



**HAL**  
open science

# Dicer, enzyme clef de l'interférence ARN : études de son intérêt clinique dans les cancers du sein et implication dans la réponse au stress réplicatif

Gaël Grelier

## ► To cite this version:

Gaël Grelier. Dicer, enzyme clef de l'interférence ARN : études de son intérêt clinique dans les cancers du sein et implication dans la réponse au stress réplicatif. Biochimie [q-bio.BM]. Université Claude Bernard - Lyon I, 1980. Français. NNT : . tel-00353077

**HAL Id: tel-00353077**

**<https://theses.hal.science/tel-00353077>**

Submitted on 14 Jan 2009

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

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

THÈSE  
présentée  
devant l'UNIVERSITÉ LYON 1 - CLAUDE BERNARD  
pour l'obtention  
du DIPLÔME DE DOCTORAT  
(arrêté du 7 août 2006)  
présentée et soutenue publiquement le 12 décembre 2008  
par M. Gaël GRELIER

TITRE :

DICER, ENZYME CLEF DE L'INTERFÉRENCE ARN :  
ÉTUDES DE SON INTÉRÊT CLINIQUE  
DANS LES CANCERS DU SEIN ET IMPLICATION DANS  
LA RÉPONSE AU STRESS RÉPLICATIF

Directeur de thèse : Dr. Caroline MOYRET-LALLE

Jury : Dr. Zdenko HERCEG (président du jury)  
Dr. Mary CALLANAN (rapporteur)  
Dr. Thierry LORCA (rapporteur)  
Dr. Jean BENHATTAR (examineur)  
Dr. Caroline MOYRET-LALLE (directeur de thèse)

## **RÉSUMÉ :**

Les cancers du sein sont la première cause de mortalité chez les femmes occidentales. Ces tumeurs solides présentent une grande hétérogénéité associée à une résistance aux traitements et un taux de rechute important à long terme. Ainsi, les cliniciens doivent personnaliser la prise en charge des patientes, ce qui est rendu possible par une meilleure connaissance des causes moléculaires qui participent à l'apparition et à la progression des tumeurs mammaires. Pour ce projet, nous avons choisi d'étudier, dans ces cancers, les éventuelles valeurs pronostique et diagnostique de *dicer*, gène codant pour la ribonucléase clef du mécanisme d'interférence ARN. Il a été montré une implication de Dicer dans la mise en place de l'hétérochromatine *de novo* des régions péricentromériques. Encore peu étudiée chez l'homme, les données disponibles chez la levure nous ont permis d'envisager que Dicer pourrait être impliquée dans la régulation de la stabilité chromosomique. Nous avons donc testé les valeurs pronostique et diagnostique de *dicer* (en PCR quantitative et en Tissue Microarray) dans une centaine d'échantillons de tumeurs, dans des lignées cellulaires et dans des modèles de progression tumorale et métastatique. Parallèlement, nous avons étudié les conséquences d'une inhibition de l'expression de *dicer* sur le cycle cellulaire et sur la réponse à un stress réplicatif. Nos résultats ont montré que l'expression de *dicer* est un facteur pronostique indépendant de rechutes métastatiques et est associé avec l'expression des récepteurs hormonaux. Par ailleurs, des cellules n'exprimant pas *dicer* ont montré des dérégulations du cycle cellulaire et de la voie de réponse aux cassures de l'ADN. En conclusion, l'altération de l'expression de *dicer* pourrait jouer un rôle dans l'apparition de l'instabilité chromosomique des cancers du sein et son analyse pourrait permettre une meilleure prise en charge des patientes à risque pour une rechute métastatique.

## **TITLE:**

Dicer, RNA interference key enzyme: studies of it clinical interest in breast cancer and of it implication in the replicative stress response

## **ABSTRACT:**

Breast cancers are the first cause of mortality in occidental women population. Breast tumours can show various forms which are frequently resistant to therapeutics and prone to late relapse. Thus, the current clinical challenge consists in refining individual therapeutics management by employing tools provided by the study of molecular basis of mammary tumorigenesis. Besides, chromosomal instability (CIN) is a hallmark of breast cancer and recent data showed that Dicer, a key ribonuclease of the RNA interference mechanism, could be a regulator of chromosomal stability in human cells. We thus hypothesized that alteration of this protein could be associated with mammary tumorigenesis. In order to test this hypothesis, we assessed *dicer* transcription and expression in breast cancer cell lines and tissues corresponding to different phases of tumor progression. We further investigated the consequences of *dicer* knock-down on cell cycle and response to replicative stress. Our results show that *dicer* expression has an independent prognostic value for metastatic relapse prediction and is correlated with hormonal receptors expression. Furthermore, cells harbouring *dicer* inactivation presented defects in cell cycle and DNA breaks response pathways. Altogether, our data indicate that *dicer* inactivation could favour CIN during mammary tumorigenesis and this feature could represent a useful tool in breast cancer management.

## **MOTS CLEFS / KEY WORDS :**

Cancers du sein, Dicer, Interférence ARN, Pronostique, Métastase, Réparation ADN, Cycle cellulaire, Réponse au stress.

Breast cancer, Dicer, RNA interference, Prognosis, Metastasis, DNA repair, Cell cycle, Stress response.

**ONCOGÈNE ET PROGRESSION TUMORALE, LABORATOIRE INSERM UMR-590, CENTRE LÉON BÉRARD, 28 RUE LAËNNEC, 69373 LYON CEDEX 08**

*C'est lorsqu'il y a trop à dire qu'il faut  
s'efforcer d'être le plus court possible.*

**Georges Bernanos**

*En science, la phrase la plus excitante  
que l'on peut entendre, celle qui annonce  
des nouvelles découvertes, ce n'est pas  
"Eureka" mais "c'est drôle".*

**Isaac Asimov**

## Remerciements

Je remercie les Docteurs Zdenko Herceg, Mary Callanan, Thierry Lorca et Jean Benhattar, membres du jury, pour avoir accepté de juger mon travail. Je remercie plus particulièrement mes deux rapporteurs, les Docteurs Mary Callanan et Thierry Lorca.

Je remercie le Professeur Alain Puisieux, directeur de l'unité INSERM U590, pour m'avoir accueilli dans son laboratoire depuis mon D.E.A. en 2003 et pour m'avoir permis de mener librement mes travaux de recherche.

Je remercie les Docteurs Zdenko Herceg et Mehmet Öztürk, membres de mon comité de thèse, pour leurs conseils à l'approche de la fin de ma thèse.

Je remercie la Ligue Nationale contre le Cancer et à travers elle les bénévoles et les donateurs de cette association pour m'avoir accordé un financement durant les quatre années de ce projet.

Je remercie ma directrice de thèse, le Docteur Caroline Moyret-Lalle, pour son encadrement technique et scientifique, pour son aide précieuse dans la rédaction de ce manuscrit et des publications issues du projet d'étude ainsi que pour son soutien amical tout au long de ces cinq années et demi passées ensemble.

Je remercie Nicolas Voirin, le Docteur Christine Lasset, le Docteur Valérie Bonadona, Sylvie Chabaud, Jean-Damien Combes et le Professeur René Ecochard pour leur collaboration fructueuse en ce qui concerne la partie statistique de ce projet.

Je remercie le Docteur Isabelle Treilleux, Sophie Léon-Goddard, le Docteur Thomas Bachelot et Thérèse Gargi pour leur collaboration tout aussi fructueuse en ce qui concerne cette fois les parties anatomo-pathologique et clinique de ce projet.

Je remercie Emilie Ruaux (Pharmacie 6<sup>ème</sup> année), Janis Vella (Pharmacie Erasmus), Cécile Montgiraud (Sciences Master 1), Emmanuelle Ruiz (Pharmacie 2<sup>ème</sup> année), Charlotte Verfaillie (Pharmacie Socrate), Mélanie Lombardo (Génie Biologique 1<sup>ère</sup> année), Pierre Levy (ENS Sciences) et Marie Gouanvic (Sciences Master 1), stagiaires que j'ai encadré et avec qui j'ai eu le plaisir de partager l'avancement du projet. Je voudrais plus particulièrement remercier Emilie qui a "subi" ma première expérience en tant qu'encadrant tout en gardant son tempérament joyeux et dynamique ainsi que Marie et Charlotte (indissociables donc indissociées) qui ont elles aussi apporté beaucoup de bonne humeur lors de leur passage au laboratoire et qui se sont beaucoup impliquées dans ce projet.

Je remercie les étudiants du laboratoire, forces vives de la recherche, pour leur esprit de groupe, leur aide et conseils aussi bien techniques que scientifiques. Merci beaucoup à Clémence Thomas, Kader Selmi, Linda Zane et le Docteur Benjamin Bouchet (tous doctorants) pour avoir fait en sorte que le quotidien au laboratoire soit tout sauf ennuyeux. Merci "très beaucoup" à Benjamin, encore, tout le monde ou presque saura pourquoi !

Je remercie Célia Payen, Virginie Belin et Julie Putelat pour leur amitié qui perdure après leur passage au laboratoire.

Je remercie Carole Audouy, technicienne au laboratoire, pour sa bonne humeur permanente et les compétences qu'elle m'a transmises.

Je remercie le Docteur Yannis Robin-Lespinasse avec qui je partage le même cursus depuis 10 ans et surtout une amitié sans faille.

Je remercie Alizé pour m'avoir encouragé et pour m'avoir aéré l'esprit en me faisant partager de très bons moments. Merci d'être auprès de moi à cette période importante (et stressante) de ma vie.

Je remercie Marjorie pour m'avoir soutenu et aidé à avoir une vie en dehors du laboratoire. Merci pour l'ouverture d'esprit que tu m'apportes.

Je remercie JP et Alex, mes colocataires, pour leur soutien et leurs encouragements tout au long de la rédaction de ce manuscrit.

Je remercie (et c'est bien peu) ma famille pour leur soutien inconditionnel et en particulier mes parents qui m'ont toujours encouragé et ainsi permis d'accéder à ce haut niveau d'étude.

# Table des matières

<b>Liste des abréviations .....</b>	<b>10</b>
<b>Liste des illustrations .....</b>	<b>12</b>
<b>I – Introduction.....</b>	<b>14</b>
I.1. Les cancers du sein .....	14
I.1.1. Épidémiologie et aspects cliniques.....	14
I.1.2. Caractérisation moléculaire des cancers du sein .....	19
I.1.3. Origine des cancers du sein .....	22
I.2. Cycle et points de contrôle cellulaire : rôle du complexe APC/C .....	25
I.2.1. Aperçu du cycle cellulaire et des points de contrôle associés.....	25
I.2.2. Le complexe APC/C : implications dans la tumorigénèse .....	29
I.3. Les réparations des cassures double-brin de l'ADN.....	31
I.3.1. Présentation des mécanismes .....	31
I.3.2. Interactions des voies de réparations et de l'interférence ARN.....	35
I.4. L'interférence ARN .....	37
I.4.1. Le mécanisme au prix Nobel.....	37
I.4.2. Les multiples facettes de l'interférence ARN .....	38
I.4.3. Implication de <i>dicer</i> dans des processus biologiques fondamentaux.....	45
I.4.4. Interférence ARN et cancers .....	46
<b>Objectifs .....</b>	<b>50</b>
<b>II – Résultats .....</b>	<b>52</b>
II.1. Recherche de facteurs pronostiques/diagnostiques des cancers du sein .....	52
II.1.1. L'analyse d'éléments clés du complexe APC/C dans les cancers du sein et du côlon révèle que APC-11 est un marqueur d'agressivité.....	52
i. Résumé .....	52
ii. Manuscrit.....	54
II.1.2. Valeur pronostique et pertinence clinique de l'expression de <i>dicer</i> dans les cancers du sein : relation avec le phénotype mésenchymateux ? .....	72
i. Résumé .....	72
ii. Manuscrit.....	73
II.2. Influence de <i>dicer</i> sur le déroulement du cycle cellulaire .....	101
II.2.1. <i>dicer</i> est régulé au cours du cycle cellulaire.....	101
II.2.2. L'inactivation de <i>dicer</i> dans les cellules HeLa retarde le cycle cellulaire..	102

II.2.3. Inactivation de DGCR8 et cycle cellulaire .....	105
II.3. Implication de <i>dicer</i> dans les voies de réponses aux cassures double-brin ...	108
II.3.1. Rôle de <i>dicer</i> dans la réponse aux irradiations .....	108
II.3.2. Rôle de <i>dicer</i> dans la réponse à l'hydroxyurée .....	109
i. cellules tumorales HeLa .....	110
ii. cellules tumorales mammaires MDA-MB 231 .....	112
iii. cellules tumorales mammaires T47D .....	113
II.3.3. Rôle de la voie des miARN dans la réponse à l'hydroxyurée.....	114
<b>III. Discussion .....</b>	<b>116</b>
<b>IV. Conclusion et perspectives.....</b>	<b>122</b>
<b>Bibliographie : .....</b>	<b>125</b>
<b>Annexe 1 : matériels et méthodes .....</b>	<b>133</b>
A1.1. Lignées cellulaires et traitements .....	133
A1.2. Tumeurs mammaires.....	134
A1.3. Extraction d'ARN et Reverse Transcription .....	135
A1.4. PCR en temps réel .....	135
A1.5. Western Blotting .....	136
A1.6. Cytométrie en flux.....	138
<b>Annexe 2 : milieux de culture.....</b>	<b>139</b>
<b>Annexe 3 : données cliniques des tumeurs mammaires humaines étudiées.....</b>	<b>139</b>
A3.1. Étude par PCR en temps réelle.....	139
A3.1. Étude par TMA.....	141
<b>Annexe 4 : données sur les lignées tumorales mammaires humaines étudiées.....</b>	<b>146</b>
<b>Annexe 5 : PCR en temps réel.....</b>	<b>147</b>
A5.1. Tableau des amorces .....	147



A5.2. Programmes d'amplification .....	147
<b>Annexe 6 : anticorps .....</b>	<b>148</b>
<b>Annexe 7 : publications .....</b>	<b>149</b>
Contribution of cell culture, RNA extraction, and reverse transcription to the measurement error in qRT-PCR-based gene expression quantification .....	149
Estimating biological variation of quantitative gene expression in real-time reverse transcription polymerase chain reaction (RT-PCR) .....	168
Génomique tumorale : l'image d'un paysage instable.....	179

## Liste des abréviations

ADN	Acide DésoxyriboNucléique
Ago	Argonaute
AIB1	Amplified In Breast cancer 1
APC-2/5/7/11	Anaphase Promoting Complex-2/5/7/11
ARN	Acide RiboNucléique
ATM	Ataxia Telangiectasia Mutated
ATP	Adenosine TriPhosphate
ATR	Ataxia Telangiectasia Mutated and Rad3-related
BER	Base Excision Repair
BRCA1	BReast Cancer 1
BRCA2	BReast Cancer 2
CA-15-3	Cancer Antigen 15-3
CDC-16/20/25C/23/27	Cell Division Cycle 16/20/25C/23/27 homolog
CDK	Cyclin-Dependent Kinase
CENP-A/C	CENtromere Protein A/C
CHK1/2	Checkpoint protein 1/2
Chp1	Chromodomain protein 1
CPB/p300	CREB Binding Protein/p300
DCL1-4	Dicer Like 1-4
DGCR8	DiGeorge syndrome Chromosomal Region 8
DNA-PK	DNA-dependent Protein Kinase
DSB	Double-Strand Break
DUF283	Domain of Unknown Function 283
EMT	Epithelial-Mesenchymal Transition
ER	Estrogen Receptor
ESC	Embryonic Stem Cell
H-RAS	Harvey RA Sarcoma viral oncogene homolog
H2A.X	Histone 2A.X
HER2 / neu / ERBB2	Human Epidermal growth factor Receptor 2
HMGA2	High Mobility Group AT-hook 2
HOXB13	HomeoboX B13
HR	Homologous Recombination
iARN	interférence ARN
IL17BR	InterLeukine 17B Receptor
Mad2	Mitotic arrest deficient 2
MDC1	Mediator of DNA Checkpoint 1
meH3K9	methyl Histone 3 Lysine 9
miARN	micro RNA
MMR	MisMatch Repair
MRE-11	Meiotic REcombination 11
MRN	MRE11-Rad50-NBS1
NBS1	Nijmegen Breakage Syndrome 1
NER	Nucleotid Excision Repair
NHEJ	Non Homologous End Joining
NLS	Nuclear Localisation Signal
PARP	Poly[ADP-Ribose]Polymerase
PAZ	Piwi-Argonaute-Zwille
PCR	Polymerase Chain Reaction
PDCD4	Programmed Cell Death 4
PIK3CA	9-Phospholinositide-3-Kinase, Catalytic
PKR	dsRNA-dependent Protein Kinase
PLK1	Polo-like kinase 1
PR	Progesteron Receptor
Rb	Rétinoblastome

RDRC	RNA-Directed RNA polymerase complex
RDRP	RNA-Dependent RNA Polymerase
RISC	RNA-Induced Silencing Complex
RITS	RNA-Induced Transcriptional Silencing complex
RT	Reverse Transcription
SBR	Scarff Bloom et Richardson
Scc1-4	Sister chromatid cohesion 1-4
SCF	SKP1-Cul1-F-box protein
shARN	short hairpin RNA
siARN	small interfering RNA
SMC1/3	Structural Maintenance of Chromosomes 1/3
TLS	TansLetion Synthesis
TMA	Tissue MicroArray
TNM	Tumor size, Nodes, distant Metastasis
TP53	Tumour Protein 53
Tpm1	Tropomyosine 1
TPR	TetratricoPeptide Repeats
TS	Template Switch
UTR	UnTranslated Region
UV	UltraViolet

## Liste des illustrations

Figure 1 : Données épidémiologiques INVS 2005 sur les cancers du sein .....	15
Figure 2 : Anato-mo-histologie de la glande mammaire.....	16
Figure 3 : Éléments de description des tumeurs de cancers du sein .....	17
Figure 4 : Relation entre profils de transcription et fonction biologique .....	20
Figure 5 : Voies d'auto-renouvellement et de différenciation des cellules souches du sein .....	23
Figure 6 : Différentes étapes du processus de transition épithélio-mésenchymateuse	20
Figure 7 : Déroulement et surveillance du cycle cellulaire.....	26
Figure 8 : Modélisation de la régulation et de la fonction du complexe APC/C lors du point de contrôle mitotique .....	28
Figure 9 : Schéma de la structure et modélisation du fonctionnement du complexe APC/C .....	29
Figure 10 : Structure et substrats du complexe APC/C .....	30
Figure 11 : Dommages à l'ADN et leurs mécanismes de réparation au cours du cycle cellulaire .....	32
Figure 12 : Voies de réponse schématique induites par une cassure double-brin et détails de la recombinaison homologue .....	34
Figure 13 : Modélisation du complexe cohésine .....	36
Figure 14 : Schémas des différents domaines protéiques de DGCR8, Drosha et Dicer et des variants transcriptionnels de <i>dicer</i> .....	39
Figure 15 : Principales voies de l'interférence ARN chez l'Homme .....	42
Figure 16 : Formation de l'hétérochromatine médiée par l'interférence ARN .....	44
Figure 17 : Expression de Dicer au cours du cycle cellulaire .....	101
Figure 18 : Les cellules inactivées pour <i>dicer</i> présentent un retard du cycle cellulaire	103
Figure 20 : Analyse de la mortalité des cellules inactivées pour <i>dicer</i> par un test annexine V .....	106
Figure 21 : Les cellules inactivées pour <i>DGCR8</i> ne présentent pas d'altération du cycle cellulaire .....	107
Figure 22 : Analyse de l'expression de SMC1 et Chk2 après rayons X en cellules HeLa .....	109
Figure 23 : Analyse de l'expression de SMC1, Chk1 et H2A.X après hydroxyurée en cellules HeLa .....	111
Figure 24 : Analyse de l'expression de SMC1 et Chk1 après hydroxyurée en cellules MDA-MB 231 .....	112

Figure 25 : Analyse de l'expression de SMC1 et Chk1 après hydroxyurée en cellules T47D .....	113
Figure 26 : Analyse de l'expression de SMC1 et Chk1 après hydroxyurée en cellules HeLa .....	115
Table 1 : Distribution des composantes du mécanisme d'interférence ARN chez les Eukaryotes .....	40
Table 2 : Données expérimentales démontrant un rôle des miARN dans les cancers .	47

## I – Introduction

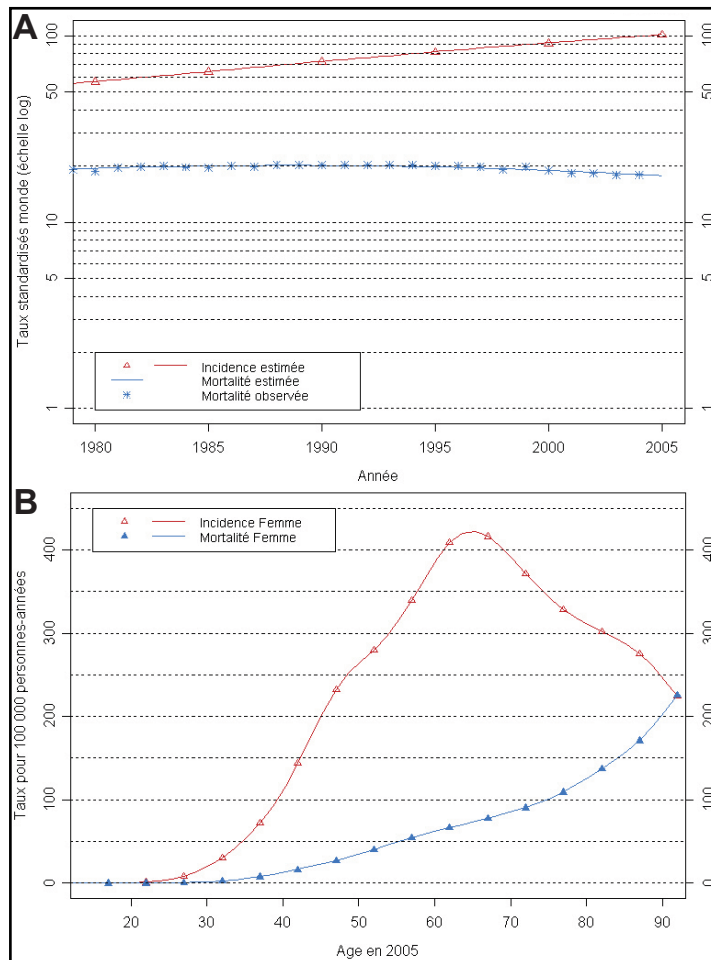
### I.1. Les cancers du sein

Parmi les défis à relever au cours de ce XXI<sup>ème</sup> siècle, se trouve le traitement des cancers qui sont une cause majeure de décès dans le monde occidental. En particulier les cancers du sein, cancers les plus fréquents chez les femmes nord-américaines et européennes. Malgré de nombreuses études aussi bien en recherche fondamentale qu'en recherches clinique et pharmaceutique, il n'existe pas encore de traitements totalement efficaces contre cette maladie si ce n'est la chirurgie, avec malgré tout un risque de rechute qui peut être élevé.

#### I.1.1. Épidémiologie et aspects cliniques

L'épidémiologie nationale des cancers du sein à partir des données de l'année 2005 montre une incidence toujours en progression par rapport aux années précédentes (INVS [http://www.invs.sante.fr/surveillance/cancers/estimations\\_cancers/default.htm](http://www.invs.sante.fr/surveillance/cancers/estimations_cancers/default.htm)). Avec 49814 dépistages positifs, cette localisation de cancers reste ainsi la plus fréquente chez la femme et continue d'être un problème majeur de santé publique. Cependant, le taux de mortalité pour ces cancers (11201 décès en 2005) a tendance à régresser depuis une décennie (Fig. 1.A). Il reste malgré tout la première cause de mortalité par cancer chez la femme. On remarque aussi que l'incidence et la mortalité varient beaucoup en fonction de l'âge mais sans être corrélées (Fig. 1.B). Ainsi, les femmes entre 60 et 70 ans sont celles qui développent le plus de cancers du sein mais ne correspondent pas à la tranche d'âge qui en meurt le plus.

Par ailleurs, il a été montré qu'il existe de nombreux facteurs de risque pour cette maladie aussi bien génétiques, environnementaux que nutritionnels. Dans cette



**Figure 1 : Données épidémiologiques INVS 2005 sur les cancers du sein**

**A.** Tendence chronologique du taux d'incidence et de mortalité en France (standardisés monde pour 100 000 personnes-années). La mortalité, qui était restée stable depuis 1980, amorce une décroissance en 2000 : le taux d'évolution annuel de la mortalité est de -0,4% sur l'ensemble de la période 1980-2005 alors qu'il est de -1,3% sur la dernière période 2000-2005. Les évolutions inverses de la mortalité et de l'incidence du cancer du sein peuvent s'expliquer en partie par l'amélioration des thérapeutiques et le diagnostic plus précoce lié au développement du dépistage en France, sans que leurs parts respectives puissent être précisées.

**B.** Incidence et mortalité par âge en 2005. Avec 49 814 nouveaux cas estimés en 2005 en France, le cancer du sein se situe au 2ème rang de tous les cancers pris dans leur globalité. Il représente 36,7% de l'ensemble des nouveaux cas de cancers chez la femme. Le taux d'incidence standardisé est de 101,5 et le taux de mortalité standardisé de 17,7. Avec 11201 décès par an, le cancer du sein est au 1er rang des décès par cancer chez la femme (18,9%).

dernière catégorie, seul l'alcool est un facteur établi avec un risque augmenté d'environ 7% pour une consommation moyenne d'une boisson alcoolisée par jour. [Collaborative group](#)

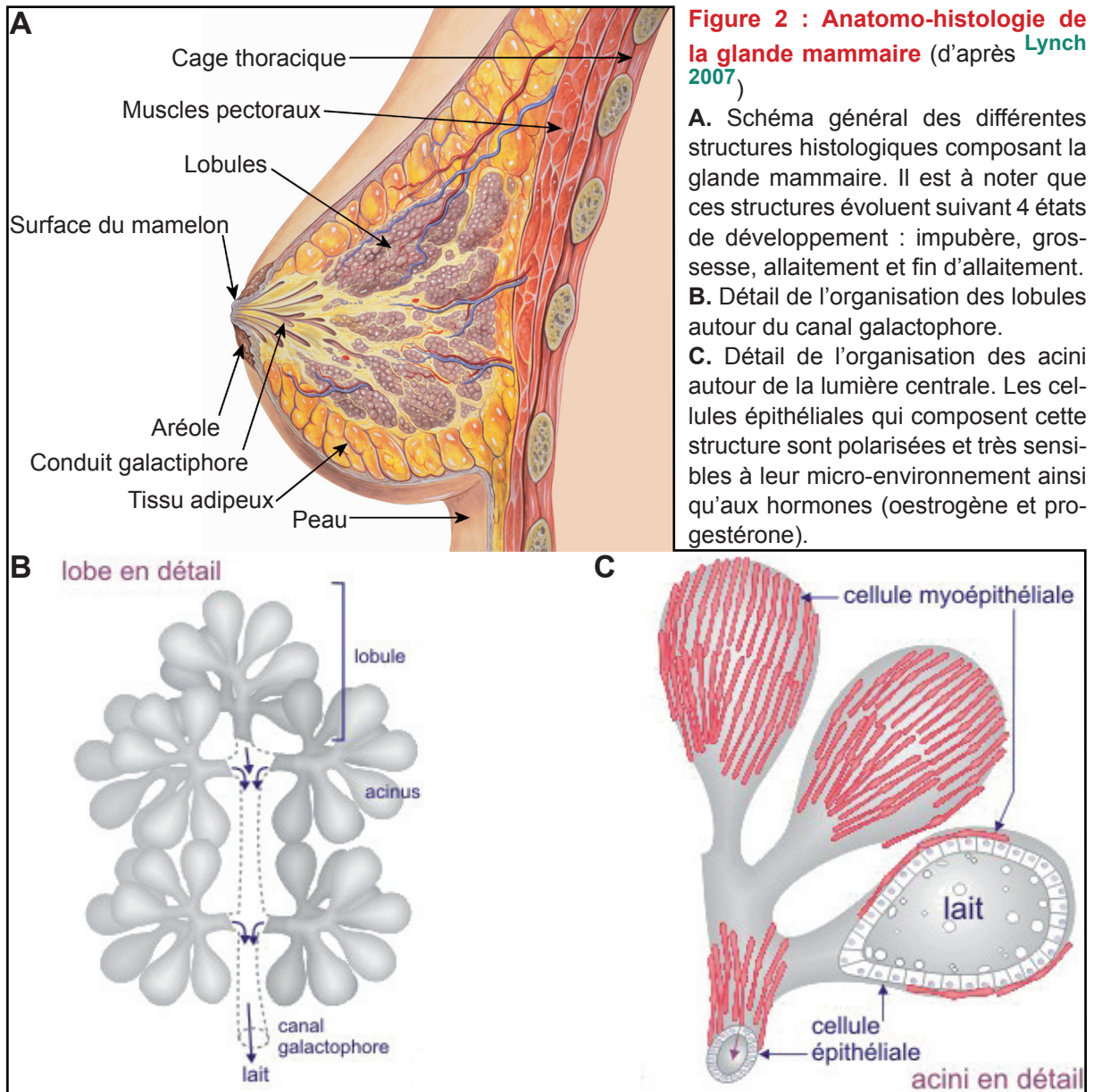
Au niveau environnemental, une exposition avant 40 ans aux radiations ionisantes pour un total de 1 Gy multiplie par 3 les risques. [Boice 1996, Little 1999](#)

Par contre, une activité physique moyenne diminue ce risque de 35% (notamment pour les femmes ménopausées). [WCRF 1997](#)

Enfin, d'un point de vue génétique, l'âge des premières règles, l'âge avancé et des mutations des gènes *BRCA1* et *BRCA2* (BReast CAncer n°1 et 2) sont trois facteurs de risque importants. [WCRF 1997, Ford 1998, Wolpert 2000, Kelsey 1996](#)

**1996**

Si les cancers du sein sont aussi fréquents, c'est aussi que la glande mammaire est en évolution constante au cours de la vie de la femme, le nombre de ses cellules en différenciation et en croissance est plus important que dans tout autre organe et la rend plus sensible aux processus de cancérisation. Au niveau histologique, le sein est une glande exocrine constituée de 2 compartiments cellulaires (**Fig. 2.A.**) : le



compartiment mésenchymateux (les vaisseaux sanguins et les nerfs) et le compartiment épithélial qui s'articule autour d'un réseau de canaux galactophores et de lobules (Fig. 2.B.) renfermant les acini (Fig. 2.C.). Ces deux compartiments sont séparés par une membrane basale de collagène, de laminine et de glycosaminoglycanes. Rudland 1995

Cette architecture se construit tout au long de la vie, du stade foetal à la ménopause, sous l'influence des hormones sexuelles (oestrogènes et progestérone) et d'un certain nombre de facteurs de croissance. Une coopération permanente s'établit entre les deux compartiments cellulaires. Ainsi, la structure assez complexe du sein et la forte influence du milieu extracellulaire sur le compartiment épithélial, rendent cet organe non seulement sensible aux cancers mais aussi difficile à traiter.



Du fait de l'hétérogénéité à la fois histopathologique, moléculaire et phénotypique de la maladie, les traitements sont proposés au cas par cas suivant de nombreux critères concernant à la fois la patiente et la tumeur. Cette dernière est classée, au niveau clinique, en deux types :

- infiltrante : souvent canalaire, rarement lobulaire, les cellules cancéreuses se disséminent vers les ganglions et des sites métastatiques privilégiés comme les os, les poumons, le foie et le cerveau,
- intracanaire ou *in situ* quand les cellules cancéreuses ne se disséminent pas.

L'adaptation du traitement fait ensuite intervenir une description détaillée de

<b>A</b>	
Dimension de la tumeur T	T1 tumeur < 2 centimètres
	T2 dimension comprise entre 2 et 4 centimètres
	T3 tumeur > 4 centimètres
	T4 extension aux structures adjacentes (os, tissus mous du cou, muscles profonds de la langue)
Atteinte des ganglions N	N0 pas de signe d'atteinte ganglionnaire régionale
	N1 1 ganglion homolatéral < 3 centimètres
	N2 1 ganglion homolatéral de dimension 3 à 6 centimètres
	N3 ganglions homolatéraux, controlatéraux ou bilatéraux < 6 centimètres
Différenciation histologique G	G0 degré élevé de différenciation histologique
	G1 degré moyen de différenciation
	G2 degré faible ou absence de différenciation
	Gx degré de différenciation non évaluable
Métastases M	M0 absence de métastases
	M1 présence de métastases
<b>B</b>	
Le grade SBR comprend trois grades I II III, obtenu par l'addition de trois critères : architecture, atypies cytonucléaires et nombre de mitoses. Les trois critères sont cotés en 1,2 et 3 :	<b>nombre de mitoses</b> : le nombre de mitoses est recherché sur 20 champs au fort grossissement en périphérie de la tumeur. Le nombre de mitose le plus important par grand champ est retenu.
<b>architecture</b> :	1 si le nombre est de 1 ou 0
1 la tumeur comprend que des tubes	2 si le nombre est de 2
2 partiellement tubulaires	3 si le nombre est de 3 ou plus
3 la tumeur ne comprend aucun tube	L'addition des trois critères permet de réaliser le
<b>atypies cytonucléaires</b> :	grade : • 3,4,5 → Grade I
1 noyaux réguliers monomorphes	• 6,7 → Grade II
2 atypies modérées	• 8,9 → Grade III
3 noyaux pléomorphes avec atypies marquées	

**Figure 3 : Éléments de description des tumeurs de cancers du sein**

**A.** Détails de la classification TNM et histologique (Gx) des tumeurs du sein. **B.** Description du grade histopronostique de Scarff Bloom et Richardson (SBR).

l'extension de la tumeur avec le stade (classification TNM, **Fig. 3.A.**), l'agressivité (grade de SBR, **Fig. 3.B.**), la sensibilité hormonale (présence des récepteurs aux oestrogènes [ER] et à la progestérone [PR]) et l'étude des ganglions (envahissement ou non).

Pour à la fois mieux classifier les cancers du sein et optimiser les traitements, de nombreuses études par puces ont été réalisées entre 2000 et 2006. **Sorlie 2001 & 2003, van de Vijver 2002, van't Veer 2002, Chang 2004 & 2005, Paik 2004, Ma 2004, Bertucci 2005, Sotiriou 2003, Wang 2005, Hu 2006** L'analyse des transcriptomes de centaines de tumeurs a ainsi permis à plusieurs équipes de proposer des combinaisons de gènes pertinents pour le pronostic, la prédiction ou les deux. On peut citer en exemple cinq de ces modèles qui ont été repris dans une publication de Fan et coll. **Fan 2006** afin d'être comparés :

- le modèle "70-gene good-versus-poor outcome" développé par van de Vijver et coll. et van't Veer et coll. **van de Vijver 2002, van't Veer 2002**,
- le modèle "wound-response" développé par Chang et coll. **Chang 2004 & 2005**,
- le modèle "recurrence-score" développé par Paik et coll. **Paik 2004**,
- le modèle "intrinsic-subtype (luminal A, luminal B, basal-like, HER2-positif [HER2+], récepteurs aux oestrogènes négatifs [ER-] et pseudo normal)" développé par Perou et coll. **Perou 2000, Sorlie 2001 & 2003, Hu 2006**,
- le modèle "two-gene-ratio (ratio du taux d'expression de homeobox 13 [HOXB13] et du récepteur à l'interleukine 17B [IL17BR])" développé par Ma et coll. **Ma 2004**

Fan et coll. mettent en évidence que si tous ces modèles de prédiction sont efficaces (au moins 70% de réussite) et comparables c'est d'une part parce qu'ils reflètent les mêmes caractéristiques des cancers du sein (comme par exemple le statut ER) et d'autre part parce que même s'ils présentent très peu de redondance dans leur sélection de gènes, ces modèles concernent les mêmes voies cellulaires (le plus souvent en rapport avec la prolifération, le cycle cellulaire, l'apoptose, etc.).

Ainsi, les cliniciens disposent d'un certain nombre de renseignements leur permettant d'orienter les traitements. Cependant, ces données restent assez parcellaires

et empiriques ce qui fait que la chirurgie reste encore aujourd'hui le meilleur moyen de soigner les cancers du sein. Les rechutes étant assez nombreuses, notamment à long terme, il est important d'approfondir ces études et d'envisager de nouvelles voies permettant une caractérisation efficace des tumeurs.

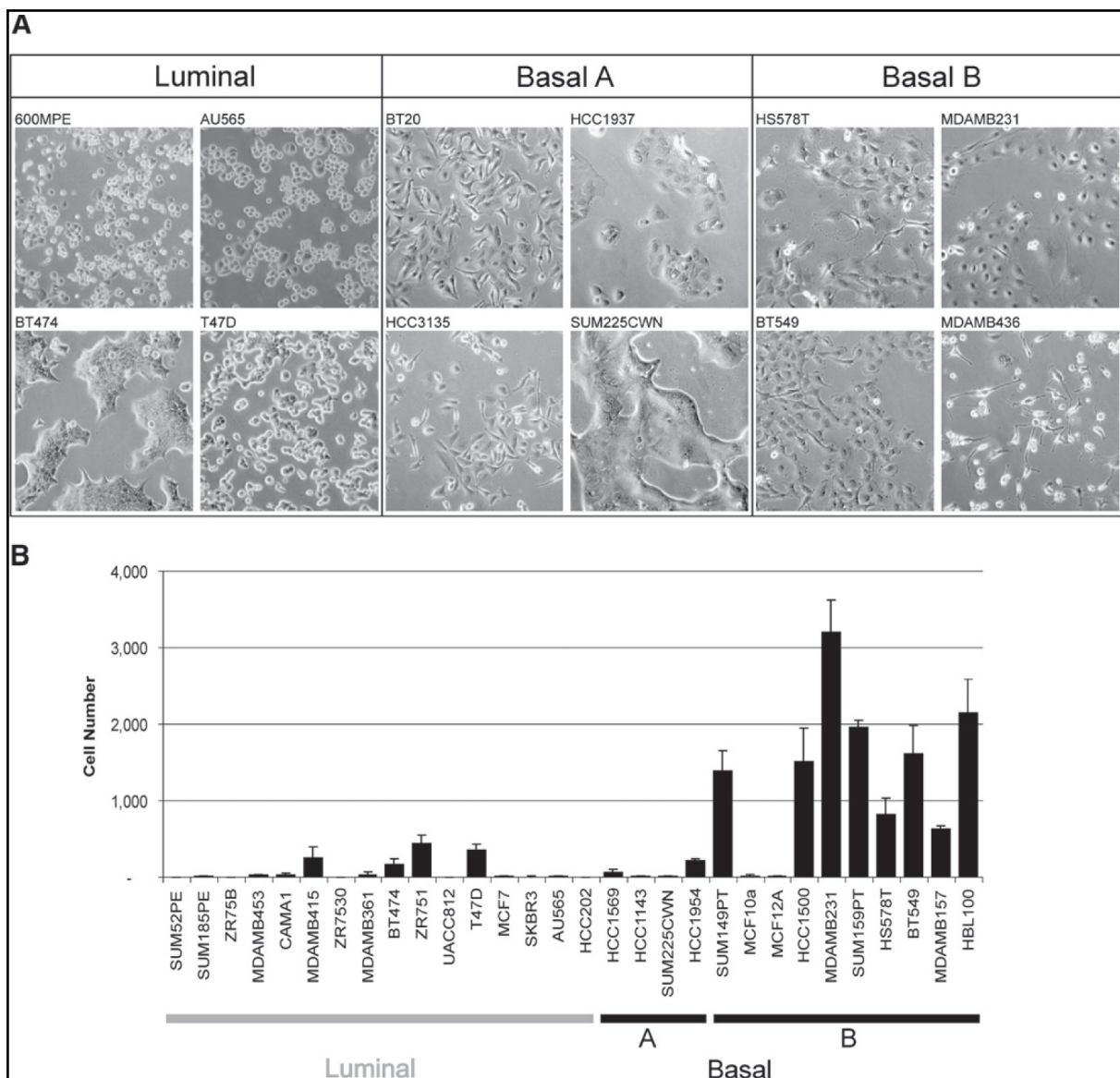
### I.1.2. Caractérisation moléculaire des cancers du sein

Contrairement à la majorité des tumeurs solides comme les cancers du colon ou du pancréas, les cancers du sein sporadiques ne présentent pas une mutation dominante conduisant à la progression tumorale. [Deramaudt 2005](#), [Segditsas 2006](#) Nombre d'études de profil d'expression de gènes, de caractérisations chromosomiques et de recherche de mutations ont montré que, très souvent, la tumorigénèse mammaire ne s'instaurait pas de façon linéaire, étape par étape, mutation par mutation. [Buerger 1999, 1999 bis & 2001](#) Il est rare que différentes tumeurs mammaires partagent la même mutation à l'exception des gènes *TP53* et *PIK3CA* retrouvés mutés dans 20 à 40% des cas suivant les études. De fait, il devient d'autant plus difficile de comprendre l'origine des tumeurs et d'obtenir ainsi une classification fiable.

A l'heure actuelle, il existe environ 18 sous-groupes "cliniques" de tumeurs mammaires sans réelle pertinence lorsque que l'on sait que 65% sont classées comme "carcinome sans type spécial" ! Cet état de fait est expliqué en partie par la méconnaissance de la hiérarchie cellulaire fine de l'épithélium mammaire. D'une nature complexe avec des types cellulaires spécifiques mais mal caractérisés, ce n'est que récemment que ce tissu épithélial commence à être pris en compte avec l'apparition, par exemple, au niveau des lignées modèles, de types cellulaires précis. Ainsi, il a été établi une classification en 5 types cellulaire des tumeurs mammaires humaines selon l'expression de l'ARNm de gènes spécifiques : luminales A, luminales B, ERBB2, basales et pseudo-normales. [Sorlie 2001 & 2003](#), [Sotiriou 2003](#), [Perou 2000](#) En

2006, l'étude publiée par Neve et coll. est venue compléter cette description par une étude très détaillée des lignées cellulaires dans la même optique. Il en résulte une classification en types luminal, basal A et basal B qui reflète assez bien, malgré quelques différences, les types tumoraux (Fig. 4.). La conclusion majeure proposée par cette équipe est que les lignées tumorales mammaires sont un très bon modèle pour, par exemple, déterminer plus précisément quel traitement proposer aux patients en fonction de leur type tumoral ou encore tester les éventuelles résistances aux traitements.

Une des principales différences relevée entre tumeurs et lignées est le type



**Figure 4 : Relation entre profils de transcription et fonction biologique** (fig. 6. Neve et coll. 2006)  
**A.** Morphologie des lignées cellulaires poussant en boîtes plastiques. **B.** Potentiel invasif de 30 lignées tumorales mammaires mesuré à l'aide d'une chambre de Boyden modifiée. Chaque histogramme représente la moyenne des mesures de trois puits  $\pm$  SD.

ERBB2. L'amplification de ce gène (20 à 30% des cancers du sein) est un critère de mauvais pronostic important et implique à l'heure actuelle un traitement spécifique au Trastuzumab (commercialisé par Roche sous le nom Herceptin, anticorps humanisés anti-HER2/neu/ERBB2). Dans les lignées, cette amplification se retrouve à la fois dans le type luminal et dans le type basal A. Cependant, comme observé pour les tumeurs, le taux de réponse à ce traitement n'est seulement que d'une lignée sur trois ce qui laisse un doute sur la validité du type ERBB2 tumoral.

Une autre caractéristique moléculaire importante dans les cancers du sein est l'association ER négatif, PR négatif et ERBB2 non amplifié. Ces "triple-négatifs", qui sont de mauvais pronostic, agressifs et ne répondant pas *a priori* aux traitements classiques, se retrouvent à 85% dans le type tumeurs basales. [Nielsen 2004](#) Pourtant, si certains sont tentés de définir ce type tumoral à partir de ces 3 marqueurs, Rakha et coll. rappellent que 15 à 45% des tumeurs dites basales expriment au moins l'un des trois. [Rakha 2008](#)

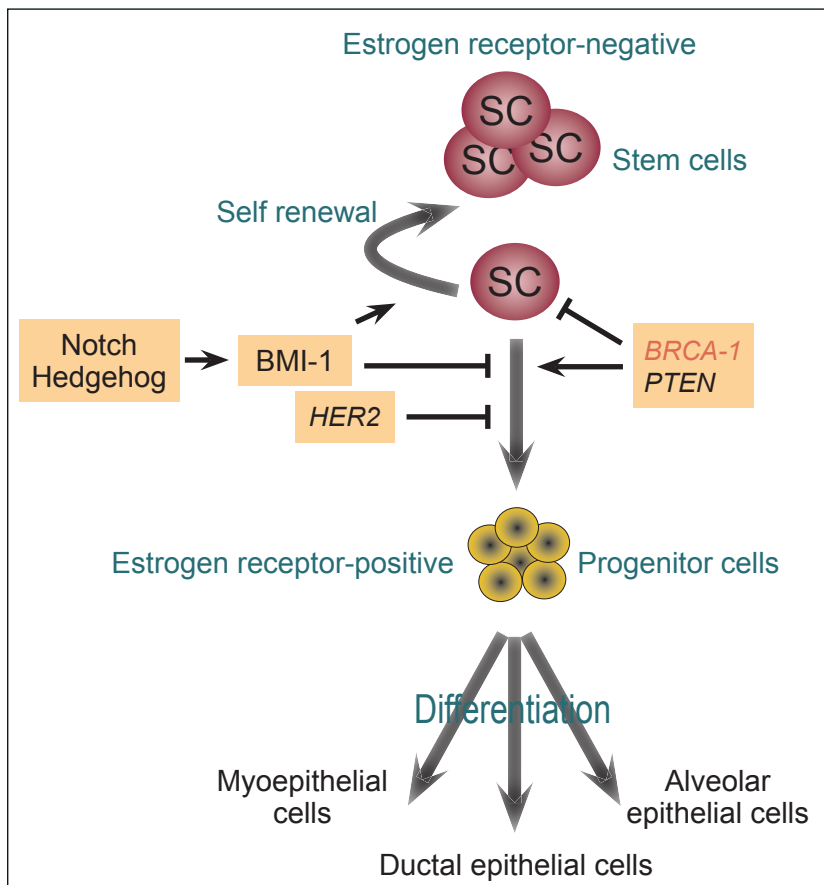
De fait, il n'existe pas encore de classification internationale consensuelle par origine histologique des tumeurs du sein malgré des efforts importants en ce sens. L'hétérogénéité de ces cancers est telle que, malgré l'analyse par puce de milliers de gènes, il reste à trouver des marqueurs moléculaires en association avec les marqueurs cliniques pour déterminer clairement des types tumoraux forts qui permettent un diagnostic, un pronostic et un traitement fiables pour les patients. Un des aspects de ce projet de thèse s'inscrit dans cette démarche de recherche de nouveaux gènes candidats afin de reclassifier plus précisément les cancers du sein. En effet, l'idée est d'étudier seulement quelques gènes (*dicer* et *APC-11* sont les deux meilleurs exemples) qui, de par leur importance dans une ou des voies cellulaires données, sont de très bon candidats pour devenir de nouveaux marqueurs (pronostiques, diagnostiques ou thérapeutiques) de la maladie.

## I.1.3. Origine des cancers du sein

Jusqu'à présent, seules les conséquences cliniques et moléculaires des cancers du sein ont été évoquées. Il importe donc maintenant de s'intéresser, quoique brièvement, aux origines des mécanismes cellulaires impliqués dans la carcinogénèse mammaire et avoir ainsi une vue d'ensemble de cette pathologie. Actuellement, deux grandes hypothèses sont avancées pour expliquer l'apparition des cancers du sein sporadiques et héréditaires. Toutes deux permettent d'expliquer l'extraordinaire hétérogénéité observées dans ces cancers. Une première théorie s'appuie sur la notion de cellules souches cancéreuses qui a fait ses preuves pour les cancers hématologiques tels que les leucémies mais aussi, plus récemment, pour des cancers solides comme les glioblastomes. [Bonnet 1997, Singh 2004](#) Ces cellules, en petit nombre car issues d'une(de) cellule(s) souche(s) de tissu normal, seraient capables à la fois d'initier et de maintenir la croissance tumorale. [Pardal 2003](#) Les raisons en sont assez simple : tout d'abord, si l'on envisage les cancers comme une accumulation séquentielle et progressive d'anomalies génétiques, les cellules souches, de par leur durée de vie longue, apparaissent comme des cibles idéales pour la transformation maligne. Ensuite, certaines propriétés des cellules souches comme le renouvellement et la différenciation (**Fig. 5**) seraient conservées par les cellules souches cancéreuses et expliquerait ainsi l'apparition d'une tumeur, présentant, de plus, des types cellulaires différents. [Vermeulen 2008, Ponti 2005](#) Des observations allant dans ce sens ont été faites dans les cancers du sein à partir de prélèvements de cellules cancéreuses issues d'effusions pleurales. Ces dernières se sont révélées hautement tumorigéniques chez la souris et possédaient des propriétés de cellules souches. Les particularités phénotypiques de ces cellules sont une expression forte du marqueur antigénique CD44 et au contraire une expression faible, voir inexistante, du marqueur antigénique CD24 (CD44<sup>+</sup>CD24<sup>-/low</sup>). [Al-Hajj 2004](#)

---

La deuxième théorie fait intervenir le processus de transition épithélio-



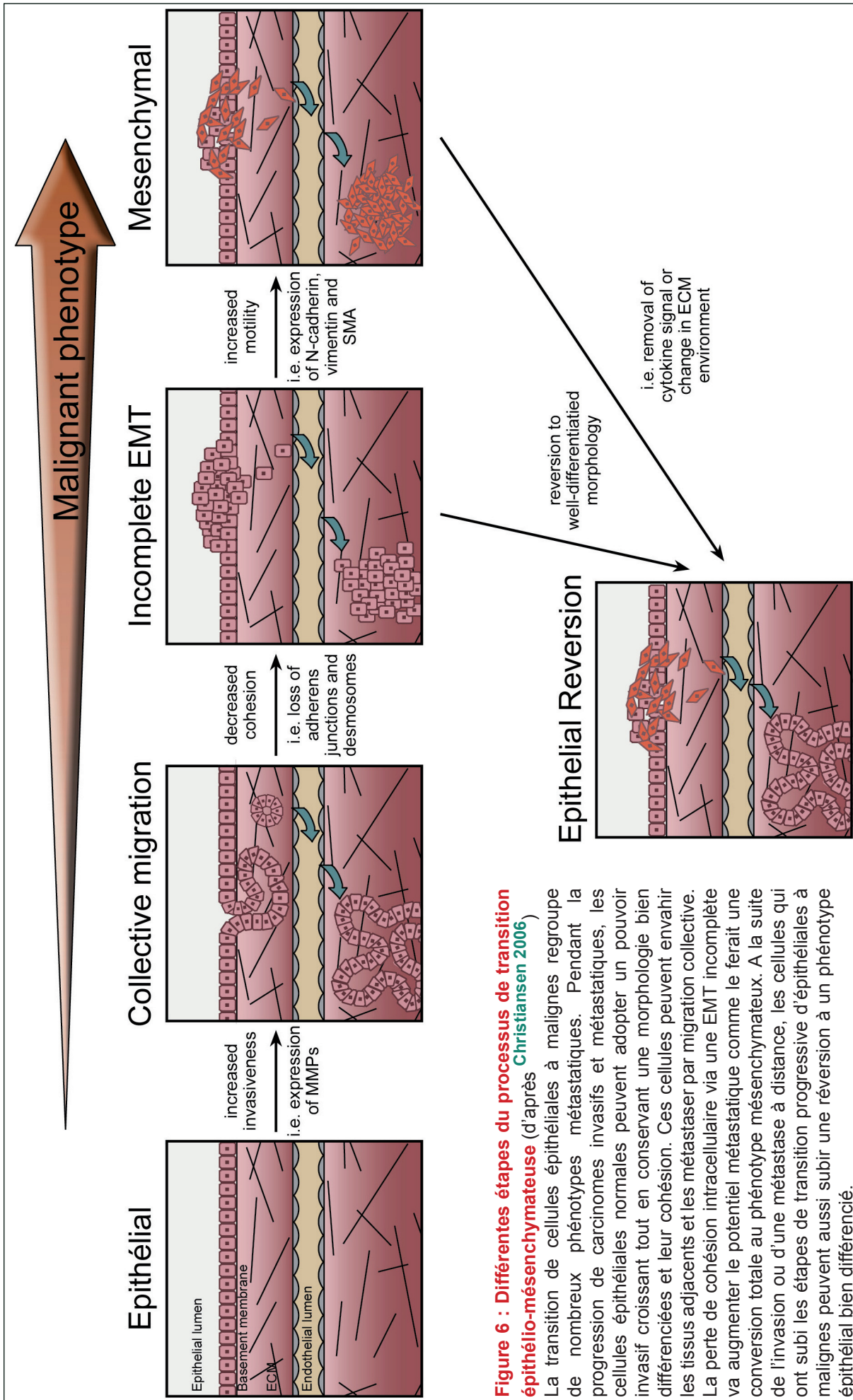
**Figure 5 : Voies d'auto-renouvellement et de différenciation des cellules souches du sein** (fig. 2. [Kakarala 2008](#)).

A la fois les cancers du sein sporadiques et héréditaires peuvent trouver leurs origines dans des cellules souches ou progénitrices mammaires par une dérégulation du processus, normalement très bien régulé, d'auto-renouvellement des cellules souches. Ce peut être le résultat de la perte de fonction du gène *BRCA1* dans les cancers du sein héréditaires. Dans les cas de cancers sporadiques, ce peut être le résultat de la perte du gène *PTEN* ou de l'activation des voies *HER2* (Human Epithelial growth factor Receptor 2), *Notch* ou encore *Hedgehog*. Les conséquences sont une expansion clonale de cellules souches constituant des cibles pour de futures événements carcinogéniques.

mésenchymateuse (EMT, [Fig. 6](#)). Ce processus se manifeste d'une part avec la perte pour les cellules épithéliales de marqueurs caractéristiques tels que la E-cadherine (jonctions serrées) et la  $\beta$ -catenine (interaction avec les cadherines) et d'autre part avec l'expression de marqueurs mésenchymateux comme la vimentine (filament intermédiaire dynamique du cytosquelette impliqué dans le maintien de l'intégrité de la cellule) et la fibronectine (protéine extracellulaire impliquée dans la motilité). [Yang 2008](#)

**pour revue** Il a été observé que cette transition était dans de nombreux cas à l'origine de la diffusion métastatique des cancers. [Christiansen 2006](#) On retrouve alors la fonction première de l'EMT nécessaire au développement embryonnaire, à savoir permettre la migration cellulaire. [Greenburg 1982](#) De plus, il a été montré récemment que les cellules subissant l'EMT pouvaient acquérir des caractéristiques de cellules souches comme notamment l'auto-renouvellement. [Mani 2008](#) Ainsi, ce processus pourrait permettre à des cellules différenciées une trans-différenciation aboutissant à l'initiation d'une tumeur.

Comme souvent en biologie, il est très probable que ces deux hypothèses soient confirmées, voir coexistent, suivant les contextes et types de cancer étudiés.



**Figure 6 : Différentes étapes du processus de transition épithélio-mésenchymateuse** (d'après Christiansen 2006)

La transition de cellules épithéliales à malignes regroupe de nombreux phénotypes métastatiques. Pendant la progression de carcinomes invasifs et métastatiques, les cellules épithéliales normales peuvent adopter un pouvoir invasif croissant tout en conservant une morphologie bien différenciées et leur cohésion. Ces cellules peuvent envahir les tissus adjacents et les métastaser par migration collective. La perte de cohésion intracellulaire via une EMT incomplète va augmenter le potentiel métastatique comme le ferait une conversion totale au phénotype mésenchymateux. A la suite de l'invasion ou d'une métastase à distance, les cellules qui ont subi les étapes de transition progressive d'épithéliales à malignes peuvent aussi subir une réversion à un phénotype épithélial bien différencié.



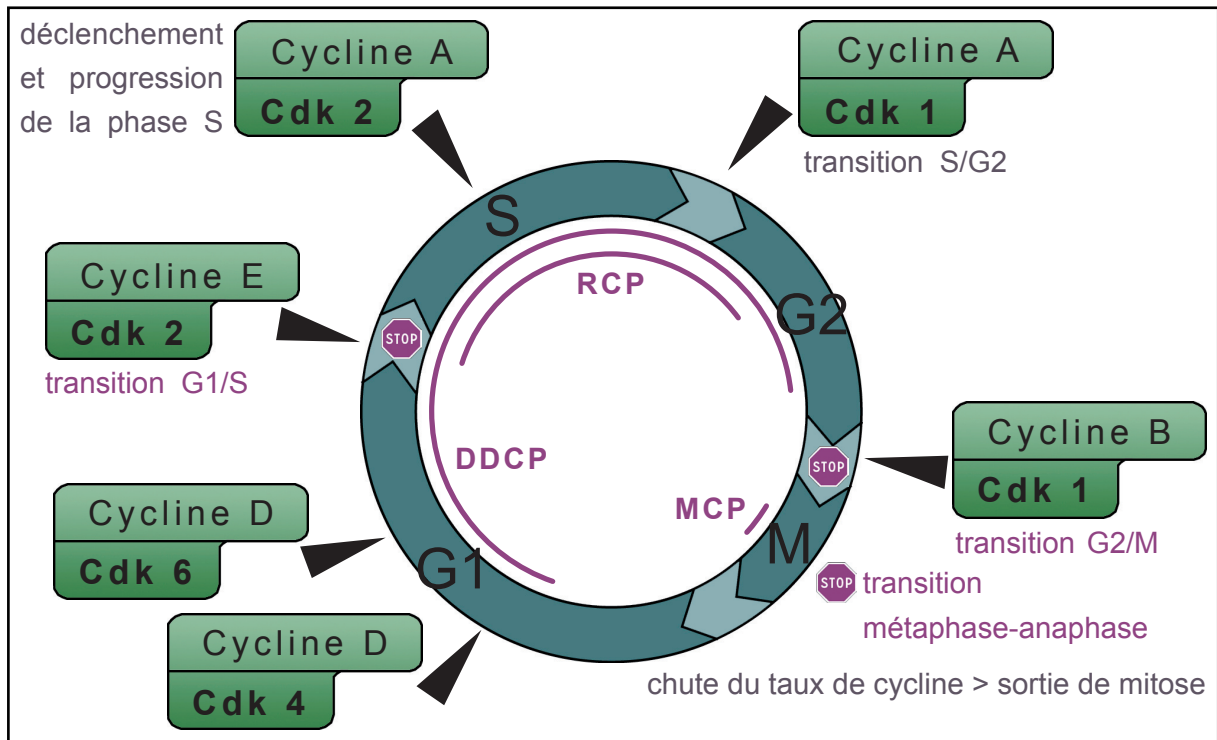
Au-delà des processus, il reste qu'une cellule, pour devenir cancéreuse, doit dériver, accumuler des anomalies de toutes sortes. On peut alors observer que certains mécanismes cellulaires, telle que la division, sont plus affectés que d'autres par des dysfonctionnements de par l'avantage sélectif important qu'ils confèrent au cours du processus de tumorigénèse.

## 1.2. Cycle et points de contrôle cellulaire : rôle du complexe APC/C

### 1.2.1. Aperçu du cycle cellulaire et des points de contrôle associés

Un des fondements des cancers est la dérégulation de la division cellulaire aboutissant à une prolifération intense et anarchique des cellules tumorales mais aussi à une instabilité chromosomique retrouvée dans presque tous les cancers du sein. [Shih 2001](#) Pourtant, ce mécanisme cellulaire primordial est un exemple de précision et de régulation grâce à de nombreux systèmes de contrôles. En explorant brièvement les différentes étapes du cycle cellulaire et en s'attardant plus particulièrement sur la transition métaphase-anaphase en cours de mitose, il devient clair que la dérégulation d'un des acteurs de cette machinerie peut être une étape majeure de la progression tumorale.

Une cellule qui se divise doit relever deux défis : la duplication conforme de son ADN (phase S) et la ségrégation de chacune des copies de son génome (phase M). Pour réussir, les phases cruciales précitées sont préparées (phases G1 et G2 respectivement) et régulées efficacement (points de contrôle du cycle cellulaire). Au niveau protéique, les acteurs majeurs de ce mécanisme sont les oscillations des cyclines associées aux kinases cycline dépendantes (Cdks) sous l'orchestration de deux complexes ubiquitine ligase SCF (SKP1-CUL1-F-box-protein) et APC/C (Anaphase Promoting Complex / Cyclosome). Brièvement, les cyclines-CDKs, suivant leur type, vont permettre les entrées et les sorties des différentes phases (**Fig. 7**) sous



**Figure 7 : Déroulement et surveillance du cycle cellulaire**

Les complexes cycline/Cdk sont les effecteurs du cycle cellulaire et permettent les entrées et sorties des différentes phases. Cependant, ils sont régulés en amont par de nombreux acteurs protéiques et notamment ceux impliqués dans la surveillance du bon déroulement du cycle. On retrouve, entre autres, ATM, ATR, p53, p21 qui garantissent l'intégrité et une réplication complète de l'ADN notamment lors de deux points de contrôles clés (STOP) (transitions G1/S et S/G2). Une autre étape primordiale de la division de la cellule est la transition métaphase-anaphase lors de la mitose. Le point de contrôle associé permet la vérification du bon alignement des chromosomes ainsi que l'attachement des kinetochores aux microtubules. L'acteur clef qui déclenche la séparation des chromatides est le complexe APC/C. DDCP : DNA Damage CheckPoint ; RCP : Replication CheckPoint ; MCP : Mitotic CheckPoint.

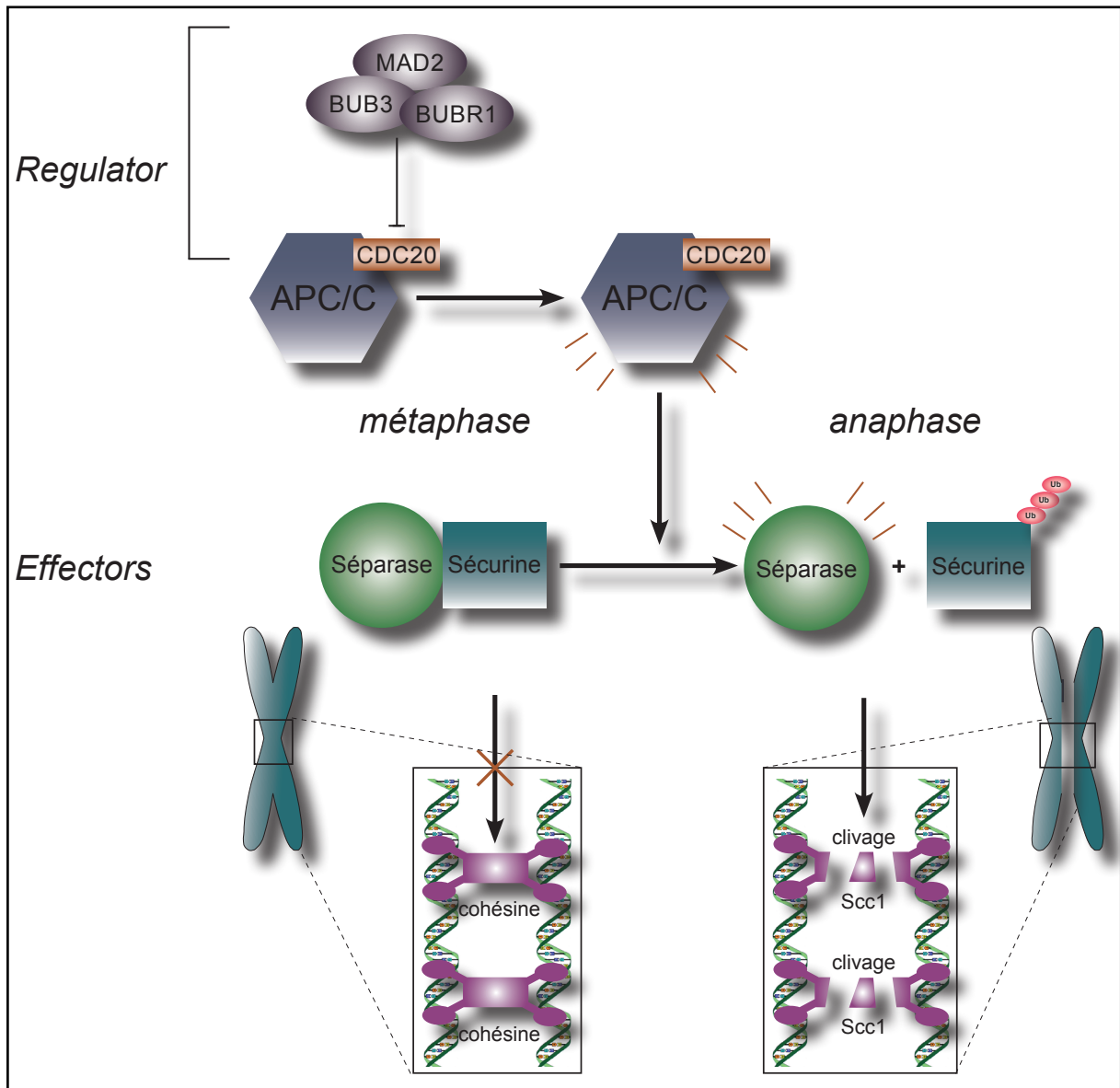
le contrôle de divers acteurs ou complexes protéiques garants du bon déroulement de chaque phase. Ces mécanismes de surveillance sont capables de détecter des anomalies (principalement des dommages à l'ADN) et d'induire un arrêt du cycle le cas échéant. [McGowan 2004](#), [Helt 2005](#), [Sancar 2004](#) Bien que ces mécanismes soient actifs en permanence, il a été mis en évidence trois points de contrôle primordiaux ([Fig. 7](#)) qui correspondent à des transitions essentielles entre phases [Lukas 2004](#) :

- entre les phases G1 et S, l'intégrité de l'ADN est vérifiée et toute cassure entraîne l'arrêt du cycle par (i) destruction de la phosphatase Cdc25A (activation de Cdk4 et 2 normalement). La voie d'activation de ce point de contrôle fait intervenir ATM et Chk2. (ii) accumulation de p21 inhibitrice de Cdk2/Cycline E et A. De même, la voie fait intervenir ATM-Chk2, cette fois pour stabiliser p53 et donc activer p21. Les complexe Cdk/Cycline étant inhibés, ils ne sont plus capables de phosphoryler

la protéine Rb (rétinoblastome) qui séquestre le facteur de transcription E2F nécessaire à la poursuite du cycle. Rb joue ainsi un rôle important en tant que gène suppresseur de tumeur en empêchant une prolifération continue des cellules.

- entre les phases G2 et M, le point de contrôle répliatif permet l'arrêt du cycle dans les cas où la réplication n'est pas achevée et s'il y a des cassures. A ce moment là, Cdc25C, qui induit normalement l'activation de Cdk1/Cycline B permettant l'entrée en mitose, est phosphorylée par Chk1 et Chk2 et ainsi séquestrée dans le cytoplasme. Lorsque la réplication est inachevée, ATR phosphoryle Chk1 qui inhibe Cdk1. De même que précédemment, ATM, via Chk2 peut induire l'accumulation de p53 et activer ainsi p21 qui inhibe Cdk1.
- entre la métaphase et l'anaphase, le point de contrôle mitotique permet de vérifier le bon alignement des chromosomes, le bon attachement des kinétochores aux microtubules ainsi que la tension du fuseau mitotique. Dans le cas d'une anomalie, certaines protéines qui forment le kinétochore activent Mad2 par formation d'un complexe avec Bub3-Mad3/BubR1 (Fig. 8). [Fang 2002](#), [Hardwick 2000](#) Mad2 est alors soluble et capable de venir se fixer sur CDC20 et inhiber le complexe APC/C. La protéolyse de la cohésine qui maintient les centromères attachés ne peut ainsi pas se faire.

Ce complexe APC/C, présent et actif tout au long du cycle cellulaire, est ainsi un élément clef du point de contrôle mitotique. Il en résulte qu'une dérégulation de cette ubiquitine ligase pourrait avoir des conséquences importantes sur la stabilité chromosomique des cellules. C'est pourquoi nous nous sommes intéressés à son fonctionnement.

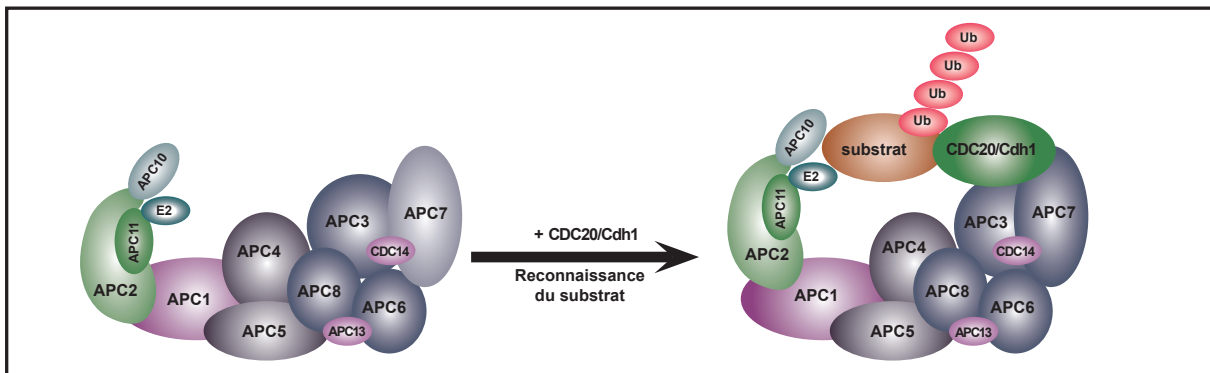


**Figure 8 : Modélisation de la régulation et de la fonction du complexe APC/C lors du point de contrôle mitotique** (d'après Nakayama 2006, Stemman 2001, Pellman 2001).

Le complexe APC/C est régulé par de nombreuses protéines au cours du cycle cellulaire. Lors du point de contrôle mitotique, pendant la métaphase, la fonction E3 ubiquitine ligase du complexe APC/C associé à CDC20 est inhibée par un signal envoyé par les kinétochores non attachés ou mal attachés aux microtubules. Les protéines Mad1-3 et Bub1-3 ont été montrées comme étant essentielles pour ce signal d'attente. [Chen 1996](#), [Li 1996](#), [Taylor 1997 & 1998](#) Cependant, Mad2 serait la seule à se lier à CDC20 et permettre son inhibition. [Dobles 2000](#), [Howell 2000](#), [Michel 2001](#) Une fois les kinétochores correctement attachés, ce signal d'attente de l'anaphase est levé et le complexe APC/C - CDC20 va induire la destruction de la securine par ubiquitination. Ainsi, la séparase initialement séquestrée par cette protéine est libérée et peut aller cliver la sous-unité Scc1 de la cohésine. Ce complexe de 4 protéines (Scc1, Scc3, SMC1 et SMC3) forme des anneaux autour des chromatides soeurs assurant leur cohésion. [Anderson 2002](#), [Weitzer 2003](#), [Haering 2002](#) Présent tout le long des chromosomes avant la mitose, ces complexes cohésine détruits pour la plus grande partie par PLK et Aurora B. [Gimenez-Abian 2004](#), [Losada 2002](#), [Sumara 2002](#) Ils ne subsistent principalement qu'au niveau des centromères jusqu'en métaphase. Le clivage de Scc1 permet alors la séparation des chromatides soeurs et leur migration aux pôles opposés de la cellule. [Waizenegger 2000](#)

## I.2.2. Le complexe APC/C : implications dans la tumorigénèse

La progression dans le cycle cellulaire est notamment régulée par la dégradation par le protéasome 26S d'acteurs clés du cycle. Cette dégradation est réalisée suite à une ubiquitination de ces substrats (**Fig. 9**). Deux complexes E3 ubiquitine ligases interviennent dans la régulation du cycle cellulaire, SCF et APC/C. Une dérégulation de la protéolyse des éléments majeurs du cycle cellulaire peut entraîner une prolifération incontrôlée et une instabilité génomique conduisant à la progression tumorale. [Tatsuka 1998](#), [Terada 1998](#), [Holtrich 1994](#), [Yuan 1997](#) Contrairement au

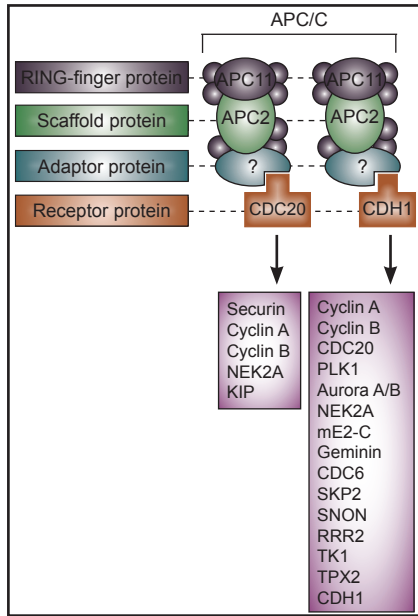


**Figure 9 : Schéma de la structure et modélisation du fonctionnement du complexe APC/C** (d'après [Acquaviva 2006](#)). Ce complexe est composé d'au minimum 11 sous-unités et exerce une fonction ubiquitine ligase permettant la dégradation de protéines du cycle cellulaire et ainsi sa régulation.

complexe SCF, peu d'études ont recherché une implication d'APC/C au cours de la progression tumorale.

Structurellement (**Fig. 10**), le complexe APC/C possède au moins 11 sous-unités chez l'Homme, dont 2 sous-unités invariables : APC11 (protéine à doigts de zinc) et APC2 (protéine de repliement). Les autres sous-unités ont des rôles moins définis. Il existe également 2 protéines activatrices du complexe : CDC20 intervenant dans le passage métaphase-anaphase et CDH1 activant APC/C pour la sortie de mitose. APC/C compte une vingtaine de substrats. On trouve notamment la cycline B1 dont la dégradation entraîne la sortie de mitose et la sécurine dont la dégradation induit la séparation des chromatides soeurs au cours de l'anaphase. Il a été montré qu'une dérégulation, dans le sens d'une sur- ou d'une sous-expression, de la sécurine pouvait

être dramatique pour la cellule. Hagting 2002, Jallepalli 2001 Cette protéine séquestre la séparase capable de cliver le complexe cohésine qui permet, entre autres, l'attachement des chromatides soeurs lors de la métaphase (Fig. 8). C'est pourquoi, dans les deux cas, il a été observé l'apparition d'instabilités chromosomiques et



**Figure 10 : Structure et substrats du complexe APC/C** (d'après Nakayama 2006). Il est à noter le nombre important de substrats.

d'aneuploïdie. Il est très probable que la dérégulation de l'activité de APC/C-CDC20 soit impliquée dans la tumorigénèse, la mutation ou la délétion de cette sous-unité activatrice étant retrouvée dans certains cancers. Li 2003, Singhal 2003, Kim 2005 Par ailleurs, notre laboratoire a montré que des mutations des sous-unités CDC16 et CDC23 sont impliquées dans la progression tumorale du cancer du côlon. Wang 2003

Ces deux sous-unités, ainsi que CDC27, présentent des motifs TPR (tetratricopeptide repeats) et jouent un rôle majeur dans l'établissement de la structure (notamment dans le positionnement des substrats) du complexe APC/C. Acquaviva 2006, Castro 2005 pour revue

De plus, la mutation de chacun de ces 3 gènes a le même effet, à savoir, un arrêt du cycle cellulaire en métaphase. Hartwell 1970 Il a été montré que d'autres sous-unités ne présentant pas le motif TPR pouvaient également intervenir dans la régulation de la structure du complexe, en particulier APC5. Passmore 2003 Récemment, il a été mis en évidence que la sous-unité APC7 était sous-exprimée dans les cancers du sein de mauvais pronostic. Park 2005 De plus, ces deux gènes ont été montrés comme interagissant avec le complexe CBP/p300 et pouvaient ainsi contrôler l'expression de p53. Turnell 2005, Liu 2006

Il est parfaitement établi aujourd'hui que des mécanismes épigénétiques peuvent altérer l'expression de certains gènes au cours de la progression tumorale. Comme, de plus, nous n'avons pas trouvé de mutation des gènes APC/C dans les

cancers du sein au cours de notre étude précédente, il nous a semblé intéressant de la compléter par une approche expressionnelle.

### I.3. Les réparations des cassures double-brin de l'ADN

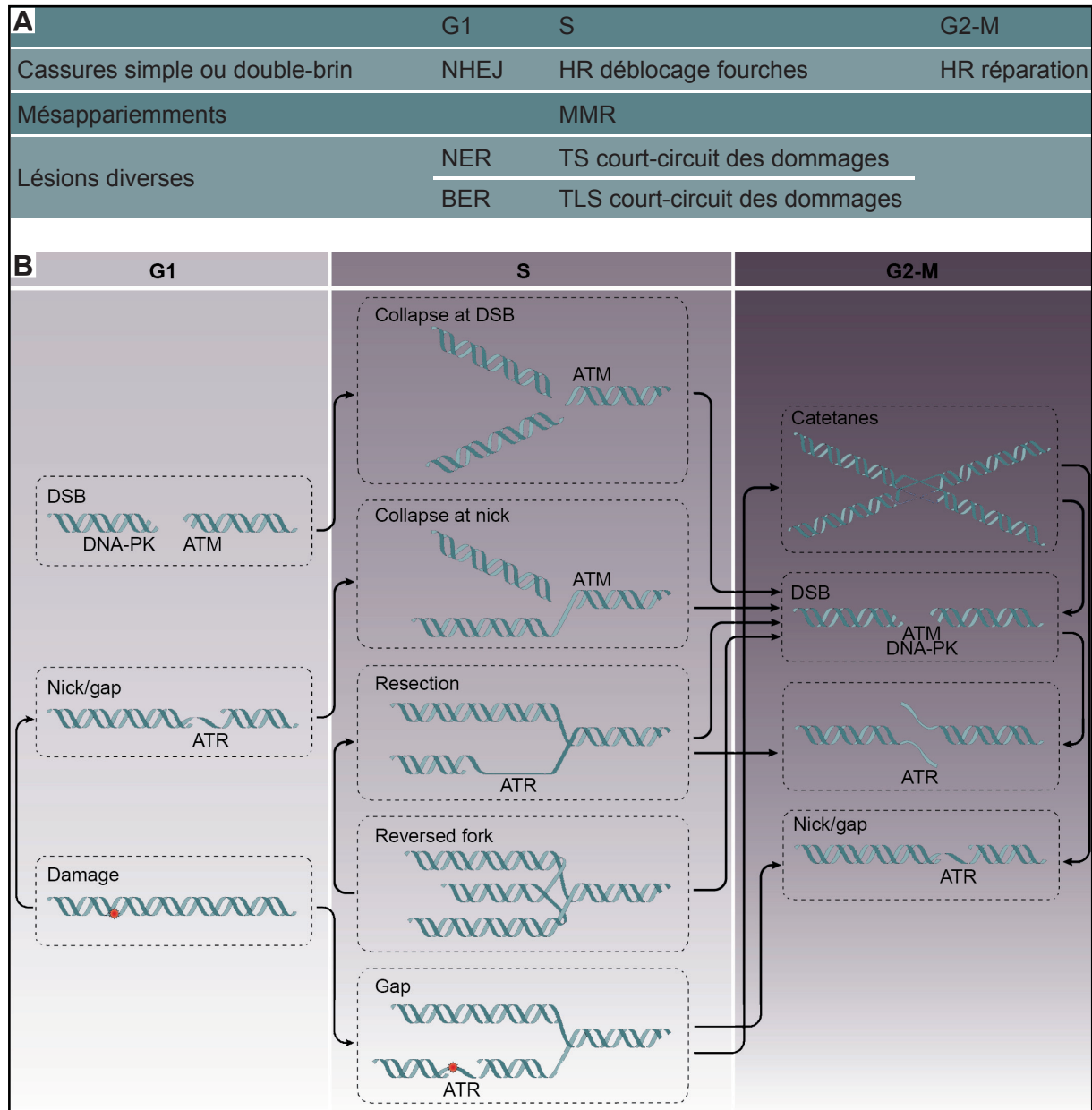
#### I.3.1. Présentation des mécanismes

**C**ycle cellulaire et réparations des cassures de l'ADN sont très liés puisqu'un enjeu majeur de la division cellulaire est de transmettre deux copies fidèles et intègres du génome parental. Ainsi, les voies de réparations sont sollicitées en permanence pour répondre aux nombreux stress endogènes et exogènes qui induisent des dommages à l'ADN. Il existe plusieurs voies de réparations de l'ADN suivant le type de lésions (dimères pyrimidiques, cassures simple ou double-brin) mais surtout suivant la phase du cycle cellulaire en cours et donc suivant le point de contrôle qui va être mis en jeu. Lorsque la cellule ne parvient pas à réparer ses dommages, elle reste arrêtée définitivement ou part en apoptose. Un des intérêts du projet de recherche présenté ici portait plus particulièrement sur la stabilité de l'ADN répliqué, je décrirai en détail la recombinaison homologue qui intervient majoritairement dans cette configuration et brièvement les autres mécanismes (Fig. 11).

Dans la plupart des cas, le dommage subit par l'ADN est reconnu par un senseur qui permet le recrutement de complexes de réparation. Ces derniers activent alors un point de contrôle dont les acteurs permettent l'amplification du signal et ainsi de mobiliser toutes les protéines de la voie de réparation adéquate et celles de la machinerie du cycle cellulaire pour en stopper le déroulement. Les protéines majeures de la signalisation/réparation des cassures sont les kinases ATM, ATR et DNA-PK. ATM et DNA-PK répondent principalement aux cassures double-brin alors que ATR est induite par l'ADN simple-brin et les "stalled replication fork" c'est à dire les fourches de réplication bloquées. [Bartek 2007](#) Le recrutement et l'activation de ces trois kinases sur les

dommages se fait par des facteurs spécifiques : NBS1 pour ATM, ATRIP pour ATR et Ku80 pour DNA-PK. [Falck 2005](#), [You 2005](#) Les principaux substrats permettant la médiation de la réparation sont les protéines kinases CHK1 et CHK2 ([Fig. 12.A.](#)), [Matsuoka 2007](#)

En phase G1, deux mécanismes ([Fig. 11](#)) interviennent en priorité du fait de la phase en cours et du type de lésions ne provoquant pas de cassures directes (souvent



**Figure 11 : Dommages à l'ADN et leurs mécanismes de réparation au cours du cycle cellulaire** (d'après [Branzei 2008](#)).

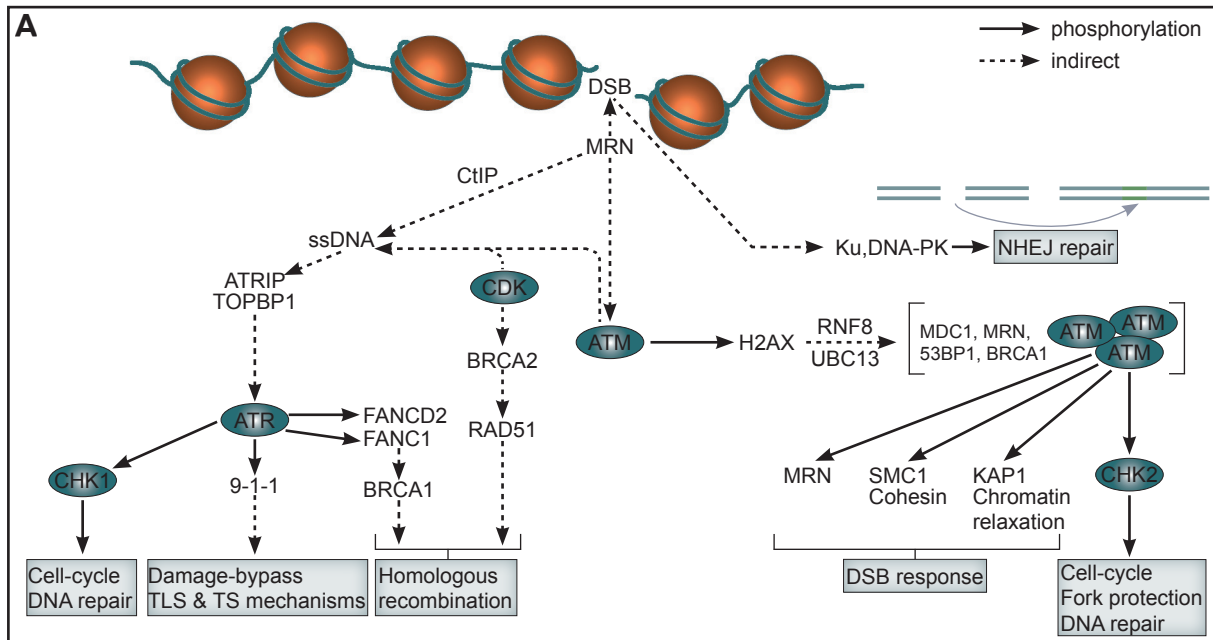
**A.** Résumé des mécanismes de réparation préférentiellement mis en jeu suivant les phases du cycle cellulaire. NHEJ pour Non Homologous End Joining ; NER pour Nucleotide-Excision Repair ; BER pour Base-Excision Repair ; HR pour Homologous Recombination ; Mismatch Repair (MMR) ; TS pour Template Switch et TLS pour Translesion Synthesis.

**B.** Vue d'ensemble des dommages à l'ADN et des kinases activées au cours du cycle cellulaire. Les flèches indiquent le devenir des lésions au cours des phases suivantes si elles ne sont pas réparées. Ainsi, les dommages de G1 aboutissent au blocage des fourches de réplication et la majeure partie des problèmes non résolus en S conduit à des cassures double-brin.



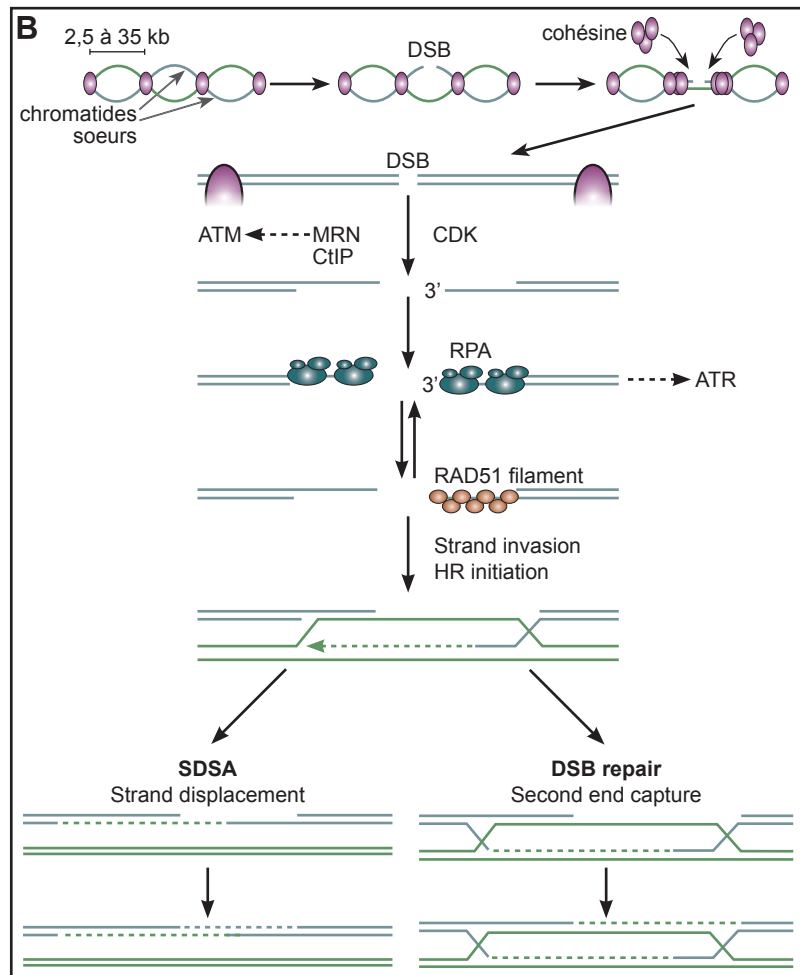
occasionnées par des radicaux libres, des agents chimiques des UV ou des radiations) : la réparation par excision de base (BER) ou de nucléotide (NER). Lors de dommages plus sévères (comme des cassures double-brin dues à des radiations ionisantes), c'est préférentiellement la réparation par ligation d'extrémités non homologues (NHEJ, Non Homologous End Joining) qui intervient (l'absence de chromatide soeur limitant très fortement la recombinaison homologue) avec la possibilité de perte d'information ou de translocation. [Sonoda 1999](#), [Paques 1999](#), [Krogh 2004](#)

En phase S ([Fig. 11](#)), les erreurs observées sont des mésappariements de nucléotides, une accumulation de coupures, des glissements sur les séquences répétées et surtout des blocages des fourches de réplication pouvant aboutir à des résections ou des cassures double-brin. Les mésappariements sont pris en charge par le système MMR (MisMatch Repair) qui élimine les mauvais nucléotides pour que le vide laissé soit comblé par une nouvelle synthèse d'ADN. [Jiricny 2007](#) Le mécanisme BER intervient sur les incorporations d'uraciles et, comme en phase G1, sur les altérations chimiques de nucléotides. [Sancar 2004](#) La phase S est aussi très propice aux coupures simple-brin de l'ADN et ces dernières ne sont généralement pas réparées avant les phases G2 et M lors desquelles la recombinaison homologue est possible. Il existe ainsi deux mécanismes de tolérance de ces coupures faisant intervenir une polymérase capable soit de continuer la réplication par dessus ces vides, soit de les remplir à partir de la nouvelle chromatide soeur en réprimant la voie de réponse due à la lésion. [Lehmann 2007](#), [Branzei 2007](#) Par contre les coupures double-brin, résultant souvent du blocage des fourches de réplication, sont préférentiellement prises en charge par le mécanisme de recombinaison homologue. Pour minimiser la compétition avec le mécanisme NHEJ, il semblerait que les protéines RAD18 et PARP (poly[ADP-ribose]polymérase) diminuent l'affinité de l'hétérodimère Ku avec les DSB. [Hochegger 2006](#), [Saber 2007](#) Le processus de recombinaison homologue nécessite tout d'abord la digestion 5'-3' des extrémités de la cassure grâce, principalement, au complexe MRN (MRE11-Rad50-NBS1) mais aussi aux CDK. [Ira 2004](#), [Aylon 2004](#) Les extrémités 3' simple-



**Figure 12 : Voies de réponse schématique induites par une cassure double-brin (DSB) et détails de la recombinaison homologue** (d'après [Watrin 2006](#), [Branzei 2008](#)).

**A.** Présentation des principales protéines intervenant potentiellement dans la réponse aux DSB. Ku et le complexe MRN (MRE11-RAD50-NBS1) sont en compétition pour se lier aux DSB et activer leur kinase respective DNA-PK et ATM. Si l'association Ku/DNA-PK l'emporte (ce qui est favorisé en phase G1), elle induit la réparation par NHEJ (ligation d'extrémités non homologues). L'activation d'ATM induit des modifications de la chromatine autour du DSB par la phosphorylation de H2A.X et entraîne ainsi une cascade de recrutement de nombreux facteurs (MDC1, MRN...). Lorsque de l'ADN simple-brin est exposé, la voie ATR est activée. La phosphorylation de CHK1 et CHK2 réduit l'activité des CDK, l'arrêt du cycle et la réparation alors que d'autres cibles de ATM et ATR favorisent la recombinaison homologue comme SMC1, FANCD2 et FANCD1.



**B.** En phases S et G2, la forte activité des CDK favorise la résection du DSB, exposant ainsi la partie 3' d'un ADN simple-brin. Ce fragment peut être recouvert de protéine RPA (Replication Protein A) et alors activé ATR. Cette protéine peut aussi être remplacée par RAD51 ce qui va permettre l'initiation de la HR par invasion de la région homologue. Une première synthèse d'ADN est ainsi possible. Ensuite, soit une hélicase ADN induit un déplacement de brin et entraîne une réaction SDSA (Synthesis-Dependent Strand Annealing), soit le second brin cassé peut être à son tour capturé et former une structure de Holliday qui sera résolue par une endonucléase ou par l'action combinée d'une hélicase et d'une topoisomérase.

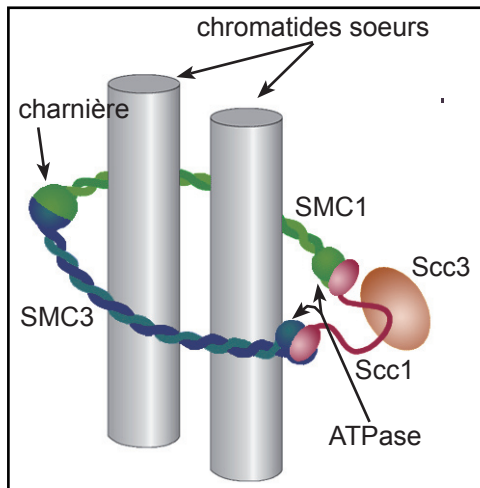
brin résultantes permettent alors l'invasion des brins requise pour la recombinaison homologue (Fig. 12.B.).

En phases G2 et M (Fig. 11), la cohésine joue un rôle très important pour permettre la recombinaison homologue. En effet, mise en place dès la phase S et répartie régulièrement, elle facilite l'étape d'appariement des chromatides soeurs nécessaire au mécanisme. [Hirano 2006](#), [Uhlmann 1998](#) Il est à noter, que les mécanismes de réparation sont très peu présents en mitose du fait de la très forte compaction de l'ADN.

Lorsque les mécanismes de la réparation des lésions subies par l'ADN et leur contrôle sont dérégulés, ils favorisent la progression tumorale. Or, le bon déroulement d'un des mécanismes principaux qu'est la recombinaison homologue, nécessite une structure physique particulière des chromosomes. Le complexe cohésine devient alors un élément clef dont la mise en place et la régulation peuvent être des cibles importantes lors de la tumorigénèse.

### I.3.2. Interactions des voies de réparations et de l'interférence ARN

Le complexe cohésine est formé de quatre sous-unités : les ATPases SMC1 et SMC3 (Structural Maintenance of Chromosomes), Scc1/Rad21 et Scc3 (Fig. 13.). [Guacci 1997](#), [Michaelis 1997](#), [Losada 1998](#), [Toth 1999](#) La mise en place de ce complexe nécessiterait l'intervention de deux autres membres de la famille, Scc2 et Scc4, et l'activité ATPasique des SMC. [Weitzer 2003](#), [Arumugam 2003](#), [Ciosk 2000](#), [Watrinn 2006](#) Cependant, des études récentes viennent bouleverser le modèle établi du fonctionnement de la cohésine. Il a en effet été montré que la liaison SMC1-SMC3 pouvait être rompue pour permettre l'enserrement de l'ADN. [Gruber 2006](#), [Milutinovich 2007](#) En ce qui concerne sa dégradation, il existerait deux voies différentes suivant que la cellule se trouve en anaphase ou pas. En effet, avant l'anaphase, la cohésine disposée tout au long des



**Figure 13 : Modélisation du complexe cohésine** (d'après Haering 2002).

Le complexe cohésine est constitué de 4 sous-unités : SMC1 et SMC3 (Structural Maintenance of Chromosome) et Rad21/ Scc1 et Scc3. Les deux premières protéines sont liées par une région charnière centrale et possèdent deux parties globulaires en C et N-ter, cette extrémité portant la fonction ATPasique. Rad21/Scc1 est une sous-unité formant un pont entre les SMC et interagissant aussi avec Scc3. Dans le modèle classique, Scc1 est clivée lors de l'anaphase par la séparase ce qui permet de libérer les chromatides sœurs *a priori* enserrées dans l'anneau de cohésine. D'autres théories proposent que deux complexes cohésine permettent le maintien des chromatides sœurs.

bras des chromosomes serait dégradée par phosphorylation de Scc3 par PLK1 (Polo-Like Kinase 1). [Sumara 2002](#) Par contre, lors de la transition métaphase-anaphase, la cohésine centromérique, seule encore présente sur les chromosomes, est dégradée par l'action de la séparase sur Scc1. [Waizenegger 2000](#)

L'importance du complexe pour la réparation a été montrée chez la levure en 1992 lorsque Birkenbihl et coll. mirent en évidence que les cellules mutées pour Scc1/ Rad21 ne pouvaient plus réparer leurs DSB correctement. [Birkenbihl 1992](#) L'hypothèse classique était de considérer que les complexes cohésine étaient positionnés au niveau des régions intergéniques sans aucune spécificité de séquence. Actuellement, la théorie est que la cohésine serait chargée non spécifiquement puis poussée par les complexes de réplication jusqu'à ces régions. [Tanaka 1999, Lengronne 2004](#) Lorsqu'une cassure double-brin apparaît, la lésion devient un site de chargement de cohésine et il se forme rapidement un "patch" sur une centaine de kilobases. Ce recrutement fait intervenir le couple Scc2 et Scc4 ainsi que des protéines de la réponse aux cassures et en particulier Mre11 du complexe senseur MRN. [Strom 2004, Unal 2004](#) Cependant, la plupart des données actuelles concernent la levure et très peu de connaissances ont été accumulées pour les vertébrés sur la localisation fine du complexe. Par contre, il est clairement démontré que SMC1 est phosphorylée sur les positions sérine 957 et 966 lors d'un dommage à l'ADN soit par ATM après irradiation ionisante, soit par ATR après irradiation UV. [Garg 2004, Kim 2002, Yazdi 2002](#) Ces phosphorylations sont parties

intégrantes de la réponse aux cassures double-brin de l'ADN or, Scc1, sous-unité du complexe cohésine avec SMC1, a été montré comme étant régulée par le mécanisme d'interférence ARN. Aussi bien chez la levure que chez l'Homme, l'inhibition de l'expression de *dicer* entraîne une mauvaise localisation de Scc1. [Fukagawa 2004](#) Il serait donc possible que cette localisation anormale ait des conséquences sur la réponse aux cassures double-brin de l'ADN.

## I.4. L'interférence ARN

### I.4.1. Le mécanisme au prix Nobel

**E**n 1990, R. Jorgensen et coll. découvrent accidentellement ce qu'ils appellent la co-suppression. [Jorgensen 1990, Napoli 1990](#) En essayant d'augmenter l'intensité de la couleur violette de leurs pétunias en utilisant des plantes transgéniques avec plusieurs copies d'un gène, ils observent l'effet inverse : une dépigmentation des fleurs. Cependant, le mécanisme d'action reste à élucider. L'équipe de M. Wassenegger approfondit les connaissances dans le domaine en découvrant en 1994 que l'introduction d'ARN double-brin chez *A. thaliana* entraîne la méthylation de l'ADN source. Le mécanisme est alors nommé TGS pour Transcriptional Gene Silencing. [Wassenegger 1994](#)

Chez l'animal, la première étape vers la découverte du mécanisme d'interférence ARN (iARN) intervient en 1995 lorsque Guo et Kemphues introduisent un ARN complémentaire (anti-sens) d'un ARN cible chez *C. elegans*. [Guo 1995](#) Le but de l'expérience est d'empêcher l'expression du gène correspondant en neutralisant son messenger par la formation d'un double-brin. La surprise provient du fait que le contrôle de l'expérience, l'ARN sens, provoque aussi l'inhibition de l'expression. Finalement, il faut attendre 1998 pour que l'iARN soit réellement découvert et compris. C'est en injectant des ARN double-brin dans les gonades et les intestins de *C. elegans* que

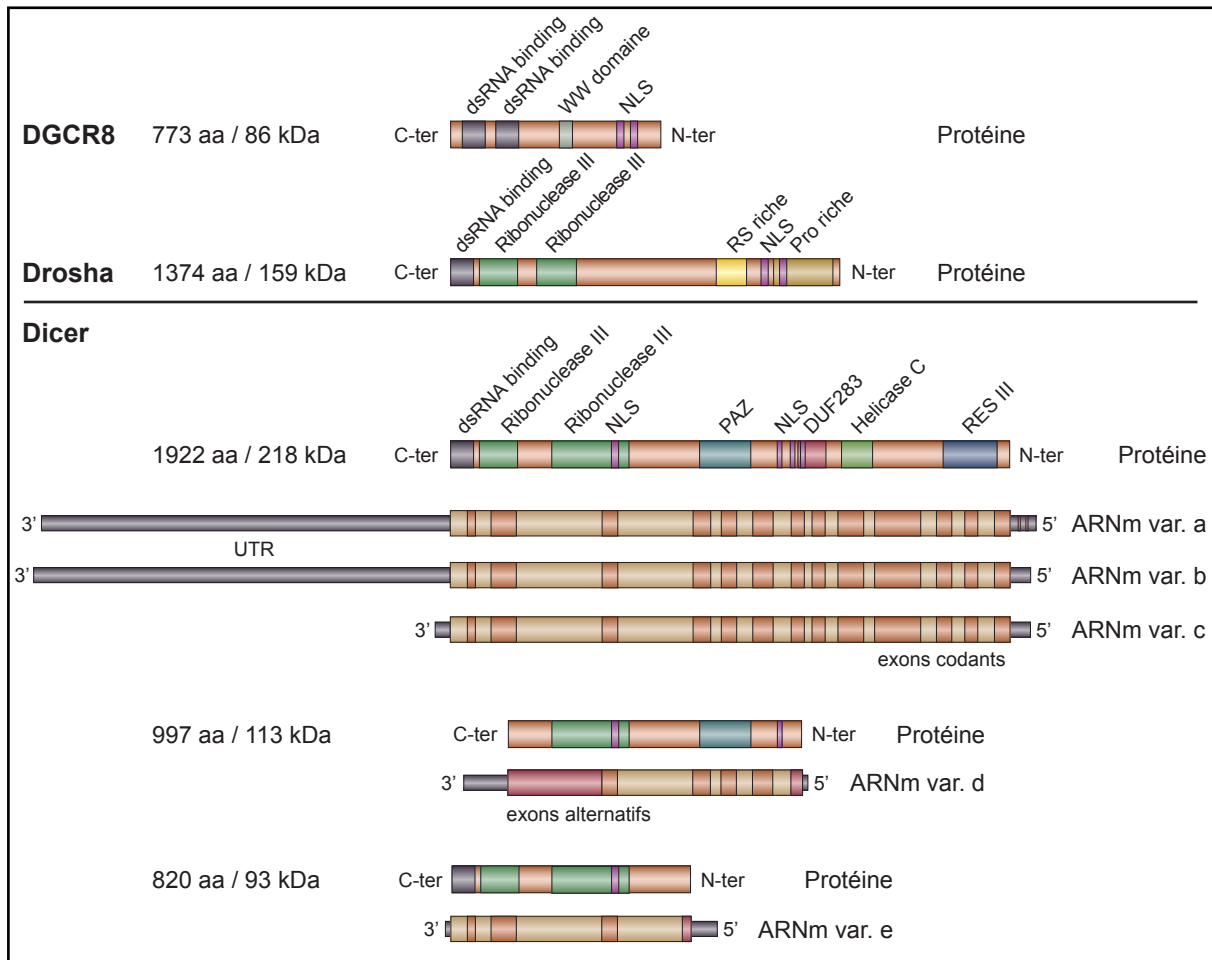
A. Fire et coll. répriment l'expression du gène homologue. **Fire 1998** Toujours la même année, L. Timmons et A. Fire font cette observation après avoir fait ingérer aux vers des bactéries produisant des ARNdb et H. Tabara et coll. après les avoir fait tremper dans une solution de ces mêmes ARN. **Timmons 1998, Tabara 1998**

Par la suite, la communauté scientifique teste rapidement et avec succès cette découverte dans un large éventail d'animaux modèles. Cependant, en ce qui concerne les cellules mammifères, un problème se pose : l'introduction de longs ARNdb induit une mort cellulaire après activation de la voie interféron. Cette voie est médiée par PKR (dsRNA-dependent Protein Kinase) qui s'autophosphoryle en présence d'ARNdb et induit alors la phosphorylation transitoire de eIF2 $\alpha$  (eukaryotic initiation factor 2  $\alpha$ ). Ce facteur entraîne l'inhibition globale de la traduction. La solution est apportée en 2001 par T. Tuschl qui eut l'idée d'utiliser des ARNdb d'au plus 25 paires de bases ce qui permet d'éviter cette réponse non désirée. **Tuschl 2001**

A. Fire et C. Mello ont reçu en 2006 le prix Nobel de médecine pour la découverte du mécanisme d'interférence ARN. A cette époque, les connaissances sur le mécanisme sont en plein essor et il est de plus en plus évident qu'il regroupe sous son nom un nombre important de voies différentes.

#### I.4.2. Les multiples facettes de l'interférence ARN

Le point commun à toutes les voies que représente l'interférence ARN est la présence de petits ARNdb. Par ailleurs, il existe aussi une protéine majeure du mécanisme, à l'origine de ces petits ARNdb, nommée Dicer. Cette protéine très conservée de 218 kDa (pour les formes complètes chez l'Homme) appartient à la famille des RNase III de classe III (**Fig. 14**). On la retrouve sous une forme très simple chez la bactérie *Aquifex aeolicus* (seulement deux domaines : liaison à l'ARN et ribonucléase) mais déjà plus complexe chez *Tetrahymena thermophila*



**Figure 14 : Schémas des différents domaines protéiques de DGCR8, Drosha et Dicer et des variants transcriptionnels de dicer**

DGCR8 (DiGeorge syndrome chromosomal region 8) a été retrouvée dans un complexe protéique contenant Drosha et appelé “microprocesseur”. Le domaine WW qu’elle possède permet sa liaison avec des protéines qui présentent une séquence riche en proline telle que Drosha. De plus, DGCR8 présente deux domaines de liaison à l’ARN ce qui fait qu’elle a été proposée comme déterminante dans la reconnaissance et le clivage des substrats de Drosha. [Han 2004, Gregory 2004](#)

Drosha est une protéine de la famille des RNase III de classe II. Ces protéines se caractérisent par un domaine de liaison à l’ARN (dsRNA binding) et deux domaines Ribonucléase III qui permettent le clivage des ARN. Cette enzyme, qui assure la maturation des pri-miRNA en pre-miRNA dans le noyau, possède aussi des séquences de localisation nucléaire (NLS) et une séquence riche en proline qui intervient dans les interactions protéine-protéine (dans ce cas avec son partenaire privilégié DGCR8). La séquence riche en arginine et sérine (RS riche) a aussi une fonction d’interaction protéique qui doit participer à la formation du complexe “microprocesseur” de 650 kDa.

Dicer est une protéine de la famille des RNase III de classe III qui, en plus des trois domaines spécifiques des classes II, possède un domaine PAZ (Piwi/Argonaute/Zwille) et un domaine hélicase. PAZ est un domaine protéique qui caractérise la famille Argonaute essentielle à l’ARN (présence des Ago dans les complexe RISC, RITS et miRNP). On observe aussi la présence potentielle de séquences de localisation nucléaire (NLS). Cependant, la localisation de Dicer est controversée et préférentiellement donnée comme cytoplasmique. Enfin, on trouve un domaine de fonction inconnue, DUF283 (Domain of Unknown Function), auquel M. Dlakic propose en 2006 une fonction de sélection du brin qui sera utilisé comme guide. [Dlakic 2006](#)

Il existe vraisemblablement plusieurs protéines Dicer issues des différents variants transcriptionnels. Ainsi, les trois variants principaux (a, b et c), dont seuls les extrémités non codantes UTR varient, coderaient pour la protéine complète. Par contre, le variant d, plus court et possédant deux exons codants alternatifs, et le variant e coderaient pour des protéines tronquées mais potentiellement fonctionnelles. Elles pourraient ainsi jouer un rôle de régulation vis à vis de la protéine complète.

(seuls manquent les domaines PAZ et DUF283). **Sasaki 2007** Chez *A. thaliana*, quatre exemplaires **Margis 2006** existent, DCL1 à DCL4, avec pour chacune un rôle clairement défini : DCL1 intervient dans la production des miARN, DCL2 dans les siARN de transcrits antisens agissant en *cis*, DCL3 dans les siARN servant de guide pour les modifications chromatiniennes et DCL4 dans les siARN agissant en *trans* dans la régulation des changement de phases végétatives. **Park 2002, Xie 2004 & 2005, Gascioli 2005, Borsani 2005** Dicer est aussi présente en un à deux exemplaires chez les champignons (une chez *Schizosaccharomyces pombe*, deux chez *Neurospora crassa*), les insectes (deux chez *Drosophila melanogaster*), les urochordates dont les nématodes (une protéine chez *Caenorhabditis elegans*) et nombre d'autres espèces (**Table 1**). Par contre, on ne retrouve, d'un point de vue phylogénique, qu'une seule protéine Dicer

Species	Genome <sup>a</sup>	Argonaute-Piwi	Dicer-like	RdRP	Role in virus, transposon, or repetitive DNA silencing <sup>b</sup>
Excavata					
<i>Giardia intestinalis</i>	Assembly	+	+	+	?
<i>Trypanosoma brucei</i>	Assembly	+	–	–	+
<i>Trypanosoma cruzi</i>	Complete	–	–	–	NA
<i>Leishmania major</i>	Assembly	–	–	–	NA
Chromalveolata					
<i>Paramecium tetraurelia</i>	In progress	+	+	+	?
<i>Tetrahymena thermophila</i>	In progress	+	+	+	+
<i>Plasmodium falciparum</i>	Assembly	–	–	–	NA
<i>Phytophthora sojae</i>	In progress	+	+	+	?
<i>Thalassiosira pseudonana</i>	Assembly	+	–	–	?
Rhizaria					
Data not available					
Archaeplastida					
<i>Cyanidioschyzon merolae</i>	Complete	–	–	–	NA
<i>Chlamydomonas reinhardtii</i>	In progress	+	+	–	+
<i>Arabidopsis thaliana</i>	Complete	+	+	+	+
<i>Oryza sativa (japonica)</i>	Complete	+	+	+	+
Amoebozoa					
<i>Dictyostelium discoideum</i>	Assembly	+	+	+	+
<i>Entamoeba histolytica</i>	Complete	+	–	+	?
Opisthokonta					
<i>Saccharomyces cerevisiae</i>	Complete	–	–	–	NA
<i>Schizosaccharomyces pombe</i>	Complete	+	+	+	+
<i>Neurospora crassa</i>	Complete	+	+	+	+
<i>Aspergillus nidulans</i>	Assembly	+	+	+	?
<i>Caenorhabditis elegans</i>	Complete	+	+	+	+
<i>Drosophila melanogaster</i>	Complete	+	+	–	+
<i>Anopheles gambiae</i>	Assembly	+	+	–	+
<i>Strongylocentrotus purpuratus</i>	In progress	+	+	–	?
<i>Ciona intestinalis</i>	Assembly	+	+	–	?
<i>Homo sapiens</i>	Complete	+	+	–	+

<sup>a</sup> Status of genome sequencing projects taken from <http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi>

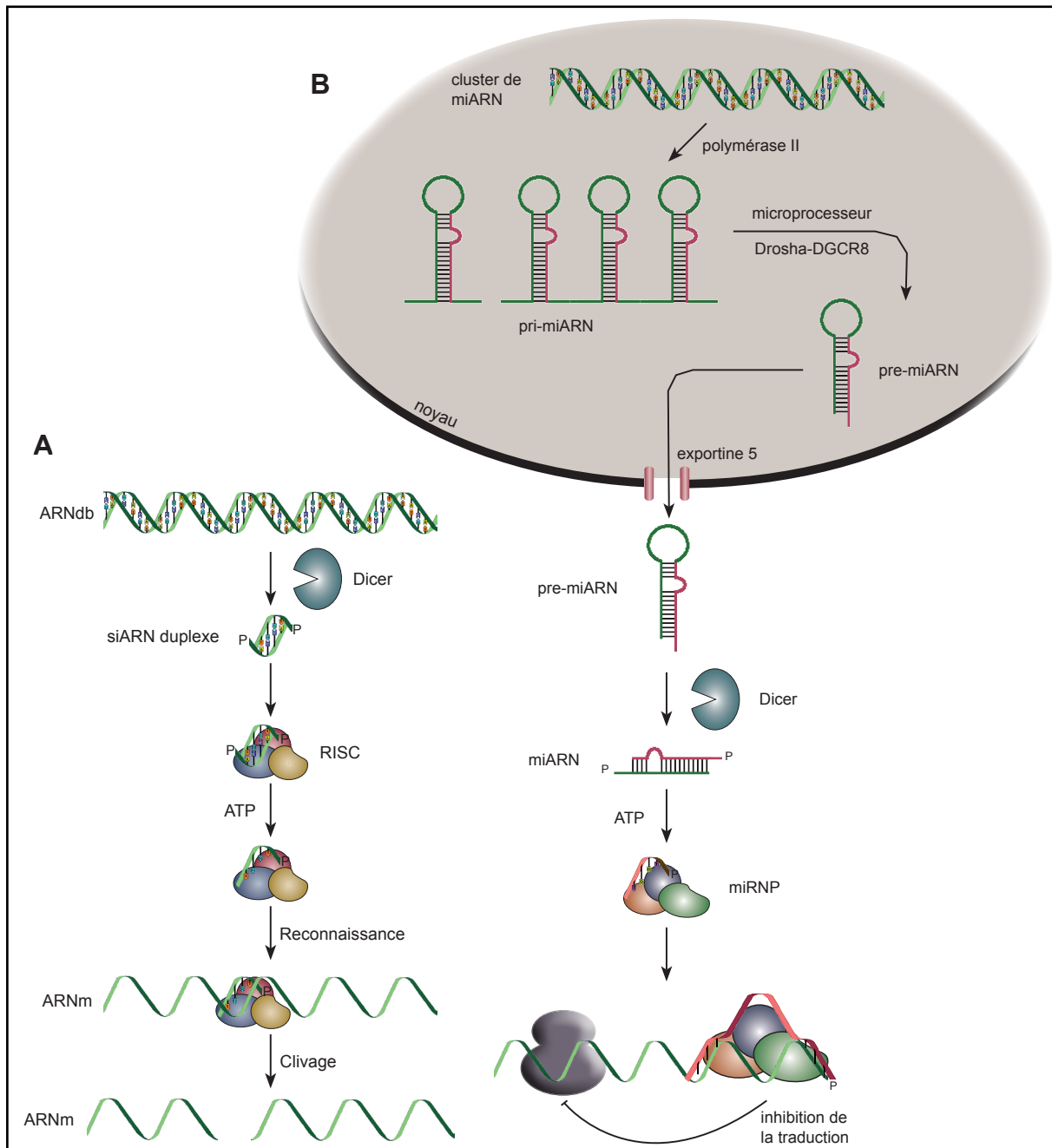
<sup>b</sup> Inferred from the phenotype of mutant or RNAi knock-down strains defective in at least one of the RNAi effectors (?), lack of evidence; NA, not applicable)

**Table 1 : Distribution des composantes du mécanisme d'interférence ARN chez les Eukaryotes (Table 1 **Cerutti 2006**)**



extrêmement conservée chez les vertébrés. Son activité requiert de l'ATP chez *C. elegans* et *D. melanogaster* mais pas chez l'Homme. [Ketting 2001](#), [Nykanen 2001](#), [Liu 2003](#), [Provost 2002](#), [Zhang 2002](#) Dicer est peut être aussi la protéine la plus importante de l'iARN du fait qu'elle est au coeur de ses trois principales voies :

- la voie des siARN (small interfering RNA, [Fig. 15.A.](#)) : les siARN (ARNdb de 21 à 23 nucléotides dont 2 sont protubérants à l'extrémité 3') sont issus du clivage cytoplasmique par Dicer de longs ARNdb provenant de la transcription bidirectionnelle de gènes, de transposons ou de virus à génome à ARNdb. [Elbashir 2001](#), [Bernstein 2001](#) Un des deux brins du siARN est ensuite incorporé dans le complexe RISC (RNA-induced silencing complex) pour servir de guide jusqu'à l'ARN messager de séquence homologue. Une fois l'appariement spécifique réalisé (100% de complémentarité dans les séquences), l'activité endonucléasique du complexe permet le clivage et donc la dégradation de l'ARNm. [Hamilton 1999](#), [Tuschl 1999](#), [Zamore 2000](#) Le complexe RISC est invariablement composé d'une protéine de la famille argonaute (Ago, sous-unité catalytique et liant l'ARN par son domaine PAZ) et d'un petit ARN guide. [Liu 2004](#), [Song 2004](#) Cependant, une grande variabilité existe pour tous les composants suivant la voie et l'espèce [Nykanen 2001](#), [Hammond 2001](#), [Caudy 2002](#), [Martinez 2002](#) : nature de l'ARN (siARN, miARN, etc.), argonaute (*a priori* uniquement Ago2 chez l'Homme car la seule à avoir une activité catalytique sur les 4 formes connues) [Liu 2004](#), [Meister 2004](#), nature et nombre des autres sous-unités (TRBP chez l'Homme). [Gregory 2005](#)
- la voie des microARN ([Fig. 15.B.](#)) : ce qui caractérise ces petits ARN est que leur structure double-brin provient du repliement en tête d'épingle d'un long transcrit initial. Il s'agit alors d'un microARN primaire (pri-miARN). Cette première structure est ensuite raccourcie (~70 nt) dans le noyau par le "microprocesseur" pour donner un microARN précurseur (pré-miARN). [Denli 2004](#) Dans ce complexe protéique, on retrouve une RNase III de classe II, Drosha, qui permet le clivage et son partenaire DGCR8 (DiGeorge syndrome chromosomal region 8) qui permet la liaison à l'ARN.



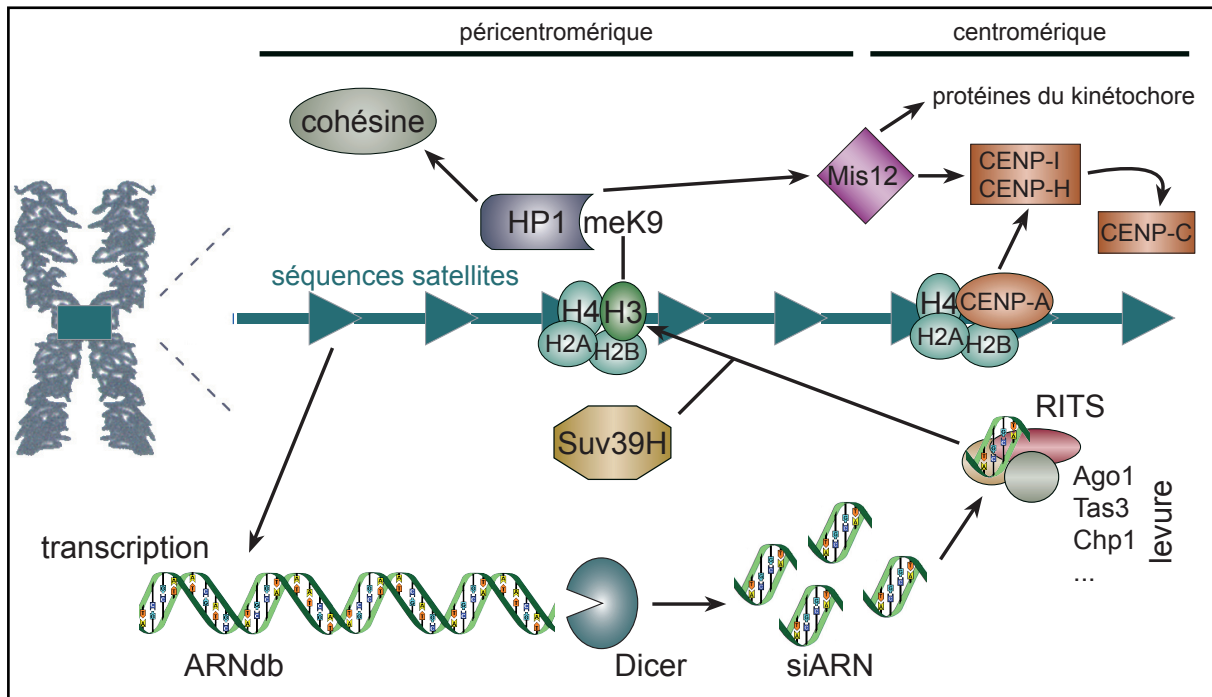
**Figure 15 : Principales voies de l'interférence ARN chez l'Homme** (d'après Dykxhoorn 2003)

**A.** La voie d'interférence ARN médiée par les siARN. Les ARNdb à l'origine de cette voie peuvent être exogènes (virus) ou endogènes (transcription bidirectionnelle de gènes ou de transposon). Ces molécules sont reconnues par la RNase III Dicer qui les clive alors en petits duplexes d'ARN de 21 nt avec 2 nt protubérants en 3' (siARN). Le siARNdb est séparé d'une manière ATP-dépendante. La différence de stabilité thermodynamique dans l'appariement des deux brins du siARN à chacune des extrémités déterminera le brin incorporé dans le RISC (RNA-induced silencing complex). Le complexe ainsi formé est capable de s'apparier spécifiquement avec l'ARNm qui partage la séquence complémentaire. Cette étape entraîne le clivage de l'ARNm donc sa déstabilisation et par la suite sa dégradation.

**B.** La voie d'interférence ARN médiée par les miARN. Cette voie débute par la transcription de cluster de miARN en pri-miARN (miARN primaires). Il s'agit à cette étape de longs ARN simple-brin qui vont se replier sur eux mêmes en formant des boucles. Ces structures sont reconnues dans le noyau par le "microprocesseur" qui contient entre autres la RNase III Drosha et son partenaire DGCR8. Ce complexe permet le clivage des pri-miARN en pre-miARN (miARN précurseurs) qui ne font alors plus qu'une trentaine de nucléotides de long. Les pre-miARN ont toujours cette structure caractéristique en épingle à cheveux et sont exportés dans le cytoplasme via l'exportine 5. Dès lors, ils sont reconnus par Dicer qui les clive en miARN matures c'est à dire en petits ARNdb aux extrémités 3' protubérantes. La particularité ici est que les deux brins ne sont pas totalement complémentaires. Comme pour les siARN, un seul brin est conservé et sert de guide au complexe RISC jusqu'à l'ARNm dont la séquence correspond partiellement. Une fois cette reconnaissance effectuée, il y a répression de la traduction.

Lee 2003, Han 2004, Gregory 2004 Le pré-miARN est alors exporté dans le cytoplasme, via l'Exportine 5 et l'hydrolyse de la Ran-GTP, où il est clivé en miARN mature par Dicer. Bohnsack 2004, Lund 2004, Yi 2003 Comme pour les siARN, un brin des miARN est par la suite chargé dans le complexe RISC et sert de guide jusqu'à une séquence cette fois partiellement complémentaire. Martinez 2002 Cette reconnaissance partielle d'un ARNm permet la répression de sa traduction. Grishok 2001

- les voies "épigénétiques" des siARN (Fig. 16) : d'une part, lorsque les ARNdb sont issus d'une transcription des séquences répétées de l'ADN centromérique, les siARN résultant du clivage par Dicer ont un autre effet. Ainsi, T.A. Volpe et coll. montrent pour la première fois (par délétion des gènes du mécanisme d'iARN) en 2002 que l'iARN joue un rôle dans la méthylation de l'histone H3 en lysine 9 (meH3K9) chez *S. pombe*. Volpe 2002 Puis, en 2004, plusieurs équipes proposent des modèles de ce processus (chez la Levure mais aussi chez l'Homme) faisant intervenir une variante du complexe RISC, le complexe nucléaire RITS (RNA-induced initiation of transcriptional gene silencing). Noma 2004, Schramke 2004, Fukagawa 2004, White 2004 Au contraire de chez l'humain, ce complexe est assez bien décrit et caractérisé chez *S. pombe* avec 3 sous-unités : Ago1 qui se lie au siARN, Chp1 qui se lie à meH3K9 et Tas3 qui lie Ago1 à Chp1. Sadaie 2004, Verdel 2004 Cette nouvelle voie d'interférence ARN serait donc à l'origine de la formation et/ou au maintien de l'hétérochromatine centromérique et péri-centromérique, primordiale pour la bonne séparation des chromosomes lors de la division cellulaire. D'autre part, lorsque les ARNdb sont issus de la transcription de régions promotrices, le mécanisme est très proche mais les conséquences diffèrent. En effet, la méthylation H3K9 induit dans ce cas l'inhibition de l'expression des séquences codantes associées au promoteur. Plus généralement, il a été montré que des cellules HCT-116 (colorectales) exprimant une protéine Dicer privée de l'exon 5 (domaine hélicase) donc inactive, présentent une déficience globale de la méthylation ; c'est à dire aussi bien la disparition presque totale de la méthylation génomique des cytosines que l'hyperméthylation



**Figure 16 : Formation de l'hétérochromatine médiée par l'interférence ARN** (d'après [White 2004](#), [Vos 2006](#))

Il a été montré au niveau des séquences répétées péricentromériques de l'ADN, autrement appelées séquences satellites, une activité de transcription bidirectionnelle. [Fukagawa 2004](#) Les ARNdb ainsi formés sont reconnus par Dicer et clivés en siARN. Ces derniers sont alors pris en charge par le complexe RITS (RNA-induced transcriptional silencing) et servent de guide jusqu'à leur locus d'origine. Le complexe recrute alors la méthyl-transférase Suv39h qui permet la méthylation des lysines 9 des histones H3 (meH3K9) avoisinantes. Cette marque de l'hétérochromatine induit le recrutement des protéines impliquées dans sa formation telles que HP1 (Hétérochromatine Protéine 1) puis le complexe cohésine par l'intermédiaire de sa sous-unité Rad21/Scc1. Chez l'Homme, et plus généralement les Vertébrés, les détails de ce mécanisme restent très flous alors que chez la Levure de nombreuses études existent. HP1 va aussi intervenir dans le recrutement de Mis12 très importante pour la mise en place d'un grand nombre de protéine du kinétochore intervenant dans le point de contrôle mitotique et la ségrégation des chromosomes. De même que Mis12, CENP-A, qui remplace l'histone H3 au sein du nucléosome et marque ainsi la spécificité de la région centromérique, permet le recrutement direct ou indirect de nombreuses protéines du kinétochore telles que le complexe CENP-I/CENP-H et CENP-C.

des îlots CpG ou encore la ré-expression de gènes déméthylés. [Ting 2008](#)

Il est à noter aussi un point très important qui différencie les mécanismes d'interférence ARN de certaines espèces, à savoir la présence de la protéine RdRP (RNA-directed RNA polymerase, [Table 1](#)). Cette enzyme est une sous-unité d'un complexe appelé RDRC (RNA-directed RNA Complex) capable de rétro-transcrire un brin ARN complémentaire à partir de n'importe quel ARN disponible. Ainsi, chez les organismes possédant ce complexe, il existe une boucle d'amplification de l'iARN qui participe grandement à l'intensité des réponses impliquant le mécanisme mais aussi à leur maintien dans le temps. [Dalmay 2000](#), [Sijen 2001](#)

I.4.3. Implication de *dicer* dans des processus biologiques fondamentaux

**E**n 2003, Bernstein et coll. observent que les souris knockout pour *dicer* ne survivent pas au-delà de 7,5 jours d'embryogénèse. [Bernstein 2003](#) Ce phénotype très délétère confirme alors ce qui a été mis en évidence chez les plantes et *C. elegans*, à savoir que *dicer* (on sait maintenant que la perte des miARN est la cause principale des effets observés) joue un rôle très important dans nombre de fonctions biologiques de la cellule souche à la différenciation. Plusieurs études ont été réalisées pour connaître les implications du gène dans différents contextes. Ainsi, l'angiogénèse embryonnaire et post-natale est limitée chez la souris lorsque *dicer* ne s'exprime plus [Yang 2005, Suarez 2008](#), la différenciation des lymphocytes T  $\alpha\beta$  est perturbée [Cobb 2005](#) et le nombre de lymphocytes T régulateurs diminue [Cobb 2006](#). Il a aussi été montré que Dicer est essentiel dans la morphogénèse de la peau [Yi 2006](#), des poumons [Harris 2006](#), des membres [Harfe 2005](#) et des follicules pileux. [Andl 2006](#) Enfin, plusieurs études se sont intéressées aux rôles de ce gène au niveau neurologique et il a été notamment mis en évidence son implication dans la survie de certains types neuronaux. [Kim 2007](#)

Par ailleurs, de nombreux miARN ont été associés aux cellules souches pour la maintenance de leurs caractéristiques et de leurs fonctions. Ainsi, des ESC (cellules souches embryonnaires) de souris privées de miARN ne sont plus capables de se différencier [Kanellopoulou 2005](#) et leur vitesse de prolifération est diminuée. [Murchison 2005](#) Plus précisément, une étude a montré que miR-302, sous contrôle des facteurs de transcription Oct4 et Sox intervenant dans l'embryogénèse et la maintenance de l'identité des ESC chez l'Homme, régulerait la cycline D1 dans ces cellules. L'inhibition de ce miARN conduit à une augmentation de l'expression de cette protéine et à pour conséquence une accumulation en phase G1 des cellules. [Card 2008](#) Par ailleurs, des travaux ont mis en évidence qu'il existerait des miARN spécifiquement exprimés, sous forme de cluster, dans les ESC humaines et murines. [Suh 2004](#) Il a aussi été montré que *dicer* serait un élément prépondérant dans le développement des cellules germinales

femelles murines. Ainsi, les oocytes n'exprimant pas ce gène ne peuvent pas terminer la méiose I du fait de défauts dans l'organisation du fuseau et dans la compaction des chromosomes. De plus, l'absence de *dicer* semble induire une forte augmentation d'éléments transposables (SINEs notamment) qui viennent alors perturber la méiose en dérégulant l'expression de nombreux transcrits. [Murchison 2007](#)

Un dernier exemple important du rôle de l'iARN dans des processus biologiques fondamentaux est la sénescence. En effet, une publication récente implique la perte de *dicer* et donc des miARN dans l'augmentation des taux de p53 et p19<sup>Arf</sup> de fibroblastes embryonnaires humains conduisant à une sénescence prématurée. [Mudhasani 2008](#) On observe donc à travers ces exemples l'importance du mécanisme d'interférence ARN dans de nombreux contextes fondamentaux dans le bon développement des organismes. Une conséquence inéluctable de cette implication dans divers mécanismes cellulaires est qu'une dérégulation de l'expression de *dicer* puisse conduire à l'apparition de maladies et plus particulièrement de cancers.

#### I.4.4. Interférence ARN et cancers

Les tumeurs solides et particulièrement les cancers du sein présentent une instabilité chromosomique forte. Cette signature est le résultat d'événements génétiques et épigénétiques sélectionnés au cours de la progression tumorale. Les dérégulations à l'origine de cette caractéristique sont cependant encore peu connues. De part son importance dans le développement [Bernstein 2003, Murchison 2005, Harris 2006](#) et son rôle central dans l'interférence ARN, il apparaît concevable d'impliquer *dicer* dans le développement de certains cancers. Ainsi, Chiosea et coll. ont montré en 2006 une sur-expression de ce gène, en analyse transcriptomique et en TMA (Tissue Micro Array), dans les cancers de la prostate. [Chiosea 2006](#) Un an après, la même équipe publie là aussi une sur-expression de *dicer* dans des lésions antérieures au cancer

du poumon et une sous-expression dans les stades les plus avancés. [Chiosea 2007](#) Les adénocarcinomes présentant quant à eux une diminution de l'expression du gène voir une perte due à la délétion du locus de *dicer*. Enfin, en 2008, Flavin et Smyth rapportent une dérégulation de *dicer* par PCR quantitative et TMA dans des cancers de l'ovaire. [Flavin 2008](#) Ils observent une sur-expression de l'enzyme dans les tissus tumoraux par rapport au tissu normal ainsi que des corrélations entre une forte expression de Dicer avec d'une part l'absence de nodules lymphatiques et d'autre part un taux de Ki-67 (marqueur de la division cellulaire) inférieur à 50%. Par ailleurs, le gène est localisé en 14q32.13 qui a été montré comme une région soumise à une certaine instabilité dans plusieurs types de cancer. [Ihara 2002](#), [Shao 2002](#), [Tzai 2003](#) Principalement, la perte d'hétérozygotie de ce locus est la plus fréquente dans le cancer des poumons [Wong 2002](#) et la région 14q43 toute proche est souvent délétée dans les cancers de l'ovaire. [Bandera 1997](#)

Beaucoup de travaux portent sur l'impact de l'expression des miARN dans les cancers, dont certains sont d'ores et déjà qualifiés d'oncogènes ou de suppresseurs de tumeurs (Table 2). [Garzon 2006](#) On peut citer, dans le premier cas, l'exemple de miR21 dans les cancers du sein dont la sur-expression est reliée à la tumorigenèse. [Si 2007](#) Ce lien proviendrait de ses activités répressives sur les gènes *Tpm1* (tropomyosine-1) [Zhu](#)

MicroRNA	Genomic location	Expression in patients	Experimental data	Function	Refs
<i>MiR-15a</i> <i>MiR-16-1</i> <i>Let-7a-2</i>	13q14 11q24	Down in CLL and in pituitary adenoma Down in lung cancer	Downmodulate <i>BCL-2</i> Downmodulates <i>RAS</i> and induces apoptosis in lung cancer cell lines	TS TS	[9,51,68] [37,38,53]
<i>MiR-155</i>	21q21	Up in bad prognosis CLL and lung cancer Up in breast cancer, lymphoma Hodgkin and pediatric BL.	Induces pre-B lymphoma and/or leukemia in mice	OG	[38,39,46-48,60]
<i>MiR-17-92</i> cluster <i>MiR-21</i>	13q14 17q23	Up in lymphomas and lung cancer Up in pancreas, glioblastoma and breast cancer	Cooperates with <i>c-MYC</i> ; modulates <i>E2F1</i> Anti-apoptotic in glioblastoma	OG OG	[38,55-57] [39-41,62]
<i>MiR-106a</i>	xq26	Up in lung, gastric and prostate cancer	Downregulates <i>RB-1</i>	OG	[41]
<i>MiR-372</i> <i>MiR-373</i> <i>MiR-142</i>	19q13 17q22	Up in testicular germ-cell tumor cell lines t(8;17) <i>c-MYC</i> is translocated downstream of the <i>miR-142</i> hairpin, resulting in B-lymphoma	Neutralize <i>p53</i> function Enhances <i>MYC</i> expression	OG OG	[61] [15]

<sup>a</sup>BCL-2, B-cell CLL/lymphoma 2; BL, Burkitt lymphoma; CLL, chronic lymphocytic leukaemia; OG: oncogene; RAS, rat sarcoma viral oncogene homolog; RB-1, retinoblastoma 1; TS, tumor suppressor.

**Table 2 : Données expérimentales démontrant un rôle des miARN dans les cancers** (Table 2 [Garzon 2006](#))

2007 et PDCD4 (Programmed Cell Death 4). Dans le deuxième cas, il a été observé, chez la souris, que plusieurs clusters de miARN se situaient au niveau de zones fragiles, correspondant à des loci perdus ou ré-arrangés dans les cancers. [Sevignani 2007](#)

Un autre exemple, dans les cancers du sein, est celui de miR-17-5p qui a été retrouvé sous-exprimé ce qui l'empêcherait d'inhiber l'expression de AIB1 (amplified in breast cancer 1) et donc d'augmenter la prolifération. [Hossain 2006](#)

Enfin, le miARN let-7 est connu comme suppresseur de tumeur dans de nombreux cancers dont ceux du sein par son action répressive de gènes tels que H-RAS et HMGA2. [Johnson 2005, Mayr 2007, Yu 2007](#)

Cependant, toutes ces études ne mettent en avant que les miARN ou les effets dûs au seul rôle de *dicer* dans la régulation des miARN. Elle font références pour cela à des études antérieures ou à des résultats qui montrent des variations des taux de certains miARN dans le même contexte. S'il est évident que cette voie du mécanisme d'interférence ARN influe fortement sur la progression tumorale lorsqu'elle est dérégulée, il nous a aussi semblé important de s'intéresser aux conséquences d'une telle dérégulation sur la stabilité chromosomique et en particulier la signalisation des cassures ADN. En effet, particulièrement chez la levure mais aussi chez l'Homme, certaines équipes ont montré des conséquences notables sur la division cellulaire lorsque *dicer* est dérégulé. Ainsi, chez *S. pombe*, la région centromérique des chromosomes perd sa structure hétérochromatique et les complexes cohésine sont délocalisés induisant des défauts de ségrégation et des retards de migration lors de l'anaphase. [Provost 2002, Hall 2003, Volpe 2003](#)

Dans une construction hybride où le chromosome 21 humain est intégré dans des cellules de poulet DT40, Fukagawa et coll. observent une mortalité très élevée avec une accumulation de cellules mitotiques anormales. Les cellules souffrent d'une séparation précoce des chromatides soeurs en lien avec une répartition totalement anormale du complexe cohésine car diffuse et non plus centromérique lors de la mitose. L'apparition en abondance de transcrits des séquences alpha-satellites humaines du chromosome 21 montre que le centromère



n'est plus sous forme hétérochromatique lorsque *dicer* ne s'exprime plus. [Fukagawa](#)

2004

Les recherches actuelles se focalisent beaucoup sur les miARN et leurs fonctions de régulation de l'expression génétique (sous PubMed, le mot clef "dicer" renvoi à 589 articles alors que "drosha" en compte 1327) mais semblent oublier cette autre fonction primordiale de l'iARN qui est une partie de l'organisation des chromosomes. A partir des rares publications disponibles, nous avons donc voulu mettre en évidence, en cellules humaines, l'impact de *dicer* sur la stabilité chromosomique.

## Objectifs

**A**u démarrage du projet, l'interférence ARN était encore un mécanisme assez confidentiel avec très peu de publications sur l'implication de ce processus dans différentes fonctions biologiques. Cependant, le principe de régulation de l'expression des gènes par le voie des siARN était connu et nous a fortement interpellé par le rôle potentiellement important qu'il pouvait jouer dans les cancers (particulièrement ceux du sein mais aussi ceux du côlon) s'il était dérégulé. Après une année de recherche dans le cadre de mon DEA (étude par PCR de l'expression des gènes *dicer*, *ago1* et *ago2* ainsi que test fonctionnel de l'activité du complexe RISC dans des lignées tumorales mammaires humaines et un modèle de progression tumorale mammaires humaine). Ces premiers résultats nous ont montré que l'enzyme Dicer semblait être impliquée dans la carcinogénèse et nous avons donc décidé de focaliser nos travaux sur cette protéine clef du mécanisme d'iARN. Nous avons alors utilisé des modèles de progression tumorale afin d'étudier l'expression de notre gène d'intérêt à différents stades de la tumorigénèse. Comme les résultats étaient encourageants, nous avons décidé d'analyser cette expression dans des échantillons tumoraux. L'objectif de cette étude était double : tout d'abord savoir si ce gène était réellement pertinent à étudier dans le contexte du cancer du sein mais aussi déterminer s'il avait un intérêt clinique.

En parallèle, les publications concernant le mécanisme d'interférence ARN ont fortement augmenté mais presque uniquement comme outil d'étude (caractérisation des rôles de gènes inactivés par iARN) et par la suite pour traiter de l'expression des miARN. Les articles portant sur le mécanisme en lui-même restant, encore aujourd'hui, assez rares chez les mammifères. Ainsi, la principale source d'information provient de la levure *S. pombe* où les découvertes originales se sont succédées au fil des années. C'est ainsi que nous avons décidés de nous intéresser au rôle de Dicer sur la mise en place de l'hétérochromatine péricentromérique et par extension sur la stabilité

chromosomique. Nous nous sommes donc demandé si *dicer* était régulé au cours du cycle cellulaire et nous avons testé les effets d'une inactivation de ce gène dans ce contexte. Nous avons également émis l'hypothèse d'une implication de *dicer* dans la signalisation des cassures de l'ADN et d'étudier l'effet de son inactivation après stress. L'objectif étant ici de savoir si au niveau moléculaire et mécanistique, *dicer* est un gène pertinent à étudier dans le cadre de la progression tumorale mammaire humaine.

## II – Résultats

### II.1. Recherche de facteurs pronostiques/diagnostiques des cancers du sein

**E**n oncologie mammaire, les applications cliniques ayant trait aux altérations génétiques somatiques sont actuellement très limitées. En effet, seuls quatre marqueurs biologiques sont utilisés en routine : ER $\alpha$ , prédictif de la réponse aux traitements par les anti-oestrogènes; l'oncogène ERBB2 pour la réponse au Trastuzumab; les marqueurs tumoraux sériques CA-15-3 et ACE pour le suivi du traitement et la détection des rechutes. Les facteurs pronostiques cliniques et histologiques se révèlent insuffisants pour prévoir avec exactitude l'évolution des cancers du sein. Il apparaît que de plus en plus de cancers sans envahissement ganglionnaire rechutent dans les dix ans qui suivent le diagnostic. Par contre, la fiabilité des prédictions de nouveaux marqueurs pronostiques moléculaires doit être recherchée. C'est dans cette perspective que nous avons réalisé deux études descriptives dans les cancers du sein : une concerne l'expression du gène *dicer* et l'autre celle de certains gènes du complexe APC/C (Anaphase Promoting Complex / Cyclosome). Cette dernière a aussi été menée dans des cancers du côlon car nous avons identifié des altérations de ce complexe dans des lignées d'adénocarcinomes coliques lors de travaux précédents.

#### II.1.1. L'analyse d'éléments clés du complexe APC/C dans les cancers du sein et du côlon révèle que APC-11 est un marqueur d'agressivité

##### i. Résumé

**L'**intérêt de cette étude sur l'expression de gènes du complexe APC/C est de compléter les premiers résultats obtenus par *Wang et col.* en 2003. Ces travaux décrivaient la présence de mutations dans des gènes codant pour des sous-unités du complexe avec une implication fonctionnelle pour celles concernant

CDC-23. Le complexe APC/C est connu pour son rôle primordial dans la régulation du cycle cellulaire et notamment la mitose. Cette E3 ubiquitine ligase composée de plus de dix sous-unités s'appuie sur deux protéines co-activatrices majeures, CDH1 et CDC-20, pour cibler la destruction séquentielle de nombreux acteurs de la division (Cyclines A et B, Sécurine, Plk1, Aurora A...) et permettre ainsi son bon déroulement. C'est pourquoi une dérégulation du complexe, du fait de mutations ou de sur-/sous-expression de sous-unités, peut induire des instabilités chromosomiques et contribuer à la progression tumorale. Par ailleurs, il a été observé plus récemment que APC/C jouerait aussi un rôle dans la réponse aux dommages à l'ADN en reconnaissant, via sa sous-unité CDC-27, la protéine MDC1 qui se lie à  $\gamma$ H2A.X. Cependant, peu d'études ont recherché une implication de ce complexe dans la tumorigenèse : une surexpression de CDC-20 est retrouvée dans les cancers du cou et de la tête, CDH1 est sousexprimée dans certains lymphomes et CDC-23 présente des mutations dans un nombre important de lignées issues de cancers du côlon. En nous intéressant à l'expression de six sous-unités dans des tumeurs et lignées cellulaires mammaires et colorectales, nous avons recherché de nouveaux marqueurs pronostiques et/ou diagnostiques. Le choix de ces sous-unités s'est fait pour leur rôle au sein du complexe. APC-11 porte l'activité catalytique, CDC-16, CDC-23 et CDC-27 possèdent toutes un motif TPR (Tetratricopeptide Repeat) cible de mutations et semblent apporter une spécificité de substrat, APC-5 et APC-7 régulent la transcription de p53 par l'intermédiaire du complexe CBP/p300.

Les résultats statistiques (utilisant un modèle original adapté à la démarche expérimentale choisie) concernant l'expression transcriptionnelle par RT-PCR quantitative de ces six gènes dans les lignées colorectales ont été significatifs pour APC-11. En effet, nous avons pu montrer une corrélation entre une surexpression d'APC-11 et un génotype aneuploïde. Si cette observation reste à confirmer par une étude fonctionnelle, elle est en tout cas cohérente avec les données publiées. Nous avons observé dans les lignées tumorales mammaires humaines une surexpression des gènes APC-

11 et CDC-23 en lien avec l'âge (supérieur à 50 ans) des patientes qui sont à l'origine des lignées. Ces deux mêmes gènes ainsi que APC-7 sont aussi surexprimés dans les lignées capables de former des tumeurs dans les souris *nude*. Enfin, les résultats obtenus avec des tumeurs primaires de cancers du sein montrent une expression accrue des sous-unités APC-7, CDC-16 et CDC-23 lorsque celles-ci forment des métastases. De plus, APC-11 est plus fortement exprimé dans les cancers de grade de SBR III que ceux de grade I. Ainsi, nous avons pu montrer des associations entre des surexpressions de sous-unités clefs du complexe APC/C et des paramètres cliniques de mauvais pronostiques. Le gène APC-11 ressort tout particulièrement de cette étude. On peut penser que la fonction catalytique de la protéine correspondante en fait une cible privilégiée lors de la tumorigenèse. Ces résultats permettent d'envisager une utilisation de drogues thérapeutiques plus précises que les actuels inhibiteurs du protéasome.

ii. Manuscrit

### **Analysis of APC/C key elements in breast and colon cancer reveals APC11 as a cancer-aggressiveness marker**

Gaël Grelier<sup>1,2,3</sup>, Nicolas Voirin<sup>4,5</sup>, Sylvie Chabaud<sup>6</sup>, Isabelle Treilleux<sup>7</sup>, Sophie Léon-Goddard<sup>7</sup>, Ruth Rimokh<sup>2,3</sup>, Alain Puisieux<sup>1,2,3</sup>, Christine Lasset<sup>4,6\*</sup> and Caroline Moyret-Lalle<sup>1,2,3\*</sup>

#### **Authors' Affiliations:**

<sup>1</sup>Université de Lyon, université Lyon 1, ISPB, Lyon, F-69003, France;

<sup>2</sup>Inserm, U590, Lyon, F-69008, France;

<sup>3</sup>Centre Léon Bérard, Lyon, F-69008, France;

<sup>4</sup>Université de Lyon, université Lyon 1, Faculté Grange Blanche, CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, Lyon, F-69373, France;

<sup>5</sup>Hospices Civils de Lyon, Hôpital Edouard Herriot, Service d'Hygiène, Epidémiologie et Prévention, Lyon, F-69437, France;

<sup>6</sup>Centre Léon Bérard, Département de Santé Publique, Lyon, F-69008, France;

<sup>7</sup>Centre Léon Bérard, Service d'Anatomopathologie, Lyon, F-69008, France.

**\*Author contributions:** C.L. and C.M.L. contributed equally to this work

**Grant support:** Ligue Nationale Contre le Cancer, Comité Départemental de la Drôme de la Ligue Nationale Contre le Cancer.

**Requests for reprints:** Caroline Moyret-Lalle, INSERM, U590, Centre Léon Bérard, Lyon, F-69008, France.

**Running title:** APC/C subunits expression in colon and breast cancers

## **Abstract**

*Purpose:* The anaphase Promoting complex is a pivotal E3 ligase with prominent role in cell cycle regulation. Deregulation of APC/C contributes to chromosomal instability and tumor progression. It was recently shown that APC/C participates in DNA damage response in mediating the DNA-damage-induced G2 cell cycle block. Deregulated APC/C subunits are increasingly evoked in cancer pathologies. So far, no studies have examined simultaneously the expression of several key subunits of the ligase complex in cancers. The expression of the catalytic subunit APC11, the tetratricopeptide repeat (TPR) motifs proteins CDC16, CDC23, CDC27, and APC/C-CBP/p300 interacting proteins APC5 and APC7 was assessed in breast and colon primary tumors and cell lines.

*Experimental Design:* In this study, the clinical significance of the APC11, CDC16, CDC23, CDC27, APC5 and APC7 expressions in patients with breast and colon primary tumors was investigated performing quantitative reverse transcription-PCR. The catalytic component of APC/C, APC11 was analysed by immunohistochemistry in both normal and cancer epithelial cells of breast tissues. Furthermore, we analysed the expression of APC11 at both mRNA and protein levels in colon and breast cancer cell lines and during the cell cycle.

*Results:* We found that APC11 expression is regulated during cell cycle with highest levels in G1 phase and a down regulation at the end of this phase, a second peak in S/G2 and a subsequent decrease at the end of the mitosis. Most notably, in colorectal cancer cell lines, the expression of APC11 was associated with chromosomal instability and in breast cancer cell lines APC11 expression is associated with tumorigenicity in nude mice. In breast primary tumours, we have found a significant association

between CDC16, CDC23, and APC7 RNA expression levels with presence of metastases. APC11 was found to be up-regulated in breast cancers of high clinical grade.

*Conclusions:* Assessment of APC/C subunits expression has a clinical significance in breast and colon cancers. APC11 appeared to be significantly associated with markers of aggressiveness in colon and breast cancer cells. Targeted therapies against APC11 would selectively stabilize its targets and have an advantage over proteasomal inhibition by avoiding important toxicity.

## **Introduction**

Chromosomal instability (CIN) is found in majority of solid cancers. It is largely admitted that missegregation of chromosomes results in aneuploidy and participates in tumour progression. Mitosis is under tight control to ensure proper chromosome segregation. Cell division including mitosis is governed by the degradation of different regulatory proteins by the ubiquitin-dependent proteolysis. The anaphase-promoting complex APC/C is a specific ubiquitin ligase essential for chromosome segregation, exit from mitosis and a stable subsequent G1 phase (Peters et al., 2006 for review). Deregulation of key cell cycle elements proteolysis might result in genomic instability and tumour progression (Keck et al., 2007; Lehman et al., 2007). Some proteins that are targeted for degradation by APC/C include the 'master regulator' of mitosis, mitotic cyclins and securin (involved in chromatid separation) were shown to be deregulated in cancer and abnormal expression of cyclin B1 or securin has been considered to be a major factor in polyploidy (Roh et al., 2005; Kim et al., 2007). Accurate substrate levels are very important for normal cell cycle oscillation; in cancer, blocked degradation and paradoxically lack of securin expression both leads to genomic instability in cultured cells (Jallepalli et al., 2001; Yu et al., 2000).

APC/C is a large multiprotein E3 ligase complex consisting of invariable core components- APC11 (RING-H2 finger protein), APC2 (CUL1-related scaffold protein)



and at least 11 other components with unknown functions, as well as a variable activator component, CDC20 and CDH1 (Nakayama et al., 2006). These activators confer substrate specificity. APC/C activity is repressed by the mitotic checkpoint complex (MCC) composed of at least four subunits with Mad2 and BubR1 inhibiting the activator subunit CDC20 (Sudakin et al., 2001). MCC orchestrates the spindle-assembly checkpoint by preventing premature advance to anaphase and is activated every cell cycle immediately upon entry into mitosis. Few years ago, mutations in BubR1 and Bub1 were reported in colon cancer and alterations in MCC genes were proposed to cause CIN and contributed to tumour progression (Cahill et al., 1998). However, this model seems today disappointing with number of reports that fail to show a correlation between these major mitotic players and cancer progression (Kops et al., 2005). Rather than driving oncogenesis, a weakened mitotic checkpoint might be a facilitator of tumorigenesis, especially in collaboration with a deregulated suppressor gene (Jelluma et al., 2008). It is assumed that CIN can occur from deregulated mitosis legitimates studies on systems involved in proteolysis of mitotic oscillators. Intriguingly, a recent work evoked the involvement of APC/C in DNA repair (Coster et al., 2007). In the ubiquitine ligase family, E1 has never been involved in tumour development and only a few reports linked an E2 enzyme to tumour development (Okamoto et al., 2003; Berlingieri et al., 2007; Hao et al., 2008). In contrast, a significant number of reports has shown deregulation of E3 ligase in cancer (for reviews: Bernassola et al., 2008; Newton and Vucic, 2007). Almost all implicated the other cell-cycle control E3 ligase complex, namely SCF (SKP1-CUL1-F-box-protein) with SKP2 as an oncogene (Nakayama et al., 2006). So far, very few reports have shown alterations of APC/C in cancers (Wang et al. 2003). The mid-mitosis APC/C activator subunit CDC20 is overexpressed in some human cancers, although some reports have shown deletions or mutations causing its inactivation (Mondal et al., 2007). CDH1 has been described with reduced levels in a B-lymphoma cell line (Wang et al., 2000). We have previously shown the presence of several heterozygote mutations of APC/C subunits genes in colon cancer cells

(Wang et al., 2003). Overexpression of a mutated CDC23 subunit leads to abnormal sustained cyclin B1 protein levels and blocked cell cycle progression. No mutations were found in other type of cancer. However, some APC/C components could relate to cancer through transcriptional deregulation. Hence, we analysed the transcriptional levels of key components of APC/C complex, in colon and breast primary tumours in comparison with those found in cell lines and correlated them with clinical and biological parameters and degree of aneuploidy present in the tissues. The APC/C subunits elected in this study were primarily chosen from our previous work reported mutations in TPR motif containing subunits (CDC16, CDC23 and CDC27) and their consequences on mitosis progression. The catalytic subunit APC11 seemed of particular interest because of its capacity alone or in an Apc2-Apc11 complex, to catalyse ubiquitylation reactions in vitro however with reduced substrate specificity, thus indicating that other APC/C subunits are not absolutely required for the E3 activity of APC/C. Finally we have included APC5 and APC7 subunits because they were recently implicated in p53-regulated transcription of p21 CIP/WAF1 in a CBP: p300-dependent manner, implying that these two subunits are potentially involved in tumorigenesis (Turnell et al., 2005). Statistical analysis was conducted with an innovative four level linear mixed model which take into account the structure of q-PCR experiments. Our results show a significant implication of APC11 subunit in cancer aggressiveness.

## **Materials and Methods**

### *Cell samples*

21 breast cancer cell lines were obtained from the American Type Culture Collection ([www.ATCC.org](http://www.ATCC.org)). Cells grown according to American Type Culture Collection protocols to 90% confluency and were collected for total RNA preparation. HMECs and HME-1 (Cambrex, Clonetics) were cultured in MEGM at 37°C with CO<sub>2</sub> as described (Stampfer, 1985). The MN-1 cell line, containing wt-p53 and the MDD2 cell line, deri-

ved from the human breast carcinoma cell line MCF-7, were generously provided by Moshe Oren (Weizmann Institute of Science, Israel). The MDD2 line is a variant derived from MCF-7 by transfection with a dominant negative mutant of p53 (pCMVDD-p53; mut-p53). This plasmid encodes a non-functional p53 miniprotein containing the first 11 residues and residues 302 to 390 of murine p53 (Shaulian et al, 1992). The MN-1 line is a control line transfected with the empty plasmid. 21 different colon cancer cell lines were included in this study. 12 cell lines were obtained from the American Type Culture Collection, EB were provided by PH. Shaw, and Co-115, Isreco1 (IS1), Isreco2 (IS2), Isreco3 (IS3), TC-7 and TC-71 were provided by R. Hamelin. The cell lines IS1, IS2 and IS3 were derived from the same patient, from a primary colon carcinoma, its corresponding liver-and peritoneal metastases (Gayet et al., 2001; Cajot et al. 1997). The other cell lines are derived from human primary colon carcinomas.

#### *Cell cycle synchronization*

Briefly, HCT116 cells were washed with DMEM (Dulbecco's modified Eagle's medium) medium and then cultured in 0.5% FCS DMEM culture conditions. After 48h, cells were released from G0/G1 block by adding 10% FCS and cells were collected at different times following addition of FCS.

#### *Clinical samples*

Following approval from the ethical and scientific committees of the participating institutions, frozen samples in the form of tissue biopsies were obtained from 117 patients with stage I/IV breast cancer and 9 patients with sporadic colorectal carcinoma, 3 different samples were obtained from the same patient, one from normal colon tissue, one from a polyp biopsy and the third one from colloid adenocarcinoma after surgical resection.

#### *Total RNA isolation and cDNA synthesis*

Total RNA from 2 normal breast and 2 normal colon adult tissues were obtained from BioChain Institute. Total RNA isolation was performed by using Tri-Reagent (Sigma Chemical Co.), largely free of contaminating DNA with additional purification

step by passage through Phase-Lock gel tube (Eppendorf). RNA concentration was determined by absorbance readings at 260 nm with Beckman UV-VIS spectrophotometer. RNA integrity was tested in 1% agarose gel MOPS 1X running buffer. cDNA was synthesized from 1 µg of total RNA with first-strand cDNA synthesis kit (Amersham Pharmacia Biotech) according to the manufacturer's instructions.

### *Real-Time PCR Analysis*

Total RNA levels of APC11, CDC16, CDC23, CDC27, APC5 and APC7 were assessed using two-step real-time PCR in a LightCycler (Roche Molecular Biochemicals). Specific primer pairs used in the PCR amplification were commercially available (Applied Biosystems). All reactions were performed in glass capillaries (Roche) using LightCycler FastStart DNA Master PLUS SYBR Green I kit. The reaction mixture contained FastStart TaqDNA polymerase, reaction buffer, deoxynucleoside triphosphate mixture, MgCl<sub>2</sub> include in kit, 0.25 µM of each primer, and 1,67 µl of a 1/60 dilution or first-strand cDNA specimen from cells or tissues. The final reaction volume was 6.67 µl. Water was used as a negative control. All samples were kept at 4°C during preparation. Thermocycling conditions were designed to program in four consecutive steps. Step 1: denaturation and polymerase activation by 10 min heating at 95°C prior amplification. Step 2 (45 cycles): denaturation for 10 sec at 95°C, annealing for 10 sec at 55°C and extension for 6 sec at 72°C. Step 3: melting which consisted of incrementing the temperature from 45°C to 95°C by steps of 0.1°C. Step 4: cooling for 30 sec at 40°C for one cycle. All standard or samples were analyzed in duplicate and repeated at least three times. All fluorescence data were analyzed by LightCycler 4.0 software (Roche) and the Ct results export to Excel sheets.

A large amount of cDNA of the HME-1 cell line was prepared prior to the experiment. It was then diluted plus aliquoted and served as the calibrator for all of the Q-PCR runs. For relative quantification and normalization, the comparative Ct (or  $E^{-\Delta\Delta C}$ , where E is the primers-dependent efficiency of the PCR) method was used. The Ct values of both the calibrator and the samples of interest were normalized to 3house-

keeping genes (PPIB,  $\beta$ -Actin and PGK).

### Statistical analysis

A log<sub>2</sub> transformation was applied to APC/C subunits RNA expression data since they did not initially meet normality requirements. Based on Kolmogorov-Smirnov test results, statistical significance for the normality hypothesis was p=0.002 and p=0.291 before and after transformation respectively. Based on the experimental design, Q-PCRs were nested within RTs, themselves nested within extractions, themselves nested within cell lines, suggesting the use of a hierarchical linear model (9). Consequently, the log<sub>2</sub> mean RNA expression for breast cancer cell lines was estimated using the following four-level linear mixed model (LMM):

$$\log_2 APC/C_{irel} = \mu + \alpha_l + \delta_{el} + \gamma_{rel} + \varepsilon_{irel}$$

where  $\mu$  was the overall mean of total RNA expression for APC/C subunit after log<sub>2</sub> transformation. For random effects,  $\alpha_l$  represented the between-cell lines variation,  $\delta_{el}$  the between-extractions or within-cell lines variation,  $\gamma_{rel}$  the between-RT or within-extraction variation and  $\varepsilon_{irel}$  the within-RT variation or measurement error. These random effects were assumed to be independent and normally distributed with zero mean and constant variance as follows:  $\alpha_l \sim N(0, \sigma_\alpha^2)$ ,  $\delta_{el} \sim N(0, \sigma_\delta^2)$ ,  $\gamma_{rel} \sim N(0, \sigma_\gamma^2)$  and  $\varepsilon_{irel} \sim N(0, \sigma_\varepsilon^2)$ .

Estimations were done using the maximum likelihood method using the nlme library of the software R (10). For likelihood ratio tests, the statistical significance level was 0.05.

### Western blotting and Antibodies

Cell pellets were disrupted with RIPA buffer (APC11 detection) for 30 min in ice and then centrifugated 30 min at 13 000 rpm at 4°C. The supernatant protein concentration was determined by a Bradford assay using Biorad Protein assay (Biorad). BSA was used as standard. 50  $\mu$ g of denatured and reduced proteins were load on 20% acrylamide-bisacrylamide gels for APC11 (9.8 kDa). Multimark multi-colored standard (Invitrogen) were used as molecular weight marker and Tris-Gly-SDS buffer 1X (Sig-

ma) for migration. Transfert was made in wet conditions with Tris-Gly 1X (Sigma), 10% methanol buffer during 1h. Then, the nitrocellulose membrane (0.2µm) was incubated 1h with TBS 1X-Tween 0.5%-milk 5%. APC11 antibodies (Abcam, ab18303) were diluted at 1/500. After 1h incubation, membranes were washed 3 times with TBS 1X-Tween 0.5%. The secondary antibodies (P0217, Dako) were diluted at 1/3000 in the same buffer and incubated 1h. APC11 protein expression was revealed by Lumi-LightPLUS (Roche).

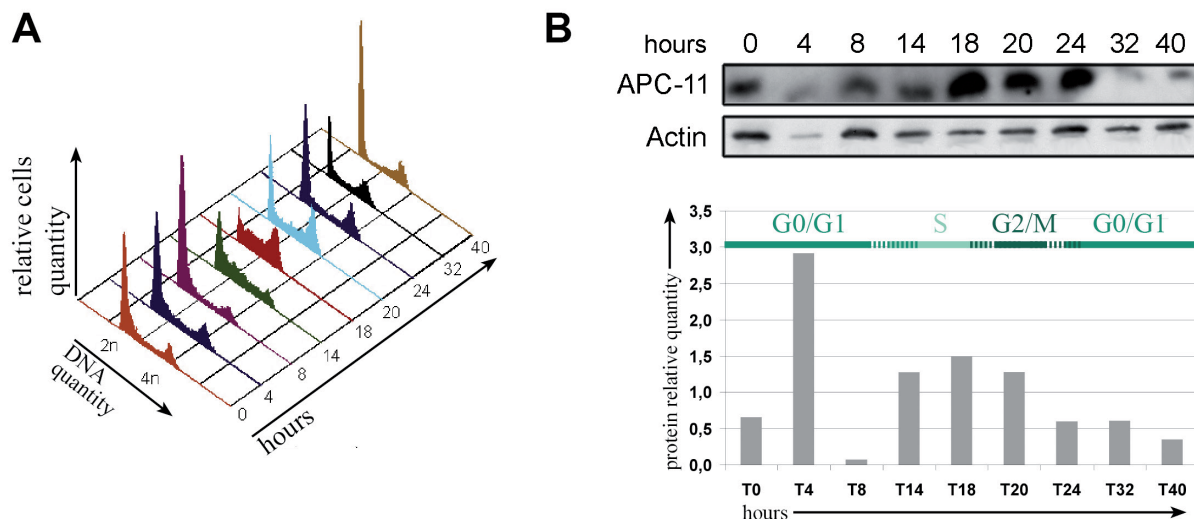
## **Results**

### *APC11 expression regulation during cell cycle*

First of all, we have analysed the expression of the key RING-H2 finger APC11 protein during cell cycle in the colon cancer cell line HCT116. Synchronization of HCT116 cells to arrest in G0/G1 by serum-deprivation was performed followed by serum-supplementation for cell cycle progression. Western blot analysis of the cell extracts prepared at regular intervals following release from serum starvation revealed oscillation of APC11 protein expression during the cell cycle with high expression in G1 phase and a dramatic decrease at the transition G1/S, APC11 levels increased in S and G2/M and decreased as cells reached the next cell cycle (fig. 1).

### *Evaluation of the biological and clinical significance of APC/C subunits expression in colon and breast cancers*

We have first analysed the expression of APC11, CDC16, CDC23, CDC27, APC5 and APC7 mRNAs in colorectal cancer cell lines. We have found a significant association between APC11 expression levels and ploidy and microsatellite status; a significantly higher APC11 expression was observed in aneuploid cell lines compared with diploid and near diploid cell lines (CIN:mean = 1.906 vs diploid:mean = 0.995, p-value = 0.0005) and MSI (Microsatellite Instable) cell lines exhibited lower levels of



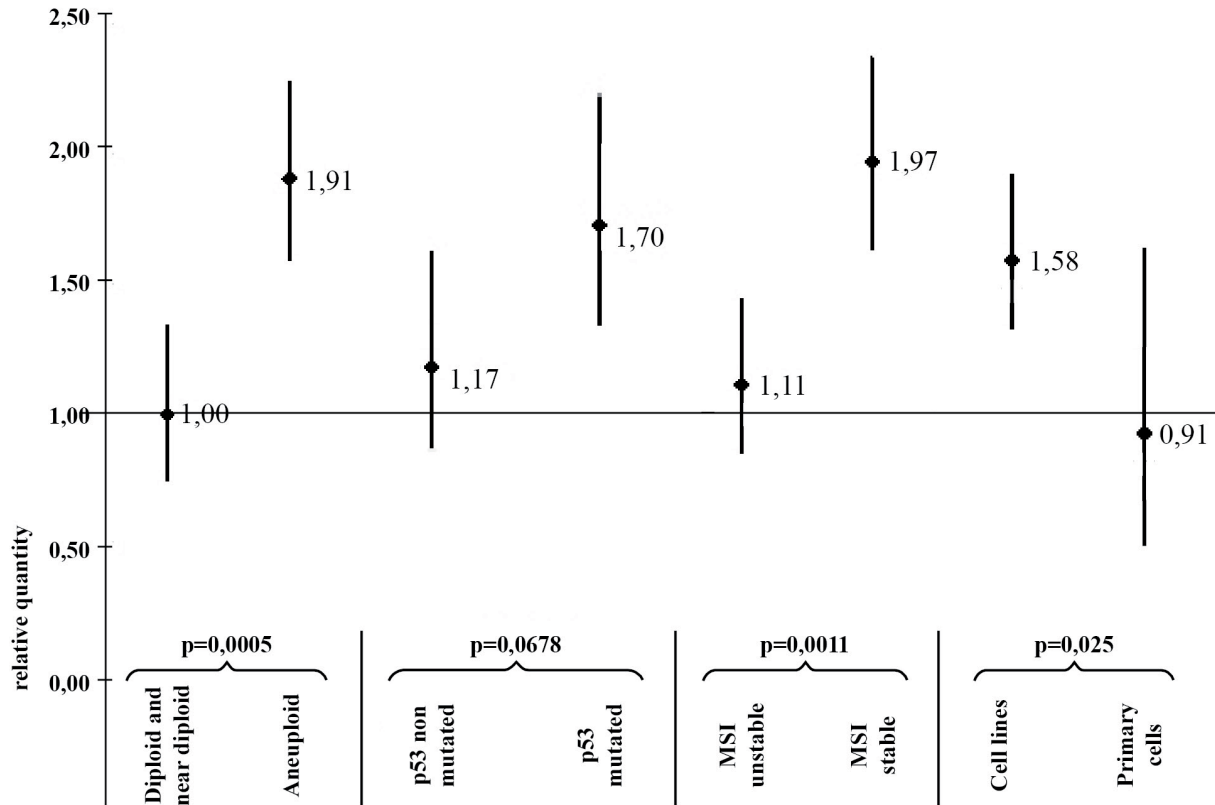
**Figure 1: Expression of APC-11 in HCT-116 cells during cell-cycle progression.**

**A.** Flow cytometry analysis (Propidium Iodide) of cell-cycle progression in HCT-116 cells after synchronization using serum deprivation method (G1 arrest). Cells were released into the cell cycle after addition of fresh medium and collected at indicated time points (hours).

**B.** Western blot analysis of APC-11 in HCT-116 cells and actin normalized relative quantification (Quatity One software, BioRad). The different cell-cycle phases noted on the graph are issued from the flow cytometry analysis (A.).

APC11 (fig. 2). We have previously found mutations in APC/C subunits only in colon cancer cell lines. The identified mutations affected APC4, CDC16, CDC23 and CDC27. Mutations were rarely found in primary colon carcinomas (Wang et al., 2003). Here we did not find a correlation between the presence of a mutation in a subunit and its expression levels except for CDC27 subunit where a tendency towards significance was found for higher levels of mRNA CDC27 in cells presenting a mutation in the coding sequence of CDC27 in comparison with levels observed in cells with type CDC27 sequence (WT CDC27: mean = 0.870 vs mt CDC27: mean = 1.973; p-value = 0.0604). In colon cancer biopsies, no clear association was found between APC/C subunits mRNA expression levels and tumour clinicopathological parameters except for type of cancer where a tendency towards significance was found for APC7 expression with lower levels in polyps compared with levels found in Lieberkühnian adenocarcinomas (polyps: mean= 0.542 vs Lieberkühnian: mean= 1.486, p-value = 0.0629).

APC11 protein levels were examined both in colon and breast cancer cell lines by western blot. Semi-quantitative measures of protein levels revealed 60% of concordance between RNA and protein levels in the two types of cancer (fig. 3).

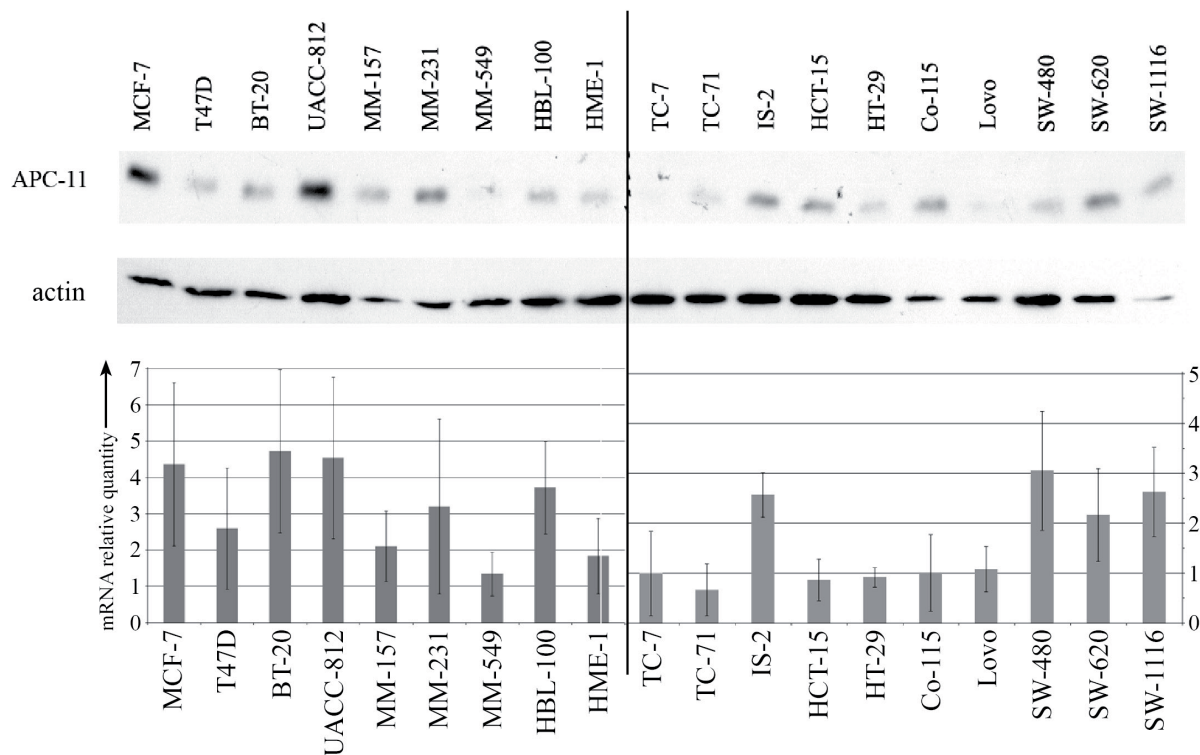


**Figure 2: Statistical correlations between APC-11 expression and molecular features in colorectal cancer cell lines**

Total RNA of each cell lines was extracted twice and the relative levels of APC-11 mRNA were measured by real-time RT-PCR (five independent RT were done). The results of the statistical analysis (see materials and methods) are displayed using a graphical representation. For each case, the number indicates the mean expression and the bars the confidence interval. Significance for the test is reach when  $p < 0,05$ . We retain four relevant correlations (three significant and one near significant):  $< p53$  mutation, MSI (microsatellite instability) and cell lines versus primary cells.

In breast cancer cell lines, the expression of APC11 and CDC23 appeared to be correlated with age, as in patients older than 50 years, the expression of the two subunits was higher in comparison to levels found in younger patients (fig. 4A). We found significant association between the levels of APC11, CDC23 and APC7 and tumorigenicity in nude mice. Interestingly, the levels of these three subunits were all higher in breast cancer cell lines able to establish a tumour in nude mice (fig. 4B). In multivariate analysis, the expression of APC11, CDC23 and CDC27 were increased respectively by 1.551 (1.022-2.353), 1.860 (1.313-2.634) and 1.442 (0.971-2.143) for cell lines established from tumours of patients older than 50 years and these effects were significant for APC11 and CDC23 (APC11 p-value = 0.0431 and CDC23 p-value = 0.0031) and presented a tendency towards significance for CDC27 (p-value = 0.0705).

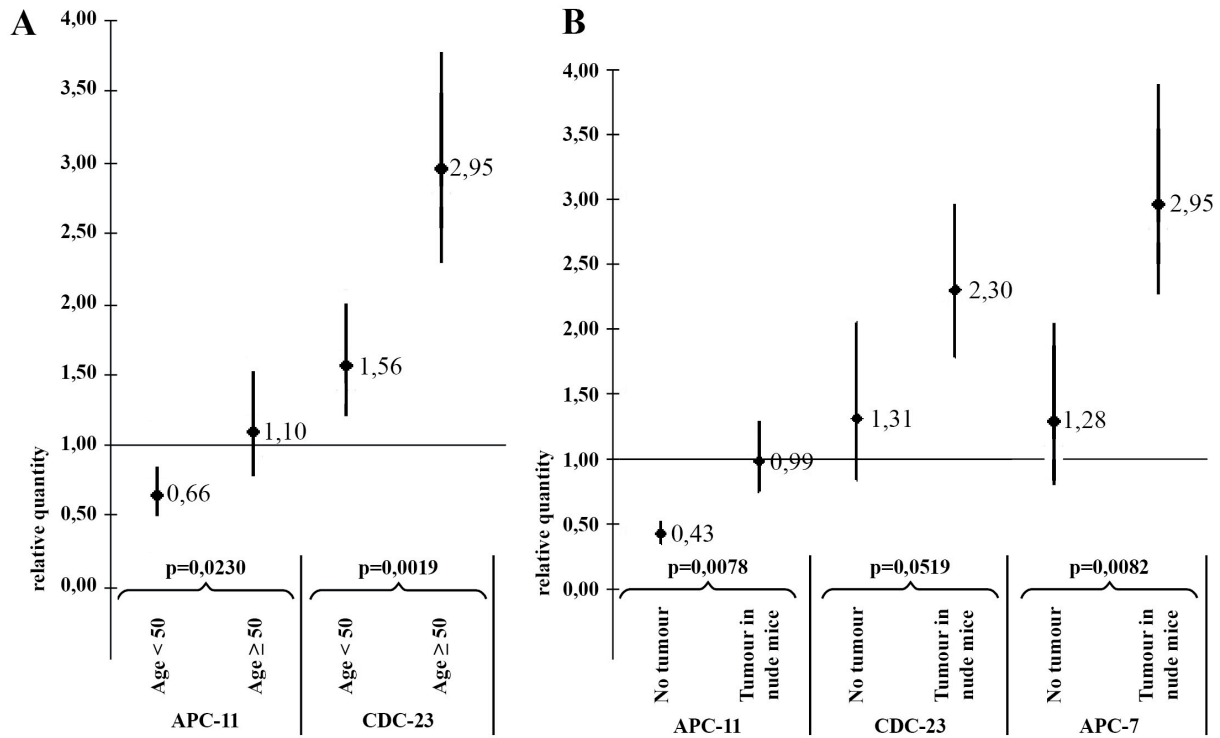




**Figure 3: Comparative analyses of APC-11 expression in breast and colon cancer cell lines**  
 Western blot analysis of APC-11 protein and mRNA relative quantification in breast/colon cancer cells. The ratios of APC-11/reference genes of five independent real time RT-PCR were expressed as mean  $\pm$  SD.

In multivariate analysis, the expression of APC11 and APC7 were increased by 1.985 (1.192-3.307) for APC11 and by 1.893 (1.095-3.373) for APC7 with significance of 0.0137 (APC11 p-value) and 0.0298 (APC7 p-value) for cell lines tumourigenic in nude mice.

In breast primary tumours, we have found a significant association between CDC16, CDC23, and APC7 RNA expression levels with presence of metastases (CDC16 p-value = 0.0476; CDC23 p-value= 0.0032; APC7 p-value = 0.0082) (fig. 5A). Significantly higher levels of these three APC/C subunits were found in tumours of patients suffering of metastatic breast cancer pathology. CDC16 expression was also found higher in breast cancer biopsies with hormone receptor status negative (ER+/PR+: mean = 1.685 vs ER-/PR-: mean = 2.427; p-value = 0.0380). Tumours of clinical grade III expressed higher levels of mRNA APC11 than tumours of clinical grade I (grade I: mean= 0.538 vs grade III: mean= 1.035; p-value = 0.0449) (fig. 5B).



**Figure 4: Statistical correlations between APC-11, APC-7 and CDC-23 expression and molecular features in breast cancer cell lines**

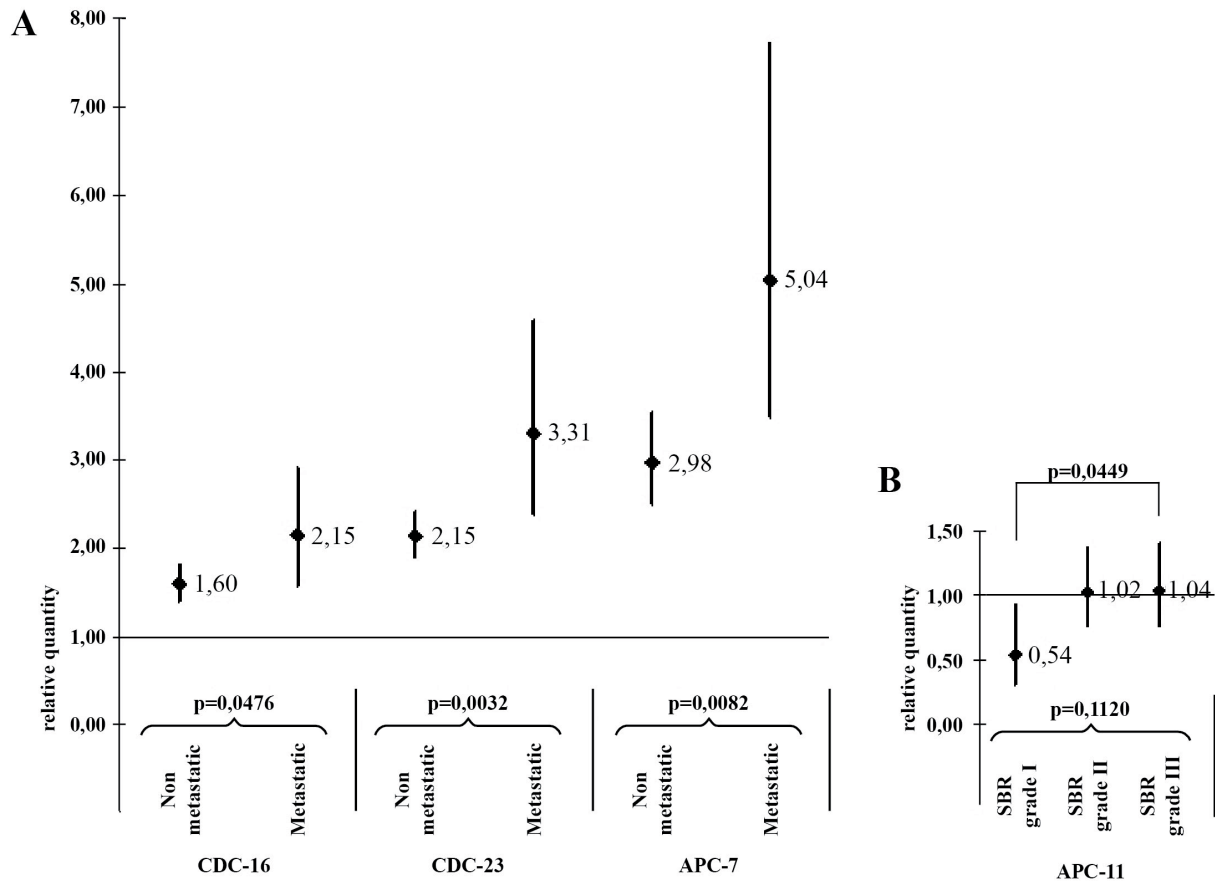
Total RNA of each cell lines was extracted twice and the relative levels of APC-11, APC-7 and CDC-23 mRNA were measured by real-time RT-PCR (five independent RT were done). The results of the statistical analysis (see materials and methods) are displayed using a graphical representation. For each case, the number indicates the mean expression and the bars the confidence interval. Significance for the test is reach when  $p < 0,05$ .

**A.** Significant correlation for each gene with the age of the patient from which the cell line was established.

**B.** Significant correlation for each gene with the ability of cells to form a tumour in nude mice.

## Discussion

To our knowledge, this study examined for the first time the mRNA expression levels of six keys APC/C subunits in two of the most frequent solid cancers occurring into worldwide population, breast and colon adenocarcinomas. We analyzed their potential relationship with clinicopathological parameters. Genetic instability is a hallmark feature of breast, colorectal and other types of cancers. One type characterized by chromosomal instability is thought to be important in the pathogenesis of many solid tumors displaying aneuploidy. This latter type of defect is named chromosomal instability (CIN). Despite its relatively common phenotype, the molecular mechanisms giving rise to CIN remain widely undefined. It is now widely accepted that tumorigenesis and cancer progression result from the progressive accumulation of genetic alterations conferring a growth advantage to the cells. The anaphase-promoting complex (cyclosome) is



**Figure 5: Statistical correlations between APC-11, APC-7, CDC-16 and CDC-23 expression and molecular features in breast tumour samples**

Total RNA of each tumour sample was extracted twice and the relative levels of APC-11, APC-7, CDC-16 and CDC-23 mRNA were measured by real-time RT-PCR (five independent RT were done). The results of the statistical analysis (see materials and methods) are displayed using a graphical representation. For each case, the number indicates the mean expression and the bars the confidence interval. Significance for the test is reach when  $p < 0,05$ .

**A.** Significant correlation for APC-7, CDC-16 and CDC-23 genes with the metastatic status of the tumours.

**B.** Significant correlations for APC-11 gene with the SBR grade of the tumours.

an unusually complicated ubiquitin ligase, composed of at least eleven core subunits in humans, and either of two associated regulatory subunits, *cdc20* and *cdh1* (Castro et al., 2005). *Cdc20* expression was previously found over-expressed in different cancer pathologies included solid tumours (Li et al., 2003; Kim et al., 2005). The other WD 40 APC/C activating subunit, *cdh1* was recently found down-regulated in breast cancers (Fujita et al., 2008). These observations highlight the importance of tight regulation of APC/C complex components in normal cells. Stoichiometry of the core subunits has been shown to be essential for cyclosome activity. Most components of APC/C have been reported to be expressed in growing tissues at fairly constant levels (Tugendreich

et al., 1995). However it is known how the expression of APC/C components is modulated according to growth or cell differentiation. Here we focused our work on key core components: APC11 (ring), three TPR subunits (CDC16, CDC23, CDC27). It was shown that the 3 TPR subunits display a sequential binding dependency, with CDC27 the most peripheral, CDC23 the most internal and CDC16 between (Thornton et al., 2006). APC5 and APC7 were shown to influence p53 expression (Liu et al., 2006; Turnell et al., 2005). APC5 and TPR subunits associate interdependently, such that loss of any one subunit reduced binding of other. It was shown that TPR subunits contributed to the binding of cdh1 to the APC/C core complex (Thornton et al., 2006). We have previously found mutations affecting APC/C subunits in colon cancer cell lines but not in other types of cancer (Wang et al., 2003). Results from epigenetic studies have further implicated APC/C subunits in tumorigenesis with deregulation in APC6, CDC16 and CDC23 in gastric and lung cancers (Singhal et al., 2003; Kim et al., 2005). Interestingly, the analysis of mRNA expression levels of APC11, CDC16, CDC23, CDC27, APC5 and APC7 revealed that APC11 expression is significantly associated with aneuploidy status in colon cancer cell lines and the other subunits failed to show association with biological and clinical parameters. In breast cancer cell lines, APC11 was also found significantly associated with aggressiveness parameters but also CDC16, CDC23 and APC7, showing a cancer specificity of deregulation. In the same way, in primary breast cancers, increased levels of these subunits were associated with bad prognosis parameters. We did not find any association for primary colon carcinomas mostly because of the limited number of tumours included in our study. Because ubiquitin-dependent proteolysis facilitates normal cell cycle progression, dysfunction of the ubiquitin-proteasome pathway could result in carcinogenesis by disrupting the balance between oncoproteins and tumour suppressor proteins (Weissman et al., 2001; Hoeller et al., 2006; Datto et al., 2005). We could hypothesize that increased levels of APC11 the catalytic subunit of cyclosome lead to increase degradation of APC/C targets like cyclin B1 causing a premature mitotic exit which have been shown to generate chromosomal

instability. Securin, another very well-known APC/C target, is required to fully activate separase for chromatid segregation. Cells that lack securing, lose chromosomes at high frequency because of abnormal anaphases, during cells make repeated unsuccessful attempts to segregate their chromosomes (Jallepalli et al., 2001). Over-expression of one the APC/C subunit most importantly APC11, probably does not inhibit the spindle checkpoint per se but rather partially uncouples APC/C activity either directly or indirectly from spindle checkpoint control. Increased levels of APC11 in cancer cells might lead to aneuploidy and increased levels of either other APC/C subunits may have different impact on cell cycle regulation. Dysfunction of subunits other than the catalytic one may affect the degradation of APC/C targets selectively. We have found only 60% of concordance between APC11 mRNA and protein levels both in colon and breast cancer cell lines. Little is known about post-transcriptional regulation of APC11 subunit in human cells. Our understanding of the individual subunits is mostly limited to the conserved structural motifs in their amino acid sequences. However it is well known that APC11 is capable of autoubiquitination (Chang et al., 2004). Increased transcription could in turn lead to activation of autoubiquitination from APC11 as a feedback control. This autocatalytic phenomenon could explain the oscillation of APC11 proteins levels we observed during cell cycle.

In the present study, poor prognostic indication such as high clinical grade, metastases and aneuploidy were found to be related to up-regulation of APC/C subunits RNA expression, which suggest that breast cancer suffering patients with high APC/C subunits levels of expression would have poor survival. It was reported (Yoshimoto et al., 1993) that clinical grade and metastases are important prognostic markers in breast cancer patients, which support the notion that over-expression of APC/C subunits could be a viable prognostic marker. We have found increased levels of APC7 in metastatic breast cancers. Very recently, Kang et al. (2008) have described high levels of APC7 in phyllodes tumours of breast (Kang et al., 2008). Interestingly, phyllodes tumours exhibit increased mitotic figures in comparison with more benign tumours.

They have also shown that APC7 levels were closely related to the proliferation index marker, Ki-67.

Therapies against the proteasome like Bortezomid were recently approved as anticancer therapies in the treatment of different cancer pathologies. Targeted therapy against APC11 in breast cancers might have the advantage of specifically targeting the most aggressive and aneuploid cells and should be less toxic than non specific proteasome inhibitors.

## **Bibliographie**

- Berlingieri, M. T., P. Pallante, et al. (2007). "UbcH10 is overexpressed in malignant breast carcinomas." *Eur J Cancer* 43(18): 2729-35.
- Bernassola, F., M. Karin, et al. (2008). "The HECT family of E3 ubiquitin ligases: multiple players in cancer development." *Cancer Cell* 14(1): 10-21.
- Cahill, D. P., C. Lengauer, et al. (1998). "Mutations of mitotic checkpoint genes in human cancers." *Nature* 392(6673): 300-3.
- Cajot, J. F., I. Sordat, et al. (1997). "Differential display cloning identifies motility-related protein (MRP1/CD9) as highly expressed in primary compared to metastatic human colon carcinoma cells." *Cancer Res* 57(13): 2593-7.
- Castro, A., C. Bernis, et al. (2005). "The anaphase-promoting complex: a key factor in the regulation of cell cycle." *Oncogene* 24(3): 314-25.
- Chang, T. S., W. Jeong, et al. (2004). "The RING-H2-finger protein APC11 as a target of hydrogen peroxide." *Free Radic Biol Med* 37(4): 521-30.
- Coster, G., Z. Hayouka, et al. (2007). "The DNA damage response mediator MDC1 directly interacts with the anaphase-promoting complex/cyclosome." *J Biol Chem* 282(44): 32053-64.
- Datto, M. and X. F. Wang (2005). "Ubiquitin-mediated degradation a mechanism for fine-tuning TGF-beta signaling." *Cell* 121(1): 2-4.
- Fujita, T., W. Liu, et al. (2008). "Dissection of the APCCdh1-Skp2 cascade in breast cancer." *Clin Cancer Res* 14(7): 1966-75.
- Gayet, J., X. P. Zhou, et al. (2001). "Extensive characterization of genetic alterations in a series of human colorectal cancer cell lines." *Oncogene* 20(36): 5025-32.
- Hao, J., A. Xu, et al. (2008). "Elevated expression of UBE2T in lung cancer tumors and cell lines." *Tumour Biol* 29(3): 195-203.
- Jallepalli, P. V., I. C. Waizenegger, et al. (2001). "Securin is required for chromosomal stability in human cells." *Cell* 105(4): 445-57.
- Jelluma, N., A. B. Brenkman, et al. (2008). "Chromosomal instability by inefficient Mps1 auto-activation due to a weakened mitotic checkpoint and lagging chromosomes." *PLoS ONE* 3(6): e2415.
- Kang, Y., J. H. Kim, et al. (2008). "Expression of anaphase-promoting complex7 in fibroadenomas and phyllodes tumors of breast." *Hum Pathol*.
- Keck, J. M., M. K. Summers, et al. (2007). "Cyclin E overexpression impairs progression through mitosis by inhibiting APC(Cdh1)." *J Cell Biol* 178(3): 371-85.
- Kim, J. M., H. Y. Sohn, et al. (2005). "Identification of gastric cancer-related genes using a cDNA microarray containing novel expressed sequence tags expressed in gastric cancer cells." *Clin Cancer Res* 11(2 Pt 1): 473-82.
- Kim, D. S., J. A. Franklyn, et al. (2007). "Securin induces genetic instability in colorectal cancer by inhibiting double-stranded DNA repair activity." *Carcinogenesis* 28(3): 749-59.
- Kops, G. J., B. A. Weaver, et al. (2005). "On the road to cancer: aneuploidy and the mitotic checkpoint." *Nat Rev Cancer* 5(10): 773-85.
- Lehman, N. L., R. Tibshirani, et al. (2007). "Oncogenic regulators and substrates of the anaphase

promoting complex/cyclosome are frequently overexpressed in malignant tumors." *Am J Pathol* 170(5): 1793-805.

- Li, D., J. Zhu, et al. (2003). "Overexpression of oncogenic STK15/BTAK/Aurora A kinase in human pancreatic cancer." *Clin Cancer Res* 9(3): 991-7.
- Liu, J. and S. Y. Fuchs (2006). "Cross-talk between APC/C and CBP/p300." *Cancer Biol Ther* 5(7): 760-2.
- Mondal, G., S. Sengupta, et al. (2007). "Overexpression of Cdc20 leads to impairment of the spindle assembly checkpoint and aneuploidization in oral cancer." *Carcinogenesis* 28(1): 81-92.
- Nakayama, K. I. and K. Nakayama (2006). "Ubiquitin ligases: cell-cycle control and cancer." *Nat Rev Cancer* 6(5): 369-81.
- Newton, K. and D. Vucic (2007). "Ubiquitin ligases in cancer: ushers for degradation." *Cancer Invest* 25(6): 502-13.
- Okamoto, Y., T. Ozaki, et al. (2003). "UbcH10 is the cancer-related E2 ubiquitin-conjugating enzyme." *Cancer Res* 63(14): 4167-73.
- Park, K. H., S. E. Choi, et al. (2005). "Downregulation of the anaphase-promoting complex (APC)7 in invasive ductal carcinomas of the breast and its clinicopathologic relationships." *Breast Cancer Res* 7(2): R238-47.
- Peters, J. M. (2006). "The anaphase promoting complex/cyclosome: a machine designed to destroy." *Nat Rev Mol Cell Biol* 7(9): 644-56.
- Roh, M., C. Song, et al. (2005). "Chromosomal instability induced by Pim-1 is passage-dependent and associated with dysregulation of cyclin B1." *J Biol Chem* 280(49): 40568-77.
- Singhal, S., K. M. Amin, et al. (2003). "Differentially expressed apoptotic genes in early stage lung adenocarcinoma predicted by expression profiling." *Cancer Biol Ther* 2(5): 566-71.
- Sudakin, V., G. K. Chan, et al. (2001). "Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2." *J Cell Biol* 154(5): 925-36.
- Thornton, B. R., T. M. Ng, et al. (2006). "An architectural map of the anaphase-promoting complex." *Genes Dev* 20(4): 449-60.
- Tugendreich, S., J. Tomkiel, et al. (1995). "CDC27Hs colocalizes with CDC16Hs to the centrosome and mitotic spindle and is essential for the metaphase to anaphase transition." *Cell* 81(2): 261-8.
- Turnell, A. S., G. S. Stewart, et al. (2005). "The APC/C and CBP/p300 cooperate to regulate transcription and cell-cycle progression." *Nature* 438(7068): 690-5.
- Wang, C. X., B. C. Fisk, et al. (2000). "Overexpression of murine fizzy-related (fzr) increases natural killer cell-mediated cell death and suppresses tumor growth." *Blood* 96(1): 259-63.
- Wang, Q., C. Moyret-Lalle, et al. (2003). "Alterations of anaphase-promoting complex genes in human colon cancer cells." *Oncogene* 22(10): 1486-90.
- Weissman, A. M. (2001). "Themes and variations on ubiquitylation." *Nat Rev Mol Cell Biol* 2(3): 169-78.
- Yoshimoto, M., G. Sakamoto, et al. (1993). "Time dependency of the influence of prognostic factors on relapse in breast cancer." *Cancer* 72(10): 2993-3001.
- Yu, R., A. P. Heaney, et al. (2000). "Pituitary tumor transforming gene causes aneuploidy and p53-dependent and p53-independent apoptosis." *J Biol Chem* 275(47): 36502-5.

II.1.2. Valeur pronostique et pertinence clinique de l'expression de *dicer* dans les cancers du sein : relation avec le phénotype mésenchymateux ?

i. Résumé

**D**icer est la protéine clef du mécanisme d'interférence ARN qui permet la maturation d'ARN double-brins principalement en siARN ou miARN. Les cibles et donc les actions de ces petits ARN sont multiples et vont de la régulation d'expression de gènes (en résumé par extinction de la transcription en ce qui concerne les siARN ou par inhibition de la traduction pour les miARN) à la mise en place d'hétérochromatine au niveau de loci bien spécifiques (régions péricentromériques notamment). Dicer joue aussi des rôles importants dans le développement embryonnaire et la différenciation cellulaire des nombreux organismes où elle a été mise en évidence. La dérégulation de son expression a donc pour conséquences des perturbations fortes de nombreuses voies cellulaires et en particulier la division. Le nombre de mitoses anormales est alors élevé, conduisant à l'apparition et à la potentielle survie de cellules au phénotype cancéreux. Dans la droite ligne de ces observations, le taux d'expression de cette enzyme ribonucléasique a été récemment corrélé à des aspects cliniques de cancers des ovaires, des poumons et de la prostate. Jusqu'à maintenant, aucune étude ne s'est intéressée au possible lien entre *dicer* et les cancers du sein alors même que cette pathologie manque de marqueurs pronostiques, diagnostiques et thérapeutiques. En effet, il existe une telle hétérogénéité des tumeurs mammaires qu'il est difficile d'une part de les classer convenablement afin d'adapter les soins en clinique et d'autre part de remonter les événements moléculaires qui ont conduit la tumorigenèse.

Dans cette étude, nous avons donc voulu savoir si *dicer* pouvait être associé aux cancers du sein. Nous avons utilisé pour cela des approches *in vitro*, en observant le taux d'expression (ARN et protéines) de notre gène d'intérêt dans des modèles et lignées cellulaires tumorales mammaires, et cliniques, en observant le taux d'expression (ARN et protéines) de biopsies tumorales en lien avec des paramètres de survie.



Les résultats ont montré qu'une faible expression de l'ARNm de *dicer* était un facteur pronostique indépendant (c'est à dire relié à aucun autre facteur déjà connu) dans les cancers du sein. Ainsi, le taux d'expression de ce gène pourrait prédire un risque de rechute métastatique pour cette pathologie. Les différentes études menées sur l'expression de *dicer* dans d'autres types de cancers ont mis en évidence des corrélations avec les phénotypes tumoraux mais n'ont pas mis en évidence d'impact de l'expression de *dicer* sur la survie. Comme il s'agit de travaux pourtant sur la protéine Dicer et que nos propres résultats de TMA ne permettent pas non plus cette conclusion, il semble important de réaliser ces études à la fois sur l'ARNm et sur les protéines d'intérêt. Par ailleurs, nous avons mis en évidence une plus faible expression de *dicer*, toujours au niveau transcriptionnel, dans des lignées tumorales mammaires de type mésenchymateux (par comparaison au type épithélial). Or, il a été montré qu'une transition épithélio-mésenchymateuse des cellules de la tumeur primaire était importante pour conférer un pouvoir métastatique à celle-ci. Ainsi, nous avons pu conclure que globalement, un faible taux d'expression de *dicer* était associé à la capacité métastatique des cellules aussi bien *in vitro* que *in vivo*.

ii. Manuscrit

**Prognostic value and clinical significance of Dicer expression in human breast cancers: any relationship with mesenchymal phenotype?**

Gaël Grelier<sup>1,2,3</sup>, Nicolas Voirin<sup>4,5</sup>, Anne-Sophie Ay<sup>1,2,3</sup>, Sylvie Chabaud<sup>6</sup>, Isabelle Treilleux<sup>7</sup>, Sophie Léon-Goddard<sup>7</sup>, Ruth Rimokh<sup>2,3</sup>, Ivan Mikaelian<sup>8</sup>, Clémence Venoux<sup>8</sup>, Alain Puisieux<sup>1,2,3</sup>, Christine Lasset<sup>4,6</sup> and Caroline Moyret-Lalle<sup>1,2,3</sup>

**Authors' Affiliations:**

<sup>1</sup>Université de Lyon, université Lyon 1, ISPB, Lyon, F-69003, France;

<sup>2</sup>Inserm, U590, Lyon, F-69008, France;

<sup>3</sup>Centre Léon Bérard, Lyon, F-69008, France;

<sup>4</sup>Université de Lyon, université Lyon 1, Faculté Grange Blanche, CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, Lyon, F-69373, France;

<sup>5</sup>Hospices Civils de Lyon, Hôpital Edouard Herriot, Service d'Hygiène, Epidémiologie et Pré-

vention, Lyon, F-69437, France;

<sup>6</sup>Centre Léon Bérard, Département de Santé Publique, Lyon, F-69008, France ;

<sup>7</sup>Centre Léon Bérard, Service d'Anatomopathologie, Lyon, F-69008, France;

<sup>8</sup>Université de Lyon, université Lyon 1, Faculté Grange Blanche, CNRS, UMR5201, Laboratoire de Génétique Moléculaire, Signalisation et Cancer, Lyon, F-69008, France.

**Grant support:** Ligue Nationale Contre le Cancer, Comité Départemental de la Drôme de la Ligue Nationale Contre le Cancer.

**Requests for reprints:** Caroline Moyret-Lalle, INSERM, U590, Centre Léon Bérard, Lyon, F-69008, France.

**Running title:** Prognostic impact of Dicer in breast cancers

## Abstract

*Purpose:* Dicer, a ribonuclease, is the key enzyme required for biogenesis of miRNAs and small interfering RNAs (siRNAs). It was previously shown that Dicer is essential for mammalian development and cell differentiation. Recent evidence indicates that Dicer may be involved in tumorigenesis. However, no studies have examined the clinical significance of Dicer at both RNA and protein levels in breast cancer.

*Experimental Design:* In this study, the biological and prognostic value of Dicer expression was assessed in breast cancer cell lines, breast cancer progression cellular-models, and in two independent and well-characterised set of breast carcinomas with a long term follow-up using tissue microarray and quantitative reverse transcription-PCR.

*Results:* We have found that Dicer protein expression is significantly associated with hormone receptors status in breast tumours. Dicer mRNA expression appeared to have an independent prognostic impact in metastatic disease. In the breast cancer cell lines, lower Dicer expression was found in cells harbouring a mesenchymal phenotype and in metastatic bone derivatives of a breast cancer cell line. These findings suggested that down-regulation of Dicer expression may be related to metastatic spread of tumours.

*Conclusions:* Assessment of Dicer expression may facilitate prediction of distant metastases for patient suffering from breast cancer.

## **Introduction**

The Dicer gene encodes a protein which functions as an RNase III type endonuclease and is required for RNA interference (RNAi) and microRNA (miRNA) pathways. It produces miRNAs and small interfering RNAs (siRNAs) from pre-miRNAs and dsRNA respectively. Dicer null mice are embryonic lethal with depletion of stem cells. Furthermore ES (Embryonic Stem) cells lacking Dicer are viable but incapable of differentiation and display severe proliferation defects (1). Dicer was shown to be essential in stem cell maintenance by its involvement in auto-renewal and proliferation (2). In humans, Dicer was very recently shown to play a key role in X-inactivation pathways (3). Because of its central role in post-transcriptional regulation of miRNAs, Dicer is increasingly evoked in cancer studies. A global decrease in miRNA expression was reported in cancer cells (4) so we hypothesize that this key cell fate regulator can be deregulated during breast cancer progression and has already been associated with ovarian, lung and prostate cancers (5-7).

Breast cancer is the leading cancer diagnosis among women in the western world. With ever-improving chemotherapeutic, radiation, hormonal treatments, and Her2 and EGFR antagonists, an improvement in overall survival had been noticed. However, the treatment of breast cancer is currently far from being optimal with patients suffering from recurrent breast carcinoma usually dying of their disease. Metastasis represents the lethal evolution of breast cancer. The natural progression of breast cancer differs greatly between patients and metastatic progression in breast cancer is a complex and largely unknown process. The different biological behaviours observed among the distinct breast cancer subtypes may suggest different mechanisms of invasion and metastasis for breast tumours. Current therapy decision-making is increasingly governed by the molecular classification of breast cancer (luminal, basal-like, HER2+). Cancer subtypes have characteristic sites to which they metastasize (8). A very common metastatic site for human breast cancer is bone. Genes involved in metastasis but not in primary tumorigenicity have been identified, particularly in bone metastases

(9). Identification of biological factors in relation with metastatic evolution and patterns is important for the optimal management of patients.

The aim of this study, using real-time reverse transcription (RT)-PCR, tissue microarray and western blotting, was to assess whether the expression levels of Dicer are altered in primary breast cancers and established cell lines, and linked to metastatic evolution. We report here for the first time an alteration in Dicer expression that could be an independent prognostic factor for metastatic evolution of breast tumours.

## **Materials and Methods**

### *Biological samples and cell lines*

Tissue specimens were purchased by Eric Tabone (Biological Resources Department) and were obtained before any therapy from 117 patients suffering from breast cancer diagnosed between 1992 and 1999 whom underwent surgery at Centre Léon Bérard. Normal breast tissue samples were also obtained from 4 individuals. HMECs-hTERT, HMECs-hTERT+ LT and HMLER were first derived and kindly provided by RA. Weinberg and M. Brooks (Withehead Institute for Biomedical Research, Ludwig Center for Molecular Oncology, MIT Department of Biology, Cambridge MA). BO2 bone derivative MDA-MB-231 clone was kindly provided by P. Clezardin (Unit Research U664, Laennec School of Medicine, INSERM, Lyon, France) SCP2 bone derivative MDA-MB-231 clone was kindly provided by Yibin Kang (Department of Molecular Biology, Princeton University, Princeton, NJ) and Joan Massague (Cancer Biology and Genetics Program, and Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Centre, New York). The 67NR, 168 FARN, 4TO7, 66c14 and 4T1 cell lines were first derived and kindly provided by F. Miller (Karmanos Cancer Centre, Detroit, Michigan) (Miller et al., 1983). The human mammary HMECs-hTERT, HMECs-hTERT+ LT and HMLER cell lines, the SCP2 and BO2 bone derivatives, and the mouse mammary tumour cell lines (67NR, 168 FARN, 4TO7, 66c14 and 4T1) were maintained as descri-

bed (Elenbaas et al., 2001; Minn et al., 2005; Garcia et al., 2008; Aslakson and Miller, 1992). 21 breast cancer cell lines were obtained from the American Type Culture Collection ([www.ATCC.org](http://www.ATCC.org)) (See supplementary Materials and Methods).

#### *RNA extraction and real-time PCR*

Total RNA extraction was performed with a phenol-chloroform method, using TriReagent (Sigma) for cells lysis and PhaseLockGel tubes (Eppendorf) for phase separation. After DNase treatment, DNA contamination of each sample was checked by electrophoresis on agarose gel. The synthesis of cDNA was done using the First-Strand cDNA Synthesis Kit (GE-Healthcare). Transcription from Dicer gene produces 14 different mRNAs, 11 alternatively spliced variants and 3 unspliced forms. We have chosen set of primers to amplify the three a, b and c unspliced isoforms coding for the entire protein. Primers sequences and real-time PCR experimental procedures using LightCycler (Roche Applied Science) are described in Supplementary Materials and Methods.

For microRNA analysis, real-time PCR was performed as above, using TaqMan microRNA assays according to the manufacturer's instructions (Applied Biosystems) on AB Prism 7000. All microRNA data are expressed relative to RNU44 a small nucleolar (sn) RNA TaqMan PCR performed on the same sample. Expression variations were calculated with the  $\Delta\Delta Ct$  method. Taqman microRNA reverse transcription Kit hsa-miR-21: 4373090, has-miR-182: 4373271, has-miR-221:4373077, sno202: 4380917, RNU44: 4373384, were used to amplify the different microRNA.

#### *Immunohistochemistry*

Formalin fixed paraffin embedded breast tumors from 120 Centre Léon Bérard patients diagnosed in 1998 with invasive breast cancer were inserted as triplicate using a 600 $\mu$ m needle (Alphelys, Plaisir, France) in 3 tissue micro-array blocks (TMA). These blocks were cut at 4  $\mu$ m. After deparaffinization and rehydration, endogenous

peroxidases were blocked by incubating the slides in 5% hydrogen peroxide in sterile water. For heat induced antigen retrieval, tissue sections were boiled for 40 minutes in 10 mM citrate buffer pH6 for 40 minutes in a water bath. Non specific binding was blocked with a protein blocking reagent (Immunotech, Marseille, France) for 5 minutes. Slides were then incubated overnight at 4°C with a mouse monoclonal anti-Dicer antibody (clone 13D6, Abcam, Cambridge, UK) diluted at 1/150 using an antibody diluent solution (Chem Mate, Dako, Trappes, France). After rinsing in phosphate buffered saline, the slides were incubated with a biotinylated secondary antibody bound to a streptavidin peroxidase conjugate [Vecastain Elite ABC reagent Vector, Abcys, Paris, France). Bound antibody was revealed by adding the substrate 3,3'-diamino benzidine (DAB substrate kit for peroxidase Vector, Abcys, Paris, France). Sections were counterstained with hematoxylin. After dehydration and mounting, they were analyzed independently by both the pathologist and the technician. Both the intensity of cytoplasmic staining (3 grades) and the percentage of positive tumor cells, was assessed. The intensity of staining was graded on a scale from 0 to 2: "0" reflected a lack of immunoreactivity, "1" reflected weak immunoreactivity and "2" reflected strong immunoreactivity. The few cases of discordance were reviewed by the two investigators to reach a consensus.

### *siRNA and transfection*

The sequences (designed by Eurogentech) of the siRNAs used in this study are: siDicer-1 5'-AGAGGUACUUAGGAAUUU-3' (recognising dicer variants a, b and c); siCt-1 5'-UUACCUCCUUAGAUACAAU-3'; siDicer-2 5'-GGAGCUUGAUUUG-CAUGAU-3' (recognising dicer variants a, b and c); siCt-2 5'-CCAAACACGGCUUU-CAAU-3'; siDicer-3 5'-UCAGGAACAUAUCAGAUUU-3' (recognising dicer variants a, b, c, d and e); siCt-3 5'-UAGUUA AUGCUAGUCAUAC-3'. Transfection of HeLa cells was performed with jetSI-endo (Polyplus Transfection), as described by the manufacturer. To examine the effect of siRNAs on target gene expression, total RNA and total

protein extracts were prepared and analysed by quantitative RT-PCR and Western Blot respectively.

#### *Immunoblot detection of Dicer*

Cell pellets were lysed with TEB150 buffer (50 mM HEPES ; 150 mM NaCl ; 2 mM MgCl<sub>2</sub> ; 5 mM EGTA ; 1 mM DTT ; 0.5% Triton X-100 ; 10% glycerol ; 1 mM Na<sub>3</sub>VO<sub>4</sub> ; 10 µL PIC-1 P2850 SigmaAldrich and 10 µL PIC-2 P5726 SigmaAldrich) for 30 min on ice. Western blot analyses were carried out as previously described (Wang et al., 2003) using a nitrocellulose membrane (BioRad). Membranes were incubated with a 1:1000 dilution of human anti-Dicer antibody (Abcam Ab14601) overnight or with a 1:750 dilution of mouse anti-Dicer antibody (Abcam Ab 13502) overnight and followed by an incubation with 1:3000 anti-mouse IgG (Dako P0260) for 1 h at room temperature. Dicer expression was detected with the ECL Western blotting detection and analysis system (GE Healthcare) according to the manufacturer's protocol. Relative protein expression was estimated using Quantity One software (BioRad). The antibody we used for human Dicer recognises full length protein (218.7 Kd) encoding by the three unspliced transcripts a, b and c and alternative proteins encoding by the d and e alternative transcripts (113.2 Kd and 97.2 Kd respectively). In order to compare Dicer RNA and protein levels we quantified only the band corresponding to the known molecular weight of Dicer protein (218.7 Kd).

#### *Clinical Database description*

Explanatory variables have been extracted from the Centre Léon Bérard institutional breast cancer database including all patients with initial surgery performed in the institution since 1996. The database has been declared to the French authorities (CNIL). Parameters of interest include: the clinical or surgical history of the patient, the histology of the tumour, the treatments applied, and some immuno-histochemical covariates (hormonal status, Her2+).

This breast cancer database was updated regularly for the follow-up data until patient's death was known (letter to the referring physician, or registrar's office). The information on the evolution of the tumour in terms of local or distance recurrences are prospectively registered.

Clinical parameters taken into account in this study were:

- clinical data (age, menopausal status, metastatic status at diagnosis)
- histological data (histological type, pT, metastatic lymph nodes status, SBR grade)
- immunohistochemical data (hormonal status, HER2 if available)

### *Statistical analysis*

#### *Immunohistochemistry expression and RNA expression:*

A comparison of the distribution of clinical parameters according to the intensity of Dicer cytoplasmic staining (0/1 versus 2) was made by using Pearson's  $X^2$  test; Fisher's exact test was used when the number of patients was too small. For each cell line and tumour, repeated measurements of mRNA expression were obtained by replicating the extraction and reverse transcription steps. The mean mRNA expression and its 95% confidence interval (95%CI) were estimated using a hierarchical model (Goldstein, 2003; Sullivan et al., 1999). This method adjusted for the within-cell line or tumour correlation of the repeated observations and estimated the mean of mRNA expression of each tumour. The median of these values was used to calculate a threshold value separating low ( $\leq 8$ ) and high ( $>8$ ) Dicer mRNA expression among tumours.

#### *Survival analysis:*

The end point of interest was Metastatic Free Survival time, defined as the time from the date of first diagnosis to the date of distant metastatic or to the date of last follow-up for living patients (censored observation).

Survival estimates were calculated using the Kaplan-Meier method (Kaplan and Meier, 1958). Differences in survival estimates between arms (i.e., Dicer levels) were assessed by the log-rank test (Peto et al., 1977) and hazard ratio were calculated



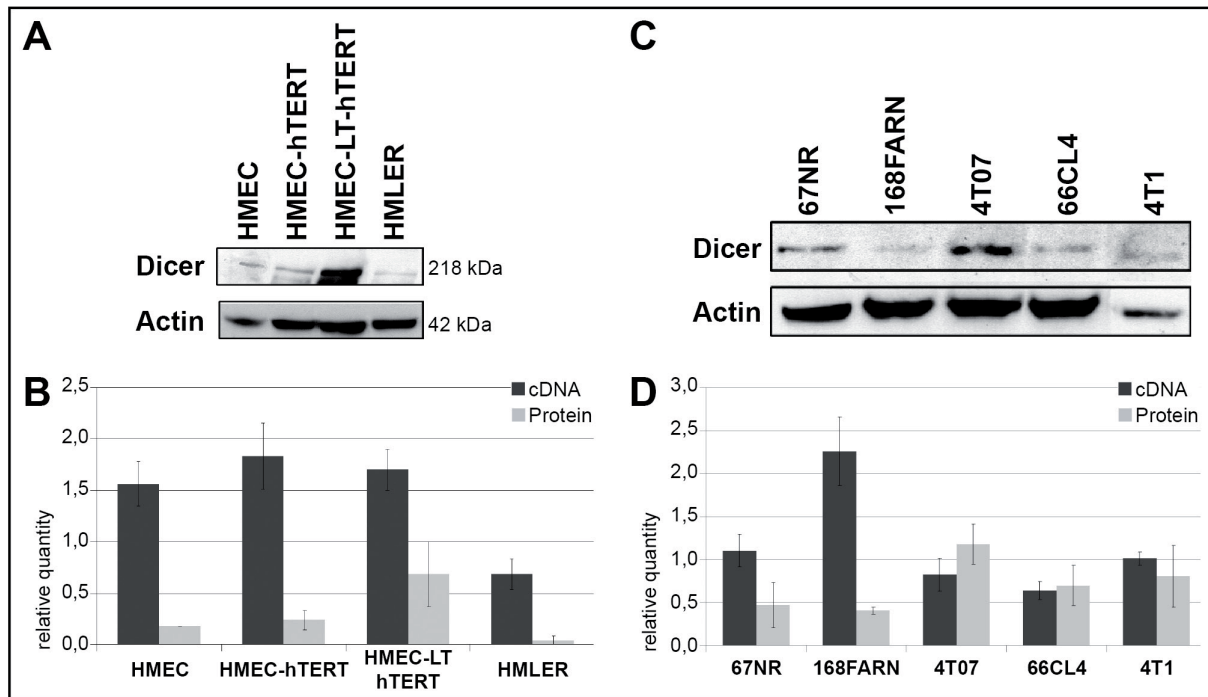
using a Cox Model.

Statistical analysis were realized using SAS software version 9.1 [SAS OnlineDoc®, Version 9, Copyright © 1999 by SAS Institute Inc., Cary, NC, USA] and R [R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.] using the “nlme” package [Jose Pinheiro, Douglas Bates, Saikat DebRoy and Deepayan Sarkar the R Core team. (2007). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-86].

## Results

*Dicer expression in tumour and metastatic progression cellular-models and in breast cancer cell lines: down-regulation in the more advanced stages and in mesenchymal phenotype breast cancer cell lines*

Several in vitro cancer progression and metastasis evolution models were developed in human and mouse cells. Elenbaas et al. (11) have developed a human breast cancer progression model from HMECs cells (normal Human Mammary Epithelial Cells). This model was thought to resume breast cancer progression with the transformation of primary human mammary epithelial cells through the introduction of hTERT (telomerase catalytic subunit), SV40 large-T antigen (LT) and H-ras. HMECs correspond to a mixture of normal luminal and basal epithelial cells (20). HMECs + hTERT and HMECs+ LT + hTERT failed to form colonies in soft agar. The cells expressing hTERT, LT and H-rasV12 (HMLER) formed colonies in soft agar and formed tumours in nude mice (11). No signs of metastatic spread were observed in tumour-bearing mice. We have assessed whether Dicer expression could be regulated transcriptionally and post-transcriptionally in this human cancer progression model (fig. 1A & 1B). Dicer mRNA levels were stable from HMECs to HMEC+ hTERT + LT and strongly decreased in HMLER (fig. 1B). In contrast, Dicer protein levels increased from HMECs



**Figure 1: Expression of Dicer in the human breast tumour progression cellular-model and in the mouse metastasis progression cellular-model**

*Expression of Dicer in the human breast tumour progression cellular-model*

**A.** Western blot analysis of Dicer expression in the four human breast cell lines HMEC, HMEC+ hTERT, HMEC+LT+hTERT and HMLER (expressing H-rasV12) cells. The signal intensity of Dicer was normalized to that of actin

**B.** The relative levels of Dicer mRNA were measured by real-time RT-PCR, and each bar represents the mean  $\pm$  SD of the PCRs in triplicate (■).

Dicer protein levels were quantified using Quantity One software (BioRad) and expressed as protein relative quantity. The ratios of Dicer/actin of three independent studies were expressed as mean  $\pm$  SD (▒).

*Expression of Dicer in the mouse metastasis progression cellular-model*

**C.** Western blot analysis of Dicer expression in the five mouse mammary cell lines 67NR, 168FARN, 4T07, 66c14 and 4T1. The signal intensity of Dicer was normalized to that of actin.

**D.** The relative levels of Dicer mRNA were measured by real-time RT-PCR, and each bar represents the mean  $\pm$  SD of the PCRs in triplicate (■).

Dicer protein levels were quantified using Quantity One software (BioRad) and expressed as protein relative quantity. The ratios of Dicer/actin of three independent studies were expressed as mean  $\pm$  SD (▒).

to HMEC+ LT + hTERT but were subsequently down-regulated in the transformed cells HMLER (fig. 1A & 1B). The cells' evolution towards tumorigenicity is accompanied by down-regulation of Dicer expression both at the mRNA and protein levels. HMLER cells were shown to be a mixed population of mesenchymal phenotype cells and epithelial phenotype cells (21). We have analysed Dicer protein expression in hTERT + LT-RAS of epithelial phenotype and in hTERT + LT-RAS of mesenchymal phenotype and found that Dicer levels in the mesenchymal population cells were lower than levels observed in epithelial phenotype cells (data not shown). As mesenchymal phenotype

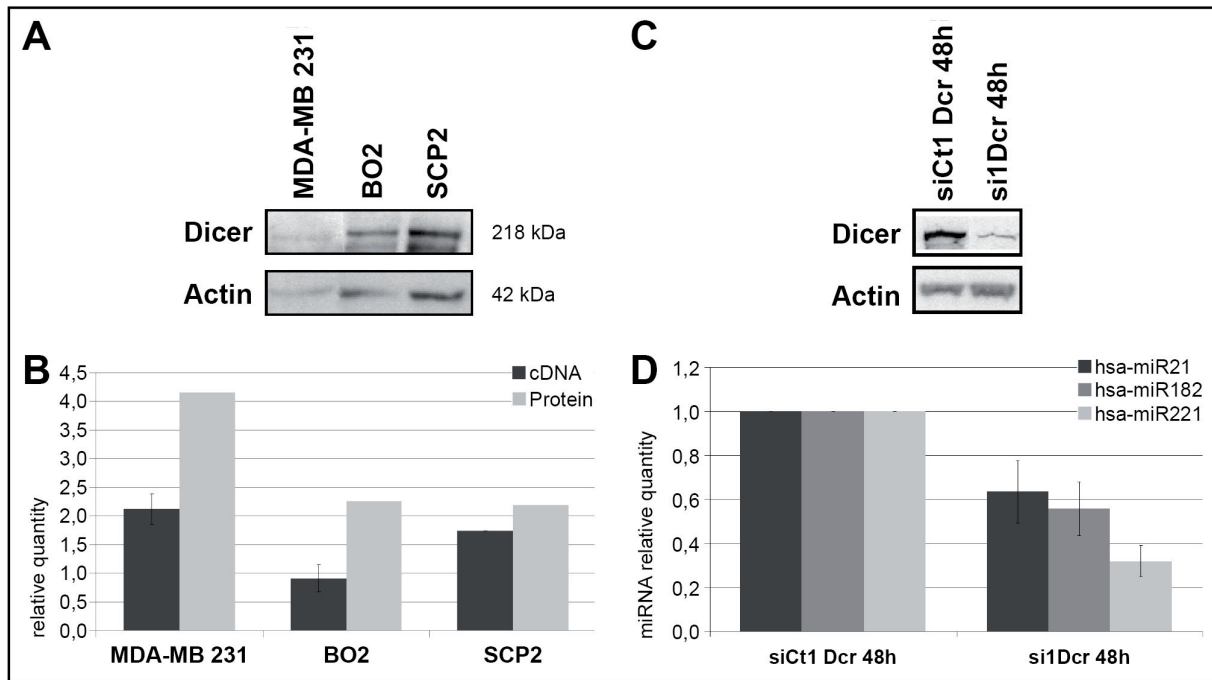
was previously evoked for metastatic potentiality, we asked whether down-regulation of Dicer could be linked to invasiveness capacity.

Because we lacked a human model for metastasis progression in mammary tumour cells, we next analysed Dicer expression in a mouse model of invasiveness (14). This model consists of five clonal tumour sublines derived from a spontaneously arising mammary tumour in a BALB/c mouse. 67NR, 168 FARN and 4TO7 sublines were considered as non metastatic and 66c14 and T1 were considered as metastatic (14). They can be classified as following from the less tumorigenic to the highest 67NR<168FARN<4TO7<66c14<4T1. 67NR cells form primary tumours but are not detectable in any distant tissue while 168 FARN can be detected in lymph nodes. 67NR cells exhibit some aspects of a mesenchymal phenotype by expressing high levels of vimentin and no E-cadherin (22). 4TO7 cells are able to spread to lung but cannot establish visible metastatic nodules. 66c14 and 4T1 are fully metastatic at distant sites (23). The behaviour of these tumour lines is thought to reflect the sequence of multi-step metastasis progression. We therefore set out to compare mRNA and protein Dicer expression profiles of the five sublines to investigate whether Dicer could be regulated during metastasis progression. Dicer at mRNA and protein levels showed similar progression with upregulation as the cells progressed from no invasion towards lymph node metastases and a subsequent decrease as the sublines were able to efficiently colonize distant organs (liver, lungs)(fig. 1C & 1D). After lymph node invasion, breast cancer commonly causes osteolytic metastases in bone. We next addressed whether Dicer could be down-regulated in bone metastasis model of breast cancer cells. Mammary tumours do not spontaneously metastasize to bone in mice, as is the case in humans, which hamper studies in animal models. We assessed Dicer expression in two distinct bone metastasis subpopulations both derived from the human breast cancer cell line MDA-MB-231. BO2 cells were obtained after injection of MDA-MB-231 cells into the tail vein (24) and SCP2 (Single Cell-derived Progeny 2) were obtained after intracardiac injection of parental cells into immune-deficient mice (12). Interestingly,

we have found a significant decreased expression of Dicer both at mRNA and protein levels in the BO2 and SCP2 cells in comparison with the parental cells (fig. 2A & 2B). Transcriptomic analyses done in MDA-MB-231 parental and in BO2 clone have revealed the same decrease in Dicer expression (13, 24) (R Bachelier and P Clézardin, personal communication).

Dicer plays a critical role in miRNA maturation, in other respects a number of studies that compared miRNA expression patterns in healthy breast tissue and breast carcinoma found differential expression for several miRNAs (25). We further examined if the down-regulation of Dicer expression in tumour cells was accompanied by reduced steady-state miRNA levels. We have studied the expression of 3 representative mature miRNA, miR-182, miR-221 and miR-21 in Dicer knocked-down human tumour cells in comparison with the levels found in the same tumour cells transfected with a specific control siRNA (fig. 2C & fig. 2D). In Dicer knocked-down cells as compared with the controlled siRNA transfected cells, we could notice a global decrease of the three mature miRNA (fig. 2D). The same results were obtained with three different control siRNA (siCt1, siCt2 and SiCt3) and with three different siRNA homologous to Dicer sequence (si1Dcr, si2Dcr and si3Dcr, data not shown).

We next investigated the clinicopathologic significance of Dicer mRNA expression in human mammary cancer cell lines and in normal mammary epithelial cells (HMECs). We analysed seven parameters and a very significant association was encountered between the levels of Dicer and cell type, as in mesenchymal phenotype the levels of Dicer was lower compared to the epithelial cell (p-value = 0.0002, Table 1). Interestingly, the clearest modification of Dicer expression was found with the cancer type, with increasing levels between immortalized cells and DCIS (ductal carcinoma in situ), and a subsequent marked decrease in IDC (invasive ductal carcinoma) and metastases (p-value < 0.0001, Table 1). In parallel, we analyzed Dicer protein levels in the breast tumour cell lines and the normal HMECs cells. Semi-quantitative measures



**Figure 2: Expression of Dicer in two bone metastatic derivatives of MDA-MB-231 cells and expression of miRNA in Dicer knocked-down human tumour cells**

*Expression of Dicer in two bone metastatic derivatives of MDA-MB-231 cells*

A. Western blot analysis of Dicer expression in the three cell lines, MDA-MB-231, BO2 and SCP2. The signal intensity of Dicer was normalized to that of actin.

B. The relative levels of Dicer mRNA were measured by real-time RT-PCR, and each bar represents the mean  $\pm$  SD of the PCRs in triplicate (■).

Dicer protein levels were quantified using Quantity One software (BioRad) and expressed as protein relative quantity. The ratios of Dicer/actin of three independent studies were expressed as mean  $\pm$  SD (■).

*Expression of mature miRNA in Dicer knocked-down human tumour cells*

C. Western blot analysis of Dicer expression in HeLa cells 48h after transfection by the control siRNA si-Ct1 or the siRNA against Dicer si1Dcr. The signal intensity of Dicer was normalized to that of actin.

D. Expression of hsa-miR-21 (■), hsa-miR-182 (■) and hsa-miR-221 (■) were measured by real-time RT-PCR in HeLa cells transfected by the si-Ct1 or the si1Dcr. Reactions were performed in triplicate and normalized to RNU44 cycle threshold values.

**Table 1** Quantitative real-time RT-PCR analysis of *dicer* mRNA expression in mammary cell lines and correlation with various biological clinicopathological characteristics

Biological and clinical Characteristics	Cases n=23	Measures 111	<i>dicer</i>	
			Mean (IC)	P-value
<b>Cell type</b>				
epithelial	11	55	4.056 (3.230-5.093)	<b>0.0002</b>
mesenchymal	7	35	1.763 (1.363-2.282)	
<b>Cancer type</b>				
Immortalized cell	3	15	2.198 (1.577-3.064)	
Intraductal carcinoma (DCIS)	4	20	4.107 (3.445-4.895)	
Invasive ductal Carcinoma (IDC)	12	60	1.524 (NA)	
Metastases	1	5	1.369 (1.045-1.794)	<b>&lt;0.0001</b>

of protein levels revealed 72% of concordance between RNA and protein levels (fig.

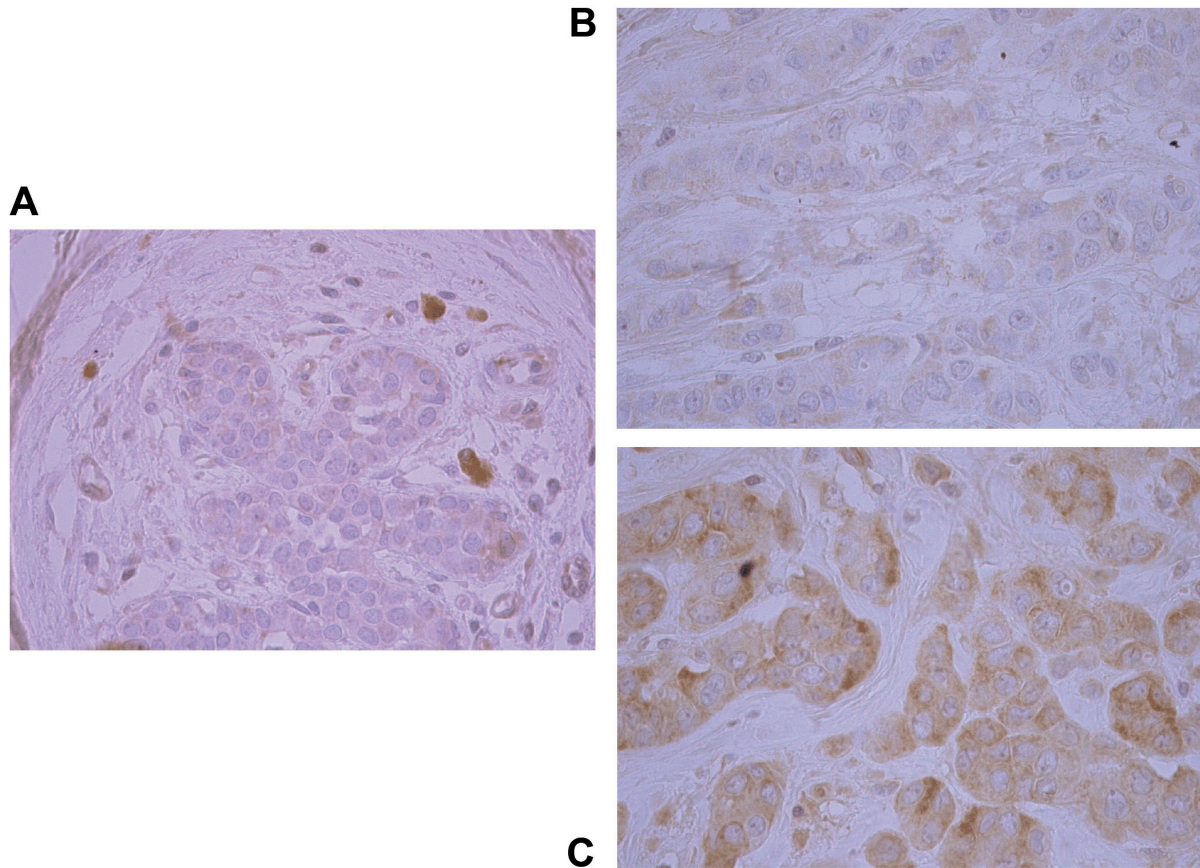
S1, see supplementary).

*Evaluation of the clinical and prognostic significance of Dicer expression in breast cancer patients: association with distant metastases.*

We further examined the potential clinical significance of mRNA Dicer and of immunohistochemically (IHC) examining Dicer protein expression in breast cancer tissues. There was no clear association between Dicer RNA expression levels and patients or tumour clinicopathological parameters we could investigate except for lymph node status: a significantly lower mRNA Dicer expression was observed in cases with lymph node metastases compared with others (N+: mean = 7.353 vs N0: mean = 12.852; p-value = 0.0204, Table S1, see supplementary).

In a preliminary IHC analysis with normal mammary gland tissues, Dicer protein was found to be expressed exclusively in the cytoplasm of the luminal cells (fig. 3A). The signal within the cytoplasm of luminal cells was weak. No positive immunoreactivity was found either in basal or in stromal cells. In mammary tumour cells, Dicer showed variable expression patterns. Some tumours were negative for Dicer staining (data not shown) while other carcinoma tissue sections exhibited a weak diffuse cytoplasmic immunoreactivity in the whole tumour sample (fig. 3B) and a third population showed stronger cytoplasmic Dicer staining in a reduced tumour cell population forming aggregates (fig. 3C). The majority of tumours in TMA analysis (65/86), 75% belonged to the intensity 0/1 and 60% (36/65) of them showed more than 60% of stained cells compared to only 29% (5/21) when intensity of staining reached 2 (p-value = 0.013). We found a significant association between Dicer protein expression and hormone receptors status and the triple negative status: the intensity 0/1 was more frequently observed in ER positive and PR positive tumours and in tumours that did not exhibit the triple negative status (ER+ 71.8% vs ER- 54.6%, p-value = 0.008; PR+ = 71.9% vs PR- 54.6%, p-value = 0.019; No "triple neg" 70.8% vs "triple neg" 50.0%, p-value = 0.036, Table S2, see supplementary data).

At last, we examined whether Dicer expression could be useful for prognostication of breast cancer patients. The probability of metastasis-free survival was signi-



**Figure 3: Immunohistochemical analysis of Dicer protein expression in normal and breast cancer tissues**

**A.** Low intensity staining in cytoplasm of epithelial luminal normal breast cells. Non specific staining of mast cells.

**B.** Low intensity staining in cytoplasm of infiltrating ductal carcinomatous cells

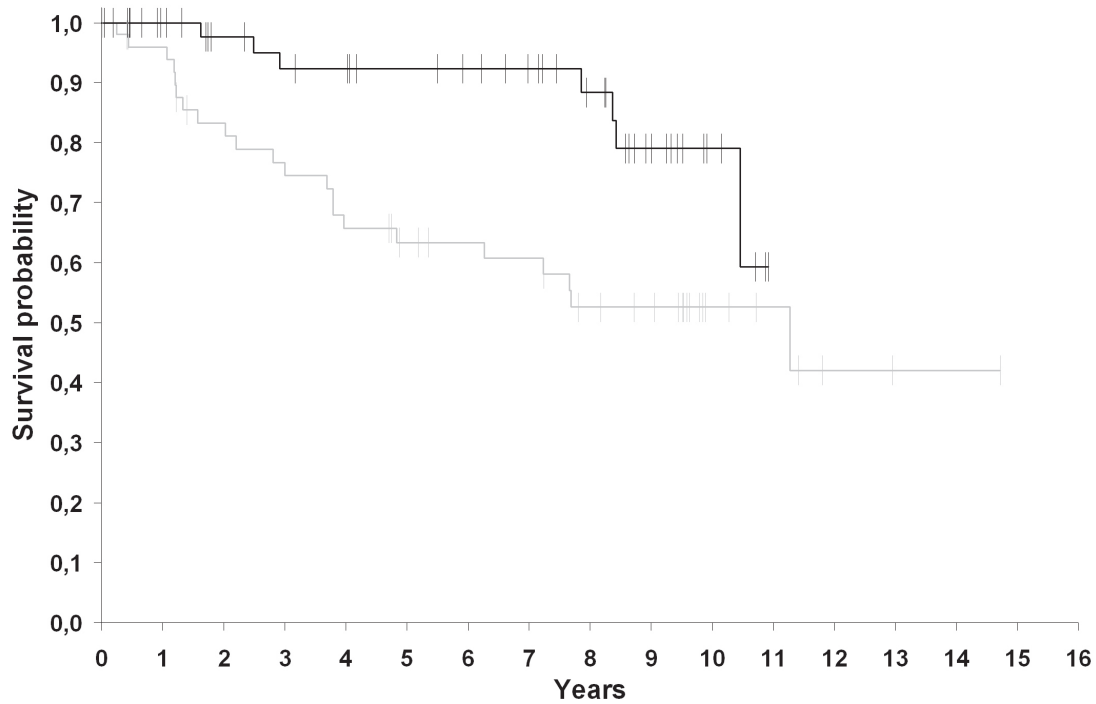
**C.** High intensity staining in cytoplasm of infiltrating ductal carcinomatous cells Magnification, x 40 (A-C).

ificantly lower for patients with levels of Dicer mRNA lower than 8 compared to those with levels higher than 8 (fig. 4, HR: 3.10, p-value = 0.0062). Interestingly, when an adjustment was performed for each clinocopathological prognosis criteria, the values of the hazard ratio remained similar and significantly higher than 1 (Table 2).

In contrast, in TMA analysis no association was observed between metastasis free survival and intensity level of Dicer protein (fig. S2, see supplementary).

## Discussion

The aim of the present study was to evaluate the expression of Dicer both at the RNA and protein levels as well as and to evaluate independently their clinical re-



**Figure 4: Kaplan-Meier estimates of metastatic free survival by dicer’s expression among the mRNA population**

— > 8 cDNA relative quantity (7/52 progressions; 8-years event free survival probability 88.3% (95%CI 77.2-99.4);

— ≤ 8 cDNA relative quantity (22/51 progressions; 8-years event free survival probability 52.6% (95%CI 37.5-67.6);

Log-rank test : P = 0.0062.

**Table 2 Metastasis Free Survival – Comparison of the stratified hazard ratios  
Metastatic progression**

Subgroups	TMA population		mRNA population	
	MetaProg /N	HR (95%CI) (0/1 vs 2)	MetaProg /N	HR (95%CI) (≤8 vs >8)
Overall	19/86	1.17 (0.38-3.55)	29/104	3.10 (1.32-7.31)
Adjusted effect on age	19/86	1.10 (0.36-3.34)	29/103	3.17 (1.35-7.47)
Adjusted effect on menopause	19/86	1.20 (0.40-3.66)	29/97	3.21 (1.6-7.60)
Adjusted effect on pT	19/86	0.77 (0.25-2.39)	29/99	3.19 (1.35-7.54)
Adjusted effect on ductal type	19/86	1.18 (0.39-3.59)	29/101	2.92 (1.24-6.90)
Adjusted effect on SBR	19/86	1.49 (0.49-4.59)	28/95	3.11 (1.32-7.38)
Adjusted effect on N	19/86	1.26 (0.41-3.89)	29/104	2.69 (1.14-6.33)
Adjusted effect on ER(%)	19/86	1.87 (0.56-6.19)	28/98	3.08 (1.30-7.33)
Adjusted effect on PR(%)	19/86	1.83 (0.57-5.90)	29/100	3.17 (1.35-7.48)
Adjusted effect on HER2	19/86	1.04 (0.34-3.24)	8/18	3.25 (0.39-27.45)
Adjusted effect on M	19/86	0.78 (0.24-2.54)	29/102	3.26 (1.38-7.67)

**Footnote** : \*

Hazard ratio definition (HR)

- HR > 1, the patient has a higher risk to have a metastatic recurrence at time t compared to the reference item.
- HR < 1, patient has a lower risk to have a metastatic recurrence at time t compared to the reference item.



levance in breast carcinomas and in breast cancer cell lines. Furthermore, we investigated whether Dicer expression could be modulated during tumour and metastasis progression using human and mouse cellular-models.

We have first showed that lower Dicer mRNA expression could be an independent prognostic factor in breast cancer. Low Dicer mRNA expression is more frequent in patients with metastatic relapse. It is noteworthy that the Dicer gene is located at the subtelomeric region 14q32.13 which carries a cluster of imprinted genes critically affected by various deletions, rearrangements and epimutations that might potentially influence methylation status of this region during tumour progression (26). It was recently observed from gene array analysis data that Dicer mRNA expression was low in the more aggressive basal-like, Her2+, luminal B types and in ER- breast tumour samples (27). We did not find such association in our RT-PCR analysis. In our study, Dicer mRNA expression appeared to be significantly correlated with occurrence of distant metastases, even when after adjusting for other prognostic parameters.

No such association was observed for Dicer protein in our TMA analysis while IHC results showed a significant association with hormone receptor status. It is noteworthy that Dicer protein detected in tissue microarray corresponded to products translated from the three RNA variants a, b and c amplified by RT-PCR but also to the e and f variants. The clinical significance of Dicer protein expression was studied in other cancer pathologies but in ovarian cancer failed to show any relationship with both disease-free and overall patient survival (5). Our results are not in contradiction with what was found in other cancer pathologies.

In breast cancer cell lines, we have found 72% of concordance between mRNA and protein levels. Thus, for 28% of the cell lines there was a discrepancy between RNA and protein levels. Very little is known about post-transcriptional and translational regulation of Dicer. Dicer protein is a part of the RISC loading, small RNA processing complex including its co-factor TRBP (TAR RNA binding protein), PACT and Ago2 which could affect its stability depending on their respective expression in cancer cells.

A number of studies have shown a deregulation of Ago2 expression in cancer cells (7, 27). TRBP was found to be amplified and overexpressed in breast cancers (28). PACT has been shown to be strongly expressed in small adenocarcinomas of the lung and has been shown to contribute to Dicer protein stabilisation (6, 29). Dicer protein was also found to be a target of caspases during apoptosis (30). Surprisingly, a very recent publication suggested a novel regulatory loop where the mature microRNA let-7 processed by Dicer, affects Dicer expression at both protein and mRNA levels (31).

In normal breast tissue, only the normal luminal epithelium expressed Dicer while no expression was found in the cells from normal basal epithelium. Interestingly, in HMECs we found a relatively low Dicer expression and Ince and collaborators have shown that in their HMEC population, the myoepithelial/basal cells were more prominent (32). The data found in lung adenocarcinomas suggested a transient upregulation of Dicer between normal and non invasive stages of carcinogenesis and a significant subsequent decrease in advanced invasive carcinomas (6). It is noteworthy that the expression of Dicer in the breast cellular-model of tumour progression likely favours the hypothesis of a preliminary expansion of the luminal cells (HMEC-hTERT and the HMEC-hTERT + LT cells were mostly of epithelial phenotype), and an EMT in the more advanced stage (HMLER). Furthermore, Dicer expression was significantly lower in the breast cell lines exhibiting a mesenchymal phenotype than the ones with a luminal phenotype. To date, no studies have shown Dicer to be involved in EMT. EMT seems to be a key step in tumour progression towards metastasis (33). Breast cancer progression may involve EMT whereby tumour cells become more invasive and metastatic.

We have found significantly lower Dicer expression in the two bone derivatives of MDA-MB-231 breast tumour cells in comparison with the levels observed in the parental cells. As it was mentioned from the transcriptomic data obtained on SCP cells, the bone-metastatic activity of MDA-MB-213 cells is accompanied by gene regulation independently of their pre-existing poor prognosis signature (12). Epithelial-to-mesenchymal transition has been linked to the acquisition of a bone metastatic phenotype in

breast cancer (34-35). We could ask whether among a mesenchymal cell population, cells which expressed very low levels of Dicer were more prone to bone homing as the two MDA-MB-231 bone derivatives showed lower Dicer expression than parental cells. To our knowledge, this work is the first to examine in parallel the expression of a carcinogenesis-associated candidate gene in primary tumours and in cellular-models of tumour and metastasis progression.

Over the past 5 years, numerous studies have reported widespread alteration in miRNA expression in human cancer cells (36). Despite some miRNA shown to be highly expressed in tumour or metastatic cells (miR-10b, cluster miR 17-92, miR-21) (37-39), a global under expression of miRNA was observed during cancer and metastatic evolution (4, 40). miRNA processing occurs in three steps with Dicer realising the final step by cleaving the pre-miRNA into the active species. In the Dicer knocked-down cancer cells, we have shown a down-regulation of 3 representative mature miRNAs. Of note, inactivation of Dicer expression did not lead to a complete elimination of mature miRNA. As it was recently shown, levels of mature miRNAs could be elevated independently of Dicer activity by Ago2 (41). Kumar et al. (42) have clearly demonstrated in mouse and human tumour cells, that concomitant inhibition of the three main miRNA machinery protagonists Drosha, DGCR8 (Drosha co-factor) and Dicer reduced steady-state miRNA levels and enhanced transformation of these tumour cells. The depletion of the Dicer co-factor PACT, has been shown to result in a strong miRNA processing decrease (29). The study realized by Kumar et al. (42) had nicely demonstrated that defective miRNA processing machinery improves transformation capacities of cancer cells but is not sufficient for de novo transformation which is in accordance with our results on Dicer expression analysis in breast tumours favouring the hypothesis of down-regulation of Dicer as a late event.

Deciphering the relationship between the RISC machinery regulation and the siRNA and miRNA production during tumour progression is a long and “unwinding” road. Because Dicer functions in both siRNA and miRNA pathways we may ask whether de-

regulation of siRNA process may be involved in cancer progression. Dicer is involved in de novo heterochromatin constitution in the pericentromeric region (43-45). In Dicer-conditional knock-out mice, loss of Dicer resulted in increased DNA damage (46) and the timing of locus replication during S-phase has been shown to be very sensitive to the influence of Dicer hemi-depletion (47). It was recently shown that Dicer deletion compromises small RNA expression during X-chromosome inactivation (XCI) (3) and lack of XCI may be a frequent phenomenon in breast tumorigenesis (48).

In conclusion, this study is to our knowledge the first to simultaneously investigate Dicer expression at mRNA and protein levels in breast primary tumours, cell lines and cancer progression cellular-models. mRNA levels but not protein expression could be predictive for clinical outcome.

We have found that Dicer expression is deregulated in breast cancers and its expression could be mandatory of predicting the breast cancers metastatic evolution.

### **Acknowledgments**

We thank Fred Miller for supplying the 67NR model; Joan Massague and Yibin Kang for providing the SCP2 bone derivative clone; Robert Weinberg and Mary Brooks for providing the HMECs transformation model; Philippe Clézardin for providing the BO2 bone derivative clone; Richard Bachelier for providing transcriptomic data on Dicer expression in BO2 and in parental MDA-MB-231 cells; Eric Tabone for providing tumour biopsies; Thérèse Gargi for helping with clinical database; Benjamin Bouchet for technical help; Thomas Bachelot and David Cox for critical review of the manuscript. This research was supported by the Comité Départemental de la Drôme de La ligue Nationale Contre le Cancer; G. G. is a recipient of Ligue National Contre Le Cancer doctoral fellowship.

### **Bibliography**

- 1.Murchison EP, Partridge JF, Tam OH, Cheloufi S, Hannon GJ. Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A* 2005; 102:12135-40.
- 2.Jin Z, Xie T. Dcr-1 maintains Drosophila ovarian stem cells. *Curr Biol* 2007;17:539-44.
- 3.Ogawa Y, Sun BK, Lee JT. Intersection of the RNA interference and X-inactivation pathways. *Science* 2008;320:1336-41.
- 4.Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. *Nature* 2005;435:834-8.
- 5.Flavin RJ, Smyth PC, Finn SP, et al. Altered eIF6 and Dicer expression is associated with clinicopathological features in ovarian serous carcinoma patients. *Mod Pathol* 2008;21:676-84.
- 6.Chiose S, Jelezcova E, Chandran U, et al. Overexpression of Dicer in precursor lesions of lung adenocarcinoma. *Cancer Res* 2007;67:2345-50.

7. Chiosea S, Jelezcova E, Chandran U, et al. Up-regulation of dicer, a component of the MicroRNA machinery, in prostate adenocarcinoma. *Am J Pathol* 2006;169:1812-20.
8. Luck AA, Evans AJ, Green AR, Rakha EA, Paish C, Ellis IO. The influence of basal phenotype on the metastatic pattern of breast cancer. *Clin Oncol (R Coll Radiol)* 2008; 20:40-5.
9. Mundy GR. Metastasis to bone: causes, consequences and therapeutic opportunities. *Nat Rev Cancer* 2002;2:584-93.
10. Miller FR, Miller BE, Heppner GH. Characterization of metastatic heterogeneity among subpopulations of a single mouse mammary tumor: heterogeneity in phenotypic stability. *Invasion Metastasis* 1983;3: 22-31.
11. Elenbaas B, Spirio L, Koerner F, Fleming MD et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev* 2001;15:50-65.
12. Minn AJ, Kang Y, Serganova I, et al. Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. *J Clin Invest* 2005;115:44-55.
13. Garcia T, Jackson A, Bachelier R, et al. A convenient clinically relevant model of human breast cancer bone metastasis. *Clin Exp Metastasis* 2008;25:33-42.
14. Aslakson CJ, Miller FR. Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res* 1992;52:1399-405.
15. Wang Q, Moyret-Lalle C, Couzon F, et al. Alterations of anaphase-promoting complex genes in human colon cancer cells. *Oncogene* 2003;22:1486-90.
16. Goldstein H, *Multilevel Statistical Models*. Arnold Publishers, 2003
17. Sullivan LM, Dukes KA, Losina E. Tutorial in biostatistics. An introduction to hierarchical linear modelling. *Stat Med* 1999;18:855-88.
18. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53:457-81.
19. Peto R, Pike MC, Armitage P, et al: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer* 1977;35:1-39.
20. Ethier SP. Human breast cancer cell lines as models of growth regulation and disease progression. *J Mammary Gland Biol Neoplasia* 1996;1:111-21.
21. Morel AP, Lièvre M, Thomas C, Hinkal G, Ansieau S, Puisieux A. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE* 2008;3:e288.
22. Lou Y, Preobrazhenska O, Auf dem Keller U, et al. Epithelial-mesenchymal transition (EMT) is not sufficient for spontaneous murine breast cancer metastasis. *Dev Dyn* 2008;237:2755-68.
23. Tester AM, Ruangpanit N, Anderson RL, Thompson EW. MMP-9 secretion and MMP-2 activation distinguish invasive and metastatic sublines of a mouse mammary carcinoma system showing epithelial-mesenchymal transition traits. *Clin Exp Metastasis* 2001;18:553-60.
24. Bellahcène A, Bachelier R, Detry C, Lidereau R, Clézardin P, Castronovo V. Transcriptome analysis reveals an osteoblast-like phenotype for human osteotropic breast cancer cells. *Breast Cancer Res Treat* 2007;101:135-48.
25. Iorio MV, Ferracin M, Liu CG, et al. MicroRNA gene expression deregulation in human breast cancer. *Cancer Res* 2005;65:7065-70.
26. Kagami M, Sekita Y, Nishimura G, et al. Deletions and epimutations affecting the human 14q32.2 imprinted region in individuals with paternal and maternal up(14)-like phenotypes. *Nat Genet* 2008;40:237-

42.

27. Blenkinson C, Goldstein LD, Thorne NP, et al. MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biol* 2007;8:R214.

28. Zhang H, Liao L, Kuang SQ, Xu J. Spatial distribution of the messenger ribonucleic acid and protein of the nuclear receptor coactivator, amplified in breast cancer-3, in mice. *Endocrinology* 2003;144:1435-43.

29. Lee Y, Hur I, Park SY, Kim YK, Suh MR, Kim VN. The role of PACT in the RNA silencing pathway. *EMBO J* 2006;25:522-32.

30. Matskevich AA, Moelling K. Stimuli-dependent cleavage of Dicer during apoptosis. *Biochem J* 2008;412:527-34.

31. Tokumaru S, Suzuki M, Yamada H, Nagino M, Takahashi T. let-7 regulates Dicer expression and constitutes a negative feedback loop. *Carcinogenesis* 2008;29:2073-7.

32. Ince TA, Richardson AL, Bell GW, et al. Transformation of different human breast epithelial cell types leads to distinct tumor phenotypes. *Cancer Cell* 2007;12:160-70.

33. Nguyen DX, Massagué J. Genetic determinants of cancer metastasis. *Nat Rev Genet* 2007;8:341-52.

34. Buijs JT, Henriquez NV, van Overveld PG, et al. Bone morphogenetic protein 7 in the development and treatment of bone metastases from breast cancer. *Cancer Res* 2007;67:8742-51.

35. Buijs JT, van der Pluijm G. Osteotropic cancers: From primary tumor to bone. *Cancer Lett* 2008; epub 14 Jul

36. Takamizawa J, Konishi H, Yanagisawa K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* 2004;64:3753-6.

37. Ma L, Weinberg RA. Micromanagers of malignancy: role of microRNAs in regulating metastasis. *Trends Genet* 2008;24:448-56.

38. Soifer HS, Rossi JJ, Saetrom P. MicroRNAs in disease and potential therapeutic applications. *Mol Ther* 2007;15:2070-9.

39. Wiemer EA. The role of microRNAs in cancer: no small matter. *Eur. J Cancer* 2007;43:1529-44.

40. Tavazoie SF, Alarcón C, Oskarsson T, et al. Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* 2008;451:147-52.

41. Diederichs S, Jung S, Rothenberg SM, Smolen GA, Mlody BG, Haber DA. Coexpression of Argonaute-2 enhances RNA interference toward perfect match binding sites. *Proc Natl Acad Sci U S A* 2008;105:9284-9.

42. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* 2007;39:673-7.

43. Verdell A, Jia S, Gerber S, et al. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 2004;303:672-6.

44. Iida T, Nakayama J, Moazed D. siRNA-mediated heterochromatin establishment requires HP1 and is associated with antisense transcription. *Mol Cell* 2008;31:178-89.

45. Folco HD, Pidoux AL, Urano T, Allshire RC. Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* 2008;319:94-7.

46. Mudhasani R, Zhu Z, Hutvagner G, et al. Loss of miRNA biogenesis induces p19Arf-p53 signaling and senescence in primary cells. *J Cell Biol* 2008;181:1055-63.

47. Jørgensen HF, Azuara V, Amoils S, et al. The impact of chromatin modifiers on the timing of locus replication in mouse embryonic stem cells. *Genome Biol* 2007;8:R169.

48. Sirchia SM, Ramoscelli L, Grati FR, Barbera F, Coradini D, Rossella F, Porta G, Lesma E, Ruggeri A, Radice P, Simoni G, Miozzo M. Loss of the inactive X chromosome and replication of the active X in BRCA1-defective and wild-type breast cancer cells. *Cancer Res* 2005;65:2139-46.

## **Supplementary Materials and Methods:**

### *Steroid receptor assay*

ER and PR status were carried out by established immunohistochemical (IHC) method of avidin-biotin-peroxidase complex (ABC) described previously ( ) using a mouse monoclonal anti-ER $\alpha$  antibody (clone 1D5, Dako) and a mouse monoclonal anti-PRA (clone PR 636, Dako). As cut-off point for positivity  $\geq 10\%$  frequency of positive stained tumor cells was used.

### *Primers sequences and real time PCR experimental procedures*

Dicer specific primer pair sequences used in RT-PCR amplification was: forward 5'TTAAC- CTTTTGGTGTGGTGGATGAGTGT and reverse 5'GCGAGGACATGATGGA-CAATT. This primers set allow to detect the three main alternatively spliced variants. Cyclophilin B (CPB),  $\beta$ -actin and phosphoglycerokinase (PGK) was used as reference genes for the results normalisation. The primer pairs and probes were designed with Universal Probe Library website Primers of each pair were located in different exons to avoid genomic amplification. All reactions were performed using LightCycler FastStart DNA Master PLUS SYBR Green I kit for Dicer or LightCycler Taqman Master for reference genes (Roche Applied Science). The reaction mixture contained FastStart TaqDNA polymerase, reaction buffer, deoxynucleoside triphosphate mixture, MgCl<sub>2</sub>, 0.25  $\mu$ M of each primer, and 1.67  $\mu$ l (5  $\mu$ L for reference genes) of a 1/60 dilution of first-strand cDNA from cells. The final reaction volume was 6.67  $\mu$ l (10  $\mu$ L for reference genes). Thermocycling conditions for dicer gene were designed in four consecutive steps (the Roche Applied Science Universal Probe Library's thermocycling conditions were used for the reference genes). Step 1: denaturation and polymerase activation by

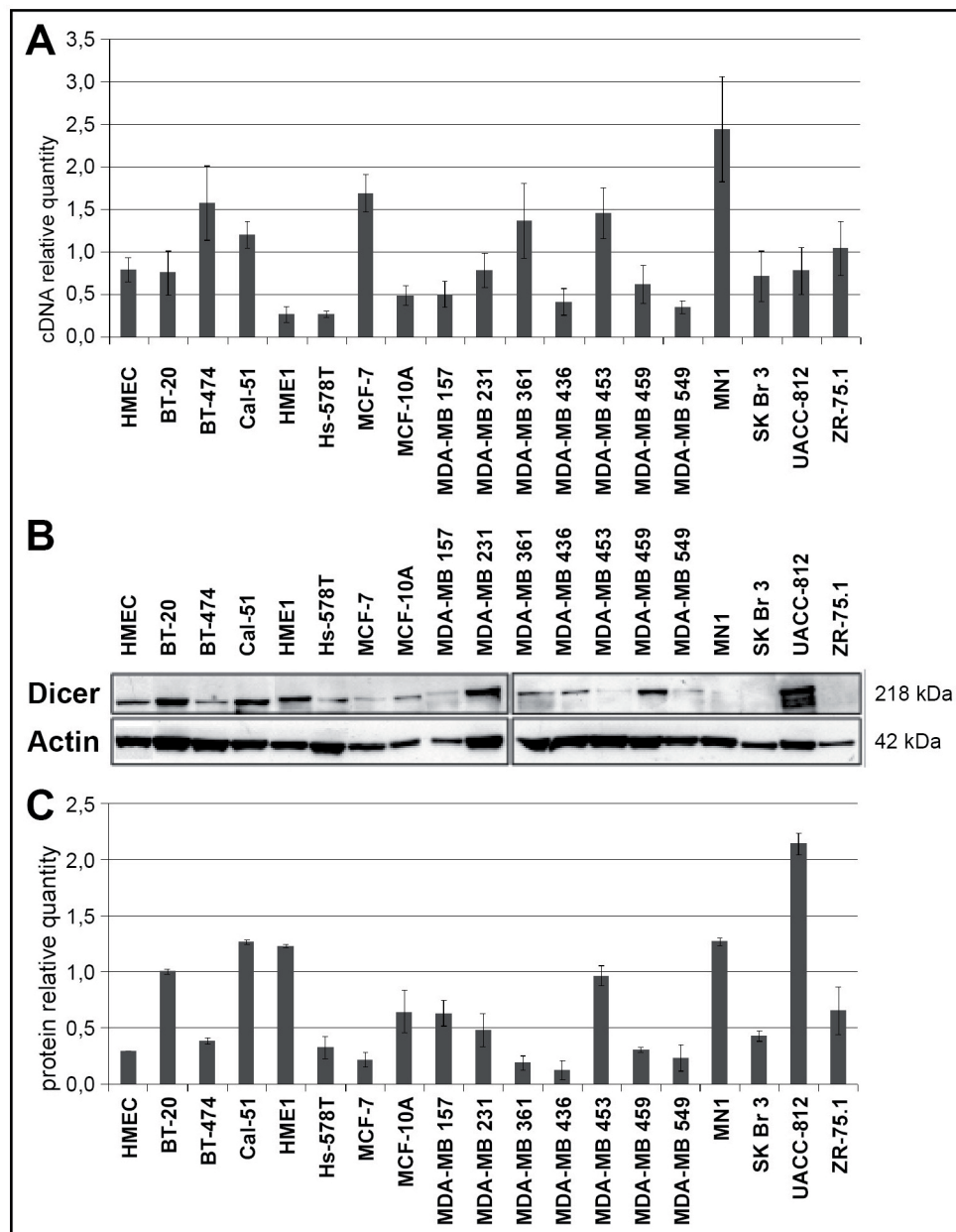
10 min heating at 95°C prior amplification. Step 2 (45 cycles): denaturation for 10 sec at 95°C, annealing for 10 sec at 55°C and extension for 6 sec at 72°C. Step 3: melting which consisted of incrementing the temperature from 45°C to 95°C by steps of 0.1°C. Step 4: cooling for 30 sec at 40°C for one cycle. This melting curve program was used to check the specificity of each dicer RT-PCR. All standards or samples were analyzed in duplicate and each experiment was repeated at least three times. All fluorescence data were analyzed by the LightCycler 4.0 software (Roche) and Ct results were exported to Excel sheets.

A large amount of cDNA was prepared from the HME-1 cell line prior to the experiment. It was then diluted, aliquoted and used as a calibrator for all RT-PCR runs. For relative quantification and normalization, the comparative Ct (or  $\text{Eff}^{\Delta\Delta\text{C}}$ ) method was used.

#### *Breast Cancer cell lines*

The following 21 breast cancer cell lines were obtained from the American Type Culture Collection<sup>9</sup>: BT20, BT474, Cal51, Cama1, HBL100, HCC1937, HS578T, MCF10A, MCF12A, MCF7, MDAMB157, MDAMB231, MDAMB361, MDAMB436, MDAMB453, MDAMB459, MDAMB549, SKBr3, T47D, UACC812, ZR75.1. Cells grown according to American Type Culture Collection protocols to 90% confluency were collected for total RNA and protein preparation. HMECs and HME-1 (Cambrex, Clonetics) were cultured in MEGM at 37°C with CO<sub>2</sub> as described (Stampfer 1985). Breast cancer cell lines clinical and biological characteristics were described in Neve et al., (2006)( Neve RM, Chin K, Fridlyand J, Yeh J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell*. 2006 Dec; 10 (6):515-27.)



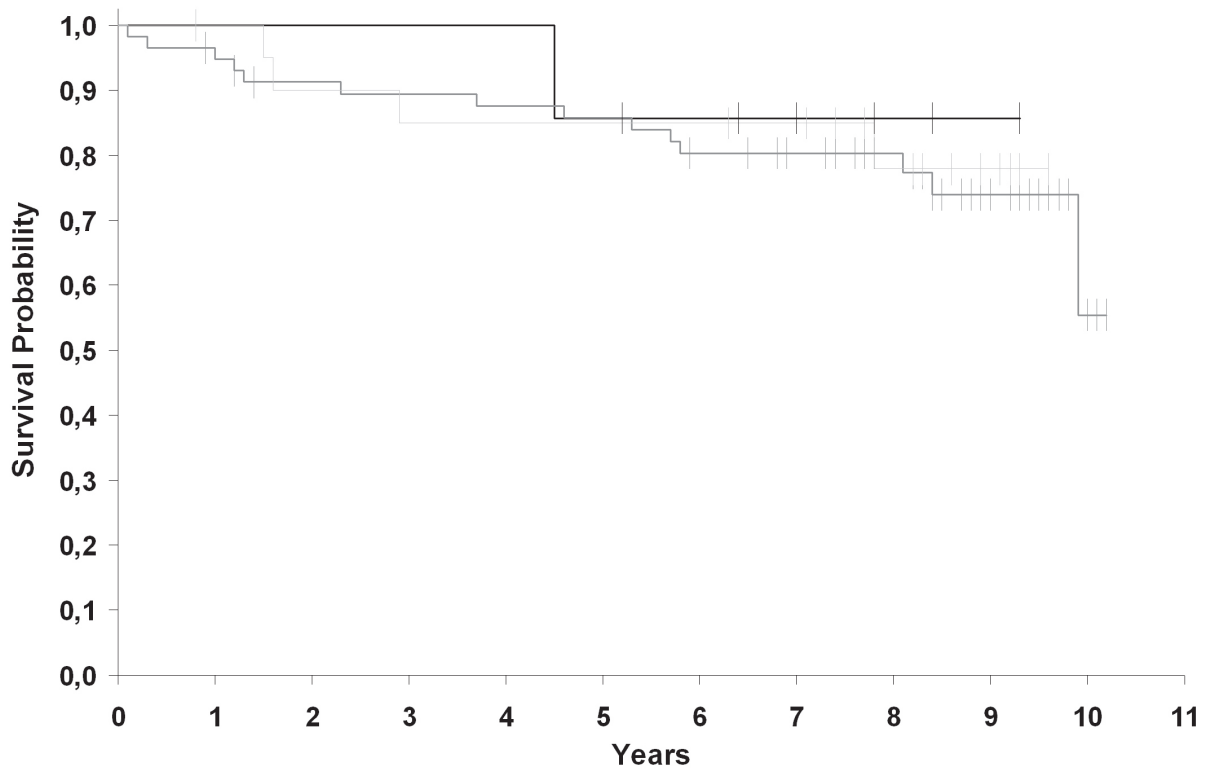


**Figure S1: Comparative analyses of Dicer expression in normal HMECs and in breast cancer cell lines**

**A.** Total RNA was extracted and Dicer mRNA relative levels were studied by real-time RT-PCR, and each bar represents the mean  $\pm$  SD of the PCRs in triplicate.

**B.** Western blot analysis was done to investigate Dicer protein levels. The signal intensity of Dicer was normalized to that of actin.

**C.** Dicer protein levels were quantified using Quantity One software (BioRad) and expressed as protein relative quantity. The ratios of Dicer/actin of three independent studies were expressed as mean  $\pm$  SD.



**Figure S2: Kaplan-Meier estimates of metastatic free survival by dicer's intensity among the TMA population**

— Intensity 0 (1/7 metastatic progressions; 8-years event free survival probability 85.7% (95%CI 33.4-97.9);

— Intensity 1 (14/58 metastatic progressions; 8-years event free survival probability 80.3% (95%CI 67.2-88.6);

— Intensity 2 (4/21 metastatic progressions; 8-years event free survival probability 77.9% (95%CI 50.4-91.3);

Log-rank test : P = 0.868

**Table S1** Quantitative real-time RT-PCR analysis of *dicer* mRNA expression in breast cancer patients samples and correlation with various clinicopathological characteristics

Tumour Characteristics		Case s n=10 9	Measure s n=410	<i>dicer</i>	
				Mean (95%IC)	P-value
<i>Tissue type</i>					
	normal	2		5.996 (NA)	0.2825
	tumour	111		4.242 (3.809-4.725)	
<i>Menopause</i>					
	yes	62	222	8.641 (7.002-10.665)	0.1029
	no	36	131	6.598 (5.161-8.8435)	
<i>Grade</i>					
	1	12	37	7.100 (4.669-10.827)	0.6177
	2	42	147	8.647 (6.713-11.139)	
	3	43	170	7.527 (5.907-9.592)	
<i>Tumour size</i>					
	< 20	14	51	6.741 (4.394-10.343)	0.6818
	20-50	56	197	7.668 (6.248-9.410)	
	>50	37	118	8.979 (6.621-12.176)	
<i>Histological type</i>					
	ductal	94	344	7.584 (6.448-8.920)	0.2239
	lobular	4	16	13.895 (NA)	
	other	4	9	10.727(NA)	
<i>N stage</i>					
	N+	87	320	7.353 (6.254-8.645)	0.0204
	N0	16	51	12.852 (8.106-20.376)	
<i>Metastatic at diagnosis</i>					
	yes	15	36	10.629 (7.451-15.163)	0.1545
	no	88	332	7.572 (6.394-8.966)	
<i>Estrogen Receptor</i>					
	ER+	79	284	8.417 (6.94-10.205)	0.1002
	ER-	21	78	6.094 (4.842-7.669)	
<i>Progesterone Receptor</i>					
	PR+	77	273	7.750 (6.468-9.285)	0.7862
	PR-	25	95	8.182 (5.966-11.223)	

**Table S2** Association between clinical data and % of Dicer marked cells among the TMA population

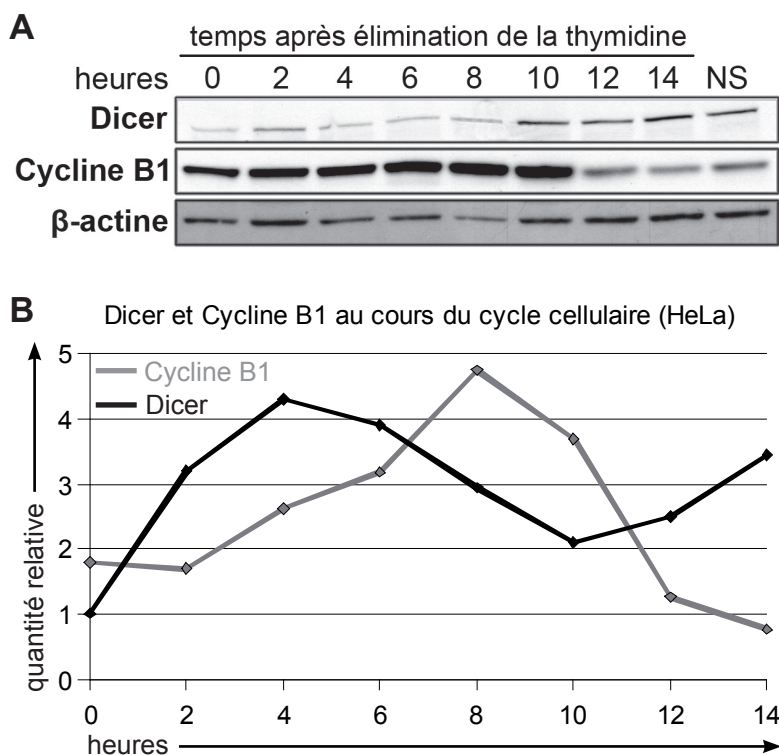
Mean ± SD (median [min,max]) (%)	Intensity 0 (n=7)	Intensity 1 (n=58)	Intensity 2 (n=21)	P-value 0/1 vs 2
Age				0.580
< 50 years	0 (0.0%)	17 (70.8%)	7 (29.2%)	
≥ 50 years	7 (11.3%)	41 (66.1%)	14 (22.6%)	
Menopausal status :				1.000
Yes	7 (12.7%)	34 (61.8%)	14 (25.5%)	
No	0 (0.0%)	20 (76.9%)	6 (23.1%)	
pT				0.361
≤ 20 mm	3 (7.1%)	26 (61.9%)	13 (31.0%)	
[20 – 50] mm	3 (7.7%)	29 (74.4%)	7 (18.0%)	
> 50 mm	1 (20.0%)	3 (60.0%)	1 (20.0%)	
Histological type:				0.347
Ductal	4 (5.6%)	48 (67.6%)	19 (26.8%)	
Lobular	3 (42.9%)	4 (57.1%)	0 (0.0%)	
Mixed	0 (0.0%)	2 (100.0%)	0 (0.0%)	
Others	0 (0.0%)	4 (66.7%)	2 (33.3%)	
Histological grade (SBR):				0.266
1	1(7.1%)	10(71.4%)	3 (21.4%)	
2	4 (11.1%)	26 (72.2%)	6 (16.7%)	
3	2 (5.6%)	22 (61.1%)	12 (33.3%)	
N status :				0.381
N0	5 (7.7%)	46 (70.8%)	14 (21.5%)	
N1	2 (9.5%)	12 (57.1%)	7 (33.3%)	
M status :				0.187
M0	6 (7.6%)	52 (65.8%)	21 (26.6%)	
M1	1 (14.3%)	6 (85.7%)	0 (1.1%)	
Estrogen receptor :				0.008
% marked cells (count):				
< 10%	0 (0.0%)	7 (46.7%)	8 (53.3%)	
≥ 10%	7 (9.9%)	51 (71.8%)	13 (18.3%)	
Progesterone receptor :				0.019
% marked cells (count)				
< 10%	0 (0.0%)	12 (54.6%)	10 (45.5%)	
≥ 10%	7 (10.9%)	46 (71.9%)	11 (17.2%)	
HER2 status				0.597
+	6 (7.7%)	53 (68.0%)	19 (24.4%)	
+++	0 (0.0%)	3 (60.0%)	2 (40.0%)	
Triple negative				0.036
Yes	0 (0.0%)	7 (50.0%)	7 (50.0%)	
No	7 (9.7%)	51 (70.8%)	14 (17.4%)	
Dicer				0.013
% marked cells (count)				
≤ 60%	7 (15.6%)	22 (48.9%)	16 (35.6%)	
> 60%	0 (0.0%)	36 (87.8%)	5 (12.2%)	

II.2. Influence de *dicer* sur le déroulement du cycle cellulaire

D'après les données bibliographiques, notamment chez *S. pombe*, il semble probable que *dicer* soit régulé au cours du cycle cellulaire. C'est pourquoi il nous a semblé important d'étudier l'expression de *dicer* dans des cellules humaines et d'analyser le cycle cellulaire lors de l'inactivation de notre gène d'intérêt.

II.2.1. *dicer* est régulé au cours du cycle cellulaire

La synchronisation des cellules HeLa est obtenue par un double blocage à la thymidine. Les cellules sont alors en fin de phase G1 et début de phase S avant d'être relancées dans le cycle par l'ajout de déoxycytidine. Des culots de cellules sont préparés toutes les deux heures pendant quatorze heures afin de permettre aux cellules de réaliser un cycle complet. L'analyse du déroulement du cycle est faite par western blot grâce à un anticorps anti-cycline B1 (Fig. 17.A). Cette cycline s'accumule pendant la phase mitotique du cycle cellulaire. On observe ainsi, après estimation de



**Figure 17 : Expression de Dicer au cours du cycle cellulaire**

**A.** Analyse par western blot d'extractions de protéines totales de cellules HeLa avec les anticorps anti-Dicer et anti-Cycline B1 à différents temps après reprise du cycle cellulaire (le temps 0h correspond au blocage des cellules en fin de phase G1 et début de phase S). Des quantités égales de protéines ont été déposées sur gel gradient Tris-Glycine 4-12% et le chargement confirmé par un anticorps anti-béta-actine.

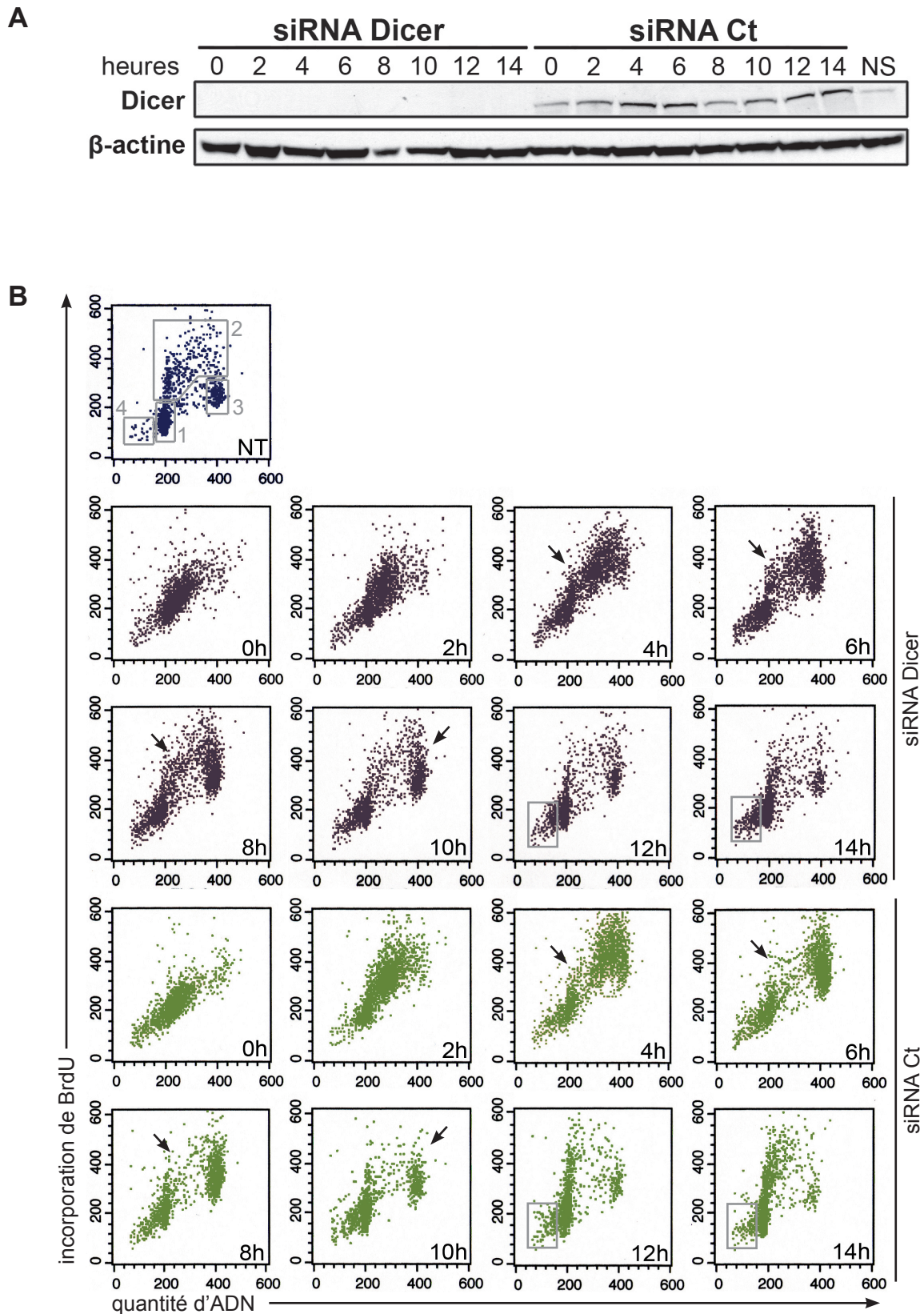
**B.** Estimation de la quantité relative au cours de la cinétique des protéines Cycline B1 (courbe grise) et Dicer (courbe noire) par le logiciel Quantity One (BioRad).

la quantité relative de protéine par rapport à la bêta-actine, que les cellules sont en mitose au temps 8h (**Fig. 17.B**). De même, après estimation de la quantité relative de protéines Dicer, on peut mettre en évidence une augmentation progressive au cours de la phase S (temps 0h à 4h) avant sa diminution en phase G2 et M (temps 4h à 10h). Le taux d'expression de Dicer augmente de nouveau des temps 10h à 14h avec le début d'un nouveau cycle (**Fig. 17**).

Ainsi, Dicer est régulé au cours du cycle cellulaire dans les cellules HeLa avec un pic d'expression en fin de phase S.

### II.2.2. L'inactivation de *dicer* dans les cellules HeLa retarde le cycle cellulaire

**A**fin d'étudier l'implication de *dicer* dans le cycle cellulaire, nous l'avons inactivé transitoirement par iARN avant de synchroniser les cellules (HeLa) par double blocage à la thymidine comme précédemment. Nous avons étudié alors deux populations de cellules : une population transfectée par un duplex siRNA#1 ciblant l'ARNm de *dicer* et une population transfectée par le contrôle négatif spécifique de ce duplex (même séquence mais dans un ordre différent), siARNCt#1 ne ciblant aucun ARNm dans les cellules. Les échantillons ont été prélevés toutes les deux heures pendant quatorze heures et analysés par cytométrie de flux (**Fig. 18.B**). L'inactivation de l'expression de *dicer* est vérifiée par western blot (**Fig. 18.A**). Sur le diagramme représentant la répartition des cellules non traitées (NT), on retrouve les différentes phase du cycle cellulaire (**Fig. 18.B**) : 1-phases G0/G1 ; 2-phase S et 3-phases G2/M ainsi que la population en sub-G1 (4). Lorsque l'on compare la progression dans le cycle cellulaire de la population où *dicer* est inactivé par rapport à la population contrôle, on observe dès le temps 4h un retard des cellules n'exprimant pas notre gène d'intérêt (flèches). Ce retard est de plus en plus marqué pour les temps qui suivent (flèches). A 10h, il reste encore des cellules HeLa siRNA#1 en fin de phase S alors que les cellules



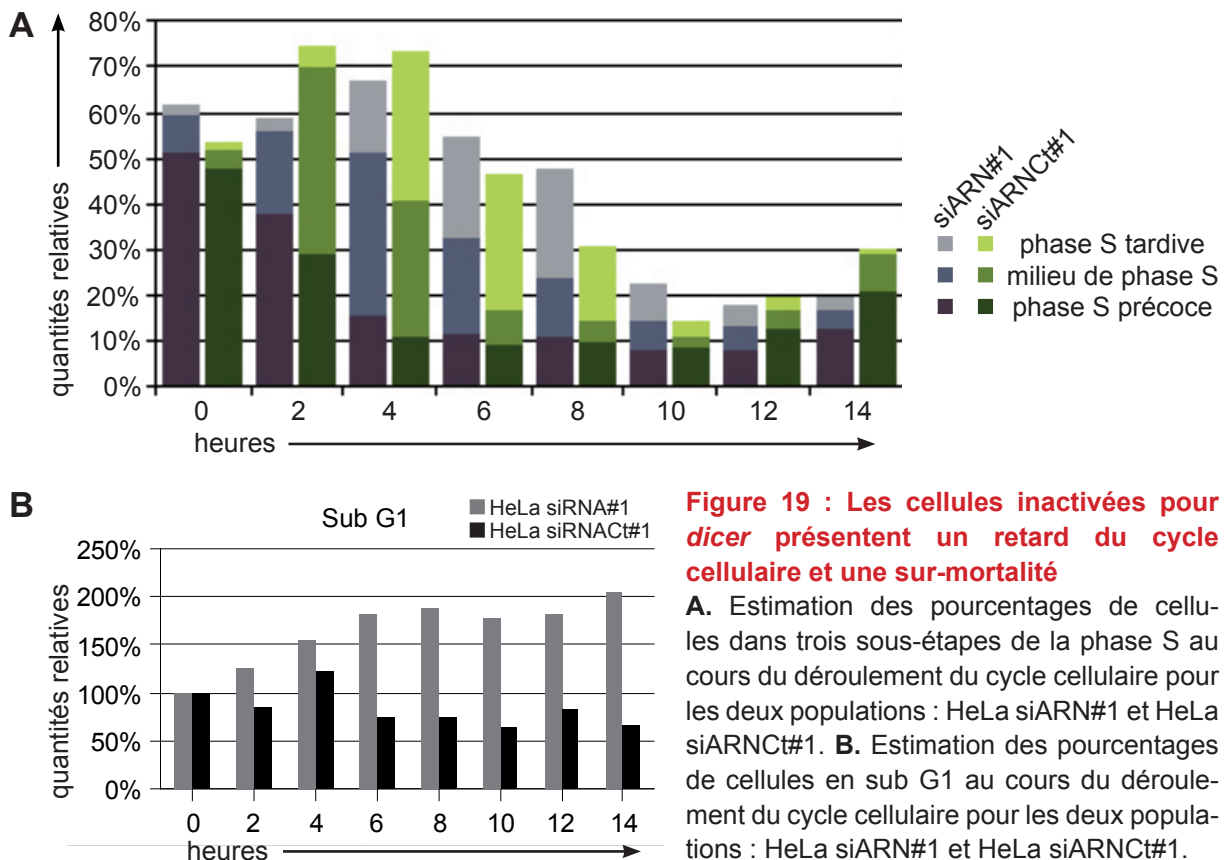
**Figure 18 : Les cellules inactivées pour *dicer* présentent un retard du cycle cellulaire**

**A.** Vérification de l'inactivation de l'expression de *dicer* par western blot. **B.** Distribution du cycle cellulaire des cellules HeLa non traitées (NT) ou après transfection par le siRNA#1 anti-*dicer* ou son contrôle siARNCt#1 et synchronisation par double blocage à la thymidine levé au temps 0h.

contrôles sont en phases G2/M pour une minorité et en phases G0/G1 pour la majorité (flèches). L'estimation des pourcentages de cellules dans la phase S artificiellement découpée en trois parties (précoce, milieu et tardive) confirme ces observations (Fig. 19.A). Par exemple, il y a au temps 4h 32% de cellules contrôles en phase S tardive contre 15% pour les cellules inactivées.

Il apparaît donc que l'inactivation de *dicer* dans des cellules HeLa synchronisées entraîne un retard dans le déroulement du cycle cellulaire.

De plus, on peut voir aux temps 12 et 14h que la mortalité cellulaire entre la population n'exprimant pas *dicer* et la population contrôle est différente, avec une surmortalité dans la première (rectangles, Fig. 18.B). On observe aussi qu'une grande partie des cellules contrôles débutent un nouveau cycle cellulaire à ces deux temps alors que ce n'est pas le cas pour les cellules inactivées pour *dicer* (flèches). L'estimation des pourcentages de cellules concernées (Fig. 19.B) confirment ces observations et montrent qu'il y a environ deux fois plus de mortalité cellulaire dans les HeLa siARN#1 des temps 6h à 14h.

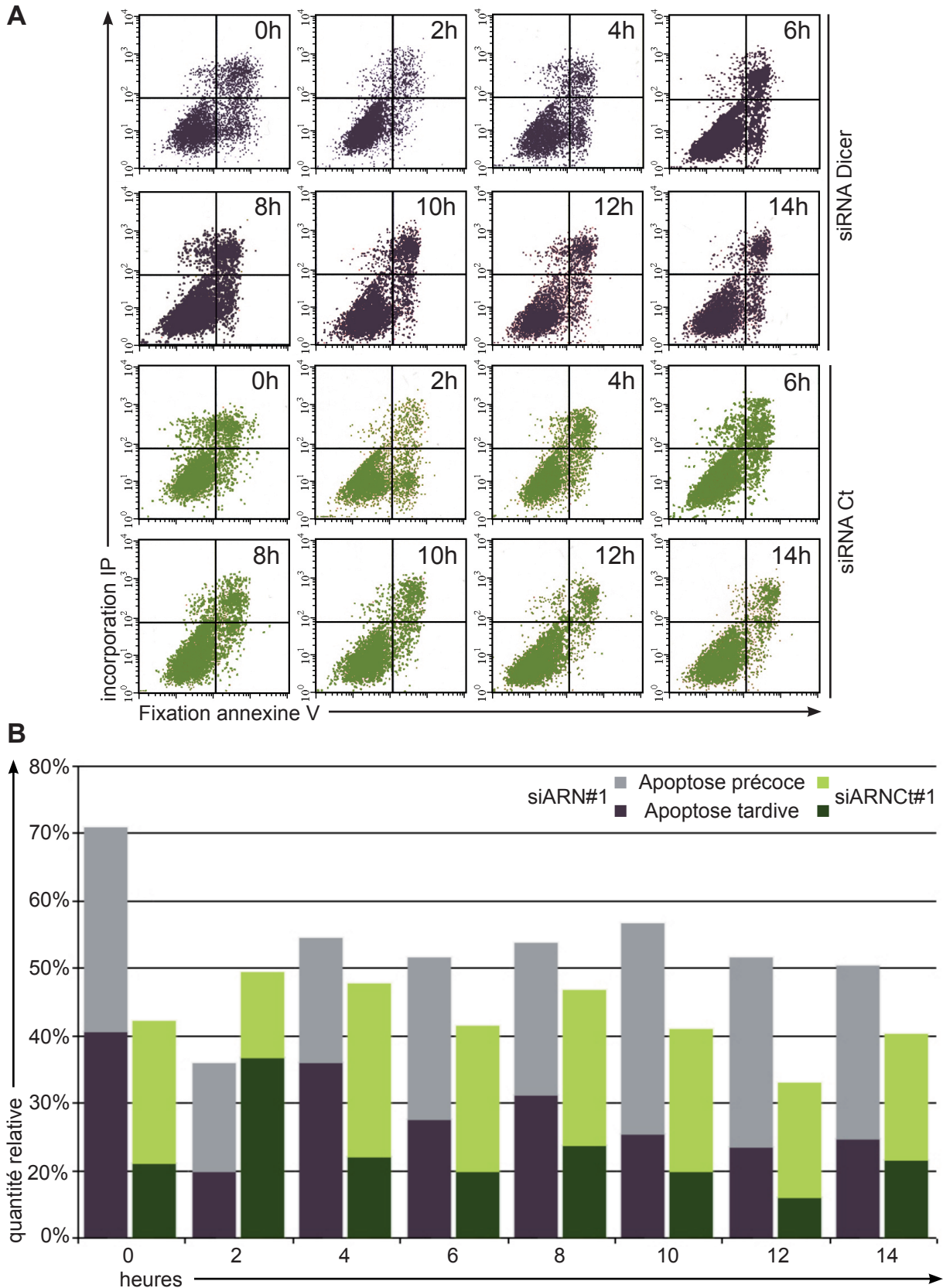




Pour déterminer plus précisément comment ces cellules meurent, nous avons utilisé un test annexine V. Ce test, réalisé en cytométrie de flux, permet de détecter un événement précoce de l'apoptose qui est l'externalisation des phosphatidylserines de la membrane interne des cellules. Ces phospholipides sont alors accessibles à l'annexine V, une protéine calcium dépendante capable de s'y fixer spécifiquement et avec une forte affinité. Les résultats obtenus sont présentés sur les diagrammes de la **figure 20.A**. Chacune des quatre fenêtres représente une population cellulaire dont les caractéristiques sont les suivantes : en bas à gauche les cellules vivantes (pas de fixation d'annexine V ni d'incorporation de l'iodure de Propidium, IP), en bas à droite les cellules en apoptose précoce (fixation d'annexine V mais pas d'incorporation de IP), en haut à droite les cellules en apoptose tardive (fixation d'annexine V et incorporation de IP) et en haut à gauche les cellules mortes autrement que par apoptose. Les résultats montrent une légère surmortalité pour la population de cellules n'exprimant pas *dicer* avec une différence plus significative aux temps 10 et 12h (**Fig. 20.B**). Les populations de cellules en sub-G1 observées lors de la synchronisation sont donc des cellules apoptotiques ce qui laisse supposer que des dommages importants de l'ADN sont survenus au cours du cycle sans qu'il soit possible de les résoudre.

### II.2.3. Inactivation de DGCR8 et cycle cellulaire

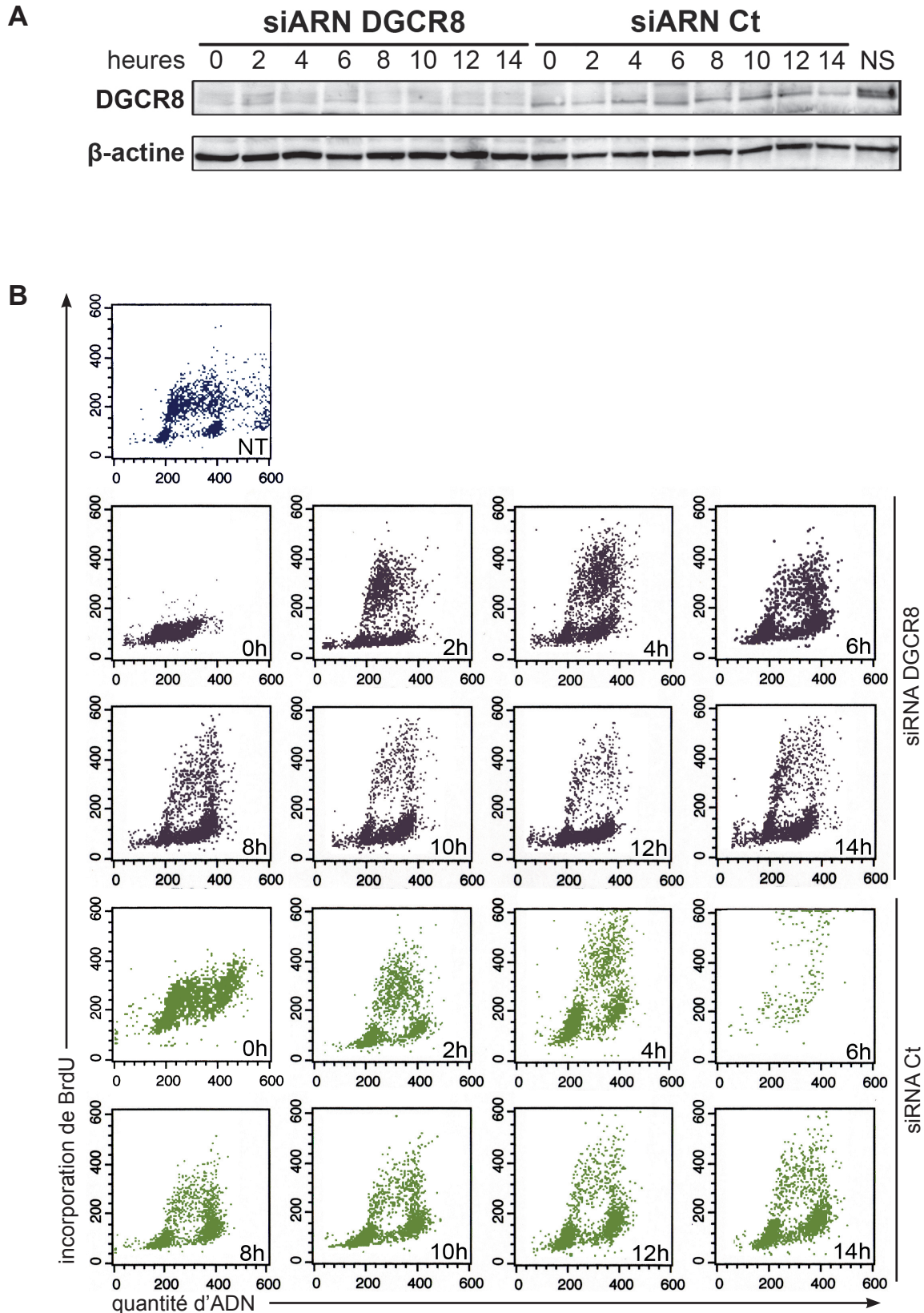
**C**omme *dicer* intervient à la fois dans la voie des siARN et des miARN, il est difficile de savoir si les conséquences observées lors de son inactivation sont principalement ou spécifiquement dues à l'une ou l'autre des voies. Cependant, il est possible de réprimer uniquement la voie des miARN en empêchant la maturation des miARN à un stade plus précoce. Ainsi, nous avons utilisé un siARN ciblant le gène DGCR8 du microprocesseur qui permet le passage des pri-miARN en pre-miARN. De même que pour les cellules HeLa siARN *dicer*, une synchronisation par double blocage



**Figure 20 : Analyse de la mortalité des cellules inactivées pour *dicer* par un test annexine V**

**A.** Distribution des cellules HeLa siARN#1 et siARN Ct#1 suivant la quantité de Iodure de Propidium incorporée (axe des ordonnées) et la quantité d'annexine V fixée (axe des abscisses). Les quatre cadrans définissent des populations de cellules aux caractéristiques différentes : en bas à gauche les cellules vivantes (IP et annexine V négatives), en bas à droite les cellules en apoptose précoce (annexine V positives et IP négatives), en haut à droite les cellules en apoptose tardive (IP et annexine V positives).

**B.** Estimation des pourcentages de cellules en apoptoses précoce et tardive avec ou sans inactivation de *dicer*.



**Figure 21 : Les cellules inactivées pour *DGCR8* ne présentent pas d'altération du cycle cellulaire**

**A.** Vérification de l'inactivation de l'expression de *DGCR8* par western blot. **B.** Distribution du cycle cellulaire des cellules HeLa non traitées (NT) ou après transfection par le siARN anti-*DGCR8* ou le contrôle siARN Ct et synchronisation par double blocage à la thymidine levé au temps 0h.

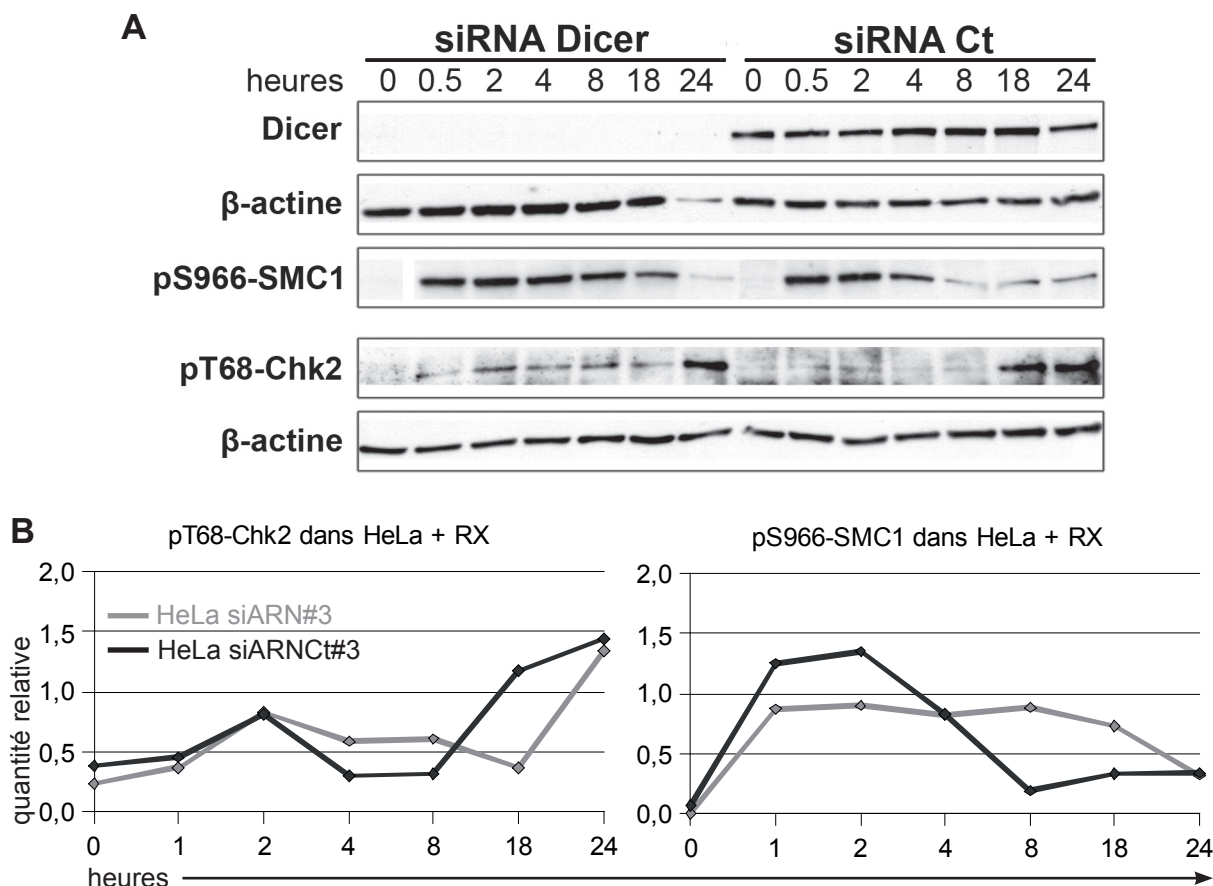
à la thymidine de cellules HeLa a été réalisée avec ce siARN DGCR8. Dans ce cas, la transfection contrôle utilise un siARN “scramble” qui ne cible aucune séquence de la cellule. Le western blot réalisé sur les échantillons montrent que la protéine DGCR8 est très peu, voire pas exprimée dans les cellules inactivées (Fig 21.A.). L'analyse du déroulement du cycle cellulaire par cytométrie en flux est présentée en figure 21.B. et montre un problème d'incorporation du BrdU dans les échantillons transfectés par le siARN DGCR8 (la forme en “fer à cheval” est fermée). Cependant, il semble que le cycle se déroule convenablement. L'échantillon contrôle à 6h n'a pu être analysé correctement par manque de cellules.

### II.3. Implication de *dicer* dans les voies de réponses aux cassures double-brin

Le retard dans le déroulement de la phase S et l'augmentation de l'apoptose dans les cellules HeLa inactivée pour *dicer* laisse penser que la réplication de l'ADN est perturbée dans ces cellules. Comme l'hydroxyurée utilisé pour synchroniser les cellules crée un léger stress réplcatif, nous nous sommes demandé si les cellules n'exprimant pas *dicer* ne présentaient pas un défaut de réponse à ce stress. Il nous a donc semblé intéressant d'étudier les voies de réponse aux cassures double-brin après irradiation ou hydroxyurée.

#### II.3.1. Rôle de *dicer* dans la réponse aux irradiations

Les cellules HeLa inactivées pour *dicer* ou non sont irradiées par rayons X (RX) à 10 grey afin de générer des cassures double-brins. La voie de réponse engagée pour réparer ces cassures est celle médiée par ATM. Il en résulte



**Figure 22 : Analyse de l'expression de SMC1 et Chk2 après rayons X en cellules HeLa**

**A.** Analyse par western blot d'extractions de protéines totales de cellules HeLa avec les anticorps anti-Dicer, anti-pS966-SMC1 et anti-pT68-Chk2 à différents temps de traitement aux rayons X (l'irradiation est réalisée à temps -1h). Des quantités égales de protéines ont été déposées sur gel gradient Tris-Glycine 4-12% et le chargement confirmé par un anticorps anti-béta-actine. **B.** Estimation de la quantité relative au cours de la cinétique des protéines pT68-Chk2 et pS966-SMC1 dans les cellules HeLa siARN#3 (courbe grise) et HeLa siARNct#3 (courbe noire) par le logiciel Quantity One (BioRad).

une cascade de phosphorylation d'un nombre important de protéines dont notamment Chk2 et SMC1 auxquelles nous nous sommes intéressés par western blot (Fig. 22.A). Les résultats obtenus montrent qu'il n'y a pas de différence significative entre les cellules HeLa siARN#3 et siARNct#3 dans leur réponse à ce stress (Fig. 22.B). Ainsi, l'absence d'expression de *dicer* ne perturbe pas la réponse médiée par ATM.

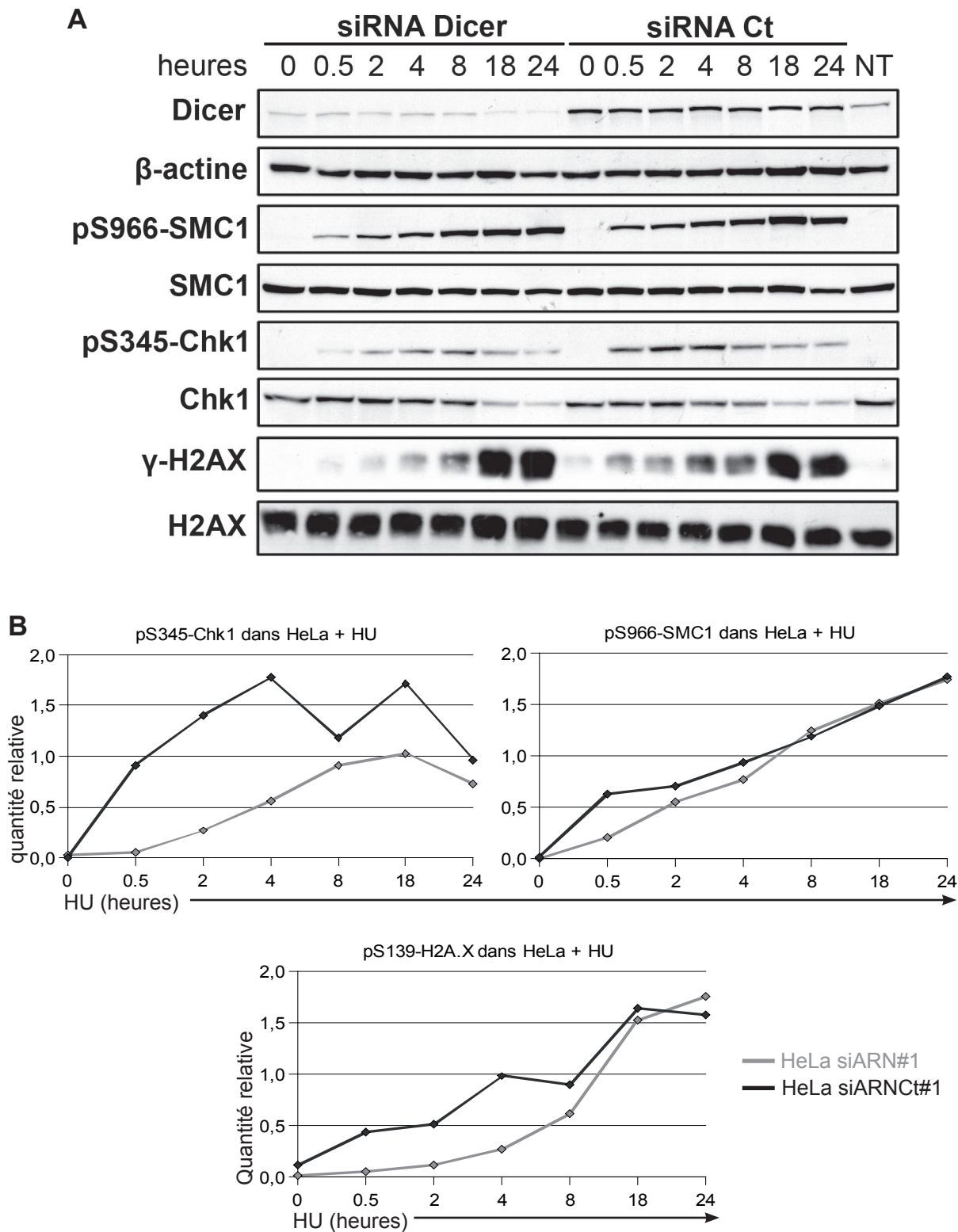
### II.3.2. Rôle de dicer dans la réponse à l'hydroxyurée

**E**n empêchant le fonctionnement de la ribonucléotide réductase et donc la synthèse des déoxynucléotides, l'hydroxyurée induit une diminution de la

réplication de l'ADN et un stress générant des cassures. Il en résulte ainsi une activation du point de contrôle répliatif des cellules en cours de phase S. Si les cassures peuvent être réparées par l'intermédiaire, dans ce cas, de la voie de réponse médiée par ATR, la phase S reprend, sinon les cellules partent en apoptose. Nous avons tout d'abord traité des cellules HeLa avec de l'hydroxyurée afin de comparer les réponses, d'une part de cellules n'exprimant pas *dicer* et d'autre part de cellules contrôles. Nous avons ensuite réalisé cette expérience dans deux lignées tumorales mammaires qui diffèrent principalement par leur type : T47D est une lignée de type épithélial alors que MDA-MB 231 est une lignée de type mésenchymateux. ATR étant la principale protéine induite par l'hydroxyurée, nous nous sommes intéressés, par western blot, à deux de ses cibles : pS345-Chk1 et pS966-SMC1. La cinétique de récupération se déroule sur 24h avec différents temps d'observation.

#### i. cellules tumorales HeLa

Les cellules ont été transfectées avec soit le siARN#1 anti-*dicer* soit le siARNct#1 contrôle, puis traitées à l'hydroxyurée. La phosphorylation de l'histone H2A.X qui intervient dans le signalment des cassures donne une estimation de leur quantité (Fig. 23.A). On observe que la phosphorylation de SMC1 n'est pas perturbée par l'inactivation de *dicer*. Par contre, il apparaît que celle de Chk1 est retardée avec une augmentation seulement au temps 1h dans les cellules inactivées par rapport au temps 0h dans les contrôles (Fig. 23.B). On remarque aussi que la phosphorylation de la protéine H2A.X est plus importante dans les cellules HeLa siARNct#1. Ainsi, l'inactivation de *dicer* dans les cellules HeLa semble influencer sur la phosphorylation de la protéine Chk1 de la voie de réponse aux cassures double brin.

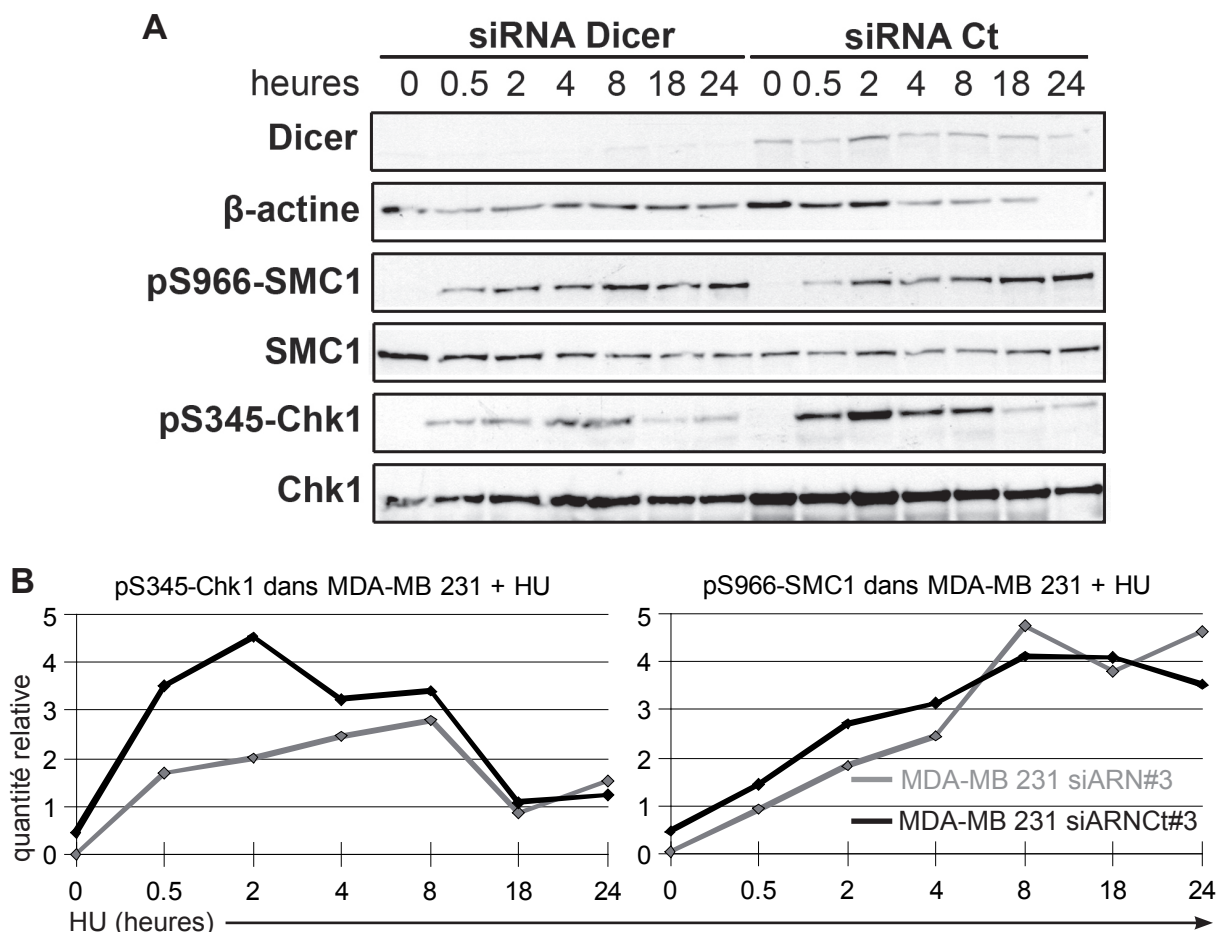


**Figure 23 : Analyse de l'expression de SMC1, Chk1 et H2A.X après hydroxyurée en cellules HeLa**

**A.** Analyse par western blot d'extractions de protéines totales de cellules HeLa avec les anticorps anti-Dicer, anti-pS966-SMC1, anti-pS345-Chk1 et anti-pS139-H2A.X à différents temps de traitement à l'hydroxyurée (le temps 0h correspond à l'ajout de HU). Des quantités égales de protéines ont été déposées sur gel gradient Tris-Glycine 4-12% et le chargement confirmé par un anticorps anti-béta-actine. Les protéines phosphorylées sont comparées à la quantité de protéines totales. **B.** Estimation de la quantité relative au cours de la cinétique des protéines pS345-Chk1, pS966-SMC1 et pS139H2A.X dans les cellules HeLa siARN#1 (courbe grise) et HeLa siARNct#1 (courbe noire) par le logiciel Quantity One (BioRad).

## ii. cellules tumorales mammaires MDA-MB 231

Ces cellules mésenchymateuses mammaires tumorales ont été transfectées avec soit le siARN#3, soit le siARNCt#3 contrôle, puis traitées à l'hydroxyurée. L'analyse des protéines est effectuée par western blot (Fig. 24.A). Comme pour les cellules HeLa, on observe que la phosphorylation de SMC1 est similaire alors que celle de Chk1 diffère : dans les cellules contrôles elle augmente fortement dès le temps 0h avec un pic à 2h. Dans les cellules inactivées pour *dicer*, l'augmentation est moins rapide et le pic est atteint au temps 8h avec environ une fois et demi moins de protéines phosphorylées par rapport au maximum du contrôle (Fig.



**Figure 24 : Analyse de l'expression de SMC1 et Chk1 après hydroxyurée en cellules MDA-MB 231**

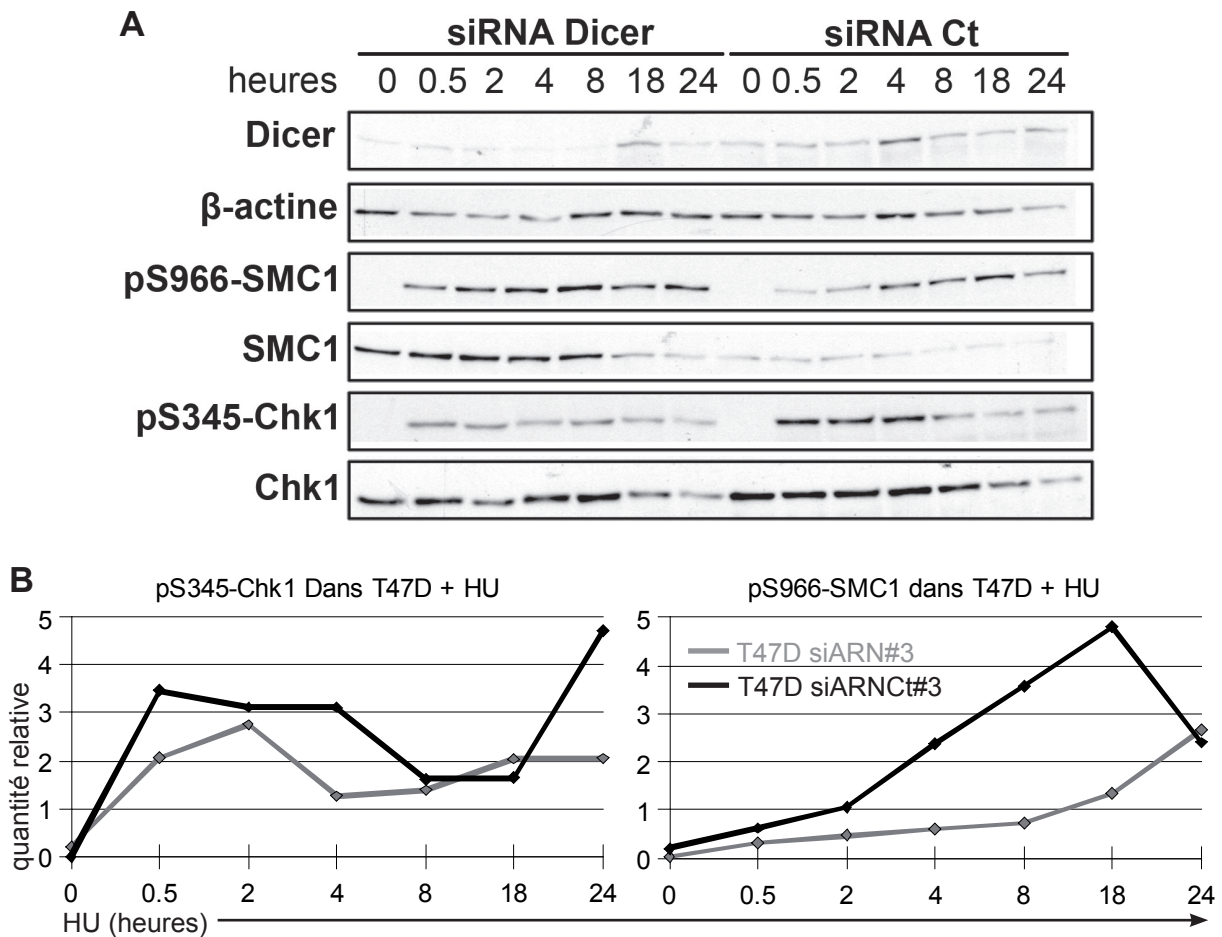
**A.** Analyse par western blot d'extractions de protéines totales de cellules HeLa avec les anticorps anti-Dicer, anti-pS966-SMC1 et anti-pS345-Chk1 à différents temps de traitement à l'hydroxyurée (le temps 0h correspond à l'ajout de HU). Des quantités égales de protéines ont été déposées sur gel gradient Tris-Glycine 4-12% et le chargement confirmé par un anticorps anti-béta-actine. Les protéines phosphorylées sont comparées à la quantité de protéines totales. **B.** Estimation de la quantité relative au cours de la cinétique des protéines pS345-Chk1 et pS966-SMC1 dans les cellules HeLa siARN#3 (courbe grise) et HeLa siARNCt#3 (courbe noire) par le logiciel Quantity One (BioRad).



24.B). La phosphorylation de Chk1 semble donc là aussi perturbée par l'absence de *dicer* dans les cellules.

iii. cellules tumorales mammaires T47D

Ces cellules épithéliales mammaires tumorales ont été transfectées avec soit le siARN#3, soit le siARNCt#3 contrôle, puis traité à l'hydroxyurée. A l'inverse des cellules MDA-MB 231, on observe dans les cellules T47D une phosphorylation similaire de Chk1 entre siARN#3 et siARNCt#3 et un retard pour la phosphorylation de SMC1 (Fig. 25.A). Cette dernière n'augmente qu'au temps 8h pour



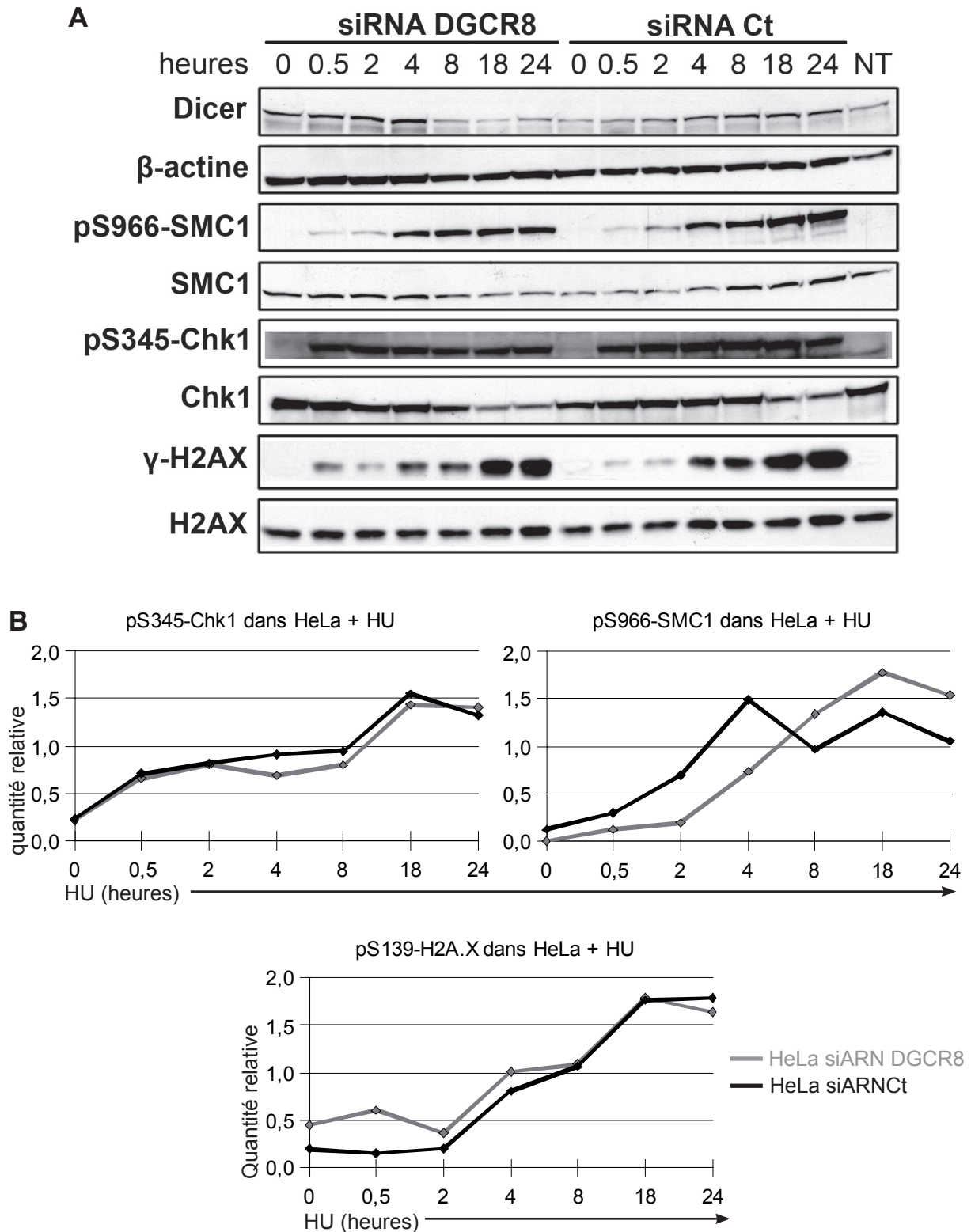
**Figure 25 : Analyse de l'expression de SMC1 et Chk1 après hydroxyurée en cellules T47D**

**A.** Analyse par western blot d'extractions de protéines totales de cellules HeLa avec les anticorps anti-Dicer, anti-pS966-SMC1 et anti-pS345-Chk1 à différents temps de traitement à l'hydroxyurée (le temps 0h correspond à l'ajout de HU). Des quantités égales de protéines ont été déposées sur gel gradient Tris-Glycine 4-12% et le chargement confirmé par un anticorps anti-béta-actine. Les protéines phosphorylées sont comparées à la quantité de protéines totales. **B.** Estimation de la quantité relative au cours de la cinétique des protéines pS345-Chk1 et pS966-SMC1 dans les cellules HeLa siARN#3 (courbe grise) et HeLa siARNCt#3 (courbe noire) par le logiciel Quantity One (BioRad).

les cellules inactivées comparé au temps 2h pour les cellules contrôles (**Fig. 25.B**). Contrairement aux lignées précédentes et en particulier à MDA-MB 231, ce n'est pas la phosphorylation de Chk1 mais celle de SMC1 qui est retardée et moins forte dans les cellules T47D.

### II.3.3. Rôle de la voie des miARN dans la réponse à l'hydroxyurée

L'inactivation de *dicer* ne permet pas de discriminer la voie prépondérante, entre celle des siARN et celle des miARN, qui intervient dans les résultats précédents. Cependant, le microprocesseur est spécifique de la voie des miARN. Ainsi, afin de déterminer plus précisément la part de chacune, nous avons réalisé la même expérience de traitement à l'hydroxyurée dans des cellules HeLa inactivées pour le gène *DGCR8*, partenaire de *Drosha* et intervenant dans la maturation des pri-miARN et pre-miARN. Les résultats obtenus sont présentés en figure 9. Le Western Blot (**Fig. 26.A.**) montre que l'expression de *Dicer* n'est pas ou peu (on observe une légère baisse aux temps 8 et 18h) perturbée par l'inactivation de *DGCR8* et les quantifications des autres protéines (**Fig. 26.B.**) montrent que la phosphorylation des protéines Chk1 et H2A.X sont inchangées entre les siARN Ct et ceux ciblant *DGCR8*. En ce qui concerne la phosphorylation de la protéine SMC1, il existe peu de différence entre ces deux catégories. Il apparaît surtout très clairement, que ces réponses sont différentes de celles observées pour l'inactivation de *dicer* (**Fig. 23**).



**Figure 26 : Analyse de l'expression de SMC1 et Chk1 après hydroxyurée en cellules HeLa**

**A.** Analyse par western blot d'extractions de protéines totales de cellules HeLa avec les anticorps anti-Dicer, anti-pS966-SMC1 et anti-pS345-Chk1 à différents temps de traitement à l'hydroxyurée (le temps 0h correspond à l'ajout de HU). Des quantités égales de protéines ont été déposées sur gel gradient Tris-Glycine 4-12% et le chargement confirmé par un anticorps anti-béata-actine. Les protéines phosphorylées sont comparées à la quantité de protéines totales. **B.** Estimation de la quantité relative au cours de la cinétique des protéines pS345-Chk1, pS966-SMC1 et pS139H2A.X dans les cellules HeLa siARN DGCR8 (courbe grise) et HeLa siARN Ct (courbe noire) par le logiciel Quantity One (BioRad).

### III. Discussion

**N**ous avons montré qu'une inhibition de *dicer* par siARN dans des cellules HeLa entraînait un retard du cycle cellulaire en phase S. En juillet 2008, Tang et coll. ont publié que l'inhibition de *dicer* par siARN n'affectait pas le cycle cellulaire des lignées HEK293T (Human Embryonic Kidney) et HepG2 (Hépatocarcinome). **Tang 2008** Ces conclusions apparaissent donc en contradiction avec le retard de 2h que nous observons dans nos résultats. Cependant, il se peut d'une part que, suivant la lignée utilisée, les conséquences sur le cycle cellulaire d'une inhibition de l'enzyme soient différentes. D'autre part, cette différence pourrait provenir de l'expérience même menée par Tang et son équipe. Brièvement, 48h après transfection de siARN dirigés contre *dicer*, ils ont analysé la répartition des cellules dans les différentes phases du cycle par marquage au iodure de propidium (G0/G1, S et G2/M). N'ayant pas synchronisé leurs cellules inactivées pour *dicer*, un décalage du cycle sans un arrêt ne peut pas être détecté. Par une synchronisation des cellules et un marquage de la phase S en BrDU, nous avons pu mettre en évidence un faible retard dans le déroulement du cycle cellulaire. Il serait aussi intéressant de multiplier l'expérience sur différents types de lignées cellulaires afin de savoir si le retard observé est dépendant du type cellulaire. Il a en effet déjà été montré que l'inactivation de *dicer* dans des cellules ES et des cellules différenciées de souris n'avait pas le même impact sur la voie des siARN.

**Kanellopoulou 2005, Babiarz 2008**

Dans le même article, Tang et coll. montrent que l'inhibition de *dicer* induit une augmentation des dommages à l'ADN, en particulier des cassures double-brin. L'inhibition de *dicer* est associée i. à une augmentation du marquage immunofluorescent de  $\gamma$ H2A.X dans les cellules HEK293T transfectées par un siDicer ii. à une augmentation de la phosphorylation de Chk1 dans leur deux types cellulaires et iii. à un pourcentage élevé de cellules ayant une queue en comète assay. Nous avons également mis en évidence, dans les cellules HeLa, une augmentation du taux de  $\gamma$ H2A.X après inacti-

vation de *dicer*. L'explication mise en avant par l'équipe est que la déstabilisation de l'hétérochromatine induite par l'inactivation de *dicer* entraîne d'une part la dérégulation du déroulement de la réplication et d'autre part l'activation de transposons créant, dans les deux cas, un stress générateur de cassures. Cependant, on observe que les résultats de Tang et coll. ne vont pas forcément dans le même sens que les nôtres. Pour être plus précis, ils mettent en évidence une augmentation de la phosphorylation de la protéine Chk1 lorsque *dicer* est inactivé dans leurs cellules alors que nous observons une diminution de cette phosphorylation dans les trois types cellulaires que nous avons étudiés. Ainsi, les résultats de Tang et coll. semblent montrer que l'inactivation de *dicer* seule conduit à l'augmentation du nombre de cassures et une activation des voies de signalisation via la phosphorylation de Chk1. D'un autre côté, nous montrons que la réponse à un stress réplicatif en absence de *dicer* entraîne un défaut de signalisation des cassures et ceci dans des types cellulaires différents. Cette observation est confortée par un retard de phosphorylation de  $\gamma$ H2A.X dans les cellules HeLa. A partir de nos résultats, nous pouvons suggérer un nouveau rôle pour Dicer au sein du cycle cellulaire. Non seulement cette protéine est impliquée dans la mise en place de l'hétérochromatine péri-centromérique mais il semblerait qu'en plus elle soit impliquée dans la signalisation des cassures qui ont lieu pendant le stress réplicatif induit par la phase S. Par contre, de façon inattendue, nous n'avons pas mis en évidence d'altération de la phosphorylation de SMC1 dans les cellules HeLa et MDA-MB 231. En effet, Dicer a été montrée comme jouant un rôle important dans la mise en place du complexe cohésine (via sa sous-unité Scc1) au niveau de l'hétérochromatine normalement formée. [Fukagawa 2004](#) Lorsque *dicer* est inactivé, la localisation du complexe cohésine et donc de la sous-unité SMC1 est grandement altérée et nous avons donc émis l'hypothèse qu'elle pourrait aussi entraîner une altération de sa phosphorylation, ce qui n'est apparemment pas le cas. Au contraire, dans les cellules T47D, nous observons une diminution de la phosphorylation de SMC1. Cependant, il faut rester prudent quand à l'interprétation de ce résultat car la quantité de la protéine totale (**Fig. 9**, ligne SMC1

du western blot) diminue progressivement sur toute la largeur de la membrane ce qui semble artéfactuel.

Le fait d'observer des dérégulations des acteurs des voies de réponse aux dommages à l'ADN après inactivation de *dicer* n'est pas anodin dans un contexte de cellules mammaires tumorales. En effet, de nombreuses études ont mis en évidence des associations entre les cancers mammaires et des altérations des gènes de la réparation des cassures double-brin. L'exemple le plus important est celui des cancers du sein de type héréditaire dont l'origine dans la majorité des cas est due à des mutations des gènes BRCA1 et BRCA2. Ces deux gènes jouent un rôle prépondérant dans la réparation des cassures double-brins. [Cortez 1999](#), Plus récemment, il a été publié que Chk2 serait également impliqué dans ces formes héréditaires. [Nevanlinna 2006](#) Par ailleurs, Rad51 (et aussi les gènes associés XRCC2, XRCC3, Rad51L3) qui se situe en aval de la recombinaison homologue présente de nombreuses dérégulations (mutations, polymorphisme, LOH et sur-expression) associées aux cancers du sein sporadiques. [Thacker 2005](#) De même pour ATM, des mutations, des sous-expressions, des hyperméthylations du promoteur et des LOH ont été observées et associées à un risque plus élevé de cancer du sein. [Cavaciuti 2005](#), [Dork 2001](#) Ces dérégulation sont d'autant plus importantes que ATM est en amont des voies de réparation et active directement p53 et BRCA1. Encore de nombreux autres acteurs (Chk2, les gènes du complexe MRN, Ku70, Ku80, DNA-PK, etc.) de ces voies sont connus pour être perturbés dans un certain pourcentage de cancers du sein. [Ralhan 2007 pour revue](#) Il apparaît que quelque soit le ou les gènes impliqués, la réparation des dommages à l'ADN soit souvent et précocement atteinte lors de la progression tumorale mammaire. [Bartkova 2005](#) C'est pourquoi on peut se poser la question de savoir si *dicer* pourrait être un dénominateur commun aux différentes dérégulations des voies de réparation, notamment en déstabilisant l'établissement de l'hétérochromatine et en modifiant la localisation de SMC1, à la fois protéine de structure des chromosomes mais aussi impliquée

dans la réparation.

**N**ous avons montré que l'inhibition de *dicer* dans des cellules HeLa synchronisées induit un excès de cellules apoptotiques dès le temps 6h c'est à dire au moment où les cellules siDicer entrent en phases G2/M. De précédentes études portant sur la levure *S. pombe* ont montré que la délétion de *dicer* conduisait à de profondes modifications de la dynamique des chromosomes et des perturbations du cycle cellulaire. Ainsi, il est reporté une croissance lente des colonies mutées, l'abrogation du point de contrôle dans la transition S-G2/M, des séparations précoces de chromatides soeurs, des retards de migration de chromosomes lors de l'anaphase, des cytokinèses partielles et autres défauts. [Provost 2002](#), [Hall 2003](#), [Volpe 2003](#) L'origine de la plupart de ces défauts est la perte de l'intégrité des centromères du fait de l'absence d'hétérochromatine. Cependant, on ne trouve aucune données sur la mortalité cellulaire qui pourrait accompagner ces événements pourtant potentiellement délétères pour une cellule. Dans la seule étude sur des cellules de vertébrés, les hybrides DT40 contenant le chromosome 21 humain, [Fukagawa et coll., 2004](#) montrent aussi une déstabilisation de l'hétérochromatine péricentromérique avec pour conséquences des séparations prématurées de chromatides soeur et une accumulation rapide de cellules aneuploïdes. L'équipe rapporte aussi la perte du point de contrôle mitotique dans leur construction. En effet, ils ne détectent pas le signal de BubR1 qui devrait être normalement présent sur les chromosomes séparés prématurément. Cette observation leur permet d'expliquer la mortalité cellulaire qu'ils observent lors de l'interphase suivant la division. [Fukagawa 2004](#) Il est très probable que l'origine de l'apoptose observée dans nos cellules HeLa siDicer soit du même ordre. De plus, on peut remarquer que, comme dans cette étude, le taux de mortalité augmente lorsque les cellules rentrent à nouveau en phase G1.

**L**es résultats obtenus après inactivation de *DGCR8* (co-facteur de la ribonucléase Drosha et jouant un rôle majeur dans la maturation des miARN)

dans les cellules HeLa synchronisées ou traitées au HU semblent montrer une différence par rapport à l'inactivation de *dicer*. Le cycle cellulaire n'apparaît pas perturbé ainsi que les voies de réponses aux cassures faisant intervenir les phosphorylation de SMC1 et Chk1. Il est important de pouvoir connaître la part d'implication respective de la voie des siARN et de la voie des miARN lorsque *dicer* est dérégulé. Les études sur le mécanisme d'interférence ARN et son implication dans la progression tumorale laissent apparaître une interrogation. En effet, alors qu'il a été montré un rôle très important de la voie des siARN chez la levure dans le maintien de l'intégrité du génome, les publications qui traitent des cellules mammifères se concentrent presque exclusivement sur la voie des miARN. Pourquoi ce mécanisme très conservé au cours de l'évolution n'aurait pas des effets similaires dans les organismes "supérieurs" ? Nos résultats, ainsi que ceux de Fugakawa et coll. et Tang et coll., tendent d'ailleurs à démontrer qu'il existe effectivement des conséquences importantes lorsque les siARN ne sont plus produits normalement. Cependant, il est assez difficile d'isoler l'une ou l'autre voie et il est aussi évident que la dérégulation des miARN entraîne la perturbation de nombreuses voies cellulaires. Ainsi, il a été montré récemment que les miARN 17 et 20a du cluster miR-17-92 jouaient un rôle important dans la transition G1-S. L'inhibition de leur expression en phase G1 conduit à l'activation du point de restriction et donc à l'arrêt du cycle. Cette activation résulte d'une accumulation trop précoce du facteur de transcription E2F1. En temps normal, miARN 17 et 20a permettraient de limiter la quantité de E2F1 jusqu'à ce que la cellule arrive à la transition G1/S et puisse le tolérer. [Pickering 2008](#) En théorie, nous aurions dû observer dans nos cellules HeLa inactivées pour *DGCR8* une accumulation de cellules en phase G1, ce qui n'est pas le cas. L'arrêt en G1 induit par la perte des deux miARN est peut-être dépendante du type cellulaire (fibroblastes de poumon embryonnaire humain dans les travaux de Pickering et coll.). Mais on peut aussi envisager que le fait d'inactiver *DGCR8* n'ai pas les mêmes effets que l'inactivation spécifique (utilisation d'oligonucléotides complémentaires des séquences des miARN étudiés) de deux miARN. Dans le premier cas,



nous avons inhibé l'étape générale de pré-maturation des miARN ce qui a été montré comme associé à une baisse globale du niveau des pre-miARN et miARN dans des cellules ES murines après plusieurs jours. [Wang 2007](#) Ainsi, en observant comme nous l'avons fait 48h après transfection des siARN anti-DGCR8, on peut penser que la durée de vie des miARN ou leur stabilisation (par un gène argonaute par exemple, comme Ago1 stabilisant miR168 chez *Arabidopsis* [Vaucheret 2006](#)) pourraient suffir à masquer les effets décrits par Pickering et coll.

## IV. Conclusion et perspectives

Nous avons mis en évidence que *dicer* est un gène pertinent à étudier dans les cancers du sein à plusieurs niveaux. Tout d'abord d'un point de vue clinique où il apparaît qu'une faible expression de Dicer est un facteur pronostique indépendant de rechute tardive. Ainsi, l'expression du gène n'est pas associée à des facteurs déjà connus de prédiction de rechute métastatique. Dans notre étude, nous montrons que le taux d'ARNm de *dicer* peut prédire un risque de récurrence aussi bien pour une tumeur de petite taille qu'une tumeur de grande taille. Les très bons odd ratios (supérieurs à 3) ayant été obtenus sur un échantillonnage d'une centaine de tumeurs primaires, il semble important de pouvoir les confirmer sur des séries plus conséquentes et aussi sur des séries provenant et examinées par d'autres centres hospitaliers. Si ces résultats sont confirmés, la recherche du taux d'expression de Dicer pourrait alors devenir une analyse réalisée en anatomie pathologique dans le but de dresser un profil précis et personnalisé des patientes risquant de rechuter. Par ailleurs, cette même étude menée en parallèle sur des sous-unités du complexe APC/C nous a permis de mettre en évidence des corrélations entre leur expression et certains critères cliniques de tumeurs mammaires. Il apparaît notamment une surexpression de la sous-unité APC-11 dans les tumeurs mammaires aneuploïdes ou de grades SBR III. La recherche de nouveaux marqueurs dans les cancers du sein est primordiale et le gène APC-11 pourrait être pertinent dans ce rôle. Poursuivre les investigations permettrait de déterminer précisément sa pertinence pronostique.

Ensuite, nous avons démontré l'implication de *dicer* dans le cycle cellulaire et les voies de réponse aux cassures double-brin après stress répliatif dans différents types cellulaires (HeLa et lignées tumorales mammaires humaine). L'inactivation de l'expression de la protéine Dicer par interférence ARN entraîne un retard d'environ 2h à partir de la phase S dans la division des cellules ainsi que des défauts de phosphorylation de protéines impliquées dans la voie ATR de réponse aux cassures. Ces effets

semblent provenir principalement de la perte de production des siARN car ils ne sont pas retrouvés lorsque le gène *DGCR8*, qui permet la maturation des miARN, est inactivé. Il serait intéressant ici de confirmer ces résultats et d'inactiver d'autres acteurs des deux voies de l'iARN (voie des siARN contre voie des miARN). Pour cela, l'utilisation d'une puce à miARN permettrait de connaître assez précisément les variations induites par ces différentes inactivations (*dicer*, *drosha*, *DGCR8*, exportine-5, *TRBP*, *PACT*, *Ago2*, etc.). Afin d'évaluer l'implication de la voie des siARN sur le retard du cycle cellulaire observé après inactivation de *dicer*, il faudrait étudier par RT-PCR l'expression des séquences satellites péri-centromériques (comme par exemple  $\alpha$ -sat1, 2 et 3 après inactivation de *dicer*). Une telle étude indiquerait le taux de transcription et donc l'état chromatinien de ces régions. Nous aurions alors la possibilité de connaître à quel(s) moment(s) précis de la phase S l'expression des séquences satellites est réprimée. Par ailleurs, une analyse par immunofluorescence, FISH ou "étalement de chromosomes" de la structure des chromosomes lors du cycle cellulaire en l'absence de *dicer* pourrait donner de nouvelles informations sur l'origine du retard observé en phase S et au moment de la mitose (mauvaise ségrégation, retard chromosomique, micro-noyaux, etc.). Il serait aussi pertinent d'inactiver plus durablement l'expression du gène *dicer* afin d'observer les conséquences d'une telle inhibition cellulaire à plus long terme (croissance cellulaire, taux de mortalité, instabilités chromosomiques, etc.). Cette étude impliquerait l'utilisation d'un shARN anti-*dicer* intégré à un vecteur de type lentivirus infecté dans les lignées choisies. Parallèlement, une surexpression de *dicer* *in vitro* permettrait de compléter les connaissances sur les effets de sa dérégulation d'autant plus qu'une trop forte expression a déjà été mise en évidence dans plusieurs types cancéreux. Il serait aussi important d'inactiver l'expression de *dicer* dans différentes lignées tumorales mammaires humaines afin d'observer si certains types (luminal, basal, etc.) sont plus sensibles que d'autres à l'inactivation de *dicer*. Nos études d'expressions tendent déjà à montrer que les cellules mésenchymateuses se comportent différemment des cellules épithéliales, il serait bien d'affiner ces observations.

A plus long terme, la réalisation de tels travaux sur d'autres types tumoraux pourrait se révéler utile d'autant plus que des dérégulations de *dicer* ont déjà été mises en évidence dans les cancers du poumon, de l'ovaire et de la prostate et qu'une expression différentielle de notre gène d'intérêt a été observée dans les différents types cellulaires les composant. Il serait aussi intéressant de se demander si la perte d'expression de *dicer* dans les cancers du sein est un évènement tardif ou encore d'étudier son expression transcriptionnelle et traductionnelle dans des cellules souches tumorales. Enfin, la génération de souris KO conditionnel pour *dicer* dans les glandes mammaires apporterait des données *in vivo* de premier ordre et pourrait permettre de savoir, par exemple, comment se comporterait des cellules inactivées pour *dicer* dans leur environnement tissulaire.

## Bibliographie :

- Acquaviva, C. and J. Pines (2006). "The anaphase-promoting complex/cyclosome: APC/C." *J Cell Sci* 119(Pt 12): 2401-4.
- Al-Hajj, M. and M. F. Clarke (2004). "Self-renewal and solid tumor stem cells." *Oncogene* 23(43): 7274-82.
- Anderson, D. E., A. Losada, et al. (2002). "Condensin and cohesin display different arm conformations with characteristic hinge angles." *J Cell Biol* 156(3): 419-24.
- Andl, T., E. P. Murchison, et al. (2006). "The miRNA-processing enzyme dicer is essential for the morphogenesis and maintenance of hair follicles." *Curr Biol* 16(10): 1041-9.
- Arumugam, P., S. Gruber, et al. (2003). "ATP hydrolysis is required for cohesin's association with chromosomes." *Curr Biol* 13(22): 1941-53.
- Aylon, Y., B. Liefshitz, et al. (2004). "The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle." *EMBO J* 23(24): 4868-75.
- Babiarz, J. E., J. G. Ruby, et al. (2008). "Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs." *Genes Dev* 22(20): 2773-85.
- Bandera, C. A., H. Takahashi, et al. (1997). "Deletion mapping of two potential chromosome 14 tumor suppressor gene loci in ovarian carcinoma." *Cancer Res* 57(3): 513-5.
- Bartek, J. and J. Lukas (2007). "DNA damage checkpoints: from initiation to recovery or adaptation." *Curr Opin Cell Biol* 19(2): 238-45.
- Bartkova, J., Z. Horejsi, et al. (2005). "DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis." *Nature* 434(7035): 864-70.
- Bernstein, E., A. A. Caudy, et al. (2001). "Role for a bidentate ribonuclease in the initiation step of RNA interference." *Nature* 409(6818): 363-6.
- Bernstein, E., S. Y. Kim, et al. (2003). "Dicer is essential for mouse development." *Nat Genet* 35(3): 215-7.
- Bertucci, F., P. Finetti, et al. (2005). "Gene expression profiling identifies molecular subtypes of inflammatory breast cancer." *Cancer Res* 65(6): 2170-8.
- Birkenbihl, R. P. and S. Subramani (1992). "Cloning and characterization of rad21 an essential gene of *Schizosaccharomyces pombe* involved in DNA double-strand-break repair." *Nucleic Acids Res* 20(24): 6605-11.
- Bohnsack, M. T., K. Czaplinski, et al. (2004). "Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs." *RNA* 10(2): 185-91.
- Boice, J. D., Jr. (1996). "Cancer following irradiation in childhood and adolescence." *Med Pediatr Oncol Suppl* 1: 29-34.
- Bonnet, D. and J. E. Dick (1997). "Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell." *Nat Med* 3(7): 730-7.
- Borsani, O., J. Zhu, et al. (2005). "Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis*." *Cell* 123(7): 1279-91.
- Branzei, D. and M. Foiani (2007). "Interplay of replication checkpoints and repair proteins at stalled replication forks." *DNA Repair (Amst)* 6(7): 994-1003.
- Branzei, D. and M. Foiani (2008). "Regulation of DNA repair throughout the cell cycle." *Nat Rev Mol Cell Biol* 9(4): 297-308.
- Buerger, H., F. Otterbach, et al. (1999). "Comparative genomic hybridization of ductal carcinoma in situ of the breast-evidence of multiple genetic pathways." *J Pathol* 187(4): 396-402.
- Buerger, H., F. Otterbach, et al. (1999 bis). "Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes." *J Pathol* 189(4): 521-6.
- Buerger, H., E. C. Mommers, et al. (2001). "Ductal invasive G2 and G3 carcinomas of the breast are the end stages of at least two different lines of genetic evolution." *J Pathol* 194(2): 165-70.
- Card, D. A., P. B. Hebbar, et al. (2008). "Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells." *Mol Cell Biol* 28(20): 6426-38.
- Castro, A., C. Bernis, et al. (2005). "The anaphase-promoting complex: a key factor in the regulation of cell cycle." *Oncogene* 24(3): 314-25.
- Caudy, A. A., M. Myers, et al. (2002). "Fragile X-related protein and VIG associate with the RNA interference machinery." *Genes Dev* 16(19): 2491-6.
- Cavaciuti, E., A. Lauge, et al. (2005). "Cancer risk according to type and location of ATM mutation in ataxia-telangiectasia families." *Genes Chromosomes Cancer* 42(1): 1-9.
- Cerutti, H. and J. A. Casas-Mollano (2006). "On the origin and functions of RNA-mediated silencing: from protists to man." *Curr Genet* 50(2): 81-99.
- Chang, H. Y., D. S. Nuyten, et al. (2005). "Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival." *Proc Natl Acad Sci U S A*

- 102(10): 3738-43.
- Chang, H. Y., J. B. Sneddon, et al. (2004). "Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds." *PLoS Biol* 2(2): E7.
  - Chen, R. H., J. C. Waters, et al. (1996). "Association of spindle assembly checkpoint component X MAD2 with unattached kinetochores." *Science* 274(5285): 242-6.
  - Chiosea, S., E. Jelezcova, et al. (2006). "Up-regulation of dicer, a component of the MicroRNA machinery, in prostate adenocarcinoma." *Am J Pathol* 169(5): 1812-20.
  - Chiosea, S., E. Jelezcova, et al. (2007). "Overexpression of Dicer in precursor lesions of lung adenocarcinoma." *Cancer Res* 67(5): 2345-50.
  - Christiansen, J. J. and A. K. Rajasekaran (2006). "Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis." *Cancer Res* 66(17): 8319-26.
  - Ciosk, R., M. Shirayama, et al. (2000). "Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins." *Mol Cell* 5(2): 243-54.
  - Cobb, B. S., T. B. Nesterova, et al. (2005). "T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer." *J Exp Med* 201(9): 1367-73.
  - Cobb, B. S., A. Hertweck, et al. (2006). "A role for Dicer in immune regulation." *J Exp Med* 203(11): 2519-27.
  - Collaborative group on hormonal factors in breast cancer. Alcohol, tobacco and breast cancer : collaborative reanalysis of individual data from 64 epidemiological studies, including 64,534 women with breast cancer and 131,348 women without breast cancer. *Br J Cancer* 2002 ; 87 : 1234-45.
  - Cortez, D., Y. Wang, et al. (1999). "Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks." *Science* 286(5442): 1162-6.
  - Dalmay, T., A. Hamilton, et al. (2000). "An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus." *Cell* 101(5): 543-53.
  - Denli, A. M., B. B. Tops, et al. (2004). "Processing of primary microRNAs by the Microprocessor complex." *Nature* 432(7014): 231-5.
  - Deramaudt, T. and A. K. Rustgi (2005). "Mutant KRAS in the initiation of pancreatic cancer." *Biochim Biophys Acta* 1756(2): 97-101.
  - Dlakic, M. (2006). "DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold." *Bioinformatics* 22(22): 2711-4.
  - Dobles, M., V. Liberal, et al. (2000). "Chromosome missegregation and apoptosis in mice lacking the mitotic checkpoint protein Mad2." *Cell* 101(6): 635-45.
  - Dork, T., R. Bendix, et al. (2001). "Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients." *Cancer Res* 61(20): 7608-15.
  - Dykxhoorn, D. M., C. D. Novina, et al. (2003). "Killing the messenger: short RNAs that silence gene expression." *Nat Rev Mol Cell Biol* 4(6): 457-67.
  - Elbashir, S. M., W. Lendeckel, et al. (2001). "RNA interference is mediated by 21- and 22-nucleotide RNAs." *Genes Dev* 15(2): 188-200.
  - Falck, J., J. Coates, et al. (2005). "Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage." *Nature* 434(7033): 605-11.
  - Fan, C., D. S. Oh, et al. (2006). "Concordance among gene-expression-based predictors for breast cancer." *N Engl J Med* 355(6): 560-9.
  - Fang, G. (2002). "Checkpoint protein BubR1 acts synergistically with Mad2 to inhibit anaphase-promoting complex." *Mol Biol Cell* 13(3): 755-66.
  - Fire, A., S. Xu, et al. (1998). "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*." *Nature* 391(6669): 806-11.
  - Flavin, R. J., P. C. Smyth, et al. (2008). "Altered eIF6 and Dicer expression is associated with clinicopathological features in ovarian serous carcinoma patients." *Mod Pathol*.
  - Ford, D., D. F. Easton, et al. (1998). "Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium." *Am J Hum Genet* 62(3): 676-89.
  - Fukagawa, T., M. Nogami, et al. (2004). "Dicer is essential for formation of the heterochromatin structure in vertebrate cells." *Nat Cell Biol* 6(8): 784-91.
  - Garg, R., S. Callens, et al. (2004). "Chromatin association of rad17 is required for an ataxia telangiectasia and rad-related kinase-mediated S-phase checkpoint in response to low-dose ultraviolet radiation." *Mol Cancer Res* 2(6): 362-9.
  - Garzon, R., M. Fabbri, et al. (2006). "MicroRNA expression and function in cancer." *Trends Mol Med* 12(12): 580-7.
  - Gascioli, V., A. C. Mallory, et al. (2005). "Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs." *Curr Biol* 15(16): 1494-500.
  - Gimenez-Abian, J. F., I. Sumara, et al. (2004). "Regulation of sister chromatid cohesion between

- chromosome arms." *Curr Biol* 14(13): 1187-93.
- Greenburg, G. and E. D. Hay (1982). "Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells." *J Cell Biol* 95(1): 333-9.
  - Gregory, R. I., K. P. Yan, et al. (2004). "The Microprocessor complex mediates the genesis of microRNAs." *Nature* 432(7014): 235-40.
  - Gregory, R. I., T. P. Chendrimada, et al. (2005). "Human RISC couples microRNA biogenesis and posttranscriptional gene silencing." *Cell* 123(4): 631-40.
  - Grishok, A., A. E. Pasquinelli, et al. (2001). "Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing." *Cell* 106(1): 23-34.
  - Gruber, S., P. Arumugam, et al. (2006). "Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge." *Cell* 127(3): 523-37.
  - Guacci, V., D. Koshland, et al. (1997). "A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in *S. cerevisiae*." *Cell* 91(1): 47-57.
  - Guo, S. and K. J. Kemphues (1995). "par-1, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed." *Cell* 81(4): 611-20.
  - Haering, C. H., J. Lowe, et al. (2002). "Molecular architecture of SMC proteins and the yeast cohesin complex." *Mol Cell* 9(4): 773-88.
  - Hagting, A., N. Den Elzen, et al. (2002). "Human securin proteolysis is controlled by the spindle checkpoint and reveals when the APC/C switches from activation by Cdc20 to Cdh1." *J Cell Biol* 157(7): 1125-37.
  - Hall, I. M., K. Noma, et al. (2003). "RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast." *Proc Natl Acad Sci U S A* 100(1): 193-8.
  - Hamilton, A. J. and D. C. Baulcombe (1999). "A species of small antisense RNA in posttranscriptional gene silencing in plants." *Science* 286(5441): 950-2.
  - Hammond, S. M., S. Boettcher, et al. (2001). "Argonaute2, a link between genetic and biochemical analyses of RNAi." *Science* 293(5532): 1146-50.
  - Han, J., Y. Lee, et al. (2004). "The Drosha-DGCR8 complex in primary microRNA processing." *Genes Dev* 18(24): 3016-27.
  - Hardwick, K. G., R. C. Johnston, et al. (2000). "MAD3 encodes a novel component of the spindle checkpoint which interacts with Bub3p, Cdc20p, and Mad2p." *J Cell Biol* 148(5): 871-82.
  - Harfe, B. D., M. T. McManus, et al. (2005). "The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb." *Proc Natl Acad Sci U S A* 102(31): 10898-903.
  - Harris, K. S., Z. Zhang, et al. (2006). "Dicer function is essential for lung epithelium morphogenesis." *Proc Natl Acad Sci U S A* 103(7): 2208-13.
  - Hartwell, L. H., J. Culotti, et al. (1970). "Genetic control of the cell-division cycle in yeast. I. Detection of mutants." *Proc Natl Acad Sci U S A* 66(2): 352-9.
  - Helt, C. E., W. Wang, et al. (2005). "Evidence that DNA damage detection machinery participates in DNA repair." *Cell Cycle* 4(4): 529-32.
  - Hirano, T. (2006). "At the heart of the chromosome: SMC proteins in action." *Nat Rev Mol Cell Biol* 7(5): 311-22.
  - Hoegger, H., D. Dejsuphong, et al. (2006). "Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells." *EMBO J* 25(6): 1305-14.
  - Holtrich, U., G. Wolf, et al. (1994). "Induction and down-regulation of PLK, a human serine/threonine kinase expressed in proliferating cells and tumors." *Proc Natl Acad Sci U S A* 91(5): 1736-40.
  - Hossain, A., M. T. Kuo, et al. (2006). "Mir-17-5p regulates breast cancer cell proliferation by inhibiting translation of AIB1 mRNA." *Mol Cell Biol* 26(21): 8191-201.
  - Howell, B. J., D. B. Hoffman, et al. (2000). "Visualization of Mad2 dynamics at kinetochores, along spindle fibers, and at spindle poles in living cells." *J Cell Biol* 150(6): 1233-50.
  - Hu, Z., C. Fan, et al. (2006). "The molecular portraits of breast tumors are conserved across microarray platforms." *BMC Genomics* 7: 96.
  - Ihara, Y., Y. Kato, et al. (2002). "Allelic imbalance of 14q32 in esophageal carcinoma." *Cancer Genet Cytogenet* 135(2): 177-81.
  - Ira, G., A. Pelliccioli, et al. (2004). "DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1." *Nature* 431(7011): 1011-7.
  - Jallepalli, P. V., I. C. Waizenegger, et al. (2001). "Securin is required for chromosomal stability in human cells." *Cell* 105(4): 445-57.
  - Jiricny, J. (2006). "The multifaceted mismatch-repair system." *Nat Rev Mol Cell Biol* 7(5): 335-46.
  - Johnson, S. M., H. Grosshans, et al. (2005). "RAS is regulated by the let-7 microRNA family." *Cell* 120(5): 635-47.
  - Jorgensen, R. (1990). "Altered gene expression in plants due to trans interactions between homologous genes." *Trends Biotechnol* 8(12): 340-4.

- Kakarala, M. and M. S. Wicha (2008). "Implications of the cancer stem-cell hypothesis for breast cancer prevention and therapy." *J Clin Oncol* 26(17): 2813-20.
- Kanellopoulou, C., S. A. Muljo, et al. (2005). "Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing." *Genes Dev* 19(4): 489-501.
- Kelsey, J. L. and L. Bernstein (1996). "Epidemiology and prevention of breast cancer." *Annu Rev Public Health* 17: 47-67.
- Ketting, R. F., S. E. Fischer, et al. (2001). "Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*." *Genes Dev* 15(20): 2654-9.
- Kim, J. M., H. Y. Sohn, et al. (2005). "Identification of gastric cancer-related genes using a cDNA microarray containing novel expressed sequence tags expressed in gastric cancer cells." *Clin Cancer Res* 11(2 Pt 1): 473-82.
- Kim, S. T., B. Xu, et al. (2002). "Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage." *Genes Dev* 16(5): 560-70.
- Kim, J., K. Inoue, et al. (2007). "A MicroRNA feedback circuit in midbrain dopamine neurons." *Science* 317(5842): 1220-4.
- Krogh, B. O. and L. S. Symington (2004). "Recombination proteins in yeast." *Annu Rev Genet* 38: 233-71.
- Lee, Y., C. Ahn, et al. (2003). "The nuclear RNase III Drosha initiates microRNA processing." *Nature* 425(6956): 415-9.
- Lehmann, A. R., A. Niimi, et al. (2007). "Translesion synthesis: Y-family polymerases and the polymerase switch." *DNA Repair (Amst)* 6(7): 891-9.
- Lengronne, A., Y. Katou, et al. (2004). "Cohesin relocation from sites of chromosomal loading to places of convergent transcription." *Nature* 430(6999): 573-8.
- Li, D., J. Zhu, et al. (2003). "Overexpression of oncogenic STK15/BTAK/Aurora A kinase in human pancreatic cancer." *Clin Cancer Res* 9(3): 991-7.
- Li, Y. and R. Benzra (1996). "Identification of a human mitotic checkpoint gene: hsMAD2." *Science* 274(5285): 246-8.
- Little, M. P., C. R. Muirhead, et al. (1999). "Relative risks of radiation-associated cancer: comparison of second cancer in therapeutically irradiated populations with the Japanese atomic bomb survivors." *Radiat Environ Biophys* 38(4): 267-83.
- Liu, J., M. A. Carmell, et al. (2004). "Argonaute2 is the catalytic engine of mammalian RNAi." *Science* 305(5689): 1437-41.
- Liu, Q., T. A. Rand, et al. (2003). "R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway." *Science* 301(5641): 1921-5.
- Liu, J. and S. Y. Fuchs (2006). "Cross-talk between APC/C and CBP/p300." *Cancer Biol Ther* 5(7): 760-2.
- Losada, A., M. Hirano, et al. (1998). "Identification of *Xenopus* SMC protein complexes required for sister chromatid cohesion." *Genes Dev* 12(13): 1986-97.
- Losada, A., M. Hirano, et al. (2002). "Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis." *Genes Dev* 16(23): 3004-16.
- Lukas, J., C. Lukas, et al. (2004). "Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time." *DNA Repair (Amst)* 3(8-9): 997-1007.
- Lund, E., S. Guttinger, et al. (2004). "Nuclear export of microRNA precursors." *Science* 303(5654): 95-8.
- Lynch, P.J. (Medical illustrator), Jaffe, C.C. (MD, cardiologist). (2007) Yale University School of Medicine, Center for Advanced Instructional Media.
- Ma, X. J., Z. Wang, et al. (2004). "A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen." *Cancer Cell* 5(6): 607-16.
- Mani, S. A., W. Guo, et al. (2008). "The epithelial-mesenchymal transition generates cells with properties of stem cells." *Cell* 133(4): 704-15.
- Margis, R., A. F. Fusaro, et al. (2006). "The evolution and diversification of Dicers in plants." *FEBS Lett* 580(10): 2442-50.
- Martinez, J., A. Patkaniowska, et al. (2002). "Single-stranded antisense siRNAs guide target RNA cleavage in RNAi." *Cell* 110(5): 563-74.
- Matsuoka, S., B. A. Ballif, et al. (2007). "ATR and ATM substrate analysis reveals extensive protein networks responsive to DNA damage." *Science* 316(5828): 1160-6.
- Mayr, C., M. T. Hemann, et al. (2007). "Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation." *Science* 315(5818): 1576-9.
- McGowan, C. H.; Russel, P. (2004). The DNA damage response: sensing and signaling (sic). *Curr Opin Cell Biol*, 16, 629-633.
- Meister, G., M. Landthaler, et al. (2004). "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs." *Mol Cell* 15(2): 185-97.



- Michaelis, C., R. Ciosk, et al. (1997). "Cohesins: chromosomal proteins that prevent premature separation of sister chromatids." *Cell* 91(1): 35-45.
- Michel, L. S., V. Liberal, et al. (2001). "MAD2 haplo-insufficiency causes premature anaphase and chromosome instability in mammalian cells." *Nature* 409(6818): 355-9.
- Milutinovich, M., E. Unal, et al. (2007). "A multi-step pathway for the establishment of sister chromatid cohesion." *PLoS Genet* 3(1): e12.
- Mudhasani, R., Z. Zhu, et al. (2008). "Loss of miRNA biogenesis induces p19Arf-p53 signaling and senescence in primary cells." *J Cell Biol* 181(7): 1055-63.
- Murchison, E. P., J. F. Partridge, et al. (2005). "Characterization of Dicer-deficient murine embryonic stem cells." *Proc Natl Acad Sci U S A* 102(34): 12135-40.
- Murchison, E. P., P. Stein, et al. (2007). "Critical roles for Dicer in the female germline." *Genes Dev* 21(6): 682-93.
- Nakayama, K. I. and K. Nakayama (2006). "Ubiquitin ligases: cell-cycle control and cancer." *Nat Rev Cancer* 6(5): 369-81.
- Napoli, C., C. Lemieux, et al. (1990). "Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans." *Plant Cell* 2(4): 279-289.
- Nevanlinna, H. and J. Bartek (2006). "The CHEK2 gene and inherited breast cancer susceptibility." *Oncogene* 25(43): 5912-9.
- Neve, R. M., K. Chin, et al. (2006). "A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes." *Cancer Cell* 10(6): 515-27.
- Nielsen, T. O., F. D. Hsu, et al. (2004). "Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma." *Clin Cancer Res* 10(16): 5367-74.
- Noma, K., T. Sugiyama, et al. (2004). "RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing." *Nat Genet* 36(11): 1174-80.
- Nykanen, A., B. Haley, et al. (2001). "ATP requirements and small interfering RNA structure in the RNA interference pathway." *Cell* 107(3): 309-21.
- Paik, S., S. Shak, et al. (2004). "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer." *N Engl J Med* 351(27): 2817-26.
- Paques, F. and J. E. Haber (1999). "Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*." *Microbiol Mol Biol Rev* 63(2): 349-404.
- Pardal, R., M. F. Clarke, et al. (2003). "Applying the principles of stem-cell biology to cancer." *Nat Rev Cancer* 3(12): 895-902.
- Park, W., J. Li, et al. (2002). "CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*." *Curr Biol* 12(17): 1484-95.
- Park, K. H., S. E. Choi, et al. (2005). "Downregulation of the anaphase-promoting complex (APC)<sup>7</sup> in invasive ductal carcinomas of the breast and its clinicopathologic relationships." *Breast Cancer Res* 7(2): R238-47.
- Passmore, L. A., E. A. McCormack, et al. (2003). "Doc1 mediates the activity of the anaphase-promoting complex by contributing to substrate recognition." *EMBO J* 22(4): 786-96.
- Pellman, D. and M. F. Christman (2001). "Separase anxiety: dissolving the sister bond and more." *Nat Cell Biol* 3(9): E207-9.
- Perou, C. M., T. Sorlie, et al. (2000). "Molecular portraits of human breast tumours." *Nature* 406(6797): 747-52.
- Pickering, M. T., B. M. Stadler, et al. (2008). "miR-17 and miR-20a temper an E2F1-induced G1 checkpoint to regulate cell cycle progression." *Oncogene*.
- Ponti, D., A. Costa, et al. (2005). "Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties." *Cancer Res* 65(13): 5506-11.
- Provost, P., R. A. Silverstein, et al. (2002). "Dicer is required for chromosome segregation and gene silencing in fission yeast cells." *Proc Natl Acad Sci U S A* 99(26): 16648-53.
- Rakha, E. A., J. S. Reis-Filho, et al. (2008). "Basal-like breast cancer: a critical review." *J Clin Oncol* 26(15): 2568-81.
- Ralhan, R., J. Kaur, et al. (2007). "Links between DNA double strand break repair and breast cancer: accumulating evidence from both familial and nonfamilial cases." *Cancer Lett* 248(1): 1-17.
- Rudland, P. S., D. G. Fernig, et al. (1995). "Growth factors and their receptors in neoplastic mammary glands." *Biomed Pharmacother* 49(9): 389-99.
- Saberi, A., H. Hohegger, et al. (2007). "RAD18 and poly(ADP-ribose) polymerase independently suppress the access of nonhomologous end joining to double-strand breaks and facilitate homologous recombination-mediated repair." *Mol Cell Biol* 27(7): 2562-71.
- Sadaie, M., T. Iida, et al. (2004). "A chromodomain protein, Chp1, is required for the establishment of heterochromatin in fission yeast." *EMBO J* 23(19): 3825-35.
- Sancar, A., L. A. Lindsey-Boltz, et al. (2004). "Molecular mechanisms of mammalian DNA repair and

- the DNA damage checkpoints." *Annu Rev Biochem* 73: 39-85.
- Sasaki, T. and N. Shimizu (2007). "Evolutionary conservation of a unique amino acid sequence in human DICER protein essential for binding to Argonaute family proteins." *Gene* 396(2): 312-20.
  - Schramke, V. and R. Allshire (2004). "Those interfering little RNAs! Silencing and eliminating chromatin." *Curr Opin Genet Dev* 14(2): 174-80.
  - Segditsas, S. and I. Tomlinson (2006). "Colorectal cancer and genetic alterations in the Wnt pathway." *Oncogene* 25(57): 7531-7.
  - Sevignani, C., G. A. Calin, et al. (2007). "MicroRNA genes are frequently located near mouse cancer susceptibility loci." *Proc Natl Acad Sci U S A* 104(19): 8017-22.
  - Shao, J., Y. Li, et al. (2002). "High frequency loss of heterozygosity on the long arms of chromosomes 13 and 14 in nasopharyngeal carcinoma in Southern China." *Chin Med J (Engl)* 115(4): 571-5.
  - Shih, I. M., W. Zhou, et al. (2001). "Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis." *Cancer Res* 61(3): 818-22.
  - Si, M. L., S. Zhu, et al. (2007). "miR-21-mediated tumor growth." *Oncogene* 26(19): 2799-803.
  - Sijen, T., J. Fleenor, et al. (2001). "On the role of RNA amplification in dsRNA-triggered gene silencing." *Cell* 107(4): 465-76.
  - Singh, S. K., C. Hawkins, et al. (2004). "Identification of human brain tumour initiating cells." *Nature* 432(7015): 396-401.
  - Singhal, S., K. M. Amin, et al. (2003). "Differentially expressed apoptotic genes in early stage lung adenocarcinoma predicted by expression profiling." *Cancer Biol Ther* 2(5): 566-71.
  - Song, J. J., S. K. Smith, et al. (2004). "Crystal structure of Argonaute and its implications for RISC slicer activity." *Science* 305(5689): 1434-7.
  - Sonoda, E., H. Hohegger, et al. (2006). "Differential usage of non-homologous end-joining and homologous recombination in double strand break repair." *DNA Repair (Amst)* 5(9-10): 1021-9.
  - Sorlie, T., C. M. Perou, et al. (2001). "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications." *Proc Natl Acad Sci U S A* 98(19): 10869-74.
  - Sorlie, T., R. Tibshirani, et al. (2003). "Repeated observation of breast tumor subtypes in independent gene expression data sets." *Proc Natl Acad Sci U S A* 100(14): 8418-23.
  - Sotiriou, C., S. Y. Neo, et al. (2003). "Breast cancer classification and prognosis based on gene expression profiles from a population-based study." *Proc Natl Acad Sci U S A* 100(18): 10393-8.
  - Stemmann, O., H. Zou, et al. (2001). "Dual inhibition of sister chromatid separation at metaphase." *Cell* 107(6): 715-26.
  - Strom, L., H. B. Lindroos, et al. (2004). "Postreplicative recruitment of cohesin to double-strand breaks is required for DNA repair." *Mol Cell* 16(6): 1003-15.
  - Suarez, Y., C. Fernandez-Hernando, et al. (2008). "Dicer-dependent endothelial microRNAs are necessary for postnatal angiogenesis." *Proc Natl Acad Sci U S A* 105(37): 14082-7.
  - Suh, M. R., Y. Lee, et al. (2004). "Human embryonic stem cells express a unique set of microRNAs." *Dev Biol* 270(2): 488-98.
  - Sumara, I., E. Vorlauffer, et al. (2002). "The dissociation of cohesin from chromosomes in prophase is regulated by Polo-like kinase." *Mol Cell* 9(3): 515-25.
  - Tabara, H., A. Grishok, et al. (1998). "RNAi in *C. elegans*: soaking in the genome sequence." *Science* 282(5388): 430-1.
  - Tanaka, T., M. P. Cosma, et al. (1999). "Identification of cohesin association sites at centromeres and along chromosome arms." *Cell* 98(6): 847-58.
  - Tang, K. F., H. Ren, et al. (2008). "Decreased Dicer expression elicits DNA damage and up-regulation of MICA and MICB." *J Cell Biol* 182(2): 233-9.
  - Tatsuka, M., H. Katayama, et al. (1998). "Multinuclearity and increased ploidy caused by overexpression of the aurora- and lpl1-like midbody-associated protein mitotic kinase in human cancer cells." *Cancer Res* 58(21): 4811-6.
  - Taylor, S. S. and F. McKeon (1997). "Kinetochores localization of murine Bub1 is required for normal mitotic timing and checkpoint response to spindle damage." *Cell* 89(5): 727-35.
  - Taylor, S. S., E. Ha, et al. (1998). "The human homologue of Bub3 is required for kinetochore localization of Bub1 and a Mad3/Bub1-related protein kinase." *J Cell Biol* 142(1): 1-11.
  - Terada, Y., M. Tatsuka, et al. (1998). "AIM-1: a mammalian midbody-associated protein required for cytokinesis." *EMBO J* 17(3): 667-76.
  - Thacker, J. (2005). "The RAD51 gene family, genetic instability and cancer." *Cancer Lett* 219(2): 125-35.
  - Timmons, L. and A. Fire (1998). "Specific interference by ingested dsRNA." *Nature* 395(6705): 854.
  - Ting, A. H., H. Suzuki, et al. (2008). "A requirement for DICER to maintain full promoter CpG island hypermethylation in human cancer cells." *Cancer Res* 68(8): 2570-5.
  - Toth, A., R. Ciosk, et al. (1999). "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication." *Genes Dev* 13(3): 320-33.

- Turnell, A. S., G. S. Stewart, et al. (2005). "The APC/C and CBP/p300 cooperate to regulate transcription and cell-cycle progression." *Nature* 438(7068): 690-5.
- Tuschl, T. (2001). "RNA interference and small interfering RNAs." *Chembiochem* 2(4): 239-45.
- Tuschl, T., P. D. Zamore, et al. (1999). "Targeted mRNA degradation by double-stranded RNA in vitro." *Genes Dev* 13(24): 3191-7.
- Tzai, T. S., H. H. Chen, et al. (2003). "Clinical significance of allelotype profiling for urothelial carcinoma." *Urology* 62(2): 378-84.
- Uhlmann, F. and K. Nasmyth (1998). "Cohesion between sister chromatids must be established during DNA replication." *Curr Biol* 8(20): 1095-101.
- Unal, E., A. Arbel-Eden, et al. (2004). "DNA damage response pathway uses histone modification to assemble a double-strand break-specific cohesin domain." *Mol Cell* 16(6): 991-1002.
- van de Vijver, M. J., Y. D. He, et al. (2002). "A gene-expression signature as a predictor of survival in breast cancer." *N Engl J Med* 347(25): 1999-2009.
- van 't Veer, L. J., H. Dai, et al. (2002). "Gene expression profiling predicts clinical outcome of breast cancer." *Nature* 415(6871): 530-6.
- Vaucheret, H., A. C. Mallory, et al. (2006). "AGO1 homeostasis entails coexpression of MIR168 and AGO1 and preferential stabilization of miR168 by AGO1." *Mol Cell* 22(1): 129-36.
- Verdel, A., S. Jia, et al. (2004). "RNAi-mediated targeting of heterochromatin by the RITS complex." *Science* 303(5658): 672-6.
- Vermeulen, L., M. R. Sprick, et al. (2008). "Cancer stem cells--old concepts, new insights." *Cell Death Differ* 15(6): 947-58.
- Volpe, T. A., C. Kidner, et al. (2002). "Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi." *Science* 297(5588): 1833-7.
- Volpe, T., V. Schramke, et al. (2003). "RNA interference is required for normal centromere function in fission yeast." *Chromosome Res* 11(2): 137-46.
- Vos, L. J., J. K. Famulski, et al. (2006). "How to build a centromere: from centromeric and pericentromeric chromatin to kinetochore assembly." *Biochem Cell Biol* 84(4): 619-39.
- Waizenegger, I. C., S. Hauf, et al. (2000). "Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase." *Cell* 103(3): 399-410.
- Wang, Q., C. Moyret-Lalle, et al. (2003). "Alterations of anaphase-promoting complex genes in human colon cancer cells." *Oncogene* 22(10): 1486-90.
- Wang, Y., J. G. Klijn, et al. (2005). "Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer." *Lancet* 365(9460): 671-9.
- Wang, Y., R. Medvid, et al. (2007). "DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal." *Nat Genet* 39(3): 380-5.
- Wassenegger, M., S. Heimes, et al. (1994). "RNA-directed de novo methylation of genomic sequences in plants." *Cell* 76(3): 567-76.
- Watrin, E., A. Schleiffer, et al. (2006). "Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression." *Curr Biol* 16(9): 863-74.
- Weitzer, S., C. Lehane, et al. (2003). "A model for ATP hydrolysis-dependent binding of cohesin to DNA." *Curr Biol* 13(22): 1930-40.
- White, S. A. and R. C. Allshire (2004). "Loss of Dicer fowls up centromeres." *Nat Cell Biol* 6(8): 696-7.
- Wolpert, N., E. Warner, et al. (2000). "Prevalence of BRCA1 and BRCA2 mutations in male breast cancer patients in Canada." *Clin Breast Cancer* 1(1): 57-63; discussion 64-5.
- Wong, M. P., W. K. Lam, et al. (2002). "Primary adenocarcinomas of the lung in nonsmokers show a distinct pattern of allelic imbalance." *Cancer Res* 62(15): 4464-8.
- World Cancer Research Fund/American Institute for Cancer Research. Expert report. Food, nutrition and the prevention of cancer : a global perspective. Washington, DC : American Institute for Cancer Research, 1997.
- Xie, Z., L. K. Johansen, et al. (2004). "Genetic and functional diversification of small RNA pathways in plants." *PLoS Biol* 2(5): E104.
- Xie, Z., E. Allen, et al. (2005). "DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*." *Proc Natl Acad Sci U S A* 102(36): 12984-9.
- Yang, W. J., D. D. Yang, et al. (2005). "Dicer is required for embryonic angiogenesis during mouse development." *J Biol Chem* 280(10): 9330-5.
- Yang, J. and R. A. Weinberg (2008). "Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis." *Dev Cell* 14(6): 818-29.
- Yazdi, P. T., Y. Wang, et al. (2002). "SMC1 is a downstream effector in the ATM/NBS1 branch of the human S-phase checkpoint." *Genes Dev* 16(5): 571-82.
- Yi, R., Y. Qin, et al. (2003). "Exportin-5 mediates the nuclear export of pre-microRNAs and short hair-

- 
- pin RNAs." *Genes Dev* 17(24): 3011-6.
- Yi, R., D. O'Carroll, et al. (2006). "Morphogenesis in skin is governed by discrete sets of differentially expressed microRNAs." *Nat Genet* 38(3): 356-62.
  - You, Z., C. Chahwan, et al. (2005). "ATM activation and its recruitment to damaged DNA require binding to the C terminus of Nbs1." *Mol Cell Biol* 25(13): 5363-79.
  - Yu, F., H. Yao, et al. (2007). "let-7 regulates self renewal and tumorigenicity of breast cancer cells." *Cell* 131(6): 1109-23.
  - Yuan, J., A. Horlin, et al. (1997). "Polo-like kinase, a novel marker for cellular proliferation." *Am J Pathol* 150(4): 1165-72.
  - Zamore, P. D., T. Tuschl, et al. (2000). "RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals." *Cell* 101(1): 25-33.
  - Zhang, H., F. A. Kolb, et al. (2002). "Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP." *EMBO J* 21(21): 5875-85
  - Zhu, S., M. L. Si, et al. (2007). "MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1)." *J Biol Chem* 282(19): 14328-36..

## Annexe 1 : matériels et méthodes

### A1.1. Lignées cellulaires et traitements

Les lignées tumorales mammaires humaines BT-20, BT-474, Cal-51, Cama-1, HBL-100, HCC-1937, Hs578T, MCF-7, MCF-10A, MCF-12A, MDA-MB 151, MDA-MB 231, MDA-MB 361, MDA-MB 436, MDA-MB 453, MDA-MB 459, MDA-MB 549, SkBr-3, T47D, UACC-812 et ZR-75.1 et la lignée mammaire humaine immortalisée HME-1 sont issues de l'ATCC (American Type Culture Collection). Les lignées MN1 et MDD2 dérivées de MCF-7 par la transfection stable d'un plasmide vide ou d'un plasmide contenant un dominant négatif du gène *tp53* respectivement ont généreusement été données par le Dr. M Oren. La lignée tumorale utérine HeLa est issue de l'ATCC. Chaque lignée a été cultivée suivant les recommandations du fournisseur (**Annexe 2, détail des milieux de culture**) à 37°C sous 5% de CO<sub>2</sub> en atmosphère humide.

La synchronisation des cellules HeLa pour l'étude du cycle cellulaire a été réalisée par double blocage à la thymidine. Brièvement, au minimum 24h après l'ensemencement, le milieu de culture est complété par une solution fraîche de thymidine (T1895, Sigma-Aldrich) à 2mM final. 16h après, les cellules sont lavées au PBS et du milieu de culture contenant 24µM de déoxycytidine (D0776, Sigma-Aldrich) est ajouté pendant 9h afin de lever le blocage (fin de phase G1, début de phase S) initié par la thymidine. Un deuxième blocage est réalisé dans les mêmes conditions afin d'obtenir un maximum de cellules synchronisées. Elles sont alors récoltées à différents temps après la levée de ce deuxième blocage.

L'inactivation de l'expression du gène *dicer* est réalisée par transfection transitoire de siRNA ciblant l'ARN messager de ce gène. Pour ce faire, 24h après l'ensemencement, un mélange de transfection contenant les siRNA, le transfectant lipidique (jetSI-endo, PolyplusTransfection) et un milieu minimum (Optimem, Gibco) est préparé et ajouté aux cellules elles aussi en milieu minimum. La concentration finale des siRNA est de 50nM. Après 4h, du milieu de culture contenant deux fois plus

de SVF est ajouté en quantité égale au milieu minimum déjà présent dans les boîtes. Cette transfection est répétée le lendemain pour augmenter son efficacité tout en évitant l'activation non souhaitée de voies cellulaires en réponse à une trop forte dose de siRNA.

Afin d'induire un stress réplicatif et ainsi des cassures double-brin de l'ADN, les cellules HeLa, T47D ou MDA-MB 231 sont traitées à l'hydroxy-urée (HU, H8627, Sigma Aldrich). La concentration finale est de 10 mM. Après un temps plus ou moins long de traitement, les cellules sont trypsinées et préparées sous forme de culots secs en vue d'une extraction protéique.

Le traitement par irradiation des cellules HeLa générant des cassures double-brin de l'ADN a consisté à exposer ces cellules aux rayons X à une dose de 10 grey puis à les récupérer à différents temps sous forme de culots secs pour en extraire les protéines.

## A1.2. Tumeurs mammaires

104 échantillons de biopsies tumorales mammaires ont été fournis par le Dr. E. TABONE du Centre de Ressource Biologique du Centre Léon Bérard de Lyon. Après avoir été réduits en poudre par cryobroyage, l'ARN de ces échantillons a été extrait par la méthode détaillée ci-après. Les données cliniques et moléculaires concernant ces tumeurs sont présentées en **annexe A3.1.**

120 échantillons de tumeurs mammaires invasives non métastatiques ont été techniquées sous forme de Tissu Micro Array (TMA) par le service d'anatomopathologie du Centre Léon Bérard (pour plus de détails, se référer à l'article). Les données cliniques et moléculaires concernant ces tumeurs sont présentées en **annexe A3.2.**

### A1.3. Extraction d'ARN et Reverse Transcription

L'extraction d'ARN total a été réalisée avec une méthode phénol-chloroforme à l'aide du tampon de lyse TriReagent (T-9424, Sigma-Aldrich) et des tubes PhaseLockGel (230-2830, Eppendorf) pour la séparation des phases. Les échantillons sont dosés par spectrophotométrie. Après traitement à la DNase (776 785, Roche Applied Science), la contamination ADN a été vérifiée sur gel d'agarose.

La synthèse des ADNc a été effectuée avec le kit First-Strand cDNA Synthesys (27926101, GE-Healthcare). Brièvement, 1 µg d'ARN dénaturé a été homogénéisé avec 1 µL de DTT à 200 mM, 1 µL d'amorces hexameriques aléatoires à 0,2 µg/µL et un mix contenant, entre autre, la murine reverse transcriptase MMLV. Après 1h d'incubation à 37°C, la réaction a été stoppée en plaçant les échantillons 3 min à 95°C.

### A1.4. PCR en temps réel

Les amorces utilisées pour les différentes PCR en temps réelle sont détaillées en **Annexe A5.1. PCR en temps réelle : tableau des amorces**. Toutes les réactions ont été réalisées sur LightCycler 2.0 (Roche Applied Science). Deux kits de préparation différents ont été utilisés suivant le type d'amorces :

- LightCycler Taqman Master pour les amorces avec sondes du système UPL (voir **Annexe 5**),
- LightCycler FastStart DNA Master PLUS SYBR Green I pour les amorces sans sonde.

Le mix de réaction est composé de FastStart DNA polymerase et son tampon, de deoxynucleoside triphosphate, de MgCl<sub>2</sub> à 180 µM, de 0,25 µM de chaque amorce (et de la sonde spécifique le cas échéant) et de cDNA dilué au 1/60 après RT. Pour les contrôles

négatifs, le cDNA est remplacé par de l'eau. Les programmes d'amplification utilisés sont décrits en **Annexe A5.2. PCR en temps réelle : programmes d'amplification.**

L'efficacité de chaque PCR a été déterminée, optimisée et validée au niveau de la reproductibilité par la réalisation de courbes standard. De l'ADNc de la lignée mammaire immortalisée HME-1 a été utilisé en dilution sérielle au dixième de 1/6 à 1/60000 et la pente de la droite obtenue permet de connaître l'efficacité. Par la suite, l'ADNc HME-1 dilué au 1/60 sert de calibrateur, c'est-à-dire qu'il est inclus dans toutes les PCR afin de savoir si le "run" s'est déroulé dans les conditions optimales (les valeurs de cycle de sortie [Cp pour Crossing point] du calibrateur ne doivent pas varier entre les PCR). Toutes les PCR ont été analysées avec le LightCycler Software 4.0 (Roche Applied Science) et les Cp obtenus transférés sous Excel (Microsoft) pour être normalisés suivant la formule suivante :

$$N = E_T^{CpT(C)-CpT(S)} \times \prod \left( E_R^{CpR(S)-CpR(C)} \right)$$

où N est le ratio normalisé, E l'efficacité, Cp (Crossing point) le nombre de cycle de sortie, T le gène cible, R un gène de référence, C le calibrateur et S l'échantillon mesuré. Les échantillons ont donc été normalisés par la moyenne géométrique du ratio des valeurs de cycle de sortie de trois gènes de référence *cyclophilin B* (PPIB), *β-actine* (ACTB) et *phosphoglyceroKinase 1* (PGK-1). Ces gènes ont été sélectionnés parmi 12 pour leur stabilité dans les échantillons à mesurer (lignées et tumeurs mammaires) grâce à l'applet geNorm (<http://medgen.ugent.be/~jvdesomp/genorm/>).

## A1.5. Western Blotting

Les culots cellulaires ont été extraits à l'aide de différents tampons de lyse suivant le type de protéines à favoriser. Ainsi, pour observer des protéines phosphorylées, le tampon utilisé est le TEB150 (50 mM HEPES, 150 mM NaCl, 2 mM MgCl<sub>2</sub>, 5 mM



EGTA, 1 mM DTT, 0,5% Triton X-100 (v/v), 10% glycerol, 1 mM Na<sub>3</sub>VO<sub>4</sub>, PIC 1 et PIC 2 (P2850 et P5726, Sigma-Aldrich)). Après estimation du volume de tampon à utiliser en fonction de la taille du culot cellulaire, les lysats sont laissés 30 min dans la glace puis centrifugés 30 min à 13000 rpm à 4°C. Les surnageants sont alors transférés dans un nouveau tube et dosés. Les protéines histones ont été extraites avec du tampon TEB (0,5% Triton X-100 (v/v), 2mM phenylmethylsulfonyl fluoride (PMSF), 0,02% (w/v) NaN<sub>3</sub> dans du PBS 1X) suivi d'une extraction acide sur la nuit avec du HCl 0,2N. Les lysats sont centrifugés 10 min à 2000 rpm à 4°C, les surnageants récupérés et dosés. Le dosage est réalisé par la méthode de Bradford à l'aide du réactif Bio-Rad Protein Assay (500-0006, Bio-Rad) suivant les recommandations du fournisseur.

Pour la plupart des western blot, 100 µg de protéines totales de chaque échantillon sont déposés sur gel après avoir été homogénéisés avec du XT Sample buffer, du XT Reducing Agent (161-0791 et 161-0792, Bio-Rad) et du tampon de lyse (qsp 20 µL) et dénaturés 5 min à 95°C. Pour les protéines histones, les gels sont dénaturants et composés de 15% d'acrylamide. Du tampon Tris-Glycine SDS a été utilisé pour la migration. Pour les autres protéines, les gels utilisés sont des gels Critériums gradients 4-12% Tris-HCl Glycine (345-0124, Bio-Rad). Du tampon XT MES (161-0789, Bio-Rad) a été utilisé pour la migration. Le marqueur de taille utilisé est soit le Multimark Multi-Colored Standard (LC5725, Invitrogen), soit le Precision Plus Protein Standards (161-0374, Bio-Rad).

Le transfert a été réalisé sur membranes 0,45 µm PVDF pour les histones (Immobilon-P, Milipore) ou nitrocellulose (Trans-Blot 162-0115, Bio-Rad) sur la nuit en tampon Tris-glycine avec 20% d'éthanol. A la fin du transfert les membranes sont saturées 15 min dans du TBS-Tween 0,05% avec 5% de lait puis incubées avec l'anticorps primaire (**Annexe 6, tableau des anticorps**).

---

## A1.6. Cytométrie en flux

Les cellules en culture destinées à être analysées par cytométrie en flux ont été mises en présence de BrdU 10  $\mu$ M final pendant 30 min. Après trypsination et lavages au PBS 1X - BSA 1%, elles ont été fixées au minimum 30 min dans de l'éthanol 70% froid. Par la suite, les cellules ont été centrifugées et re-suspendues en PBS 1X avant d'être perméabilisées pendant 30 min dans du HCl 4N à température ambiante. L'effet du HCl a été stoppé par un passage de 10 min en borax 0,1 M. Après centrifugation, les cellules ont été lavées puis re-suspendues dans du PBS 1X - Tween 20 0,5% pour une incubation de 20 min avec un anticorps anti-BrdU (247583, Becton Dickinson). Après lavages, une autre incubation a été réalisée avec un anticorps secondaire FITC (F0261, Dako) au 1/150 pendant 20 min. Enfin, une fois les cellules lavées, elles ont été re-suspendues dans 1 mL d'une solution PBS 1X - Iodure de Propidium 20  $\mu$ g/mL (P4170, Sigma Aldrich) - RNAse 1 mg/mL (10109142001, Roche Applied Science). Les échantillons sont alors prêts à être analysés sur le cytomètre de flux (FACs).

## Annexe 2 : milieux de culture

MEM	McCoy's 5A	RPMI 1640	DMEM		DMEM/Ham's F12 1:1
Péni/Strepto 2% SVF 10% Sodium pyruvate 1.0 mM NEAA 0.1 mM	Péni/Strepto 2% SVF 10%	Péni/Strepto 2% SVF 10% Sodium pyruvate 1.0 mM	Péni/Strepto 2% SVF 10%	SVF 20%	Péni/Strepto 2% Horse serum 5% Insulin 0.01 mg/ml EGF 20 ng/ml Hydrocortisone 500 ng/ml Cholera toxin 100 ng/ml
BT-20 CAMA-1 MCF-7 + Insulin 0.01 mg/ml	HBL-100 SK-BR3	HCC1937 ZR-75-1 T47D + Insulin 0.01 mg/ml	MDA-MB 157 MDA-MB 231 MDA-MB 453 MDA-MB 459 MDA-MB 549 Hs 578T + Insulin 0.01 mg/ml MDA-MB 436 + Insulin 0.01 mg/ml + Glutathione 16µg/mL BT-474 + Sodium pyruvate 1mM + Oxaloacetic acid 1.2 mM + Insulin 0.01 mg/mL + NCTC 135 medium 10% (v/v)	Cal-51 Lat-CLB 7mois MDA-MB 361 ZAN-CLB UACC 812 + EGF 20 ng/ml	MCF-10A MCF-12A

merci au Dr. Benjamin BOUCHET pour son aide précieuse dans la culture de ces différentes lignées et pour la réalisation de ce tableau.

## Annexe 3 : données cliniques des tumeurs mammaires humaines étudiées

## A3.1. Étude par PCR en temps réelle

N° CRB	Ménopause	Méta d'emblée	Méta à 3 ans	Méta à 5 ans	pT	Env. gang.	Grade SBR	ER	PgR	Ploïdie	origine	nb de mitoses
52	2	non	non	non	T3	-	lob	+	+		2	
56	2	non	non	non	T2	+	2	+	+		1	
62	1	oui	oui	oui	T2	+	2	+	-		1	
88	1	non	non	non	T3	+	lob	+	+		2	
97	1	non	non	non	T2	+	1	+	+		1	1
105	2	oui	oui	oui	T3	+	2				1	1
143	-9	oui	oui	oui	T3	+	3	-	-		1	2
149	2	non	oui	oui	T2	+	2	+	+		1	2
151	2	non	oui	oui	T2	+	1	+	+		1	1
155	1	oui	oui	oui	T2	+	2	+	+		1	2
164	1	oui	oui	oui	T2	+	1	+	+		1	1
169	2	non	non	non	T2	+	3	+	+		1	3
174	1	non	non	non	T1	+	3	+	+		1	2
210	1	non	non	non	T2	+	2	+	+		1	1
213		non	non	non		-		-	-	D	5	
245	1	non	non	non	T2	+	2	+	+		1	1
260	1	oui	oui	oui	T3		2	+	+		1	1
266	1	oui	oui	oui	T2	+	3	+	+		1	
273	1	non	non	non	T3	-		-	-	A	3	
285	55 ans	non	oui	oui	T2	+	3	+	+		1	3
291	54 ans	oui	oui	oui		+		-	-	A	1	
295	2	non	non	non	T3	+	lob	-	+		2	
322	2	non	non	non	T3	+	2	-	+	A	1	2
332	1	non	oui	oui	T3	+	2	-	+		1	1
344	1	non	non	non	T2	+	2	+	+		1	1
350	3	non	non	non	T2	+	2	+	+		1	3
359	2	oui	oui	oui		+	3	+	+	A	1	3
375	1	non	non	non	T2	+	2	+	+	A	1	2
382	2	non	non	non	T3	+	2	+	+	A	1	2

N° CRB	Ménopause	Méta d'emblée	Méta à 3 ans	Méta à 5 ans	pT	Env. gang.	Grade SBR	ER	PgR	Ploïdie	origine	nb de mitoses
385	2	non	non	non	T2	+	2	+	+	A	1	1
388	2	non	non	non	T3	+	2	+	+	A	1	1
400	1	non	non	non	T1	-	2	+	+	A	1	1
416	2	non	non	non	T2	+	1	+	+	A	3	1
427	1	non	non	non	T3	+	3	+	+	A	1	2
429	1	non	non	non	T2	+	2	+	+		1	2
435	2	non	non	non	T2	+	3	-	-		1	2
442	2	non	non	non	T1	+	2	+	+		1	1
443	50 ans	non	non	non	T3	-	lob		+		1	
456	2	non	oui	oui	T3	+	2	+	-	A	1	2
460	2	non	non	non	T2	+	3	+	+		1	3
472	1	non	oui	oui	T3	+	2	+	+		1	1
474	1	non	non	non	T2	+	2	+	+	A	1	1
483	1	non	non	non	T3	+	1	+	+		1	1
486	1	non	non	non	T2	+	3	+	+	A	1	3
491	2	non	oui	oui	T3	+	3		+	A	1	2
499	2	non	non	non	T1	+	1	+	+		1	1
503	2	non	oui	oui	T3	+	3	-	+		1	3
511	1	non	non	oui	T2	+	4	+	-		1	2
515	1	non	non	non	T3	+	2	-	+	A	1	2
521	2	non	non	non		+				A	4	
524	1	non	non	non	T2	-	1	+	+	D	1	1
526	2	non	non	non	T3	+	2	-	+	A	1	2
532	-9	oui	oui	oui	T2	+	3	+	+		1	3
535	2	non	non	oui	T3	+	2	+	+		1	1
540	1	oui	oui	oui	T3	+	2	+	+	A	2	2
541	1	oui	oui	oui	T2	+	3	+	-	A	1	3
558	4	non	oui	oui	T2	+	3	+	+		1	2
616	1	non	non	non	T2	+	2	+	+		1	3
633	1	non	non	non	T2	+	1	+	+		1	2
638	2	non	non	non	T3	+	3	+	+		1	3
664	1	non	non	non	T1	+	3	-	-		1	3
669	1	non	non	oui	T1	+	1	+	-		1	1
671	1	non	non	non	T2	+	1	+	+		1	1
673	2	non	non	non	T2	+	3	-	+		1	3
676	1	non	oui	oui	T1	+	3	+	+		1	3
683	1	non	oui	oui	T2	+	2	+	-		1	3
695	1	non	oui	oui	T2	+	2	+	-		1	1
706	1	oui	oui	oui	T3	+	2	+	-		1	2
725	1	non	non	non	T2	+	2	+	+		1	1
743	1	non	non	non	T3	+	3	+	-		1	3
754	1	non	non	non	T2	+	2	+	+		1	2
803	2	non	non	non	T2	+	3	+	+		1	3
810	1	non	non	oui	T1	+	3	+	+		1	2
814	1	non	non	non	T2	-	3	+	-	A	1	2
816	2	non	non	non	T2	+	2	+	+		1	1
826	1	non	non	non	T2	-	3	-	-	A	1	3
852	1	non	non	non	T2	-	2	+	+	A	1	1
853	1	non	non	non	T2	+	3	+	+		1	3
914	1	non	non	non	T1	-	2	+	+		?	1
943	1	non	non	non	T1	+	3	+	+		1	3
945	2	non	oui	oui	T1	+	3	+	+		1	2
952	1	non	non	oui	T2	+	3	+	+		1	2
961	1	non	non	non	T2	+	3	+	-		1	2
963	1	non	non	non	T2	+	1	+	+		1	1
1018	1	non	oui	oui	T2	+	3	-	+		1	3
1019	2	non	oui	oui	T2	+	3	+	+		1	3
1046	1	non	non	non	T2	+	2	+	+		1	2

N° CRB	Ménopause	Méta d'emblée	Méta à 3 ans	Méta à 5 ans	pT	Env. gang.	Grade SBR	ER	PgR	Ploïdie	origine	nb de mitoses
1064	1	oui	oui	oui	T2	+	3	+	+		1	3
1073	1	non	oui	oui	T3	+	3	+	+		1	3
1081	1	non	non	non	T1	+	3	+	+		1	2
1089	2	oui	oui	oui	T2	+	2	+	+	A	1	1
1093	1	non	non	non	T2	-	3	+	-	A	1	2
1119	3	non	non	non	T1	-	3	-	-	A	1	2
1159	2	non	non	non	T2	-	3	-	-		1	3
1248	2	non	non	non	T2	+	1	+	+		1	1
1271	1	non	non	non	T2	-	3	+	+	D	3	3
1279	1	non	non	non	T1	+	2	+	+		1	2
1306	1	non	non	non	T2	+	3	-	-		1	3
1317	1	non	non	non	T2	+	2	+	+		3	2
1331	2	non	non	non	T2	-	2	+	+		1	2
1523	2	non	non	non	T2	+	2	+	+		2	
1589	2	oui	oui	oui	T3	+	3	-	-		1	3
1652	1	oui	oui	oui	T3	+	3	+	+		2	
1981	3		non	non	T3	-	2	+	+		2	1
2103	1	non	non	non	T3	+	3	+	+		1	3
2171		non	non	non		-		-	-	D	5	
2262	2	non	non	non	T3	+	2	+	+		1	1
2266	-9	non	non	non	T3	+	3	-	-		1	3
2268	1	non	non	non	T3	+	3	-	-		1	3
2501	1	non	non	non	T2	-	3	-	-		1	3

Ménopause : 1 = oui ; 2 = non ; 3 = hystérectomie ; 4 = période ménopausique

pT (taille de la tumeur) : T1 ≤ 20mm ; 20mm < T2 ≤ 50mm ; T3 > 50mm

ploïdie : D = diploïde ; A = aneuploïde

Origine : 1 = canalaire ; 2 = lobulaire ; 3 = autre ; 4 = *in situ* ; 5 = bénin

nombre de mitoses (composante SBR) : 1 = "≤ 9" ; 2 = "10 à 19" ; 3 = "≥ 20"

### A3.1. Étude par TMA

voir pages suivantes.

n° TMA	Mén	Méta d'emblée	Méta à 3 ans	Méta à 5 ans	pT	Env gang	Grade SBR	%-ER	I-ER	%-PgR	I-PgR	ori	nb de mitose	% DICER	I DICER	% ATM	I ATM	% MRE11	I MRE11	% CERB2	I CERB2	% cell tum
T1	2	non	non	non	T1	-	3	3	2	3	2	1	3	80	1	80	1	90	2	0	0	60
T10	1	non	non	non	T1	-	2	5	2	5	1	1	2	80	1	60	1	90	1	0	0	40
T101	2	non	non	non	T3	+	2	5	3	5	2	3	1	80	1	40	1	90	2	0	0	-9
T102	1	non	non	non	T1	-9	1	5	2	5	3	1	1	80	1	70	1	90	2	0	0	-9
T104	1	non	non	non	T2	+	3	5	2	5	2	1	3	0	0	60	1	90	2	0	0	-9
T105	1	non	non	non	T2	+	3	4	2	3	3	1	3	40	1	60	1	60	2	40	-9	-9
T106	1	non	non	non	T2	-	3	3	1	1	-9	1	2	80	1	20	1	-9	-9	80	-9	60
T107	1	non	non	non	T1	-	1	5	3	4	2	1	1	60	1	80	1	90	2	60	-9	-9
T108	1	non	non	non	T2	+	3	4	1	4	1	1	3	-9	-9	60	1	90	1	0	0	60
T108	1	non	oui	oui	T2	+	3	4	1	4	1	1	3	-9	-9	60	1	90	1	0	0	-9
T109	1	non	non	non	T1	-	1	5	3	5	3	1	1	70	2	60	1	90	2	0	0	-9
T11	2	non	non	non	T1	-	2	2	-9	5	3	2	1	-9	-9	-9	-9	-9	-9	0	0	20
T110	1	non	non	non	T1	+	2	5	2	5	2	1	1	-9	-9	60	1	-9	-9	0	0	-9
T111	1	non	non	non	T1	-	1	5	2	5	2	1	1	40	2	80	1	90	2	0	0	-9
T112	1	non	non	non	T1	-	3	4	1	3	1	1	3	-9	-9	-9	-9	-9	0	0	0	-9
T113	1	non	non	non	T1	+	2	5	2	3	2	2	1	50	1	70	1	90	2	0	0	-9
T114	2	non	non	non	T1	-	3	4	1	3	2	1	3	60	1	90	2	90	2	0	0	-9
T115	1	non	non	non	T1	-	3	4	2	2	-9	1	3	40	2	90	2	90	3	0	0	-9
T116	2	non	non	non	T2	+	2	5	2	4	1	1	1	80	1	60	1	80	1	0	0	-9
T117	1	non	non	non	T1	+	1	5	3	4	2	1	1	80	1	70	1	80	3	0	0	-9
T118	1	non	non	non	T1	-	3	5	2	4	2	1	2	60	2	-9	-9	90	2	0	0	-9
T119	1	oui	non	non	T3	-	3	1	-9	1	-9	3	3	70	1	0	0	90	1	0	0	-9
T12	2	non	non	non	T1	-	1	5	2	5	1	1	1	-9	-9	70	1	90	1	0	0	-9
T120	3	non	non	non	T1	-	2	4	2	4	3	1	2	70	1	70	1	50	2	0	0	-9
T13	2	non	non	non	T2	+	2	5	2	5	3	1	2	80	1	60	1	90	2	0	0	-9
T14	1	non	non	non	T2	+	2	5	3	3	1	1	1	0	0	60	1	90	2	0	0	-9
T16	2	non	non	non	T2	-	3	4	1	5	3	1	2	60	1	30	1	90	2	80	-9	-9
T18	1	non	oui	oui	T3	-	3	1	-9	2	-9	1	3	90	2	60	2	70	1	0	0	-9
T19	1	oui	non	non	T2	-9	2	5	3	4	2	1	1	70	1	-9	-9	90	2	0	0	-9
T2	2	non	non	non	T1	+	3	2	-9	2	-9	3	3	80	2	70	1	90	2	0	0	40
T20	1	non	non	non	T1	+	2	4	2	4	2	1	1	80	1	20	1	90	2	0	0	-9
T21	1	non	non	non	T2	-	2	5	3	5	2	1	1	70	1	50	1	90	2	0	0	-9
T22	2	non	non	non	T1	-	3	2	-9	2	-9	1	3	70	1	70	1	90	2	0	0	50
T23	2	non	non	oui	T2	+	2	5	2	5	2	1	3	80	1	40	1	90	2	0	0	35

n° TMA	Mén	Méta d'emblée	Méta à 3 ans	Méta à 5 ans	pT	Env gang	Grade SBR	%-ER	I-ER	%-PgR	I-PgR	ori	nb de mitose	% DICER	I DICER	% ATM	I ATM	% MRE11	I MRE11	% CERB2	I CERB2	% cell tum
T24	1	non	non	non	T2	+	2	5	3	4	3	1	1	80	1	0	0	0	2	0	0	30
T25	1	non	non	non	T2	+	3	5	2	5	2	1	3	60	1	0	0	0	2	50	-9	50
T26	1	oui	non	non	T2	-9	3	5	3	3	2	1	3	60	1	0	0	0	1	0	0	-9
T27	2	non	non	non	T1	-	1	5	2	4	2	1	1	-9	-9	-9	-9	-9	-9	0	0	-9
T28	1	non	non	non	T1	-	1	5	3	2	-9	1	1	60	2	0	0	0	2	0	0	60
T29	4	non	non	non	T2	+	2	5	2	2	-9	2	1	40	1	0	0	0	2	0	0	-9
T3	1	oui	non	non	T1	+	2	5	3	3	2	1	1	0	0	70	1	0	3	0	0	60
T30	1	non	non	non	T2	-	1	5	3	5	3	1	1	60	1	0	0	0	2	0	0	-9
T31	2	non	non	non	T1	+	3	5	2	4	2	1	3	40	2	0	0	0	1	0	0	-9
T32	1	non	non	non	T1	-	2	5	2	4	2	1	1	40	2	10	1	0	2	0	0	60
T34	1	non	non	non	T2	-	2	5	3	5	2	3	1	70	1	0	0	0	2	0	0	40
T35	1	non	non	non	T1	-	2	5	2	3	1	1	1	60	1	0	0	0	2	0	0	-9
T36	1	non	non	non	T1	-	2	5	2	5	1	2	1	60	1	30	1	0	2	0	0	40
T38	2	non	non	non	T2	+	2	5	2	4	2	1	1	-9	-9	0	0	0	2	0	0	60
T4	1	non	non	non	T2	-	2	5	1	5	3	1	1	60	1	80	1	0	2	0	0	70
T40	2	non	non	non	T2	-	3	2	-9	2	-9	1	3	80	2	0	0	0	1	0	0	60
T41	1	oui	non	non	T3	-9	2	2	-9	2	-9	1	2	60	1	60	1	0	1	0	0	-9
T42	1	non	non	non	T1	+	2	5	2	3	2	1	2	30	2	0	0	0	2	0	0	-9
T43	1	non	non	oui	T2	+	3	4	2	3	1	1	2	70	1	-9	-9	2	60	-9	-9	-9
T44	1	non	non	non	T1	-	1	4	1	4	1	1	1	-9	-9	40	1	-9	0	0	0	-9
T45	2	non	non	non	T2	+	3	5	3	4	3	1	3	70	1	50	1	0	1	0	0	70
T46	2	non	non	non	T2	+	3	2	-9	2	-9	1	3	60	2	30	1	0	2	0	0	-9
T47	1	non	non	non	T2	+	3	5	3	2	-9	1	2	70	1	0	0	0	2	30	-9	60
T48	2	non	non	non	T1	-	3	2	-9	2	-9	3	3	70	1	60	1	0	2	0	0	-9
T49	1	non	non	non	T2	-	2	5	2	3	3	1	2	30	1	40	1	0	2	0	0	70
T50	1	non	non	non	T1	-	2	5	3	3	2	1	3	40	2	0	0	0	3	0	0	-9
T51	1	non	oui	oui	T2	+	3	2	-9	2	-9	1		80	2	70	2	0	1	80	-9	60
T52	1	non	non	non	T2	-	1	5	3	3	3	1	1	40	1	80	1	0	2	0	0	-9
T53	1	non	non	oui	T2	+	3	5	3	4	2	1	3	0	0	80	1	0	2	0	0	50
T55	1	non	non	non	T1	+	2	5	2	4	2	1	2	50	1	60	2	0	2	0	0	30
T56	1	non	non	non	T1	-	2	4	2	3	1	2	1	0	0	60	1	0	1	0	0	-9
T59	2	non	non	non	T1	+	1	5	3	4	2	1	1	70	1	30	1	0	2	0	0	40
T6	1	non	non	non	T2	+	3	5	2	5	3	1	3	60	1	70	1	0	2	0	0	70
T6	1	non	non	non	T1	+	2	5	2	4	2	1	2	60	1	70	1	0	2	0	0	-9

n° TMA	Mén	Méta d'emblée	Méta à 3 ans	Méta à 5 ans	pT	Env gang	Grade SBR	%-ER	I-ER	%-PgR	I-PgR	ori	nb de mitose	% DICER	I DICER	% ATM	I ATM	% MRE11	I MRE11	% CERB2	I CERB2	% cell tum
T61	1	non	oui	oui	T3	-9	2	5	2	4	3	2	1	0	0	60	1	90	2	0	0	10
T62	1	non	non	non	T1	-	2	2	-9	2	-9	1	2	40	1	40	2	60	2	0	0	-9
T63	2	non	non	non	T2	+	2	5	3	4	2	1	1	50	1	80	2	90	2	0	0	35
T64	2	non	non	non	T1	+	1	4	1	5	2	2	1	30	1	60	1	90	3	0	0	-9
T65	2	non	non	non	T1	-	3	5	2	5	2	1	3	60	1	0	0	90	1	80	-9	-9
T66	2	non	non	non	T1	-	1	4	1	5	2	1	1	80	1	80	1	90	2	0	0	35
T67	2	non	non	non	T1	+	2	4	1	5	2	1	1	0	0	0	0	-9	-9	0	0	-9
T68	1	oui	non	non	T3	+	3	5	1	3	2	2	3	60	1	0	0	90	2	0	0	-9
T7	4	non	non	non	T1	-	2	5	1	3	2	1	2	30	1	70	1	90	1	0	0	-9
T70	2	non	oui	oui	T2	-	3	2	-9	2	-9	1	3	50	2	80	1	90	2	0	0	60
T71	2	non	non	non	T1	+	1	4	1	4	2	1	1	70	1	0	0	90	1	0	0	-9
T72	1	non	non	non	T1	+	2	5	3	4	2	1	2	80	1	70	1	80	2	0	0	60
T74	1	non	non	non	T2	+	2	5	3	4	2	1	2	80	1	40	1	90	2	0	0	-9
T75	1	non	non	non	T1	-	2	5	1	5	2	1	1	80	1	20	1	80	1	0	0	-9
T77	1	non	non	non	T2	+	2	5	2	5	2	1	2	40	2	50	1	90	2	0	0	70
T79	2	non	non	non	T2	+	3	5	2	5	2	3	3	70	1	0	0	90	2	0	0	-9
T8	1	non	oui	oui	T1	-	3	2	-9	2	-9	1	3	80	1	-9	-9	1	0	0	0	30
T80	3	non	non	non	T2	-	2	5	2	5	2	1	1	80	1	0	0	90	2	0	0	-9
T81	2	non	non	non	T1	+	3	4	1	2	-9	1	2	60	1	60	1	80	2	0	0	-9
T82	2	non	non	non	T1	-	3	2	-9	2	-9	1	3	70	1	30	1	-9	-9	0	0	-9
T83	1	non	oui	oui	T2	-	3	5	3	2	-9	1	3	70	1	-9	-9	1	-9	-9	0	-9
T84	1	oui	oui	oui	T2	+	3	3	1	2	-9	3	3	70	1	10	1	80	2	50	-9	-9
T85	1	non	non	non	T2	-	2	5	3	4	1	2	1	0	0	0	0	90	1	-9	-9	-9
T86	1	non	non	non	T2	-9	2	2	-9	2	-9	1	1	70	1	70	1	90	3	0	0	-9
T87	1	non	non	non	T2	-	3	2	-9	2	-9	1	3	50	2	0	0	70	2	0	0	-9
T88	1	non	non	non	T2	+	3	2	-9	2	-9	3	2	60	2	70	1	90	2	0	0	-9
T89	1	non	non	non	T1	-	2	5	3	5	2	1	1	40	2	60	1	90	2	0	0	-9
T9	1	non	non	non	T1	-9	1	4	2	5	3	1	1	0	0	80	1	90	2	0	0	-9
T90	1	non	non	non	T2	-	3	5	2	4	3	1	2	80	1	0	0	90	2	0	0	-9
T91	1	non	non	non	T1	-	2	2	-9	1	-9	1	2	80	1	-9	-9	-9	-9	80	-9	-9
T92	2	non	non	non	T1	-	2	4	1	4	2	1	2	-9	-9	-9	-9	-9	-9	0	0	-9
T93	1	non	non	non	T2	+	3	4	1	3	1	1	3	80	1	70	1	80	2	-9	-9	50
T94	2	oui	non	non	T2	+	3	5	1	5	2	1	2	80	1	70	1	90	2	-9	-9	-9
T96	1	oui	non	non	T1	+	3	5	2	3	3	1	2	-9	-9	0	0	90	1	0	0	70



n° TMA	Mén	Méta d'emblée	Méta à 3 ans	Méta à 5 ans	pT	Env gang	Grade SBR	%-ER	I-ER	%-PgR	I-PgR	ori	nb de mitose	% DICER	I DICER	% ATM	I ATM	% MRE11	I MRE11	% CERB2	I CERB2	% cell tum
T97	2	non	non	non	T1	-	3	4	2	4	2	1	3	40	2	30	1	80	2	60	-9	
T98	3	non	non	non	T1	-	2	5	2	5	3	1	1	60	2	30	1	60	3	0	-9	
T99	2	non	non	non	T1	+	2	5	1	3	2	1	1	80	1	0	0	80	2	0	-9	

Ménopause : 1 = oui ; 2 = non ; 3 = hystérectomie ; 4 = période ménopausique

pT (taille de la tumeur) : T1 ≤ 20mm ; 20mm < T2 ≤ 50mm ; T3 > 50mm

Origine : 1 = canalaire ; 2 = lobulaire ; 3 = autre ; 4 = *in situ* ; 5 = bénin

nombre de mitoses (composante SBR) : 1 = "≤ 9" ; 2 = "10 à 19" ; 3 = "≥ 20"

## Annexe 4 : données sur les lignées tumorales mammaires humaines étudiées

<i>Lignée</i>	<i>métastatique</i>	<i>source</i>	<i>type</i>	<i>âge</i>	<i>Tumeurs nude</i>	<i>C-myc Amp</i>	<i>p53</i>	<i>ER</i>	<i>PR</i>	<i>ERBB2 Amp</i>
BT-20	non	primaire	mixte	>40	oui	oui	mut	-	-	non
BT-474	non	primaire	épithélial	>40	oui	oui	mut	+	+	oui
Cal-51	oui	metasta	épithélial	>40	oui	oui	wt	-9	-	non
Cama-1	oui	metasta	épithélial	>40	oui	non	mut	+	-	non
HBL-100	oui	tissu	-9	<40	non	non	wt	-	-	-9
HCC-1937	non	primaire	épithélial	<40	-9	oui	wt	-	-	-9
HME-1	non	tissu	-9	-9	non	oui	wt	-9	-9	non
HMEC	non	tissu	-9	>40	non	non	wt	-9	-9	-9
HS-578T	non	primaire	mesen	>40	non	oui	mut	-	-	non
MCF-7	oui	metasta	épithélial	>40	oui	oui	wt	+	+	non
MN1	oui	metasta	épithélial	>40	oui	oui	wt	-9	-9	non
MDD2	oui	metasta	épithélial	>40	oui	oui	mut	-9	-9	non
MCF-10A	non	tissu	mixte	<40	non	-9	-9	-	-	non
MCF-12A	non	tissu	basal	>40	non	++	wt	-	-	non
MDA-MB 157	oui	metasta	mésen	>40	oui	oui	mut	-9	-9	non
MDA-MB 231	oui	metasta	mésen	>40	oui	oui	mut	-	-	non
MDA-MB 361	oui	metasta	épithélial	<40	oui	non	mut	+	-	oui
MDA-MB 436	oui	metasta	mésen	>40	non	oui	mut	-	-	non
MDA-MB 453	oui	metasta	épithélial	>40	non	non	mut	-	-	oui
MDA-MB 459	-9	-9	mésen	-9	-9	oui	mut	-9	-9	-9
MDA-MB 549	oui	-9	mésen	-9	-9	oui	mut	-9	-9	-9
SK Br3	oui	metasta	épithélial	>40	oui	oui	mut	-	-	oui
T47D	oui	metasta	épithélial	>40	-9	non	mut	+	+	non
UACC 812	non	primaire	épithélial	>40	-9	oui	-9	+	-	oui
ZR 75.1	oui	metasta	épithélial	>40	oui	-9	wt	+	+	non

La majeure partie des informations ci-dessus proviennent des fiches techniques de l'ATCC (American Type Culture Collection) mais aussi d'informations internes au laboratoire INSERM U590.

## Annexe 5 : PCR en temps réel

### A5.1. Tableau des amorces

Nom	Séquence	Sonde	Kit / Référence
dicer Hum f744	ttaaccttttggtgttgatgagtgt	non	
dicer Hum r844	gcgaggacatgatggacaatt	non	
APC-11 F	inconnue	non	Assay on demand Hs00212858
APC-11 R	inconnue	non	Applied Biosystem
APC-5 F	inconnue	non	Assay on demand Hs00212120
APC-5 R	inconnue	non	Applied Biosystem
APC-7 F	inconnue	non	Assay on demand Hs00212139
APC-7 R	inconnue	non	Applied Biosystem
CDC-16 F	inconnue	non	Assay on demand Hs00187430
CDC-16 R	inconnue	non	Applied Biosystem
CDC-23 F	inconnue	non	Assay on demand Hs00190562
CDC-23 R	inconnue	non	Applied Biosystem
CDC-27 F	inconnue	non	Assay on demand Hs00265810
CDC-27 R	inconnue	non	Applied Biosystem
PPIB F	acttcaccaggggagatgg	#20	UPL - Roche Applied Science
PPIB R	agccgttggtgtcttgc	#20	UPL - Roche Applied Science
ACTB F	attggcaatgagcggctc	#11	UPL - Roche Applied Science
ACTB R	ggatgccacaggactccat	#11	UPL - Roche Applied Science
PGK-1 F	ctgtggcttctggcacaact	#42	UPL - Roche Applied Science
PGK-1 R	cttgctgcttcaggacca	#42	UPL - Roche Applied Science
dicer Mus #63F	agatggaggcggaggtcag	#63	UPL - Roche Applied Science
dicer Mus #63R	caatgagcaggttggtct	#63	UPL - Roche Applied Science
36B4 F (Mus)	gctgatgggcaagaacacca	non	
36B4 R (Mus)	cccggatgtgaggcagcagtt	non	

UPL = Universal Probe Library (<https://www.roche-applied-science.com/sis/rtpcr/upl/center.jsp?id=030000>).

### A5.2. Programmes d'amplification

Le programme d'amplification des gènes du complexe APC/C comprend 4 étapes. Etape 1 : dénaturation et activation de la polymérase à 95°C pendant 10 min. Etape 2 répétée 45 fois : dénaturation à 95°C pendant 5 sec suivie de l'appariement et l'extension à 60°C pendant 20 sec. Etape 3 : courbe de fusion consistant à l'augmentation de la température de 58°C à 95°C par paliers de 0,1°C. Etape 4 : refroidissement pendant 30 sec à 40°C. Les courbes de fusion ont été utilisées pour vérifier la spécificité de chaque amplification d'échantillon pour un gène donné grâce à

la présence d'un seul et unique pic et l'absence de dimères d'amorces.

Le programme The Roche Applied Science Universal Probe Library's thermocycling conditions were used for the reference genes: 10 min at 95°C then 10 sec at 95°C, 30 sec at 59°C and 1 sec at 72°C for 40 cycles and cooling.

## Annexe 6 : anticorps

Protéine	Marque	Référence	Dilution	Incubation	Ac Ilr
Dicer (Humain)	Abcam	ab14601	1/1000	nuit à 4°C	P0260
Dicer (Souris)	Abcam	ab13502	1/750	nuit à 4°C	P0217
SMC1	Abcam	ab16147	1/1000	1h amb	P0260
pSer966 SMC1	Abcam	ab1276	1/1500	nuit à 4°C	P0217
Chk1	Cell Signaling	#2360	1/1000	1h amb	P0260
pSer345 Chk1	Cell Signaling	#2348	1/1000	nuit à 4°C	P0217
Chk2	Sigma Aldrich	C 9233	1/1000	nuit à 4°C	P0260
pThr68 Chk2	Cell Signaling	#2661	1/1000	nuit à 4°C	P0217
H2A.X	Cell Signaling	#2595	1/1000	nuit à 4°C	P0217
pSer139 H2A.X	Upstate	#05-636	1/1000	nuit à 4°C	P0260
cycline B1	Upstate	#05-373	1/3000	1h amb	P0260
béta-actine	Sigma Aldrich	A5441	1/10000	30min amb	P0260

Anticorps secondaires (Ac Ilr) :

P0260 Dako anti-souris, 1/3000, 1h à température ambiante.

P0217 Dako anti-lapin, 1/3000, 1h à température ambiante.

## Annexe 7 : publications

### Contribution of cell culture, RNA extraction, and reverse transcription to the measurement error in qRT-PCR-based gene expression quantification

Jean-Damien COMBES<sup>1,2</sup>, Gaël GRELIER<sup>3</sup>, Matthieu LAVERSANNE<sup>1</sup>, Nicolas VOIRIN<sup>2</sup>, Sylvie CHABAUD<sup>2</sup>, Christine LASSET<sup>2,4\*</sup>, René ECOCHARD<sup>1</sup> and Caroline MOYRET-LALLE<sup>3</sup>

<sup>1</sup> Hospices Civils de Lyon, Service de Biostatistique, Lyon, F-69424, France ; Université de Lyon ; Université Lyon I, Villeurbanne, F-69622, France ; CNRS ; UMR 5558, Laboratoire Biostatistique Santé, Pierre-Bénite, F-69495, France

<sup>2</sup> Centre Léon Bérard, Département de Santé Publique, Lyon, F-69373, France

<sup>3</sup> Inserm U590, Lyon, F-69008, France ; Université de Lyon, Université Lyon I, Institut des sciences pharmaceutiques et biologiques (ISPB), Lyon, F-69373, France

<sup>4</sup> Université de Lyon ; Université Lyon I, Villeurbanne, F-69622, France ; CNRS ; UMR 5558, Equipe Epidémiologie et Santé Publique, Villeurbanne, F-69622, France

\* Corresponding author

Email addresses:

JDC: jean-damien.combes@chu-lyon.fr

GG: grelier@lyon.fnclcc.fr

ML: matthieu.laversanne@chu-lyon.fr

NV: nicolas.voirin@recherche.univ-lyon.fr

SC: chabauds@lyon.fnclcc.fr

CL: lasset@lyon.fnclcc.fr

RE: rene.ecochard@chu-lyon.fr

CML: moyret@lyon.fnclcc.fr

#### Abstract

*Background:* Quantitative PCR instruments are known to be reliable. However, there is still uncertainty about the reliability of the procedures that precede the measurement of gene expression; i.e., cell culture, RNA extraction, and reverse transcription. The objective is to assess the measurement errors due to those steps and to estimate the correction that would accrue from replicating either of those steps.

*Results:* Measuring the relative expression of APC-11 gene (the catalytic Anaphase Promoting complex / cyclosome subunit, suspected to be involved in breast cancer) with step replication in 18 breast cancer cell lines, the final quantitative PCR step was found reproducible (SD=0.26) while, in comparison with the between-cell-line variability (SD=1.7), the variability due the previous steps (cell culture, RNA extraction, reverse transcription) seemed to be of the same order of magnitude (SD=1.2 to 2.0). The

measurement error was also estimated using misclassification rates. A mixed model provided unbiased predictors for the gene expression value in each cell line. Those predictors were compared two-by-two to select the pairs of cell lines whose ratio of unbiased predicted values was equal or higher than 2. Within the latter selected lines, two-by-two comparisons were made between measurements to obtain misclassification rates (rate less than one). The misclassification rates improved with replication of cell culture, or of RNA extraction, or of reverse transcription (90%, 60.9%, and 61.1% decrease, respectively).

*Conclusion:* This study points out an unexpectedly high error levels due to reverse transcription, RNA extraction, and cell culture. The technical variability was as high as the biological variability. Overall, the reliability of the qRT-PCR-based gene expression quantification would be best improved by replicating the first step i.e. cell culture.

## **Background**

Quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR) is widely used in genetic and medical research to quantify gene expression. In its main application fields, such as oncology and microbiology, qRT-PCR remains the technique of reference. In addition, since their inception, quantitative PCR (qPCR) measuring instruments have undergone enough technical evolutions to be recognized now as accurate and reliable. However, reproducibility concerns have been observed and widely discussed [1-8] and numerous innovative methodological strategies have been proposed [2, 9-21].

Several studies have analyzed either the accuracy [22-27] or reliability [1, 23, 28-32] of qPCR measuring instruments. However, many manual procedures are required before instrument quantification of the amount of cDNA present in a sample and although many researchers have pointed out the variability introduced by those procedures [1, 3, 5, 24, 33-40], no studies have investigated their impact on the reliability of

---

qRT-PCR-based gene expression quantification.

This study is questioning the reliability of gene expression quantification when cell culture, RNA extraction, and RT procedures are required. Original experiments were carried out to estimate the part of variability due to each of those procedures involved in qRT-PCR-based gene expression quantification. Misclassification rates were used to account for measurement errors and the impact of replicating each manual procedure was assessed.

## **MATERIALS AND METHODS**

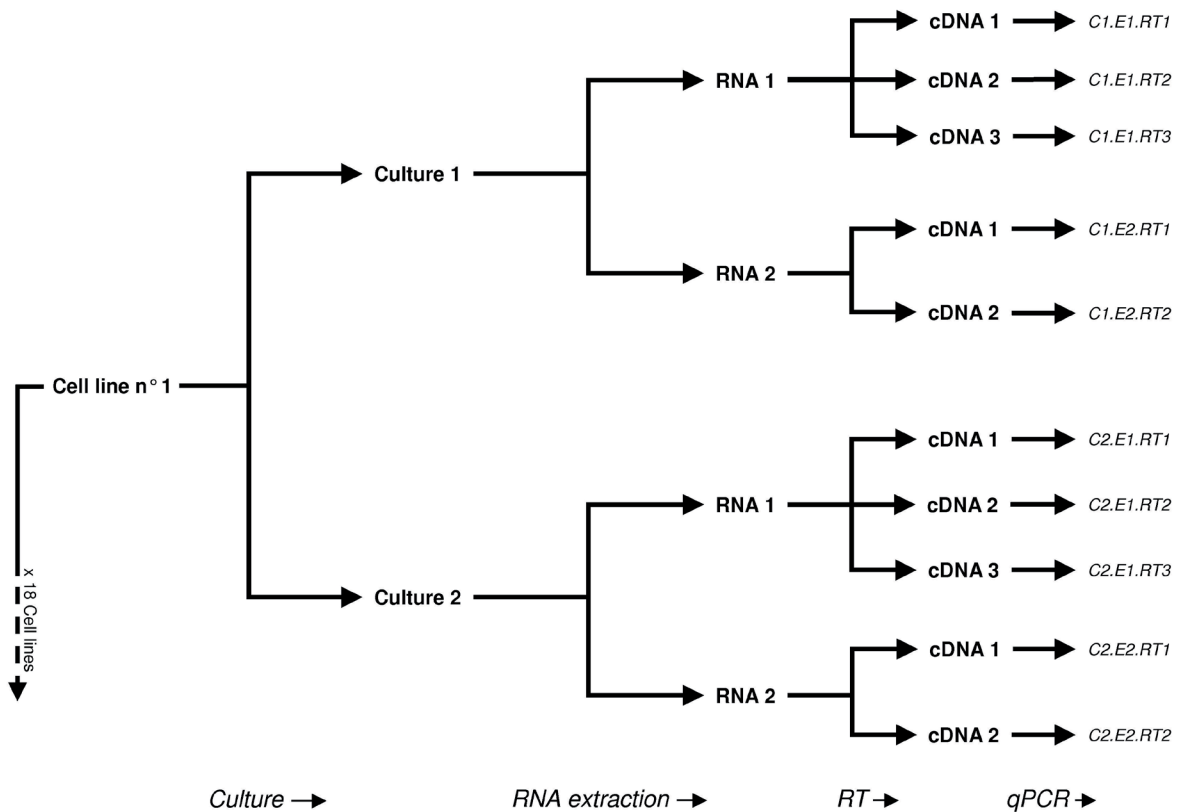
### *Data source and key experiments*

This study used data from two original experiments carried out in an INSERM laboratory (U590, Lyon, France) specialized in the study of genes involved in tumor progression.

Experiment 1 was designed to measure the variability of each procedure from cell culture to qPCR measurement. The relative expression of gene APC-11 (that codes for the catalytic subunit of the E3 ubiquitin ligase Anaphase Promoting Complex/Cyclosome) was measured in 18 breast cancer cell lines obtained from the American Type Culture Collection (ATCC): BT-20, BT-474, Cal-51, Cama-1, HCC-1937, HS-578T, MCF-7, MDA-MB 157, MDA-MB 231, MDA-MB 361, MDA-MB 436, MDA-MB 453, MDA-MB 459, MDA-MB 549, SK Br3, T47D, UACC 812 and ZR 75.1.

All the procedures were replicated: i) each cell line was cultured twice; ii) RNA extractions from those cultures provided two RNA extracts per culture; iii) each RT was also replicated (Figure 1). Because of sample scarcity, some RTs were carried out twice, some thrice. After these three procedures, each cDNA solution was transferred to the qPCR measuring instrument. Finally, 180 measurements of the APC-11 gene expression were obtained.

Experiment 2 was designed concomitantly to measure the variability due to the qPCR measuring instrument. Seven cell lines; Cal-51, Cama-1, MCF-12A, MDA-MB



**Figure 1 - Successive steps and replications relative to qRT-PCR measurement from each cell line (Experiment 1)**

453, MDA-MB 459, MN1, T47D, were selected for three cDNA measurements of each cDNA, leading to 21 measurements of the APC-11 gene expression.

### Experimental details

In Experiment 1, each cell line was cultured twice: once according to the ATTC standardized protocol and again according to the usual laboratory protocol. In the standardized protocol, cell proliferation was controlled using an optical microscope and stopped when proliferation reached approximately 70% confluence. In the usual laboratory protocol, there was no confluence control. The values stemming from each culture protocol were compared using a paired Student's t-test. As the differences were not statistically significant ( $p=0.65$ ), the two culture protocols were considered as replicates.

Total RNA extraction was performed using the phenol-chloroform method with TriReagent (Sigma-Aldrich, France) for cell lysis and PhaseLockGel tubes (Eppen-



dorf, France) for phase separation. After DNase treatment, DNA contamination of each sample was checked by electrophoresis on agarose gel.

cDNA synthesis was performed using the First-Strand cDNA Synthesis Kit (GE-Healthcare, France). Precisely, 1  $\mu\text{g}$  of denatured total RNA was mixed with dithiothreitol (DTT), random hexadeoxynucleotides, and a bulk reaction mix containing Murine Reverse Transcriptase. After 1 hour incubation at 37°C, the reaction was stopped by heating at 95°C for 3 minutes.

The APC-11-specific PCR assay used for RT-PCR amplification was obtained from Applied Biosystems (Foster City, USA). Samples were normalized by geometric averaging using Cyclophilin B (PPIB),  $\beta$ -actin (ACTB), and Phospho-glycerol kinase (PGK-1), selected with geNorm software [13] from 12 initial housekeeping genes. The primer pairs and probes were designed using the Universal Probe Library website (Roche Applied Science). Primers of each pair were located in different exons to avoid genomic amplification. All reactions were performed in glass capillaries (Roche Applied Science) using LightCycler FastStart DNA Master PLUS SYBR Green I kit for APC-11 or LightCycler Taqman Master for reference genes (Roche Applied Science). The reaction mixture contained FastStart TaqDNA polymerase, reaction buffer, deoxynucleoside triphosphate mixture, 180  $\mu\text{M}$  of  $\text{MgCl}_2$ , 0.25  $\mu\text{M}$  of each primer, and 1/60 diluted cDNA. Water was used as negative control. During preparation, all samples were kept at 4°C. Thermocycling conditions for the APC-11 gene include four consecutive steps. Step 1: denaturation and polymerase activation by 10 min heating at 95°C prior to amplification. Step 2: denaturation for 5 seconds at 95°C, followed by 20 sec of annealing/extension at 60°C for 45 cycles. Step 3: melting; this consisted of incrementing the temperature from 58°C to 95°C by steps of 0.1°C. Step 4: cooling for 30 sec at 40°C in a single cycle. The melting curve was used to check the specificity of each APC-11 run with a single peak and no primer dimer. Roche Applied Science Universal Probe Library's thermocycling conditions were used for the reference genes: 10 min at 95°C then 10 sec at 95°C, 30 sec at 59°C, and 1 sec at 72°C for 40 cycles and cooling. Each run

was prepared with the calibrator, 14 different samples, and a no-template control, all in duplicates. For any given sample, the five cDNA issued from each RT were used in independent runs. All fluorescence data were analyzed by the LightCycler 4.0 software (Roche) and the results were exported to Microsoft Excel.

A large amount of cDNA was prepared from the HME-1 cell line prior to the experiments. It was then diluted, aliquoted, and used as a calibrator for all RT-PCR runs. For relative quantification and normalization, the following calculation was used:

$$N = E_T^{CpT(C) - CpT(S)} \times \prod \left( E_R^{CpR(S) - CpR(C)} \right)$$

where N is the normalized ratio, E the efficiency, Cp the crossing point, T the target gene (APC-11 with E=1.763), R the reference genes (PPIB with E=2.000, ACTB with E=1.920, and PGK-1 with E=1.994), C the calibrator (HME-1) and S the unknown sample. PCR efficiencies were calculated using relative standard curves derived from serial dilutions of the calibrator HME-1, from 1/6 to 1/60,000, with three measurements.

#### *Assessment of variability at each procedure of gene expression quantification*

Two types of variability were considered: i) Biological variability that stems from differences in gene expression between cell lines (or patients) and assumed governed by a “genotype factor”; and ii) Technical variability that stems from the procedure succession and assumed governed by the following factors: a “culture factor”, an “extraction factor” (inter-RNA variability), a “RT+qPCR instrument factor” (inter-cDNA variability), and a “qPCR instrument factor” (inter-measurement variability).

Regarding the genotype factor, ten measurements were performed on each cell line. The mean value of those measurements represents an estimation of the APC-11 gene expression in a given cell line. The range of variation of these means among the 18 cell lines (18 means) reflects the extent of the variability due to the genotype factor.

Regarding the culture factor, two cell cultures were performed with each cell

---

line. Five measurements were made in each culture. The mean value of those five measurements represents an estimation of the APC-11 gene expression in that culture. The difference between the two estimations (the two means of two sets of five measurements) informs us about the variability linked to the culture factor. The range of differences over the 18 cell lines (18 differences) reflects the extent of variability due to the “phenotype factor”.

Regarding the extraction factor (inter-RNA variability), two RNA extractions were performed from each cell culture. The mean value of the measurements corresponding to each extraction (sometimes two, sometimes three measurements) reflects RNA concentration in the extract. The difference between the two means informs us about the variability linked to the extraction factor. The range of this difference over the 18 cell lines (36 differences) reflects the extent of variability due to the “extraction factor”.

Regarding the “RT+qPCR instrument factor” (inter-cDNA variability), one RNA extraction gave sometimes two, sometimes three cDNA solutions. Thus, the “RT+qPCR instrument factor” was sometimes represented by the difference between two values and sometimes by the three possible differences between the three values. Finally, the range of the 144 differences reflects the extent of the variability due to the “RT+qPCR instrument factor”.

The qPCR instrument factor was represented by the three possible differences between the three values obtained for each cDNA. Thus, the range of these 21 differences reflects the extent of the variability due to the “qPCR instrument factor”.

The standard deviation (SD) is the easiest indicator to compare the effects of the above-cited factors, especially that it does not depend on the sample size. The standard deviation at each step provides a positioning of the importance of the biological variability with regard to the importance of the four factors of the technical variability. In contrast, comparing the ranges leads to difficult interpretations because the ranges depend on the sample size.

---

*Gene expression values corrected with a mixed effect model*

A linear mixed-effect regression model was performed to compute the corrected APC-11 gene expression value for each cell line.

To assess the mean-variance relationship, means, variances, and standard deviations were computed over the 10 measures performed on each of the 18 cell lines. As mixed-effect models are based on the assumption of a constant variance whatever the measurement values, that assumption was checked with two different plots [41, 42]. The first one analyzed the relationship between the mean and the standard deviation. The second analyzed the relationship between the mean and the variance through a log-scale plot, allowing a more in-depth analysis. The linearity of those relationships was checked graphically and the goodness-of-fit was estimated by R-squared values.

A proportional relationship between the mean and the standard deviation motivates a log-transformation of the measurements. In the mean-variance log-scaled plot, a proportional relationship between the mean and the variance would agree with a mean-variance power relationship. If not, a quadratic relationship in the same plot would indicate a negative binomial distribution of the measurement error. According to these results, the appropriate parameter of mean-variance relationship was introduced into the regression.

The observed mean value of the ten measurements is a raw estimation of the APC-11 gene expression value in a cell line. Thus, unbiased predictors were estimated for each of the 18 cell lines. According to the Best Linear Unbiased Predictor method, the mixed effect model allows considering both inter-cell line and intra-cell line variances when estimating the gene expression value from replicated measurements [43]. Briefly, the unbiased predictor is weighted by the reliability assigned to the observed value. When the observed ten values for a given cell line are distributed with a high variance, the unbiased predictor value is close to the overall mean of all 180 measurements because of the weak reliability of the ten values. Conversely, when the intra-cell

line variance is low, the unbiased predicted value is close to the mean of the ten values for that cell line. Thus, estimating the “real” gene expression value in each cell line, the unbiased predictor may be considered as a better estimation than the mean. Likewise, the mixed effect model allows providing a better estimation of the mean-variance relationship. The parameters of the mean-variance relationship equation were estimated with the same regression model.

#### *Misclassification rate in differential gene expression quantification*

Another expression of the measurement error was the misclassification rate. In a first step, the unbiased predicted values for the 18 cell lines were compared two-by-two in order to select the pairs of cell lines whose differential predicted value (the ratio of the unbiased predicted values) was equal or higher than 2.

In a second step, within each retained pair, the two-by-two comparisons were made between the two sets of 10 observed measurements from Experiment 1; that is, between each of the 10 measurements obtained from one cell line and each of the 10 obtained from the other cell line. There were thus 100 possible differential gene expression values. In these comparisons, a differential gene expression value was considered misclassified when ratio of the measurements was less than 1 and the rate of misclassification computed.

The same process was used to compute the misclassification rates considering gene expression as the mean value of two measurements stemming from the replicata due to replicating cell culture, RNA extraction, or RT. With replication of each of the three procedures and each pair of cell lines, there were respectively 625, 144, and 64 possible comparisons or differential gene expression values.

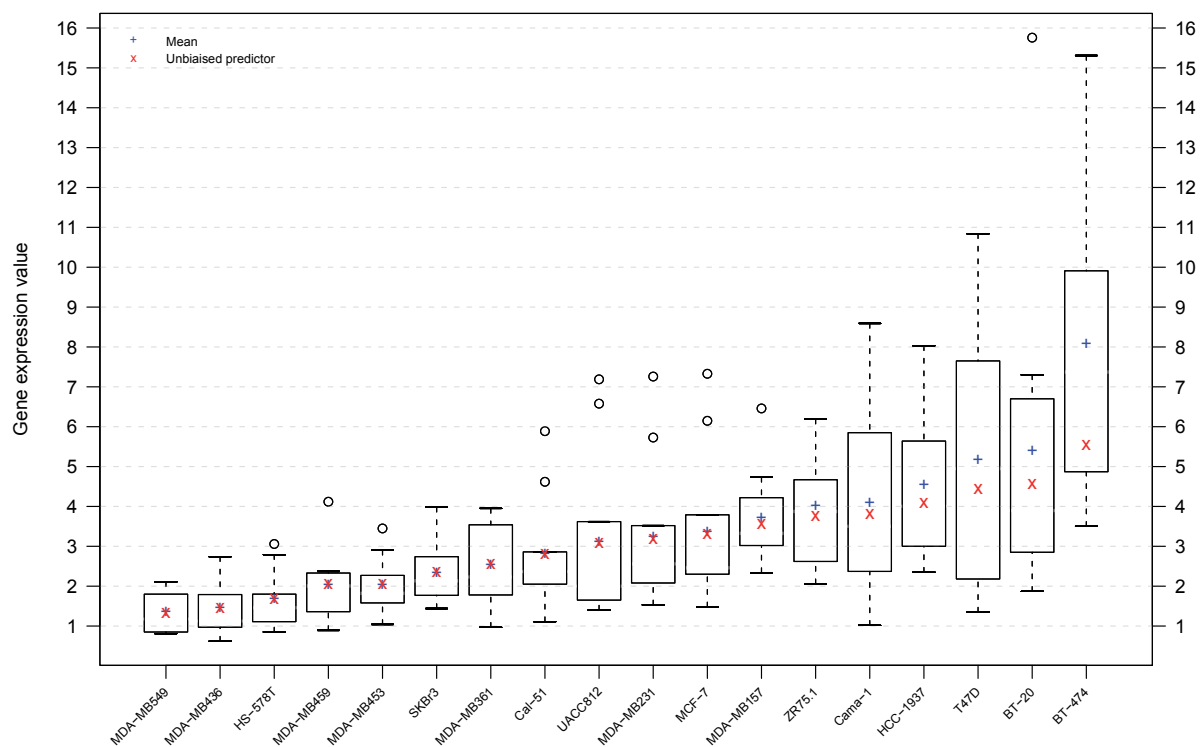
The statistical analyses were performed with R statistical software version 2.6.0. [44], and “nlme” package was used for mixed effect modeling [45].

## **Results**

## Data description

In Experiment 1, 180 measurements of the APC-11 gene expression were made on the 18 cell lines, with values ranging from 0.62 to 15.76. A first look at the plot (Figure 2) shows an increase of the variance with the increase of the measurement values.

In Experiment 2, 21 measurements of the APC-11 gene expression were made on 7 cell lines, with values ranging from 1.2 to 4.9.



**Figure 2 - Distribution of the 10 measurements for each cell line involved in Experiment 1**

The observed values range from 0.62 to 15.76, although the estimated “real” gene expression values range from 1.31 to 8.04 according to the mean values and from 1.34 to 5.55 according to the unbiased predictor. Besides, the higher is the gene expression level (mean or unbiased predictor), the higher is the measurement error (variance).

## Assessment of variability at each process of gene expression quantification

The measures of variability found at each procedure are shown in Table 1. The “RT+qPCR instrument factor” was the most important source of variability (SD=2.0), followed by the “genotype factor” that represents the biological variability (SD=1.7). The “culture factor” and the “extraction factor” were less important sources of variability. However, compared to the biological variability, that of the three other above-cited factors of technical variability were roughly of the same order of magnitude (i.e., com-

prised within one SD unit: 1.2 to 2.0).

The difference between the variability related to the “qPCR instrument factor” (SD=0.26) and the variability related to the “RT + qPCR instrument factor” (SD=2.0) reflects the “RT factor” and shows the negligible contribution of the “qPCR instrument factor” vs. the “RT + qPCR instrument factor”.

**Table 1**

Source of variability	Number of values	[min - max]	Range	SD
<b>Biological</b>				
Genotype	18 means	1.31 - 8.04	6.7	1.7
<b>Technical</b>				
Culture	18 differences	0.01 - 5.06	5.1	1.3
Extraction	36 differences	0.02 - 5.48	5.5	1.2
RT+qPCR instrument	144 differences	0.01 - 11.85	11.8	2.0
qPCR instrument*	21 differences	0 - 1.02	1.02	0.26

\* Experiment 2

**Table 1 - Assessment of the variability in the procedures of gene expression quantification**

#### *Gene expression values corrected the with mixed effect model*

Plotting the standard deviation versus the mean as well as the log-scaled variance versus the mean showed linear relationships (not shown). However, the R-squared values were different, equal to 0.78 and 0.85, respectively. Backed by the graphical analysis, these results tend to detect a better linear relationship between the variance and the mean. This result supports the hypothesis of a power relationship between the mean and the variance. Thus, a mean-variance power parameter (varPower) was included into the mixed effect model.

The unbiased predicted values given by the mixed model are presented in Figure 2. According to those values, the variability due to the genotype effect (i.e. biological variability) was estimated as SD=1.2.

The diagnostic plots indicated no major deficiencies of the model: the random effects were almost normally distributed and the variance was stable whatever the mean.

As secondary result, the mixed-effect model estimated the variance-mean relationship as:

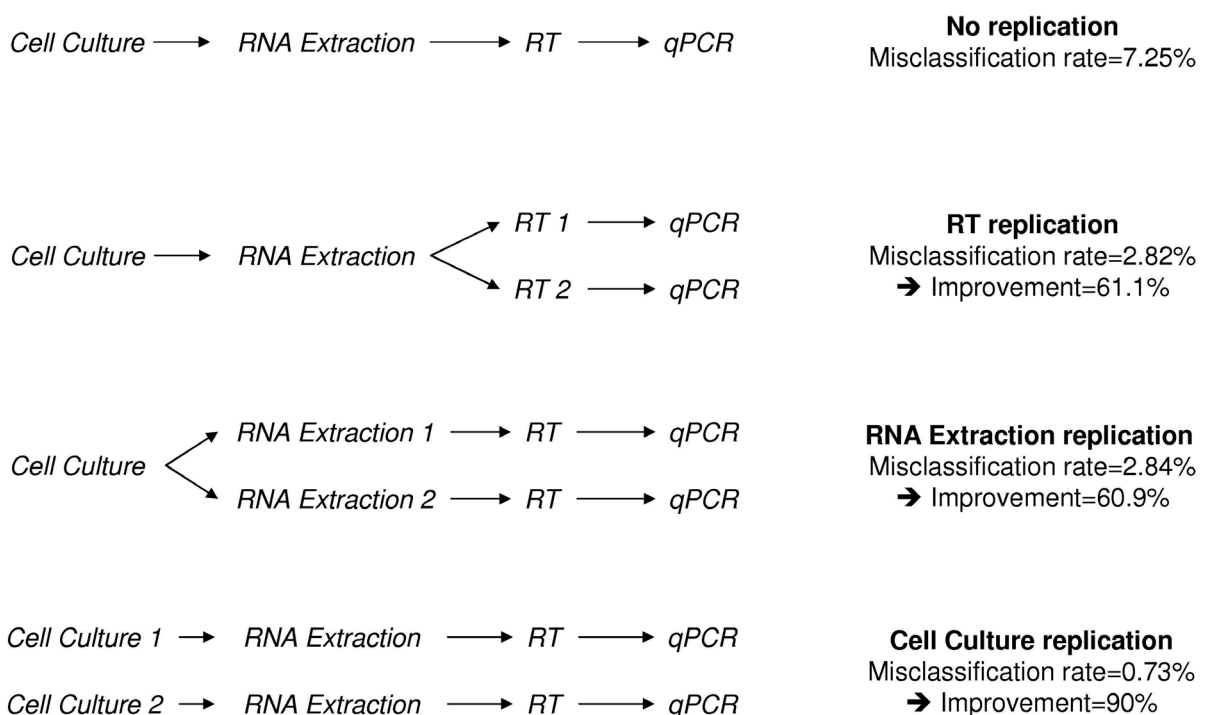
$$\text{var} = 0.10 [0.05 - 0.17] \times \text{mean}^{3[2.46-3.53]}$$

corresponding to the relationship between the measurement error range (variance) and the gene expression level (mean).

#### *Misclassification rate in differential gene expression quantification*

Quantifying differential gene expression with only one measurement resulted in 7.25% of misclassified. When gene expression quantification was the mean value of two measurements obtained by replication of cell culture, the misclassification rate decreased to 0.73%, which corresponded to a 90% improvement. Replicating RNA extraction and RT, the misclassification rates were respectively 2.84 and 2.82%, which corresponded to 60.9 and 61.1% improvements (Figure 3).

Detailed misclassification rates for each selected pairs of cell-lines are presented in Table 2.



**Figure 3 - Improvement of misclassification rates with replication of each cDNA preparation step**



**Table 2**

Selected pairs of cell lines	Differential gene expression*	Misclassification rates (%)			
		No replication	Replication (two measurements)		
			RT	Extraction	Cell culture
MDA-MD157 / HS-578T	2.1	5	0	0	0.2
UACC-812 / MDA-MD436	2.1	18	12.5	16.7	3.8
Cal-51 / MDA-MD549	2.1	8	0	2.1	1.9
MDA-MD231 / MDA-MD549	2.1	7	3.1	1.4	0.6
ZR 75.1 / HS-578T	2.2	6	1.6	0	0.6
BT-474 / MDA-MD361	2.2	3	0	0	0
MDA-MD231 / MDA-MD436	2.2	15	4.7	6.9	2.7
BT-20 / MDA-MD453	2.2	9	0	1.4	2.9
T47D / MDA-MD453	2.2	23	14.1	18.1	0
BT-20 / MDA-MD459	2.2	10	3.1	0	2.9
T47D / MDA-MD459	2.2	21	12.5	16	1
Cama-1 / HS-578T	2.3	15	9.4	1.4	3.8
MCF-7 / MDA-MD436	2.3	16	4.7	5.6	1.3
UACC-812 / MDA-MD549	2.3	15	7.8	4.2	1.3
HCC-1937 / HS-578T	2.4	4	0	0.7	0
MDA-MD157 / MDA-MD436	2.4	4	0	0	0
BT-474 / SK Br3	2.4	1	0	0	0
MCF-7 / MDA-MD549	2.5	7	4.7	2.1	0.2
T47D / HS-578T	2.6	16	6.3	11.8	0
Cama-1 / MDA-MD436	2.6	12	6.3	0.7	1.8
ZR 75.1 / MDA-MD436	2.6	5	0	0	0
BT-20 / HS-578T	2.7	5	0	1.4	0.8
BT-474 / MDA-MD453	2.7	0	0	0	0
BT-474 / MDA-MD459	2.7	1	0	0	0
MDA-MD157 / MDA-MD549	2.7	0	0	0	0
HCC-1937 / MDA-MD436	2.8	2	0	0	0
ZR 75.1 / MDA-MD549	2.8	1	0	0	0
Cama-1 / MDA-MD549	2.9	9	7.8	0	0.6
T47D / MDA-MD436	3	11	3.1	8.3	0
BT-20 / MDA-MD436	3.1	3	0	0	0
HCC-1937 / MDA-MD549	3.1	0	0	0	0
BT-474 / HS-578T	3.3	0	0	0	0
T47D / MDA-MD549	3.3	8	0	3.5	0
BT-20 / MDA-MD549	3.4	1	0	0	0
BT-474 / MDA-MD436	3.8	0	0	0	0
BT-474 / MDA-MD549	4.1	0	0	0	0
<b>Mean</b>	2.63	<b>7.25</b>	<b>2.82</b>	<b>2.84</b>	<b>0.73</b>

\* Differential gene expression: ratio of gene expression values estimated by best linear unbiased predictor

**Table 2 - Detailed misclassification rates with and without replication for each selected pairs of cell lines**

---

## Discussion

This study points out to an unexpectedly high error level in the current quantification of gene expression. Errors were due to cell culture, RNA extraction, and RT steps. Fortunately, a significant part of this measurement error might be managed by replicating cell culture.

Compared to the biological variability, the technical variability appeared of the same order of magnitude. In other words, the measurement error was almost as high as the measured value. However, the qPCR measuring instrument was found reliable. Thus, comparing the variability due to the qPCR instrument and the other three technical factors might lead to consider useless replicating measurements of the same cDNA sample.

Expressing the measurement error by the misclassification rate was in accordance with the error level determined above. For the selected 36 pairs with estimated differential expression higher or equal to two, 7.25% of misclassified measurements were found. This selection was based on experts' opinion that considered biologically relevant any differential gene expression value greater or equal to two. Although a rate of 7.25% might not be impressive, this misclassification rate reflects only major errors that are perfectly opposite results. However, this misclassification rate was used to check improvements due to procedure replication.

There is no doubt that the larger the number of replications is, the better the improvement in misclassification rate would be. However, replicating a procedure is a time-consuming and costly procedure. Thus, we estimated exclusively the improvement due to the duplication of one of the cDNA preparation procedures, which is conceivable in routine practice.

The best improvement was found with replication of cell culture (90%). Even if the major source of variability was attributable to RT, the variabilities relative to cell culture and RNA extraction were found of the same order of magnitude. Thus, replicating the first chronological step; that is, cell culture, allows taking into account the two other

---

variabilities. If the measurement error due to cell culture and RNA extraction were negligible compared to that due to RT, a simple replication of RT would considerably improve the misclassification rate. Obviously, but unfortunately, cell culture replication means finally the need for two cell cultures, two extractions, and two transcriptions (one extraction and one transcription for each culture).

As secondary result, the relationship between the mean and the variance was found dependent of the cell line. In other words, the higher was the gene expression, the higher was the measurement error. Besides, the mean-variance power relationship implies that the usual log-transformation of the measurement value is not an optimal way to compensate for an unstable variance. The mixed-effect model estimation of the measurement error supports the hypothesis of an under-dispersed gamma distribution of that error. This descriptive result might help understanding the origins of the measurement error, even if there is currently no theoretical explanation for that form of distribution.

All the data stemmed from an original experimental plan carried out by an experienced team in a laboratory where qRT-PCR-based gene expression quantification is routine and we cannot exclude that those results are specific to the processes of that laboratory. However, to our present knowledge, there have been no similar studies in the literature to which we may compare our results. Nevertheless, several authors have underlined the importance of the cDNA preparation steps. Our results agree with this view, as we have shown that cell culture, RNA extraction and RT are all sources of variability.

The reliability of gene expression measurement will still be debated as long as those procedures remain poorly reproducible. Detecting or quantifying micro-organism DNAs with PCR instruments do not seem as arguable. Indeed, with DNAs, the preliminary procedures (sample processing) are obviously less source of errors even if improvements are still needed [27, 46, 47]. On the contrary, confirmation of micro-array results is more questionable. In fact, confirming micro-array results is not a problem

when the same cDNA solution is used for micro-array and for qRT-PCR; here, the only source of error would be the measuring instrument. When another cell culture, extraction, or RT is needed, the result becomes more disputable. In the same way, a gene expression-based cancer prognostic or diagnostic test might be questionable. Even if there is no cell culture needed, the variability due to biopsy sampling [48] or even to a blood sample processing [1, 49] should be assessed before quantifying gene expression. Fortunately, recent studies have shown some gene expression-based classifications that seem clinically applicable [50-52]. These positive results might be explained by the use of a great number of genes. If the correlation between a specific gene and a given diagnosis or prognosis seems questionable due to the extent of the measurement error, screening for a panel of genes might reduce the impact of these errors. In 2005, a study designed to assess the reliability of gene expression measurement using laser-microdissection and oligonucleotide arrays showed that technical variability was lower than biological variability [53]. Our study has shown that technical variability was as high as biological variability. Indeed, the improvement of automated high-technology instruments, such as qPCR, might have gone too far compared to the improvements of manual procedures; i.e., cDNA preparation.

## Conclusions

If the results of this study are confirmed, we will be able to significantly improve the reliability of qRT-PCR-based gene expression quantification by replicating the cell culture procedure. Anyhow, we should do our best to improve the reproducibility of RT, but also of RNA extraction and of cell culture conditions. Thus, we might be able to take full advantage of the qPCR high accuracy and improve the clinical applications of transcriptomic research.

## Authors' contributions

GG, CML and NV designed the experiment. GG performed the gene expression measurements and drafted a section of the manuscript. JDC, ML and RE implemented the statistical methods and analyzed the data. JDC drafted the manuscript. CL, SC, RE and CML supervised the work. All authors read and

approved the final manuscript.

### Acknowledgments

We thank Jean Iwaz, PhD and Scientific Advisor, for revising and editing the manuscript as well as Fabien Subtil, PhD student, for statistical and computing advice.

### References

1. Keilholz U, Willhauck M, Rimoldi D, Brasseur F, Dummer W, Rass K, de Vries T, Blaheta J, Voit C, Lethe B et al: Reliability of reverse transcription-polymerase chain reaction (RT-PCR)-based assays for the detection of circulating tumour cells: a quality-assurance initiative of the EORTC Melanoma Cooperative Group. *Eur J Cancer* 1998, 34(5):750-753.
2. Vu HL, Troubetzkoy S, Nguyen HH, Russell MW, Mestecky J: A method for quantification of absolute amounts of nucleic acids by (RT)-PCR and a new mathematical model for data analysis. *Nucleic Acids Res* 2000, 28(7):E18.
3. Bustin SA: Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* 2002, 29(1):23-39.
4. Ginzinger DG: Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Hematol* 2002, 30(6):503-512.
5. Skern R, Frost P, Nilsen F: Relative transcript quantification by quantitative PCR: roughly right or precisely wrong? *BMC Mol Biol* 2005, 6(1):10.
6. Wong ML, Medrano JF: Real-time PCR for mRNA quantitation. *Biotechniques* 2005, 39(1):75-85.
7. Levesque-Sergerie JP, Duquette M, Thibault C, Delbecchi L, Bissonnette N: Detection limits of several commercial reverse transcriptase enzymes: impact on the low- and high-abundance transcript levels assessed by quantitative RT-PCR. *BMC Mol Biol* 2007, 8:93.
8. Vlems FA, Ladanyi A, Gertler R, Rosenberg R, Diepstra JH, Roder C, Nekarda H, Molnar B, Tulassay Z, van Muijen GN et al: Reliability of quantitative reverse-transcriptase-PCR-based detection of tumour cells in the blood between different laboratories using a standardised protocol. *Eur J Cancer* 2003, 39(3):388-396.
9. Gibson UE, Heid CA, Williams PM: A novel method for real time quantitative RT-PCR. *Genome Res* 1996, 6(10):995-1001.
10. Morrison TB, Weis JJ, Wittwer CT: Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques* 1998, 24(6):954-958, 960, 962.
11. Pfaffl MW: A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001, 29(9):e45.
12. Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001, 25(4):402-408.
13. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002, 3(7):RESEARCH0034.
14. Tichopad A, Dilger M, Schwarz G, Pfaffl MW: Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic Acids Res* 2003, 31(20):e122.
15. Andersen CL, Jensen JL, Orntoft TF: Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 2004, 64(15):5245-5250.
16. Szabo A, Perou CM, Karaca M, Perreard L, Quackenbush JF, Bernard PS: Statistical modeling for selecting housekeeper genes. *Genome Biol* 2004, 5(8):R59.
17. Larionov A, Krause A, Miller W: A standard curve based method for relative real time PCR data processing. *BMC Bioinformatics* 2005, 6:62.
18. Yuan JS, Reed A, Chen F, Stewart CN, Jr.: Statistical analysis of real-time PCR data. *BMC Bioinformatics* 2006, 7:85.
19. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J: qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 2007, 8(2):R19.
20. Spiess AN, Feig C, Ritz C: Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. *BMC Bioinformatics* 2008, 9:221.
21. Yang L, Liang W, Jiang L, Li W, Cao W, Wilson ZA, Zhang D: A novel universal real-time PCR system using the attached universal duplex probes for quantitative analysis of nucleic acids. *BMC Mol Biol* 2008, 9:54.
22. Bustin SA: Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 2000, 25(2):169-193.
23. Gentle A, Anastasopoulos F, McBrien NA: High-resolution semi-quantitative real-time PCR without the use of a standard curve. *Biotechniques* 2001, 31(3):502, 504-506, 508.

24. Klein D, Leutenegger CM, Bahula C, Gold P, Hofmann-Lehmann R, Salmons B, Lutz H, Gunzburg WH: Influence of preassay and sequence variations on viral load determination by a multiplex real-time reverse transcriptase-polymerase chain reaction for feline immunodeficiency virus. *J Acquir Immune Defic Syndr* 2001, 26(1):8-20.
25. Schmitt Y: Performance characteristics of quantification assays for human immunodeficiency virus type 1 RNA. *J Clin Virol* 2001, 20(1-2):31-33.
26. Palmer S, Wiegand AP, Maldarelli F, Bazmi H, Mican JM, Polis M, Dewar RL, Planta A, Liu S, Metcalf JA et al: New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol* 2003, 41(10):4531-4536.
27. Dionisi HM, Harms G, Layton AC, Gregory IR, Parker J, Hawkins SA, Robinson KG, Sayler GS: Power analysis for real-time PCR quantification of genes in activated sludge and analysis of the variability introduced by DNA extraction. *Appl Environ Microbiol* 2003, 69(11):6597-6604.
28. Yue H, Eastman PS, Wang BB, Minor J, Doctolero MH, Nuttall RL, Stack R, Becker JW, Montgomery JR, Vainer M et al: An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. *Nucleic Acids Res* 2001, 29(8):E41-41.
29. Ikeda T, Murakami M, Funaba M: Reliability of RT-PCR methods for measuring relative gene expression in mast cells. *Vet Immunol Immunopathol* 2004, 100(1-2):1-5.
30. Pagliarulo V, George B, Beil SJ, Groshen S, Laird PW, Cai J, Willey J, Cote RJ, Datar RH: Sensitivity and reproducibility of standardized-competitive RT-PCR for transcript quantification and its comparison with real time RT-PCR. *Mol Cancer* 2004, 3:5.
31. Raggi CC, Verderio P, Pazzagli M, Marubini E, Simi L, Pinzani P, Paradiso A, Orlando C: An Italian program of external quality control for quantitative assays based on real-time PCR with Taq-Man probes. *Clin Chem Lab Med* 2005, 43(5):542-548.
32. Arikawa E, Sun Y, Wang J, Zhou Q, Ning B, Dial SL, Guo L, Yang J: Cross-platform comparison of SYBR Green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray Quality Control (MAQC) study. *BMC Genomics* 2008, 9(1):328.
33. Freeman WM, Walker SJ, Vrana KE: Quantitative RT-PCR: pitfalls and potential. *Biotechniques* 1999, 26(1):112-122, 124-115.
34. Gabert J, Beillard E, van der Velden VH, Bi W, Grimwade D, Pallisgaard N, Barbany G, Cazzaniga G, Cayuela JM, Cave H et al: Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. *Leukemia* 2003, 17(12):2318-2357.
35. Stahlberg A, Hakansson J, Xian X, Semb H, Kubista M: Properties of the reverse transcription reaction in mRNA quantification. *Clin Chem* 2004, 50(3):509-515.
36. Suslov O, Steindler DA: PCR inhibition by reverse transcriptase leads to an overestimation of amplification efficiency. *Nucleic Acids Res* 2005, 33(20):e181.
37. Bustin SA, Benes V, Nolan T, Pfaffl MW: Quantitative real-time RT-PCR--a perspective. *J Mol Endocrinol* 2005, 34(3):597-601.
38. Fleige S, Pfaffl MW: RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med* 2006, 27(2-3):126-139.
39. Max N, Willhauck M, Wolf K, Thilo F, Reinhold U, Pawlita M, Thiel E, Keilholz U: Reliability of PCR-based detection of occult tumour cells: lessons from real-time RT-PCR. *Melanoma Res* 2001, 11(4):371-378.
40. Calogero A, Hospers GA, Timmer-Bosscha H, Koops HS, Mulder NH: Effect of specific or random c-DNA priming on sensitivity of tyrosinase nested RT-PCR: potential clinical relevance. *Anticancer Res* 2000, 20(5B):3545-3548.
41. Dagnelie P: *Théorie et méthodes statistiques: applications agronomiques*, vol. II. Gembloux; 1975.
42. Li J: *Statistical inference* (2 vol.). Ann Arbor; 1964.
43. Robinson G: That BLUP is a good thing: the estimation of random effects. *Statist Sci* 1991(6):15-32.
44. RDC T: R: A language and environment for statistical computing. In., 2.6.0 edn. Vienna, Austria: R Foundation for Statistical Computing;; 2007.
45. Pinheiro J, Bates D: *Mixed-Effects Models in S and Splus*. New York; 2000.
46. Barbi M, MacKay WG, Binda S, van Loon AM: External quality assessment of cytomegalovirus DNA detection on dried blood spots. *BMC Microbiol* 2008, 8:2.
47. Cankar K, Stebih D, Dreo T, Zel J, Gruden K: Critical points of DNA quantification by real-time PCR--effects of DNA extraction method and sample matrix on quantification of genetically modified organisms. *BMC Biotechnol* 2006, 6:37.
48. Jochumsen KM, Tan Q, Holund B, Kruse TA, Mogensen O: Gene expression in epithelial ovarian cancer: a study of intratumor heterogeneity. *Int J Gynecol Cancer* 2007, 17(5):979-985.
49. Ladanyi A, Soong R, Tabiti K, Molnar B, Tulassay Z: Quantitative reverse transcription-PCR compa-

ribose of tumor cell enrichment methods. *Clin Chem* 2001, 47(10):1860-1863.

50. Li Y, Elashoff D, Oh M, Sinha U, St John MA, Zhou X, Abemayor E, Wong DT: Serum circulating human mRNA profiling and its utility for oral cancer detection. *J Clin Oncol* 2006, 24(11):1754-1760.

51. Morris SR, Carey LA: Molecular profiling in breast cancer. *Rev Endocr Metab Disord* 2007, 8(3):185-198.

52. Pusztai L: Current status of prognostic profiling in breast cancer. *Oncologist* 2008, 13(4):350-360.

53. King C, Guo N, Frampton GM, Gerry NP, Lenburg ME, Rosenberg CL: Reliability and reproducibility of gene expression measurements using amplified RNA from laser-microdissected primary breast tissue with oligonucleotide arrays. *J Mol Diagn* 2005, 7(1):57-64.

---

**Estimating biological variation of quantitative gene expression in real-time reverse-transcription polymerase chain reaction (RT-PCR)**

Nicolas VOIRIN<sup>1</sup>, Gaël GRELIER<sup>2</sup>, Caroline MOYRET-LALLE<sup>2,3</sup>, Valérie BONADONA<sup>1</sup>, Sylvie CHABAUD<sup>1</sup>, Alain PUISIEUX<sup>2,3</sup>, Benjamin-Pierre BOUCHET<sup>2</sup> and Christine LASSET<sup>1,4</sup>

<sup>1</sup>Centre Léon Bérard, Département de Santé Publique, Lyon, France

<sup>2</sup>Centre Léon Bérard, INSERM U590, Lyon, France

<sup>3</sup>Université Claude Bernard Lyon 1, Faculté de Pharmacie, Lyon, France.

<sup>4</sup>Université Claude Bernard Lyon 1, Faculté Grange Blanche, Lyon, France.

**Corresponding author**

Dr Christine Lasset – Leon Berard Cancer Centre and Claude Bernard Lyon 1 University –  
28, rue Laennec – 69008 Lyon – Tel: +33 4 78 78 27 52 – Fax: +33 4 78 78 27 15 –  
[lasset@lyon.fnclcc.fr](mailto:lasset@lyon.fnclcc.fr)

**Address for correspondence**

Dr Christine Lasset – Leon Berard Cancer Centre and Claude Bernard Lyon 1 University –  
28, rue Laennec – 69008 Lyon – Tel: +33 4 78 78 27 52 – Fax: +33 4 78 78 27 15 –  
[lasset@lyon.fnclcc.fr](mailto:lasset@lyon.fnclcc.fr)



---

**Abstract**

Real-time reverse transcription polymerase chain reaction (RT-PCR) is widely used in genetic and medical research for measuring the expression of genes. The technical variability occurring during each step of the RT-PCR has been well studied in order to improve the interpretation of RT-PCR results. However, the biological variability that may occur when the biological material is issued from a number of different samples received less interest. A random effects model was applied to RT-PCR data with the aim to estimate the biological variation included in the relative RNA expression when multiple biological samples are used. Because biological variation was found to be significant, this approach is proposed as a statistical framework when the biological relevance of mRNA quantification is of interest.

---

## Introduction

Real-time reverse transcription polymerase chain reaction (RT-PCR) is widely used in genetic and medical research for measuring the expression of genes. The quantification of mRNA using real-time RT-PCR is based on four technical steps (1,2) and problems related to reproducibility and variability have been observed (3). However innovative methodological strategies have been proposed to improve both absolute (3) and relative (4) quantifications as well as for selecting the appropriate stably expressed genes for normalization (5-7). Some software which include these new methods have been recently proposed to process the data generated by real time RT-PCR (8,9).

The previous methods allow taking into account the variability occurring during each step of the RT-PCR (i.e. technical variability) however they did not allow the analysis of biological replicates (i.e. biological variability) that may occur when the biological material is issued from a number of different samples. When multiple measures are made within a same sample, a specific statistical processing may be needed to assess the biological relevance of the results following the generation of RT-PCR data if however biological variability is detected.

In this study, a random effects model was applied to RT-PCR data with the aim to estimate simultaneously the relative RNA expression and the biological variation when multiple biological samples are used. This approach is proposed as a statistical framework to assess the biological relevance in studies using mRNA quantification.

## Material and Methods

Relative RNA expression was determined using APC-11, a gene belonging to the Anaphase Promoting Complex/Cyclosome (APC/C). Human APC/C is implicated in the regulation of the cell cycle and is composed of at least 12 different proteins that are

---

constitutively present in the complex. The APC-11 unit contains a RING finger domain involved in the ubiquitin ligase activity of the APC/C complex (10).

*Cell culture, total RNA extraction, cDNA synthesis and RNA quantification method*

Eighteen breast cancer cell lines (BT-20, BT-474, Cal-51, Cama-1, HCC-1937, HS-578T, MCF-7, MDA-MB 157, MDA-MB 231, MDA-MB 361, MDA-MB 436, MDA-MB 453, MDA-MB 459, MDA-MB 549, SK Br3, T47D, UACC 812 and ZR 75.1) and the immortalized breast cell line HME-1 were obtained from the American Type Culture Collection (ATCC) and cultured according to the manufacturer's recommendations (ATCC). In each sample, cells were collected at a confluence status of 70% ensuring standardization of the cells proliferation level.

Total RNA extraction was performed with a phenol-chloroform method, using TriReagent (Sigma) for cells lysis and PhaseLockGel tubes (Eppendorf) for phase separation. After DNase treatment, DNA contamination of each sample was checked by electrophoresis on agarose gel. The synthesis of cDNA was done using the First-Strand cDNA Synthesis Kit (GE-Healthcare). Briefly, 1 µg of denatured total RNA was mixed with dithiothreitol (DTT), random hexadeoxynucleotides and a bulk reaction mix containing Murine Reverse Transcriptase. After 1 hour incubation at 37°C, the reaction was stopped by heating for 3 minutes at 95°C.

APC-11 specific primer pair used in RT-PCR amplification was commercial primers from Applied Biosystems. Cyclophilin B (CPB),  $\beta$ -actin and phosphoglycerokinase (PGK) was used as reference genes for the results normalisation. The primer pairs and probes were designed with Universal Probe Library website (Roche Applied Science). Primers of each pair were located in different exons to avoid genomic amplification. All reactions were performed in glass capillaries (Roche Applied Science) using LightCycler FastStart DNA Master PLUS SYBR Green I kit for APC11 or LightCycler Taqman Master for reference genes (Roche

Applied Science). The reaction mixture contained FastStart TaqDNA polymerase, reaction buffer, deoxynucleoside triphosphate mixture, MgCl<sub>2</sub>, 0.25 μM of each primer, and 1.67 μl (5 μL for reference genes) of a 1/60 dilution of first-strand cDNA from cells. The final reaction volume was 6.67 μl (10 μL for reference genes). Water was used as a negative control. All samples were kept at 4°C during preparation. Thermocycling conditions for APC-11 gene were designed in four consecutive steps (the Roche Applied Science Universal Probe Library's thermocycling conditions were used for the reference genes). Step 1: denaturation and polymerase activation by 10 min heating at 95°C prior amplification. Step 2: denaturation for 5 sec at 95°C, followed by 20 sec of annealing-extension at 60°C for 45 cycles. Step 3: melting which consisted of incrementing the temperature from 58°C to 95°C by steps of 0.1°C. Step 4: cooling for 30 sec at 40°C for one cycle. This melting curve program was used to check the specificity of each APC-11 RT-PCR. All standards or samples were analyzed in duplicate and each experiment was repeated at least three times. All fluorescence data were analyzed by the LightCycler 4.0 software (Roche) and Ct results were exported to Excel sheets.

A large amount of cDNA was prepared from the HME-1 cell line prior to the experiment. It was then diluted, aliquoted and used as a calibrator for all RT-PCR runs. For relative quantification and normalization, the comparative C<sub>t</sub> (or Eff<sup>ΔΔC</sup>) method was used.

#### *Statistical analysis*

Five RT-PCR were carried out for each of the 18 breast cancer cell lines for a total of 90 measures. A log<sub>2</sub> transformation was applied to APC11 RNA expression data since they did not initially meet normality requirements. Because the RT-PCR will tend to be correlated within cell lines, the use of a random effects model was appropriate (11). Consequently, the log<sub>2</sub> mean RNA expression for breast cancer cell lines was estimated using the following random effects model:

$$\log_2 (Y_{ij}) = \mu + \alpha_i + \varepsilon_{ij}$$

where  $Y_{ij}$  was the total APC11 RNA expression for measurement  $j$  of cell line  $i$ ,  $\mu$  was the overall mean after  $\log_2$  transformation,  $\alpha_i$  represented the random effects associated with cell lines and  $\varepsilon_{ij}$  the unexplained error. The random effects were assumed to be independent and normally distributed with zero mean and constant variance as follows:  $\alpha_i \sim N(0, \sigma_\alpha^2)$  and  $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ .

The statistical significance of the  $\alpha_i$  term (i.e. biological variation) was estimated using a forward likelihood profile analysis based on the likelihood ratio (LR) test. This was done by comparing the likelihood of the 2 following models:

$$\log_2 (Y_{ij}) = \mu + \varepsilon_{ij} \quad (\text{model 1})$$

and

$$\log_2 (Y_{ij}) = \mu + \alpha_i + \varepsilon_{ij} \quad (\text{model 2}).$$

Estimations were done using the maximum likelihood method using the *nlme* library of the software R (12). For likelihood ratio tests, the statistical significance level was 0.05.

## Results

Parameter estimations are presented in Table 1. The addition of the between cell lines variation improved significantly the likelihood of the model with an increase of the log-likelihood from -120 for model 1 to -114 for model 2 ( $p=0.0008$ ). In model 1, the measurement error was initially estimated at 0.92. Considering model 2, the measurement error was explained by a between-cell lines variation of  $\sigma_\alpha = 0.50$  (95% confidence interval (CI), 0.31 – 0.81) with a final residual standard deviation of  $\sigma_\varepsilon = 0.77$  (95% CI, 0.65 – 0.90). Conformity of the models to the assumptions for random effects and residual error were assessed and no evidence of deviation from these hypotheses was found (data not shown).

---

The estimates of APC11 RNA expression after 2<sup>h</sup> transformation were 2.81 (95% CI, 2.45 – 3.20) and 2.79 (95% CI, 2.30 – 3.41) in model 1 and 2 respectively.

## Discussion

The results showed that variation due to cell lines is not negligible and should be considered in the analysis of quantitative gene expression data.

More generally, failing to take into account the possible biological variability due to the choice of biological samples may expose to uncertainty regarding the interpretation of the results. In model 1, the confidence interval of the relative RNA expression was weaker than those in model 2. This is explained by the presence of correlation between measurements within cell lines. Because model 2 was found to be more appropriate than model 1 for estimating relative RNA expression, this suggested that failing to deal with biological variations in real-time RT-PCR may lead to a misinterpretation of the data.

One advantage of random effects models is that they allow estimating a fixed effect (e.g. a mean gene expression ratio) by simultaneously estimating and therefore adjusting for the variability identified by other variables in the dataset. The addition of supplementary covariates is possible and may help to further model the biological variation. The choice of the cell lines for RNA extraction as well as the problem of the culture of the cell lines used for gene expression analysis might be crucial. Cell line heterogeneity might reflect a particular proliferative advantage of certain cell lines under culture conditions. Expression may vary in response to alterations of culture conditions *in vitro*, such as variations in the composition of serum-supplemented media. Some cancer sublines have been derived from a single cell line, depending on their growth capacities in the presence or absence of some growth factors like EGF (13). Therefore, longer culture times may lead to a random genetic drift from the original cancer cell line. It would be interesting to further study the culture characteristics of

---

the cell lines. These characteristics could be introduced in a random effect model and may help to better estimate the relative RNA expression.

Interestingly random effects models could be applied to the analysis of tumor samples. Indeed, *in vivo* RNA extractions carried out from biopsies contain malignant cells but may also include adjacent, normal or abnormal, cells surrounding the tumor. It is expected that these different cells express different genotypes. Replicating measurements on a same tumor sample could allow considering the various expression profiles of the different cells within the tissue. In addition, some characteristics of the tumor or of the patient such as the tumor size or the patient age could be introduced in a random effect model in order to better explain the variation in gene expression.

We cannot rule out that variations in the comparative  $C_t$  method may have occurred. However, because this method was used throughout the study for each cell lines, it seems unlikely that it could confuse our conclusion about the biological variability observed between cell lines.

In conclusion, the biological heterogeneity associated with the biological material could be an important parameter that should be taken into account to correctly interpreting gene expression results in real-time RT-PCR. Using random effects models for data analysis focusing on biological processes could improve the interpretation of quantitative gene expression data as well as the comparability between studies. This statistical framework could be extended to various study designs and could be easily implemented in the software specialized in RT-PCR data processing.

### **Acknowledgments**

This study was partially funded by the “Comité Départemental de la Ligne Nationale contre le Cancer de la Drôme”. G. Grelier is the recipient of a PhD scholarship from the “Ligue

---

Nationale Contre le Cancer". The authors would like to thank René Ecochard for very helpful advice. The authors are indebted to M. D. Reynaud for editorial assistance.

## References

1. Wong ML, Medrano JF. (2005) Real-time PCR for mRNA quantitation. *Biotechniques*, **39**:75-85.
2. Bustin SA. (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol.*, **29**:23-39.
3. Vu HL, Troubetzkoy S, Nguyen HH, Russell MW, Mestecky J. (2000) A method for quantification of absolute amounts of nucleic acids by (RT)-PCR and a new mathematical model for data analysis. *Nucleic Acids Res.*, **28**:E18.
4. Pfaffl MW. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, **29**:e45.
5. Szabo A, Perou CM, Karaca M, Perreard L, Quackenbush JF, Bernard PS. (2004) Statistical modeling for selecting housekeeper genes. *Genome Biol.*, **5**:R59.
6. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.*, **3**:RESEARCH0034.
7. Andersen CL, Jensen JL, Orntoft TF. (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.*, **64**:5245-50.
8. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* 2007;8(2):R19



- 
9. Simon P. Q-Gene: processing quantitative real-time RT-PCR data. *Bioinformatics*. 2003 Jul 22;19(11):1439-40
  10. Peters JM. (2002) The anaphase-promoting complex: proteolysis in mitosis and beyond. *Mol Cell.*, **9**:931-43.
  11. Sullivan LM, Dukes KA, Losina E. (1999) Tutorial in biostatistics. An introduction to hierarchical linear modelling. *Stat Med.*, **18**:855-88.
  12. R Development Core Team (2004). R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
  13. Moyret C, Madsen MW, Cooke J, Briand P, Theillet C. (1994) Gradual selection of a cellular clone presenting a mutation at codon 179 of the p53 gene during establishment of the immortalized human breast epithelial cell line HMT-3522. *Exp Cell Res.*, **215**:380-5.

**Table 1.** Likelihood profile analysis. Model 2 was the more appropriate model for estimating relative RNA expression.

	$\mu$ (95%CI)	$\sigma_{\alpha}$ (95%CI)	$\sigma_{\epsilon}$ (95%CI)	Log Likelihood (degree of freedom)	LR test*
<b>Model 1</b>	1.49 (1.29 – 1.68)	-	0.92 (-)	-120 (2)	-
<b>Model 2</b>	1.48 (1.20 – 1.77)	0.50 (0.31 – 0.81)	0.77 (0.65 – 0.90)	-114 (3)	0.0008

\**P*-value for the likelihood ratio test (3-2=1 degree of freedom) between model 1 and model 2.



FORMATION SFC  
© John Libbey Eurotext

Bull Cancer 2008 ; 95 (10) : 1-8

# Génomique tumorale : l'image d'un paysage instable

## *Tumour genomics: an unstable landscape*

Caroline MOYRET-LALLE<sup>1,3</sup>  
Nicole FALETTE<sup>2</sup>  
Gaël GRELIER<sup>1,3</sup>  
Alain PUISIEUX<sup>1,2,3</sup>

<sup>1</sup> Inserm U590, 69008 Lyon  
<sup>2</sup> Centre Léon-Bérard, Unité d'oncologie moléculaire, 28, rue Laennec, 69008 Lyon  
<sup>3</sup> Université de Lyon-I, ISPB, 69008 Lyon  
<puisieux@lyon.fnclcc.fr>

**Résumé.** Le processus d'oncogenèse et de progression tumorale repose sur l'accumulation progressive d'anomalies génétiques et épigénétiques qui confèrent aux cellules pré-cancéreuses puis cancéreuses des capacités anormales de prolifération et de survie. Les avancées technologiques récentes, en particulier le développement des approches de CGH arrays et des techniques de séquençage à haut débit, permettent aujourd'hui de définir le profil génétique d'une tumeur. Les programmes de caractérisation moléculaire ont ainsi pour objectifs d'améliorer la classification des tumeurs, de développer de nouvelles approches thérapeutiques et d'identifier des biomarqueurs permettant d'optimiser la prise en charge des patients et leurs suivis cliniques. L'instabilité génétique des cellules cancéreuses et l'hétérogénéité intra-tumorale qui en résulte restent cependant des contraintes importantes dans le contexte de l'émergence des thérapeutiques ciblées. ▲

**Mots clés :** génomique tumorale, oncogenèse, instabilité génétique

**Abstract.** *Oncogenesis and tumour progression are caused by the progressive accumulation of genetic and epigenetic abnormalities in pre-cancerous and cancerous cells, conferring increased capabilities of proliferation and survival. Recent technological advances, including the development of CGH arrays and high-throughput sequencing, have made it possible to map the genetic landscape of human cancers. Molecular characterisation studies have provided key insights into the disease mechanisms that can be used for the design of tailored therapies and have led to the identification of specific biomarkers for guiding patient management. Nevertheless, the genetic instability of cancer cells and the consecutive intra-tumoral heterogeneity remain critical constraints in the context of the emergence of targeted therapies.* ▲

**Key words:** tumour genomics, oncogenesis, genetic instability

## Altérations génétiques et cancers

Le processus de cancérogenèse est un processus complexe, dicté par l'accumulation progressive d'altérations génétiques et épigénétiques affectant des facteurs contrôlant la différenciation, la division et la mort cellulaire. Il s'agit d'un phénomène dynamique qui, à l'image de la théorie de l'évolution des espèces selon Darwin, repose sur une succession de phases de sélections et d'expansions clonales, chacune d'entre elles faisant suite à l'acquisition par une cellule d'un avantage de prolifération ou de survie. La plupart des cancers sont d'origine monoclonale. Cela a pu être démontré initialement chez les femmes par l'analyse de l'origine paternelle ou maternelle du chromosome X inactivé dans les cellules tumorales qui montre que toutes les cellules tumorales ont le même chromosome X inactivé. Plus récemment, l'analyse de la séquence nucléotidique de sites de translocations chromosomiques spécifiques a confirmé cette observation. C'est le cas de la translocation t(9,22) de la leucémie myéloïde chronique, toujours identique dans une même tumeur chez un individu, mais dont les points de cassure diffèrent le plus souvent dans deux tumeurs de même histologie provenant de

deux individus différents. Origine monoclonale ne signifie pas homogénéité des cellules cancéreuses au sein d'une tumeur donnée. Son développement reposant sur une succession de vagues d'expansions clonales, une tumeur est en effet composée de différentes sous-populations de cellules cancéreuses ayant pour points communs les altérations survenues précocement. Comme nous le verrons par la suite, cette hétérogénéité intra-tumorale est favorisée par l'instabilité génétique des cellules cancéreuses.

Une anomalie unique n'est pas suffisante pour permettre à une cellule normale d'acquérir les propriétés caractéristiques des cellules cancéreuses : capacités accrues de division et insensibilité aux signaux inhibiteurs de prolifération, échappement au processus de vieillissement cellulaire (sénescence répliquative) et à la mort cellulaire programmée (apoptose), capacité à induire la formation de nouveaux vaisseaux (néo-angiogenèse) et capacités d'invasion et de dissémination à distance (métastase). Sur la base d'études *in vitro*, on estime que 5 à 10 événements génétiques affectant des oncogènes et des gènes suppresseurs de tumeur sont nécessaires à la transformation maligne d'une cellule humaine [1]. Plusieurs types

C. Moyret-Lalle, et al.

d'altérations génétiques sont observés dans les cellules cancéreuses [2] :

a) *Mutations subtiles*, de types substitutions de bases, petites délétions ou insertions de quelques nucléotides. Elles sont fréquentes dans les cancers humains. Un exemple classique est la mutation des codons 12, 13 ou 61 des oncogènes H-Ras ou K-Ras. De même, le gène suppresseur de tumeur *p53* est muté dans plus de 50 % des cancers humains, la mutation d'un allèle étant généralement associée à la perte de l'autre allèle par délétion génique.

b) *Translocations chromosomiques*. Ces altérations sont détectées par analyse cytogénétique comme des fusions de différents chromosomes ou de deux régions non-contigües d'un même chromosome. Dans certains cas, ces translocations sont à l'origine de protéines chimères associant le produit d'un gène activement transcrit et une partie des séquences codantes d'un oncogène, conduisant à la synthèse d'une protéine anormalement active. Un exemple de ce type de fusion est la protéine *bcr-abl*, produit de la translocation t(9,22) de la leucémie myéloïde chronique. Dans d'autres cas, la translocation place l'oncogène sous la dépendance du promoteur d'un gène activement transcrit, comme le gène des chaînes lourdes des immunoglobulines pour la t(8,14) des lymphomes de Burkitt. Les translocations chromosomiques sont fréquentes dans les tumeurs pédiatriques et les tumeurs adultes d'origine hématopoïétiques. Récemment, certaines translocations récurrentes ont été mises en évidence dans des tumeurs solides adultes, en particulier dans les cancers de la prostate [3].

c) *Amplifications géniques*. Elles sont détectées en cytogénétique sous forme de fragments extra-chromosomiques (chromosomes double minute) ou comme des enchaînements de copies se traduisant par des régions de coloration homogène (HSR). Les amplifications géniques constituent des mécanismes récurrents d'activation oncogénique. À titre d'exemple, l'amplification de *ErbB2* est fréquente dans les cancers du sein alors que l'amplification des oncogènes de la famille *Myc* est un événement récurrent dans une grande variété de cancers humains.

d) *Délétions géniques*. Très fréquentes dans les cancers humains, elles sont souvent impliquées dans l'inactivation des gènes suppresseurs de tumeur (*p53*, *Rb*...), en association avec des mutations ponctuelles affectant l'allèle complémentaire. Dans ce contexte, la délétion est alors appelée perte d'hétérozygotie (LOH).

e) *Altérations du nombre des chromosomes* : aneuploïdie. L'aneuploïdie provient du gain ou de la perte de chromosomes entiers. Elle caractérise la majorité des cancers humains. À titre d'exemple, la perte du chromosome 10 est fréquente dans les glioblastomes, et reflète l'inactivation du gène suppresseur de tumeur *PTEN*.

## Rôle et mécanismes de l'instabilité génétique au cours de la progression tumorale

Comme précisé précédemment, il est généralement admis que la transformation maligne d'une cellule nécessite l'altération de 5 à 10 gènes différents, conduisant à la dérégulation de voies de signalisation particulières. Les analyses récentes par séquençage de tumeurs humaines confortent l'hypothèse « multi-génique » des cancers, une tumeur donnée présentant, en moyenne, une quinzaine de mutations délétères,

accompagnées de dizaines de mutations dites passagères dans des régions codantes [4]. En prenant en compte l'intégrité du génome, ces observations suggèrent que les cellules cancéreuses présentent des milliers de mutations différentes. Dès 1974, Lawrence Loeb avait émis l'hypothèse selon laquelle le processus de développement tumoral nécessitait la perte de la stabilité du génome [5]. Deux observations sont en faveur de cette hypothèse. Tout d'abord, le taux de mutations spontanées est très significativement supérieur dans les cellules cancéreuses que dans les cellules normales [6]. Par ailleurs, plusieurs syndromes de prédisposition au cancer sont liés à l'altération spécifique de systèmes de réparation de l'ADN. C'est le cas du syndrome *Xeroderma pigmentosum* dont les personnes atteintes présentent une anomalie de gènes codant pour une des protéines impliquées dans la réparation des lésions de l'ADN dues aux ultraviolets (système NER, Réparation par Excision de Nucléotides). Ces personnes présentent un risque 200 fois supérieur à la population générale de développer des cancers cutanés. De même, la dérégulation des mécanismes de réparation de l'ADN de types NER ou BER (Réparation par Excision de Base), ou des systèmes de réparation des cassures double-brin peut être à l'origine de certains cancers. Enfin, le syndrome de prédisposition héréditaire du colon non polyposique (syndrome HNPCC) est lié à la transmission d'une anomalie affectant des gènes (*hMSH2*, *hMLH1* ou plus rarement *hMSH6*) impliqués dans la réparation des mésappariements de l'ADN.

Ces notions ont conduit Kinzler et Vogelstein à définir deux catégories de gènes dérégulés dans les cancers et pouvant générer une instabilité génétique : les « caretakers » et les « gatekeepers » [7]. Les « gatekeepers » participent aux mécanismes de contrôle du cycle cellulaire et de l'homéostasie cellulaire dans les tissus. Les « caretakers » codent pour des protéines qui stabilisent le génome en prévenant l'accumulation de mutations aussi bien au niveau nucléotidique qu'au niveau chromosomique. Certains modèles mathématiques de la cancérogenèse proposent que des mutations affectant les gènes « caretakers » conduisent à des mutations dans les gènes « gatekeepers ».

### Les différents types d'instabilité génétique

L'instabilité génétique traduit une augmentation du taux d'altérations génétiques dans les cellules cancéreuses. La présence d'une ou plusieurs mutations (ou d'une aneuploïdie) dans une cellule ne signifie donc pas que cette cellule est génétiquement instable. Ces mutations peuvent en effet être la conséquence d'une instabilité passée ou le résultat d'un processus de sélection et d'expansion clonale sur la base d'un taux normal de mutations spontanées. À ce jour, trois types principaux d'instabilité génétique ont été décrits : l'instabilité microsatellitaire l'instabilité chromosomique et, très récemment, l'instabilité de type mutation ponctuelle.

a) *L'instabilité microsatellitaire* (MSI ou MIN) a été initialement identifiée dans le syndrome de prédisposition héréditaire aux cancers du colon non polyposique ou syndrome HNPCC [8]. Un phénotype MSI est également observé dans 15 % des cancers colorectaux sporadiques, des cancers de l'estomac et de l'endomètre. Il est plus rare dans les autres types de cancers. Les cellules tumorales des patients affectés sont le plus souvent diploïdes mais présentent une altération du nombre de répétitions des séquences microsatellites. Cette instabilité génétique est due à l'inactivation du mécanisme de

réparation des mésappariements de l'ADN, par mutation ou hyperméthylation des gènes *hMLH1*, *hMSH2* ou *hMSH6*. Ce système est impliqué dans la réparation des erreurs commises par l'ADN polymérase au cours de la réplication de l'ADN. Cette enzyme faisant essentiellement des erreurs au niveau des séquences répétitives de l'ADN (« microsatellites »), cette instabilité est caractérisée par une accumulation de mutations subtiles (le plus souvent de type micro-délétions ou micro-insertions) au niveau des microsatellites. Ces séquences sont localisées préférentiellement dans des régions non codantes du génome. Toutefois, des mutations délétères ont été identifiées dans des segments monotones de G, C ou A localisés au sein de plusieurs gènes impliqués dans le contrôle de la croissance cellulaire tels que le gène du récepteur II du TGF $\beta$  (*TGF $\beta$ RII*, segment A<sub>10</sub>), le gène *MSH6* (segment C<sub>8</sub>), le gène du récepteur du facteur de croissance *insulin-like* (*IGF-R*; segment G<sub>8</sub>) ou le gène *Bax* (segment G<sub>8</sub>). Ces mutations provoquent l'inactivation du gène concerné, et participent au processus de transformation. Les cellules tumorales présentant une instabilité microsatellitaire restent le plus souvent diploïdes et n'ont pas ou peu d'altérations chromosomiques visibles au niveau cytogénétique.

b) L'*instabilité chromosomique* (CIN) est observée dans la majorité des cancers épithéliaux. Elle constitue un type d'instabilité génomique plus globale avec anomalie du nombre des chromosomes (aneuploïdie), cassures chromosomiques, pertes ou amplifications de portions de chromosomes. Cette forme d'instabilité apparaît être un caractère dominant, qui persiste au cours de toute la vie de la cellule tumorale. Alors que longtemps l'aneuploïdie et les altérations structurales des chromosomes ont été considérées comme des processus secondaires à la progression tumorale, de nombreuses observations ont conduit à les présenter dorénavant comme des processus moteurs de la cancérogenèse. En accord avec cette hypothèse, l'analyse de tumeurs bénignes (adénomes) a permis de montrer que ces altérations chromosomiques intervenaient précocement au cours de la progression tumorale. Sur le modèle des cancers colorectaux, les instabilités microsatellitaires et chromosomiques ont été présentées comme mutuellement exclusives et toutes les tumeurs non MSI étaient considérées comme CIN. La réalité est probablement plus complexe puisque certaines tumeurs ne présentent ni MSI, ni CIN, alors que certaines tumeurs MSI présentent une instabilité chromosomique modérée.

c) L'*instabilité de type mutation ponctuelle* (PIN) correspond à une augmentation du taux de mutations spontanées dans les cellules cancéreuses. L'équipe de Lawrence Loeb a estimé récemment ce taux à  $210 \times 10^{-8}$  par paire de base, soit un taux 200 fois plus élevé que celui des cellules normales ( $1 \times 10^{-8}$  par paire de base) [6]. Cette augmentation globale du taux de mutations augmenterait la probabilité de dérégulation d'un gène impliqué dans le maintien de la stabilité génétique, favorisant ainsi l'émergence d'une instabilité de type MIN ou CIN. Par ailleurs, l'instabilité de type PIN jouerait un rôle crucial dans l'apparition de sous-clones de cellules cancéreuses au cours de la progression maligne, favorisant ainsi l'hétérogénéité intra-tumorale.

### Mécanismes d'instabilité chromosomique

Les mécanismes moléculaires à l'origine de l'instabilité chromosomique restent mal connus. De nombreuses observations suggèrent cependant qu'elle serait secondaire à la dérégulation, par mécanisme génétique ou épigénétique, de certains

facteurs qui régulent les points de contrôle mitotiques et les fonctions télomériques.

a) Points de contrôle du cycle cellulaire et instabilité chromosomique

Dans les cellules eucaryotes, le maintien de la stabilité génétique est assuré par des systèmes de contrôle complexes appelés « points de contrôle du cycle cellulaire ». Ces points de contrôle sont activés au cours de la réponse cellulaire aux altérations de l'ADN, afin de prévenir une réplication et/ou une ségrégation de chromosomes anormaux ; ils régulent la réplication de l'ADN pour prévenir la mitose de génomes imparfaitement dupliqués et ils contrôlent les processus essentiels de la mitose tels que la ségrégation des chromosomes, l'assemblage du fuseau mitotique et le positionnement approprié des chromosomes au cours de la métaphase (point de contrôle du fuseau mitotique). Des altérations de facteurs directement impliqués dans ces points de contrôle mitotiques (Bub1, BubR1, Mad2, sous-unités du complexe APC/C, sécurine, STK15/Aurora2...) ont été identifiées dans quelques cancers solides. Expérimentalement, l'introduction des formes mutées de ces gènes dans des cellules diploïdes provoque l'apparition d'une instabilité chromosomique et d'une aneuploïdie, suggérant qu'elles pourraient être causales. À titre d'exemple, la perte d'expression de la sécurine, protéine qui assure la cohésion des chromatides sœurs en plaque métaphasique en inhibant l'enzyme responsable de la migration des chromatides sœurs aux pôles opposés de la cellule, provoque une erreur de ségrégation des chromatides sœurs en anaphase et une instabilité numérique du génome tumoral [9]. La dérégulation du point de contrôle du fuseau mitotique défectueux suite à la perte d'expression des protéines des familles Mad et Bub, induit quant à elle une endoreduplication, ou un dysfonctionnement de la cytodièrese.

Le rôle de l'inactivation de la protéine onco-suppressive *p53* dans la genèse de l'instabilité génétique reste mal déterminé. *In vitro*, l'inactivation de *p53* et l'aneuploïdie apparaissent généralement de façon simultanée, suggérant un lien de cause à effet. De même les cellules déficientes pour le gène *p53* présentent des fuseaux mitotiques anormaux [10]. Cependant, de nombreuses lignées cellulaires humaines présentent une mutation de *p53*, sans instabilité chromosomique. De même, l'inactivation de *p53* par recombinaison homologue dans des cellules de cancers du côlon avec une instabilité microsatellitaire, ne conduit pas à une instabilité chromosomique [11]. Enfin, alors que l'instabilité chromosomique semble apparaître de façon précoce au cours de la progression tumorale, les mutations affectant *p53* sont fréquemment de survenue tardive dans les cancers humains. Ces différentes observations suggèrent que la perte de *p53* permettrait d'exacerber une instabilité préexistante, mais ne constituerait pas l'élément initiateur.

b) Érosion télomérique, défauts de réparation de l'ADN et amplifications géniques

L'amplification est le mécanisme principal d'activation des oncogènes dans les tumeurs solides. Ainsi, un grand nombre d'oncogènes sont des cibles d'amplifications géniques récurrentes : *ErbB2* dans les cancers du sein, *AKT2* dans les cancers ovariens, *N-Myc* dans le neuroblastome, *Rel* dans le lymphome de Burkitt, *EGFR* dans les gliomes et des cancers du poumon non-à petite cellule et *c-Myc* dans un grand nombre de pathologies tumorales [12]. Les amplifications sont observées sous forme de fragments extra-chromosomiques acentriques (chromosomes double minute) ou

C. Moyret-Lalle, et al.

correspondent à des enchaînements de copies se traduisant par des régions de coloration homogène (HSR). Selon le modèle proposé par Barbara McClintock, l'événement initiateur de l'HSR serait constitué par des cassures double brin (DSB, double strand break) ou par l'érosion télomérique, conduisant à la formation de pontages (BFB : breakage-fusion-bridge) [13]. Depuis, il a été montré que l'amplification chromosomique peut être générée à partir d'échange inégal de chromatides sœurs, d'une cassure au niveau d'une boucle de réplication inappropriée, ou d'une amplification extra-chromosomique réintégrée dans le génome [14]. Un certain nombre de facteurs extrinsèques et intrinsèques peuvent induire des cassures double brin. À titre d'exemple, au cours du vieillissement, la formation de radicaux libres et le raccourcissement télomérique sont à l'origine de cassures DSB. De même différents facteurs environnementaux et des carcinogènes tels que le tabac et l'amiante ou des virus tels que le virus HIV peuvent induire des cassures conduisant à une amplification chromosomique [15]. Plusieurs systèmes de réparation de l'ADN sont impliqués dans la réparation des cassures DSB. Des cellules déficientes pour *hMLH1* ou *hMSH6* (réparation des mésappariements, MMR) traitées au méthotrexate donnent naissance à 15 fois plus de colonies résistantes au méthotrexate que les cellules ayant un système MMR fonctionnel. Cette résistance est due à une amplification du gène *DHFR* et à l'accumulation de chromosomes double minute [16]. Une déficience du mécanisme de réparation par NHEJ (Réparation par Jonction des Extrémités Non -Homologues) est également responsable d'une augmentation de la fréquence de survenue d'amplifications chromosomiques.

#### Impact de l'instabilité épigénétique sur l'instabilité génétique ?

Plusieurs observations suggèrent qu'il existe un lien entre certains processus épigénétiques et l'instabilité génétique. À titre d'exemple une hypométhylation des régions centromériques est à l'origine d'instabilités chromosomiques dans les cellules sanguines dérivées de patients atteints du syndrome ICF (Immunodeficiency - Centromeric region instability - Facial anomalies) [17]. Le syndrome ICF est lié à une anomalie de méthylation dans les régions hétérochromatiniennes. Cette déficience, qui prédispose à certains types de cancers, est associée à des anomalies chromosomiques majeures dans les régions péricentromériques des chromosomes 1 et 16. À l'exception d'associations télomériques, aucune autre anomalie affectant d'autres chromosomes n'a été mise en évidence. Cependant, à l'heure actuelle, la nature des dommages génomiques liés à l'hypométhylation globale de l'ADN dans les cellules cancéreuses reste mal connue (figures 1-3). Il manque les appels des figures. Merci de vérifier l'emplacement des figures .

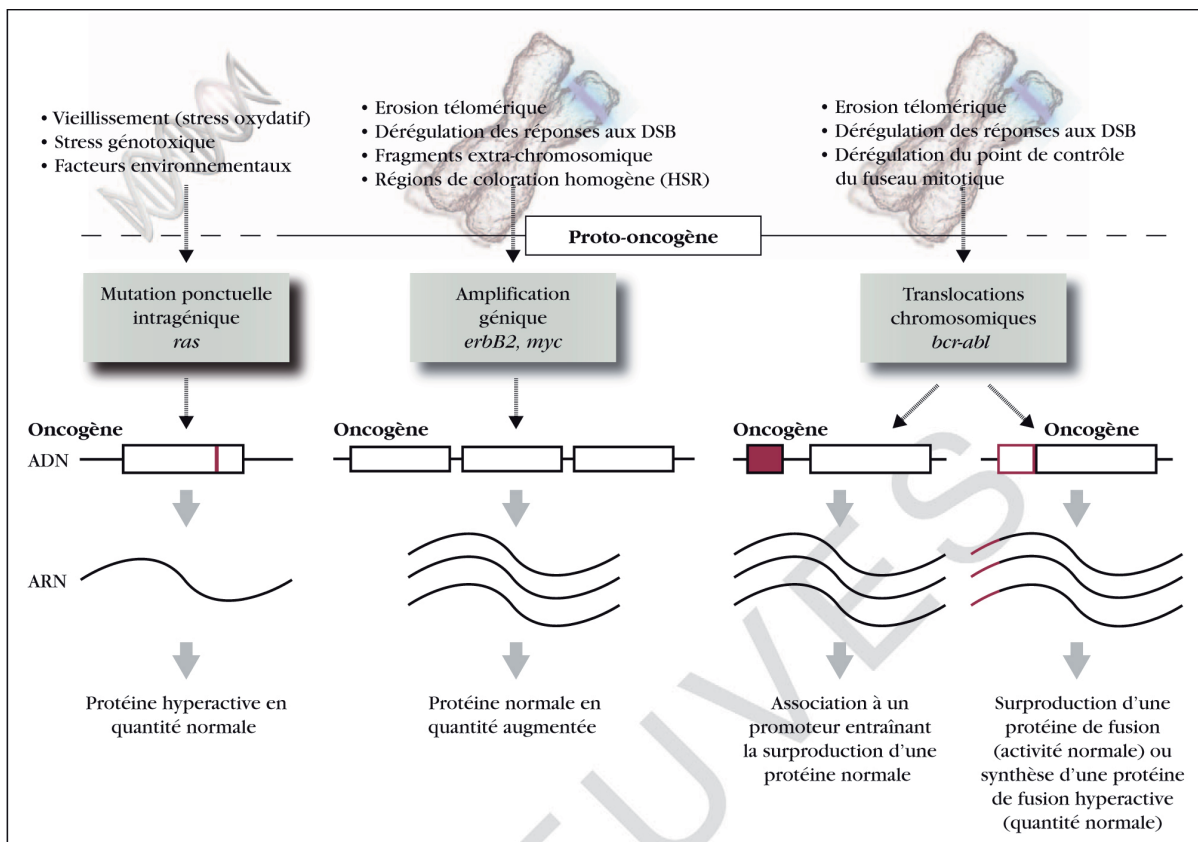
#### MicroARNs et instabilité génomique

Au cours de ces dernières années, une nouvelle classe de molécules a été impliquée dans les processus de cancérogenèse et de progression tumorale : les microRNAs (miRNAs). Les miRNAs sont des ARNs non codants transcrits par l'ARN polymérase II, dont l'expression est régulée au niveau cellulaire et tissulaire. Ils sont transcrits sous forme de précurseurs (pri-miRNA), avant de subir une maturation conduisant à des miRNAs matures d'une taille comprise entre 20 et 24 nucléo-

tides. Leur fonction est de réguler l'expression génique. Dans le contexte d'un complexe protéique, ils vont se lier à des ARNm cibles et provoquer leur dégradation ou inhiber leur traduction. Les miRNAs ont été impliqués dans le contrôle de nombreux processus fondamentaux, tels que le développement, la différenciation cellulaire, la prolifération, l'apoptose et le maintien des cellules souches [18]. Du fait de ces fonctions majeures, il n'est pas surprenant de constater qu'un certain nombre de miRNAs peut se comporter comme des oncogènes ou des gènes suppresseurs de tumeur. Ces miRNAs, appelés oncomirs sont exprimés de façon anormale dans les cancers et sont généralement localisés dans des régions chromosomiques altérées dans les cellules cancéreuses. De façon intéressante, plusieurs études récentes ont montré que des miRNA étaient localisés dans des régions d'instabilité génétique ou des sites fragiles [19]. Trente-sept pour cent des miRNA, chez l'Homme, sont rassemblés sous forme de clusters. À ce jour, le cluster le plus grand identifié est le cluster hsa-miR-127, qui contient 50 miRNAs. Il réside dans une région d'empreinte génomique sur le chromosome 14q32. Cette région est fréquemment la cible de translocations et d'inversions dans des leucémies T [20]. De même, un membre de la famille hsa-let-7, cluster en 9q22, a été montré comme inhibant l'expression de *Ras* [21]. Ce cluster de 6 miRNAs est localisé dans une région de 1.46 Mb qui recouvre une région fréquente de délétion. Ces observations suggèrent fortement que l'instabilité génétique des cellules cancéreuses peut avoir des conséquences significatives sur les miRNAs, et donc sur le contrôle de l'expression génique. Il reste à déterminer si cette dérégulation peut, en retour, exacerber une instabilité génétique préexistante.

#### Instabilité génétique et cancers : un sujet toujours controversé

Selon l'hypothèse initialement émise par Lawrence Loeb, l'instabilité chromosomique est secondaire à l'inactivation, par mutation, de gènes impliqués dans la régulation des points de contrôle du cycle cellulaire et des processus de ségrégation des chromosomes au cours de la mitose. Elle pourrait ainsi être secondaire à l'apparition d'une instabilité de type PIN. Malgré les arguments cités précédemment, cette hypothèse reste cependant un sujet de débat et de controverses. Certains défendent ainsi une théorie, dite « non mutationnelle », qui place l'aneuploïdie au centre du processus de progression tumorale. Selon cette hypothèse, l'aneuploïdie provient de l'action d'agents non mutagènes agissant sur des protéines régulatrices du processus mitotique. L'aneuploïdie provoquerait alors un déséquilibre d'expression génique à la base de la transformation maligne. L'instabilité chromosomique ne serait qu'un phénomène secondaire, acquis au cours de la progression tumorale. D'autres défendent la théorie selon laquelle l'instabilité génétique n'est qu'une conséquence très secondaire du développement tumoral. Selon cette hypothèse, les mutations apparaîtraient à un taux normal mais la succession de multiples cycles de sélection et d'expansion serait à l'origine du processus de cancérogenèse. Au cours de la progression tumorale, le taux anormal de divisions des cellules cancéreuses serait responsable de l'apparition d'une aneuploïdie. En lien indirect avec cette dernière hypothèse, il est nécessaire de souligner l'importance des interactions cellules-cellules au cours de la progression tumorale. Ainsi, il est maintenant clairement établi que l'interaction des cellules cancéreuses avec les cellules stromales



**Figure 1.** Principaux mécanismes d'activation des oncogènes. La transition de proto-oncogène à oncogène provient de différents mécanismes qui impliquent l'ADN aussi dans bien sa structure que dans sa composition. On retrouve ainsi 3 grands mécanismes que sont les mutations ponctuelles, l'amplification génique et les translocations chromosomiques. Quelle que soit la voie engagée, le résultat de cette transition est soit une protéine hyperactive en quantité normale, soit une protéine normale en grande quantité.

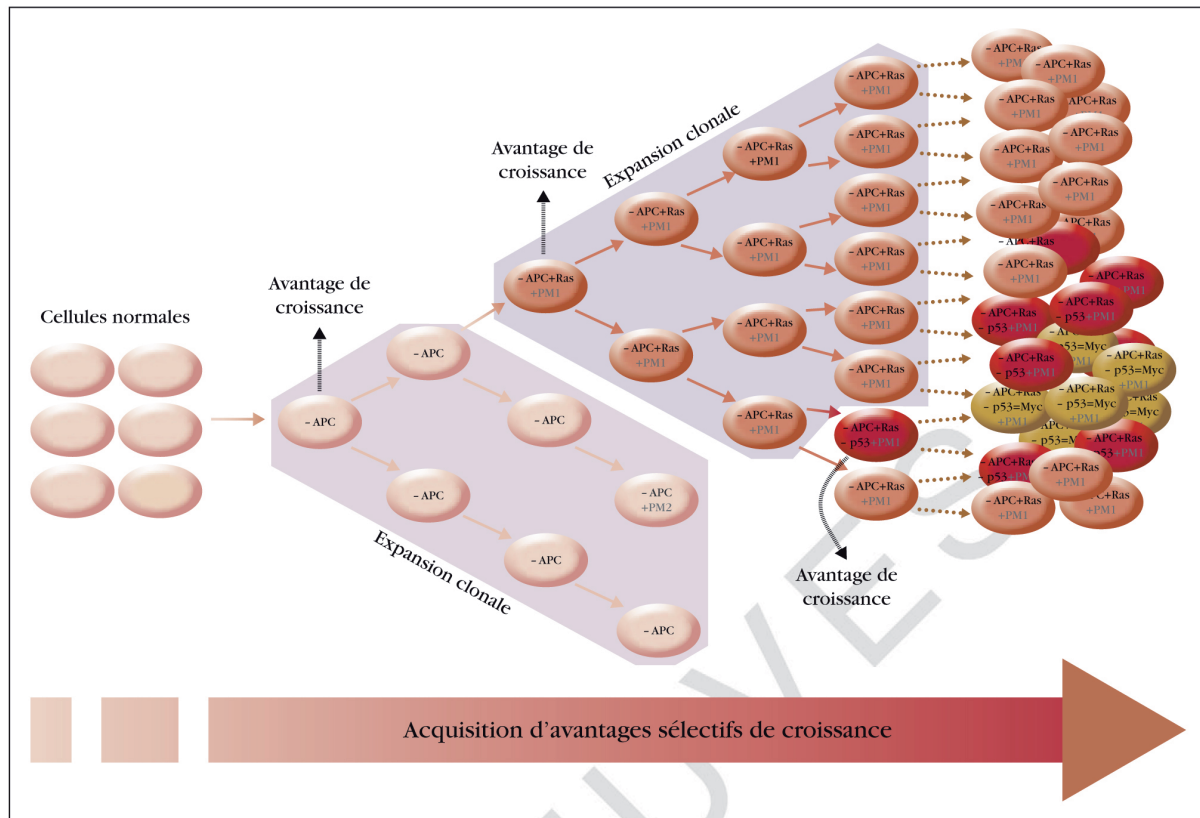
environnementales et les cellules endothéliales sont des événements cruciaux pour obtenir les facteurs de croissance nécessaires et pour développer le processus d'angiogenèse. Au-delà des notions de « caretakers » et de « gatekeepers », cela a conduit à la proposition d'une troisième catégorie de gènes impliqués dans le processus de progression tumorale : les « landscapers ». Leur dérégulation induit des changements dans l'environnement cellulaire, au niveau du stroma, conduisant à une prolifération accrue des cellules épithéliales [22].

### Vers la définition de « paysages mutationnels » des cancers : l'exemple des cancers du sein et des cancers du côlon

Les avancées technologiques récentes en matière de séquençage ont permis d'établir de façon plus précise les événements mutationnels intervenant au cours de la progression tumorale. Les équipes de M. Stratton en Grande-Bretagne (Wellcome Trust Sanger Institute, Hinxton) et de B. Vogelstein aux États-Unis (Howard Hughes Medical Institute, Baltimore) ont été les premières à développer des programmes très ambitieux de séquençage systématique de tumeurs humaines. À ce jour, le programme mis en place au Wellcome Trust

Sanger Institute (*Cancer Genome Project*) a permis de mettre en évidence des mutations ponctuelles récurrentes de 10 gènes (*BRAF*, *ErbB2*, *FGFR2*, *FGFR3*, *FLT3*, *HRas*, *KRas2*, *Met*, *NRas*, *PDGFRA*) dans différents types de cancers [23]. L'équipe du Howard Hughes Medical Institute a dans un premier temps entrepris le séquençage systématique des kinases et phosphatases dans les cancers colorectaux, conduisant notamment à l'identification de mutations récurrentes du gène *PI3KCA* [24-26]. Très récemment, cette même équipe a entrepris l'analyse de 20 857 transcrits correspondant à 18 191 gènes dans 11 cancers colorectaux et 11 cancers du sein, avec, dans chaque cas, une analyse parallèle du tissu adjacent correspondant [4]. Chaque gène retrouvé muté dans une tumeur était analysé dans 24 tumeurs supplémentaires du type du cancer correspondant. Au total, 1 718 gènes, correspondant à 9,4 % des gènes analysés, étaient affectés par au moins une mutation non silencieuse dans une tumeur. Plus de 92 % des mutations observées correspondaient à des substitutions d'une base, le reste étant des insertions, des délétions ou des duplications. 82 % des substitutions étaient responsables de faux-sens, 6 % généraient un codon Stop, et 4 % affectaient les jonctions intron-exon ou les régions non traduites immédiatement adjacentes au codon d'initiation ou au codon Stop. Une difficulté majeure de ce type d'analyse est de différencier les mutations jouant un rôle causal au cours de

C. Moyret-Lalle, et al.



**Figure 2.** « Driver mutations » et « Passenger mutations » Illustration, à travers l'exemple des principaux gènes retrouvés mutés dans le cancer du côlon, de la notion de « driver mutation » jouant un rôle causal au cours de la progression tumorale et de mutations passagères (PM1, PM2...). Ces dernières sont sans effets sur la tumorigénèse mais sont favorisées par l'instabilité génétique et retenues de façon passive par les cellules.

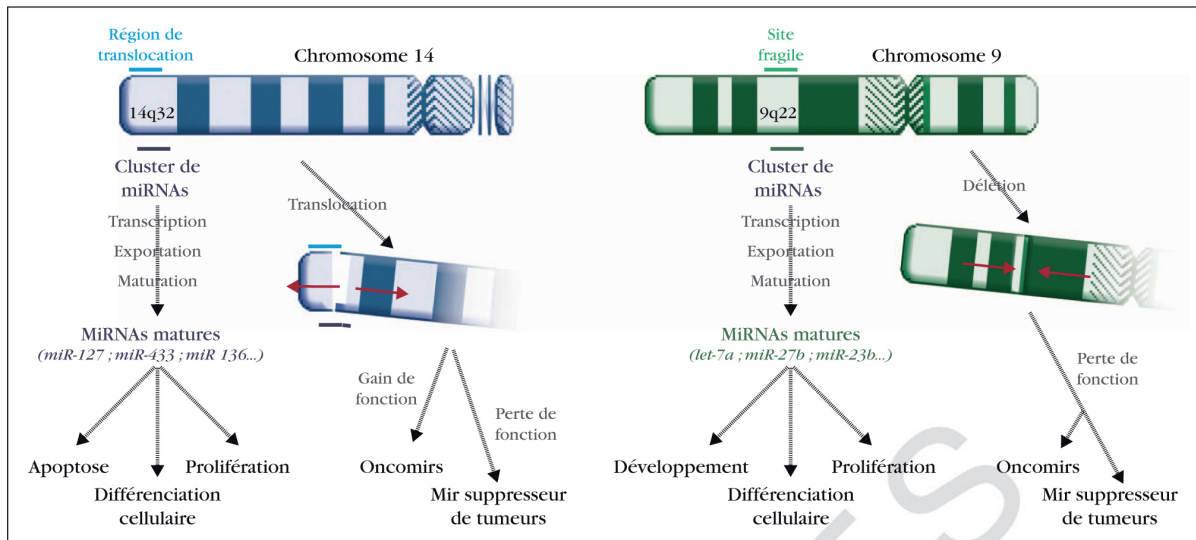
la progression tumorale (« driver mutations »), des mutations dites passagères (« passenger mutations »), sans effet sur le processus de cancérogenèse mais dont la survenue est facilitée par l'instabilité génétique. Les premières sont sélectionnées de façon positive durant la tumorigénèse alors que les secondes sont retenues de façon « passive » au cours des différentes étapes du cycle cellulaire et de l'expansion clonale. Sur la base du nombre et du type de mutations affectant chacun des gènes, les auteurs ont identifié 280 gènes différents, appelés « gènes de cancer », potentiellement impliqués dans le développement des tumeurs étudiées. Parmi ceux-ci, la majorité présentait des mutations dans moins de 5 % des cancers étudiés, confirmant l'hypothèse selon laquelle un grand nombre de mutations, chacune d'elle étant associée à un faible avantage de croissance, est impliqué dans le processus de progression tumorale. L'analyse fonctionnelle des protéines affectées montre cependant que si le nombre de gènes potentiellement mutés est élevé, le nombre de voies de signalisation altérée par ces mutations est relativement faible (10 à 20). Cette interprétation est confortée par de multiples analyses d'organismes uni et multi-cellulaires qui démontrent que le même phénotype peut être obtenu par altération d'une grande diversité de gènes. Aussi, ces résultats montrent que l'anomalie de voies de signalisation particulières, plutôt que de gènes particuliers, constitue la force motrice commune au processus de progression tumorale. Selon cette même étude,

le nombre de mutations détectées dans une tumeur donnée était relativement similaire dans les cancers colorectaux étudiés (variant de 49 à 111), alors qu'une variation plus importante était observée dans les tumeurs mammaires (de 38 à 193). En moyenne, le nombre de mutations affectant des « gènes du cancer » était de 15 dans les cancers colorectaux et de 14 dans les cancers du sein. Ces mutations pathogéniques sont accompagnées de plusieurs dizaines de mutations passagères au niveau des régions codantes, suggérant la présence de plusieurs milliers d'altérations dans l'ensemble du génome des cellules cancéreuses.

## Conclusions et perspectives

Les programmes de caractérisation moléculaire des cancers ont démontré amplement la complexité moléculaire et l'hétérogénéité des tumeurs humaines. Les progrès technologiques dans le domaine du séquençage et des approches de caractérisation moléculaire de type CGH array, ont contribué à identifier certains événements récurrents dans les principaux types de cancers. Bien que l'observation d'un très grand nombre de gènes mutés à une fréquence faible dans les cancers rende plus délicate une parfaite compréhension mécanistique, ces avancées ouvrent des perspectives importantes dans le domaine de la prise en charge personnalisée des patients atteints de cancer. La caractérisation des altérations





**Figure 3.** Hypothèse de l'implication des miRNAs dans la progression tumorale. La découverte récente de l'organisation en clusters de nombreux miRNAs et de leur localisation au niveau de sites fragiles a conduit à l'émergence d'une nouvelle hypothèse sur leur fonction. Ils seraient ainsi susceptibles de se comporter, de par leurs fonctions régulatrices, comme des oncogènes ou des gènes suppresseurs de tumeur. L'altération de la structure des clusters et/ou leur translocation générées par l'instabilité génétique modifieraient les quantités de miRNAs produits dans la cellule aboutissant à la dérégulation de gènes cibles clés.

structurales du génome tumoral par CGH array a permis d'améliorer le diagnostic et l'évaluation pronostique de certains types cancéreux. Le séquençage permettra, par l'identification des gènes mutés, de développer de nouvelles approches de thérapies ciblées, et de sélectionner de façon fiable les patients susceptibles de bénéficier de tel ou tel traitement. L'importance des processus d'instabilité génétique dans les cellules cancéreuses est cependant à prendre en compte dans le contexte de l'émergence de ces nouvelles thérapies. En effet, combinée à la présence de multiples sous-clones de cellules cancéreuses, cette instabilité pourrait être à l'origine d'un réservoir de mutations disponibles pour permettre l'émergence de résistances. Cette hypothèse conforte le concept d'associations de molécules anticancéreuses, une cellule donnée ayant une probabilité plus faible de présenter des mutations lui conférant une résistance à des agents doués de propriétés cytotoxiques distinctes. ▼

## RÉFÉRENCES

- Hahn WC, Counter CM, Lundberg AS, Beijersbergen RL, Brooks MW, Weinberg RA. Creation of human tumour cells with defined genetic elements. *Nature* 1999 ; 400 : 464-8.
- Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* 1998 ; 396 : 643-9.
- Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent Fusion of *TMPRSS2* and *ETS* Transcription Factor Genes in Prostate Cancer. *Science* 2005 ; 310 : 644-8.
- Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancer. *Science* 2007 ; 318 : 1108-13.
- Loeb LA, Springgate CF, Battula N. Errors in DNA replication as a basis of malignant changes. *Cancer Res* 1974 ; 34 : 2311-21.

- Bielas JH, Loeb KR, Rubin BP, True LD, Loeb LA. From the Cover : human cancers express a mutator phenotype. *PNAS* 2006 ; 103 : 18238-42.
- Kinzler KW, Vogelstein B. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* 1997 ; 386 : 761-3.
- Hamelin R, Chalastanis A, Colas C, El Bchiri J, Mercier D, Schreurs AS, et al. Clinical and molecular consequences of microsatellite instability in human cancers. *Bull Cancer* 2008 ; 95(1) : 121-32.
- Kops GJ, Weaver BA, Cleveland DW. On the road to cancer : aneuploidy and the mitotic checkpoint. *Nat Rev Cancer* 2005 ; 5 : 773-85.
- Cross SM, Sanchez CA, Morgan CA, Schimke MK, Ramel S, Idzerda RL, et al. A p53-dependent mouse spindle checkpoint. *Science* 1995 ; 267 : 1353-6.
- Bunz F, Fauth C, Speicher MR, Dutriaux A, Sedivy JM, Kinzler KW, et al. Targeted inactivation of p53 in human cells does not result in aneuploidy. *Cancer Res* 2002 ; 62 : 1129-33.
- Futreal PA, Coin L, Marshall M, Hubbard T, Wooster R, et al. A census of human cancer genes. *Nat Rev Cancer* 2004 ; 4 : 177-83.
- McClintock B. Chromosome organization and genic expression. *Cold Spring Harbor Symp Quant Biol* 1951 ; 16 : 13-47.
- Stark GR. Regulation and mechanisms of mammalian gene amplification. *Adv Cancer Res* 1993 ; 61 : 87-113.
- Shimura M, Tanaka Y, Nakamura S, et al. Micronuclei formation and aneuploidy induced by Vpr, an accessory gene of human immunodeficiency virus type 1. *FASEB J* 1999 ; 13 : 621-37.
- Lin CT, Lyu YL, Xiao H, Lin WH, Whang-Peng J. Suppression of gene amplification and chromosomal DNA integration by the DNA mismatch repair system. *Nucleic Acids Res* 2001 ; 29 : 3304-10.
- Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, et al. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 1999 ; 402 : 187-91.
- Jeanteur P. RNA and cancer. *Bull Cancer* 2008 ; 95(1) : 33-41.
- Calin GA, Liu CG, Sevignani C, Ferracin M, Felli N, Dumitru CD, et al. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci USA* 2004 ; 101 : 11755-60.

C. Moyret-Lalle, *et al.*

20. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* 2006 ; 103 : 2257-61.
21. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, *et al.* RAS is regulated by the let-7 microRNA family. *Cell* 2005 ; 120 : 635-47.
22. Michor F, Iwasa Y, Rajagopalan H, Lengauer C, Nowak MA. Linear model of colon cancer initiation. *Cell Cycle* 2004 ; 3 : 358-62.
23. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, *et al.* Mutations of the *BRAF* gene in human cancer. *Nature* 2002 ; 417 : 949-54.
24. Bardelli A, Parsons DW, Silliman N, Ptak J, Szabo S, Saha S, *et al.* Mutational Analysis of the Tyrosine Kinome in Colorectal Cancers. *Science* 2003 ; 300 : 949.
25. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, *et al.* High frequency of mutations of the *PIK3CA* gene in human cancers. *Science* 2004 ; 304 : 554.
26. Wang Z, Shen D, Parsons DW, Bardelli A, Sager J, Szabo S, *et al.* Mutational analysis of the tyrosine phosphatome in colorectal cancers. *Science* 2004 ; 304 : 1164-6.

ÉPREUVES