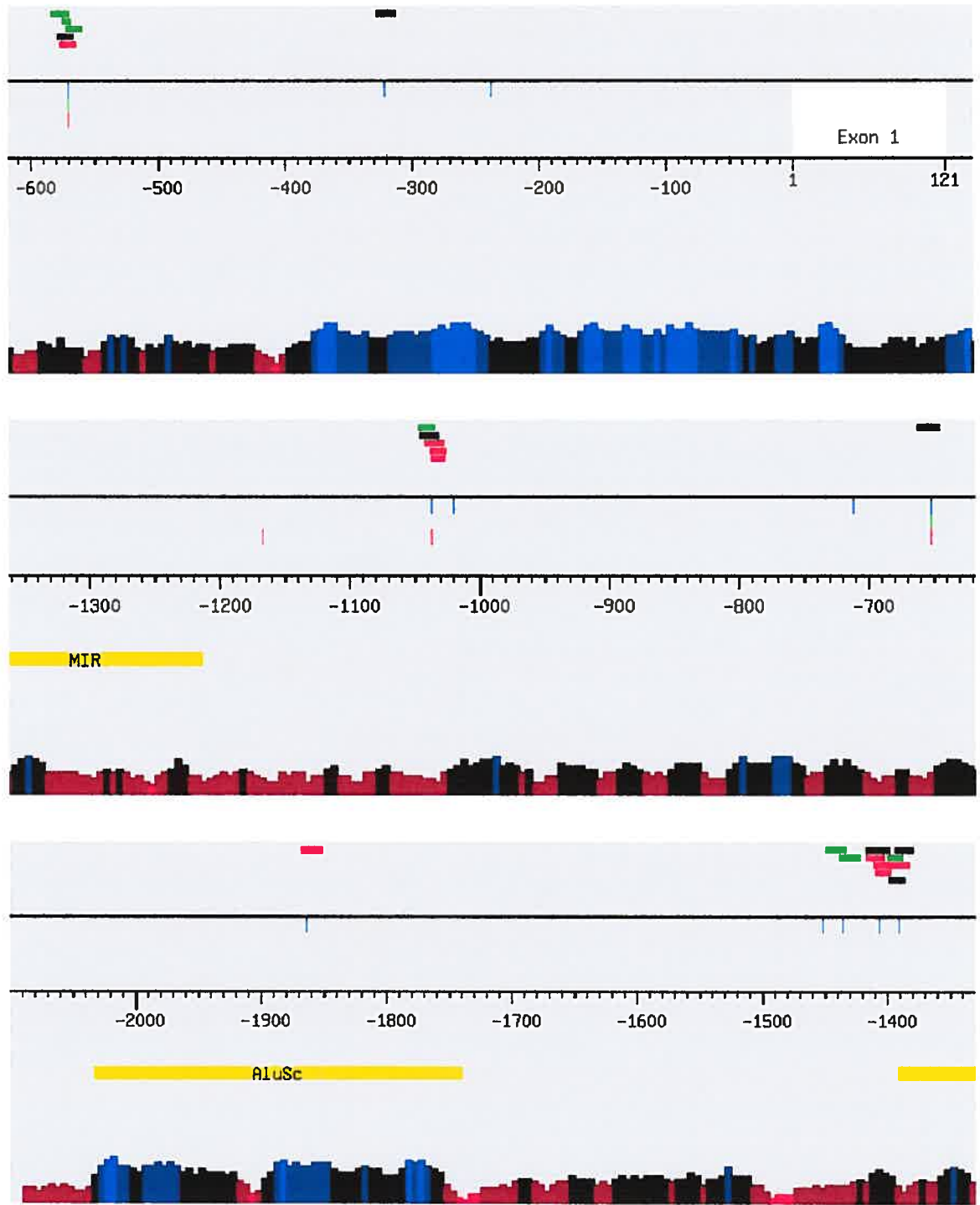


Figure 44. Représentation graphique de la région régulatrice du gène XRCC4.

Tableau XXXIX. Description détaillée des polymorphismes identifiés dans le gène XRCC4

Gène	Position	Type de polymorphisme	Base de référence	Pourcentage GC	Sites de liaison allèle majeur	Sites de liaison allèle mineur	Région conservée	Répétition
XRCC4	Ex2.-1032	A->G transition	A	45	cut-like homeodomain protein->0.754  Ecotropic viral integration site 1 encoded factor->0.998  GATA-binding factor 1->0.997 GATA-binding factor 1->0.982 Runt-related transcription factor 2 / CBFA1 (core-binding factor, runt domain, alpha subunit 1)->0.851	cut-like homeodomain protein->0.754  Ecotropic viral integration site 1 encoded factor->0.885  GATA-binding factor 1->0.997		
XRCC4	Ex1.-1020	T->G transversion	T	39				
XRCC4	Ex1.-1037	A->T transversion	A	35	homeo domain factor Nkx-2.5/Csx, tinman homolog low affinity sites->0.971  octamer-binding factor 1->0.826	GATA-binding factor 3->0.942  homeo domain factor Nkx-2.5/Csx, tinman homolog low affinity sites->0.971  octamer-binding factor 1->0.851 PAR-type chicken vitellogenin promoter-binding protein->0.873		
XRCC4	Ex1.-1391	A->C transversion	A	39	MyT1 zinc finger transcription factor involved in primary neurogenesis->0.888  octamer-binding factor 1->0.935  prostate-specific homeodomain protein NKX3.1->0.871	barbiturate-inducible element->0.904  MyT1 zinc finger transcription factor involved in primary neurogenesis->0.892		MIR
XRCC4	Ex1.-1436	A->C	A	31	Ecotropic viral integration site 1 encoded factor->0.875	Ecotropic viral integration site 1 encoded factor->0.875		

		transversion			transcriptional repressor CDP->0.938	transcriptional repressor CDP->0.938	
<u>XRCC4</u>	<u>Ex1.-1864</u>	T->C transition	T	59		Activator protein 4->0.976	AluSc
<u>XRCC4</u>	<u>Ex1.-1452</u>	T->C transition	T	29			
<u>XRCC4</u>	<u>Ex1.-1407</u>	G->C transversion	G	39	signal transducer and activator of transcription 3->0.757	AREB6 (Atp1a1 regulatory element binding factor 6)->0.975  complex of Lmo2 bound to Tal-1, E2A proteins, and GATA-1, half-site 1->0.983  E2F, involved in cell cycle regulation, interacts with Rb p107 protein->0.849	
<u>XRCC4</u>	<u>Ex1.-322</u>	G->C transversion	G	63	Runt-related transcription factor 2 / CBFA1 (core-binding factor, runt domain, alpha subunit 1)->0.852		
<u>XRCC4</u>	<u>Ex1.-238</u>	T->C transition	T	61			
<u>XRCC4</u>	<u>Ex1.-712</u>	G->A transition	G	39			
<u>XRCC4</u>	<u>Ex1.-652</u>	G->T transversion	G	43	Ecotropic viral integration site 1 encoded factor->0.787		
<u>XRCC4</u>	<u>Ex1.-571</u>	C->T transition	C	41	Homeobox protein MEIS1 binding site->0.828  MZF1->0.99  NF-kappaB->0.863  Zinc finger transcription factor ZBP-89->0.939	Homeobox protein MEIS1 binding site->0.83  Myc associated zinc finger protein (MAZ)->0.908  MZF1->0.985  NF-kappaB->0.863	

## CURRICULUM VITAE

Caroline Meloche

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

### Formation

- 2001 à aujourd'hui **M.Sc.** Biologie Moléculaire. Université de Montréal, Date prévue d'obtention mai 2003. Directeur : Daniel Sinnett. Titre : Détection de polymorphismes au niveau des régions régulatrices de gènes impliqués dans la réparation de l'ADN
- 1998-2001 **B.Sc.** Biologie Médicale, avec mention spéciale. Université du Québec à Trois-Rivières, mai 2001.

### Expérience

- Janvier 2001- Mai 2001 **Stagiaire de recherche** dans le cadre de mon B.Sc. dans le laboratoire du Dr. D. Sinnett au centre de recherche de l'hôpital Ste-Justine. Titre de mon projet : Polymorphismes dans les gènes XRCC1, XRCC3, hOGG1 et MnSOD et susceptibilité à la leucémie lymphoblastique aiguë.
- Mai 2000- Août 2000 **Technicienne de laboratoire au contrôle de la qualité**, Les Breuvages Cott. Effectuer des analyses microbiologiques des produits bruts et finis et analyser la qualité des produits finis.
- Mai 1999- Août 1999 **Préposée à la réception et à l'expédition**, Burns Philp Food Limited. Planifier l'horaire de l'expédition, confirmer les commandes, entrer la production de la journée, superviser les chauffeurs de chariots-élévateurs.
- Mai 1998- Août 1998 **Commis de bureau**, Burns Philp Food Limited. Entrer la production, classement de dossiers, confirmer les commandes.
- Mai 1997- Août 1997 **Télémarketing**, Les entreprises Gilles Sénécal. Vente de revues par téléphone.

## Bourses et prix

**2001-2003**

**CRNSG ESA (M.Sc.)** Lieu d'utilisation : Université de Montréal

**Juin 2002**

**Deuxième prix présentation orale par un étudiant à la maîtrise**, au Congrès annuel des Étudiants du Centre de Recherche de l'Hôpital Ste-Justine, Montréal

## Publications

Mathonnet, G., D. Labuda, C. Meloche et D. Sinnett (2002) Polymorphisms in the coding region of DNA-repair genes. *Mutat Res.* (soumis)

Meloche, C., J. Simard et D. Sinnett (2003) Role of ELAC2 and variants in the susceptibility to childhood acute lymphoblastic leukemia. (en preparation).

Simard, J. M. Dumont, D. Labuda, D. Sinnett, C. Meloche, M. El-alfy, L. Berger, E. Lees, F. Labrie et SV Tavtigian (2003) Prostate cancer susceptibility genes : lessons learned and challenges posed. (accepté par Endocrine-related cancer)

## Communications

Meloche, C., S. Drouin, D. Labuda et D. Sinnett. Détection de rSNPs dans des gènes de réparation de l'ADN, IXe Journée scientifique des programmes de biologie moléculaire, IRCM, 25 avril 2003. (affiche)

Meloche, C. et D. Sinnett. Polymorphismes dans les gènes XRCC1, XRCC3, hOGG1 et MnSOD et susceptibilité à la leucémie chez l'enfant, XVII<sup>e</sup> Congrès annuel de la recherche des étudiants gradués du Centre de recherche de l'Hôpital Ste-Justine, Montréal, 7 juin 2002. (oral)

Pastinen, T., R. Sladek, A. Beck, A. Sammak, B. Ge, S. Gurd, A. Verner, J. Kingsley, K. Dewar, C. Greenwood, E. Harmsen, D. Labuda, P. Lepage, C. Meloche, K. Morgan, A. Peterson, D. Sinnett, M-C. Vohl, A.D. Long, T.J. Hudson. Regulatory polymorphism discovery using allelic imbalance detection, Cold Spring Harbor Laboratory, Mai 2002. (affiche)

Meloche, C. et D. Sinnett. Role of XRCC1, XRCC3 and hOGG1 genetic polymorphisms in the susceptibility to childhood acute lymphoblastic leukemia, American Association for Cancer Research 2002, San Francisco, Avril 2002 (affiche)

Meloche, C. et D. Sinnett. Polymorphismes dans les gènes XRCC1, XRCC3, hOGG1 et MnSOD et susceptibilité à la leucémie chez l'enfant, III<sup>e</sup> Congrès des stagiaires de recherche de la Faculté de médecine et des centres hospitaliers de l'université de Montréal, Montréal, 26 janvier 2002 (oral)

### **Autres activités**

#### **2002 à aujourd'hui**

**Représentante du service d'héματο-oncologie**, association des étudiants du centre de recherche de l'Hôpital Ste-Justine

**Participation au projet SEUR**, Université de Montréal, accueillir des étudiants du secondaire pour leur faire découvrir la carrière en recherche

### **Langues**

Français et anglais

Handwritten text at the bottom right of the page, possibly a signature or date.