

Université de Montréal

Impacts fonctionnels des polymorphismes dans les promoteurs
des gènes de l'apoptose

par
Marie-Eve Lalonde

Département de biochimie
Faculté de médecine

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de maîtrise en biochimie

Avril, 2009

Copyright, Marie-Eve Lalonde, 2009

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :
Impacts fonctionnels des polymorphismes dans les promoteurs
des gènes de l'apoptose

Présenté par :
Marie-Eve Lalonde

a été évalué par un jury composé des personnes suivantes :

Normand Brisson
Président-rapporteur

Daniel Sinnett
Directeur de recherche

Estelle Schmitt
Membre du jury

Résumé

La susceptibilité ou la résistance aux cancers peuvent impliquer plusieurs mécanismes, incluant l'apoptose, la croissance cellulaire et la différenciation, la réplication et la réparation de l'ADN. Mon projet porte plus particulièrement sur l'apoptose. Une dérégulation dans les voies d'activation de l'apoptose entraîne une accumulation de cellules dérégulées, créant ainsi un environnement propice à l'instabilité génétique et au développement du cancer. Comme l'apoptose est une voie biologique hautement régulée, nous proposons l'hypothèse que des polymorphismes « fonctionnels » dans les régions de régulations des gènes (rSNPs) perturberaient cette voie à cause de taux variables de transcrits et des protéines correspondantes dû à la modification des sites de reconnaissances des facteurs de transcription. Les principaux objectifs de mon projet sont : (i) identifier les SNPs présents dans la région promotrice des gènes d'apoptose; (ii) déterminer les haplotypes de promoteurs les plus fréquents présents dans la population générale; (rHaps) (iii) vérifier leurs impacts fonctionnels sur l'expression génique par des essais *in vitro* (gène rapporteur et retard sur gel). Cette étude permettra d'identifier des rSNPs et rHaps ayant un impact sur le niveau d'expression des gènes d'apoptose, au moins dans un contexte *in vitro*. Ces différences alléliques au niveau de l'expression de ces gènes d'apoptose pourraient contribuer à la susceptibilité interindividuelle de développer un cancer.

Mots clés : Polymorphismes, promoteurs, apoptose, impacts fonctionnels, cancer.

Abstract

Cancer susceptibility can involve many different mechanisms, including apoptosis, cell growth, cell differentiation, DNA replication and DNA repair. My project focuses on the apoptosis pathway. Decreased apoptosis level can lead to the accumulation of dysregulated cells, creating a favourable environment for genetic instability and tumorigenesis. Since apoptosis is a highly regulated pathway, the hypothesis of this project is that functional polymorphisms (rSNPs) in the regulatory regions of apoptosis genes such as promoters could create or disrupt transcription factor binding sites and modify their transcription levels. The main objectives of my project are: (i) Identify rSNPs in promoter regions of apoptosis genes; (ii) Calculate the frequent promoter haplotypes (rHaps) in general population; (iii) Validate their functional impact on gene expression with *in vitro* assays (gene reporter and gel shift). This study will allow identification of rSNPs and rHaps that could influence apoptosis gene expression levels, at least in the *in vitro* context. These allelic differences of expression in apoptosis genes could contribute to interindividual cancer susceptibility.

Key Words: Polymorphisms, promoters, apoptosis, functional impact, cancer.

Table des Matières

RÉSUMÉ	III
ABSTRACT	IV
TABLE DES MATIÈRES	V
LISTE DES TABLEAUX	VII
LISTE DES FIGURES	VIII
LISTE DES SIGLES ET ABRÉVIATIONS	IX
REMERCIEMENTS	XI
INTRODUCTION	1
Apoptose	1
Caspases	1
La voie extrinsèque	3
La voie intrinsèque	3
Autres voies de l'apoptose	6
Dérégulation de l'apoptose et cancer	7
Variabilité génétique	11
Catégories de polymorphismes	11
Variation du nombre de copies (Copy Number Variations)	12
SNPs régulateurs	13
Maladies complexes et SNPs régulateurs	14
SNPs régulateurs dans les gènes d'apoptose	15
Haplotypes et déséquilibre de liaison	17
Hypothèses	19
Objectifs du projet	20
MATÉRIEL ET MÉTHODES	21
Identification des polymorphismes et génotypage	21
Détermination des haplotypes fréquents	22

Étude de l'impact fonctionnel des rSNPs/rHaplotypes	23
Essai de gène rapporteur	23
Expériences de Retard sur gel	24
Identification et priorisation de nouveaux gènes candidats	26
Approche qRT-PCR	26
Approche - Réanalyse de données de Microarrays	28
RÉSULTATS	30
Préambule	30
MDM2	33
Détermination des gènes candidats	40
qRT-PCR	40
Analyse de données de microarrays de patients LLA	42
Analyse de données de microarrays d'individus contrôles	43
BCL2	45
PARP1	49
DISCUSSION	55
Rappel du projet	55
Sommaire des résultats obtenus	56
Analyse des gènes candidats	58
MDM2	59
BCL2	63
PARP1	66
Limitations de l'étude	68
Projets futurs	71
CONCLUSION	72
BIBLIOGRAPHIE	73
ANNEXES	XII

Liste des tableaux

Tableau I. Polymorphismes des régions régulatrices des gènes d'apoptose associés à diverses maladies dans la littérature.

Tableau II. Gènes d'apoptose dont l'expression est mesurée par qRT-PCR par la trousse de Super Array.

Tableau III. Résumé des résultats obtenus pour les promoteurs des 12 gènes étudiés de la voie intrinsèque d'apoptose.

Tableau IV. SNPs des promoteurs de MDM2 avec leur position génomique et leur fréquence allélique observée dans notre cohorte.

Tableau V. Haplotypes fréquents de la région promotrice de MDM2.

Tableau VI. Gènes candidats obtenus par analyse de données d'expression de qRT-PCR et de microarrays.

Tableau VII. SNPs du promoteur de BCL2 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

Tableau VIII. Haplotypes fréquents de la région promotrice de BCL2.

Tableau IX. SNPs du promoteur de PARP1 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

Tableau X. Haplotypes fréquents de la région promotrice de PARP1.

Liste des figures

Figure 1. Les différentes caspases.

Figure 2. Schéma des différentes voies d'apoptose.

Figure 3. Évènements critiques menant au développement de tous les cancers.

Figure 4. La protéine suppresseur de tumeur p53 agit à plusieurs niveaux pour bloquer la progression du cancer.

Figure 5. Types de polymorphismes, haplotypes et déséquilibre de liaison.

Figure 6. Schéma des promoteurs de MDM2 avec leurs SNPs.

Figure 7. Impact fonctionnel des haplotypes du promoteur P1 de MDM2 sur l'activité promotrice.

Figure 8. Retard sur gel montrant une liaison différentielle au niveau du rSNP - 182C>G de MDM2.

Figure 9. Schéma du promoteur de BCL2. Le gène est orienté de façon anti-sens dans le génome.

Figure 10. Impact fonctionnel des haplotypes du promoteur de BCL2 sur l'activité promotrice.

Figure 11. Retard sur gel montrant une liaison différentielle au niveau du rSNP - 225G>T de BCL2.

Figure 12. Schéma de la région promotrice de PARP1. Le gène est orienté de façon anti-sens dans le génome.

Figure 13. Impact fonctionnel des haplotypes du promoteur de PARP1 sur l'activité promotrice.

Figure 14. Retard sur gel montrant une liaison différentielle au niveau du rSNP - 1491A>G de PARP1.

Liste des sigles et abréviations

ADN	Acide désoxyribonucléique
AIF	Apoptosis Inducing Factor
APAF1	Apoptotic Protease Activating Factor-1
ARN	Acide ribonucléique
ARNm	ARN messenger
ATP	Adénosine triphosphate
BAG4	BCL2-associated athanogene 4
BAK	BCL2-antagonist/killer
BAX	BCL2-associated X protein
BCL2	B-cell CLL/lymphoma 2
BCL2L1	BCL2-like 1
BCL2L10	BCL2-like 10
BCL2L11	BCL2-like 11
Bcl-XL	BCL2-like 1
BH	Bcl-2 homology
Bid	BH3 interacting domain death agonist
BIRC4	Baculoviral IAP repeat-containing protein 4
BIRC5	Baculoviral IAP repeat-containing protein 5
bp	base pair
Ca ²⁺	Calcium
CARD	Caspase Recruitment Domains
CASP3	caspase 3
CASP9	caspase 9
cDNA	coding DNA
CNV	Copy Number Variation
EDTA	ethylenediaminetetraacetic acid
ENDO G	endonucléase G
fmol	fento mole
IAP	Inhibitor of Apoptosis
kb	kilo base
LLA	leucémie lymphoblastique aïgue
LLC	leucémie lymphoblastique chronique
MCL-1	Myeloid Cell Leukemia-1
MDM2	Mouse Double Minute 2
miRNA	micro ARN
mM	millimolaire
NaOH	Hydroxide de sodium
NF-KB	Nuclear Factor-kappa B

PARP1	Poly-ADP-ribose Polymerase
PCR	Polymerase Chain Reaction
qRT-PCR	quantitative Real Time PCR
rSNP	regulatory SNP
SIDA	syndrome de l'immunodéficience acquise
SNP	Single Nucleotide Polymorphism
t-Bid	truncated-Bid
TNF	Tumor Necrosis Factor
UPR	Unfolded Protein Response
UTR	Untranslated Region
UV	ultraviolet
VNTR	Variable Number of Tandem Repeats
XIAP	X-linked Inhibitor of Apoptosis
µg	microgramme
µL	microlitre

Remerciements

Je tiens à remercier le Dr Sinnett ainsi que tous les membres de son laboratoire et ceux du laboratoire du Dr Labuda pour leur aide. Un merci particulier à Nina N. et Mathieu L., pour vos idées, votre support et nos discussions que j'ai beaucoup appréciés, je vous en suis très reconnaissante.

Introduction

Apoptose

La mort cellulaire est un processus biologique hautement régulé et nécessaire pour le maintien de l'homéostasie cellulaire (1). Il existe deux types de mort cellulaire, l'apoptose et la nécrose. La nécrose survient suite à de grands chocs environnementaux comme des températures extrêmes et une hypoxie/ischémie sévère qui détruisent la membrane cytoplasmique de la cellule, provoquant la destruction de toute la structure cellulaire et la dilatation des organelles (2).

L'apoptose, aussi appelée mort cellulaire programmée, est un mécanisme plus ordonné et survient en réponse à des stimuli menant à une cascade de signaux intracellulaires bien régulés et conservés au cours de l'évolution (3). L'intégrité de la membrane est conservée lors de l'apoptose et c'est d'ailleurs de cette façon que l'apoptose a été identifiée, dû aux différences physiologiques observées entre les cellules apoptotiques et les cellules nécrosées (4). Lors du développement embryonnaire, l'apoptose constitue un mécanisme développemental essentiel pour la morphogenèse, alors que chez l'adulte il est plutôt perçu comme aidant à l'homéostasie des tissus et organes (5, 6).

Caspases

Les caspases sont des protéases à cystéine clivant au niveau des résidus aspartates (7). Ce sont les protéines majeures de l'apoptose, menant à la protéolyse et à la mort cellulaire. Le plus connu des substrats des caspases est PARP (poly-ADP-ribose polymerase), une enzyme de réparation de l'ADN (8). Sur les 14 caspases

identifiées jusqu'à ce jour, les caspases 1, 4, 5 et 11 à 14 ne jouent pas de rôle dans l'apoptose, mais sont plutôt impliquées au niveau de la réponse inflammatoire (1). Ces protéines sont synthétisées comme des proenzymes inactives qui peuvent s'activer par autoprotéolyse ou par protéolyse par d'autres caspases, engendrant ainsi une cascade d'activation des caspases. Ce sont les caspases ayant les plus longs pro-domaines qui sont les caspases appelées initiatrices de l'apoptose (caspases 2, 8, 9 et 10) et qui activent par protéolyse les caspases effectrices (caspases 3, 6 et 7) (9, 10) (Figure 1). Il existe principalement deux voies d'activation de l'apoptose, menant à l'activation des caspases : la voie extrinsèque et la voie intrinsèque (Figure 2).

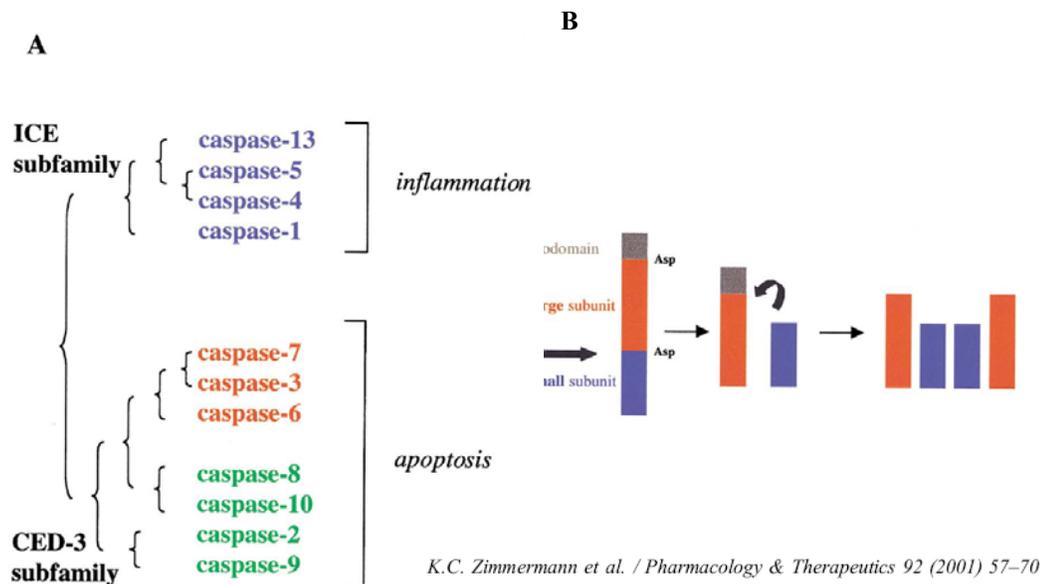


Figure 1. Les différentes caspases. (A) Les classes et sous-classes de caspases et leur rôle, (B) Activation des caspases par clivage des prodomaines.

La voie extrinsèque

La voie extrinsèque est déclenchée via des récepteurs de mort transmembranaires (*death receptors*) appartenant à la superfamille des TNF-récepteurs (*tumor necrosis factor receptor*). Les plus connus de ces récepteurs sont Fas et TNF-R1 qui sont liés respectivement par le ligand Fas (Fas-L) et le TNF- α , menant à la trimérisation de ces récepteurs (11-13). Des protéines adaptatrices sont ensuite recrutées pour activer les caspases 8 et 2 qui à leur tour activeront d'autres caspases en aval (9). Cette voie est empruntée, entre autre, pour l'élimination des cellules infectées par un virus, cellules sur lesquelles le ligand Fas des cellules T cytotoxiques va se lier pour y déclencher l'apoptose.

La voie intrinsèque

La voie intrinsèque (ou mitochondriale) survient, quant à elle, en réponse à un stress intracellulaire (dommages à l'ADN, stress oxydatif, etc.). Ce stress provoque des signaux internes dans la cellule qui déclencheront l'apoptose. Les protéines pro-apoptotiques de la famille de BCL2 telles que BAX et BAK sont activées et migrent alors vers la surface de la mitochondrie où elles permettent d'inhiber d'autres membres anti-apoptotiques de la famille des BCL-2, soit BCL2 et BCL-XL. Ces protéines pro ou anti-apoptotiques possèdent toutes de un à quatre domaines BH (*Bcl-2 homology*) (14-19). Les membres de cette famille peuvent homo ou hétéro-dimériser et c'est le ratio entre les quantités de protéines pro et anti-apoptotiques qui détermine la sensibilité des cellules à l'apoptose (20, 21). En présence d'une plus grande quantité de protéines pro-apoptotiques, il y a ouverture de pores à la surface de la mitochondrie permettant le relâchement dans le

cytoplasme du cytochrome c mitochondrial (22). Cela entraîne également le relargage de protéines telles que les caspases 3 et 9 ainsi que l'AIF (*apoptosis inducing factor*), l'endonucléase G et Smac/Diablo (23-26). Ces dernières sont responsables de l'inhibition de protéines inhibitrices de l'apoptose (IAP) dont fait partie XIAP (*X-linked inhibitor of apoptosis*). L'endonucléase G, pour sa part, est transloquée au noyau et induit la fragmentation de l'ADN (27). Une fois le cytochrome c relâché, il active APAF-1 (*apoptotic protease activating factor-1*) qui, en présence d'ATP, lie la procaspase-9 par ses domaines CARD (*caspase recruitment domain*) (28). Le complexe résultant est appelé apoptosome et peut maintenant activer les caspases effectrices 3, 6 et 7 (29).

Il existe par contre un lien entre les voies extrinsèque et intrinsèque, au niveau de la caspase 8. Cette dernière peut en effet cliver la protéine Bid, une protéine de la famille des BCL-2 (6, 30). Une fois clivée, la protéine tBid (*truncated Bid*) peut migrer à la surface de la mitochondrie et activer le déclenchement de l'apoptose de la même façon que les protéines BAX et BAK (31, 32).

Autres voies de l'apoptose

Il existe d'autres voies secondaires pouvant mener à l'apoptose. La synthèse de protéines dans le réticulum endoplasmique implique certains changements post-traductionnels et un repliement des protéines avant d'atteindre leurs conformations finales. Il arrive parfois que le repliement dans le réticulum endoplasmique ne se fasse pas de façon appropriée, ce qui active une voie de signalisation menant au rétablissement du repliement (UPR, *unfolded protein response*). Par contre, si la situation est trop critique, l'apoptose sera alors déclenchée via le Ca^{2+} stocké dans la lumière du réticulum (33, 34). C'est ce Ca^{2+} qui une fois relâché dans le cytoplasme joue sur la perméabilité de la membrane mitochondriale et permet ainsi le relâchement du cytochrome c (35). Des protéines de la famille de BCL-2 se retrouvent d'ailleurs à la surface interne de la membrane du réticulum et influencent également la perméabilité de celle-ci (35). Le Ca^{2+} peut aussi activer la caspase 12 qui est spécifiquement activée lors de l'apoptose induite par stress du réticulum endoplasmique (35, 36). Il existe également certaines voies d'activation de l'apoptose qui ne passent pas par les caspases, mais plutôt par l'activation d'autres protéases telles que les calpains, les cathepsines et les protéases à sérine comme Granzyme A et B (37, 38). Celles-ci peuvent fonctionner de pair avec les caspases pour activer l'apoptose ou activer l'apoptose de façon caspase-indépendante (39).

Dérégulation de l'apoptose et cancer

Puisque l'apoptose est un processus important et bien régulé, il est évident qu'une dérégulation de cette voie peut mener à différentes maladies. Plusieurs maladies sont liées à une apoptose trop élevée, comme le SIDA, l'Alzheimer et les syndromes myélodysplasiques. À l'opposé, certaines maladies sont liées à une diminution de l'apoptose : l'athérosclérose, certaines maladies auto-immunes et les cancers (40, 41).

Le cancer est décrit comme une prolifération excessive de cellules dérégulées. Bien que tous les cancers soient un peu différents les uns des autres, certains évènements critiques tels que la perte d'apoptose et l'accumulation de mutations sont des éléments communs à tous les cancers (Figure 3) (42). Le développement du cancer peut donc être influencé par une surexpression des gènes anti-apoptotiques ou encore par une plus faible expression des gènes pro-apoptotiques, qui mènent alors à une apoptose insuffisante et favorisent la croissance de cellules malignes (42).

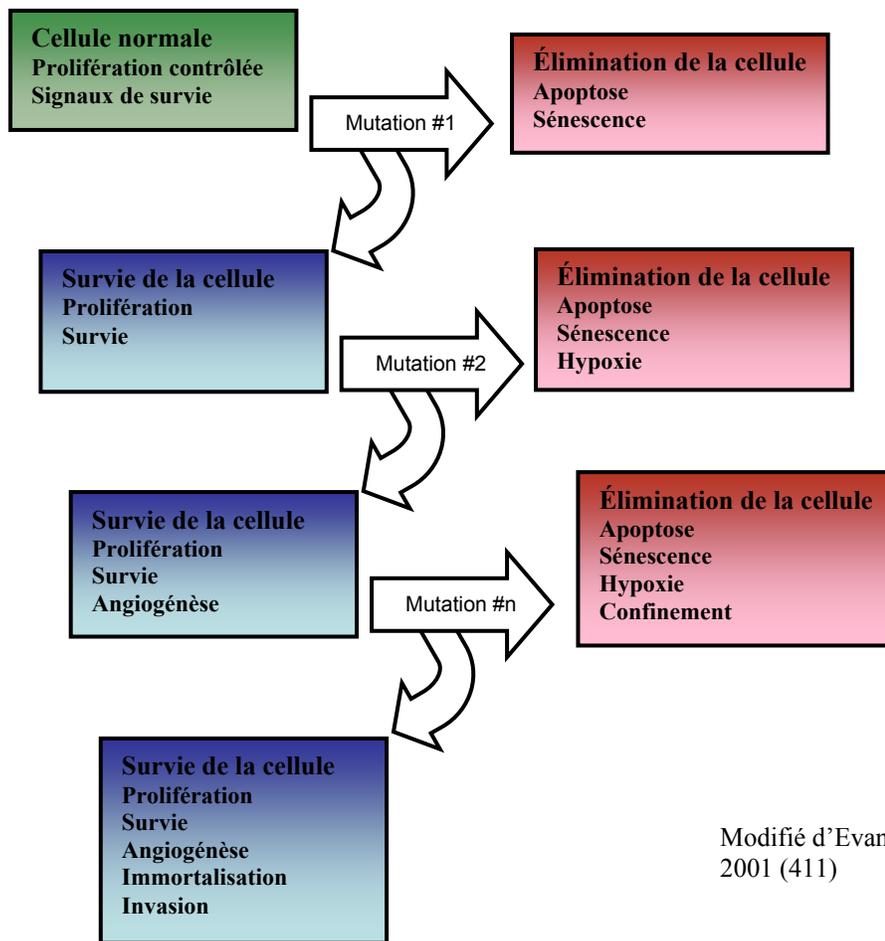


Figure 3. Évènements critiques menant au développement de tous les cancers.

L'altération des niveaux d'expression de protéines impliquées dans l'apoptose a été observée dans plusieurs types de cancers. La plupart de ces protéines se retrouvent dans la voie intrinsèque de l'apoptose, probablement parce que cette voie qui est liée aux stress intracellulaires comme les dommages à l'ADN. Une surexpression de la protéine anti-apoptotique Bcl-2 a été montrée dans plusieurs tumeurs solides, leucémies ou lymphomes, et a été corrélée avec un faible pronostic (43, 44). La surexpression de Mcl-1 et Bcl-XL (d'autres membres de la famille des BCL-2) a été démontrée dans certaines leucémies (45-47). La protéine XIAP (BIRC4) montre une surexpression dans pratiquement tous les types de

cancers, permettant ainsi de court-circuiter le point de contrôle G₂/M (48-51). Les protéines pro-apoptotiques BAX et BAK ont montré une faible expression dans plusieurs cancers dont le cancer du colon et certaines leucémies (52, 53). Un ratio élevé de BCL-2 sur BAX est d'ailleurs associé avec un faible pronostic des leucémies myéloïde et lymphoblastique aigües (54-57). Certains mélanomes ont aussi montré une perte d'expression d'APAF-1(58).

La protéine suppresseur de tumeur p53 peut être activée en réponse à différents stress cellulaires tels que les dommages à l'ADN, l'hypoxie ou la prolifération excessive (Figure 4). Ce facteur de transcription a un rôle central dans plusieurs processus tel que l'apoptose, ce qui en fait une protéine très importante dans l'oncogenèse. Cette protéine régule la transcription de plusieurs gènes d'arrêt du cycle cellulaire (p21, Gadd25) et de gènes pro-apoptotiques (CASP9, BAX, APAF-1) (59, 60). La dérégulation de p53 peut donc avoir des conséquences catastrophiques pour la cellule. D'ailleurs, plus de 50% des tumeurs solides montrent une dérégulation de p53 (61, 62). Certaines protéines qui régulent le niveau d'expression de p53 ont également démontré des dérégulations dans certains cancers. C'est le cas de MDM2 (*mouse double minute 2 homolog*), une ubiquitine ligase directe de p53 qui est surexprimée dans la leucémie lymphoblastique aigüe (63, 64).

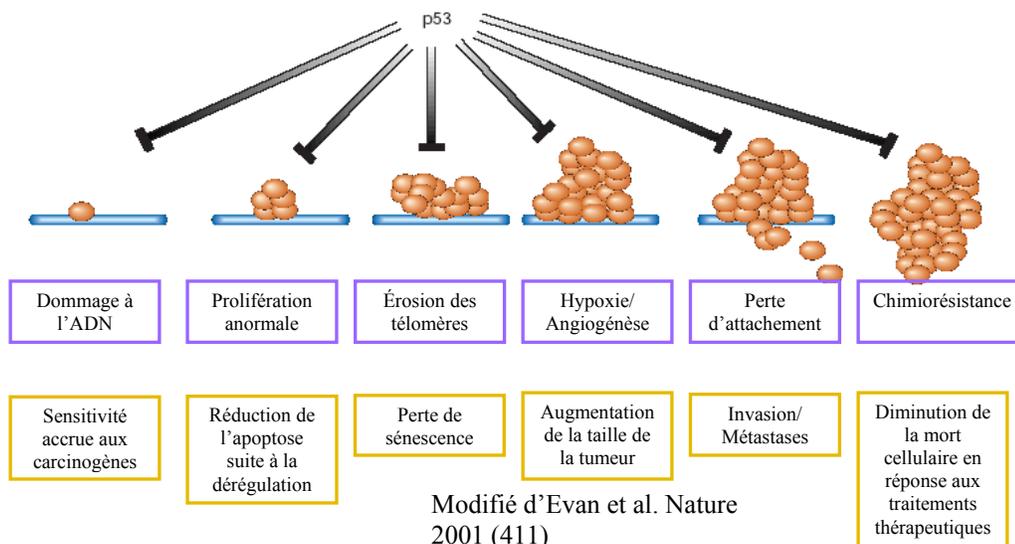


Figure 4. La protéine suppresseur de tumeur p53 agit à plusieurs niveaux pour bloquer la progression du cancer.

Plusieurs médicaments anti-cancer ciblant la voie d'apoptose ont été développés. Également, plusieurs inhibiteurs des protéines anti-apoptotiques de la famille des BCL-2 ont été développés dont un oligonucléotide anti-sens de BCL2 qui peut cibler plusieurs types de cancer (65-71). Des inhibiteurs du facteur de transcription NF- κ B, qui normalement induit la transcription de gènes anti-apoptotiques ont également été développés (72-76).

On observe que des patients montrent une résistance à certains médicaments anti-cancer due justement à une dérégulation au niveau des gènes impliqués dans la régulation de l'apoptose. Par exemple, la résistance au dexaméthasone, un glucocorticoïde utilisé dans le traitement du cancer, observée chez certains patients a été corrélée avec une surexpression de BCL-2 ou BCL-XL dans les cellules tumorales (77-80).

Variabilité génétique

Catégories de polymorphismes

Le génome humain est identique chez tous les individus à près de 99% (81). Il reste donc 1% de la séquence génomique qui diffère entre les individus, d'où la notion de variabilité génétique. On définit souvent les polymorphismes comme étant des variations présentes dans l'ADN avec une fréquence allélique $\geq 1\%$ dans la population (82). Il existe plusieurs types de polymorphismes. Les plus connus sont les SNPs (*single nucleotide polymorphisms*) où un seul nucléotide est affecté. Ce nucléotide peut être supprimé, inséré ou substitué par n'importe quel des trois autres nucléotides. Ces substitutions sont assez fréquentes puisqu'elles surviennent en moyenne 1 fois par 800 nucléotides dans le génome complet (82). Il existerait d'ailleurs plus de 12 millions de SNPs répertoriés à ce jour chez l'humain (83).

D'autres types de polymorphismes, les VNTR (*variable number of tandem repeats*), sont des répétitions variables de séquences de longueurs variables. Ces VNTR sont appelées microsatellites lorsqu'elles sont de 1 à 6 pb, alors que les minisatellites sont des répétitions un peu plus longues (environ 10 pb) (83). Les microsatellites ne sont généralement pas répétés plus d'une vingtaine de fois et sont localisés un peu partout dans le génome (84). Les minisatellites peuvent pour leur part présenter des centaines de répétition étalées sur de grandes régions (20 kb). Ceux-ci sont généralement riches en GC et surtout localisés dans les régions sub-téломériques chez l'humain (85, 86). Pour ces polymorphismes, c'est le nombre de répétitions qui varie entre les individus et qui est transmis de façon

mendélienne. Il existe par contre une exception de transmission, puisque certains microsatellites sont constitués de trinuécléotides en tandem qui peuvent parfois créer une séquence instable dans le génome. Cette instabilité peut mener à des modifications chez les descendants, avec extension ou délétion de ces trinuécléotides en tandem. Un exemple bien connu est celui de la maladie de l'X fragile, où le motif CGG qui est répété de 6 à 54 fois chez un individu normal est répété jusqu'à 1000 fois chez les patients (87). Ces extensions modifient non seulement la séquence du génome, mais également la structure de la chromatine, menant ainsi à l'inactivation de gènes voisins (87).

Variation du nombre de copies (Copy Number Variations)

Une autre grande source de variabilité génétique est le polymorphisme structural comme les CNVs (*copy number variations*) (88, 89). La définition actuelle des CNVs est qu'il s'agit de répétitions de régions de plus de 1 kb et allant jusqu'à quelques mégabases (83). Bien que ceux-ci soient beaucoup moins fréquents que les SNPs, ils affectent une région beaucoup plus grande dans l'ADN. Il a d'ailleurs été démontré que plus de 12% du génome humain serait en fait constitué de CNVs (81). Ils représentent donc à eux seuls plus de 90% de la variabilité des séquences entre les individus. Les scientifiques croient que la définition actuelle des CNVs sera bientôt révisée, car ceux-ci pourraient comprendre des répétitions plus petites (83). Ces CNVs pourraient non seulement agir de façon cis comme les SNPs (i.e. sur la copie du gène présente sur le même allèle que le SNP) mais également de façon trans (83).

SNPs régulateurs

Pendant plusieurs années suivant la découverte des SNPs, seuls les SNPs localisés dans des exons codant pour une protéine étaient étudiés. Ces cSNPs (*coding SNPs*) peuvent avoir ou non un impact prédit sur la séquence en acide aminés de la protéine. On croyait alors que seuls les cSNPs qui modifiaient la séquence de la protéine avaient une importance et devaient être étudiés. En affectant la séquence de la protéine, ces SNPs peuvent altérer son activité, son repliement, sa stabilité ou sa modification post-traductionnelle (phosphorylation, etc.). Ils peuvent aussi mener à la production d'une protéine tronquée dans le cas d'un codon stop prématuré.

La plupart des SNPs se retrouvent par contre dans les séquences non codantes, soit les régions intergéniques (gSNPs), introniques (iSNPs) ou encore dans des régions régulatrices (rSNPs). Au cours de mes travaux de maîtrise, mon attention a porté plus particulièrement sur les rSNPs. Les rSNPs ne modifient pas la séquence de la protéine, mais peuvent influencer le niveau d'expression de celle-ci de plusieurs façons, dépendamment de la région dans laquelle ils se retrouvent (82). Les rSNPs les plus étudiés sont ceux des régions promotrices, qui peuvent influencer le niveau de transcription d'un gène en altérant (gain ou perte de liaison) le site de liaison à un facteur de transcription. Il peut également y avoir des rSNPs dans les régions *enhancer*, empêchant le recrutement de protéines qui replient correctement la région *enhancer* vers le promoteur. Il existe aussi des rSNPs dans les régions 5'UTR et 3'UTR. Ceux-ci peuvent agir au niveau de la traduction de l'ARNm, soit en ne permettant pas à la machinerie traductionnelle de s'y lier (ou au contraire de

se lier en créant un nouvel AUG d'initiation), soit en altérant un site de liaison de miRNA potentiel dans la région 3'UTR (ou créant un nouveau site de liaison pour un miRNA) (90, 91). Ces rSNPs dans les sites de liaisons de miRNA sont de plus en plus étudiés, car ils empêchent la dégradation de l'ARNm ciblé. Les rSNPs en 3' UTR peuvent également influencer la stabilité de l'ARNm, sans nécessairement se retrouver dans un site de liaison de miRNA (92). On peut retrouver des rSNPs aux sites d'épissage de l'ARNm, empêchant un épissage normal ou créant différents isoformes de protéines (93, 94). Il faut par contre préciser que la majorité des gènes pour lesquels une différence allélique est observée ne possèdent pas de différence d'épissage de leur ARNm. L'explication de la différence d'expression se retrouve donc principalement au niveau de la transcription du gène (92). On parlera alors d'un changement quantitatif, c'est-à-dire un effet de dosage.

Maladies complexes et SNPs régulateurs

Contrairement aux maladies mendéliennes monogéniques classiques, les maladies complexes ne sont pas dues à l'altération d'un seul gène. Ce sont des maladies multifactorielles qui impliquent des facteurs génétiques et environnementaux (92). La théorie du « *common disease/common variant* » proposée il y a plusieurs années, s'appuie sur le fait que les maladies complexes sont le résultat d'une combinaison de plusieurs allèles de prédisposition qui sont fréquentes dans la population (95-97). Dans un tel modèle, les rSNPs, en influençant le niveau d'expression génique, contribueraient à la création d'allèles de prédisposition pour certaines maladies d'où leur impact sur la susceptibilité interindividuelle de développer une maladie complexe. Les personnes ayant ces allèles ne

développeraient pas nécessairement la maladie, mais seraient plus à risque d'être affectés. Il existerait d'ailleurs une grande quantité de gènes présentant une différence d'expression allèle-spécifique. Certaines études ont rapporté qu'environ 20% des gènes présentaient une expression allèle-spécifique dans des lignées lymphoblastiques humaines, alors que d'autres études montrent plutôt une différence dans 50% des gènes (92, 98, 99).

Le cancer étant une maladie complexe, il est envisageable qu'une dérégulation dans le niveau d'expression de certains gènes clés créerait un environnement oncogénique favorable. Une étude a d'ailleurs montré qu'une petite baisse dans le niveau d'expression d'un seul gène suppresseur de tumeur prédisposerait à la tumorigenèse (100). Certains rSNPs pourraient donc influencer la susceptibilité interindividuelle d'avoir le cancer en altérant, même modestement, le niveau d'expression de certains gènes (92).

SNPs régulateurs dans les gènes d'apoptose

Puisque notre attention porte sur la voie intrinsèque de l'apoptose, il est important de répertorier les rSNPs connus pour les gènes de cette voie (Tableau I). Le gène MDM2 possède un SNP (SNP 309) situé dans sa région promotrice qui semble influencer la susceptibilité à plusieurs types de cancer (101-107). Un second SNP dans la région promotrice de MDM2, C197G, influencerait le risque de développer le cancer de la vessie (108). Le gène CASP9 possède différents rSNPs dans sa région promotrice qui augmenteraient le risque de développer le cancer du poumon (109). Le gène CASP3 possède un SNP dans sa région 5' UTR qui est

associé avec les carcinomes à cellules squameuses et le cancer du poumon (108, 110, 111). Le gène BIRC5 a un SNP dans sa région promotrice (-31C>G) qui augmente le risque de développer un cancer du poumon (112). Le gène BAX a un rSNP dans sa région promotrice associé au développement de la leucémie lymphoblastique chronique (LLC) (113, 114) alors que BCL2 possède un polymorphisme dans sa région promotrice et un autre dans sa région 3' UTR, associés à la même maladie (115, 116). Finalement, le gène APAF1 posséderait un SNP dans sa région 3'UTR qui est associé à une augmentation du risque de développer la LLC (115). Malgré ces quelques exemples tirés de la littérature, nos connaissances demeurent limitées sur la diversité génétique au niveau des promoteurs ainsi que l'impact fonctionnel de ces rSNPs sur l'activité des promoteurs des gènes de la voie intrinsèque d'activation de l'apoptose.

Tableau I. Polymorphismes des régions régulatrices des gènes d'apoptose associés à diverses maladies dans la littérature.

Gène	# rs	Région régulatrice	Références*
MDM2	rs2279744	Promotrice	Bond et <i>al</i>
MDM2	rs937282	Promotrice	Wang et <i>al</i>
CASP9	rs4645978	Promotrice	Park et <i>al</i>
CASP9	rs4645981	Promotrice	Park et <i>al</i>
CASP3	rs4647603	5'UTR	Jang et <i>al</i>
BIRC5	rs17884799	Promotrice	Jang et <i>al</i>
BAX	rs4645878	Promotrice	Saxena et <i>al</i>
BCL2	rs2279115	Promotrice	Nuckel et <i>al</i>
BCL2	rs4987853	3' UTR	Enjuanes et <i>al</i>
APAF1	rs17028658	3' UTR	Enjuanes et <i>al</i>

* Voir texte ci-haut pour références détaillées ou autres références.

Haplotypes et déséquilibre de liaison

Puisque le nombre de SNPs dans le génome est très élevé, il est presque impensable de vouloir tester l'impact fonctionnel de chacun sur l'expression génique. On peut se servir de la notion de déséquilibre de liaison qui est une mesure de la coexistence non aléatoire des allèles de deux loci. Deux allèles en déséquilibre de liaison sont trouvés ensemble sur le même haplotype plus souvent que ce qui peut être prévu par le hasard (83). Un déséquilibre de liaison total survient lorsque deux allèles de deux loci coségrègent toujours, n'ayant jamais subi de recombinaison entre ces deux loci. Puisque le taux de recombinaison dans le génome est assez faible (environ 1 par 100 Mb par génération), les allèles associées à un phénotype sont souvent en déséquilibre de liaison avec quelques SNPs à proximité, permettant l'utilisation de « tagging SNPs » (117). Le déséquilibre de liaison crée donc des « blocs » d'haplotypes dans le génome qui regroupent les régions n'ayant pas subi de recombinaison (Figure 5). Il est donc important dans une étude d'association, de considérer la région étudiée comme un haplotype (un regroupement de SNPs), plutôt que d'étudier l'impact individuel de chacun des SNPs. Cela permet de se rapprocher davantage du contexte génomique en étudiant l'impact que peut avoir un ensemble de SNPs sur l'expression génique. Il a d'ailleurs été démontré qu'un SNP pouvait jouer un rôle différent dans un contexte haplotypique comparativement à un contexte individuel (81).

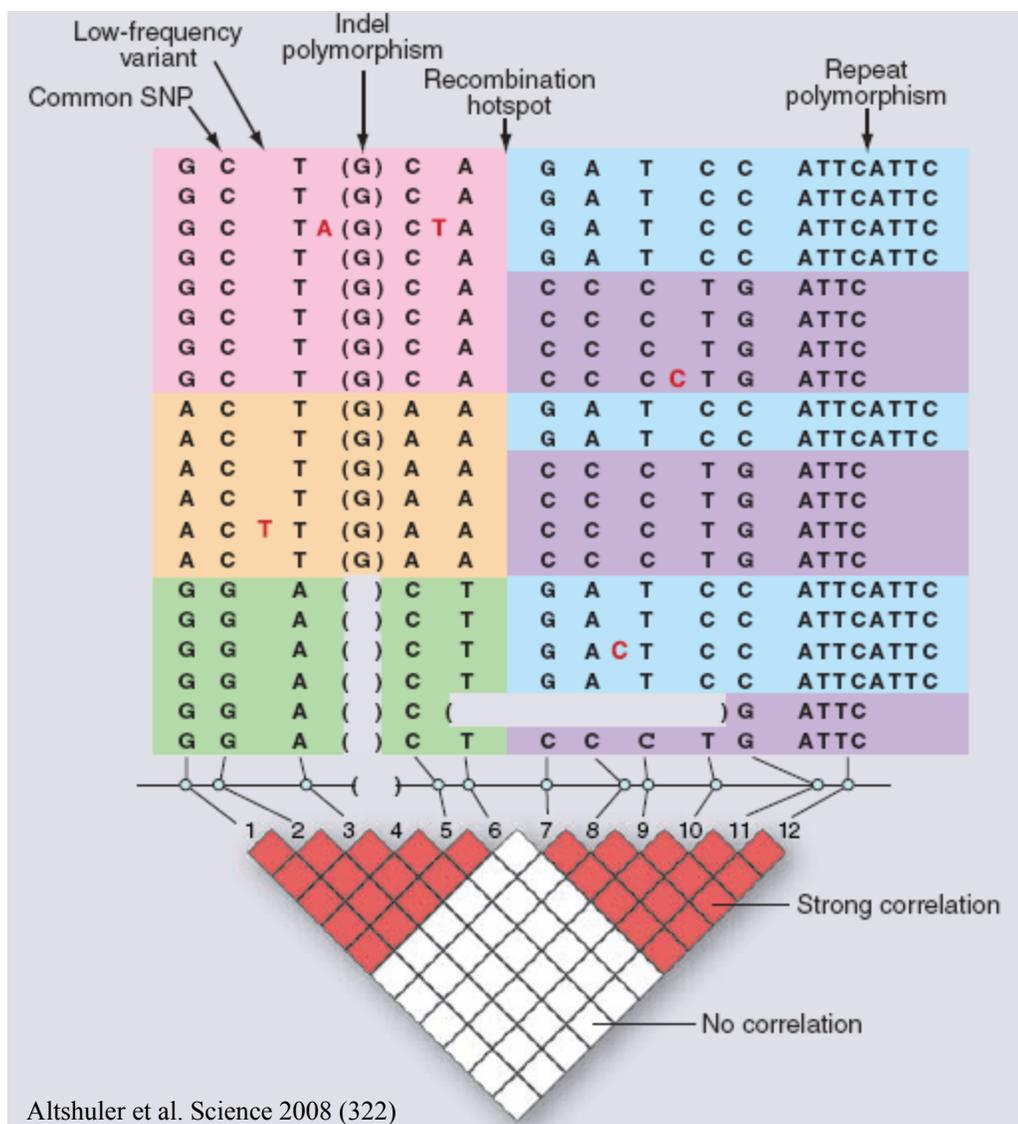


Figure 5. Types de polymorphismes, haplotypes et déséquilibre de liaison.

Hypothèses

Comme l'apoptose est une voie biologique hautement régulée, nous proposons que des SNPs dans les régions promotrices (rSNP) des gènes régulateurs de l'apoptose perturberaient cette voie en créant des taux variables de transcrits et des protéines correspondantes. La présence de rSNPs modifierait la séquence de reconnaissance de certains facteurs de transcription au niveau du promoteur (gain ou perte de liaison), altérant ainsi le niveau d'expression des gènes. De plus, il se peut qu'un seul polymorphisme n'ait pas d'impact, mais que ce soit la combinaison de plusieurs polymorphismes (i.e. haplotype) qui modifie l'expression du gène. Nous proposons donc que ces rSNPs génèrent une variabilité interindividuelle dans les mécanismes contrôlant l'apoptose, d'où un risque variable de développer certaines maladies complexes tel que le cancer.

Objectifs du projet

L'objectif général est d'identifier des rSNPs et haplotypes fonctionnels qui seraient responsables de la variation des niveaux d'expression des gènes impliqués dans la voie intrinsèque d'activation de l'apoptose.

Les objectifs spécifiques du projet de recherche sont donc :

- (1) Identifier les différents polymorphismes présents dans la région promotrice proximale des gènes régulateurs de l'apoptose;
- (2) Déterminer les haplotypes les plus fréquents dans la population pour chacun des promoteurs étudiés;
- (3) Mesurer l'impact fonctionnel des rSNPs et des haplotypes correspondants à l'aide d'une combinaison d'approches *in vitro*.

Matériel et Méthodes

Identification des polymorphismes et génotypage

Les polymorphismes des gènes MDM2 et CASP3 ont été identifiés à partir de la base de données dbSNPs (build 126). Afin de vérifier la présence de ces polymorphismes, une cohorte de 40 individus européens a été génotypée par réaction de polymérase en chaîne (PCR). Les amorces, les conditions et les programmes de PCR utilisés pour l'amplification du promoteur de MDM2 et de CASP3 sont décrits en annexe aux tableaux I et II.

Les fragments de PCR (environ 2 kb en amont de l'exon 1) ainsi générés ont été dénaturés avec une solution de NaOH 0,4 N, puis transférés sur membranes chargées positivement (Amersham), à l'aide d'un appareil à dot blot (S&S minifold). L'ADN a été fixé sur la membrane aux UVs par un UV Stratalinker 1800 (Stratagene). Ces membranes ont été incubées en présence de sondes oligonucléotides simple-brin spécifiques à chaque allèle (voir Annexes Tableau III pour séquences) marquées au ^{32}P avec une T4 kinase (Invitrogen). Ces sondes oligonucléotidiques peuvent s'hybrider aux fragments d'ADNs qui leur sont complémentaires et chevauchant un SNP. L'incubation a été faite sur une étendue de température variant de 70°C à 37°C et en présence d'un compétiteur non marqué correspondant à l'oligonucléotide spécifique à l'autre allèle du SNP. Les membranes étaient ensuite exposées sur film Kodak 16 heures à -80°C. Les films étaient développés par l'appareil Micro-Max (Hope). L'analyse visuelle des films a permis de déterminer le génotype de chacun des 40 individus pour les SNPs étudiés. Seule la délétion de la région promotrice de MDM2 a été génotypée par

PCR suivi d'une séparation électrophorétique (conditions, programmes et amorces dans les tableaux I et II en annexe). Suite à la migration (gel d'agarose 3%) des fragments de PCR couvrant la région de délétion, nous pouvons distinguer les individus homozygotes avec la délétion (1 bande migrant rapidement) des individus hétérozygotes (2 bandes) et des homozygotes sans délétion (1 bande migrant moins rapidement).

Pour tous les autres gènes étudiés, l'identification des SNPs dans les régions promotrices a été faite par reséquençage direct des régions promotrices de ces gènes chez 40 individus caucasiens dont 20 individus sains (contrôles) et 20 patients atteints de leucémie lymphoblastique aigüe (LLA). Ces derniers ont été inclus dans cette étude à cause des intérêts de notre laboratoire pour la leucémie de l'enfant. Pour procéder au reséquençage, des régions promotrices amplifiées par PCR ont été envoyées à la plate-forme de séquençage de Génome Québec à l'Université McGill. Les amorces, les conditions et programmes de PCR utilisés pour les gènes XIAP, BCL2, BAX, BCL2L10, BCL2L11, PARP1, BAG4, ENDOG, BCL2L1 et TP73 sont décrits en annexe dans les tableaux I et II. Les polymorphismes sont ensuite identifiés en comparant les séquences des individus étudiés avec la séquence de référence de NCBI et les SNPs identifiés sont comparés à ceux répertoriés dans dbSNPs.

Détermination des haplotypes fréquents

Une fois les polymorphismes identifiés dans une région promotrice, il est possible de déterminer la structure des haplotypes, correspondant (c'est-à-dire les combinaisons de rSNPs) Pour ce faire, on utilise le logiciel PHASE (version 2,

University of Washington, Seattle) qui calcule les haplotypes à partir des données de fréquence génotypique. Les haplotypes des promoteurs ayant une fréquence égale ou supérieure à 5% ont été conservés pour l'étude de leur impact fonctionnel. Le déséquilibre de liaison entre les polymorphismes est calculé par les valeurs de D' et de R^2 (118, 119).

Étude de l'impact fonctionnel des rSNPs/rHaplotypes

L'impact fonctionnel des rSNPs sur l'activité promotrice est mesuré *in vitro* par des essais de gènes rapporteurs et par des essais de retard sur gel.

Essai de gène rapporteur

Les régions promotrices d'intérêts ont été amplifiées à partir de l'ADN génomique des individus ayant les haplotypes désirés. Le sous-clonage de chaque haplotype à tester se fait à l'aide du système Gateway (Invitrogen) qui nécessite une série de deux réactions de recombinaison, utilisant les enzymes BP et LR. La première recombinaison s'effectue entre les différents amplicons et le vecteur pDONRTM (Invitrogen). Une deuxième recombinaison est faite entre les vecteurs pDONRTM contenant les différents inserts de la région promotrice et le vecteur sans promoteur pGL3basic (Promega), exprimant le gène de la luciférase *Firefly*. Cette deuxième recombinaison permet de placer les inserts des différents haplotypes en amont du gène de la luciférase *Firefly* et les constructions générées seront utilisées pour nos essais de gènes rapporteurs. Les réactions de recombinaison sont utilisées pour la transformation de bactéries compétentes DH5 α . L'intégrité des clones, purifiés avec le Quiagen mini kit, a été vérifiée par séquençage pour confirmer la présence de l'haplotype désiré avant d'effectuer les expériences de transfection transitoire.

Pour ces dernières, on a utilisé un pool de trois clones afin de minimiser l'impact d'une erreur potentielle de séquence sur l'activité promotrice de la construction.

Les pools de clones sont ensuite transfectés dans trois lignées cellulaires différentes (HeLa, HepG2 et JEG3) en utilisant la lipofectamine selon le protocole du manufacturier (Invitrogen). Le choix de 3 lignées cellulaires est justifié afin de vérifier si l'impact observé est tissu-spécifique ou bien s'il est généralisé. Les constructions générées ainsi que le vecteur pRL-CMV (Promega) exprimant la luciférase *Renilla* sous contrôle d'un promoteur SV40 sont cotransfectés (ratio 100 :1). Le vecteur pGL3basic (Invitrogen) est utilisé comme contrôle négatif alors que le vecteur pGL3SV40 (Invitrogen) est utilisé comme contrôle positif. Les cellules en suspension ont été transfectées, puis déposées dans des plaques de 96 puits (environ 6×10^4 cellules/puits). Vingt-quatre heures après la transfection transitoire, la lyse passive des cellules est effectuée et l'activité du gène rapporteur (la luciférase) est mesurée avec la trousse Dual-Luciferase Reporter Assay (Promega) avec un luminomètre Spectra Max 190 (Molecular Devices) selon le protocole du manufacturier. L'activité *Firefly* des constructions contenant les divers haplotypes est normalisée par l'activité de contrôle interne de la *Renilla*. Les résultats sont exprimés en ratio de l'activité *Firefly/Renilla* et correspondent à la moyenne de quatre réplicats. Les analyses statistiques effectuées sont le Student's *t*-test non païré pour déterminer les probabilités (*p-values*).

Expériences de retard sur gel

L'impact des différents SNPs sur la liaison de protéines nucléaires (activateurs ou répresseurs transcriptionnels) est déterminé par essais de retard sur gel pour les

gènes MDM2, BCL2 et PARP1. Des oligonucléotides double-brins marqués au ^{32}P correspondant à la région entourant le site polymorphique d'intérêt (liste des séquences des oligonucléotides utilisés en annexe tableau IV) sont incubés avec les extraits nucléaires de cellules HeLa, HepG2 et JEG3 en utilisant la trousse de Gel Shift assay System (Promega). La séquence de ces sondes est déterminée en utilisant le logiciel MatInspector (Genomatix Software GmbH) pour prédire les sites de liaison de facteurs de transcription autour de chacun des SNPs (120). Tous les sites de liaison qui chevauchent le site putatif du SNP sont conservés dans la séquence de la sonde en plus de 3 nucléotides supplémentaires à chacune des extrémités. Le dosage des protéines dans les extraits nucléaires est effectué par essai Bradford (BioRad). Cinq μg de chaque extrait nucléaire sont incubés séparément avec 35 fmol de sonde marquée dans un tampon contenant 50 mM Tris-HCl (pH 7,5), 5 mM MgCl_2 , 2,5 mM EDTA, 2,5 mM DTT, 250 mM NaCl, 0,25 $\mu\text{g}/\mu\text{L}$ poly deoxyinosinate-deoxycytidylate et 20% glycérol dans un volume final de 10 μL . Ces incubations ont été effectuées à température ambiante pendant 20 minutes. Les complexes sont ensuite séparés par gel de polyacrylamide non dénaturant (6% acrylamide :bisacrylamide, 60 :1) avec du tampon 0,5X Tris-borate-EDTA à 4°C et à 190V. La compétition avec des sondes non marquées est faite dans un ratio de 50 :1 (non marquée : marquée). Une sonde non-spécifique et non marquée appelée AP2 est utilisée pour contrôler la spécificité de liaison. Sa séquence correspond à la séquence mutée de liaison du facteur de transcription AP2.

Identification et priorisation de nouveaux gènes candidats

Approche qRT-PCR

Du sang périphérique (≈ 7 mL) a été prélevé chez 21 individus normaux d'origine européenne. Les globules blancs, isolés des échantillons de sang par centrifugation à 2000 g pendant 10 minutes, ont été mélangés immédiatement avec une solution RNA later (Ambion). L'ARN total de ces cellules a été extrait à l'aide de la trousse Ribopure Blood d'Ambion. Un traitement de l'ARN purifié à la DNase I (Ambion) est fait pendant 30 minutes à 37°C. L'ARN de chaque individu contrôle est ensuite transformé en cDNA par transcription inverse avec la trousse PCR Array Apoptosis First Strand (SA Biosciences). Chaque cDNA obtenu est ensuite déposé dans chacun des 96 puits (10 ng/puits) d'une plaque de qRT-PCR de la trousse PCR Array Apoptosis (SA Biosciences), où chaque puits correspond à des amorces capables d'amplifier un gène d'apoptose différent (voir Tableau II). Les courbes d'amplification obtenues sont ensuite analysées pour déterminer le Ct, c'est-à-dire l'endroit où la courbe commence son amplification détectable par la fluorescence du SYBR[®] Green. Chaque Ct chez un individu est ensuite divisé par la moyenne des Ct d'un groupe de gènes contrôles (GAPDH, B2M, HPRT1, RPL3A, ACTB et HGDC) à des fins de normalisation. Les ratios de Ct des 21 individus contrôles sont ensuite comparés entre eux pour chacun des gènes afin de voir la variabilité interindividuelle dans les niveaux d'expression des gènes régulateurs de l'apoptose. Cette variabilité interindividuelle est calculée en faisant l'écart interquartile entre les ratios de Ct pour chaque gène.

Tableau II. Gènes d'apoptose dont l'expression est mesurée par qRT-PCR par la trousse de Super Array.

Gènes d'apoptose et leur voie			
ABL1	extrinsèque	CIDEA	extrinsèque
AKT1	intrinsèque	CIDEB	extrinsèque
APAF1	intrinsèque	CRADD	extrinsèque
BAD	intrinsèque	DAPK1	extrinsèque
BAG1	intrinsèque	DFFA	intrinsèque
BAG3	intrinsèque	FADD	extrinsèque
BAG4	intrinsèque	FAS	extrinsèque
BAK1	intrinsèque	FASLG	extrinsèque
BAX	intrinsèque	GADD45A	extrinsèque
BCL10	intrinsèque	HRK	intrinsèque
BCL2	intrinsèque	IGF1R	extrinsèque
BCL2A1	intrinsèque	LTA	extrinsèque
BCL2L1	intrinsèque	LTBR	extrinsèque
BCL2L10	intrinsèque	MCL1	intrinsèque
BCL2L11	intrinsèque	NOL3	extrinsèque
BCL2L2	intrinsèque	PYCARD	extrinsèque
BCLAF1	intrinsèque	RIPK2	extrinsèque
BFAR	intrinsèque	TNF	extrinsèque
BID	intrinsèque	TNFRSF10A	extrinsèque
BIK	intrinsèque	TNFRSF10B	extrinsèque
NAIP	intrinsèque	TNFRSF11B	extrinsèque
BIRC2	intrinsèque	TNFRSF1A	extrinsèque
BIRC3	intrinsèque	TNFRSF21	extrinsèque
BIRC4	intrinsèque	TNFRSF25	extrinsèque
BIRC6	intrinsèque	CD27	extrinsèque
BIRC8	intrinsèque	TNFRSF9	extrinsèque
BNIP1	intrinsèque	TNFSF10	extrinsèque
BNIP2	intrinsèque	CD70	extrinsèque
BNIP3	intrinsèque	TNFSF8	extrinsèque
BNIP3L	intrinsèque	TP53	intrinsèque
BRAF	extrinsèque	TP53BP2	intrinsèque
NOD1	intrinsèque	TP73	intrinsèque
CARD6	extrinsèque	TRADD	extrinsèque
CARD8	extrinsèque	TRAF2	extrinsèque
CASP1	extrinsèque	TRAF3	extrinsèque
CASP10	extrinsèque	TRAF4	extrinsèque
CASP14	extrinsèque	CFLAR	extrinsèque
CASP2	extrinsèque	CASP8	extrinsèque
CASP3	intrinsèque	CASP9	intrinsèque
CASP4	extrinsèque	CD40	extrinsèque
CASP5	extrinsèque	CD40LG	extrinsèque
CASP6	intrinsèque	CASP7	intrinsèque

Approche - Réanalyse de données de Microarrays

Les données brutes de Microarrays de 132 patients LLA ont été obtenues à partir de l'étude de Ross et al. (121) qui regroupe les données générées à l'aide de puces Affymetrix (U133A et U133B) couvrant le génome humain. Puisqu'on ne s'intéresse qu'aux données concernant les gènes régulateurs de l'apoptose, les différentes sondes de la puce sont filtrées en fonction de leur appartenance à un gène de la voie intrinsèque d'activation de l'apoptose. Une fois ces sondes regroupées, on calcule l'écart interquartile entre les 132 patients pour chaque sonde. Dans cette étape de criblage, nous avons utilisé une valeur $p=0,10$ appliquée sur la distribution des écarts interquartiles. Cela permet de conserver les sondes les plus significatives, le plus grand écart étant la sonde variant le plus chez les 132 patients leucémiques.

Nous avons aussi utilisé un ensemble de données de microarrays générées à partir d'individus contrôles. Dans cette étude d'Emilsson et *al.* (122), ils ont utilisé l'ARN du sang de 1021 individus islandais non-malades pour mesurer l'expression du génome humain en entier sur la plateforme GPL3991 (Rosetta/Merck) (122). Comme précédemment, les données ont été triées afin de conserver les sondes liées à des gènes de la voie intrinsèque de l'apoptose. Les données obtenues pour chaque gène chez chaque individu sont sous forme de p-values qui découlent de la comparaison entre le signal pour l'ARN d'un individu et le signal pour un pool d'ARN référence composés de 85 individus islandais (43 hommes, 42 femmes)

(123, 124). L'analyse de ces p-values se fait selon la méthode décrite dans Emilsson et *al.* (122) en prenant tout d'abord le $-\log_{10}$ (p-value) pour toutes les données. On calcule ensuite l'écart-type entre les différents $-\log_{10}$ (p-value) des individus pour chacune des sondes. Les écart-types de toutes les sondes sont ensuite classés en ordre décroissant où les sondes tombant dans le haut de la liste correspondent aux gènes ayant la plus grande variabilité interindividuelle. Les premiers 10% ayant la plus grande variabilité sont priorisés comme gènes candidats.

Résultats

Préambule

Le projet sur l'étude de l'impact fonctionnel des polymorphismes dans les promoteurs des gènes de l'apoptose a débuté alors que nous avions peu de moyens de les prioriser. Dans un tel contexte, les premiers gènes, MDM2, CASP3, BAG4, BAX et ENDOG, ont été sélectionnés en fonction de leur importance dans la voie intrinsèque de l'apoptose et des données disponibles dans la littérature. Les autres gènes, PARP1, BCL2L10, XIAP, BCL2, BCL2L1, BCL2L11, TP73, ont été choisis en fonction de la variabilité interindividuelle observée au niveau de leur taux d'expression. À cet effet, des analyses de données d'expression de plusieurs gènes impliqués dans la régulation de l'apoptose ont été effectuées. Ces données d'expression ont été soit générées par qRT-PCR à partir du sang d'individus contrôles ou encore par Microarrays sur deux types de cohorte : des lymphocytes d'individus atteints de LLA ou du sang d'individus contrôles. Il s'agissait dans chaque cas de trouver les gènes ayant la plus grande variabilité interindividuelle.

De plus, le génotypage des SNPs des premiers gènes, MDM2 et CASP3, a d'abord été fait par hybridation avec des sondes oligonucléotides allèle-spécifiques, alors que les autres gènes ont été génotypés par reséquençage direct. Le séquençage direct permet d'éliminer les SNPs qui sont présents dans les bases de données mais absents de notre cohorte, dû à leur trop faible fréquence. Cela permet aussi d'inclure certains SNPs qui n'auraient pas encore été répertoriés et qui sont présents dans les régions promotrices que nous étudions et dans notre cohorte.

Au total, nous avons étudié les promoteurs de 12 gènes de la voie intrinsèque de l'apoptose. Nous avons identifiés 52 SNPs dans les promoteurs proximaux que nous avons définis comme étant une région de 2kb en amont du premier exon. Parmi les gènes étudiés, les promoteurs de MDM2, BCL2 et PARP1 ont montré une activité promotrice différentielle ainsi que des liaisons ADN-protéines différentielles entre les différents haplotypes. Le tableau III résume l'étude de l'ensemble des promoteurs. Les gènes MDM2, BCL2 et PARP1 seront par la suite décrits en détails. On trouve en annexe toutes les données sur les SNPs, les haplotypes et les essais luciférase des gènes n'ayant pas montré d'impact fonctionnel.

Tableau III. Résumé des résultats obtenus pour les promoteurs des 12 gènes étudiés de la voie intrinsèque de l'apoptose.

gènes**	Nb. de SNPs	Nb. d'haplotypes	Essais gènes rapporteurs	Essais de retard sur gel*
MDM2	4	3	expression différentielle	liaison différentielle
CASP3	9	3	expression non-différentielle	-
BAX	6	2	expression non-différentielle	-
ENDO G	9	3	expression non-différentielle	-
BAG4	2	2	expression non-différentielle	-
BCL2L1	2	2	expression non-différentielle	-
BCL2	2	3	expression différentielle	liaison différentielle
BCL2L10	5	4	expression non-différentielle	-
BCL2L11	6	3	expression non-différentielle	-
TP73	1	2	expression non-différentielle	-
XIAP	6	3	expression non-différentielle	-
PARP1	6	3	expression différentielle	liaison différentielle

*: plusieurs gènes n'ont pas d'essais de retard sur gel à cause de l'absence d'activité promotrice différentielle.

**Les gènes en gris sont ceux ayant un impact fonctionnel.

MDM2

Le gène anti-apoptotique MDM2 est une ubiquitine ligase E3 impliquée dans la dégradation de p53, une protéine suppresseur de tumeurs. MDM2 possède deux promoteurs, un en amont de l'exon 1 (P1) et un autre dans le premier intron (P2) (125). Le promoteur P2 est sous le contrôle du facteur de transcription p53 alors que le promoteur P1 est p53-indépendant (107). La base de données dbSNPs contient 9 SNPs pour la région P1 et un SNP pour la région P2 (voir Tableau IV). Ce dernier (SNP601) a déjà été associé au risque de développer certains cancers (101-107). Suite au génotypage de chacun de ces SNPs avec des oligonucléotides allèle-spécifiques, il s'avère que seuls les SNPs -1494A>G, la délétion de 40pb, -182C>G et le SNP +601T>G sont présents dans notre cohorte. Le schéma du promoteur de MDM2 avec les SNPs présents dans la cohorte étudiée est illustré à la Figure 6. Les SNPs -737C>T, -683A>G et -151C>T ne sont pas détectés dans notre cohorte peut-être parce qu'ils sont spécifiques à certaines populations ou encore qu'ils correspondent à des singletons, i.e. qu'ils se retrouvent chez un seul individu dans une population. D'autre part, les SNPs -1164C>G et -1166T>G ne sont pas de vrais SNPs. Ce sont les individus ayant la délétion qui donnent l'impression d'avoir 2 polymorphismes supplémentaires aux positions -1166 et -1164. En réalité, la délétion dans cette région est celle d'une répétition de 40 pb identiques aux 40 pb suivantes à l'exception de 2 nucléotides, donnant l'impression d'être en présence de deux SNPs à ces positions.

Tableau IV. SNPs des promoteurs de MDM2 avec leur position génomique et leur fréquence allélique observée dans notre cohorte.

SNP ^θ	Position chromosomique *	# rs	Fréquence allélique
-1494 G>A	chr12:67,486,752	rs1144944	25%
40 pb indel (-1208 to-1169)	chr12:67,487,038_67,487,077	rs3730485	37%
-1166 T>G	chr12:67,487,080	rs2904506	-
-1164 C>G	chr12:67,487,082	rs3930427	-
-737 C>T	chr12:67,487,509	rs3730486	0%
-683 A>G	chr12:67,487,563	rs3730487	0%
-182 C>G	chr12:67,488,064	rs937282	50%
-151 C>T	chr12:67,488,095	rs3730491	0%
+601 [‡] T>G	chr12:67,488,847	rs2279744	35%

^θ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

[‡] Ce SNP est connu dans la littérature sous le nom de SNP309 de MDM2.

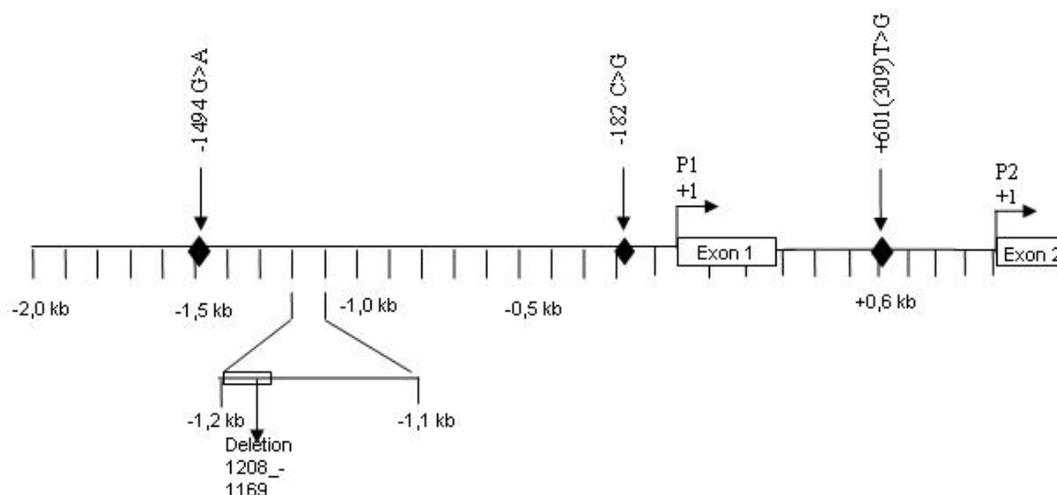


Figure 6. Schéma des promoteurs de MDM2 avec leurs SNPs.

Puisque l'étude d'un seul SNP peut être insuffisante pour expliquer l'expression allélique due à une activité promotrice différentielle, nous avons considéré les promoteurs sous la forme d'haplotypes qui regroupent un certain nombre de SNPs. Les différents haplotypes ont été inférés à l'aide du logiciel PHASE. Seulement les haplotypes ayant une fréquence de plus de 5% dans notre cohorte ont été conservés pour les études fonctionnelles. La région contenant les deux promoteurs de MDM2 possèdent 3 haplotypes fréquents (tableau V).

Tableau V. Haplotypes fréquents de la région promotrice de MDM2.

Haplotype	-1494 G>A	40 pb indel	-182 C>G	+601T>G	Fréquence
1A	A	Non del	C	G	36,8%
1B	A	Non del	C	T	17,0%
2	G	del	G	T	38,5%

Les haplotypes 1A et 1B sont identiques pour la région promotrice P1 alors qu'ils diffèrent pour le SNP+601 de la région P2. Pour ce qui est de l'haplotype 2, il est complètement différent des deux autres. Les trois haplotypes couvrent à eux seuls 92,3% des haplotypes observés. On peut voir la présence d'un déséquilibre de liaison entre les polymorphismes -1494, la délétion de 40 pb ainsi que -182. Le déséquilibre de liaison est illustré par des valeurs de R^2 et de D' qui se rapprochent de 1. La valeur de R^2 est de 0,934 alors que le D' est de 0,977 entre ces trois polymorphismes.

Pour vérifier l'impact de ces trois haplotypes sur l'activité promotrice de MDM2, des essais de gène rapporteur ont été effectués en sous-clonant les différents haplotypes de la région promotrice P1 en amont du gène de la luciférase dans le

vecteur pGL3basic. Des transfections transitoires ont été effectuées pour chaque construction haplotype-spécifique dans 3 différentes lignées cellulaires (HeLa, HepG2 et JEG3). Puisque la région amplifiée pour MDM2 ne contient pas le SNP +601, on ne peut distinguer les haplotypes 1A et 1B. On voit premièrement que la région sous-clonée possède bel et bien une activité promotrice, ayant des valeurs de luciférase 13 à 70 fois plus élevées que le contrôle négatif (pGL3b). On arrive à distinguer une différence d'activité promotrice entre les haplotypes 1 et 2 (Figure 7). Le niveau de luciférase de l'haplotype 1 est de 2 à 3 fois plus élevé que celui de l'haplotype 2, indiquant une différence allélique entre les haplotypes étudiés. Cette différence est statistiquement significative dans les 3 lignées cellulaires étudiées. On observe donc une activité promotrice différentielle entre les haplotypes communs du promoteur P1 de MDM2.

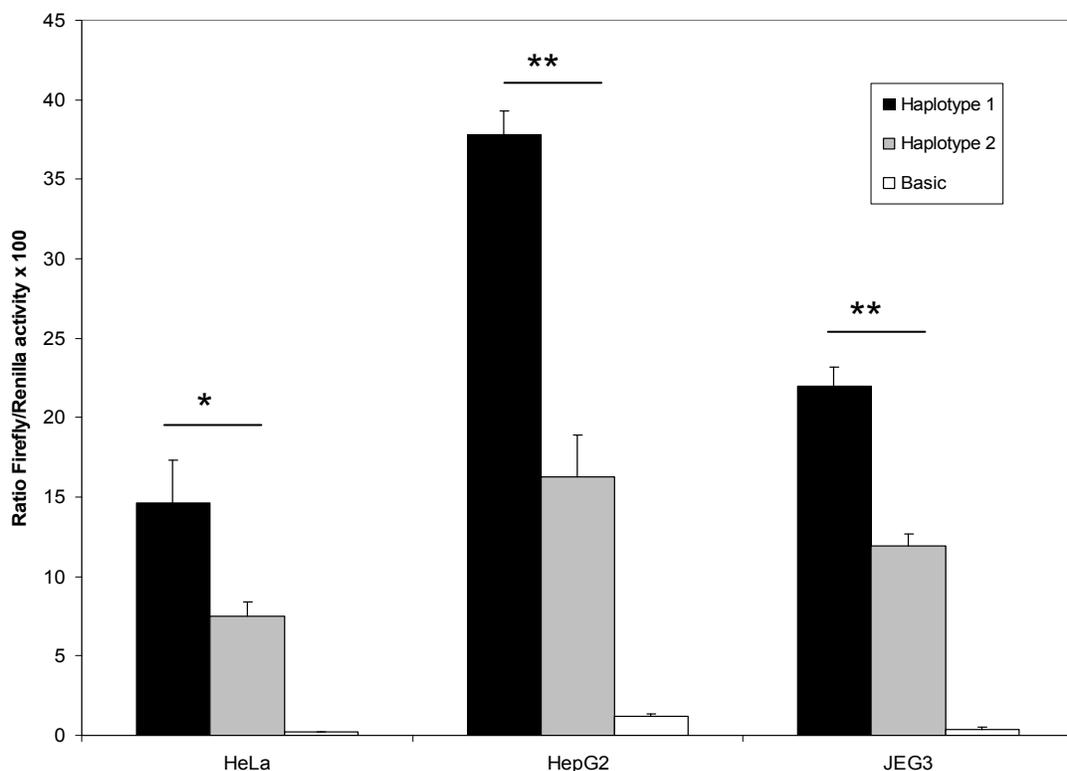


Figure 7. Impact fonctionnel des haplotypes du promoteur P1 de MDM2 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont tirés d'une expérience représentative parmi 3 expériences indépendantes. Les résultats sont exprimés en ratio de l'activité de la *Firefly / Renilla* multiplié par 100. Les valeurs de p sont calculées à partir de 4 réplicats avec le test de *t* de Student non païré (*: $p < 3 \times 10^{-3}$; **: $p < 8 \times 10^{-6}$). Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

Puisque les SNPs des promoteurs sont susceptibles de modifier un site de liaison de facteurs de transcription, l'impact de ceux-ci sur la liaison des protéines a été vérifié par des expériences de retard sur gel. Ceci permet de distinguer la présence ou non de liaisons différentielles de protéines nucléaires au niveau des sites

polymorphiques et ainsi de déterminer quels sont les SNPs susceptibles de créer ou d'abolir un site de liaison expliquant l'activité promotrice différentielle observée en essais de gènes rapporteurs. Pour ce faire, des sondes marquées double-brins (voir tableau IV en annexe) sont incubées avec les extraits nucléaires des lignées cellulaires utilisées pour les essais de gènes rapporteurs (HeLa, HepG2 et JEG3). Les complexes formés sont séparés par électrophorèse sur gel de polyacrylamide non-dénaturant. Ici, seule l'expérience du SNP -182C>G avec la lignée HepG2 est montrée à titre d'exemple (Figure 8). Le complexe 1 montre la présence d'une liaison différentielle. On voit, en comparant les puits 2 et 3, que la compétition est plus forte avec le compétiteur contenant l'allèle C (le complexe disparaît davantage). De plus, on ne retrouve pas de complexe dans les puits 5 à 8. Ces observations laissent suggérer que le complexe 1 de protéines préfère la liaison à la sonde ayant l'allèle C. C'est également le cas pour le complexe 2, qui est davantage compétitionné par la sonde contenant l'allèle C non marquée. De plus, on voit que la liaison de ce complexe est spécifique à la sonde puisque la compétition avec la sonde non marquée AP2 ne fait pas disparaître la bande (puits 4). Les polymorphismes testés (-1494, -182 et +601) ont tous montré une liaison différentielle de protéines, et ce dans les 3 lignées cellulaires testées à l'exception du SNP -182C>G qui a montré de liaison différentielle seulement avec l'extrait nucléaire de la lignée JEG3. Malheureusement, dû à la nature du SNP de la délétion de 40 pb, il est impossible de tester l'impact sur la liaison de protéines par cette méthode.

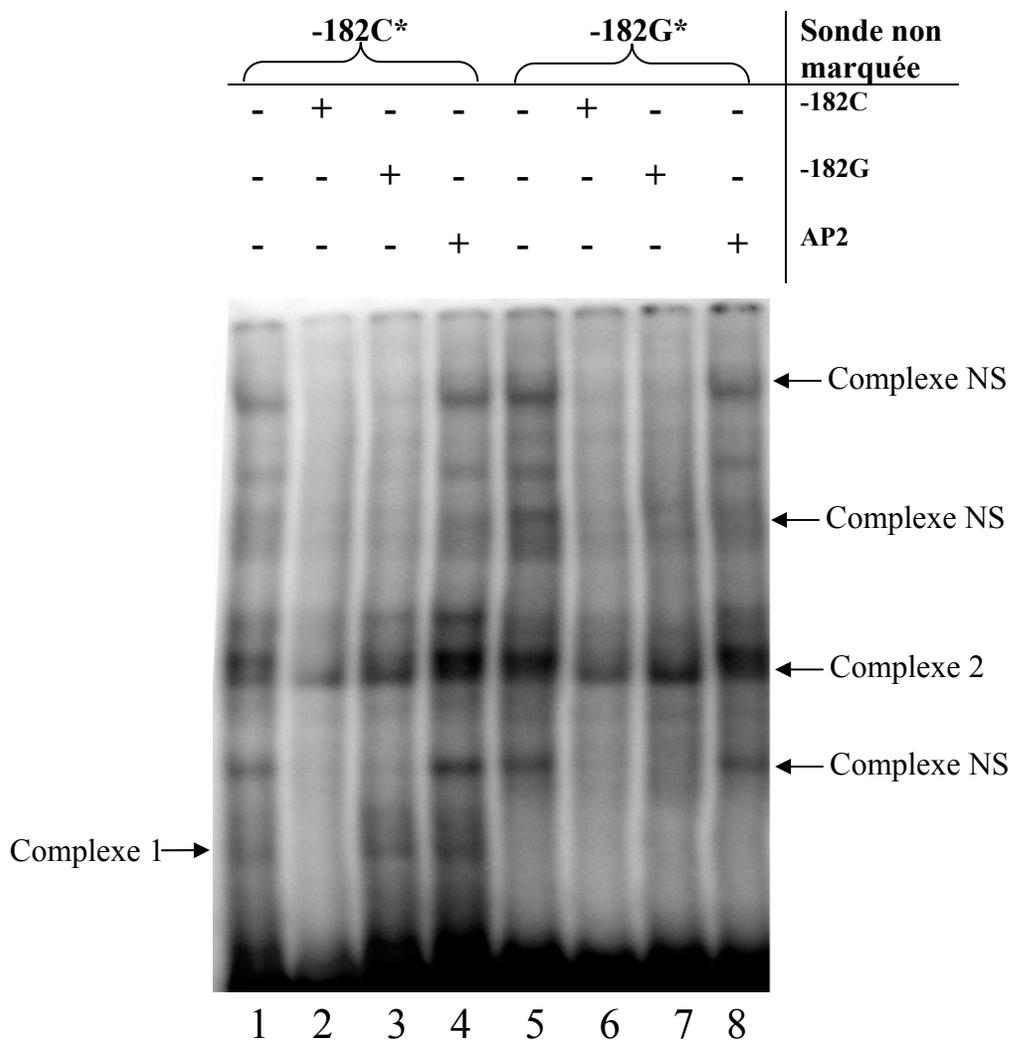


Figure 8. Retard sur gel montrant une liaison différentielle au niveau du rSNP -182C>G de MDM2. Des oligonucléotides double-brins et marqués (*) correspondant aux différents allèles du SNP-182C>G sont incubés avec l'extrait nucléaire de HepG2. Les puits 1 à 4 contiennent la sonde marquée -182C alors que les puits 5 à 8 contiennent la sonde -182G marquée. Les sondes non marquées utilisées pour compétitionner la liaison de protéines à l'ADN sont en excès dans un ratio de 50/1. Des compétiteurs spécifiques correspondant aux sondes non marquées des différents allèles sont utilisés dans les puits 2, 3, 6 et 7 alors qu'un compétiteur non-spécifique est utilisé dans les puits 4 et 8. Les séquences des

Détermination des gènes candidats

Afin de prioriser l'étude de certains promoteurs, la détermination des gènes d'apoptose candidats s'est faite par des réanalyses de données d'expression de gènes de l'apoptose : l'analyse de données de qRT-PCR et l'analyse de données de microarrays.

qRT-PCR

L'ARN total provenant des échantillons de sang de 21 individus contrôles d'origine européenne a été utilisé sur des plaques de PCR Array Apoptosis (SA Biosciences). Ces plaques ont permis de mesurer par qRT-PCR l'expression de 84 gènes d'apoptose chez chacun des 21 individus contrôles. La comparaison des niveaux d'expression (Ct) entre les individus pour chaque gène est effectuée pour déterminer quels sont les gènes ayant la plus grande variabilité interindividuelle. La mesure de cette variabilité est faite avec l'écart interquartile. Un seuil assez permissif de $p=0,1$ est appliqué sur la distribution des écarts interquartiles afin de récupérer les gènes ayant les plus grands écarts (les plus inter-variables) (Tableau VI). Six gènes qui se retrouvent sur la plaque ont des valeurs de p

Figure 8 (suite) : sondes sont décrites dans le tableau IV en annexe. Les sondes non liées migrent rapidement au bas du gel alors que les complexes protéines-sondes migrent moins rapidement. Les complexes non-spécifiques sont identifiés « complexe NS ».

inférieurs à $p=0,1$, dont 3 gènes appartenant à la voie extrinsèque d'activation de l'apoptose (TNFRSF11B, CIDEA et TRAF3) et 3 autres à la voie intrinsèque (BCL2L10, TP73 et BCL2L1). Puisque seule la voie intrinsèque nous intéresse, ces trois derniers gènes ont été priorisés comme gènes candidats. Les haplotypes de ces trois gènes ont été testés par essais gènes rapporteurs et n'ont montré malheureusement aucune activité différentielle (annexe Figure 5, 6 et 8). On peut expliquer cette situation surtout par le fait que les constructions pour les gènes BCL2L10 et TP73 n'ont montré aucune activité promotrice significative. Ces faibles activités ne sont pas dues à un échec de l'essai gène rapporteur, car les contrôles positifs pGL3SV40 (non illustrés) ont bel et bien une activité pour TP73 et BCL2L10. Elles sont donc attribuables à l'absence d'éléments stimulant la transcription. Seules les constructions de BCL2L1 possédaient une activité promotrice (annexe Figure 5). Si on regarde en détails les différents haplotypes de BCL2L10, on s'aperçoit que les niveaux d'activité semblent varier entre les haplotypes, malgré le fait que les promoteurs n'ont pas d'activité significative et que nous sommes peut-être seulement en présence d'un bruit de fond (annexe Figure 6). Il aurait donc peut-être fallu placer ces fragments de 2 kb dans un vecteur en aval d'un promoteur fort comme pGL3SV40 afin d'obtenir une activité « enhancer » différentielle.

D'un autre côté, il se peut que ce soit la façon d'analyser les données de RT-PCR qui ne soit pas tout à fait appropriée et qui ait faussé la liste des candidats. Nous avons choisi d'utiliser l'écart interquartile pour déterminer les gènes ayant la plus grande variabilité interindividuelle. L'écart interquartile permet en effet de

mesurer la dispersion des valeurs d'un ensemble de données, mais ne comprend que 50% des données de la distribution (de Q1 à Q3) et élimine donc les valeurs extrêmes. Dans notre cas, ces valeurs extrêmes auraient peut-être dû être prises en considération et nous auraient permis de trouver des gènes avec une plus grande dispersion des valeurs d'expression entre les individus. Il faudrait peut-être dans ce cas employer l'écart moyenne-médiane qui tient compte de ces valeurs extrêmes. Nous avons appliqué cet écart moyenne-médiane sur nos données. Les gènes suivant sont alors significatifs à $p=0,1$ sont BCL2L11, BCL2 et APAF1. On remarque que ces gènes sont aussi significatifs dans nos analyses de données de microarrays (Tableau VI) et que le promoteur de BCL2 a effectivement donné une activité différentielle entre ses haplotypes (Figure 10).

Analyse de données de microarrays de patients LLA

L'analyse de données d'expression provenant d'ARN extraits d'échantillons de moelles osseuses de 132 patients LLA a été effectuée de façon similaire à l'analyse par qRT-PCR (121). Seules les données des sondes hybridant dans des gènes de la voie intrinsèque de l'apoptose ont été conservées sur les puces U133A et U133B (Affymetrix). Les écarts interquartiles entre les individus pour chaque sonde sont ensuite calculés, et on attribue un seuil encore assez large de $p=0,1$ sur la distribution de ces écarts. Avec ce seuil, 15 gènes sortent significatifs et ont été priorisés comme gènes candidats (Tableau VI).

Analyse de données de microarrays d'individus contrôles

Les données d'individus contrôles sont obtenus à partir de l'ARN du sang de 1021 individus islandais (122). Comme précédemment, seules les données appartenant à des sondes hybridant dans des régions contenant un gène de la voie intrinsèque d'activation de l'apoptose sont conservées. Les p-values obtenues pour chaque gène sont analysées avec la même méthode que celle décrite dans l'étude, de façon à classer les gènes en ordre décroissant des écart-types calculés sur le $-\log_{10}$ (p-value) de tous les individus. Une fois ordonnés, les premiers 10% ayant les plus grands écart-types sont conservés comme étant les gènes ayant la plus grande variabilité interindividuelle et deviennent donc des gènes candidats à leur tour. On obtient donc 4 gènes candidats avec cette analyse, dont BCL2 qui a aussi été identifié par l'analyse de données de microarrays de patients LLA (Tableau VI).

Au total nous avons prioriser 21 nouveaux gènes candidats qui seront plus amplement étudiés.

Tableau VI. Gènes candidats obtenus par analyse de données d'expression de qRT-PCR et de microarrays.

Gène	Méthode	Chromosome	Plate-forme	Cohortes	Tissus analysés
BCL2L10 TP73 BCL2L1	qRT-PCR	15q21 1p36.3 20q11.21	Super Array	Individus sains	Sang périphérique
BCL2L11 YWHAQ YWHAB DYNLL1 BCL2	Microarrays	2q13 2p25.1 20q13.1 12q24.23 18q21.3	Affymetrix U133A, U133B Affymetrix U133A, U133B et GPL3991 Rosetta/Merck	Patients LLA Individus sains et patients LLA	Moelle osseuse Sang périphérique et moelle osseuse
PMAIP1 PARP1 XIAP BIRC6 MCL1 MAPK1 NOD1 CYCS CD47 PECAM1	Microarrays	18q21.32 1q41-q42 Xq25 2p22-p21 1q21 22q11.21 7p15-p14 7p15.2 3q13.1-q13.2 17q23	Affymetrix U133A, U133B	Patients LLA	Moelle osseuse
BCL2A1 BIRC5 APAF1	Microarrays	15q24.3 17q25 12q23	GPL3991 Rosetta/Merck	Individus sains	Sang périphérique

BCL2

BCL2 est un gène anti-apoptotique qui code pour une protéine située à la surface de la membrane mitochondriale. BCL2 empêche la dimérisation de protéines pro-apoptotiques telles que BAX et BAK, bloquant ainsi la libération du cytochrome c de la mitochondrie. Le reséquençage de la région promotrice de BCL2 chez 40 individus a permis d'identifier deux rSNPs : -1534G>A et -225G>T (Tableau VII). Le SNP -1534G>A qui n'est pas répertorié dans dbSNP montre une fréquence plus élevée chez les patients LLA que chez les contrôles. Le schéma du promoteur de BCL2 est illustré à la Figure 9.

Tableau VII. SNPs du promoteur de BCL2 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1534G>A	chr18:59,139,126	non répertorié	35,7%	50%
-225G>T	chr18:59,137,817	rs2279115	65,6%	60,5%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

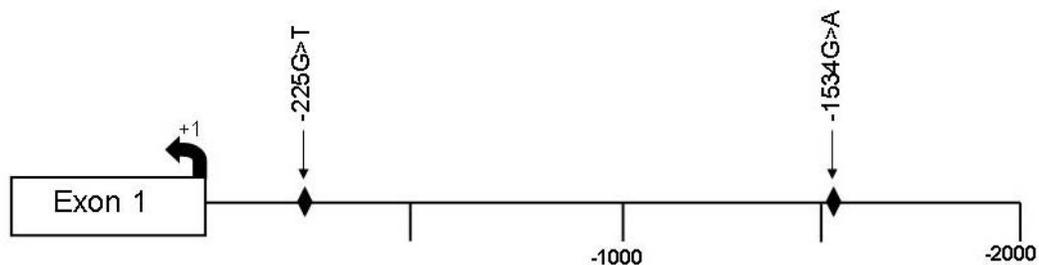


Figure 9. Schéma du promoteur de BCL2. Le gène est orienté de façon anti-sens dans le génome.

Ces deux SNPs de BCL2 se retrouvent sous la forme de trois haplotypes fréquents (tableau VIII). La combinaison -1534A -225G n'a pas été observée chez les individus génotypés.

Tableau VIII. Haplotypes fréquents de la région promotrice de BCL2.

Haplotype	-1534G>A	-225G>T	Fréquence*
1	G	G	37,8%
2	G	T	16,2%
3	A	T	46,0%

*Ces trois haplotypes forment la totalité des haplotypes observés (100%).

L'activité promotrice de ces trois haplotypes a été testée par des essais de gène rapporteur pour voir leur impact sur l'activité promotrice de BCL2 (Figure 10). On constate que le promoteur sous-cloné est bel et bien fonctionnel puisque ses valeurs de luciférase sont de 2 à 10 fois plus élevées que le contrôle négatif (pGL3b). Une activité différentielle est observée ($p < 0,02$ dans les trois lignées) où l'haplotype 2 montre une activité promotrice plus faible que les deux autres : l'haplotype 1 a une activité intermédiaire (de 1,2 à 1,6 fois l'haplotype 2) et l'haplotype 3 est le promoteur le plus fort (de 1,4 à 2,2 fois l'haplotype 2). Le même profil d'expression a été observé dans toutes les lignées cellulaires.

On peut donc confirmer la présence d'une activité promotrice différentielle *in vitro* selon l'haplotype du promoteur de BCL2 présent.

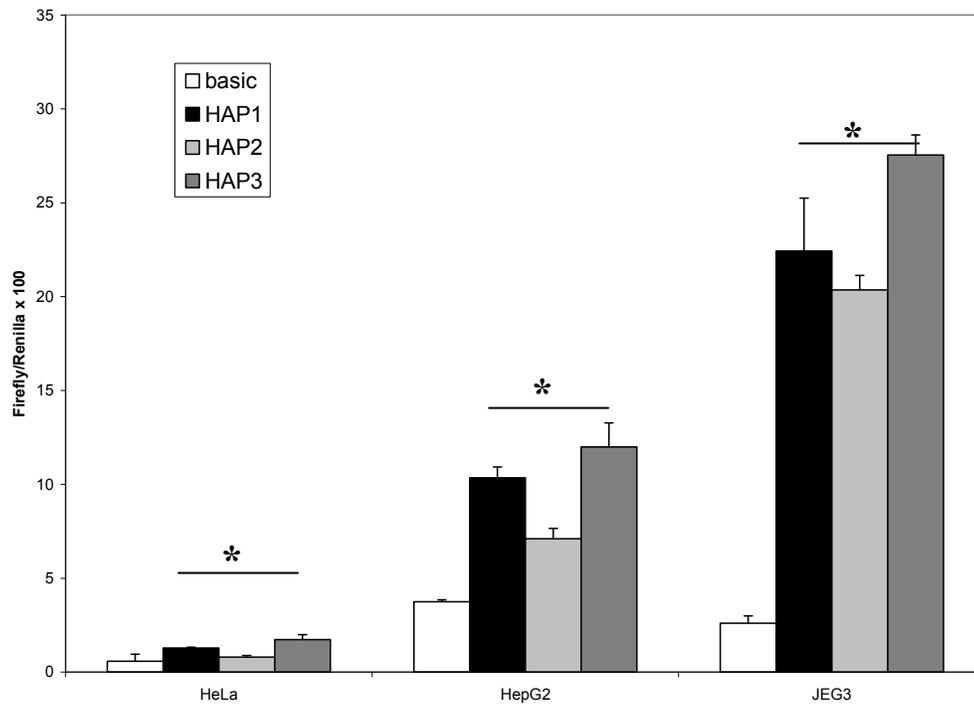


Figure 10. Impact fonctionnel des haplotypes du promoteur de BCL2 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés proviennent d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Les p-values sont calculées à partir de 4 réplicats avec le test de *t* de Student non pairé (*: $p < 0,02$). Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

Des essais de retard sur gel ont été effectués pour vérifier la présence de liaisons différentielles au niveau des SNPs de cette région promotrice. Le gel du SNP -225 avec l'extrait nucléaire de HepG2 est montré à titre d'exemple (Figure 11). On peut y voir la présence d'une liaison différentielle au niveau du complexe 1, qui lie uniquement -225T (puits 1 vs puits 5). La compétition dans les puits 6 et 7

confirme que le complexe se lie davantage à la sonde contenant le T, car la bande ne disparaît qu'en présence du compétiteur -225T non marqué. La compétition avec la sonde AP2 montre aussi une faible disparition de la bande, ce qui indique que la liaison n'est peut-être pas très spécifique. Toutes les lignées cellulaires pour ce polymorphisme ont montré une liaison différentielle alors qu'elles n'en ont pas montrée pour le SNP -1534G>A. Il semble donc que ce soit le SNP -225G>T qui serait responsable de l'activité promotrice différentielle observée.

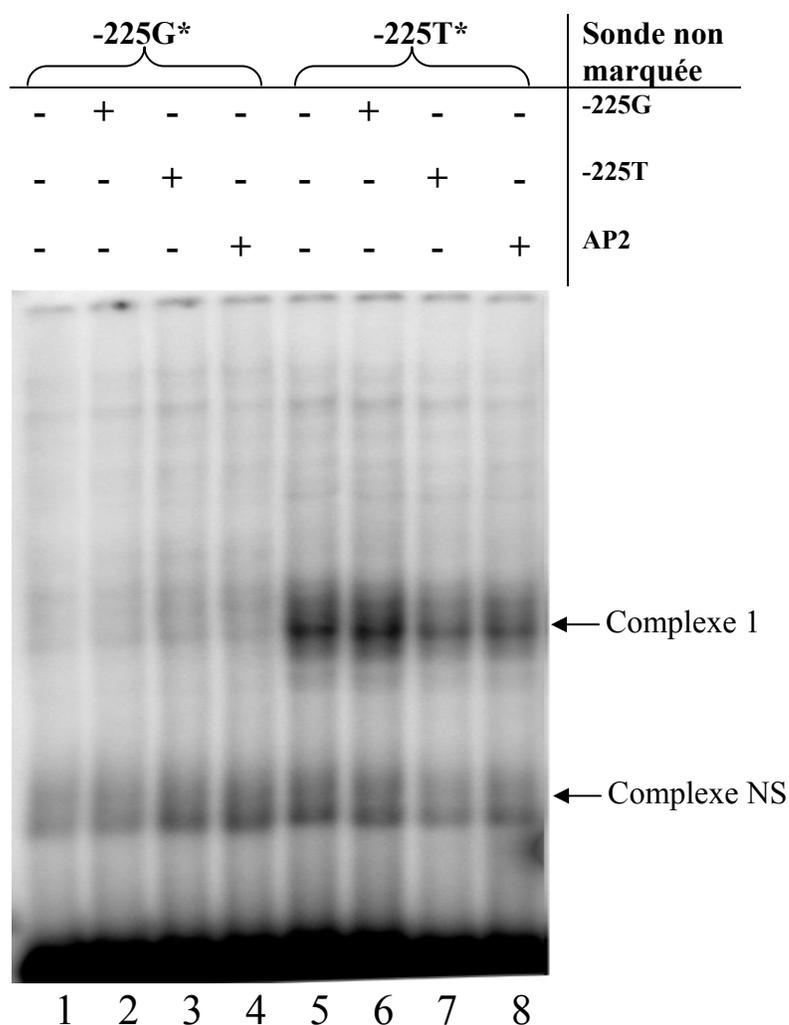


Figure 11. Retard sur gel montrant une liaison différentielle au niveau du rSNP -225G>T de BCL2. Des oligonucléotides double-brins et marqués (*)

PARP1

Le gène PARP1 encode une poly (ADP-ribosyl) transférase qui possède une activité anti-apoptotique qui est clivée par les caspases lors du déclenchement de l'apoptose. Le reséquençage de la région promotrice de PARP1 chez 40 individus a permis d'identifier six rSNPs (tableau IX) : Les SNPs -589A>G, -279C>T, -239C>T et -219A>T sont un peu plus fréquents dans le groupe contrôle que dans le groupe de patients alors que la fréquence des SNPs -886 T>A et -1491 A>G est similaire dans les deux groupes. Le schéma du promoteur de PARP1 est illustré à la Figure 12.

Figure 11 (suite) : correspondant aux différents allèles du SNP-225G>T sont incubés avec l'extrait nucléaire de HepG2. Les puits 1 à 4 contiennent la sonde marquée -225G alors que les puits 5 à 8 contiennent la sonde -225T marquée. Les sondes non marquées utilisées pour compétitionner la liaison de protéines à l'ADN sont en excès dans un ratio de 50/1. Des compétiteurs spécifiques correspondant aux sondes non marquées des différents allèles sont utilisés dans les puits 2, 3, 6 et 7 alors qu'un compétiteur non-spécifique est utilisé dans les puits 4 et 8. Les séquences des sondes sont décrites dans le tableau IV en annexe. Les sondes non liées migrent rapidement au bas du gel alors que les complexes protéines-sondes migrent moins rapidement. Les complexes non-spécifiques sont identifiés « complexe NS ».

Tableau IX. SNPs du promoteur de PARP1 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1491A>G	chr1:224,663,964	rs2793378	35,3%	34,6%
-886 T>A	chr1:224,663,309	rs2793379	18,8%	23,1%
-589 A>G	chr1:224,663,012	rs1341336	37,5%	22,7%
-279 C>T	chr1:224,662,702	rs7527192	20,6%	16,7%
-239 C>T	chr1:224,662,662	rs2077197	21,9%	13,6%
-219 A>T	chr1:224,662,642	rs7531668	21,9%	13,6%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

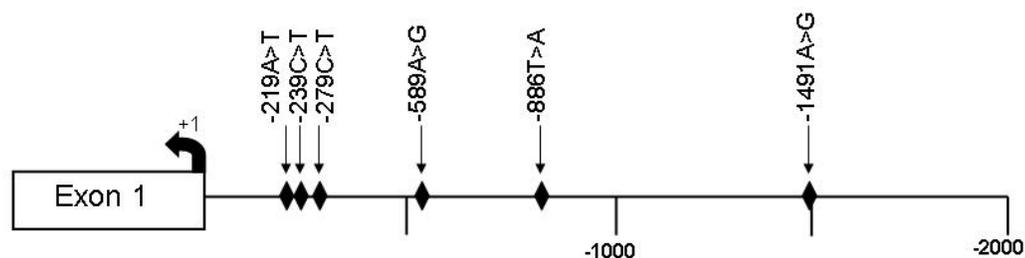


Figure 12. Schéma de la région promotrice de PARP1. Le gène est orienté de façon anti-sens dans le génome.

Cette région promotrice possède trois haplotypes fréquents (Tableau X). On constate que les polymorphismes -279C>T, -239C>T et -219A>T sont en déséquilibre de liaison car on ne retrouve que deux combinaisons possible pour ces SNPs soit « CCA » et « TTT ».

Tableau X. Haplotypes fréquents de la région promotrice de PARP1.

Haplotype	-1491 A>G	-886 T>A	-589 A>G	-279 C>T	-239 C>T	-219 A>T	Fréquence
1	A	T	A	C	C	A	59,7%
2	G	T	G	T	T	T	16,1%
3	G	A	G	C	C	A	17,7%

Les essais de gènes rapporteurs effectués avec les différentes constructions de ces différents haplotypes ont montré que le promoteur sous-cloné est bel et bien fonctionnel dû au niveau de luciférase qui est de 4 à 10 fois supérieur au pGL3b. De plus, on note la présence d'une activité promotrice différentielle (Figure 13). L'haplotype 1 possède une activité promotrice de 1,8 à 2,4 fois moins élevée que les haplotypes 2 et 3, et ce dans les trois lignées cellulaires étudiées. Si l'on compare les différences entre les allèles de ces différents haplotypes on remarque que seul les SNPs -1491A>G et -589A>G sont différents chez l'haplotype 1 comparativement aux autres haplotypes. Les positions -886, -279, -239 et -219 sont différentes entre les haplotypes 2 et 3 mais ne semblent pas affecter l'activité promotrice *in vitro*.

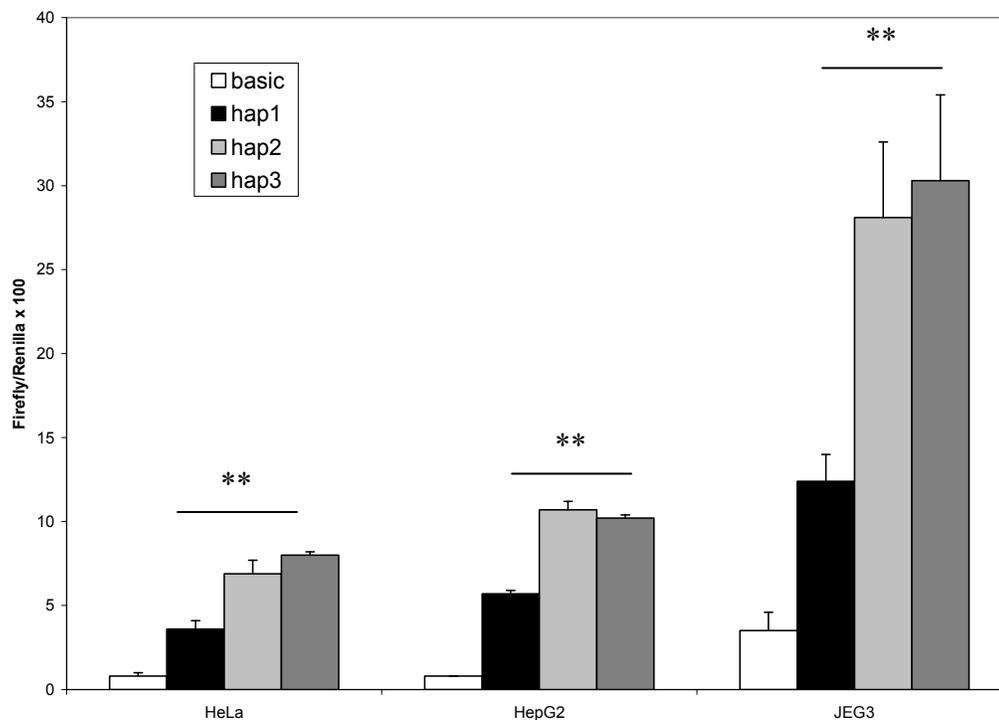


Figure 13. Impact fonctionnel des haplotypes du promoteur de PARP1 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés proviennent d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic). Les p-values sont calculées à partir de 4 réplicats avec le test de *t* de Student non païré (**: $p < 6,0 \times 10^{-5}$)

Afin de voir quels seraient les SNPs responsables de cette expression différentielle de luciférase, des essais de retard sur gel ont été effectués. Seul les SNPs -1491A>G et -589A>G sont testés, car ce sont eux qui semblent responsable de l'activité différentielle observée. Le SNP -1491A>G en présence de l'extrait nucléaire de HeLa est illustré comme expérience représentative (Figure 14). Le complexe 1 lie davantage la sonde contenant l'allèle A que la sonde contenant

l'allèle G (puits 1 et 5). Par ailleurs, on voit que lorsqu'on compétitionne avec une sonde non marquée -1491A, la bande disparaît davantage qu'avec la compétition avec la sonde non marquée -1491G (puits 2 et 3). La liaison semble être spécifique puisque la compétition avec la sonde AP2 ne fait pas disparaître la bande (puits 4). Le SNP -1491A>G a montré une liaison différentielle dans toutes les lignées cellulaires de même que le SNP -589A>G qui a montré une liaison différentielle dans 2 lignées sur 3.

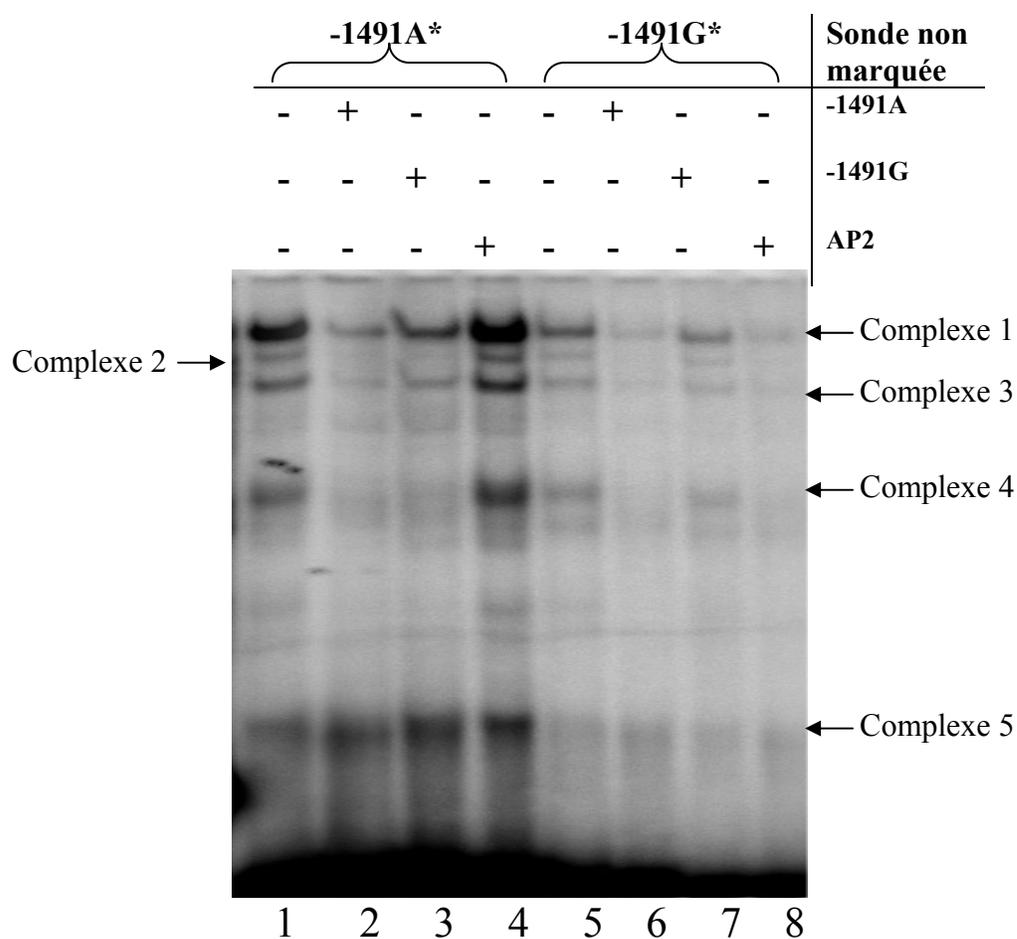


Figure 14. Retard sur gel montrant une liaison différentielle au niveau du rSNP -1491A>G de PARP1. Des oligonucléotides double-brins et marqués (*) correspondant aux différents allèles du SNP-1491A>G sont incubés avec l'extrait

nucléaire de HeLa. Les puits 1 à 4 contiennent la sonde marquée -1491A alors que les puits 5 à 8 contiennent la sonde -1491G marquée. Les sondes non marquées utilisées pour compétitionner la liaison de protéines à l'ADN sont en excès dans un ratio de 50/1. Des compétiteurs spécifiques correspondant aux sondes non marquées des différents allèles sont utilisés dans les puits 2, 3, 6 et 7 alors qu'un compétiteur non-spécifique est utilisé dans les puits 4 et 8. Les séquences des sondes sont décrites dans le tableau IV en annexe. Les sondes non liées migrent rapidement au bas du gel alors que les complexes protéines-sondes migrent moins rapidement.

Discussion

Rappel du projet

Contrairement aux maladies monogéniques, où la présence d'une mutation suffit pour développer la maladie, les maladies complexes s'expliquent plutôt par l'interaction étroite entre des facteurs génétiques et environnementaux. La variabilité interindividuelle de la susceptibilité de développer une maladie complexe est, en partie, due à la diversité génétique (95-97). Une des principales sources de cette variabilité génétiques sont les SNPs qui se retrouvent dispersés dans le génome. Mon étude a porté plus particulièrement sur les SNPs localisés dans les régions promotrices des gènes. Ces SNPs peuvent altérer la transcription des gènes, en modifiant des sites de liaison à des facteurs de transcription. Puisque l'apoptose est une voie bien contrôlée et essentielle pour l'homéostasie cellulaire, les SNPs qui modifieraient les niveaux de transcrits des gènes de l'apoptose pourraient engendrer un déséquilibre de cette voie et ainsi empêcher la cellule dérégulée d'entrer en apoptose. Ceci favoriserait alors l'accumulation de cellules dérégulées et contribuerait à créer un environnement propice à la tumorigenèse.

Sommaire des résultats obtenus

L'étude des polymorphismes dans les promoteurs des gènes de la voie intrinsèque d'activation de l'apoptose a permis d'identifier 52 rSNPs et 33 haplotypes fréquents répartis dans 12 gènes (Tableau III). La densité des polymorphismes observée est de 2,2 SNPs / kb en moyenne, ce qui est un peu plus faible comparativement à ce qui a pu être observé précédemment (environ 4,5 SNPs/kb)(126). Cela peut s'expliquer, entre autre, par le génotypage d'une seule population d'origine européenne. En d'autres termes, les SNPs présents dans d'autres populations mais absents chez les européens n'étaient pas détectés, faisant ainsi baisser la densité de SNPs / kb. Il est aussi possible que les gènes étudiés aient été davantage conservés au cours de l'évolution, dû à leur rôle dans une voie biologique importante. Parmi les SNPs identifiés, 13,5% de ceux-ci n'étaient pas présents dans la base de données dbSNP. Pour ce qui est de la structure et la diversité des haplotypes (2,8 haplotypes/ promoteur), elle correspond à la fréquence attendue (i.e. 2-3 haplotypes majeurs) lorsque l'on considère uniquement les haplotypes fréquents.

Sur les 12 gènes étudiés, les promoteurs de MDM2, BCL2 et PARP1 ont montré une activité promotrice différentielle en fonction de leurs haplotypes. Ces résultats supportent notre hypothèse initiale qui proposait que la variabilité génétique au niveau des promoteurs puisse mener à une activité promotrice différentielle, au moins dans un contexte *in vitro*. Pour les neuf autres promoteurs étudiés, il est possible que les rSNPs présents ne soient pas fonctionnels ou qu'ils ne modifient

pas l'activité du promoteur dans le contexte des lignées cellulaires étudiées. C'est entre autre pour cette raison que nous avons décidé pendant le projet de sélectionner de nouveaux gènes candidats en se basant sur la variabilité interindividuelle de l'expression des gènes de l'apoptose en utilisant des données d'expression de microarrays et de qRT-PCR. Ces études ont permis d'identifier 21 gènes nouveaux candidats (Tableau VI). Les gènes qui ont été testés avant nos analyses de données d'expression (MDM2, CASP3, BAG4, BAX et ENDOG) n'ont pas montré d'activité différentielle entre leurs haplotypes à l'exception de MDM2 (discuté en détails plus loin).

Analyse des gènes candidats

En réanalysant les données de microarrays de patients LLA, nous avons sélectionnés 15 nouveaux gènes candidats dont 4 qui ont été testés dans mon projet (BCL2L11, BCL2, PARP1 et XIAP). Parmi ces gènes, BCL2 et PARP1 ont montré une activité différentielle entre les haplotypes de leur promoteur (Figure 10 et 13). Ces gènes sont discutés en détails plus loin. Même si ces résultats sont encourageants, il faudra tester davantage de gènes avant de pouvoir conclure que cette stratégie permet de sélectionner des gènes ayant des activités promotrices différentielles.

Nous devons souligner que l'analyse des données d'expression de patients LLA comporte certains pièges, car il est difficile de déterminer si la variabilité de l'expression est due à des polymorphismes constitutifs (variation germinale) ou à des mutations somatiques causant la dérégulation de ces gènes dans les cellules leucémiques. Par exemple, le gène XIAP, qui a montré une grande variabilité interindividuelle dans les données de microarrays de patients LLA, est bien connu pour être dérégulé dans plusieurs cancers (48-51). Par contre, nous n'avons pas observé que la variabilité génétique au niveau du promoteur de ce gène altérait les niveaux de transcription (annexe Figure 9). Cette observation nous a motivés à réanalyser, en parallèle, des données de microarrays d'individus contrôles. Sur les 4 gènes candidats ainsi identifiés, nous avons eu l'occasion d'en tester un seul, BCL2, qui a montré une activité promotrice différentielle (Tableau VIII et Figure 10). Il ne faut pas oublier qu'une expression inter-variable mesurée sur les ARNm

des gènes de l'apoptose soit par qRT-PCR ou par microarrays n'est pas nécessairement due à une différence au niveau de la transcription de ces ARNm (pourrait être stabilité/dégradation). Cette limitation sera donc toujours présente peu importe la façon d'analyser les données. De plus, les ARNm mesurés ici proviennent toujours des cellules du sang ou de la moelle. Il est bien connu que chaque type cellulaire possède des quantités et des sortes spécifiques de facteurs de transcription et répresseurs. Il se peut donc que la variabilité observée dans la moelle ou dans le sang soit due à des SNPs auxquels se lient des facteurs de transcription présents uniquement dans ces types cellulaires et absents des lignées cellulaires utilisées pour la transfection.

MDM2

Puisque MDM2 joue un rôle important dans le contrôle de l'apoptose en modulant les niveaux de p53, un niveau d'expression dérégulé de ce gène pourrait devenir propice au développement du cancer (127, 128). D'ailleurs, la surexpression de MDM2 est retrouvée dans plusieurs types de cancers (63, 64). On ne connaît rien sur le lien entre MDM2 et la tumorigenèse, indépendamment de la voie de p53. En effet, il se pourrait très bien que ce soit l'interaction de MDM2 avec d'autres protéines impliquées dans le cancer qui affecte le développement des tumeurs, sans passer par p53 (129-132). Deux régions promotrices sont bien définies pour MDM2, soit P1 qui est en amont de l'exon 1 et P2 qui est localisé dans l'intron 1 (125). La région P2 contient deux éléments imparfaits de réponse à p53 et une boîte TATA qui sont nécessaires pour activer la transcription à partir de ce promoteur (p-53 dépendant), alors que le promoteur P1 ne contient aucun de ces

éléments et est donc p53 indépendant (107, 133). Les deux régions promotrices sont riches en GC ce qui peut provoquer de multiples sites d'initiation de la transcription dans ces promoteurs, expliquant l'impressionnante variété des transcrits observés pour ce gène (134). De plus, les deux régions sont également constitutivement libres en nucléosome, les rendant constamment accessibles aux protéines nucléaires (135).

Notre étude a porté sur l'impact des haplotypes du promoteur P1 de MDM2 (Figure 6). Les résultats obtenus font d'ailleurs l'objet d'un article qui est présentement en révision et présenté en annexe (136). Nous avons démontré que les différents haplotypes de ce promoteur basal ont une activité promotrice différentielle et que les rSNPs retrouvés dans ces haplotypes contribuent seul ou ensemble aux différences observées (Figure 7). De plus, tous les rSNPs présents ont montré une liaison différentielle de protéines au niveau de leur site putatif et ce dans pratiquement toutes les lignées (Figure 8).

Une étude précédente avait déjà démontré une corrélation entre le SNP -182C>G et une expression allélique différentielle (137). Par contre, cette étude montrait que c'était l'allèle G du rSNP qui avait une activité promotrice élevée comparativement à son allèle C. Dans notre étude, l'allèle G du rSNP -182C>G de MDM2 fait partie de l'haplotype 2 (Tableau V). Cet haplotype est celui qui montre la plus faible activité promotrice, allant à l'encontre des résultats de l'étude de Wang *et al.* (108). Cette différence observée entre les deux études peut provenir du fait que la région sous-clonée dans le vecteur pGL3basic (avec le gène rapporteur *Firefly*) n'est pas la même. En effet, l'étude de Wang *et al.* (108) cible une région

de 400 pb entourant le rSNP -182C>G (de -25 à -424) alors que notre étude cible une région plus grande, allant de -27 à -2173. Il se peut que certains sites de liaisons à des facteurs de transcription soient présents dans notre étude mais absents de celle de Wang *et al.* (108). Notre design expérimental nous a également permis de capter l'impact de l'ensemble des rSNPs du promoteur (haplotype). Ce contexte « haplotypique » se rapproche davantage du contexte génomique. Plusieurs autres études ont été faites sur le rSNP +601T>G, montrant des corrélations entre ce polymorphisme et différents types de cancers (101, 102, 104-106, 127, 137, 138). Pourtant, d'autres études ont tenté d'obtenir une telle corrélation avec d'autres types de cancers et se sont avérées un échec (139, 140). Plusieurs facteurs pourraient expliquer ces résultats contradictoires. Tout d'abord, il est possible que le rSNP +601 ne soit pas l'explication au phénomène observé, mais qu'un autre rSNP en déséquilibre de liaison avec celui-ci en soit responsable. Nous avons montré dans notre étude que l'allèle G du rSNP +601 était uniquement associé à l'haplotype 1 du promoteur P1 de MDM2, créant ainsi l'haplotype 1A qui combine les promoteurs P1 et P2 (Tableau V). La combinaison de l'allèle G du rSNP+601 avec l'haplotype 1 du promoteur P1 (grande activité promotrice) augmenterait davantage les niveaux de MDM2 et appuierait les études qui ont montré une corrélation entre le rSNP +601G et certains cancers. Des études *in vitro* ont d'ailleurs montré que l'allèle G du rSNP +601 augmente l'affinité du facteur de transcription sp1 au site du polymorphisme, menant ainsi à l'augmentation de MDM2 et à la diminution de p53 (102, 103). Une autre explication pour ces études contradictoires serait que les facteurs de transcription

liant le rSNP +601T>G sont spécifiques à certains types cellulaires et absents dans d'autres, expliquant les différences observées entre les types de cancers. Aussi, une étude du rSNP +601T>G dans différentes populations a montré que l'association de ce polymorphisme avec le risque de développer un cancer était présent chez les Asiatiques, mais absent chez les Caucasiens et les Africains (140). Cette différence peut alors s'expliquer par la fréquence variable du rSNP entre les différentes populations (50% chez les Asiatiques, 10% chez les Caucasiens et les Africains)(140). Notre fréquence chez les Caucasiens était un peu plus élevée que celle observée dans la littérature, soit de 35% (Tableau IV).

Finalement, nous avons fait des analyses *in silico* pour prédire l'impact de ces rSNPs sur les sites (i.e. gain ou perte) de liaison à des facteurs de transcription. Lorsque l'on compare la séquence délétée avec la séquence non délétée, la différence n'est pas au niveau des sortes de facteurs de transcription qui se lient, mais plutôt au niveau de leur quantité. Comme dit précédemment, la délétion est celle d'une séquence répétée de 40 pb. Les facteurs qui se retrouvent donc dans la région de la délétion lient également la région adjacente qui contient les mêmes 40 pb à 2 nucléotides près. Les prédictions pour le rSNP -182C>G ne montrent pas de différence de liaison de protéine sur son site, alors que celles du rSNP -1494G>A montrent une liaison différentielle. L'allèle A du rSNP contribuerait en effet à former une « boîte GC », à laquelle se lierait le facteur de transcription Sp1. L'allèle G ne formerait pas cette boîte et ne lierait donc pas le facteur de transcription Sp1 qui est connu pour lier cette « boîte GC » et activer la transcription (141). Les individus qui auraient l'haplotype 1 (avec l'allèle A du

rSNP -1494G>A), pourraient avoir une liaison de la protéine Sp1 sur leur promoteur et ainsi voir leur niveau de transcription de MDM2 augmenté par rapport aux individus qui ont l'haplotype 2 et qui n'ont pas cette liaison de la protéine. Cela appuierait nos résultats d'essais de gènes rapporteurs qui montrent que l'haplotype 1 a une plus grande activité promotrice que l'haplotype 2.

BCL2

Le gène BCL2 code pour une protéine anti-apoptotique qui est surexprimée dans plusieurs cancers (43, 44). Notre étude a démontré la présence de deux rSNPs au niveau du promoteur de BCL2 soit -1534G>A et -225G>T (Tableau VII et Figure 9). Elle illustre aussi qu'il existe une activité promotrice différentielle entre les trois haplotypes du promoteur de BCL2 (Tableau VIII et Figure 10). L'haplotype 2 a la plus faible activité promotrice alors que l'haplotype 3 a la plus forte et l'haplotype 2 a une activité intermédiaire.

Une étude précédente a déjà établi une corrélation entre le rSNP -225G>T et la leucémie lymphoblastique chronique (116). Dans cette étude, on constate que l'allèle G du rSNP -225 est associé à une augmentation de l'activité promotrice de BCL2. Par contre, malgré l'augmentation de l'activité promotrice, le niveau de BCL2 est diminué chez les patients ayant le génotype GG par rapport aux patients avec les phénotypes GT ou TT. Cette observation s'explique par la présence de deux régions différentes du promoteur pour le gène BCL2 (142). La première région promotrice (P1) localisée plus en amont, agit comme un promoteur classique, c'est-à-dire qu'elle stimule la transcription. La seconde région (P2), qui

est plus proche du site d'initiation de transcription, contient un « élément de régulation négatif » (143). La région P2 agit sur la région P1 en diminuant son activité promotrice via cet élément régulateur. Il se peut donc qu'une augmentation de l'activité de la région P2 diminue en bout de ligne le niveau d'expression de BCL2, tel qu'observé dans l'étude de Nüchel *et al.* (116). En effet, lorsqu'ils ont testé l'impact du rSNP -225G>T, ils n'ont utilisé qu'une courte région entourant ce rSNP (-228 à +23) pour tester l'activité promotrice, qui ne contient qu'une partie de P2. Cela a permis de mesurer une augmentation de l'activité de la région P2 en présence de l'allèle G par essais de gènes rapporteurs, qui corrèle avec une diminution globale de l'expression du gène avec ce même allèle. Deux autres études ont aussi établi un lien entre ce polymorphisme et le développement du cancer de la peau (108, 144). Il est difficile de savoir si notre étude semble aller dans une direction similaire. En effet, si on compare les haplotypes 1 et 2 de BCL2, la seule différence est au niveau du rSNP -225G>T. On voit alors que l'haplotype 1 qui contient l'allèle G du rSNP -225, a une activité promotrice plus forte que l'haplotype 2 (qui a l'allèle T). Par contre, il ne faut pas oublier que notre design expérimental contient une partie de P1+P2, ce qui signifie que l'expression plus élevée observée dans l'haplotype 1, résultera probablement en un niveau également plus élevé de BCL2. Nos expériences ne montrent donc pas l'allèle G du rSNP -225 comme augmentant une activité répressive, tel que décrit par Nüchel *et al.* (116). Une explication pour cette différence entre nos résultats vient probablement du fait que nous n'avons pas sous-cloné la région P2 au complet, ce qui fait que nous avons peut-être perdu une partie de l'élément répresseur de P2.

Pour ce qui est de l'haplotype 3, il diffère seulement de l'haplotype 2 à la position -1534G>A. Ainsi, la présence de l'allèle A de ce rSNP semble rétablir et même surpasser le niveau d'activité promotrice de l'haplotype 1. Ce rSNP est localisé dans la région promotrice P1. Donc l'allèle A de ce rSNP qui augmente ainsi l'activité promotrice de P1+P2 dans l'haplotype 3 semble augmenter l'activité promotrice de la région P1 et contribuera probablement à augmenter la transcription de BCL2. L'impact de l'allèle A du rSNP -1534 semble important car il doit compenser pour la diminution de l'activité créée par l'allèle T du rSNP-225G>T.

Par ailleurs, on note que le rSNP-225G>T affecte la liaison de protéines nucléaires sur les essais de retard sur gel, et ce dans les 3 lignées cellulaires utilisées (Figure 11). Si l'on teste *in silico* les prédictions de liaisons de facteurs de transcription pour ce polymorphisme, on voit que l'allèle G pourrait lier 4 facteurs de transcription et que 3 de ses facteurs ne pourraient lier l'allèle T du rSNP -225. Cela vient appuyer notre observation de l'activité promotrice plus grande de l'haplotype 1 (ayant l'allèle G) par rapport à l'haplotype 2 (ayant l'allèle T). Malgré le fait que nous n'observons pas de liaison différentielle de protéines pour le rSNP -1534G>A par les essais de retard sur gel, les prédictions *in silico* montrent qu'il y aurait une perte de liaison d'un répresseur transcriptionnel (CTCF) au promoteur en présence de l'allèle A de ce rSNP. La protéine CTCF est impliquée dans une activité « insulator » et empêche l'interaction d'un élément « enhancer » avec son promoteur. Elle permet également de bloquer la propagation de l'hétérochromatine (145). La perte de cette liaison qui pourrait avoir lieu chez

les individus ayant l'haplotype 3 pourrait donc expliquer pourquoi leur activité promotrice est compensée par la présence de l'allèle A du rSNP -1534.

Une augmentation de la quantité de BCL2 (anti-apoptotique) favoriserait l'apparition du cancer, car nous aurions une diminution de l'apoptose. Les individus qui posséderaient donc l'haplotype 3 seraient plus susceptibles de développer un cancer. À cet égard, il est important de souligner que l'allèle A du rSNP -1534 est plus fréquent dans le groupe de patients atteints de LLA que dans les contrôles (Tableau VII). Il sera très pertinent de pousser plus loin cette analyse en étudiant une plus grande cohorte de patients.

PARP1

PARP1 est un substrat des caspases qui est impliqué dans la réparation de l'ADN et l'élongation des télomères (146, 147). Le promoteur de ce gène contient 6 rSNPs, qui ensemble forment 3 haplotypes fréquents (Tableaux IX et X). Nous avons montré qu'il existe une activité promotrice différentielle entre les haplotypes du promoteur de PARP1 (Figure 13). En effet, l'haplotype 1 possède une activité promotrice plus faible que les deux autres haplotypes. En faisant une comparaison de séquence des haplotypes, seuls les polymorphismes -1491A>G et -589A>G ont des allèles différents chez l'haplotype 1 comparativement aux deux autres haplotypes. Cela veut donc dire que se serait un des deux (ou les deux) rSNPs qui seraient responsables de l'activité différentielle observée. Les essais de retard sur gel ont montré qu'il y avait des liaisons différentielles au niveau de ces deux rSNPs et ce dans au moins deux des trois lignées cellulaires testées (Figure 14).

Il n'y a aucune littérature sur des polymorphismes du promoteur de PARP1. Le seul SNP connu dans ce gène comme étant associé au cancer se trouve dans sa région codante (Val762Ala) (148). Ce SNP est également connu pour diminuer l'activité enzymatique de la protéine (149). Nous ne pouvons donc pas comparer nos résultats avec ceux de la littérature. Par contre, nous pouvons proposer que des individus possédant l'haplotype 1 du promoteur, auraient un niveau plus faible de PARP1. Ce niveau plus faible défavorisera la réparation de l'ADN et donc augmentera le taux de mutation, favorisant ainsi la formation de tumeurs. On peut d'ailleurs noter que les allèles G du rSNP-589, T du rSNP -239 et T du rSNP -219, qui sont présents dans l'haplotype 1, sont plus fréquents chez le groupe de patients LLA que dans le groupe contrôles.

Les prédictions *in silico* de liaisons de facteurs de transcription du rSNP -1491A>G ont montré que l'allèle A du rSNP pouvait lier deux familles de facteurs de transcriptions (« NKX homeodomain factors » et « Homeodomain transcription factors ») alors que l'allèle G perdait la liaison de ces facteurs. Le gène PARP1 n'est pas un gène connu pour être régulé par des « homeobox » et n'est pas impliqué dans la morphogenèse et le développement. Comme ce sont des prédictions, celles-ci doivent être confirmées à l'aide d'expériences. Pour ce qui est du rSNP -589A>G, l'allèle A lie 4 facteurs de transcription qui peuvent également se lier à l'allèle G. Toutefois, l'allèle G pourrait former un « core promoter element » appelé XCPE1, qui a récemment été découvert (150). Cet élément aurait une faible activité transcriptionnelle constitutive, mais aurait une beaucoup plus grande activité lorsque combiné à d'autres activateurs. Il serait

également présent dans environ 1% des promoteurs des gènes humains (150). Cet élément pourrait donc contribuer à augmenter l'activité transcriptionnelle tel qu'observé dans les haplotypes 2 et 3.

Limitations de l'étude

Comme notre étude cherche à démontrer qu'il existe une variabilité génétique pouvant influencer le niveau d'expression des gènes et qu'il a été démontré que la plupart de cette variabilité se trouve au niveau de la transcription des gènes, nous avons concentré nos efforts sur la région promotrice des gènes de l'apoptose (92). Nous sommes par contre conscients que cette variabilité peut être expliquée par plusieurs autres facteurs. Elle peut être due à d'autres polymorphismes localisés à l'extérieur de la région promotrice du gène. Elle peut avoir lieu au niveau de la traduction, soit en ayant un niveau de stabilité différent de l'ARNm, soit en ayant un taux de traduction différent (SNPs dans les régions 5'UTR ou 3'UTR). De plus, il se peut que nos régions ciblées (2 kb en amont de l'exon 1) ne correspondent pas tout à fait au promoteur du gène. Nous pouvons donc avoir manqué des polymorphismes qui sont présents dans la région promotrice non étudiée.

Bien que nous observions parfois une différence allélique d'activité promotrice dans nos essais luciférase, nous savons que cela reste des essais *in vitro* et que ces résultats ne reflètent pas nécessairement ce que nous pourrions observer *in vivo*. La régulation génique implique plusieurs facteurs et cofacteurs transcriptionnels ainsi que plusieurs éléments régulateurs qui ne sont pas nécessairement présents dans notre contexte expérimental. De plus, le contexte épigénétique du promoteur

influence en temps normal l'accessibilité de la région pour ces différents facteurs, ce qu'on ne retrouve pas avec nos essais de gènes rapporteurs. Il faut également préciser qu'une différence au niveau de la transcription ne résulte pas nécessairement en une différence de la quantité de protéines, car il existe d'autres niveaux de régulation que ce soit au niveau de la traduction, des modifications post-transcriptionnelles ou post-traductionnelles. Par ailleurs, les essais de gènes rapporteurs introduisent plusieurs copies du vecteur, ce qui nous éloigne du contexte cellulaire normal. L'utilisation de la luciférase *Firefly* comme gène rapporteur n'est pas non plus parfaite, car il a été démontré récemment que ce gène possédait un promoteur cryptique (151). Toutefois, l'activité de ce promoteur est assez faible. Cela peut par contre expliquer en partie le niveau basal d'activité observé avec le contrôle négatif, le vecteur pGL3basic seul.

Même constat pour nos études de retard sur gel qui sont également effectuées dans un contexte *in vitro*. Les protéines de l'extrait nucléaire qui se lient ou non à la sonde, peuvent montrer une liaison totalement différente dans le contexte cellulaire. Il se pourrait qu'*in vivo*, le recrutement de la protéine (ou du complexe protéique) à l'ADN nécessite une interaction avec d'autres facteurs de transcription pour venir se lier au promoteur. Les sites de liaisons de ces autres facteurs de transcription ne sont pas nécessairement inclus dans la séquence de la sonde utilisée dans nos expériences. La perte de ces sites de liaison empêche à son tour le recrutement de la protéine ou du complexe protéique pouvant normalement se lier au niveau de la sonde. De plus, certaines protéines se lient faiblement à l'ADN dans le contexte cellulaire. Ces liaisons faibles entre les protéines et notre

sonde peuvent être brisées dans nos essais de retard sur gel simplement dû au fait que les complexes doivent physiquement migrer sur gel. Aussi, il faut préciser que les conditions utilisées pour tous les essais de retard sur gel sont fixes, ce qui doit également nous faire perdre certaines liaisons. En effet, la quantité de protéines incubées avec la sonde est fixe, les concentrations des tampons utilisés aussi. Ceux-ci contiennent des sels qui peuvent nuire à l'interaction des protéines avec les sondes. Le ratio de compétition entre les sondes froides et marquées est aussi fixe à 50/1, ce qui peut nuire à des liaisons faibles et déplacer rapidement la protéine liée de la sonde marquée. L'utilisation de plusieurs différents ratios de compétition aurait permis de qualifier la liaison (forte ou faible). Finalement, il se peut que nous perdions certains complexes ou certaines protéines lors de notre préparation des extraits nucléaires, dû à la demi-vie assez courte de certaines protéines.

Projets futurs

La présente étude s'inscrit dans le cadre d'un projet visant à identifier les déterminants génétiques de la LLA chez l'enfant. La première étape est donc d'identifier les SNPs régulateurs qui peuvent influencer le niveau d'expression des gènes de l'apoptose par des études fonctionnelles (projet de ce mémoire). Tous les gènes candidats déterminés au cours de cette étude seront éventuellement analysés par d'autres membres du laboratoire. Suite à l'identification des SNPs régulateurs, des études de cas-contrôles seront menées sur 350 patients atteints de LLA ainsi que 350 contrôles d'origine canadienne-française. Ces études permettront d'établir une corrélation entre les différents haplotypes ou SNPs régulateurs observés dans les études fonctionnelles et le risque de développer ce cancer. De plus, l'utilisation de trios parentaux (mère, père, patient) permettra d'étudier l'effet de la génétique des parents sur le risque de développer la leucémie pour l'enfant. Finalement, un autre volet de l'étude sera d'analyser les interactions gène-gène entre les différents gènes de l'apoptose. Ces interactions apportent plus de puissance à l'étude, permettant de regrouper plusieurs rSNPs communs à faible pénétrance présent dans différents gènes. Ces regroupements permettront donc d'établir une association encore plus forte entre la variabilité génétique et la susceptibilité de développer la maladie.

Conclusion

Cette étude permettra d'identifier des polymorphismes ou des haplotypes des promoteurs des gènes d'apoptose qui peuvent avoir un impact sur le niveau d'expression de ces gènes (rSNPs / rHAPs). Cette variabilité pourrait ainsi contribuer à la susceptibilité interindividuelle de développer le cancer. Suite au génotypage de cohortes de patients, il se pourrait qu'on observe une fréquence plus élevée d'un ou de plusieurs rSNPs/rHaps chez ces individus, ce qui permettrait de cibler ces rSNPs comme marqueurs génétiques aidant à prédire la susceptibilité interindividuelle à développer la maladie.

Bibliographie

1. Vermeulen, K., Van Bockstaele, D.R. and Berneman, Z.N. (2005) Apoptosis: mechanisms and relevance in cancer. *Ann Hematol*, **84**, 627-639.
2. Farber, J.L. and El-Mofty, S.K. (1975) The biochemical pathology of liver cell necrosis. *Am J Pathol*, **81**, 237-250.
3. Yuan, J. (1996) Evolutionary conservation of a genetic pathway of programmed cell death. *J Cell Biochem*, **60**, 4-11.
4. Kerr, J.F., Wyllie, A.H. and Currie, A.R. (1972) Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer*, **26**, 239-257.
5. Scavo, L.M., Ertsey, R., Chapin, C.J., Allen, L. and Kitterman, J.A. (1998) Apoptosis in the development of rat and human fetal lungs. *Am J Respir Cell Mol Biol*, **18**, 21-31.
6. Schittny, J.C., Djonov, V., Fine, A. and Burri, P.H. (1998) Programmed cell death contributes to postnatal lung development. *Am J Respir Cell Mol Biol*, **18**, 786-793.
7. Chowdhury, I., Tharakan, B. and Bhat, G.K. (2008) Caspases - an update. *Comp Biochem Physiol B Biochem Mol Biol*, **151**, 10-27.
8. Duriez, P.J. and Shah, G.M. (1997) Cleavage of poly(ADP-ribose) polymerase: a sensitive parameter to study cell death. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **75**, 337-349.
9. Thornberry, N.A. and Lazebnik, Y. (1998) Caspases: enemies within. *Science*, **281**, 1312-1316.
10. Thornberry, N.A., Rano, T.A., Peterson, E.P., Rasper, D.M., Timkey, T., Garcia-Calvo, M., Houtzager, V.M., Nordstrom, P.A., Roy, S., Vaillancourt, J.P. *et al.* (1997) A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. *The Journal of biological chemistry*, **272**, 17907-17911.
11. Ashkenazi, A. and Dixit, V.M. (1998) Death receptors: signaling and modulation. *Science*, **281**, 1305-1308.
12. Pitti, R.M., Marsters, S.A., Ruppert, S., Donahue, C.J., Moore, A. and Ashkenazi, A. (1996) Induction of apoptosis by Apo-2 ligand, a new member of the tumor necrosis factor cytokine family. *The Journal of biological chemistry*, **271**, 12687-12690.

13. Smith, C.A., Farrah, T. and Goodwin, R.G. (1994) The TNF receptor superfamily of cellular and viral proteins: activation, costimulation, and death. *Cell*, **76**, 959-962.
14. Adams, J.M. and Cory, S. (1998) The Bcl-2 protein family: arbiters of cell survival. *Science*, **281**, 1322-1326.
15. Gross, A., McDonnell, J.M. and Korsmeyer, S.J. (1999) BCL-2 family members and the mitochondria in apoptosis. *Genes & development*, **13**, 1899-1911.
16. Kelekar, A. and Thompson, C.B. (1998) Bcl-2-family proteins: the role of the BH3 domain in apoptosis. *Trends Cell Biol*, **8**, 324-330.
17. Kuwana, T. and Newmeyer, D.D. (2003) Bcl-2-family proteins and the role of mitochondria in apoptosis. *Curr Opin Cell Biol*, **15**, 691-699.
18. Wu, X. and Deng, Y. (2002) Bax and BH3-domain-only proteins in p53-mediated apoptosis. *Front Biosci*, **7**, d151-156.
19. Minn, A.J., Swain, R.E., Ma, A. and Thompson, C.B. (1998) Recent progress on the regulation of apoptosis by Bcl-2 family members. *Adv Immunol*, **70**, 245-279.
20. Burlacu, A. (2003) Regulation of apoptosis by Bcl-2 family proteins. *J Cell Mol Med*, **7**, 249-257.
21. Chen, L., Willis, S.N., Wei, A., Smith, B.J., Fletcher, J.I., Hinds, M.G., Colman, P.M., Day, C.L., Adams, J.M. and Huang, D.C. (2005) Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Mol Cell*, **17**, 393-403.
22. Mignotte, B. and Vayssiere, J.L. (1998) Mitochondria and apoptosis. *Eur J Biochem*, **252**, 1-15.
23. Du, C., Fang, M., Li, Y., Li, L. and Wang, X. (2000) Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell*, **102**, 33-42.
24. Mancini, M., Nicholson, D.W., Roy, S., Thornberry, N.A., Peterson, E.P., Casciola-Rosen, L.A. and Rosen, A. (1998) The caspase-3 precursor has a cytosolic and mitochondrial distribution: implications for apoptotic signaling. *The Journal of cell biology*, **140**, 1485-1495.
25. Susin, S.A., Lorenzo, H.K., Zamzami, N., Marzo, I., Brenner, C., Larochette, N., Prevost, M.C., Alzari, P.M. and Kroemer, G. (1999) Mitochondrial release of caspase-2 and -9 during the apoptotic process. *J Exp Med*, **189**, 381-394.
26. Verhagen, A.M., Ekert, P.G., Pakusch, M., Silke, J., Connolly, L.M., Reid, G.E., Moritz, R.L., Simpson, R.J. and Vaux, D.L. (2000) Identification of DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins. *Cell*, **102**, 43-53.

27. Li, L.Y., Luo, X. and Wang, X. (2001) Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature*, **412**, 95-99.
28. Zou, H., Henzel, W.J., Liu, X., Lutschg, A. and Wang, X. (1997) Apaf-1, a human protein homologous to *C. elegans* CED-4, participates in cytochrome c-dependent activation of caspase-3. *Cell*, **90**, 405-413.
29. Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S.M., Ahmad, M., Alnemri, E.S. and Wang, X. (1997) Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell*, **91**, 479-489.
30. Li, H., Zhu, H., Xu, C.J. and Yuan, J. (1998) Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell*, **94**, 491-501.
31. Desagher, S., Osen-Sand, A., Nichols, A., Eskes, R., Montessuit, S., Lauper, S., Maundrell, K., Antonsson, B. and Martinou, J.C. (1999) Bid-induced conformational change of Bax is responsible for mitochondrial cytochrome c release during apoptosis. *The Journal of cell biology*, **144**, 891-901.
32. Zamzami, N., El Hamel, C., Maise, C., Brenner, C., Munoz-Pinedo, C., Belzacq, A.S., Costantini, P., Vieira, H., Loeffler, M., Molle, G. *et al.* (2000) Bid acts on the permeability transition pore complex to induce apoptosis. *Oncogene*, **19**, 6342-6350.
33. Herr, I. and Debatin, K.M. (2001) Cellular stress response and apoptosis in cancer therapy. *Blood*, **98**, 2603-2614.
34. Jaattela, M. (2004) Multiple cell death pathways as regulators of tumour initiation and progression. *Oncogene*, **23**, 2746-2756.
35. Breckenridge, D.G., Germain, M., Mathai, J.P., Nguyen, M. and Shore, G.C. (2003) Regulation of apoptosis by endoplasmic reticulum pathways. *Oncogene*, **22**, 8608-8618.
36. Momoi, T. (2004) Caspases involved in ER stress-mediated cell death. *J Chem Neuroanat*, **28**, 101-105.
37. Green, D.R. and Kroemer, G. (2004) The pathophysiology of mitochondrial cell death. *Science*, **305**, 626-629.
38. Johnson, D.E. (2000) Noncaspase proteases in apoptosis. *Leukemia*, **14**, 1695-1703.
39. Leist, M. and Jaattela, M. (2001) Four deaths and a funeral: from caspases to alternative mechanisms. *Nat Rev Mol Cell Biol*, **2**, 589-598.
40. Fadeel, B., Orrenius, S. and Zhivotovsky, B. (1999) Apoptosis in human disease: a new skin for the old ceremony? *Biochemical and biophysical research communications*, **266**, 699-717.

41. Granville, D.J., Carthy, C.M., Hunt, D.W. and McManus, B.M. (1998) Apoptosis: molecular aspects of cell death and disease. *Lab Invest*, **78**, 893-913.
42. Evan, G.I. and Vousden, K.H. (2001) Proliferation, cell cycle and apoptosis in cancer. *Nature*, **411**, 342-348.
43. Reed, J.C., Miyashita, T., Takayama, S., Wang, H.G., Sato, T., Krajewski, S., Aime-Sempe, C., Bodrug, S., Kitada, S. and Hanada, M. (1996) BCL-2 family proteins: regulators of cell death involved in the pathogenesis of cancer and resistance to therapy. *J Cell Biochem*, **60**, 23-32.
44. Wojcik, I., Szybka, M., Golanska, E., Rieske, P., Blonski, J.Z., Robak, T. and Bartkowiak, J. (2005) Abnormalities of the P53, MDM2, BCL2 and BAX genes in acute leukemias. *Neoplasma*, **52**, 318-324.
45. Kaufmann, S.H., Karp, J.E., Svingen, P.A., Krajewski, S., Burke, P.J., Gore, S.D. and Reed, J.C. (1998) Elevated expression of the apoptotic regulator Mcl-1 at the time of leukemic relapse. *Blood*, **91**, 991-1000.
46. Amarante-Mendes, G.P., McGahon, A.J., Nishioka, W.K., Afar, D.E., Witte, O.N. and Green, D.R. (1998) Bcl-2-independent Bcr-Abl-mediated resistance to apoptosis: protection is correlated with up regulation of Bcl-xL. *Oncogene*, **16**, 1383-1390.
47. Michels, J., Johnson, P.W. and Packham, G. (2005) Mcl-1. *The international journal of biochemistry & cell biology*, **37**, 267-271.
48. Velculescu, V.E., Madden, S.L., Zhang, L., Lash, A.E., Yu, J., Rago, C., Lal, A., Wang, C.J., Beaudry, G.A., Ciriello, K.M. *et al.* (1999) Analysis of human transcriptomes. *Nat Genet*, **23**, 387-388.
49. Ambrosini, G., Adida, C. and Altieri, D.C. (1997) A novel anti-apoptosis gene, survivin, expressed in cancer and lymphoma. *Nat Med*, **3**, 917-921.
50. Li, F. (2005) Role of survivin and its splice variants in tumorigenesis. *Br J Cancer*, **92**, 212-216.
51. Li, F., Ambrosini, G., Chu, E.Y., Plescia, J., Tognin, S., Marchisio, P.C. and Altieri, D.C. (1998) Control of apoptosis and mitotic spindle checkpoint by survivin. *Nature*, **396**, 580-584.
52. Brimmell, M., Mendiola, R., Mangion, J. and Packham, G. (1998) BAX frameshift mutations in cell lines derived from human haemopoietic malignancies are associated with resistance to apoptosis and microsatellite instability. *Oncogene*, **16**, 1803-1812.
53. Rampino, N., Yamamoto, H., Ionov, Y., Li, Y., Sawai, H., Reed, J.C. and Perucho, M. (1997) Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. *Science*, **275**, 967-969.

54. Ong, Y.L., McMullin, M.F., Bailie, K.E., Lappin, T.R., Jones, F.G. and Irvine, A.E. (2000) High bax expression is a good prognostic indicator in acute myeloid leukaemia. *Br J Haematol*, **111**, 182-189.
55. Hogarth, L.A. and Hall, A.G. (1999) Increased BAX expression is associated with an increased risk of relapse in childhood acute lymphocytic leukemia. *Blood*, **93**, 2671-2678.
56. Prokop, A., Wieder, T., Sturm, I., Essmann, F., Seeger, K., Wuchter, C., Ludwig, W.D., Henze, G., Dorken, B. and Daniel, P.T. (2000) Relapse in childhood acute lymphoblastic leukemia is associated with a decrease of the Bax/Bcl-2 ratio and loss of spontaneous caspase-3 processing in vivo. *Leukemia*, **14**, 1606-1613.
57. Srinivas, G., Kusumakumary, P., Nair, M.K., Panicker, K.R. and Pillai, M.R. (2000) Mutant p53 protein, Bcl-2/Bax ratios and apoptosis in paediatric acute lymphoblastic leukaemia. *J Cancer Res Clin Oncol*, **126**, 62-67.
58. Soengas, M.S., Capodici, P., Polsky, D., Mora, J., Esteller, M., Opitz-Araya, X., McCombie, R., Herman, J.G., Gerald, W.L., Lazebnik, Y.A. *et al.* (2001) Inactivation of the apoptosis effector Apaf-1 in malignant melanoma. *Nature*, **409**, 207-211.
59. Miyashita, T. and Reed, J.C. (1995) Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell*, **80**, 293-299.
60. Soengas, M.S., Alarcon, R.M., Yoshida, H., Giaccia, A.J., Hakem, R., Mak, T.W. and Lowe, S.W. (1999) Apaf-1 and caspase-9 in p53-dependent apoptosis and tumor inhibition. *Science*, **284**, 156-159.
61. Amundson, S.A., Myers, T.G. and Fornace, A.J., Jr. (1998) Roles for p53 in growth arrest and apoptosis: putting on the brakes after genotoxic stress. *Oncogene*, **17**, 3287-3299.
62. Levine, A.J. (1997) p53, the cellular gatekeeper for growth and division. *Cell*, **88**, 323-331.
63. Zhou, M., Yeager, A.M., Smith, S.D. and Findley, H.W. (1995) Overexpression of the MDM2 gene by childhood acute lymphoblastic leukemia cells expressing the wild-type p53 gene. *Blood*, **85**, 1608-1614.
64. Bull, E.K., Chakrabarty, S., Brodsky, I. and Haines, D.S. (1998) mdm2-P2 transcript levels predict the functional activity of the p53 tumor suppressor in primary leukemic cells. *Oncogene*, **16**, 2249-2257.
65. Degtarev, A., Lugovskoy, A., Cardone, M., Mulley, B., Wagner, G., Mitchison, T. and Yuan, J. (2001) Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-xL. *Nat Cell Biol*, **3**, 173-182.
66. Lugovskoy, A.A., Degtarev, A.I., Fahmy, A.F., Zhou, P., Gross, J.D., Yuan, J. and Wagner, G. (2002) A novel approach for characterizing

- protein ligand complexes: molecular basis for specificity of small-molecule Bcl-2 inhibitors. *J Am Chem Soc*, **124**, 1234-1240.
67. Wang, J.L., Liu, D., Zhang, Z.J., Shan, S., Han, X., Srinivasula, S.M., Croce, C.M., Alnemri, E.S. and Huang, Z. (2000) Structure-based discovery of an organic compound that binds Bcl-2 protein and induces apoptosis of tumor cells. *Proc Natl Acad Sci U S A*, **97**, 7124-7129.
 68. Wang, J.L., Zhang, Z.J., Choksi, S., Shan, S., Lu, Z., Croce, C.M., Alnemri, E.S., Korngold, R. and Huang, Z. (2000) Cell permeable Bcl-2 binding peptides: a chemical approach to apoptosis induction in tumor cells. *Cancer research*, **60**, 1498-1502.
 69. Banerjee, D. (2001) Genasense (Genta Inc). *Curr Opin Investig Drugs*, **2**, 574-580.
 70. Keith, F.J., Bradbury, D.A., Zhu, Y.M. and Russell, N.H. (1995) Inhibition of bcl-2 with antisense oligonucleotides induces apoptosis and increases the sensitivity of AML blasts to Ara-C. *Leukemia*, **9**, 131-138.
 71. Koty, P.P., Zhang, H. and Levitt, M.L. (1999) Antisense bcl-2 treatment increases programmed cell death in non-small cell lung cancer cell lines. *Lung cancer (Amsterdam, Netherlands)*, **23**, 115-127.
 72. Burke, J.R., Pattoli, M.A., Gregor, K.R., Brassil, P.J., MacMaster, J.F., McIntyre, K.W., Yang, X., Iotzova, V.S., Clarke, W., Strnad, J. *et al.* (2003) BMS-345541 is a highly selective inhibitor of I kappa B kinase that binds at an allosteric site of the enzyme and blocks NF-kappa B-dependent transcription in mice. *The Journal of biological chemistry*, **278**, 1450-1456.
 73. Castro, A.C., Dang, L.C., Soucy, F., Grenier, L., Mazdiyasni, H., Hottelet, M., Parent, L., Pien, C., Palombella, V. and Adams, J. (2003) Novel IKK inhibitors: beta-carbolines. *Bioorg Med Chem Lett*, **13**, 2419-2422.
 74. Dai, Y., Pei, X.Y., Rahmani, M., Conrad, D.H., Dent, P. and Grant, S. (2004) Interruption of the NF-kappaB pathway by Bay 11-7082 promotes UCN-01-mediated mitochondrial dysfunction and apoptosis in human multiple myeloma cells. *Blood*, **103**, 2761-2770.
 75. Kishore, N., Sommers, C., Mathialagan, S., Guzova, J., Yao, M., Hauser, S., Huynh, K., Bonar, S., Mielke, C., Albee, L. *et al.* (2003) A selective IKK-2 inhibitor blocks NF-kappa B-dependent gene expression in interleukin-1 beta-stimulated synovial fibroblasts. *The Journal of biological chemistry*, **278**, 32861-32871.
 76. Rossi, A., Kapahi, P., Natoli, G., Takahashi, T., Chen, Y., Karin, M. and Santoro, M.G. (2000) Anti-inflammatory cyclopentenone prostaglandins are direct inhibitors of IkappaB kinase. *Nature*, **403**, 103-108.
 77. Campos, L., Rouault, J.P., Sabido, O., Oriol, P., Roubi, N., Vasselon, C., Archimbaud, E., Magaud, J.P. and Guyotat, D. (1993) High expression of bcl-2 protein in acute myeloid leukemia cells is associated with poor response to chemotherapy. *Blood*, **81**, 3091-3096.

78. Minn, A.J., Rudin, C.M., Boise, L.H. and Thompson, C.B. (1995) Expression of bcl-xL can confer a multidrug resistance phenotype. *Blood*, **86**, 1903-1910.
79. Miyashita, T. and Reed, J.C. (1993) Bcl-2 oncoprotein blocks chemotherapy-induced apoptosis in a human leukemia cell line. *Blood*, **81**, 151-157.
80. Simonian, P.L., Grillot, D.A. and Nunez, G. (1997) Bcl-2 and Bcl-XL can differentially block chemotherapy-induced cell death. *Blood*, **90**, 1208-1216.
81. Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W. *et al.* (2006) Global variation in copy number in the human genome. *Nature*, **444**, 444-454.
82. Knight, J.C. (2005) Regulatory polymorphisms underlying complex disease traits. *J Mol Med*, **83**, 97-109.
83. Beckmann, J.S., Estivill, X. and Antonarakis, S.E. (2007) Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet*, **8**, 639-646.
84. Karlin, S. and Burge, C. (1995) Dinucleotide relative abundance extremes: a genomic signature. *Trends Genet*, **11**, 283-290.
85. Jarman, A.P. and Wells, R.A. (1989) Hypervariable minisatellites: recombinators or innocent bystanders? *Trends Genet*, **5**, 367-371.
86. Bois, P.R. (2003) Hypermutable minisatellites, a human affair? *Genomics*, **81**, 349-355.
87. Oostra, B.A. and Chiurazzi, P. (2001) The fragile X gene and its function. *Clin Genet*, **60**, 399-408.
88. Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W. and Lee, C. (2004) Detection of large-scale variation in the human genome. *Nat Genet*, **36**, 949-951.
89. Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M. *et al.* (2004) Large-scale copy number polymorphism in the human genome. *Science*, **305**, 525-528.
90. Kanaji, T., Okamura, T., Osaki, K., Kuroiwa, M., Shimoda, K., Hamasaki, N. and Niho, Y. (1998) A common genetic polymorphism (46 C to T substitution) in the 5'-untranslated region of the coagulation factor XII gene is associated with low translation efficiency and decrease in plasma factor XII level. *Blood*, **91**, 2010-2014.
91. Saunders, M.A., Liang, H. and Li, W.H. (2007) Human polymorphism at microRNAs and microRNA target sites. *Proc Natl Acad Sci U S A*, **104**, 3300-3305.

92. Yan, H. and Zhou, W. (2004) Allelic variations in gene expression. *Curr Opin Oncol*, **16**, 39-43.
93. Pagani, F. and Baralle, F.E. (2004) Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet*, **5**, 389-396.
94. Kwan, T., Benovoy, D., Dias, C., Gurd, S., Provencher, C., Beaulieu, P., Hudson, T.J., Sladek, R. and Majewski, J. (2008) Genome-wide analysis of transcript isoform variation in humans. *Nat Genet*, **40**, 225-231.
95. Cargill, M., Altshuler, D., Ireland, J., Sklar, P., Ardlie, K., Patil, N., Shaw, N., Lane, C.R., Lim, E.P., Kalyanaraman, N. *et al.* (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet*, **22**, 231-238.
96. Halushka, M.K., Fan, J.B., Bentley, K., Hsie, L., Shen, N., Weder, A., Cooper, R., Lipshutz, R. and Chakravarti, A. (1999) Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet*, **22**, 239-247.
97. Lander, E.S. (1996) The new genomics: global views of biology. *Science*, **274**, 536-539.
98. Pastinen, T., Sladek, R., Gurd, S., Sammak, A., Ge, B., Lepage, P., Lavergne, K., Villeneuve, A., Gaudin, T., Brandstrom, H. *et al.* (2004) A survey of genetic and epigenetic variation affecting human gene expression. *Physiological genomics*, **16**, 184-193.
99. Yan, H., Yuan, W., Velculescu, V.E., Vogelstein, B. and Kinzler, K.W. (2002) Allelic variation in human gene expression. *Science*, **297**, 1143.
100. Yan, H., Dobbie, Z., Gruber, S.B., Markowitz, S., Romans, K., Giardiello, F.M., Kinzler, K.W. and Vogelstein, B. (2002) Small changes in expression affect predisposition to tumorigenesis. *Nat Genet*, **30**, 25-26.
101. Bond, G.L., Hirshfield, K.M., Kirchhoff, T., Alexe, G., Bond, E.E., Robins, H., Bartel, F., Taubert, H., Wuerl, P., Hait, W. *et al.* (2006) MDM2 SNP309 accelerates tumor formation in a gender-specific and hormone-dependent manner. *Cancer research*, **66**, 5104-5110.
102. Bond, G.L., Hu, W., Bond, E.E., Robins, H., Lutzker, S.G., Arva, N.C., Bargonetti, J., Bartel, F., Taubert, H., Wuerl, P. *et al.* (2004) A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell*, **119**, 591-602.
103. Arva, N.C., Gopen, T.R., Talbott, K.E., Campbell, L.E., Chicas, A., White, D.E., Bond, G.L., Levine, A.J. and Bargonetti, J. (2005) A chromatin-associated and transcriptionally inactive p53-Mdm2 complex occurs in mdm2 SNP309 homozygous cells. *The Journal of biological chemistry*, **280**, 26776-26787.

104. Bond, G.L. and Levine, A.J. (2007) A single nucleotide polymorphism in the p53 pathway interacts with gender, environmental stresses and tumor genetics to influence cancer in humans. *Oncogene*, **26**, 1317-1323.
105. Copson, E.R., White, H.E., Blaydes, J.P., Robinson, D.O., Johnson, P.W. and Eccles, D.M. (2006) Influence of the MDM2 single nucleotide polymorphism SNP309 on tumour development in BRCA1 mutation carriers. *BMC cancer*, **6**, 80.
106. Swinney, R.M., Hsu, S.C., Hirschman, B.A., Chen, T.T. and Tomlinson, G.E. (2005) MDM2 promoter variation and age of diagnosis of acute lymphoblastic leukemia. *Leukemia*, **19**, 1996-1998.
107. Zauberman, A., Flusberg, D., Haupt, Y., Barak, Y. and Oren, M. (1995) A functional p53-responsive intronic promoter is contained within the human mdm2 gene. *Nucleic acids research*, **23**, 2584-2592.
108. Chen, K., Hu, Z., Wang, L.E., Sturgis, E.M., El-Naggar, A.K., Zhang, W. and Wei, Q. (2007) Single-nucleotide polymorphisms at the TP53-binding or responsive promoter regions of BAX and BCL2 genes and risk of squamous cell carcinoma of the head and neck. *Carcinogenesis*, **28**, 2008-2012.
109. Park, J.Y., Park, J.M., Jang, J.S., Choi, J.E., Kim, K.M., Cha, S.I., Kim, C.H., Kang, Y.M., Lee, W.K., Kam, S. *et al.* (2006) Caspase 9 promoter polymorphisms and risk of primary lung cancer. *Human molecular genetics*, **15**, 1963-1971.
110. Jang, J.S., Kim, K.M., Choi, J.E., Cha, S.I., Kim, C.H., Lee, W.K., Kam, S., Jung, T.H. and Park, J.Y. (2008) Identification of polymorphisms in the Caspase-3 gene and their association with lung cancer risk. *Molecular carcinogenesis*, **47**, 383-390.
111. Chen, K., Zhao, H., Hu, Z., Wang, L.E., Zhang, W., Sturgis, E.M. and Wei, Q. (2008) CASP3 polymorphisms and risk of squamous cell carcinoma of the head and neck. *Clin Cancer Res*, **14**, 6343-6349.
112. Jang, J.S., Kim, K.M., Kang, K.H., Choi, J.E., Lee, W.K., Kim, C.H., Kang, Y.M., Kam, S., Kim, I.S., Jun, J.E. *et al.* (2008) Polymorphisms in the survivin gene and the risk of lung cancer. *Lung cancer (Amsterdam, Netherlands)*, **60**, 31-39.
113. Starczynski, J., Pepper, C., Pratt, G., Hooper, L., Thomas, A., Milligan, D., Bentley, P. and Fegan, C. (2005) Common polymorphism G(-248)A in the promoter region of the bax gene results in significantly shorter survival in patients with chronic lymphocytic Leukemia once treatment is initiated. *J Clin Oncol*, **23**, 1514-1521.
114. Saxena, A., Moshynska, O., Sankaran, K., Viswanathan, S. and Sheridan, D.P. (2002) Association of a novel single nucleotide polymorphism, G(-248)A, in the 5'-UTR of BAX gene in chronic lymphocytic leukemia with disease progression and treatment resistance. *Cancer letters*, **187**, 199-205.

115. Enjuanes, A., Benavente, Y., Bosch, F., Martin-Guerrero, I., Colomer, D., Perez-Alvarez, S., Reina, O., Ardanaz, M.T., Jares, P., Garcia-Orad, A. *et al.* (2008) Genetic variants in apoptosis and immunoregulation-related genes are associated with risk of chronic lymphocytic leukemia. *Cancer research*, **68**, 10178-10186.
116. Nuckel, H., Frey, U.H., Bau, M., Sellmann, L., Stanelle, J., Durig, J., Jockel, K.H., Duhrsen, U. and Siffert, W. (2007) Association of a novel regulatory polymorphism (-938C>A) in the BCL2 gene promoter with disease progression and survival in chronic lymphocytic leukemia. *Blood*, **109**, 290-297.
117. Altshuler, D., Daly, M.J. and Lander, E.S. (2008) Genetic mapping in human disease. *Science*, **322**, 881-888.
118. Lewontin, R.C. (1964) The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. *Genetics*, **49**, 49-67.
119. Hudson, R.R. (1985) The sampling distribution of linkage disequilibrium under an infinite allele model without selection. *Genetics*, **109**, 611-631.
120. Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. and Werner, T. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics (Oxford, England)*, **21**, 2933-2942.
121. Ross, M.E., Zhou, X., Song, G., Shurtleff, S.A., Girtman, K., Williams, W.K., Liu, H.C., Mahfouz, R., Raimondi, S.C., Lenny, N. *et al.* (2003) Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood*, **102**, 2951-2959.
122. Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S. *et al.* (2008) Genetics of gene expression and its effect on disease. *Nature*, **452**, 423-428.
123. He, Y.D., Dai, H., Schadt, E.E., Cavet, G., Edwards, S.W., Stepaniants, S.B., Duenwald, S., Kleinhanz, R., Jones, A.R., Shoemaker, D.D. *et al.* (2003) Microarray standard data set and figures of merit for comparing data processing methods and experiment designs. *Bioinformatics (Oxford, England)*, **19**, 956-965.
124. Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., Bennett, H.A., Coffey, E., Dai, H., He, Y.D. *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell*, **102**, 109-126.
125. Juven, T., Barak, Y., Zauberman, A., George, D.L. and Oren, M. (1993) Wild type p53 can mediate sequence-specific transactivation of an internal promoter within the mdm2 gene. *Oncogene*, **8**, 3411-3416.
126. Labuda, D., Labbe, C., Langlois, S., Lefebvre, J.F., Freytag, V., Moreau, C., Sawicki, J., Beaulieu, P., Pastinen, T., Hudson, T.J. *et al.* (2007)

- Patterns of variation in DNA segments upstream of transcription start sites. *Hum Mutat*, **28**, 441-450.
127. Bond, G.L., Hu, W. and Levine, A.J. (2005) MDM2 is a central node in the p53 pathway: 12 years and counting. *Current cancer drug targets*, **5**, 3-8.
 128. Zhang, X., Miao, X., Guo, Y., Tan, W., Zhou, Y., Sun, T., Wang, Y. and Lin, D. (2006) Genetic polymorphisms in cell cycle regulatory genes MDM2 and TP53 are associated with susceptibility to lung cancer. *Hum Mutat*, **27**, 110-117.
 129. Dobbstein, M., Wienzek, S., Konig, C. and Roth, J. (1999) Inactivation of the p53-homologue p73 by the mdm2-oncoprotein. *Oncogene*, **18**, 2101-2106.
 130. Xiao, Z.X., Chen, J., Levine, A.J., Modjtahedi, N., Xing, J., Sellers, W.R. and Livingston, D.M. (1995) Interaction between the retinoblastoma protein and the oncoprotein MDM2. *Nature*, **375**, 694-698.
 131. Zhang, Y., Xiong, Y. and Yarbrough, W.G. (1998) ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell*, **92**, 725-734.
 132. Zhang, Z. and Zhang, R. (2005) p53-independent activities of MDM2 and their relevance to cancer therapy. *Current cancer drug targets*, **5**, 9-20.
 133. Barak, Y., Gottlieb, E., Juven-Gershon, T. and Oren, M. (1994) Regulation of mdm2 expression by p53: alternative promoters produce transcripts with nonidentical translation potential. *Genes & development*, **8**, 1739-1749.
 134. Jones, S.N., Ansari-Lari, M.A., Hancock, A.R., Jones, W.J., Gibbs, R.A., Donehower, L.A. and Bradley, A. (1996) Genomic organization of the mouse double minute 2 gene. *Gene*, **175**, 209-213.
 135. Xiao, G., White, D. and Bargonetti, J. (1998) p53 binds to a constitutively nucleosome free region of the mdm2 gene. *Oncogene*, **16**, 1171-1181.
 136. Lalonde, M., Lariviere, M., Sinnett, D. (2009) Functional variation in the constitutive promoter region of MDM2. *Human Genomics*, In revision.
 137. Wang, M., Zhang, Z., Zhu, H., Fu, G., Wang, S., Wu, D., Zhou, J., Wei, Q. and Zhang, Z. (2008) A novel functional polymorphism C1797G in the MDM2 promoter is associated with risk of bladder cancer in a Chinese population. *Clin Cancer Res*, **14**, 3633-3640.
 138. Bond, G.L., Hu, W. and Levine, A. (2005) A single nucleotide polymorphism in the MDM2 gene: from a molecular and cellular explanation to clinical effect. *Cancer research*, **65**, 5481-5484.
 139. Wilkening, S., Hemminki, K., Rudnai, P., Gurzau, E., Koppova, K., Forsti, A. and Kumar, R. (2007) No association between MDM2 SNP309 promoter polymorphism and basal cell carcinoma of the skin. *The British journal of dermatology*, **157**, 375-377.

140. Zhibin Hu GJ, L.W., Feng Chen, Xinru Wang, and Hongbing Shen (2007) MDM2 Promoter Polymorphism SNP309 Contributes to Tumor Susceptibility: Evidence from 21 Case-Control Studies *Cancer Epidemiol Biomarkers Prev*, **16**, 2717-2723.
141. Lania, L., Majello, B. and De Luca, P. (1997) Transcriptional regulation by the Sp family proteins. *The international journal of biochemistry & cell biology*, **29**, 1313-1323.
142. Seto, M., Jaeger, U., Hockett, R.D., Graninger, W., Bennett, S., Goldman, P. and Korsmeyer, S.J. (1988) Alternative promoters and exons, somatic mutation and deregulation of the Bcl-2-Ig fusion gene in lymphoma. *The EMBO journal*, **7**, 123-131.
143. Young, R.L. and Korsmeyer, S.J. (1993) A negative regulatory element in the bcl-2 5'-untranslated region inhibits expression from an upstream promoter. *Molecular and cellular biology*, **13**, 3686-3697.
144. Lehnerdt, G.F., Franz, P., Bankfalvi, A., Grehl, S., Kelava, A., Nuckel, H., Lang, S., Schmid, K.W., Siffert, W. and Bachmann, H.S. (2009) The regulatory BCL2 promoter polymorphism (-938C>A) is associated with relapse and survival of patients with oropharyngeal squamous cell carcinoma. *Ann Oncol*.
145. Dunn, K.L. and Davie, J.R. (2003) The many roles of the transcriptional regulator CTCF. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **81**, 161-167.
146. Sugimura, K., Takebayashi, S., Taguchi, H., Takeda, S. and Okumura, K. (2008) PARP-1 ensures regulation of replication fork progression by homologous recombination on damaged DNA. *The Journal of cell biology*, **183**, 1203-1212.
147. Beneke, S., Cohausz, O., Malanga, M., Boukamp, P., Althaus, F. and Burkle, A. (2008) Rapid regulation of telomere length is mediated by poly(ADP-ribose) polymerase-1. *Nucleic acids research*, **36**, 6309-6317.
148. Berndt, S.I., Huang, W.Y., Fallin, M.D., Helzlsouer, K.J., Platz, E.A., Weissfeld, J.L., Church, T.R., Welch, R., Chanock, S.J. and Hayes, R.B. (2007) Genetic variation in base excision repair genes and the prevalence of advanced colorectal adenoma. *Cancer research*, **67**, 1395-1404.
149. Wang, X.G., Wang, Z.Q., Tong, W.M. and Shen, Y. (2007) PARP1 Val762Ala polymorphism reduces enzymatic activity. *Biochemical and biophysical research communications*, **354**, 122-126.
150. Tokusumi, Y., Ma, Y., Song, X., Jacobson, R.H. and Takada, S. (2007) The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. *Molecular and cellular biology*, **27**, 1844-1858.

151. Vopalensky, V., Masek, T., Horvath, O., Vicenova, B., Mokejs, M. and Pospisek, M. (2008) Firefly luciferase gene contains a cryptic promoter. *RNA*, **14**, 1720-1729.

Annexes

Tableau I. Amorces pour l'amplification PCR des promoteurs des différents gènes d'apoptose.

Gène	Amorces Sens (F) et anti-sens (R) ¹
MDM2	GGGACAAGTTTGTACAAAAAAGCAGGCTAAAGCCCAAATTCCTTGCT (F) GGGACCACCTTTGTACAAGAAAGCTGGGTCTCCATCTTCCGACACACA (R)
Délétion MDM2	TTTCCTTTCTGGTAGGCTGG (F) CACCTACTTTCCACAGAGA(R)
CASP3	GGGACCACCTTTGTACAAGAAAGCTGGGTCTCTGGGAAGAAGAGGAGCA (F) GGGACAAGTTTGTACAAAAAAGCAGGCTAGTCGAATAGGCCAAGTGT (R)
XIAP	GGGACAAGTTTGTACAAAAAAGCAGGCTCTGGGGTTTCACCGTGTTAG (F) GGGACCACCTTTGTACAAGAAAGCTGGGTCAGCCTAGGTGAAGG AAACG (R)
BCL2	GGGACCACCTTTGTACAAGAAAGCTGGGTGGCGGCAGATGAATTACAA (F) GGGACAAGTTTGTACAAAAAAGCAGGCTTTTGCCTCGAGCCCTATTA (R)
BAX	P1 = GGGACAAGTTTGTACAAAAAAGCAGGCTGAGAAGTGCCACTAGAGTTC (F) GGGACCACCTTTGTACAAGAAAGCTGGGTCCGGCCGGGTATACACATA (R) P3= GGGACAAGTTTGTACAAAAAAGCAGGCTTCGCTTGAGCTGGGAGTTC (F) GGGACCACCTTTGTACAAGAAAGCTGGGTCTGAACGTGCGTCTTAC (R)
BCL2L10	GGGACCACCTTTGTACAAGAAAGCTGGGTGCCCCGGTATATTGTTTT (F) GGGACAAGTTTGTACAAAAAAGCAGGCTGGACAGCCTTGAGTTTCTGC (R)
BCL2L11	GGGACAAGTTTGTACAAAAAAGCAGGCTCCCCGACTCTTCTTTTTCAA (F) GGGACCACCTTTGTACAAGAAAGCTGGGTTCCAACAACTGCAGACCAG (R)
PARP1	GGGACCACCTTTGTACAAGAAAGCTGGGTCCGTTCCCTGATAGATTGCT (F) GGGACAAGTTTGTACAAAAAAGCAGGCTGAGTGGCCATTTAGGGATGA (R)
BAG4	GGGACAAGTTTGTACAAAAAAGCAGGCTGATGAGGGCAAAAAGAGCATC (F) GGGACCACCTTTGTACAAGAAAGCTGGGTGTTTCTCGCCATTTATCCA (R)
ENDO G	GGGACAAGTTTGTACAAAAAAGCAGGCTCTGAGGCCACAGATCACTT (F) GGGACCACCTTTGTACAAGAAAGCTGGGTTTTAAGACGCCGGTATGAGG (R)
BCL2L1	GGGACCACCTTTGTACAAGAAAGCTGGGTGAATTGCGAAGCTCAGGAAC (F) GGGACAAGTTTGTACAAAAAAGCAGGCTCACCACCCTCAAAGAAAGA (R)
TP73	GGGACCACCTTTGTACAAGAAAGCTGGGTATACGGTGTCTGTGCGCAT (F) GGGACAAGTTTGTACAAAAAAGCAGGCTGGCGTTAGCGCTTTTTAG (R)

¹ Les nucléotides en **gras** sont des queues servant au système de recombinaison

Gateway (Invitrogen)

Tableau II. Conditions et programmes de PCR pour l'amplification des régions promotrices.

Gène	Conditions de PCR*	Programme de PCR
MDM2	Promoteur total =B Délétion= A	Promoteur =95°C/10', 10 cycles (95°C/30'', 55-50°C/30'', 68°C/2'), 25 cycles (95°C/30'', 50°C/30'', 68°C/2'), 68°C/7' Délétion = 95°C/3', 32 cycles (95°C/30'', 52°C/30'', 72°C/20''), 72°C/7'
CASP3	C	95°C/10', 14 cycles (95°C/45'', 60-53°C/45'', 72°C/1'20''), 25 cycles (95°C/45'', 53°C/45'', 72°C/1'20''), 72°C/7'
XIAP	B	95°C/10', 14 cycles (95°C/30'', 60-53°C/30'', 68°C/2'), 25 cycles (95°C/30'', 53°C/30'', 68°C/2'), 68°C/7'
BCL2	B	95°C/10', 14 cycles (95°C/30'', 60-53°C/30'', 68°C/2'), 25 cycles (95°C/30'', 53°C/30'', 68°C/2'), 68°C/7'
BAX	P1= B P3= C	95°C/10', 14 cycles (95°C/30'', 62-55°C/30'', 68°C/1'), 25 cycles (95°C/30'', 55°C/30'', 68°C/1'), 68°C/7'
BCL2L10	C	95°C/10', 14 cycles (95°C/30'', 62-55°C/30'', 68°C/2'), 25 cycles (95°C/30'', 55°C/30'', 68°C/2'), 68°C/7'
BCL2L11	A	95°C/10', 14 cycles (95°C/30'', 62-55°C/30'', 68°C/2'), 25 cycles (95°C/30'', 55°C/30'', 68°C/2'), 68°C/7'
PARP1	C	95°C/10', 14 cycles

		(95°C/30'', 62-55°C/30'', 68°C/2'), 25 cycles (95°C/30'', 55°C/30'', 68°C/2'), 68°C/7'
BAG4	A	95°C/10', 14 cycles (95°C/30'', 60-53°C/30'', 68°C/2'), 25 cycles (95°C/30'', 53°C/30'', 68°C/2'), 68°C/7'
ENDOG	A	95°C/10', 14 cycles (95°C/30'', 60-53°C/30'', 68°C/2'), 25 cycles (95°C/30'', 53°C/30'', 68°C/2'), 68°C/7'
BCL2L1	C	95°C/10', 14 cycles (95°C/30'', 62-55°C/30'', 68°C/2'), 25 cycles (95°C/30'', 55°C/30'', 68°C/2'), 68°C/7'
TP73	A	95°C/10', 14 cycles (95°C/30'', 62-55°C/30'', 68°C/2'), 25 cycles (95°C/30'', 55°C/30'', 68°C/2'), 68°C/7'

* A = 20 pmoles des amorces F et R, 1,5mM de MgCl₂, 0,2 mM dNTPs, 1x de tampon Fast Start Taq (Roche), 1X de tampon GC rich (Roche), 2 unités d'ADN polymérase Fast Start Taq (Roche) et 15 ng d'ADN génomique dans un volume total de 50µL.

B = 20 pmoles des amorces F et R, 2mM de MgCl₂, 0,2 mM dNTPs, 1x de tampon Fast Start Taq (Roche), 1X de tampon GC rich (Roche), 2 unités d'ADN polymérase Fast Start Taq (Roche) et 15 ng d'ADN génomique dans un volume total de 50µL.

C = 20 pmoles des amorces F et R, 3mM de MgCl₂, 0,2 mM dNTPs, 1x de tampon Fast Start Taq (Roche), 1X de tampon GC rich (Roche), 2 unités d'ADN polymérase Fast Start Taq (Roche) et 15 ng d'ADN génomique dans un volume total de 50µL.

Tableau III. Séquences des oligonucléotides allèle-spécifiques utilisés pour le génotypage de MDM2 et CASP3.

Gène	SNP*	séquence
MDM2	-1494 G -1494 A	TCGTGATCCGCCACCTCA CTCGTGATCCACCCACCTCA
	-1166 T -1166 G -1164 C -1164 G	AAAAGCTGCAGAAGGGAAGGA AAAAGCTGGAGAAGGGAAGGA AAAAGCGGCAGAAGGGAAGGA AAAAAAAAAAAAAGCGGGAGAAGGG
	-737 C -737 T	CCACCACGCCGACTAATTTT CCACCACGCCTGACTAATTTT
	-683 G -683 A	TGGCCAGGTTGGTCTTGAAC TGGCCAGGTTAGTCTTGAAC
	-182 C -182 G	ACCCCGCCTCACAGCCC ACCCCGCCTCAGAGCCC
	-151 C -151 T	GACACCCCCACTCCATCAT GACACCCCTACTCCATCAT
CASP3	-1912 CTTT -1912 -	AGGCCTTTCTTTATCCCTCC AGGCCTTTATCCCTCCTGAA
	-1764 A -1764 G	CGCAATCTCAGCTCACTGCA CGCAATCTCGGCTCACTGCA
	-1689 A -1689 G	AGTGGTGGGCACCTGTAATC AGTGGTGGGCGCCTGTAATC
	-1629 A -1629 G -1624 C -1624 T	GCCTGGTCAAGATGGTGAAA GCCTGGTCAAGGTGGTGAAA GCCTGGCCAAGATGGTGAAA GCCTGGCCAAGGTGGTGAAA
	-1309 G -1309 C	CCATGGCAGAGCCTGTTTC CCATGGCACAGCCTGTTTC
	-1147 C -1147 T	ACTGCCCTACCGATCCCCTT ACTGCCCTACTGATCCCCTT
	-928 C -928 T	GCCGAAGGCCCATCTTTTAT GCCGAAGGCCTCATCTTTTAT
	-494 C -494 T	TCAGGTCTCTCCCTAAATC TCAGGTCTCTTCCCTAAATC

*Les positions des SNPs sont calculées en prenant le début de l'exon 1 comme étant la position +1.

Tableau IV. Séquences des oligonucléotides utilisés pour les essais de retard sur gel.

Gène	SNP*	Séquence**
MDM2	-1494 A	CTCGTGATCCACCCACCTCAG
	-1494 G	CTCGTGATCCGCCACCTCAG
	-182 C	GCCAACACCCCCACCCCGCCTCACAGCCCGCCGC GCCCCGCGGGGCGACA,
-182 G	GCCAACACCCCCACCCCGCCTCAGAGCCCGCCGC GCCCCGCGGGGCGACA	
+309 G	CCGCTGCGGCGCGGGAGGTCCGGA,	
+309 T	CCGCTTCGGCGCGGGAGGTCCGGA	
BCL2	-1534 G	CCGCCCCCACC GCCCCAGCAGTGGGACCTCAG CGCTGCCCTGCTGTGA,
	-1534 A	CCGCCCCCACC GCCCCAGCAGTGGGACCTCAG CGCTGCCCTGCTGTGA
	-225 G	GAGAGAGACAGGGGAGAGGGGACGATGAAGG,
-225 T	GAGAGAGACAGGGGAGATGGGACGATGAAGG	
PARP1	-1491 A	CGGGAGGCTGAAGTGGGAGGATCACC,
	-1491 G	CGGGAGGCTGAGGTGGGAGGATCACC
	-589 A	CAGCTGGGTCCGGGAAGCGCAGGCCCCCGCCTC GGGAATATA,
-589 G	CAGCTGGGTCCGGGAAGCGCGGGCCCCCGCCTC GGGAATATA	
contrôle de spécificité	AP2	GATCGAACTGACCGCTTGCGGCCCGT

* Les positions des SNPs sont calculées en prenant le début de l'exon 1 comme étant la position +1

**Seule la séquence du brin sens est donnée.

Tableau V. SNPs du promoteur de CASP3 avec leur position génomique et leur fréquence allélique observée dans notre cohorte.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
-1912 CTTT>delCTTT	chr4:185,809,535	rs10560954	0%	
-1764A>G	chr4:185,809,387	rs2696040	0%	
-1689A>G	chr4:185,809,312	rs2705880	0%	
-1629 A>G	chr4:185,809,252	rs7693625	0%	
-1624 C>T	chr4:185,809,247	rs2720383	0%	
-1309 G>C	chr4:185,809,159	rs35741846	18%	
-1147 C>G	chr4:185,808,770	rs12507711	15%	
-928C>T	chr4:185,808,551	rs12108497	33%	
-494C>T	chr4:185,808,117	rs12506750	18%	

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

Tableau VI. SNPs du promoteur de BAX avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1916T>G	chr19:54,148,013	non répertorié	5%	12,5%
-1837T>A	chr19:54,148,092	rs11671610	7,9%	11,8%
-1820G>A	chr19:54,148,109	rs12983717	2,6%	11,8%
-1766G>A	chr19:54,148,163	non répertorié	2,6%	2,9%
-1276C>T	chr19:54,148,653	non répertorié	0%	2,9%
-179A>G	chr19:54,149,750	rs4645878	2,5%	13,6%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

Tableau VII. SNPs du promoteur d'ENDOG avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1598C>T	chr9:130,619,003	rs3115875	52,4%	50%
-1572A>G	chr9:130,619,029	non répertorié	5%	2,5%
-1259A>C	chr9:130,619,342	non répertorié	5%	2,5%
-1206G>C	chr9:130,619,395	rs56940912	5%	5%
-1201A>G	chr9:130,619,400	rs10988130	38,1%	50%
-469T>A	chr9:130,620,132	rs2997923	38,1%	33,3%
-378A>G	chr9:130,620,223	rs12380423	40,2%	38,5%
-293G>A	chr9:130,620,308	non répertorié	20%	30%
-210T>C	chr9:130,620,391	rs2977998	55%	61,1%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1

Tableau VIII. SNPs du promoteur de BAG4 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1103A>C	chr8:38,152,161	rs17435276	21,1%	22,2%
-118G>C	chr8:38,153,146	rs2270376	22,5%	25%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

Tableau IX. SNPs du promoteur de BCL2L1 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-741A>C	chr20:29,775,057	rs45443696	21,1%	22,7%
-603G>C	chr20:29,774,919	rs6121209	25%	21,9%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

Tableau X. SNPs du promoteur de BCL2L10 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1687C>T	chr15:50,193,950	rs2414132		
-954A>G	chr15:50,193,217	rs12912428	87,5%	83,3%
-692C>T	chr15:50,192,955	rs2231287	12,5%	22,2%
-493T>C	chr15:50,192,756	rs2231289	20%	11,1%
-330A>G	chr15:50,192,593	rs2231290	10,5%	0%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

Tableau XI. SNPs du promoteur de BCL2L11 avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1728G>A	chr2:111,593,234	rs7577824	38,2%	50%
-1550A>C	chr2:111,593,412	rs7593638	38,2%	50%
-1387C>T	chr2:111,593,575	rs56952027	38,2%	54,5%
-1194T>C	chr2:111,593,768	rs4848393	41,2%	54,5%
-1140->insG	chr2:111,593,822	rs11378324	41,2%	54,5%
-274G>A	chr2:111,594,698	rs7582030	8,8%	22,7%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

Tableau XII. SNP du promoteur de TP73 avec sa position génomique et sa fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-968G>T	chr1:3,558,022	rs1004650	5,2%	5,5%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

Tableau XIII. SNPs du promoteur de XIAP avec leur position génomique et leur fréquence allélique observée dans notre cohorte de contrôles et de patients.

SNP ⁰	Position chromosomique*	# rs	Fréquence allélique	
			Contrôles	Patients
-1352C>G	chrX:122,820,377	rs5956578		
-1282C>T	chrX:122,820,447	rs28382699	65,7%	57,9%
-1257A>G	chrX:122,820,472	non répertorié	2,6%	0%
-1226T>C	chrX:122,820,503	rs5958318	5,2%	0%
-1170C>T	chrX:122,820,559	rs12687176	55,3%	38,8%
-303C>T	chrX:122,821,426	rs28382701	2,6%	0%

⁰ Position calculée en considérant le début de la transcription comme étant le +1.

* NCBI Build 36.1.

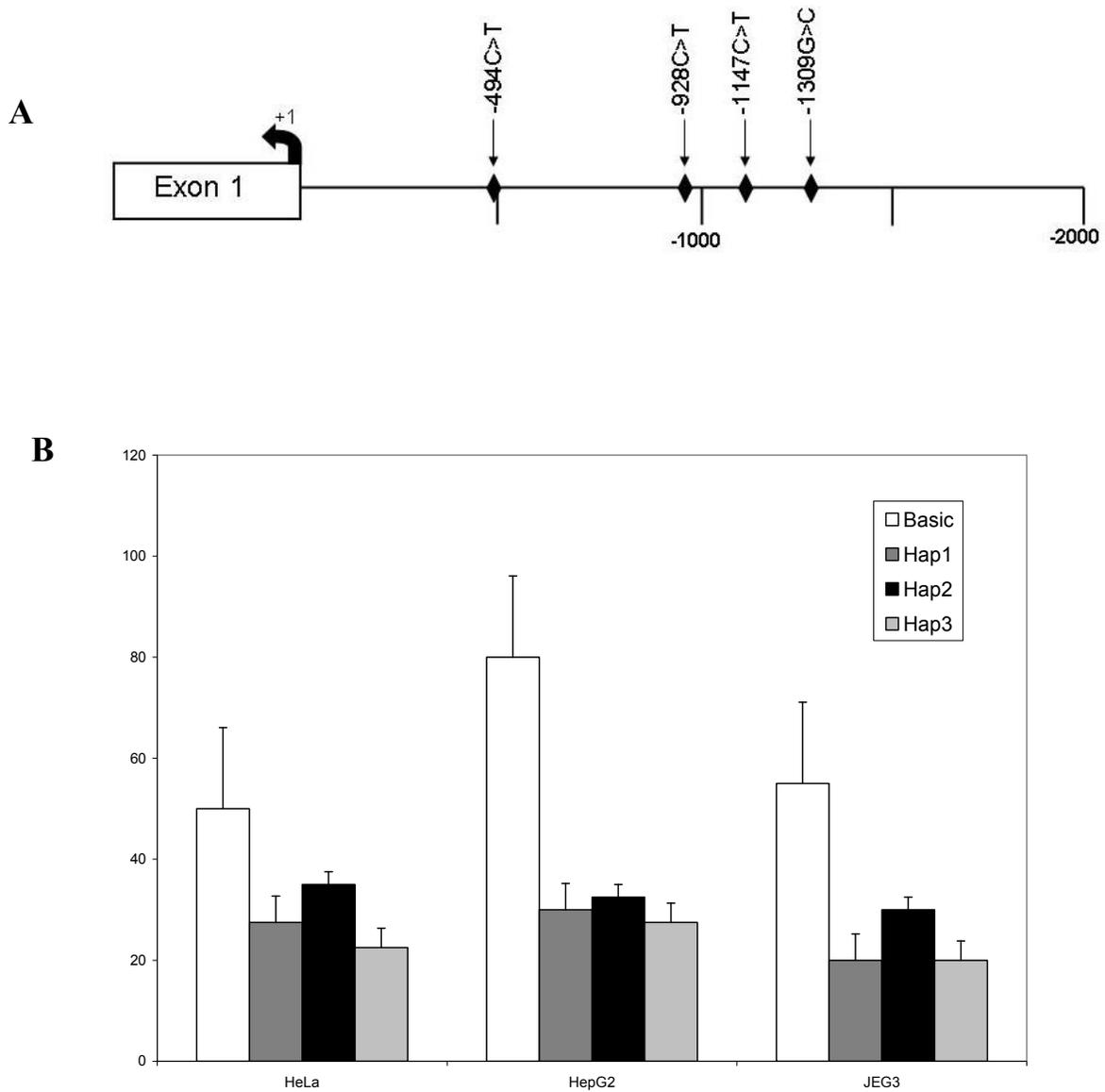


Figure 1. (A) Schéma du promoteur de CASP3 (B) Impact fonctionnel des haplotypes du promoteur de CASP3 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

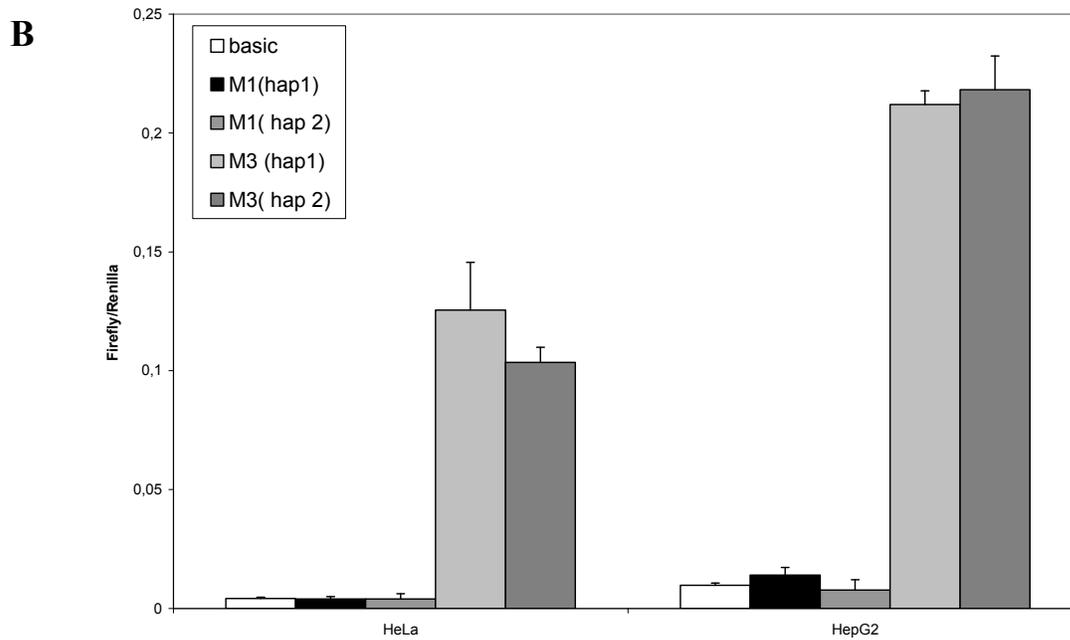
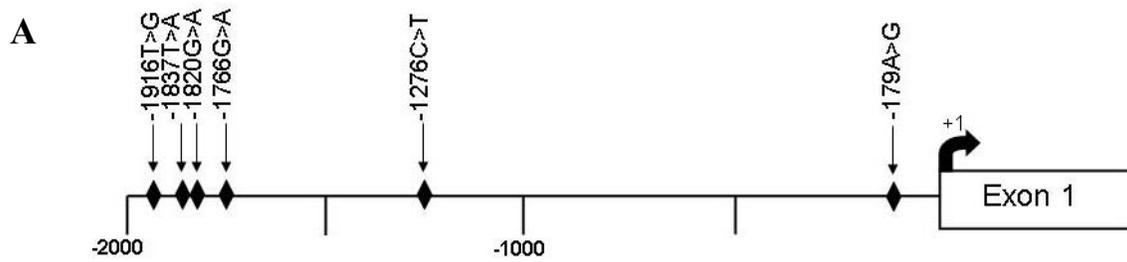


Figure 2. (A) Schéma du promoteur de BAX (B) Impact fonctionnel des haplotypes du promoteur de BAX sur l'activité promotrice. La transfection est faite dans 2 lignées cellulaires différentes (HeLa et HepG2). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

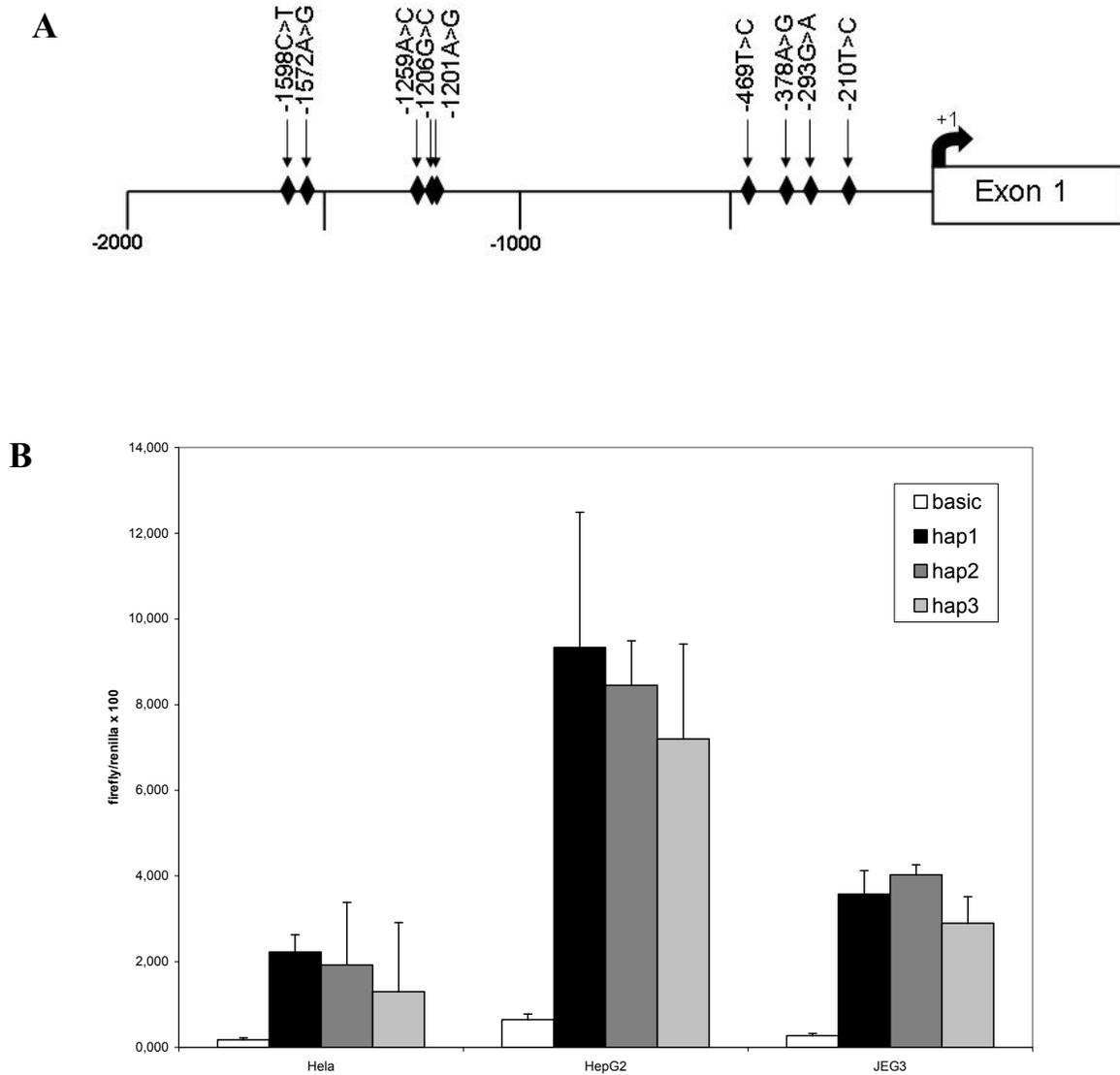


Figure 3. (A) Schéma du promoteur d'ENDOG (B) Impact fonctionnel des haplotypes du promoteur d'ENDOG sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

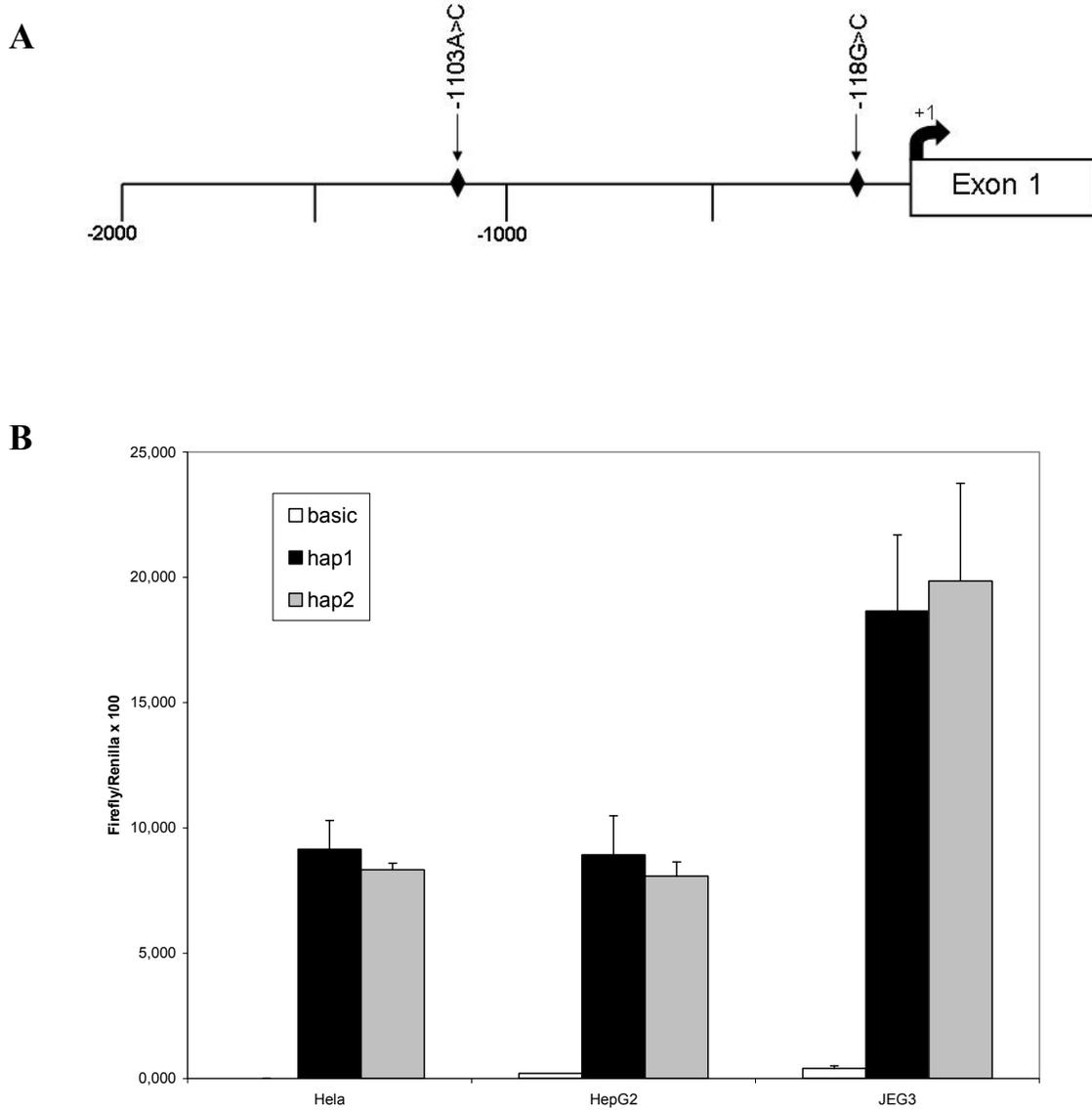


Figure 4. (A) Schéma du promoteur de BAG4 (B) Impact fonctionnel des haplotypes du promoteur de BAG4 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

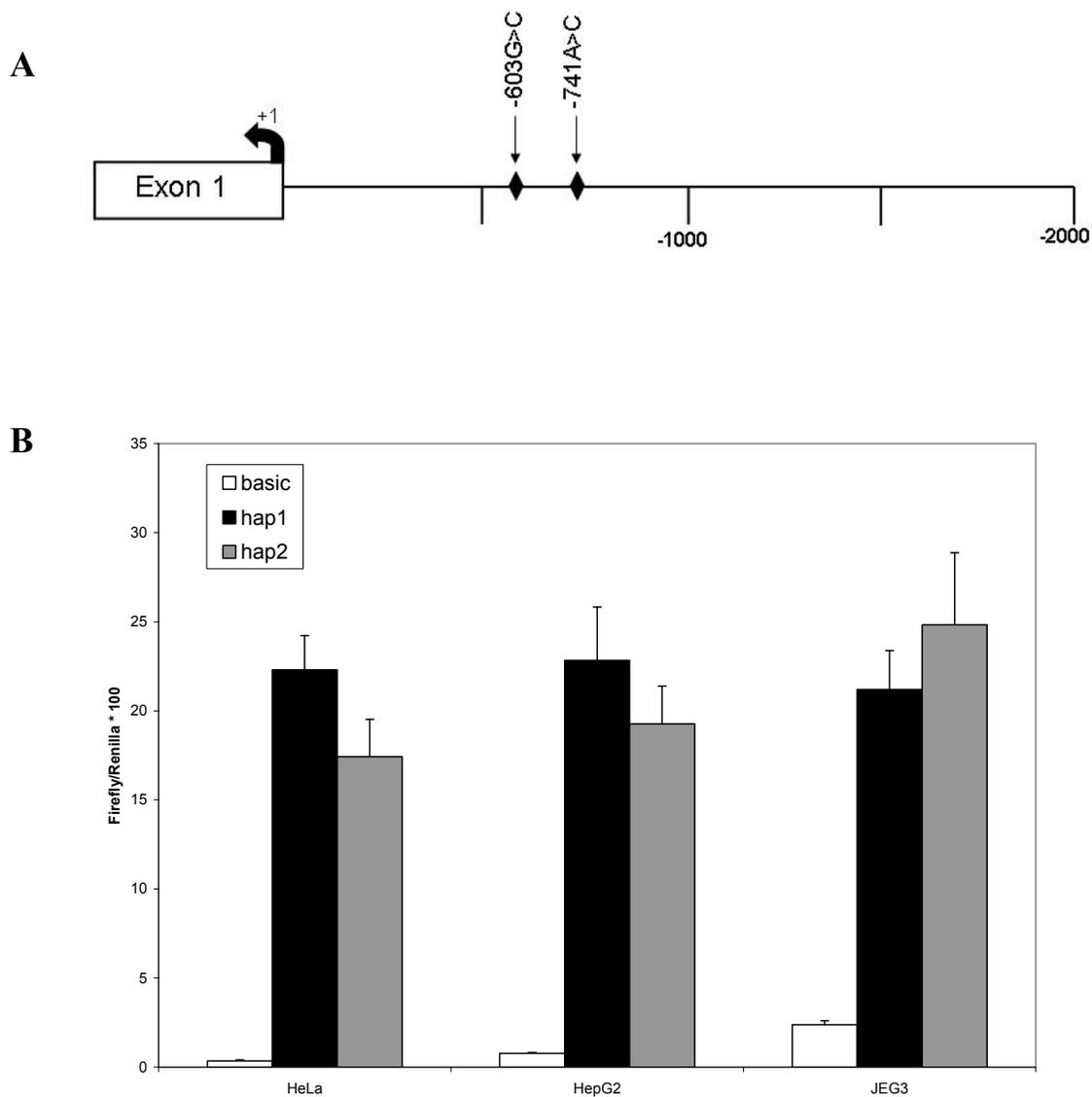


Figure 5. (A) Schéma du promoteur de BCL2L1 (B) Impact fonctionnel des haplotypes du promoteur de BCL2L1 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

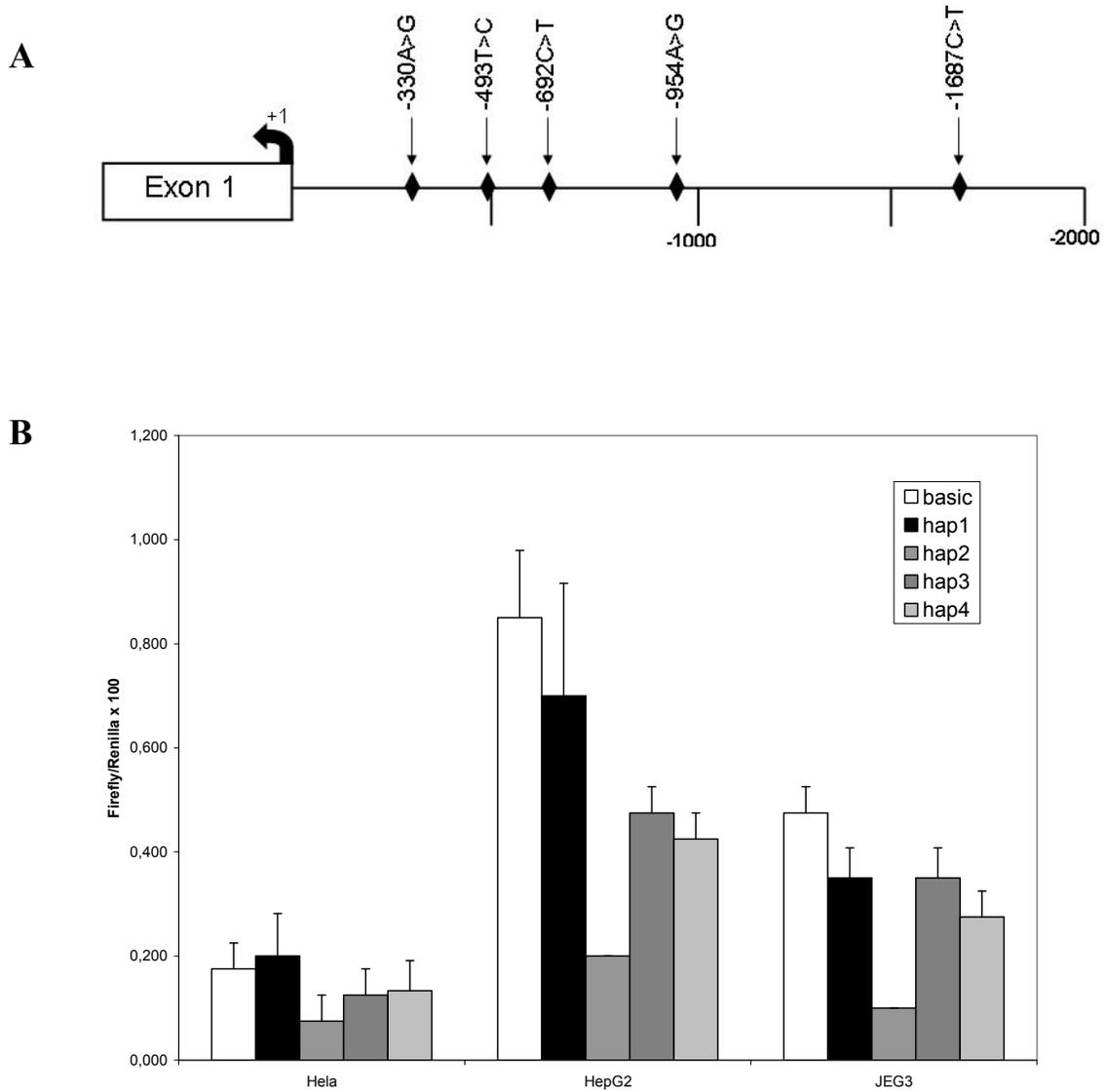


Figure 6. (A) Schéma du promoteur de BCL2L10 (B) Impact fonctionnel des haplotypes du promoteur de BCL2L10 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

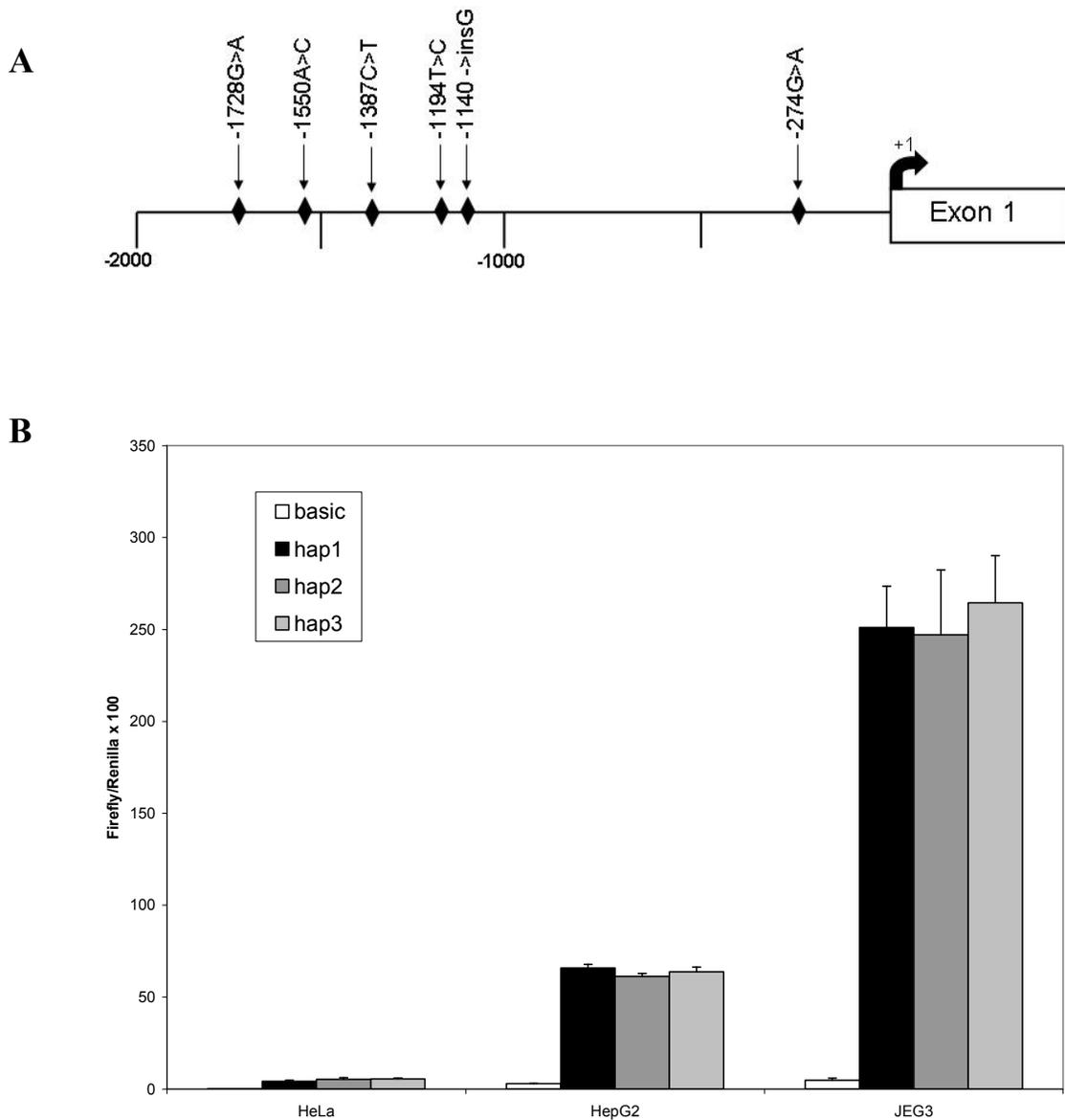


Figure 7. (A) Schéma du promoteur de BCL2L11 (B) Impact fonctionnel des haplotypes du promoteur de BCL2L11 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

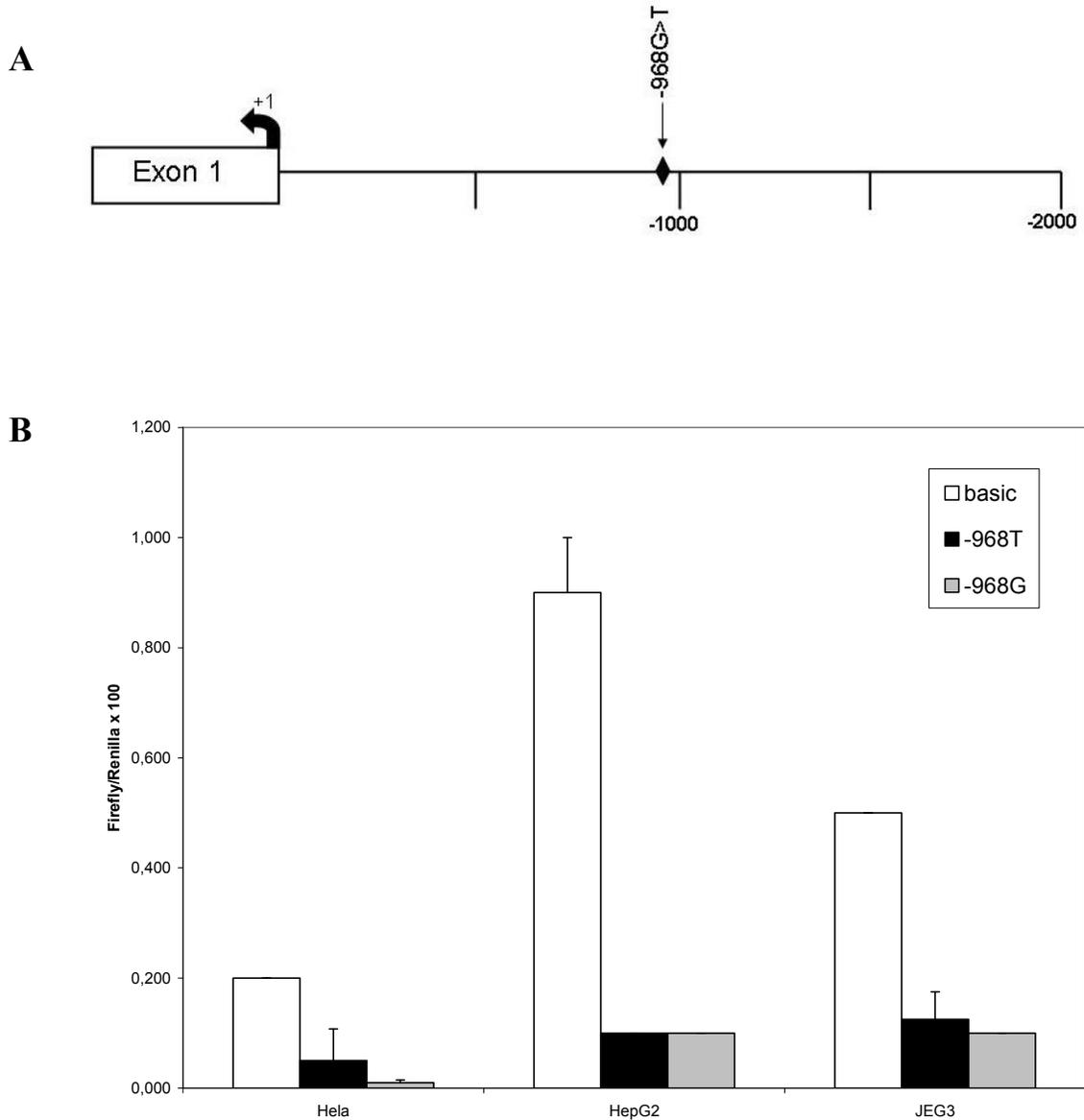


Figure 8. (A) Schéma du promoteur de TP73 (B) Impact fonctionnel des haplotypes du promoteur de TP73 sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla multiplié par 100. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

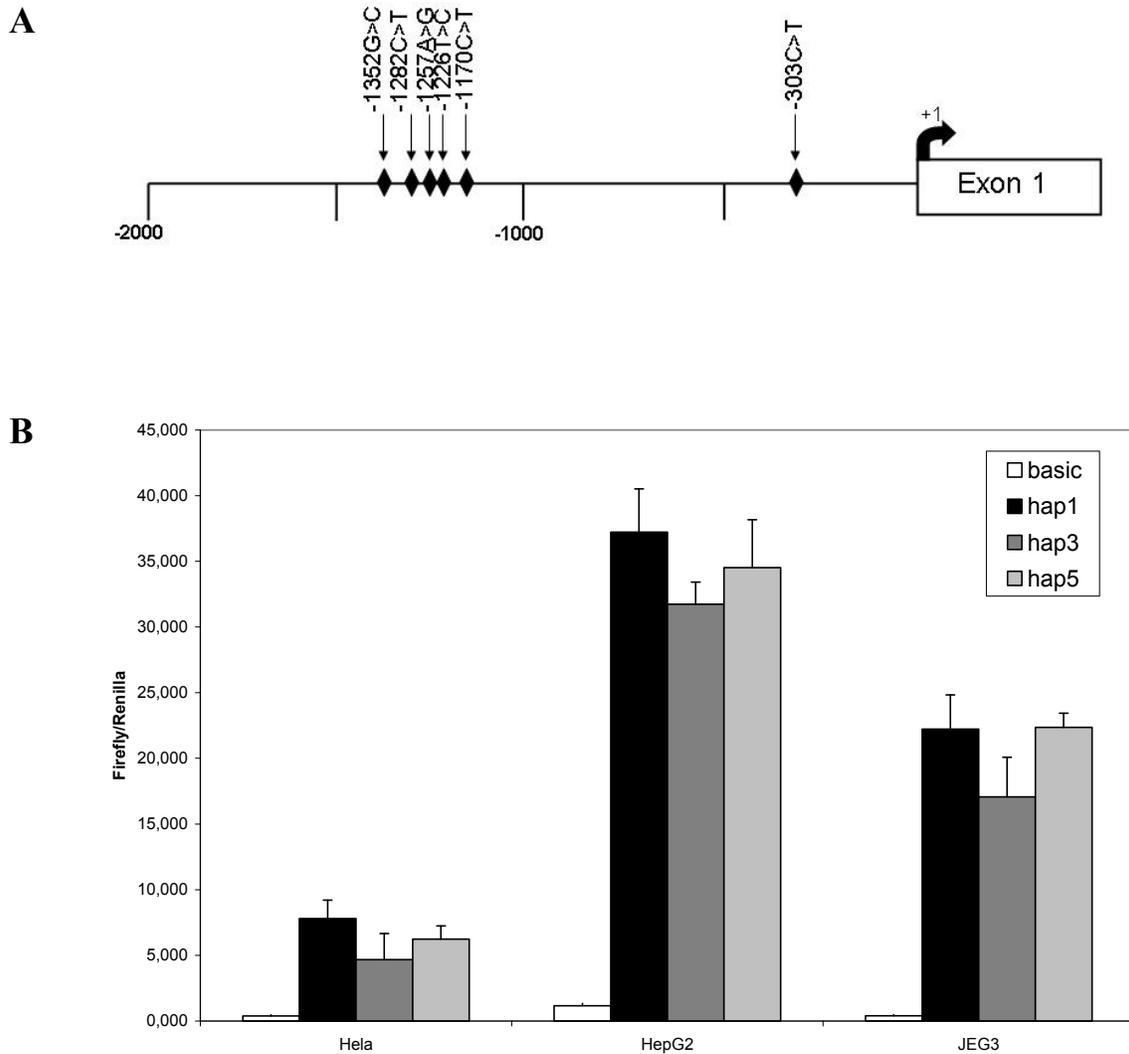


Figure 9. (A) Schéma du promoteur de XIAP (B) Impact fonctionnel des haplotypes du promoteur de XIAP sur l'activité promotrice. La transfection est faite dans 3 lignées cellulaires différentes (HeLa, HepG2 et JEG3). Les résultats illustrés sont d'une expérience représentative de 3 expériences indépendantes. Les résultats sont exprimés comme ratio de l'activité de la Firefly / Renilla. Le vecteur pGL3basic seul est utilisé comme contrôle négatif (Basic).

Tableau XIV. Haplotypes fréquents de la région promotrice de CASP3.

Haplotype	-1309G>C	-1147C>T	-928C>T	-494C>T	Fréquence
1	G	C	C	C	58%
2	G	T	T	T	17%
3	C	C	T	C	13%

Tableau XV. Haplotypes fréquents de la région promotrice de BAX.

Haplotype	-1916T>G	-1837A>T	-1820C>T	-179G>A	Fréquence
1	T	A	G	G	86%
2	G	T	A	A	8%

Tableau XVI. Haplotypes fréquents de la région promotrice d'ENDOG.

Haplotype	-1598 C>T	-1201 A>G	-469 T>A	-378 A>G	-293 G>A	-210 T>C	Fréquence
1	T	A	A	G	G	C	38,8%
2	C	G	T	A	A	T	21,3%
3	C	G	T	A	G	T	26,3%

Tableau XVII. Haplotypes fréquents de la région promotrice de BAG4.

Haplotype	-1103A>C	-118G>C	Fréquence
1	A	G	79,5%
2	C	C	17,9%

Tableau XVIII. Haplotypes fréquents de la région promotrice de BCL2L1.

Haplotype	-741A>C	-603G>C	Fréquence
1	A	G	78,3%
2	C	C	21,6%

Tableau XIX. Haplotypes fréquents de la région promotrice de BCL2L10.

Haplotype	-1687 C>T	-954 A>G	-692 C>T	-493 T>C	-330 A>G	Fréquence
1	C	A	C	T	A	64,5%
2	C	G	C	T	A	5,3%
3	T	A	C	C	G	14,5%
4	T	A	T	T	A	10,5%

Tableau XX. Haplotypes fréquents de la région promotrice de BCL2L11.

Haploype	-1728 G>A	-1550 A>C	-1387 C>T	-1194 T>C	-1140 ->insG	-274 G>A	Fréquence
1	A	C	T	C	insG	G	29,3%
2	A	C	T	C	insG	A	13,8%
3	G	A	C	T	-	G	55,2%

Tableau XXI. Haplotypes fréquents pour la région promotrice de XIAP.

Haploype	-1352C>G	-1226T>C	Fréquence
1	C	T	41,4%
2	G	T	7,1%
3	G	C	42,8%

Functional variation in the constitutive promoter region of MDM2.

Marie-Eve Lalonde¹, Mathieu Larivière¹, and Daniel Sinnett^{1,2}

¹Division of Hematology-Oncology, Research Center, Sainte-Justine Hospital;²Department of Pediatrics, Faculty of Medicine, University of Montreal, Montreal, Quebec, H3T 1C5, Canada.

Correspondence: D Sinnett, Ph.D., Division of Hematology-Oncology, Research Center, Sainte-Justine Hospital, 3175 chemin de la Côte-Sainte-Catherine, Montreal, H3T 1C5, Canada. Phone: (514) 345-4931 # 2990 Fax: (514) 345-4731.

Abstract

Mutations in the oncoprotein MDM2 are rare, but overexpression of this gene is observed in a number of human tumors. Because it is suggested that even modest modifications of MDM2 levels might affect the p53 pathway, we propose that sequence variation in the promoter region of MDM2 could lead to abnormal expression and variation in gene dosage. Two promoters have been reported for MDM2, an internal promoter (P2) located near the end of intron 1 that is p53-responsive, and an upstream constitutive promoter (P1) that is p53-independent. Both MDM2 promoter regions contain DNA variants that could influence its expression levels, including the well studied SNP309, which is located in the p53-dependant promoter P2. In this report we screened the promoter P1 region for DNA variants and assessed the functional impact of the promoter SNPs. The search in the public database dbSNP and genotyping validation in individuals of European descent led to the identification of three common SNPs (-1494G>A; indel 40bp; -182C>G). Three major promoter haplotypes were inferred by using the 3 promoter SNPs together with SNP309. When tested by electrophoretic mobility shift assay, the SNPs -1494G>A, -182C>G and +309T>G all showed differential allele-specific DNA-protein binding in the cell lines tested. Following subcloning into a gene reporter system, we found that the two major P1 promoter haplotypes significantly influenced MDM2 transcriptional activity in a haplotype-specific manner. This study suggests that part of the variability in the *in vivo* MDM2 expression levels could be explained by allelic p53-independent P1 promoter activity.

Introduction

The transcription factor p53 is mutated in more than 50% human tumors.¹ In a nonstressed environment, p53 is kept at low concentration in cells. When cells are exposed to certain stress, for instance DNA damage, p53 is activated and subsequently regulates transcription of target genes.² that are involved in the regulation of biological processes including cell division, growth arrest or apoptosis.³ MDM2 is a key component of the p53 pathway and even modest modifications of MDM2 levels might affect the p53 pathway.⁴ The gene MDM2 encodes an ubiquitine E3 ligase that is a negative regulator of p53.^{5,6} Although mutations in MDM2 are rare, MDM2 overexpression is observed in a number of human tumors due to various mechanisms including gene amplification.⁷⁻⁹ and increased transcription.^{10,11} In addition, MDM2 overexpression predisposed transgenic mice to spontaneous tumor formation.¹² and therefore overexpression of MDM2 may substitute for inactivating mutations in p53.⁸ In other words, a deregulation in MDM2 expression levels might create a favourable environment for genetic instability, oncogene activation and tumorigenesis. Functional sequence variants in promoter regions will lead to a variable expression level of the gene.¹³⁻¹⁵ and especially SNPs in promoters of apoptosis genes could be implicated in individual susceptibility to develop cancer.¹⁶⁻²¹ Because MDM2 is a key component of apoptosis, promoter SNPs in this gene might influence this highly regulated pathway by modifying cellular MDM2 protein level.²² The MDM2 gene has a basal promoter (P1) and an alternative promoter (P2) present in the intron 1.²³ The promoter P2 contains a p53-responsive element, controlling MDM2 levels

in stressed cells while promoter P1 acts principally in a nonstressed environment.^{23,24} The SNP309 in the MDM2 intronic promoter has been shown to attenuate apoptotic activity and to accelerate tumor formation.^{4,25,26} Several studies reported associations between SNP309 and the risk of cancers.²⁷⁻³⁰ although many others failed to reveal such association.³¹⁻³³ Such discrepancies between studies might be explained by distinct designs, tissue-specificity or from the fact that only one SNP was investigated. In this report, we determined the SNP content and the haplotypes structure of the constitutive P1 promoter of MDM2. We found that distinct P1 promoter haplotypes can influence the p53-independent promoter activity in an allele-specific manner.

Materials and Methods

SNP discovery in MDM2 proximal promoter region

The initial search for promoter SNPs (pSNPs) in MDM2 proximal promoter defined as the ~2.0kb upstream of the transcription start site was done using the dbSNP database.³⁴ (build 128). Seven SNPs were selected for genotyping in a panel of 91 individuals of European descent. The corresponding promoter region was amplified in one PCR fragment in a 50 μ L reaction volume, using the following conditions: 20 pmoles of 5'AAAGCCCAAATTCCTTGCT3' (forward) and 5'CTCCATCTTTCCGACACACA3' (reverse) primers, 2mM MgCl₂, 0.2 mM dNTPs, 1x Fast Start Taq DNA polymerase buffer and GC rich buffer, 2U Fast Start Taq DNA polymerase (Roche Diagnostics) and 15 ng of genomic DNA. PCR program was 95°C for 3 min; 10 cycles with a denaturation at

95°C for 15 sec; annealing at 55-50°C (each cycle decreases by 0.5°C) for 20 sec and elongation at 68°C for 2 min; followed by 25 cycles at 50°C for annealing. The amplicons were dot-blotted in duplicate on a nylon membrane and were hybridized with allele-specific oligonucleotides (ASOs) as described by Labuda et al.³⁵ Oligonucleotide probes specific for each promoter SNP were used for ASO analysis and are given in Table 1S. A 40bp insertion/deletion (indel) polymorphism was genotyped by amplification of a 260 bp fragment containing the indel region and the electrophoresis of the resulting amplicons on a 3% agarose gel to detect of one (homozygous) or two bands (heterozygous). PCR conditions were: 20 pmoles of 5'TTCCTTTCTGGTAGGCTGG3' (forward) and 5'CACCTACTTTCCCACAGAGA3' (reverse) primers, 1.5 mM MgCl₂, 0.2 mM dNTPs, 1x Fast Start Taq DNA polymerase buffer and GC rich buffer, 1U Fast Start Taq DNA polymerase (Roche Diagnostics) and 15 ng of genomic DNA. PCR program was 95°C for 3 min; 32 cycles with a denaturation at 95°C for 30 sec; annealing at 52°C for 30 sec and elongation at 72°C for 20 sec. Hardy-Weinberg equilibrium was tested with a χ^2 test for goodness of fit. Haplotypes were generated by PHASE software (version 2; University of Washington, Seattle).³⁶

Electrophoretic mobility shift assay

Differential binding of DNA-nuclear proteins for the different alleles of the promoter SNPs was assessed by incubating 5' end radiolabelled ds-DNA oligonucleotide probes corresponding to the surrounding region of each SNP (top strand sequences of oligos are listed in table 2S) with nuclear extracts derived from HeLa, HepG2 and JEG-3 cell lines.³⁷ according to the manufacturer's protocol (Promega gel shift assay system). Quantification of protein extracts was done with the Bradford protein assay (BioRad). Nuclear extracts (5µg) were incubated independently with 35 fmol radiolabelled ds-DNA probes and in a buffer containing 50 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 2.5 mM EDTA, 2.5 mM DTT, 250 mM NaCl, 0.25 µg/µL poly (deoxyinosinate-deoxycytidylate) and 20% glycerol in a final volume of 10 µL. Incubations were performed at room temperature for 20 min. The resulting complexes were separated in a non-denaturing polyacrylamide gel (6% acrylamide:bisacrylamide, 60:1) with 0,5X Tris-borate-EDTA buffer at 4°C, 190V. Competitions with unlabelled probes were done in a 50:1 ratio (unlabelled: labelled). Non specific unlabelled probe is a mutated sequence of the AP2 transcription factor binding site.

Gene reporter assay

Constructs. The two major promoter haplotypes (~2.0kb region) were amplified from genomic DNA of known homozygous individuals and cloned individually in the promoterless pGL3basic firefly luciferase vector (Promega)

using the Gateway Technology (Invitrogen). Clones chosen for transfection were sequenced to confirm presence of the SNPs and then purified using Quiagen plasmid mini kit prior transfection.

Transfection. The resulting constructs were used to transiently transfect 3 cell lines (HeLa, HepG2, JEG-3) using lipofectamine reagent according to the manufacturer's protocol (Invitrogen). Constructs (99 ng) and SV40-driven (1ng) (ratio 100:1) *Renilla* luciferase pRL-CMV was co-transfected to control transfection efficiency. The pGL3basic promoterless plasmid (Promega) was used as negative control and the pGL3SV40 plasmid (Promega) was used as positive control. Transfected cells were plated in 96-well plates with approximately 6×10^4 cells per well. The cells were harvested 24 hours following transfection and luciferase reporter gene activity was measured with Dual-Luciferase Reporter Assay System (Promega) in a Spectra Max 190 luminometer according to the manufacturer's protocol (Molecular Devices). Firefly luciferase activities of the allelic constructs were normalized using the *Renilla* luciferase pRL-CMV activity. Results were expressed as the ratio of firefly luciferase activity divided by the pRL-CMV internal control activity and expressed as relative luciferase (means \pm sd) of four replicates. Three independent experiments were carried out for each cell line. Statistical analyses were performed using unpaired Student's *t*-test to determine *p*-values. Global *p*-value is calculated with Fischer's inverse Chi-squared test.³⁸

***In silico* predictions of putative TFBS (transcription factor binding site)**

MatInspector program from Genomatix Software GmbH (Germany, www.genomatix.de) was used to determine the presence of putative binding sites for known transcription factors. The predicted gain or loss of putative TFBS due to a given SNP was determined by the optimized matrix threshold as defined in the MatInspector program.

Results

The search for SNPs in the constitutive P1 promoter of MDM2 led to the identification of 8 pSNPs, including a 40bp indel (see Table 1). In addition to these pSNPs, we included the SNP309 located in the intronic promoter for haplotype analysis. The schematic representation of MDM2 promoters is shown in Fig.1. Only 4 promoter SNPs (-1494G>A, rs1144944; indel 40bp, rs3730485; -182C>G, rs937282; SNP309/601T>G, rsrs2279744) were found to be polymorphic by genotyping a panel of 91 unrelated Europeans. For SNP309/601T>G, the observed minor allele frequency of 35% was similar to the one reported for Caucasians in other studies (Zhibin Hu *et al*, 2007). Among the five non-polymorphic SNPs, both -1166T>G and -1164C>G are located in the 40 bp indel sequence thus creating in some individuals a near identical (except for 2 bps) tandem duplication. Therefore, individuals carrying the deletion behave like they have different alleles at these two positions. Because single variants might not be sufficient to capture the genetic variability relative to a given phenotype, we constructed haplotypes using these 4 promoter SNPs. Haplotype phase was estimated and corresponding

frequencies are given in Table 2. The three most common promoter haplotypes represented 92.3% of the observed haplotypes in Europeans. Haplotypes 1A and 1B differ only at SNP309, while haplotype 2 is totally different.

To assess the functional impact of the major promoter haplotypes 1 and 2, both promoter haplotypes were subcloned in the promoterless pGL3-Basic firefly luciferase reporter vector and transient transfection experiments were carried out for each haplotype-specific constructs (Fig. 2). Because our constructs did not include SNP309 but only the proximal promoter, difference in promoter activity between haplotype 1A and 1B could not be tested. Transient transfections of these constructs into 3 different cell lines were carried out to verify the differential promoter activities between both constructs (Fig. 2). The promoter haplotype H1 has a strong promoter activity in all cell lines, whereas H2 showed weaker promoter activity. Significant differences were found between H1 and H2, across all three cell lines (Fig. 2). Relative luciferase activity driven by H1 was up to 2,3-fold higher than the luciferase levels driven by its counterpart H2, indicating variable haplotype-specific expression levels of MDM2. The 309G allele was only present in 1.0% of individuals carrying haplotypes other than H1A (Data not shown), European individuals carrying allele G of this SNP would thus have more chances to possess the high P1 promoter activity haplotype because of this linkage disequilibrium. In an attempt to identify the *cis*-acting elements responsible for the observed changes in MDM2 promoter activity, haplotype SNPs (-1494G>A, -182C>G, SNP309/601T>G) were further validated by EMSAs. Toward this goal, ds-DNA radiolabelled oligonucleotides containing the SNP and its neighboring

sequence were constructed for each pSNPs (see Table 2S). After incubating these probes with different nuclear extracts (HeLa, HepG2, JEG3), complexes were separated on gel electrophoresis. Representative EMSA results for the SNP -1494G>A with JEG3 cell nuclear extract is shown in Fig. 3. Two DNA-protein complexes were observed. Complex 1 is a slow migrating complex only observed with -1494G radiolabelled probes (Lanes 5-8), whereas the fast migrating complex 2 is only observed with -1494A radiolabelled probe (Lanes 1-4). These results suggest differential DNA protein binding at the polymorphic site -1494G>A. Competition with the nonspecific unlabelled probe shows binding specificity to the ds-oligonucleotide sequences (Lanes 4 and 8). Similar results were obtained in HeLa and HepG2 cell extracts (Table 2S). Similarly, we observed allelic DNA-protein binding for the other promoter SNPs tested (Table 2S). Using *in silico* predictive tools, none of these SNPs seem to affect the putative binding of known transcription factors. These results, together with EMSA DNA-protein differential bindings support the functional impact of MDM2 promoter haplotypes on promoter activity.

Discussion

Overexpression of MDM2 is common in a variety of human tumours. Deregulation of apoptosis pathway can lead to an uncontrolled proliferation of cells, generating a favourable environment for genetic instability, oncogene activation and tumorigenesis. Because MDM2 plays a key role in apoptosis regulation, modification of its expression levels might indeed lead to such instable environment.²² MDM2 may play an important role in cancer development, mostly

through inactivation of the p53 pathway.³⁹ In contrary, the p53-independent tumorigenicity of MDM2 is less understood. MDM2 might affect cancer risk through its interaction with other key cancer genes with various functions.⁴⁰⁻⁴³ In this report, we characterized the two major haplotypes corresponding to the upstream p53-independent constitutive P1 promoter. We have shown that the constitutive expression levels of MDM2 might at least be partly regulated by distinct promoter SNPs and the corresponding promoter haplotypes. Previous work has shown correlation of the SNP-182C>G with allelic difference in promoter activity, with the allele -182G showing high promoter activity.⁴⁴ In our study, this allele was associated with the low-activity P1 promoter haplotype. This discrepancy could be explained by cell-specific differences or by the fact the extended promoter P1 haplotype was not determined.

Most previous studies have been focused on the impact of the SNP309 (T>G), which is located in the p53-dependent promoter P2. *In vitro* studies have shown that allele SNP309G increased the affinity of sp1 transcription activator for a putative binding site increasing the steady-state levels of MDM2, which in turn reduces the basal p53 levels.^{26,45} Many association studies have assessed the relationship between SNP309T>G and different cancer types. Although some studies demonstrated such association others showed no association between this pSNP and certain cancer types.e.g.^{30,46,47} These conflicting rather than conclusive results might be explained by several reasons, including linkage disequilibrium between SNP309 and another, yet unknown, functional SNP in MDM2 with nearby loci. In this report, we showed that SNP309G was associated with the high

P1 promoter activity haplotype. This linkage disequilibrium could also contribute to cancer associations with SNP 309 suggesting that haplotypes constructions of MDM2 pSNPs would add force to these association studies. We believe that looking at the impact of haplotypes rather than individual SNPs on promoter activity is a more suitable approach since it takes into account the putative interaction between SNPs. In a recent meta-analysis, Zhibin et al reported a clear association between SNP309 and cancer risk in Asians but not in Europeans and in Africans.⁴⁷ The explanation for this observation is unclear but could be explained by genetic heterogeneity because the observed SNP309 frequencies are variable in different populations ranging from ~50% in Asians to ~ 33% and ~ 10% in Caucasians and Africans, respectively.⁴⁷ Finally, the transcription factor influenced by SNP309 might be cell-specific so that this variant does not affect MDM2 expression in certain tissues.⁴

In conclusion, this study revealed differential constitutive P1 promoter activities, at least *in vitro*. This observation implies that individuals carrying distinct p53-independent P1 promoter haplotypes might have a modified risk of cancer development. Association studies in patient cohorts will further determine their importance in cancer development.

Acknowledgements

This study was supported by research funds provided by the CIHR as well as Genome Quebec/Canada. MEL is the recipient of a NSERC Canada Graduate's scholarship. DS holds the François-Karl Viau Chair in Pediatric Oncogenomics and is a scholar of the Fonds de la Recherche en Santé du Québec (FRSQ).

References

1. Bennett, W.P., Hussain, S.P., Vahakangas, K.H., Khan, M.A., Shields, P.G. & Harris, C.C. Molecular epidemiology of human cancer risk: gene-environment interactions and p53 mutation spectrum in human lung cancer. *J Pathol* 1999; 187: 8-18.
2. Hu, W., Feng, Z., Ma, L. et al. A single nucleotide polymorphism in the MDM2 gene disrupts the oscillation of p53 and MDM2 levels in cells. *Cancer Res* 2007; 67: 2757-65.
3. Harris, S.L. & Levine, A.J. The p53 pathway: positive and negative feedback loops. *Oncogene* 2005; 24: 2899-908.
4. Bond, G.L., Hu, W. & Levine, A. A single nucleotide polymorphism in the MDM2 gene: from a molecular and cellular explanation to clinical effect. *Cancer Res* 2005; 65: 5481-4.
5. Haupt, Y., Maya, R., Kazaz, A. & Oren, M. Mdm2 promotes the rapid degradation of p53. *Nature* 1997; 387: 296-9.
6. Kubbutat, M.H., Jones, S.N. & Vousden, K.H. Regulation of p53 stability by Mdm2. *Nature* 1997; 387: 299-303.
7. Momand, J., Jung, D., Wilczynski, S. & Niland, J. The MDM2 gene amplification database. *Nucleic Acids Res* 1998; 26: 3453-9.
8. Oliner, J.D., Kinzler, K.W., Meltzer, P.S., George, D.L. & Vogelstein, B. Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature* 1992; 358: 80-3.
9. Meddeb, M., Valent, A., Danglot, G. et al. MDM2 amplification in a primary alveolar rhabdomyosarcoma displaying a t(2;13)(q35;q14). *Cytogenet Cell Genet* 1996; 73: 325-30.
10. Bueso-Ramos, C.E., Yang, Y., deLeon, E., McCown, P., Stass, S.A. & Albitar, M. The human MDM-2 oncogene is overexpressed in leukemias. *Blood* 1993; 82: 2617-23.
11. Watanabe, T., Hotta, T., Ichikawa, A. et al. The MDM2 oncogene overexpression in chronic lymphocytic leukemia and low-grade lymphoma of B-cell origin. *Blood* 1994; 84: 3158-65.
12. Jones, S.N., Hancock, A.R., Vogel, H., Donehower, L.A. & Bradley, A. Overexpression of Mdm2 in mice reveals a p53-independent role for Mdm2 in tumorigenesis. *Proc Natl Acad Sci U S A* 1998; 95: 15608-12.
13. King, M.C. & Wilson, A.C. Evolution at two levels in humans and chimpanzees. *Science* 1975; 188: 107-16.
14. Knight, J.C. Regulatory polymorphisms underlying complex disease traits. *J Mol Med* 2005; 83: 97-109.

15. Pastinen, T., Sladek, R., Gurd, S. et al. A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics* 2004; 16: 184-93.
16. Park, J.Y., Park, J.M., Jang, J.S. et al. Caspase 9 promoter polymorphisms and risk of primary lung cancer. *Hum Mol Genet* 2006; 15: 1963-71.
17. Harris, S.L., Gil, G., Robins, H. et al. Detection of functional single-nucleotide polymorphisms that affect apoptosis. *Proc Natl Acad Sci U S A* 2005; 102: 16297-302.
18. Ho, S.Y., Wang, Y.J., Chen, H.L. et al. Increased risk of developing hepatocellular carcinoma associated with carriage of the TNF2 allele of the -308 tumor necrosis factor-alpha promoter gene. *Cancer Causes Control* 2004; 15: 657-63.
19. Wagner, K., Hemminki, K., Grzybowska, E. et al. Polymorphisms in genes involved in GH1 release and their association with breast cancer risk. *Carcinogenesis* 2006; 27: 1867-75.
20. Ma, X., Ruan, G., Wang, Y. et al. Two single-nucleotide polymorphisms with linkage disequilibrium in the human programmed cell death 5 gene 5' regulatory region affect promoter activity and the susceptibility of chronic myelogenous leukemia in Chinese population. *Clin Cancer Res* 2005; 11: 8592-9.
21. Moshynska, O., Sankaran, K. & Saxena, A. Molecular detection of the G(-248)A BAX promoter nucleotide change in B cell chronic lymphocytic leukaemia. *Mol Pathol* 2003; 56: 205-9.
22. Bond, G.L., Hu, W. & Levine, A.J. MDM2 is a central node in the p53 pathway: 12 years and counting. *Curr Cancer Drug Targets* 2005; 5: 3-8.
23. Zauberman, A., Flusberg, D., Haupt, Y., Barak, Y. & Oren, M. A functional p53-responsive intronic promoter is contained within the human mdm2 gene. *Nucleic Acids Res* 1995; 23: 2584-92.
24. Ries, S., Biederer, C., Woods, D. et al. Opposing effects of Ras on p53: transcriptional activation of mdm2 and induction of p19ARF. *Cell* 2000; 103: 321-30.
25. Bond, G.L., Hirshfield, K.M., Kirchhoff, T. et al. MDM2 SNP309 accelerates tumor formation in a gender-specific and hormone-dependent manner. *Cancer Res* 2006; 66: 5104-10.
26. Bond, G.L., Hu, W., Bond, E.E. et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell* 2004; 119: 591-602.
27. Dharel, N., Kato, N., Muroyama, R. et al. MDM2 promoter SNP309 is associated with the risk of hepatocellular carcinoma in patients with chronic hepatitis C. *Clin Cancer Res* 2006; 12: 4867-71.

28. Bond, G.L., Menin, C., Bertorelle, R., Alhopuro, P., Aaltonen, L.A. & Levine, A.J. MDM2 SNP309 accelerates colorectal tumour formation in women. *J Med Genet* 2006; 43: 950-2.
29. Ohmiya, N., Taguchi, A., Mabuchi, N. et al. MDM2 promoter polymorphism is associated with both an increased susceptibility to gastric carcinoma and poor prognosis. *J Clin Oncol* 2006; 24: 4434-40.
30. Bond, G.L. & Levine, A.J. A single nucleotide polymorphism in the p53 pathway interacts with gender, environmental stresses and tumor genetics to influence cancer in humans. *Oncogene* 2007; 26: 1317-23.
31. Wilkening, S., Hemminki, K., Rudnai, P. et al. No association between MDM2 SNP309 promoter polymorphism and basal cell carcinoma of the skin. *Br J Dermatol* 2007; 157: 375-7.
32. Petenkaya, A., Bozkurt, B., Akilli-Ozturk, O., Kaya, H.S., Gur-Dedeoglu, B. & Yulug, I.G. Lack of association between the MDM2-SNP309 polymorphism and breast cancer risk. *Anticancer Res* 2006; 26: 4975-7.
33. Pine, S.R., Mechanic, L.E., Bowman, E.D. et al. MDM2 SNP309 and SNP354 are not associated with lung cancer risk. *Cancer Epidemiol Biomarkers Prev* 2006; 15: 1559-61.
34. Sherry, S.T., Ward, M.H., Kholodov, M. et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001; 29: 308-11.
35. Labuda, D., Krajcinovic, M., Richer, C. et al. Rapid detection of CYP1A1, CYP2D6, and NAT variants by multiplex polymerase chain reaction and allele-specific oligonucleotide assay. *Anal Biochem* 1999; 275: 84-92.
36. Stephens, M., Smith, N.J. & Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001; 68: 978-89.
37. Schreiber, E., Matthias, P., Muller, M.M. & Schaffner, W. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res* 1989; 17: 6419.
38. Idelman, G., Taylor, J.G., Tongbai, R. et al. Functional profiling of uncommon VCAM1 promoter polymorphisms prevalent in African American populations. *Hum Mutat* 2007; 28: 824-9.
39. Zhang, X., Miao, X., Guo, Y. et al. Genetic polymorphisms in cell cycle regulatory genes MDM2 and TP53 are associated with susceptibility to lung cancer. *Hum Mutat* 2006; 27: 110-7.
40. Dobbstein, M., Wienzek, S., Konig, C. & Roth, J. Inactivation of the p53-homologue p73 by the mdm2-oncoprotein. *Oncogene* 1999; 18: 2101-6.
41. Xiao, Z.X., Chen, J., Levine, A.J. et al. Interaction between the retinoblastoma protein and the oncoprotein MDM2. *Nature* 1995; 375: 694-8.

42. Zhang, Y., Xiong, Y. & Yarbrough, W.G. ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell* 1998; 92: 725-34.
43. Zhang, Z. & Zhang, R. p53-independent activities of MDM2 and their relevance to cancer therapy. *Curr Cancer Drug Targets* 2005; 5: 9-20.
44. Wang, M., Zhang, Z., Zhu, H. et al. A novel functional polymorphism C1797G in the MDM2 promoter is associated with risk of bladder cancer in a Chinese population. *Clin Cancer Res* 2008; 14: 3633-40.
45. Arva, N.C., Gopen, T.R., Talbott, K.E. et al. A chromatin-associated and transcriptionally inactive p53-Mdm2 complex occurs in mdm2 SNP309 homozygous cells. *J Biol Chem* 2005; 280: 26776-87.
46. Wilkening, S., Bermejo, J.L. & Hemminki, K. MDM2 SNP309 and cancer risk: a combined analysis. *Carcinogenesis* 2007; 28: 2262-7.
47. Zhibin Hu GJ, L.W., Feng Chen, Xinru Wang, and Hongbing Shen MDM2 Promoter Polymorphism SNP309 Contributes to Tumor Susceptibility: Evidence from 21 Case-Control Studies *Cancer Epidemiol Biomarkers Prev* 2007; 16: 2717-2723.

Titles and legends to figures

Figure 1. Schematic illustration of MDM2 basal (P1) and internal (P2) promoters with the position of the investigated pSNPs. The promoter positions were numbered with respect to the first nucleotide of the first exon as +1, and the nucleotide immediately upstream as -1.

Figure 2. Gene reporter assays evaluating the functional impact of the most frequent MDM2 basal promoter haplotypes. **A.** Constructs tested for luciferase gene reporter assays in pGL3basic vector. **B.** Relative luciferase activity of MDM2 promoter haplotypes was measured following transient transfection in HeLa, Jeg-3 and HepG2 cells. The empty promoterless pGL3-Basic vector was used as negative control. Results are expressed in a ratio of Firefly/*Renilla* activity multiplied by 100. Promoter haplotype H2 was used as reference against which relative expression was compared. Haplotype H1 showed significantly higher expression levels across all four cell lines. P values are calculated from 4 replicates with unpaired Student's t-test. Significant differences are marked with an asterisk (*: $P < 3 \times 10^{-3}$; **: $P < 8 \times 10^{-6}$).

Figure 3. Representative EMSA analysis showing allelic DNA-protein interactions in the promoter region of MDM2. Labeled double-stranded oligonucleotide probes (marked with an asterisk) corresponding to the -1494 G>A alleles were incubated with Jeg3 nuclear extracts. Lanes 1 to 4 represent radiolabelled ds-DNA probe -1494A; lanes 5 to 8, radiolabelled ds-DNA probe -

1494G. The unlabeled probes used to compete DNA-protein interactions (in 50-fold molar excess) are indicated (+) at the top of each lane. Specific competitors corresponding to the unlabeled allele-specific probes were used in lanes 2, 3, 6 and 7 and a non-specific double-stranded oligonucleotide competitor was used in lanes 4 and 8. Probe sequences are listed in the Table 4 (supplementary materials). Fast migrating unbound probes are found at the bottom of the gel while protein-DNA complexes have slower mobility (arrows). In this experiment, 2 distinct complexes were found following incubation of the probes with Jeg-3 nuclear extract. Allele- the -1494G (lane 5 vs. 1) and complex 2 which showed high binding affinity for the -1494A (lane 1 vs. 5). Lane 4 has slightly more radioactivity than lane 1-3. A nonspecific complex is observed at the centre of the gel center.

List of supplementary materials

Table 1S. List of oligonucleotide probes used for PCR-ASO genotyping.

Table 2S. Summary of EMSA results for selected pSNPs in MDM2 promoters.

Table 1. List of the SNPs found in dbSNP database for MDM2 promoter.

rs number [†]	Position [‡]	SNPs ID [§]	MAF [¶]
rs1144944	g.67,486,752G>A	-1494	25%
rs3730485	g.67,487,038_67,487,077del	40 bp indel (-1208 to-1169)	37%
rs2904506	g.67,487,080T>G	-1166	-
rs3930427	g.67,487,082C>G	-1164	-
rs3730486	g.67,487,509C>T	-737	0%
rs3730487	g.67,487,563A>G	-683	0%
rs937282	g.67,488,064C>G	-182	50%
rs3730491	g.67,488,095C>T	-151	0%
rs2279744 [¥]	g.67,488,847T>G	+601 [‡]	35%

[†]From dbSNP build 128

[‡] NCBI Build 36.1

[§]Position relative to the transcription start site (based on Ref Seq mRNA).

[¶]Minor allele frequencies (MAF) were calculated with 91 unrelated European individuals genotyped by ASO.

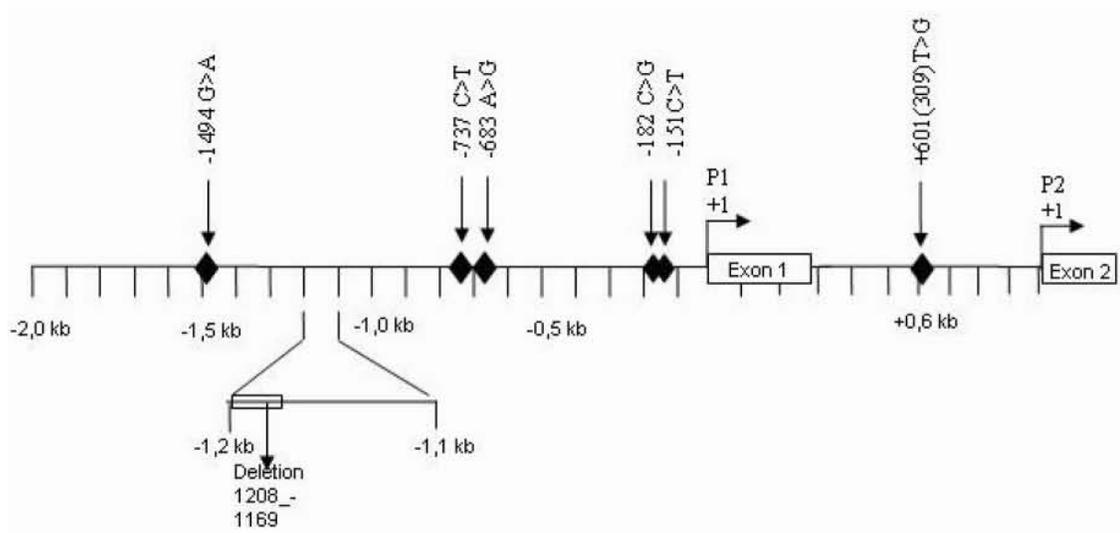
[¥]For comparison purpose we included the SNP309 associated with the internal promoter P2.

Table 2. Most frequent MDM2 promoter haplotypes.

Haplotype	-1494 G>A	40 bp indel	-182 C>G	+309T>G	Frequency [†]
1A	A	No deletion	C	G	36.8%
1B	A	No deletion	C	T	17.0%
2	G	deletion	G	T	38.5%

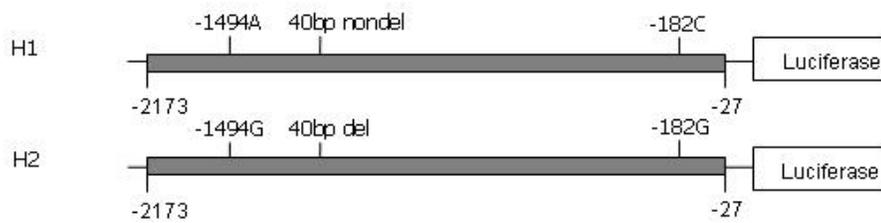
[†]Frequencies calculated with genotyping results of the 91 unrelated European individuals.

Lalonde, ME. Figure 1

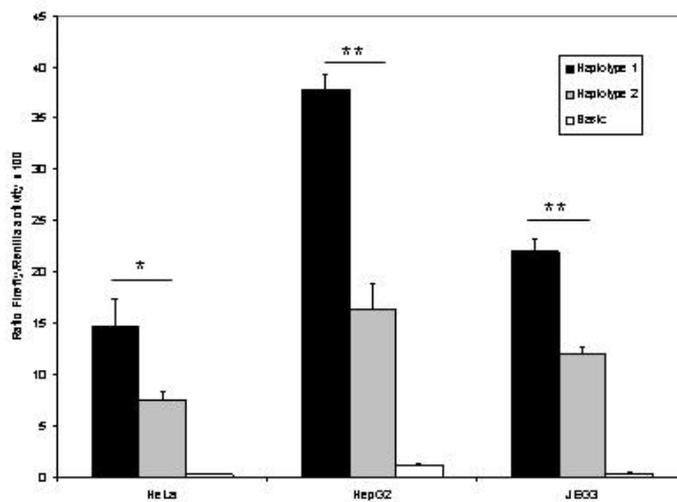


Lalonde et al. Figure 2

A



B



Lalonde, ME. Figure 3

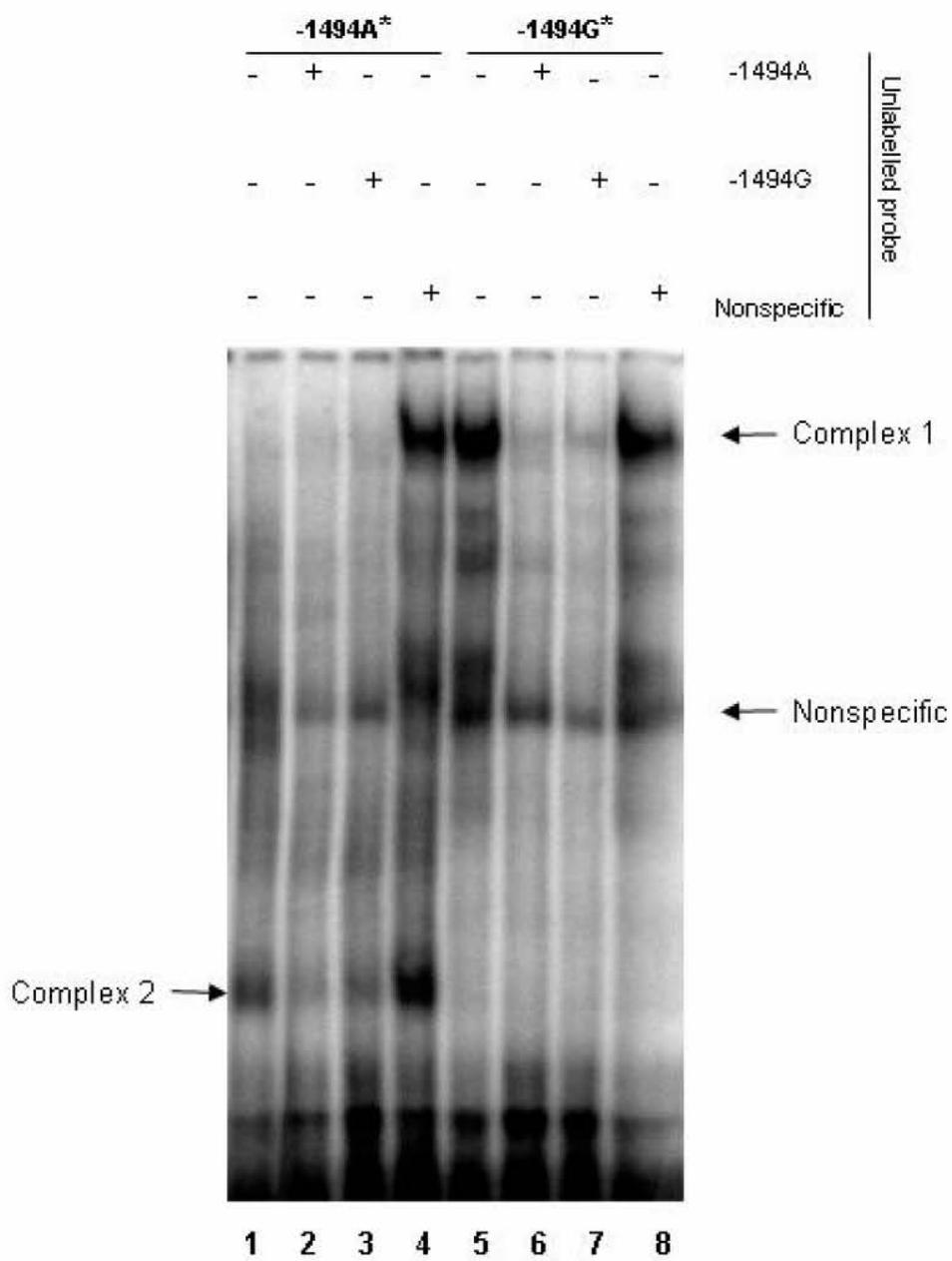


Table 1S. List of oligonucleotide probes used for PCR-ASO genotyping.

Rs number¹	Major Allele Probe	Minor Allele Probe
rs1144944	TCGTGATCCGCCCACCTCA	CTCGTGATCCACCCACCTCA
rs3730486	CCACCACGCCCGACTAATTTT	CCACCACGCCTGACTAATTTT
rs3730487	TGGCCAGGTTAGTCTTGAAC	TGGCCAGGTTGGTCTTGAAC
rs937282	ACCCCGCCTCACAGCCC	ACCCCGCCTCAGAGCCC
rs3730491	GACACCCCCCACTCCATCAT	GACACCCCCTACTCCATCAT
rs2279744	CGCTGCGGCGCGGGA	CGCGGCGGCGCGGGA

¹From dbSNP build 128.

Table 2S. Summary of EMSA results for selected pSNPs in MDM2 promoters.

rSNPs	nuclear cell extract¹			Oligonucleotide used for EMSA²
	HeLa	HepG2	JEG3	
-1494	Yes	Yes	Yes	CTCGTGATCC(A/G)CCCACCTCAG GCCAACACCCCCACCCCGCCTCA(C/G)AGCCC
-182	Yes	Yes	No	GCCGCGCCCCGCGGGGCGACA
+309	Yes	Yes	Yes	CCGCT(T/G)CGGCGCGGGAGGTCCGGA

¹Presence (Yes) or absence (No) of differential protein-DNA interactions at allelic position.

²Only the forward strand of the ds-DNA oligonucleotides used for EMSA are shown here. SNPs are in bold characters.

