

Université de Sherbrooke

**Caractérisation de l'expression des snoRNA et recherche de biomarqueurs dans le cancer séreux de l'ovaire**

Par  
Audrey Roy  
Programme de maîtrise en biochimie

Mémoire présenté à la Faculté de médecine et des sciences de la santé  
en vue de l'obtention du grade de maître ès sciences (M. Sc.)  
en biochimie

Sherbrooke, Québec, Canada  
Mai 2019

Membres du jury d'évaluation  
Michelle Scott, département de biochimie, FMSS, Université de Sherbrooke  
François Bachand, département de biochimie, FMSS, Université de Sherbrooke  
Dr Maxime Richer, département d'anatomopathologie, FMSS, Université de Sherbrooke

© Audrey Roy, 2019

## RÉSUMÉ

### Caractérisation de l'expression des snoRNA et recherche de biomarqueurs dans le cancer séreux de l'ovaire

Par  
Audrey Roy  
Programme de maîtrise en biochimie

Mémoire présenté à la Faculté de médecine et des sciences de la santé en vue de l'obtention du diplôme de maître ès sciences (M.Sc.) en biochimie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4

Le cancer séreux de l'ovaire est le cancer gynécologique le plus mortel, et est au cinquième rang de mortalité par cancer au Canada. Il représente un fardeau important dû aux difficultés reliées à son diagnostic précoce et au manque de biomarqueurs de dépistage sensibles et spécifiques. En effet, le taux de survie à cinq ans est de 15% à un stade avancé, alors qu'il est de 90% s'il est traité au stade précoce. Or, seulement 25% des cancers de l'ovaire sont diagnostiqués à un stade précoce. Il est donc primordial de trouver des nouveaux biomarqueurs. Les snoRNA (petits ARN nucléolaires), une famille d'ARN non-codants impliqués dans la modification de l'ARN ribosomal, apparaissent comme de potentiels acteurs dans la carcinogénèse, certains d'entre eux ayant été identifiés comme biomarqueurs diagnostics et pronostics dans des cancers spécifiques. Des études phénotypiques et mécanistiques ont également montré un impact de certains snoRNA dans le développement et la progression du cancer. Ils n'ont en revanche jamais été étudiés dans le cancer de l'ovaire. Ils constituent donc une source potentielle de biomarqueurs. Le but de ce mémoire est donc de caractériser l'expression des snoRNA dans le cancer séreux de l'ovaire ainsi que d'identifier des biomarqueurs potentiels. Pour ce faire, le séquençage à haut débit de l'ARN (RNA-Seq) de trois tissus cancéreux ovariens de haut grade (HGSC) et trois tissus de tumeurs séreuses limites (SBT) a été réalisé grâce à une nouvelle méthode avec une rétrotranscriptase thermostable d'intron de groupe II (TGIRT) qui permet de mieux détecter les snoRNA, traditionnellement mal représentée dans les séquençages classiques. Une analyse des données de séquençage adaptée aux caractéristiques propres à cette famille d'ARN avec l'outil de quantification CoCo a également été appliquée afin d'obtenir des données fiables. Les résultats obtenus par séquençage ont été validés par qRT-PCR, méthode de référence, dans des tissus indépendants du RNA-Seq, soit cinq HGSC et quatre SBT. Ces méthodes d'analyse ont permis d'identifier 36 snoRNA candidats significativement surexprimés par RNA-Seq dans les HGSC, permettant de les distinguer des SBT. SNORA2 ressort comme un biomarqueur potentiel surexprimé dans les données de séquençage par RNA-Seq et par qRT-PCR. Des validations ultérieures sur un plus grand nombre de tissus permettront de confirmer le pouvoir diagnostique de ces snoRNA candidats, et des analyses phénotypiques pourront nous éclairer sur les fonctions générales et oncogéniques des snoRNA.

Mots clés : snoRNA, cancer de l'ovaire, séquençage à haut débit, RNA-Seq, transcriptomique, biomarqueurs, dépistage

## SUMMARY

### **Characterization of snoRNA expression and biomarker discovery in serous ovarian cancer**

By  
Audrey Roy  
Biochemistry Program

Thesis presented at the Faculty of medicine and health sciences for the obtention of Master degree diploma maitre ès sciences (M.Sc.) in biochemistry, Faculty of medicine and health sciences, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4

Ovarian cancer is the 5th leading cause of mortality by cancer in Canada, making it the deadliest gynecologic cancer. The main difficulty lies in our inability to detect it at an early stage, in major part because of the lack of reliable screening tools. If the mean survival is of 90% at five years for early stage cancer, it is of only 15% for late stage disease. Furthermore, only 25% of ovarian cancers are detected at an early stage. It is thus a priority to find new biomarkers. snoRNAs (small nucleolar RNA), non-coding RNA playing a role in the maturation of ribosomal RNA, appear to be potential actors in tumorigenesis, some of them identified as diagnostic or prognostic biomarkers in specific cancers. Phenotypic and mechanistic studies have also shown the role of some snoRNAs in cancer development and progression. However, they have not been studied in ovarian cancer yet, making them a source of potential biomarkers for the screening of this disease. The goal of this thesis is to characterise the expression of snoRNAs in serous ovarian cancer and to identify potential biomarkers. We used high throughput RNA sequencing of three high grade serous carcinoma (HGSC) tissues and three serous borderline tumor (SBT) tissues with a new sequencing method using a thermostable group II intron reverse transcriptase (TGIRT), making possible the accurate detection of snoRNAs, traditionally badly represented in datasets. The analysis of the sequencing included the quantification tool CoCo, adapted for the specific characteristics of this RNA family. The sequencing results were then validated by qRT-PCR in independent tissues (five HGSC and 4 SBT). These methods have identified 36 candidate snoRNAs significantly overexpressed in HGSC, distinguishing them from SBT. SNORA2 is a potential biomarker for HGSC, since it is overexpressed by RNA-Seq and qRT-PCR. More validations on a greater number of tissues will be conducted and phenotypic studies will be done to make light on the general and oncogenic functions of snoRNAs.

Keywords : snoRNA, ovarian cancer, RNA-Seq, deep sequencing, transcriptomic, biomarkers, screening

## TABLE DES MATIÈRES

<b>Introduction.....</b>	<b>1</b>
<b>Cancer de l’ovaire.....</b>	<b>1</b>
Épidémiologie et présentation clinique.....	1
Histologie de l’ovaire et types de cancers de l’ovaire .....	2
Grades et stades du cancer de l’ovaire.....	4
Génétique du cancer de l’ovaire.....	5
Développement.....	6
Facteurs de risque et facteurs protecteurs .....	9
Marqueurs du cancer de l’ovaire .....	11
<b>Biologie des snoRNA.....</b>	<b>15</b>
Biogénèse des snoRNA .....	15
Fonctions des snoRNA .....	16
<b>snoRNA et cancers.....</b>	<b>17</b>
<b>Découverte de biomarqueurs.....</b>	<b>18</b>
Paramètres requis pour le développement de marqueurs de dépistage.....	18
Étapes du développement de biomarqueurs .....	20
<b>Séquençage à haut débit de l’ARN .....</b>	<b>20</b>
<b>Analyses bioinformatiques des résultats de séquençage .....</b>	<b>21</b>
<b>Hypothèse/problématique.....</b>	<b>23</b>
<b>Objectifs .....</b>	<b>23</b>
<b>Matériel et Méthodes.....</b>	<b>24</b>
Méthodologie de la recherche de biomarqueurs .....	24
Tissus ovariens .....	25
Extraction de l’ARN et préparation des bibliothèques de séquençage.....	26
Séquençage et analyse du séquençage .....	27
Analyse de l’expression différentielle entre les tissus SBT et HGSC.....	29
Corrélation entre le RNA-Seq et le qRT-PCR .....	30
Diagrammes de quartiles ( <i>boxplots</i> ) .....	30
Production des heatmaps.....	30
RT-PCR quantitatif (qRT-PCR) .....	30
Analyses phénotypiques.....	32

<b>Résultats.....</b>	<b>34</b>
<b>Le séquençage de l'ARN avec TGIRT permet la détection des snoRNA dans des tissus ovariens .....</b>	<b>34</b>
<b>La régulation de l'abondance des snoRNA dans les HGSC diffère selon le type C/D ou H/ACA.....</b>	<b>36</b>
<b>Les snoRNA sont enrichis lors de la sélection des candidats significativement surexprimés dans les HGSC .....</b>	<b>39</b>
<b>Les snoRNA candidats significativement surexprimés dans les HGSC sont principalement des snoRNA H/ACA connus pour modifier une position dans l'ARN ribosomal .....</b>	<b>41</b>
<b>Les snoRNA sélectionnés permettent de discriminer entre les HGSC et les SBT séquençés par RNA-Seq.....</b>	<b>43</b>
<b>L'expression des snoRNA candidats par qRT-PCR dans des tissus HGSC et SBT indépendants montrent une hétérogénéité .....</b>	<b>45</b>
<b>SNORA2 permet de discriminer entre les HGSC et les SBT par qRT-PCR.....</b>	<b>46</b>
<b>Discussions et conclusions .....</b>	<b>47</b>
<b>Une nouvelle méthode de séquençage permet de détecter les snoRNA dans des tissus humains .....</b>	<b>47</b>
<b>36 snoRNA surexprimés dans le cancer de l'ovaire découverts par RNA-Seq .....</b>	<b>48</b>
<b>SNORA2, un snoRNA surexprimé dans le cancer de l'ovaire détecté par qRT-PCR .....</b>	<b>49</b>
<b>Rôle des snoRNA dans le cancer .....</b>	<b>50</b>
<b>Prochaines étapes de validation de biomarqueurs dans le cancer de l'ovaire .....</b>	<b>52</b>
<b>Défis de la recherche de biomarqueurs pour le cancer de l'ovaire .....</b>	<b>55</b>
<b>Applications futures des snoRNA biomarqueurs.....</b>	<b>55</b>
<b>Conclusions .....</b>	<b>56</b>
<b>Liste des références .....</b>	<b>57</b>

## LISTE DES TABLEAUX

Tableau 1 : Description des tissus ovariens utilisés .....	26
Tableau 2 : Amorces utilisées pour le qRT-PCR des snoRNA .....	32
Tableau 3 : Amorces utilisées pour le qRT-PCR des gènes constitutifs .....	32
Tableau 4 : Séquences des ASO utilisés pour la déplétion de snoRNA.....	33
Tableau 5 : Nombre de lectures aux différentes étapes du traitement des données de séquençage .....	34
Tableau 6 : Corrélations entre l'abondance détectée par RNA-Seq et par qRT-PCR .....	35
Tableau 7 : Liste des snoRNA candidats et leurs caractéristiques.....	42
Tableau 8 : Expression relative de 11 snoRNA candidats par qRT-PCR dans des tissus indépendants (5 HGSC et 4 SBT) .....	45
Tableau 9 : Description des tissus ovariens disponibles pour la validation et les tests à l'aveugle des biomarqueurs potentiels.....	52

## LISTE DES FIGURES

Figure 1 : Distribution des types histologiques et des grades des cancers malins ovariens ...	4
Figure 2 : Schéma du développement du cancer séreux de l’ovaire .....	9
Figure 3 : Méthodologie de la recherche de biomarqueurs .....	25
Figure 4 : Pipeline d’analyse RNA-Seq .....	27
Figure 5 : Abondance des snoRNA dans les tissus ovariens séquencés par RNA-Seq et qRT-PCR.....	35
Figure 6 : Répartition des lectures des familles d’ARN séquencées par RNA-Seq .....	36
Figure 7 : Distribution des snoRNA C/D et H/ACA dans les tissus séquencés par RNA-Seq .....	37
Figure 8 : Distribution de l’abondance des snoRNA dans les tissus HGSC (A) et SBT (B)	38
Figure 9 : Diagramme de quartiles ( <i>boxplot</i> ) du $\log_2$ fold change (Log2FC) pour tous les gènes séquencés et en fonction de la famille de snoRNA .....	39
Figure 10 : Nombre de candidats sélectionnés aux différentes étapes de sélection.....	40
Figure 11 : Distribution des familles d’ARN lors des différentes étapes de sélection des candidats discriminant les HGSC des SBT .....	41
Figure 12 : <i>Heatmap</i> des 36 snoRNA candidats dans les tissus HGSC et SBT séquencés par RNA-Seq.....	44
Figure 13 : Expression relative moyenne par qRT-PCR de SNORA2 (ENSG00000206647) dans 5 tissus HGSC et 4 tissus SBT .....	46
Figure 14 : Courbe de croissance par mesure de l’impédance par xCELLingence de lignées cellulaires SKOV3.ip1 avec déplétion de SNORA19 .....	51

## LISTE DES ABRÉVIATIONS ET DES SIGLES

ADN : Acide désoxyribonucléique, *Deoxyribonucleic acid*  
ARN : Acide ribonucléique, *Ribonucleic acid*  
ARNm : ARN messenger, *Messenger RNA*  
ARNr : ARN ribosomal, *Ribosomal RNA*  
ARNt : ARN de transfert, *Transfer RNA*  
ARTB : *Alberta Research Tumor Bank*  
ASO : Oligonucléotide antisens, *Antisens oligonucleotide*  
BRCA : *Breast cancer gene*  
ER : Expression relative, *Relative expression*  
FRQS : Fonds de recherche du Québec – Santé  
GPEC : *The Genetic Pathology Evaluation Centre*  
HGSC : Carcinome séreux de haut grade, *High grade serous carcinoma*  
LGSC : Carcinome séreux de bas grade, *Low grade serous carcinoma*  
log<sub>2</sub>FC : Changement logarithmique base 2, *log<sub>2</sub> fold change*  
lncRNA : Long ARN non-codant, *Long non-coding RNA*  
miARN : Micro ARN, *Micro RNA*  
OTB : *Ontario Tumor Bank*  
PCR : Réaction en chaîne par polymérase, *Polymerase chain reaction*  
qRT-PCR : PCR quantitatif par rétrotranscriptase, *quantitative reverse transcriptase PCR*  
RNA-Seq : Séquence de l'ARN, *RNA sequencing*  
SBT : Tumeur séreuse limite, *Serous borderline tumor*  
SD : Déviation standard, *Standard deviation*  
scaRNA : Petit ARN du corps de Cajal, *Small Cajal body*  
snoRNA : Petit ARN nucléolaire, *Small nucleolar RNA*  
snRNA : Petit ARN nucléaire, *Small nuclear RNA*  
TGIRT : Rétrotranscriptase thermostable d'intron de groupe II, *Thermostable Group II intron reverse transcriptase*



## REMERCIEMENTS

Je tiens à remercier ma directrice de recherche, Michelle Scott, pour le soutien apporté tout au long de mon parcours. Les données expérimentales de séquençage et de qRT-PCR ont été rendues possibles grâce au laboratoire de Sherif Abou Elela, en particulier Sonia Couture et Laurence Faucher Giguère, à la plateforme Rnominique de l'Université de Sherbrooke, ainsi qu'au laboratoire d'Alan Lambowitz de la *University of Texas*. Merci également à mon mentor, Dr François Corbin, pour son support et ses conseils. Finalement, merci à tous les membres de mon laboratoire pour leur collaboration, en particulier Gabrielle Deschamps-Francoeur et Vincent Boivin pour leur travail sur l'analyse de séquençage.

## INTRODUCTION

### Cancer de l'ovaire

#### *Épidémiologie et présentation clinique*

Le cancer de l'ovaire est le 8<sup>e</sup> cancer en incidence et le 5<sup>e</sup> en mortalité par cancer en 2017 au Canada, ce qui en fait le cancer gynécologique le plus mortel (Navaneelan et Lawrence, 2015) (Comité consultatif de la Société canadienne du cancer, 2018). Ce cancer touche principalement les femmes en périménopause, avec un âge médian de 55 à 64 ans (Barakat RR, 2009). Une des raisons principales de la mortalité élevée de ce cancer découle de sa détection tardive. En effet, au Canada, environ 65% des cas de cancer de l'ovaire sont détectés alors qu'ils sont à un stade avancé (stade III ou IV), où la maladie s'est déjà répandue ailleurs dans le (Comité consultatif de la Société canadienne du cancer, 2018). Du fait de sa localisation anatomique, les symptômes du cancer de l'ovaire apparaissent tardivement et sont non-spécifiques, les patientes et les médecins les reconnaissant alors la plupart du temps seulement lorsque la tumeur est devenue importante. Ces symptômes incluent un inconfort abdominal, des ballonnements, de la constipation, des manifestations urinaires, de l'ascite, des irrégularités menstruelles, des maux de dos et de la fatigue (Barakat RR, 2009) (Goff *et al*, 2004). Or, ces symptômes sont plus fréquemment dus à d'autres conditions, pour la plupart bénignes ou passagères. Même avec une recherche systématique des symptômes par le médecin traitant, le cancer de l'ovaire n'est pas mieux diagnostiqué (Sarojini *et al*, 2012). Ainsi, le cancer de l'ovaire est souvent découvert fortuitement lors de tests prodigués pour d'autres raisons médicales. Un autre élément défavorable à la détection du cancer de l'ovaire est l'absence de biomarqueur sensible et spécifique qui permettrait de pallier la problématique présentée par cette symptomatologie fruste.

Une autre raison pour la mortalité élevée du cancer de l'ovaire est la difficulté à traiter ce cancer. Les ovaires étant libres dans l'abdomen, les cellules tumorales peuvent établir des métastases sans passer par la voie hématogène, par simple détachement des cellules tumorales. Des implants tumoraux sont donc précocement retrouvés dans le pelvis ainsi que sur les intestins et les autres organes de la cavité abdominale, qui peuvent ensuite envahir ces

tissus (Lengyel, 2010). Or, l'étendue du cancer est un des facteurs pronostiques les plus importants (Barakat RR, 2009). Le traitement déterminant pour le pronostic demeure donc à ce jour la chirurgie de réduction tumorale qui vise à obtenir une maladie résiduelle nulle macroscopiquement et dont le résultat dépend en grande partie de l'habileté du chirurgien et de la charge tumorale initiale, et donc du stade (Winter *et al*, 2007) (Winter *et al*, 2008) (Eisenhauer *et al*, 2006) (Engelen *et al*, 2006) (Bristow *et al*, 2002). Cependant, il reste toujours des foyers cancéreux microscopiques, qu'on tente alors de traiter par chimiothérapie. Or, bien que dans la plupart des cas la réponse initiale à la chimiothérapie soit bonne (environ 85%), jusqu'à 75% des cancers de l'ovaire deviennent chimiorésistants (Colombo *et al*, 2014). Ainsi, les patientes atteintes d'un cancer de l'ovaire de stade avancé possèdent un taux de survie à 5 ans inférieur à 45%; or, si détecté au stade I, il est de 90% (Comité consultatif de la Société canadienne du cancer, 2018). Il devient donc primordial de développer des outils afin d'améliorer la détection du cancer de l'ovaire. Il s'agit d'ailleurs d'une priorité établie par l'organisme Cancer de l'ovaire Canada (Ovarian Cancer Canada, 2013).

### ***Histologie de l'ovaire et types de cancers de l'ovaire***

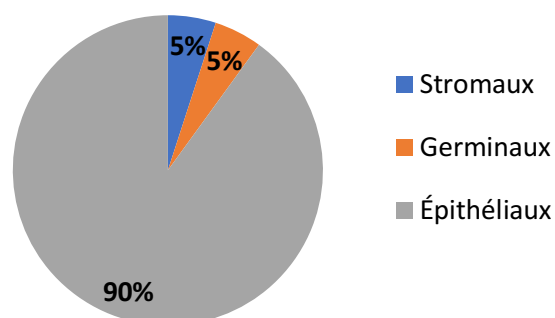
L'ovaire sain présente 3 types cellulaires pouvant chacun être à l'origine d'une néoplasie (Ovalle *et al*, 2013). Les cellules stromales sont les plus abondantes et servent au soutien des autres cellules de l'ovaire. On retrouve deux types de cellules stromales, les cellules de la thèque et de la granulosa, qui sécrètent la progestérone et l'œstrogène (Barakat RR, 2009). Les tumeurs stromales sont très rares et peuvent sécréter des taux anormaux d'hormones entraînant des symptômes détectables comme la masculinisation, des saignements anormaux ou une puberté précoce, ce qui favorise leur détection (Tavassoli, 1994). Ces tumeurs sont généralement de bas grade lorsque détectées et ainsi généralement moins agressives et létales (Malmstrom *et al*, 1994).

Les deuxièmes cellules en abondance sont les cellules germinales, aussi appelés oocytes, destinées à former les ovules. Un sous-groupe de ces cellules subit une maturation de façon mensuelle, un seul oocyte atteignant une pleine maturité dans un but de reproduction (Ovalle *et al*, 2013). Les tumeurs provenant de ce type de cellules, comme les tératomes ou les

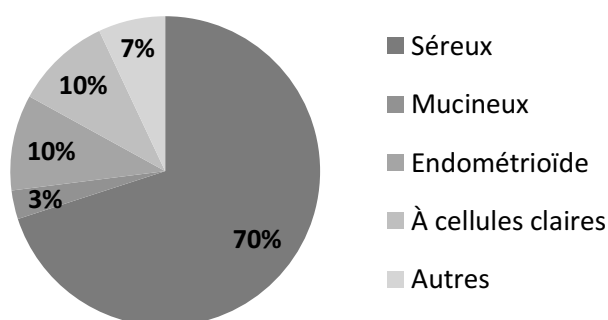
dysgerminomes, affectent principalement les femmes plus jeunes et ont en général un bon pronostic, étant bénignes ou curables dans plus de 90% des cas (Poynter *et al*, 2010).

Le dernier type cellulaire est formé des cellules épithéliales, qui sont peu nombreuses et forment une couche simple à la surface de l'ovaire (Ovalle *et al*, 2013). Les cellules épithéliales forment les tumeurs ovariennes les plus fréquentes et mortelles et affectent surtout les femmes en péri-ménopause (Barakat RR, 2009). Le cancer épithélial se divise par la suite en sous-types histologiques, les principaux étant à cellules claires, mucineux, endométrioïdes et séreux (figure 1) (Köbel *et al*, 2010) (Prat, 2012). Le carcinome séreux étant responsable de la majorité de l'incidence et de la mortalité du cancer de l'ovaire, il est celui qui fait l'objet de mes recherches.

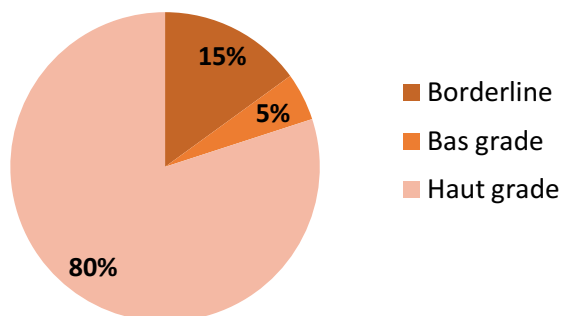
#### A. Tous cancers malins ovariens



#### B. Cancers épithéliaux



### C. Carcinomes séreux



**Figure 1 : Distribution des types histologiques et des grades des cancers malins ovariens**

### *Grades et stades du cancer de l'ovaire*

#### *a) Grades*

Le grade d'un cancer est l'analyse histopathologique des cellules cancéreuses qui décrit le degré de différenciation de celles-ci, basée entre autres sur l'activité mitotique et la présence d'atypies nucléaires. Plus la tumeur acquiert des mutations, moins elle se différencie, et plus elle a tendance à être agressive et à se propager rapidement (National Cancer Institute, 2013). Dans le cas du grade du cancer séreux de l'ovaire, les carcinomes sont divisés en deux groupes, soit de bas grade (LGSC, *low grade serous carcinoma*) et de haut grade (HGSC, *high grade serous carcinoma*). On retrouve également une troisième catégorie, les tumeurs séreuses dites *borderline*, ou limite (SBT, *serous borderline tumor*), quelquefois encore appelées tumeurs à faible potentiel de malignité (Kurman *et al*, 2014). Ces cellules cancéreuses présentent des atypies nucléaires légères à modérées, mais aucune invasion et une activité proliférative faible. De plus, ces tumeurs progressent différemment des HGSC, car les SBT n'évoluent que rarement en carcinome et ne sont pas associées à une hausse de la mortalité (Seidman et Kurman, 2000) (Longacre *et al*, 2005) (Sherman *et al*, 2004). Elles ne sont donc pas considérées comme faisant typiquement partie du continuum de progression tumorale de la lésion bénigne en lésion maligne. De plus, lorsqu'elles récidivent, elles apparaissent souvent plusieurs années plus tard et sont rarement invasives (Morice *et al*,

2012). Les SBT présentent un profil génomique différent des HGSC, et se rapprochent donc plus du profil des cellules épithéliales saines (Shih et Kurman, 2004), ce pourquoi elles sont fréquemment utilisées comme tissu de comparaison pour la recherche de biomarqueurs du HGSC. La répartition entre ces trois types de tumeurs séreuses est présentée à la figure 1C.

### *b) Stades*

Le stade d'un cancer décrit son étendue dans le corps. Dans le cas du cancer de l'ovaire, on retrouve 4 stades définis par la Fédération Internationale de Gynécologie et d'Obstétrique (FIGO) (Prat, 2014). Le stade I se définit par une tumeur confinée à un ou aux deux ovaires seulement. Lorsque la tumeur s'étend au pelvis (par exemple à l'utérus), on parle d'un stade II. Lorsqu'on retrouve des implants tumoraux ailleurs dans le péritoine, le cancer est de stade III. Finalement, lorsque le cancer s'est propagé à l'extérieur du péritoine (par exemple aux poumons ou aux ganglions) ou a infiltré un viscère comme le foie ou la rate, le cancer a atteint le stade le plus avancé, soit le stade IV. Un stade avancé est décrit comme un stade III ou IV.

### ***Génétique du cancer de l'ovaire***

Le carcinome séreux de l'ovaire est une maladie hétérogène au niveau génétique, ce qui se traduit par des réponses aux traitements, des évolutions et des pronostics très variés (Lalwani *et al*, 2011). Les carcinomes de haut et de bas grade possèdent des profils génétiques distincts, la plupart des mutations se retrouvant exclusivement dans un de ces deux types de cancer. Le HGSC présente invariablement une instabilité génétique résultant en une altération du nombre de copies d'ADN, contrairement au LGSC (Kuo *et al*, 2009). Cette instabilité est causée en grande partie par l'acquisition de mutations dans le gène p53, un suppresseur de tumeur. Une mutation de p53 est présente dans plus de 96% des cas. Les autres gènes fréquemment mutés sont ceux impliqués dans la réparation homologue de l'ADN, dont les mutations de BRCA 1 et 2, héréditaires ou somatiques, présentes chez près du tiers des patientes (Cancer Genome Atlas Research Network, 2011). Cette instabilité génomique favorise l'hétérogénéité génétique de la tumeur, dans laquelle on retrouve donc une population de sous-clones qui répondent différemment à la chimiothérapie, ce qui contribue

au développement de résistances et à l'échec thérapeutique presque invariablement observé après la réponse initiale à la chimiothérapie (Salomon-Perzyński *et al*, 2017).

Les mutations les plus souvent retrouvées dans le cancer de bas grade sont celles de KRAS (19%), BRAF (5-38%) et ERBB2 (6%). Ces mutations sont mutuellement exclusives (Hunter *et al*, 2015). Les tumeurs SBT, quant à elles, présentent plus souvent des mutations de BRAF (28-48%). Ces trois gènes sont liés à la voie des MAP kinases, leurs mutations causant une activation constitutive de cette voie mitogène. Les cancers de bas grade présentent une meilleure stabilité génomique et sont diploïdes. Une mutation de p53 est absente dans la très grande majorité des cancers « *borderline* » et de bas grade, contrairement à ceux de haut grade (Cancer Genome Atlas Research Network, 2011).

### ***Développement***

Le modèle de développement du cancer de l'ovaire proposé actuellement est dualistique. En effet, il a été observé que les tumeurs séreuses de l'ovaire se divisaient en deux types d'après leur évolution et leurs mutations (Shih et Kurman, 2004). Le type I se distingue par une détection à un bas grade et une évolution lente, tandis que le type II est détecté à un grade déjà avancé et est très agressif. Les deux types de cancer sont également différents quant à leurs mutations : les cancers de type II ont typiquement une mutation de p53, tandis qu'elle est absente dans les cancers de type I, qui présentent plutôt des mutations de BRAF ou KRAS dans la majorité des cas (Singer *et al*, 2005) (Mayr *et al*, 2006). Il est à noter que ces deux types sont deux maladies différentes : le type I, de bas grade, est rarement le précurseur du type II, de haut grade (Dehari *et al*, 2007). De plus, les carcinomes de bas grade demeurent en général indolents lors de récurrence (Gershenson *et al*, 2006) (Plaxe, 2008).

Le premier modèle de développement du cancer de l'ovaire, qui décrit bien les cancers séreux de type I, soit de bas grade, propose que la tumeur se développe à partir de cellules saines de l'ovaire, qui acquièrent graduellement des mutations permettant leur transformation en tumeur bénigne, puis maligne. Ce modèle propose donc un continuum de la cellule saine jusqu'au cancer, du cystadénome bénin à la tumeur *borderline* jusqu'au carcinome de bas grade (Vang *et al*, 2009). Étant donné le lien épidémiologique avec le nombre de cycles

ovulatoires, il a été proposé que l'environnement de l'ovaire, avec les hormones, les facteurs de croissance et l'inflammation reliés aux ovulations, est propice pour créer des dommages à l'ADN (Fathalla, 1971). Ce modèle est également appuyé par l'observation de lésions intermédiaires dans ce continuum ainsi que la similitude histologique entre les précurseurs bénins et le cancer de bas grade. De plus, des régions de tumeurs « borderline » sont fréquemment retrouvées en contiguïté avec des carcinomes de bas grade, ce qui appuie l'hypothèse de progression tumorale (Vang *et al*, 2009).

Or, dans le cas des cancers séreux de type II, ou de haut grade, aucune lésion précurseur ne pouvait être trouvée, créant ainsi le besoin de modifier ce modèle (Barakat *et al*, 2000). C'est en examinant les trompes de Fallope de patientes portant la mutation BRAC1/2 et ayant subi une salpingo-ovariectomie prophylactique que les chercheurs ont découvert des lésions précancéreuses dans les fimbriae des trompes de Fallope, appelées carcinome séreux intraépithélial tubaire, alors qu'il n'y avait pas encore de cancer ovarien (Piek *et al*, 2001) (Carlson *et al*, 2008). Subséquemment, en étudiant les trompes de patientes atteintes de cancer de l'ovaire, il a été découvert que ces lésions tubaires présentent le même profil de mutations de p53 que la tumeur ovarienne, appuyant son origine tubaire (Kurman et Shih, 2010) (Kurman et Shih, 2011) (Kuhn *et al*, 2012). Des mutations de p53 ont également été notées dans les cellules épithéliales des trompes de Fallope de morphologie normale, désignées « signature p53 », suggérant l'acquisition de mutations dans les cellules des trompes de Fallope précédant la transformation tumorale (Folkins *et al*, 2008). En plus de mutations de p53, le profil génétique des HGSC est similaire à celui des trompes de Fallope et diffère de l'épithélium ovarien (Marquez *et al*, 2005). Le deuxième modèle élaboré propose donc que le cancer ovarien séreux de haut grade provient plutôt de lésions précancéreuses des trompes de Fallope qui, par contiguïté, s'implantent dans l'ovaire, où le milieu inflammatoire riche en cytokines et hormones est propice au développement du cancer.

Or, pour 40 à 50% des HGSC sporadiques, aucune lésion tubulaire ne peut être trouvée (Chen *et al*, 2017). Certains chercheurs soutiennent donc l'idée que le HGSC pourrait également provenir de kystes d'inclusion ou encore de transformation néoplasique des cellules épithéliales ovariennes elles-mêmes, favorisée par l'environnement inflammatoire, riche en

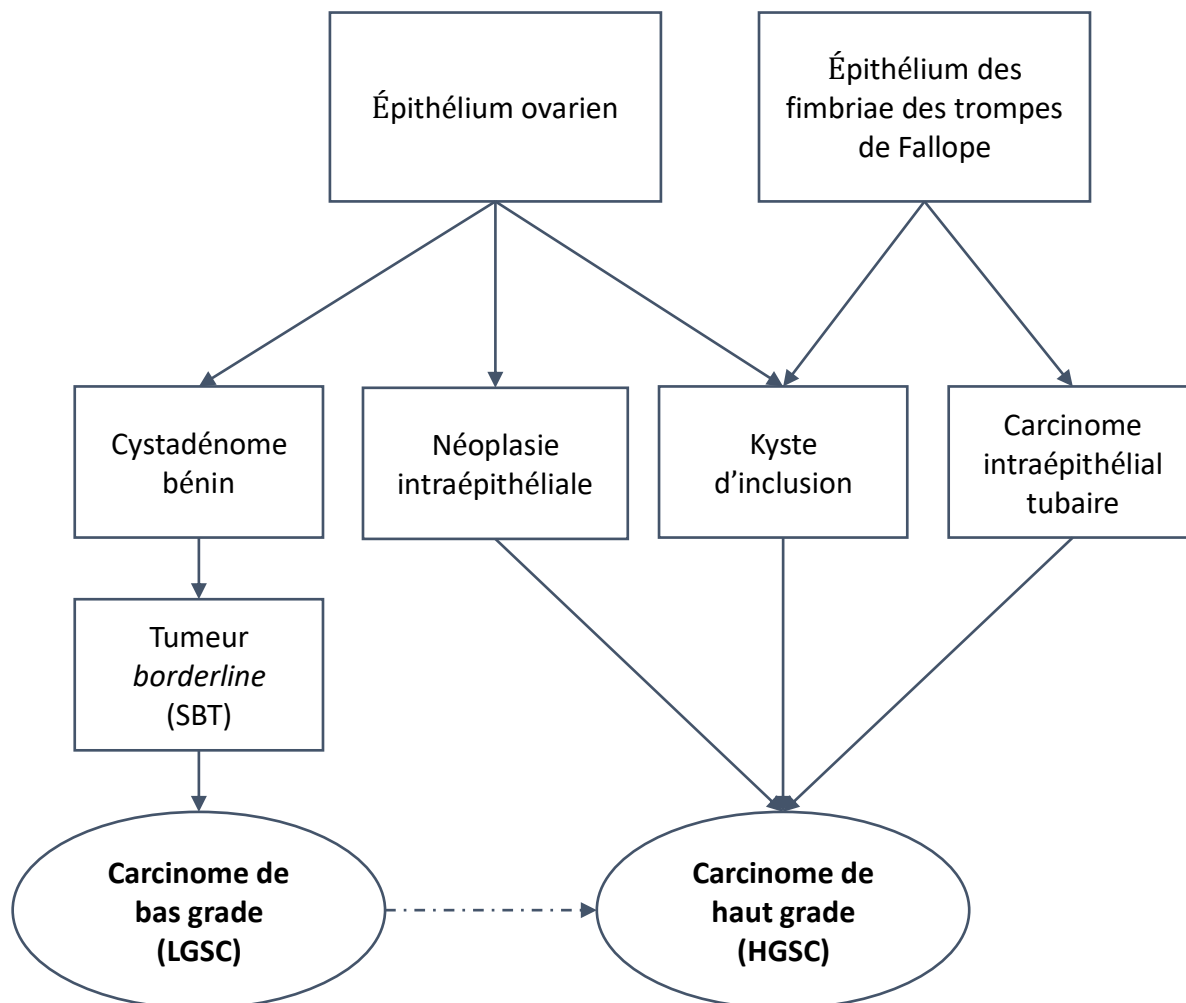


cytokines et hormones dû aux cycles ovulatoires mensuels (Auersperg, 2013). Lors de l'ovulation, marquée par la rupture du follicule et la réparation qui s'ensuit, il peut se créer des kystes d'inclusion, où les cellules épithéliales normalement retrouvées à la surface de l'ovaire se trouvent alors dans un espace fermé à l'intérieur du stroma. Des hypothèses contradictoires quant aux cellules incluses dans ces kystes persistent encore aujourd'hui : certains proposent que ce sont des cellules épithéliales ovariennes qui s'invaginent, tandis que d'autres soutiennent que ce sont des cellules tubaires saines qui s'implantent dans l'ovaire (Blaustein et Kurman, 2011). Cette dernière hypothèse est également appuyée par le fait que les cellules formant les kystes sont d'origine mésothéliales, comme les cellules tubaires. Les cellules épithéliales sont plutôt d'origine müllérienne, et doivent donc subir une métaplasie pour expliquer cette observation, ce qui peut effectivement se produire lorsque les cellules ne sont pas dans leur environnement habituel. Une autre hypothèse concernant l'origine purement ovarienne du cancer de l'ovaire se base sur le potentiel néoplasique de cellules souches de l'épithélium ovarien (Flesken-Nikitin *et al*, 2013). Finalement, dans de rares cas, des régions « *borderline* » ou de bas grade ont été décrites dans des tumeurs de haut grade, proposant une évolution de bas à haut grade pour certains cas, plus rares (Kurman et Shih, 2010). La figure 2 résume les différentes voies par lesquelles le cancer de l'ovaire peut se développer.

Ces différentes origines et leur évolution distincte souligne l'importance de bien choisir les tissus de référence dans la recherche de biomarqueur. En effet, il faut pouvoir distinguer les cancers de haut grade (HGSC) de ceux *borderline* (SBT) ou de bas grade (LGSC) afin d'éviter de traiter trop agressivement des cancers qui n'en ont pas besoin (Jacobs et Menon, 2004). De plus, puisque les cellules épithéliales ovariennes saines sont en très faible minorité dans un ovaire normal, il n'est pas possible de les utiliser comme tissu de comparaison. Les chercheurs utilisent donc généralement des tumeurs *borderline*, qui ont un profil similaire aux cellules épithéliales saines, ainsi que l'épithélium des trompes de Fallope étant donné son rôle précurseur dans certains cancers ovariens de haut grade.

Une meilleure compréhension de l'évolution du cancer de l'ovaire de haut grade permettrait d'identifier une lésion précancéreuse pouvant servir pour la prévention secondaire, avant l'apparition de symptômes. En effet, en sachant exactement d'où provient le cancer de

l'ovaire, la recherche de biomarqueurs de dépistage pourrait alors être ciblée à cette lésion précurseur, augmentant ainsi les chances de découvrir des marqueurs ayant un potentiel de réduire la mortalité liée au cancer de l'ovaire.



**Figure 2 : Schéma du développement du cancer sévère de l'ovaire**

### ***Facteurs de risque et facteurs protecteurs***

#### ***a) Facteurs hormonaux***

Plusieurs facteurs de risque et facteurs protecteurs du cancer de l'ovaire sont liés au cycle ovulatoire, appuyant ainsi les modèles de son développement. En effet, le nombre de grossesses, l'allaitement et la prise de contraceptifs oraux sont des facteurs réduisant le nombre total d'ovulations auxquelles la femme est exposée au cours de sa vie qui sont

associés à une baisse du risque de développer un cancer de l'ovaire (Wentzensen *et al*, 2016). L'effet protecteur de la contraception hormonale se poursuit d'ailleurs plusieurs années après son arrêt et est relié à sa durée d'utilisation (Iversen *et al*, 2017). À l'inverse, l'hormonothérapie de remplacement à la ménopause est associée à une hausse du risque de cancer de l'ovaire (Beral *et al*, 2015).

### *b) Facteurs génétiques*

Seulement près de 5% des cancers de l'ovaire sont d'origine héréditaire, principalement causés par des mutations des gènes BRCA 1 et 2 ainsi que le syndrome de Lynch, tous deux transmis de façon autosomale dominante (Barakat RR, 2009). Des mutations des gènes BRCA1 et BRCA2, des suppresseurs de tumeurs intervenant dans la réparation de l'ADN également associées au cancer du sein, augmentent le risque de développer un cancer de l'ovaire par 39-46% et 12-20%, respectivement. Ces mutations sont particulièrement retrouvées chez les Juifs Ashkenazi ainsi que les Canadiens français.

Le syndrome de Lynch, causé principalement par une mutation dans un des 4 gènes impliqués dans la réparation mésappariée de l'ADN (MLH1, MSH2, MSH6 ou PMS2), augmente surtout le risque de cancer colorectal héréditaire sans polypose (40-80%). Or, il augmente également le risque de développer un cancer de l'endomètre (20-50%) et dans une moindre mesure celui de l'ovaire (5-10%), plus particulièrement le sous-type endométrioïde (Helder-Woolderink *et al*, 2016).

Lorsqu'une histoire familiale de cancer du sein ou de l'ovaire est révélée, des tests de dépistage génétique sont offerts aux familles. Des mesures préventives et prophylactiques, comme la salpingo-ovariectomie, sont proposées aux patientes possédant ces mutations. Des antécédents familiaux au premier degré de cancer du sein ou de l'ovaire, même sans mutations génétiques connues détectées, sont également associés à un plus grand risque de développer un cancer de l'ovaire (Andrews et Mutch, 2017).

### *c) Autres facteurs*

Le tabac et l'indice de masse corporelle ne sont pas associés au cancer de l'ovaire. Les études sur l'utilisation de poudre de talc restent contradictoires au sujet de son lien avec le cancer de l'ovaire (Wentzensen *et al*, 2016). La salpingectomie seule, avec conservation des ovaires, est associée à une diminution du risque de cancer de l'ovaire, en particulier chez les femmes

porteuses de mutations BRAC1/2, appuyant l'hypothèse des trompes de Fallope comme origine du cancer de l'ovaire (Chohan et Richardson, 2019) (Ely et Truong, 2017). Un âge avancé est également relié au cancer de l'ovaire et présente une distribution différente selon le type : les LGSC se présentent à un âge moyen de 45 ans, tandis que les HGSC apparaissent dans la soixantaine (Barakat RR, 2009).

### ***Marqueurs du cancer de l'ovaire***

#### ***a) CA-125***

Le premier marqueur utilisé pour le cancer de l'ovaire est le CA-125 (*cancer-antigen 125*), une protéine sérique provenant du gène MUC16 (Yin et Lloyd, 2001) contenant deux domaines antigéniques contre lesquels des anticorps monoclonaux sont dirigés afin de la détecter (Bast *et al*, 1981) (Nustad *et al*, 1996). Il est encore aujourd'hui le seul biomarqueur utilisé en clinique. Cette protéine est exprimée normalement durant le développement fœtal, puis par plusieurs tissus chez l'adulte incluant la plèvre, le péritoine et l'épithélium endométrial. Par contre, elle n'est pas exprimée par l'ovaire normal et est surexprimée dans le cancer de l'ovaire, mais aussi lors de métaplasie ou de kystes ovariens (Kabawat *et al*, 1983). Ainsi, ce marqueur manque de spécificité, car il est influencé par plusieurs conditions bénignes ou physiologiques fréquentes comme l'endométriose et le cycle menstruel, de même que par le statut ménopausique et la présence ou non d'une hystérectomie. Il manque également de spécificité pour distinguer les tumeurs bénignes de malignes (Alagoz *et al*, 1994) (Sjovall *et al*, 2002) (Miralles *et al*, 2003). De plus, ce marqueur est peu sensible pour les cancers de l'ovaire de stade précoce. En effet, bien qu'il soit élevé chez plus de 90% des cancers ovariens de stade avancé, moins de 50% des cancers de stade I présentent des taux de CA-125 élevés (Mann *et al*, 1988) (Sölétormos *et al*, 2016). De plus, 15% à 20% des cancers de l'ovaire n'expriment pas cet antigène (Cramer *et al*, 2010). Ainsi, il n'est pas recommandé pour le dépistage dans la population asymptomatique puisque l'objectif est de détecter le cancer au stade précoce afin d'améliorer les possibilités de cure et donc réduire la mortalité (U.S. Preventive Services Task Force, 2013). Son usage se limite donc pour l'instant à l'aide au diagnostic d'une masse pelvienne (Tingulstad *et al*, 1996) ainsi qu'au suivi de la réponse aux traitements : par des prises de sang sériées, il est possible de

déterminer si la tumeur régresse ou récidive (Hempling *et al*, 1993) (Rustin *et al*, 1996). De plus, puisqu'elle reflète le volume de la maladie, sa valeur au diagnostic est également corrélée au pronostic (Steffensen *et al*, 2011).

### *b) Échographie transvaginale*

L'échographie transvaginale est un outil utilisé en clinique pour visualiser les masses pelviennes lorsqu'il y a une suspicion de tumeur selon les symptômes rapportés par la patiente ou lors d'anomalies détectées à l'examen physique par le médecin. Cette méthode est capable de déterminer si la masse est d'allure bénigne ou si celle-ci mérite des investigations plus poussées selon des critères échographiques précis (Timmerman *et al*, 2016). Par contre, la bonne sensibilité de ce test, qui permet de s'assurer qu'une masse désignée comme bénigne ait peu de probabilités d'être maligne, se fait au détriment de la spécificité, ce qui entraîne un haut taux de faux positifs qui doivent subir des investigations plus poussées ou une chirurgie, causant par le fait même un stress psychologique et des risques opératoires inutiles chez ces patientes en réalité en bonne santé (Henderson *et al*, 2018). De plus, c'est une procédure coûteuse et invasive pour la patiente comparativement à une prise de sang. L'échographie transvaginale n'est donc pas applicable comme programme de dépistage à grande échelle (U.S. Preventive Services Task Force, 2013) (Buhling *et al*, 2017).

### *c) Combinaison de marqueurs*

Afin d'augmenter la sensibilité et la spécificité du CA-125, des études ont tenté de combiner ce marqueur avec l'échographie transvaginale. Une première étude sur plus de 78 000 femmes de 55 à 74 ans évaluant l'impact sur la mortalité du dépistage simultané par CA-125 pendant 6 ans et échographie pendant 4 ans n'a pas pu montrer une réduction du risque (Buys *et al*, 2011). Une étude clinique importante plus récente, le *UK Collaborative Trial of Ovarian Cancer Screening* (UKCTOCS), regroupant plus de 200 000 femmes de 50 à 74 ans post-ménopausées de 2001 à 2014 a évalué le potentiel de réduction de la mortalité du dépistage soit par l'échographie seule ou par une approche multimodale de mesures sériées du CA-125 suivie par l'échographie selon le risque de cancer établi par le résultat du CA-125 (Jacobs *et al*, 2016). Aucun impact sur la survie des patientes n'a pu être révélé par cette étude clinique randomisée d'envergure, mais un suivi à plus long terme pourrait faire ressortir

un bénéfice du dépistage. De plus, cette étude a montré une capacité à détecter les cancers de l'ovaire à un stade plus précoce, montrant ainsi un potentiel de prévenir près de 20% des décès par cancer de l'ovaire (Menon *et al*, 2009). Par contre, une revue systématique de ces études n'a pas pu montrer de bénéfices sur la survie, et le dépistage entraîne des événements indésirables reliés aux investigations et aux chirurgies engendrées (Henderson *et al*, 2018).

#### *d) Recherche de nouveaux marqueurs*

De nombreuses études se sont penchées ces dernières années à découvrir de nouveaux biomarqueurs pour détecter et diagnostiquer le cancer de l'ovaire. Or, aucun de ceux proposés jusqu'à maintenant ne présente une meilleure sensibilité et spécificité que le CA-125, même lorsque combinés (Moore *et al*, 2012) (Bastani *et al*, 2017). Le consensus est donc qu'un meilleur dépistage du cancer de l'ovaire pourra être atteint avec des nouveaux biomarqueurs combinés avec le CA-125. Un seul marqueur protéique a réellement retenu l'attention pour l'instant, HE4 (Granato *et al*, 2015). La recherche de biomarqueurs protéiques ayant montré ses limites, d'autres types de molécules suscitent également l'intérêt, soient les miRNAs et l'ADN tumoral.

- Protéines : HE4

*Human epididymis protein 4* (HE4) est une glycoprotéine normalement exprimée par l'épithélium de l'épididyme qui est augmentée dans le cancer de l'ovaire. Bien qu'elle soit moins sensible que le CA-125 pour détecter les cancers de stade précoce, son intérêt réside dans sa meilleure sensibilité et spécificité à distinguer les tumeurs bénignes et malignes (Molina *et al*, 2011) (Kristjansdottir *et al*, 2013) (Granato *et al*, 2015).

- miARN

Les microARN (miARN) sont des petits ARN non-codants généralement de 19 à 25 nucléotides qui jouent principalement un rôle dans le silençage de l'expression génique (Bartel, 2004). Plusieurs études ont montré qu'ils étaient dérégulés dans plusieurs cancers et qu'ils pouvaient ainsi modifier la régulation de gènes reliés au cancer, soulignant leur rôle potentiel dans le diagnostic et le traitement des cancers (Lu *et al*, 2005) (Zhang *et al*, 2007) (Farazi *et al*, 2013) (Di Leva *et al*, 2014). Un aspect important de leur potentiel diagnostique et thérapeutique réside dans le fait que les miARN se retrouvent dans le sang à l'intérieur

d'exosomes en compagnie de protéines qui les protègent de la dégradation (Mitchell *et al*, 2008) (Cortez *et al*, 2011). Ainsi, les miARN provenant de la tumeur sont détectables avec une prise de sang (Ortiz-Quintero, 2016). Ces exosomes auraient par ailleurs un rôle dans l'évolution du cancer grâce à la communication cellulaire (Valadi *et al*, 2007) (Whiteside, 2016).

Dans le cas du cancer de l'ovaire, des études ont trouvé des signatures d'expression de miARN permettant de distinguer le HGSC (Nam *et al*, 2008) (Yokoi *et al*, 2018). Cependant, les résultats sont contradictoires entre différentes études et manquent de reproductibilité et de validation (Nakamura *et al*, 2016) (Prahm *et al*, 2016). Ainsi, aucun biomarqueur de miARN n'a été retenu à ce jour, mais des validations plus poussées pourraient faire ressortir des miARN applicables à la clinique.

- ADN et cellules tumorales circulants

L'ADN tumoral présente une piste intéressante de marqueur. En effet, celui-ci est fréquemment hyperméthylé dans les néoplasies et circule dans le sang. Les profils de méthylation de l'ADN montrent une corrélation avec l'activité de la maladie, la réponse au traitement et la survie (Teschendorff *et al*, 2009) (Gloss et Samimi, 2014). Par contre, il existe des difficultés techniques liées à la très faible concentration de l'ADN tumoral ainsi qu'à la fragmentation de cet ADN, rendant les résultats encore difficilement applicables à la clinique (Warton et Samimi, 2015). De plus, étant donné que la quantité d'ADN tumoral relâchée dépend de la taille de la tumeur, il reste à déterminer si l'ADN des tumeurs au stade précoce, donc de petite taille, est suffisant pour permettre la détection précoce.

Il est également possible de détecter des cellules tumorales dans le sang (Zhang *et al*, 2018). Par contre, comme pour l'usage actuel du CA-125 en clinique, l'utilité de la détection des cellules tumorales demeurent surtout pour le suivi de la maladie. Elle semble toutefois présenter l'avantage de détecter davantage de cancers de stade I et II par rapport au CA-125.

- snoRNA

Les petits ARN nucléolaires (snoRNA, *small nucleolar RNAs*) sont une famille d'ARN non-codants qui, bien qu'ils soient reliés à plusieurs cancers (Thorenoor et Slaby, 2015), n'ont

pas encore été étudiés dans le cancer de l’ovaire. Les prochains paragraphes en discutent davantage.

### **Biologie des snoRNA**

Les snoRNA sont des petits ARN non-codants de 60 à 300 nucléotides de long retrouvés dans le nucléole. Leur abondance se situe de  $2 \text{ à } 3 \times 10^5$  molécules par cellule, ce qui est dans le même ordre de grandeur que les miARN ( $1\text{--}3 \times 10^5$ ) et les ARN messagers (ARNm) ( $3\text{--}10 \times 10^5$ ) (Palazzo et Lee, 2015). Les snoRNA servent classiquement à la maturation de l’ARN ribosomal (ARNr) (Reichow *et al*, 2007). Un sous-groupe de snoRNA a comme cibles les petits ARN nucléaires (snRNA, *small nuclear RNA*) (Tycowski *et al*, 1998) et un autre sous-groupe sert de guide pour la maturation moléculaire du pré-ARNr en facilitant son clivage et son repliement (Henras *et al*, 2015). Les snoRNA se divisent en deux familles selon des éléments conservés, les boîtes C/D (C : RUGAUGA, D : CUGA) et les boîtes H/ACA (H : ANANNA) (Balakin *et al*, 1996). Les snoRNA C/D, mesurant de 50 à 100 nucléotides, ont leurs extrémités 3’ et 5’ appariées ce qui permet de former un *kink turn* (Kiss-László *et al*, 1998), tandis que les snoRNA H/ACA, de 120 à 140 nucléotides, sont formés de deux tiges-boucles reliées par la boîte H (Henras *et al*, 2004) (Reichow *et al*, 2007).

### **Biogénèse des snoRNA**

Les snoRNA sont majoritairement retrouvés dans les introns de gènes codants, une grande proportion de ces gènes encodant des protéines ribosomales ou reliées aux ribosomes ou à la synthèse protéique (Filipowicz et Pogacić, 2002) (Boivin *et al*, 2018). Certains snoRNA sont également encodés dans des gènes non-codants comme les lncRNA (*long non-coding RNA*) (Maxwell et Fournier, 1995) (Dieci *et al*, 2009). Dans ces deux cas, ils sont produits subséquemment à la transcription de leur gène hôte par l’ARN polymérase II. À partir de l’épissage des introns, le snoRNA est libéré par le débranchement du lasso d’épissage suivi de la dégradation exonucléolytique des extrémités (Kiss et Filipowicz, 1995). Un autre mécanisme, présent chez l’humain seulement pour SNORD16, est par clivage endonucléolytique, également suivi de la dégradation exonucléolytique des extrémités (de Turrís *et al*, 2004). SNORD3, SNORD118 et SNORD13, tous trois impliqués dans le



repliement et le clivage du préARNr, sont quant à eux transcrits de façon indépendante par l'ARN polymérase II (Filipowicz et Pogacić, 2002). Ainsi, l'expression des snoRNA est principalement dépendante de la transcription de leur gène hôte. Toutefois, leur accumulation n'est pas entièrement reliée à l'abondance de leur gène hôte, suggérant une régulation post-transcriptionnelle qui leur est propre (Boivin *et al*, 2018).

Afin d'accomplir leurs fonctions, les snoRNA sont associées avec des protéines qui permettront leur protection contre la dégradation ainsi que l'activité enzymatique nécessaire pour les modifications chimiques des ARN cibles (Bachellerie *et al*, 2002). Les snoRNA C/D sont associés aux protéines fibrillarine (une méthyltransférase), Nop56p, Nop58p et 15.5 kDa (Lafontaine et Tollervey, 1999) (Filipowicz et Pogacić, 2002) (Kiss *et al*, 2006). Les snoRNA H/ACA sont quant à eux associés aux protéines dyskérine (une pseudouridine synthase), Nhp2, Nhp10 et Gar1 (Bousquet-Antonelli *et al*, 1997) (Henras *et al*, 1998) (Watkins *et al*, 1998).

### ***Fonctions des snoRNA***

Les snoRNA C/D guident la méthylation de leur ARN cible tandis que les snoRNA H/ACA guident la pseudouridilation. Leur rôle de guide se fait par complémentarité de séquence avec leur ARN cible. Dans le cas des snoRNA C/D, la région guide se situe en amont de la boîte D, et pour les H/ACA elle se situe dans les boucles (Kiss, 2001) (Kiss, 2002). De nombreuses positions sont modifiées dans l'ARNr (Maden, 1990) et celles-ci sont majoritairement situées aux domaines fonctionnels importants et conservés pour la traduction (Brimacombe *et al*, 1993) (Lapeyre, 2004). Ces modifications permettent la fonction optimale du ribosome ainsi que la régulation de la fonction du ribosome par rapport à sa fidélité et sa processivité, entre autres (Baxter-Roshek *et al*, 2007). En effet, la pseudouridilation permet d'ajouter un donneur de pont hydrogène, augmentant entre autres la rigidité de l'ARNr (Kierzek *et al*, 2014), tandis que la méthylation permet entre autres de protéger l'ARNr de l'hydrolyse (Williams *et al*, 2001). Bien que la déplétion globale de la méthylation et de la pseudouridilation dans une cellule affecte négativement la fonction des ribosomes, peu est encore connu sur le rôle individuel des positions modifiées, étant donné que la déplétion individuelle de snoRNA ne montre pas d'effet sur les ribosomes (Decatur et Fournier, 2002).

Il a par contre été montré qu'environ le tiers des positions sur l'ARNr sont différenciellement méthylées, proposant les modifications chimiques comme mécanisme de réglage de précision de la fonction du ribosome (Krogh *et al*, 2016). Cependant, le lien entre l'abondance d'un snoRNA et le niveau de modification de sa position cible n'a pas été établi. Sachant que la fonction des ribosomes est altérée dans certains cancers en faveur de la traduction de protéines pro-tumorales, il a été proposé que la modulation des modifications chimiques de l'ARNr pourrait jouer un rôle sur la tumorigenèse (Ruggero et Pandolfi, 2003).

Bien que la fonction canonique des snoRNA, c'est-à-dire la maturation de l'ARNr et des snRNA, soit bien connue, il existe des snoRNA orphelins pour lesquels on ne trouve pas de cible complémentaire (Dupuis-Sandoval *et al*, 2015). De plus, de nouvelles fonctions diverses en lien avec l'expression génique s'ajoutent aux fonctions de snoRNA canoniques ou orphelins, telles que la régulation de l'épissage alternatif, le remodelage de la chromatine, des fonctions « miRNA-like », la modulation du stress oxydatif et la modification d'ARN autre que ribosomal (messenger, de transfert, snoRNA) (Makarova *et al*, 2013) (Falaleeva *et al*, 2016). Ceci étend donc leurs fonctions physiologiques et peut expliquer leurs rôles dans le cas de certaines maladies (Stepanov *et al*, 2015) (Mleccko et Bakowska-Zywicka, 2016).

### **snoRNA et cancers**

Les snoRNA ont été étudiés dans plusieurs cancers, où leur dérégulation et leurs rôles dans le développement ou la progression du cancer ont été montrés (Deogharia et Majumder, 2018). De nombreuses études mentionnent les gènes hôtes des snoRNA comme étant dérégulés dans une multitude de cancers. Par exemple, une revue de la littérature a trouvé que SNHG1, qui contient 8 snoRNA boîte C/D, est dérégulé dans de nombreux cancers et pourrait être un promoteur de la prolifération cellulaire (Thin *et al*, 2019). SNHG20 aurait également un rôle dans la carcinogénèse de plusieurs cancers (Zhao *et al*, 2019). Or, on sait maintenant que le niveau d'expression des snoRNA ne corrèle pas nécessairement avec celui de son gène hôte (Boivin *et al*, 2018). Ainsi, les études sur les gènes hôtes de snoRNA ne permettent pas d'établir de lien direct entre leurs snoRNA et la carcinogénèse.

Plusieurs des snoRNA trouvés comme surexprimés ou sous-exprimés sont retrouvés dans des régions chromosomiques connues pour être fréquemment amplifiées ou délétées, appuyant ainsi leur valeur potentielle dans la tumorigenèse (Tanaka *et al*, 2000) (Goeze *et al*, 2002) (Dong *et al*, 2008). Récemment, plusieurs expériences *in vitro* et *in vivo* de déplétion et de surexpression de snoRNA viennent également souligner leur implication comme oncogènes ou suppresseurs de tumeurs. Par exemple, SNORD 126 a été montré comme étant surexprimé dans des échantillons cliniques de cancers hépatocellulaires et colorectaux, et des analyses *in vitro* et *in vivo* ont montré sa capacité oncogénique en lien avec la voie de signalisation PI3K-AKT (Fang *et al*, 2016). SNORD105b, quant à lui, a été relié au cancer de l'estomac comme oncogène et agirait dans la voie de c-Myc (Zhang *et al*, 2018). SNORD47 est dans son cas identifié comme suppresseur de tumeur dans le glioblastome, et dont l'expression est reliée au pronostic des patients (Xu *et al*, 2017). SNORD50A et SNORD50B sont également dérégulés dans plusieurs cancers et agirait par la voie K-Ras, connue pour être impliquée dans plusieurs cancers (Siprashvili *et al*, 2016).

Du côté des snoRNA HACA, SNORA55 a été identifié comme oncogène dans le cancer de la prostate, en lien avec la progression métastatique (Crea *et al*, 2016). Une étude sur le cancer du poumon a permis d'identifier, par séquençage de l'ARN, 3 snoRNA HACA (SNORA47, SNORA68 et SNORA78) permettant de prédire le pronostic (Gao *et al*, 2015).

Les snoRNA ont également été étudiés dans des échantillons humains plus accessibles dans le contexte clinique que les biopsies de tissus dans une visée de test diagnostique. Par exemple, l'analyse d'expectorations a permis d'identifier SNORD66 et SNORD78 comme biomarqueurs diagnostiques pour le cancer du poumon (Su *et al*, 2016).

Cependant, aucune donnée sur l'accumulation des snoRNA n'est encore disponible dans le cas du cancer de l'ovaire.

## **Découverte de biomarqueurs**

### ***Paramètres requis pour le développement de marqueurs de dépistage***

D'après le « Biomarkers Definitions Working Group », un biomarqueur est une caractéristique qui peut être objectivement mesurée et évaluée comme indicateur d'un processus biologique normal ou pathogénique ou d'une réponse à une intervention

thérapeutique (Biomarkers Definitions Working Group, 2001). Les fonctions des biomarqueurs sont diverses : dépistage, diagnostic, pronostic, classification, suivi, etc. Ils sont utilisés en recherche *in vitro* et en clinique.

Pour que le dépistage par biomarqueur d'un cancer soit possible, plusieurs paramètres doivent être remplis. L'objectif principal est de réduire la mortalité reliée à la maladie dépistée. L'Organisation mondiale de la santé a établi des critères pour guider la pertinence d'un dépistage (Wilson et Jungner, 1968). Premièrement, la maladie doit représenter un fardeau pour la société et le traitement du stade précoce doit pouvoir améliorer le pronostic, ce qui est le cas pour le cancer de l'ovaire, comme discuté plus tôt. De plus, pour pouvoir effectuer le test de dépistage à des intervalles raisonnables, il doit s'écouler un temps suffisamment long entre la détection clinique du cancer par le marqueur et son évolution en cancer de stade avancé ou métastases. Pour le cancer de l'ovaire de stade I/II, le développement du cancer *in situ* en cancer cliniquement détectable prend 4 ans, tandis que les cancers de stade III/IV apparaissent en 1 an, rendant ainsi possible un programme de dépistage (Brown et Palmer, 2009).

Ensuite, le test doit être acceptable pour la population en termes de coûts, d'accessibilité et d'effets secondaires. Ainsi, les échantillons sanguins, obtenus par une prise de sang, sont un moyen à privilégier car entraînant peu de désagréments pour les patients et requérant peu de ressources humaines et financières additionnelles. En effet, la prise de sang peut se faire à l'occasion d'autres bilans sanguins souvent requis pour le suivi des patients, augmentant ainsi les chances d'adhésion au dépistage.

Finalement, le test de dépistage doit être fiable, c'est-à-dire avoir de bonnes sensibilité, spécificité et valeurs prédictives positive et négative. Dans le cas du cancer de l'ovaire, la conséquence d'un dépistage positif est la chirurgie définitive, car une biopsie n'est pas possible. Le test doit détecter le plus grand nombre de cancers possible, tout en évitant de faire subir des chirurgies inutiles à celles qui n'ont pas de cancer. Étant donné la relativement faible prévalence du cancer de l'ovaire, afin d'obtenir une valeur prédictive positive de 10%, ce qui résulterait en 10 chirurgies pour découvrir 1 vrai cancer, valeur jugée acceptable dans le domaine, la spécificité doit être supérieure à 99,6% et la sensibilité supérieure à 75% (Bast, 2004) (Menon *et al*, 2014).

Les snoRNA sont des candidats ayant un bon potentiel de biomarqueurs. En effet, ils sont présents en bonne proportion parmi le transcriptome, et certains d'entre eux font partie des gènes les plus abondants (Deschamps-Francoeur *et al*, 2018) (Boivin *et al*, 2018).

De plus, ils sont détectables dans le sang (Freedman *et al*, 2016) (Umu *et al*, 2018), ce qui permettra de les utiliser comme biomarqueurs dans une prise de sang, un moyen privilégié pour le dépistage.

### ***Étapes du développement de biomarqueurs***

Plusieurs étapes sont nécessaires afin d'aboutir à un biomarqueur utilisé à grande échelle en clinique (Pepe *et al*, 2001). La première étape, qui est celle couverte par ce mémoire, est l'étude préclinique, grâce à laquelle les biomarqueurs potentiels sont identifiés et priorisés à partir d'échantillons de tissus. Par la suite, un essai clinique permet d'estimer la sensibilité et la spécificité de ces biomarqueurs sélectionnés sur le spécimen visé pour le programme de dépistage, comme le sang. Ensuite, une étude rétrospective est réalisée à partir d'échantillons antérieurs de patientes atteintes du cancer de l'ovaire afin d'évaluer la capacité du biomarqueur à détecter le cancer avant l'apparition clinique. Cette étape permet également d'établir les valeurs de référence pour la prochaine étape, soit le dépistage prospectif, avec lequel on peut évaluer le taux de détection et le taux de faux positifs ainsi qu'évaluer le stade auquel on détecte le cancer. Finalement, à plus long terme, des études du contrôle du cancer permettent d'estimer le gain en termes de diminution de la mortalité issu du dépistage.

### **Séquençage à haut débit de l'ARN**

Le séquençage à haut débit, ou de nouvelle génération, de l'ARN (RNA-Seq) est une méthode puissante pour étudier l'expression génique. En effet, le RNA-Seq permet de quantifier simultanément l'abondance globale des ARN dans un échantillon, ce qui permet ensuite de quantifier des différences d'expression entre différentes conditions (Wang *et al*, 2009). Cette méthode se prête donc bien aux phases initiales de découverte de biomarqueurs. Effectivement, avec une seule expérience, il est possible de considérer tous les snoRNA présents dans un tissu. De plus, son coût diminuant et les capacités informatiques augmentant au fil des années, cette technologie devient de plus en plus accessible (Metzker, 2010). De

nombreuses données de séquençage sont donc disponibles publiquement pour diverses organismes et conditions, par exemple *The Cancer Genome Atlas* (TCGA) (<http://cancergenome.nih.gov/>). Or, le séquençage à haut débit des snoRNA présentent des défis supplémentaires, rendant les données générées à présent inutilisables pour étudier cette famille d'ARN en particulier (Veneziano *et al*, 2016). En effet, les données générées pour quantifier l'expression des ARNm étant classiquement obtenues avec la technique des queues poly-A et les snoRNA étant introniques donc majoritairement dépourvus de queue poly-A, ces derniers se retrouvent en faible abondance dans ces ensembles de données (Zhao *et al*, 2014). D'autres ensembles de données sont générés dans le but de quantifier les petits ARN, en particulier les miARN, avec la technique de sélection par la taille, qui isole les ARN plus petits que 200 nucléotides. Il est donc possible a priori de détecter les snoRNA avec cette méthode (Bai et Laiho, 2016). Or, les analyses effectuées précédemment dans notre laboratoire montrent que cette technique présente un biais par rapport à la représentativité des familles de snoRNA en lien avec la longueur des snoRNA (Nottingham *et al*, 2016) (Boivin *et al*, 2018). L'hypothèse proposée pour ce biais est que la longueur plus grande et la structure secondaire complexe des snoRNA H/ACA entrave leur transcription inverse en ADN complémentaire (ADNc) par la transcriptase inverse virale classiquement utilisée lors de la préparation de la librairie. Ceci résulte en une presque absence des snoRNA H/ACA par rapport aux C/D. Le biais que représente ce résultat a été confirmé par des expériences de RT-PCR quantitatif (qRT-PCR), considéré comme l'étalon de référence. Une autre transcriptase inverse bactérienne, la transcriptase inverse thermostable d'intron de groupe II (TGIRT, *thermostable group II intron reverse transcriptase*), développée récemment par le laboratoire d'Alan Lambowitz, apparaît comme une meilleure transcriptase inverse pour séquencer les snoRNA, car elle est plus processive, donc plus apte à passer à travers la longueur et la structure secondaire des snoRNA (Nottingham *et al*, 2016). Nos résultats sur des lignées cellulaires montrent effectivement une meilleure représentation des snoRNA avec cette technique (Boivin *et al*, 2018).

### **Analyses bioinformatiques des résultats de séquençage**

Différents outils bioinformatiques sont disponibles pour obtenir des résultats d'expression différentielle entre deux groupes d'échantillons d'ARN à partir des données de séquençage.

Étant donné les caractéristiques particulières des snoRNA, il faut bien choisir ces outils à chacune des étapes afin de s'assurer de la validité des résultats (Conesa *et al*, 2016).

Tout d'abord, le séquençage produit un fichier fastq, qui contient la séquence de nucléotides, chacun associé avec un score de qualité. L'outil FastQC permet ensuite de vérifier la qualité de ces séquences (Andrews S, 2010). Pour effectuer le séquençage de l'ARN, des adaptateurs dont la séquence est connue sont ajoutés à chaque brin d'ARN présent dans la librairie soumise ; l'outil Cutadapt permet de retirer ces séquences d'adaptateurs (Martin, 2011). Les extrémités des lectures ont tendance à être de moins bonne qualité; l'outil Trimmomatic (Bolger *et al*, 2014) permet de retirer les nucléotides de basse qualité aux extrémités (*trimming*). La prochaine étape est l'alignement des lectures à un génome de référence. Plusieurs outils existent pour réaliser l'alignement (Baruzzo *et al*, 2016). Certains tels STAR (Dobin *et al*, 2013), tiennent compte des différentes possibilités d'épissage. D'autres outils tels Bowtie v2 (Langmead et Salzberg, 2012), possèdent une bonne performance pour les lectures de courte taille, ce qui est important pour récupérer les snoRNA. Dans les outils conventionnels disponibles, les lectures qui s'alignent sur plus d'un gène sont classées comme étant ambiguës, et généralement éliminées des analyses subséquentes. Or, ceci peut poser un problème pour certains snoRNA qui existent en plusieurs copies dans le génome ou dont les séquences sont presque identiques (Deschamps-Francoeur *et al*, 2018). De plus, lorsqu'une lecture incluse dans un gène hôte dépasse l'annotation du plus petit gène, même d'un seul nucléotide, celle-ci est assignée au gène hôte. Sachant que l'annotation des snoRNA est encore inexacte par rapport à la longueur réelle des transcrits (Deschamps-Francoeur *et al*, 2014) (Kishore *et al*, 2013), les outils disponibles risquent donc de perdre un certain nombre de snoRNA et de surestimer le nombre de lectures de leurs gènes hôtes. Afin de remédier à ce problème, l'outil CoCo (Deschamps-Francoeur *et al*, 2018), tient compte de ces situations afin de correctement annoter les lectures correspondant aux snoRNA.

Ces étapes permettent d'obtenir, pour chaque tissu séquençé, le nombre de lectures pour chaque gène. La prochaine étape est donc d'identifier les gènes qui sont différentiellement exprimés entre les deux groupes de comparaison. Encore une fois, différents outils existent (Labaj et Kreil, 2016) (Seyednasrollah *et al*, 2015). L'un des outils les plus souvent utilisés est DESeq2 (Love *et al*, 2014) puisqu'il est approprié pour un nombre petit d'échantillons,

présente le taux de faux positifs le plus faible, est consistant dans ses résultats, possède un temps computationnel raisonnable (quelques minutes) et fournit des ressources didactiques accessibles et bien détaillées (Seyednasrollah *et al*, 2015).

## **Hypothèse/problématique**

La détection précoce du cancer de l’ovaire est la pierre angulaire pour diminuer la mortalité reliée à ce cancer, au 5<sup>e</sup> rang des causes de mortalité par cancer chez les femmes. Les snoRNA apparaissent comme des biomarqueurs potentiels pour le cancer. Or, aucune donnée sur l’expression des snoRNA n’est disponible à ce jour pour le cancer de l’ovaire.

Étant donné le besoin de mieux détecter le cancer de l’ovaire afin d’en réduire la mortalité et le potentiel des snoRNA comme biomarqueurs, il est proposé d’étudier l’abondance des snoRNA dans le HGSC afin de découvrir des biomarqueurs ainsi qu’éventuellement comprendre leurs rôles dans la tumorigenèse.

## **Objectifs**

### **Objectif #1 : Caractérisation de l’expression des snoRNA dans des tissus ovariens**

Le séquençage à haut débit avec une méthode adaptée à la détection des snoRNA permettra de caractériser l’expression de cette famille d’ARN dans des tissus humains.

### **Objectif #2 : Découverte de biomarqueurs du cancer séreux ovarien de haut grade de type snoRNA par séquençage d’ARN à haut débit**

À partir de tissus ovariens séreux *borderline* et cancéreux de haut grade, le séquençage à haut débit de l’ARN adapté à la détection des snoRNA permettra d’obtenir une vue d’ensemble de l’expression des snoRNA et ainsi d’identifier ceux qui sont dérégulés dans le HGSC.

### **Objectif #3 : Validation de biomarqueurs du cancer séreux ovarien de haut grade de type snoRNA par RT-PCR**

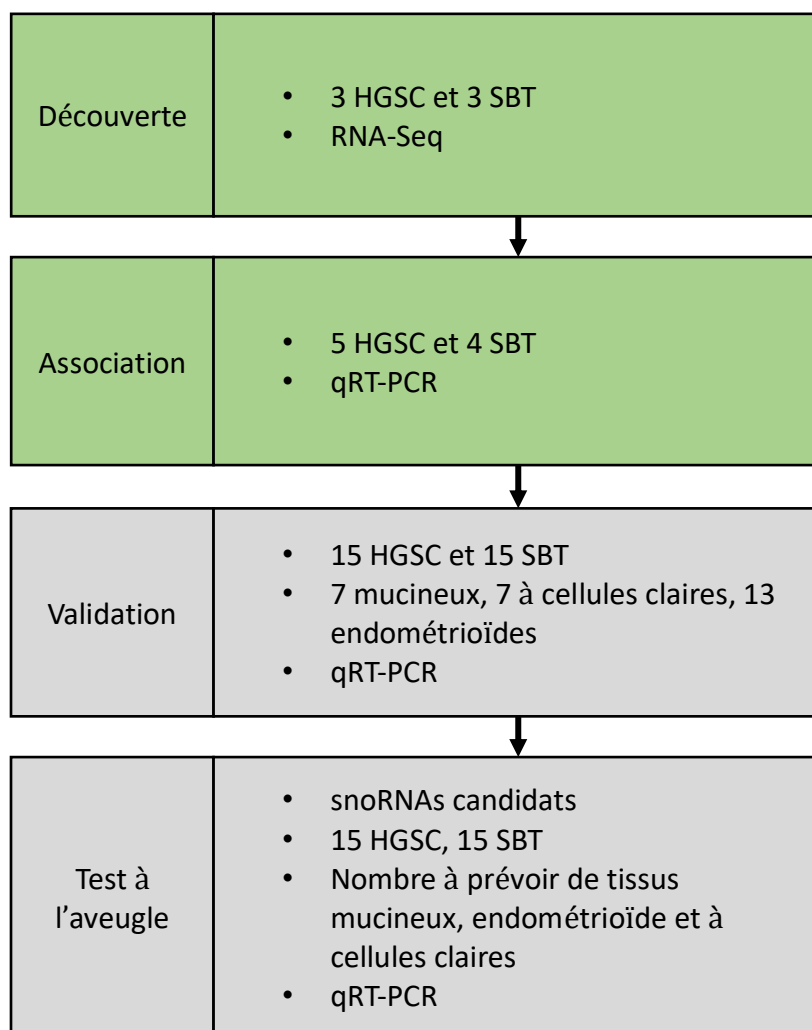
La validation des snoRNA dans des tissus ovariens par qRT-PCR permettront ensuite d’identifier des biomarqueurs potentiels pour le HGSC.



## MATÉRIEL ET MÉTHODES

### Méthodologie de la recherche de biomarqueurs

La recherche de biomarqueurs se divise en quatre étapes (figure 3). La première étape de découverte consiste au séquençage de l'ARN (*RNA-Seq*) sur 3 tissus HGSC et 3 tissus SBT, ce qui est considéré comme un nombre suffisant de répliques pour avoir une bonne détection de gènes différentiellement exprimés (Conesa *et al*, 2016). Les snoRNA significativement surexprimés dans les tissus HGSC issus de cette étape de découverte sont ensuite testés par RT-PCR quantitatif (qRT-PCR) sur 5 tissus HGSC et 4 tissus SBT pour l'étape d'association. Ces deux étapes sont celles couvertes par ce mémoire. Par la suite, les snoRNA significativement surexprimés d'après les expériences de qRT-PCR passeront à l'étape de validation par qRT-PCR sur 15 HGSC et 15 SBT ainsi que 7 tissus cancéreux ovariens de d'autres types histologiques, soit mucineux (4 *borderline*, 2 grade II, 1 grade III), 7 à cellules claires (3 grade II, 4 grade III) et 13 endométrioïdes (4 grade I, 3 grade II, 6 grade III). Les snoRNA permettant une discrimination des HGSC seront finalement testés à l'aveugle par qRT-PCR sur 15 HGSC et 15 SBT, et éventuellement sur des tissus ovariens cancéreux de types mucineux, endométrioïdes et à cellules claires dans la suite du projet. Les tissus sont utilisés à une seule étape, c'est-à-dire qu'il s'agit de tissus différents à chaque étape.



**Figure 3 : Méthodologie de la recherche de biomarqueurs**  
**En vert, les étapes couvertes par ce mémoire, en gris les étapes ultérieures**

### Tissus ovariens

Les tissus utilisés sont présentés dans le tableau 1. Il s'agit de tissus frais congelés (*fresh frozen*) obtenus auprès de diverses banques de tissus avec le diagnostic associé. Les critères d'exclusion sont : antécédent de cancer, traitement de chimiothérapie antérieur à l'acquisition du spécimen, origine de la tumeur inconnue ou autre que l'ovaire, tumeur mixte.

**Tableau 1 : Description des tissus ovariens utilisés**

Nom du tissu	Provenance	Type histologique	Grade	Utilisé pour quelle expérience
<b>SBT1</b>	ARTB	Séreux	SBT	RNA-Seq
<b>SBT2</b>	ARTB	Séreux	SBT	RNA-Seq
<b>SBT3</b>	ARTB	Séreux	SBT	RNA-Seq
<b>HGSC1</b>	OTB	Séreux	HGSC	RNA-Seq
<b>HGSC2</b>	FRQS	Séreux	HGSC	RNA-Seq
<b>HGSC3</b>	OTB	Séreux	HGSC	RNA-Seq
<b>D03113</b>	OTB	Séreux	HGSC	qRT-PCR association
<b>D02960</b>	OTB	Séreux	HGSC	qRT-PCR association
<b>B02779</b>	OTB	Séreux	HGSC	qRT-PCR association
<b>B02610</b>	OTB	Séreux	HGSC	qRT-PCR association
<b>B01779</b>	OTB	Séreux	HGSC	qRT-PCR association
<b>B00873</b>	OTB	Séreux	SBT	qRT-PCR association
<b>B01547</b>	OTB	Séreux	SBT	qRT-PCR association
<b>D02795</b>	OTB	Séreux	SBT	qRT-PCR association
<b>D03512</b>	OTB	Séreux	SBT	qRT-PCR association
<b>OVC458</b>	FRQS	Séreux	SBT	qRT-PCR validation
<b>OVC 596</b>	FRQS	Séreux	SBT	qRT-PCR validation
<b>RT266</b>	ARTB	Séreux	SBT	qRT-PCR validation
<b>RT148</b>	ARTB	Séreux	SBT	qRT-PCR validation
<b>VOA 123</b>	GPEC	Séreux	SBT	qRT-PCR validation
<b>VOA 245</b>	GPEC	Séreux	SBT	qRT-PCR validation
<b>VOA 391</b>	GPEC	Séreux	SBT	qRT-PCR validation
<b>B01368</b>	OTB	Séreux	SBT	qRT-PCR validation

OTB : *Ontario Tumor Bank*

ARTB : *Alberta Research Tumor Bank*

FRQS : Fonds de recherche du Québec – Santé

GPEC: *The Genetic Pathology Evaluation Centre, University of British Columbia*

### **Extraction de l'ARN et préparation des bibliothèques de séquençage**

L'ARN des tissus ovariens a été extrait avec le *RNEasy Mini Kit* de Qiagen. Les bibliothèques sont non-fragmentées et ribodéplétées avec le *Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat)* d'Illumina afin de mieux enrichir les petits ARN dont les snoRNA. L'ARN ribodéplété a ensuite été purifié avec le kit *RNA clean and concentrator* de Zymo selon un protocole modifié permettant de mieux retenir les petits ARN (soit avec huit fois le volume de EtOH 95-100% au lieu de trois fois selon le protocole du produit). La transcriptase inverse utilisée pour produire l'ARN complémentaire est la TGIRT (Mohr *et al*, 2013).

## Séquençage et analyse du séquençage

Le séquençage de l'ARN a été effectué sur un séquenceur Illumina HiSeq2500. Il s'agit d'un séquençage en paires, permettant d'avoir des bonnes qualités de lectures aux deux extrémités.

Les étapes de l'analyse de ce séquençage sont présentées à la figure 4.

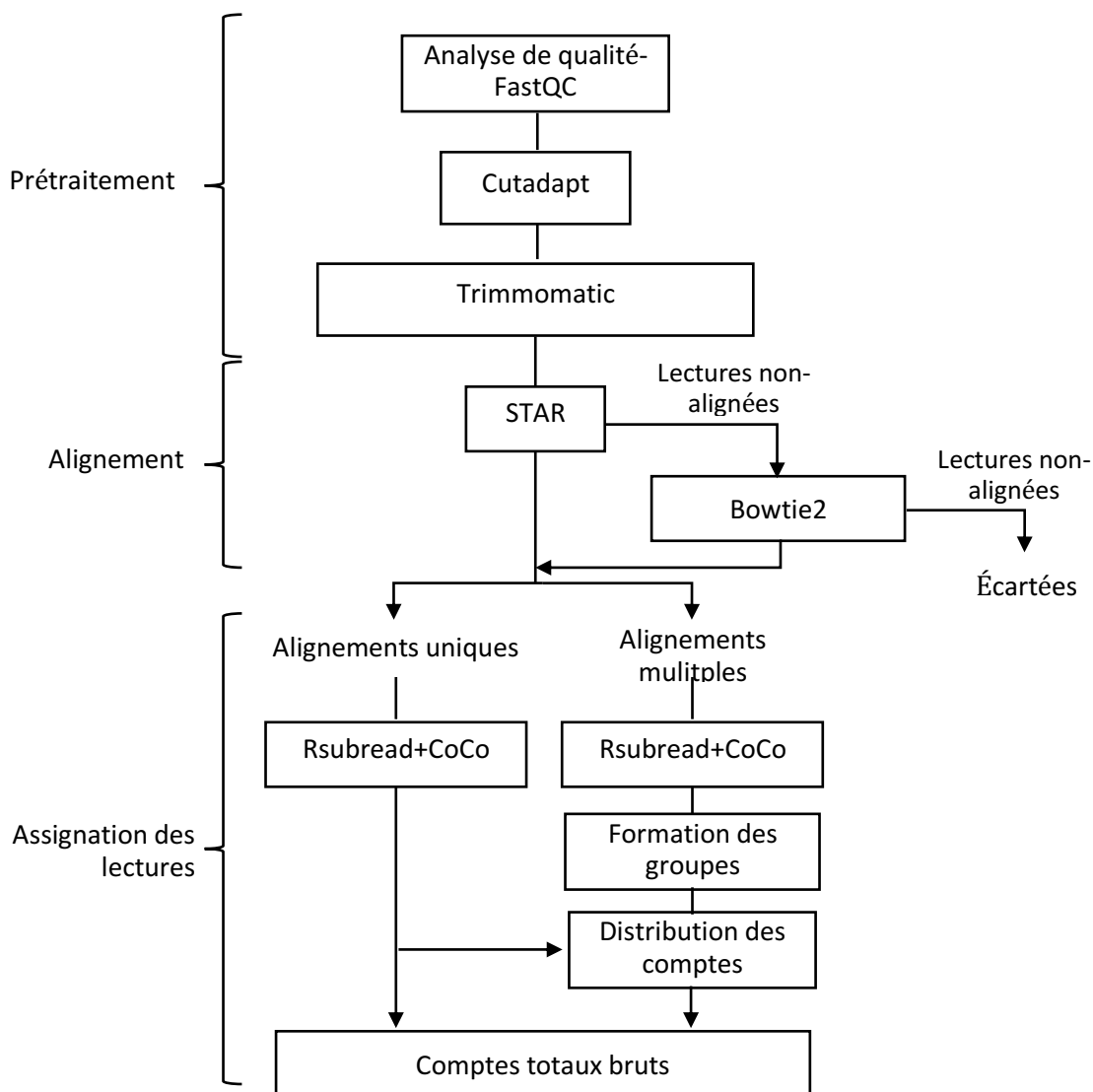


Figure 4 : Pipeline d'analyse RNA-Seq

Le contrôle de la qualité des lectures a été effectué avec FastQC (Andrews et Mutch, 2017). Les séquences des adaptateurs ont été retirés avec Cutadapt (Martin, 2011) avec les paramètres suivants :

```
-g GATCGTCGGACTGTAGA AACTCTGAACGTGTAGATCTCGGTGGTCGCCGTATCATT  
-a AGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCCGTCTT  
CTGCTTG  
--minimum-length 2  
--match-read-wildcards  
-q 3  
--paired-output
```

Les positions de mauvaise qualité ont été retirées avec Trimmomatic (Bolger *et al*, 2014) avec les paramètres suivants :

```
-XX:ParallelGCThreads=1  
-Xmx30G  
-jar trimmomatic-0.35.jar PE  
-threads 22  
-phred33  
TRAILING:30
```

Le génome de référence utilisé est hg38 version 87 d'Ensembl (Yates *et al*, 2016). L'alignement des lectures s'est d'abord fait avec STAR (Dobin *et al*, 2013) avec les paramètres par défaut sauf pour les paramètres suivants :

```
--outSAMtype BAM SortedByCoordinate,  
--outSAMprimaryFlag AllBestScore  
--alignIntronMax 1250000
```

Les lectures non alignées par STAR ont été alignées avec Bowtie2 (Langmead et Salzberg, 2012) avec les paramètres suivants :

```
--local  
-p 24  
-q  
-I 13
```

Les lectures demeurant non alignées ont été écartées. Pour les lectures alignées, leur assignation à leur gène s'est fait avec *featureCounts* (Liao *et al*, 2013) avec les paramètres suivants :

```
-C  
--larg-estOverlap  
-p  
-B  
-s 1  
--minOverlap 10
```

Le paramètre `-M` est spécifié pour les lectures avec alignements multiples, et ne l'est pas pour les alignements uniques.

Pour l'assignation des lectures, le fichier d'annotation soumis à *featureCounts* est un fichier fait sur mesure tel que décrit dans (Deschamps-Francoeur *et al*, 2018) avec le fichier d'annotation d'Ensembl hg38 version 87 (Yates *et al*, 2016) auquel a été ajouté celui de RefSeq (O'Leary *et al*, 2016) pour 20 snoRNA manquants dans Ensembl, ainsi que les ARN de transfert (ARNt) de GtRNADB (Chan et Lowe, 2009). Par la suite, CoCo permet de corriger l'assignation des gènes retrouvés au sein d'un gène hôte ainsi que d'attribuer les lectures avec alignements multiples (Deschamps-Francoeur *et al*, 2018). Toutes ces étapes permettent d'obtenir les comptes bruts pour chaque gène dans chaque tissu.

### **Analyse de l'expression différentielle entre les tissus SBT et HGSC**

Les analyses d'expression différentielle ont été effectuées avec le progiciel DESeq2 version 3.8 (Love *et al*, 2014) avec le logiciel R version 3.5.3 (R Development Core Team, 2019). Les comptes bruts obtenus par l'analyse de séquençage ont été soumis à DESeq2 afin d'obtenir les comptes normalisés et évaluer l'expression différentielle. Les paramètres par défaut ont été utilisés. La fonction *collapse technical replicates* a été utilisée pour les tissus SBT2 et SBT3 qui ont chacun quatre répliques techniques. Les ARN étant surexprimés dans les HGSC, ayant une valeur P ajustée (avec la correction Benjamin-Hochberg) inférieure à 0,0001 et une accumulation moyenne supérieure à 200 en abondance normalisée par DESeq2 ont été sélectionnés comme étant les candidats significativement surexprimés.

## Corrélation entre le RNA-Seq et le qRT-PCR

La corrélation entre le logarithme base 2 des comptes normalisés par DESeq2 des données de RNA-Seq et l'abondance mesurée par qRT-PCR a été calculée avec R avec la fonction *corr.test* avec les méthodes « spearman » et « pearson », successivement.

## Diagrammes de quartiles (*boxplots*)

Les diagrammes de quartiles ont été créés avec la fonction *boxplot* de la librairie *ggplot2* (Wickham, 2016.) de R version 3.5.3 (R Core Team, 2019). Les gènes retenus pour l'analyse sont ceux ayant au moins 10 comptes par million dans au moins un tissu.

## Production des heatmaps

Les *heatmaps* des snoRNA surexprimés candidats ont été produits avec la fonction *heatmap.2* de R provenant du progiciel *gplots* version 3.0.1.1 (Warnes *et al*, 2019). Les comptes bruts de DESeq2 ont été transformés avec la fonction « *rlog* » de DESeq2 afin d'obtenir des logarithmes base 2 régularisés. Les *heatmaps* ont été produits par agrégation hiérarchique (*hierarchical clustering*) et la distance calculée par la méthode euclidienne avec les paramètres suivants :

```
Rowv = TRUE,
Colv = TRUE,
distfun = function(x) dist(x,method = "euclidean"),
hclustfun = function(x) hclust(x, method = "complete"),
dendrogram="both"
```

## RT-PCR quantitatif (qRT-PCR)

L'extraction de l'ARN total a été effectuée sur des culots de cellules avec TRIzol (Invitrogen) et chloroforme selon le protocole du manufacturier. La phase aqueuse a été récupérée et mélangée avec un volume d'éthanol 70% et appliqué directement sur une colonne RNeasy Mini Kit (Qiagen). Un traitement à la DNase sur la colonne et la récupération de l'ARN total ont été faits selon le protocole du manufacturier. La qualité de l'ARN et la présence de contamination par de l'ADN génomique ont été vérifiées tel que décrit précédemment

(Brosseau *et al*, 2010). L'intégrité de l'ARN a été évaluée avec un *Agilent 2100 Bioanalyzer* (Agilent Technology). La transcription inverse s'est effectuée sur 1.1 µg d'ARN total avec une transcriptase inverse Transcriptor, des hexamères aléatoires, des dNTP (Roche Diagnostics) et 10 unités de RNaseOUT (Invitrogen) selon le protocole du manufacturier, dans un volume total de 10µg. Toutes les amorces sens et anti-sens ont été individuellement resuspendues dans 20–100 µM de solution mère dans un tampon de Tris-EDTA (IDT) et diluées en pair d'amorces dans 1 µM d'eau exempte de RNase et DNase (IDT). Les réactions de qRT-PCR ont été effectuées dans 10 µl sur des plaques de 96 puits plates sur un thermocycleur CFX-96 (BioRad) avec 5 µL de *2X iTaq Universal SYBR Green Supermix* (BioRad), 10 ng (3 µl) de l'ADNc, et 200 nM final (2 µl) de solution de paires d'amorces. Les conditions des cycles sont les suivantes : 3 min à 95°C; 50 cycles: 15 sec à 95°C, 30 sec à 60°C, 30 sec à 72°C. Les niveaux d'expression relative ont été calculés avec le logiciel qBASE (Hellemans *et al*, 2007), d'après les gènes constitutifs YWHAZ, PUM1 et MRPL19 pour l'ADNc. Le design et la validation des amorces ont été faits tel que décrit précédemment (Brosseau *et al*, 2010). Pour chaque expérience de qPCR, un contrôle sans matrice a été effectué pour chaque paire d'amorce et ont été uniformément négatifs. Les amorces utilisées pour le qRT-PCR des snoRNAs et des gènes constitutifs sont présentés dans les tableaux 2 et 3, respectivement. Le design des amorces s'est fait en tenant compte des lectures réelles séquencées et non des séquences de référence étant donné des disparités par rapport aux fins des extrémités de certains snoRNAs. Les snoRNA significativement surexprimés dans les HGSC ont été déterminés par test t de Student unilatéral avec alpha=0,05 sur les moyennes des expressions relatives des HGSC par rapport aux SBT.



**Tableau 2 : Amorces utilisées pour le qRT-PCR des snoRNA**

snoRNA	Identifiant Ensembl	Amorce sens	Amorce anti-sens
<b>SNORA19</b>	ENSG00000207468	TGACCTGCTTTCTTTTAT GTGAGTAGTGT	TGTCTGAAGGAGACTGG CAGCA
<b>SNORA51</b>	ENSG00000271798	GGCTGTGTTCTTACACT GACTG	GGGCCTGAGCTGAGGT GTAT
<b>SNORA2</b>	ENSG00000206647	AGTGTGTTGCTGTAGCT GTTGG	GCCACTGACACTGGGAT ACT
<b>SNORA44</b>	ENSG00000252840	GCCATGGGATCTCCAAC TGCAT	GGGATTGTGCTGACCTG TTCT
<b>SNORA65</b>	ENSG00000201302	CCAGGTCTCTGTTGGCT GGTGC	AAGAACAACCCTGTTTC CTAAGCTGG
<b>SNORA56</b>	ENSG00000206693	CCTTTCTAGTCTGGCTC GTGGGA	ACGCTGGACAGAAGAA CTCACCA
<b>SNORA2</b>	ENSG00000199959	TCCTGATTCAAGGCCAG CAG	AGATGGCCAACAGACCA TGA
<b>SNORA5B</b>	ENSG00000200656	GCAGCCATGTCAAATTC AGTGCC	TTTATCCCTGAGCCTGG GACAC
<b>SNORA16B</b>	ENSG00000201544	GCTGCAGCTGCTTCCAT GTAG	GCAAACAGAACTATGAC CATCAAGG
<b>SNORA30B</b>	ENSG00000202189	TAGCATTTTCACAGGTC CTCCCTCA	AGGGCAAGGATGCAAT CAAGCG
<b>SNORA46</b>	ENSG00000212371	TGTAGCACAGAAGAATT CCTGGCTTC	ACTCCACACTATAGCAG CAGAATGGC
<b>SNORA48</b>	ENSG00000212626	CTGGGTAGAGTGGCAT CTGGCT	CCAGGCATGGGCACAG ACAAA
<b>SNORA72</b>	ENSG00000201898	AGTCAGGACAGGATAAA CAGTCGC	GGGAAAGAACTAGGTCT AGGGACACG

**Tableau 3 : Amorces utilisées pour le qRT-PCR des gènes constitutifs**

Gène de référence	Amorce sens	Amorce anti-sens
<b>YWHAZ</b>	TCCCCAATGCTTCACAAGCAGA	TCTTGTGCATCACCAGCGGCAA
<b>MRPL19</b>	AAGGAGAAAAGTACTCCACATTCCA GAG	TGGGTCTAGCTGTAGTAACACG A
<b>PUM1</b>	TGAGGTGTGCACCATGAAC	CAGAATGTGCTTGCCATAGG

### Analyses phénotypiques

Les analyses phénotypiques ont été effectuées sur la lignée cellulaire SKOV3.ip1, un modèle de cancer séreux ovarien. Les déplétions de snoRNA ont été effectuées par des ASO (*antisense oligonucleotide*) (Liang *et al*, 2011). Le design et la production de ces ASO a été fait par *Integrated DNA technologies* (IDT) et les séquences sont présentées dans le tableau 4. Les ASO ont été transfectés avec la Lipofectamine 2000. 2 ASO différents ont été conçus

par snoRNA. Les contrôles négatifs sont des cellules transfectées avec Lipofectamine seulement, des cellules transfectées avec un ASO avec la séquence aléatoire de SNORD31 qui n'est complémentaire à aucun transcrit annoté (NCSR2) ainsi que des cellules non-traitées (NT). Kif11 a été utilisé comme contrôle positif. Les courbes de croissance ont été mesurées par impédance avec le système xCELLigence (Ke *et al*, 2011).

**Tableau 4 : Séquences des ASO utilisés pour la déplétion de snoRNA**

snoRNA	ASO
<b>SNORA19 #1</b>	mA*mG*mC*mA*mG*G*T*C*A*A*T*G*A*A*A*mU*mG*mU*mG*mC
<b>SNORA19 #2</b>	mG*mU*mC*mU*mG*A*A*G*G*A*G*A*C*T*G*mG*mC*mA*mG*mC
<b>NCSR2</b>	mC*mA*mU*mG*mA*C*T*C*G*C*A*G*A*A*C*G*mA*mU*mA*mG*mU

\* : squelette phosphorothioate

m : bases d'ARN 2'O-methyl

NCSR2 : séquence aléatoire de SNORD31 (contrôle négatif)

## RÉSULTATS

### **Le séquençage de l'ARN avec TGIRT permet la détection des snoRNA dans des tissus ovariens**

Pour détecter les snoRNA différentiellement exprimés entre les HGSC et les SBT, l'ARN de trois tissus HGSC et trois tissus SBT a été extrait, ribodéplété et préparé en librairies pour séquençage à haut débit de type TGIRT. Aucune étape de fragmentation de l'ARN n'a été effectuée de sorte à ne considérer que les petits ARN, dont les snoRNA font partie. Une fois le séquençage effectué, la qualité des lectures a été évaluée et les lectures de basse qualité ont été retirées (par *trimming*), puis les lectures résultantes ont été alignées au génome humain. Le tableau 5 montre le nombre de lectures suivants les différentes étapes du pipeline d'analyse de RNA-seq. Les nombres de lectures des quatre répliques techniques ont été additionnés pour les tissus SBT2 et SBT3.

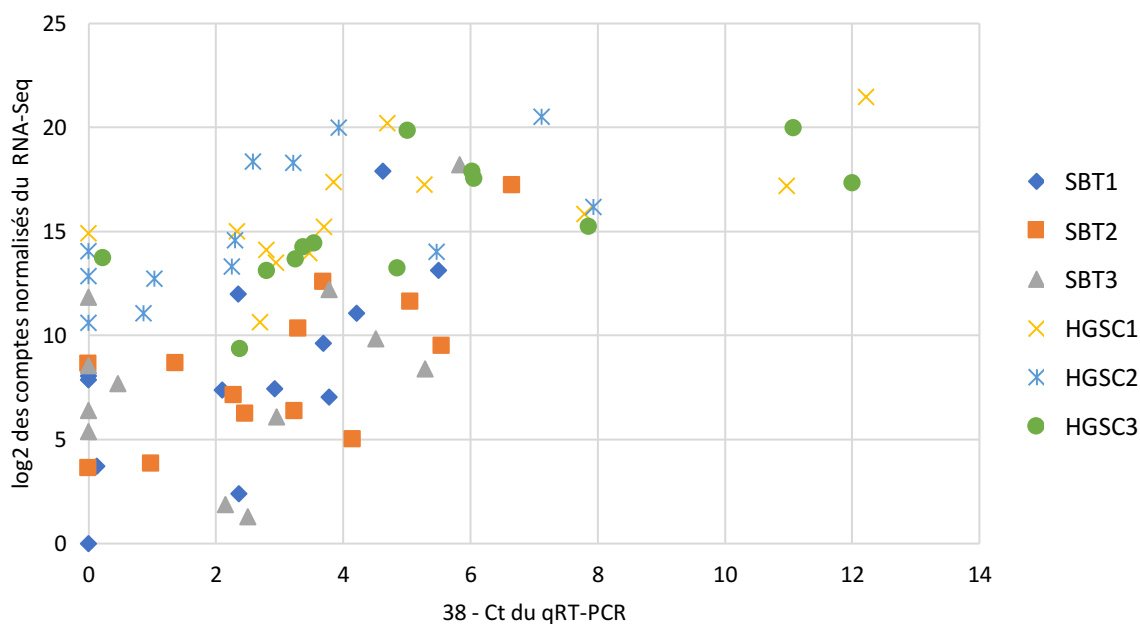
**Tableau 5 : Nombre de lectures aux différentes étapes du traitement des données de séquençage**

<b>Tissu</b>	<b>Nb de paires de lectures avant <i>trimming</i></b>	<b>Nb de paires de lectures après <i>trimming</i></b>	<b>Nb de lectures alignées par STAR</b>	<b>Nb de lectures alignées par Bowtie2</b>
<b>SBT1</b>	22552724	13466599	10863422	3736160
<b>SBT2</b>	101542703	72179526	58886819	17825934
<b>SBT3</b>	136403191	94970881	67953397	20657711
<b>HGSC1</b>	100384826	86027123	68638652	32557141
<b>HGSC2</b>	58199650	54180237	45327680	210899971
<b>HGSC3</b>	72937112	62493128	47188848	22949250

Suivant l'alignement au génome, les transcrits détectés ont été quantifiés en les assignant à leur gène correspondant en utilisant l'outil CoCo, qui tient compte des lectures s'alignant aux gènes imbriqués et aux lectures multi-alignées.

Afin de s'assurer que cette nouvelle méthode de séquençage et d'alignement est valide pour la détection des snoRNA, 13 snoRNA ont été quantifiés par qRT-PCR, jugé comme méthode

de référence, dans les mêmes tissus qui ont été séquencés par RNA-Seq, afin d'observer la corrélation entre ces deux méthodes. Le Ct du qRT-PCR (*cycle threshold*) représente le nombre de cycles d'amplification avant que l'ARN d'intérêt soit détecté. Ainsi, il représente une échelle logarithmique base 2 puisque la quantité d'ARN double à chaque cycle. Plus le Ct est élevé, moins l'ARN est abondant. En appliquant un logarithme base 2 sur les comptes d'abondance quantifiés par RNA-Seq et en soustrayant d'un chiffre plus élevé que tous les Ct obtenus (dans ce cas 38) le Ct obtenu par qRT-PCR, la relation entre les deux méthodes de séquence est linéaire, et on peut donc appliquer les calculs de corrélation de Pearson et de Spearman. La figure 5 et le tableau 6 montrent en effet qu'il y a une corrélation entre les deux méthodes.

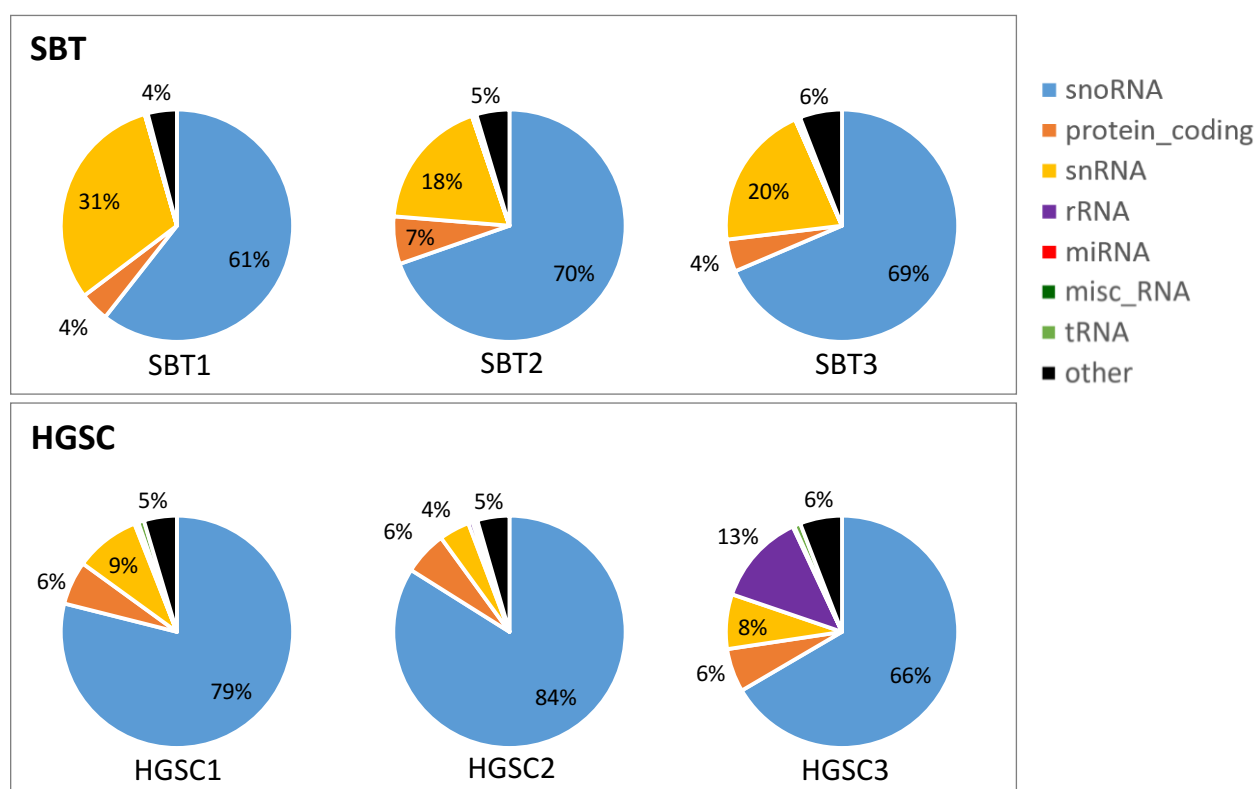


**Figure 5 : Abondance des snoRNA dans les tissus ovariens séquencés par RNA-Seq et qRT-PCR**

**Tableau 6 : Corrélations entre l'abondance détectée par RNA-Seq et par qRT-PCR**

	SBT1	SBT2	SBT3	HGSC1	HGSC2	HGSC3
<b>Pearson</b>	0.67	0.69	0.42	0.64	0.65	0.66
<b>Spearman</b>	0.57	0.65	0.48	0.75	0.73	0.77

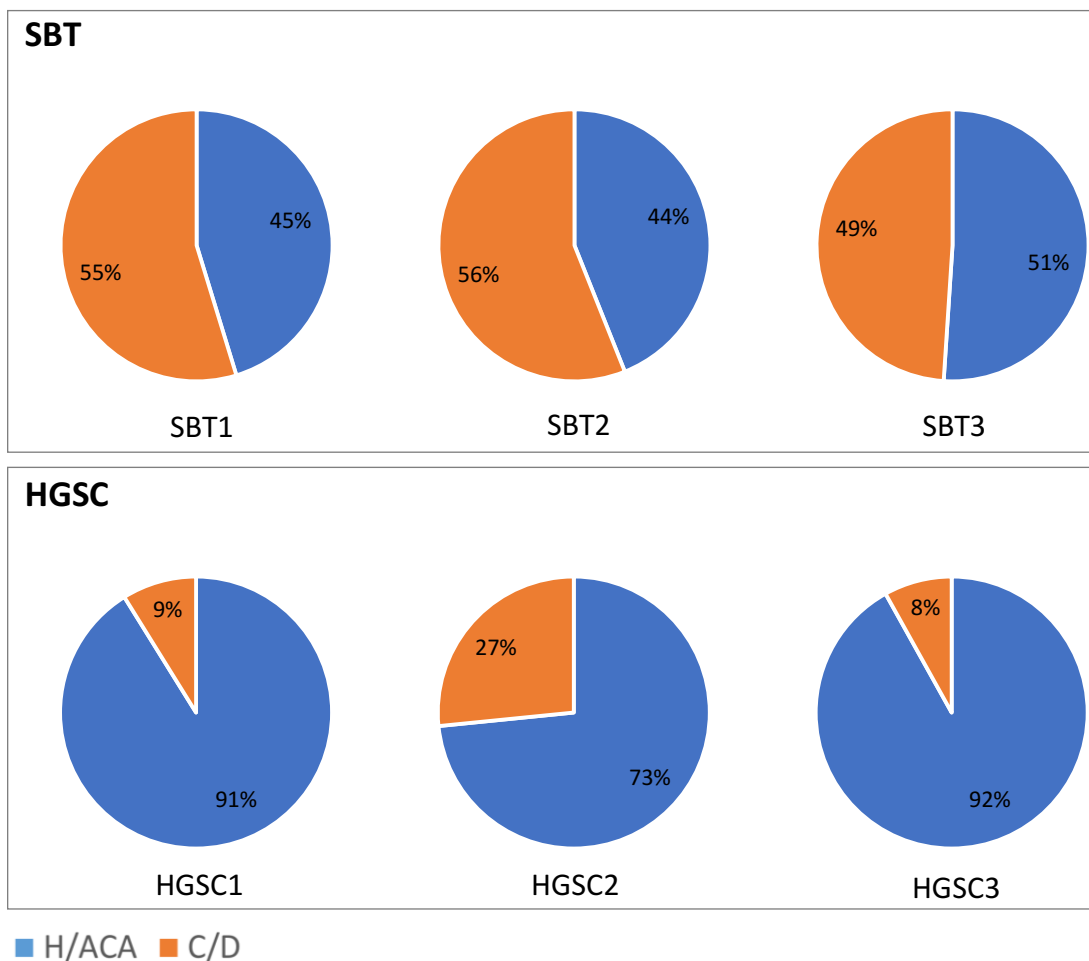
On peut donc maintenant étudier les résultats du séquençage. La distribution d'abondance des transcrits par biotype (figure 6) montre que la méthodologie utilisée détecte principalement des snoRNA, des snRNA (petits ARN nucléaires, *small nuclear RNA*), des fragments d'ARN codants (protein\_coding) et des ARN ribosomiaux (rRNA). Entre 61 et 84% des lectures totales des tissus séquencés provient de snoRNA, ce qui en fait la famille d'ARN la plus représentée dans ces séquençages.



**Figure 6 : Répartition des lectures des familles d'ARN séquencées par RNA-Seq**

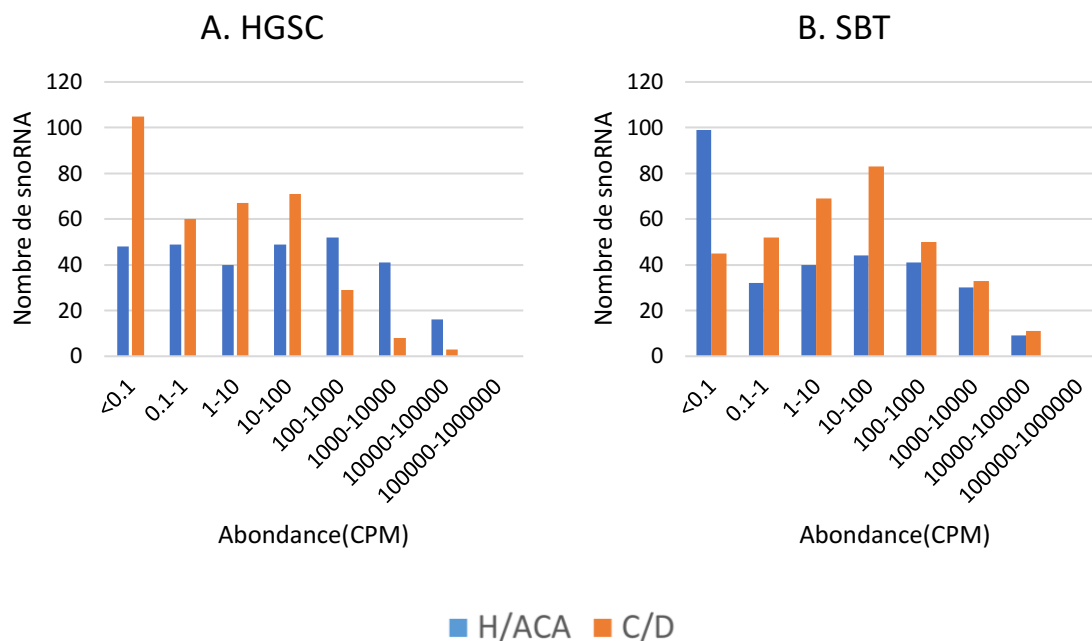
### **La régulation de l'abondance des snoRNA dans les HGSC diffère selon le type C/D ou H/ACA**

Des séquençages de type TGIRT-seq de lignées cellulaires, dont les SKOV3IP1, une lignée de cancer de l'ovaire, indiquent une proportion similaire de snoRNA de type C/D et H/ACA parmi les gènes détectés (Boivin *et al*, 2018). Dans notre séquençage de tissus ovariens, nous détectons une proportion similaire des deux types de snoRNA dans les tissus SBT. Or, les snoRNA H/ACA sont plus abondants que les snoRNA C/D dans les tissus HGSC (figure 7).



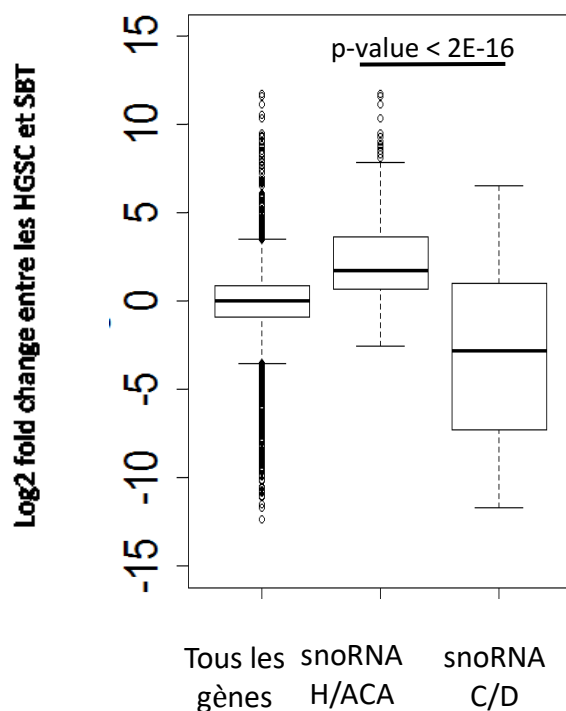
**Figure 7 : Distribution des snoRNA C/D et H/ACA dans les tissus séquencés par RNA-Seq**

En observant la distribution de l'abondance des snoRNA selon leur type (C/D ou H/ACA) et selon les tissus HGSC ou SBT (figure 8), on remarque qu'il y a moins de snoRNA avec une abondance inférieure à 1 compte par million (CPM) dans les HGSC par rapport aux SBT, tandis qu'on retrouve davantage de snoRNA H/ACA ayant une abondance supérieure à 100 CPM dans les HGSC. Pour ce qui est des snoRNA C/D, on observe la tendance inverse.



**Figure 8 : Distribution de l'abondance des snoRNA dans les tissus HGSC (A) et SBT (B)**

Par la suite, pour identifier les gènes différentiellement exprimés entre les HGSC et les SBT, l'outil DESeq2 a été employé. Parmi tous les gènes séquencés dont l'abondance est d'au moins 10 comptes par million dans un moins tissu, la distribution est plutôt uniforme entre les gènes surexprimés, sous-exprimés et ceux ne montrant pas de différences entre les tissus HGSC et SBT. Or, on observe une différence significative entre la distribution des snoRNA C/D et H/ACA : les snoRNA C/D sont globalement significativement sous-exprimés dans les HGSC, tandis que les snoRNA H/ACA sont globalement significativement surexprimés (figure 9). Ceci concorde avec les figures 7 et 8, où les H/ACA sont retrouvés en plus grande proportion dans les HGSC.



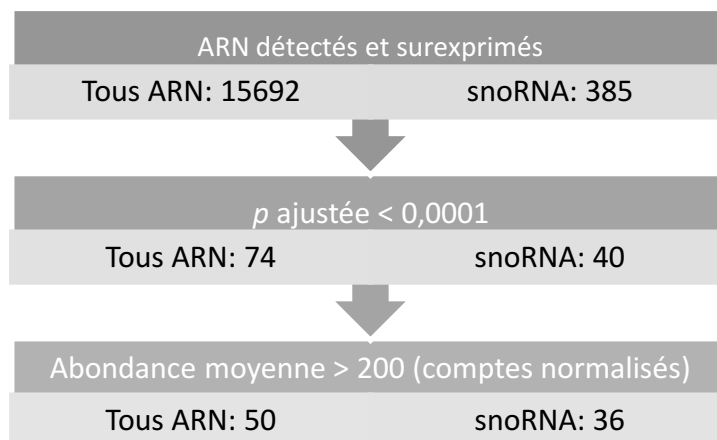
**Figure 9 : Diagramme de quartiles (boxplot) du  $\log_2$  fold change (Log2FC) pour tous les gènes séquencés et en fonction de la famille de snoRNA**

### **Les snoRNA sont enrichis lors de la sélection des candidats significativement surexprimés dans les HGSC**

Pour identifier les candidats les plus prometteurs pour servir de biomarqueurs pour les HGSC, plusieurs filtres ont été appliqués (figure 10). À partir de tous les ARN détectés par séquençage, seuls les ARN surexprimés ont été gardés pour la sélection des biomarqueurs. En effet, on recherche idéalement une molécule qui est différenciellement exprimée entre les HGSC et les SBT, et qui est facilement détectable dans les HGSC. Des 15692 ARN détectés et surexprimés dans le séquençage, 385 sont des snoRNA. Un seuil de  $p < 0,0001$  a été employé pour ne conserver que les transcrits différenciellement surexprimés supportés par une bonne significativité statistique. Parmi les 74 transcrits passant ce seuil, 40 sont des

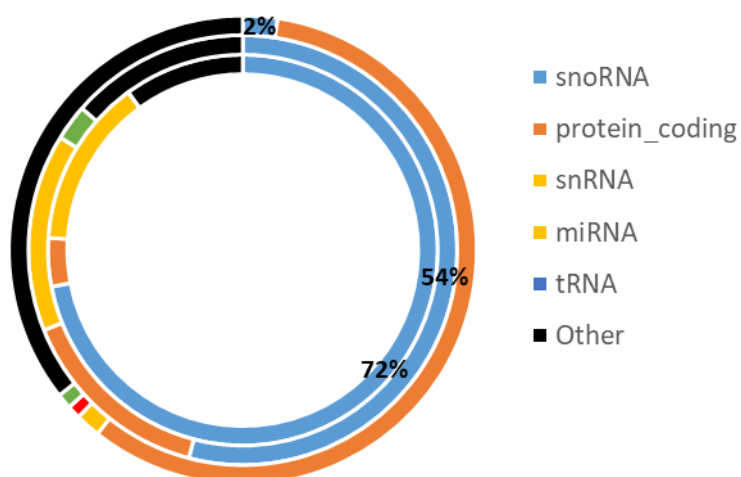


snoRNA, démontrant un important enrichissement des snoRNA. Finalement, un bon biomarqueur sera abondant pour permettre une détection facile. Nous avons donc choisi un seuil minimal d'abondance moyenne de 200 (comptes normalisés par DESeq2). On obtient donc 36 snoRNA candidats significativement surexprimés dans les HGSC comme biomarqueurs potentiels.



**Figure 10 : Nombre de candidats sélectionnés aux différentes étapes de sélection**

La figure 11 montre l'enrichissement des snoRNA par rapport aux autres familles d'ARN lors de l'application des filtres pour la sélection des candidats. Parmi les ARN les plus significativement surexprimés dans les HGSC et ayant une abondance moyenne de plus de 200, on observe un important enrichissement des snoRNA, qui, alors qu'ils représentent 2% de tous les ARN surexprimés (significativement ou non), représentent 72% des ARN considérés suffisamment abondants et significativement surexprimés à  $p < 0,0001$ .



**Figure 11 : Distribution des familles d'ARN lors des différentes étapes de sélection des candidats discriminant les HGSC des SBT**

Extérieur : ARN détectés et surexprimés; Milieu : ARN différentiellement surexprimés à  $p$  ajusté  $< 0,0001$  ; Intérieur : ARN significatifs ayant une abondance moyenne supérieure à 200.

### **Les snoRNA candidats significativement surexprimés dans les HGSC sont principalement des snoRNA H/ACA connus pour modifier une position dans l'ARN ribosomal**

Le tableau 7 présente les caractéristiques des snoRNA candidats. Des 36 snoRNA significativement surexprimés dans les HGSC et ayant une abondance moyenne de plus de 200, seulement 2 sont de la famille C/D, les 34 autres étant tous de la famille H/ACA. Certains snoRNA se retrouvent en plus d'une copie (SNORA2, SNORA31, SNORA48, SNORA51, SNORA70, SNORA72), car ils sont retrouvés à différents endroits dans le génome et leurs séquences varient suffisamment pour qu'on puisse les distinguer par RNA-Seq. En regroupant ces copies, seulement deux snoRNA parmi les candidats sont orphelins, c'est-à-dire qu'ils n'ont pas de fonction connue, soit SNORA16B et SNORA51. Tous les autres sont connus pour modifier une position dans l'ARN ribosomal.

**Tableau 7 : Liste des snoRNA candidats et leurs caractéristiques**

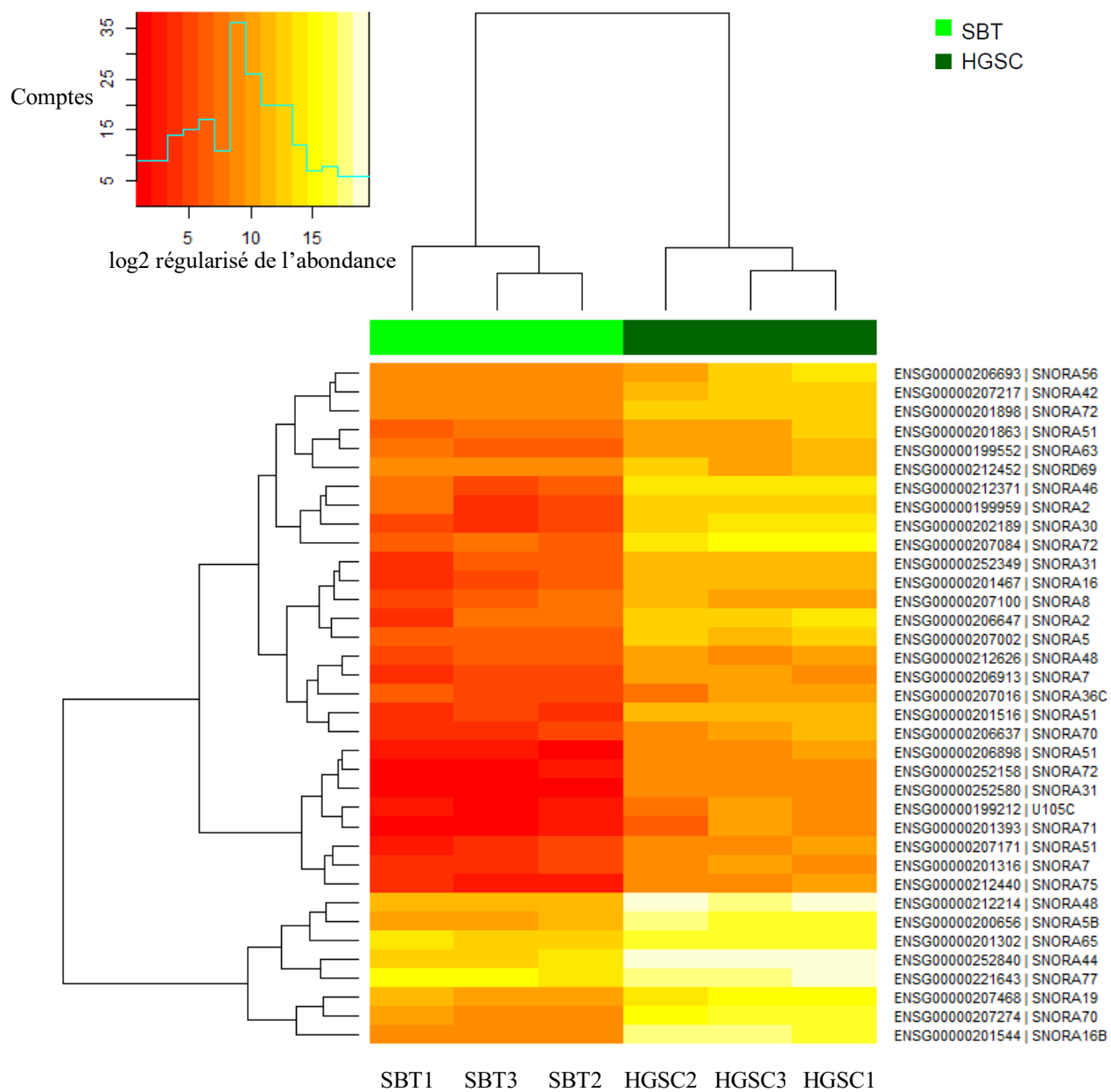
Nom	Identifiant Ensembl	Famille	Fonction	Abondance moyenne	log2FC	padj
SNORA16	ENSG00000201467	H/ACA	28S rRNA U4412	2390	7,72	2,37E-13
SNORA16B	ENSG00000201544	H/ACA	orphelin	125498	11,12	1,49E-41
SNORA19	ENSG00000207468	H/ACA	28S rRNA U3618 28S rRNA U3709	19521	4,90	1,18E-06
SNORA2	ENSG00000206647	H/ACA	28S rRNA U4263 28S rRNA U4282	5752	7,65	1,16E-09
SNORA2	ENSG00000199959	H/ACA	28S rRNA U4263 28S rRNA U4282	5408	7,47	3,20E-07
SNORA30	ENSG00000202189	H/ACA	28S rRNA U4643	11962	11,62	7,05E-32
SNORA31	ENSG00000252580	H/ACA	18S rRNA U218 28S rRNA U3713	480	11,71	2,32E-14
SNORA31	ENSG00000252349	H/ACA	18S rRNA U218 28S rRNA U3713	2644	6,84	2,08E-08
SNORA36C	ENSG00000207016	H/ACA	18S rRNA U1244	774	5,35	1,26E-05
SNORA42	ENSG00000207217	H/ACA	18S rRNA U572 18S rRNA U109	2820	3,93	8,03E-05
SNORA44	ENSG00000252840	H/ACA	18S rRNA U822 18S rRNA U686	532955	7,84	5,51E-23
SNORA46	ENSG00000212371	H/ACA	18S rRNA U649	13845	8,70	6,05E-13
SNORA48	ENSG00000212214	H/ACA	28S rRNA U3797	306039	9,04	1,58E-25
SNORA48	ENSG00000212626	H/ACA	28S rRNA U3797	760	5,06	3,08E-06
SNORA5	ENSG00000207002	H/ACA	18S rRNA U1238 18S rRNA U1625	3361	7,24	8,76E-19
SNORA51	ENSG00000201516	H/ACA	orphelin	2333	8,81	1,30E-16
SNORA51	ENSG00000206898	H/ACA	orphelin	472	9,50	2,97E-15
SNORA51	ENSG00000207171	H/ACA	orphelin	598	6,71	5,70E-09
SNORA51	ENSG00000201863	H/ACA	orphelin	1828	5,39	3,61E-07
SNORA56	ENSG00000206693	H/ACA	28S rRNA U1664	7880	5,48	3,40E-06
SNORA5B	ENSG00000200656	H/ACA	18S rRNA U1238	111831	8,11	2,05E-19
SNORA63	ENSG00000199552	H/ACA	28S rRNA U4390	947	4,26	1,21E-05
SNORA65	ENSG00000201302	H/ACA	28S rRNA U4373 28S rRNA U4427	67620	4,53	2,72E-06
SNORA7	ENSG00000201316	H/ACA	28S rRNA U4393	398	6,06	1,75E-09
SNORA7	ENSG00000206913	H/ACA	28S rRNA U4393	808	5,98	5,32E-09
SNORA70	ENSG00000207274	H/ACA	18S rRNA U1692	62407	8,54	1,61E-17
SNORA70	ENSG00000206637	H/ACA	18S rRNA U1692	1699	8,34	3,33E-14
SNORA71	ENSG00000201393	H/ACA	18S rRNA U406	462	9,33	3,25E-10
SNORA72	ENSG00000207084	H/ACA	5.8S rRNA U55	21790	10,33	4,44E-27
SNORA72	ENSG00000252158	H/ACA	5.8S rRNA U55	295	8,72	2,53E-14
SNORA72	ENSG00000201898	H/ACA	5.8S rRNA U55	5968	5,35	3,25E-10
SNORA75	ENSG00000212440	H/ACA	18S rRNA U93	488	7,85	9,05E-14
SNORA77	ENSG00000221643	H/ACA	18S rRNA U814	209411	4,51	1,40E-05

<b>SNORA8</b>	ENSG00000207100	H/ACA	18S rRNA U1056 18S rRNA U1081 28S rRNA U3832	1504	5,60	8,65E-07
<b>SNORD69</b>	ENSG00000212452	C/D	28S rRNA G4464	3455	4,53	8,71E-05
<b>U105C</b>	ENSG00000199212	C/D	18S rRNA U799	475	9,36	1,11E-11

rRNA, ARN ribosomal; log<sub>2</sub>FC, *log2FoldChange*; padj, valeur *p* ajustée

### **Les snoRNA sélectionnés permettent de discriminer entre les HGSC et les SBT séquencés par RNA-Seq**

Pour visualiser l'expression des 36 snoRNA candidats dans les tissus HGSC et SBT, un *heatmap* a été produit à partir des logarithmes régularisés de DESeq2. La fonction d'agrégation (*clustering*) permet également de confirmer si l'expression des snoRNA candidats permet de distinguer les deux groupes de tissus. Comme on le voit à la figure 12, ils sont effectivement séparés par l'agrégation hiérarchique.



**Figure 12 : Heatmap des 36 snoRNA candidats dans les tissus HGSC et SBT séquencés par RNA-Seq**

## L'expression des snoRNA candidats par qRT-PCR dans des tissus HGSC et SBT indépendants montrent une hétérogénéité

11 des 36 snoRNA candidats comme significativement surexprimés dans les HGSC par séquençage de l'ARN (RNA-Seq) ont été séquencés par qRT-PCR dans des tissus indépendants de ceux utilisés pour le RNA-Seq (5 HGSC et 4 SBT). Les snoRNA candidats qui avaient une identité de séquences trop similaire avec un autre paralogue empêchant le design d'une amorce spécifique ont été exclus. Les résultats du qRT-PCR des 11 candidats testés sont présentés dans le tableau 8. Seulement SNORA2 (ENSG00000206647) montre une différence significative par qRT-PCR. Pour les autres snoRNA testés, on remarque que les déviations standards sont importantes par rapport à l'expression relative, témoignant une hétérogénéité d'expression au sein même de chacun des 2 groupes de tissus.

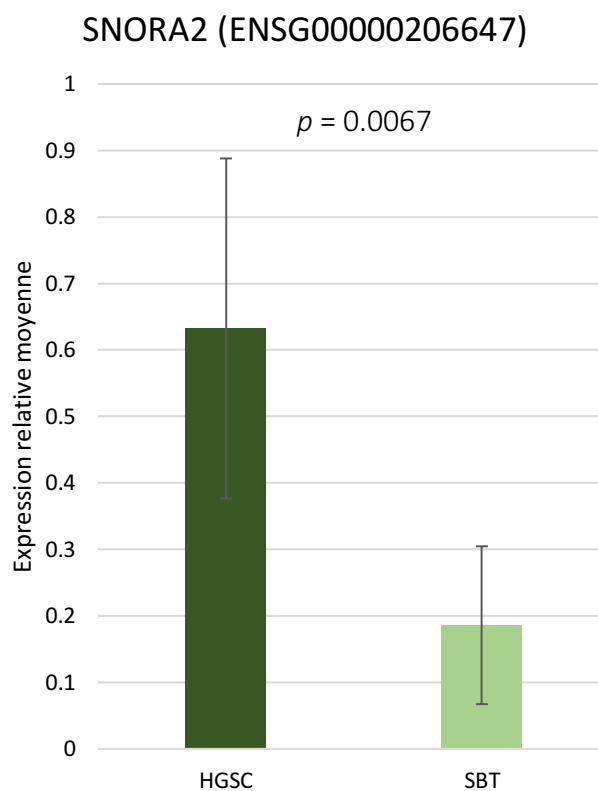
**Tableau 8 : Expression relative de 11 snoRNA candidats par qRT-PCR dans des tissus indépendants (5 HGSC et 4 SBT)**

snoRNA	Identifiant Ensembl	ER moyenne HGSC	ER moyenne SBT	SD HGSC	SD SBT	Test t unilatéral
SNORA16B	ENSG00000201544	0,577	1,053	0,328	0,527	0,093
SNORA19	ENSG00000207468	0,576	0,546	0,269	0,171	0,422
SNORA2	ENSG00000206647	0,632	0,186	0,256	0,119	0,007
SNORA2	ENSG00000199959	26,74	1,311	55,35	0,821	0,181
SNORA30B	ENSG00000202189	2,08	2,71	1,78	1,65	0,310
SNORA44	ENSG00000252840	1,214	1,185	0,799	0,624	0,060
SNORA46	ENSG00000212371	0,567	0,666	0,305	0,227	0,311
SNORA51	ENSG00000206898	0,550	0,498	0,348	0,138	0,386
SNORA56	ENSG00000206693	0,647	0,804	0,363	0,401	0,343
SNORA5B	ENSG00000200656	0,446	0,758	0,373	0,220	0,105
SNORA65	ENSG00000201302	1,674	4,36	0,776	2,90	0,085

ER, Expression relative; SD, déviation standard

## SNORA2 permet de discriminer entre les HGSC et les SBT par qRT-PCR

La quantification par RT-PCR de SNORA2 (ENSG00000206647) montre une différence significative entre les HGSC et les SBT, dans des tissus indépendants de ceux séquencés par RNA-Seq (figure 13). Nous avons donc un snoRNA candidat significativement surexprimé dans les HGSC avec deux méthodes différentes et deux groupes de tissus indépendants.



**Figure 13 : Expression relative moyenne par qRT-PCR de SNORA2 (ENSG00000206647) dans 5 tissus HGSC et 4 tissus SBT**

## DISCUSSIONS ET CONCLUSIONS

### **Une nouvelle méthode de séquençage permet de détecter les snoRNA dans des tissus humains**

La présente étude constitue la première application de cette nouvelle méthode de séquençage dans le cancer de l'ovaire. Il s'agit également de la première caractérisation globale de l'expression des snoRNA dans le cancer de l'ovaire. En effet, la préparation des bibliothèques pour le séquençage avec TGIRT et l'application de CoCo pour l'analyse des données de séquençage permet d'obtenir une vue d'ensemble représentative de cette famille d'ARN précédemment mal représentée dans les séquençages à haut débit traditionnels (Veneziano *et al*, 2016). Les corrélations avec le qRT-PCR, méthode de référence, obtenues dans cette étude (tableau 6) sont comparables à celles obtenues dans (Boivin *et al*, 2018) à la figure supplémentaire S7, qui montre que cette méthode de séquençage est supérieure aux méthodes traditionnelles, assurant ainsi une bonne confiance dans les résultats de séquençage obtenus.

Quant à la distribution des types de snoRNA, des séquençages de type TGIRT de lignées cellulaires, dont les SKOV3.ip1, une lignée de cancer de l'ovaire, indiquent une proportion similaire de snoRNA de type C/D et H/ACA parmi les gènes détectés (Boivin *et al*, 2018). On obtient également une proportion d'environ 50% pour chaque type de snoRNA dans nos échantillons de tissus SBT, tandis que les HGSC présentent une proportion plus importante de snoRNA H/ACA (figure 8), concordant avec les résultats montrant une surexpression significative de ce type de snoRNA dans les HGSC alors que les C/D sont globalement sous-exprimés (figure 9). Cette première étape de découverte de biomarqueur est primordiale pour que les analyses subséquentes soient valables.

L'hétérogénéité dans la carcinogénèse du cancer séreux de l'ovaire est considérée comme un aspect majeur complexifiant la découverte de biomarqueurs (Lee *et al*, 2013). Le séquençage à haut débit de l'ARN est proposé comme stratégie optimale pour obtenir une vue d'ensemble des événements différentiels caractérisant les HGSC afin d'avoir une meilleure compréhension de cette maladie mortelle grâce à une sous-catégorisation des types de HGSC



et au développement de biomarqueurs de dépistage. Ainsi, la méthode de séquençage avec TGIRT et CoCo utilisée dans ce projet est tout indiquée pour répondre à cette problématique.

Les études précédentes sur les snoRNA dans le cancer utilisent principalement des méthodes comme le qRT-PCR et les micropuces, qui peuvent seulement étudier un nombre restreint d'ARN à la fois et choisis subjectivement. Ces méthodes ne sont donc pas optimales pour la recherche de biomarqueurs. De surcroît, les études sur les snoRNA dans le cancer qui utilisent le séquençage à haut débit avec les méthodes traditionnelles risquent d'avoir des biais par rapport aux résultats d'abondance et donc d'expression différentielle des snoRNA, comme montré dans (Deschamps-Francoeur *et al*, 2018) et (Boivin *et al*, 2018). Ainsi, de potentiels biomarqueurs peuvent avoir été manqués dans ces études. De plus, de nombreuses études se sont intéressées aux gènes hôtes des snoRNA. Or, on sait maintenant que les méthodes d'analyse traditionnelles attribuent faussement l'abondance des gènes imbriqués au gène hôte (Deschamps-Francoeur *et al*, 2018). De plus, une déplétion ou une surexpression des gènes hôtes affectent tous les gènes imbriqués, donc on ne peut pas savoir le rôle précis des snoRNA individuels dans les études leur attribuant un phénotype ou une voie de signalisation. Les résultats de ces études ont donc une signification discutable.

### **36 snoRNA surexprimés dans le cancer de l'ovaire découverts par RNA-Seq**

Le séquençage à haut débit de l'ARN et l'étude de l'expression différentielle dans des tissus ovariens a permis de détecter 36 snoRNA comme étant significativement surexprimés et suffisamment abondants dans les HGSC. Avec la méthode de séquençage utilisée, ceux-ci représentent la famille d'ARN la plus abondante parmi les ARN significativement surexprimés et abondants dans les HGSC. Ceci confirme donc que cette famille d'ARN présente un intérêt particulier dans la recherche de biomarqueurs.

Parmi les snoRNA surexprimés découverts, on retrouve des candidats qui ont déjà été décrits dans la littérature comme dérégulés dans différents cancers. Entre autre, SNORA42 a été étudié dans le cancer du poumon (Mei *et al*, 2012), de la prostate (Yi *et al*, 2018) et du côlon (Okugawa *et al*, 2017). Ces études ont également montré le rôle de SNORA42 dans la

tumorigenèse grâce à des analyses phénotypiques de prolifération, de migration et d'invasion. U105B a également été découvert comme surexprimé dans le cancer de l'estomac (Zhang *et al*, 2018). En plus des analyses phénotypiques, des analyses mécanistiques démontrent un lien dans la voie C-Myc par une liaison spécifique à la protéine ALDOA. Ils ont également pu démontrer que U105B était détectable et surexprimé dans le sang des gens atteints du cancer de l'estomac, offrant ainsi une méthode d'échantillonnage pratique pour une application clinique de recherche de biomarqueurs. SNORA65 a quant à lui été décrit parmi une signature de snoRNA pour l'améloblastome de la mâchoire (Davarian *et al*, 2017).

2 snoRNA détectés comme candidats, SNORA16B et SNORA51, sont orphelins, c'est-à-dire qu'ils n'ont pas de fonction canonique connue, soit la modification de l'ARN ribosomal. Des fonctions non-canoniques émergent de plus en plus pour les snoRNA, qu'ils soient orphelins ou comme fonction supplémentaire à leur fonction canonique (Dupuis-Sandoval *et al*, 2015). Or, il n'y a pas à ce jour de fonction proposée pour ces deux snoRNA orphelins, surexprimés et abondants dans le cancer de l'ovaire. Leur étude dans le cancer de l'ovaire pourrait donc éventuellement permettre de découvrir des voies dans lesquelles ils sont impliqués et de leur trouver de nouvelles fonctions non-canoniques. Il est également possible de découvrir de nouvelles fonctions supplémentaires aux fonctions canoniques des autres snoRNA candidats.

### **SNORA2, un snoRNA surexprimé dans le cancer de l'ovaire détecté par qRT-PCR**

SNORA2 est un snoRNA candidat par RNA-Seq qui a également été confirmé par qRT-PCR sur des tissus ovariens indépendants. Il constitue donc un biomarqueur potentiel pouvant à lui seul distinguer un cancer séreux ovarien de haut grade (HGSC). Il a une fonction canonique connue, soit la pseudouridilation de l'ARN ribosomal 28S aux positions 4263 et 4282. Il n'y a pas de littérature sur ce snoRNA à propos du cancer ou d'autres fonctions non-canoniques supplémentaires. Des expériences supplémentaires, discutées plus loin, pourront permettre de valider ce biomarqueur sur plus de tissus et de cibler son rôle dans la carcinogénèse.

Le faible nombre de snoRNA candidats par RNA-Seq détectés également comme significativement surexprimés dans les HGSC par qRT-PCR peut s'expliquer par l'hétérogénéité des tissus cancéreux ovariens. En effet, la variance de l'expression est importante parmi les tissus testés (tableau 8). Un plus grand échantillon de tissus pourrait permettre de détecter des différences significatives. De plus, l'utilisation de plusieurs snoRNA comme signature au lieu d'un seul pourrait permettre de classer les tissus entre HGSC et SBT.

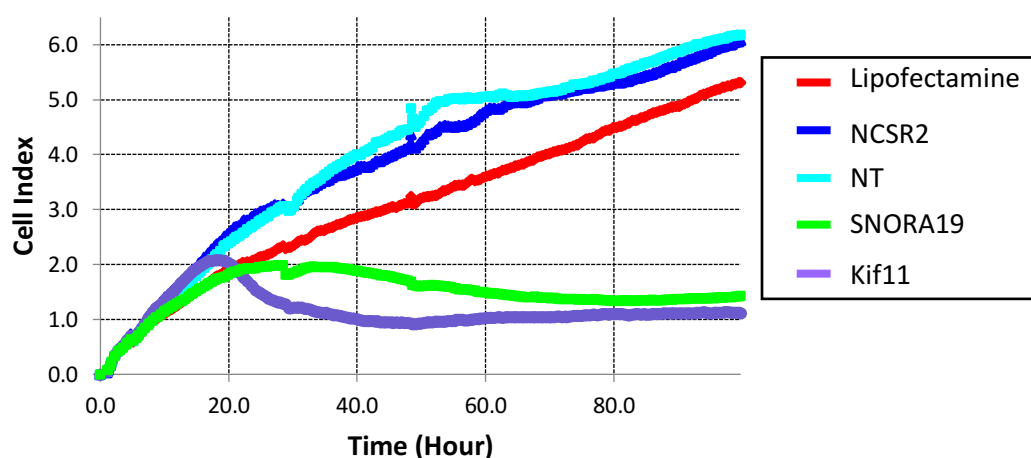
De plus, l'analyse subséquente des résultats du qRT-PCR a montré des discordances au sein des résultats de RNA-Seq. Cette méthode de quantification utilisant tout comme le RNA-Seq une rétrotranscriptase, il a été essayé d'utiliser TGIRT au lieu de la rétrotranscriptase, pour les mêmes avantages que pour le RNA-Seq présentés en introduction. Une fois la méthode de qRT-PCR optimisée, les snoRNA candidats pourront être testés de nouveau par qRT-PCR et un plus grand nombre de snoRNA avec une expression différentielle significative par qRT-PCR pourrait être obtenu.

### **Rôle des snoRNA dans le cancer**

Différentes hypothèses existent quant aux rôles des snoRNA dans le cancer. Étant donné que les cellules cancéreuses ont une activité mitotique et métabolique plus élevée, il a été proposé que ceci entraîne une augmentation de la production de ribosomes, et donc une augmentation des snoRNA car ces derniers participent à la ribogénèse (Montanaro *et al*, 2008). Or, la présente étude montre que les snoRNA C/D, dont ceux qui ont une fonction canonique, sont globalement sous-exprimés dans les HGSC (figure 9). Une autre hypothèse rapportée suggère que la modulation de snoRNA spécifiques permet de contrôler différemment les modifications aux différentes positions sur l'ARNr et ainsi permet un contrôle différentiel de la traduction au bénéfice des protéines oncogènes (Ruggero et Pandolfi, 2003) (Belin *et al*, 2009). Ceci pourrait expliquer pourquoi certains snoRNA sont significativement sous-exprimés dans nos données de séquençage de HGSC. Une autre possibilité émergente est la présence de fonctions non-canoniques pour les snoRNA orphelins ou qui viendraient s'ajouter aux fonctions canoniques. Certaines études citées plus tôt en introduction font

effectivement la démonstration de l'implication de snoRNA dérégulés dans différents cancers dans différentes voies de signalisation, suggérant des rôles distincts de la modification de l'ARNr.

Dans le cadre de ce projet, une fois des snoRNA surexprimés dans le cancer détectés, l'étude de leur rôle dans le développement du cancer est intéressante pour, d'une part, mieux comprendre leurs fonctions générales et, d'autre part, identifier des voies ayant un potentiel thérapeutique. Des analyses phénotypiques préliminaires ont été effectuées avec SNORA19 dans des lignées cancéreuses ovariennes. La déplétion de ce snoRNA avec des ASO entrave la croissance cellulaire, suggérant que la seule surexpression de ce snoRNA dans les cellules cancéreuses peut leur conférer un avantage de croissance, une condition nécessaire à la tumorigénèse (figure 14). Ainsi, même si SNORA19 n'a pas été validé par qRT-PCR, ceci suggère un rôle dans la carcinogénèse malgré une expression hétérogène dans les tissus cancéreux ovariens.



**Figure 14 : Courbe de croissance par mesure de l'impédance par xCELLingence de lignées cellulaires SKOV3.ip1 avec déplétion de SNORA19**

Lipofectamine : cellules transfectées avec la Lipofectamine seulement (contrôle négatif)

NCSR2 : cellules transfectées avec un ASO de séquence aléatoire (contrôle négatif)

NT : cellules non-traitées (contrôle négatif)

SNORA19 : cellules transfectées avec un ASO contre la séquence de SNORA19

Kif11 : cellules transfectées avec Kif11 (contrôle positif)

D'autres expériences sont également en cours afin d'élucider les rôles des snoRNA surexprimés dans le cancer de l'ovaire. Des tests de prolifération et d'apoptose seront réalisés avec la déplétion de snoRNA. Les analyses phénotypiques seront également réalisées dans d'autres lignées cellulaires ovariennes et de types tissulaires différents afin de déterminer si les effets observés sont spécifiques au cancer ou bien au type tissulaire. D'autre part, afin de comprendre si les mécanismes des snoRNA sont dus à leur fonction canonique ou bien à des fonctions non-canoniques supplémentaires, leur effet sur la biogénèse des ribosomes sera étudié grâce à des expériences de *Northern blot*.

### **Prochaines étapes de validation de biomarqueurs dans le cancer de l'ovaire**

Tel que décrit à la section « Méthodologie de la recherche de biomarqueurs » (figure 3), d'autres étapes sont nécessaires pour valider les snoRNA candidats comme potentiels biomarqueurs, soit la validation de ces snoRNA par qRT-PCR dans un plus grand nombre de tissus ovariens puis les tests à l'aveugle des snoRNA sélectionnés pour confirmer leur capacité d'identifier les HGSC. Des tissus sont déjà disponibles pour ces étapes de validation (tableau 9).

**Tableau 9 : Description des tissus ovariens disponibles pour la validation et les tests à l'aveugle des biomarqueurs potentiels**

<b>Nom du tissu</b>	<b>Provenance</b>	<b>Type histologique</b>	<b>Grade</b>
<b>OVC 265</b>	FRQS	Séreux	HGSC
<b>VOA485</b>	GPEC	Séreux	SBT
<b>VOA535</b>	GPEC	Séreux	SBT
<b>E00230</b>	OTB	Séreux	SBT
<b>D00921</b>	OTB	Séreux	SBT
<b>B00840</b>	OTB	Séreux	SBT
<b>OSC 329 BT</b>	FRQS	Séreux	HGSC
<b>OVC 341</b>	FRQS	Séreux	HGSC
<b>OSC 382 BT</b>	FRQS	Séreux	HGSC
<b>OSC 392 BT</b>	FRQS	Séreux	HGSC
<b>OSC 446 BT</b>	FRQS	Séreux	HGSC
<b>OSC 543 BT</b>	FRQS	Séreux	HGSC
<b>A01360105</b>	OTB	Séreux	HGSC
<b>D0070110</b>	OTB	Séreux	HGSC
<b>OSC 456 BT</b>	FRQS	Séreux	HGSC
<b>OVC328 FT</b>	FRQS	Séreux	HGSC
<b>VOA 130</b>	GPEC	Mucineux	Borderline

<b>VOA 257</b>	GPEC	Mucineux	Borderline
<b>VOA 422</b>	GPEC	Mucineux	Borderline
<b>VOA 595</b>	GPEC	Mucineux	Borderline
<b>RT491</b>	ARTB	Mucineux	II
<b>VOA 321</b>	GPEC	Mucineux	III
<b>VOA 439</b>	GPEC	Mucineux	II
<b>E00215105</b>	OTB	À cellules claires	III
<b>A00045101</b>	OTB	À cellules claires	III
<b>D00816107</b>	OTB	À cellules claires	III
<b>D00831107</b>	OTB	À cellules claires	III
<b>E00548102</b>	OTB	À cellules claires	II
<b>D00175107</b>	OTB	À cellules claires	II
<b>D00528107</b>	OTB	À cellules claires	II
<b>B00096101</b>	OTB	Endométrioïde	I
<b>B00334101</b>	OTB	Endométrioïde	I
<b>D00552107</b>	OTB	Endométrioïde	I
<b>D01034115</b>	OTB	Endométrioïde	I
<b>OVC 410</b>	FRQS	Endométrioïde	II
<b>RT234</b>	ARTB	Endométrioïde	II
<b>E00157101</b>	OTB	Endométrioïde	II
<b>OVC 17</b>	FRQS	Endométrioïde	III
<b>MT114</b>	ARTB	Endométrioïde	III
<b>RT591</b>	ARTB	Endométrioïde	III
<b>RT385</b>	ARTB	Endométrioïde	III
<b>B00536101</b>	OTB	Endométrioïde	III
<b>B00679101</b>	OTB	Endométrioïde	III
<b>OVC189</b>	FRQS	Séreux	HGSC
<b>OVC203 FT</b>	FRQS	Séreux	HGSC
<b>B01582103</b>	OTB	Séreux	HGSC
<b>B01183101</b>	OTB	Séreux	HGSC
<b>B01772</b>	OTB	Séreux	HGSC
<b>D03608</b>	OTB	Séreux	HGSC
<b>B02207</b>	OTB	Séreux	HGSC
<b>D03009</b>	OTB	Séreux	HGSC
<b>D02760</b>	OTB	Séreux	HGSC
<b>B02047</b>	OTB	Séreux	HGSC
<b>B01517</b>	OTB	Séreux	HGSC
<b>B01669</b>	OTB	Séreux	HGSC
<b>B02522</b>	OTB	Séreux	HGSC
<b>D00104</b>	OTB	Séreux	HGSC
<b>D02591</b>	OTB	Séreux	HGSC
<b>B03842</b>	OTB	Séreux	HGSC
<b>D03467</b>	OTB	Séreux	HGSC
<b>B03501</b>	OTB	Séreux	HGSC
<b>B04069</b>	OTB	Séreux	HGSC
<b>D00642</b>	OTB	Séreux	HGSC
<b>B01562</b>	OTB	Séreux	HGSC
<b>B01778</b>	OTB	Séreux	HGSC
<b>B03467</b>	OTB	Séreux	HGSC

<b>B01238</b>	OTB	Séieux	HGSC
<b>B02497</b>	OTB	Séieux	HGSC
<b>B01170</b>	OTB	Séieux	SBT
<b>B01593</b>	OTB	Séieux	SBT
<b>B03135</b>	OTB	Séieux	SBT

OTB : *Ontario Tumor Bank*

ARTB : *Alberta Research Tumor Bank*

FRQS : Fonds de recherche du Québec – Santé

GPEC: *The Genetic Pathology Evaluation Centre, University of British Columbia*

Grâce aux résultats du RNA-Seq et d'association par qRT-PCR réalisés dans ce projet de recherche, il a été possible d'estimer les moyennes d'expression et les écarts-types. Pour s'assurer d'une haute probabilité de détection, le niveau de significativité a été établi à 1% et la puissance à 95% aux deux étapes de validation et de test à l'aveugle par qRT-PCR, ce qui résulte en un besoin de 15 tissus par groupe (SBT et HGSC) à chacune de ces étapes.

Le nombre de tissus nécessaires pour la validation par qRT-PCR a été déterminé par un calcul de la puissance avec R :

$$nB=2*(sd*(qnorm(1-alpha/2)+qnorm(1-beta))/(muA-muB))^2$$

où :

nB = Puissance

sd = écart-type

qnorm = fonction R déterminant la cote Z au quantile spécifié d'une distribution normale

muA = moyenne du groupe HGSC

muB = moyenne du groupe SBT

Comme on a pu constater avec les résultats de qRT-PCR à l'étape d'association que les tissus cancéreux ovariens présentent une hétérogénéité d'expression des snoRNA testés, il est possible qu'un seul snoRNA ne permette pas de bien détecter les HGSC. Cependant, étant donné qu'il y a plusieurs snoRNA potentiels découverts par RNA-Seq, une signature de plusieurs snoRNA pourrait permettre de détecter les HGSC. Plusieurs études sur la recherche d'ARN biomarqueurs utilisent effectivement une signature allant de 2 à une vingtaine d'ARN (Liao *et al*, 2010) (Ren *et al*, 2016) (Yang *et al*, 2018).

## **Défis de la recherche de biomarqueurs pour le cancer de l’ovaire**

Étant donné que la recherche de biomarqueurs a été effectuée sur des échantillons cliniques d’ovaires chez des patientes ayant déjà reçu un diagnostic de masse ovarienne, il se peut que les biomarqueurs trouvés ne soient pas valables au niveau pré-diagnostique, donc pour le dépistage (Jacobs et Menon, 2011). Il faudra donc par la suite valider ces biomarqueurs par des essais cliniques sur des échantillons sains et cancéreux accessibles comme le sang puis procéder à des études prospectives à plus grande échelle pour en arriver à un biomarqueur de dépistage (Pepe *et al*, 2001). Une autre limitation possible réside dans les tissus disponibles comparés, soit les cancers sévères de haut grade (HGSC) et les tumeurs de potentiel malin limité (SBT). En effet, bien que les SBT possèdent un profil génétique similaire aux cellules ovariennes saines (Malpica et Wong, 2016), il ressort que les HSGC proviendraient de cellules précancéreuses des trompes de Fallope (Perets et Drapkin, 2016). Les tumeurs SBT et HGSC semblent en effet se développer par des processus distincts et se présentent à des stades différents (Yemelyanova *et al*, 2008). Il n’est donc pas possible d’affirmer à ce stade-ci que les biomarqueurs distinguant ces deux entités permettront de dépister le HGSC plus tôt. Il serait donc pertinent d’inclure également des tissus de trompes de Fallope saines pour tester les biomarqueurs découverts. Il reste que les tissus provenant de tumeurs SBT demeurent une source plus accessible de matériel et un modèle acceptable de comparaison pour la recherche de biomarqueurs.

Malgré ces limitations, les biomarqueurs peuvent avoir d’autres utilités que le dépistage, comme les biomarqueurs diagnostiques pouvant aider à distinguer une tumeur bénigne d’une tumeur maligne, ou encore les biomarqueurs prédictifs pouvant aider à orienter le traitement. Différentes applications des biomarqueurs potentiels découverts peuvent donc être possibles, selon les résultats subséquents.

## **Applications futures des snoRNA biomarqueurs**

Les applications éventuelles de la découverte de snoRNA différenciellement exprimés vont au-delà du dépistage du cancer. En effet, de façon similaire aux déplétions avec les ASO effectuées sur les lignées cellulaires dans ce projet, des médicaments à base d’ASO contre des ARN commencent à être développés comme agents anti-cancer (MacLeod et Crooke,



2017). Ainsi, la caractérisation des snoRNA surexprimés dans les HGSC et ayant un rôle dans la tumorigénèse pourront servir à développer des thérapies ciblées. Sachant que le cancer de l’ovaire devient classiquement résistant à la chimiothérapie conventionnelle, ces opportunités de nouveaux traitements sont d’autant plus intéressantes.

## **Conclusions**

Ce projet constitue le tout début de la recherche de snoRNA biomarqueurs pour le cancer de l’ovaire. La problématique majeure du cancer de l’ovaire réside dans la difficulté de le diagnostiquer à un stade précoce, la découverte de biomarqueurs de dépistage aiderait grandement à diminuer le fardeau de cette maladie, qui possède un taux de survie de 90% à 5 ans lorsque diagnostiqué à un stade précoce. Cependant, les biomarqueurs actuels n’ont pas pu améliorer la survie des patientes jusqu’à maintenant. Les snoRNA sont une nouvelle source de bons candidats potentiels par leur abondance dans les cellules et leurs fonctions émergentes dans la carcinogénèse, et l’amélioration des méthodes de séquençage avec TGIRT et CoCo permet de diminuer les biais de détection et d’analyse de cette famille d’ARN. Ce projet a donc permis d’identifier SNORA2 comme un snoRNA H/ACA surexprimé dans les HGSC, validé par séquençage à haut débit et par qRT-PCR. Des analyses supplémentaires permettront d’optimiser la méthode de qRT-PCR avec TGIRT au lieu de la rétrotranscriptase classique, puis de compléter la validation des snoRNA biomarqueurs candidats dans un plus grand nombre de tissus. Un autre aspect prometteur sera l’étude des fonctions de ces snoRNA dans la carcinogénèse, afin d’en apprendre plus sur le développement du cancer mais aussi sur les fonctions des snoRNA en soi. Le séquençage à haut débit de l’ARN est une méthode de choix pour la détection à grande échelle de biomarqueurs. L’utilisation de méthodes permettant une détection adéquate simultanée de toutes les familles d’ARN permettrait de profiter pleinement du matériel de tissus humains, une ressource limitée et précieuse, en produisant des ensembles de données utilisables par différents groupes de chercheurs selon leurs intérêts et expertise. La recherche de biomarqueurs en sera alors facilitée et permettra de meilleurs soins aux patients.

## LISTE DES RÉFÉRENCES

Alagoz T, Buller RE, Berman M, Anderson B, Manetta A, DiSaia P (1994) What is a normal CA125 level? *Gynecol Oncol*;53(1):93-7.

Andrews L, Mutch DG (2017) Hereditary Ovarian Cancer and Risk Reduction. *Best Practice & Research Clinical Obstetrics & Gynaecology*;41:31-48.

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Accessible à : <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.

Auersperg N (2013) Ovarian surface epithelium as a source of ovarian cancers: unwarranted speculation or evidence-based hypothesis? *Gynecol Oncol*;130(1):246-51.

Bachellerie J-P, Cavallé J, Hüttenhofer A (2002) The expanding snoRNA world. *Biochimie*;84(8):775-90.

Bai B, Laiho M (2016) Deep Sequencing Analysis of Nucleolar Small RNAs: RNA Isolation and Library Preparation. *Methods in molecular biology (Clifton, NJ)*;1455:231-41.

Balakin AG, Smith L, Fournier MJ (1996) The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell*;86(5):823-34.

Barakat RR BA, Markman M, Randall ME (2009) *Principles and Practice of Gynecologic Oncology*. Baltimore, USA: Lippincott Williams and Wilkins.

Barakat RR, Federici MG, Saigo PE, Robson ME, Offit K, Boyd J (2000) Absence of premalignant histologic, molecular, or cell biologic alterations in prophylactic oophorectomy specimens from BRCA1 heterozygotes. *Cancer*;89(2):383-90.

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*;116(2):281-97.

Baruzzo G, Hayer KE, Kim E, Camillo B, FitzGerald GA, Grant GR (2016) Simulation-based comprehensive benchmarking of RNA-seq aligners. *Nature Methods*;14(2).

Bast RC, Jr. (2004) Early detection of ovarian cancer: new technologies in pursuit of a disease that is neither common nor rare. *Transactions of the American Clinical and Climatological Association*;115:233-47; discussion 47-8.

Bast RC, Jr., Feeney M, Lazarus H, Nadler LM, Colvin RB, Knapp RC (1981) Reactivity of a monoclonal antibody with human ovarian carcinoma. *The Journal of clinical investigation*;68(5):1331-7.

Bastani A, Asghary A, Heidari MH, Karimi-Busheri F (2017) Evaluation of the sensitivity and specificity of serum level of prostatic acid phosphatase, CA125, LDH, AFP, and hCG+beta in epithelial ovarian cancer patients. *Eur J Gynaecol Oncol*;38(3):418-24.

Baxter-Roshek JL, Petrov AN, Dinman JD (2007) Optimization of ribosome structure and function by rRNA base modification. *PLoS one*;2(1).

Belin S, Beghin A, Solano-Gonzalez E, Bezin L, Brunet-Manquat S, Textoris J, et al. (2009) Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. *PLoS One*;4(9):e7147.

Beral V, Gaitskell K, Hermon C, Moser K, Reeves G, Peto R (Collaborative Group on Epidemiological Studies of Ovarian Cancer) (2015) Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. *Lancet*;385(9980):1835-42.

Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical pharmacology and therapeutics*;69(3):89-95.

Blaustein A, Kurman RJ (2011) *Blaustein's pathology of the female genital tract*. 6th ed. New York, NY: Springer; xv, 1246 p. p.

Boivin V, Deschamps-Francoeur G, Couture S, Nottingham RM, Bouchard-Bourelle P, Lambowitz AM, et al. (2018) Simultaneous sequencing of coding and noncoding RNA reveals a human transcriptome dominated by a small number of highly expressed noncoding genes. *RNA (New York, NY)*;24(7):950-65.

Boivin V, Deschamps-Francoeur G, Scott MS (2018) Protein coding genes as hosts for noncoding RNA expression. *Seminars in cell & developmental biology*;75:3-12.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*;30(15):2114-20.

Bousquet-Antonelli C, Henry Y, G'Elugne JP, Caizergues-Ferrer M, Kiss T (1997) A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *The EMBO journal*;16(15):4770-6.

Brimacombe R, Mitchell P, Osswald M, Stade K, Bochkariov D (1993) Clustering of modified nucleotides at the functional center of bacterial ribosomal RNA. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*;7(1):161-7.

Bristow RE, Tomacruz RS, Armstrong DK, Trimble EL, Montz FJ (2002) Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: a meta-analysis. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*;20(5):1248-59.

Brousseau J-PP, Lucier J-FF, Lapointe E, Durand M, Gendron D, Gervais-Bird J, et al. (2010) High-throughput quantification of splicing isoforms. *RNA (New York, NY)*;16(2):442-9.

Brown PO, Palmer C (2009) The preclinical natural history of serous ovarian cancer: defining the target for early detection. *PLoS medicine*;6(7).

Buhling KJ, Lezon S, Eulenburg C, Schmalfeldt B (2017) The role of transvaginal ultrasonography for detecting ovarian cancer in an asymptomatic screening population: a systematic review. *Arch Gynecol Obstet*;295(5):1259-68.

Buys SS, Partridge E, Black A, Johnson CC, Lamerato L, Isaacs C, et al. (2011) Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA*;305(22):2295-303.

Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature*;474(7353):609-15.

Carlson JW, Miron A, Jarboe EA, Parast MM, Hirsch MS, Lee Y, et al. (2008) Serous tubal intraepithelial carcinoma: its potential role in primary peritoneal serous carcinoma and serous cancer prevention. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*;26(25):4160-5.

Chan PP, Lowe TM (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic acids research*;37(Database issue):7.

Chen F, Gaitskell K, Garcia MJ, Albukhari A, Tsaltas J, Ahmed AA (2017) Serous tubal intraepithelial carcinomas associated with high-grade serous ovarian carcinomas: a systematic review. *BJOG : an international journal of obstetrics and gynaecology*.

Chohan L, Richardson DL (2019) ACOG Committee Opinion No. 774: Opportunistic Salpingectomy as a Strategy for Epithelial Ovarian Cancer Prevention. *Obstetrics and gynecology*;133(4).

Colombo P-E, Fabbro M, Theillet C, Bibeau F, Rouanet P, Ray-Coquard I (2014) Sensitivity and resistance to treatment in the primary management of epithelial ovarian cancer. *Critical Reviews in Oncology/Hematology*;89(2):207-16.

Comité consultatif de la Société canadienne du cancer (2018) Statistiques canadiennes sur le cancer 2019. Toronto, Société canadienne du cancer. Accessible à : [cancer.ca/Canadian-Cancer-Statistics-2018-FR](http://cancer.ca/Canadian-Cancer-Statistics-2018-FR).

Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. (2016) A survey of best practices for RNA-seq data analysis. *Genome Biology*;17(1):13.

Cortez M, Bueso-Ramos C, Ferdin J, Lopez-Berestein G, Sood AK, Calin GA (2011) MicroRNAs in body fluids—the mix of hormones and biomarkers. *Nature Reviews Clinical Oncology*;8(8):467-77.

Cramer DW, O'Rourke DJ, Vitonis AF, Matulonis UA, Dijohnson DA, Sluss PM, et al. (2010) CA125 immune complexes in ovarian cancer patients with low CA125 concentrations. *Clinical chemistry*;56(12):1889-92.

Crea F, Quagliata L, Michael A, Liu HH, Frumento P, Azad AA, et al. (2016) Integrated analysis of the prostate cancer small-nucleolar transcriptome reveals SNORA55 as a driver of prostate cancer progression. *Molecular oncology*;10(5):693-703.

Davianian H, Balasiddaiah A, Heymann R, Sundström M, Redenström P, Silfverberg M, et al. (2017) Ameloblastoma RNA profiling uncovers a distinct non-coding RNA signature. *Oncotarget*;8(3):4530-42.

de Turrís V, Leva G, Caldarola S, Loreni F, Amaldi F, Bozzoni I (2004) TOP Promoter Elements Control the Relative Ratio of Intron-encoded snoRNA Versus Spliced mRNA Biosynthesis. *Journal of Molecular Biology*;344(2):383-94.

Decatur WA, Fournier MJ (2002) rRNA modifications and ribosome function. *Trends in biochemical sciences*;27(7):344-51.

Dehari R, Kurman RJ, Logani S, Shih I-Me (2007) The development of high-grade serous carcinoma from atypical proliferative (borderline) serous tumors and low-grade micropapillary serous carcinoma: a morphologic and molecular genetic analysis. *The American journal of surgical pathology*;31(7):1007-12.

Deogharia M, Majumder M (2018) Guide snoRNAs: Drivers or Passengers in Human Disease? *Biology*;8(1).

Deschamps-Francoeur G, Boivin V, Elela S, Scott MS (2018) CoCo: RNA-seq Read Assignment Correction for Nested Genes and Multimapped Reads. *bioRxiv*:477869.

Deschamps-Francoeur G, Garneau D, Dupuis-Sandoval F, Roy A, Frappier M, Catala M, et al. (2014) Identification of discrete classes of small nucleolar RNA featuring different ends and RNA binding protein dependency. *Nucleic Acids Res*;42(15):10073-85.

Di Leva G, Garofalo M, Croce CM (2014) MicroRNAs in cancer. *Annual review of pathology*;9:287-314.

Dieci G, Preti M, Montanini B (2009) Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics*;94(2):83-8.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*;29(1):15-21.

Dong X-YY, Rodriguez C, Guo P, Sun X, Talbot JT, Zhou W, et al. (2008) SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. *Human molecular genetics*;17(7):1031-42.

Dupuis-Sandoval F, Poirier M, Scott MS (2015) The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdiscip Rev RNA*;6(4):381-97.

Eisenhauer EL, Abu-Rustum NR, Sonoda Y, Levine DA, Poyner EA, Aghajanian C, et al. (2006) The addition of extensive upper abdominal surgery to achieve optimal cytoreduction improves survival in patients with stages IIIc-IV epithelial ovarian cancer. *Gynecologic oncology*;103(3):1083-90.

Ely KL, Truong M (2017) The Role of Opportunistic Bilateral Salpingectomy vs Tubal Occlusion or Ligation for Ovarian Cancer Prophylaxis. *Journal of Minimally Invasive Gynecology*;24(3):371-8.

Engelen MJ, Kos HE, Willems PH, Aalders JG, de Vries EG, Schaapveld M, et al. (2006) Surgery by consultant gynecologic oncologists improves survival in patients with ovarian carcinoma. *Cancer*;106(3):589-98.

Falaleeva M, Pages A, Matuszek Z, Hidmi S, Agranat-Tamir L, Korotkov K, et al. (2016) Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. *Proceedings of the National Academy of Sciences of the United States of America*;113(12):34.

Fang X, Yang D, Luo H, Wu S, Dong W, Xiao J, et al. (2016) SNORD126 promotes HCC and CRC cell growth by activating the PI3K-AKT pathway through FGFR2. *Journal of molecular cell biology*.

Farazi TA, Hoell JI, Morozov P, Tuschl T (2013) MicroRNAs in human cancer. *Advances in experimental medicine and biology*;774:1-20.

Fathalla MF (1971) Incessant ovulation--a factor in ovarian neoplasia? *Lancet*;2(7716):163.

Filipowicz W, Pogacić V (2002) Biogenesis of small nucleolar ribonucleoproteins. *Current opinion in cell biology*;14(3):319-27.

Flesken-Nikitin A, Hwang C-II, Cheng C-YY, Michurina TV, Enikolopov G, Nikitin AY (2013) Ovarian surface epithelium at the junction area contains a cancer-prone stem cell niche. *Nature*;495(7440):241-5.

Folkins AK, Jarboe EA, Saleemuddin A, Lee Y, Callahan MJ, Drapkin R, et al. (2008) A candidate precursor to pelvic serous cancer (p53 signature) and its prevalence in ovaries and fallopian tubes from women with BRCA mutations. *Gynecologic oncology*;109(2):168-73.

U.S. Preventive Services Task Force (2013) Screening for ovarian cancer: reaffirmation recommendation statement. *American family physician*;87(10).

Freedman JE, Gerstein M, Mick E, Rozowsky J, Levy D, Kitchen R, et al. (2016) Diverse human extracellular RNAs are widely detected in human plasma. *Nature communications*;7:11106.

Gao L, Ma J, Mannoor K, Guarnera MA, Shetty A, Zhan M, et al. (2015) Genome-wide small nucleolar RNA expression analysis of lung cancer by next-generation deep sequencing. *International journal of cancer*;136(6):E623-9.

Gershenson DM, Sun CC, Lu KH, Coleman RL, Sood AK, Malpica A, et al. (2006) Clinical behavior of stage II-IV low-grade serous carcinoma of the ovary. *Obstetrics and gynecology*;108(2):361-8.

Gloss BS, Samimi G (2014) Epigenetic biomarkers in epithelial ovarian cancer. *Cancer letters*;342(2):257-63.

Goeze A, Schlüns K, Wolf G, Thäsler Z, Petersen S, Petersen I (2002) Chromosomal imbalances of primary and metastatic lung adenocarcinomas. *The Journal of pathology*;196(1):8-16.

Goff BA, Mandel LS, Melancon CH, Muntz HG (2004) Frequency of symptoms of ovarian cancer in women presenting to primary care clinics. *JAMA*;291(22):2705-12.

Granato T, Porpora M, Longo F, Angeloni A, Manganaro L, Anastasi E (2015) HE4 in the differential diagnosis of ovarian masses. *Clinica Chimica Acta*;446:147-55.

Helder-Woolderink JM, Blok EA, Vasen HF, Hollema H, Mourits MJ, De Bock GH (2016) Ovarian cancer in Lynch syndrome; a systematic review. *Eur J Cancer*;55:65-73.

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*;8(2):R19.

Hempling RE, Piver MS, Natarajan N, Baker TR, Thompson JM, Hicks ML, et al. (1993) Predictive value of serum CA125 following optimal cytoreductive surgery during weekly cisplatin induction therapy for advanced ovarian cancer. *Journal of surgical oncology*;54(1):38-44.

Henderson JT, Webber EM, Sawaya GF (2018) Screening for Ovarian Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA*;319(6):595-606.

Henras A, Henry Y, Bousquet-Antonelli C, Noaillac-Depeyre J, Gélugne JP, Caizergues-Ferrer M (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *The EMBO Journal*;17(23):7078-90.

Henras AK, Dez C, Henry Y (2004) RNA structure and function in C/D and H/ACA s(no)RNPs. *Current opinion in structural biology*;14(3):335-43.

Henras AK, Plisson-Chastang C, O'Donohue M-FF, Chakraborty A, Gleizes P-EE (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley interdisciplinary reviews RNA*;6(2):225-42.

Hunter SM, Anglesio MS, Ryland GL, Sharma R, Chiew Y-E, Rowley SM, et al. (2015) Molecular profiling of low grade serous ovarian tumours identifies novel candidate driver genes. *Oncotarget*;6(35):37663-77.

Iversen L, Sivasubramaniam S, Lee AJ, Fielding S, Hannaford PC (2017) Lifetime cancer risk and combined oral contraceptives: the Royal College of General Practitioners' Oral Contraception Study. *Am J Obstet Gynecol*;216(6):580 e1- e9.

Jacobs I, Menon U (2011) The sine qua non of discovering novel biomarkers for early detection of ovarian cancer: carefully selected preclinical samples. *Cancer prevention research (Philadelphia, Pa)*;4(3):299-302.

Jacobs IJ, Menon U (2004) Progress and challenges in screening for early detection of ovarian cancer. *Molecular & cellular proteomics : MCP*;3(4):355-66.

Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, et al. (2016) Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet (London, England)*;387(10022):945-56.

Kabawat SE, Bast RC, Jr., Bhan AK, Welch WR, Knapp RC, Colvin RB (1983) Tissue distribution of a coelomic-epithelium-related antigen recognized by the monoclonal antibody OC125. *Int J Gynecol Pathol*;2(3):275-85.

Ke N, Wang X, Xu X, Abassi YA (2011) The xCELLigence system for real-time and label-free monitoring of cell viability. *Methods in molecular biology (Clifton, NJ)*;740:33-43.

Kierzek E, Malgowska M, Lisowiec J, Turner DH, Gdaniec Z, Kierzek R (2014) The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Research*;42(5):3492-501.

Kishore S, Gruber AR, Jedlinski DJ, Syed AP, Jorjani H, Zavolan M (2013) Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. *Genome Biol*;14(5):R45.

Kiss-László Z, Henry Y, Kiss T (1998) Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *The EMBO Journal*;17(3):797-807.

Kiss T (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *The EMBO journal*;20(14):3617-22.

Kiss T (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*;109(2):145-8.

Kiss T, Fayet E, JÁDy BE, Richard P, Weber M (2006) Biogenesis and Intranuclear Trafficking of Human Box C/D and H/ACA RNPs. *Cold Spring Harbor Symposia on Quantitative Biology*;71(0):407-17.



Kiss T, Filipowicz W (1995) Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes & development*;9(11):1411-24.

Köbel M, Kalloger SE, Huntsman DG, Santos JL, Swenerton KD, Seidman JD, et al. (2010) Differences in tumor type in low-stage versus high-stage ovarian carcinomas. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists*;29(3):203-11.

Kumar V, Abbas AK, Aster JC (2015) Grading and staging of neoplasia dans *Robbins and Cotran pathologic basis of disease*, Philadelphie, Elsevier/Saunders (9e éd., 1ère éd. 1944), 332.

Kristjansdottir B, Levan K, Partheen K, Sundfeldt K (2013) Diagnostic performance of the biomarkers HE4 and CA125 in type I and type II epithelial ovarian cancer. *Gynecologic oncology*;131(1):52-8.

Krogh N, Jansson MD, Häfner SJ, Tehler D, Birkedal U, Christensen-Dalsgaard M, et al. (2016) Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic acids research*;44(16):7884-95.

Kuhn E, Kurman RJ, Vang R, Sehdev AS, Han G, Soslow R, et al. (2012) TP53 mutations in serous tubal intraepithelial carcinoma and concurrent pelvic high-grade serous carcinoma--evidence supporting the clonal relationship of the two lesions. *The Journal of pathology*;226(3):421-6.

Kuo K-TT, Guan B, Feng Y, Mao T-LL, Chen X, Jinawath N, et al. (2009) Analysis of DNA copy number alterations in ovarian serous tumors identifies new molecular genetic changes in low-grade and high-grade carcinomas. *Cancer research*;69(9):4036-42.

Kurman RJ, Carcangiu ML, Herrington CS, Young RH (2014) WHO Classification of Tumours of Female Reproductive Organs 4th edn. *IARC Press: Lyon*.

Kurman RJ, Shih I-Me (2010) The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. *The American journal of surgical pathology*;34(3):433-43.

Kurman RJ, Shih I-Me (2011) Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer--shifting the paradigm. *Human pathology*;42(7):918-31.

Labaj PP, Kreil DP (2016) Sensitivity, specificity, and reproducibility of RNA-Seq differential expression calls. *Biol Direct*;11(1):66.

Lafontaine DL, Tollervey D (1999) Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA (New York, NY)*;5(3):455-67.

Lalwani N, Prasad SR, Vikram R, Shanbhogue AK, Huettner PC, Fasih N (2011) Histologic, molecular, and cytogenetic features of ovarian cancers: implications for diagnosis and

treatment. *Radiographics : a review publication of the Radiological Society of North America, Inc*;31(3):625-46.

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*;9(4):357-9.

Lapeyre B (2004) Conserved ribosomal RNA modification and their putative roles in ribosome biogenesis and translation. Dans Grosjean H (Éds) *Fine-Tuning of RNA Functions by Modification and Editing. Topics in Current Genetics*, Berlin, Springer (vol 12), 263-284.

Lee J-YY, Kim HS, Suh DH, Kim M-KK, Chung HH, Song Y-SS (2013) Ovarian cancer biomarker discovery based on genomic approaches. *Journal of cancer prevention*;18(4):298-312.

Lengyel E (2010) Ovarian cancer development and metastasis. *The American journal of pathology*;177(3):1053-64.

Liang XH, Vickers TA, Guo S, Crooke ST (2011) Efficient and specific knockdown of small non-coding RNAs in mammalian cells and in mice. *Nucleic Acids Res*;39(3):e13.

Liao J, Yu L, Mei Y, Guarnera M, Shen J, Li R, et al. (2010) Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. *Molecular cancer*;9(1):198.

Liao Y, Smyth GK, Shi W (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic acids research*;41(10).

Longacre TA, McKenney JK, Tazelaar HD, Kempson RL, Hendrickson MR (2005) Ovarian serous tumors of low malignant potential (borderline tumors): outcome-based study of 276 patients with long-term (> or =5-year) follow-up. *The American journal of surgical pathology*;29(6):707-23.

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*;15(12):550.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. (2005) MicroRNA expression profiles classify human cancers. *Nature*;435(7043):834-8.

MacLeod RA, Crooke ST (2017) RNA Therapeutics in Oncology: Advances, Challenges, and Future Directions. *The Journal of Clinical Pharmacology*;57(S10).

Maden BE (1990) The numerous modified nucleotides in eukaryotic ribosomal RNA. *Progress in nucleic acid research and molecular biology*;39:241-303.

Makarova JA, Ivanova SM, Tonevitsky AG, Grigoriev AI (2013) New functions of small nucleolar RNAs. *Biochemistry (Moscow)*;78(6):638-50.

Malmstrom H, Hogberg T, Risberg B, Simonsen E (1994) Granulosa cell tumors of the ovary: prognostic factors and outcome. *Gynecol Oncol*;52(1):50-5.

Malpica A, Wong KK (2016) The molecular pathology of ovarian serous borderline tumors. *Annals of oncology : official journal of the European Society for Medical Oncology*;27 Suppl 1:i16-i9.

Mann WJ, Patsner B, Cohen H, Loesch M (1988) Preoperative serum CA-125 levels in patients with surgical stage I invasive ovarian adenocarcinoma. *J Natl Cancer Inst*;80(3):208-9.

Marquez RT, Baggerly KA, Patterson AP, Liu J, Broaddus R, Frumovitz M, et al. (2005) Patterns of gene expression in different histotypes of epithelial ovarian cancer correlate with those in normal fallopian tube, endometrium, and colon. *Clinical cancer research : an official journal of the American Association for Cancer Research*;11(17):6116-26.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*;17(1).

Maxwell ES, Fournier MJ (1995) The small nucleolar RNAs. *Annual review of biochemistry*;64:897-934.

Mayr D, Hirschmann A, Löhrs U, Diebold J (2006) KRAS and BRAF mutations in ovarian tumors: a comprehensive study of invasive carcinomas, borderline tumors and extraovarian implants. *Gynecologic oncology*.

Mei YP, Liao JP, Shen J, Yu L, Liu BL, Liu L, et al. (2012) Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. *Oncogene*;31(22):2794-804.

Menon U, Gentry-Maharaj A, Hallett R, Ryan A, Burnell M, Sharma A, et al. (2009) Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *The Lancet Oncology*;10(4):327-40.

Menon U, Griffin M, Gentry-Maharaj A (2014) Ovarian cancer screening--current status, future directions. *Gynecol Oncol*;132(2):490-5.

Metzker ML (2010) Sequencing technologies - the next generation. *Nature reviews Genetics*;11(1):31-46.

Miralles C, Orea M, Espana P, Provencio M, Sanchez A, Cantos B, et al. (2003) Cancer antigen 125 associated with multiple benign and malignant pathologies. *Annals of surgical oncology*;10(2):150-4.

Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proceedings of the National Academy of Sciences of the United States of America*;105(30):10513-8.

Mleczko AM, Bakowska-Zywicka K (2016) When small RNAs become smaller: emerging functions of snoRNAs and their derivatives. *Acta biochimica Polonica*;63(4):601-7.

Mohr S, Ghanem E, Smith W, Sheeter D, Qin Y, King O, et al. (2013) Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA*;19(7):958-70.

Molina R, Escudero JM, Augé JM, Filella X, Foj L, Torné A, et al. (2011) HE4 a novel tumour marker for ovarian cancer: comparison with CA 125 and ROMA algorithm in patients with gynaecological diseases. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*;32(6):1087-95.

Montanaro L, Trere D, Derenzini M (2008) Nucleolus, ribosomes, and cancer. *Am J Pathol*;173(2):301-10.

Moore LE, Pfeiffer RM, Zhang Z, Lu KH, Fung ET, Bast RC (2012) Proteomic biomarkers in combination with CA 125 for detection of epithelial ovarian cancer using prediagnostic serum samples from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. *Cancer*;118(1):91-100.

Morice P, Uzan C, Fauvet R, Gouy S, Duvillard P, Darai E (2012) Borderline ovarian tumour: pathological diagnostic dilemma and risk factors for invasive or lethal recurrence. *The Lancet Oncology*;13(3):15.

Nakamura K, Sawada K, Yoshimura A, Kinose Y, Nakatsuka E, Kimura T (2016) Clinical relevance of circulating cell-free microRNAs in ovarian cancer. *Mol Cancer*;15(1):48.

Nam EJ, Yoon H, Kim SW, Kim H, Kim YT, Kim JH, et al. (2008) MicroRNA expression profiles in serous ovarian carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research*;14(9):2690-5.

National Cancer Institute (2013) Tumor grade. Accessible à : <https://www.cancer.gov/about-cancer/diagnosis-staging/prognosis/tumor-grade-fact-sheet>.

Navaneelan T, Lawrence E (2015) Ovarian cancer: Survival statistics. In: Canada S, editor.

Nottingham RM, Wu DC, Qin Y, Yao J, Hunicke-Smith S, Lambowitz AM (2016) RNA-seq of human reference RNA samples using a thermostable group II intron reverse transcriptase. *RNA*;22(4):597-613.

Nustad K, Bast RC, Brien TJ, Nilsson O, Seguin P, Suresh MR, et al. (1996) Specificity and affinity of 26 monoclonal antibodies against the CA 125 antigen: first report from the ISOBM TD-1 workshop. International Society for Oncodevelopmental Biology and Medicine. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*;17(4):196-219.

O'Leary NA, Wright MW, Brister JR, Ciuffo S, Haddad D, McVeigh R, et al. (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids research*;44(D1):45.

Okugawa Y, Toiyama Y, Toden S, Mitoma H, Nagasaka T, Tanaka K, et al. (2017) Clinical significance of SNORA42 as an oncogene and a prognostic biomarker in colorectal cancer. *Gut*;66(1):107.

Ortiz-Quintero B (2016) Cell-free microRNAs in blood and other body fluids, as cancer biomarkers. *Cell proliferation*;49(3):281-303.

Ovalle WK, Nahirney PC, Netter FH (2013) *Netter's essential histology*. 2nd ed. Philadelphia, PA: Elsevier/Saunders; xv, 517 pages p.

Ovarian Cancer Canada (2013) Contributions to Ovarian Cancer Research in Canada. Toronto. Accessible à : <https://ovariancanada.org/OCC/media/Content/PDFs/Reports/OvarianCancerCanada-ResearchImpact-2013.pdf>.

Palazzo AF, Lee ES (2015) Non-coding RNA: what is functional and what is junk? *Frontiers in genetics*;6:2.

Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. (2001) Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute*;93(14):1054-61.

Perets R, Drapkin R (2016) It's Totally Tubular....Riding The New Wave of Ovarian Cancer Research. *Cancer research*;76(1):10-7.

Piek JM, van Diest PJ, Zweemer RP, Jansen JW, Poort-Keesom RJ, Menko FH, et al. (2001) Dysplastic changes in prophylactically removed Fallopian tubes of women predisposed to developing ovarian cancer. *The Journal of pathology*;195(4):451-6.

Plaxe SC (2008) Epidemiology of low-grade serous ovarian cancer. *American journal of obstetrics and gynecology*;198(4):4590.

Poynter JN, Amatruda JF, Ross JA (2010) Trends in incidence and survival of pediatric and adolescent patients with germ cell tumors in the United States, 1975 to 2006. *Cancer*;116(20):4882-91.

Prahm KP, Novotny GW, Hogdall C, Hogdall E (2016) Current status on microRNAs as biomarkers for ovarian cancer. *APMIS*;124(5):337-55.

Prat J (2012) Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features. *Virchows Archiv*;460(3):237-49.

Prat J (2014) Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics*;124(1):1-5.

R Development Core Team R (2006) A language and environment for statistical computing (version 3.5.3). R Foundation for Statistical Computing, Vienna. Accessible à : <https://www.R-project.org/>

Reichow SL, Hamma T, Ferre-D'Amare AR, Varani G (2007) The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res*;35(5):1452-64.

Ren W, Zhang J, Li W, Li Z, Hu S, Suo J, et al. (2016) A Tumor-Specific Prognostic Long Non-Coding RNA Signature in Gastric Cancer. *Med Sci Monit*;22:3647-57.

Ruggero D, Pandolfi PP (2003) Does the ribosome translate cancer? *Nature reviews Cancer*;3(3):179-92.

Rustin GJ, Nelstrop AE, Tuxen MK, Lambert HE (1996) Defining progression of ovarian carcinoma during follow-up according to CA 125: a North Thames Ovary Group Study. *Annals of oncology : official journal of the European Society for Medical Oncology*;7(4):361-4.

Salomon-Perzyński A, Salomon-Perzyńska M, Michalski B, Skrzypulec-Plinta V (2017) High-grade serous ovarian cancer: the clone wars. *Archives of gynecology and obstetrics*;295(3):569-76.

Sarojini S, Tamir A, Lim H, Li S, Zhang S, Goy A, et al. (2012) Early detection biomarkers for ovarian cancer. *Journal of oncology*;2012:709049.

Seidman JD, Kurman RJ (2000) Ovarian serous borderline tumors: A critical review of the literature with emphasis on prognostic indicators. *Human Pathology*;31(5):539-57.

Seyednasrollah F, Laiho A, Elo LL (2015) Comparison of software packages for detecting differential expression in RNA-seq studies. *Briefings in bioinformatics*;16(1):59-70.

Sherman ME, Mink PJ, Curtis R, Cote TR, Brooks S, Hartge P, et al. (2004) Survival among women with borderline ovarian tumors and ovarian carcinoma: a population-based analysis. *Cancer*;100(5):1045-52.

Shih I-Me, Kurman RJ (2004) Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. *The American journal of pathology*;164(5):1511-8.

Singer G, Stöhr R, Cope L, Dehari R, Hartmann A, Cao D-FF, et al. (2005) Patterns of p53 mutations separate ovarian serous borderline tumors and low- and high-grade carcinomas and provide support for a new model of ovarian carcinogenesis: a mutational analysis with immunohistochemical correlation. *The American journal of surgical pathology*;29(2):218-24.

Siprashvili Z, Webster DE, Johnston D, Shenoy RM, Ungewickell AJ, Bhaduri A, et al. (2016) The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nature genetics*;48(1):53-8.

Sjovall K, Nilsson B, Einhorn N (2002) The significance of serum CA 125 elevation in malignant and nonmalignant diseases. *Gynecol Oncol*;85(1):175-8.

Sölétormos G, Duffy MJ, Hassan S, Verheijen R, Tholander B, Bast RC, et al. (2016) Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer: Updated Guidelines From the European Group on Tumor Markers. *International Journal of Gynecological Cancer*;26(1):43.

Steffensen KD, Waldstrøm M, Brandslund I, Jakobsen A (2011) Prognostic impact of prechemotherapy serum levels of HER2, CA125, and HE4 in ovarian cancer patients. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society*;21(6):1040-7.

Stepanov GA, Filippova JA, Komissarov AB, Kuligina EV, Richter VA, Semenov DV (2015) Regulatory Role of Small Nucleolar RNAs in Human Diseases. *BioMed Research International*;2015:1-10.

Su J, Liao J, Gao L, Shen J, Guarnera MA, Zhan M, et al. (2016) Analysis of small nucleolar RNAs in sputum for lung cancer diagnosis. *Oncotarget*;7(5):5131-42.

Tanaka R, Satoh H, Moriyama M, Satoh K, Morishita Y, Yoshida S, et al. (2000) Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. *Genes to Cells*;5(4):277-87.

Tavassoli FA (1994) Ovarian tumors with functioning manifestations. *Endocrine Pathology*;5(3):137-48.

Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Gayther SA, Apostolidou S, et al. (2009) An epigenetic signature in peripheral blood predicts active ovarian cancer. *PloS one*;4(12).

Thin K, Tu J, Sudheesh C (2019) Long non-coding SNHG1 in cancer. *Clinica Chimica Acta*.

Thorenoor N, Slaby O (2015) Small nucleolar RNAs functioning and potential roles in cancer. *Tumour Biol*;36(1):41-53.

Timmerman D, Van Calster B, Testa A, Savelli L, Fischerova D, Froyman W, et al. (2016) Predicting the risk of malignancy in adnexal masses based on the Simple Rules from the International Ovarian Tumor Analysis group. *American journal of obstetrics and gynecology*;214(4):424-37.

Tingulstad S, Hagen B, Skjeldestad FE, Onsrud M, Kiserud T, Halvorsen T, et al. (1996) Evaluation of a risk of malignancy index based on serum CA125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses. *British journal of obstetrics and gynaecology*;103(8):826-31.

Tycowski KT, You ZH, Graham PJ, Steitz JA (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs. *Molecular cell*;2(5):629-38.

Umu SU, Langseth H, Bucher-Johannessen C, Fromm B, Keller A, Meese E, et al. (2018) A comprehensive profile of circulating RNAs in human serum. *RNA biology*;15(2):242-50.

Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO (2007) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nature cell biology*;9(6):654-9.

Vang R, Shih I-Me, Kurman RJ (2009) Ovarian low-grade and high-grade serous carcinoma: pathogenesis, clinicopathologic and molecular biologic features, and diagnostic problems. *Advances in anatomic pathology*;16(5):267-82.

Veneziano D, Di Bella S, Nigita G, Laganà A, Ferro A, Croce CM (2016) Noncoding RNA: Current Deep Sequencing Data Analysis Approaches and Challenges. *Human mutation*;37(12):1283-98.

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*;10(1):57-63.

Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B (2019) ggplots: Various R Programming Tools for Plotting Data. Accessible à : <https://CRAN.R-project.org/package=ggplots>.

Warton K, Samimi G (2015) Methylation of cell-free circulating DNA in the diagnosis of cancer. *Frontiers in Molecular Biosciences*;2.

Watkins NJ, Gottschalk A, Neuer G, Kastner B, Fabrizio P, Mann M, et al. (1998) Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA*;4(12):1549-68.

Wentzensen N, Poole EM, Trabert B, White E, Arslan AA, Patel AV, et al. (2016) Ovarian Cancer Risk Factors by Histologic Subtype: An Analysis From the Ovarian Cancer Cohort Consortium. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*.

Whiteside TL (2016) Tumor-Derived Exosomes and Their Role in Cancer Progression. *Advances in clinical chemistry*;74:103-41.

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York. Accessible à : <https://ggplot2.tidyverse.org>.

Williams DJ, Boots JL, Hall KB (2001) Thermodynamics of 2'-ribose substitutions in UUCG tetraloops. *RNA (New York, NY)*;7(1):44-53.



Wilson JM, Jungner YG (1968) WHO principles and practice of screening for disease. *Geneva: World Health Organization*;65(4):281-393.

Winter WE, Maxwell GL, Tian C, Carlson JW, Ozols RF, Rose PG, et al. (2007) Prognostic factors for stage III epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*;25(24):3621-7.

Winter WE, Maxwell GL, Tian C, Sundborg MJ, Rose GS, Rose PG, et al. (2008) Tumor residual after surgical cytoreduction in prediction of clinical outcome in stage IV epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*;26(1):83-9.

Xu B, Ye MH, Lv SG, Wang QX, Wu MJ, Xiao B, et al. (2017) SNORD47, a box C/D snoRNA, suppresses tumorigenesis in glioblastoma. *Oncotarget*;8(27):43953-66.

Yang H, Lin P, Wu HY, Li HY, He Y, Dang YW, et al. (2018) Genomic analysis of small nucleolar RNAs identifies distinct molecular and prognostic signature in hepatocellular carcinoma. *Oncol Rep*;40(6):3346-58.

Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, et al. (2016) Ensembl 2016. *Nucleic acids research*;44(D1):6.

Yemelyanova AV, Cosin JA, Bidus MA, Boice CR, Seidman JD (2008) Pathology of stage I versus stage III ovarian carcinoma with implications for pathogenesis and screening. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society*;18(3):465-9.

Yi C, Wan X, Zhang Y, Fu F, Zhao C, Qin R, et al. (2018) SNORA42 enhances prostate cancer cell viability, migration and EMT and is correlated with prostate cancer poor prognosis. *The International Journal of Biochemistry & Cell Biology*;102(Nat. Genet. 25 2000):138-50.

Yin BW, Lloyd KO (2001) Molecular cloning of the CA125 ovarian cancer antigen: identification as a new mucin, MUC16. *The Journal of biological chemistry*;276(29):27371-5.

Yokoi A, Matsuzaki J, Yamamoto Y, Yoneoka Y, Takahashi K, Shimizu H, et al. (2018) Integrated extracellular microRNA profiling for ovarian cancer screening. *Nat Commun*;9(1):4319.

Zhang B, Pan X, Cobb GP, Anderson TA (2007) microRNAs as oncogenes and tumor suppressors. *Developmental biology*;302(1):1-12.

Zhang C, Zhao L-m, Wu H, Tian G, Dai S-l, Zhao R-y, et al. (2018) C/D-Box Snord105b Promotes Tumorigenesis in Gastric Cancer via ALDOA/C-Myc Pathway. *Cellular Physiology and Biochemistry*;45(6):2471-82.

Zhang C, Zhao LM, Wu H, Tian G, Dai SL, Zhao RY, et al. (2018) C/D-Box Snord105b Promotes Tumorigenesis in Gastric Cancer via ALDOA/C-Myc Pathway. *Cell Physiol Biochem*;45(6):2471-82.

Zhang X, Li H, Yu X, Li S, Lei Z, Li C, et al. (2018) Analysis of Circulating Tumor Cells in Ovarian Cancer and Their Clinical Value as a Biomarker. *Cell Physiol Biochem*;48(5):1983-94.

Zhao W, He X, Hoadley KA, Parker JS, Hayes D, Perou CM (2014) Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics*;15(1):1-11.

Zhao W, Ma X, Liu L, Chen Q, Liu Z, Zhang Z, et al. (2019) SNHG20: A vital lncRNA in multiple human cancers. *J Cell Physiol*.