

UNIVERSITÉ CHEIKH ANTA DIOP



ECOLE DOCTORALE SCIENCES DE LA VIE, DE LA SANTÉ ET DE **L'ENVIRONNEMENT**



FACULTE DES SCIENCES ET TECHNIQUES  
DEPARTEMENT DE BIOLOGIE ANIMALE

Année : 2018

N° d'ordre : 201867

# Caractérisation et évolution génétiques du cancer du sein chez les femmes nigérianes et sénégalaises

**THESE DE DOCTORAT**

Spécialité : **GENETIQUE DES POPULATIONS**

Présentée par :

**Aniefiok John UDOAKANG**

Soutenue le 11 février 2019 à 15h devant le jury composé de :

Président :	M. Bhen Sikina TOGUEBAYE,	Professeur titulaire, FST/UCAD
Rapporteurs :	M. Malick FALL	Professeur assimilé, FST/UCAD
	M. Nnodimele ATULOMAH,	Professeur titulaire, Université BABCOCK
	M. Jean Yves CORNO	Directeur de recherche Univ. MONTP
Examineur :	M. Ganiyu O. ARINOLA,	Professeur titulaire, UCH/UI
	M. Gora DIOP,	Maître de Conférences titulaire, FST/UCAD
	Mme. Fatimata MBAYE,	Maître de Conférences assimilé, FST/UCAD

---

Directeur de Thèse : M. Mbacké SEMBENE Professeur titulaire, FST/UCAD  
Co-Directeur de Thèse : M. Temitayo O. OGUNDIRAN, Professeur titulaire, UCH/UI

## DEDICACES

Cet ouvrage est dédié à la Trinité :

Au Dieu tout puissant pour son infinie miséricorde, sa grâce et le privilège qu'il m'a donné pour entreprendre cette étude.

Au Seigneur Jésus pour m'avoir sauvée de l'esclavage du péché, de la mort, Au Saint-Esprit, mon partenaire principal, pour m'avoir guidée sur le chemin interminable de la vie, et d'être mon aide infaillible en tout temps.

Je t'aime et je m'abandonne à toi.

Ma vie m'est empruntée,

Et je la vivrai selon les termes de Celui qui me l'a donnée.

Mais où serai-je sans ta GRACE ?

Impuissante et sans espoir. Perdue sans tes miséricordes.

Sans ton amour et ta grâce, je serai perdue dans la misère, Ô Seigneur.

Je te donne mon tout et je les dépose au pied de la croix.

## **REMERCIEMENTS**

Je remercie la Communauté Économique des États de l'Afrique de l'Ouest (CEDEAO) pour le parrainage de cette étude doctorale et l'Université Cheikh Anta Diop de Dakar, au Sénégal, pour l'opportunité d'être admise. Je voudrais remercier les membres de l'équipe "GENGESPOP", qui m'ont accueillie et avec qui ce travail a été effectué jusqu'à la fin.

Mes remerciements les plus sincères à mon superviseur, le professeur Mbacké SEMBENE, pour m'avoir confié la possibilité de faire ce travail et d'avoir patiemment supervisé le processus jusqu'à la fin. Je vous suis reconnaissante.

En outre, je remercie le professeur Temidayo O. Ogundiran qui, malgré son emploi du temps, a volontairement accepté de co-diriger ce travail particulièrement pendant la période de prélèvement de l'échantillon et aussi pour m'avoir mis en contact avec le Professeur Arinola qui m'a aidée dans le processus de collecte des échantillons au Nigéria, je vous remercie pour votre soutien.

Dr. Fatimata Mbaye SY, qui n'est pas seulement un mentor mais une sœur pour moi dans ce pays étranger. Je suis si heureuse de vous avoir rencontrée et que nos chemins se soient croisés à ce moment de ma vie. Je n'aurais pu demander plus. Votre patience, vos conseils et vos innombrables contributions ont mené à bien ce travail.

Dr. MamaRacky Ndiaye, Mlle Binta Keneme, Mlle Diarra, Dr. Cheikh Abdou Khadre DIA, M. Adjouma SARR, Dr. Ndèye Penda, je vous remercie pour tous les moments passés ensemble. Je vous apprécie tous.

Mes parents : M. et Mme John UDOAKANG, pour votre patience et votre soutien pendant la durée de ce programme.

Mes frères Prince-Wealth et God'swill Udoakang, les moments de joie partagés ont dissipé une partie du stress du doctorat.

Mes chers amis : M. et Mme Simon Gondo et leur famille (Ivande, Ngusha et Sister Tee), je vous remercie pour votre soutien et vos encouragements, en particulier pendant la période de prélèvement d'échantillons au Nigéria et tout au long de ce programme. Votre soutien ne peut être quantifié. M. et

Mme Olarewanju Ojelabi, je suis reconnaissante pour votre immense soutien, vos encouragements et vos appels malgré la distance. Au Frère Charles Okotie et sa famille, avec qui Dieu a permis que l'on soit de nouveau en contact au moment opportun pour me soutenir dans les prières et les encouragements.

Ma famille ecclésiastique à Dakar, Madame Deborah Dione, ma rencontre avec vous a été l'une des meilleures choses qui me soit arrivée ici au Sénégal. Vous avez été comme une mère pour moi à travers vos conseils, votre soutien, vos prières et vos encouragements. Que Dieu vous bénisse. M. et Mme Julia Munyai (ma confidente) et sa famille (Vhahangwele et Musiki), votre soutien ne saurait être quantifié, car vous m'avez accueillie dans votre famille et je vous en suis extrêmement reconnaissante. Hope Odinakachi, qui a été comme un frère pour moi. Au Pasteur Jacques et sa famille, Tosin Onyinlade, Stanley Njigwum, Thomas Aristide, Philippe et Niassa Wazou, Jean Eudes, Lloyd, Pierre, Victoire, Blessing Onyeka, Ogbulu, Dr. Franklin Ani, M. et Mme Oshi, Fred, Aniefon, et la famille Emeruwa.

A Mary Oboh, Thomas Colbe, Deborah Obisesan, Haddy Lowe née Crookes et Mazo Kone qui étaient des collègues de travail et qui, plus tard, sont devenus des amis et une famille. Je vous remercie pour votre soutien, vos rires et vos encouragements. Dieu vous bénisse. Aux voisins qui étaient là au moment des besoins : Mme Ka Fatou Diouma et sa famille et Dr. Grace Sikuli, je suis reconnaissante. M. Abdoulaye Niang, dont les encouragements et le soutien ont été formidable. Madame Chibuzor Emmanuel et Madame Afolabi, vous avez rendu mon séjour à l'Université College Hospital inoubliable. J'en suis là grâce à votre immense contribution à ce travail. Pr. Adekunle A. Bakare, mon mentor, sur qui je peux toujours compter. Merci à M. Julius, Mme Esther, Mme Diongo, Abiket, Oluwadunsin et sa mère, et à tous ceux que je n'ai pas pu mentionner - vos appels, vos prières et votre soutien sont grandement appréciés. Que Dieu vous bénisse et vous le rende au centuple. Amen.

## LISTE DES ABREVIATIONS

ADN :	Acide désoxyribonucléique
ADNcl :	d'ADN Cellules Libre
ADNct :	ADN Tumoral Circulant
ADP :	Adénosine diphosphate
CfADN :	ADN libre circulant
ANCP :	Antigène Nucléaire de Cellules en Prolifération
ARNm:	ARN messenger
ATL:	Animal Tissue Lysis
ATP :	Adénosine triphosphate
BC :	Breast Cancer
BIOPASS :	Biologie des Populations Animales Sahélo-Soudaniennes
CCI :	Carcinome Canalaire Infiltrant
CIRC :	Centre International de Recherche sur le Cancer
CHU :	Centre Hospitalier Universitaire
CRC :	Colorectal Cancer
CTC :	Cellules Tumourales Circulantes
CTD :	Cellules Tumourales Disséminées
dNTP :	désoxyribonucléotides
FCT :	Facteur de Croissance Transformant
FNA :	Aspiration à l'aiguille fine
IDH :	Indice de Développement Humain
INC :	Instabilité de Chromosome
IMRAT:	Institute for Medical Research and Training
IMP :	Instabilité de la Mutation Ponctuelle
IRD :	Institut de Recherche pour le Développement
ISRA :	Institut Sénégalais de Recherche Agricole
MACEp :	Molécule d'adhésion des cellules épithéliales
MACM :	Molécule d'Adhesion Cellulaire du Mélanome
HML 1:	Homologue mutL 1
HNPCC:	Hereditary Non-Polypopsy Colorectal Cancer
LOH:	Loss of heterozygosity
LSB :	Liaison à l'ADN simple brin

MMR :	Mismatch Repair
MSI :	Microsatellite instability
MSI-H :	Instabilité Microsatellite Haute (Microsatellite Instability-High)
MSI-L :	Instabilité microsatellite bas (Microsatellite Instability-Low)
MSS :	Microsatellites stables (Stable Microsatellite)
OMS :	Organisation Mondiale de la Santé
PMIC :	Phénotype méthylateur de l'îlot CpG
PCR :	Polymerase Chain Reaction
RMN :	Répétitions mononucléotidiques
SSD :	Sum of Squared Deviation
TEM :	Transition Epithéliale-Mésenchymale
TNM :	Tumour Node Metastases
TNBC:	Triple negative breast cancer (cancer du sein triple négatif)
UCAD :	Université Cheikh Anta Diop
UCH:	University College Hospital
UTR :	Untranslated region
WHO :	World Health Organization

## LISTE DES FIGURES

<b>Figure 1</b> : Diagramme en secteur illustrant la répartition de l'incidence et des décès par cancer par région du monde en 2018 pour les hommes et les femmes.....	7
<b>Figure 2</b> : Diagramme en secteur montrant la proportion d'incidence du cancer et de décès chez les femmes par zone mondiale en 2018.....	7
<b>Figure 3</b> : Graphique à secteurs montrant le nombre de nouveaux cas de cancer et de décès dans le monde pour les huit cancers les plus fréquents et autres .....	8
<b>Figure 4</b> : Carte globale présentant le type d'incidence du cancer le plus courant chez les femmes en 2018 dans chaque pays .....	8
<b>Figure 5</b> : Diagramme à barres illustrant le nombre de nouveaux cas et de décès dus au cancer du sein chez les femmes de tous âges dans différentes régions d'Afrique.....	12
<b>Figure 6</b> : Graphique en camembert montrant le pourcentage de nouveaux cas de cancer et de décès chez toutes les femmes au Nigeria en 2018 .....	12
<b>Figure 7</b> : Diagramme à secteurs illustrant le pourcentage de nouveaux cas de cancer et de décès chez toutes les femmes au Sénégal en 2018 .....	13
<b>Figure 8</b> : Préparations complètes portant un lobule de type 1 (Lob 1) et un carcinome du canal in situ impliquant plusieurs canaux terminaux .....	15
<b>Figure 9</b> : Représentation schématique des voies postulées des actions des œstrogènes sur les cellules épithéliales du sein .....	17
<b>Figure 10</b> : Principes concernant les origines, les types et les conséquences de l'hétérogénéité néoplasique .....	21
<b>Figure 11</b> : Hiérarchie de l'épithélium mammaire normal et des cellules de sous-types intrinsèques du cancer du sein .....	22
<b>Figure 12</b> : Hétérogénéité intra-tumorale dans le diagnostic et le traitement du cancer .....	24
<b>Figure 13</b> : Voies vers la cancérogenèse.....	27
<b>Figure 14</b> : Représentation du système MMR chez les eucaryotes, la plupart des bactéries et chez E. coli .....	37
<b>Figure 15</b> : Le système DNA MMR fonctionne en plusieurs étapes .....	39
<b>Figure 16</b> : Régions des interactions protéine-protéine parmi les membres du système MMR. ATP, adénosine triphosphate; PMS, ségrégation post-méiotique .....	43
<b>Figure 17</b> : Carte de l'Afrique montrant les deux pays de la population étudiée.....	51
<b>Figure 18</b> : Phénotype MSI NR-21 pour le Nigeria et le Sénégal.....	73
<b>Figure 19</b> : Phénotype MSI NR-24 pour le Nigeria et le Sénégal.....	73

<b>LISTE DES TABLEAUX</b>
---------------------------

Tableau I : Résumé des facteurs pronostiques et prédictifs.....	32
Tableau II : Terminologie des microsatellites.....	41
Tableau III : Comparaison des caractéristiques cliniques des cas de cancer du sein au Nigéria et au Sénégal.....	53
Tableau IV : Cas de la comparaison de certaines caractéristiques du cancer du sein au Nigéria et au Sénégal.....	54
Tableau V : Séquences d'amorces pour les tests d'instabilité des microsatellite.....	65
Tableau VI : Composition du mélange de réaction PCR.....	65
Tableau VII : Étapes et conditions des thermocycleur.....	66
Tableau VIII : Résultats du phénotype MSI chez les patientes atteintes d'un cancer du sein au Nigeria et au Sénégal.....	72
Tableau IX : Tailles et fréquences alléliques du <i>NR-21</i> au Nigeria et au Senegal.....	74
Tableau X : Liste des haplotypes et des motifs <i>NR-24</i> .....	75
Tableau XI : Tailles et fréquence d'allèles du <i>NR-24</i> au Nigéria et au Sénégal.....	76
Tableau XII : Paramètres de variabilité génétique des séquences de la <i>NR-21</i> et <i>NR-24</i> des tissus cancéreux.....	78
Tableau XIII : Analyse des différences inter-populations et intra population pour le <i>NR-21</i> par AMOVA.....	79
Tableau XIV : Analyse des différences inter-populations et intra population pour le <i>NR-24</i> par AMOVA.....	79
Tableau XV : Test de diversité moléculaire.....	80
Tableau XVI : Test de neutralité selective.....	80
Tableau XVII : Analyse des mismatch distribution pour l'expansion génétique.....	81
Tableau XVIII : Relation entre les paramètres cliniques et le phénotype MSI pour <i>NR-21</i> chez les patientes atteintes d'un cancer du sein au Nigéria et au Sénégal.....	82
Tableau XIX : Relation entre les paramètres cliniques et le phénotype MSI pour <i>NR-24</i> chez les patientes atteintes d'un cancer du sein au Nigéria et au Sénégal.....	82
Tableau XX : Relation entre les paramètres cliniques et le phénotype MSI pour les tumeurs MSS, MSI-L et MSI-H chez les patientes atteintes d'un cancer du sein au Nigéria et au Sénégal.....	83

# TABLE DES MATIERES

TITRES	N° DE PAGE
<b>DEDICACES</b> .....	i
<b>REMERCIEMENTS</b> .....	ii
<b>LISTE DES ABREVIATIONS</b> .....	iv
<b>LISTE DES FIGURES</b> .....	vi
<b>LISTE DES TABLEAUX</b> .....	vii
<b>INTRODUCTION GENERALE</b> .....	1
<b>CHAPITRE I. SYNTHESE BIBLIOGRAPHIQUE</b> .....	6
<b>INTRODUCTION</b> .....	6
<b>I.1. Développement du cancer du sein</b> .....	10
<b>I.2. Evolution génétique et hétérogénéité moléculaire du génome du cancer</b> .....	16
<b>I.2.1. Hétérogénéité moléculaire dans le cancer du sein</b> .....	19
<b>I.2.1.1. Hétérogénéité inter-tumorale dans le cancer du sein</b> .....	19
<b>I.2.1.1.1. Sous-types moléculaires du cancer du sein</b> .....	19
<b>I.2.1.2. Hétérogénéité intra-tumorale du cancer du sein</b> .....	22
<b>I.2.1.3. Cellules tumorales circulantes et disséminées (CTC)</b> .....	23
<b>I.2.1.3.3. Transition épithélio-mésenchymale dans les CTC</b> .....	26
<b>I.2.1.3.2. ADN libre et circulant des tumeurs</b> .....	28
<b>I.3. Facteurs de risque épidémiologique du cancer du sein</b> .....	29
<b>I.3.1. Facteurs de risque de reproduction associés au cancer du sein</b> .....	30
<b>I.3.1.1. Age à la ménarche et âge à la ménopause</b> .....	30
<b>I.3.1.2. Âge à la première grossesse à terme et à la première naissance vivante</b> .....	30
<b>I.3.1.3. Parité, allaitement et cancer du sein</b> .....	31
<b>I.4. Caractéristiques cliniques et pathologiques du cancer du sein</b> .....	31
<b>I.5. Instabilité génétique</b> .....	33
<b>I.5.1 Microsatellites</b> .....	39
<b>I.5.2. Utilisation de deux marqueurs pour la détection de l'instabilité des microsatellites dans les tumeurs</b> .....	46
<b>I.5.2.1. Procédure de réaction en chaîne par polymérase</b> .....	47

<b>CHAPITRE II. COMPARAISON DES PARAMÈTRES ÉPIDÉMIOLOGIQUES DU CANCER DU SEIN ENTRE LES PATIENTES NIGÉRIANES ET SÉNÉGALAISES.....</b>	<b>49</b>
<b>INTRODUCTION.....</b>	<b>49</b>
<b>II.1. Matériel et méthodologie.....</b>	<b>50</b>
II.1.1. Approbation éthique.....	50
II.1.2. Critères d'inclusion et d'exclusion.....	50
II.1.3. Collecte de données et analyses statistiques.....	51
<b>II.2. Résultats et Discussion.....</b>	<b>52</b>
<b>II.2.1. Résultats.....</b>	<b>52</b>
<b>II.2.2. Discussion.....</b>	<b>55</b>
Conclusion partielle.....	57
<b>CHAPITRE III : POLYMORPHISME ET EVOLUTION GENETIQUE DES MICRO-SATELLITES NR-21 ET NR-24.....</b>	<b>59</b>
<b>INTRODUCTION.....</b>	<b>59</b>
<b>III.1. Matériel et méthodes.....</b>	<b>61</b>
III.1.1. Considération éthique.....	61
III.1.2. Sites d'étude, Critères d'Inclusion et d'Exclusion.....	61
III.1.2.1. Nigéria.....	61
III.1.2.2. Sénégal.....	62
III.1.3. Étude du Phénotype des Loci NR-21 et NR-24.....	63
III.1.3.1. Choix de Marqueurs dans l'Etude.....	63
III.1.3.2. Extraction de l'ADN.....	63
III.1.3.3. Réaction en chaîne par polymérase (PCR).....	64
III.1.3.4. Séquençage.....	66
III.1.4. Nomenclature et classification de variantes de séquence.....	66
III.1.4.1. Variante DELETION.....	67
III.1.4.2. Variante INSERTION.....	67
III.1.4.3. Variante SUBSTITUTION.....	67
III.1.4.4. Variante DUPLICATION.....	67
III.1.4.5. Variante INSERTION PAR DELETION.....	68
III.1.4.6. Variante INVERSION.....	68

III.1.5. Analyses moléculaires pour la génétique des populations à l'aide de microsatellites...	68
III.1.5.1. Analyse du polymorphisme et de la diversité génétique.....	68
III.1.5.2. Analyse de l'évolution génétiques.....	69
III.1.5.2.1. Analyse de la structuration génétique.....	69
III.1.5.2.2. Test de neutralité pour la variabilité intra-spécifique et inter-spécifique .....	70
III.1.5.3. Analyse d'expansion génétique.....	71
III.1.6. Analyse Statistique.....	71
III.2. Résultats et Discussion.....	72
III.2.1. Résultats.....	72
<b>III.2.1.1</b> Phénotype d'instabilité des microsatellites <i>NR-21</i> et <i>NR-24</i> chez les patients atteints du cancer du sein au Nigéria et au Sénégal.....	72
<b>III.2.1.2.</b> Paramètres de la diversité génétique.....	77
<b>III.2.1.3.</b> Structuration génétique.....	78
<b>III.2.1.4.</b> Test de l'hypothèse de mutation neutre.....	80
III.2.2. Les relations entre les marqueurs étudiés ( <i>NR-21</i> et <i>NR-24</i> ) et les caractéristiques clinico-pathologiques.....	81
III.2.3. Discussion.....	83
Conclusion partielle.....	90
Conclusion générale et recommandations.....	92
ANNEXES	

## INTRODUCTION GENERALE

Le cancer, également appelé tumeur maligne, est un nom générique qui représente un grand groupe de maladies pouvant toucher n'importe quelle partie du corps (WHO, 2018). C'est un trouble pathologique majeur qui provoque la morbidité et la mortalité dans diverses populations dans travers le monde. L'une des principales caractéristiques des tissus cancéreux est qu'ils sont constitués de cellules ayant une croissance et une multiplication anormales, au-delà de la taille normale (WHO, 2018). Pour cette raison, ces cellules envahissent rapidement les tissus voisins atteignant d'autres parties du corps (WHO, 2018). Au cours de la progression de la tumeur, certaines cellules acquièrent un potentiel oncogène. Ainsi, ils prolifèrent rapidement, échappent aux mécanismes de régulations cellulaires (apoptose) et immunologiques (immunosuppression) et entre en compétition avec les cellules saines pour l'exploitation des ressources de l'organisme (Hanahan et Weinber, 2011 ; Vogelstein *et al.*, 2013 ; Greaves, 2015).

En 2008, le Centre International de Recherche sur le Cancer (CIRC) a signalé 715 000 nouveaux cas de cancer et 542 000 décès dûs au cancer en Afrique (Ferlay *et al.*, 2008). Dix ans plus tard, Ferlay *et al.*, (2018) ont enregistré une estimation de 18,1 millions de nouveaux cas de cancer et de 9,6 millions de décès avec une incidence de 5,8% et un taux de mortalité de 7,3% en Afrique (Ferlay *et al.*, 2018). Cette forte augmentation pourrait s'expliquer en partie, selon l'Organisation Mondiale de la Santé (WHO, 2008), par l'adoption de comportements ou de modes de vie à risque (tabagisme, alimentation mal saine, inactivité physique...) consécutive au processus de développement économique.

Concernant le cancer du sein, il est devenu le deuxième cancer le plus diagnostiqué chez les hommes après le cancer du poumon et le cancer le plus fréquemment diagnostiqué chez les femmes (Ferlay, 2018). Ce même rapport a estimé en 2018 le nombre de nouveaux cas de cancer du sein à 2 088 849 (11,6%) et de décès 626 679 (6,6%) dans le monde. Bien qu'il soit plus répandu dans les pays développés et dans certaines régions industrialisées, son taux de mortalité est en baisse dans ces pays (Ogundiran et Adebamowo, 2015). En Afrique, il est devenu le cancer le plus répandu, avec des taux d'incidence et de mortalité en croissance et une grande variabilité selon les régions géographiques (Ogundiran et Adebamowo, 2015). Sa survenue semble aussi de plus en plus précoce chez les patientes. En effet, des études faites en milieu hospitalier ont mis en évidence que de plus en plus de patientes âgées entre 20 et 30 étaient affectées par ce cancer (Akarolo-Anthony *et al.*, 2010 ; Mbaye *et al.*, 2012). Ces études ont aussi montré que la plupart des cas intervenait principalement entre 40 et 44 ans avec souvent une forme de survenue agressive (Akarolo-Anthony *et al.*, 2010 ; Mbaye *et al.*, 2012).

Plus précisément, au Nigéria et au Sénégal, le cancer du sein constitue un réel problème de santé publique représentant respectivement 28,7% et 26,1% des nouveaux cas de cancer chez les femmes de tous les âges dans les deux pays (Ferlay *et al.*, 2018). Alors qu'il est le cancer le plus répandu chez les femmes nigérianes avec une incidence croissante de 34,3 pour 100 000 femmes (21 750 nouveaux cas par an) à 41,7 pour 100 000 entre 1970 et 2018, il seconde le cancer du col de l'utérus au Sénégal (Adebamowo et Ajayi, 2000 ; Ferlay *et al.*, 2018). De plus, il est principalement diagnostiqué comme une tumeur faiblement différenciée au stade avancé (III et IV) chez les femmes en pré-ménopause (41 à 50 ans), présentant des tumeurs agressives et le Carcinome Canalaire Infiltrant (CCI) comme type histologique le plus courant (Dem *et al.*, 2008 ; Ntekim *et al.*, 2009 ; Mbaye *et al.*, 2012 ; Agbo *et al.*, 2014).

L'impossibilité d'identifier un agent étiologique spécifique, le moment précis de l'initiation et tous les mécanismes moléculaires responsables de l'initiation et de la progression du cancer du sein sont aujourd'hui un obstacle majeur à l'éradication de cette maladie (Russo *et al.*, 2000).

L'effet du mode de vie et des facteurs prédisposants à l'environnement révèlent les différences ethniques et géographiques observées dans l'incidence du cancer du sein (Mehrgou et Akouchekian, 2016). Trois à dix pour cent de tous les cas de cancer du sein et environ 30% des cancers du sein de stade précoce résultent de facteurs héréditaires, tandis que le cancer du sein sporadique représente environ 90% de la maladie (Tereschenko *et al.*, 2002 ; Calderon-Garciduenas *et al.*, 2005). Les causes réelles du cancer du sein ne sont pas actuellement connues mais plusieurs facteurs épidémiologiques, tels que : l'âge, le sexe, l'exposition aux radiations, la diminution de l'activité physique, le mode de vie sédentaire, les antécédents familiaux de cancer du sein et d'autres cancers, l'hérédité, l'alimentation (y compris l'alcool et le tabac, etc.), les facteurs reproductifs et hormonaux tels que les antécédents menstruels lointains (la nulliparité, le manque d'allaitement, ménarche précoce et ménopause tardive), l'utilisation de contraceptifs et les caractéristiques anthropométriques (poids ; taille ; hanche ; hauteur) ont été incriminés comme facteurs étiologiques possibles (Hamajima *et al.*, 2002 ; Colditz *et al.*, 2006 ; Okobia *et al.*, 2006 ; Lord *et al.*, 2008 ; Schulz *et al.*, 2008 ; Datta et Biswas, 2009 ; Ayub *et al.*, 2014). C'est donc un trait multifactoriel avec une prédisposition génétique et des facteurs de risques environnementaux et hormonaux (Chlebowski *et al.*, 2013).

En dépit des progrès importants réalisés dans le diagnostic et le traitement du cancer du sein, il subsiste certains problèmes cliniques et scientifiques non résolus à ce jour. Certains d'entre eux incluent : la prévention, le diagnostic (trouver des méthodes plus spécifiques et plus sensibles),

la progression et la récurrence de la tumeur, le traitement et la résistance thérapeutique (Polyak, 2007). En effet, les différences moléculaires (sous-types) du cancer du sein conduisent à un résultat clinique et à une réponse aux traitements différents. Par conséquent, en raison du taux élevé de la mortalité et du coût financier du traitement, le cancer du sein revêt une grande importance sociale et économique en Afrique. De ce fait, la mise au point de moyens efficaces de prévention et de détection précoce est actuellement la solution la plus viable.

Dans cette optique, de plus en plus de recherches sont menées pour déterminer les facteurs les plus associées à la survenue du cancer du sein et des techniques de détection précoce ont été expérimentées. Une bonne partie de ces études sont menées sur des marqueurs moléculaires. Bien que le cancer du sein soit un trait multifactoriel avec une prédisposition génétique, deux autres facteurs majeurs identifiés pour ce type de cancer incluent les facteurs environnementaux et les facteurs hormonaux. En dépit des progrès importants réalisés dans le diagnostic et le traitement du cancer du sein, il subsiste certains problèmes cliniques et scientifiques non résolus. Certains incluent et sont liés à : la prévention (qui en a besoin et quand), le diagnostic (méthodes plus spécifiques et plus sensibles), la progression et récurrence de la tumeur (les causes et comment les prévoir), le traitement (qui doit être traité et comment ?) et la résistance thérapeutique (comment prédire, prévenir et vaincre ?) (Polyak, 2007). Les différences moléculaires (sous-types) du cancer du sein conduisent à un résultat clinique et à une réponse au traitement distincts ; le sous-type de tumeur basale ressemblant au moins à un résultat clinique et à la réponse au traitement les plus défavorables. Ainsi, en raison du taux élevé de mortalité et du coût financier du traitement, le cancer du sein revêt une grande importance sociale et économique en Afrique. Par conséquent, la mise au point de moyens efficaces de prévention et de détection précoce est actuellement la solution la plus viable.

Le système MMR est un système de réparation de l'ADN hautement conservé qui contribue principalement au maintien de la stabilité génomique par la correction de paires de bases non appariées (Friedberg *et al.*, 2005). Le phénotype d'instabilité des microsatellites, l'un des dysfonctionnements causés par le dysfonctionnement du MMR, a été rapporté dans différents types de tumeurs sporadiques et est bien normalisé pour le cancer colorectal (Yaza *et al.*, 2017). Il conduit à la génération de nouveaux allèles de tailles très différentes du type sauvage et est associé à un pronostic amélioré, une survie supérieure du patient, un bon prédicteur de la réponse au cancer et une résistance à certaines chimiothérapies (Rubic *et al.*, 2003). La technique actuelle d'analyse de l'état des microsatellites implique l'utilisation de kits

disponibles dans le commerce basés sur l'amplification de cinq "marqueurs" de microsatellites. Ces cinq marqueurs sont: *BAT-25*, *BAT-26*, *NR-21*, *NR-22* et *NR-24* et sont appelés quasi-monomorphes car ils se caractérisent par un certain nombre de tailles répétitives homogènes dans une population donnée (entre les individus et les deux allèles du même individu).

L'utilisation de ces marqueurs est née de la nécessité de contourner l'utilisation de tissus normaux (non cancéreux) appariés à des tissus cancéreux. Ceci pour n'utiliser que des tissus cancéreux. Ce qui permettra de minimiser les coûts et d'éviter la confusion des échantillons lors de l'analyse. Il a été prouvé que, lorsque ces cinq marqueurs mononucléotidiques étaient utilisés ensemble dans une réaction d'amplification unique appelée PCR pentaplex, il suffisait d'identifier correctement les cas de MSI dans les tissus tumoraux, bien que *BAT-25* et *NR-21* se soient révélés polymorphes dans certaines populations en Afrique (Pyatt *et al.*, 1999; Buecher *et al.*, 2011). Cette option figurait dans les Lignes directrices Bethesda révisées de 2004 pour HNPCC et MSI, bien que le panel initial de mononucléotides et de marqueurs dinucléotides demeurait inchangé, de même que la définition de l'instabilité microsatellite haute (MSI-H), instabilité microsatellite faible (MSI-L) et les tumeurs microsatellites stables (MSS) (Umar *et al.*, 2004a).

Malgré le fait que les dinucléotides et les marqueurs trinucléotidiques aient été étudiés et recommandés (Boland *et al.*, 1998; Perucho, 1999; Sutter *et al.*, 1999; Loukola *et al.*, 2001; Ozer *et al.*, 2002), les répétitions de dinucléotides contiennent entre 60% et 80% de MSI-H et sont parfois classées de manière erronée (Houng *et al.*, 1997; Sutter *et al.*, 1999; Loukola *et al.*, 2001). Par conséquent, la sensibilité et la spécificité de ces cinq marqueurs mononucléotidiques ont été recommandées pour le diagnostic de l'instabilité des microsatellites (Suraweera *et al.*, 2002). Ces marqueurs mononucléotidiques sont situés dans le génome comme suit: *BAT-25* (situé dans l'intron 16 du *c-kit*); *BAT-26* (situé dans l'intron 5 du gène *MSH2*); *NR-21* (situé dans la région non traduite en 5' du gène *SLC7A8*); *NR-24* (situé dans la région non traduite en 5' du gène *ZNF-2*) et *NR-27-MONO-27* (situé dans la région non traduite en 5' du gène *PAI-1*). En 2013, Esmailnia et ses collègues ont étudié les cinq mononucléotides quasimonomorphes : *BAT-25*, *BAT-26*, *NR-21*, *NR-24*, *NR-27* dans le cancer colorectal et ont découvert une instabilité de 45% dans *NR-21* et *NR-24* (36 cas sur 80). Le *NR-21* était le marqueur instable le plus fréquemment observé avec une instabilité respective de 53% et 25,6% dans les cancers colorectaux sporadiques (CRC) et le cancer colorectal héréditaire sans polypopsie (HNPCC) (Shemirani *et al.*, 2011).

L'hypothèse posée dans ces travaux de recherche est que les caractéristiques épidémiologiques et les processus évolutifs du cancer du sein sont identiques entre les femmes nigérianes et sénégalaises. Cette étude visait à contribuer à l'acquisition d'une meilleure compréhension des facteurs épidémiologiques et du processus évolutif du cancer en Afrique. Quatre objectifs spécifiques ont été pris en compte :

- déterminer et comparer les données épidémiologiques associées au cancer du sein pour la population nigériane et sénégalaise ;
- identifier le polymorphisme génétique de deux loci microsatellites (*NR-21* et *NR-24*) du cancer du sein des deux pays ;
- comparer et contraster les similitudes et les différences entre l'évolution du cancer du sein chez les femmes nigérianes et sénégalaises ;
- déterminer la relation entre les polymorphismes de *NR-21* et *NR-24* et les caractéristiques clinico-pathologiques du cancer du sein.

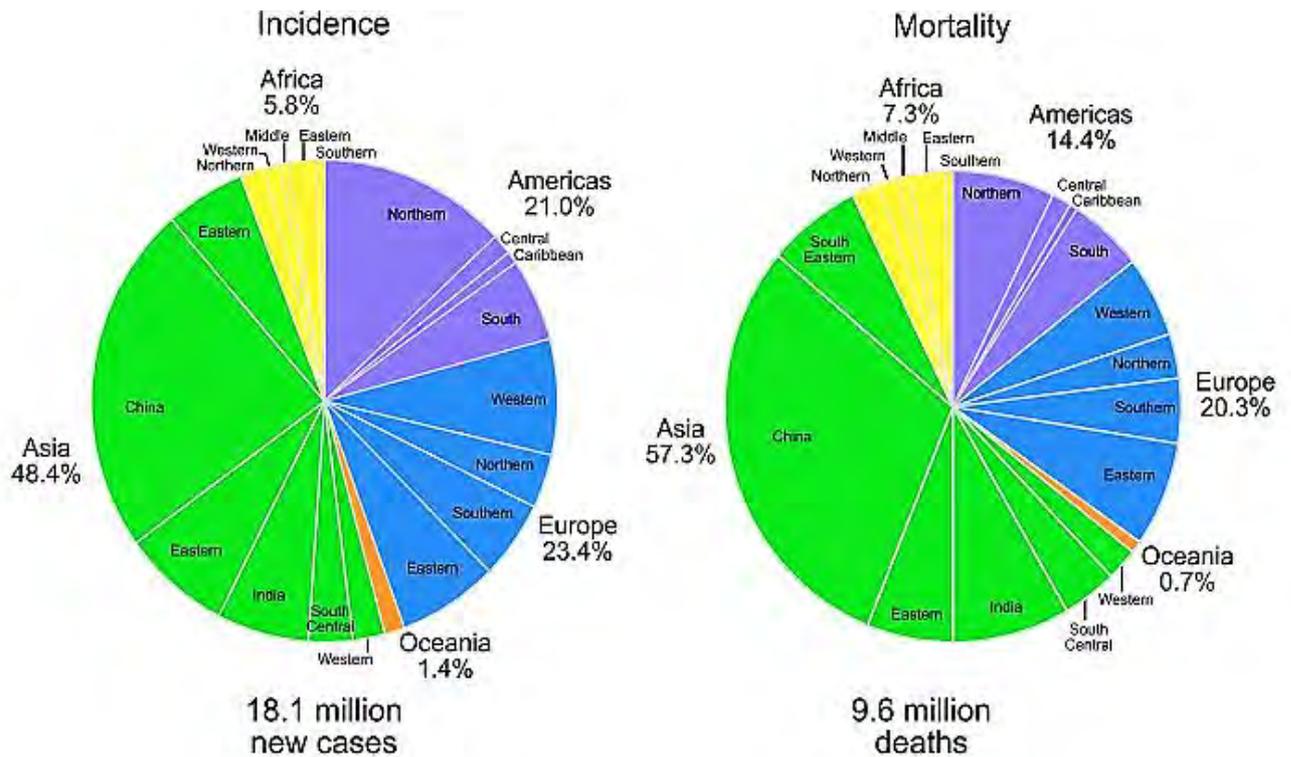
Ce document est scindé en trois chapitres. Le premier renferme une synthèse des principales notions et connaissances acquises sur les cancers (plus particulièrement le cancer du sein). Le deuxième chapitre est consacré à la comparaison des paramètres épidémiologiques entre les patientes nigérianes et sénégalaises. Le troisième chapitre traite du processus d'évolution et les caractéristiques génétiques du cancer du sein chez les femmes nigérianes et sénégalaises à l'aide des marqueurs microsatellites *NR-21* et *NR-24* et cherche les corrélations entre le polymorphisme des marqueurs étudiés et les caractéristiques épidémiologiques du cancer du sein dans les deux populations de notre étude.

# CHAPITRE I. SYNTHÈSE BIBLIOGRAPHIQUE

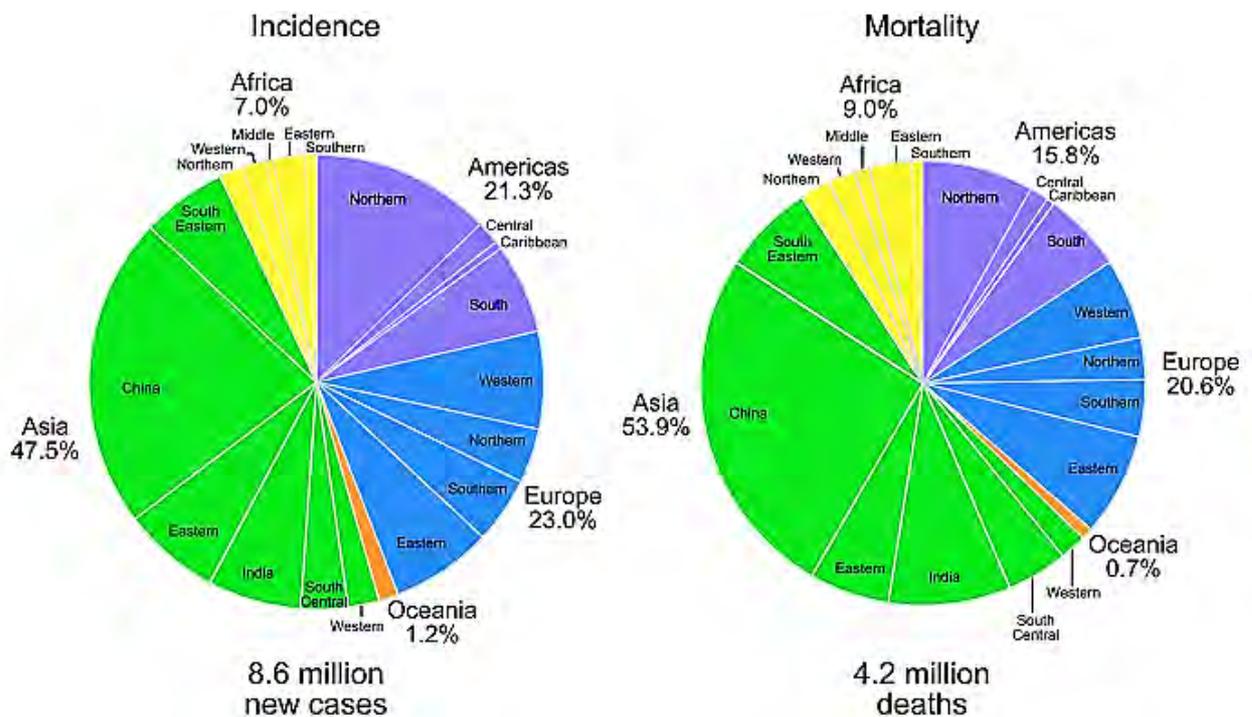
## INTRODUCTION

Au 21<sup>ème</sup> siècle, le cancer devrait être considéré comme la principale cause de décès et le seul obstacle majeur à la croissance de l'espérance de vie dans tous les pays du monde (Bray *et al.*, 2018). L'Organisation Mondiale de la Santé (OMS) a estimé en 2015 que le cancer était la première ou la deuxième cause de décès avant l'âge de 70 ans dans 91 pays sur 172 et la troisième ou la quatrième dans 22 pays supplémentaires. La croissance des incidences et des taux de mortalité par cancer à l'échelle mondiale en fait une source de préoccupation constante pour le monde, allant de paire avec la croissance démographique constante, le vieillissement, l'évolution de la prévalence et la répartition des principaux facteurs de risque, dont certains sont liés au développement socio-économique (Omran, 1971 ; Gersten et Wilmoth, 2002). Les économies émergentes assistent à un remplacement progressif des cancers liés à l'infection et à la pauvreté par les cancers dus à des modes de vie occidentalisés et aux pays à revenu élevé comme en Europe, en Amérique du Nord, en Asie et en Océanie (Gersten et Wilmoth, 2002 ; Maule et Merletti, 2012 ; Bray, 2014). Néanmoins, les différences de profils observées entre différents pays et régions du monde montrent une diversité géographique marquée associée à des facteurs de risques locaux (Bray, 2014). Les figure 1 et 2 montrent la distribution de l'incidence du cancer et des décès par cancer et par région du monde aussi bien chez les hommes que chez les femmes, augmentant les taux d'incidence de 1,5% et de 2% supplémentaires lorsque les chiffres comparés concernent uniquement les femmes.

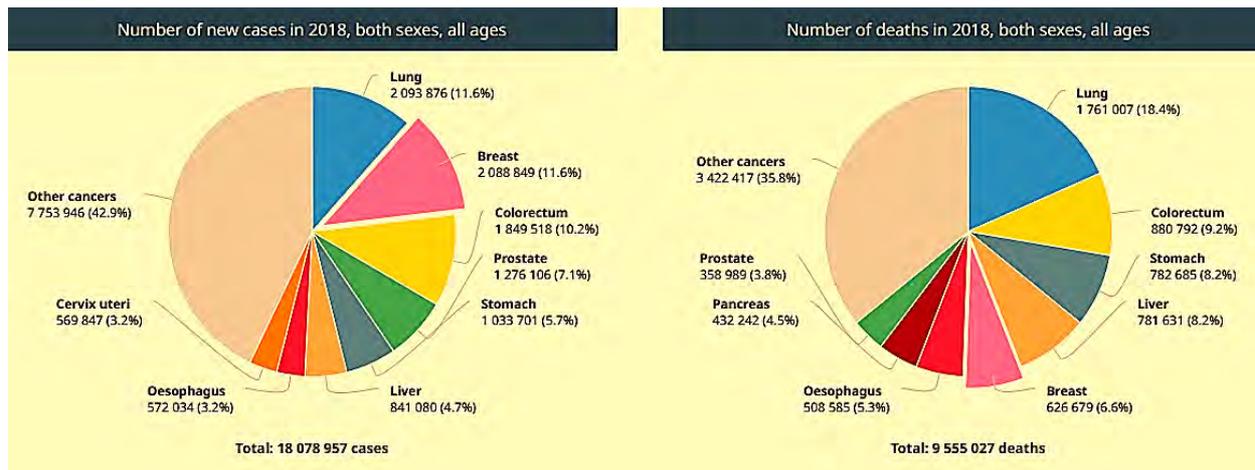
A l'échelle mondiale, le cancer du sein est la cinquième cause de mortalité par le cancer (626 679) ; alors que l'Afrique occupe le cinquième rang en termes d'incidence et de prévalence sur cinq ans, elle se situe au troisième rang en termes de taux de mortalité (74 072). La fourchette de taux de mortalité entre les régions du monde est inférieure à celle de l'incidence en raison de la survie plus favorable du cancer du sein dans les régions développées (à incidence élevée), les taux allant de 8,6 pour 100 000 en Asie orientale à 25,5 pour 100 000 en Mélanésie (Ferlay *et al.*, 2018). La figure 3 montre le nombre de nouveaux cas de cancer et de décès dans le monde pour les huit cancers les plus fréquents en 2018, tandis que la carte de la figure 4 illustre clairement l'incidence des cinq types de cancer les plus répandus chez les femmes dans le monde. Le cancer du sein étant le plus prévalent dans 154 pays, suivis des cancers du col utérin, du foie, du poumon et de la thyroïde. Cela montre l'urgence de s'attaquer à la menace du cancer du sein dans le monde entier (Ferlay *et al.*, 2018).



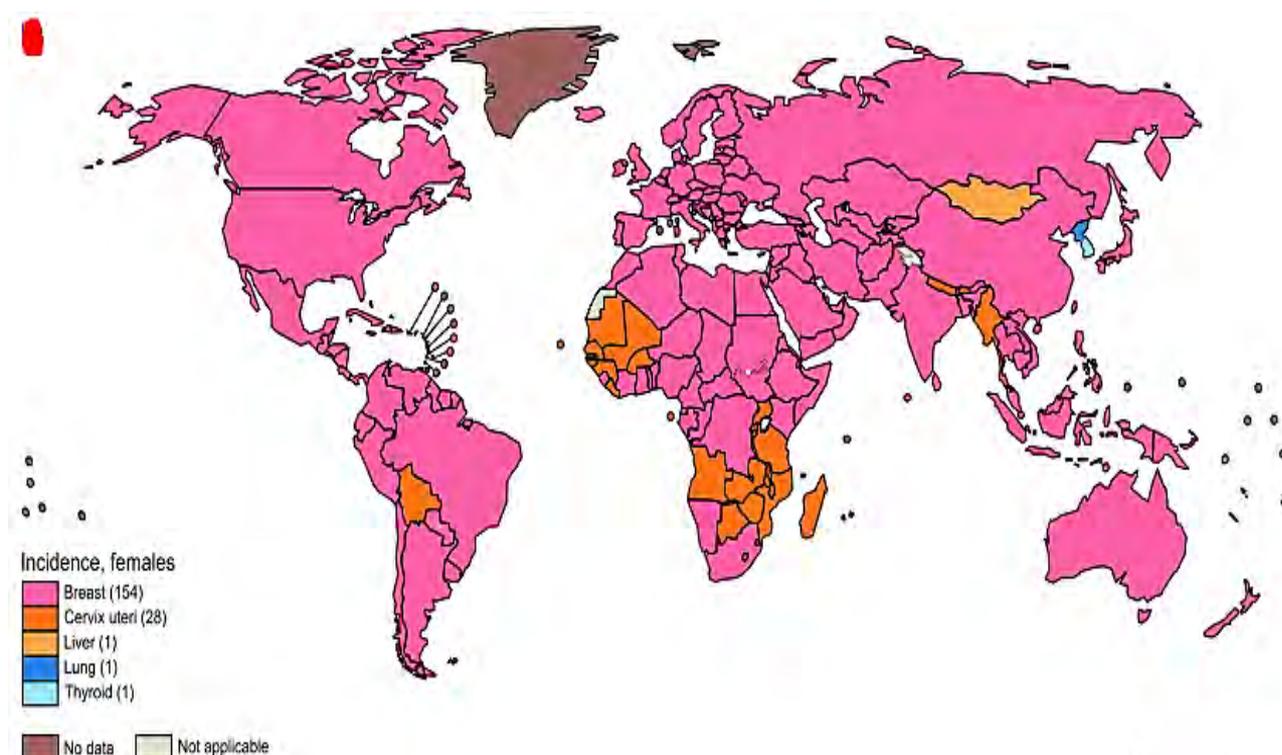
**Figure 1:** Diagramme en secteur illustrant la répartition de l'incidence et des décès par cancer et par région du monde en 2018 pour les hommes et les femmes (Source : GLOBOCAN, 2018).



**Figure 2 :** Diagramme en secteur montrant la proportion d'incidence du cancer et de décès chez les femmes par zone mondiale en 2018 (Source : GLOBOCAN, 2018).



**Figure 3** : Graphique à secteurs montrant le nombre de nouveaux cas de cancer et de décès dans le monde pour les huit cancers les plus fréquents et autres (Source : GLOBOCAN, 2018).



**Figure 4** : Carte globale présentant le type d'incidence du cancer le plus courant chez les femmes en 2018 dans chaque pays (Source : GLOBOCAN, 2018).

En Afrique, avec une population extrêmement diversifiée qui affecte les occurrences des maladies, le traitement et ses résultats, le cancer est de plus en plus reconnu par les professionnels et les experts en santé critique, suscitant de graves inquiétudes quant à sa progression, sa gestion et la réponse des patients au traitement (Parkin *et al.*, 2003 ; Akarolo-

Anthony *et al.*, 2010). L'augmentation de l'espérance de vie, des soins de santé et des maladies non transmissibles ; les changements de régime et de mode de vie ; la diminution de la charge de morbidité liée aux maladies transmissibles, contribuent à la charge croissante du cancer sur le continent. Comparée à d'autres régions du monde, les ressources disponibles pour le dépistage précoce des cancers en Afrique sont associées à un accès plus difficile au traitement et au manque de registres de cancer bien établis et de haute qualité, basés sur la population. En outre, le cancer en Afrique varie considérablement selon le type, l'incidence, le stade du diagnostic, le taux de survie et le taux de mortalité. Cette différence est dûe en grande partie aux différences d'expositions aux principaux facteurs de risque, aux pratiques de détection (disponibilité des services de diagnostic et de dépistage), à la sensibilisation aux signes et symptômes précoces et à la disponibilité du traitement. Le cancer du sein au cours des dernières années a dépassé le cancer du col utérin dans la plupart des pays d'Afrique subsaharienne en tant que cancer le plus souvent diagnostiqué chez les femmes (Mackay *et al.*, 2006). Ceci est le résultat de la prise de conscience croissante de la maladie et de la prévalence de facteurs de risque de cancer du sein tels que la ménarche précoce, l'accouchement tardif, la parité faible et le poids corporel élevé associés au développement économique (Bray et Moller, 2006).

Les cancéreux vivant dans les pays plus développés représentent environ la moitié des cas de cancer du sein et 38% des décès, les taux étant généralement élevés en Amérique du Nord, en Australie, en Nouvelle-Zélande et dans le nord-ouest de l'Europe ; intermédiaires en Europe centrale et orientale, en Amérique latine et dans les Caraïbes ; et faibles dans la plupart des pays d'Afrique et d'Asie (Torre *et al.*, 2015). Cependant, ces chiffres ont changé avec la publication en 2018 du document du Centre international de recherche sur le cancer (CIRC) par l'Organisation mondiale de la santé (OMS) sur l'observation mondiale du cancer. Certains chiffres et résumés sont présentés plus loin dans ce document. Entre 1980 et la fin des années 1990, les taux d'incidence du cancer du sein ont augmenté d'environ 30% dans les pays occidentaux, probablement en raison de l'évolution des facteurs de reproduction et de l'utilisation du traitement hormonal ménopausique, puis dans les années 2000, en raison de l'accroissement du dépistage (Althuis *et al.*, 2005). Néanmoins, au début des années 2000, les taux d'incidence dans ces pays ont diminué, ce qui a été attribué à la réduction de l'utilisation du traitement hormonal ménopausique. La variation internationale des taux d'incidence du cancer du sein reflète les différences de disponibilité du dépistage précoce ainsi que des facteurs de risque (Torre *et al.*, 2015). En outre, des études régionales menées aux États-Unis montrent une variation du cancer du sein par race/ethnicité entre les femmes blanches et les

Afro-Américaines (DeSantis *et al.*, 2008 ; Iqbal *et al.*, 2015). Les résultats concernant la mortalité ont révélé une réduction des taux de mortalité chez les femmes blanches atteintes de la maladie dans les 50 états, contre 11 dans le cas des femmes afro-américaines, avec un taux accru dans deux états et un taux nivelé dans 24 états (DeSantis *et al.*, 2008).

Bien que le cancer du sein soit le plus répandu chez les femmes dans les pays développés, les taux de mortalité sont plus élevés dans les pays en développement que dans les régions développées, ce qui reflète les niveaux nationaux de développement social et économique tirés de la comparaison de l'Indice de Développement Humain (IDH) (Bray *et al.*, 2018). Les taux d'incidence varient de près de 3,5 fois dans les régions du monde, allant de 27,9% sur 100 000 en Afrique centrale à 48,9% sur 100 000 en Afrique du Nord (Figure 4) et à 94,2% en Australie/Nouvelle-Zélande (Ferlay *et al.*, 2018). Le graphique de la figure 5 illustre le nombre de nouveaux cas de cancer et de décès par cancer dans tous les groupes d'âge de femmes en Afrique.

En Afrique de l'Ouest, elle a une incidence de 37,3% et un taux de mortalité de 17,8%, bien supérieur au taux de mortalité de 12,6% en Australie et en Nouvelle-Zélande, avec le taux d'incidence le plus élevé de 94,2% (Bray, 2018).

Au Nigéria, l'incidence du cancer du sein a augmenté, en raison des changements démographiques tels qu'une espérance de vie plus longue, une meilleure notification de la maladie et un meilleur accès aux soins cliniques. Néanmoins, selon les dernières statistiques de GLOBOCAN pour 2018, le pays enregistre 37% de nouveaux cas de cancer du sein et 28% de décès par rapport aux autres cancers (la Figure 6 montre l'incidence et le taux de mortalité des cancers au Nigéria). Au Sénégal, le cancer du sein vient en second lieu avec 26% et 18% des nouveaux cas de cancer et des décès en 2018 par rapport aux autres types de cancer (la Figure 7).

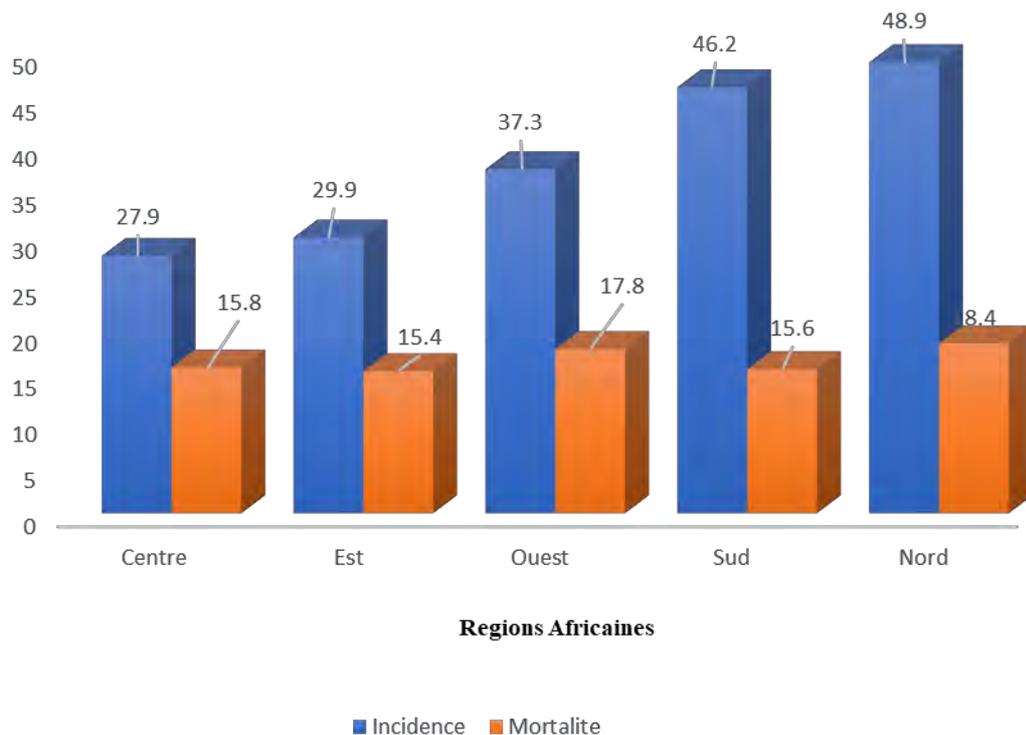
### **I.1. Développement du cancer du sein**

L'apparition du carcinome est un processus en plusieurs étapes qui commence par la transformation progressive d'une cellule normale en une cellule maligne et implique l'acquisition de divers phénotypes cellulaires par l'accumulation de multiples mutations génétiques (Abdel-Rahman, 2008). Les cancers sont des maladies chroniques à caractères changeants, ponctuées d'épisodes rares de populations accélérées, essentiellement à évolution lente, qui se précipitent parfois vers de nouveaux paramètres plus performants. Avant, les renouvellements tissulaires, ils grossissent et se diversifient (diversité clonale des cellules

cancéreuses) suffisamment vite pour acquérir les propriétés de migration et d'invasion nécessaires à la métastase. Dans certains cas, il se forme des sous-clones résistants au traitement ou pouvant se développer après le traitement (Song *et al.*, 2015). Il existe de nombreuses mutations qui permettent à une cellule cancéreuse une adaptation à un effet spécifique dans un environnement tumoral variable. Ainsi, chaque tumeur subit sa propre évolution darwinienne en suivant des mécanismes complexes et élaborés qui la différencient de toutes les autres tumeurs (Ogino *et al.*, 2012).

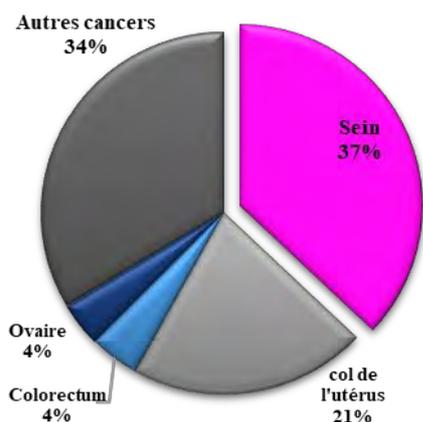
Une évolution convergente se produit, de sorte qu'un nombre limité de chemin évolutif soit parcouru de manière répétée par une tumeur provenant de patients différents. Ceci conduit à des repères majeurs reconnaissables se reproduisant dans une tumeur différente. Mais dans les cancers du sein, de la prostate et du pancréas, il a été démontré que ces cellules se disséminent bien avant que la colonisation métastatique ne devienne cliniquement évidente (Husemann *et al.*, 2008 ; Rhim *et al.*, 2012). Par conséquent, les métastases et la réponse au traitement prolongeraient, diversifieraient ou remanieraient les trajectoires évolutives du cancer du sein (Ding *et al.*, 2010).

Le cancer du sein est une maladie multifactorielle qui manifeste clairement l'hétérogénéité du comportement clinique et de la réponse thérapeutique (Yates et Campbell, 2012). Cela se reflète dans les variations morphologiques et moléculaires sous-jacentes, avec tout un éventail de caractéristiques histologiques et de marqueurs pathologiques moléculaires utiles pour prédire les résultats cliniques et sélectionner le traitement approprié (Subramaniam et Isaacs, 2005). Malgré diverses incertitudes concernant l'initiation du cancer du sein, des études génétiques, épidémiologiques et cliniques ont montré que certaines caractéristiques biologiques et sociales étaient des facteurs de risque liés à la maladie (Pike *et al.*, 1993 ; Bernstein et Ross, 1993 ; Mettlin, 1999 ; Hu *et al.*, 2000). Environ 50% de l'incidence du cancer du sein peut être attribuée à des facteurs de risque physiologiques ou comportementaux. Les facteurs de risque génétiques connus qui représentent 5 à 10% des cas de cancer du sein, présentent des mutations dans les gènes prédisposants rares du cancer du sein à forte pénétrance tels que *BRCA-1* et *BRCA-2*. Ces gènes sont responsables de 16 à 25% de la composante héréditaire du cancer du sein (Madigan *et al.*, 1995 ; Tonin, 2000 ; Beggs et Hodgson, 2009).

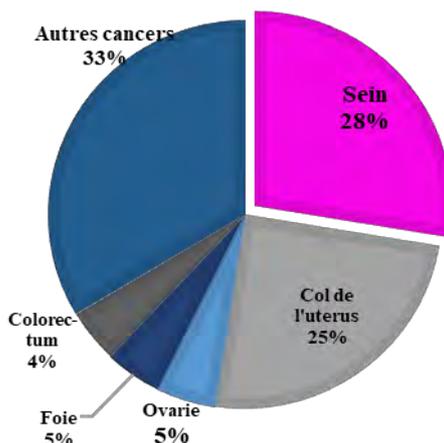


**Figure 5:** Diagramme en barres illustrant le nombre de nouveaux cas et de décès dûs au cancer du sein chez les femmes de tout âge dans différentes régions d'Afrique (Source : Globocan, 2018).

**Figure 6a : Nouveaux cas**

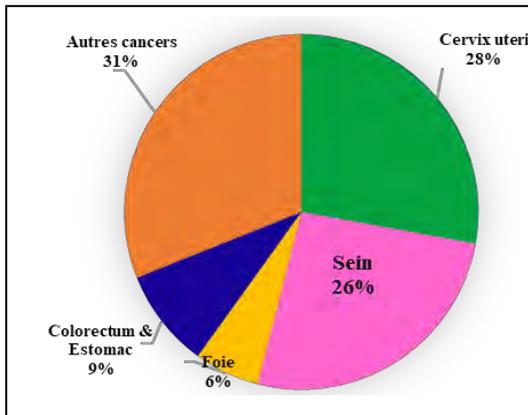


**Figure 6b : Décès**

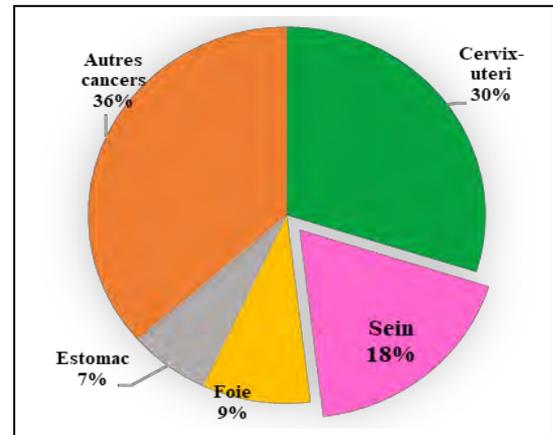


**Figure 6:** Graphique en camembert montrant le pourcentage de nouveaux cas de cancer et de décès chez toutes les femmes au Nigeria en 2018 (Source : Globocan, 2018).

**Figure 7a : Nouveaux cas**



**Figure 7b : Décès**



**Figure 7:** Diagramme à secteurs illustrant le pourcentage de nouveaux cas de cancer et de décès chez toutes les femmes au Sénégal en 2018 (Source : Globocan, 2018).

Certains des effets génétiques, environnementaux ou combinés (combinaison d'interactions génétiques et environnementales) associés au risque de cancer du sein incluent : les antécédents familiaux du cancer du sein, des ovaires ou de l'endomètre ; l'histoire individuelle de certains types de maladies mammaires bénignes ; la vieillesse ; le statut socio-économique plus élevé ; l'exposition excessive aux rayonnements ionisants ; la taille adulte ; la consommation d'alcool. En outre, divers facteurs endocrinologiques et reproductifs sont également associés à des facteurs de risque du cancer du sein, notamment la ménarche précoce, la nulliparité ou le premier accouchement tardif, l'allaitement au sein de courte durée, la ménopause tardive, l'obésité post-ménopausique, l'utilisation prolongée de contraceptifs oraux et l'utilisation prolongée d'œstrogènes substitutifs (Bernstein et Ross, 1993 ; Pike *et al.*, 1993 ; Mettlin, 1999 ; Hu *et al.*, 2000). D'autres facteurs sont aussi incriminés tels que le tabagisme, certains aspects nutritionnels (consommation de viande et de graisse), l'activité physique et le stress psychologique (Bleiker et Ploeg, 1999 ; Gerber *et al.*, 2003).

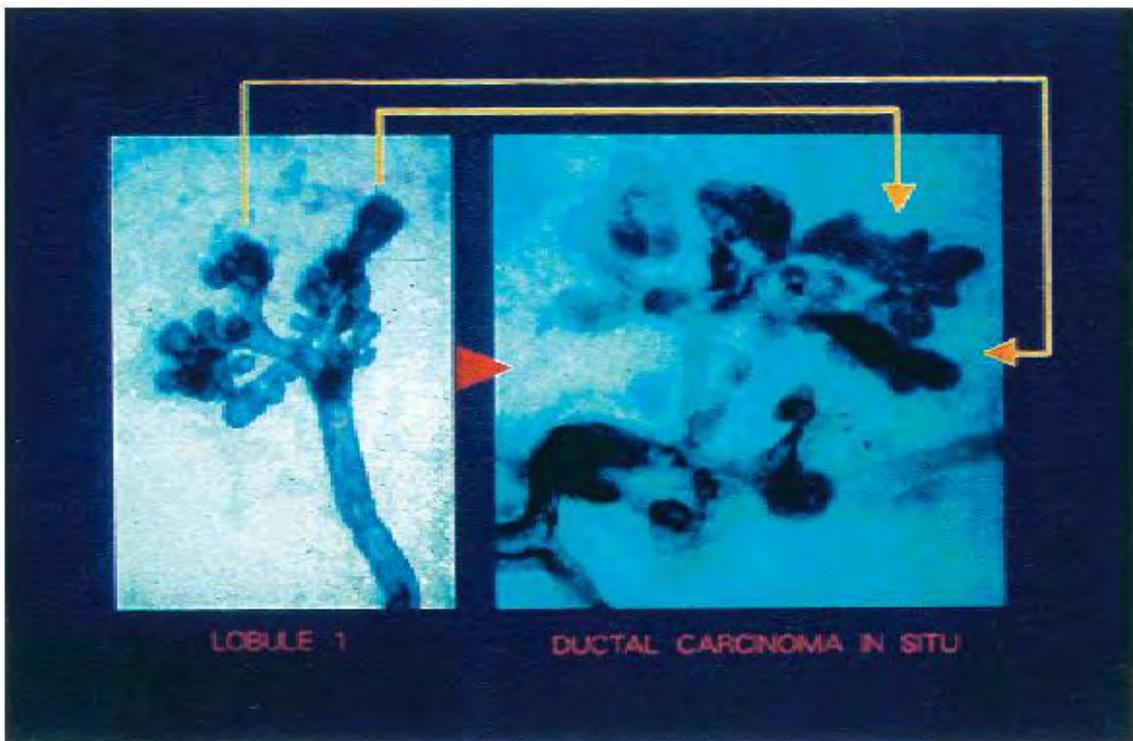
Concernant les facteurs hormonaux, les œstrogènes ont été identifiés comme un facteur attribuable majeur dans l'étiologie du cancer du sein avec trois mécanismes considérés comme responsables de son rôle (Henderson *et al.*, 1988 ; Russo *et al.*, 2000). Le premier est l'activité hormonale médiée par le récepteur associé à la stimulation de la prolifération cellulaire entraînant une accumulation accrue de dommages génétiques conduisant à une carcinogénèse (Adlercreutz *et al.*, 1994 ; Nandi *et al.*, 1995) ; puis il y a l'activation métabolique médiée par le cytochrome P450 qui stimule les effets génotoxiques directs par le biais de taux de mutation accrus (Schneider *et al.*, 1982 ; Bradlow *et al.*, 1986 ; Liehr *et al.*, 1986 ; Niwa *et al.*, 1989 ;

Malins et Haimanot, 1991 ; Malins *et al.*, 1993 ; Osborne *et al.*, 1993 ; Spink *et al.*, 1994 ; Liehr et Ricci, 1996) ; et enfin l'instabilité génomique dans le système de réparation de l'ADN, entraînant une mutation génomique cruciale pour la tumorigène induite par les œstrogènes (Yan et Roy, 1995). Ainsi, la teneur en récepteur d'œstrogène alpha (ER $\alpha$ ) d'une tumeur est un facteur d'importance pronostique (Molino *et al.*, 1997 ; Khan *et al.*, 1998). Les cancers ER $\alpha$  positifs et ER $\alpha$  négatifs sont de différentes entités, bien que les cancers du sein ER $\alpha$  négatifs puissent résulter de la perte de la capacité des cellules à produire des ER $\alpha$  au cours de l'évolution clinique des tumeurs ER $\alpha$  positives (Moolgavkar *et al.*, 1980 ; Habel et Stanford, 1993).

L'initiation du cancer du sein provient d'une prolifération cellulaire incontrôlée et/ou d'une défaillance des mécanismes de mort cellulaire programmée (apoptose) en raison de dommages génétiques cumulatifs. Celles-ci conduisent à des altérations génétiques activant les proto-oncogènes et inactivant les gènes suppresseurs de tumeurs (Van de Vijver, 1993 ; Black, 1994). Ces altérations sont héritées ou acquises sous forme de mutations germinales et somatiques. Bien que les mutations somatiques puissent résulter d'une exposition à des cancérrogènes environnementaux (physiques ou chimiques) et/ou biologiques (par exemple, des virus) (Hu *et al.*, 2000), l'agent causal et l'heure du début du cancer du sein sont inconnus dans la population humaine (Russo *et al.*, 1993a ; 1996). Des études associant le développement normal du sein et la pathogenèse de la carcinogenèse mammaire tant expérimentale que spontanée ont montré que la structure la plus indifférenciée trouvée dans le sein des jeunes femmes nullipares est équivalente à l'unité lobulaire du canal terminale qui est le site d'origine des carcinomes (Figure 8) (Wellings *et al.*, 1975 ; Russo et Russo, 1987 ; 1996). Ces structures indifférenciées sont à l'origine des néoplasmes les plus agressifs et sont plus nombreuses chez les femmes nullipares qui présentent un risque plus élevé de développer un cancer du sein (Russo *et al.*, 1992).

Selon Russo *et al.*, (2000), les tissus mammaires des femmes qui ont accouché étaient principalement constitués de Lob 3 et d'un très faible pourcentage de Lob 1. Cependant, les femmes ayant développé un cancer du sein avaient également un Lob 1 comme leur structure principale qui était similaire à celles des femmes nullipares. De plus, les femmes qui avaient développé un cancer du sein avaient des antécédents de première grossesse à terme tardive ou des antécédents familiaux du cancer du sein. L'augmentation du taux d'œstrogènes joue un rôle dans le développement du cancer du sein en induisant un stress génotoxique et des mutations dans le tissu mammaire (Yager et Davidson, 2006 ; Germain, 2011). De même, il a été observé

que Lob 1 contenait au moins trois types de cellules : les cellules positives pour ER $\alpha$  qui ne proliféraient pas, les cellules négatives pour ER $\alpha$  capables de proliférer et une faible proportion de cellules positives pour ER $\alpha$  pouvant proliférer. Cela montre que l'œstrogène pourrait stimuler les cellules positives pour ER $\alpha$  à produire un facteur de croissance susceptible de stimuler la prolifération de cellules voisines positives pour ER $\alpha$  (Figure 8) (Russo *et al.*, 1993b). La théorie de la hiérarchie cellulaire au niveau de la glande mammaire stipule qu'une cellule souche mammaire ER non différenciée donnera lieu à des cellules qui se différencieront davantage en cellules matures luminales, lobulaires ou myoépithéliales (Almendro et Fuster, 2011).



**Figure 8:** Préparations complètes portant un lobule de type 1 (Lob 1) et un carcinome du canal *in situ* impliquant plusieurs canaux terminaux (Russo *et al.*, 2000).

Ceci est montré sur la Figure 9, indiquant que trois types différents de cellules sont considérées comme présentes dans l'épithélium mammaire, à savoir : les cellules proliférantes négatives pour le récepteur des œstrogènes (ER-) (Ki67 +), les cellules ER positives (ER +) qui ne prolifèrent pas (Ki67-) et une faible proportion de celles-ci et cellules Ki67 +. L'œstrogène pourrait stimuler les cellules ER + à produire un facteur de croissance qui, à son tour, stimule les cellules ER- voisines capables de proliférer. Les cellules ER + Ki67 + qui peuvent proliférer et qui pourraient être stimulées par l'œstrogène sont à l'origine des cellules filles ER

+ ou probablement des tumeurs. Les cellules ER- peuvent se convertir en cellules ER + lors de la transformation néoplasique (Russo *et al.*, 1993b).

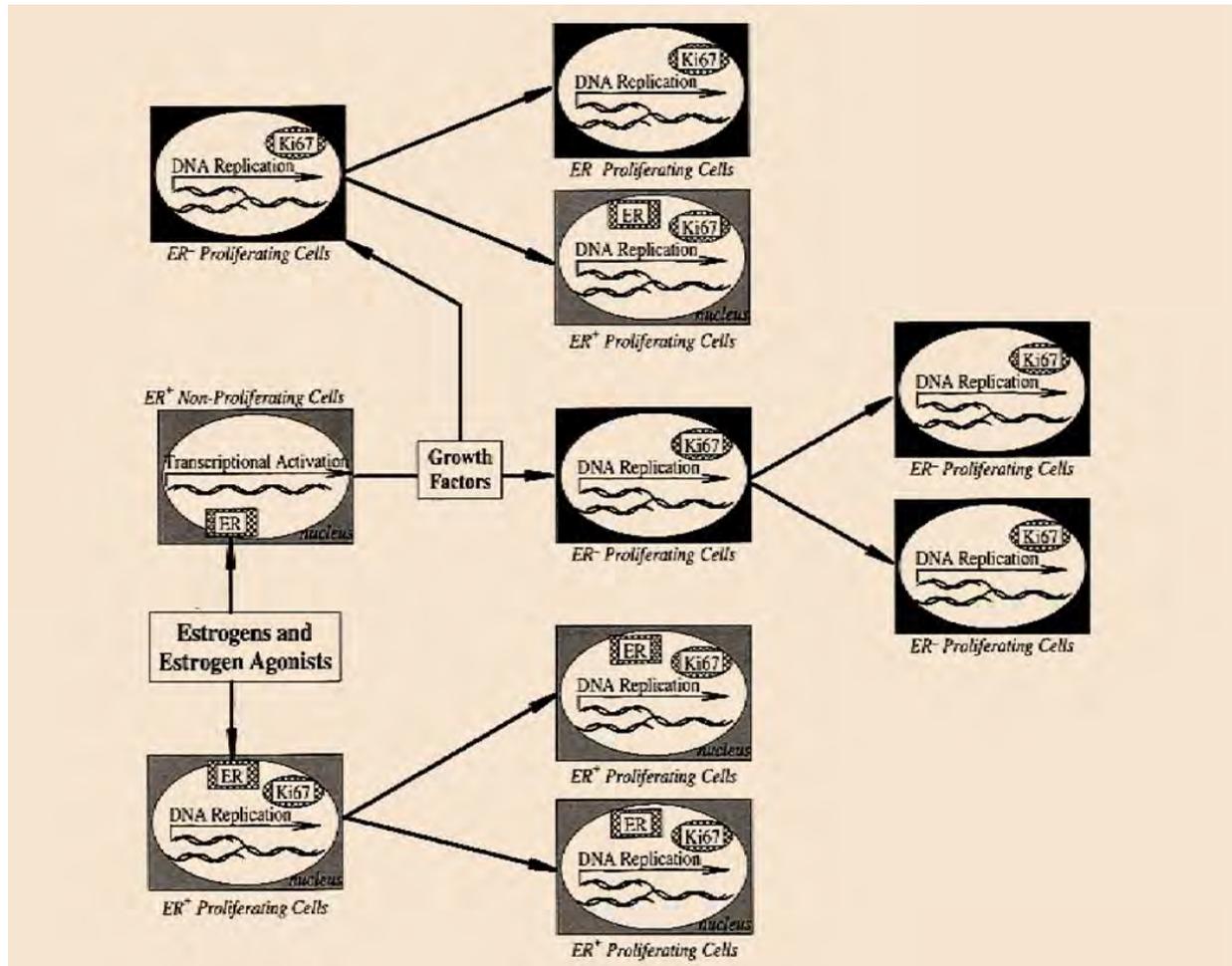
## **I.2. Evolution génétique et hétérogénéité moléculaire du génome du cancer**

La progression du cancer est un processus évolutif et dynamique d'adaptation clonale à des microenvironnements en évolution. Au cours de ce processus des cellules tumorales dotées d'un avantage sélectif commencent à apparaître lorsque des cellules clonales dominantes et défavorisées sont sélectionnées, donnant ainsi à la tumeur une composition unique (Waghorne *et al.*, 1988 ; Price *et al.*, 1990 ; Swanton, 2012). Lors du développement de la tumeur, les cellules cancéreuses acquièrent diverses aberrations, notamment les mutations passagères (neutres) et conductrices (avantageuses). En raison de la sélection et de l'expansion clonale, plusieurs sous-clones génétiquement distincts apparaissent et évoluent selon un schéma d'évolution ramifiée décrit pour divers types de tumeurs solides (Diaz-Cano *et al.*, 2000 ; Campbell *et al.*, 2010 ; Ding *et al.*, 2010 ; Yachida *et al.*, 2010 ; Gerlinger *et al.*, 2012 ; Nik-Zainal *et al.*, 2012a et 2012b ; Shah *et al.*, 2012 ; Baca *et al.*, 2013 ; Bashashati *et al.*, 2013 ; Gerlinger *et al.*, 2014 ; Nordentoft *et al.*, 2014). Cette évolution ramifiée correspond à plusieurs sous-clones qui présentent un avantage phénotypique dans un environnement spécifique et évoluent simultanément, ce qui conduit à une hétérogénéité intra-tumorale (Hiley *et al.*, 2014).

De plus, au cours des métastases tumorales, les cellules tumorales se détachent de la tumeur primaire vers des sites d'organes distants pour former une tumeur métastatique (Valastyan et Weinberg, 2011). Il a été observé dans les cancers du sein, de la prostate et du pancréas que ces cellules se propagent bien avant que la colonisation métastatique ne devienne cliniquement évidente (Husemann *et al.*, 2008 ; Rhim *et al.*, 2012). Un ensemencement polychoral et bidirectionnel entre différents sites tumoraux et une évolution parallèle ont été décrits lors de la propagation précoce et tardive de la tumeur (Klein, 2009 ; Gudem *et al.*, 2015). Les différents sites tumoraux sont constitués de paysages évolutifs distincts qui conduisent à une hétérogénéité inter métastatique (Bashashati *et al.*, 2013 ; Haffner *et al.*, 2013 ; Gudem *et al.*, 2015 ; Murtaza *et al.*, 2015).

L'existence de sous-populations distinctes de cellules cancéreuses au sein de tumeurs présentant une tumorigénicité, une résistance au traitement et une aptitude à métastaser différentes a été décrite dans des études pionnières sur l'hétérogénéité tumorale ainsi que sur leur implication clinique (Fidler et Kripke, 1977 ; Fidler, 1978 ; Miller *et al.*, 1983 ; Beca et Polyak, 2016). Bien que l'initiation de la plupart des tumeurs humaines provienne d'une seule

cellule, elles présentent de manière surprenante une hétérogénéité en ce qui concerne leur potentiel angiogénique et métastatique, leur degré de prolifération, leur expression des molécules de surface cellulaire et leur morphologie cellulaire au moment du diagnostic (Heppner, 1984 ; Heppner et Miller, 1983). Depuis la cellule fondatrice individuelle jusqu'à la détection clinique des tumeurs, il existe des expansions clonales consécutives et une acquisition constante d'altérations génétiques et épigénétiques ; événements qui contribuent à la génération d'une hétérogénéité intra-tumorale (Almendro et Fuster, 2011).



**Figure 9:** Représentation schématique des voies postulées des actions des œstrogènes sur les cellules épithéliales du sein (Russo *et al.*, 1993a).

La plasticité morphologique et épigénétique avec une variabilité semblable à celle des cellules souches et des caractéristiques cellulaires plus différenciées initie partiellement l'hétérogénéité de la tumeur avec de fortes preuves de clones génétiquement distincts dans la même tumeur (Park *et al.*, 2010). L'identification d'une hétérogénéité génétique et clonale importante dans les tumeurs et les carcinomes du sein, en particulier, pose un défi important en raison de la résistance thérapeutique et de la difficulté à réaliser un diagnostic efficace du cancer, à concevoir des schémas thérapeutiques efficaces et aussi à réaliser une médecine personnalisée

(Heppner et Miller, 1983 ; Merlo *et al.*, 2006 ; Marusyk et Polyak, 2010 ; 2016 ; Ellsworth *et al.*, 2017). Une hétérogénéité tumorale non génétique telle que l'hétérogénéité épigénétique et une hétérogénéité cellulaire (non héréditaire pour divers phénotypes) ont également été identifiées (Marusyk *et al.*, 2012 ; Almendro *et al.*, 2013).

Par ailleurs, des caractéristiques phénotypiques montrant une variabilité intra-tumorale significative d'une cellule à l'autre ont également été identifiées. Il s'agit de : l'activation des voies de signalisation, l'évasion de l'immunité anti tumorale, l'induction de la sénescence, la production de facteurs sécrétés, la migration, le potentiel métastatique, la capacité angiogénique, la constitution génétique, les médicaments et l'activation des voies métaboliques (Beca et Polyak, 2016). Par conséquent, une meilleure compréhension de la composition des tumeurs, de leur évolution au cours de la progression de la maladie et du traitement offrira de meilleures méthodes de diagnostic et de conception thérapeutique.

En 2016, Welch a présenté une perspective historique de quatre types d'hétérogénéité tumorale à partir des travaux de Hepper (1984), à savoir :

- L'hétérogénéité de la population ou différences entre les tumeurs de patients différents : on parle souvent d'hétérogénéité inter tumorale en faisant référence au caractère unique des tumeurs provenant du même type de cellule chez différents patients. Cela montre que même si la tumeur de chaque patient est unique, elle partage cependant certaines propriétés.
- L'hétérogénéité intra-tumorale, génétique ou intrinsèque : il s'agit d'une hétérogénéité spatiale au sein d'une masse tumorale unique et plus communément appelée hétérogénéité intra-tumorale.
- L'hétérogénéité de position : il s'agit principalement des locus de traits quantitatifs qui contrôlent les propriétés cellulaires (facteurs d'instabilité oncogéniques et génétiques) qui contribuent à la variabilité entre les patients.
- L'hétérogénéité temporelle : elle reflète l'évolution de la tumeur au fil du temps (croissance et développement) ou en réponse au traitement (Welch, 2016).

Néanmoins, les hétérogénéités inter-tumorale, intra-tumorale, positionnelle et temporelle peuvent avoir des origines et des conséquences cliniques différentes dans le cancer (Almendro et Fuster, 2011).

### **I.2.1. Hétérogénéité moléculaire dans le cancer du sein**

Le cancer du sein est une maladie extrêmement hétérogène. Cette hétérogénéité se manifeste dans presque toutes les composantes qui le caractérisent. Il se reflète dans les cellules tumorales à partir des caractéristiques moléculaires, de leurs caractéristiques fonctionnelles au moment du diagnostic, ainsi qu'au niveau histopathologie, clinique et moléculaire (Perou *et al.*, 2000 ; Sorlie *et al.*, 2001 ; Almendro et Fuster, 2011). Le cancer du sein est la première tumeur solide dans laquelle la classification de son hétérogénéité phénotypique pour les phénotypes cellulaires a été établie, ainsi que ses implications cliniques et thérapeutiques (Foote et Stewart, 1946 ; Hawkins *et al.*, 1980 ; Teixeira *et al.*, 1996 ; Klein *et al.*, 2002 ; Shipitsin *et al.*, 2007 ; Torres *et al.*, 2007 ; Young et Louis, 2011).

#### **I.2.1.1. Hétérogénéité inter-tumorale du cancer du sein**

L'hétérogénéité inter-tumorale du cancer du sein peut être observée dans les sous-types intrinsèques de cancer du sein observés chez différents patients ; luminal A, luminal B, luminal C, HER2 +, les sous-types normaux et basaux (Perou *et al.*, 2000 ; Sorlie *et al.*, 2001), et plus récemment, le sous-type de claudine faible (Herschkowitz *et al.*, 2007) (Figure 10). En raison de leur longue durée de vie et de la possibilité d'accumuler les aberrations génétiques nécessaires à la transformation, les cellules souches mammaires adultes ont été proposées comme cibles de la cancérogenèse. En conséquence, les sous-types moléculaires pourraient provenir de cellules d'origine similaires, mais une accumulation supplémentaire de modifications génétiques et épi génétiques se développerait en un sous-type ou un autre (Sorlie *et al.*, 2001).

##### **I.2.1.1.1. Sous-types moléculaires du cancer du sein**

Les marqueurs tumoraux sont les outils de classification des sous-types de cancer du sein à des fins de diagnostic et de traitement. Ces marqueurs sont les récepteurs aux œstrogènes (ER), à la progestérone (PR) et à HER2/neu (Daly et Olopade, 2015). Les tumeurs lumineales A sont reconnues par l'expression du récepteur aux œstrogènes (ER), la présence ou l'absence de récepteur à la progestérone (PR) et l'absence d'amplification de HER2. Les tumeurs lumineales B sont reconnues par l'expression du récepteur des œstrogènes (ER), la présence ou l'absence du récepteur de la progestérone (PR) et l'expression du récepteur du facteur de croissance épidermique humain 2 (HER2). Les tumeurs HER2 + sont identifiées par le manque d'expression de ER et de PR et par la présence d'une amplification 17q, la région

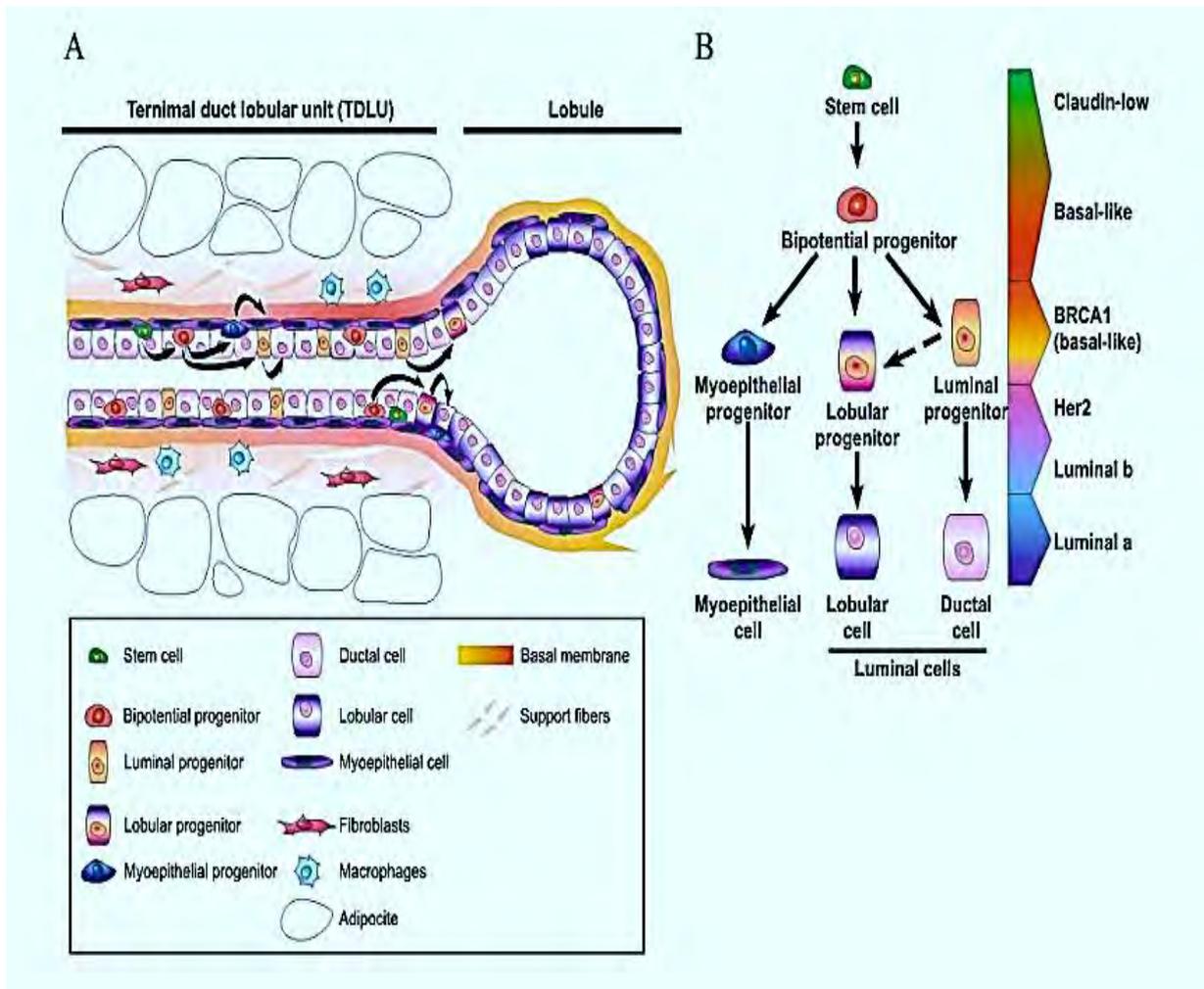
chromosomique qui code pour le récepteur de la tyrosine kinase HER2 qui favorise la croissance cellulaire. Les tumeurs basales du sein, également connues sous le nom de triple négatif, sont identifiées par l'absence de ER, de PR et de HER2, ainsi que par la surexpression de Her1 / récepteur du facteur de croissance épidermique (EGFR) et des cytokératines 5, 6, 14 et 17 (Perou *et al.*, 2000 ; Sorlie *et al.*, 2001 ; Carey *et al.*, 2006 ; Ademuyiwa *et al.*, 2011), tandis que le sous-type "claudin-low" enrichit les gènes associés au mésenchyme et aux cellules souches (Herschkowitz *et al.*, 2007; Prat *et al.*, 2010).

La présence de récepteurs aux œstrogènes et à la progestérone (ER / PR) décrit un sous-type de tumeur qui peut être ciblé par une thérapie de modulation hormonale et par des inhibiteurs de l'aromatase conduisant à un pronostic et à un traitement favorable. L'amplification ou la surexpression de HER2 renvoie à un sous-type distinct qui peut être ciblé par les antagonistes de HER2 (Carey *et al.*, 2006). Les sous-types de tumeurs triple négatifs (ER, PR et HER2) ont un pronostic et une survie médiocres, car ils ne répondent pas aux thérapies ciblées (Bauer *et al.*, 2007 ; Reis-Filho et Tutt, 2008 ; Podo *et al.*, 2010).

La figure 11 montre les voies menant à l'initiation de chaque sous-type moléculaire du cancer du sein. Le « A » montre la composition cellulaire de l'épithélium mammaire qui contient une couche luminale de cellules (canales ou lobulaires) entourée d'une couche basale de cellules myoépithéliales contractiles. Les différents progéniteurs restreints à la lignée qui existent dans l'épithélium mammaire se différencient en cellules matricielles canalaire, lobulaire et myoépithéliales matures (la différenciation cellulaire est indiquée par les flèches noires) qui sont entourées de différents types cellulaires et d'une membrane basale riche en fibres ainsi que d'un stroma. « B » illustre la hiérarchie épithéliale proposée dans la glande mammaire, qui indique qu'au cours du développement, une cellule souche du cancer du sein donnera naissance à des progéniteurs à potentiel double, qui se différencieront davantage en progéniteurs à lignée restreinte pour la génération de cellules canalaire, lobulaire et myoépithéliales. Sur la base des similitudes moléculaires entre les cellules normales et les cellules cancéreuses, les cellules de la glande mammaire à différents stades de différenciation seront les cellules d'origine des différents sous-types de BC (Almendo et Fuster, 2011).

<b>Generators of heterogeneity</b>
Genetic instability Stem cell renewal
<b>Types of heterogeneity</b>
Population Genetic/intrinsic Position Epigenetic
<b>Influencers of heterogeneity</b>
Cell-cell interactions tumor-tumor tumor-host Cell-matrix interactions Cell-therapy interactions
<b>Consequences of heterogeneity</b>
Neoplastic evolution and progression Requirement for multi-modal therapy Development of resistance Immune escape

**Figure 10:** Principes concernant les origines, les types et les conséquences de l'hétérogénéité néoplasique (Welch, 2016).



**Figure 11:** Hiérarchie de l'épithélium mammaire normal et des cellules de sous-types intrinsèques du cancer du sein (Almendro et Fuster, 2011).

### I.2.1.2. Hétérogénéité Intra-tumorale du Cancer du Sein

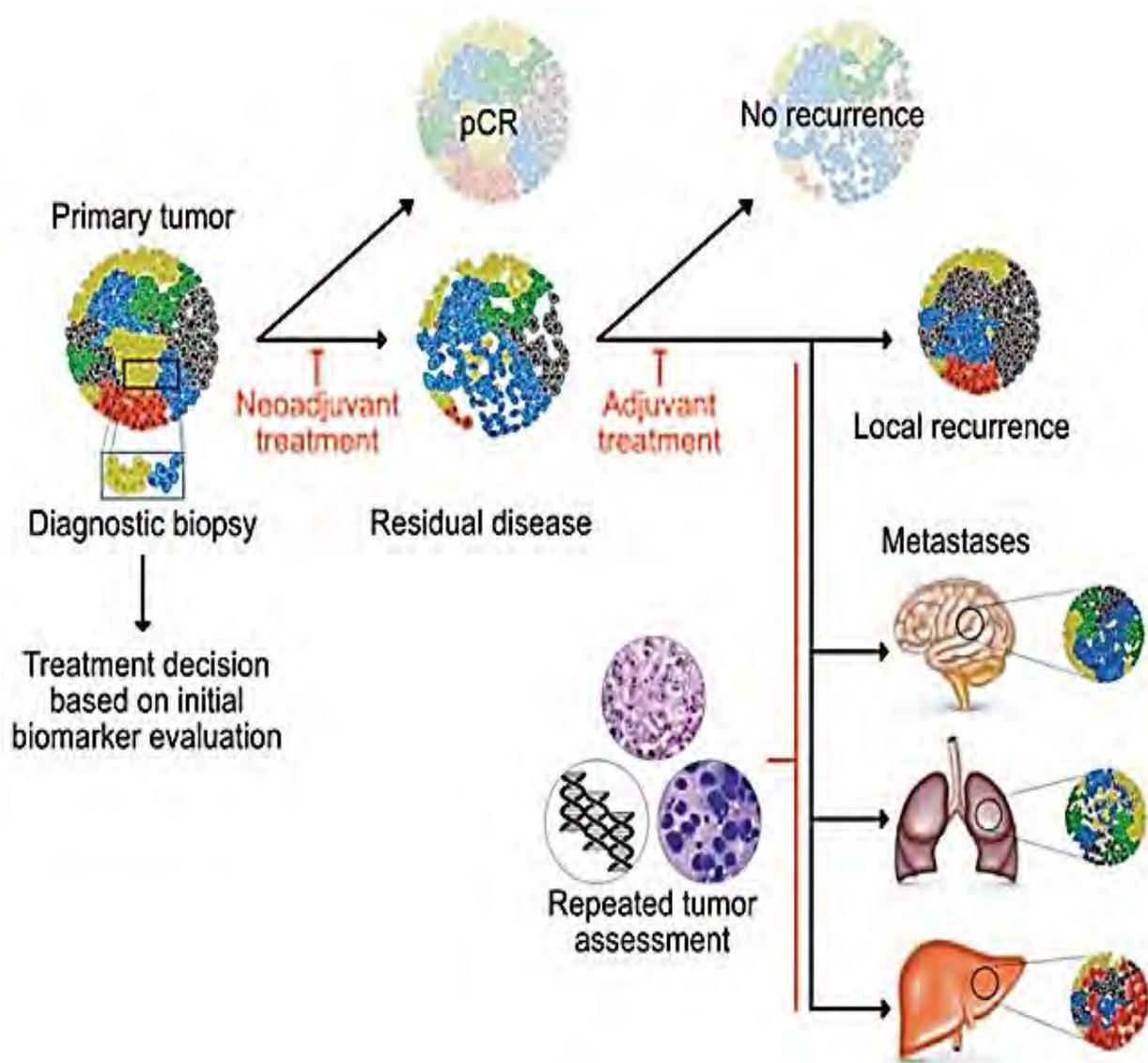
Selon Brown et Fee (2006), la découverte de l'hétérogénéité intra-tumorale au XIXe siècle peut être attribuée à Rudolph Carl Virchow, qui a observé la présence de phénotypes cellulaires distincts dans les tumeurs. L'affichage de l'hétérogénéité intra-tumorale est observée dans la majorité des tumeurs au moment du diagnostic et la figure 12 montre l'hétérogénéité intra-tumorale dans le diagnostic et le traitement du cancer. La sélection clonale et l'hypothèse des cellules souches ont été identifiées comme étant à l'origine de l'hétérogénéité intra-tumorale (Almendro et Fuster, 2011). L'hypothèse de sélection clonale stipule que les clones offrant la meilleure forme physique seront sélectionnés pour un microenvironnement tumoral particulier au cours de l'évolution de la tumeur, tandis que l'hypothèse relative aux cellules souches stipule que seule une petite partie des cellules tumorales (cellules souches tumorales) est responsable du maintien et de la progression des tumeurs, quelle que soit leurs cellules d'origine, en raison de leur capacité unique d'auto-renouvellement et de différenciation des

cellules tumorales différenciées (Marusyk et Polyak, 2010). Ainsi, l'hétérogénéité intra-tumorale est acceptée comme le résultat de la différenciation de cellules de type tige du cancer du sein en liaison avec la sélection clonale de sous-populations cellulaires particulières en raison des études montrant la présence de ces deux processus dans la voie le cancer du sein (Marusyk et Polyak, 2010).

### **I.2.1.3. Cellules Tumorales Circulantes et Disséminées (CTC)**

Les cellules cancéreuses viables, libérées d'un carcinome primitif et circulant dans le système vasculaire, portant les altérations génétiques trouvées dans la tumeur primitive, sont appelées CTC (Fehm *et al.*, 2002). Parce que les CTC sont supposées initier la croissance de dépôts métastatiques dans des organes distants, la capacité de les détecter et de les quantifier dans le sang périphérique des patientes atteintes d'un cancer du sein peut fournir un marqueur facilement accessible avec une valeur prédictive et pronostique élevée. La possibilité de quantifier les CTC par une collecte de sang peu invasive au cours du traitement a conduit à l'idée d'utiliser les tests de CTC comme « biopsie liquide ». En 2004, une étude menée par Allard *et al.*, (2004) auprès de 422 patientes atteintes d'un cancer du sein métastatique et visant à évaluer la prévalence des CTC dans le sang des patientes métastatiques par rapport aux individus en bonne santé a montré une nette distinction de la quantité de CTC entre les patients et les témoins.

Le diagnostic absolu du cancer et l'action sur les thérapies sont souvent compliqués par la complexité complexe de l'hétérogénéité intra-tumorale des phénotypes cellulaires, selon laquelle une seule biopsie peut ne pas montrer de manière adéquate la composition phénotypique de la tumeur entière en raison de l'hétérogénéité spatiale présente dans la tumeur. Par conséquent, il est nécessaire d'échantillonner successivement les caractéristiques de la tumeur par biopsies, FNA (aspiration à l'aiguille fine) ou ctDNA pour guider une prise de décision équilibrée au cours du traitement, car le diagnostic posé au stade initial de la présentation risque de ne pas être représentatif de l'évolution incessante de la tumeur au cours du traitement. Nous avons besoin de biomarqueurs fiables et robustes pouvant montrer avec précision l'hétérogénéité intra-tumorale menant à de meilleures estimations prédictives des réponses au traitement et des voies probables d'évolution de la tumeur [Beca et Polyak dans Stearns (eds.), 2016].



**Figure 12:** Hétérogénéité intra-tumorale dans le diagnostic et le traitement du cancer [Beca et Polyak dans Stearns (eds.), 2016].

La figure 12 illustre l'hétérogénéité intra-tumorale dans le diagnostic et le traitement du cancer. Globalement, une moyenne de  $84 \pm 885$  CTC était observée dans 7,5 ml de sang. 26% des patients avaient un nombre de CTC supérieur ou égale à 5 CTC dans 7,5 ml de sang alors que les témoins avaient tous un nombre de CTC dans 7,5 ml de sang inférieur à 5. D'autres études ont montré que la présence et / ou l'abondance de CTC dans le sang total constituait un facteur de prévision indépendant et fiable de la morbidité, de la mortalité par maladie, de la réponse défavorable au traitement et de la récurrence précoce de nombreux types de cancer, y compris le cancer du sein, (Xenidis *et al.*, 2006 ; Smerage *et al.*, 2014 ; Gong *et al.*, 2016 ; Janni *et al.*, 2016).

Actuellement, dans le cadre des pratiques cliniques standard, les statuts ER, PR et HER2 sont évalués dans les carcinomes mammaires primaires, mais les données du séquençage du génome

entier suggèrent l'évolution des cellules tumorales disséminées (DTC) à partir de populations sous-clonales de cellules dans la tumeur primaire après diffusion (Møller *et al.*, 2013). Par conséquent, les bio-marqueurs du cancer et la variation génétique peuvent évoluer au cours de la progression de la maladie et une discordance moléculaire importante avec des implications thérapeutiques considérables peut se développer entre la tumeur primitive et les CTC correspondants (Fehm *et al.*, 2010 ; Sieuwerts *et al.*, 2011). La présence de CTC exprimant HER2 a été associée à une augmentation du risque de mortalité, ce qui permet de prévoir une survie sans maladie et une survie globale médiocres (Ellsworth *et al.*, 2017). L'hétérogénéité moléculaire et fonctionnelle observées parmi les CTC est la raison pour laquelle les comptages de CTC n'ont pas été cliniquement utiles pour prédire les avantages potentiels du traitement (Bidard *et al.*, 2016). L'hétérogénéité génomique, y compris la perte d'hétérozygotie et de variantes de l'ADN du gène phosphatidylinositol-4,5-bisphosphate 3-kinase, sous-unité catalytique alpha (PIK3CA), a été détectée chez des patientes atteintes d'un cancer du sein métastatique individuel (Pestrin *et al.*, 2015).

Des études, réalisées avec les outils de séquençage de nouvelle génération, portant sur des gènes liés au cancer, ont également observé chez des CTC du même patient et entre des CTC de tumeur primitive appariée des tumeurs du sein ER + / HER2 et triple négatives (Luca *et al.*, 2016). Des études examinant l'expression de gènes et de protéines dans des CTC simples ou groupés ont révélé une hétérogénéité marquée dans l'expression de bio-marqueurs tumoraux connus, même en comparant les CTC d'une seule ponction veineuse (Onstenk *et al.*, 2013). À partir des données observées, nous pouvons déduire que des informations utiles pour la conception de thérapies ciblées et le suivi de la réponse thérapeutique peuvent être obtenues à partir de l'utilisation d'une biopsie liquide pour évaluer les signatures génomiques des cellules cancéreuses en circulation chez tous les patients (Ellsworth *et al.*, 2017). De même, l'hétérogénéité des cellules tumorales circulantes et disséminées peut constituer un moyen par lequel les patients développent une résistance à l'endocrinothérapie ; par conséquent, des informations cliniques vitales pour un traitement amélioré peuvent être obtenues à partir de l'analyse des cellules tumorales circulantes et disséminées (Ellsworth *et al.*, 2017). On peut donc observer que la charge tumorale totale peut être plus précisément quantifiée en examinant plusieurs CTC/DTC chez chaque patient en raison de la grande hétérogénéité des cellules individuelles (Powell *et al.*, 2012 ; Babayan *et al.*, 2013 ; Gasch *et al.*, 2015).

Au-delà du nombre de CTC pour prédire la survie et les récurrences, les caractéristiques fonctionnelles des CTC sont entrain d'être identifiées pour adapter le traitement et surveiller la

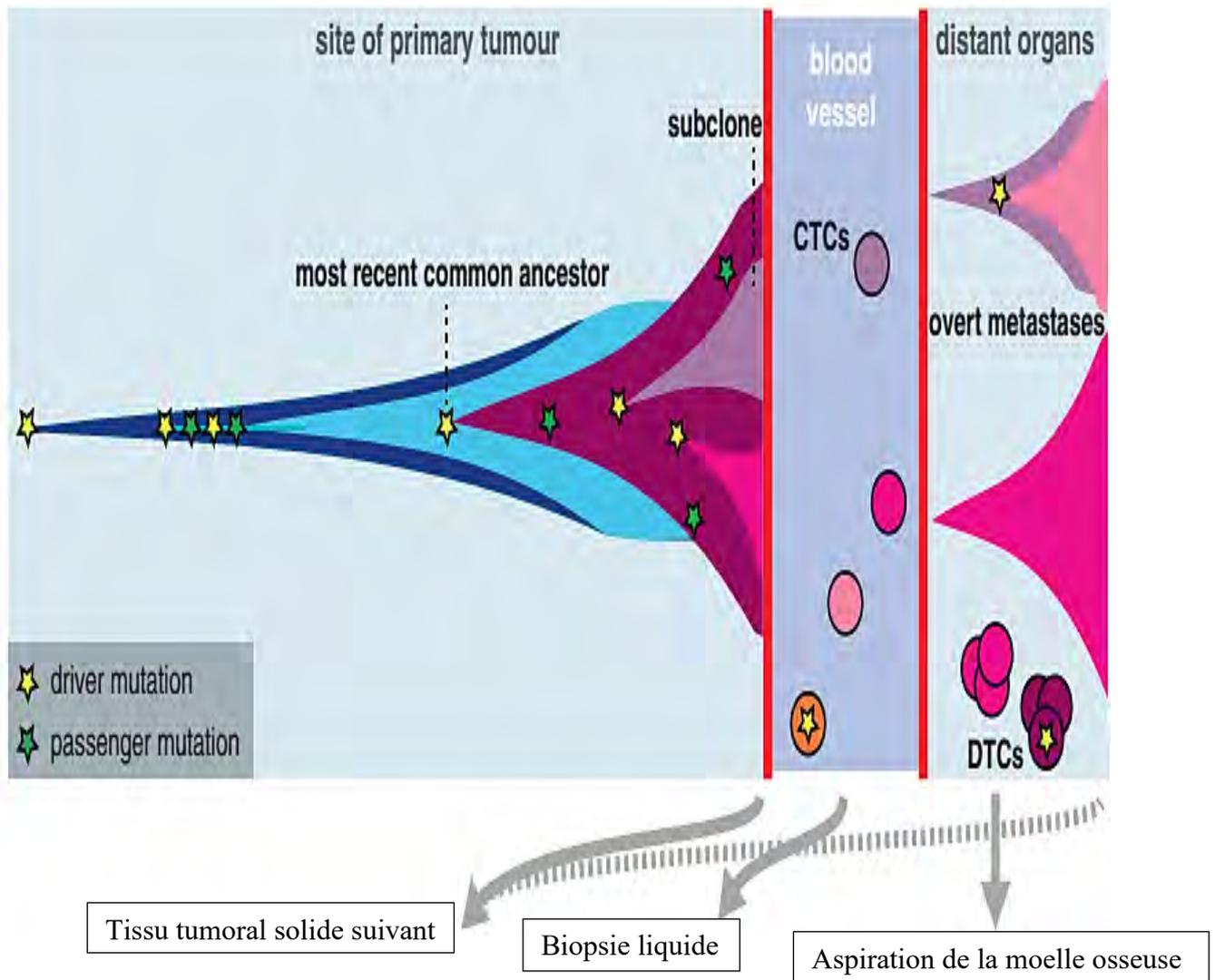
progression de la maladie (Paoletti *et al.*, 2015). La plupart des études sur les CTC dans le sang périphérique et/ou les DTC dans la moelle osseuse ont mis l'accent sur l'importance pronostique de la charge de CTC/DTC, mais des rapports récents montrent que la caractérisation moléculaire de ces cellules est essentielle à l'optimisation des régimes thérapeutiques. La complexité génomique globale et les mutations spécifiques du conducteur présentes dans les cellules disséminées peuvent constituer un facteur important d'échec du traitement. Par conséquent, l'évaluation des signatures génomiques de nombreuses cellules cancéreuses en circulation chez tous les patients peut être plus informative que les biopsies traditionnelles de la tumeur primitive pour la conception de thérapies ciblées et le suivi de la réponse thérapeutique (Ellsworth *et al.*, 2017).

La figure 13 illustre le processus de progression tumorale. Les cancers sont dûs à l'acquisition de mutations du conducteur, qui entraînent des expansions clonales successives de cellules tumorales naissantes. Les mutations des pilotes qui se produisent après l'émergence d'un ancêtre commun le plus récent donneront lieu à des sous-clones de tumeurs. Les tumeurs solides libèrent également des cellules dans le sang du patient (cellules tumorales circulantes ou CTC) et des cellules se disséminant dans des organes distants (cellules tumorales disséminées ou DTC), susceptibles de provoquer des métastases apparentes (Van Loo et Voet, 2014).

#### **1.2.1.3.1. Transition Epithélio-Mésenchymale dans les CTC**

La transition épithélio-mésenchymateuse (EMT) est un processus par lequel les cellules épithéliales subissent des modifications biochimiques leur faisant perdre leur polarité cellulaire et leur adhésivité cellule-cellule, acquérant ainsi les propriétés couramment observées dans les cellules mésenchymateuses, telles que la motilité, l'envahissement et la résistance à l'apoptose (Kalluri et Weinberg, 2009). Les EMT jouent un rôle important dans la progression des cancers épithéliaux tels que le cancer du sein et peuvent être induits par un large éventail de cytokines et de facteurs de croissance tels que le facteur de croissance transformant (TGF) - $\beta$  (Singh et Settleman, 2010). La progression tumorale est facilitée par l'activation du programme EMT à travers la génération de cellules stromales activées qui dirigent la progression du cancer par le biais de modifications de la structure et de la biochimie du microenvironnement tumoral. L'augmentation de la malignité du phénotype des cellules souches peut être stimulée par les médiateurs de l'EMT ainsi que par le déclenchement du mécanisme invasif et anti-apoptotique (Nistico *et al.*, 2012). Les cellules qui se disséminent à partir du carcinome primaire subissent

une activation aberrante de l'EMT, perdent leurs caractéristiques épithéliales et acquièrent davantage de phénotypes de type mésenchymateux (Bonnomet *et al.*, 2010).



**Figure 13:** Voie vers la cancérogénèse (Van Loo et Voet, 2014).

Les CTC isolés de patients atteints d'un cancer du sein présentent une variabilité importante dans l'expression de la molécule d'adhésion des cellules épithéliales (EpCAM) et d'autres marqueurs épithéliaux. Les patients atteints d'un cancer du sein métastatique peuvent être porteurs de plusieurs populations de CTC définies par la présence de marqueurs épithéliaux, et de marqueurs mésenchymateux. L'analyse des CTC capturées dans le sang total de patientes atteintes d'un cancer du sein a défini cinq catégories de cellules présentant exclusivement des caractéristiques épithéliales, intermédiaires et exclusivement mésenchymateuses (Yu *et al.*, 2013). La fraction de CTC co-exprimant les marqueurs épithéliaux et mésenchymateux a été associée de manière significative à une survie sans progression et à une survie globale plus médiocre (Bulfoni *et al.*, 2016). Les CTC en cours de TEM ou en transition vers un phénotype

de cellules mésenchymateuses peuvent fournir des informations importantes sur la progression tumorale et les métastases, mais les technologies actuelles d'enrichissement du CTC qui ciblent spécifiquement les marqueurs épithéliaux pour différencier les CTC des autres types de cellules peuvent sous-estimer le nombre de CTC et potentiellement manquer des sous-populations critiques de cellules en cours ou ayant terminé la TEM (Hyun *et al.*, 2016).

En raison de l'impossibilité de détecter complètement les populations hétérogènes dans les CTC, une protéine vimentine, à la surface des cellules, a montré des résultats prometteurs dans les CTC EMT chez les patients atteints d'un cancer épithélial visant à améliorer la détection et l'isolement des CTC présentant des caractéristiques mésenchymateuses (Satelli *et al.*, 2015). La détection des CTC chez les patientes atteintes d'un cancer du sein et soumises à une chimiothérapie néoadjuvante a été considérablement améliorée par la combinaison de la molécule d'adhésion cellulaire du mélanome (MCAM) enrichie en EpCAM que lorsque EpCAM est utilisé seul (Onstenk *et al.*, 2015). Par conséquent, dans le cadre d'une méthode saine de biopsie liquide, ces études montrent des sous-populations cliniquement pertinentes de CTC (Ellsworth *et al.*, 2017).

#### **1.2.1.3.2. ADN libre et tumoral circulants des tumeurs**

Les fragments d'ADN non encapsulés qui sont libérés dans la circulation sanguine et d'autres liquides corporels au cours de l'apoptose, de la nécrose et de la phagocytose des macrophages sont connus sous le nom d'ADN libre et circulant (cfADN) (Ellsworth *et al.*, 2017). Une partie de l'ADNc appelé ADN tumoral circulant (ADNct) est rapidement éliminée du corps quelques heures après sa production, bien qu'il y soit plus d'abondance que les CTC et que les cas observés chez les patientes atteintes d'un cancer du sein représentent la composition mutationnelle de CTC individuels (Shaw *et al.*, 2016). Leon *et al.*, (1977) ont observé que le cfDNA était plus présent chez les patients cancéreux, en particulier ceux atteints de métastases ( $209 \pm 39$  ng / ml) que chez les patients non métastatiques ( $100 \pm 30$  ng/ml,  $p < 0,02$ ) que chez les patients en bonne santé.

Bien qu'il y ait eu cohérence entre la mutation du gène *PIK3CA* dans l'ADN circulant de tumeur primaire correspondant et dans le plasma provenant de patients atteints d'un cancer du sein (Board *et al.*, 2010 ; Higgins *et al.*, 2012), l'hétérogénéité génomique avec des conséquences cliniques potentiellement importantes entre les carcinomes et l'ADNc ont été identifiés via le séquençage de nouvelle génération d'ADNc (Butler *et al.*, 2015). Par conséquent, cela peut indiquer un outil précieux pour surveiller la progression du cancer du

sein en vue d'un échantillonnage en temps réel non invasif de la réponse au traitement chez les femmes atteintes d'un cancer du sein métastatique (Dawson *et al.*, 2013 ; Canzoniero et Park, 2016).

### **I.3. Facteurs de risque épidémiologiques du cancer du sein**

Des études épidémiologiques ont été menées pour identifier les facteurs de risque potentiels du cancer du sein, car il n'existait aucun traitement efficace capable de guérir ou de prévenir la maladie (Adebamowo *et al.*, 2005 ; Becker *et al.*, 2007). On sait que le cancer du sein est un processus complexe pouvant résulter d'un ensemble de facteurs pouvant être biologiques ou gynécologiques, d'antécédents génétiques/familiaux, d'hormones, d'environnement et de modes de vie (Gross, 2000 ; Jemal *et al.*, 2011). En 2004, l'Organisation mondiale de la santé a documenté que les taux de cancer du sein avaient augmenté de 0,5% depuis 1987, dont environ 10% à 15% étaient attribués à une prédisposition génétique héréditaire. Les facteurs de risque associés positivement au risque de cancer du sein comprennent : l'obésité (en particulier l'obésité post-ménopausique), l'adiposité centrale, l'augmentation de la taille, des antécédents familiaux de cancer du sein et d'autres cancers, des antécédents d'affection mammaire bénigne (hyperplasie atypique), des antécédents personnels de cancer du sein, la consommation d'alcool, les antécédents de reproduction (âge précoce de la ménarche, âge avancé de la ménopause, nulliparité, âge tardif de la première grossesse à terme - premier enfant après l'âge de 30 ans), les mutations des gènes BRCA, CHEK2 et PALB2, certaines conditions génétiques (telles que Ataxia telangiectasia, les syndromes de Li-Fraumeni, Cowden et de Peutz-Jeghers), l'exposition aux rayonnements ionisants, le traitement hormonal substitutif, les contraceptifs oraux et une activité physique réduite (Schatzkin *et al.*, 1987; Adebamowo et Adekunle, 1999 ; Anyanwu, 2000 ; Adebamowo *et al.*, 2003 ; Okobia *et al.*, 2006 ; Huo *et al.*, 2008 ; Ogundiran *et al.*, 2012 ; Qian *et al.*, 2014 ; Hou *et al.*, 2014).

Bien que les effets protecteurs de niveaux d'œstrogènes endogènes plus bas au cours de la vie dûs à l'âge tardif à la ménarche, l'âge précoce à la ménopause, l'âge précoce à la maternité (premier enfant avant l'âge de 30 ans), la multiparité et l'allaitement prolongé semblent avoir été bien établis à partir d'études menées ; chez les femmes africaines, la multiparité et l'allaitement prolongé tendent à être des facteurs de risque de cancer du sein dans certains pays africains tels que le Sénégal, où cela a été observé chez plus de 50% des patientes ayant au moins cinq enfants et ayant allaité leurs enfants (Mbaye *et al.*, 2014 ).

Il convient toutefois de noter que les facteurs de risque de cancer du sein peuvent varier en fonction des sous-types de cancer du sein, car il s'agit d'une maladie hétérogène pouvant varier d'une population à l'autre (Huo *et al.*, 2008).

### **I.3.1. Facteurs de risque de reproduction associés au cancer du sein**

#### **I.3.1.1. Âge à la ménarche et âge à la ménopause**

La ménarche est le début de la menstruation. Lorsqu'elle se produit chez un individu avant l'âge de douze ans, la ménarche est dite précoce. La ménopause, en revanche, est la fin du cycle menstruel, caractérisée par une chute des niveaux hormonaux (en particulier les œstrogènes et la progestérone) dans le corps lorsque les ovaires arrêtent la production des hormones. Les deux sont des marqueurs qui représentent respectivement le début et la fin de l'activité ovarienne et l'activité endocrinienne associée à la reproduction chez la femme. La ménarche précoce et la ménopause tardive soumettent la femme à une exposition prolongée aux hormones stéroïdiennes produites par les ovaires qui influent directement sur le développement et la fonction du sein, ce qui augmente le risque de cancer du sein (Adebamowo *et al.*, 2003 ; Kapil *et al.*, 2014). L'âge tardif à la ménarche a été associé à une diminution du risque de cancer du sein chez les femmes non ménopausées que chez les femmes ménopausées (Kelsey *et al.*, 1993 ; Clavel-Chapelon et Gerber, 2002 ; Al-Ajmi *et al.*, 2018). En exemple, une étude de Huo *et al.*, (2008) a montré qu'un retard de 2 ans dans les premières règles produisait une réduction de 7% du risque, ce qui correspond à une réduction de 10% observée dans une étude multicentrique internationale (Hsieh *et al.*, 1990).

#### **I.3.1.2. Âge à la première grossesse à terme et à la première naissance vivante**

L'âge à la première grossesse à terme est significativement associé au risque de cancer du sein (Okobia *et al.*, 2006). Le nombre total de cycles menstruels au cours de la vie d'une femme est réduit par la grossesse qui soumet les cellules du sein à l'exposition à des œstrogènes en circulation. Avoir une grossesse à terme à moins de 30 ans ou à un âge précoce a été associé à une probabilité réduite de développer un cancer du sein, même parmi les porteurs du gène BRCA, tandis qu'une première grossesse à terme à un âge plus avancé (plus de 30 ans) peut initier la croissance de cellules tumorales existantes (Evans *et al.*, 2018). Pour ce qui est du nombre de grossesses à terme, il existe une controverse concernant son aspect protecteur (Kelsey *et al.*, 1993). L'analyse de 47 études épidémiologiques a montré que chaque accouchement réduisait le risque relatif du cancer du sein de 7% en l'absence d'allaitement et que chaque enfant allaité correspondait à une diminution de 3,4% du risque (Collaborative

Group on Hormonal Factors in Breast Cancer, 2002). Dans une étude menée au Nigéria, il n'y avait pas d'association significative entre l'âge à la première naissance vivante et le risque de cancer du sein, ce qui s'oppose à l'observation d'études menées principalement auprès de populations caucasiennes et asiatiques (Wohlfahrt *et al.*, 2001 ; Clavel-Chapelon et Gerber, 2002 ; Albrektsen *et al.*, 2005 ; Neale *et al.*, 2005 ; Tamakoshi *et al.*, 2005 ; Huo *et al.*, 2008). Kotsopoulos et ses collaborateurs (2007) ont également observé que l'âge à la première naissance vivante ne protégeait pas les porteurs de la mutation du gène BRCA du risque de cancer du sein. De même, une augmentation transitoire du risque de cancer du sein après la grossesse a été également rapportée (Chie *et al.*, 2000).

### **I.3.1.3. Parité, allaitement et cancer du sein**

La parité est depuis longtemps associée au risque de cancer du sein : elle a été multipliée par 2,2 chez les femmes nullipares par rapport aux femmes parovivantes et le nombre de grossesses associé à un risque réduit ou accru de maladie (Gajalakshmi *et al.*, 1998 ; Rao *et al.*, 1994 ; Chang-Claude *et al.*, 2000 ; Huo *et al.*, 2008). Une parité supérieure de trois ou plus était positivement associée au cancer du sein inférieur, ce qui n'a pas été observé dans une étude sur des patientes atteintes du cancer du sein au Nigéria et au Sénégal (Rao *et al.*, 1994 ; Sweeney *et al.*, 2004 ; Neale *et al.*, 2005).

De même, l'allaitement prolongé a été associé à une réduction du risque de cancer du sein car les substances cancérigènes chimiques potentielles sont éliminées de l'organisme pendant l'allaitement (Furber *et al.*, 1999 ; Bernier *et al.*, 2000).

## **I.4. Caractéristiques cliniques et pathologiques du cancer du sein**

La majorité des patientes présentant un cancer du sein sont souvent traitées avec une thérapie ciblée qui cible une voie moléculaire spécifique, inhibant ainsi l'activité de la tumeur dans cette voie clé. Parmi ces traitements figurent la thérapie néo-adjuvante, la thérapie adjuvante et la polychimiothérapie. En dépit du risque associé, des adjuvants et une polychimiothérapie sont administrés aux patientes afin de réduire le risque de récurrence et de décès, dans la mesure où cette maladie vise à éliminer les dépôts micro métastatiques à distance. Il sera donc utile de sélectionner avec soin les patientes qui bénéficieront le plus de ces traitements : en identifiant celles qui peuvent abriter des micrométastases silencieuses (Early Clinical Trials Research Group for Breast Cancer, 1998). Ces sélections de patientes sont réalisables en identifiant des facteurs pronostiques et prédictifs bien établis pour chaque patiente. Bien que les facteurs prédictifs soient des mesures identifiables associées à la réponse de la patiente à un traitement particulier

et utilisés pour déterminer le traitement approprié pour chaque patiente, les facteurs pronostiques sont des mesures présentes au moment de la chirurgie, qui sont corrélées au fait que la maladie soit libre / globale survie d'une patiente en l'absence de traitement adjuvant systémique en corrélation avec l'histoire naturelle de la maladie, ils permettent de mieux sélectionner les patientes dont la maladie risque davantage de survenir sans traitement adjuvant (Cianfrocca et Goldstein, 2004). Le tableau I présente un résumé des facteurs pronostiques et prédictifs du cancer du sein.

**Tableau I : Résumé des facteurs pronostiques et prédictifs**

Facteurs	Pronostique	Prédictif
Lymphes	Oui	-
Tumeur	Oui	-
Invasion lympho-vasculaire	Oui	-
Marqueurs de prolifération	Oui	-
Origine ethnique	Peut-être	-
Âge	Oui	-
Statut ER / PR	Oui	Oui
HER2 / neu	Oui	Oui
uPA/PAI	Oui	Oui
Profilage génétique	Oui	Oui

Les génomes de tumeurs subissent des changements qui modifient leurs structures et affectent leurs fonctions à la suite de mutations conduisant à la formation et à la progression de croissances néoplasiques ou métastatiques (Theillet, 2010). Le cancer étant une maladie multifactorielle, outre sa prédisposition génétique, plusieurs facteurs clinico-pathologiques influent sur l'évolution du cancer du sein.

Hankinson *et al.*, (2004) ont déclaré que le cancer du sein était principalement une maladie post-ménopausique, plus des trois quarts des tumeurs réagissant aux hormones. Cette tendance est en train de changer rapidement car il a été observé qu'elle prévalait au cours des dernières années chez les femmes non ménopausées (Udoakang *et al.*, données non publiées). Cette dépendance hormonale provoque des facteurs environnementaux et génétiques qui déterminent l'incidence et la progression de la maladie en fonction de plusieurs facteurs cliniques et pathologiques tels que l'âge au moment de la présentation, le stade de la tumeur, le grade tumoral, le type histologique, l'atteinte des ganglions lymphatiques, les métastases, la taille de la tumeur.

Bien que le cancer du sein soit moins fréquent chez les femmes plus jeunes dans les pays développés (11% des cas de cancer du sein chez les jeunes femmes) (Marina, 2010), il est

généralement diagnostiqué chez des femmes plus jeunes en Afrique subsaharienne dont l'âge moyen est compris entre 40 et 50 ans et ils se présentent également à des stades avancés III ou IV avec des ganglions axillaires et des métastases distantes (Huo *et al.*, 2009). Les tumeurs du sein chez les jeunes femmes présentent généralement des caractéristiques plus agressives, notamment une taille de tumeur plus importante, une incidence plus élevée de tumeurs mal différenciées, des ganglions lymphatiques positifs, des taux de prolifération élevés, une expression plus élevée du récepteur du facteur de croissance épidermique (HER2) et une absence de récepteurs endocriniens (Bleiker *et al.*, 1999).

### **I.5. Instabilité génétique**

La compréhension des mécanismes moléculaires à la base de l'évolution du cancer a été de mieux en mieux connue, car l'activation des oncogènes et l'inactivation des gènes suppresseurs de tumeurs sous-tendent la carcinogenèse. Les tumeurs se développent à la suite de plusieurs altérations génétiques et le taux de mutation des cellules somatiques s'avère trop faible pour expliquer le nombre de mutations nécessaires à la carcinogenèse d'une cellule (Batistatou *et al.*, 2004).

Le terme instabilité génomique (ou génétique) indique la tendance accrue des cellules tumorales à acquérir de nouvelles mutations lors de la division cellulaire. La simple observation de multiples mutations dans une tumeur est un « état » alors que l'instabilité génomique est un « taux » bien qu'elles aient été observées simultanément dans la plupart des cas dans la même tumeur (Lengauer *et al.*, 1997). Le terme « instabilité » est utilisé dans de nombreux cas pour indiquer « l'état » de multiples mutations dans les cancers car il est difficile et exigeant expérimentalement de prouver les changements de « taux ». Le développement de l'instabilité génétique est une première étape critique vers la cancérogenèse et les études moléculaires fournissent des preuves permettant d'expliquer le développement et la progression du cancer au niveau cellulaire. Cela a conduit à l'identification d'aberrations moléculaires spécifiques et de biomarqueurs appelés agents ciblés (Niederacher *et al.*, 1996 ; Kim *et al.*, 2004 ; Dienstmann *et al.*, 2015 ; Rubio-Perez *et al.*, 2015). Le phénotype d'instabilité génétique, qui caractérise la plupart des tumeurs solides, est essentiellement classé en au moins trois formes :

- L'instabilité des microsatellites (MSI) associée au phénotype de mutateur en raison de la mutation de gènes de réparation incompatibles ;

- L'instabilité des chromosomes (CIN) reconnue par des anomalies chromosomiques globales et le signe distinctif des tumeurs aneuploïdes associé à l'activation d'oncogènes et / ou à l'inactivation de gènes suppresseurs de tumeurs (Modrich, 1994 ; Charames et Bapat, 1998 ; Lengauer *et al.*, 1998).
- Le phénotype méthylateur de l'îlot CpG (CIMP) qui chevauche généralement l'ISM sporadique et se retrouve dans la plupart des tumeurs présentant des mutations dans l'oncogène BRAF (Jass, 2005 ; 2007).

Bielas *et al.* (2006) ont suggéré une autre forme d'instabilité appelée « instabilité de mutation ponctuelle » (PIN) basée sur l'estimation du taux d'acquisition de mutations ponctuelles aléatoires dans une tumeur par opposition à des cellules normales. Ils ont rapporté que les mutations ponctuelles se produisent 200 fois plus souvent dans les cancers que dans les cellules normales. En 2006, le nombre moyen de mutations nucléotidiques dans les cancers du sein et colorectal a été estimé à environ 100, dont 20 pourraient jouer un rôle causal (Sjoblom *et al.*, 2006).

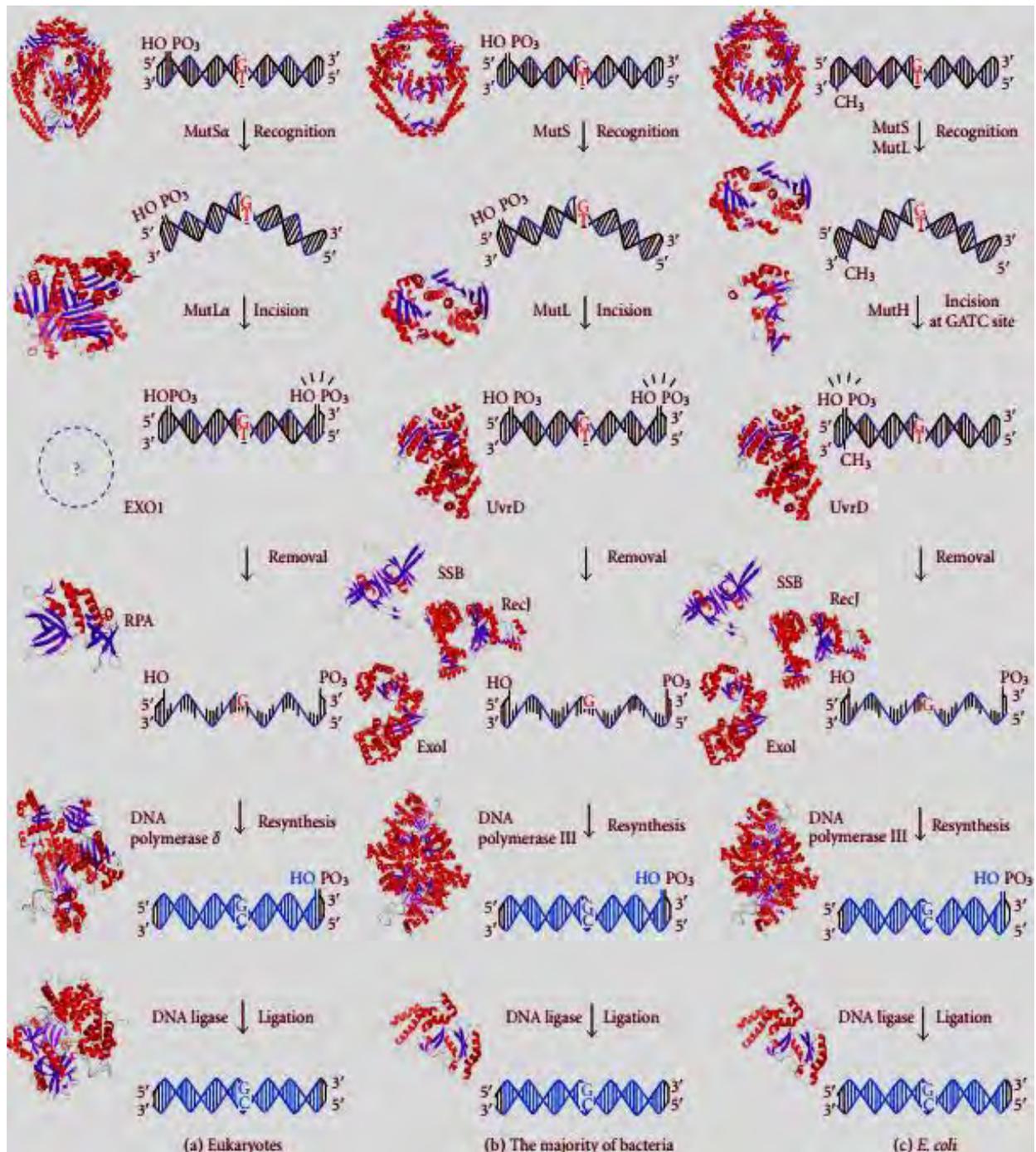
Le système MMR est un système de réparation de l'ADN hautement conservé qui contribue principalement au maintien de la stabilité génomique par la correction de paires de bases non appariées (Friedberg *et al.*, 2005). Le phénotype d'instabilité des microsatellites, l'un des dysfonctionnements causés par le dysfonctionnement du MMR, a été rapporté dans différents types de tumeurs sporadiques et est bien normalisé pour le cancer colorectal. Il conduit à la génération de nouveaux allèles de tailles très différentes du type sauvage et est associé à un pronostic amélioré, une survie supérieure du patient, un bon prédicteur de la réponse au cancer et une résistance à certaines chimiothérapies (Rubic *et al.*, 2003). Le système revêt une importance capitale pour la rectification des inadéquations de séquences d'ADN lors de la réplication de l'ADN (Gelsomino *et al.*, 2016). Ce mécanisme de réparation de mésappariement post-répliatif, responsable du maintien de l'intégrité génomique, reconnaît et répare les nucléotides désappariés au cours de la synthèse de l'ADN qui émergent du glissement du brin de l'ADN polymérase, ce qui conduit à une incorporation erronée. Ce système de réparation est composé principalement de quatre des six protéines humaines (homologue mutL 1 - MLH1, homologue PMS1 2 - PMS2, homologue mutS 2 - MSH2 et homologue mutS 6 - MSH6 ou homologue mutS 3 - MSH3) interagissant entre eux pour détecter les inadéquations et couper afin que l'ADN polymérase et l'ADN ligase puissent resynthétiser et relier le brin d'ADN correct (Vilar et Gruber, 2010).

Le fait que le système DNA MMR ne répare pas les erreurs qui se produisent lors de la réplication de l'ADN entraîne un phénomène appelé instabilité des microsatellites (MSI). Ce phénotype est caractérisé par une accumulation accélérée de mutations nucléotidiques uniques et par des altérations de la longueur de séquences microsatellites simples et répétitives qui surviennent de manière omniprésente dans le génome des organismes (Schlotterer et Harr, 2014). Par conséquent, la présence de MSI dans le tissu tumoral est associée à certaines caractéristiques cliniques et pathologiques uniques. Une meilleure compréhension de la composition et de l'évolution des tumeurs au cours de la progression de la maladie et du traitement améliorera le diagnostic du cancer et la conception des thérapies. (Boland *et al.*, 1998). La figure 14 présente l'illustration du processus du système ROR chez les eucaryotes, la plupart des bactéries et *E. coli*.

Chez les eucaryotes, l'incorporation erronée d'une base lors de la réplication de l'ADN entraîne une inadéquation de l'ADN. MutS $\alpha$  reconnaît les mésappariements base-base et MutL $\alpha$  entaille le côté 3' ou 5' de la base non appariée sur le brin discontinu (Kadyrov *et al.*, 2006). Le segment d'ADN résultant est excisé par l'exonucléase EXO1 en coopération avec la protéine de liaison à l'ADN simple brin RPA. Le brin d'ADN est encore synthétisé par l'ADN polymérase  $\delta$  et l'ADN ligase 1. Dans la plupart des bactéries, les bases mal appariées sont reconnues par MutS. Après l'incision du brin discontinu par MutL, le brin d'ADN contenant des erreurs est éliminé par les fonctions coopératives des ADN hélicases, telles que UvrD, les exonucléases RecJ et ExoI et la protéine de liaison à l'ADN simple brin, SSB. L'ADN polymérase III et l'ADN ligase combrent le vide pour achever la réparation. MutS reconnaît des bases incompatibles dans *E. coli*. Le MutL interagit avec le complexe et le stabilise. L'endonucléase MutH est activée pour inciser le site GATC non méthylé afin de créer un point d'entrée pour la réaction d'excision. L'ADN hélicase, une protéine de liaison à l'ADN simple brin, et plusieurs exonucléases sont impliquées dans la réaction d'excision. Les PDB ID des structures cristallines dans cette figure sont 2O8B (MutS $\alpha$  humain), 1H7S (MutL $\alpha$  humain), 1L10 (RPA humain), 3IAY (ADN polymérase humaine  $\delta$ ), 1X9N (ADN ligase 1 humaine), 1E3M (MutS bactérienne), 1B63 (MutL bactérienne), 2AZO (*E. coli* MutH), 2ISI (UvrD bactérienne), 2ZXO (RecJ bactérienne), 3C95 (ExoI bactérienne), 2CWA (SSB bactérienne), 2HQA (ADN polymérase III bactérienne) et 2OWO (ADN ligase bactérienne) (Fukui, 2010).

Parmi les gènes MMR impliqués dans le système MMR, le gène MSH2 donne les instructions pour la production d'une protéine de réparation de l'ADN qui corrige les erreurs commises lors de la réplication de l'ADN afin de préparer la division cellulaire. La protéine MSH2 s'associe à

l'une des deux autres protéines, MSH6 ou MSH3 (chacune produite à partir d'un gène différent), pour former un complexe protéique qui identifie les sites d'erreur de réplication de l'ADN sur l'ADN. Le complexe protéique hétérodimère MLH1-PMS2 répare les erreurs. Le gène *MLH1* fournit des instructions pour fabriquer la protéine MLH1 essentielle au mécanisme de réparation de l'ADN. La protéine corrige les erreurs commises lors de la réplication de l'ADN afin de préparer les cellules à la division cellulaire. La protéine MLH1 ainsi que la protéine PMS2 qui est produite à partir du gène *PMS2* forment un complexe protéique hétérodimère. Le complexe synchronise les activités d'autres protéines de réparation des erreurs de réplication d'ADN. Une section de l'ADN contenant l'erreur de réplication est retirée et remplacée par une séquence d'ADN corrigée (Genetics Home Reference, 2018). Les taux de mutations microsatellites sont significativement élevés en l'absence de protéines MMR : hMLH1, hMSH2, hMSH6 et hPMS2 (Nakagawa *et al.*, 2004 ; Baudhuin *et al.*, 2005 ; Bellizzi et Frankel, 2009 ; Akoum *et al.*, 2009).

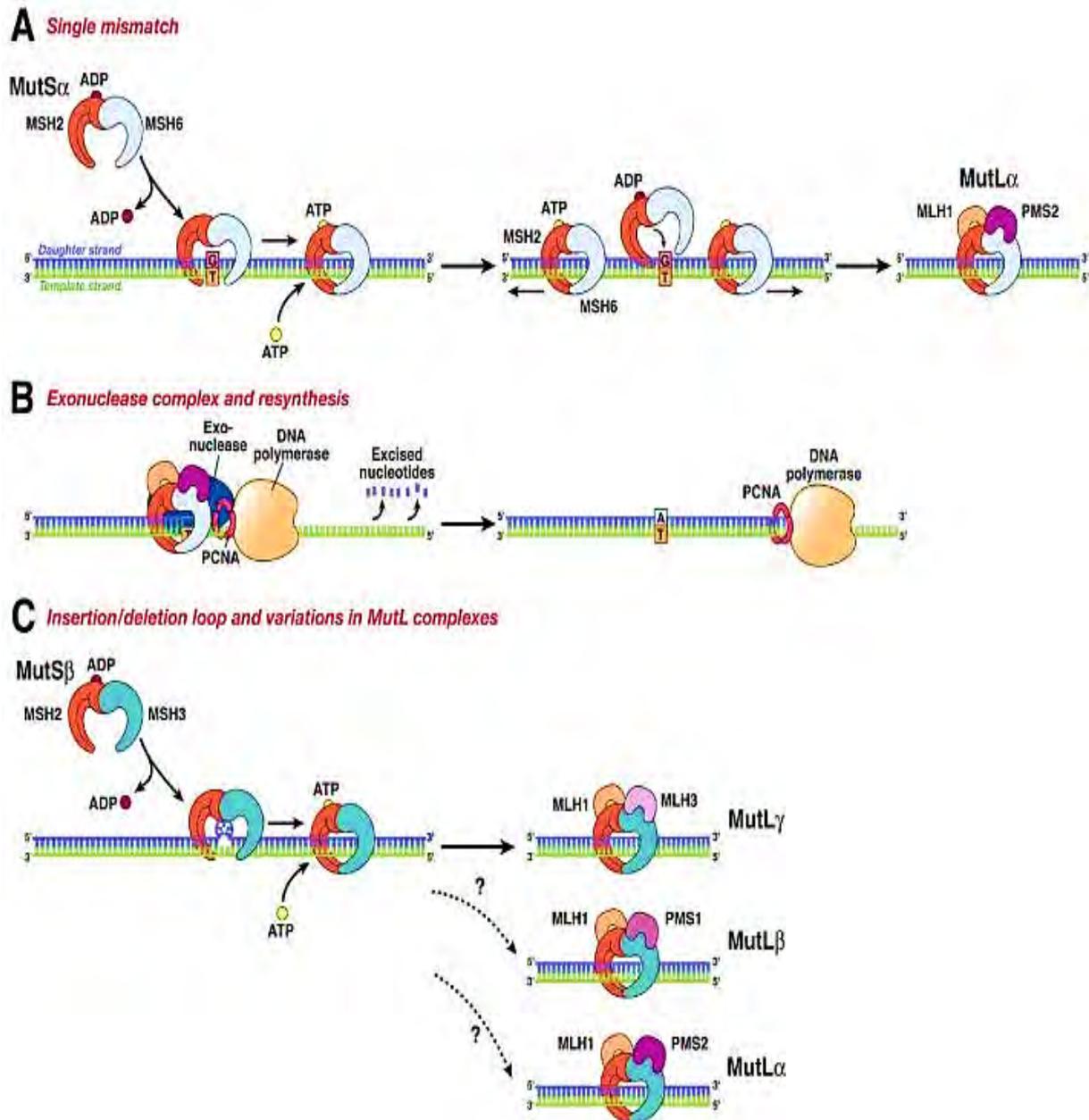


**Figure 14:** Représentation du système MMR chez les eucaryotes, la plupart des bactéries et chez *E. coli* (Fukui, 2010).

La figure 15 est une représentation schématique des étapes impliquées dans la fonction de réparation d'ADN mal apparié. «A» montre une incompatibilité unique où MSH2 – MSH6 (MutS $\alpha$ ) reconnaît les incompatibilités simple paire de bases que l'ADN polymérase a mal appariées (correspondance inappropriée de la base G avec la base T sur le brin de matrice (montré de gauche à droite), et crée une pince glissante autour de l'ADN : l'échange de l'adénosine triphosphate (ATP) contre l'adénosine diphosphate (ADP) par MSH2 (et non par MSH6 ou MSH3) est nécessaire à cette étape. Le complexe MLH1-PMS2 (MutL $\alpha$ ) se lie

ensuite au complexe précédemment formé sur le site mal assorti (à droite). Ce complexe de « matchmaker » se déplace le long de la nouvelle chaîne d'ADN jusqu'à ce qu'il rencontre le complexe ADN polymérase. Le complexe exonucléase et la resynthèse sont illustrés en « B », où la pince glissante de la protéine DNA MMR interagit avec l'exonucléase 1, l'antigène nucléaire de cellules en prolifération (PCNA) et l'ADN polymérase. Ce complexe excise ensuite le brin fille jusqu'au site du mésappariement et tombe de l'ADN menant à la survenue d'une resynthèse qui corrige. La « C » indique la boucle d'insertion /suppression et les variations dans les complexes MutL. Alors que MSH2 – MSH6 reconnaît les asymétries d'une paire unique et de petites IDL, MSH2 – MSH3 (MutS $\beta$ ) reconnaît les IDL plus grandes (illustrées à gauche). Le côté droit montre les interactions possibles avec différents dimères de MutL, car MLH1 peut se dimériser avec PMS2, PMS1 ou MLH3. L'interaction préférée avec MSH2 – MSH3 est MLH1 – MLH3 (MutL $\gamma$ ), mais le rôle précis des autres hétérodimères de MutL dans cette réaction n'est pas entièrement compris (Boland et Goel, 2010).

Le cancer du sein a de nombreuses entités biologiquement diverses, des caractéristiques pathologiques distinctes et des implications cliniques et est devenu un problème croissant pour les femmes des pays à ressources faibles et moyennes. Les preuves accumulées dans la littérature suggèrent que l'instabilité génomique joue un rôle critique dans son étiologie et que les tumeurs instables présentent des comportements distincts qui conduisent à des réponses de traitement différentes (Ding *et al.*, 2010). L'instabilité des microsatellites qui découle d'une erreur dans le système de réparation inadéquat est l'une des cibles actuellement explorées, en particulier avec les succès actuels du cancer du sein chez les patientes présentant un sous-type de cancer du sein triple négatif (Kok *et al.*, 2017 ; Hu et McArthur, 2018).



**Figure 15:** Le système DNA MMR fonctionne en plusieurs étapes (Boland and Goel, 2010).

### I.5.1 Microsatellites

La plupart des erreurs de réplication qui se produisent lors de la synthèse de l'ADN sont corrigées par le système de réparation de mésappariement, qui reconnaît et corrige les insertions, les délétions et les substitutions de bases. Ce processus commence par la liaison du complexe protéique MutS à la séquence de mésappariement. Les deux homologues de MutS ; MutS $\alpha$  (MSH2-MSH6) et MutS $\beta$  (MSH2-MSH3) se sont révélés être respectivement impliqués dans la réparation de l'insertion-délétion (correspondances) et des mésappariements base-base uniques. Ils ont aussi la capacité de réparer l'insertion-délétion d'un brin (Kunkel et Erie, 2005 ; Kadyrov, *et al.*, 2007). La liaison de MutS à la séquence de mésappariement est le

signale pour que le complexe protéique MutL analyse la région voisine du brin d'ADN nouvellement répliqué, à la recherche d'une entaille (rupture d'un brin). Le processus de réparation est activé par un processus de dégradation d'une portion du brin néosynthétisé contenant l'erreur, suivi par la synthèse du brin d'ADN (phase de correction) (Bergfors, 2014).

Les microsatellites sont généralement des répétitions de 10 à 60 fois de la même base ou séquence de bases, avec une longueur unitaire allant d'une à six bases. Ils sont retrouvés dans les régions codantes et non codantes du génome. Les ADN polymérases sont plus enclins à commettre des erreurs dans ces régions, soit en insérant des bases supplémentaires en cas de glissement sur le brin synthétisé, soit en les retirant en cas de glissement sur le brin matrice, ce qui conduit à des brins d'ADN incompatibles.

Ils peuvent être classés en fonction du type ou du nombre de séquences répétées. Le tableau II présente les différentes terminologies utilisées pour décrire les microsatellites. Les répétitions mononucléotidiques sont des répétitions des mêmes nucléotides dont certains effets néfastes possibles ont été découverts, en particulier pour ceux trouvés dans la région codante de l'ADN (Wagner *et al.*, 1990). Ces répétitions sont susceptibles de glisser lors de la transcription et de la traduction et entraînent donc une perturbation de la fonction des produits géniques pour lesquelles elles sont codées (Gurvich *et al.*, 2003 ; Baranov *et al.*, 2005). Le risque d'augmentation de la mutation du décalage de cadre ayant des conséquences graves lors de la réplication de l'ADN est fortement associé aux répétitions mononucléotidiques et à la présence d'insertion / délétion (indels), car de longues répétitions mononucléotidiques peuvent échapper au système de ROM menant à une instabilité génomique (Kroutil *et al.*, 1996 ; Tran *et al.*, 1997 ; Strauss, 1999). De même, les répétitions mononucléotidiques C/G sont plus instables que celles A/T chez *Escherichia coli*, les levures et les cellules de mammifères (Sagher *et al.*, 1999 ; Harfe et Jinks-Robertson, 2000 ; Boycheca *et al.*, 2003). Les autres répétitions mentionnées ci-dessus comprennent : les dinucléotides, les trinucleotides, les tétranucleotides et les pentanucleotides.

Tableau II : Terminologie des microsatellites

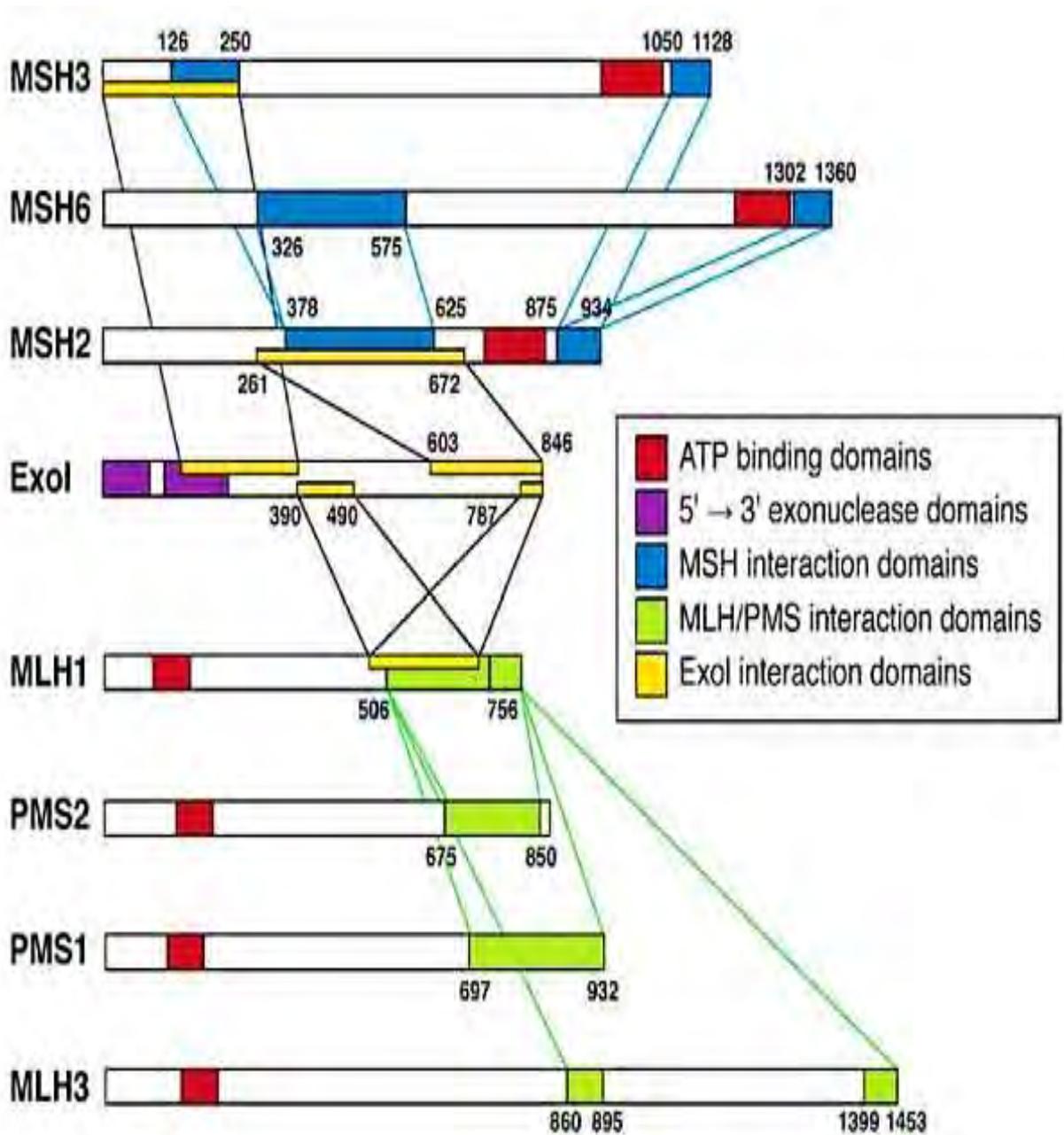
Nom de Microsatellites	Motifs de Microsatellite
Mononucleotide: (T) <sub>21</sub>	TTTTTTTTTTTTTTTTTTTTTTT
Dinucleotide: (TA) <sub>9</sub>	TATATATATATATATATA
Trinucleotide: (CTG) <sub>5</sub>	CTGCTGCTGCTGCTG
Tetratanucleotide: (AAAG) <sub>10</sub>	AAAGAAAGAAAGAAAG...AAAG <sub>10</sub> .
Pentanucleotide: (AAAAT) <sub>8</sub>	AAAATAAAAATAAAAAT...AAAAT <sub>8</sub>
Hexanucleotide: (GGGGCC) <sub>3</sub>	GGGGCCGGGGCCGGGGCC
Microsatellite interrompu	TTTTTTTCCCCTTTTTTTTTTTT
Microsatellite imparfait	TTTTTTTTTTTTTTTTTTTGTTT
Microsatellite compose (CG) <sub>5</sub> (AC) <sub>18</sub>	CGCGCGCGCGACACACACAC...AC <sub>18</sub>

La classification des microsatellites selon les types comprend : microsatellite interrompu, parfait, imparfait et composé. Un microsatellite parfait a une base ininterrompue qui répète celle de la séquence de motifs tandis qu'un microsatellite imparfait a une paire de bases non appariées n'appartenant pas à la séquence entre les motifs de séquences répétées (Oliveira *et al.*, 2006). Un microsatellite interrompu a une petite répétition de séquence sans correspondance dans les séquences de motifs répétés lorsque les motifs de séquence d'un microsatellite composé contiennent deux répétitions de séquences adjacentes et distinctives (Oliveira *et al.*, 2006).

Les cancers MSI sont associés à des taux de mutation multipliés de 100 à 1 000 fois. Des mutations faux-sens qui résultent de défauts dans le système de MMR d'ADN (Modrich, 1987 ; Fishel *et al.*, 1993 ; Bronner *et al.*, 1994 ; Liu *et al.*, 1994 ; Papadopoulos *et al.*, 1994 ; Eshleman *et al.*, 1996). Après la réplication de l'ADN, la machinerie du MMR glisse le long de l'ADN et cible les inadéquations pour être corrigées lorsqu'elle les rencontre. Les patients présentant un défaut dans l'un de ces composantes ou dans un gène en amont de *MSH2* qui code pour EPCAM et conduit à une perte d'expression de *MSH2* lors d'une mutation (Ligtenberg et Kuiper, 2009) développent des mutations de décalage en chaîne conduisant à des produits protéiques tronqués. Les gènes les plus souvent affectés sont *TGFBRII*, *ACVR2A*, *TCF4*, *IGF2R*, *BAX*, *MSH6* et *MSH3* (Markowitz *et al.*, 1995 ; Umar *et al.*, 2004a ; Cancer Genome Atlas Network, 2012). La figure 16 est une représentation schématique des étapes impliquées dans la fonction de réparation d'ADN mal apparié.

Les MSI conduisent à un phénotype d'erreur de réplication positive (RER +), caractérisé par une instabilité entraînant des mutations de décalage de petites séquences d'ADN répétitives (microsatellites) dispersées dans le génome humain, en séquences codantes et non codantes (Modrich, 1994 ; Karran, 1996 ; Lengauer *et al.*, 1998). Par conséquent, l'expansion ou la contraction des microsatellites due à l'augmentation ou à la diminution du nombre d'unités répétées donne lieu à ce phénotype (Al-Sukhni *et al.*, 2008). Cela se produit souvent en raison de l'incapacité des gènes de réparation de mésappariement (MMR) à réparer les mésappariements d'ADN conduisant à la génération de bases insérées et/ou supprimées lors de la réplication (Ashktorab *et al.*, 2003 ; 2005 ; 2008 ; Azadeh *et al.*, 2008). Ce phénomène produit de nouveaux allèles qui sont détectés comme des changements de taille d'allèle entre l'ADN normal et l'ADN tumoral.

La MSI est donc l'une de ces voies d'aberration moléculaire par lesquelles les tumeurs sont à l'origine (Abdel-Rahman, 2008 ; McMurray, 2010). Bien que les MSI soient classiquement attribuées à des défauts génétiques, il existe également des facteurs épigénétiques qui sont apparus tout aussi dommageables pour le contrôle du cycle cellulaire, tels que l'hyperméthylation de la région promotrice du gène de l'ADN MMR en tant que mécanisme épigénétique qui réduit au silence les gènes qui contribuent et peuvent conduire à la tumorigène (Kane *et al.*, 1997 ; Cunningham *et al.*, 2001 ; Lindor *et al.*, 2002). Par conséquent, l'instabilité microsatellite (MSI) ou l'instabilité des short tandem repeats (STR) est un phénotype moléculaire dû à un système de réparation des mismatch (MMR) défectueux (Pagin *et al.*, 2013). L'étude de l'instabilité des microsatellites (MSI) a fourni les preuves pour soutenir une voie séquentielle et progressive du développement du cancer. La Figure 16 montre l'intrication protéine-protéine parmi les familles du système MMR avec des interactions définies entre MutS, MutL ainsi qu'entre MutS et Mut L avec Exo1. Elle montre comment les fonctions incompatibles risquent d'être tronquées à la suite de mutations dans ces régions. (Boland et Goel, 2010).



**Figure 16:** Régions des interactions protéine-protéine parmi les membres du système MMR. ATP, adénosine triphosphate ; PMS, ségrégation post-méiotique (Boland et Goel, 2010).

L'instabilité des microsatellites et la perte d'hétérozygotie (LOH) sont les deux aberrations connues pour être impliquées dans les premières étapes de la voie de la tumorigénèse (Nobili *et al.*, 2011 ; Laurent-Puig *et al.*, 2011). Bien que de nombreuses recherches aient été menées sur le cancer du côlon, qui représente un bon modèle de progression du cancer d'un adénome à un carcinome, accompagné de mutations génétiques et de LOH, en raison de l'hétérogénéité morphologique et biologique du cancer du sein, il a été difficile de définir une série d'aberrations génétiques impliquées dans sa progression (Kim *et al.*, 2004). Environ 15% des cancers colorectaux sporadiques, 20% des cancers endométriaux sporadiques, ainsi qu'une

proportion non négligeable de cancers provenant d'autres sites présente une MSI (Ionov *et al.*, 1993 ; Thibodeau *et al.*, 1993 ; Salovaara *et al.*, 2000).

L'analyse par microsatellite est une méthode fréquemment utilisée pour la détection des délétions chromosomiques par perte d'hétérozygotie et pour la détection de l'instabilité des microsatellites (Wild *et al.*, 2005). Au niveau clinique, on sait que les patients atteints de tumeurs colorectales MSI ont un meilleur pronostic ajusté en fonction du stade et peuvent réagir différemment à la chimiothérapie adjuvante (Gryfe *et al.*, 2000 ; Ribic *et al.*, 2003 ; Sargent *et al.*, 2010). Pour des analyses microsatellites fiables, une teneur en cellules tumorales d'au moins 80% est requise et divers marqueurs de répétition ont été proposés pour déterminer le MSI (Wild *et al.*, 2005).

Le modèle génétique à plusieurs étapes de la carcinogenèse colorectale a ouvert une nouvelle ère pour les chercheurs qui recherchent de nouveaux gènes suppresseurs de tumeurs grâce à l'utilisation de nouvelles techniques telles que la réaction en chaîne de la polymérase (PCR) amorcée de manière arbitraire et utilisée par Perucho (1999). Les chercheurs ont extrait et amplifié des échantillons d'ADN appariés provenant de tumeurs colorectales et d'ADN de tissu du colon normal adjacent, ont séparé les produits de PCR par électrophorèse, puis ont séquencés à l'aide d'un petit nombre d'amorces de PCR choisies au hasard, produisant des signatures électrophorétiques uniques pour chaque échantillon (Vogelstein *et al.*, 1989).

L'atelier international sur le HNPCC, parrainé par le National Cancer Institute en 1997, a confirmé l'utilité de l'analyse de l'instabilité des microsatellites comme première étape de l'identification des patients devant faire l'objet d'un test de dépistage de mutations du gène de réparation des mésappariements, critères qui incitent à réaliser des tests d'instabilité des microsatellites (Rodriguez-Bigas *et al.*, 1997). La conférence a recommandé l'utilisation d'un panel de cinq marqueurs microsatellites pour évaluer l'instabilité. Ce panel était constitué de deux séquences répétées mono nucléotidiques (*BAT-25* et *BAT-26*) et trois di nucléotides (*D2S123*, *D5S346* et *D17S250*) (Boland *et al.*, 1998). Cependant, certaines limitations sont apparues, principalement en raison de l'utilisation de marqueurs di nucléotidiques présentant une sensibilité et une spécificité moindres par rapport aux marqueurs mono nucléotidiques (Perucho, 1999 ; Suraweera *et al.*, 2002).

En 2002, Hamelin *et al.* (2002) ont révisé les lignes directrices de Bethesda afin de prendre en compte certaines caractéristiques pathologiques communes et d'inclure les patients pouvant être plus âgés mais susceptibles de présenter des mutations du gène de réparation de

mésappariements (Umar *et al.*, 2004a). Ils ont proposé un nouveau panel de cinq marqueurs mono nucléotidiques quasi-monomorphes (le même nombre de répétitions étant retrouvé dans une proportion écrasante d'humains), connu sous le nom de panel pentaplex, qui permettait une identification précise des tumeurs MSI sans qu'il soit nécessaire de rechercher des tumeurs normales (Suraweera *et al.*, 2002). Ce panel a montré une sensibilité et une spécificité accrues pour la détection des tumeurs MSI-H et éviterait la nécessité de réaliser des tests séparés sur l'ADN à la fois de la tumeur et du tissu normal d'un patient par rapport au panel Bethesda et a donc été recommandé pour les tests MSI (Suraweera *et al.*, 2002 ; Buhard *et al.*, 2004 ; Umar *et al.*, 2004a ; Wong *et al.*, 2006 ; Xicola *et al.*, 2007 ; Goel *et al.*, 2010 ; You *et al.*, 2010). Si deux ou plusieurs des cinq séquences microsatellites étaient mutées, la tumeur serait classée dans la catégorie des tumeurs hautement instable (MSI-H) ; si seulement un mutait, ce serait une instabilité microsatellite faible (MSI-L) ; et si aucun n'était muté, ce serait microsatellite stable (MSS) (Boland *et al.*, 1998).

La plupart des patients atteints de HNPCC présentent une mutation héréditaire dans l'un des quatre gènes (*MLH1*, *MSH2*, *MSH6* et *PMS2*) impliqués dans le système de réparation de mésappariements de l'ADN. Ces gènes sont impliqués dans la réparation des mismatch d'ADN (MMR). Les tumeurs qui se forment chez les patients porteurs de ces mutations héréditaires présentent un défaut dans le MMR qui se manifeste par un phénotype cellulaire caractéristique appelé « instabilité des microsatellites (MSI) (REF) ». Étant donné que la MSI est présente dans la grande majorité des tumeurs de patients atteints de HNPCC, son absence rend très improbable le diagnostic de HNPCC. Bien que la MSI puisse également survenir chez des patients ne présentant pas le syndrome HNPCC mais ayant plutôt un cancer sporadique non familial, environ 15% des cas de cancer du côlon sporadique sont atteints de MSI. Ainsi, la présence de MSI ne distingue pas le cancer familial (HNPCC) du cancer sporadique, mais le test MSI peut être utile pour déterminer le pronostic de la maladie et choisir le meilleur protocole de traitement chimio thérapeutique pour les patients.

Plusieurs chercheurs ont également démontré la présence de tumeurs sans forme apparente d'instabilité génomique (Georgiades *et al.*, 1999 ; Chan *et al.*, 2001 ; Jones *et al.*, 2005 ; Abdel-Rahman *et al.*, 2005). Les tumeurs malignes ne montrent pas une instabilité génomique évidente (Abdel-Rahman, 2008). Ces résultats suggèrent qu'il existe des formes d'instabilité non identifiées qui sous-tendent le développement de tels sous-ensembles de tumeurs ou qu'une forme d'instabilité n'est pas requise par toutes les tumeurs. Le résultat naturel du phénotype d'instabilité est la génération de multiples clones tumoraux à un taux supérieur à celui habituel.

Cela se reflète sur l'étendue de l'hétérogénéité tumorale observée et il est supposé qu'une hétérogénéité importante des génomes tumoraux a des implications importantes pour le traitement du cancer grâce à la génération de clones résistants (Loeb *et al.*, 2008).

### **I.5.2. Utilisation de deux marqueurs pour la détection de l'instabilité des microsatellites dans les tumeurs**

Afin d'établir l'approche la plus simple et la plus fiable pour la détection des tumeurs présentant une instabilité des microsatellites, Xicola *et al.*, (2016) ont examiné si une combinaison d'un nombre réduit de marqueurs serait aussi efficace que l'ensemble du panel pentaplex de répétitions mono-nucléotidiques. Ils ont identifié toutes les tumeurs MSI-H comportant un ou deux marqueurs stables et ont observé que l'association de *BAT-25* ou *BAT-26* stable avec *NR-21* ou *NR-24* stable n'a été observée dans aucune tumeur avec perte d'expression de la protéine de réparation de mismatch. Ainsi, ils ont déterminé que l'évaluation de deux marqueurs, *BAT-25* ou *BAT-26* et *NR-21* ou *NR-24*, serait suffisante pour détecter toutes les tumeurs présentant une perte d'expression de la protéine de réparation de mésappariements. Étant donné que le *NR-24* est systématiquement monomorphe dans la plupart des populations du monde et qu'une seule tumeur présentant un déficit de réparation de mésappariements exclusivement instable pour le *BAT-26* a été observée dans leur série, ils ont décidé d'évaluer les performances d'un panel composé uniquement des *BAT-26* et *NR-24*.

Cette méthode produisait des valeurs de sensibilité et de spécificité similaires aux valeurs obtenues lorsque l'ensemble des marqueurs pentaplex mono nucléotidiques était testé et que les produits de PCR issus de l'amplification de ces deux marqueurs pouvaient être séparés en une seule cycle de séquençage. Par conséquent, ils ont recommandé ces deux marqueurs fiables pour la détection des tumeurs déficientes en réparation de mésappariement (Xicola *et al.*, 2007). De plus, cela permet de réduire les coûts, comparativement à l'utilisation du panel complet.

Des modèles et des stratégies de mesure sont nécessaires pour lutter contre le cancer du sein. Les taux de létalité plus élevés du cancer du sein dans les pays à faibles ressources ont été attribués à un manque de sensibilisation aux avantages du dépistage et du traitement précoces, à un diagnostic tardif et à un accès limité à des soins appropriés. Des progrès importants ont été réalisés ces dernières années dans le traitement du cancer, ce qui représente un espoir pour l'avenir dans les pays moins développés et en développement mais qui n'est pas encore devenu une réalité.

Afin de réduire cette disparité mondiale des résultats du cancer du sein chez les femmes, d'ici 2025, de meilleures façons de mettre en œuvre des stratégies abordables, appropriées, acceptables et réalisables reposant sur des bases factuelles sont explorées. Il est particulièrement important d'identifier, de documenter et de promouvoir de meilleurs moyens de propager des approches nouvelles et novatrices de la prestation des soins de santé du cancer du sein mis au point dans les pays à faibles ressources. Par conséquent, des efforts combinés d'experts, de scientifiques et de ressources d'individus et d'institutions engagés dans les soins du cancer du sein sont nécessaires pour atteindre cet objectif ultime (Dvaladze *et al.*, 2016).

#### **I.5.2.1. Procédure de la réaction en chaîne par polymérase (PCR) et Séquençage**

La PCR est une méthode utilisée pour l'amplification de courtes séquences d'ADN appelée séquence cible à partir d'une petite quantité d'ADN génomique. Il s'agit de l'amplification non sélective d'une séquence de matrice d'ADN particulière par extension de deux amorces (l'une directe et l'autre inverse) par une ADN polymérase en présence de désoxyribonucléotides (dNTP) et d'ions  $MgCl_2^+$ . C'est une méthode relativement rapide d'amplification du fragment d'ADN cible qui est exponentielle. Les deux amorces utilisent un cadrage de polarité opposée des fragments à amplifier (Saiki *et al.*, 1985 ; Mullis *et al.*, 1986).

Trois étapes de dénaturation, d'annealing et d'extension sont effectuées pendant la PCR et sont répétées plusieurs fois dans le thermocycleur pour obtenir les amplicots (pour le séquençage). La dénaturation de l'ADN se fait par élévation de la température. Un recuit, également appelé hybridation des amorces, est réalisé par évaluation de la température sous la forme  $T_m = 2AT + 4GC$ , et l'allongement, également appelé synthèse du brin complémentaire, est effectué par l'ADN polymérase thermostable (Bio basic ou Dream Taq).

La méthode de séquençage de Sanger a été développée et proposée pour la première fois par Frederick Sanger et ses collègues en 1977 et est utilisée depuis plus de 40 ans maintenant. Il permet de mesurer la diversité fondamentale de tous les marqueurs en comparant les séquences du même gène chez différents individus de la même espèce ou d'espèces différentes afin de déterminer les différences ou les similitudes au sein ou entre une population donnée. Il est basé sur l'utilisation de l'incorporation sélective de didésoxyribonucléotides à terminaison de chaîne qui bloquent la synthèse de l'ADN par les ADN polymérases après leur incorporation. Cette méthode utilise des nucléotides modifiés (didésoxyribonucléotides - ddNTP) en plus des composés habituels (amorces, dNTP, ADN matrice, taq polymérase et  $Mg^{2+}$ ). Les ddNTP sont couplés aux fluorochromes et il est difficile pour les nucléotides de former des liaisons

phosphodiesteres les uns avec les autres en raison de l'absence du groupe hydroxyle sur le carbone 3'.

## CHAPITRE II. COMPARAISON DES PARAMETRES EPIDEMIOLOGIQUES DU CANCER DU SEIN ENTRE LES PATIENTES NIGERIANES ET SENEGALAISES

### INTRODUCTION

Le cancer du sein est le cancer le plus répandu et la principale cause de mortalité due au cancer chez les femmes dans le monde (Bray *et al.*, 2018). Alors que le taux de survie de cinq ans du cancer du sein aux États-Unis est de 86%, il est inférieur à 40% en Afrique subsaharienne. En Afrique de l'Ouest, le cancer du sein représente 35% des nouveaux cas de cancer et est la principale cause de décès par cancer ; 28% chez les femmes (Global Cancer Facts and figures, 2015 ; Cumber et Tsoka-Gwegweni, 2017). Bien que des similitudes aient été observées entre les facteurs de risque dans différentes populations, des différences dans les caractéristiques tumorales sont observées chez les noirs comparativement aux populations blanches Caucasiennes (Alami *et al.*, 2003 ; Huo *et al.*, 2009 ; Brinton *et al.*, 2014 ; Breyer *et al.*, 2018). Les meilleurs résultats dans les populations Caucasiennes ont été attribués à des facteurs démographiques tels que l'espérance de vie plus longue, la sensibilisation accrue, le système amélioré de déclaration de la maladie et l'accès à de meilleurs services de santé (Alami *et al.*, 2003 ; Breyer *et al.*, 2018).

Le cancer du sein est la malignité la plus fréquente chez les femmes nigérianes et est le deuxième après le cancer du col utérin chez les femmes au Sénégal (Alami et Ajayi, 2000 ; Dem *et al.*, 2008 ; Huo *et al.*, 2009 ; Gueye *et al.*, 2013 ; Gueye *et al.*, 2016). Les changements séculaires dans les modes de vie, les facteurs hormonaux et reproductifs ont été associés à l'augmentation de l'incidence dans les pays en développement comme le Nigéria et le Sénégal (Otis *et al.*, 2010 ; Gueye *et al.*, 2016). Certains facteurs de risque connus et identifiés dans différents groupes de populations comprennent l'âge à la ménarche, l'âge à la première grossesse, l'âge à la ménopause, l'allaitement maternel, la parité et les antécédents familiaux du cancer du sein (Alami et Adehunle, 1999 ; Chlebowski *et al.*, 2005 ; Krieger *et al.*, 2010 ; Gueye *et al.*, 2016 ; Cumber et Tsoka-Gwegweni, 2017).

En Afrique subsaharienne, il y a eu moins de mesures modérées pour la prévention et le contrôle du cancer du sein, mais il y a eu une amélioration de la détection et du traitement précoce (OMS, 2007 ; Yip *et al.*, 2008). Des mesures spécifiques à la population pour contrôler et prévenir la maladie ainsi que des mesures de dépistage précoce dans la population plus jeune qui sont souvent présentes à un stade tardif font défaut. Malgré l'augmentation relative de la recherche sur le cancer du sein au Nigéria et dans certains pays d'Afrique, d'autres études sont

nécessaires pour comprendre la répartition des facteurs de risque de pathogenèse du cancer du sein dans les différentes régions du continent. Cela contribuera à l'élaboration de pratiques efficaces fondées sur la population afin de maîtriser son incidence et d'accroître les stratégies préventives en Afrique (Okobia *et al.*, 2006 ; Ferlay *et al.*, 2013 ; Willet *et al.*, 2014 ; Breyer *et al.*, 2018). Dans ce chapitre, certains facteurs épidémiologiques, des facteurs de risque et les caractéristiques clinico-pathologiques du cancer du sein chez les femmes de tous les groupes d'âges ont été comparés entre la population nigériane et sénégalaise.

## **II.1. MATERIEL ET METHODOLOGIE**

### **II.1.1. Approbation éthique**

L'approbation éthique de cette étude a été obtenue auprès des commissions d'examen institutionnel du University College Hospital (CHU) d'Ibadan au Nigeria (HREC attribuée au protocole numéro : UI/EC/16/0334) et de l'université Cheikh Anta Diop, Dakar Comité d'éthique de la recherche du Sénégal (numéro de référence : protocole 0269/2015/CER/UCAD). Le consentement éclairé a été obtenu de toutes les participantes à l'étude après une explication détaillée de l'objectif.

### **II.1.2. Critères d'inclusion et d'exclusion**

Toutes les patientes atteintes du cancer du sein qui fréquentaient consécutivement les cliniques d'oncologie chirurgicale et radiothérapie du CHU d'Ibadan (Nigéria) entre novembre 2016 et juillet 2017 ont été recrutés dans cette étude. Cet hôpital sert une population d'environ 3,5 millions de personnes, et c'est un hôpital de référence majeur dans la troisième plus grande ville du pays. Ibadan est situé sur 7,38° de latitude N et de 3,91° de longitude E, avec une élévation de 181 mètres au-dessus du niveau de la mer. La majorité des patientes atteintes du cancer du sein diagnostiquées sur référence dans la région ont été observés au CHU.

Les participantes ont été interviewées le même jour et se sont présentées à la clinique d'oncologie avec un diagnostic clinique de cancer du sein. Les objectifs de l'étude, le protocole, les risques, les avantages, la confidentialité et leurs droits ont été expliqués aux participantes potentielles et chaque personne avait la possibilité de refuser sa participation à l'étude. L'opportunité a été accordée aux participantes de poser des questions et celles qui étaient disposées à participer ont signé le consentement éclairé après quoi, un questionnaire structuré a été dirigé par le chercheur. La confirmation tissulaire du cancer du sein a été faite par incision,

excision, aux aiguilles de biopsie et de ponction et toute participante qui avait un diagnostic autre que le cancer du sein a été retirée de l'étude.



**Figure 17** : Carte de l’Afrique montrant les deux pays de la population étudiée

Au Sénégal, toutes les patientes, de tout âge confondu, atteintes de cancer du sein et ayant fréquenté, entre 2014 et 2015, le centre de cancérologie de l’Institut Joliot-Curie de l’hôpital Aristide le Dantec, Dakar (Sénégal), ont été approchées pour l’étude. La majorité d’entre elles ont accepté de participer à l’étude avec un taux de refus de 4%. Le consentement éclairé a été obtenu des participantes. L’Hôpital Aristide le Dantec est situé à Dakar qui est la capitale du Sénégal et la plus grande ville. Il est un important et le seul hôpital de référence pour les patientes cancéreuses dans le pays au service d’une population immédiate d’environ 2.450.000 personnes. Dakar est situé sur la latitude 1,690 N° et de longitude 17,440 W° avec une élévation de 12 mètres au-dessus du niveau de la mer.

### **II.1.3. Collecte de données et analyses statistiques**

Le questionnaire de collecte de données et d’analyses statistiques contient des facteurs démographiques tels que l’âge, l’âge au diagnostic, l’état matrimonial, les antécédents familiaux du cancer du sein et d’autres cancers et l’ethnicité. Les variables reproductrices incluaient l’âge à ménarche, la grossesse [nombre de grossesses, l’âge à la première grossesse à terme (FFTP)], l’historique de l’accouchement, y compris la parité, l’allaitement et le statut de la ménopause. Les informations sur les caractéristiques tumorales ont été extraites des notes de cas et du rapport de pathologie des participantes et ils incluaient la date du diagnostic, le site et la taille

des tumeurs, l'histologie, l'étape des métastases des ganglions tumoraux (TNM), la catégorie Scarff-Bloom-Richardson (SBR), le ganglion lymphatique et les métastases.

Les données relatives aux caractéristiques tumorales disponibles chez les participantes sont les suivantes : type histologique (200), grade SBR (163) et stade TNM (188). Bien que les caractéristiques reproductrices examinées étaient : l'âge à ménarche, le nombre de grossesses, l'âge à la première grossesse à terme, l'historique de l'allaitement maternel et le statut de la ménopause ; l'âge à ménarche a été étudié dans quatre catégories : 10-12, 13-15, 16-18 et  $\geq 19$  ; le nombre de grossesses dans cinq catégories : 0, 1-3, 4 – 6, 7 – 9 et  $\geq 10$  ; l'âge à FFTP : <20 ans, 21-30 ans, 31-40 ans et  $\geq 41$ . L'histoire de l'allaitement maternel et le statut ménopausique ont été classés en tant que : jamais allaité (oui/non) et si vous avez atteint la ménopause ? (oui/non). La cohorte a été plus tard divisée en deux groupes : jeunes femmes  $\leq 50$  ans et les femmes âgées  $> 50$  ans pour comparer la prévalence du cancer du sein chez les femmes dans ces deux groupes d'âge au Nigéria et au Sénégal.

Des variables quantitatives ont été analysées à l'aide des écarts moyens et standards et des t-tests indépendants pour comparer les moyennes entre les deux groupes ; tandis que les variables catégorielles ont été décrites en utilisant des fréquences et des pourcentages absolus et le test Chi-2 pour comparez les proportions. Les données ont été analysées à l'aide de la version 20,0 de Statistical Packages for Social Sciences (SPSS) (SPSS, Inc. Chicago IL, USA).

## **II.2. Résultats et Discussion**

### **II.2.1. Résultats**

Deux cent une (201) patientes qui ont été atteintes du cancer de sein histologiquement confirmé et qui ont accepté de participer ont été incluses dans l'étude : cent deux (102) patientes du Nigéria et quatre dix-neuf (99) patientes du Sénégal. Cent cinquante-quatre des cas (85% des patientes) présentent un stade avancé (stades III et IV) : 84 des cas (90%) du Nigéria et 70 des cas (80%) du Sénégal.

Le type histologique le plus fréquent parmi les participants des deux pays a été le carcinome canalaire invasif/infiltration (88,1%). D'autres inclus : carcinome lobulaire, phylloïdes et sarcome (Tableau III).

L'âge au diagnostic, le statut de la ménopause, l'allaitement maternel et les antécédents familiaux du cancer n'étaient pas statistiquement significatifs entre les patientes nigérianes et

sénégalaises. L'âge moyen au moment du diagnostic pour les participantes du Nigeria était de  $49.1 \pm 12,4$  ans, variant de 29 à 80 ans, tandis que les participantes du Sénégal étaient  $46,9 \pm 14,0$  ans, entre 15 et 80 ans. Les patientes nigérianes avaient une parité moyenne plus faible ( $3,7 \pm 2,1$ ) et différente significativement de celle des patientes sénégalaises ( $4,5 \pm 3,1$ ) (Tableau IV). Environ quatre-vingt-huit pour cent (88,7%) des participantes de l'étude avaient eu des enfants, dont 71,5% des femmes ont eu au moins 3 naissances vivantes. L'utilisation des contraceptifs était significativement plus fréquente chez les nigérianes (31%) que chez les sénégalaises (11,5%) ( $p = 0,001$ ). Une proportion significative ( $p = 0,032$ ) de sénégalaises (57,1%) que de nigérianes (41,8%) ont atteint la ménopause. Les deux pays différaient également dans l'état matrimonial. De plus, il y avait une différence significative d'âge à la ménarche ( $p < 0,001$ ) et d'âge à la première grossesse à terme ( $p < 0,001$ ), car les paramètres étaient plus ou moins profonds dans un pays que l'autre (Tableau IV).

**Tableau III** : Comparaison de certaines caractéristiques cliniques des cas de cancer du sein au Nigeria et au Sénégal

Caractéristiques	Tous les cas N (%)	Nigeria N (%)	Sénégal N (%)	$\chi^2$	p-value
<b>Type histologique :</b>					
Invasif/infiltré	195(97)	100(98)	95(96)	4.084	0.395
Carcinome invasif lobulaire	2(1.0)	-	2(2.0)		
Carcinome lobulaire in-situ	1(0.5)	1(1.0)	-		
<b>Autres :</b>					
Tumeur Phylloides	2(1.0)	1(1.0)	1(1.0)		
Sarcome	1(0.5)	-	1(1.0)		
<b>Catégorie SBR :</b>					
Bien différencié	29(17.9)	20(22.0)	9(12.7)	2.596	0.273
Modérément différencié	95(58.6)	52(57.1)	43(60.6)		
Mal différencié	38(23,5)	19(20.9)	19(26.8)		
<b>Stade</b>					
1	8(4.4)	3(3.2)	5(5.7)	4.596	0.204
2	20(11.0)	7(7.4)	13(14.8)		
3	91(50.0)	53(56.4)	38(43.2)		
4	63(34.6)	31(33.0)	32(34.6)		

Tableau IV : Cas de la comparaison de certaines caractéristiques du cancer du sein au Nigéria et au Sénégal

Caractéristiques	Tous les cas (No= 201)	Nigeria No= 102	Sénégal No. = 99	$\chi^2$ /t	p-Valeur
<b>L'Age au diagnostic, No (%)</b>					
≤50 y	124(62)	60(58.8)	64(65.3)	0.8915 <sup>a</sup>	0.345
>50 y	76(38)	42(41.2)	34(34.7)		
<b>L'Age à la ménarche, No. (%)</b>					
10-12 y	9(5.1)	5(5.3)	4(4.8)	20.59 <sup>a</sup>	<0.001**
13-15 y	122(68.5)	5s2(54.7)	70(84.3)		
16-18 y	43(24.2)	34(35.8)	9(10.8)		
≥19 y	4(2.2)	4(4.2)	0		
<b>Marié ? No. (%)</b>					
Oui	165(87.3)	85(94.4)	80(80.8)	7.908 <sup>b</sup>	0.005*
Non	24(12.7)	5(5.6)	19(19.2)		
<b>L'Age à FFTP, No (%)</b>					
≤20 y	71(49.7)	33(36.3)	38(73.1)	19.409 <sup>a</sup>	<0.001**
21-30 y	64(44.8)	50(54.9)	14(26.9)		
31-40 y	7(4.9)	7(7.7)	0		
≥41 y	1(0.7)	1(1.1)	0		
<b>Parité, No (%)</b>					
0	22(11.3)	4(4.1)	18(18.8)	35.450	<0.001**
1-3	46(23.7)	27(27.6)	22(19.8)		
4-6	72(37.1)	51(52.0)	21(21.9)		
7-9	36(18.6)	14(14.3)	23(22.9)		
≥10	18(9.3)	2(2.0)	18(16.7)		
<b>Allaitement, No (%)</b>					
Oui	171(96.6)	85(95.5)	86(97.7)	0.533 <sup>b</sup>	0.466
Non	6(3.4)	4(4.5)	2(2.3)		
<b>Contraception, No (%)</b>					
Oui	41(21.4)	30(31.3)	11(11.5)	11.1956 <sup>b</sup>	0.001*
Non	151(78.6)	66(68.8)	85(88.5)		
<b>Statut ménopausique, No (%)</b>					
Pré ménopause	112(57.1)	57(58.2)	55(56.1)	0.0833b	0.773
Ménopause	84(42.9)	41(41.8)	43(43.9)		
<b>Antécédents familiaux du cancer, No (%)</b>					
Oui	22(12.9)	11(11.2)	11(14.6)	0.0007 b	0.980
Non	173(87.1)	87(88.8)	86(85.4)		

## II.2.2. DISCUSSION

Le cancer du sein est le cancer le plus fréquent chez les femmes et la tumeur la plus étudiée dans le monde. Les études épidémiologiques ont révélé des variations géographiques importantes. Alors que le taux de mortalité et d'incidence du cancer du sein reste plus élevé dans les pays industrialisés, ces taux augmentent rapidement dans les pays en développement comme ceux de l'Afrique de l'Ouest. Bien qu'il y ait eu plusieurs études sur les facteurs de risque de cancer du sein au Nigéria et au Sénégal (Adebamowo et Adekunle, 1999 ; Okobia *et al.*, 2006 ; Huo *et al.*, 2008 ; Gueye *et al.*, 2013), le manque de données rend difficile la comparaison des facteurs de risque entre les deux populations (Huo *et al.*, 2009). Quelques différences apparentes, ont été observées sur certains facteurs analysés dans cette étude. Il s'agit de l'âge à la puberté, de l'âge à la première grossesse à terme, de l'état matrimonial, de l'utilisation des contraceptifs, du statut ménopausique et de la parité.

La précocité de l'âge moyen au diagnostic chez les patientes Nigériennes et Sénégalaises a été observée dans d'autres études en Afrique ou à travers le monde (Saghir *et al.*, 2002 ; Ouedraogo-Yuen et Alison, 2012 ; Mollah-Karim *et al.*, 2015). L'âge jeune au diagnostic a été identifié comme l'une des caractéristiques du cancer du sein chez les populations étudiées par rapport aux populations caucasiennes (Fregene et Newman, 2005 ; Gueye *et al.*, 2016). Cinquante neuf pour cent (59 %) des patientes du Nigéria et 65% du Sénégal avaient 50 ans ou plus. La capacité de tirer des conclusions précises sur le risque de cancer du sein chez les femmes âgées est légèrement entravée en raison de l'espérance de vie relativement courte observée dans ces pays (49,1 ans au Nigeria et 60,7 ans au Sénégal), entraînant une population fortement constituée de jeunes (Ntekim *et al.*, 2009 ; Huo *et al.*, 2009). La constatation des cas de cancer du sein est faible chez les femmes africaines plus âgées en raison de leur taux d'alphabétisation inférieur, de leur sensibilisation au cancer du sein et de leur faible statut socioéconomique (Amir *et al.*, 1994).

L'âge à la ménarche différait significativement chez les patientes des deux pays ( $p= 0,001$ ) avec un âge moyen de 15,3 ans (1,9) pour le Nigéria et de 14,3 ans (1,3) au Sénégal. Nos résultats ont révélé une association positive entre l'âge à la ménarche et le cancer du sein. Des études ont montré que l'âge précoce à la menarche a été lié à l'apparition précoce du cancer du sein et la ménopause a été systématiquement associée à un risque réduit (Claven-Chapelon et Gerber, 2002 ; Huo *et al.*, 2008).

Les courbes âge-incidence du cancer du sein en Afrique pourraient être fonction des risques liés à la population (modes de vie des sociétés pré-industrialisées) avec une augmentation de

l'incidence du cancer du sein chez les femmes pré-ménopausées (Fregene et Newman, 2005). Le statut ménopausique n'était pas différent entre les deux populations et environ les deux tiers de la population étudiée étaient en pré ménopause au diagnostic. Une prédominance du diagnostic du cancer du sein chez les femmes pré-ménopausées a été observée à travers le monde (Alami et Ajayi, 2000 ; Lopez *et al.*, 2002 ; Fregene et Newman, 2005 ; Gueye *et al.*, 2016). Les ménarches précoces avec la ménopause tardive sont corrélées positivement avec le risque de cancer du sein puisque les individus sont exposés à des hormones femelles circulant pendant une période prolongée (Mollah *et al.*, 2015).

Bien que le type de contraceptif utilisé n'ait pas été défini dans notre étude, l'utilisation de contraceptifs était plus fréquente chez les patientes nigérianes. Environ un tiers, a indiqué qu'elles avaient utilisé des contraceptifs au cours de leur vie. L'utilisation de contraceptifs hormonaux augmente en Afrique et dans les études cas-témoins, plus de participants à l'étude du cancer du sein que les contrôles utilisent des contraceptifs oraux/hormonaux. Cela a été identifié comme un facteur de risque pour le cancer du sein dans différentes populations, surtout à des années ultérieures de la vie reproductive d'une femme (Nkondjock et Ghadirian, 2005) et un risque accru de 24,8% dans les cas qui avaient utilisé des contraceptifs oraux pendant plus de 6 mois (Laamiri *et al.*, 2015). Presque toutes les patientes des deux cohortes (toutes les sénégalaises et plus de 90% de nigérianes) ont eu leur première grossesse à terme avant leurs 30 ans, malgré la grande variabilité entre les deux pays.

Des études dans le passé ont montré que le risque à vie de cancer du sein est plus faible chez les femmes qui ont eu des grossesses multiples avec un jeune âge à la première grossesse à terme, mais une légère augmentation du risque de cancer du sein au cours de la période de 6 à 12 mois post-partum qui peut durer environ trois à quatre ans (Wohlfahrt *et al.*, 1999 ; Plante et Osuch, 2000 ; Althuis *et al.*, 2004). Ce risque d'augmentation temporaire est attribuable aux manques d'effets protecteurs de la parité chez les femmes en pré-ménopause diagnostiquées avant leur 50 ans (Miller et Bulbrook, 1986 ; Bruzzi *et al.*, 1988 ; Williams *et al.*, 1990 ; Miller, 1993 ; Lambe *et al.*, 1994 ; Albrektsen *et al.*, 1995 ; Leon *et al.*, 1995 ; Liu *et al.*, 2002 ; Palmer *et al.*, 2003). Cependant, des études plus récentes chez des africains autochtones, y compris des femmes de la même cohorte, n'ont trouvé aucune différence dans le risque de cancer du sein en vieillissant (Huo *et al.*, 2008 ; Sighoko *et al.*, 2015). Ces différences dans les modèles reproductifs peuvent peut-être dûes à des facteurs génétiques ou à des variations de la biologie tumorale primaire et de l'agressivité tumorale, ce qui suggère que le modèle de traitement peut différer pour différentes populations géographiques.

La multiparité a été identifiée avec un faible risque de cancer du sein et les femmes africaines commencent à porter des enfants à un âge relativement plus jeune. Ce rôle protecteur de la parité augmente proportionnellement en fonction du nombre d'enfants et de l'âge à la première naissance vivante (Hinkula *et al.*, 2001). Une étude qui a encore analysée 47 études épidémiologiques a observé qu'en absence d'allaitement maternel, chaque naissance réduit le risque relatif de cancer du sein de 7% avec un risque réduit de 3,4% pour chaque enfant qui est allaité (Hinkula *et al.*, 2001). La majorité de nos participantes à l'étude (71,5%) avaient eu au moins 3 enfants. La parité différait entre les deux populations, bien que la plupart des patientes des deux populations aient eu au moins de trois naissances vivantes. De plus, l'allaitement maternel et les antécédents familiaux du cancer étaient similaires dans les deux pays, 96,6% de la population allaitaient leurs enfants et 87,1% n'avaient pas d'antécédents familiaux du cancer. L'état matrimonial était différent entre les deux pays, car plus de femmes sénégalaises ne s'étaient jamais mariées. La plupart des participantes de l'étude ont été diagnostiquées avec un carcinome invasif ou infiltré, et il n'y a pas eu de différence entre les stades et le niveau des tumeurs entre les pays. La présentation en fin de stage était commune aux deux pays, ce qui n'est pas inattendu étant donné l'accès limité aux soins de santé et aux infrastructures sanitaires médiocres dans ces pays.

Les limitations de ce travail incluent la petite taille d'échantillon et l'absence de témoins. Bien que les tendances puissent être données, des inférences et des associations de cause à effet ne peuvent être faites. Nous reconnaissons la possibilité de diagnostiquer la maladie et l'effet du biais de sélection des cas dans les populations étudiées. Bien que la collecte des données reproductrices ait été auto-déclarée et qu'elle soit sujette à un préjugé de rappel, la validité des antécédents reproductifs auto-déclarés est élevée. En outre, il peut y avoir un faible niveau d'écart dans le rapport de parité et le statut de la ménopause, l'erreur systématique en raison de préjugés de rappel pour la ménarche et l'âge à la première naissance vivante peut exister, mais probablement pas différenciellement rappelé par les patientes. Étant donné que toutes les participantes avaient déjà un cancer du sein au moment de l'entrevue, il est peu probable qu'il y ait un biais différentiel pour la détermination des cas, comme on le voit dans les cas-témoins. Malgré cela, et en connaissance des travaux antérieurs dans l'un des deux pays et dans d'autres milieux similaires, les résultats représentent un véritable reflet du cancer du sein dans la majeure partie du continent et fournissent des informations de base pour les études futures, en particulier dans la moins documentée des deux populations.

## CONCLUSION PARTIELLE

Dans cette étude, nous avons recherché des similitudes et des différences pour certains facteurs de risque connus pour le cancer du sein et les caractéristiques cliniques entre les patientes atteintes du cancer du sein au Nigéria et au Sénégal. L'âge au diagnostic, l'étape du cancer au diagnostic et les caractéristiques tumorales étaient semblables chez les patientes des deux pays. Nous avons observé des disparités subtiles à la ménarche, à l'âge de la première grossesse à terme, à la parité, à l'état matrimonial et à l'usage de contraceptifs. En raison du jeune âge au diagnostic en particulier chez les femmes en pré-ménopause et du diagnostic à un stade avancé, nous sommes d'avis que l'introduction de programmes de dépistage organisés axés sur la population ciblant cette catégorie de femmes permettrait d'améliorer le traitement. Cette étude a des implications cliniques et politiques importantes pour le contrôle et la prévention du cancer du sein dans la sous-région Ouest Africaine. Cependant, d'autres études, englobant des paramètres cliniques, environnementaux et biologiques sont nécessaires pour mieux déceler l'impact du cancer du sein et contrôler sa menace grandissante dans le continent africain.

## **CHAPITRE III. POLYMORPHISME ET EVOLUTIO GENETIQUES DES MICRO-SATELLITES NR-24 ET NR-24 ET IMPACT DES CARACTERISTIQUES CLINICO-PATHOLOGIQUES**

### **INTRODUCTION**

Les différences génétiques observées dans les populations modernes ont été façonnées par l'évolution du génome africain à travers des facteurs démographiques tels que la migration (distance et fermeture), les fluctuations de la population, la structure de la sous-population, la dérive et le brassage génétique (Tishkoff et Verrelli, 2003 ; Campbell et Tishkoff, 2008). Ceci a conduit à la grande diversité génomique / génétique observée au sein des populations (éthnies et pays) et entre celles-ci sur le continent par rapport à la population non africaine. Une étude d'association pangénomique de 848 polymorphismes de short tandem repeats (STRP), de 476 délétions d'insertion (INDEL), et de 3 polymorphismes nucléotidiques simples (SNP) génotypés chez environ 2400 individus de 121 populations géographiquement diverses, a montré 14 groupes de populations ancestrales génétiquement divergentes en Afrique révélant une vaste structure de population sur le continent (Tishkoff *et al.*, 2009). En raison de cette grande diversité génétique en Afrique, les traitements conçus pour les populations d'ascendance européenne ne sont souvent pas efficaces et produisent souvent de faibles résultats/succès dans le contexte africain.

Le poids le plus lourd de maladies avec des phénotypes variés peut être observé en Afrique et, par conséquent, le continent est riche en échantillons potentiels avec des phénotypes intéressants pour élucider les bases génétiques et environnementales des maladies transmissibles et non transmissibles. L'Afrique subsaharienne a une incidence plus faible de cancer du sein mais un taux de mortalité élevé et le pire résultat par rapport à la population de race blanche. En raison de la variation observée, il est nécessaire de comprendre et de caractériser la diversité génomique et l'évolution génétique de cette population afin d'identifier les variantes qui jouent un rôle dans la progression et la susceptibilité de la maladie, ce qui peut nous aider à trouver des solutions possibles à ces questions liées à la prévention, au contrôle et au diagnostic des maladies (Campbell et Tishkoff, 2010 ; Campbell *et al.*, 2014). Comprendre les processus biologiques et appliquer des méthodes innovantes de diagnostic et de traitement aidera à éliminer la fracture raciale en termes d'incidence et de résultats et permettra d'améliorer les résultats du traitement de la maladie.

L'utilisation de marqueurs moléculaires a été normalisée pour l'étude de l'histoire de l'évolution humaine et la détection de schémas de variations génomiques parmi les populations (Agrafioti et Stumpf, 2007). Les polymorphismes mononucléotidiques (SNP) et les microsatellites sont deux marqueurs moléculaires largement utilisés pour étudier ces processus. Un polymorphisme peut être identifié en tant que SNP lorsque plus d'un nucléotide d'un site de séquence est présent dans une population, généralement avec un seuil de 1%. Les microsatellites sont des motifs d'ADN courts (1 à 10 paires de bases) qui se présentent sous la forme de répétitions en tandem réparties de manière non aléatoire sur de nombreux locus dans le génome, dans les régions codantes et non codantes ainsi que dans les régions non traduites 5' et 3'. (Ellegren, 2004). Ils sont particulièrement susceptibles aux délétions et aux insertions pendant la réplication en fonction de leur longueur - plus la séquence répétée est longue, plus l'instabilité est fréquente (Sagher *et al.*, 1999). Par conséquent, ils peuvent être utilisés pour retracer des événements démographiques récents en raison de leur très faible taux de mutation estimé de  $10^{-2}$  à  $10^{-6}$  nucléotides par locus par génération (Sia *et al.*, 2000 ; Tishkoff et Verrelli, 2003).

Le phénotype d'instabilité des microsatellites (MSI) est la présence de microsatellites de taille alternative dans l'ADN tumoral différents de ceux observés dans l'ADN germlinal correspondant (European Patent Specification, 2010). MSI augmente le taux de mutation de la base, qui est l'un des mécanismes responsables du « phénotype mutant » dans les cancers, entraînant une prolifération, une évolution et une survie démesurées des cellules cancéreuses Hampel *et al.*, 2005 ; Morgan et Reuter, 2006).

Les répétitions de microsatellites dans les régions codantes sont hautement conservées, tandis que celles dans les régions non codantes sont polymorphes, ce qui signifie un taux élevé de mutations spontanées (Suraweera *et al.*, 2001). Les séquences conservées montrées dans certaines des régions non codantes, en particulier des UTR 5' et 3' qui flanquent des séquences codant un gène, sont révélatrices de leur rôle dans l'expression génique (Sonenberg, 1994 ; Pesole *et al.*, 1997 ; Conne *et al.*, 2000). Les microsatellites peuvent être des mononucléotides, des dinucléotides, des trinucléotides, des tétranucléotides et des penta nucléotides. Parmi ceux-ci, les répétitions mononucléotidiques (RMN) représentent le type de microsatellites le plus intéressant avec des longueurs conservées pour celles situées dans les régions codantes, tandis que la longueur de celles contenues dans les régions non codantes peut varier fortement d'un individu à l'autre. (Metzgar *et al.*, 2000). Il existe également dans les régions non codantes un certain nombre de RMN quasi-monomorphes de longueurs comprises entre 20 et 40 pb. Ces RMN présentent des variations de longueur significativement réduites au sein de la population humaine, ce qui peut indiquer la possibilité de leur pertinence fonctionnelle (Woerner *et al.*,

2009). En raison du niveau élevé de variabilité et du taux de mutations qui modifient leur longueur, les microsatellites sont utiles pour étudier la structure de la population pendant des milliers de générations (Bowcock *et al.*, 1994 ; Roy *et al.*, 1994 ; Goldstein *et al.*, 1995 ; Underhill *et al.*, 1996 ; Goldstein et Pollock, 1997 ; Reich et Goldstein, 1998 ; Goldstein *et al.*, 1999).

Le phénotype d'instabilité des microsatellites a été identifié à la fois comme marqueur pronostique et prédictif du cancer colorectal, dans lequel les patientes atteintes de tumeurs MSI-H sont mieux sélectionnées pour certains types de traitement et ont une meilleure survie. Trois facteurs principaux justifient l'utilisation de facteurs pronostiques et prédictifs : l'identification des patients dont le pronostic est bon et qui pourraient bénéficier du traitement systémique adjuvant, ceux dont le pronostic est mauvais pour valider le bien-fondé d'une approche adjuvante plus agressive et enfin pour sélectionner les patients atteints de tumeur qui sont plus ou moins susceptibles de bénéficier de différentes formes de thérapie (Clark, 1994).

Par conséquent, dans ce chapitre, nous avons évalué et comparé la distribution des microsatellites *NR-21* (situé dans la région non traduite en 5' du gène *SLC7A8*) et *NR-24* (situé dans la région non traduite en 5' du gène *ZNF-2*) chez les patientes atteintes d'un cancer du sein au Nigéria et au Sénégal, ainsi que l'impact des caractéristiques clinicopathologiques. Cette étude visait à examiner la distribution évolutive des tumeurs du sein dans les deux pays (si les tumeurs suivent un modèle d'évolution identique ou différent) et à vérifier également l'utilisation possible des marqueurs étudiés pour le diagnostic clinique dans ces populations.

### **III.1. MATERIEL ET METHODES**

#### **III.1.1. Considération Ethique**

Toutes les procédures de traitement des participantes à cette étude ont été réalisées dans le respect des normes et de l'approbation éthique de l'Université d'Ibadan / University College Hospital, Ibadan, Nigéria, et de l'Université Cheikh Anta Diop, Comité d'éthique de la recherche, Dakar (Sénégal) comme indiqué dans le chapitre précédent. L'étude a été menée conformément aux recommandations de la déclaration d'Helsinki sur la recherche biomédicale.

#### **III.1.2. Sites d'étude, Critères d'Inclusion et d'Exclusion**

##### **III.1.2.1. NIGERIA**

Tous les cas consécutifs de cancer du sein ayant fréquenté l'unité d'oncologie chirurgicale du département de chirurgie de l'University College Hospital d'Ibadan, au Nigéria (UCH) entre novembre 2016 et juillet 2017 étaient éligibles pour cette étude. Cet hôpital couvre une

population d'environ 3,5 millions de personnes et constitue un hôpital de référence majeur dans le pays, situé dans la troisième ville la plus peuplée du monde. Ibadan est situé à une latitude de 7,38° et de 3,91° et à 181 mètres d'altitude. La majorité des patientes atteintes d'un cancer du sein et diagnostiquées dans la région, en particulier sur la base de leur référence, sont examinées à l'hôpital de l'University Health Unit. Les cas ont été recrutés lors de la première présentation. Après avoir expliqué le sujet de la recherche et obtenu le consentement éclairé, les patients ont été interrogés à l'aide d'un questionnaire bien structuré. Une biopsie a été réalisée par les chirurgiens, tandis que les mesures de la taille, du poids, de la hanche et de la taille étaient faites par les infirmières. Les patients recrutés avaient entre vingt-neuf (29) et quatre-vingt (80) ans. Il y a eu 2 refus et cinquante (50) cas ont été utilisés pour les analyses. Les échantillons de tissus ont été placés dans des bouteilles ordinaires contenant de l'alcool 100% comme fixateur et conservés à -200° C dans le laboratoire de recherche sur le cancer du sein de l'Institut de recherche médicale avancée (IMRAT) UCH, Ibadan. Les échantillons ont ensuite été transportés dans de la glace en vertu de l'Accord de transfert de matériel au laboratoire commun UCAD-IRD-ISRA/ BIOPASS du Centre de Biologie pour la Gestion des Populations (CBGP) de Bel-Air, Dakar, au Sénégal, pour les différentes étapes de l'analyse moléculaire.

### **III.1.2.2. SENEGAL**

Les cas présentés à l'Institut d'oncologie de l'hôpital Aristide Le Dantec de Dakar au Sénégal, qui ont accepté de participer à l'étude ont signé un consentement éclairé et ont été recrutés pour l'étude. L'hôpital Aristide le Dantec se trouve à Dakar, capitale du Sénégal et plus grande ville du pays. Il s'agit d'un important et du seul hôpital de référence pour les patients atteints de cancer dans le pays, desservant une population immédiate d'environ 2,45 millions de personnes. Dakar est situé entre la latitude 14.690°N et la longitude -17.440°W et se situe à une altitude de 12 mètres au-dessus du niveau de la mer. Les oncologues chirurgicaux ont prélevé un échantillon de tissu cancéreux par biopsie ou par mastectomie chez chaque patient après consentement éclairé. Un questionnaire standardisé a également été utilisé pour collecter les données relatives au patient. Les échantillons de tissus ont été placés dans un flacon contenant de l'alcool 100% comme fixateur. Ils ont ensuite été transférés dans de la glace au laboratoire commun UCAD-IRD-ISRA/BIOPASS du Centre de Biologie pour la Gestion des Populations (CBGP) de Bel-Air, à Dakar, au Sénégal pour les différentes étapes de l'analyse moléculaire.

### III.1.3. Étude du Phénotype les Loci *NR-21* et *NR-24*

#### III.1.3.1. Choix de marqueurs dans l'étude

Les deux répétitions mono-nucléotidiques quasi-monomorphes analysées dans cette étude, *NR-21* et *NR-24*, sont conservées dans l'ADN germinale des Hommes et sont impliqués dans le processus de transcription. Ils sont très sensibles à la délétion somatique dans les tumeurs MSI-H. Le marqueur *NR-21* est une séquence répétée 21T située dans la position 5'UTR du gène *SLC7A8* à la position 11 du bras long (q) du chromosome 14 (14q11.2). Le marqueur *NR-24* est une répétition 24T trouvée sur le bras long (q) du chromosome 2 dans le 3'UTR du gène *Zinc finger-2* (2q11.1).

L'utilisation de marqueurs *BAT-25* et *BAT-26* en tant que combinaison a été proposée pour la détection de MSI (Hoang *et al.*, 1997 ; Zhou *et al.*, 1998). Ils sont quasi-monomorphes dans les populations de race blanche, ce qui leur permet d'être facilement classés en fonction des importantes variations de taille alléliques observées dans l'ADN de tumeur MSI en raison d'une altération somatique (Hoang *et al.*, 1997 ; Zhou *et al.*, 1997). La majorité des analyses des tumeurs colorectales des *BAT-25* et *BAT-26* établissent leur statut MSI sans référence à l'ADN germinale (Zhou *et al.*, 1998), mais certaines études ont identifié leurs allèles de taille alternatifs chez environ 19% des Afro-Américains (Pyatt *et al.*, 1999 ; Samowitz *et al.*, 1999). Par conséquent, l'analyse de répétitions supplémentaires a été proposée pour éviter les résultats faux positifs occasionnels provenant de ces polymorphismes de la lignée germinale. En réponse à ce besoin, nous avons décidé de mener cette étude pour identifier le phénotype MSI de *NR-21* et *NR-24* chez des patientes atteintes d'un cancer du sein au Nigéria et au Sénégal et de comparer les variations polymorphes de ces marqueurs entre les deux populations.

#### III.1.3.2. Extraction de l'ADN

L'ADN obtenu à partir du tissu tumoral de patients atteints de cancer du sein par biopsie à l'aiguille fine, biopsie tru-cut ou mastectomie a été extrait à l'aide du kit standard BioBasic, codé et stocké à -20°C pour les analyses PCR.

L'extraction de l'ADN a été réalisée selon le protocole des kits BioBasic avec un léger ajustement effectué par l'équipe de Génétique pour la Gestion des Populations, BIOPASS, Dakar, Sénégal. Pour ce faire, 20 mg de chaque échantillon de tissu ont été pesés, coupés et taquinés pour faciliter la digestion. Le tissu pesé a été taquiné et placé dans un tube eppendorf de micro-centrifugeuse de 1,5 ml correctement étiqueté contenant 100 µl de tampon de digestion du LCA pour chaque échantillon. Le tampon ATL est un détergent qui provoque la

dissociation tissulaire et l'individualisation des cellules. Un volume de 20 µl de protéinase K a été ajouté au mélange pour dissocier l'association protéine-ADN-histone, puis le mélange a été vortexé et incubé à 55°C pendant une nuit jusqu'à ce que le tissu soit complètement lysé. Après l'incubation, les artefacts ont été éliminés après centrifugation rapide à 12 000 radian/minute (1256,6371 radian/seconde) et le surnageant a été récupéré. Puis, 200 µl de solution AB ont été ajoutés au mélange, agités au vortex et incubés à 70°C pendant 10 minutes. Ensuite, 200 µl d'éthanol absolu ont été ajoutés au mélange pour précipiter l'ADN et vortexés vigoureusement pendant 20 secondes.

Le lysat a ensuite été transféré sur une colonne de membrane de silice préalablement placée sur un tube collecteur de 2 ml (fourni dans le kit) ; puis centrifugé à 13 000 tr/min (1361,35688 tr/sec) pendant 3 minutes pour retenir l'ADN chargé négativement au niveau de la membrane de la colonne de silice chargée positivement. Par conséquent, les protéines, les lipides et les polysaccharides ont été éliminés, laissant un précipité au fond du tube. L'ADN obtenu a été encore purifié pour éliminer toutes les formes de contaminants. Le processus de lavage a été effectué deux fois en ajoutant ensuite 500 µl du tampon de lavage et en centrifugeant le mélange à 13 000 tr/min (1361,35688) pendant 1 et 3 minutes, respectivement. Enfin, les colonnes contenant chaque échantillon ont été placées dans un autre tube eppendorf de 1,5 ml. 50 µl de tampon d'élution préchauffé à 70°C pour augmenter le rendement de 15 à 20% ont été ajoutés directement à la membrane, suivie d'une incubation à la température ambiante pendant 2 min pour la remise en suspension de l'ADN et d'une centrifugation à 13 000 tr/min pendant 1 min. Les échantillons d'ADN obtenus ont ensuite été stockés à 4°C pour être utilisés pour l'analyse PCR.

### **III.1.3.3. Réaction en Chaîne par Polymérase (PCR)**

Les conditions de la PCR sont optimisées pour chaque paire d'amorces et appliquées uniformément pour tous les échantillons, ce qui permet de comparer les résultats des analyses. Les amorces utilisées pour cette étude sont consignées dans le tableau V.

**Tableau V : Séquences d'amorces pour le test d'instabilité de microsatellites (Buhard *et al.*, 2006)**

Microsatellites	Gènes	Numéro de Genbank	Répétitipns	Les séquences d'amorces	Taille de l'amplicon (pb)
<i>NR-21</i>	<i>SLC7A8</i>	XM_033393	21T 5'UTR	F : GAGTCGCTGGCACAGTTCTA R : CTGGTCACTCGCGTTACAA	107-110
<i>NR-24</i>	<i>Zinc-finger 2</i>	X60152	24T 3'UTR	F : GCTGAATTTTACCTCCTGAC R : ATTGTGCCATTGCATTCCAA	126-129

Les amplifications furent effectuées dans un volume réactionnel de 50 µl et les conditions de la réaction PCR pour chacun des tubes à échantillons sont consignées dans le tableau VI. Enfin, 49 µl du mélange réactionnel total sont distribués dans chaque tube d'échantillon et 1 µl de chaque extrait d'ADN échantillon y est ajouté. Pour quelques-uns des échantillons contenant de l'ADN très concentré, celui-ci a été dilué au 1/10.

**Tableau VI : Composition du mélange de réaction PCR**

Composition du mélange réactionnel	Volume pour chaque marqueur	
	<i>NR-21</i>	<i>NR-24</i>
<b>Marqueurs</b>		
<b>Eau Milli-Q</b>	34.9 µl	34.9 µl
<b>Tampon</b>	5 µl	5 µl
<b>MgCl<sub>2</sub></b>	2 µl	2 µl
<b>Chaque dNTP</b>	2 µl	2 µl
<b>Amorce forward</b>	2.5 µl	2.5 µl
<b>Amorce reverse</b>	2.5 µl	2.5 µl
<b>Taq polymérase</b>	0.1 µl	0.1 µl
<b>Extrait DNA</b>	1 µl	1 µl
<b>Total</b>	50 µl	50 µl

La PCR a été réalisée dans un thermocycleur de type Eppendorf dans les conditions indiquées dans le tableau VII et les conditions étaient les mêmes pour les deux marqueurs microsatellites.

Il a été réalisé en 35 cycles de dénaturation et en maintien pendant 10 minutes avant la sortie des amplicons pour électrophorèse sur gel afin de vérifier la qualité de la PCR (Remarque : le maintien peut aller à l'infini, mais dans cette étude, les réactions ont été laissées en attente pendant au moins 10 minutes).

**Tableau VII : Étapes de la PCR et conditions d'amplification**

<b>Conditions du thermocycleur PCR</b>	
Dénaturation (initiale))	94 <sup>0</sup> C pendant 5 minutes
Dénaturation (35 cycles)	94 <sup>0</sup> C pendant 30 seconds
Hybridation	55 <sup>0</sup> C pendant 30 seconds
Elongation	72 <sup>0</sup> C pendant 30 secondes
Extension finale	72 <sup>0</sup> C pendant 7 minutes
Conservation (Hold)	10 <sup>0</sup> C ~ infinie

Les produits de PCR sont ensuite soumis à la migration électrophorétique sur un gel d'agarose à 1,5% provenant de 7 µl de chaque échantillon et de 3 µl de bleu de bromophénol.

#### **III.1.3.4. Séquençage**

Les amplicons (30 µl chacun) ont ensuite été placés dans une microplaque de microtitration à 96 puits avec chaque amorce (15 µl de 10 µm) et pour chaque échantillon et envoyés à la société MacroGen, South Korea et New York pour le séquençage Sanger (Mbaye *et al.*, 2015).

#### **III.1.4. Nomenclature et classification de variantes de séquences**

Antonarakis et le groupe de travail sur la nomenclature en 1998 ont suggéré un système de nomenclature pour la description des mutations et des polymorphismes observés dans les séquences d'ADN et de protéines. Cette numérotation est connue sous le nom de recommandations / nomenclature de la Société de la variation du génome humain (HGVS) (den Dunnen et Antonarakis, 2000). Par conséquent, la nomenclature des variantes de séquences constitue les modifications observées dans une séquence spécifique lorsque cette séquence est comparée à une séquence de référence. Par la suite, afin de remédier aux erreurs et aux incohérences et d'obtenir une description stable, précise, significative, non équivoque et sans ambiguïté mais souple, le groupe de travail sur la description des

variantes de séquences (SVD-WG) a achevé sa procédure en 2015. Ce groupe comprend Organisation (HUGO), le Projet du variome humain (HVP) et le HGVS (den Dunnen *et al.*, 2016). Les variantes les plus prioritaires sont : inversion, suppression, duplication, insertion et conversion. Les autres sont de type substitution, suppression-insertion, translocation, transposition, conversion génique. Les définitions suivantes des variantes de séquences sont telles que rapportées par den Dunnen et Antonarakis, 2001 ; den Dunnen *et al.*, 2016).

#### **III.1.4.1. Variante DELETION**

La délétion est utilisée pour identifier un changement dans une séquence spécifique au cours de laquelle un ou plusieurs nucléotides sont excisés par rapport à la séquence de référence. Ils sont décrits en utilisant "del" après une indication du premier et du dernier nucléotide (s) délétés qui sont séparés par un trait de soulignement ("\_").

#### **III.1.4.2. Variante INSERTION**

Une insertion est un changement de séquence où, par rapport à la séquence de référence, un ou plusieurs nucléotides sont insérés et où l'insertion n'est pas une copie d'une séquence immédiatement en 5'. Les séquences de nucléotide (s) insérées sont ajoutées entre deux nucléotides adjacents dans la séquence et sont désignées par "ins" après les nucléotides encadrant l'insertion.

#### **III.1.4.3. Variante SUBSTITUTION**

Un changement de séquence où, comparé à une séquence de référence, un nucléotide est remplacé par un autre nucléotide. Les substitutions sont désignées par un caractère ">" après le numéro du nucléotide affecté.

#### **III.1.4.4. Variante DUPLICATION**

En duplication, une séquence change lorsque, comparée à une séquence de référence, une copie d'un ou plusieurs nucléotides est insérée directement en 3' de la copie originale de la séquence. Ils sont désignés par "dup" après une description du segment dupliqué qui est le premier (et le dernier) nucléotide (s) dupliqué (s). Ceci est applicable même pour les mononucléotides.

### **III.1.4.5. Variante INSERTION PAR DELETION**

La variante d'insertion par délétion est parfois décrite comme une variante complexe et est un changement de séquence dans une séquence spécifique qui, comparée à une séquence de référence, remplace un ou plusieurs nucléotides par un ou plusieurs autres nucléotides et n'est pas une substitution, une inversion ou une substitution conversion. Aussi connu sous le nom ("indels"), ils sont décrits comme une suppression ("del"), suivie d'une insertion ("ins"). Exemple g.6775delinsT

### **III.1.4.6. Variante INVERSION**

Cela se produit lorsqu'un changement de plus d'un nucléotide dans une séquence spécifique remplace la séquence d'origine par rapport à la référence. Il est souvent représenté par "inv".

## **III.1.5. Analyses moléculaires pour la génétique des populations à l'aide de microsatellites**

### **III.1.5.1 : Analyse du polymorphisme et de la diversité génétique**

Les séquences de *NR-21* et *NR-24* obtenues après le séquençage ont été analysées à l'aide de plusieurs logiciels bioinformatiques incorporant des approches d'analyses de génétique des populations. Il s'agit d'une étape importante dans le processus d'analyse des données car elle met en évidence les similitudes (homologies) et les différences entre les séquences utilisant le format de fichier fasta en recherchant les positions de délétion et d'insertion. Pour tester l'association entre le cancer du sein et l'instabilité des marqueurs microsatellites dans la population étudiée, les séquences obtenues ont été alignées avec les séquences de référence de chaque marqueur en utilisant le numéro de la banque de gènes comme dans Buhard *et al.*, (2006) et indiqué dans le tableau V. Les séquences ont été alignées à l'aide de l'algorithme Clustal de la version 7.1.9 de BioEdit Sequence Alignment Editor (Thompson *et al.*, 1994 ; Hall, 1999). L'identification d'une mutation après l'alignement d'une séquence d'ADN a été suivie par une nouvelle visualisation manuelle du chromatogramme de la séquence d'ADN pour valider si la mutation observée était présente ou non.

Par la suite, pour chacun des marqueurs, le nombre des Séquences (N), le nombre d'haplotypes (H), le nombre de sites polymorphes ( $\theta_s$ ), le nombre moyen de différences par paire de séquences (K), les diversités haplotypique (h) et nucléotidique ( $\pi$ ) ont été estimés pour la population étudiée en utilisant les logiciels DnaSP (pour les trois premiers) et

l'Arlequin (pour les trois derniers) (Nei, 1987 ; Rozas *et al.*, 2012). Le nombre moyen de différences nucléotidiques par site par paire était estimé par  $\pi$  (Tajima, 1989), tandis que la valeur  $\theta_w$  est basée sur le nombre de sites de ségrégation dans une population (Watterson, 1975). La diversité haplotypique prend en compte le nombre d'individus et la fréquence des haplotype ( $h$ ) et montre la probabilité que deux séquences sélectionnées de manière aléatoire au sein d'une population spécifique soient identiques (Nei, 1987). Par conséquent, une grande diversité d'haplotypes montre une probabilité plus grande d'observer différents haplotypes parmi deux individus sélectionnés au hasard. La diversité des nucléotides, bien qu'une mesure de la diversité génétique, incorpore le degré réel de différences entre les séquences étudiées. Ainsi, plus les haplotypes présents dans une population sont distincts les uns des autres en termes de nombre de sites polymorphes les différenciant, plus la diversité en nucléotides est élevée.

Pour vérifier l'instabilité, les différences observées pour chaque microsatellite dans la population étudiée par rapport à la référence en termes de taille et de configuration des variantes (délétion, insertion, transition ou transversion) ont été considérées comme représentant l'instabilité. Etant donné que l'ADN normal n'a pas été examiné, chaque échantillon présentant un polymorphisme dans l'une des répétitions mono nucléotidiques *NR-21* et *NR-24* a été évalué positif pour l'instabilité (Wong *et al.*, 2006). Les tumeurs ont été définies comme étant des microsatellites stables, si elles étaient stables pour les deux marqueurs ; une instabilité microsatellite faible (MSI-L), si elles étaient instables pour un marqueur et une instabilité microsatellite élevée (MSI-H) si elles étaient instables pour les deux marqueurs.

### **III.1.5.2 : Analyse de l'évolution génétiques**

#### **III.1.5.2.1 : Analyse de la Structuration génétique**

Les  $F_{ST}$ , qui mesure la différenciation génétique et l'analyse de la variance moléculaire (AMOVA) qui indique la source de la variation moléculaire, ont été déterminés globalement pour la population d'étude et indépendamment pour chaque pays. Les  $F_{ST}$  et leurs probabilités associées ont été calculées à l'aide du logiciel Arlequin version 3.0 (Weir et Hill, 2002 ; Excoffier *et al.*, 2005). Le  $F_{ST}$  pour le test de neutralité montre la dérive génétique et le motif de différenciation de la tumeur dans la même direction sur tous les loci le long de la voie du génome pour une population donnée (Beaumont et Baldoing, 2004 ; Excoffier *et al.*, 2009). Le

$F_{ST}$  détecte les événements de sélection positifs récents et les événements spécifiques aux SNP, ce qui permet de cibler directement la variante en cours de sélection.

Ces valeurs de différenciation génétique ont été estimées entre le Nigéria et le Sénégal et leurs probabilités associées ont été obtenues avec leur  $p$  - *value*. La différence observée entre les deux pays a été acceptée comme vraie si la  $p$  - *value* du  $F_{ST}$  était significative à ( $P < 0,05$ ) ou rejetée comme fausse lorsque la  $p$  - *value* du  $F_{ST}$  global n'était pas significative ( $P > 0,05$ ). Les loci soumis à une sélection divergente devraient avoir des valeurs de  $F_{ST}$  plus élevées que les locus neutres, tandis que les locus soumis à une sélection d'équilibre devraient avoir des valeurs de  $F_{ST}$  inférieures (Lewontin et Krakauer, 1973). Sous neutralité, le  $F_{ST}$  est uniquement déterminé par la dérive génétique qui affectera tous les loci situés le long du génome dans une voie similaire.

### **III.1.5.2.2 : Test de Neutralité pour la variabilité intra-spécifique et inter-spécifique**

Les tests de neutralité intra-spécifiques sont basés sur le polymorphisme observé chez les populations humaines. Différentes approches statistiques ont été utilisées pour déterminer l'évolution moléculaire et estimer la diversité des haplotypes et des nucléotides. Les paramètres statistiques ont été calculés à l'aide de la version 3.0 d'Arlequin et du niveau de signification évalué après 10000 simulations sur coalescence (Excoffier et Lischer, 2010). Le polymorphisme InDel a été calculé sur DnaSP en utilisant Kimura 2-parameter (K2P) (Librado et Rozas, 2009).

Le test D de Tajima évalue une déviation générale du spectre de fréquences alléliques par rapport aux prévisions dans des conditions neutres. Ce test prend en compte l'ensemble du spectre : il mesure le nombre moyen de différences observées entre les paires de séquences et le nombre total de sites polymorphes. Les valeurs attendues du D de Tajima et du  $F_s$  de Fu sous neutralité seront zéro. Les valeurs seront négatives sous sélection positive ou sous sélection négative en raison d'un excès de mutations rares, elles seront positives sous sélection équilibrée en raison d'un excès d'allèles de fréquence intermédiaire. Les autres tests qui mesurent également un excès ou un déficit en mutations rares incluent les statistiques  $D^*$  et  $F^*$  de Fu et Li (Fu et Li, 1993). Ces tests sont similaires à ceux du D de Tajima et mesurent quant à eux l'excès ou le déficit en mutations présentes qu'une seule fois dans l'échantillon (singletons). Les valeurs négatives et positives indiquent respectivement une sélection positive / négative et une sélection équilibrée. Toutes ces statistiques doivent être proches de zéro en neutralité.

### III.1.5.3 : Analyse d'expansion Génétique

L'expansion génétique, qui se caractérise par l'existence d'un écart par rapport à l'hypothèse de neutralité et d'une différence de nucléotide par paire dans les tissus cancéreux, a été étudiée en calculant le D de Tajima et le Fs de Fu (Tajima, 1989, Fu, 1997). Si la *p - value* pour le D de Tajima et le Fs de Fu est significative à  $p < 0,05$  ou  $p < 0,02$  respectivement, nous acceptons l'hypothèse d'une expansion tumorale pour la population évaluée mais le niveau est non significatif à  $p > 0,05$  pour le D de Tajima et  $p > 0,02$  pour Fs de Fu. La valeur de Fs de Fu est plus pertinente pour une petite population (taille de l'échantillon) comme dans cette étude.

L'analyse de Mismatch distribution, qui apprécie les distances génétiques entre chaque paire de tissus cancéreux a été déterminé en considérant l'écart entre les distributions observées et attendues afin de tester la validité du modèle d'expansion avec le logiciel DnaSP (Rogers et Harpending 1992). Une courbe unimodale indique une absence d'expansion, alors qu'une courbe multimodale soutient l'expansion de la tumeur dans la population donnée. La somme de la déviation des carrés (SSD) et l'indice de raggedness de harpending (RI) ont été calculées à l'aide du logiciel Arlequin (Schneider et Excoffier 1999 ; Excoffier *et al.*, 2005). Une valeur statistiquement significative du SSD à  $p = 0,005$  soutient le modèle d'expansion.

Bien que tous ces tests reposent sur le même principe, ils couvriront différentes parties du spectre des fréquences alléliques.

### III.1.6 Analyses Statistiques

L'analyse statistique a été réalisée à l'aide de SPSS version 20 (SPSS, Inc. Chicago IL, États-Unis). Les résultats obtenus ont été enregistrés en tant que valeurs du khi deux (Chi-2) pour les variables qualitatives et du t-test pour les variables continues. La valeur statistiquement significative a été placée à *p - value*  $< 0,05$ .

Les paramètres utilisés pour étudier la corrélation entre les marqueurs étudiés et les paramètres clinico-pathologiques dans ce chapitre sont : l'âge au diagnostic (50 ans,  $\geq 50$  ans) ; le stade de la tumeur (1 à 4), le grade SBR (grade 1 - 3 représentant respectivement une tumeur bien différenciée, une tumeur moyennement différenciée et une tumeur faiblement différenciée) ; et si la tumeur était invasive ou non invasive pour chaque cas. Pour chaque population étudiée, les caractéristiques étudiées ont été regroupées en tant que variables catégorielles afin de déterminer la corrélation entre chaque paramètre et le

phénotype MSI de *NR-21* et *NR-24*. Un t-test sur échantillon indépendant a également été réalisé pour comparer les moyennes de certaines caractéristiques continues chez ces patientes. Les variables continues étaient : l'âge moyen au diagnostic, l'âge à la ménarche, l'âge à la première grossesse à terme et la parité, et les résultats sont présentés en moyenne  $\pm$  S.D.

## III.2. Résultats et Discussion

### III.2.1. Résultats

#### III.2.1.1. Phénotype d'instabilité des microsatellites *NR-21* et *NR-24* chez les patients atteints de cancer du sein au Nigéria et au Sénégal

Quatre-vingt-seize échantillons au total ont été chacun séquencés pour les deux marqueurs microsatellites. Après alignement, quatre-vingt-treize séquences ont été obtenues pour le *NR-21* au Nigéria et au Sénégal (47 séquences par pays) et deux séquences qui n'étaient pas propres ont été retirées des l'analyses. Pour *NR-24*, quatre-vingt-sept séquences claires et correctement alignées ont été obtenues (44 séquences pour le Nigéria et 43 séquences pour le Sénégal), dix séquences qui n'étaient pas propres et ont été retirées des l'analyses. Par conséquent, les données analysées dans cette étude sont celles des séquences propres et correctement alignées. Pour éviter tout biais, l'alignement de chaque marqueur microsatellite étudié a ensuite été confirmé manuellement avec le chromatogramme de chaque individu. Les résultats pour chaque pays et pour chaque marqueur et pour les deux marqueurs sont résumés dans le tableau VIII.

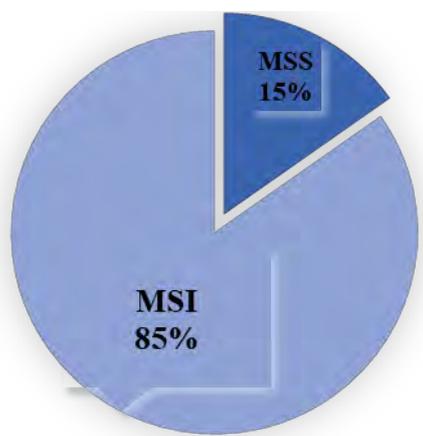
**Tableau VIII :** Résultats du phénotype MSI chez les patientes atteintes d'un cancer du sein au Nigeria et au Sénégal

Marqueurs	MSI status	Nigeria : No. (%)	Senegal : No. (%)	Total : No. (%)
<i>NR-21</i>	MSS	7 (15.2)	7 (14.9)	14 (15.1)
	MSI	39 (84.8)	40 (85.1)	79 (84.9)
<i>NR-24</i>	MSS	13 (29.5)	8 (18.6)	21 (24.1)
	MSI	31 (70.5)	35 (81.4)	66 (75.9)
Deux marqueurs Instables	MSS	2 (4.3)	2 (4.7)	4 (4.5)
	MSI-L	20 (43.5)	12 (27.9)	32 (35.9)
	MSI-H	24 (52.2)	29 (67.4)	53 (59.6)

Pour le Nigéria, deux tumeurs (4,3%) ont été stables, vingt (43,5%) ont présenté une MSI-L et vingt-quatre (52,2%) ont présenté une tumeur MSI-H. Au Sénégal, deux patientes (4,7%)

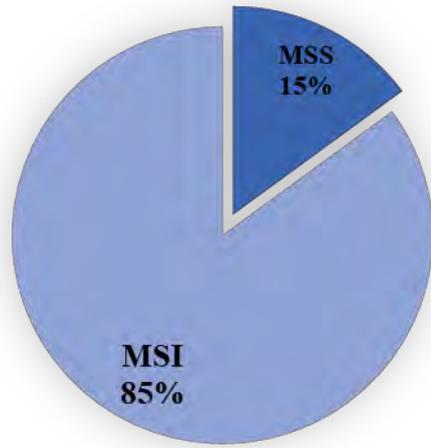
atteintes de tumeurs MSS, douze (27,9%) et vingt-neuf individus (59,6%) atteints de tumeurs MSI-L et MSI-H ont été décelés, respectivement. Les résultats obtenus n'étaient pas statistiquement significatifs pour les deux marqueurs (données non présentées). Les figures 18 et 19 montrent les pourcentages d'individus stables et instables dans chaque pays pour *NR-21* et *NR-24*. Le pourcentage d'individus stables et instables pour le *NR-21* était le même pour chaque pays (MSS - 15% ; MSI - 85%). Mais pour *NR-24*, 81% de la population à d'étude étaient instables pour le Sénégal, 70% des Nigériennes présentaient une instabilité. Ces données représentent également les pourcentages d'allèles variantes pour chaque marqueur.

Figure 18a



Nigeria\_NR-21

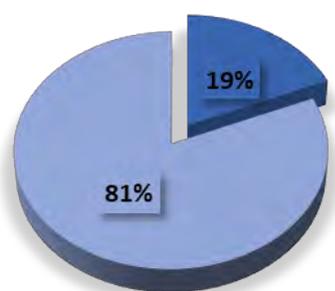
Figure 18b



Senegal\_NR-21

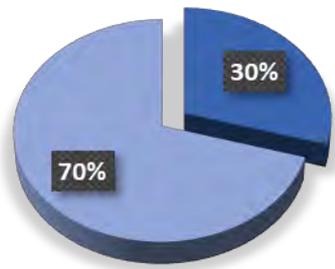
Figure 18 : Phénotype MSI *NR-21* pour le Nigéria et le Sénégal

Figure 19a



Senegal\_NR-24

Figure 19b



Nigeria\_NR-24

Figure 19 : Phénotype MSI *NR-24* pour le Nigéria et le Sénégal

Le tableau IX montre la distribution des haplotypes et motifs pour *NR-21* pour le Nigéria et le Sénégal. Il y avait huit haplotypes différents au total pour les deux pays avec le plus grand nombre d'haplotypes (8) présents au Sénégal tandis que le Nigéria avait 4 haplotypes. L'haplotype 6 avec les répétitions 22T présentait la fréquence la plus élevée pour l'ensemble de la population étudiée (15 au Nigéria et 23 au Sénégal). Il y avait quatre haplotypes de *NR-21* pour le Nigéria avec le plus grand nombre de fréquences d'haplotypes (16) ayant des répétitions 23T et la plus basse fréquence d'haplotypes pour les allèles variantes (8) ayant des répétitions 24T. Le nombre de cas avec les répétitions de type sauvage (21T) était le même (7) pour chaque pays (tableau IX). Pour le Sénégal, il y avait deux haplotypes avec la fréquence minimale ayant chacun 16T et 17T répétées (tableau IX).

Les tableaux X et XI montrent les motifs et la distribution des haplotypes pour *NR-24* respectivement pour le Nigéria et le Sénégal. *NR-24* avait le plus grand nombre d'haplotypes (23 haplotypes) que *NR-21*. Les répétitions 24T de type sauvage avaient la fréquence de distribution la plus élevée (Hap 1), le Nigéria ayant le nombre le plus élevé (13) et la fréquence 8 pour le Sénégal. Parmi les déviations par rapport au type sauvage, Hap 4 avec une substitution de G (24T>G) était le plus présent avec 10 haplotypes tous trouvés dans la population sénégalaise. Huit haplotypes au total avaient une fréquence d'un. Le Sénégal avait également le plus grand nombre d'haplotypes pour le *NR-24* (16 haplotypes), tandis que le Nigéria comptait 12 haplotypes. La variante la plus commune observée ici était la duplication (73 séquences) de la répétition thymine jusqu'à 3 nucléotides.

Tableau IX : Tailles et fréquences allele du *NR-21* au Nigeria et au Senegal

HAPLO -TYPES	Tailles des allèles	Motifs	NIG : (Fréq.)	SEN : (Fréq.)	Total (Fréq.)
HAP 1	1_5delT (16T)	-----TTTTTTTTTTTTTTTTTT	-	1	1
HAP 2	1_4delT (17T)	----TTTTTTTTTTTTTTTTTTT	-	1	1
HAP 3	1_3delT (18T)	---TTTTTTTTTTTTTTTTTTTT	-	2	2
HAP 4	1delT (20T)	-TTTTTTTTTTTTTTTTTTTTTT	-	2	2
HAP 5	21T (type sauvage)	TTTTTTTTTTTTTTTTTTTTTTT	7	7	14
HAP 6	22T	TTTTTTTTTTTTTTTTTTTTTTTT	15	23	38
HAP 7	23T	TTTTTTTTTTTTTTTTTTTTTTTTT	16	9	25
HAP 8	24T	TTTTTTTTTTTTTTTTTTTTTTTTT	8	2	10
<b>TOTAL</b>			46	47	93

Les différents motifs d'haplotype pour *NR-21* et *NR-24* sont présentés respectivement dans les tableaux IX et X.

**Tableau X** : Liste des haplotypes et des motifs *NR-24*

Haplotypes	Fréquences	Motifs
H1	21	TTTTTTTTTTTTTTTTTTTTTTTTTT
H2	4	TTTTTTTTTTTTTTTTTTTTTTTTTT
H3	10	TTTTTTTTTTTTTTTTTTTTTTTTTG
H4	1	TTTTTTTTTTTTTTTTTTTTTTTT-TT
H5	4	TTTTTTTTTTTTTTTTTTTTTTTT-TG
H6	1	-TTTTTTTTTTTTTTTTTTTTGTTT
H7	2	TTTTTTTTTTTTTTTTTTTTTT-GTTT
H8	1	TTTTTTTTTTTTTTTTTTTTGTTTT
H9	3	TTTTTTTTTTTTTTTTTTTTGTTTG
H10	1	TTTTTTTTTTTTTTTTTTTTGT-TG
H11	7	TTTTTTTTTTTTTTTTTTTTGTG--
H12	8	TTTTTTTTTTTTTTTTTTTTGT---
H13	2	--TTTTTTTTTTTTTTTTTTTTGT--
H14	2	TTTTTTTTTTTTTTTTTTTTGT--T-
H15	5	TTTTTTTTTTTTTTTTTTTT-TTTG-
H16	1	TTTTTTTTTTTTTTTTTTTT-TT-G-
H17	1	TTTTTTTTTTTTTTTTTTTT--TTG-
H18	2	TTTTTTTTTTTTTTTTTTTT--GTT-
H19	2	TTTTTTTTTTTTTTTTTTTT--GTG-
H20	1	TTTTTTTTTTTTTTTTTTTT--GT--
H21	4	TTTTTTCCCTTTTTTTTTTTTTTT-
H22	3	-TTTTTTTTTTTTTTTTTTTTTTG-
H23	1	TTTTTTTTTTTTTTTTGTTG-----

**Tableau XI : Tailles et fréquence des'allèles du NR-24 au Nigéria et au Sénégal**

Haplotypes	Tailles des allèles	Nigéria: No. (Fréq. relative.)	Sénégal: No. (Fréq.relative)	Total
Hap 1	<b>24T (type sauvage)</b>	<b>13 (0.295)</b>	<b>8 (0.186)</b>	21
Hap 2	24dupT	3 (0.0682)	1 (0.0233)	4
Hap 3	24T>G	- (0)	10 (0.233)	10
Hap 4	22delT	- (0)	1 (0.233)	1
Hap 5	22delT; 24T>G	1 (0.0227)	3 (0.0698)	4
Hap 6	1delT; 22delinsG; 24dupT	- (0)	1 (0.0233)	1
Hap 7	21delT; 22delinsG; 24dupT	- (0)	2 (0.0465)	2
Hap 8	20delinsG	- (0)	1 (0.0233)	1
Hap 9	20delinsG; 24T>G	1 (0.0227)	2 (0.0465)	3
Hap 10	20delinsG; 22delT; 24T>G	1 (0.0227)	- (0)	1
Hap 11	20delinsG;22T>G; 23_24delT	7 (0.159)	- (0)	7
Hap 12	20delinsG; 22_24delT	7 (0.159)	1 (0.0233)	8
Hap 13	1_2delT; 22delinsG; 24delT	2 (0.0455)	- (0)	2
Hap 14	20delins; 22_23delT	2 (0.0455)	- (0)	2
Hap 15	20delT; 24T>G	- (0)	5 (0.116)	5
Hap 16	20delT; 23delT; 24T>G	3 (0)	1 (0.0233)	1
Hap 17	20_21delT; 24T>G	4 (0)	1 (0.0233)	1
Hap 18	20_21delT; 22delinsG	5 (0)	2 (0.0465)	2
Hap 19	20_21delT; 22delinsG; 24T>G	3 (0.0455)	6 (0)	2
Hap 20	20_21delT; 22delinsG; 24delT	1 (0.0227)	4 (0)	1
Hap 21	8_10T>C	4 (0.0909)	2 (0)	4
Hap 22	1delT; 24T>G	5 (0)	3 (0.0698)	3
Hap 23	17T>G; 20delinsG; 21_24delT	6 (0)	1 (0.0233)	1
<b>Total</b>		<b>44</b>	<b>43</b>	<b>87</b>

### III.2.1.2. Paramètres de la diversité génétique

Le tableau XII présente un résumé des paramètres de la diversité génétique pour chaque marqueur montrant la diversité InDel, la diversité haplotypique InDel  $k(i)$  et la diversité InDel par site  $P(i)$ .

Pour le *NR-21*, nous avons observé une diversité haplotypique ( $h$ ) élevée (Nigéria :  $h = 0,7353 \pm 0,0277$  ; Sénégal :  $0,7105 \pm 0,0557$ ) et une diversité nucléotidique ( $\pi$ ) faible pour chaque pays (Nigéria :  $\pi = 0,0000 \pm 0,0000$  ; Sénégal :  $0,0069 \pm 0,0121$ ). Il y avait une différence de 0,0248 dans la diversité des haplotypes entre les deux pays, avec une plus grande diversité d'haplotypes parmi les tumeurs du Nigéria que celles du Sénégal. Le Sénégal a montré une plus grande diversité nucléotidique dans chaque tumeur alors que pour le Nigéria, nous n'avons pas observé de diversité dans les tumeurs pour le *NR-21*. De plus, le nombre de différences par paires moyennes était de  $0,0000 \pm 0,0000$  et  $0,152 \pm 0,196$  respectivement au Nigéria et au Sénégal, avec une différence plus grande observée au Sénégal qu'au Nigéria. Ceci corrobore avec la diversité de nucléotides nulle observée pour le Nigéria.

Pour le *NR-24*, les résultats étaient similaires avec une haute diversité haplotypique de  $0,8605 \pm 0,0320$  et  $0,8981 \pm 0,0268$  respectivement pour le Nigéria et le Sénégal et une diversité nucléotidique faible (Nigéria :  $0,0298 \pm 0,0286$  ; Sénégal :  $0,0026 \pm 0,0072$ ), avec une différence moyenne correspondante de  $0,5074 \pm 0,4384$  et  $0,0465 \pm 0,1165$  respectivement pour le Nigéria et le Sénégal. Une plus grande diversité haplotypique a été observée au Sénégal par rapport au Nigéria, avec une différence de 0,0376 entre les deux pays. Alors que la différence observée en nucléotides était de 0,0272 entre les deux populations, une différence plus grande a été observée au Nigéria. De même, la différence moyenne par paire était plus grande au Nigéria qu'au Sénégal avec une différence de 0,4609 entre les deux pays.

Tableau XII : Paramètres de diversité génétique des séquences des microsatellites *NR-21* et *NR-24* des tissus cancéreux

	<i>NR-21</i>		<i>NR-24</i>	
	Nigéria	Sénégal	Nigéria	Sénégal
Nombre total de séquences utilisés	46	47	44	43
Nombre total de sites (hors sites présentant des lacunes)	21	16	17	18
Nombre de sites polymorphes	3	8	11	8
Nombre total de sites InDels	3	8	8	7
Nombre total d'évènements de sites InDels	3	7	8	9
Nombre d'haplotypes InDel	4	8	9	11
Diversité Haplotypique (InDel)	0,735	0,710	0,786	0,742
Diversité InDel, k(i) (InDel)	1,177	1,488	1,581	1,200
Diversité InDel par site P(i)	0,0490	0,0619	0,0633	0,0480
Diversité haplotypique (h±SD)	0,7353 ± 0,0277	0,7105 ± 0,0557	0,8605 ± 0,0320	0,8981 ± 0,0268
Diversité nucléotidique (π±SD)	0,0000 ± 0,0000	0,0069 ± 0,0121	0,0298 ± 0,0286	0,0026 ± 0,0072
Nombre moyen de différences par paire (k)	0,0000 ± 0,0000	0,1258 ± 0,1962	0,5074 ± 0,4384	0,0465 ± 0,1165

### III.2.1.3. Structuration génétique

Les tableaux XIII et XIV montrent le niveau de structuration des résultats d'AMOVA par locus et le degré de différenciation génétique à la fois globalement et spécifiquement par pays pour *NR-21* et *NR-24* respectivement.

Pour *NR-21*, le pourcentage de variation au sein de chaque pays était de 91,5214 et entre les deux pays 8,4786. Bien que la valeur du  $F_{ST}$  soit faible (0,0848), elle était néanmoins significative (0,0019), indiquant que la structuration observée entre le Nigéria et le Sénégal est déterminante. Egalement, le *NR-24* a montré plus de variabilité entre les tumeurs de patientes du même pays (au sein d'une population : 88,5986) qu'entre les deux pays (11,4015) avec une valeur du  $F_{ST}$  statistiquement significative (0,0000) entre les deux populations mais non significative au sein de chacune d'elles.

La composition en nucléotides pour *NR-21* était celle d'un motif microsatellite parfait de répétitions de thymine (100%) et de 96,42% de thymine avec 3,58% de guanine principalement sous forme d'indels ou de substitution de T > G pour *NR-24*.

Tableau XIII : Analyse des différences inter-populations et intra population pour le NR-21 par AMOVA

Source de variation	Somme des carrés	Variance	Pourcentage de variation
Entre la population	3,453	0,0603	8,4786
Selon la population	59,213	0,6507	91,5214
Total	62,667	0,7109	
<b><math>F_{ST}</math> par loci : 0,0848 ; <math>p</math> - value = 0,0019</b>			
Composition nucléotidique : T : 100% G : 0% C : 0% A : 0%			
<b><math>F_{ST}</math> spécifique à la population</b>			
Nigeria	0,0709		
Sénégal	0,0305		
$p$ - value = 0,1256			

Tableau XIV : Analyse des différences inter-populations et intra population pour le NR-24 par AMOVA

Source of variation	Somme des carrés	Variance	Pourcentage of variation
Entre population	9,033	0,1762	11,4015
Selon la population	116,381	1,3692	88,5986
Total	125,414	1,5454	
<b><math>F_{ST}</math> par loci : 0,11401 ; <math>p</math> - value = 0,0000</b>			
Composition nucleotide : T : 96,42% G : 3,58% C : 0% A : 0%			
<b><math>F_{ST}</math> spécifique à la population</b>			
Nigeria	0,0458		
Sénégal	0,0813		
$p$ - value = 0,0811			

Le paramètre de mutation estimé  $\theta$  ( $\theta$ ) calculé pour le nombre de sites de ségrégation ( $\theta_s$ ) et le nombre moyen de différences de nucléotides ( $\theta_\pi$ ) de la comparaison par paires montre plus de diversités pour NR-24 que pour NR-21.

NR-21 avait des estimations zero des deux paramètres pour le Nigéria alors que pour le Sénégal les paramètres étaient zero pour  $\theta_s$  et  $\theta_\pi$  (0,1258 + 0,2178). Comme le montre le tableau XV, ces paramètres sont plus diversifiés au Nigéria qu'au Sénégal pour le NR-24.

**Tableau XV : Test de diversité moléculaire**

Marqueurs		<i>NR-21</i>		<i>NR-24</i>	
<b>Paramètres</b>	<b>Pays</b>	Sénégal	Nigéria	Sénégal	Nigéria
Theta (S)		0,0000	0,0000	0,2311	0,6897
S.D. Theta (S)		0,0000	0,0000	0,2311	0,4285
Theta ( $\pi$ )		0,1258	0,0000	0,0465	0,5074
S.D. Theta ( $\pi$ )		0,2178	0,0000	0,1294	0,4868
Theta (k)		0,8542	2,5141	8,7749	5,0899
95% C.I.		0.2923 ; 2,2498	1.1331, 5.2626	4,6659 ; 16,1789	2,5580 ; 9,7821

**III.2.1.4. Test de l'hypothèse des mutations neutres**

L'observation du TD était égale à 0 alors que la valeur p était égale à 1 ; par conséquent, nous ne rejetons pas l'hypothèse nulle selon laquelle tout le polymorphisme de délétion/insertion observé est maintenu sans sélection pour les deux pays au niveau de 5%. En outre, l'hypothèse de mutations neutres peut expliquer le polymorphisme de l'ADN pour le marqueur *NR-21*. D'autre part, une valeur négative a été obtenue pour le Fs de Fu au Nigéria et était significativement différent de 0, probablement en raison davantage de signes de croissance tumorale ou d'auto-stop dans cette population qu'au Sénégal où le résultat était positif et non statistiquement différent de 0. Ceci est montré sur le tableau.

Le test de neutralité pour le *NR-24* est également illustré au tableau XVI : des valeurs négatives ont été obtenues pour les TD et les FF dont seul FF était significativement différent de 0. Si le TD est pris en soi, cela impliquera que l'hypothèse de mutation neutre peut expliquer le polymorphisme des séquences obtenues pour ce marqueur. Néanmoins, nous suggérerions davantage de croissance tumorale ou de hitchhiking que de sélection pour ce gène (Fu, 1997), étant donné que seul le facteur FF est significativement différent de 0.

**Tableau XVI : Tests de neutralité selective**

Marqueurs		<i>NR-21</i>		<i>NR-24</i>	
<b>Paramètres</b>	<b>Pays</b>	Sénégal	Nigéria	Sénégal	Nigéria
D de Tajima		0,0000	0,0000	-1,1175	- 0,5627
<i>p – value</i>		1,0000	1,0000	0,1290	0,3130
Fs de Fu		-14,6989	0,3402	-0,3403	-14,6505
<i>p – value</i>		0,0000	1,0000	0,0000	0,0000

**Tableau XVII : Analyse des mismatch distribution pour l'expansion génétique**

Marqueurs		<i>NR-21</i>		<i>NR-24</i>	
<b>Paramètres</b>	<b>Pays</b>	Sénégal	Nigéria	Sénégal	Nigéria
	SSD	0.0021	0.0000	0.0000	0.0413
	<i>p - value</i>	0.1700	0.0000	0.1510	0.0160
	Raggedness	0.7643	0.0000	0.8248	0.7476
	<i>p - value</i>	0.7200	0.0000	0.8950	0.7010

### III.2.1.2 : Les relations entre les marqueurs étudiés (*NR-21* et *NR-24*) et les caractéristiques clinico-pathologiques

Les résultats concernant la relation entre le *NR-21* et les caractéristiques clinico-pathologiques des patients du Nigéria et du Sénégal sont représentés au tableau XVIII. D'après les résultats, plus de la moitié des patientes (65%) du Sénégal âgés de moins de 50 ans présentaient une instabilité pour le *NR-21* contre 41% des Nigériennes de la même catégorie d'âge. Les patientes présentant une tumeur du stade 3 étaient plus nombreuses que celles diagnostiquées aux autres stades pour les deux pays et étaient instables pour *NR-21* (50% pour le Nigeria et 43% pour le Sénégal). Plus de patientes avec une tumeur de grade intermédiaire (grade 3) dans les deux pays étaient instables pour *NR-21* (44% pour le Nigeria et 57% pour le Sénégal) que de patientes avec tumeurs bien ou mal différenciées. Alors que toutes les patientes des deux pays sauf une du Sénégal avaient un carcinome canalaire invasif (IDC), 84% et 87,5% des Nigériennes et des Sénégalaises respectivement étaient instables pour le *NR-21*.

Le tableau XIX présente les résultats des paramètres cliniques en association avec le *NR-24*. La plupart des patients atteints de tumeurs instables étaient également âgés de moins de 50 ans, avec la fréquence la plus élevée au Sénégal (41,3%). Les patientes avec une tumeur de stade 3 ainsi que ceux de grade 2 avaient plus de tumeurs instables que toutes les patientes ayant un CDI dans les deux pays.

Tableau XVIII : Relations entre les paramètres cliniques et le phénotype MSI pour *NR-21* chez les patientes atteintes d'un cancer du sein au Nigéria et au Sénégal

		Nigéria		Sénégal	
Paramètres cliniques		MSS	MSI	MSS	MSI
Age(années) :	< 50	5	19	3	27
	≥ 50	2	20	2	9
Stade	1	0	0	0	3
	2	0	5	1	3
	3	2	21	3	16
	4	4	10	1	10
Grade SBR	1	1	8	0	4
	2	4	20	2	19
	3	2	10	3	5
Histologie	IDC	7	39	5	35
	Autres	0	0	0	1

Tableau XIX : Relations entre les paramètres cliniques et le phénotype MSI pour *NR-24* chez les patientes atteintes d'un cancer du sein au Nigéria et au Sénégal

		Nigéria		Sénégal	
Paramètres cliniques		MSS	MSI	MSS	MSI
Age (années)	< 50	11	16	4	26
	≥ 50	2	15	4	7
Stade	1	0	0	1	2
	2	2	3	0	4
	3	6	18	3	16
	4	2	9	4	7
Grade SBR	1	3	6	2	2
	2	9	16	2	18
	3	1	8	4	14
Histologie	IDC	13	31	8	32
	Autres	0	0	0	0

Le tableau XX montre la proportion d'individus présentant une instabilité microsatellite, des tumeurs hautes et basses pour les deux marqueurs des deux pays par rapport aux patients stables pour les deux marqueurs et leur corrélation avec les paramètres cliniques. Ce résultat est similaire aux résultats présentés individuellement pour chaque marqueur avec plus de patientes de chaque pays ayant des tumeurs MSI-H présentant également, au stade 3, des tumeurs modérément différenciées en termes d'âge au moment du diagnostic. Pour ce dernier, alors que le résultat reste le même pour le Sénégal avec 53% des patientes atteintes de tumeur MSI-H ayant moins de 50 ans, l'inverse a été le cas pour les patients du Nigéria avec 30% des

patientes présentant une tumeur MSI-H étant de moins de 50 ans et 28% ayant eu une tumeur MSI-L étaient âgés de 50 ans ou plus.

Parmi tous les résultats, seul l'âge au diagnostic (valeur  $p = 0,05$ ) était associé de manière significative au *NR-24* chez les Nigérianes, tandis que le grade SBR ( $p - value = 0,041$ ) était significatif chez les Sénégalaises (données non présentées).

**Tableau XX : Relations entre les paramètres cliniques et le phénotype MSI pour les tumeurs MSS, MSI-L et MSI-H chez les patientes atteintes du cancer du sein au Nigéria et au Sénégal.**

		Nigéria			Sénégal		
Paramètres cliniques		MSS	MSI-L	MSI-H	MSS	MSI-L	MSI-H
Age (années)	< 50	1	13	10	1	7	22
	≥ 50	1	7	14	1	4	6
Stade	1	0	0	0	0	1	2
	2	0	3	2	0	1	3
	3	1	6	16	1	6	12
	4	1	7	6	1	3	7
Niveau	1	1	3	5	0	2	2
	2	1	11	12	0	5	16
	3	0	6	6	2	4	2
Histologie	IDC	2	20	24	2	11	27
	Autres	0	0	0	0	0	1

### III.2.2 : DISCUSSION

Le cancer du sein, qui a maintenant dépassé le cancer du col de l'utérus dans la plupart des pays, constitue un défi de plus en plus préoccupant pour la santé publique en Afrique subsaharienne, étant responsable d'un diagnostic de cancer sur quatre et d'un décès sur cinq chez les femmes (Akarolo-Anthony *et al.*, 2010 ; Sylla et Wild, 2012 ; Ferlay *et al.*, 2018).

Les études sur le cancer du sein en Afrique offrent des possibilités uniques d'évaluation de nouveaux facteurs de risque pouvant jouer un rôle important dans l'identification des loci susceptibles une maladie, compte tenu de la fraction de la diversité génétique en Afrique (Campbell et Tishkoff, 2008). L'instabilité chromosomique et microsatellitaire sont deux formes distinctes d'instabilités génomiques qui ont été reconnues dans les cancers humains. Les défauts dans le système de réparation non apparié de l'ADN qui maintient l'intégrité des

génomés, ont été associés aux mutations des gènes de réparation de l'ADN et peuvent augmenter jusqu'à 700 fois dans les cellules de mammifères. (Kolodner et Marsischky, 1999).

Les microsatellites ont été largement utilisés dans les études d'association en raison de leur grand pouvoir discriminant, leur association à des maladies humaines ainsi que leur utilisation pratique dans les études de génétique des populations (Mahadevan *et al.*, 1992 ; Stallings, 1994 ; Wright et Bentzen, 1994 ; Schlotterer, 2000 ; O'Donnell et Warren, 2002). De plus, il a été rapporté que dans les tumeurs avec le phénotype MSI-H, l'instabilité des marqueurs de répétitions mononucléotidiques longs et non codants, telle que présentée dans *BAT-25*, *BAT 26*, *NR-21*, *NR-24* et *NR-27* sont progressifs et indépendants (Buhard *et al.*, 2014). On pensait autrefois que les microsatellites étaient sélectivement neutres et non affectés par des pressions sélectives. On a découvert qu'ils étaient associés à des tumeurs humaines telles que la HNPCC.

Dans ce chapitre, nous avons étudié le polymorphisme du phénotype MSI de *NR-21* et *NR-24* entre des patientes atteintes du cancer du sein au Nigéria et au Sénégal. Nous avons émis l'hypothèse qu'il n'y avait pas de différences dans le processus évolutif du développement de la tumeur entre les femmes Nigériennes Sénégalaises. Les deux marqueurs microsatellites utilisés dans cette étude font partie d'un panel de cinq mononucléotides quasi-monomorphes du panel de Bethesda, qui a été réévalué en décembre 2002 lors de la réunion de consensus internationale. Le dinucléotidique répété qui dans le panel précédent été remplacés par trois répétitions mononucléotidiques (Umar *et al.*, 2004b). Buhard et ses collègues en 2014 ont proposé l'application du panel à la population non blanche en raison de la nature quasi monomorphe des marqueurs et ils les ont analysés avec l'ADN de cinquante-cinq populations à l'échelle mondiale.

Dans leur étude, 44,8% des individus en Afrique (y compris le Nigéria et le Sénégal) ont montré peu ou pas de polymorphisme avec une variation de 0,1% pour le *NR-24*, qui était également le marqueur le plus stable. En outre, la population sénégalaise était le seul pays africain qui ne présentait pas de fréquence élevée d'allèles variantes tandis que la population nigérienne présentait un polymorphisme pour seulement *BAT-26* et *NR-27*. La population nigérienne ne montre pas de fréquence élevée d'allèles variantes. Bien que le *NR-21* ait montré un polymorphisme dans d'autres populations, ceci était absent pour l'Afrique. Par conséquent, parce que le *NR-24* ne peut être utilisé seul pour l'évaluation de MSI car il n'était pas à 100% monomorphe dans cette étude, nous avons décidé de l'utiliser avec *NR-21*.

Lorsqu'un locus microsatellite subit une mutation, il gagne ou perd une répétition. Cela implique que deux allèles ne différant que par un motif sont plus liés (c'est-à-dire un ancêtre commun plus récent) que des allèles différant de plusieurs répétitions. Dans cette étude, nous avons génotypé deux marqueurs microsatellites sur un total de quatre-vingt-treize patientes atteintes de cancer du sein (46 Nigériens et 47 Sénégalais). Au total, les pourcentages d'allèles variantes pour chaque marqueur étaient de 85% pour le *NR-21* (15% pour chaque population) et le *NR-24* (50,57% pour le Nigeria et 49,43% pour le Sénégal) (tableau IX et X).

Un total de 8 haplotypes a été généré pour le *NR-21*, le Sénégal ayant le plus grand nombre d'haplotypes (8 haplotypes) et le Nigeria (4 haplotypes). L'analyse d'AMOVA avec une valeur de  $F_{ST}$  significative a suggéré que la majeure partie de la variation était aux différences au sein des populations (91,52%), tandis que la différenciation entre les populations représentait 8,48% du total (Tableau XIII). La faible différenciation génétique observée entre les pays conforte l'opinion selon laquelle l'évolution intra-tumorale (selon la population) est dictée par la sélection darwinienne. Par conséquent, l'évolution du cancer du sein au Nigeria et au Sénégal suit le mode d'évolution darwinien (Sottoriva *et al.*, 2015 ; Tao *et al.*, 2015) et n'est pas différent entre les deux pays. Ce processus de diversité génétique inclut la mutation, la dérive génétique, la migration et la sélection naturelle et parce que les tumeurs sont considérées comme une population de cellules asexuées, on pense qu'elles suivent le schéma d'évolution naturelle et la structure de population de la sélection positive (Nowell, 1976 ; Wen-Hsiung, 1997 ; Hartl et Clark, 2006). ; Maley *et al.*, 2006 ; Merlo *et al.*, 2006 ; Ewens, 2010). Cela est également vrai pour le *NR-24* (tableau XIV) avec une valeur de  $F_{ST}$  statistiquement significative avec un niveau de variation élevé dans chaque pays (88,59%) et faible entre les deux pays (11,40%). De même, le niveau de diversité et la différenciation tumorale affectent également la réponse des cellules tumorales aux perturbations environnementales naturelles ou médicales (Maley *et al.*, 2006 ; Merlo *et al.*, 2006 ; Marusyk et Polyak, 2010). Néanmoins, le résultat obtenu ici est en contradiction avec le résultat de Ling *et al.* (2016), qui discutent de la petite taille de l'échantillon utilisé pour l'analyse de l'hétérogénéité intra-tumorale conduisant à un faible pouvoir de rejeter le modèle non darwinien (Tao *et al.*, 2011 ; Gerlinger *et al.*, 2012 ; Sottoriva *et al.*, 2013 ; de Bruin *et al.*, 2014 ; Zhang *et al.*, 2014 ; Wang *et al.*, 2014). Ling et ses collègues ont évalué des échantillons multirégionaux d'une seule tumeur à l'aide du séquençage du génome entier (WGS) et ont observé une diversité génétique élevée, conforme au modèle non darwinien de l'évolution de la tumeur, sans aucune preuve de sélection positive (Ling *et al.*, 2016).

Un modèle de MSI indique que la perte et le gain en nucléotides résultent d'un glissement dans la matrice et dans les brins d'amorce nouvellement synthétisés (Sagher, 1999). La grande diversité haplotypique et la faible diversité de nucléotides pour les deux marqueurs suggère une expansion récente de la population après une période de faible taille effective de la population. Cela montre que le temps qui suit l'expansion de la population est suffisamment long pour opérer le changement d'haplotypes résultant de la mutation, mais il n'est pas assez long pour accumuler de grandes différences parmi les séquences (Avisé, 2000). Ce n'est pas rare pour le cancer du sein, en particulier en Afrique subsaharienne où les patientes se présentent plus jeunes avec une tumeur agressive et à un stade tardif. Aussi dans cette étude, comme similaires à d'autres, la plupart des patientes sont en pré ménopause et sont donc toujours au stade rapide de la division cellulaire et de la maladie active. Des études ont démontré des niveaux plus élevés de diversité de nucléotides et d'haplotypes, à la fois dans les génomes nucléaires et mitochondriaux de la population africaine ainsi que davantage d'allèles et d'haplotypes privés par rapport aux non-Africains qui semblent posséder un sous-ensemble de la diversité génétique de l'Afrique subsaharienne (Tishkoff *et al.*, 1996 ; Tishkoff et Williams, 2002 ; Tishkoff et Verrelli, 2003 ; Tishkoff et Kidd, 2004 ; Épouse. Int. HapMap, 2005 ; Ramachandran *et al.*, 2005 Conrad *et al.*, 2006 ; Zhao *et al.*, 2006 ; Garrigan *et al.*, 2007 ; Jakobsson *et al.*, 2008 ; Li *et al.*, 2008).

Pour *NR-24*, les valeurs des tests de neutralité pour les  $F_s$  de  $F_u$  et le  $D$  de Tajima sont négatives, ce qui indique la sensibilité à l'expansion de la population telle qu'établie précédemment ainsi qu'aux effets d'embouteillage (Tajima, 1993 ; 1996 ; 1996 ; Arris-Brosou et Excoffier, 1996). Le test  $F_s$  de  $F_u$  est également sensible à la croissance démographique récente et applicable aux résultats des statistiques du test de *NR-21* pour cette étude (Fu, 1997).

Le cancer du sein est le cancer le plus répandu chez les femmes dans le monde. Il est en augmentation dans les pays en développement où la plupart des cas sont diagnostiqués à un stade avancé (Agarwal *et al.*, 2007 ; Bray *et al.*, 2018). Il est, à l'instar des autres cancers sporadiques, un processus en plusieurs étapes au cours duquel des mutations dans les gènes de réparation incompatibles entraînent une augmentation des taux de variation génétique des cellules. La question de l'effet de la mutation dans les gènes de réparation incompatibles pouvant conduire à une instabilité génomique telle que l'instabilité des microsatellites a particulièrement préoccupé les oncologues, en particulier à la lumière de l'immunothérapie pour le traitement des tumeurs sporadiques documentées pour le cancer colorectal. De plus, la

thérapie ciblée étant généralement transitoire et souvent suivie d'une rechute inévitable, l'identification de marqueurs pronostiques et prédictifs afin de mieux sélectionner les patients pour le traitement est nécessaire et la distinction entre tumeur haute MSI revêt une importance clinique.

Ozer et ses collègues en 2002 ont indiqué que le choix des marqueurs microsatellites pouvait expliquer plusieurs divergences dans les études concernant le statut MSI chez les patientes atteintes de cancer du sein et ont donc suggéré que le choix des marqueurs informatifs soit impératif pour l'utilisation du MSI dans le diagnostic du cancer du sein. En raison de la variabilité différentielle des marqueurs mononucléotidiques d'une population à l'autre, notre objectif était de comparer les caractéristiques cliniques et pathologiques des patientes atteintes d'un cancer du sein au Nigéria et au Sénégal avec les marqueurs *NR-21* et *NR-24*. Ces deux marqueurs font partie du panel de mono nucléotides quasi-monomorphes qui ont été substitués aux marqueurs di nucléotides et ont été décrits par Buhard et ses collègues (Umar *et al.*, 2004a ; Buhard *et al.*, 2006).

Les résultats de cette étude montrent des MSI parmi les patientes atteintes d'un cancer du sein au Nigéria et au Sénégal, comme indiqué au chapitre II. Les tumeurs MSI-H ont des caractéristiques cliniques distinctes ; par conséquent, certaines de ces caractéristiques ont été évaluées dans notre étude. Les caractéristiques clinico-pathologiques des patientes de cette étude étaient les mêmes pour les deux pays et pour les deux marqueurs, sauf pour l'âge où une différence significative ( $p - value = 0,005$ ) a été observée pour le marqueur *NR-24* entre les phénotypes MSS et MSI chez les patientes nigériennes ce qui n'était pas le cas pour le groupe sénégalais ( $p - value = 0,178$ ) ; et le grade SBR qui était significativement différent ( $p - value = 0,041$ ) chez les patientes sénégalaises.

L'âge moyen au diagnostic chez les patientes incluses dans l'analyse du phénotype d'instabilité des microsatellites était de 49 ans pour les patientes Nigériennes et de 44 ans pour les patientes Sénégalaises. Ce n'était pas le cas avec l'étude MSI dans le cancer du sein à début précoce où aucune différence n'a été observée pour l'âge dans la population étudiée (Ozer *et al.*, 2002). Néanmoins, des caractéristiques cliniques distinctes ont été observées chez les patientes atteintes d'un cancer du sein atteintes de tumeurs instables microsatellites (Thibodeau *et al.*, 1993). En outre, le plus jeune âge observé de la plupart des patientes dans cette étude pourrait expliquer pourquoi la mammographie sera moins précise pour ce groupe de patientes car la densité élevée du tissu mammaire au cours de cet âge peut réduire la sensibilité et la détection des microcalcifications (Foxcroft *et al.*, 2004). Ce résultat peut également refléter la population plus jeune des Africains subsahariens, qui résulte d'une fertilité plus élevée et d'une espérance

de vie plus courte, ainsi que de la plus faible prévalence des facteurs de risque de cancer du sein parmi la population plus âgée. L'âge moyen du diagnostic du cancer du sein entre 40 et 50 ans au moment du diagnostic a été observé dans plusieurs études menées dans le monde (Vaziri *et al.*, 2001 ; Brose *et al.*, 2002 ; Huo *et al.*, 2008 ; Huo *et al.*, 2009 ; Mavaddat *et al.*, 2012 ; Tannouri *et al.*, 2018).

La corrélation entre le carcinome canalaire invasif et le MSI était élevée mais non significative entre les deux pays, car il s'agissait du type histologique le plus répandu dans cette étude pour tous les patients sauf un. Il s'agit d'un type histologique courant observé dans la plupart des études au Nigéria et au Sénégal (Udoakang *et al.*, données non publiées ; Huo *et al.*, 2009). De plus, les tumeurs de grade intermédiaire étaient plus prévalentes chez les individus porteurs de marqueurs instables. Parmi les patientes atteintes de tumeurs MSI-H, 51% des cas du Sénégal et 26% du Nigéria avaient des tumeurs de grade intermédiaire (grade 2). Dans une étude ghanéenne également, 60,8% des patients ont présenté une tumeur de grade 2 comme étant la tumeur la plus élevée (Seshie *et al.*, 2015).

Dans une analyse multi-population (incluant le Nigéria et le Sénégal) des cinq mononucléotides, le NR-24 avait une fourchette de taille plus stable de 0,1% dans la population étudiée et en Afrique subsaharienne, le NR-21 et le NR-24 n'ont montré aucun allèle variant approchant les 10% (Buhard *et al.*, 2006). En outre, ils n'ont pas observé d'allèles variantes pour ces deux marqueurs chez les populations nigériane et sénégalaise. Il existe plusieurs divergences concernant l'importance pronostique de la MSI dans le cancer du sein. Certaines études ont rapporté de mauvais pronostics de cancer du sein pour les tumeurs MSI-H (Contegiacomo *et al.*, 1995 ; Paulson *et al.*, 1996; Saurvinos *et al.*, 1997) et d'autres n'en trouvant aucune corrélation entre le phénotype MSI et le cancer du sein (Yee *et al.*, 1994 ; Aldaz *et al.*, 1995 ; Toyama *et al.*, 1996 ; DeMarchis, *et al.*, 1997 ; Sourvinos *et al.*, 1997 ; Rush *et al.*, 1997 ; Anbazhagan *et al.*, 1999 ; Benachenhou *et al.*, 1999; Tomita *et al.*, 1999 ; Ando *et al.*, 2000 ; Bock *et al.*, 2000 ; Tokunaga *et al.*, 2000 ; Laqmani *et al.*, 2003 ; Chapgar *et al.*, 2004).

En outre, la plupart des études réalisées avec le panel Bethesda de mono nucléotides quasi-monomorphes ont été réalisées avec les *BAT-25* et *BAT-26*, qui ont été identifiées comme étant les marqueurs les plus connus de l'évaluation MSI-H bien que polymorphes parmi les populations afro-américaines et africaines y compris le Nigéria (Pyatt *et al.*, 1999 ; Buhard *et al.*, 2008). Par conséquent, la disparité entre les résultats des MSI dans le cancer du sein peut être liée aux marqueurs utilisés pour chaque étude, qui vont des mono nucléotides aux répétitions de tri nucléotides et à la taille des échantillons utilisés pour chaque étude. Il y a eu

des erreurs de classification des tumeurs MSI-H lors de l'utilisation du panel de référence Bethesda incluant les deux marqueurs di nucléotides, ainsi que lors de l'utilisation de répétitions dinucléotides pour l'évaluation des tumeurs MSI-H, car elles sont hautement polymorphes et nécessitent une analyse de l'ADN germinale correspondant qui en résulte d'une analyse coûteuse, prenant du temps et susceptible d'introduire des erreurs dues à la confusion dans l'échantillon.

Le MSI dans le cancer du sein a été associé à une agressivité de la maladie et à une récurrence prédictive particulièrement précoce, comme observé chez la population jeune de notre étude, mais son apparition tardive a été moins agressive avec un taux de croissance plus lent (Paulso *et al.*, 1996 ; De Marchis *et al.*, 1997 ; Pizzi *et al.*, 2002). Un rôle mineur des gènes de réparation incompatibles a été suggéré dans l'oncogenèse du cancer du sein en raison des résultats de MSI observés dans certaines études (Lacave et Larsen, 2005). Par conséquent, étant donné le haut niveau d'instabilité des deux marqueurs étudiés par rapport au stade avancé qui prédit l'implication des ganglions lymphatiques et une plus grande taille de la tumeur, nous suggérons que *NR-21* et *NR-24* puissent être des marqueurs d'intérêt pour l'identification des tumeurs MSI, notant la forme agressive de la maladie, en particulier chez les jeunes de moins de 50 ans au Nigéria et au Sénégal. De plus, l'association de l'instabilité des deux marqueurs dans la même tumeur montre que la mutation dans les gènes de réparation de mésappariement pourrait entraîner l'apparition et la progression du cancer du sein chez les patientes de cette région. Les résultats obtenus ne montrent aucune différence entre le phénotype MSI au Nigéria et au Sénégal, ce qui soutient l'analyse moléculaire selon laquelle le cancer du sein dans les deux pays présente un schéma évolutif similaire.

## **CONCLUSION PARTIELLE**

À partir des résultats que nous avons obtenus dans ce chapitre, nous pouvons donc confirmer que le schéma de développement du cancer du sein au Nigéria et au Sénégal est identique et que les deux marqueurs utilisés pour cette étude suivent le même pattern de l'instabilité. Des études sur le continent pourraient également fournir des informations supplémentaires sur le rôle des conditions des immunodéprimées chez les patients cancéreux, un domaine d'intérêt croissant dans la biologie du cancer, étant donné le rôle des marqueurs inflammatoires dans l'étiologie du cancer. Les processus démographiques peuvent également causer des biais similaires dans la fréquence des polymorphismes. Par exemple, lorsque la taille de la population augmente rapidement, la dérive génétique a moins d'effet sur une population en expansion rapide, ce qui entraîne un excès de polymorphismes rares (imitant le schéma observé lors d'une sélection positive ou purifiante). Des études comparatives de populations humaines diverses sur le plan ethnique, en particulier en Afrique, sont importants pour reconstruire l'histoire de l'évolution humaine et pour comprendre les bases génétiques de l'adaptation phénotypique et d'une maladie complexe.

Une analyse plus approfondie est nécessaire pour mieux comprendre la tumorigenèse de MSI et pour obtenir une meilleure compréhension de la pathogenèse, tout en développant de nouvelles approches thérapeutiques pour cibler ces altérations pathogéniques essentielles.

En caractérisant la variation génétique parmi les individus et les populations, nous pouvons mieux comprendre la susceptibilité différentielle à la maladie, la réponse différentielle aux agents pharmacologiques, l'historique de l'évolution humaine et l'interaction complexe des facteurs génétiques et environnementaux dans la production de phénotypes. L'Afrique est une région importante pour l'étude de la diversité génétique humaine en raison de l'histoire complexe de sa population et des variations dramatiques du climat, du régime alimentaire et de l'exposition aux maladies infectieuses, qui entraînent des niveaux élevés de variation génétique et phénotypique au sein des populations africaines.

La nécessité d'une méthode simple, sensible et spécifique pour la détermination de la MSI dans le cancer du sein devient de plus en plus pertinente pour les cliniciens en raison de ses implications génétiques, pronostiques et diagnostiques. Comme indiqué précédemment, l'évaluation du statut MSI peut aider à établir un pronostic clinique, à prédire la réponse tumorale à la chimiothérapie et à servir de marqueur pronostique pour la détection du cancer du sein (Bacher *et al.*, 2004 ; Yamamoto *et al.*, 2006). Il peut également servir de cible thérapeutique pour le diagnostic précoce afin de prédire l'agressivité de la maladie et de

sélectionner un traitement pour chaque patient, ainsi que de prédire leur réponse au traitement. L'immunothérapie est une nouvelle frontière de traitement qui est explorée pour le traitement du cancer du sein, en particulier des tumeurs triples négatives. Etant donné que le cancer du sein n'est pas intrinsèquement immunogène, la plupart des patientes ne bénéficient pas de ce traitement. Il est donc impératif d'identifier des biomarqueurs pour sélectionner les patientes pour le traitement (Kok *et al.*, 2017 ; Hu et McArthur, 2018).

Bien que la fréquence des réparations incompatibles ait été controversée dans le cancer du sein, la plupart des études ont rapporté que 0 à 3% de personnes atteintes de MSI étaient présentes (Dudley *et al.*, 2016). Néanmoins, la présence de MSI dans d'autres populations vaut également la peine de faire un pas en avant dans l'application du test de MSI à la population de patients susceptibles de bénéficier de son application, en particulier pour l'application de la thérapie combinatoire. Les approches diagnostiques ont évolué depuis le début des années 90, passant exclusivement de critères cliniques à l'intégration de caractéristiques pathologiques, MSI joue ici un double rôle de bio marqueur prédictif et pronostique. Les efforts de recherche doivent également viser à explorer son application clinique au Nigéria, au Sénégal et par extension en Afrique où la plupart des patientes présentent d'un cancer du sein triple négatif (TNBC).

## CONCLUSION GÉNÉRALE ET RECOMMANDATIONS

Il existe de fortes preuves d'instabilité génomique résultant de la fréquence élevée d'anomalies chromosomiques et de mutations dans les cancers humains. Cette instabilité est caractérisée par une succession de mutations dans la tumeur résultant de l'augmentation continue de mutations endogènes dans les cellules. Un dysfonctionnement du système de réparation de l'ADN, entraînant l'instabilité des microsatellites, a été identifié comme un mécanisme par lequel la tumorigenèse se produit. Cette étude a évalué le phénotype MSI de deux marqueurs mononucléotidiques quasimonomorphes : *NR-21* et *NR-24* ; comparé et caractérisé leur évolution au sein des patientes atteintes du cancer du sein au Nigéria et au Sénégal. À titre préliminaire, nous avons évalué la tendance épidémiologique de la maladie dans les deux pays afin d'évaluer leurs similitudes ou leurs différences dans chaque population. Le résultat obtenu indique de légères différences entre plusieurs facteurs de risque liés à la reproduction et aux facteurs hormonaux (âge à la ménarche, âge à la première grossesse à terme, parité et utilisation de contraceptifs) et à l'état matrimonial. Malgré les différences apparentes, les tendances observées basculent dans la même direction, montrant probablement des différences spécifiques à la population.

En outre, l'analyse des paramètres cliniques en association avec le MSI dans les deux pays n'était pas statistiquement significative, mais montrait clairement l'évolution de l'effet du système de réparation inadéquat sur l'étiologie du cancer du sein dans notre population étudiée. Cela suggère que les marqueurs (*NR-21* et *NR-24*) peuvent être utilisés comme outils de diagnostic pour la prédiction de la réponse tumorale au traitement dans cette population. Avec les plus hauts niveaux de diversité génétique trouvés en Afrique, on pourrait prédire que les Africains auront des quantités relativement élevées d'adaptation locale au sein des populations en raison du large éventail de différences environnementales ainsi que de la forte exposition aux maladies infectieuses. Néanmoins, les signatures de sélection dans les populations africaines peuvent être omises du fait d'études menées principalement sur une ou deux populations africaines. Par conséquent, pour mieux comprendre les adaptations génétiques et phénotypiques en Afrique, il est important de rechercher les signatures génétiques de sélection dans un large éventail de populations africaines ethniquement diverses, vivant dans des environnements distincts et d'identifier clairement les variantes fonctionnelles susceptibles d'être la cible de sélection afin de vérifier leur impact sur les variations phénotypiques telles que le cancer du sein.

En conclusion, les facteurs génétiques sous-jacents à des maladies complexes sont encore mal compris, en particulier sous l'angle des interactions génétiques et environnementales, ainsi que des interactions épistatiques entre les loci susceptibles d'influencer la susceptibilité aux maladies complexes et avec seulement une petite fraction des études effectuées en Afrique, une analyse approfondie des variations génétiques doit être effectuée si nous voulons répondre à certaines des questions scientifiques sur la susceptibilité au cancer du sein, l'hétérogénéité, les relances de traitement et la survie. Bien que les études sur les populations ouest-africaines soient utiles pour identifier les marqueurs, d'autres études conduites sur le continent, notamment dans les populations est-africaines, permettront de tester des modèles d'origine et des modèles de dispersion d'hommes modernes dehors de l'Afrique. Enfin, nous pouvons conclure de cette étude qu'il n'existe aucune différence entre le modèle évolutif et le potentiel de cancérogenèse du cancer du sein au Nigéria et au Sénégal. Par conséquent, le modèle de traitement et les traitements mis au point pour cibler une population peuvent servir à l'autre. Néanmoins, nous suggérons de mener une vaste cohorte d'analyses cas-témoins afin de dissiper la notion de polymorphisme de *NR-21* et *NR-24* dans les deux pays.

## REFERENCES BIBLIOGRAPHIQUES

### A

1. **Abdel-Rahman, W. M., Ollikainen, M., Kariola, R., Jarvinen, H. J., Mecklin, J. P., Nystrom-Lahti, M., Knuutila, S. et Peltomaki, P. 2005.** Comprehensive characterization of HNPCC-related colorectal cancers reveals striking molecular features in families with no germline mismatch repair gene mutations. *Oncogene*. 24: 1542-1551
2. **Abdel-Rahman, W. 2008.** Genomic Instability and Carcinogenesis: An Update. *Current Genomics*. 9(8): 535-541
3. **Adebamowo, C. A. et Ajayi, O. O. 2000.** Breast cancer in Nigeria. *West African Journal of Medicine*. 19: 179-191
4. **Adebamowo, C., Hu, F., Cho, E., Spiegelman, D., Holmes, M., et Willett, W. 2005.** Dietary patterns and the risk of breast cancer. *Annals of Epidemiology*. 15(10): 789-795.
5. **Adlercreutz, H., Gorbach, S. L., Goldin, B. R., Woods, M. N., Dwyer, J. T. et Hamalainen, E. 1994.** Estrogen metabolism and excretion in Oriental and Caucasian women. *Journal of National Cancer Institute* 86:1076-1082.
6. **Agbo, P. S., Khalid, A. et Oboirien, M. 2014.** Clinical presentation, Prevalence and Management of Breast Cancer in Sokoto, Nigeria. *Journal of Women's Health Care*. 3:149
7. **Agrafioti, I. et Stumpf, M. P. 2007.** SNPSTR: a database of compound microsatellite-SNP markers *Nucleic Acids Research*, Vol. 35, Database issue D71–D75. doi:10.1093/nar/gkl806
8. **Akarolo-Anthony, S. N., Ogundiran, T. O. et Adebamowo, C. A. 2010.** Emerging breast cancer epidemic: evidence from Africa. *Breast Cancer Research*. 12(Suppl 4):58
9. **Akoum, R., Ghaoui, A., Brihi, E., Ghabash, M. et Hajjar, N. 2009.** Early-onset breast cancer in a Lebanese family with Lynch syndrome due to MSH2 gene mutation. *Hereditary Cancer Clinical Practice*. 7(1):10.
10. **Al-Ajmi, K., Lophatananon, A., Ollier, W. et Muir, K. R. 2018.** Risk of breast cancer in the UK biobank female cohort and its relationship to anthropometric and reproductive factors. *PLoS One*. 26;13(7): e0201097.
11. **Albrektsen, G., Heuch, I., Hansen, S., Kvåle, G. 2005.** Breast cancer risk by age at birth, time since birth and time intervals between births: Exploring interaction effects. *British Journal of Cancer*. 92:167-75.
12. **Aldaz, C. M., Chen, T., Sahin, A., Cunningham, J. et Bondy, M. 1995.** Comparative allelotype of in situ and invasive human breast cancer: high frequency of microsatellite instability in lobular breast cancer. *Cancer Research* 55: 3976-3981.
13. **Al-Sukhni, W., Aronson, M. et Gallinger, S. 2008.** Hereditary colorectal cancer syndromes: Familial adenomatous polyposis and lynch syndrome. *Surgical Clinic of North America*. 88(4): 819-844.
14. **Althuis, M. D., Fergenbaum, J. H., Garcia-Closas, M., Brinton, L. A., Madigan, M. P. et Sherman, M. E. 2004.** Etiology of hormone receptor-defined breast cancer: a systematic review of the literature. *Cancer Epidemiology Biomarkers & Prevention*. 13: 1558 –1568
15. **Althuis, M. D., Dozier, J. M., Anderson, W. F., Devesa, S. S. et Brinton, L. A. 2005.** Global trends in breast cancer incidence and mortality 1973-1997. *International Journal of Epidemiology*. 34:405-412.
16. **Ando, Y., Iwase, H., Ichihara, S., Toyoshima, S., Nakamura, T., Yamashita, H., Toyama, T., Omoto, Y., Karamatsu, S., Mitsuyama, S., Fujii, Y. et Kobayashi, S.**

2000. Loss of heterozygosity and microsatellite instability in ductal carcinoma in situ of the breast. *Cancer Letters* 156: 207-214.
17. Antonarakis, S. E. et the Nomenclature Working Group. 1998. *Recommendations for a nomenclature system for human gene mutations. Human Mutations.* 11: 1-3
  18. Ashktorab, H., Smoot, D. T., Carethers, J. M., Rahmanian, M., Kittles, R., Vosganian, G., Doura, M., Nidhiry, E., Naab, T., Momen, B., Shakhani, S. et Giardiello, F. M. 2003. High incidence of microsatellite instability in colorectal cancer from African Americans. *Clinical Cancer Research.* 9(3): 1112-1117.
  19. Ashktorab, H., Smoot, D. T., Farzanmehr, H., Fidelia-Lambert, M., Momen, B., Hylind, L., Iacozio-Dononue, C., Carethers, J., M., Goel, A., Boland, C. R. et Giardiello, F. M. 2005. Clinicopathological features and microsatellite instability (MSI) in colorectal cancers from African Americans. *International Journal of Cancer.*; 116(6): 914-919.
  20. Ashktorab, H., Brim, H., Al-Riyami, M. Date, A., Al-Mawaly, K., Kashoub, M., Al-Mjeni, R., Smoot, D. T., Al-Moundhri, M, Al-Hashemi, S., Ganguly, S. S. et Raeburn, S. 2008. Sporadic colon cancer Mismatch repair immunohistochemistry and microsatellite instability in Omani subjects. *Digestive Disease and Sciences.*; 53(10): 2723-2731.
  21. Ayub, S. G., Rasool, S., Ayub, T., Khan, S. N., Wani, K. A. et Andrabi, K. I. 2014. Mutational analysis of the BRCA2 gene in breast carcinoma patients of Kashmiri descent. *Molecular medicine reports.* 9(2):749–53.
  22. Azadeh, S., Moghimi-Dehkordi, B., Fatem, S. R., Pourhoseingholi, M. A., Ghiasi, S. et Zali, M. R. 2008. Colorectal cancer in Iran: An epidemiological study. *Asian Pacific Journal of Cancer Prevention*; 9(1): 123-126.

## B

23. Baca, S. C., Prandi, D., Lawrence, M. S., Mosquera, J. M., Romanel, A., Drier, Y., Park, K., Kitabayashi, N., MacDonald, T. Y., Ghandi, M., Van Allen, E., Kryukov, G. V., Sboner, A., Theurillat, J. P., Soong, T. D., Nickerson, E., Auclair, D., Tewati, A., Beltran, H., Onofrio, R. C., Boysen, G., Guiducci, C., Barbieri, C. E., Cibulskis, K., Sivachenko, A., Carter, S. L., Saksena, G., Voet, D., Ramos, A. H., Winckler, W., Cipicchio, M., Ardlie, K., Kantoff, P. W., Berger, M. F., Gabriel, S. B., Golub, T. R., Meyerson, M., Landers, E. S., Elemento, O., Getz, G., Demichelis, F., Rubin, M. A. et Garraway, L. A. 2013. Punctuated evolution of prostate cancer genomes. *Cell.* 153: 666–677.
24. Baranov, P. V., Hammer, A. W., Zhou, J., Gesteland, R. F. et Atkins, J. F. 2005. Transcriptional slippage in bacteria : distribution in sequenced genomes and utilization in IS element gene expression. *Genome Biology.* 6: R25
25. Bashashati, A., Ha, G., Tone, A., Ding, J., Prentice, L. M., Roth, A., Rosner, J., Shumansky, K., Kalloger, S., Senz, J., Yang, W., McConechy, M., Melnyk, N., Anglesio, M., Luk, M. T., Tse, K., Zeng, T., Moore, R., Zhao, Y., Marra, M. A., Gilks, B., Yip, S., Huntsman, D. G., McAlpine, J. N. et Shah, S. P. 2013. Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. *Journal of Pathology.* 231:21–34.
26. Baudhuin, L. M., Burgart, L. J., Leontovich, O. et Thibodeau, S. N. 2005. Use of microsatellite instability and immunohistochemistry testing for the identification of individuals at risk for Lynch syndrome. *Familial Cancer,* 4(3):255–265.
27. Bauer, K. R., Brown, M., Cress, R. D., Parise, C. A. et Caggiano, V. 2007. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2- negative invasive breast cancer, the so-call triple-negative

- phenotype: a population-based study from the California cancer registry. *Cancer*. 109:1721-1728
28. **Beca, F. et Polyak, K. 2016.** Novel Biomarkers in the Continuum of Breast Cancer. Vered Stearns (Eds.), *Advances in Experimental Medicine and Biology*, Vol. 882, 33-68, 2016
  29. **Beggs, A. D. et Hodgson, S. V. 2009.** Genomics and breast cancer: the different levels of inherited susceptibility. *European Journal of Human Genetics* 17: 855–856. PMID:19092772
  30. **Bellizzi, A. M. et Frankel, W. L. 2009.** Colorectal cancer due to deficiency in DNA mismatch repair function: *a review of Advances in Anatomic Pathology*. 16(6):405–417.
  31. **Benachenhou, N., Guiral, S., Gorska-Flipot, I., Labuda, D. et Sinnett, D. 1999.** Frequent loss of heterozygosity at the DNA mismatch-repair loci hMLH1 and hMSH3 in sporadic breast cancer. *Br J Cancer*. 79(7-8): 1012–1017
  32. **Bergfors, M. 2014.** Evaluation of Microsatellite Instability Analysis as a Diagnostic Tool to Identify Lynch Syndrome in Endometrial Cancer Patients
  33. **Bernier, M. O., Plu-Bureau, G., Bossard, N., Ayzac, L. Thalabard, J. C. 2000.** Breastfeeding and risk of breast cancer: A metaanalysis of published studies. *Human Reproduction Update*. 6: 374-86
  34. **Bernstein, L., et Ross, R. K. 1993.** Endogenous hormones and breast cancer risk. *Epidemiologic Reviews*. 15(1): 48–65.
  35. **Black, D. M. 1994** The genetics of breast cancer. *European Journal of Cancer*. 30A:1957–61.
  36. **Bleiker E. et van der Ploeg H. 1999.** Psychosocial factors in the etiology of breast cancer: review of a popular link *Patient Education and Counseling*. 37: 201-214. 10.1016/S0738-3991(99)00027-0.
  37. **Bock, N., Meden, H., Regenbrecht, M., Junemann, B., Wangerin, J. et Marx, D. 2000:** Expression of the mismatch repair protein hMSH2 in carcinoma in situ and invasive cancer of the breast. *Anticancer Research*. 20: 119-124.
  38. **Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W., Meltzer, S. J., Rodriguez-Bigas, M. A., Fodde, R., Ranzani, G. N. et Srivastava, S. 1998.** A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Research* 58(22): 5248–5257
  39. **Boland, C. R., et Goel, A. 2010.** Microsatellite instability in colorectal cancer. *Gastroenterology*, 138(6): 2073-2087.
  40. **Bonnomet, A., Brysse, A., Tachsidis, A., Waltham, M., Thompson, E., Polette M., et Gilles., C. 2010.** Epithelial to Mesenchymal Transitions and Circulating Tumor Cells. *Journal of Mammary Gland Biology Neoplasia*.
  41. **Bowcock, A. M., Linares, R. A., Tomfohrde, J., Minch, E., Kidd J. R., et Cavalli-Sforza, L. L. 1994.** High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 368: 455–457.
  42. **Bradlow, H. L., Hershcopf, J., Martucci, C. et Fishman, J. 1986.** 16-hydroxylation of estradiol: a possible risk marker for breast cancer. *Annals of the New York Academy of Sciences*. 464:138–51.
  43. **Bray, F. et Moller, B. 2006.** Predicting the future burden of cancer. *Nat Review of Cancer*.;6: 63-74
  44. **Bray, F. 2014.** Transitions in human development and the global cancer burden. In: Stewart B W, Wild CP, eds. *World Cancer Report Lyon: IARC Press; 2014:42-55*

45. **Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L. et Jemal, A. 2018.** **Global Cancer Statistics 2018:** GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*. 0:5-31
46. **Bronner, C. E., Baker, S. M., Morrison, P. T., Warren, G., Smith, L. G., Lescoe, M. K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., Tannergard, P., Bollag, R. J., Godwin, A. R., Ward, D. C., Nordenskjold, M., Fishel, R., Kolodner, R. et Liskay, M. 1994.** Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature*. 368: 258 – 261.
47. **Buecher, B., de Pauw, A., Freneaux, P. et Rouleau, E. 2011.** Instabilité des microsatellites et cancers colorectaux. *Cancéro digest*. 3(3): 200-203
48. **Buhard, O., Suraweera, N., Lectard, A., Duval, A. et Hamelin, R. 2004.** Quasimonomorphic mononucleotide repeats for high-level microsatellite instability analysis. *Dis Markers* 20(4-5): 251–257
49. **Buhard, O., Cattaneo, F., Wong, Y. F., Yim, S. F., Friedman, E., Flejou, J. F., Duval, A. et Hamelin, R. 2006.** Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumours. *Journal Clinic of Oncology* 24(2): 241–251

## C

50. **Calderon-Garciduenas, A. L., Ruiz-Flores, P., Cerda-Flores, R. M et Barrera-Saldana, H. A. 2005.** Clinical follow up of mexican women with early onset of breast cancer and mutations in the BRCA1 and BRCA2 genes. *Salud publica de Mexico*. 47(2):110–5.
51. **Campbell, M. C. et Tishkoff, S. A. 2008.** African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annual Review Genomics Human Genetics*.9:403–433. [PubMed: 18593304]
52. **Campbell, M. C. et Tishkoff, S. A. 2010.** The evolution of human genetic and phenotypic variation in Africa. *Current Biology*. R166–R173. [PubMed: 20178763]
53. **Campbell, P., Yachida, S., Mudie, L., Stephens, P., Pleasance, E., Stebbings, L., Morsberger, L., Latimer, C., McLaren, S., Lin, M., McBride, D., Varela, I., Nik-Zainal S. A., Leroy, C., Jia, M., Menzies, A., Butler, A. P., Teague, J. W., Griffin, C. A., Burton, J., Swerdlow, H., Quail, M. A., Stratton, M. R., Iacobuzio-Donahue, C. et Futreal, A. 2010.** The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature*. 467:1109–1113.
54. **Campbell, M., Hirbo, J., Townsend, J. et Tishkoff, S. 2014.** The peopling of the African continent and the diaspora into the new world. *Current Opinion in Genetics and Development*. 29: 120–132. doi: 10.1016/j.gde.2014.09.003.
55. **Cancer Genome Atlas Network. 2012.** Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 487:330 – 337
56. **Canfell, K., Banks, E., Moa, A. M. et Beral, V. 2008.** Decrease in breast cancer incidence following a rapid fall in use of hormone replacement therapy in Australia. *Medical Journal of Australia*. 188: 641-644.
57. **Carey, L. A., Perou, C. M., Livasy, C. A, Dressler, L. G., Cowan, D., Conway, K., Karaca, G., Troester, M. A., Tse, C. K. Edmiston, S., Deming, S. L., Geradts, J., Cheang, M. C., Nielsen, T. O., Moorman, P. G., Earp, H. S. et Millikan, R. C. 2006.** Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA*. 295: 2492-2502
58. **Chan, T.L., Curtis, L.C., Leung, S.Y., Farrington, S.M., Ho, J.W.C., Chan, A.S.Y., Lam, P.W.Y., Tse, C.W., Dunlop, M.G., Wyllie, A.H. et Yuen, S.T. 2001.** Early-onset

- colorectal cancer with stable microsatellite DNA and near-diploid chromosomes. *Oncogene*. 20: 4871-4876.
59. **Chang-Claude, J., Eby, N., Kiechle, M., Bastert, G. et Becher, H. 2000.** Breastfeeding and breast cancer risk by age 50 among women in Germany. *Cancer Causes Control*. 11: 687-95.
  60. **Chapgar, A., Magliocco, A., Kerviche, A., Tan, L., Walley, B. et DeCoteau, J. F. 2004:** The replication erMMR phenotype is associated with the development of distant metastases in hormonally treated patients with breast carcinoma. *Cancer*. 100: 913-919.
  61. **Charames, G. S. et Bapat, B. 2004. Genomic instability and cancer.** *BJU International* 93: 19-24.
  62. **Chie, W. C., Hsieh, C., Newcomb, P. A., Longnecker, M. P., Mittendorf, R., Greenberg, E. R., Clapp, R. W., Burke, K. P., Titus-Ernstoff, L., Trentham-Dietz, A. et MacMahon, B. 2000.** Age at any full-term pregnancy and breast cancer risk. *American Journal of Epidemiology*. 1;151(7):715-722
  63. **Chlebowski, R. T., Manson, J. E., Anderson, G. L., Cauley, J. A., Aragaki, A. K., Stefanick, M. L., Lane, D. S., Johnson, K. C., Wactawski-Wende, J., Chen, C., Qi, L., Yasmeen, S., Newcomb, P. A. et Prentice, R. L. 2013.** Estrogen plus progestin and breast cancer incidence and mortality in the Women's Health Initiative Observational Study. *Journal of the National of Cancer Institute*. 105:526-535.
  64. **Clavel-Chapelon, F. et Gerber, M. 2002.** Reproductive factors and breast cancer risk. Do they differ according to age at diagnosis? *Breast Cancer Research Treatment*. 72:107-115
  65. **Colditz, G. A., Baer, H. J. et Tamimi, R. M. 2006.** Breast cancer In: Schottenfeld D, Fraumeni JF Jr, eds. *Cancer Epidemiology and Prevention*. 3rd ed. New York: Oxford University
  66. **Collaborative Group on Hormonal Factors in Breast Cancer and breastfeeding 2002:** collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet*. 360:187-195.
  67. **Conne, B., Stutz, A. et Vassalli, J. D. 2000.** *Nature Medicine* 6: 637-641.
  68. **Conrad, D. F., Jakobsson, M., Coop, G., Wen, X., Wall, J. D., Rosenberg, N. A. et Pritchard, J. K. 2006.** A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nature Genetics*. 38:1251-1260.
  69. **Consort. Int. HapMap. 2005.** A haplotype map of the human genome. *Nature*. 437:1299-1320.
  70. **Contegiacomo A, Palmirotta R, DeMarchis L, Pizzi C, Mastranzo P, Delrio P, Patrella F, Figliolini M, Bianco AR, Frati L et al. 1995.** Microsatellite instability and pathologic aspects of breast cancer. *International Journal of Cancer*. 64: 264-268
  71. **Cunningham, J. M., Kim, C. Y., Christensen, E. R., Tester, D. J., Parc, Y., Burgart, L. J., Halling, K. C., McDonnell, S. K., Schaid, D. J., Walsh, Vockley, C. W., Kubly, V., Nelson, H., Michels, V. V. et Thibodeau, S. N. 2001:** The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *American Journal of Human Genetics*. 69(4): 780-790

## D

72. **Daly, B. et Olopade, O. 2015.** A perfect storm: How tumor biology, genomics, and health care delivery patterns collide to create a racial survival disparity in breast cancer and proposed interventions for change... *CA: A Cancer Journal for Clinicians*. 65: 221-238. DOI: 10.3322/caac.21271

73. **Datta, K. et Biswas, J. 2009.** Influence of dietary habits, physical activity and affluence factors on breast cancer in East India: a case-control study. *Asian Pacific journal of cancer prevention: APJCP.* 10(2):219–22
74. **Dem, A., Traoré, B., Dieng, M. M., Ouajdi, T., Lalani, M. T., Diop, M., Dangou, J. M. et Touré, P. 2008.** Les cancers gynécologiques et mammaires à l'institut du cancer de Dakar. *Cahiers santé.* 18 (1): 25-29
75. **DeMarchis, L., Contegiacomo, A, D'Amico, C., Palmirotta, R., Pizzi, C., Ottini, L., Mastranzo, P., Figliolini, M., Petrella, G., Amanti, C. et al 1997:** Microsatellite instability is correlated with lymph node positive breast cancer. *Clinic Cancer Research* 3: 241-248.
76. **den Dunnen, J. T. et Antonarakis, S. E. 2001.** Nomenclature for the description of human sequence variations. *Human Genetics.* 109(1) : 121-124
77. **den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., Roux, A. F., Smith, T., Antonarakis, S. E. et Taschner, P. E. 2016.** HGVS Recommendations for the description of sequence variants: 2016 Update. *Human Mutation.* 37(6): 564–569. doi: 10.1002/humu.22981
78. **DeSantis, C., Jemal, A., Ward, E., Thun, M. J. 2008.** Temporal trends in breast cancer mortality by state and race. *Cancer Causes Control.* 19:537-545
79. **Diaz-Cano, S. J., Blanes, A., Rubio, J., Matilla, A. et Wolfe, H. J. 2000.** Molecular evolution and intratumour heterogeneity by topographic compartments in muscle-invasive transitional cell carcinoma of the urinary bladder. *Laboratory Investigation.* 80:279–289.
80. **Dienstmann, R., Jang, I., Bot, B., Friend, S. et Guinney, J. 2015.** Database of genomic biomarkers for cancer drugs and clinical targetability in solid tumours. *Cancer Discovery.* 5:118–123
81. **Ding, L., Ellis, M. J, Li, S., Larson, D. E, Chen, K., Wallis, J. W., Harris, C. C., McLellan, M. D., Fulton, R. S., Fulton, L. L., Abbott, R. M., Hoog, J., Dooling, D. J., Koboldt, D. C., Schmidt, H., Kalicki, J., Zhang, Q., Chen, L., Lin, L., Wendl, M. C., McMichael, J. F., Magrini, V. J., Cook, L., McGrath, S. D., Vickery, T. L., Appelbaum, E., Deschryver, K., Davies, S., Guintoli, T., Lin, L., Crowder, R., Tao, Y., Snider, J. E., Smith, S. M., Dukes, A. F., Sanderson, G. E., Pohl, C. S., Delehaunty, K. D., Fronick, C. C., Pape, K. A., Reed, J. S., Robinson, J. S, Hodges, J. S., Schierding, W., Dees, N. D., Shen, D., Locke, D. P., Wiechert, M. E., Eldred, J. M., Peck, J. B., Oberkfell, B. J., Lolofie, J. T., Du, F., Hawkins, A. E., O'Laughlin, M. D., Bernard, K. E., Cunningham, M., Elliott, G., Mason, M. D., Thompson, D. M. Jr., Ivanovich, J. L., Goodfellow, P. J., Perou, C. M., Weinstock, G. M., Aft, R., Watson, M., Ley, T. J., Wilson, R. K. et Mardis. E. R. 2010.** Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature.* 464(7291): 999–1005.
82. **Dvaladze, A., Duggan, C., Gralow, J. R. et Anderson, B. O. 2016.** Breast Cancer Initiative 2.5. (BC12.5): A Global Campaign to Reduce Disparity in Breast Cancer Outcomes. *Journal of Global Oncology.*

## E

83. **Ellegren, H. 2004.** Microsatellites: simple sequences with complex evolution. *National Review Genetics.* 5: 435–445.
84. **Ellsworth, R., Blackburn, H., Shriver, C., Soon-Shiong, P. et Ellsworth, D. 2017.** Molecular heterogeneity in breast cancer: State of the science and implications for patient care features. *Seminars in Cell & Developmental Biology* 64 (2017) 65–72

85. **Eshleman, J. R. et Markowitz, S. D. 1996.** Mismatch repair defects in human carcinogenesis. *Human Molecular Genetics*. 5:1489–94.
86. **European patent specification. 2010.** Mononucleotide repeats microsatellite markers for detecting microsatellite instability. WO 2003/072822 (04.09.2003 Gazette 2003/36).
87. **Evans, D. G., Harkness, E. F., Howel, S., Woodward, E. R., Howell, A., et Lalloo, F. 2018.** Young age at first pregnancy does protect against early onset breast cancer in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Research and Treatment*. 167(3): 779-785. doi: 10.1007/s10549-017-4557-1.
88. **Excoffier, L. G. Laval, and S. Schneider. 2005.** Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 4750
89. **Excoffier, L., Hofer, T. and Foll, M. 2009.** Detecting loci under selection in a hierarchically structured population. *Heredity (Edinb)*. 103(4): 285-298
90. **Excoffier, L, and Lischer, H.E.L. 2010.** Arlequin Suite ver 3.5, a New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Molecular Ecology Resources*. 10: 564-567.

## F

91. **Ferlay, J., Soerjomataram, I, Ervik, M., Dikshit, R., Eser, S., Mathers, C. Rebelo, M., Parkin, D. M., Forman, D. et Bray, F. 2013.** GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 11. Lyon, France: International Agency for Research on Cancer. Available from <http://globocan.iarc.fr>, Accessed on 01 Dec 2017
92. **Ferlay, J, Colombet M, Soerjomataram I, Mathers, C., Parkin, D. M., Pineros, Marion, Znaor, A. et Bray, F. 2018.** Global and Regional Estimates of the Incidence and Mortality for 38 Cancers: GLOBOCAN 2018. Lyon: International Agency for Research on Cancer/World Health Organization
93. **Fidler, I. J. et Kripke, M. L. 1977.** Metastasis Results from Preexisting Variant Cells Within a Malignant Tumor. *Science*. 197(4306): 893 - 895
94. **Fidler, I. J. 1978.** Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Research*. 38(9): 2651 – 2660.
95. **Fishel, R., Lescoe, M. K., Rao, M. R., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M. et Kolodner, R. 1993.** The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*. 75(5):1027-1038.
96. **Friedberg, E. C., Lehmann, A. R. Fuchs, R. P. 2005.** Trading places: How do DNA polymerases switch during translesion DNA synthesis? *Mol Cell* 18: 499–50
97. **Fu, Y. X. 1997.** Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*. 147(2): 915-925
98. **Fukui, K. 2010.** DNA Mismatch Repair in Eukaryotes and Bacteria. *Journal of Nucleic Acids*. 2010:1-16. Doi: 10.4061/2010/260512

## G

99. **Gajalakshmi, C. K., Shanta, V. et Hakama, M. 1998.** Risk factors for contralateral breast cancer in Chennai (Madras), India. *International Journal of Epidemiology*. 27:743 - 50.
100. **Garrigan, D., Kingan, S. B., Pilkington, M. M., Wilder, J. A., Cox, M.P., Soodyall, H., Strassmann, B., Destro-Bisol, G., De Knijff, P., Vovelleto, A., Friedlaender, J. et Hammer, M. F. 2007.** Inferring human population sizes, divergence times and rates of

- gene flow from mitochondrial, X and Y chromosome resequencing data. *Genetics*. 177: 2195–2207.
101. **Gelsomino, F., Barbolini, M., Spallanzani, A., Pugliese, G. et Cascinu, S. 2016.** The evolving role of microsatellite instability in colorectal cancer: A review. *Cancer Treatment Reviews*. 51:19–26
  102. **Georgiades, I. B., Curtis, L. J., Morris, R. M., Bird, C. C. et Wyllie, A. H. 1999.** Heterogeneity studies identify a subset of sporadic colorectal cancers without evidence for chromosomal or microsatellite instability. *Oncogene*. 18: 7933-7940.
  103. **Gerber, B., Muller, H., Reimer, T., Krause, A. et Friese, K. 2003.** Nutrition and lifestyle factors on the risk of developing breast cancer. *Breast Cancer Research Treatment*. 2003, 79: 265-276. 10.1023/A:1023959818513.
  104. **Gerlinger, M., Rowan, A. J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., Varela, I., Phillimore, B., Begum, S., McDonald, N. Q., Butler, A., Jones, D., Raine, K., Latimer, C., Santos, C. R., Nohadani, M., Eklund, A. C., Spencer-Dene, B., Clark, G., Pickering, L., Stamp, G., Gore, M., Szallasi, Z., Downward, J., Futreal, A. et Swanton, C. 2012** Intratumour heterogeneity and branched evolution revealed by multiregion sequencing. *New England Journal of Medicine*. 366(10): 883 – 892.
  105. **Gerlinger, M., Horswell, S., Larkin, J., Rowan, A. J., Salm, M. P., Varela, I., Fisher, R., McGranahan, N., Matthews, N., Santos, C. R., Martinez, P., Phillimore, B., Begum, S., Rabinowitz, A., Spencer-Dene, B., Gulati, S., Bates, P. A., Stamp, G., Pickering, L., Gore, M., Nicol, D. L., Hazell, S., Futreal, A., Stewart, Aengus, S. et Swanton. 2014.** Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature Genetics*. 46(3): 225 - 233
  106. **Germain D. 2011.** Estrogen carcinogenesis in breast cancer. *Endocrinology metabolism Clinic of North America*; 40: 473-484
  107. **Gersten, O. et Wilmoth, J. R. 2002.** The cancer transition in Japan since 1951. *Demographic Research*. 7 : 271-306
  108. **Goel, A., Nagasaka, T., Hamelin, R. et Boland, C. R. 2010.** An optimized pentaplex PCR for detecting DNA mismatch repair-deficient colorectal cancers. *PLoS One* 5(2): e9393
  109. **Goldstein, D. B., Ruiz-Linares, A., Cavalli-Sforza, L. L., et Feldman, M. W. 1995.** Genetic absolute dating based on microsatellites and modern human origins. *Proceedings of the National Academy of Science. USA* 92: 6723–6727.
  110. **Goldstein, D. B. et Pollock, D. D. 1997.** Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*. 88: 335–342.
  111. **Goldstein, D. B., Roemer, G. W., Smith, D. A., Reich, D. E., Bergman A. et Wayne, R. K. 1999.** The use of microsatellite variation to infer population structure in a natural model system. *Genetics* 151: 797–801.
  112. **Greaves, M. 2015.** Evolutionary determinants of cancer. *Cancer Discovery*. 5(8):806–20.
  113. **Gross, R.E. 2000.** Breast Cancer Risk Factors, Screening, and Prevention. *Seminars in Oncology Nursing*, 16(3) :176-184
  114. **Gryfe, R., Kim, H., Hsieh, E. T., Aronson, M. D., Holowaty, E. J., Bull, S. B., Redston, M. et Gallinger, S. 2000.** Tumour microsatellite instability and clinical outcome in young patients with colorectal cancer. *New England Journal of Medicine* 342(2): 69–77
  115. **Gundem, G., Van Loo, P., Kremeyer, B., Alexandrov, L. B., Tubio, J. M., Papaemmanuil, E., Brewer, D. S., Kallio, H. M., Hognas, G., Annala, M., Kivinummi, K., Goody, V., Latimer, C., O'Meara, S., Dawson, K. J., Isaacs, W., Emmert-Buck, M. R., Nykter, M., Foster, C., Kote-Jarai, Z., Easton, D., Whitaker,**

H. C. ICGC Prostate Group, Neal, D. E., Cooper, C. S., Eeles, R. A., Visakorpi, T., Campbell, P. J., McDermott, U., Wedge, D. C., Bova, G. S. 2015. The evolutionary history of lethal metastatic prostate cancer. *Nature*. 520(7547): 353–357.

116. Gurvich, O. L., Baranov, P. V., Zhou, J., Hammer, A. W., Gesteland, R. F. et Atkins, J. F. 2003. Sequences that direct significant levels of frameshifting are frequent in coding regions of *Escherichia coli*. *EMBO Journal* 22: 5941-5950.

## H

117. Habel, L. A. et Stanford, J. L. 1993. Hormone receptors and breast cancer. *Epidemiology Review*. 15: 209 – 19
118. Haffner, M. C., Mosbrugger, T., Esopi, D. M., Fedor, H., Heaphy, C. M., Walker, D. A., Adejola, N., Gurel, M., Hicks, J., Meeker, A. K., Halushka, M. K., Simons, J. W., Isaacs, W. B., Demarzo, A. M., Nelson, W. G. et Yegnasubramanian, S. 2013. Tracking the clonal origin of lethal prostate cancer. *Journal of Clinical Investigation*. 123(11): 4918 – 4922.
119. Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium. Series*. 41:95-98.
120. Hamajima, N., Hirose, K., Tajima, K., Rohan, T., Calle, E. E., Heath, C. W., Coates, R. J., Liff, J. M., Talamini, R., Chantarakul, N., Koetsawang, S., Rachawat, D., Morabia, A., Schuman, L., Stewart, W., Szklo, M., Bain, C., Schofield, F., Siskind, V., Band, P., Coldman, A. J., Gallagher, R. P., Hislop, T. G., Yang, P., Kolonel, L. M., Nomura AMY., Hu J., Johnson KC., Mao Y., De Sanjosé S., Lee N., Marchbanks, P., Ory, H. W., Peterson, H. B., Wilson, H. G., Wingo, P. A., Ebeling, K., Kunde, D., Nishan, P., Hopper, J. L., Colditz, G., Gajalanski, V., Martin, N., Pardthaisong, T., Silpisornkosol, S., Theetranont, C., Boosiri, B., Chutivongse, S., Jimakorn, P., Virutamasen, P., Wongsrichanalai, C., Ewertz, M., Adami, H. O., Bergkvist, L., Magnusson, C., Persson, I., Chang-Claude, J., Paul, C., Skegg, D. C. G., Spears, G. F. S., Boyle, P., Evstifeeva, T., Daling, J. R., Hutchinson, W. B., Malone, K., Noonan, E. A., Stanford, J. L., Thomas, D. B., Weiss, N. S., White, E., Andrieu, N., Brêmond, A., Clavel, F., Gairard, B., Lansac, J., Piana, L., Renaud, R., Izquierdo, A., Viladiu, P., Cuevas, H. R., Ontiveros, P., Palet, A., Salazar, S. B., Aristizabel, N., Cuadros, A., Tryggvadottir, L., Tulinius, H., Bachelot, A., Lê, M. G., Peto, J., Franceschi, S., Lubin, F., Modan, B., Ron, E., Wax, Y., Friedman, G. D., Hiatt, R. A., Levi, F., Bishop, T., Kosmelj, K., Primic-Zakelj, M., Ravnihar, B., Stare, J., Beeson, W. L., Fraser, G., Bullbrook, R. D., Cuzick, J., Duffy, S. W., Fentiman, I. S., Hayward, J. L., Wang, D. Y., McMichael, A. J., McPherson, K., Hanson, R. L., Leske, M. C., Mahoney, M. C., Nasca, P. C., Varma, A. O., Weinstein, A. L., Moller, T. R., Olsson, H., Ranstam, J., Goldbohm, R. A., van den Brandt, P. A., Apelo, R. A., Baens, J., de la Cruz, J. R., Javier, B., Lacaya, L. B., Ngelangel, C. A., La Vecchia, C., Negri, E., Marubini, E., Ferraroni, M., Gerber, M., Richardson, S., Segala, C., Gatei, D., Kenya, P., Kungu, A., Mati, J. G., Brinton, L. A., Hoover, R., Schairer, C., Spirtas, R., Lee, H. P., Rookus, M. A., van Leeuwen, F. E., Schoenberg, J. A., McCredie, M., Gammon, M. D., Clarke, E. A., Jones, L., Neil, A., Vessey, M., Yeates, D., Appleby, P., Banks, E., Beral, V., Bull, D., Crossley, B., Goodill, A., Green, J., Hermon, C., Key, T., Langston, N., Lewis, C., Reeves, G., Collins, R., Doll, R., Peto, R., Mabuchi, K., Preston, D., Hannaford, P., Kay, C., Rosero-Bixby, L., Gao, Y. T., Jin, F., Yuan, J-M., Wei, H. Y., Yun, T., Zhiheng, C., Berry, G., Cooper Booth, J., Jelihovsky, T., MacLennan, R., Shearman, R., Wang, Q-S., Baines, C-J., Miller, A. B., Wall, C., Lund, E., Stalsberg, H., Shu, X. O., Zheng, W., Katsouyanni, K., Trichopoulou, A., Trichopoulos, D., Dabancens, A.,

- Martinez, L., Molina, R., Salas, O., Alexander, F. E., Anderson, K., Folsom, A. R., Hulka, B. S., Bernstein, L., Enger, S., Haile, R. W., Paganini-Hill, A., Pike, M. C., Ross, R. K., Ursin, G., Yu, M. C., Longnecker, M. P., Newcomb, P., Bergkvist, L., Kalache, A., Farley, T. M. M., Holck, S., Meirik, O., Collaborative Group on Hormonal Factors in Breast Cancer None. 2002. Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *British journal of cancer*. 87(11):1234 – 1245.
121. Hampel, H., Frankel, W. L., Martin, E., Arnold, M., Khanduja, K., Kuebler, P., Nakagawa, H., Sotamaa, K., Prior, T. W., Westman, J., Panescu, J., Fix, D., Lockman, J., Comeras, I, et de la Chapelle, A. 2005. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *New England Journal of Medicine*. 352(18): 1851–1860
  122. Hanahan, D. et Weinberg, R. A. 2011. Hallmarks of cancer: the next generation. *Cell*. 144(5):646–74.
  123. Hankinson, S., Colditz G. et Willett W. 2004. Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones. *Breast Cancer Research* 6:(5): 213-218
  124. Harfe, B. D., et Jinks-robertson, S. 2000. Sequence composition and context effects on the generation and repair of frameshift intermediates in mononucleotide runs in *Saccharomyces cerevisiae*. *Genetics* 156: 571-578.
  125. Henderson, B. E., Ross, R. et Bernstein, L. 1988. Estrogens as a cause of human cancer: the Richard and Hinda Rosenthal Foundation award lecture. *Cancer Research*. 48:246–53
  126. Heppner, G. H. et Miller, B. E. 1983. Tumour heterogeneity: Biological implication and therapeutic consequences. *Cancer Metastasis Rev.*, 2, 5
  127. Heppner, G. 1984. Tumour heterogeneity. *Cancer Research*. 44(6): 2259 - 2265
  128. Herschkowitz, J. I., Simin, K., Weigman, V. J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K. E., Jones, L., Assefnia, S., Chandrasekharan, S., Backlund, M. G. Yin, Y., Khramtsov, A. I., Bastein, R., Quackenbush, J., Glazer, R. I., Brown, P. H., Green, J. E., Kopelovich, L., Furth, P., Palazzo, J. P., Olopade, O. I., Bernard, P. S., Churchill, G., van Dyke, T. Et Perou, C. M. 2007. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biology* 8: R76
  129. Hiley, C., de Bruin, E. C., McGranahan, N. et Swanton, C. 2014. Deciphering intratumour heterogeneity and temporal acquisition of driver events to refine precision medicine. *Genome Biology*. 15:453.
  130. Hoang, J., Cottu, P., Thuille, B., Salmon, R., Thomas, G., et Hamelin, R. 1997. BAT-26, an indicator of the Replication ErMMR Phenotype in Colorectal Cancers and Cell Lines. *Cancer Research*. 57: 300-303
  131. Hu, Y. F., Russo, I. H. et Russo, J. 2000. Prevention of human breast cancer. *Journal of Women's Cancer*. In press
  132. Hu, Z. I. et McArthur, H. L. 2018. Immunotherapy in Breast Cancer: The New Frontier. *Current Breast Cancer Reports*. 10(2): 35 - 40
  133. Huo, D., Adebamowo, C. A., Ogundiran, T. O., Akang, E. E., Campbell, O., Adenipekun, A., Cummings, S., Fackenthal, J., Ademuyiwa, F., Ahsan, H., Olopade, O. I. 2008. Parity and breastfeeding are protective against breast cancer in Nigerian women. *British journal of cancer*, 98(5), 992-6.
  134. Huo, D., Ikpatt, F., Khramtsov, A., Dangou, J-M., Nanda, R., Dignam, J., Zhang, B., Grushko, T., Zhang, C., Oluwasola, O., Malaka, D., Malami, S., Odetunde, A., Adeoye, A. O., Iyare, F., Falusi, A., Perou, C. M. et Olopade, O. I. 2009. Population

Differences in Breast Cancer: Survey in Indigenous African Women Reveals Over-Representation of Triple-Negative Breast Cancer. *Journal of Clinical Oncology*. 27(27): 4515-4521

135. **Husemann, Y., Geigl, J. B., Schubert, F., Musiani, P., Meyer, M., Burghart, E., Forni, G., Eils, R., Fehm, T., Riethmuller, G. et Klein, C. A. 2008.** Systemic spread is an early step in breast cancer. *Cancer Cell*; 13:58–68.

## I

136. **Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. et Perucho, M. 1993.** Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363(6429): 558–561
137. **Iqbal, J., Ginsburg, O., Rochon, P. A., Sun, P., Narod, S. A. 2015.** Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. *Journal of the American Medical Association*. 313: 165-173

## J

138. **Jakobsson, M., Scholz, S. W., Scheet, P., Gibbs, J. R., VanLiere, J. M., Fung, H. C., Szpiech, Z. A., Degnan, J. H., Wang, K., Guerreiro, R., Bras, J. M., Schymick, J. C., Hernandez, D. G., Traynor, B. J., Simon-Sanchez, J., Matarin, M., Britton, A., van de Leemput, J., Rafferty, I., Bucan, M., Cann, H. M., Hardy, J. A., Rosenberg, N. A. et Singleton, A. B. 2008.** Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*. 451(7181): 998–1003.
139. **Jass, J. R. Serrated adenoma of the colorectum and the DNA methylator phenotype. 2005.** *National. Clinic of Practical. Oncology.*, 2: 398-405.
140. **Jass, J. R. Heredity and DNA methylation in colorectal cancer. 2007** *Gut*, 56: 154-155.
141. **Jemal, A., Freddie, B., Center, M., Ferlay, J., Ward, E. et Forman, D. 2011.** Global cancer statistics. *CA Cancer J Clin* 61: 69–90.
142. **Jones, A. M., Douglas, E. J., Halford, S. E. R., Fiegler, H., Gorman, P. A., Roylance, R. R., Carter, N. P. et Tomlinson, I. P.M. 2005.** Array-CGH analysis of microsatellite-stable, near-diploid bowel cancers and comparison with other types of colorectal carcinoma. *Oncogene*. 24: 118-129.
143. **Jonsson, M., Johansson, O. et Borg, A. 1995.** Infrequent occurrence of microsatellite instability in sporadic and familial breast cancer. *European Journal of Cancer* 31A: 2330-2334,

## K

144. **Kadyrov, F., Dzantiev, L., Constantin, N. et Modrich, P. 2006.** Endonucleolytic function of mutL $\alpha$  in human mismatch repair,” *Cell*. 126(2): 297–308.
145. **Kadyrov, F. A., Holmes, S. F., Arana, M. E., Lukianova, O., O'Donnell, M., Kunkel, T. A. et Modrich, P. 2007** “*Saccharomyces cerevisiae* MutL $\alpha$  is a mismatch repair endonuclease,” *Journal of Biological Chemistry*. 282(51): 67181 - 37190
146. **Kalluri, R. et Weinberg, R. 2009.** The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation*. 119:1420–1428
147. **Kane, M. F., Loda, M., Gaida, G. M., Lipman, J., Mishra, R., Goldman, H., Jessup, J. M. et Kolodner, R. 1997.** Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumours and mismatch repair-defective human tumour cell lines. *Cancer Research* 57: 808 811.

148. **Kapil, U., Bhadoria, A., Sareen, N., Singh, P. et Dwivedi, S. 2014.** Reproductive factors and risk of breast cancer: *A Review. Indian Journal of Cancer* 51:571-6
149. **Karran, P. 1996:** Microsatellite instability and DNA mismatch repair in human cancer. *Seminars in Cancer Biology* 7: 15-24,
150. **Kelsey, J. L., Gammon, M. D. et John, E. M. 1993.** Reproductive factors and breast cancer. *Epidemiology Review* 15:36-47.
151. **Khan, S., Rogers, M., Khurana, K., Meguid, M. et Numann, P. 1998.** Estrogen receptor expression in benign breast epithelium and breast cancer risk. *Journal of National Cancer Institute.* 90: 37 – 41.
152. **Kim, D-J., Park, J-Y, Lee, M-H, Sohn, Y-K. 2004.** The Role of Microsatellite Instability at Chromosome 11p15.5 in the Progression of Breast Ductal Carcinoma. *Journal of Korean Medical Science;* 19: 698-703 ISSN 1011-8934
153. **Kim, G. P., Colangelo, L. H., Wieand, H. S., Paik, S., Kirsch, I., Wolmark, N. et Allegra, C. J. 2007.** Prognostic and predictive roles of high-degree microsatellite instability in colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project Collaborative Study. *Journal of Clinical Oncology;* 25(7): 767 - 772
154. **Klein, C. A. 2009.** Parallel progression of primary tumours and metastases. *Nature Reviews Cancer.* 9: 302–312.
155. **Kotsopoulos, J., Lubinski, J., Lynch, H. T., Klijn, J., Ghadirian, P., Neuhausen, S. L., Kim-Sing, C., Foulkes, W. D., Moller, P., Isaacs, C., Domchek, S., Randall, S., Offit, K., Tung, N., Ainsworth, P., Gershoni-Baruch, R., Eisen, A., Daly, M., Karlan, B., Saal, H. M., Couch, F., Pasini, B., Wagner, T., Friedman, E., Rennert, G., Eng, C., Weitzel, J., Sun, P. et Narod, S. A. Hereditary Breast Cancer Clinical Study Group, Garber, J., Osborne, M., Fishman, D., McLennan, J., McKinnon, W., Merajver, S., Olsson, H., Provencher, D., Pasche, B., Evans, G., Meschino, W. S., Lemire, E., Chudley, A., Rayson, D. et Bellati, C. 2007.** Age at first birth and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Research and Treatment.* 105(2): 221 - 8.
156. **Kroutil, L. C., Register, K., Bebenek, K et Kunkel, T. A. 1996.** Exonucleolytic proofreading during replication of repetitive DNA. *Biochemistry* 35: 1046-1053.
157. **Kunkel, T. et Erie, D. 2005.** DNA Mismatch repair. *Annual review of Biochemistry.* 74: 681-710.
158. **Kvale G. et Heuch I. 1988.** Menstrual factors and breast cancer risk. *Cancer;* 62: 1625-31

## L

159. **Laqmani, Y. A., Mathew, M. et Lobo, S. 2003.** Frequent microsatellite alterations in a predominantly younger age of onset breast cancer population. *International Journal of Molecular Medicine* 11: 117-124,
160. **Laurent-Puig, P., Agostini, J. et Maley, K. 2011.** Colorectal oncogenesis. *Bull Cancer.* 97(11): 1311–1321.
161. **Lengauer, C., Kinzler, K.W., Vogelstein, B. 1997.** Genetic instability in colorectal cancers. *Nature.* 386: 623-627.
162. **Lengauer, C., Kinzler, K.W., Vogelstein, B. 1998.** Genetic instabilities in human cancers. *Nature.* 396: 643-649
163. **Lewontin, R. C. et Krakauer, J. 1973.** Distribution of Gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics.* Vol.74(1):175-195
164. **Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., Cann, H. M., Barsh, G. S., Feldman, M., Cavalli-Sforza, L. L. et Myers, R. M.**

2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*. 319(5866): 1100–1104.
165. **Librado, P. et Rozas, J. 2009.** DnaSP v6: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 25: 1451–1452.
166. **Liehr, J. G., Ulubelen, A. A. et Strobel, H. W. 1986.** Cytochrome P-450-mediated redox cycling of estrogens. *Journal of Biological Chemistry*. 261:16865–70.
167. **Liehr, J. G. et Ricci, M. J. 1996-** Hydroxylation of estrogens as a marker of human mammary tumors. *Proceedings of the National Academy of Sciences U S A*. 93:3294–6
168. **Ligtenberg, M. J., Kuiper, R. P., Chan, T. L., Goossens, M., Hebeda, K. M., Voorendt, M., Lee, T. Y. H., Bodmer, D., Hoenselaar, E., Hendriks-Cornelissen, S. J. B., Tsui, W. Y., Kong, C. K., Brunner, H. G., van KESSEL, A. G., Yuen, S. T., van Krieken, H., Leung, S. Y. et Hoogerbrugge, N. 2009.** Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 30 exons of TACSTD1. *Nature Genetics*. 41: 112–7.
169. **Lindor, N. M., Burgart, L. J., Leontovich, O., Goldberg, R. M., Cunningham, B. A., Young, J. P., Barker, M. A., Jass, J. R., Hopper, J., Gallinger, S., Bapat, B., Redston, M. et Thibodeau, S. N. 2002.** Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumours. *Journal of Clinical Oncology*. 20: 1043 - 1048.
170. **Liu, B., Parsons, R. E., Hamilton, S. R., Petersen, G. M., Lynch, H. T., Watson, P., Markowitz, S., Wilson, J. K. V., Green, J. S. et de la Chapelle, A. 1994.** hMSH2 mutations in hereditary nonpolyposis colorectal cancer kindreds. *Cancer Research*. 54(17): 4590 – 4594.
171. **Liu, Q., Wu, J., Lambe, M., Hsieh, S-F., Ekbom, A. et Hsieh, C-C. 2002.** Transient increase in breast cancer risk after giving birth: postpartum period with the highest risk. *Cancer Causes Control*. 13:299–305
172. **Loeb, L. A., Bielas, J. H. et Beckman, R. A. 2008.** Cancers exhibit a mutator phenotype: clinical implications. *Cancer Research*. 68:3551–3557.
173. **Lord, S. J., Bernstein, L., Johnson, K. A., Malone, K. E., McDonald, J. A., Marchbanks, P. A., Simon, M. S., Strom, B. L., Press, M. F., Folger, S. G., Burkman, R. T., Deapen, D., Spirtas, R. et Ursin, G. 2008.** Breast cancer risk and hormone receptor status in older women by parity, age of first birth, and breastfeeding: a case-control study. *Cancer epidemiology, biomarkers and prevention*. 17(7):1723–1730
174. **Loukola A, Eklin K, Laiho P, Salovaara R, Kristo P, Jarvinen H, Mecklin JP, Launonen V, Aaltonen LA. 2001.** Microsatellite marker analysis in screening for hereditary nonpolyposis colorectal cancer (HNPCC). *Cancer Research*. 61:4545– 4549.

## M

175. **Mackay, J., Jemal, A., Lee, N. C. et Parkin, D. M. 2006.** The Cancer Atlas. Atlanta: *American Cancer Society*
176. **Madigan, M., Ziegler, R., Benichou, J., Byrne C., et Hoover R. 1995.** Proportion of breast cancer cases in the United States explained by well-established risk factors. *Journal National Cancer Institute*. 87: 1681-1685. 10.1093/jnci/87.22.1681.
177. **Malins, D. C. et Haimanot, R. 1991.** Major alterations in the nucleotide structure of DNA in cancer of the female breast. *Cancer Research*. 51:5430–2.
178. **Malins, D. C., Holmes, E. H., Polissar, N. L. et Gunselman, S. J. 1993.** The etiology of breast cancer. Characteristic alteration in hydroxyl radical-induced DNA base lesions during oncogenesis with potential for evaluating incidence risk. *Cancer*. 71:3036–43.

179. **Marina, P. 2010.** Epidemiology of breast cancer in young women. *Breast Cancer Research and Treatment, Springer Verlag*, 123 (s1), pp.3-6. <10.1007/s10549-010-1098-2>. <hal 00565563
180. **Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R., Zborowska, E., Kinzler, K., Vogelstein, B., Brattain, M. et Willson, J. 1995** Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science*. 268:1336–1338.
181. **Maule, M. Merletti, F. 2012.** Cancer transition and priorities for cancer control. *The Lancet Oncology*. 13:745-746
182. **Mbaye, F., Dem, A., Fall, M. et Sembène, M. 2012.** Implication of the Cytochrome B nucleotide and protein mutations in the occurrence of breast cancer in Senegal. *International Journal of Applied Biology and Pharmaceutical Technology*. 3 (2): 107-114
183. **Mbaye, F., Dem, A., KA, S., Balloucouné, G., Fall, M., Kane, M. et Sembene, M. 2014.** Mutation screening in the mitochondrial D-Loop region of tumoral and non tumoral breast tissues in Senegalese patients. *International Journal of Cancer Research*. 48 (2): 1564-1573.
184. **Mbaye, F., Ka, S., Dem, A., Kane, M., et Sembène, M. 2015.** Clinicopathological and survival significance of *BAT-25* and *BAT-26* instability in breast cancer among Senegalese patients. *International Journal of Cancer Research and Experimental Oncology*. 7. 13-19. 10.5897/JCREO2015.0122.
185. **McMurray, C. T. 2010.** Mechanisms of trinucleotide repeat instability during human development. *Nature Review Genetics* 11(11):786–799.
186. **Mehrgou, A. et Akouchekian, M. 2016.** The importance of BRCA1 and BRCA2 genes mutations in breast cancer development. *Medical Journal of the Islamic Republic of Iran*. 30: 369.
187. **Mettlin, C. 1999.** Global breast cancer mortality statistics. *CA Cancer of Journal Clinicians* 49:138–44
188. **Metzgar, D., Bytof, J. et Wills, C. 2000.** Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Research*, 10: 72–80.
189. **Miller, B. E., Roi, L. D., Howard, L. M. et Miller, F. R. 1983.** Quantitative selectivity of contact-mediated intercellular communication in a metastatic mouse mammary tumour line. *Cancer Research*. 43, 4102
190. **Miller, A. B. et Bulbrook, R. D. 1986.** UICC multidisciplinary project on breast cancer: the epidemiology, aetiology and prevention of breast cancer. *International Journal of Cancer*. 37: 173–177.
191. **Miller, W. R. 1993.** Hormonal factors and risk of breast cancer. *Lancet*. 341: 25 – 26.
192. **Modrich, P. 1987** DNA mismatch correction. *Annual Review of Biochemistry*. 56: 435–66.
193. **Modrich, P: 1994.**Mismatch repair, genetic instability and cancer. *Science* 266: 1959-1960,
194. **Modrich, P. 2006.** “Mechanisms in eukaryotic mismatch repair,” *Journal of Biological Chemistry*, vol. 281, no. 41, pp. 30305–30309.
195. **Molino, A., Micciolo, R., Turazza, M., Bonetti, F., Piubello, Q., Corgnati A, Sperotto, L., Recaldin, E., Spagnoli, R., Manfrin, E., Bonetti, A., Nortilli, R., Pollini, G. P., Modena, S. et Cetto, G. L. 1997.** Prognostic significance of estrogen receptors in 405 primary breast cancers: a comparison of immunohistochemical and biochemical methods. *Breast Cancer Research Treatment*. 45:241–9
196. **Moolgavkar, S. H., Day, N. E. et Stevens, R. G. 1980.** Two-stage model for carcinogenesis: epidemiology of breast cancer in females. *Journal of National Cancer Institute*; 65:559–69.

197. **Morgan, M. A. et Reuter, C. W. 2006** Molecularly targeted therapies in myelodysplastic syndromes and acute myeloid leukemias. *Annals of Hematology*. 85(3):139–63
198. **Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. et Erlich, H. 1986.** Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on quantitative biology*. 51 (1): 263-273
199. **Murata, H., Khattar, N. H., Kang, Y., Gu, L. et Li, G. M. 2002.** Genetic and epigenetic modification of mismatch repair genes hMSH2 and hMLH1 in sporadic breast cancer with microsatellite instability. *Oncogene* 21: 5696-5703.
200. **Murtaza, M., Dawson, S. J., Pogrebnik, K., Rueda, O. M., Provenzano, E., Grant, J., Chin, S. F., Tsui, D. W., Marass, F., Gale, D., Ali, H. R., Shah, P., Contente-Cuomo, T., Farahani, H., Shumansky, K., Kingsbury, Z., Humphray, S., Bentley, D., Shah, S. P., Wallis, M., Rosenfeld, N. et Caldas, C. 2015.** Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. *Nature Communications*. 6:8760.

## N

201. **Nakagawa, H., Lockman, J. C., Frankel, W. L., Hampel, H., Steenblock, K., Burgart, L. J., Thibodeau, S. N. et de la Chapelle, A. 2004.** Mismatch repair gene PMS2: diseasecausing germline mutations are frequent in patients whose tumors stain negative for PMS2 protein, but paralogous genes obscure mutation detection and interpretation. *Cancer Research*. 64(14):4721–4727
202. **Nandi, S., Guzman, R. C. et Yang, J. 1995.** Hormones and mammary carcinogenesis in mice rats and humans: a unifying hypothesis. *Proceedings of National Academy of Sciences U S A*; 92:3650–7.
203. **Neale, R. E., Darlington, S., Murphy, M. F., Silcocks, P. B., Purdie, D. M. et Talbäck, M. 2005.** The effects of twins, parity and age at first birth on cancer risk in Swedish women. *Twin Research Human Genetics*. 8:156-62.
204. **Niederacher, D., Schnurch, H. G., An H. X., Ellenberger, I., Dall, P., Van Roeyen, C. R., Kupperts, V. et Beckmann, M. W. 1996.** Detection of sequential genetic alterations relevant for breast cancer development. *European Journal of Cancer Prevention*; 5: 497-503.
205. **Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen, L., Leroy, C., Ramakrishna, M., Rance, R., Lau, K. W., Mudie, L. J., Varela, I., McBride, D. J., Bignell, G. R., Cooke, S. L., Shlien, A., Gamble, J., Whitmore, I., Maddison, M., Tarpey, P. S., Davies, H. R., Papaemmanuil, E., Stephens, P., J., McLaren, S., Butler, A. P., Teague, J. W., Jönsson, G., Garber, J. E., Silver, D., Miron, P., Fatima, A., Boyault, S., Langerød, A., Tutt, A., Martens, J. W. M., Aparicio, S. A. J. R., Borg, Å., Salomon, A. V., Thomas, G., Børresen-Dale, A-L., Richardson, A. L., Campbell, P. J., Breast Cancer Working Group of the International Cancer Genome Consortium. 2012.** Mutational processes molding the genomes of 21 breast cancers. *Cell*; 149:979–993.
206. **Nik-Zainal, S., Loo, P. V., Wedge, D. C., Alexandrov, L. B., Greenman, C. D., Lau, K. W., Raine, K., Jones, D., Marshall, J., Ramakrishna, M., Shlien, A., Cooke, S. L., Hinton, J., Menzies, A., Stebbings, L. A., Leroy, C., Jia, M., Rance, R., Mudie, L. J., Gamble, S. J., Stephens, P. J., McLaren, S., Tarpey, P. S., Papaemmanuil, E., Davies, H. R., Varela, I., McBride, D. J., Bignell, G. R., Leung, Adam P. Butler, Jon W. Teague, Sancha Martin, Goran Jönsson, Odette Mariani, K., Boyault, S., Miron, P., Fatima, A., Langerød, A., Aparicio, A. J. R., Tutt, A., Sieuwerts, A. M., Borg, A.,**

- Thomas, G., Salomon, A. V., Richardson, A., Borresen-Dale, A. L., Futreal, P. A., Stratton, M. R. et Campbell, P. J. 2012. The life history of 21 breast cancers. *Cell*; 149:994–1007.
207. Niwa, T., Bradlow, H. L., Fishman, J. et Swaneck, G. E. 1989. Determination of estradiol 2- and 16-alpha hydroxylase activities in MCF-7 human breast cancer cells in culture using radiometric analysis. *Journal of Steroid Biochemistry and Molecular Biology*; 33: 311–4.
208. Nobili, S., Bruno, L., Landini, I., Napoli, C., Bechi, P., Tonelli, F., Rubio, C. A., Mini, E. et Nesi, G. 2011. Genomic and genetic alterations influence the progression of gastric cancer. *World Journal of Gastroenterology*. 17(3):290–299.
209. Nordentoft, I., Lamy, P., Birkenkamp-Demtroder, K., Shumansky, K., Vang, S., Hornshøj, H., Juul, M., Villesen, P., Hedegaard, J., Roth, A., Thorsen, K., Hoyer, S., Borre, M., Reinert, T., Fristrup, N., Dyrskjot, L., Shah, S., Pedersen, J. S. et Orntoft, T. F. 2014. Mutational Context and Diverse Clonal Development in Early and Late Bladder Cancer. *Cell Reports*. 7(5): 1649–1663.
210. Ntekim, A., Nufu, F. et Campbell, O. 2009. Breast cancer in young women in Ibadan Nigeria. *African health sciences*. 9(4): 242-246.

## O

211. Ogino, S., Fuchs, C. S. et Giovannucci, E. 2012. How many molecular subtypes? Implications of the unique tumour principle in personalized medicine. *Expert Review of Molecular Diagnostics*. 12(6):621–8.
212. Ogundiran, T. et Adebamowo, C. 2015. Tropical Hemato-Oncology. J.-P. Droz *et al.* (eds.). *Springer International Publishing Switzerland*. (293-302).
213. Okobia, M., Bunker, C., Zmuda, J., Kammerer, C., Vogel, V., Uche, E., Anyanwu, S., Ezeome, E., Ferrell, R. et Kuller, L. 2006. Case – control study of risk factors for breast cancer in Nigerian women. *International Journal of Cancer* 119:2179– 2185
214. Omran, A. R. 1971. The epidemiologic transition. A theory of the epidemiology of population change. *Milbank Memorial Fund Q*. 49:509-538.
215. Oliveira, E. J., Padua, J. G., Zucchi, M. I., Vencovsky, R. et Viera, M. L. C. 2006. Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology*. 29(2):
216. Osborne, M. P., Bradlow, H. L., Wong, G. Y. et Telang, N. T. 1993. Upregulation of estradiol C-16 hydroxylation in human breast tissue: a potential biomarker of human breast cancer risk. *Journal of National Cancer Institute* 85:1917–20.
217. Ottini, L., Palli, D., Falchetti, M., D’Amico, C., Noviello, C., Saieva, C., Russo, A., Corsi, A. M., Masala, G., Paglierani, M., Vezzosi, V., Bianchi, S. et Mariani-Costantini, R. 2000. Instability at dinucleotide and trinucleotide repeats in breast cancer. *International Journal of Oncology*. 17: 819-826,.
218. Ozer, E., Yuksel, E., Kizildag, S., Sercan, O., Ozen, E., Canda, T. et Sakilzi, M. 2002. Microsatellite instability in early-onset breast cancer. *Pathology Research Practice* 198: 525-530.

## P

219. Papadopoulos, N., Nicolaidis, N. C., Wei, Y. F., Ruben, S. M., Carter, K. C., Rosen, C. A., Haseltine, W., Fleishmann, R. D., Fraser, C. M et Adams, M. D. 1994 Mutation of a mutL homolog in hereditary colon cancer. *Science*. 263:1625–9.
220. Parkin, D., Ferlay, J., Hamdi-Cherif, M., Sitas, F., Thomas, J., Wabinga, H. et Whelan, S. 2003. *Cancer in Africa: Epidemiology and Prevention*. Lyon: IARC

221. Paulson, T. G., Wright, F. A., Parker, B. A., Russack, V. et Wahl, G. M. 1996. Microsatellite instability correlates with reduced survival and poor disease prognosis in breast cancer. *Cancer Research* 56: 4021-4026.
222. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A-L., Brown, P. O. et Botstein, D. 2000. Molecular portraits of human breast tumours. *Nature* 406:747–752
223. Perucho, M. 1999. Correspondence ref.: Boland C. C. *et al.*, 1998. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Research*. 58: 5248–5257, 1998. *Cancer Research* 59(1): 249–256
224. Pesole, G., Liuni, S., Grillo, G. et Saccone, C. 1997. Structural and compositional features of untranslated regions of eukaryotic mRNAs. *Gene*. 205: 95 – 102.
225. Pike, M. C., Spicer, D. V., Dahmouch, L. et Press, M. F. 1993. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiology Review*. 15:17–35.
226. Podo, F., Buydens, L. M., Degani, H., Hilhorst, R., Klipp, E., Gribbestad, I. S., van Huffel, S., van Laarhoven, H. W., Luts, J., Monleon, D., Postma, J. G., Schneiderhan-Marra, N., Santoro, F., Wouters, H., Russnes, H. G., Sørli, T., Tagliabue, E. et Børresen-Dale, A. L. 2010. Triple negative breast cancer: Present challenges and new perspectives. *Molecular. Oncology*. 4 (3), 209e229.
227. Polyak, K. 2007. Breast cancer: Origins and evolution. *Journal of Clinical Investigations*. 117(11): 3155-3163
228. Prat, A., Parker, J. S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J. I., He, X. et Perou, C. M. 2010. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research* 12: R68 Press: 995-1012.
229. Price, J. E., Bell, C. et Frost, P. 1990. The use of a genotypic marker to demonstrate clonal dominance during the growth and metastasis of a human breast carcinoma in nude mice. *International Journal of Cancer*; 45:968-971
230. Pyatt, R., Chadwick, R. B., Johnson, C. K., Adebamowo, C., Chapelle, A. et Prior, T. W. 1999. Polymorphic Variation at the *BAT-25* and *BAT-26* Loci in Individuals of African Origin. *American Journal of Pathology*. 155(2) : 349-353.

## R

231. Rao, D. N., Ganesh, B. et Desai, P. B. 1994. Role of reproductive factors in breast cancer in a low-risk area: A case-control study. *British Journal of Cancer*. 70:129-132
232. Reich, D. E. et Goldstein D. B. 1998. Genetic evidence for a Paleolithic human population expansion in Africa. *Proceedings of National Academy of Sciences. USA* 95: 8119–8123.
233. Reis-Filho, J. S. et Tutt, A. N., 2008. Triple negative tumours: a critical review. *Histopathology* 52(1): 108-118
234. Rhim, A. D., Mirek, E. T., Aiello, N. M., Maitra, A., Bailey, J. M., McAllister, F., Reichert, M., Beatty, G. L., Rustgi, A. K., Vonderheide, R. H., Leach, S. D. et Stanger, B. Z. 2012. EMT and dissemination precede pancreatic tumour formation. *Cell*. 148:349–361.
235. Ribic, C. M., Sargent, D. J., Moore, M. J., Thibodeau, S. N., French, A. J., Goldberg, R. M., Hamilton, S. R., Laurent-Puig, P., Gryfe, R., Shepherd, L. E., Tu, D., Redston, M. et Gallinger, S. 2003. Tumour microsatellite instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *New England Journal of Medicine*. 349(3): 247 – 257.

236. **Rodriguez-Bigas, M. A., Boland, C. R., Hamilton, S. R., Henson, D. E., Jass, J. R., Khan, P. M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L et Srivastava, S. 1997.** A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *Journal of National Cancer Institute.* 89(23): 1758-1762
237. **Rogers, A. R. et Harpending, H. 1992.** Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution.* 9(3):552-69.
238. **Rozas J., Librado P., Sánchez-Del Barrio J.C., Messeguer X. et Rozas R. 2012.** DnaSP Version 5 Help Contents [Help File]. Available with the program at <http://www.ub.edu/dnasp/>
239. **Roy M. S., Geffen E., Smith D., Ostrander E., et Wayne R. K. 1994.** Patterns of differentiation and hybridization in North American wolflike Canids. *Molecular Biology and Evolution.* 11: 553–570.
240. **Rubio-Perez, C., Tamborero, D., Schroeder, M., Antolin, A., Deu-Pons, J., Perez-Llamas, C., Mestres, J., Gonzalez-Perez, A., et Lopez-Bigas, N. 2015.** In silico prescription of anticancer drugs to cohorts of 28 tumour types reveals targeting opportunities. *Cancer Cell.* 27:382–396.
241. **Rush, E. B., Calvano, J. E., Van Zee, K. J., Zelenetz, A. D. et Borgen, P. I. 1997.** Microsatellite instability in breast cancer. *Annals of Surgical Oncology* 4: 310-315.
242. **Russo, J. et Russo, I. H. 1987.** Development of human mammary gland. In: Neville MC, Daniel CW, editors. The mammary gland development regulation and function. New York (NY): *Plenum Publishing Corporation.* p. 67–93
243. **Russo, J., Rivera, R. et Russo, I. H. 1992.** Influence of age and parity on the development of the human breast. *Breast Cancer Research Treatment.* 23:211–8
244. **Russo, J., Calaf, G., et Russo, I. H. 1993a.** A critical approach to the malignant transformation of human breast epithelial cells with chemical carcinogens. *Critical Reviews in Oncogenesis* 4:403–17.
245. **Russo, J. et Russo, I. H. 1993b.** Development pattern of human breast and susceptibility to carcinogenesis. *European Journal of Cancer Prevention;* 2(suppl 3):85–100.
246. **Russo, J., Romero, A. L. et Russo, I. H. 1994.** Architectural pattern of the normal and cancerous breast under the influence of parity. *Cancer Epidemiology Biomarker Prevention.* 3:219–24
247. **Russo, I. H. et Russo J. 1996.** Mammary gland neoplasia in long-term rodent studies. *Environmental Health Perspectives.* 104:938–67.
248. **Russo, J., Barnabas, N., Zhang, P. L. et Adesina, K. 1996.** Molecular basis of breast cell transformation. *Radiations Oncology Investigations.* 3:424–9.
249. **Russo, J., A. X., Grill, C. et Russo, I. H. 1999.** Pattern of distribution of cells positive for estrogen receptor  $\alpha$  and progesterone receptor in relation to proliferating cells in the mammary gland. *Breast Cancer Research Treatment;*53: 217–27
250. **Russo, J., Hu, Y. E., Yang, X. et Russo, I. H. 2000.** Developmental, Cellular, and Molecular Basis of Human Breast Cancer. *Journal of the National Cancer Institute Monographs* 27: 17-37

## S

251. **Sagher, D., Hsu, A. et Strauss, B. 1999.** Stabilization of the intermediate in frameshift mutation. *Mutation Research.*, 423, 73–77.
252. **Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. et Arnheim, N. 1985.** Enzymatic amplification of  $\beta$ -globin genomic sequences and

- restriction site analysis for diagnosis of sickle cell anemia. *Science*. 230 (4732): 1350-1354.
253. **Salovaara R, Loukola A, Kristo P, Kaariainen H, Ahtola H, Eskelinen M, Harkonen N, Julkunen R, Kangas E, Ojala S, Tulikoura J, Valkamo E, Jarvinen H, Mecklin JP, Aaltonen LA, de la Chapelle A (2000)** Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *Journal of Clinic Oncology* 18(11): 2193–2200
  254. **Samowitz, W. S., Slattery, M. L., Potter, J. D. et Leppert, M. F. 1999.** *BAT-26* and *BAT-40* instability in colorectal adenomas and carcinomas and germline polymorphisms. *The American Journal of Pathology*. 154(6): 1637-1641.
  255. **Sargent, D. J., Marsoni, S., Monges, G., Thibodeau, S. N., Labianca, R., Hamilton, S. R., French, A. J., Kabat, B., Foster, N. R., Torri, V., Ribic, C., Grothey, A., Moore, M., Zaniboni, A., Seitz, J. F., Sinicrope, F. et Gallinger, S. 2010.** Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *Journal Clinic of Oncology* 28(20): 3219–3226
  256. **Schmitt, F. C., Soares, R., Gobbi, H., Milanezzi, F., Santos-Silva, F., Cirnes, L., Costa, C. et Seruca, R. 1999.** Microsatellite instability in medullary breast carcinomas. *International Journal of Cancer* 82: 6440647
  257. **Schneider, J., Kinne, D., Fracchia, A., Pierce, V., Anderson, K. E., Bradlow, H. L. et Fishman, J. 1982.** Abnormal oxidative metabolism of estradiol in women with breast cancer. *Proceedings of National Academy of Sciences U. S. A.* 79(9):3047–3051.
  258. **Schneider, S. et Excoffier, L. 1999.** Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics*. 152: 1079–1089.
  259. **Schulz, M., Hoffmann, K., Weikert, C., Nothlings, U., Schulze, M. B. et Boeing, H. 2008.** Identification of a dietary pattern characterized by high-fat food choices associated with increased risk of breast cancer: the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. *The British Journal of Nutrition*. 100(5): 942 – 946.
  260. **Shah, S. P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., Turashvili, G., Ding, J., Tse, K., Haffari, G., Bashashati, A., Prentice, L. M., Khattra, J., Burleigh, A., Yap, D., Bernard, V., McPherson, A., Shumansky, K., Crisan, A., Giuliany, R., Heravi-Moussavi, A., Rosner, J., Lai, D., Birol, I., Varhol, R., Tam, A., Dhalla, N., Zeng, T., Ma, K., Chan, S., Griffith, M., Moradian, A., Cheng, S-W. G., Chai, S., Chin, S-F., Curtis, C., Rueda, O., Pharoah, P. D., Damaraju, S., Mackey, J., Hoon, K., Harkins, T., Tadigotla, V., Sigaroudinia, M., Gascard, P., Tlsty, T., Costello, J. F., Meyer, I. M., Eaves, C., Wasserman, W. W., Jones, S., Huntsman, D., Hirst, M., Caldas, C., Marra, M. A. et Aparicio, S. 2012.** The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*. 486:395–399.
  261. **Shemirani A., Haghghi M. et Zadeh S. 2011.** Simplified MSI marker panel for diagnosis of colorectal cancer. *Asian Pacific Journal of Cancer Prevention.*; 12:2101–4.
  262. **Sia, E. A., Butler, C. A., Dominska, M., Greenwell, P., Fox T. D., et Thomas, D. P. 2000.** Analysis of microsatellite mutations in the mitochondrial DNA of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Science USA* 97: 250–255
  263. **Singh, A. et Settleman, J. 2010.** EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* 29: 4741–4751
  264. **Sjoblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., Szabo, S., Buckhaults, P., Farrell, C., Meeh, P., Markowitz, S.D., Willis, J., Dawson, D., Willson, J.K., Gazdar, A.F., Hartigan, J., Wu, L., Liu, C., Parmigiani, G., Park, B.H., Bachman, K.E., Papadopoulos, N.,**

- Vogelstein, B., Kinzler, K.W. et Velculescu, V.E. 2006. The consensus coding sequences of human breast and colorectal cancers. *Science*. 314: 268-274
265. Sonenberg, N. 1994. mRNA translation: influence of the 5' and 3' untranslated regions. *Current Opinion in Genetics & Development*. 4(2):310-315 A.
266. Song, Q., Merajver, S. et Li, J. 2015. Cancer classification in the genomic era: five contemporary problems. *Human Genomics* 9 : 27 DOI 10.1186/s40246-015-0049-8
267. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffery, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P., Botstein, D., Lonning, P. E., et Borresen-Dale, A. L. 2001. Gene expression patterns of breast carcinomas distinguish tumour sub-classes with clinical implications. *Proceedings of National Academy of Sciences U. S. A.* 98(19): 10869 – 10874
268. Sourvinos, G., Kiaris, H., Tsikkinis, A., Vassilaros, S. et Spandidos, D. A. 1997. Microsatellite instability and loss of heterozygosity in primary breast cancers. *Tumour Biology*. 18: 157-166,
269. Spink, D. C., Hayes, C. L., Young, N. R., Christou, M., Sutter, T. R., Jefcoat, C. R. et Gierthy, J. F. 1994. The effects of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin on estrogen metabolism in MCF-7 breast cancer cells: evidence for induction of a novel 17-estradiol 4-hydroxylase. *Journal of Steroid Biochemistry and Molecular Biology*. 51(5-6): 251–258.
270. Strauss, B. S. 1999. Frameshift mutation, microsatellites and mismatch repair. *Mutation Research* 437: 195-203.
271. Subramaniam, D. S. et Isaacs, C. 2005. Utilizing prognostic and predictive factors in breast cancer. *Current Treatment Options in Oncology*. 6 : 147e159
272. Suraweera, N., Iacopetta, B., Duval, A., Compoin, A., Tubacher, E. et Hamelin, R. 2001. Conservation of mononucleotide repeats within 30 and 50 untranslated regions and their instability in MSI-H colorectal cancer. *Oncogene*, 20, 7472–7477.
273. Suraweera, N., Duval, A., Reperant, M., Vaury, C., Furlan, D., Leroy, K., Seruca, R., Iacopetta, B. et Hamelin, R. 2002. Evaluation of tumour microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* 123(6): 1804–1811
274. Sutter, C., Gebert, J., Bischoff, P., Herfarth, C. et von Knebel, D. M. 1999. Molecular screening of potential HNPCC patients using a multiplex microsatellite PCR system. *Molecular and Cellular Probes*. 13: 157–165
275. Swanton, C. 2012. Intratumor heterogeneity: evolution through space and time. *Cancer Research*. 72: 4875–82
276. Sweeney, C., Blair, C. K., Anderson, K. E., Lazovich, D. et Folsom, A. R. 2004. Risk factors for breast cancer in elderly women. *The American Journal of Epidemiology*. 160: 868 - 875

## T

277. Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 123 (3): 585-95
278. Tamakoshi, K., Yatsuya, H., Wakai, K., Suzuki, S., Nishio, K., Lin, Y., Niwa, Y., Kondo, T., Yamamoto, A., Tokudome, S., Toyoshima, H. et Tamakoshi, A. 2005. Impact of menstrual and reproductive factors on breast cancer risk in Japan: Results of the JACC study. *Cancer Science*. 96(1): 57-62.
279. Tereschenko, I. V., Basham, V. M., Ponder, B. A. et Pharoah, P. D. 2002. BRCA1 and BRCA2 mutations in Russian familial breast cancer. *Human mutation*. 19(2): 184

280. **Theillet, C. 2010.** Vers un inventaire des mutations oncogéniques dans les cancers humains. *Bull Cancer*. 97(11):1223-1229
281. **Thibodeau, S. N., Bren, G. et Schaid, D. 1993.** Microsatellite instability in cancer of the proximal colon. *Science* 260(5109): 816–819
282. **Tishkoff, S. A., Dietzsch, E., Speed, W., Pakstis, A. J., Kidd, J. R., Bonn -Tamir, B., Santachiara-Benerecetti, A. S., Moral, P., Watson, E., Krings, M., P  bo, S., Risch, N., Jenkins, T. et Kidd, K. K. 1996.** Global patterns of linkage disequilibrium at the *CD4* locus and modern human origins. *Science*. 271: 1380 –1387.
283. **Tishkoff, S. A. et Kidd, K. K. 2004.** Implications of biogeography of human populations for ‘race’ and medicine. *Nature Genetics*. 36: S21–S27.
284. **Tishkoff, S. A. et Verrelli, B. C. 2003.** Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annual Review of Genomics and Human Genetics*. 4:293–340
285. **Tishkoff, S. A. et Williams, S. M. 2002.** Genetic analysis of African populations: human evolution and complex disease. *Nature Review Genetics* 3:611–621.
286. **Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J. B., Awomoyi, A. A., Bodo, J. M., Dumbo, O., Ibrahim, M., Juma, A. T., Kotze, M. J., Lema, G., Moore, J. H., Mortensen, H., Nyambo, T. B., Omar, K. Powell, G.S. Pretorius, M.W. Smith, M.A. Thera, C. Wambebe, J.L. Weber, S. A. et Williams, S. M. 2009.** The genetic structure and history of Africans and African Americans. *Science* 324(5930): 1035-1044.
287. **Tokunaga, E., Oki, E., Oda, S., Kataoka, A., Kitamura, K., Olno, S., Maehara, Y. et Sugimachi, K. 2000.** Frequency of microsatellite instability in breast cancer determined by high resolution fluorescent microsatellite analysis. *Oncology* 59: 44-49.
288. **Tomita, S., Deguchi, S., Miyaguni, T., Muto, Y., Tamamoto, T. et Toda, T. 1999.** Analyses of microsatellite instability and the transforming growth factor factor-, receptor type II gene mutations in sporadic breast cancer and their correlation with clinicopathological features. *Breast Cancer Research and Treatment*. 53(1): 33-39
289. **Tonin, P. 2000.** Genes implicated in hereditary breast cancer syndromes. *Seminars in Surgical Oncology*. 18:281-286.
290. **Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J. et Jemal, A. 2015.** Global Cancer Statistics, 2012. *CA: A Cancer Journal for Clinicians*. 65(2): 87 - 108
291. **Toyama, T., Iwase, H., Yamashita, H., Iwata, H., Yamashita, T., Ito, K., Hara, Y., Suchi, M., Kato, T., Nakamura, T. et Kobayashi, S. 1996.** Microsatellite instability in sporadic human breast cancers. *International Journal of Cancer*. 68: 447-451.
292. **Tran, H. T., Keen, J. D., Kricker, M., Resnick M. A. et Gordenin, D. A. 1997.** Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Molecular Cell Biology*. 17: 2859-2865

## U

293. **Umar, A., Risinger, J. I., Hawk, E. T. et Barrett, J. C. 2004(a).** Testing guidelines for hereditary non-polyposis colorectal cancer. *National Review of Cancer*. 4:153–8.
294. **Umar, A., Boland, C. R., Terdiman, J. P., Syngal, S., de la Chapelle, A., Ruschoff, J., Fishel, R., Lindor, N. M., Burgart, L. J., Hamelin, R., Hamilton, S. R., Hiatt, R. A., Jass, J., Lindblom, A., Lynch, H. T., Peltomaki, P., Ramsey, S. D., Rodriguez-Bigas, M. A., Vasen, H. F., Hawk, E. T., Barrett, J. C., Freedman, A. N., Srivastava, S. 2004(b).** Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *Journal of National Cancer Institute* 96(4): 261–268.

295. Underhill, P. A., Jin, L., Zemans, R., Ofener, P. J. et Cavalli-Sforza, L. L. 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proceedings of National Academy of Sciences. USA* 93: 196–200.

## V

296. Valastyan, S. et Weinberg, R. 2011. Tumour metastasis: molecular insights and evolving paradigms. *Cell*. 147:275–292.
297. Van de Vijver, M. J. 1993. Molecular genetic changes in breast cancer. In: Weinberg RA, editor. *Oncogenes and the molecular origins of cancer*. Cold Spring Harbor (NY): Cold Spring Harbor Press. p. 25–6.
298. Vilar, E. et Gruber, S. B. 2010. Microsatellite instability in colorectal cancer-the stable evidence. *National Review of Clinic Oncology*. 7(3):153–62.
299. Vogelstein, B., Fearon, E. R., Kern, S. E., Hamilton, S. R., Preisinger, A. C., Nakamura, Y. et White, R. 1989 Allelotype of colorectal carcinomas. *Science*. 244(4901): 207–211.
300. Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, Jr L. A. et Kinzler, K. W. 2013. Cancer genome landscapes. *Science*. 339(6127): 1546–58.

## W

301. Waghorne, C., Thomas, M., Lagarde, A. E., Kerbel, R. S., Breitman, M. L. 1988. Genetic evidence for progressive selection and overgrowth of primary tumours by metastatic cell subpopulations. *Cancer Research*. 48: 6109–6114.
302. Wagner, L. A., Weiss, R. B., Driscoll, R., Dunn, D. S. et Gesteland, R. F. 1990. Transcriptional slippage occurs during elongation at runs of adenine or thymine in *Escherichia coli*. *Nucleic Acids Research* 18 : 3529-3535.
303. Welch, D. 2016. Tumour heterogeneity-a ‘contemporary concept’ founded on historical insights and predictions, *Cancer Research*; 76 (1) 4–6.
304. Wellings SR, Jensen HM, Marcum RG. 1975. An atlas of sub-gross pathology of the human breast with special reference to possible precancerous lesions. *Journal of National Cancer Institute*. 55:231–73
305. Wild, P., Stoehr, R., Knuechel, R., Hartmann, A. et Dietmaier, W. 2005. Laser microdissection for microsatellite analysis in colon and breast cancer. *Methods in Molecular Biology*. 293:93-101.
306. Woerner, S., Yuan, Y., Benner, A., Korff, S., Doeberitz, M. et Bork, P. 2009. SelTarbase, a database of human mononucleotidemicrosatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic Acids Research*; 1–8
307. Wohlfahrt, J., Andersen, P. K., Mouridsen, H. T., Adami, H. O. et Melbye, M. 1999. Reproductive history and stage of breast cancer. *American Journal of Epidemiology*. 150: 1325–1330
308. Wohlfahrt, J., Anderson, P., Mouridsen, H. et Melbye, M. 2001. Risk of late-stage breast cancer after a childbirth. *American Journal of Epidemiology*. 153: 1079-1084
309. Wong, Y. F., Cheung, T. H., Lo, K. W., Yim, S. F., Chan, L. K., Buhard, O., Duval, A., Chung, T. K. et Hamelin, R. 2006. Detection of microsatellite instability in endometrial cancer: advantages of a panel of five mononucleotide repeats over the National Cancer Institute panel of markers. *Carcinogenesis* 27(5): 951–955
310. World Health Organization: WHO MONICA Project Manual. 1990. Geneva: World Health Organization MONICA Project.
311. World Health Organization (WHO). 2008. *World Cancer Report*. Lyon: International Agency for Research on Cancer.

## X

312. **Xicola, R. M., Llor, X., Pons, E., Castells, A., Alenda, C., Pinol, V., Andreu, M., Castellvi-Bel, S., Paya, A., Jover, R., Bessa, X., Giros, A., Duque, J. M., Nicolas-Perez, D., Garcia, A. M., Rigau, J. et Gassull, M. A. 2007.** Performance of different microsatellite marker panels for detection of mismatch repair deficient colorectal tumours. *Journal of National Cancer Institute* 99(3): 244–252

## Y

313. **Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R. H., Eshleman, J. R., Nowak, M. A., Velculescu, V. E., Kinzler, K. W., Vogelstein, B. et Iacobuzio-Donahue, C. A. 2010.** Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*. 467: 1114 – 1117.
314. **Yager, J. et Davidson, N. 2006.** Estrogen Carcinogenesis in breast cancer. *The New England Journal of Medicine*. 354: 270-282
315. **Yan, Z. J. et Roy, D. 1995.** Mutations in DNA polymerase RNA of stilbene estrogen-induced kidney tumors in Syrian hamster. *Biochemistry and Molecular Biology International*. 37:175–83
316. **Yates, L. R. et Campbell, P. J. 2012.** Evolution of the cancer genome. *Nature Reviews Genetics*. 13(11): 795-806.
317. **Yee, C. J., Roodi, N., Verrier, C. S. et Parl, F. F. 1994.** Microsatellite instability and loss of heterozygosity in breast cancer. *Cancer Research*. 54: 1641-1644.
318. **You, J. F., Buhard, O., Ligtenberg, M. J., Kets, C. M., Niessen, R. C., Hofstra, R. M., Wagner, A., Dinjens, W. N., Colas, C., Lascols, O., Collura, A., Flejou, J. F., Duval, A. et Hamelin, R. 2010.** Tumours with loss of MSH6 expression are MSI-H when screened with a pentaplex of five mononucleotide repeats. *British Journal of Cancer* 103(12): 1840–1845
319. **Yuza, K., Nagahashi, M., Watanabe, S., Takabe, K., et Wakai, T. 2017.** Hypermutation and Microsatellite instability in gastrointestinal cancers. *Oncotarget*. 8:67

## Z

320. **Zhao, Z., Yu, N., Fu, Y. X. et Li, W. H. 2006.** Nucleotide variation and haplotype diversity in a 10-kb noncoding region in three continental human populations. *Genetics*. 174: 399 – 409.
321. **Zhou, X. P., Hoang, J-M., Cottu, P., Thomas, G. et Hamelin, R. 1997.** Allelic profiles of mononucleotide repeat microsatellites in control individuals and in colorectal tumors with and without replication erMMRs. *Oncogene*, 15, 1713–1718
322. **Zhou, X. P., Hoang, J. M, Li, Y. J., Seruca, R., Carneiro, F., Sobrinho-Simoes, M., Lothe, R. A., Gleeson, C. M., Russell, S. E., Muzeau, F., Flejou, J. F., Hoang-Xuan, K., Lidereau, R., Thomas, G. et Hamelin, R. 1998.** Determination of the replication error phenotype in human tumors without the requirement for matching normal DNA by analysis of mononucleotide repeat microsatellites." *Genes Chromosomes Cancer*. 21(2): 101-107.

## Référence Web

323. **Canadian Cancer Society. 2018.** <http://www.cancer.ca/en/cancer-information/cancer-type/breast/risks/?region=bc>

324. **Genetics Home Reference - NIH. 2018.** Disponible sur <https://ghr.nlm.nih.gov/gene>, Consulté le 1er novembre, 2018.
325. **Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C. D. et Parkin, D.** Cancer Incidence and Mortality Worldwide: *IARC Cancer Base No.10 [Internet]*. **Globocan. 2008.** Lyon, France: International Agency for Research on Cancer. 2010; Available from: <http://globocan.iarc.fr>.
326. **International Agency for Research on Cancer (2015) Globocan 2012:** estimated cancer incidence, mortality and prevalence worldwide. IARC Press, Lyon. Available at: [http://globocan.iarc.fr/Pages/fact\\_sheets\\_population.aspx](http://globocan.iarc.fr/Pages/fact_sheets_population.aspx)
327. **World Health Organization. 12 September 2018.** Disponible sur <https://www.who.int/news-room/fact-sheets/detail/cancer>, Consulté le 15 octobre, 2018 à 12:00pm.

**ANNEXES**

<b>CLINICAL RECORD</b>	
<b>PATIENT IDENTIFICATION (CASES)</b>	
1. Code No: .....	
2. Hospital No: .....	
3. City/State of residence: .....	
4. Date of birth: (dd/mm/yy): .....	
5. Age (in Years): .....	
6. Place of birth (City/State): .....	
7. Height (cm): .....	
8. Weight (kg): .....	
9. WHR: .....	
10. Ethnicity: (i) Igbo (ii) Yoruba (iii) Hausa, (iv) Others, please specify .....	
11. Occupation: (i) Trader; (ii) Housewife; (iii) Farmer; (iv) Teacher; (v) Artisan (vi) Student (vii) Others, Please Specify: .....	
12. Marital status: (i) Single; (ii) Married (iii) Divorced (iv) Widowed	
13. Birth order: .....out of ..... children	
<b>ANTECEDENTS</b>	
14. Age of first menstruation .....	
15. Date of last menstruation: .....	
16. Gravidity (Number of pregnancy): .....	
17. Parity (Number of Live birth/s): .....	
18. Have you taken any fertility drug before? Yes (.....) No (.....)	
19. Have you taken any birth control pills before? Yes (.....) No (.....)	
<b>PLEASE TICK WHERE APPROPRIATE:</b>	
20. Menopausal status: (i) Premenopausal: Yes (...) No (...); (ii) Postmenopausal Yes (...); No (.....) (iii) Age at menopause .....	
21. Past medical history: (i) Hypertension (.....) (ii) Diabetes (.....) (iii) Heart disease (.....) (iv) Asthma (...)	
22. Similar family history (of past medical history): Yes (.....) No (.....) If 'Yes', which disease	

(.....).

23. Degree of relationship .....

**Risk factor for cancer:**

24. Have you ever smoked before? Yes (.....) No (.....). If 'NO', Go to question 28 (iii)

25. If yes, for how long? .....

26. Do you still smoke? Yes (.....), No (.....)

27. If 'No', When did you stop? .....

28. Which of these do you take? : (i) Snuff (....) How often do you smoke in a day? (.....)

(ii) Cigarette (...). How many stick(s)/packet(s) do you smoke in a day (.....)

(iii) Smoke/fume: I cook with (Stove)..... (Gas) ..... (Firewood).....

29. Do you have any family history of cancer: (i) Yes ..... (ii) No .....

30. If 'Yes' to 27, please specify which cancer: .....

31. Degree of relationship: Ist degree (.....), 2<sup>nd</sup> degree (.....).

32. Which family member? (i) Father (.....), (ii) Mother (.....), (iii) Brother (.....), (iv)

Sister (.....) (v) Aunt (.....) (iv) Others, Please specify: .....

(Paternal....., Maternal.....)

33. Do you have any other type of cancer? Yes (.....), No (.....).

34. If 'Yes', which cancer .....

35. Has it been treated before? Yes ( ), No ( );

36. If 'Yes' to 35, what type of treatment? (i) Chemotherapy ( ), (ii) Radiotherapy ( ), (iii)

Surgery ( ). 37.

Have you had/treated any breast disease before? Yes (.....); No (.....). If 'Yes', What type? (i)

Lump (ii) tumour

**CLINICAL**

38. Tumour site/Localization (Quadrant): Right breast: (i). UOQ (ii). UIQ (iii). LOQ

(iv). LIQ (v). NAC (vi). AT

Left breast: (i). UOQ (ii). UIQ (iii). LOQ (iv). LIQ (v). NAC (vi). AT

39. T.....N.....M.....

**HISTOLOGY**

40. Type of tumour: DCIS (.....), IDC (.....), LCIS (.....), ILC (.....)

39. Grade: .....

41. Immuno-histochemical molecular sub-type ER + / - ; PR + / - HER2 + / - 42.

Ganglionic invasion: Yes ..... No.....

## **Publication I**

## **Microsatellite Instability of BAT-26 and NR-24 among Breast Cancer Patients in Senegal**

Aniefiok John Udoakang, Fatimata Mbaye, Mbacke Sembene

*Equipe Génétique et Gestion pour les Populations, Département de Biologie Animale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop de Dakar, BP. Box 5005 Dakar, Sénégal.*  
Corresponding Author: Aniefiok John Udoakang

**Abstract:** Microsatellite instability high (MSI-H) tumours are known to have better prognosis than microsatellite stable (MSS) tumours and diagnostic characterization of these tumours reveal differences in clinical, pathological and molecular characteristics. Different populations have variable results for MSI tumours due to varied environmental factors that influence diverse outcomes. Therefore, with the successes in the exploration of immunotherapy in cancers and currently in breast cancer, this study aims to evaluate the microsatellite phenotypes of breast cancer patients in Senegal to better understand the pattern of tumour progression and design therapies for improved treatment outcome. Sixty-five breast tumours were genotyped for microsatellite instability of BAT-26 and NR-24 markers. The results showed MSI in more than 50% of the tumours with variant alleles of 92.5% and 81.4% in BAT-26 and NR-24 respectively. Statistical significance ( $p$ -value = 0.041) was observed for SBR grade between MSS and MSI tumours. This study shows the MSI phenotype of BAT-26 and NR-24 markers among breast cancer patients in Senegal with a positive correlation between MSI and SBR grade in NR-24 marker. We propose that NR-24 which often shows little or no variation among healthy individuals in different populations may be a marker of interest in the determination of MSI status as a prognostic factor among breast cancer patients in Senegal.

Date of Submission: 30-12-2018

Date of acceptance: 14-01-2019

### **I. Introduction**

Globally, female breast cancer (FBC) ranks second (11.6%) and fifth (6.6%) in terms of incidence and death respectively. Additionally, FBC along with lung and colorectal cancers are responsible for one third of the burden of cancer incidence and death (Bray, 2018). In Africa, it ranks fifth in incidence and five-year prevalence, as well as third in mortality. Likewise, in Senegal, FBC is second to cervical cancer in number of new cases (18.97%) and deaths (12.46%) (Ferlay *et al.*, 2018).

Being a multifactorial disease that results from the association between various genetic, environmental, hormonal and lifestyle factors, breast cancer development involves a multistep process that is associated with various genetic alterations (Mehrgou and Akouchekian, 2016). Only about 5 – 10 percent of breast cancer have clearly defined genetic predispositions while about 90 – 95 percent of the disease have no defined germline mutation but may develop through epigenetic inactivation or acquired mutations that may arise from a defective mismatch repair (dMMR) system (Walsh *et al.*, 2006; Chacon and Costanzo, 2010; Guler *et al.*, 2011; Prat and Perou, 2011).

The mismatch repair (MMR) system maintains the integrity of the genome by correcting accumulated mutations that occur as a result of DNA polymerase slippage errors during DNA replication (Levinson and Gutman, 1987; Ellegren, 2000; Scarpa *et al.*, 2016). The MMR system in humans is regulated by MutS $\alpha$  (a heterodimer complex of *hMSH2* and *hMSH6* that binds to base-base mismatches and the loops of one or several nucleotides) and MutS $\beta$  (a MutS-related complex composed of the heterodimers *hMSH2* and *hMSH3* that repairs the heteroduplexes with two or several extrahelical bases (Acharya *et al.*, 1996; Genschel *et al.*, 1998). Together, MutS $\alpha$  and MutS $\beta$  then binds to MutL $\alpha$  (a heterodimer of MutL-homologues *hMLH1* and *hPMS2*), excising the DNA strand with the mismatch (Buermeier *et al.*, 1999; Kolodner and Marsischky, 1999; Jiricny, 2000). The inactivation of any of these genes due to mutation results in dMMR leading to an increased rate of mutation especially in the microsatellite regions and this is referred to as microsatellite instability (MSI) (Modrich and Lahue, 1996; Geurts-Giele *et al.*, 2014; Joly *et al.*, 2015; Richman, 2015).

The microsatellites region consists of tandem nucleotide repeat motifs of 1 to 6 bp and are interspersed throughout the eukaryotic genome in the coding and non-coding regions of the DNA (Subramanian *et al.*, 2003; Cuadrado *et al.*, 2008; Richman, 2015). They are considered to be mutational hot-spots because of the high level of polymorphisms observed among individuals and populations and therefore are widely used as molecular markers in evolutionary studies, genetic mapping, bone marrow engraftment monitoring and DNA

fingerprinting (Lee *et al.*, 2001; Schlotterer and Har, 2001; De Bustos *et al.*, 2016). Although, microsatellites are stable in each individual, their mutation rates are significantly elevated not only in the dMMR system, but also in tumour suppressor genes resulting in instability of the microsatellite repeats (Imai and Yamamoto, 2008; Shia, 2015). MSI has been identified in about 10% - 15% of common human tumours such as colorectal, gastric, stomach and endometrium (Boland *et al.*, 1998; Hamelin *et al.*, 2008; Zhu *et al.*, 2015; Ratti *et al.*, 2018). The slippage errors that occur during DNA replication often lead to frameshift mutations in target genes and encourage the growth of tumour cells. These drive the process of carcinogenesis in microsatellite instability high (MSI-H) tumours and are more likely to occur in long runs of G:C than A:T repeats (Sagher *et al.*, 1999). The evaluation of MSI in tumours has proven to be an efficient tool for the prognosis of various cancers and they are more sensitive in mononucleotide repeats than dinucleotide repeats (Imai and Yamamoto, 2008). Thus, there is the need to utilize poly-T or poly-A mononucleotides microsatellite markers for the analysis of MSI phenotype in tumour cells (Ionov *et al.*, 1993; Boland *et al.*, 1998).

The evaluation of MSI is standardized in some cancers especially in hereditary non-polyposis colorectal cancers (HNPCC) in which MSI-H tumours are known to have better prognosis than microsatellite stable (MSS) tumours (Popat *et al.*, 2005; Shokal and Sharma, 2012). These tumours exhibit specific pattern of gene expression in their phenotype while their diagnostic characterization reveal differences in clinical, pathological and molecular characteristics (Abida *et al.*, 2018). Therefore, it is necessary to analyze different tumours in diverse populations with the vast demographic differences from varied geographical locations. MSI has been reported in various frequencies of between 0 – 33% in patients with breast cancer although the differences observed may be due to the non-uniform selection of microsatellite markers, samples, and the use of different criteria for analysis (De Marchis *et al.*, 1997; Arzimanoglou *et al.*, 1998; Anbazhagan *et al.*, 1999; Dudley *et al.*, 2016).

After its discovery in 1993, several highly variable markers were initially used for the evaluation of MSI but to have a consensus and facilitate comparison between studies, the NCI in 1997 proposed a panel of two mononucleotides (*BAT-25* and *BAT-26*) and three dinucleotides (*D5S346*, *D2S123* and *D17S250*) repeats (Ionov *et al.*, 1993; Boland *et al.*, 1998; Akiyama *et al.*, 1997; Perucho, 1999). However, apart from the highly polymorphic nature of dinucleotide repeats requiring the use of corresponding germline DNA (which can be expensive and time consuming), there have also been evidence of misclassification of MSI-H tumours when using dinucleotide repeats or the Bethesda panel (Hoang *et al.*, 1997; Loukola *et al.*, 2001; Suraweera *et al.*, 2002). Hence, to circumvent these challenges, an optimized pentaplex panel of five quasimonomorphic mononucleotide repeats has been opined to be more sensitive and specific for the detection of MSI-H tumours (Suraweera *et al.*, 2002; Bacher *et al.*, 2004; Umar *et al.*, 2004). *BAT-25* and *BAT-26* which are common to both panels are the most commonly used mononucleotide markers and it has been suggested that testing them alone would be sufficient to establish the MSI status of a tumour without the need for the analysis of corresponding normal tissues (Hatch *et al.*, 2005; Malesci *et al.*, 2007; Xicola *et al.*, 2007). Also, the analysis of *BAT-26* alone has been suggested to be sufficient for the detection of the MSI-H phenotype in most cases (Hoang *et al.*, 1997). Furthermore, polymorphic variations of these markers (*BAT-25* and *BAT-26*) have been observed in several populations depending on ethnicity in sub-Saharan Africa including Senegal, showing variant alleles corresponding to 5% and more (Samowitz and Slattery, 1999; Pyatt *et al.*, 1999; Suraweera *et al.*, 2002; Buhard *et al.*, 2006). Besides, the use of *BAT-26* alone might possibly lead to the determination of false positive MSI-H or may allow for underestimation of the true levels of MSI due to its failure to recognize rare cases with biallelic *MSH2* deletions since it lies in the intragenic region of *MSH2*. Hence, the use of at least one other mononucleotide marker (e.g., *BAT-25* or *NR-24*) has been suggested (Laghi *et al.*, 2008). Since diverse populations of the world have variable results for MSI tumours due to various environmental factors ensuing dissimilar outcomes (Shokal and Sharma, 2012), and considering the current success in the exploration of immunotherapy in breast cancer (Hu and McArthur, 2018), it becomes imperative to evaluate the microsatellite phenotypes of *BAT-26* and *NR-24* among breast cancer patients in Senegal for a better understanding of the disease progression to tailor diagnosis and design of therapies for a better treatment outcome.

## II. Materials And Methods

### Sample collection

All consecutive breast cancer patients who visited the Institut Joliot-Curie Cancer Centre, of the Aristide le Dantec Hospital Dakar, Senegal between 2014 and 2015 were approached for the study. Majority of them agreed to participate in the study with a refusal rate of 4%. Aristide le Dantec Hospital is the major and only referral hospital for cancer patients in the country serving an immediate population of about 2.45 million people. Sixty-five participants between ages 15 and 80 years old were recruited for the study. Cancerous tissue samples were collected by the surgical oncologists through biopsy or later by mastectomy from each patient after informed consent and a standardized questionnaire was used to collect the patients' data. All tissue samples were histologically confirmed to be tumour samples before they were taken in for molecular analysis.

**Ethical approval**

This study was conducted in compliance with the ethical approval and standards of the Universite Cheikh Anta Diop de Dakar Research Ethics Committee (Reference number: Protocole 0269/2015/CER/UCAD). The recommendations guidelines of the Declaration of Helsinki for biomedical research was followed and informed consent was obtained from all participants after detailed explanation of the objectives, benefits, seeming inconveniences (in terms of sample collection) and contribution of the study.

**Choice of markers used in the study and Laboratory Analysis**

The markers used in this study are as outlined, location and primer sequences are as outlined in in Table 1. DNA was extracted from the sixty tumour tissues using the standard protocol for the Qiagen kit and thereafter, they were coded and stored at -20°C for PCR analysis. The amplifications were carried out in a monoplex reaction volume of 50 µl for each repeat and the PCR reaction was made up of 2.5 µl of the forward and reverse primers, 2 µl of each dNTP, 2 µl of MgCl<sub>2</sub> and 0.1 µl of *Taq* polymerase. The PCR was performed in an Eppendorf-type thermal cycler and the conditions were as follows: Initial denaturation at 94°C for 5 minutes, 35 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds, extension at 72°C for 30 seconds and final extension at 72°C for 7 minutes. Thereafter, the PCR products were run on gel electrophoresis using 1.5% agarose. For the sequencing, 30µl of each amplicon was placed in a 96-well microtiter microplates along with each primer (15µl of 10Mm) for each sample and sent to Macrogen, South Korea for Sanger sequencing (Mbaye *et al.*, 2015).

**Table 1:** Primer Sequences for the Microsatellite Instability Assay (Suraweera *et al.*, 2002)

Name	Gene	Genbank number	Repeat	Primer Sequences	Amplicon size (bp)
<i>BAT-26</i>	hMSH2	U41210	26(A) Intron 5	F: TGACTACTTTTGACTTCAGCC R: AACCATCAACATTTTAACCC	120
<i>NR-24</i>	Zinc finger 2 (ZNF-2)	X60152	24T 3'UTR	F: CCATTGCTGAATTTACCTC R: ATTGTGCCATTGCATCCAA	132

**Sequence variant nomenclature and classification**

Antonarakis and the Nomenclature Working Group (1998) suggested a nomenclature system for the description of mutations and polymorphisms observed in DNA and protein sequences. This nomenclature is known as the Human Genome Variation Society (HGVS) recommendations/ nomenclature (den Dunnen and Antonarakis, 2000). The sequence variant nomenclature constitutes the changes observed in a specific sequence when compared to a reference sequence. This gives a stable, accurate, meaningful, unequivocal, unambiguous but flexible description (den Dunnen *et al.*, 2016).

**Molecular and Statistical Analysis**

Data obtained after capillary electrophoresis were analyzed with Bio-Edit Sequence Alignment Editor version 7.1.9 (Thompson *et al.*, 1994; Hall, 1999). To test the association between the association of breast cancer and instability of the studied markers, the sequences were cleaned, corrected and aligned to the reference sequence from the GenBank (Table 1) using the Clustal algorithm (Thompson *et al.*, 1994; Hall, 1999). Since normal DNA was not examined in this study, every sample with a polymorphic variation in any of the mononucleotide repeats (*BAT-26* and *NR-24*) in terms of size and pattern of variations (deletion, insertion, transition or transversion) was scored positive for instability (Wong *et al.*, 2006). The Statistical Package for Social Sciences (SPSS) version 20.0 (SPSS, Inc. Chicago IL, USA) was used to evaluate the frequencies of different variables as well as the chi-square values to determine the statistical differences or otherwise among test variables. The statistically significant value was placed at p-value < 0.05. Tumours were classified as MSI-H tumours if they were unstable in the two markers, MSI-low (MSI-L) if only one marker unstable and MS-stable (MSS) if none of the markers was unstable.

**III. Results**

**Frequency of MSI and their haplotypes**

Of the 65 tumours genotyped for microsatellite instability status among breast cancer patients in Senegal, only 40 of the sequences for *BAT-26* and 43 of the sequences for *NR-24* were clean after alignment and confirmation of each sequence with their chromatogram. Based on the classification of MSI (as earlier stated), 7 (10.8%) were MSS, 45 (69.2%) were MSI-L, and 13 (20%) were MSI-H tumours were identified from the analyzed sequence. The composition of the allele sizes and variants for each marker are presented in Table 2 for *BAT-26* and Table 3 for *NR-24*. The percentage of variant alleles observed for each marker was 92.5% and 81.4% for *BAT-26* and *NR-24* respectively.

For *BAT-26* with 26A repeats, shortened alleles of 1 to 3 bp was observed as well as INDELS (hereafter denoted as ‘delins’) ‘delins’ of up to 3 bp with deletion of the A and insertion of G. Three haplotype groups had substitution of A with G (represented as > for substitution). The highest haplotype with frequency of 10 had a

deletion of A at the 25th position and A>G at the 26th position of the microsatellite repeat. Eight tumours (18.6%) were stable for NR-21 having the wildtype allele of 24T repeat. The variants alleles observed included deletions, substitutions, deletion-insertion and duplication. The highest number of deletions was 4 bp and the haplotype with the highest frequency of 10 had substitution of T>G.

**Table 2:** Composition of allele sizes and variants for BAT-26

Haplotypes	Allele Sizes	Allele variants	Senegal : No.
Hap 1	26A (Wild type)	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	3
Hap 2	26delA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA -	8
Hap 3	24 26delinsG	AAAAAAAAAAAAAAAAAAAAAAAAAAGGG	5
Hap 4	25delA; 26A>G	AAAAAAAAAAAAAAAAAAAAAAAAAAAA-G	10
Hap 5	24 26delA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA --	2
Hap 6	25 26>G	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG	6
Hap 7	24delA; 25>G; 26delA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA -G -	1
Hap 8	25delA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA -A	1
Hap 9	24 25delA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA -A	1
Hap 10	26>G	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	3
<b>Total</b>			<b>40</b>

(del = deletion, > = substitution, delins = deletion-insertion)

**Table 3:** Composition of allele sizes and variants for NR-24

Haplotypes	Allele Sizes	Allele variants	Senegal : No.
Hap 1	24T (Wild type)	TTTTTTTTTTTTTTTTTTTTTTTTTT	8
Hap 2	24dupT	TTTTTTTTTTTTTTTTTTTTTTTTTT	1
Hap 3	24T>G	TTTTTTTTTTTTTTTTTTTTTTTTTG	10
Hap 4	22delT	TTTTTTTTTTTTTTTTTTTTTT-TT	1
Hap 5	22delT; 24T>G	TTTTTTTTTTTTTTTTTTTTTT-TG	3
Hap 6	1delT; 22delinsG; 24dupT	-TTTTTTTTTTTTTTTTTTTTGTTT	1
Hap 7	21delT; 22delinsG; 24dupT	TTTTTTTTTTTTTTTTTTTT-GTTT	2
Hap 8	20delinsG	TTTTTTTTTTTTTTTTTTTTGTTT	1
Hap 9	20delinsG; 24T>G	TTTTTTTTTTTTTTTTTTTTGTTG	2
Hap 10	20delinsG; 22 24delT	TTTTTTTTTTTTTTTTTTTTGT---	1
Hap 11	20delT; 24T>G	TTTTTTTTTTTTTTTTTTTTTTTG	5
Hap 12	20delT; 23delT; 24T>G	TTTTTTTTTTTTTTTTTTTTTT-TG	1
Hap 13	20 21delT; 24T>G	TTTTTTTTTTTTTTTTTTTTTT - TTG	1
Hap 14	20 21delT; 22delinsG	TTTTTTTTTTTTTTTTTTTTTT - GTT	2
Hap 15	1delT; 24T>G	-TTTTTTTTTTTTTTTTTTTTTG	3
Hap 16	17T>G; 20delinsG; 21 24delT	TTTTTTTTTTTTTTTTTTGTTG-----	1
<b>Total</b>			<b>43</b>

(del = deletion, dup = duplication, > = substitution, delins = deletion-insertion)

**Correlation between studied markers and Age at diagnosis, SBR grade and tumour stage of patients.**

Statistical analysis of some clinic-pathological characteristics (age at diagnosis, SBR grade and TNM stage) showed no significant difference between MSS and MSI tumours for the two markers except in the SBR grade for NR-24 (p-value = 0.041). As presented in Table 4, analysis of BAT-26 showed that cases with MSI were more frequent in the age group that are 50 years old (64%) than in those above who were 50 years and above (28%), 73.3% had grade 2 tumour (SBR grading) and 75.7% were either at stage 3 or stage 4 (TNM staging). Likewise, for NR-24, majority of the cases with MSI tumours were of younger age (86.2% < 50). They were more among those who presented either at stage 3 or 4 (62.2%) and had tumour grade two (56%) (Table 5).

**Table 4:** Comparison of Age, SBR grade and TNM stage between MSS and MSI BAT-26 phenotype among breast cancer patients in Senegal

Characteristics		BAT-26 (MSS)	BAT-26 (MSI)	p-value
Age	< 50	2 (5%)	25 (64%)	0.680 <sup>a</sup>
	≥ 50	1 (3%)	11 (28%)	
SBR Grade	1	-	3 (10%)	0.075 <sup>b</sup>
	2	-	22 (73.3%)	
	3	1 (3.3%)	4 (13.3%)	
TNM Stage	1 & 2	1 (2.7%)	6 (16.2%)	0.477 <sup>b</sup>
	3 & 4	2 (5.4%)	28 (75.7%)	

p-value<sup>a</sup> = fisher's exact test; <sup>b</sup> =  $\chi^2$

**Table 5:** Comparison of Age, SBR grade and TNM stage between MSS and MSI NR-24 phenotype among breast cancer patients in Senegal

Characteristics		NR-24 (MSS)	NR-24 (MSI)	p-value
Age	< 50	4 (13.8)	25 (86.2%)	0.202 <sup>a</sup>
	≥ 50	4 (33.3%)	8 (66.7%)	
SBR Grade	1	2 (6%)	2 (6%)	0.041 <sup>b</sup>
	2	2 (6%)	18 (56%)	
	3	4 (13%)	4 (13%)	
TNM Stage	1 & 2	1 (2.7%)	6 (16.2%)	1.000 <sup>b</sup>
	3 & 4	7 (18.9%)	23 (62.2%)	

p-value<sup>a</sup> = fisher's exact test; <sup>b</sup> =  $\chi^2$

#### IV. Discussion

The utilization of microsatellite markers as diagnostic tools to determine the prognosis of certain cancers in order to properly tailor treatment and predict patients' response is increasing rapidly for most solid tumours especially with the FDA approval of drugs for the treatment of patients with MSI-H or dMMR tumours (FDA, 2017). Although breast cancer was not known to be inherently immunogenic, current successes in clinical trials especially for triple negative breast cancer is fast making immunotherapy a possible therapeutic strategy as monotherapy or in combination with the currently used conventional therapies (Williams *et al.*, 2017; De La Cruz and Czerniecki, 2018). Nevertheless, the constant challenge has been to select patients who will most benefit from this treatment and therefore, the need for the evaluation of molecular markers especially for the African population which is plagued with the menace of triple negative breast cancer. Although the use of *BAT-25* and *BAT-26* has already been proposed (de la Chapelle, 1999), polymorphism of these markers exist in individuals from Africa (Pyatt *et al.*, 1999) and this may lead to misrepresentation of the MSI-H phenotype if the two markers are used together. Hence, the need for the evaluation of another mononucleotide marker which can be used alongside either of *BAT-25* and *BAT-26*.

In this study, we evaluated the microsatellite instability status of sixty-five breast tumours from Senegal using *BAT-26* and *NR-24*. 20% of the tumours showed instability in the two markers while *BAT-26* and *NR-24* alone were unstable in 92.5% and 81.4% of the tumours respectively. In a similar study carried out by Mbaye *et al.*, (2015) using *BAT-25* and *BAT-26*, they observed that 70.6% of the tumours were unstable for the *BAT-26* phenotype. Also, the shortening of the *BAT-26* as observed in this study have been positively associated with MSI-H tumours. Consistent with this, *BAT-26* also showed variant alleles of up to 14 alleles out of its quasi-monomorphic variation range (QMVR) in a Brazilian population analysis of the pentaplex panel while no variant alleles were observed for *NR-24* out of its QMVR (Campanella *et al.*, 2014). *NR-24* has been shown to be a remarkably stable marker with little or no variation among several studies that have evaluated mononucleotide repeats in human cancers (Buhard *et al.*, 2006; Mamdouh *et al.*, 2015). Results of microsatellite instability in breast cancer vary for different studies and in different populations with rates as low as 2.9% (Glebov *et al.*, 1994; Patel *et al.*, 1994; Yee *et al.*, 1994; Bergthorsson *et al.*, 1995; Paulso *et al.*, 1996; Risinger *et al.*, 1996; De Marchis *et al.*, 1997; Sourvinos *et al.*, 1997; Fujii *et al.*, 1998; Walsh *et al.*, 1998; Zhou *et al.*, 1998; Siah *et al.*, 1999; Tomita *et al.*, 1999; Siah *et al.*, 2000). It has also been suggested that the choice of markers used for the analysis, ranging from mononucleotides to tetranucleotide repeats and the sample sizes may have influenced these varying results (Siah *et al.*, 2000).

Tumours with MSI-H phenotype have distinct clinical and pathological features from MSS tumours. In this study, statistical significance was observed for only SBR grade and *NR-24* marker between MSS and MSI tumours. However, younger patients are more likely to be unstable for the microsatellite phenotype as recorded in our study where more than half (64%) of individuals below 50 years were unstable for both markers. This has been observed in studies involving other tumours with instability in about 58% of tumours from younger individuals (Lamberti *et al.*, 1999; Gryfe *et al.*, 2000; Boland and Goel, 2010). We found that MSI was more frequent in grade 2 tumours for both markers (73.3% for *BAT-26* and 56% for *NR-24*), with statistical significance ( $p = 0.041$ ) for grade and *NR-24*. Similarly, Ward and colleagues (2001), observed that MSI-H tumours were associated with high grade tumours. The percentage of MSI varied with tumour stage (which was grouped into 2 due to the small sample size and data). About 63% of the microsatellite instable tumours were for individuals who were either at stage 3 or stage 4 of the disease with presence of lymph nodes status. This is in agreement with other studies (Soreide *et al.*, 2009; Eveno *et al.*, 2010; Belt *et al.*, 2012;) and may indicate an association between MSI and aggressiveness of the disease leading to increase likelihood for metastasis as observed in this population where patients present at late stage with aggressive tumour. However, further studies need to be carried out to access if MSI leads to aggressive tumours or vice versa. Nevertheless, our result in contrast to Merok *et al.*, (2012) who observed high frequency of MSI in stage two tumours. It was suggested that this is reflexive of tumours that are less aggressive with lower metastatic tendency as MSI may induce larger lymph nodes.

Two modes of microsatellite instability have been proposed to occur in human cancers: Type A alterations with length changes of  $\leq 6$  bp and Type B alterations with length changes of  $\geq 8$  bp (Oda *et al.*, 2005).

In our study, the Type A mode of MSI was observed as no haplotype had an alteration that was  $\geq 8$  bp. Being that Type A MSI is connected with a dMMR system, we can therefore propose that the MSI observed here is as a result of the inactivation and loss of function of the DNA MMR system. To establish this theory, analysis that identifies particular mutations in the MMR genes such as immunohistochemistry (IHC) must be carried out. Because of the missense or in-frame insertion/deletion mutations that occur in microsatellite unstable tumours, which do not result in the truncation of a protein, assessment of the MSI status could help to ascertain the true functional consequences of these variant mutations especially in some DNA MMR genes like the *MLH1* and *MSH6* genes. This may not be true for the routine IHC analysis which may give a false negative result stemming from biological or technical reasons, therefore, evaluation of MSI allows for the elimination of confusion in case of inconclusive result. MSI analysis provides a functional analysis of the deficient MMR activity elucidating functionality of the MMR system (Zhang and Li., 2013).

### V. Conclusion

Finally, the current success recorded in the use of immune checkpoint blockade molecules in the treatment of cancers that are resistant to chemotherapy such as TNBC requires the efficacious utilization of a molecular marker that will better predict immunotherapeutic response in patients. This study demonstrates microsatellite instability of *BAT-26* and *NR-24* markers among breast cancer patients in Senegal with a positive correlation between MSI and SBR grade in *NR-24*. Also, *NR-24* which shows little or no variation was shown to exhibit instability in our study and may be a better marker for the diagnosis of MSI status among breast cancer patients unlike the previously suggested *BAT-26* which is very polymorphic. Nevertheless, both markers could work effectively as a pair due to their varying polymorphic phenotype in different populations.

### References

- [1] Abida, W., Cheng, M. L. and Armenia, J., Middha, S., Autio, K. A., Vargas, H. A., Rathkopf, D., Morris, M. J., Danila, D. C., Slovin, S. F., Carbone, E., Barnett, E. S., Hullings, M., Hechtman, J. F., Zehir, A., Shia, J., Jonsson, P., Stadler, Z. K., Srinivasan, P., Laudone, V. P., Reuter, V., Wolchok, J. D., Succi, N. D., Taylor, B. S., Berger, M. F., Kantoff, P. W., Sawyers, C. L., Schultz, N., Solit, D. B., Gopalan, A. and Scher, H. I. 2018. Analysis of the prevalence of Microsatellite Instability in Prostate Cancer and Response to Immune Check-point Blockade. *JAMA Oncology*.
- [2] Aclary, S., Wilson, T., Gradia, S., Kame, M. F., Guerrette, S., Marsischky, G. T., Kolodner, R., and Fishel, R. 1996. hMSH2 forms specific mismatch binding complexes with hMSH3 and hMSH6. *Proceedings of the National Academy of Sciences of the United States of America*. 93: 13629–13634
- [3] Akiyama OY, Sato H, Yamada T, Nagasaki H, Tsuchiya A, Abe R, Yuasa Y. 1997. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. *Cancer Research* 57:3920–3923
- [4] Anbazhagan R, Fujii H, Gabrielson E. 1999. Microsatellite instability is uncommon in breast cancer. *Clin Cancer Res* 5: 839-844
- [5] Antonarakis, S.E. and the Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. *Hum. Mut.* 11: 1-3
- [6] Arzimanoglou, I. I., Gilbert, F. and Barber, H. R. K. 1998. Microsatellite Instability in solid human tumors. *Cancer: Interdisciplinary International Journal of the American Cancer Society*. 82(10): 1808-1820.
- [7] Bacher J, Flanagan L, Smalley R, et al. 2004. Development of a fluorescent multiplex assay for detection of MSI-High tumors. *Disease Markers*. 20:237-50.
- [8] Belt E, J, te Velde EA, Krijgsman O, et al. 2012. High lymph node yield is related to microsatellite instability in colon cancer. *Annals of Surgical Oncology*. 19:1222–1230.
- [9] Bergthorsson J, T, Egiłsson V, Gudmundsson J, Arason A, Ingvarsson S. 1995. Identification of a breast tumor with micro satellite instability in a potential carrier of the hereditary non-polyposis colon cancer trait. *Clin Genet* 47: 305–310.
- [10] Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S. 1998. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Research* 58:5248–5257.
- [11] Boland, C. R. and Goel, A. 2010. Microsatellite instability in colorectal cancer. *Gastroenterology*. 138(6): 2073-2087.
- [12] Bonneville, R., Krook, M. A., Kautto, E. A., Miya, J., Wing, M. R., Chen, H.-Z., Reeser, J. W., Yu, L. and Roychowdhury, S. 2017. Landscape of Microsatellite Instability Across 39 Cancer Types. *JCO Precision Oncology*.
- [13] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L., Jemal, A. 2018. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*. 0:5-31
- [14] Buermeyer, A. B., Descheines, S. M., Baker, S. M., and Liskay, M. 1999. Mammalian DNA mismatch repair. *Annu. Rev. Genet.* 33: 533–564.
- [15] Bulhard, O et al. 2006. Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J. Clin. Oncol.*, 24, 241–251.
- [16] Campanella, N. C., Berardinelli, G. N., Scapulatempo-Neto, C., Viana, D., Palermo, E. I., Pereira, R., & Reis, R. M. 2013. Optimization of a pentaplex panel for MSI analysis without control DNA in a Brazilian population: correlation with ancestry markers. *European journal of human genetic E/HG*. 22(7), 875-80.
- [17] Chacon, R. D. and Costanzo, M. V. 2010. Triple-negative breast cancer. *Breast Cancer Research*. 10 Suppl 2: S3
- [18] Cuadrado, A., Cardoso, M. and Jouve, N. 2008. Physical organization of simple sequence repeats (SSRs) in Triteaceae: structural, functional and evolutionary implications. *Cytogenet. Genome Res.* 120, 210–219.
- [19] De Bustos, A., Cuadrado, A. and Jouve, N. 2016. Sequencing of long stretches of repetitive DNA. *Scientific Reports* 6: 36665
- [20] De La Cruz, I. M. and Czerniecki, B. J. 2018. Immunotherapy for Breast Cancer is Finally at the Doorstep: Immunotherapy in Breast Cancer. *Annals of Surgical Oncology*. 25:2852-2857
- [21] De Marchis L, Contegiacomo A, D'Amico C, Palmirota R, Pizzi C, Ottini L, Mastranzo P, Figliolini M, Petrella G, Amanti C, Battista P, Bianco AR, Frati L, Cama A, Mariani-Costantini R. 1997. Microsatellite instability is correlated with lymph node-positive breast cancer. *Clin Cancer Res* 3: 241–248.

- [22]. den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., et al. 2016. HGVS Recommendations for the description of sequence variants: 2016 Update. *Hum. Mutat.* 37, 564–569. doi: 10.1002/humu.22981
- [23]. den Dunnen, JT and Antonarakis, SE. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum.Mut.* 15: 7-12
- [24]. Dudley JC, Lin MT, Le DT, et al. 2016. Microsatellite Instability as a Biomarker for PD-1 Blockade. *Clin Cancer Res*; 22: 813-820
- [25]. Duval, A. and Hamelin, R. 2003. Réparation des erreurs de réplication, microsatellites et cancer. *19(1): 55-62.*
- [26]. Ellegren, H. 2000. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*; 16: 551–8.
- [27]. Eveno C, Nemeth J, Soliman H, et al. 2010. Association between a high number of isolated lymph nodes in T1 to T4 N0M0 colorectal cancer and the microsatellite instability phenotype. *Arch Surg.* 145:12–17.
- [28]. Ferlay, J. GLOBOCAN 2018. Lyon: World Health Organization International Agency for Research on Cancer. 2018.
- [29]. Ferlay J, Colombet M., Soerjomataram I, Mathers, C., Parkin, D M., Piñeros, M., Znaor, A. and Bray, F. 2018. Global and Regional Estimates of the Incidence and Mortality for 38 Cancers: GLOBOCAN 2018. Lyon: International Agency for Research on Cancer/World Health Organization.
- [30]. Fujii S, Takeshima Y, Arihiro K, Kaneko M, Inai K. 1998. Microsatellite instability in breast cancers with special reference to patients' age and bilaterality. *Hiroshima J Med Sci* 47: 89–97.
- [31]. Genschel, J., Littman, S. J., Drummond, J. T., and Modrich, P. 1998. Isolation of MutS from human cells and comparison of the mismatch repair specificities of MutS and MutS. *J. Biol. Chem.*, 273: 19895–19901.
- [32]. Geurts-Giele, W. R., Leenen, C. H., Dubbink, H. J., Meijssen, I. C., Post, E., Sleddens, H. F., Kuipers, E. J., Goverde, A., van den Ouweland, A. M., van Lier, M. G., Steyerberg, E. W., van Leerdam, M. E., Wagner, A., Dinjens, W. N. 2014. Somatic aberrations of mismatch repair genes as a cause of microsatellite unstable cancers. *TheJournal of Pathology*, 234: 548–59
- [33]. Glebov OK, McKenzie KE, White CA, Sukumar S. 1994. Frequent p53 gene mutations and novel alleles in familial breast cancer. *Cancer Res* 54: 3703–3709
- [34]. Gryfe R, Kim H, Hsieh ET, et al. 2000. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *N Engl J Med.* 342:69–77.
- [35]. Güler, G., Himmethoglu, C., Jimenez, R. E., Geyer, S. M., Wang, W. P., Costineanu, S., Pilarski, R. T., Morrison, C., Suren, D., Liu, J., Chen, J., Kamal, J., Shapiro, C. L. and Huebner, K. 2011. Aberrant expression of DNA damage response proteins is associated with breast cancer subtype and clinical features. *Breast Cancer Research and Treatment*. 129: 421-432.
- [36]. Hall, T. A. 1999. "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT". *Nucleic Acids Research*, vol. 44, pp. 211-232.
- [37]. Hamelin R<sup>1</sup>, Chalastanis A, Colas C, El Bchiri J, Mercier D, Schreurs AS, Simon V, Svrcek M, Zaanen A, Borie C, Buhard O, Capel E, Zouali H, Praz F, Mulenis M, Fléjou JF, Duval A. 2008. Clinical and molecular consequences of microsatellite instability in human cancers. *Bulletin of Cancer*. 95(1): 121-132.
- [38]. Hatch SB, Lightfoot HM, Garwacki CP, et al. 2005. Microsatellite instability testing in colorectal carcinoma: choice of markers affects sensitivity of detection of mismatch repair-deficient tumors. *Clin. Cancer Res.* 11(6):2180–2187
- [39]. Hoang JM, Cotru PH, Thuille B, Salmon RJ, Thomas G, Hamelin R. 1997. *BAT-26*, an indicator of the replication error phenotype in colorectal cancers and cell lines. *Cancer Res* 57: 300–303.
- [40]. Hu, Z. I., and McArthur, H. L. 2018. Immunotherapy in Breast Cancer: the New Frontier. *Current Breast Cancer Reports*. 10(2): 35-40.
- [41]. Imai, K. and Yamamoto, H. 2008. Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. *Carcinogenesis* 29:673-680.
- [42]. Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. and Perucho, M. 1993. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*. 363(6429): 558-561
- [43]. Juricny, J. Mediating mismatch repair. *Nat. Genet.*, 24: 6–8, 2000.
- [44]. Joly, M. O., Atignon, V., Saurin, J. C., Desseigne, F., Leroux, D., Martin-Denavit, T., Giraud, S., Bonnet-Dupeyron, M. N., Favier, L., Auclair, J., Grand-Masson, C., Audouy, C. and Wang, Q. 2015. Somatic MMR gene mutations as a cause for MSI-H sebaceous neoplasms in Muir-Torre Syndrome-Like patients. *Human Mutation* 36: 292–5
- [45]. Kolodner, R. D., and Marsischky, G. T. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.* 9: 89–96, 1999
- [46]. Laghi, L., Bianchi, P. and Malesci, A. 2008. Differences and evolution of the methods for the assessment of microsatellite instability. *Oncogene*. 27(49):6313–6321.
- [47]. Lamberti, C., Kruse, R., Ruelfs, C., Caspari, C., Wang, Y., Jungck, M., Mathiak, M., Malayeri, H. R. H., Friedl, W., Sauerbruch, T. and Propping, P. 1999. Microsatellite instability – a useful diagnostic tool to select patients at high risk for hereditary non-polyposis colorectal cancer: a study in different groups of patients. *Gut*. 44(6): 839-843
- [48]. Lee S.C., Berg K.B., Sherman M.E., Griffin C.A. & Eshleman J.R. 2001. Microsatellite Instability is frequent in Medullary Breast Cancer. *American Journal of Clinical Pathology*. 115 (6): 823-827.
- [49]. Levinson, G. and Gutman, G. A. 1987. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage ml3 in *Escherichia coli* k-12. *Nucleic Acids Research*; 15: 5323–38.
- [50]. Loeb LA. 1994. Microsatellite instability: marker of a mutator phenotype in cancer. *Cancer Res*54:5059-5063
- [51]. Loeb, K. R. and Loeb, L. A. 1999. Genetic Instability and the Mutator Phenotype Studies in Ulcerative Colitis. *American Journal of Pathology*. 154(6): 1621-1626
- [52]. Loukola, A., Eklun, K., Salovaara, R., Kristo, P., Jarvinen, H., Mecklin, J. P., Launonen, V., Aaltonen, L. A. 2001. Microsatellite marker analysis in screening for hereditary nonpolyposis colorectal cancer (HNPCC). *Cancer Research*. 61(11): 4545-4549
- [53]. Ly, M., Antoine, M., André, F., Callard, P., Bernaudin, J. F. and Diallo, D.A. 2011. Le cancer du sein chez la femme de l'Afrique subsaharienne: Etatactuel des connaissances. *Bulletin du Cancer*.98: 797-806
- [54]. Malesci A, Laghi L, Bianchi P, et al. Reduced likelihood of metastases in patients with microsatellite-unstable colorectal cancer. *Clin. Cancer Res.* 2007;13(13):3831–3839.
- [55]. Mamdouh, S, Metwally, A. M., Aref, A. M., Khaled, H. M., Hamman, O. and Saber, M. A. *BAT-26* is Associated with Clinical Stage and Lymph Node Status in Schistosomiasis Associated Bladder Cancer. *American Journal of Biochemistry*. 5(1): 15-21
- [56]. Mehrgou, A., and Akouchekian, M. 2016. The importance of BRCA1 and BRCA2 genes mutations in breast cancer development. *Medical Journal of the Islamic Republic of Iran*, 30, 369.
- [57]. Merok, M. A., Ahlquist, T., Royrvik, E. C., Tufeland, K. F., Hektoen, M., Sjo, O. H., Mala, T., Svindaland, A., Lothe, R. A. and Nesbakken, A. 2012. Microsatellite instability has a positive prognostic impact on stage II colorectal cancer after complete resection: results from a large, consecutive Norwegian series. *Annals of Oncology*. 24(5): 1274-1282.
- [58]. Modrich, P. and Lahue, R. 1996. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annual Review of Biochemistry*. 65: 101-133
- [59]. Oda, S., Maehara, Y., Ikeda, Y., Oki, E., Egashira, A., Okamura, Y., Takahashi, I., Kakeji, Y., Miyashita, K., Yamada, Y., Zhao, Y., Hattori, H., Taguchi, K., Ikeuchi, T., Tsuzuki, T., Sekiguchi, M., Karran, P. and Yoshida, M. A. 2005. Two modes of microsatellite

- instability in human cancer: differential connection of defective DNA mismatch repair to dinucleotide repeat instability. *Nucleic Acids Research*, 33(5): 1628-1636
- [60]. Patel U, Grundfest-Broniatowski S, Gupta M, Banerjee S. 1994. Microsatellite instabilities at five chromosomes in primary breast tumors. *Oncogene* 9: 3695-3700
- [61]. Paulson TG, Wright FA, Parker BA, Russack V, Wahl GM. 1996. Microsatellite instability correlates with reduced survival and poor disease prognosis in breast cancer. *Cancer Res* 56: 4021-4026
- [62]. Perucho, M. 1999. Correspondence re. Boland CR et al. (1998) A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.*, 58, 5248-5257.
- [63]. Popat, S., Hubner, R. and Houlston, R. S. 2005. Systematic review of microsatellite instability and colorectal cancer prognosis. *Journal of Clinical Oncology* 23: 609-618
- [64]. Prat, A. and Perou C. M. 2011. Deconstructing the molecular portraits of breast cancer. *Molecular Oncology*, 5: 5-23
- [65]. Pyatt, R., Chadwick, R. B., Johnson, C. K., Adebamowo, C., Chapelle, A. and Prior, T. W. 1999. Polymorphic Variation at the *BAT-25* and *BAT-26* Loci in Individuals of African Origin. Implications for Microsatellite Instability Testing. *The American Journal of Pathology*, 155(2): 349-353
- [66]. Ratu, M., Lampis, A., Jens, H. C., Passalacqua, R. and Valeri, N. 2018. Microsatellite instability in gastric cancer: molecular bases, clinical perspectives, and new treatment approaches. *Cell Mol Life Sci*, 75(22): 4151-4162.
- [67]. Richman, S. 2015. Deficient mismatch repair: Read all about it (Review). *International Journal of Oncology*, 47: 1189-1202
- [68]. Risinger JI, Barrett JC, Watson P, Lynch HT, Boyd J. 1996. Molecular genetic evidence of the occurrence of breast cancer as an integral tumor in patients with the hereditary nonpolyposis colorectal carcinoma syndrome. *Cancer* 77: 1836-1843
- [69]. Sapher, D., Hsu, A. and Strauss, B. 1999. Stabilization of the intermediate in frameshift mutation. *Mutation Research*, 423(1-2): 73-77
- [70]. Samowitz, W. S. and Slattery, M. L. 1999. Regional reproducibility of microsatellite instability in sporadic colorectal cancer. *Genes, Chromosomes and Cancer*, 26(2): 106-114
- [71]. Scarpa, A., Cataldo, I. and Salvatore, L. 2016. Microsatellite Instability (MSI) indicates a defective mismatch repair (dMMR) system. ESMO Biomarker Factsheet
- [72]. Schlotterer, C. and Harr, B. 2001. Microsatellite Instability. Encyclopedia of Life Sciences. Nature Publishing Group
- [73]. Shia J, Klimstra DS, Nafa K, et al. 2005. Value of immunohistochemical detection of DNA mismatch repair proteins in predicting germline mutation in hereditary colorectal neoplasms. *Am J Surg Pathol* 29: 96-104.
- [74]. Shia J. 2015. Evolving approach and clinical significance of detecting DNA mismatch repair deficiency in colorectal carcinoma. *Semin Diagn Pathol* 32:352-361.
- [75]. Shokal, U. and Sharma, P. C. 2012. Implication of microsatellite instability in human gastric cancer cancers. *The Indian Journal of Medical Research*, 135(5): 599-613.
- [76]. Siah S-P, Quinn DM. 1999. Microsatellite markers in breast cancer. <http://pmrserve1/microsatellite-markers>
- [77]. Soreide K, Nedrebo BS, Soreide J, A, et al. 2009. Lymph node harvest in colon cancer: influence of microsatellite instability and proximal tumor location. *World J Surg*.
- [78]. Sourvinos G, Kiaris H, Tsikkinis A, Vassilaros S, Spandidos DA. 1997. Microsatellite instability and loss of heterozygosity in primary breast tumors. *Tumour Biol* 18: 157-166
- [79]. Subramanian, S., Mishra, R. K. and Singh, L. 2003. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biology*, 4R13
- [80]. Suraweera N, Duval A, Reperant M, et al. 2002. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* 123:1804-11
- [81]. Thompson J.D., Higgins D.G. and Gibson T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22 (22): 4673-4680
- [82]. Tomita S, Deguchi S, Miyagami T, Muto Y, Tamamoto T, Toda T. 1999. Analyses of microsatellite instability and the transforming growth factor-beta receptor type II gene mutation in sporadic breast cancer and their correlation with clinicopathological features [In Process Citation]. *Breast Cancer Res Treat* 53: 33-39
- [83]. U.S. Food and Drug Administration. 2017. FDA approves first cancer treatment for any solid tumor with a specific genetic feature. <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm560167.htm>. Accessed 25 Dec. 2018.
- [84]. Walsh T, Chappell SA, Shaw JA, Walker RA. 1998. Microsatellite instability in ductal carcinoma in situ of the breast. *J Pathol* 185: 18-24.
- [85]. Walsh, T., Casadei, S., Coats, K. H., Swisher, E., Stray, S. m., Higgins, J., Roach, K. C., Mandell, J., Lee, M. K., Ciernikova, S., Foretova, L., Soucek, P. and King, M. C. 2006. Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *JAMA*, 295: 1379-1388.
- [86]. Ward, R., Meagher, A., Tomlinson, I., O'Conner, T., Norrie, M., Wu, R. and Hawkins, N. 2001. Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. *Gut*, 48(6): 821-829
- [87]. Williams, A. D., Payne, K. K., Posey, A. D., Hill, C., Garcia, J. C., June, C. H. and Tchou, J. 2017. Immunotherapy for Breast Cancer: Current and Future Strategies. *Current Surgery Reports*.
- [88]. Wong YF, Cheung TH, Lo KW, Yim SF, Chan LK, Buhard O, Duval A, Chung TK, Hamelin R. 2006. Detection of microsatellite instability in endometrial cancer: advantages of a panel of five mononucleotide repeats over the National Cancer Institute panel of markers. *Carcinogenesis* 27(5): 951-955
- [89]. Xicola RM, Llor X, Pons E, et al. 2007. Performance of different microsatellite marker panels for detection of mismatch repair-deficient colorectal tumors. *J. Natl Cancer Inst*, 99(3):244-252
- [90]. Yee CJ, Roodi N, Verrier CS, Parl FF. 1994. Microsatellite instability and loss of heterozygosity in breast cancer. *Cancer Res* 54: 1641-1644
- [91]. You, J-F., Buhard, O., Ligtenberg, M. J. L., Kets, C. M., Niessen, R. C., Hofstra, R. M. W., Wagner, A., Dinjens, W. N. M., Colas, C., Lascols, O., Collura, A., Flejou, J-F., Duval, A. and Hamelin, R. 2010. Tumours with loss of MSH6 expression are MSI-H when screened with a pentaplex of five mononucleotide repeats. *British Journal of Cancer*, 103: 1840-1845
- [92]. Zhang X, Li J. 2013. Era of universal testing of microsatellite instability in colorectal cancer. *World J Gastrointest Oncol* 5: 12-19
- [93]. Zhou XP, Hoang JM, Li YJ, Seruca R, Carneiro F, Sobrinho-Simoes M, Løthe RA, Gleeson CM, Russell SE, Muzeau F, Flejou JF, Hoang-Xuan K, Lidereau R, Thomas G, Hamelin R. 1998. Determination of the replication error phenotype in human tumors without the requirement for matching normal DNA by analysis of mononucleotide repeat microsatellites. *Genes Chromosomes Cancer* 21: 101-107

- [98]. Zhu, L., Li, Z., Wang, Y., Zhang, C., Liu, Y. and Qu, X. 2015. Microsatellite Instability and Survival in gastric cancer: A systemic review and meta-analysis. *Mol. Clin. Oncol.* 3(3): 699-705

Aniefiok John Udoakang. "Microsatellite Instability of Bat-26 and Nr-24among Breast Cancer Patients in Senegal." IOSR Journal of Dental and Medical Sciences (IOSR-JDMS), vol. 18, no. 1, 2019, pp 71-78.

**Manuscripts under review**

**MANUSCRIPT 1 (Under Review): PAN-AFRICAN MEDICAL JOURNAL**

**Title of the Manuscript:**

SURVEY OF BREAST CANCER PATIENTS IN NIGERIA AND SENEGAL

**Authors full Names:**

*Aniefiok John UDOAKANG<sup>1</sup>, & Fatimata MBAYE<sup>1</sup>, Bineta KENEME<sup>1</sup>, Adeyinka ADEMOLA<sup>2</sup>, Ganiyu Olatunbosun ARINOLA<sup>3</sup>, Mbacke SEMBENE<sup>1</sup>, Temidayo Olusade OGUNDIRAN<sup>2</sup>.*

**Institutions of Origin:**

<sup>1</sup>*Equipe Génétique et Gestion pour les Populations, Département de Biologie Animale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop de Dakar, BP. Box 5005 Dakar, Sénégal.*

<sup>2</sup>*Department of Surgery, College of Medicine, University of Ibadan, Ibadan, Nigeria.*

<sup>3</sup>*Department of Chemical Pathology and Immunology, College of Medicine, University of Ibadan, Ibadan, Nigeria.*

**&Corresponding author:** Aniefiok John UDOAKANG, Equipe Génétique et Gestion pour les Populations, Département de Biologie Animale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop de Dakar, BP. Box 5005 Dakar, Sénégal.

Telephone: +221778215047

Email: [mandith2004@yahoo.com](mailto:mandith2004@yahoo.com); [aniefiokjohn.udoakang@ucad.edu.sn](mailto:aniefiokjohn.udoakang@ucad.edu.sn))

**Word Count:**

Abstract: 246

Main Text: 2,556

## ABSTRACT

### Background

Despite the increasing incidence and research on breast cancer in sub-Saharan Africa, few studies have been performed to compare risk factors for breast cancer among patients in the region. In this study, we examined the risk factors and clinical characteristics of breast cancer patients in Nigeria and Senegal.

### Methods

This was a case-only analysis of eligible females between 15 and 80 years who were diagnosed with breast cancer. In-person interviews were conducted using structured questionnaires to collect demographic and clinical parameters from 201 consecutive patients in Nigeria and Senegal.

### Results

The mean ( $\pm$  standard deviation) age of the cases from Nigeria and Senegal was 49.1( $\pm$ 12.4) and 46.0( $\pm$ 14.0) years respectively ( $p=0.244$ ), peak age was between 41 to 50 years for both countries. Most Nigerian (89.4%) and Senegalese (79.5%) patients presented at stages III and IV. Histological type ( $p=0.395$ ), Scarff-Bloom-Richardson grade ( $p=0.237$ ) and stage ( $p=0.204$ ) did not vary significantly between the two countries. Although the trend was in the same direction, the magnitude of some reproductive risk factors and marital status differed between the two populations: age at menarche ( $p<0.001$ ), age at first full term pregnancy ( $p<0.001$ ), parity ( $p<0.001$ ), contraceptive use ( $p=0.001$ ), and marital status ( $p=0.005$ ).

### Conclusion

The findings from this study demonstrate variable magnitude of differences in socio-demographic and reproductive characteristics between breast cancer patients in Nigeria and Senegal. We propose more studies, enlarged to include un-researched parts of Africa, encompassing clinical, environmental and biological parameters to fully unravel breast cancer and control its growing menace in the African continent.

**Key words:** Breast cancer, Risk factors, Nigeria, Senegal

### Background

Breast cancer is the most common cancer and the leading cause of cancer related deaths in women worldwide [1]. While the five-year survival rate of breast cancer in the United States is 86%, it is less than 40% in sub-Saharan Africa. In West Africa, breast cancer accounts for 35%

of new cancer cases and is the leading cause of cancer deaths; 28% among women [2]. Although similarities have been observed among risk factors in different populations, a difference in tumour features: higher grade, advanced stage, younger age; outcome: poor clinical and pathological prognosis and survival have been observed in black compared to white Caucasian populations. The better outcome in Caucasian populations have been attributed to demographic factors such as longer life expectancy, increased awareness, improved system of reporting the disease and access to better health services [3-6].

Breast cancer is the commonest malignancy among Nigerian women and is second to cervical cancer among women in Senegal [7-9]. Secular changes in lifestyles, hormonal and reproductive factors have been associated with increasing incidence in developing countries like Nigeria and Senegal [10,11]. Some known risk factors identified in different racial groups and populations include age at menarche, age at first childbirth, age at menopause, breastfeeding, parity and family history of breast cancer [12-14].

In sub-Saharan Africa, there have been less than model measures for the prevention and control of breast cancer but rather some improvement in early detection and treatment [15,16]. Population specific measures to control and prevent the disease as well as measures for early detection in the younger population who often present at a late stage are lacking. In spite of the relative increase in breast cancer research in Nigeria and some countries of Africa, more studies are needed to understand the distribution of risk factors for pathogenesis of breast cancer in different regions of the continent. This will aid in the development of effective population-based practices to control its incidence and increase preventive strategies in the African continent [17-19]. The aim of this study was to compare some reproductive risk factors and clinical parameters of breast cancer among women of all age groups with the disease in Nigeria and Senegal.

## **Methods**

All consecutive female breast cancer patients of all ages attending the Surgical Oncology and Radiotherapy clinics of the University College Hospital (UCH), Ibadan, Nigeria between November, 2016 and July, 2017 were approached to participate in the study. Most of the patients who were eligible agreed to participate in the study with a refusal rate of one per cent. In Senegal, all consecutive female breast cancer patients of all ages who attended the Institut Joliot-Curie Cancer Centre, Dakar, Hopital Aristide le Dantec, Dakar, Senegal between 2014 and 2015 were approached for the study. Majority of them agreed to participate in the study with a refusal rate of 4%. Both hospitals are major referral hospitals where most breast cancer patients from these countries are referred to for treatment. This study was approved by the

Institutional Review Board of the University of Ibadan/University College Hospital, Ibadan, Nigeria (HREC Assigned Protocol number: UI/EC/16/0334) and the Universite Cheikh Anta Diop de Dakar, Senegal Research Ethics Committee (Protocole: 0269/2015/CER/UCAD).

All the patients who presented with a clinical diagnosis of breast cancer at the study sites in Nigeria and Senegal were eligible for the study. The study objectives, protocol, risks, benefits, confidentiality and rights of participants were explained to all potential participants and each person had an option to refuse participation in the study. Opportunity was granted for participants to ask questions and those that were willing to participate signed a written informed consent after which a structured questionnaire was administered by two of the investigators (Aniefiok John Udoakang for Nigeria) and (Fatimata Mbaye for Senegal). Study participants were recruited and interviewed on the same day that they presented at the oncology clinic with a clinical diagnosis of breast cancer. Tissue confirmation of breast cancer was made by incision, excision, and fine-needle or core-needle biopsy and from the result; any participant who had a diagnosis other than breast cancer was removed from the study.

The questionnaire covered demographic factors such as age, age at diagnosis, marital status, family history of breast cancer and other cancers in first- and second-degree relatives and ethnicity. Reproductive variables included age at menarche, pregnancy [number of pregnancies, age at first full term pregnancy (FFTP)], childbirth history including parity and breastfeeding and menopausal status. Information on tumour characteristics was extracted from participants' case notes and pathology report which included date of diagnosis, tumour site and tumour size, histology, Tumour Node Metastases (TNM) stage and Scarff-Bloom-Richardson (SBR) grade. Data for the tumour characteristics were available for participants as follows: Histological type (200), SBR grade (163), and TNM stage (188). In this study, the reproductive characteristics examined were age at menarche, number of pregnancies, age at first full-term pregnancy, breastfeeding history and menopausal status. Age at menarche was considered in four categories: 10 - 12, 13 - 15, 16 - 18 and  $\geq 19$ ; number of pregnancies was categorized into five: 0, 1 - 3, 4 - 6, 7 - 9 and  $\geq 10$ ; age at FFTP:  $\leq 20$  years, 21 - 30 years, 31 - 40 years and  $\geq 41$ . Breastfeeding history and menopausal status were categorized as: ever breastfed (Yes/No) and Have you reached menopause? (Yes/No). The cohort was later divided into two groups: younger women  $\leq 50$  years and older women  $> 50$  years to compare the prevalence of breast cancer among women in these two age groups in Nigeria and Senegal.

Continuous variables were analysed using mean and standard deviations and the Independent sample t-tests to compare the means of the two groups while categorical variables were

described using absolute frequencies and percentages and the chi-square test to compare the proportions. The data was analysed using SPSS version 20.0.

## **Results**

Two hundred and one patients who had histologically confirmed breast cancer and agreed to participate in the study were included: One hundred and two patients from Nigeria and ninety-nine patients from Senegal. One hundred and fifty-four cases (85% of the patients) presented at an advanced stage (Stages III and IV): 84 cases, (90%) of the patients from Nigeria and 70 cases, (80%) of patients from Senegal. The most common breast cancer histological type among the participants from both countries was invasive/infiltrating ductal carcinoma (88.1%). Others included: lobular carcinoma, phylloides and sarcoma and (Table 1) shows the comparison of clinical characteristics of breast cancer cases in Nigeria and Senegal. Age at diagnosis, menopausal status, breastfeeding and family history of cancer were not statistically significant between Nigerian and Senegalese patients. The mean age at diagnosis for participants from Nigeria was 49.1(12.4) and ranged from 29 to 80 years while the mean age for participants from Senegal was 46.9(14.0) and ranged between 15 and 80 years ( $p = 0.244$ ). Mean parity was 3.7(2.1) for Nigerian patients and 4.5(3.1) for Senegalese patients which differed significantly ( $p = 0.044$ ) between the two populations. 88.7 per cent of the study participants were parous women of which 71.5% had at least 3 live births. Contraceptive use was more prevalent among Nigerian than Senegalese patients ( $p = 0.001$ ). There was significant difference in age at menarche ( $p < 0.001$ ) and age at first full term pregnancy ( $p < 0.001$ ) as the parameters were more or less profound in one country than the other. Table 2 shows the comparison of selected characteristics of breast cancer cases in Nigeria and Senegal.

## **Discussion**

Breast cancer is one of the most studied neoplasm of which epidemiological studies have revealed marked geographical variation in its burden worldwide. While breast cancer incidence and mortality rates remain highest in industrialized countries, it is rapidly increasing in developing countries such as West Africa. Although there have been several studies on breast cancer in Nigeria and Senegal [20,21], paucity of data exists that compare breast cancer risk factors between the two population. Some apparent differences, mainly in magnitude, were observed in some of the factors that were analysed from the study participants in the two populations. They were: age at menarche, age at first full term pregnancy, parity, contraceptive use and marital status.

The young mean age at diagnosis in Nigerian and Senegalese patients have been observed in other studies from within and outside Africa and has been identified as one of the

characteristics of breast cancer in the studied populations compared to other populations [22-24]. Fifty-nine per cent of patients from Nigeria and 65% of patients from Senegal were 50 years and below. The ability to make definite conclusions regarding breast cancer risk in older women is slightly hindered due to the relatively short life expectancy observed in these countries (49.1 years in Nigeria and 60.7 years in Senegal) resulting in a large population of young people [25, 26]. Breast cancer case ascertainment is low among older African women because of their lower literacy rates, awareness of breast cancer and poor socioeconomic status [27].

Age at menarche differed significantly in the patients in both countries ( $p = 0.001$ ) with mean age of 15.3(1.9) for Nigerian and 14.3(1.3) years in Senegal. Age at menarche has an inverse association with breast cancer as early menarcheal age has been linked with early onset of breast cancer and late age at menarche has been consistently associated with decreased risk [28].

The age-incidence curves of breast cancer in Africa could be a function of the population-based risks lifestyles of pre-industrialized societies in which there is increased incidence of premenopausal breast cancer [29].

Menopausal status was not different between the two populations and about two-thirds of the studied population were premenopausal at diagnosis. The predominance of breast cancer diagnosis in premenopausal women have been observed in studies around Africa. Early menarche and late menopause are positively correlated with breast cancer risk as individuals are exposed to circulating female hormones for a prolonged period of time as observed in this study with early age at menarche [30,31].

Although the type of contraceptive used was not defined in our study, the use of contraceptive was more prevalent in the Nigerian patients. About a third of the participants indicated that they had used contraceptives during their lifetime. The use of hormonal contraceptives is increasing in Africa and in case-control studies, more breast cancer study participants than controls use oral/hormonal contraceptives. This has been identified as a risk factor for breast cancer in different populations especially at later years of a woman's reproductive life [32, 33] and an increased risk of 24.8% in cases who had used oral contraceptives for more than 6 months [34].

Almost all the patients in the two cohorts (all Senegalese and over 90% Nigerians) had their first full term pregnancy below 30 years of age in spite of the difference in magnitude between the two countries. Studies in the past showed that the lifetime risk of breast cancer is lower in women that have multiple pregnancies with young age at first full term pregnancy, but a slight

increase in breast cancer risk had been observed during the six to twelve months postpartum period which may last for about three to four years [35-37]. This temporary increase risk was attributable to the lack of protective effects of parity in premenopausal women diagnosed under the age of 50 [38-46]. However, later studies among indigenous Africans, including women from the same cohort as for our present study, found no difference in breast cancer risk with an increasing age [47,48]. These differences in reproductive patterns may perhaps be due to genetic factors or variations in primary tumour biology and tumour aggressiveness, suggesting that treatment pattern may differ for different geographical populations.

Multiparity has been identified with low breast cancer risk and African women begin child bearing at relatively younger age. This protective role of parity increases in proportion with the number of children and the early age at first live birth [49]. Result from a collaborative reanalysis of 47 epidemiologic studies showed that in the absence of breastfeeding, each birth reduces the relative risk of breast cancer by 7% with a decreased risk of 3.4% for each child that is breastfed [50]. The majority of our study participants were parous women and 71.5 per cent had at least 3 children. Parity differed between the two populations and most patients from the two populations had at least three live births. Breastfeeding and family history of cancer were similar in both countries as 96.6 per cent of the population breastfed their children and 87.1 per cent had no family history of cancer. Marital status was different between the two countries as more Senegalese women were never married.

Most of the study participants were diagnosed with invasive/infiltrating carcinoma, and there was no difference in tumour stage and grade between the countries. Late stage presentation was common to both countries and this is not unexpected given the limited access to health care and poor health infrastructure in these countries.

The limitations of this work include small sample size and its case-only nature. While trends may be feasible, inferences and cause-effect associations cannot be made. The apparent differences in some parameter between the two countries may be due to selection (as per hospital-based study) since the trend pointed in the same direction in most of the studied parameters in the two cohorts. Therefore, we acknowledge the possibility of under diagnosis of the disease and the effect of the selection bias of cases in the study populations. Although reproductive data collection was self-reported and would be subject to recall bias, the validity of self-reported reproductive history is high. Also, there may be low level of discrepancy in the report of parity and menopausal status, systematic error due to recall bias for menarche and age at first live birth may exist but presumably not differentially recalled by the patients. Since all participants already had breast cancer at the time of the interview, it is unlikely that there

would be differential bias for case ascertainment as seen in case-control. In spite of these however, and in cognizance of previous work in one of the two countries and in other similar settings, the findings represent a true reflection of breast cancer in most of the continent and provide baseline information for future studies, especially in the less researched of the two populations.

### **Conclusion**

In this study, we have documented similarities and seeming differences in some known reproductive risk factors for breast cancer and in clinical characteristics between breast cancer patients in Nigeria and Senegal. Age at diagnosis, cancer stage at presentation and tumour characteristics were similar among the patients in the two countries. We observed subtle disparities in age at menarche, age at first full term pregnancy, parity, marital status and contraceptive use although the trend pointed in the same direction. Due to the young age at diagnosis especially in premenopausal women and presentation at an advanced stage, we are of the opinion that introduction of organized population-based screening programs targeting this category of women would improve the outlook and the outcome of breast cancer in them. This study has important clinical and policy implications for the control and prevention of breast cancer in the West African sub-region and the whole of Africa. However, more studies, enlarged to include un-researched parts of Africa, and encompassing clinical, environmental and biological parameters are needed to fully unravel breast cancer and control its growing menace in the African continent.

### **What is known about this topic?**

- The incidence of breast cancer is increasing in sub-Saharan Africa and several reproductive factors have been associated with low or high risk of breast cancer.
- Similarities exist for breast cancer risk factors in different populations in Africa.

### **What this study adds?**

- That there exist some subtle disparities in risk factors for breast cancer in Nigeria and Senegal although the trend pointed in the same direction.
- That there should be population-based studies that compare breast cancer risk factors especially in un-research parts of Africa to allow for a full understanding of the disease in the sub-region.

### **Acknowledgements**

The authors are indebted to all the study participants for providing the data for the study by their agreement to participate in this study. We acknowledge the Economic Community of

West African States (ECOWAS) department of education, science and culture for awarding AJU scholarship for her PhD.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

AJU, FM, TOO, AA and GOA were involved in the acquisition of the data. AJU cleaned the data, ran the statistical analysis, interpreted the data and drafted the manuscript. FM, BK, MS, TOO and GOA revised the manuscript. MS, FM and TOO supervised the manuscript preparation. MS and FM conceived the study and supervised the statistical analysis. All the authors contributed to the design of the study, read and approved the final manuscript.

### References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, *et al.* GLOBOCAN 2013 v1.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 11. Lyon, France: International Agency for Research on Cancer; 2013. Available from: <http://globocan.iarc.fr>, Accessed on 01 Dec 2017
2. Cumber S, Nchanji K, Tsoka-Gwegweni J. Breast cancer among women in sub-Saharan Africa: prevalence and a situational analysis. *Southern African Journal of Gynaecological Oncology*. 2017. 1(1):1–3. doi: [10.1080/20742835.2017.1391467](https://doi.org/10.1080/20742835.2017.1391467)
3. American Cancer Society. Global Cancer Facts & Figures 3rd Edition. Atlanta: American Cancer Society. 2015
4. Adebamowo C, Ogundiran T, Adenipekun A, Oyeseun R, Campbell O, Akang E, *et al.* Waist-hip ratio and breast cancer risk in urbanized Nigerian women. *Breast Cancer Research*. 2003. 5(2): 18 - 24
5. Brinton L, Figueroa JD, Awuah B, Yarney J, Wiafe S, Wood S, *et al.* Breast Cancer in Sub-Saharan Africa: Opportunities for Prevention. *Breast Cancer Research Treatment*. 2014. 144(3): 467–478. doi:10.1007/s10549-014-28
6. Breyer J, Wendland E, Kops N, Calef M, Hammes L. Assessment of potential risk factors for breast cancer in a population in Southern Brazil. *Breast Cancer Research and Treatment*. 2018. <https://doi.org/10.1007/s10549-017-4655-0>
7. Adebamowo CA, Ajayi OO. Breast cancer in Nigeria. *West African Journal of Medicine*. 2000. 19 :179-191.
8. Dem A, Traore B, Dieng M, Ouajdi T, Lalani M, Diop M, *et al.* Gynaecological and breast cancers at the Dakar Institute. *Cah. Sant*. 2008. 18(1): 25-29
9. Guèye M, Guèye SMK, Mbaye M, Dia FN, Diémé MEF, Niang MM, *et al.* Clinical Features and Prognosis of Triple Negative Breast Cancer at the Senology Unit of Aristide-Le-Dantec Teaching Hospital. *Journal Africain du Cancer*. 2013. 5, 42-47.
10. Gueye M, Kane GSM, Ndiaye GMD, Gueye L, Moreau JC. Profile of black woman in Senegal with breast cancer. *Med Sante Trop*. 2016. 26(2):165-9
11. Ogundiran T, Huo D, Adenipekun A, Campbell O, Oyeseun R, Akang E, *et al.* Case-Control Study of Body Size and Breast Cancer Risk in Nigerian Women. *American Journal of Epidemiology*. 2010. 172(6):682–690
12. Adebamowo CA, Adekunle OO. Case-controlled study of the epidemiological risk factors for breast cancer in Nigeria. *British Journal of Surgery*. 1999. 86(5):665–668

13. Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, *et al.* Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *Journal of National Cancer Institute.* 2005. 16; 97(6):439-48.
14. Krieger N, Chen JT, Waterman PD. Decline in US breast cancer rates after the Women's Health Initiative: socioeconomic and racial/ethnic differentials. *Am J Public Health.* 2010. 100 Suppl 1: S132-9.
15. World Health Organization. Cancer control: knowledge into action: WHO guide for effective programmes: early detection. 2007.
16. Yip CH, Smith RA, Anderson BO, Miller AB, Thomas, DB, Caffarella RS, *et al.* Guideline implementation for breast healthcare in low- and middle-income countries: Early detection resource allocation. *Cancer.* 2008. 113(suppl):2244–2256.
17. Okobia M, Bunker C, Zmuda J, Kammerer C, Vogel V, Uche E, *et al.* Case – control study of risk factors for breast cancer in Nigerian women. *International Journal of Cancer.* 2006. 119:2179– 2185
18. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C. Rebelo M. *et al.* Cancer incidence and mortality worldwide: Sources methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer.* 2015. 136:359–386. doi: 10.1002/ijc.29210.
19. Willett WC, Tamimi R, Hankinson SE, Hazra A, Eliassen AH, Colditz GA. Chapter 18: Non-genetic Factors in the Causation of Breast Cancer, in Harris JR, Lippman ME, Morrow M, Osborne CK. *Diseases of the Breast.* 2014. 5th edition, Lippincott Williams & Wilkins
20. Huo D, Adebamowo CA, Ogundiran TO, Akang EE, Campbell O, Adenipekun A, *et al.* Parity and breastfeeding are protective against breast cancer in Nigerian women. *Br J Cancer.* 2008. 11; 98(5):992-6. doi: 10.1038/sj.bjc.6604275
21. Gueye Mamour, Kane-Gueye Serigne, Ndiaye-Gueye Mame Diarra, Gassama Omar, Diallo Moussa, Moreau Jean. Inflammatory breast cancer: features and outcomes in a breast unit in Dakar, Senegal. *International Journal of Reproduction, Contraception, Obstetrics and Gynecology.* 2016 Feb; 5(2):361-366
22. Saghir N, Shamseddin A, Geara F, Bikhazi K, Rahal B, Salem Z, *et al.* Age Distribution of Breast Cancer in Lebanon: Increased Percentages and Age Adjusted Incidence Rates of Younger-Aged Groups at Presentation. *Le Journal médical libanais. The Lebanese medical journal* 2002. 50(1-2):3-9
23. Ohene-Yeboah M, Adjei E. Breast cancer in Kumasi, Ghana. *Ghana Medical Journal.* 2012. 46(1):8-13
24. Anders CK, Hsu DS, Broadwater G, Acharya CR, Foekens JA, Zhang Y, Wang Y, Marcom PK, Marks JR, Febbo PG, Nevins JR, Potti A, Blackwell KL. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. *J Clin Oncol.* 2008; 26:3324–3330. doi: 10.1200/JCO.2007.14.2471.
25. Huo D, Ikpatt F, Khramtsov A, Dangou J, Nanda R, Dignam J, *et al.* Population Differences in Breast Cancer: Survey in Indigenous African Women Reveals Over-Representation of Triple-Negative Breast Cancer. *Journal of Clinical Oncology.* 2009. 27:4515-4521
26. Ntekim A, Nufu F, Campbell, O. Breast cancer in young women in Ibadan Nigeria. *African health sciences.* 2009. 9(4): 242-246.
27. Amir H, Kitinya JN, Parkin DM. A comparative study of carcinoma of the breast in an African population. *East Afr Med J.* 1994. 71:215–218
28. Clavel-Chapelon F, Gerber M. Reproductive factors and breast cancer risk. Do they differ according to age at diagnosis? *Breast Cancer Res Treat,* 2002. 72 (2):107-15

29. Fregene A, Newman LA. Breast cancer in sub-Saharan Africa: how does it relate to breast cancer in African-American women? *Cancer*. 2005. 103(8):1540-50
30. Molah Karim SA, Ali Ghalib HH, Mohammed SA, Rahim Fattah FH. The incidence, age at diagnosis of breast cancer in the Iraqi Kurdish population and comparison to some other countries of Middle-East and West. *International Journal of Surgery*. 2015. 13: 71-75
31. Lopez AD, Ahmad OB, Guillot M, Ferguson BD, Salomon JA, Murray CJL, *et al.* *World Mortality in 2000: Life Tables for 191 Countries*. Geneva: World Health Organization 2002.
32. Collaborative Groups on Hormonal Factors in Breast Cancer. Breast Cancer and Hormonal Contraceptives: Collaborative Reanalysis of Individual Data on 53,297 Women with Breast Cancer and 100,239 Women without Breast Cancer from 54 Epidemiological Studies. *Lancet*, 1996. 347:1713-1727
33. Nkondjock A, Ghadirian P. Facteurs de risque du cancer du sein. *Med Sci*. 2005. 21: 175-80
34. Laamiri FZ, Bouayad A, Hasswane N, Ahid S, Mrabet M, Amina, B. Risk Factors for Breast Cancer of Different Age Groups: Moroccan Data? *Open Journal of Obstetrics and Gynecology*. 2015. 5:79-87.
35. Wohlfahrt J, Andersen PK, Mouridsen HT, Adami HO, Melbye M. Reproductive history and stage of breast cancer. *American Journal of Epidemiology*. 1999. 150:1325–1330
36. Pathak DR, Osuch JR, He J. Breast carcinoma etiology: current knowledge and new insights into the effects of reproductive and hormonal risk factors in black and white populations. *Cancer*. 2000. 88 (5 Suppl):1230 –1238
37. Althuis MD, Fergenbaum JH, Garcia-Closas M, Brinton LA, Madigan MP, *et al.*, Etiology of hormone receptor-defined breast cancer: a systematic review of the literature. *Cancer Epidemiol Biomarkers Prev*. 2004. 13: 1558 –1568
38. Miller AB, Bulbrook RD. UICC Multidisciplinary Project on Breast Cancer: the epidemiology, etiology and prevention of breast cancer. *International Journal of Cancer*. 1986. 37:173–7.
39. Bruzzi P, Negri E, La Vecchia C, Decarli A, Palli D, Parazzini F, *et al.* Short-term increase in risk of breast cancer after full term pregnancy. *BMJ*. 1988. 297:1096–8.
40. Williams EM, Jones L, Vessey MP, McPherson K. Short term increase in risk of breast cancer associated with full term pregnancy. *BMJ*. 1990. 300: 578–9.
41. Miller WR. Hormonal factors and risk of breast cancer. *Lancet*. 1993. 341:25–6.
42. Lambe M, Hsieh C, Trichopoulos D, Ekbohm A, Pavia M, Adami H. Transient increase in the risk of breast cancer after giving birth. *N Engl J Med*. 1994. 331:5–9.
43. Albrektsen G, Heuch I, Kvale G. The short-term and long-term effect of a pregnancy on breast cancer risk: a prospective study of 802,457 parous Norwegian women. *Br J Cancer* 1995. 72: 480–4.
44. Leon DA, Carpenter LM, Broeders MJ, Gannarskog MJ, Murphy MF. Breast cancer in Swedish women before age 50: evidence of a dual effect of completed pregnancy. *Cancer Causes Control*. 1995. 6:283–91.
45. Liu Q, Wu J, Lambe M, Hsieh SF, Ekbohm A, Hsieh CC. Transient increase in breast cancer risk after giving birth: postpartum period with the highest risk. *Cancer Causes Control*. 2002. 13:299–305
46. Palmer J, Wise LA, Horton NJ, Adams-Campbell LL, Rosenberg L. Dual Effect of Parity on Breast Cancer Risk in African-American Women. *Journal of the National Cancer Institute*. 2003. 95:6
47. Sighoko D, Ogundiran T, Ademola A, Adebamowo C, Chen L, Odedina S, Anetor I, Ndom P, Gakwaya A, Ojengbede O, Huo D, Olopade OI. Breast cancer risk after full-

- term pregnancies among African women from Nigeria, Cameroon, and Uganda. *Cancer*. 2015 Jul 1; 121(13):2237-43
48. Ewertz M, Duffy SW, Adami HO, Kvåle G, Lund E, Meirik O, *et al.* Age at First Birth Parity and Risk of Breast Cancer: A Meta-Analysis of 8 Studies from Nordic Countries. *International Journal of Cancer*, 1990. 46:597-603  
195
49. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet*. 2002. 360: 187 –
50. Hinkula M, Pukkala E, Kyyronen P, Kauppila, A. Grand Multiparity and the Risk of Breast Cancer: Population-Based Study in Finland. *Cancer Causes Control*. 2001. 12:491-500

## TABLES

<b>Table 1: Comparison of Clinical Characteristics of Breast Cancer cases in Nigeria and Senegal</b>					
<b>Characteristics</b>	<b>All cases</b>	<b>Nigeria</b>	<b>Senegal</b>	$\chi^2$	<i>p</i> - value
	<b>No.: (%)</b>	<b>No.: (%)</b>	<b>No.: (%)</b>		
<b>Histological Type</b>					
Invasive/Infiltrating carcinoma	195 (97)	100 (98)	95 (96)	4.084	0.395
Invasive lobular carcinoma	2 (1.0)	-	2 (2.0)		
Lobular carcinoma In-situ	1 (0.5)	1 (1.0)	-		
<b>Others</b>					
Phylloides tumour	2 (1.0)	1 (1.0)	1(1.0)	2.596	0.273
Sarcoma	1 (0.5)	-	1 (1.0)		
<b>SBR Grade</b>					
Well differentiated	29 (17.9)	20 (22.0)	9 (12.7)	4.596	0.204
Moderately Differentiated	95 (58.6)	52 (57.1)	43 (60.6)		
Poorly differentiated	38 (23.5)	19 (20.9)	19 (26.8)		
<b>Stage</b>					
1	8 (4.4)	3(3.2)	5 (5.7)	4.596	0.204
2	20 (11.0)	7 (7.4)	13 (14.8)		
3	91 (50)	53 (56.4)	38 (43.2)		
4	63 (34.6)	31 (33.0)	32 (34.6)		
Clinical characteristics were the same for both countries					

<b>Table 2: Comparison of Selected characteristics of breast cancer cases in Nigeria and Senegal</b>					
<b>Characteristics</b>	<b>All Cases (No= 201)</b>	<b>Nigeria No= 102</b>	<b>Senegal No.=99</b>	<b><math>\chi^2</math></b>	<b><i>p</i> - value</b>
Age at diagnosis, No. (%)					
≤50 y	124(62)	60(58.8)	64(65.3)	0.8915	0.345
>50 y	76(38)	42(41.2)	34(34.7)		
Age at Menarche, No. (%)				20.59	<0.001**
10-12 y	9 (5.1)	5(5.3)	4(4.8)		
13-15 y	122 (68.5)	52(54.7)	70(84.3)		
16-18 y	43 (24.2)	34(35.8)	9 (10.8)		
≥19 y	4 (2.2)	4 (4.2)	0		
Ever Married, No. (%)				7.908	0.005*
Yes	165(87.3)	85(94.4)	80(80.8)		
No	24(12.7)	5(5.6)	19(19.2)		
Age at FFTP, No. (%)				19.409	<0.001**
≤20 y	71(49.7)	33(36.3)	38(73.1)		
21-30 y	64(44.8)	50(54.9)	14(26.9)		
31-40 y	7(4.9)	7(7.7)	0		
≥41 y	1(0.7)	1(1.1)	0		
Parity, No. (%)				35.450	<0.001**
0	22(11.3)	4(4.1)	18(18.8)		
1-3	46(23.7)	27(27.6)	22(19.8)		
4-6	72(37.1)	51(52.0)	21(21.9)		
7-9	36(18.6)	14(14.3)	23(22.9)		
≥10	18(9.3)	2(2.0)	18(16.7)		
Ever breastfed, No. (%)				0.533	0.466
Yes	171(96.6)	85(95.5)	86(97.7)		
No	6(3.4)	4(4.5)	2(2.3)		
Contraceptive use, No. (%)				11.1956	0.001*
Yes	41(21.4)	30(31.3)	11(11.5)		

No	151(78.6)	66(68.8)	85(88.5)		
Menopausal status, No. (%)					
Premenopausal	112(57.1)	57(58.2)	55(56.1)	0.0833	0.773
Menopausal	84(42.9)	41(41.8)	43(43.9)		
Family history of cancer, No. (%)					
Yes	22(12.9)	11(11.2)	11(14.6)	0.0007	0.980
No	173(87.1)	87(88.8)	86(85.4)		
The significance at $p < 0.05$ also exist for the values when computed as continuous variables					

## MANUSCRIPT II: Pre-Submission Enquiry

### DIVERSITY AND GENETIC EVOLUTION OF BREAST CANCER AMONG NIGERIAN AND SENEGALESE WOMEN

#### Authors full Names:

Aniefiok John UDOAKANG<sup>1,&</sup>, Fatimata MBAYE<sup>1</sup>, Adeyinka ADEMOLA<sup>2</sup>, Ganiyu Olatunbosun ARINOLA<sup>3</sup>, Temidayo Olusade OGUNDIRAN<sup>2</sup>, Mbacke SEMBENE<sup>1</sup>.

#### Institutions of Origin:

<sup>1</sup>*Equipe Génétique et Gestion pour les Populations, Département de Biologie Animale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop de Dakar, BP. Box 5005 Dakar, Sénégal.*

<sup>2</sup>*Department of Surgery, College of Medicine, University of Ibadan, Ibadan, Nigeria.*

<sup>3</sup>*Department of Chemical Pathology and Immunology, College of Medicine, University of Ibadan, Ibadan, Nigeria.*

**&Corresponding author:** Aniefiok John UDOAKANG, Equipe Génétique pour la Gestion des Populations, Département de Biologie Animale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop de Dakar, BP. Box 5005 Dakar, Sénégal.

Telephone: +221778215047

Email: [mandith2004@yahoo.com](mailto:mandith2004@yahoo.com); [aniefiokjohn.udoakang@ucad.edu.sn](mailto:aniefiokjohn.udoakang@ucad.edu.sn))

#### ABSTRACT

High tumour clones, a common feature of triple negative breast cancer which is prevalent in West Africa can result from genomic instability and although there are several therapies for treatment, there have been the constant challenge of resistance, relapse, tumour growth and expansion as well as defining the patient population who will most benefit from these treatments. To detect and compare the patterns of genomic variations and evolutionary history of breast cancer among patients with the disease in West Africa, we genotyped 100 cases (50 cases each from Nigeria and Senegal) using two microsatellite markers (*NR-21* and *NR-24*) to evaluate the instability of the markers in breast cancer in the study populations. For the MSI status, the result showed MSI-H tumours of 52.2% and 67.4% for Nigeria and Senegal respectively. There was more genetic variability for each marker within (*NR-21* - 91.52 and *NR-24* - 88.59) than between (*NR-21*- 8.48 and *NR-24* - 11.40) the two countries with high haplotype and low nucleotide diversity [(*NR-21* - haplotype diversity:  $0.7353 \pm 0.0277$ ; nucleotide diversity:  $0.0069 \pm 0.0121$ ); (*NR-24* haplotype diversity:  $0.8981 \pm 0.0268$ ; nucleotide diversity:  $0.0026 \pm 0.0072$ )]. These indicates rapid tumour growth and expansion, positive selection, non-equilibrium between mutation and genetic drift and tumours that follow the Darwinian mode of evolution with no difference between the two populations. Therefore, we suggest the utilization of these markers in Nigeria and Senegal and more studies to include germline DNA in our population and other African countries.

#### INTRODUCTION

Breast cancer is the most common cancer among women and the second most frequent cancer worldwide representing 11.6% of the global cancer burden (GLOBOCAN, 2018). Although it is more common among women in developed countries, the death rates are more in developing countries than in the developed regions reflecting the national levels of social and economic development from the comparison of the Human Development Index (HDI) (Bray *et al.*, 2018). Incidence rates vary nearly 3.5 times across world regions ranging from 27.9 per 100,000 in

Middle Africa to 48.9% in Northern Africa and 94.2% in Australia/New Zealand (Ferlay *et al.*, 2018).

Breast cancer ranks as the fifth cause of death from cancer overall (626,679) and while Africa ranks fifth in incidence and five-year prevalence, it ranks third in the level of death rates globally (74,072) (Ferlay *et al.*, 2018). The range in mortality rate between world regions is less than that for incidence because of the more favourable survival of breast cancer in (high-incidence) developed regions, with rates ranging from 8.6 per 100,000 in eastern Asia to 25.5 per 100,000 in Melanesia (Ferlay *et al.*, 2018). The highest burden of disease with varying phenotypes can be seen in Africa with rich potential samples and interesting phenotypes that unravel the genetic and environmental basis for diseases. In West Africa, it has an incidence of 37.3% with death rate of 17.8% far exceeding the mortality rate of Australia/New Zealand 12.6% which has the highest incidence rate of 94.2% (Bray, 2018). The rise in incidence of breast cancer in Nigeria from 34 in every 100,000 women (21,750 new cases per year) to 42 per 100,000 between 1970s and 2018 has made it the most common cancer among women in the country while it is second to cervical cancer among women in Senegal (Adebamowo and Ajayi, 2000; Ferlay *et al.*, 2018). Tumour suppressor genes or oncogenes with microsatellite repeats accumulate mutations that lead to variable forms of frame-shift mutations in the microsatellite sequences which can cause truncations in proteins as well as progression of tumours. To detect the patterns of genomic variations and evolutionary history among populations, molecular markers such as microsatellites are widely used to study these processes because they are particularly susceptible to deletions and insertions during replication depending on their length with the frequency of instability recurring more in longer repeats than shorter repeat sequences (Sagher *et al.*, 1999; Agrafioti et Stumpf, 2007). Therefore, they can be used to trace recent demographic events because of their low estimated mutation rate of nucleotides per locus per generation (Tishkoff and Verrelli, 2003).

The genetic differences observed in modern populations have been shaped by the evolution of the African genome through demographic factors such as migration (distance and close), fluctuations in population, sub-population structure, genetic drift and admixture (Tishkoff and Verrelli, 2003; Campbell and Tishkoff, 2008). This has led to the vast genomic/genetic diversity observed within and between different populations (ethnicities and countries) in the continent. A genome-wide association study of 848 short tandem repeat polymorphisms (STRPs), 476 insertion-deletions (INDELs) and 3 single nucleotide polymorphisms (SNPs) genotyped in ~2400 individuals from 121 geographically diverse populations showed 14 genetically divergent ancestral population cluster in Africa revealing a wide population structure on the continent (Tishkoff *et al.*, 2009). Due to this large genetic diversity in Africa, therapeutics designed for populations of European descent are often not effective and often produce low results/successes in the African setting (UChicago Medicine, 2005).

Cancer cells undergo invasion-metastasis cascade, a discrete multistep process of disease progression where neighbouring cells are invaded and others migrate to distant cells, tissues and organs through the functional abilities that permit survival, growth and metastasis (Hanahan and Weinberg, 2011). This is reflected on the extent of the tumour genome heterogeneity which have important implications for cancer treatment and is made possible through the enabling characteristics of genomic instability and tumour promoting inflammation that naturally generate multiple resistant tumour clones at higher rate than usual (Loeb *et al.*, 2008; Talmadge and Fidler, 2010). Oncogenes and tumour suppressor genes which

respectively are activated by alterations in the DNA and inactivated during tumorigenesis are cancer genes that are causally implicated in the development and conferring of selective growth advantage of tumour cells (Greenman *et al.*, 2007; Boland and Goel, 2010).

The term genomic instability indicates the increased tendency of tumour cells to acquire new mutations with each cell division and the concept has become increasingly obvious from studies of chromosomal instability substantiated by the establishment of “microsatellite instability” due to mismatch repair deficiency (Lengauer *et al.*, 1997). Chromosomal and microsatellite instability are two distinct forms of genomic instability identified in human cancers. One model of MSI indicates that the loss and gain of nucleotides result from slippage in the template and newly synthesized primer strands respectively which can occur either in germline or sporadic tumours (Sagher, 1999).

The DNA mismatched repair system maintains the integrity of the genomes and failure of the system to correct and repair mutations in the short tandem repeats lead to microsatellite instability (Kolodner and Marsischky, 1999). Microsatellite instability (MSI) phenotype is the presence of alternative sized microsatellites in tumour DNA that are different from that seen in the corresponding germline DNA (European Patent Specification, 2010). MSI increases the base mutation rate which is one of the mechanisms that brings about the ‘mutator phenotype’ in cancers leading to inordinate proliferation, evolution and survival of cancerous cells (Hampel *et al.*, 2005; Morgan and Reuter, 2006). The MSI phenotype has become an increasingly relevant tool in genetic and immuno-oncology exploration of molecular markers for the treatment of hematological and solid tumours (U.S. FDA, 2017; Hempelmann *et al.*, 2018).

Although this procedure is well standardized for colorectal cancers with MSI tumours having better prognosis than microsatellite stable (MSS) tumours, the results differ considerably in breast cancer; 0 - 30% and 0 – 83% respectively in studies involving sporadic and familial breast cancers with a higher frequency in lobular than medullary carcinomas (Yee *et al.*, 1994; Aldaz *et al.*, 1995; Jonsson *et al.*, 1995; Contegiacomo *et al.*, 1995; Paulson *et al.*, 1996; Toyama *et al.*, 1996; Rush *et al.*, 1997; DeMarchis *et al.*, 1997; Sourvinos *et al.*, 1997; Schmitt *et al.*, 1999; Benachenhou *et al.*, 1999; Tomita *et al.*, 1999; Bock *et al.*, 2000; Ottini *et al.*, 2000; Tokunaga *et al.*, 2000; Ando *et al.*, 2000; Ozer *et al.*, 2002; Murata *et al.*, 2002; Laqmani *et al.*, 2003; Chapgar *et al.*, 2004). Nevertheless, although breast cancer is not an inherently immunogenic disease, recent efforts exploring its use have shown promising results for its utilization especially among triple negative breast cancer tumour subtype (Kok *et al.*, 2017; Hu and McArthur, 2018). But a constant challenge has been to define the patient population that would most benefit from immunotherapy.

Microsatellite repeats in the coding regions are highly conserved while those in the non-coding regions are polymorphic signifying a high rate of spontaneous mutation (Suraweera *et al.*, 2001). The conserved sequences shown in some of the non-coding regions especially of the 5’ and 3’ UTRs that flank gene coding sequences are indicative of their role in gene expression (Sonenberg, 1994; Pesole *et al.*, 1997; Conne *et al.*, 2000). Mononucleotide repeats (MNRs) represent the most interesting kind of microsatellites with conserved lengths for those within the coding regions while the length of those in the non-coding regions can vary highly from individual to individual. (Metzgar *et al.*, 2000). There also exist within the non-coding regions a number of quasimonomorphic MNRs of lengths 20 - 40bp. These quasimonomorphic MNRs show significantly restricted length variation within the human population which may indicate the possibility of their functional relevance (Woerner *et al.*, 2009). Shortening or elongation of MNRs within 5’ UTRs on the transcription level of its mRNA and of the 3’ UTRs on the transcript stability of its mRNA (Suraweera *et al.*, 2001). Due to the high level of variability and rate of mutations that alter their length, microsatellites are useful for studying population structure on the time scale of thousands of generations (Bowcock *et al.*, 1994; Roy *et al.*, 1994;

Goldstein *et al.*, 1995; Underhill *et al.*, 1996; Goldstein and Pollock, 1997; Reich and Goldstein 1998; Goldstein *et al.*, 1999).

The current technique of microsatellite status analysis involves the use of commercially available kits based on the amplification of five microsatellites "markers". These five markers are: *BAT-25*, *BAT-26*, *NR-21*, *NR-22* and *NR-24* and are called quasi-monomorphic because they are characterized by a number of homogeneous repetitive sizes in a given population (between individuals and the two alleles of the same individual). They are located in the genome as follow: *BAT-25* (located in intron 16 of the *c-kit*); *BAT-26* (located in intron 5 of *MSH2* gene); *NR-21* (located in the 5' untranslated region of the *SLC7A8* gene); *NR-24* (located in the 5' untranslated region of the *ZNF-2* gene) and *NR-27* -*MONO-27* (located in the 5' untranslated region of the *PAI-1* gene).

The utilization of these markers arose from the need to circumvent the use of matched normal tissues with cancer tissues and instead use cancer tissues alone which will avoid the mixing up of samples during analysis and less expensive. It was discovered that when these five mononucleotide markers were used together in a unique amplification reaction named pentaplex PCR, it was sufficient to correctly identify MSI cases by tumour tissues only although *BAT-25* and *NR-21* have been shown to be polymorphic in some populations and linked to Africa (Pyatt *et al.*, 1999 ; Buecher *et al.*, 2011). Although dinucleotides and trinucleotide markers have been studied and recommended (Boland *et al.*, 1998; Perucho *et al.*, 1999; Sutter *et al.*, 1999; Loukola *et al.*, 2001; Ozer *et al.*, 2002), dinucleotide repeats have shown 60%– 80% of MSI-H tumours (Sutter *et al.*, 1999) and misclassification of MSI-H tumours (Hoang *et al.*, 1997; Loukola *et al.*, 2001). Therefore, due to the sensitivity and specificity of these five mononucleotide markers, they have been recommended for the diagnosis of microsatellite instability status in tumours (Suraweera *et al.*, 2002).

## **METHODOLOGY**

### **Ethical consideration**

All procedures for the handling of study participants in this study were performed in compliance with the ethical approval and standards of the University of Ibadan/University College Hospital, Ibadan, Nigeria (HREC Assigned Protocol number: UI/EC/16/0334) and the Cheikh Anta Diop University, Dakar, Senegal Research Ethics Committee (Reference number: Protocole 0269/2015/CER/UCAD). The study was conducted according to the recommendations of the Declaration of Helsinki for biomedical research and informed consent was obtained from all participants after detailed explanation of the objectives, benefits and contribution of the study to public health.

### **Study sites**

All consecutive cases of breast cancer who visited the surgical oncology unit in the Department of Surgery of the University College Hospital, Ibadan, Nigeria (UCH) from November 2016 to July 2017 were eligible for this study. This hospital serves an immediate population of about 3.5 million people and it is a major referral hospital in the country located in the third most populous city. The majority of breast cancer patients diagnosed in the region especially based on referral are seen in the hospital at the UCH. After obtaining informed consent, cases were recruited at first presentation, using a well-structured questionnaire and biopsy was done by the surgeons. 50 cases ages 29 years to 80 years old were recruited for the study. In Senegal, cases were recruited in 2015 at the Institute of oncology of Aristide Le Dantec hospital, Dakar,

Senegal. Aristide le Dantec Hospital in Dakar, Senegal is the major and only referral hospital for cancer patients in the country serving an immediate population of about 2.45 million people. Participants were between 15 years and 80 years old. Cancerous tissue sample was collected by the surgical oncologists through biopsy or later through mastectomy from each patient after informed consent and a standardized questionnaire was used to collect the patients' data.

### **Choice of markers used in the study**

The two quasi-monomorphic mononucleotide repeats analysed in this study *NR-21* and *NR-24* are conserved in germline DNA from Caucasian and African subjects. Both are involved in the transcription process and are similar to *NR-22*, *BAT-25* and *BAT-26*. They are highly sensitive to somatic deletion in MSI-H tumours. *NR-21* marker is a 21T repeat that is located in the 5' UTR of the *SLC7A8* gene on position 11 of the long arm (q) of chromosome 14 (14q11.2). The *NR-24* marker is a 24T repeat found on the long arm (q) of chromosome 2 in the 3' UTR of the Zinc finger-2 gene (2q11.1). The use of a *BAT-25* and *BAT-26* markers as a combination has been proposed for the detection of MSI (Hoang *et al.*, 1997; Zhou *et al.*, 1997). They are quasi-monomorphic in Caucasian populations which allows them to be readily classified by the large allelic size variations seen in MSI tumour DNA as being due to somatic alteration (Hoang *et al.*, 1997; Zhou *et al.*, 1997). Majority of colorectal tumours analysis of *BAT-25* and *BAT-26* establish their MSI status without reference to the germline DNA (Zhou *et al.*, 1997) but some studies identified their alternative size alleles in about 19% of African Americans (Pyatt *et al.*, 1999; Samowitz *et al.*, 1999). Therefore, analysis of additional repeats has been proposed to avoid the occasional false positive results that originate from these germline polymorphisms (European Patent Specification, 2003). In line with this need, we decided to carry out this study to identify the MSI phenotype of *NR-21* and *NR-24* among breast cancer patients in Nigeria and Senegal and compare the polymorphic variations of these markers between the two populations.

### **Laboratory Analysis**

The DNA extracted from the tumour tissues of breast cancer obtained through fine needle biopsy, tru-cut biopsy or mastectomy was extracted using the standard protocol for the BioBasic kit and thereafter, coded and stored at -20°C for PCR analysis. The primers used for the amplification of the two microsatellites markers (*NR-21* and *NR-24*) in this study are as recorded on Table 1 (Buhard *et al.*, 2006). The amplifications were carried out in a reaction volume of 50 µl and the PCR reaction for each repeat contained 2.5 µl of the forward and reverse primers, 2 µl of each dNTP, 2 µl of MgCl<sub>2</sub> and 0.1 µl of *Taq* polymerase. The PCR was performed in an Eppendorf-type thermal cycler and the conditions were as follows: Initial denaturation at 94°C for 5 minutes, 35 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds, extension at 72°C for 30 seconds and final extension at 72°C for 7 minutes. Thereafter, the PCR products were run on gel electrophoresis using 1.5% agarose. For the sequencing, 30 µl of each amplicot was placed in a 96-well microtiter microplates along with each primer (15 µl of 10Mm) for each sample and sent to Macrogen, South Korea for Sanger sequencing (Mbaye *et al.*, 2015).

Table 1. Primer Sequences for the Microsatellite Instability Assay (Buhard *et al.*, 2006)

Name	Gene	Genbank number	Repeat	Primer Sequences	Amplification size (bp)
NR-21	SLC7A8	XM_033393	21T 5'UTR	F: GAGTCGCTGGCACAGTTCTA R: CTGGTCACTCGCGTTTACAA	107-110
NR-24	Zinc finger 2	X60152	24T 3'UTR	F: GCTGAATTTTACCTCCTGAC R: ATTGTGCCATTGCATTCCAA	126-129

### Sequence variant nomenclature and classification

**Antonarakis and the Nomenclature Working Group** (1998) suggested a nomenclature system for the description of mutations and polymorphisms observed in DNA and protein sequences. This nomenclature is known as the Human Genome Variation Society (HGVS) recommendations/ nomenclature (den Dunnen and Antonarakis, 2000). The sequence variant nomenclature constitutes the changes observed in a specific sequence when that sequence is compared to a reference sequence and for a stable, accurate, meaningful, unequivocal, unambiguous but flexible description, the Sequence Variant Description Working Group (SVD-WG) was established to address errors and inconsistencies (den Dunnen *et al.*, 2016). The variants given the most priority were Inversion, Deletion, Duplication, Insertion and Conversion. Others include Substitution, Deletion-Insertion, Translocation, Transposition, Gene conversion.

### Population genetics diversity for the study of variability and differentiation

The alignment of the sequences of NR-21 and NR-24 obtained after the sequencing were analyzed using the BioEdit Sequence Alignment Editor version 7.1.9 to clean, correct and aligned using the Clustal algorithm (Thompson *et al.*, 1994; Hall, 1999). To test the association between the association of breast cancer and instability of the microsatellite markers among the studied population, the sequences obtained were aligned along with the reference sequences for each marker using the GenBank number as in Buhard *et al.*, 2006 and stated in Table 1. The differences observed in the study population for each microsatellite sequences to the reference in terms of size and pattern of variations (deletion, insertion, transition or transversion) were considered to represent instability. Since normal DNA was not examined in this study, every sample with a polymorphism in any of the mononucleotide repeats NR-21 and NR-24, was scored positive for instability (Wong *et al.*, 2006). Tumours were defined as Microsatellite stable (MSS) tumours if they were stable for the two markers, Microsatellite Instability low (MSI-L) if they were unstable for one marker and Microsatellite Instability high (MSI-H) if they were unstable for the two markers (Lee *et al.*, 2001).

Arlequin version 3.0 (Excoffier and Lischer, 2005) was used to calculate the statistical parameters (gene diversity, nucleotide diversity and pairwise differences). For each marker, the number of polymorphic sites, haplotype ( $\theta_w$ ) and nucleotide ( $\pi$ ) diversity were estimated for each study population (Nei, 1987; Rozas *et al.*, 2012, Tamura *et al.*, 2013). Population expansion and the existence of a deviation from the hypothesis of neutrality in cancerous tissues were studied by calculating  $[P_i (\pi) \pm SD]$  and haplotype  $[(\theta_w) \pm SD]$  diversities, Tajima's D and Fu's  $F_s$  (Tajima, 1989; Fu, 1997). The test measures the difference in the average number of differences observed between the pairs of sequences and the total number of polymorphic sites and the level of significance was evaluated at  $p = 0.05$  for Tajima's D and 0.02 for Fu's  $F$  (Excoffier and Lischer, 2010). The values will be

negative under directional selection (positive or negative) or expansion of population size (tumour) due to an excess of rare mutations and low frequency polymorphism, and will be positive under balanced selection or decrease in population size (tumour) due to an excess of intermediate frequency alleles. The genetic differentiation ( $F_{ST}$ ) between Nigeria and Senegal and their associated probabilities were calculated using the Arlequin version 3.0 software (Weir and Hill, 2002; Excoffier *et al.*, 2005). Since sequence data contain information on the evolutionary relationship between haplotypes, we performed Analysis of Molecular variance (AMOVA) on the sequenced data where the matrix of pairwise differences between haplotypes was used to calculate the  $\theta_{ST}$  statistics (Excoffier *et al.*, 1992). This is because  $\theta_{ST}$  is independent of the mutation rate when calculated for sequenced data (Kronholm *et al.*, 2010). There is greater sensitivity in detecting diversity when microsatellite markers are used in comparison to other markers. This favourable aspect is also observed when populations are subdivided and total diversity is split into components between and within subpopulations, as proposed by Nei (1973).

### Statistical analysis

The Statistical Package for Social Sciences (SPSS) version 20.0 (SPSS, Inc. Chicago IL, USA) was used to evaluate the frequencies of different variables as well as the chi-square values to determine the statistical differences or otherwise among test variables. The statistically significant value was placed at  $p$  - value < 0.05.

## RESULTS

### Microsatellite instability phenotype of *NR-21* and *NR-24* among breast cancer patients in Nigeria and Senegal

The results herein presented are those obtained only from the region of the markers where the microsatellites repeats are present and to avoid any bias, the alignment for each microsatellite marker studied were subsequently confirmed manually with the chromatogram for each individual sequence. A total of ninety-six samples were each sequenced for both microsatellite markers but after alignment, ninety-three sequences were obtained for *NR-21* in Nigeria and Senegal (46 for Nigeria and 47 for Senegal). The results are as summarized in Table 2 which shows the percentages of individuals in each country that were stable and unstable for *NR-21* and *NR-24*. The percentage of individuals that were stable and unstable for *NR-21* were the same for each country (MSS - 15%; MSI - 85%). For *NR-24* marker, 81% of the study population were unstable for Senegal and 70% of cases from Nigeria showed instability. These data also represent the percentages of variant alleles for each marker. For Nigeria, two cases (4.3%) were MSS tumours, twenty cases (43.5%) showed MSI-L and twenty-four cases (52.2%) showed MSI-H tumours. Cases from Senegal had two individuals (4.7%) with MSS tumours, 12 (27.9%) with MSI-L and twenty-nine individuals (67.4%) that had MSI-H tumours. The result comparing the two countries was not statistically significant for both markers (data not shown).

<b>Table 2:</b> Results of MSI phenotype among breast cancer patients in Nigeria and Senegal				
Marker	MSI status	Nigeria : No. (%)	Senegal : No. (%)	Total : No. (%)
<i>NR-21</i>	MSS	7 (15.2)	7 (14.9)	14 (15.1)
	MSI	39 (84.8)	40 (85.1)	79 (84.9)
<i>NR-24</i>	MSS	13 (29.5)	8 (18.6)	21 (24.1)
	MSI	31 (70.5)	35 (81.4)	66 (75.9)

Table 3 and Table 4 shows the distribution of haplotypes for *NR-21* and *NR-24* for Nigeria and Senegal which result from the mutations cause by the instability of the microsatellite repeats. Of the two markers, *NR-24* had the highest number of haplotypes (23 haplotypes) than *NR-21* (8 haplotypes). Of the 44 sequences for Nigeria, there were only 13 cases which had the wild type allele for *NR-24* and 8 cases for Senegal. For *NR-21*, Senegal had the highest number of haplotypes (8 haplotypes) with the highest frequency having a sequence of 22T repeats (23 cases) and two haplotypes with the lowest frequency of one each having 16T and 17T repeats (Table 3). There were four haplotypes of *NR-21* for Nigeria with the highest number of haplotype frequency (16) having 23T repeats and the lowest haplotype frequency for the variant alleles (8) having 24T repeats. The number of cases with the wildtype (21T) repeats was the same (7) for each country. The allele variants for each marker are also presented in Table 3 for *NR-21* and Table 4 for *NR-24*.

For *NR-21* 6 (12.8%) tumours from Senegal had deletions of up to 5 nucleotides (nts) while 87.2% of the tumours had duplication of up to 3 nts but tumours from Nigeria had no deletions but rather, all had duplications of variable numbers of thymine (T) for this marker. *NR-24* had substitutions of Guanine (G) nucleotide between position 17 and 24 of the microsatellite sequence.

Table 3: *NR-21* allele sizes and frequency for Nigeria and Senegal

Haplotypes	Allele Sizes	Allele variants	Nigeria : (Frequencies)	Sénégal : (Frequencies)
Hap 1	1_5delT (16T)	-----TTTTTTTTTTTTTTTTTT	-	1
Hap 2	1_4delT (17T)	----TTTTTTTTTTTTTTTTTTTT	-	1
Hap 3	1_3delT (18T)	---TTTTTTTTTTTTTTTTTTTTT	-	2
Hap 4	1delT (20T)	-TTTTTTTTTTTTTTTTTTTTTTT	-	2
Hap 5	<b>21T(Wildtype)</b>	TTTTTTTTTTTTTTTTTTTTTTTT	7	7
Hap 6	22T	TTTTTTTTTTTTTTTTTTTTTTTT	15	23
Hap 7	23T	TTTTTTTTTTTTTTTTTTTTTTTT	16	9
Hap 8	24T	TTTTTTTTTTTTTTTTTTTTTTTT	8	2
Total			46	47

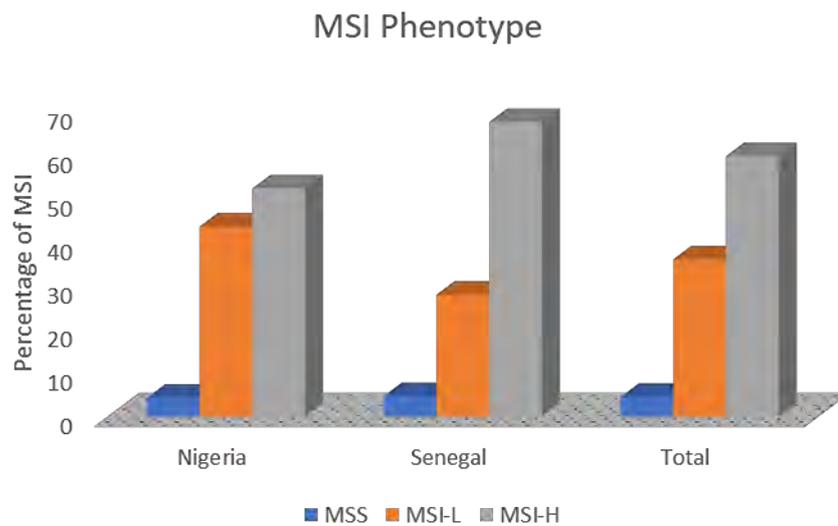


Figure 1: MSI Phenotype for Nigeria, Senegal and both countries

Table 4: NR-24 allele sizes and frequencies for Nigeria and Senegal

Haplo- types	Allele Sizes	Allele variants	Nigeria: No.	Senegal: No.
Hap 1	24T (Wild type)	TTTTTTTTTTTTTTTTTTTTTTTTTTTT	13	8
Hap 2	24dupT	TTTTTTTTTTTTTTTTTTTTTTTTTT	3	1
Hap 3	24T>G	TTTTTTTTTTTTTTTTTTTTTTTTTTG	-	10
Hap 4	22delT	TTTTTTTTTTTTTTTTTTTTTTTTTT	-	1
Hap 5	22delT; 24T>G	TTTTTTTTTTTTTTTTTTTTTTTTTTG	1	3
Hap 6	1delT; 22delinsG; 24dupT	-TTTTTTTTTTTTTTTTTTTTTTGTTT	-	1
Hap 7	21delT; 22delinsG; 24dupT	TTTTTTTTTTTTTTTTTTTTTTTTTTGTTT	-	2
Hap 8	20delinsG	TTTTTTTTTTTTTTTTTTTTTTGTTTT	-	1
Hap 9	20delinsG; 24T>G	TTTTTTTTTTTTTTTTTTTTTTGTTTG	1	2
Hap 10	20delinsG; 22delT; 24T>G	TTTTTTTTTTTTTTTTTTTTTTGT-TG	1	-
Hap 11	20delinsG; 22T>G; 23_24delT	TTTTTTTTTTTTTTTTTTGTG--	7	-
Hap 12	20delinsG; 22_24delT	TTTTTTTTTTTTTTTTTTGT---	7	1
Hap 13	1_2delT; 22delinsG ; 24delT	--TTTTTTTTTTTTTTTTTTGT-	2	-
Hap 14	20delins; 22_23delT	TTTTTTTTTTTTTTTTTTGT-T	2	-
Hap 15	20delT; 24T>G	TTTTTTTTTTTTTTTTTTT-TTG	-	5
Hap 16	20delT; 23delT ; 24T>G	TTTTTTTTTTTTTTTTTTT-TT-G	-	1
Hap 17	20_21delT; 24T>G	TTTTTTTTTTTTTTTTTTT--TTG	-	1
Hap 18	20_21delT ; 22delinsG	TTTTTTTTTTTTTTTTTTT--GTT	-	2
Hap 19	20_21delT; 22delinsG; 24T>G	TTTTTTTTTTTTTTTTTTT--GTG	2	-
Hap 20	20_21delT; 22delinsG; 24delT	TTTTTTTTTTTTTTTTTTT--GT-	1	-
Hap 21	8_10T>C	TTTTTTTCCCTTTTTTTTTTTTTTT	4	-
Hap 22	1delT; 24T>G	-TTTTTTTTTTTTTTTTTTTTTTG	-	3
Hap 23	17T>G; 20delinsG; 21_24delT	TTTTTTTTTTTTTTTTTTGTTG----	-	1
<b>Total</b>			44	43

The percentages of the microsatellite instability phenotype (MSS, MSI-L and MSI-H) for Nigeria and Senegal is represented in figure 1. MSI-H tumours (instability in the 2 markers) were most frequent in Senegalese 29 (67.4%) than Nigerian 24 (52.2%) tumours while MSI-L tumours were more in Nigerian 20 (43.5) than Senegalese 12 (27.9) cases and MSS tumours were the same for both countries. In total, the result showed that MSI-H tumours were the most frequent tumours.

Table 5: Parameters of genetic variability for NR-24

<b>Intra-population variability</b>			
<b>Source of variation</b>	<b>Sum of square</b>	<b>Variance</b>	<b>Percentage of variance</b>
<b>Between Population</b>	9.033	0.1762	11.4015
<b>Within Population</b>	116.381	1.3692	88.5986
<b>Total</b>	125.414	1.5454	
$F_{ST} : 0.1140 ; p - value = 0.0000$			
Haplotype diversity : $0.8981 \pm 0.0268$			
Nucleotide diversity : $0.0026 \pm 0.0072$			
Mean number of pairwise differences : $0.0465 \pm 0.1165$			
Nucleotide composition	T : 96.42%	G : 3.58%	C : 0%    A : 0%
<b>Inter-population variability</b>			
$F_{ST} p - values$			
	Senegal	Nigeria	
Senegal	*		
Nigeria	$0.1016 \pm 0.0108$	*	
Test of molecular diversity	Senegal	Nigeria	
Theta (S)	0.2311	0.6897	
S.D. Theta (S)	0.2311	0.4285	
Theta ( $\pi$ )	0.0465	0.5074	
S.D. Theta ( $\pi$ )	0.1294	0.4868	
Theta (k)	8.7749	5.0899	
95% C.I.	4.6659, 16.1789	2.5580, 9.7821	
<b>Test of Selective Neutrality</b>			
Tajima's D	-1.1175	-0.5627	
<i>P - value</i>	0.1290	0.3130	
Fu's Fs	-0.3403	-14.6505	
<i>P - value</i>	0.0000	0.0000	

Table 6: Parameters of genetic variability for *NR-21*

<b>Intra-population variation</b>				
Source of variation	Sum of square	Variance	Percentage of variance	
<b>Between Population</b>	3.453	0.0603	8.4786	
<b>Within Population</b>	59.21	0.6507	91.5214	
<b>Total</b>	62.6670	0.7109		
F <sub>ST</sub> : 0.0848; <i>P</i> - value = 0.0019				
Haplotype diversity: 0.7353 ± 0.0277				
Nucleotide diversity: 0.0069 ± 0.0121				
Mean number of pairwise differences: 0.1258 ± 0.1962				
Nucleotide composition		T: 100%	G: 0%	C: 0% A: 0%
<b>Inter-population variation</b>				
F <sub>ST</sub> <i>p</i> - values				
	Senegal	Nigeria		
Senegal	*			
Nigeria	0.1256 ± 0.0056	*		
Test of molecular diversity	Senegal	Nigeria		
Theta (S)	0.0000	0.0000		
S.D. Theta (S)	0.0000	0.0000		
Theta (π)	0.1258	0.0000		
S.D. Theta ((π)	0.2178	0.0000		
Theta (k)	0.8542	2.5141		
95% C.I.	0.2923, 2.2498	1.1331, 5.2626		
<b>Test of Selective Neutrality</b>				
Tajima's D	0.0000	0.0000		
<i>P</i> - value	1.0000	1.0000		
Fu's Fs	-14.6989	0.3402		
<i>P</i> - value	0.0000	1.0000		

Table 5 and 6 shows the result of the analysis of molecular variance (AMOVA) and the degree of differentiation between Nigeria and Senegal. For the intra-population variability, the percentage of variability for each marker was greater within tumours from each country (*NR-21* - 91.52 and *NR-24* - 88.59) than between them (*NR-21*- 8.48 and *NR-24* - 11.40) with low F<sub>ST</sub> value and significant *p* - value [*NR-21* (F<sub>ST</sub> = 0.8048, *p* - value = 0.0019); *NR-24* (F<sub>ST</sub> = 0.1140, *p* - value = 0.0000)]. The result also showed high haplotype diversity ± S.D and low nucleotide diversity ± S.D for each marker in the two populations [(*NR-21* - haplotype diversity: 0.7353 ± 0.0277; nucleotide diversity: 0.0069 ± 0.0121); (*NR-24* haplotype diversity: 0.8981 ± 0.0268; nucleotide diversity: 0.0026 ± 0.0072)]. Being only thymine repeats, the nucleotide composition was 100% and 96.4% T for *NR-21* and *NR-24* microsatellites respectively while the remaining 3.6% of *NR-24* were insertions and substitutions of guanine (G) nucleotides. The result for the inter-population pairwise difference and F<sub>ST</sub> was significant in Senegal but not in Nigeria for both markers. The test of neutrality shows negative values for both Tajima's D and Fu's Fs in the two markers for the two populations but significant (*p* - value = 0.0000) for *NR-24* for both and only in Senegal for *NR-21*.

### 3.4.2 DISCUSSION

Breast cancer which has now surpassed cervical cancer in most countries is a growing public health challenge worldwide with lower incidence but worst outcome and high mortality rates in Sub-Saharan Africa (Akarolo-Anthony *et al.*, 2010; Sylla and Wild, 2012; GLOBOCAN, 2018). Understanding the biological processes and application of innovative methods of diagnosis and treatment will help eliminate the racial divide in incidence and outcome and lead to improve outcome for the disease. Studies of breast cancer in Africa presents unique opportunities for identifying markers and novel alleles that may play important roles in the identification of population specific disease susceptible loci given the huge fraction of genetic diversity on the continent (Campbell and Tishkoff, 2008). Therefore, it is important to characterize the genomic diversity and evolution of tumour cells in this population to identify variants that are responsible for disease progression and susceptibility (Campbell and Tishkoff, 2010; Campbell *et al.*, 2014). A better understanding of the molecular factors involved in the development of breast cancer is needed to identify new biomarkers to diagnose breast cancer in early stages and develop more effective therapeutic strategies. This biomarker needs to be easily quantifiable in clinical samples, detected in high level in cancer patients, show progressive function of the disease and provide prognostic or diagnostic information on the disease (Gilad *et al.*, 2008; Etheridge *et al.*, 2011). Microsatellites were once thought to be selectively neutral and not affected by selective pressures but have been widely used in association studies because of their high discriminating power, association with human diseases and practical use in population genetics studies (Mahadevan *et al.*, 1992; Stallings, 1994; Wright and Bentzen, 1994; Schlotterer, 2000; O'Donnell and Warren, 2002). Mutation in a microsatellite locus leads to a gain or loss of a repeat which implies that two alleles differing by only one motif are more related by sharing a more recent common ancestor than alleles differing by several repeats. Tumours with MSI-H phenotype, the instability of long, non-coding mononucleotide repeats markers as present in *BAT-25*, *BAT-26*, *NR-21*, *NR-24* and *NR-27* are progressive and independent (Buhard *et al.*, 2006).

In this study, we evaluated the polymorphism of the MSI phenotype of *NR-21* and *NR-24* between breast cancer patients in Nigeria and Senegal and we hypothesized that there was no difference in the evolutionary process of tumour development between breast cancer in Nigeria and Senegal. The two microsatellite markers utilized in this study are part of a panel of five quasimonomorphic mono-nucleotides that were added to the Bethesda panel when the panel was re-evaluated in December 2002 at the international consensus meeting. The dinucleotide repeats which were in the earlier panel were substituted with three mononucleotide repeats (Umar *et al.*, 2004). Buhard and colleagues in 2014 proposed the application of the panel to non-white population due to the quasimonomorphic nature of the markers and they analyzed them with DNA from fifty-five populations worldwide. Their study revealed no polymorphism in these markers among 44.8% of individuals from Africa including Nigeria and Senegal with *NR-24* been the most stable of the markers with and having 0.1% of polymorphic variants in the study. High frequency of variant alleles was not observed in Senegalese tumours while tumours from Nigerian showed polymorphism in only *BAT-26* and *NR-27*.

Here, we genotyped two microsatellite markers from a total of ninety-three (46 Nigerians and 47 Senegalese) breast tumours. In total, the percentage of variant alleles for each marker were: 85% for *NR-21* (15% for each population) and *NR-24* (50.57% for Nigeria and 49.43% for Senegal). The result obtained in this study is higher than that observed for microsatellites markers in breast cancer for most study but this may be attributed to the type or inappropriate choice of microsatellite markers used, the number of loci, histologic subtypes and the criteria for defining MSI (Jonsson *et al.*, 1995; Contegiacomo *et al.*, 1995; Toyama *et al.*, 1996; Shaw *et al.*, 1996; Paulson *et al.*, 1996; De Marchis *et al.*, 1997; Rush

*et al.*, 1997; Schmitt *et al.*, 1999; Lee *et al.*, 200; Yuksel *et al.*, 2002). Therefore, there is need for the careful selection of informative microsatellite markers for the evaluation of the microsatellite instability phenotype in breast cancer (Bacher *et al.*, 2004). Most studies have utilized *BAT-25*, *BAT-26*, as well as dinucleotide, trinucleotide and tetranucleotide markers and although MSI have been commonly observed in dinucleotide and trinucleotide repeats, there have been the challenge of misclassification of MSI-H tumours with these markers (Hoang *et al.*, 1997; Sutter *et al.*, 1999; Loukola *et al.*, 2001). The most common histological subtype in this study in Nigeria and Senegal was invasive ductal carcinoma (data not shown).

A total of 8 haplotypes was generated for *NR-21* with Senegal having the greatest number of haplotypes (8 haplotype) and Nigeria (4 haplotypes). AMOVA with a significant  $F_{ST}$  value suggested that most of the variation was attributed to variation within populations (91.52%), while differentiation of among populations contributed 8.48% of the total (Table 3.11). The low genetic differentiation observed between the two countries supports the view that intra-tumour evolution is driven by Darwinian selection, therefore the evolution of breast cancer in Nigeria and Senegal follow the Darwinian mode of evolution (Sottoriva *et al.*, 2015; Tao *et al.*, 2015) and is not different between the two countries. This process of genetic diversity include mutation, genetic drift, migration, and natural selection and because tumours are regarded as asexual population of cells, they are believed to follow the natural evolutionary pattern and population structure of positive selection which allow them to adapt to unfavourable environmental conditions and grow rapidly (Nowell, 1976 ; Nei, 1987 ; Gillespie, 1991 ; Wen-Hsiung, 1997 ; Hartl and Clark, 2006 ; Maley *et al.*, 2006 ; Merlo *et al.*, 2006 ; Ewens, 2010). The is also true for *NR-24* (Table 3.10) with a statistically significant  $F_{ST}$  value and a high level of differentiation within (88.59%) than between (11.40%) the two countries. Likewise, the level of diversity and tumour differentiation also affects the response of tumour cells to environmental disturbances whether natural or medical leading often to evolutionary differences observed between primary and metastatic tumours (Maley *et al.*, 2006; Merlo *et al.*, 2006; Marusyk and Polyak, 2010). Nevertheless, the result obtained here is in contrast to the result of Ling *et al.*, (2016), who argues about the small sample size used for the analysis of intra tumour heterogeneity leading to a low power to reject the non-Darwinian model (Tao *et al.*, 2011; Gerlinger *et al.*, 2012; Sottoriva *et al.*, 2013; de Bruin *et al.*, 2014; Zhang *et al.*, 2014; Wang *et al.*, 2014). Ling and Colleagues evaluated multiregional samples from a single tumour using whole genome sequencing (WGS) and they observed high genetic diversity which is in line with the non-Darwinian model of tumour evolution without any evidence of positive selection (Ling *et al.*, 2015). The haplotypic diversity considers the number of samples selected and the haplotype frequency ( $h$ ) shows the probability that two randomly selected sequences within a specific population are identical. The high haplotype diversity seen here is an indication of the different haplotypes observed among two randomly selected tumours. The nucleotide diversity measures the genetic diversity that incorporates the actual degree of differences between the sequences studied. Thus, the more the haplotypes present in a population are distinct from each other in terms of the number of polymorphic sites differentiating them, the higher the nucleotide diversity. Therefore, the high haplotype diversity and low nucleotide diversity for both markers suggests a recent population expansion after a period of low effective population size. This shows that the time after the tumour expansion in each population is long enough to examine the change in haplotypes that resulted from mutation, but it is not long enough to accumulate big differences among sequences because the diversity of a tumour depends solely on the length of time that it takes to grow to a given size (Avisé, 2000; Ling *et al.*, 2015). This is not uncommon for breast cancer especially in sub-Saharan Africa where patients present at younger age with aggressive tumour and at a late stage. Also, in this study as similar to others, most patients are premenopausal and

therefore are still at the stage of active and rapid cell division. Studies have demonstrated higher levels of nucleotide and haplotype diversity both in the nuclear and mitochondrial genomes of the African population as well as more private alleles and haplotypes when compared to non-Africans, seem to have a subset of sub-Saharan Africa's genetic diversity (Tishkoff *et al.*, 1996; Tishkoff and Williams, 2002; Tishkoff and Verrelli, 2003; Tishkoff and Kidd, 2004; Consort. Int. HapMap, 2005; Ramachandran *et al.*, 2005; Conrad *et al.*, 2006; Zhao *et al.*, 2006; Garrigan *et al.*, 2007; Jakobsson *et al.*, 2008; Li *et al.*, 2008). Intra-specific neutrality tests are based on the polymorphism observed in human populations. The existence of pairwise nucleotide difference per site ( $\pi$ ) in the hypothesis of neutrality was estimated by calculating the Tajima's D (TD) and Fu's F (Tajima, 1989; Fu and Li, 1993) while the  $\theta_w$  value is based on the number of segregating sites in a population (Watterson, 1975). The Tajima's D test evaluates a general deviation of the allele frequency spectrum from the expected under neutral conditions taking into account the whole spectrum: it measures the difference between the average of the number of differences observed between pairs of sequences and the total number of polymorphic sites observed. The TD test calculates the normalized difference between  $\theta_w$  and  $\pi$ . A negative value explains an excess of low frequency polymorphism and rare mutations under positive selection/directional selection or population size expansion. When Tajima's D and Fu's F are significant and negative, it is also an indication of a non-equilibrium between mutation and genetic drift which may be due to demographic or population expansion as earlier stated. A positive value depicts balancing selection or population size reduction due to excess of intermediate frequency alleles. For *NR-24*, the value of the neutrality tests for Tajima's D and Fu's F<sub>s</sub> are negative which indicates sensitivity to population expansion and bottleneck effects encouraging tumour growth and metastasis as well as an imbalance between mutation and drift in the tumour population among individual cases in the two studied populations (Tajima, 1993 ; Tajima, 1996 ; Arris-Brosou and Excoffier, 1996). The sensitivity of Fu's F to population expansion is shown by its significant *p* - value for Nigeria and Senegal (Fu, 1997). Also, the *NR-21* test statistics result output for Senegal shows a significant negative value further reinforcing the presence of more tumour expansion for tumours from Senegal than Nigeria (Fu, 1997).

The  $F_{ST}$  for the neutrality test shows genetic drift and the pattern of the tumour differentiation in the same direction on all loci along the genome pathway for each population. Here, it also further supports the hypothesis of recent positive selection of tumour cells and polymorphism between them which makes it possible to directly target the variant under selection which in this case are the microsatellite markers (Beaumont and Balding, 2004; Excoffier *et al.*, 2009). Nonetheless, it will be necessary to study the matching normal of the germline DNA in both countries to ascertain that the observed microsatellite instability of the two markers are not as a result of common polymorphic variations in these populations as numerous studies have shown that there exist a high level of nucleotide and haplotype diversities in the African population (Tishkoff *et al.*, 1996; Tishkoff and Williams, 2002; Tishkoff and Kidd, 2005; Consort. Int. HapMap, 2005; Garrigan *et al.*, 2007; Jakobsson *et al.*, 2008; Li *et al.*, 2008). The loci under divergent selection are expected to have higher  $F_{ST}$  values than neutral loci while loci under balancing selection are expected to have lower  $F_{ST}$  values (Lewontin and Krakauer, 1973). The low  $F_{ST}$  values for the two markers between Nigeria and Senegal been statistically significant shows that the tumours in each population evolve under a balancing selection with only a moderate genetic differentiation between the two countries (Hartl and Clark (1997).

## CONCLUSION

There is strong evidence of genomic instability as a result of high frequencies of chromosomal abnormalities and mutations in human cancers which is characterized by succession of mutations in the tumour as a result of continuous increase of endogenous mutations in the cells. To understand this process in tumour development and resistance to therapy, it is necessary to study the evolutionary process of cancer and its population dynamics to provide biomarkers for development of therapeutic agents and strategies, accurate intervention and risk management. In this study to compare and characterize the evolutionary pattern of breast cancer patients in Nigeria and Senegal, the results suggest that the evolutionary pattern of breast cancer development is the same among patients with the disease in both countries following a process of balancing selection, neutrality, genetic drift and recent tumour expansion. The evolution of breast tumour cells often produce resistant clones which can be prevented by halting the initiation of the tumour growth and expansion through effective therapy. With the highest levels of genetic diversity found in Africa, it might be predicted that Africans would have relatively high amounts of local adaptation contain within the populations because of the wide range of environmental differences as well as high exposure to infectious disease. Nevertheless, signatures of selection in African populations may be missed due to studies carried out mainly on one or two African populations. Therefore, to have a better understanding of genetic and phenotypic adaptations in Africa, it is important to scan for genetic signatures of selection in a broad range of ethnically diverse African populations living in distinct environments and clearly identify functional variants that are likely to be targets of selection so as to verify their impact on phenotypic variations of diseases such as breast cancer. Studies on the continent may also allow further assessment of the role of immune-compromised states of interest given the increasing recognition for a role of inflammatory markers in the etiology of breast cancer and the promising results from the exploration of immunotherapy in treatment. In conclusion, the genetic factors underlying complex diseases are still poorly understood especially in the aspect of genetic and environment interactions as well as epistatic interactions among loci likely to influence complex disease susceptibility. This limitation of this study is the lack of match normal germline DNA and controls to eliminate polymorphic variations of the studied markers in Nigeria and Senegal. Therefore, further studies to establish this fact are needed not only in our population but in Africa as a whole to discover her population or regional specific susceptibility alleles underlying some complex diseases. Also, we only evaluated the region of the tandem repeats which may result in the high haplotype diversity and frequencies observed. We suggest that an extensive analysis of the entire pentaplex panel of the mononucleotide markers *NR-21*, *NR-24*, *NR-27*, *BAT-25* and *BAT-26* be carried out among breast cancer patients to answer some of the scientific questions on breast cancer susceptibility, heterogeneity, treatment, relapse. and survival as studies in West African populations may be informative for the identification of markers necessary for the treatment of breast cancer. Finally, because there is no difference in the evolutionary pattern and the pathway of carcinogenesis of breast cancer in Nigeria and Senegal and therefore treatment pattern and therapies developed to target one population can serve for the other.

## REFERENCES

1. Adebamowo CA, Ajayi OA. 2000. Breast cancer in Nigeria. *West Afr J Med.*; 19:179–191
2. Agrafioti, I & Stumpf, MP. 2007. SNPSTR: a database of compound microsatellite-SNP markers *Nucleic Acids Research*, Vol. 35, Database issue D71–D75.
3. Akarolo-Anthony, S. N, Ogundiran, T. O., Adebamowo, C. A. 2010. Emerging breast cancer epidemic: evidence from Africa. *Breast Cancer Research*. 12(Suppl 4):58
4. Aldaz CM, Chen T, Sahin A, Cunningham J and Bondy M. 1995. Comparative allelotype of in situ and invasive human breast cancer: high frequency of microsatellite instability in lobular breast cancer. *Cancer Res* 55: 3976-3981
5. Ando Y, Iwase H, Ichihara S, Toyoshima S, Nakamura T, Yamashita H, Toyama T, Omoto Y, Karamatsu S, Mitsuyama S *et al* 2000: Loss of heterozygosity and microsatellite instability in ductal carcinoma in situ of the breast. *Cancer Letters* 156: 207-214.
6. Antonarakis, S.E. and the Nomenclature Working Group. 1998. *Recommendations for a nomenclature system for human gene mutations*. *Hum. Mut.* 11: 1-3]
7. Avise, J.C. 2000. *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, MA
8. Bacher, J. W., Flanagan, L. A., Smalley, R. L., Nassif, N. A., Burgart, L. J., Halberg, R. B. *et al.*, 2004. Development of a fluorescent multiplex assay for detection of MSI-high tumours. *Disease Markers*. 20:237-250. doi:10.1155/2004/136734.
9. Beaumont, M. A, and Balding, D. J. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13, 969–980.
10. Benachenhou N, Guiral S, Gorska-Flipot I, Labuda D, Sinnett D. 1999. Frequent loss of heterozygosity at the DNA mismatch-repair loci hMLH1 and hMSH3 in sporadic breast cancer. *Br J Cancer*. 79(7-8):1012–1017
11. Bock N, Meden H, Regenbrecht M, Junemann B, Wangerin J and Marx D. 2000. Expression of the mismatch repair protein hMSH2 in carcinoma in situ and invasive cancer of the breast. *Anticancer Res* 20: 119-124.
12. Boland CR, Thibodeau SN, Hamilton SR *et al.* 1998. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res*; 58: 5248–57
13. Boland, C. R., & Goel, A. 2010. Microsatellite instability in colorectal cancer. *Gastroenterology*, 138(6), 2073-2087.e3.
14. Bowcock A. M., Linares R. A., Tomfohrde J., Munch E., Kidd J. R., *et al.* 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*. 368: 455–457.
15. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L., Jemal, A. 2018. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*. 0:5-31
16. Buecher B., de Pauw A., Freneaux P. and Rouleau E. 2011. Instabilité des microsatellites et cancers colorectaux. *Cancéro digest*. 3(3): 200-203
17. Buhard O, Cattaneo F, Wong YF, Yim SF, Friedman E, Flejou JF, Duval A, Hamelin R 2006. Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumours. *Journal Clinic of Oncology* 24(2): 241–251
18. Campbell MC, Tishkoff SA. 2008. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annual Review Genomics Human Genetics*.9:403–433.

19. Campbell P., Yachida S., Mudie L., Stephens P., Pleasance E., Stebbings L., Morsberger L., Latimer C., McLaren S., Lin M., McBride D., Varela I., Nik-Zainal S., *et al.* 2010. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature*. 467:1109–1113.
20. Campbell, M., Hirbo, J., Townsend, J., & Tishkoff, S. 2014. The peopling of the African continent and the diaspora into the new world. *Current Opinion in Genetics and Development*. 29: 120–132
21. Chapgar A, Magliocco A, Kerviche A, Tan L, Walley B and DeCoteau JF. 2004. The replication error phenotype is associated with the development of distant metastases in hormonally treated patients with breast carcinoma. *Cancer* 100: 913-919.
22. Conne B, Stutz A and Vassalli JD. 2000. The 3' untranslated region of messenger RNA: A molecular 'hotspot' for pathology? *Nature Med.*, 6(6) 637-641.
23. Conrad, D. F., Andrews, T. D., Carter, N. P., Hurler, M. E., and Pritchard, J. K. 2006. A high-resolution survey of deletion polymorphism in the human genome. *Nature Genetics*. 38(1):75-81
24. Consort. Int. HapMap. 2005. A haplotype map of the human genome. *Nature*. 437:1299–1320.
25. Contegiacomo A, Palmirotta R, De Marchis L, *et al.* 1995. Microsatellite instability and pathological aspects of breast cancer. *Int J Cancer*. 64:264-268.
26. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, Jamal-Hanjani M, Shafi S, Murugaesu N, Rowan AJ, Grönroos E, Muhammad MA, Horswell S, Gerlinger M, Varela I, Jones D, Marshall J, Voet T, Van Loo P, Rasi DM, Rintoul RC, Janes SM, Lee SM, Forster M, Ahmad T, Lawrence D, Falzon M, Capitanio A, Harkins TT, Lee CC, Tom W, Teefe E, Chen SC, Begum S, Rabinowitz A, Phillimore B, Spencer-Dene B, Stamp G, Szallasi Z, Matthews N, Stewart A, Campbell P, Swanton C. 2014. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science*. 346(6206):251-256.
27. De Marchis L, Contegiacomo A, D'Amico C, *et al.* 1997. Microsatellite instability is correlated with lymph node–positive breast cancer. *Clin Cancer Res*. 3:241-248.
28. den Dunnen, J. T and Antonarakis, S. E. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Human mutation*. 15: 7-12
29. den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., *et al.* 2016. HGVS Recommendations for the description of sequence variants: 2016 Update. *Hum. Mutat*. 37, 564–569. doi: 10.1002/humu.22981
30. Etheridge, A., Lee, I., Hood, L., Galas, D. and Wang, K. 2011. “Extracellular microRNA: a new source of biomarkers,” *Mutation Research*, vol. 717, no. 1-2, pp. 85–90,
31. European patent specification. 2010. Mononucleotide repeats microsatellite markers for detecting microsatellite instability. WO 2003/072822 (04.09.2003 Gazette 2003/36.
32. Ewens, W. J. 2010. *Mathematical Population Genetics 1: Theoretical Introduction*. Springer, New York.
33. Excoffier L., Laval G. & Schneider S. 2005. Arlequin (version 3): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*. 1: 47-50
34. Excoffier L., Smouse P.E. & Quattro J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 131 (2): 479-491
35. Excoffier, L., Hofer, T. and Foll, M. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity (Edinb)*. 103(4):285-298
36. Excoffier, L. and Lischer, H.E.L. 2010. Arlequin Suite ver 3.5, a New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564-567.

37. Ferlay J, Colombet M, Soerjomataram I, *et al.* 2018. Global and Regional Estimates of the Incidence and Mortality for 38 Cancers: GLOBOCAN 2018. Lyon: International Agency for Research on Cancer/World Health Organization
38. Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*. 147(2): 915-925
39. Fu Y.X. and Li W.H. 1993. Statistical tests of neutrality of mutations. *Genetics*. 133 (3): 693-709.
40. Garrigan D, Kingan SB, Pilkington MM, Wilder JA, Cox MP, *et al.* 2007. Inferring human population sizes, divergence times and rates of gene flow from mitochondrial, X and Y chromosome resequencing data. *Genetics* 177:2195–2207.
41. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, *et al.* 2012 Intratumour heterogeneity and branched evolution revealed by multiregion sequencing. *New England Journal of Medicine.*; 366:883–892
42. Gilad, S., Meiri, E., Yogeve, Y., *et al.*, 2008. Serum microRNAs are promising novel biomarkers. *PLoS ONE*, vol. 3(9).
43. Gillespie, J. H. 1991. The cause of molecular evolution. Oxford University Press, Oxford.
44. Goldstein D. B., Pollock D. D. 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *J. Hered.* 88: 335–342.
45. Goldstein D. B., Ruiz-Linares A., Cavalli-Sforza L. L., Feldman M. W. 1995. Genetic absolute dating based on microsatellites and modern human origins. *Proc. Natl. Acad. Sci.* 92: 6723–6727.
46. Goldstein D. B., Roemer G. W., Smith D. A., Reich D. A., Bergman A., *et al.*, 1999 The use of microsatellite variation to infer population structure in a natural model system. *Genetics* 151: 797–801.
47. Greenman, C., Stephens, P., Smith, R., Dalgliesh, G. L., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., Stevens, C., Edkins, S., O'Meara, S., Vastrik, I., Schmidt, E. E., Avis, T., Barthorpe, S., Bhamra, G., Buck, G., Choudhury, B., Clements, J., Cole, J., Dicks, E., Forbes, S., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Jenkinson, A., Jones, D., Menzies, A., Mironenko, T., Perry, J., Raine, K., Richardson, D., Shepherd, R., Small, A., Tofts, C., Varian, J., Webb, T., West, S., Widaa, S., Yates, A., Cahill, D. P., Louis, D. N., Goldstraw, P., Nicholson, A. G., Brasseur, F., Looijenga, L., Weber, B. L., Chiew, Y. E., DeFazio, A., Greaves, M. F., Green, A. R., Campbell, P., Birney, E., Easton, D. F., Chenevix-Trench, G., Tan, M. H., Khoo, S. K., Teh, B. T., Yuen, S. T., Leung, S. Y., Wooster, R., Futreal, P. A. and Stratton, M. R. 2007. Patterns of somatic mutation in human cancer genomes. *Nature*, 446(7132), 153-8.
48. Hall, T. A. 1999. "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT", *Nucleic Acids Research*, vol. 44, pp. 211-232.
49. Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, *et al.* 2005. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *New England Journal of Medicine.*;352:1851–60
50. Hanahan D, and Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell*. 144(5):646–74.
51. Hartl, D. L. and Clark, A. G. 1997. Principles of population genetics. Sinauer associates Sunderland, vol. 116.
52. Hartl, D. L. and Clark, A. G. 2006. Principle of Population Genetics. Sinauer, Sunderland, MA, 4th Ed
53. Hempelmann, J. A., Lockwood, C. M., Konnick, E. Q., Schweizer, M. T., Antonarakis, E. S., Lotan, T. L., Montgomery, B., Nelson, P. S., Klemfuss, N., Salipante, S. J., and Pritchard, c. c. 2018. Microsatellite Instability in prostate cancer by PCR or next-generation sequencing. *Journal for Immunotherapy of Cancer*. 6(29).

54. Hoang J.M., Cottu P.H., Benedicte T., Salmon R.J., Thomas G. & Hamelin R. 1997. *BAT-26*, an indicator of the replication error phenotype in colorectal cancers and cell lines.
55. Hu, Z. I., and McArthur, H. L. 2018. Immunotherapy in Breast Cancer: the New Frontier. *Curr Breast Cancer Rep.* 10(2): 35-40. doi: 10.1007/s12609-018-0274-y.
56. Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, *et al.* 2008. Genotype, haplotype and copy number variation in worldwide human populations. *Nature.* 451:998–1003.
57. Jonsson M, Johansson O and Borg A. 1995. Infrequent occurrence of microsatellite instability in sporadic and familial breast cancer. *Eur J Cancer* 31A: 2330-2334.
58. Kok M, Horlings HM, van de Vijver K, *et al.* 2017. LBA14 Adaptive phase II randomized non-comparative trial of nivolumab after induction treatment in triple negative breast cancer. TONIC-trial. *Ann Oncol.* 28: mdx440.006-mdx440.006.
59. Kok, M., Horlings, H. M., Snaebjornsson, P., Chalabi, M., Schumacher, T. N., Blank, C. U., Linn, S. C., and Dieren, J. 2017. Profound Immunotherapy Response in Mismatch Repair-Deficient Breast Cancer. *American Society of Clinical Oncology Journal.*
60. Kolodner and Marsischky, 1999. Eukaryotic DNA mismatch repair. *Current Opinion in Genetics and Development.* Vol.9 (1) Pages 89-96
61. Kronholm, I., O. Loudet, and J. de Meaux. 2010. Influence of mutation rate on estimators of genetic differentiation—lessons from *Arabidopsis thaliana*. *BMC Genet.* 11:33
62. Laqmani YA, Mathew M and Lobo S. 2003. Frequent microsatellite alterations in a predominantly younger age of onset breast cancer population. *Int J Mol Med* 11: 117-124.
63. Lee, S. C. Berg, K. B., Sherman, M. E. Griffin, C. A. and Eshleman, J. R. 2001. "Microsatellite Instability is frequent in Medullary Breast Cancer", *American Journal of Clinical Pathology*, vol. 115, no. 6, pp. 823-827
64. Lengauer, C., Kinzler, K. W., and Vogelstein, B. 1997. Genetic instability in colorectal cancers. *Nature.* 386: pp. 623-627
65. Lewontin, R. C. and Krakauer, J. 1973. Distribution of Gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics.* Vol.74(1):175-195
66. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, *et al.* 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science.* 319:1100–1104.
67. Ling, S., Hu, Z., Yang, Z., Yang, F., Li, Y., Lin, P., Chen, K., Dong, L., Cao, L., Tao, Y., Hao, L., Chen, Q., Gong, Q., Wu, D., Li, W., Zhao, W., Tian, X., Hao, C., Hungate, E.A., Catenacci, D. V-T., Hudson, R. R., Li, W-H., Lu, X. and Wu, C-I. 2015. "Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution". *Proceedings of the National Academy of Sciences* 112.47.
68. Loeb LA, Bielas JH, and Beckman RA. 2008. Cancers exhibit a mutator phenotype: clinical implications. *Cancer Research.* 68:3551–3557. discussion 3557
69. Loukola, A., Eklin, K., Laiho, P., Salovaara, R., Kristo, P., Jarvinen, H., Mecklin, J. P., Launonen, V., Aaltonen, L. A. 2001. Microsatellite marker analysis in screening for hereditary nonpolypopsis colorectal cancer (HNPCC). *Cancer research.* 61(11):4545-4549
70. Mahadevan M<sup>1</sup>, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barceló J, O'Hoy K, *et al.* 1992. Myotonic dystrophy mutation : an unstable CTG repeat in the 3' untranslated region of the gene. *Science.* 255(5049) :1253-5
71. Maley, C. C., Galipeau, P. C., Finley, J. C., Wongsurawat, V. J., Li, X., Sanchez, C. A., Paulson, T. G., Blount, P. L., Risques, R. A., Rabinovitch, P. S. and Reid, B. J. 2006. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nature Genetics.* 38(4): 468-473
72. Marusyk, A. and Polyak, K. 2010. Tumour heterogeneity: Causes and consequences. *Biochimica et Biophysica Acta.* 1: 105-117

73. Mbaye, F., Ka, S., Dem, A., Kane, M., and Sembéne, M. 2015. Clinicopathological and survival significance of *BAT-25* and *BAT-26* instability in breast cancer among Senegalese patients. *International Journal of Cancer Research and Experimental Oncology*. 7. 13-19. 10.5897/JCREO2015.0122.
74. Merlo, L. M. F., Pepper, J. W., Reid, B. J. and Maley, C. C. 2006. Cancer as an evolutionary and ecological process. *Nature reviews. Cancer* 6(12): 924-935
75. Metzgar, D., Bytof, J. and Wills, C. 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res.*, 10, 72–80.
76. Morgan MA, Reuter CW. 2006 Molecularly targeted therapies in myelodysplastic syndromes and acute myeloid leukemias. *Annals of Hematology*.;85(3):139–63
77. Murata H, Khattar NH, Kang Y, Gu L and Li GM: 2002. Genetic and epigenetic modification of mismatch repair genes hMSH2 and hMLH1 in sporadic breast cancer with microsatellite instability. *Oncogene* 21: 5696-5703
78. Nei, M. 1973. Analysis of Gene Diversity in Subdivided Populations. *PNAS*. 70(12): 3321-3323
79. Nei, M. 1987. Molecular Evolutionary Genetics. *Columbia University Press*. New York
80. Nowell, P. C. 1976. The Clonal Evolution of Tumour Cell Populations. *Science*. Vol. 194(4260), pp. 23-28
81. O'Donnell, W. T. and Warren, S. T. 2002. A decade of molecular studies of fragile X syndrome. *Annu Rev Neurisci*. 25:315-338
82. Ottini, L, Palli D, Falchetti M, D'Amico C, Noviello C, Saieva C, Russo A, Corsi AM, Masala, G, Paglierani M *et al*. 2000. Instability at dinucleotide and trinucleotide repeats in breast cancer. *Int J Oncol* 17: 819-826.
83. Ozer, E., Yuksel, E., Kizildag, S., Sercan, O., Ozen, E., Canda, T. and Sakilzi, M. 2002. Microsatellite instability in early-onset breast cancer. *Pathology Research Practice* 198: 525-530.
84. Paulson TG, Wright FA, Parker BA, *et al*. 1996. Microsatellite instability correlates with reduced survival and poor disease prognosis in breast cancer. *Cancer Res*. 56:4021-4026.
85. Perucho, M. 1999. Correspondence ref.: Boland C. C. *et al.*, 1998. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.*, 58: 5248–5257; *Cancer Research* 59(1): 249–256
86. Pesole, G., Liuni, S., Grillo, G. and Saccone, C. 1997. Structural and compositional features of untranslated regions of eukaryotic mRNAs. *Gene*. 205, 95 – 102.
87. Pyatt, R., Chadwick, R. B., Johnson, C. K., Adebamowo, C., Chapelle, A. and Prior, T. W. 1999. Polymorphic Variation at the *BAT-25* and *BAT-26* Loci in Individuals of African Origin. *American Journal of Pathology* 155(2), 349-353.
88. Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W. and Cavalli-Sforza, L. L. 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *PNAS*. 102(44):15942-15947
89. Reich D. E. & Goldstein D. B. 1998. Genetic evidence for a Paleolithic human population expansion in Africa. *Proc. Natl. Acad. Sci. USA* 95: 8119–8123.
90. Roy M. S., Geffen E., Smith D., Ostrander E., Wayne R. K. 1994. Patterns of differentiation and hybridization in North American wolflike Canids. *Molecular Biology and Evolution*. 11: 553–570.
91. Rozas J., Librado P., Sánchez-Del Barrio J.C., Messeguer X. and Rozas R. 2012. DnaSP Version 5 Help Contents [Help File]. Available with the program at <http://www.ub.edu/dnasp/>

92. Rush EB, Calvano JE, van Zee KS, *et al.* 1997. Microsatellite instability in breast cancer. *Ann Surg Oncol.* 4:310-315
93. Sagher, D., Hsu, A. and Strauss, B. 1999. Stabilization of the intermediate in frameshift mutation. *Mutat. Res.*, 423, 73–77
94. Samowitz WS, Slattery ML, Potter JD, Leppert MF 1999. *BAT-26* and *BAT-40* instability in colorectal adenomas and carcinomas and germline polymorphisms. *American Journal of Pathology.* 154:1637-1641
95. Schlötterer, C. 2000. Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109 : 365-371.
96. Schmitt FC, Soares R, Gobbi H, Milanezzi F, Santos-Silva F, Cirnes L, Costa C and Seruca R. 1999. Microsatellite instability in medullary breast carcinomas. *International Journal of Cancer* 82: 6440647
97. Shaw JA, Walsh T, Chappell SA, *et al.* 1996. Microsatellite instability in early sporadic breast cancer. *Br J Cancer.* 73:1393-1397.
98. Siah, SP, Quinn DM, Bennett GD, Casey, G., Flower, RL., Suthers, G. and Rudzki, Z. 2000. Microsatellite instability markers in breast cancer: a review and study showing MSI was not detected at ‘*BAT 25*’ and ‘*BAT 26*’ microsatellite markers in early-onset breast cancer. *Breast Cancer Res Treat.*; 60:135-142.
99. Sonenberg N . 1994. mRNA translation: influence of the 5' and 3' untranslated regions. *Current Opinion in Genetics & Development.* 4(2):310-315 A.
100. Sottoriva *et al.*, Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. 2013. *Proc. Natl. Acad. Sci. U. S. A.* 110, 4009-4014.
101. Sottoriva, A., Kang, H., Ma, Z., Graham, T. A., Salomon, M. P., Zhao, J., Marjoram, P., Siegmund, K., Press, M. F., Shibata, D., ... Curtis, C. 2015. A Big Bang model of human colorectal tumor growth. *Nature genetics*, 47(3), 209-16.
102. Sourvinos G, Kiaris H, Tsikkinis A, Vassilaros S and Spandidos DA 1997: Microsatellite instability and loss of heterozygosity in primary breast cancers. *Tumour Biology* 18: 157-166,
103. Stallings R. L. 1994. *Genomics* 21, 116.
104. Suraweera, N., Iacopetta, B., Duval, A., Compoint, A., Tubacher, E. and Hamelin, R. 2001. Conservation of mononucleotide repeats within 30 and 50 untranslated regions and their instability in MSI-H colorectal cancer. *Oncogene*, 20, 7472–7477.
105. Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, Seruca R, Iacopetta B, Hamelin R. 2002. Evaluation of tumour microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* 123(6): 1804–1811
106. Sutter C, Gebert J, Bischoff P, Herfarth C, von Knebel Doeberitz M. 1999. Molecular screening of potential HNPCC patients using a multiplex microsatellite PCR system. *Molecular and Cellular Probes*;13: 157–165
107. Sylla, B.S. and Wild, C.P. 2012. A Million Africans a Year Dying from Cancer by 2030 : What Can Cancer Research and Control Offer to the Continent ? *International Journal of Cancer.* 130, 245-250.
108. Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics.* 123 (3) : 585-95
109. Tajima, F. 1993. Measurement of DNA polymorphism. Pages 37–59 in N. Takahata and A. G. Clark (editors). *Mechanisms of molecular evolution. Introduction to molecular paleopopulation biology.* Japan Scientific Societies Press, Sinauer Associates, Sunderland, Massachusetts.
110. Tajima, F. 1996. The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics* 143 : 1457–1465.

111. Talmadge, J.E. and Fidler, I.J. 2010. AACR centennial series: The biology of cancer metastasis: Historical perspective. *Cancer Research*. Volume 70: 14 (5649-5669).
112. Tamura K., Stecher G., Peterson D., Filipski A. & Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics version 6.0. *Molecular Biology and Evolution*. 30 (12): 2725-2729
113. Tao, Y., Hu, Z., Ling, S., Yeh, S-H., Zhai, W., Chen, K., Li, C., Wang, Y., Wang, K., Wang, H-Y., *et al.* 2015. Further genetic diversification in multiple tumors and an evolutionary perspective on therapeutics. <https://www.biorxiv.org/content/early/2015/08/25/025429>
114. Thompson J.D., Higgins D.G. & Gibson T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research*. 22 (22): 4673-4680
115. Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, *et al.* 1996. Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science*; 271:1380–1387
116. Tishkoff SA, and Williams SM. Genetic analysis of African populations: human evolution and complex disease. 2002. *Nat. Rev. Genetics* 3:611–621.
117. Tishkoff SA, Verrelli BC. 2003. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annual Review of Genomics and Human Genetics*. 4:293–340
118. Tishkoff SA, Kidd KK. 2004. Implications of biogeography of human populations for ‘race’ and medicine. *Nat. Genet.* 36: S21–S27.
119. Tishkoff, S.A., F.A. Reed, F.R. Friedlaender, C. Ehret, A. Ranciaro, A. Froment, J.B. Hirbo, A.A. Awomoyi, J.M. Bodo, O. Doumbo, M. Ibrahim, A.T. Juma, M.J. Kotze, G. Lema, J.H. Moore, H. Mortensen, T.B. Nyambo, S.A. Omar, K. Powell, G.S. Pretorius, M.W. Smith, M.A. Thera, C. Wambebe, J.L. Weber, S.M. Williams. 2009. The genetic structure and history of Africans and African Americans. *Science* 324(5930): 1035-1044.
120. Tokunaga E, Oki E, Oda S, Kataoka A, Kitamura K, Oino S, Maehara Y and Sugimachi K. 2000. Frequency of microsatellite instability in breast cancer determined by high resolution fluorescent microsatellite analysis. *Oncology* 59: 44-49.
121. Tomita S, Deguchi S, Miyaguni T, Muto Y, Tamamoto T and Toda T. 1999. Analyses of microsatellite instability and the transforming growth factor factor-, receptor type II gene mutations in sporadic breast cancer and their correlation with clinicopathological features. *Breast Cancer Res Treat* 53: 33-39.
122. Toyama T, Iwase H, Yamashita H, Iwata, H., Yamashita, T., Ito, K., Hara, Y., Suchi, M., Kato, T., Kobayashi, S. 1996. Microsatellite instability in sporadic human breast cancers. *Int J Cancer*. 68:447-451.
123. U.S. Food and Drugs Administration. 2017. FDA approves first cancer treatment for any solid tumour with a specific genetic feature. <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm560167.htm>. Accessed 20 Nov 2018.
124. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN, Srivastava S. 2004. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *Journal of National Cancer Institute* 96(4): 261–268
125. University of Chicago Medicine, 2005. Study reveals dramatic difference between breast cancers in Caucasian and African women. April 18, 2005. Available from <https://www.uchicagomedicine.org/forefront/news/2005/april/study-reveals-dramatic-differences-between-breast-cancers-in-caucasian-and-african-women>. Accessed March 15, 2018.

126. Underhill P. A., Jin L., Zemans R., Oefner P. J., Cavalli-Sforza L. L. 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc. Natl. Acad. Sci. USA* 93: 196–200
127. Wang, Y., Waters, J., Leung, M. L., Unruh, A., Roh, W., Chen, K., Scheet, P., Vattathil, S., Liang, H., Multani, A., Zhang, H., Zhao, R., Michor, F., Meric-Bernstam, F. and Vavin, N. E. 2014. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 512(7513):155–160
128. Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol.* 7(2): 256–276
129. Weir B.S. and Hill W.G. 2002. Estimating F-statistics. *Annual Reviews Genetics.* 36 (1): 721–750.
130. Wen-Hsiung, L. 1997. Molecular evolution. Sinauer Associates Inc., Sunderland, MA.
131. Woerner, S., Yuan, Y., Benner, A., Korff, S., Doeberitz, M., Bork, P. 2009. SelTarbase, a database of human mononucleotidemicrosatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic Acids Research*; 1–8 doi:10.1093/nar/gkp839
132. Wong YF, Cheung TH, Lo KW, Yim SF, Chan LK, Buhard O, Duval A, Chung TK, Hamelin R (2006). Detection of microsatellite instability in endometrial cancer: advantages of a panel of five mononucleotide repeats over the National Cancer Institute panel of markers. *Carcinogenesis* 27(5): 951–955
133. Wright, J. M., and Bentzen, P. 1994. Microsatellites: Genetic markers for the future. *Reviews in Fish Biology and Fisheries.* 4: 384–388
134. Yee CS, Roodi N, Verrier CS, *et al.* 1994. Microsatellite instability and loss of heterozygosity in breast cancer. *Cancer Res.* 54:1641–1644.
135. Yuksel, E., Ozer, E., Kizildag, S., Sercan, O., Canda, T., and Sakizli, M. 2002. Microsatellite instability is not a common feature in medullary breast cancer. *The Breast.* Vol. 11(6): 473–477.
136. Zhang, J., Fujimoto, J., Zhang, J., Wedge, D. C., Song, X., Zhang, J., Seth, S., Chow, C-W., Cao, Y., Gumbs, C., Gold, K. A., Kalhor, N., Little, L., Mahadeshwar, C. M., Protopopov, A., Sun, H., Tang, J., Wu, X., Ye, Y., William, W. N., Lee, J. J., Heymach, J. V., Hong, W. K., Swisher, s., Wistuba, I. I., Futreal., P. A. 2014. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science.* Vol. 346(6206), pp. 256–259.
137. Zhao, Z., Yu, N., Fu, Y. X. and Li, W. H. 2006. Nucleotide variation and haplotype diversity in a 10-kb noncoding region in three continental human populations. *Genetics.* 174: 399–409.
138. Zhou X.P., Hoang J.M., Cottu P., Thomas G. and Hamelin R. 1997. Allelic profiles of mononucleotide repeat microsatellites in control individuals and colorectal tumors with and without replication errors. *Oncogene.* 20 (15): 1713–1718.

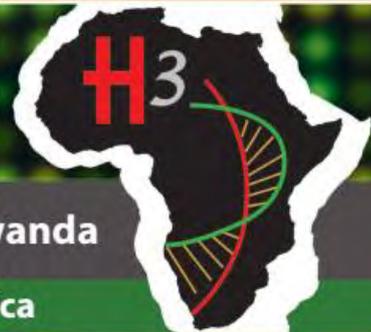
# POSTER



## Environmental Health in Africa: Opportunities to Expand Research Capacity in the H3Africa Consortium

Sunday, September 16, 2018 • Kigali, Rwanda

H3Africa — Human Heredity and Health in Africa



## RESUME :

Le cancer du sein est le cancer le plus répandu et la principale cause de décès de cancer chez les femmes dans le monde avec un taux d'incidence et de mortalité de 37,3% et de 17,8% respectivement en Afrique de l'Ouest. Au Nigéria, il s'agit du cancer le plus répandu et de la principale cause de décès par cancer chez les femmes, tandis qu'au Sénégal, il occupe le deuxième rang après le col de l'utérus, ce qui pose un problème de santé publique majeur. Dans cette étude, caractérisé et comparé l'évolution génétique du cancer du sein chez les femmes nigériennes et sénégalaises atteintes de la maladie à l'aide des marqueurs mononucléotidiques *NR-21* et *NR-24*, ainsi que leur association avec des caractéristiques clinico-pathologiques ainsi que les paramètres de reproduction de ces populations. Deux cent une (201) patientes atteintes d'un cancer du sein et confirmées histologiquement ont été recrutées dans deux hôpitaux de référence de chaque pays (102 patientes du Nigéria et 99 patientes du Sénégal) et cent échantillons de tissus mammaires (50 de chaque pays) ont été sélectionnés pour les analyses moléculaires. Les résultats des paramètres épidémiologiques ont indiqué une différence significative de l'âge à la ménarche ( $p < 0,001$ ) de l'âge de la première grossesse à terme ( $p < 0,001$ ), de la parité ( $p < 0,001$ ), de la contraception ( $p < 0,001$ ) et de l'état matrimonial ( $p = 0,005$ ) entre les deux pays. Nous avons observé un nombre d'haplotypes plus élevé au Sénégal (8 et 16 haplotypes pour *NR-21* et *NR-24* respectivement) qu'au Nigéria (4 et 12 haplotypes de *NR-21* et *NR-24* respectivement). 15% des cas de chaque pays étaient stables pour le marqueur *NR-21*, tandis que 70% et 82% des cas étaient instables pour le marqueur *NR-24* au Nigéria et au Sénégal. Une diversité haplotypique élevée et une diversité de nucléotides faible ont été observées pour chaque marqueur, ainsi que des valeurs négatives pour le test de neutralité qui indiquent une expansion récente de la population (cellules tumorales à croissance rapide). Ce phénomène n'est pas rare pour le cancer du sein en Afrique subsaharienne avec un type de tumeur agressive. L'analyse de la variance moléculaire (AMOVA) a révélé une forte variation génétique (91,5% et 88,6% pour *NR-21* et *NR-24*) et une faible différenciation génétique entre les populations (8,5% et 11,4% pour *NR-21* et *NR-24*) indiquant l'absence de différenciation génétique dans l'évolution des cancers du sein entre les deux populations étudiées. De même, l'association d'instabilité de MSI avec des paramètres cliniques n'a révélé aucune différence statistiquement significative sauf l'âge pour les Nigériennes ( $p < 0,05$ ) et le grade SBR pour les Sénégalaises ( $p = 0,04$ ) pour le marqueur *NR-24*. Cette étude met en exergue l'implication des instabilités des microsatellites (MSI) *NR-21* et *NR-24* dans l'évolution du tumeurs cancer du sein dans ces populations. En conclusion, des résultats montrer qu'il n'existe aucune différence entre le développement et la progression du cancer du sein au Nigeria et au Sénégal et que, le MSI joue un double rôle de biomarqueur prédictif et pronostique. L'instabilité observée ici devrait renforcer les efforts de recherche vers son exploration et son application clinique au Nigeria et au Sénégal.

**Mots clés :** *sein, cancer, microsatellite, instabilité, evolution*

## ABSTRACT

Breast cancer is the most prevalent cancer and the leading cause of cancer deaths among women worldwide with an incidence and mortality rate of 37.3% and 17.8% respectively in Africa. In Nigeria, it is the most common cancer and the leading cause of cancer deaths in women, while in Senegal it ranks second after cervix posing a major public health problem. In this study, characterized and compared the genetic evolution of breast cancer among Nigerian and Senegalese women with the disease using the *NR-21* and *NR-24* mononucleotide markers as well as their association with clinico-pathological features and also the reproductive parameters of these populations. Two hundred and one (201) histologically confirmed breast cancer patients were recruited from two reference hospitals in each country (102 female patients from Nigeria and 99 female patients from Senegal) and 100 breast tissue samples (50 from each country) were selected for molecular analyzes. The results of the epidemiological parameters indicated a significant difference between age at menarche ( $p < 0.001$ ), age of first term pregnancy ( $p < 0.001$ ), parity ( $p < 0.001$ ), contraception ( $p < 0.001$ ) and marital status ( $p = 0.005$ ) between the two countries. We observed a high number of haplotypes in Senegal (8 and 16 haplotypes for *NR-21* and *NR-24* respectively) than in Nigeria (4 and 12 haplotypes of *NR-21* and *NR-24*, respectively). 15% of the cases in each country were stable for the *NR-21* marker, while 70% and 82% of the cases were unstable for the *NR-24* marker in Nigeria and Senegal. High haplotypic diversity and low nucleotide diversity were observed for each marker, as well as negative values for the neutrality test that indicate a recent expansion of the population (fast-growing tumor cells). This phenomenon is not uncommon for breast cancer in sub-Saharan Africa with an aggressive type of tumor. Molecular Variance Analysis (AMOVA) revealed high genetic variation (91.5% and 88.6% for *NR-21* and *NR-24*) and low genetic differentiation between populations (8.5% and 11.6%). 4% for *NR-21* and *NR-24*) indicating the absence of genetic differentiation in the evolution of breast cancer between the two populations studied. Similarly, the association of MSI instability with clinical parameters revealed no statistically significant difference except for Nigerian women ( $p < 0.05$ ) and SBR grade for Senegalese women ( $p = 0.04$ ) for the *NR-24* marker. In conclusion, the results show no differences in the development and progression of breast cancer in Nigeria and Senegal and the involvement of MSI as predictive and prognostic biomarkers. The study will strengthen research efforts towards its exploration and clinical application of these markers in Nigeria and Senegal.

**Keywords:** *breast, cancer, microsatellite, instability, evolution*