

Titre: Stratégie d'analyse d'images pour l'évaluation de biopsies
cancéreuses marquées par nanoparticules plasmoniques

Auteur: Marjolaine Malgéri

Date: 2025

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Malgéri, M. (2025). Stratégie d'analyse d'images pour l'évaluation de biopsies
cancéreuses marquées par nanoparticules plasmoniques [Mémoire de maîtrise,
Polytechnique Montréal]. PolyPublie. <https://publications.polymtl.ca/71722/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/71722/>
PolyPublie URL:

**Directeurs de
recherche:** Michel Meunier
Advisors:

Programme: Génie biomédical
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Stratégie d'analyse d'images pour l'évaluation de biopsies cancéreuses marquées
par nanoparticules plasmoniques**

MARJOLAINE MALGÉRI

Institut de génie biomédical

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie biomédical

Décembre 2025

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Stratégie d'analyse d'images pour l'évaluation de biopsies cancéreuses marquées
par nanoparticules plasmoniques**

présenté par **Marjolaine MALGÉRI**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Benjamin DE LEENER, président

Michel MEUNIER, membre et directeur de recherche

Lucien WEISS, membre

REMERCIEMENTS

Je tiens à remercier mon directeur de recherche, le Professeur Michel Meunier, pour m'avoir permis de participer à ce beau projet. Merci pour sa confiance et son soutien tout au long de cette aventure. Un immense merci à Jennyfer d'être venue prêter main forte sur ce projet. Merci pour sa bonne humeur, sa bienveillance, et la qualité de son travail qui auront marqué cette dernière année. Le travail d'équipe est facile et agréable lorsque l'on a la chance d'avoir une telle coéquipière! Merci à Cécile d'avoir bâti les fondations de ce projet auquel je suis heureuse d'avoir contribué. Merci pour les belles opportunités offertes et les précieux conseils partagés.

Merci à toutes les personnes qui ont contribué à ce travail et m'ont permis d'en faire ce qu'il en est. Merci à Paul pour son aide avec tout ce qui commence par I et finit par A. Merci à Hubert pour toute l'aide qu'il m'a apporté avec Dragonfly. Merci à Jie pour l'automatisation du microscope. Merci à Flavie pour ses supers alliages. Merci à Julie pour sa disponibilité et son enthousiasme à réaliser de multiples déparaffinages et IHC. Merci à Charline pour son travail et sa résilience avec nos NPs parfois capricieuses. Merci à l'équipe de l'Université Libre de Bruxelles (ULB) pour leur collaboration et l'apport de leur calixarène sur ce projet. Merci à la Dre Dominique Trudel et son équipe du Centre de Recherche du Centre Hospitalier de Montréal (CR-CHUM) pour leur collaboration sur ce projet. Merci tout particulièrement à la Dre Roula Albadine pour son implication et son intérêt pour mon travail. J'ai été chanceuse de pouvoir apprendre de son expertise.

Plus largement, je souhaite remercier toute l'équipe du Laboratoire de Plasmonique et de Procédés par Laser (LP2L) avec qui j'ai eu beaucoup de plaisir à travailler. Merci à Isabelle et Amélie qui ont toujours su répondre à mes questions. Merci à Guillaume et Julien pour leur accueil à mon arrivée au LP2L et merci à Rémi et Grégoire pour ces derniers mois passés ici. Merci également à Yves pour son super travail auprès de nous. Merci aux voisins du JAB, et plus particulièrement à Nicolas et Annalisa pour les pauses agrémentées de bon café avec l'un, et de café un peu plus expérimental avec l'autre.

Finalement, merci à ma famille qui a su être présente malgré la distance et à Mehdi pour sa bienveillance et son soutien sans faille. Merci d'avoir eu la patience de regarder mes images à chaque fois que je rentrais fièrement à la maison en répétant «*Regarde comme c'est beau!*».

RÉSUMÉ

Le cancer du sein est une maladie hétérogène dont le diagnostic repose sur plusieurs étapes, incluant la détermination du sous-type moléculaire. Dans les cas de cancers HER2 positifs, caractérisés par une surexpression de la protéine HER2 responsable de la régulation et la prolifération cellulaire, les patientes peuvent bénéficier de thérapies ciblées anti-HER2 qui améliorent considérablement leur pronostic. Depuis 2019, un nouveau traitement est disponible pour les patientes présentant une faible surexpression de HER2, désignées comme cas HER2-low. Cependant, le diagnostic clinique actuel, basé sur la coloration par immunohistochimie (IHC) de prélèvements tissulaires, est particulièrement subjectif et incertain lorsqu'il s'agit d'évaluer ces cas de faible expression.

Dans ce contexte, le laboratoire du Professeur Michel Meunier propose d'utiliser des nanoparticules (NPs) plasmoniques bioconjuguées pour cibler spécifiquement une protéine d'intérêt et en quantifier l'expression dans divers échantillons biologiques. Grâce à leurs propriétés de diffusion optique, les nanoparticules plasmoniques apparaissent sous la forme de points brillants en microscopie optique, permettant leur dénombrement précis. Cette stratégie, nommée immunoplasmonique (IP) promet une évaluation quantitative de l'expression protéique, comblant ainsi une lacune importante des méthodes actuellement utilisées en clinique.

Ce mémoire s'inscrit dans un projet réunissant les efforts de plusieurs étudiants ayant pour objectif d'utiliser la méthode IP pour évaluer l'expression de HER2 dans divers échantillons biologiques. Ce travail vise en particulier à proposer une méthodologie d'analyse d'images de biopsies marquées par nanoparticules plasmoniques. L'imagerie des échantillons dans deux modalités d'illumination, permet d'évaluer l'attachement des NPs en connaissance du contexte biologique de chaque biopsie. L'utilisation d'un modèle d'apprentissage profond (*Deep Learning*, DL) pour la détection des NPs assure une détection robuste surpassant les performances d'algorithmes naïfs. La segmentation du tissu, puis des noyaux et de leur classification selon leur type cellulaire sont aussi assurés grâce à des modèles de DL et permettent d'extraire le contexte biologique. Enfin, chaque NP est analysée en fonction de la surface sur laquelle elle s'est fixée pour établir plusieurs métriques permettant d'évaluer l'attachement non spécifique, témoin de la qualité du marquage, et l'attachement spécifique permettant la quantification de la protéine HER2. Grâce à cette méthodologie et aux métriques qu'elle génère, il devient possible d'analyser chaque biopsie marquée par IP, d'une part pour assister les utilisateurs de la méthode IP à l'optimisation de leurs expériences, et d'autre part pour prouver le potentiel de la méthode IP pour l'évaluation quantitative de l'expression protéique.

À l'avenir, la méthodologie développée pourra être réutilisée lors des études de validation de la technique IP sur des échantillons de biobanque.

ABSTRACT

Breast cancer is a heterogeneous disease whose diagnosis relies on several steps, including the determination of the molecular subtype. In cases of HER2-positive cancers, characterized by overexpression of the HER2 protein involved in cell regulation and proliferation, patients can benefit from targeted anti-HER2 therapies that significantly improve their prognosis. Since 2019, a new treatment has also become available for patients exhibiting low HER2 expression, classified as HER2-low cases. However, the current clinical diagnosis, based on immunohistochemistry (IHC) staining of tissue samples, remains subjective and is particularly uncertain when evaluating these low-expression cases.

In this context, Professor Michel Meunier’s laboratory proposes using bioconjugated plasmonic nanoparticles (NPs) to specifically target a protein of interest and quantify its expression in various biological samples. Thanks to their optical scattering properties, plasmonic nanoparticles appear as bright spots under optical microscopy, allowing them to be counted accurately. This strategy, known as immunoplasmonics (IP), promises quantitative assessment of protein expression, thus filling a significant gap in the methods currently used in clinical practice.

This work is part of a project bringing together the efforts of several students with the aim of using the IP method to evaluate HER2 expression in various biological samples. More specifically, this project proposes a methodology for analyzing images of biopsies labeled with plasmonic nanoparticles. Imaging the samples in two illumination modes enables the assessment of NPs attachment while accounting for the biological context of each biopsy. The use of a Deep Learning (DL) model for NPs detection ensures a robust detection that surpasses the performance of naive algorithms. The segmentation of the tissue, then of the nuclei and their classification according to their cell type are also ensured thanks to DL models and allow the biological context to be extracted. Finally, each NP is analyzed according to the surface to which it has attached in order to establish several metrics for evaluating non-specific attachment, which indicates the quality of the labelling, and specific attachment, which allows quantification of the HER2 protein. Thanks to this methodology and the metrics it generates, it becomes possible to analyze each IP-labeled sample both to help users of the IP method optimize their experiments and to demonstrate the potential of IP for the quantitative evaluation of protein expression. In the future, the methodology developed can be reused in studies to validate the IP technique on biobank samples.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
RÉSUMÉ	iv
ABSTRACT	vi
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES SIGLES ET ABRÉVIATIONS	xvi
CHAPITRE 1 INTRODUCTION	1
CHAPITRE 2 REVUE DE LITTÉRATURE	5
2.1 Cancer du sein	5
2.1.1 Hétérogénéité des cancers du sein	5
2.1.2 Procédure de diagnostic	7
2.1.3 Statut HER2 : intérêt clinique et défis à relever	12
2.2 Immunoplasmonique	14
2.2.1 Propriétés optiques des nanoparticules plasmoniques	14
2.2.2 Fonctionnalisation des nanoparticules plasmoniques	16
2.2.3 Techniques d'imagerie et de quantification des nanoparticules plasmoniques	19
2.3 Intelligence Artificielle en pathologie digitale	22
2.3.1 Concepts fondamentaux de l'Intelligence Artificielle	22
2.3.2 Modèles de Deep Learning	24
2.3.3 Outils de segmentation et de classification cellulaire	26
CHAPITRE 3 MATÉRIEL ET MÉTHODES POUR LA PRÉPARATION DES ÉCHANTILLONS ET L'ACQUISITION DES IMAGES À ANALYSER	31
3.1 Préparation des échantillons marqués par immunoplasmonique	31
3.1.1 Nanoparticules utilisées	31
3.1.2 Échantillons utilisés	36
3.2 Acquisition des images	43
3.2.1 Imagerie des TMA	44
3.2.2 Imagerie des CMA	47

3.2.3	Imagerie des produits BCS	49
CHAPITRE 4 OUTILS DE SEGMENTATION DES IMAGES		50
4.1	Génération des masques de vérité terrain	50
4.1.1	Méthodologie pour la génération des masques de vérité terrain	50
4.1.2	Vérité terrain pour la segmentation des NPs	51
4.1.3	Vérité terrain pour la segmentation du tissu	51
4.1.4	Jeux de données générés	52
4.1.5	Méthodologie pour l'évaluation des outils de segmentation	54
4.2	Limites de l'algorithme naïf ImageJ	57
4.3	Modèles de Deep Learning développés avec Dragonfly	61
4.3.1	Entraînement des modèles	61
4.3.2	Evaluation des modèles	62
4.4	Segmentation et classification des noyaux cellulaires	68
4.4.1	Segmentation des noyaux	69
4.4.2	Classification des noyaux	70
4.4.3	Évaluation de la classification des noyaux	70
CHAPITRE 5 ANALYSE DES MATRICES DE TISSUS		74
5.1	Méthodologie sur TMA	74
5.1.1	Flux de travail	74
5.1.2	Métriques générées	77
5.2	Résultats sur TMA	90
5.2.1	Résultats généraux des 7 lames analysées	90
5.2.2	Utilisation du marquage IP pour affiner le score IHC	98
5.3	Translation préliminaire aux CMA	99
5.3.1	Méthodologie	99
5.3.2	Exemple de résultats	101
CHAPITRE 6 IMMUNOPLASMONIQUE SUR LES STANDARDS DE LA COM- PAGNIE BCS		103
6.1	Preuve de concept - Attachement des NPs aux contrôles BCS	103
6.2	Utilisation des contrôles BCS comme contrôle qualité	106
6.2.1	Contrôle qualitatif	106
6.2.2	Vers un contrôle quantitatif	108
6.3	Vers la caractérisation de la plage dynamique de la méthode IP avec les cali- brateurs	110

6.3.1	Réponse dynamique de la coloration IHC	110
6.3.2	Vers la réponse dynamique du marquage IP	112
CHAPITRE 7	DISCUSSION	113
7.1	Outils de segmentation	113
7.2	Gestion des agrégats	117
7.3	Réflexion sur les métriques	117
7.3.1	Importance de l’aspect qualitatif	117
7.3.2	Non exhaustivité des métriques	118
7.4	Réflexion sur les produits BCS	118
7.4.1	Limites liées à l’utilisation d’un produit commercial	118
7.4.2	Approche future pour l’évaluation de la plage dynamique	119
CHAPITRE 8	CONCLUSION	120
8.1	Synthèse des travaux	120
8.2	Limitations	122
8.2.1	Imagerie avec focus manuel	122
8.2.2	Manque d’automatisation	122
8.2.3	Faible intégration	122
8.3	Améliorations futures	123
8.3.1	Proposer une solution entièrement automatisée et intégrée	123
8.3.2	Ajouter la classification des NPs selon leur couleur	123
8.3.3	Intégrer une métrique pour décrire la présence de motifs membranaires	123
RÉFÉRENCES		124

LISTE DES TABLEAUX

Tableau 3.1	Scores IHC attribués par deux pathologistes.	37
Tableau 3.2	Paramètres optiques en fonction des objectifs utilisés.	43
Tableau 3.3	Paramètres d’acquisition utilisés pour imager les TMA.	44
Tableau 3.4	Paramètres d’acquisition utilisés pour imager les CMA.	48
Tableau 4.1	Description des jeux de données générés.	52
Tableau 4.2	Résultat de l’évaluation du modèle PreTrained100_UNet obtenus avec les deux jeux de d’évaluation, comparé aux performances de <i>Find Maxima</i>	63
Tableau 4.3	Performances des trois modèles après un entraînement sur 10 images du jeu de données B uniquement.	63
Tableau 4.4	Performances des trois modèles à la fin de l’entraînement sur les 300 images du jeu de données B.	65
Tableau 5.1	Moyenne du nombre de NPs détectées par carotte pour chaque lame.	91
Tableau 6.1	Combinaisons de NPs et de PEG investiguées pour la fonctionnalisation des NPs avec EPR.	103
Tableau 6.2	Clones utilisés pour tenter de fonctionnaliser les trois types de NPs, et cibler les protéines HER2, ER et PR à la surface des billes de contrôle.	105

LISTE DES FIGURES

Figure 1.1	Corrélation entre le marquage immunoplasmonique et immunohisto-chimique pour un tissu de sein.	2
Figure 2.1	Schéma de l'anatomie macroscopique du sein.	6
Figure 2.2	Coupes successives de biopsie de cancer du sein, colorées avec deux techniques différentes.	8
Figure 2.3	Schéma du fonctionnement de l'IHC et coloration visible en microscopie optique.	9
Figure 2.4	Représentation schématique d'une sonde utilisée pour le FISH et vi-sualisation des sondes en microscopie à fluorescence.	10
Figure 2.5	Consignes d'évaluation du statut HER2 de l'ASCO/CAP, avec ajout des catégories émergentes HER2-low et HER2-ultralow.	13
Figure 2.6	Effet du diamètre de la NP sur la dépendance spatiale du champ incident.	15
Figure 2.7	Illustration d'un SH-PEG-COOH dans les conformations champignon et brosse.	18
Figure 2.8	Structure du calixarène utilisé pour fonctionnaliser les NPs d'or. . . .	19
Figure 2.9	Modalités d'illumination permettant la visualisation des structures bio-logiques et des NPs.	21
Figure 2.10	Architecture du U-Net.	25
Figure 2.11	Architecture du ViT.	26
Figure 2.12	Méthode utilisée par StarDist pour segmenter les noyaux grâce au cal-cul de la probabilité d'appartenance et des distances radiales.	27
Figure 2.13	Méthode utilisée par CellPose pour segmenter les cellules grâce à la prédiction de champs vectoriels horizontaux et verticaux.	28
Figure 2.14	Méthode utilisée par CellViT pour segmenter et classer les noyaux. .	30
Figure 3.1	Représentation schématique d'une Au90NP-NanoBrand-X4C4-EPR. .	33
Figure 3.2	Représentation schématique d'une Au90NP-Nanopartz-LAPEG-EPR.	33
Figure 3.3	Absorbances normalisées des Au90NPs-Nanopartz-LAPEG-EPR et des Au90NPs-NanoBrand-X4C4-EPR mesurées par spectroscopie UV-visible. La demi-largeur à mi-hauteur (<i>Half Width at Half Maximum</i> , HWHM) a été calculée en utilisant le plateau situé entre 350 et 450 nm comme base pour le calcul de la prominence.	34
Figure 3.4	Absorbances normalisées des alliages mesurées par spectroscopie UV-visible.	35

Figure 3.5	Représentation schématique des alliages.	35
Figure 3.6	Carte du TMA utilisé.	36
Figure 3.7	Carottes de cancer du sein colorées par H&E montrant la diversité de morphologies observables.	37
Figure 3.8	Carotte de cancer du sein (F3) colorée par H&E.	38
Figure 3.9	Images des IHC et scores HER2 des sept carottes du CMA HER2Low Array Science.	40
Figure 3.10	Résultats attendus et exemple d'analyse pour les IHControls®.	41
Figure 3.11	Exemple de résultats attendus et d'analyse pour les IHCalibrators®.	42
Figure 3.12	Représentation schématique d'une section de 4µm dans un bloc de TMA.	44
Figure 3.13	Reconstructions de la carotte B2 (Prostate) acquise en BF et en SI.	45
Figure 3.14	Lame 14 acquise dans les deux modalités d'illumination.	46
Figure 3.15	Représentation schématique d'une section de 4µm dans un bloc de CMA.	47
Figure 3.16	Exemple d'images BF (A) et SI (B) pour une carotte de CMA (C7 - lignée BT-474)	48
Figure 3.17	Paramètres d'acquisition pour l'imagerie des billes BCS et images obtenues.	49
Figure 4.1	Génération du masque de vérité terrain pour la segmentation des NPs.	51
Figure 4.2	Génération du masque de vérité terrain pour la détection du tissu.	52
Figure 4.3	Exemples de parcelles issues des jeux de données A et B.	53
Figure 4.4	Exemple de parcelles issues du jeu de données C pour la segmentation du tissu.	54
Figure 4.5	Représentation schématique de la classification en VP, FP et FN	55
Figure 4.6	Évaluation de la détection du tissu.	56
Figure 4.7	Exemple d'images pour les 4 groupes de densité des données d'évaluation.	57
Figure 4.8	Détections de l'algorithme <i>Find Maxima</i> comparées aux positions vérité terrain.	58
Figure 4.9	Sensibilité, précision et F1-score évalués pour chaque groupe de densité en fonction de la prominence.	59
Figure 4.10	Prominence optimale permettant de maximiser la sensibilité, la précision et le F1-score pour chaque densité.	60
Figure 4.11	Sensibilité, précision et F1-score moyens de la détection <i>Find Maxima</i> sur les 100 images d'évaluation, en fonction de la prominence utilisée.	60
Figure 4.12	Évolution des métriques de performance pour les différents modèles	64
Figure 4.13	Détection des NPs dans l'image SI grâce au modèle SegNPs.	66
Figure 4.14	Évolution du score Dice au cours de l'entraînement pour les deux modèles.	67

Figure 4.15	Évolution du score IoU au cours de l'entraînement pour les deux modèles.	67
Figure 4.16	Génération du masque de tissu grâce au modèle TissuDetection appliqué sur l'image BF.	68
Figure 4.17	Résultat de la segmentation des noyaux par les modèles Stardist et CellPose-SAM.	69
Figure 4.18	Résultat de la segmentation des noyaux et de leur classification selon leur type cellulaire par le modèle CellViT++.	71
Figure 4.19	Corrélation entre la tumeur surexprimant HER2 visible en IHC (coloration brune) et les noyaux des cellules épithéliales néoplasiques prédits par CellViT++ (en vert).	72
Figure 5.1	Carotte imagée en champ clair avec la superposition des noyaux détectés par CellViT, et l'expansion des cytoplasmes extrapolée par watershed.	75
Figure 5.2	Histogramme montrant la distribution du diamètre des NPs détectées dans les 7 lames analysées.	76
Figure 5.3	Positions des NPs détectées sur les cellules.	76
Figure 5.4	Graphe montrant le taux d'agrégation pour toutes les carottes d'une lame.	78
Figure 5.5	Image montrant en rose les NPs sur le tissu et en bleu les NPs sur le substrat.	79
Figure 5.6	Graphe montrant la densité de NPs sur le substrat pour toutes les carottes d'une lame.	79
Figure 5.7	Images montrant les NPs sur les zones cancéreuse et non cancéreuse.	80
Figure 5.8	Graphe montrant la densité de NPs sur la zone non cancéreuse pour toutes les carottes d'une lame.	81
Figure 5.9	Graphe montrant la densité de NPs sur les noyaux pour toutes les carottes d'une lame.	81
Figure 5.10	Graphe montrant l'indice d'affinité nucléaire calculé pour toutes les carottes d'une lame.	82
Figure 5.11	Densités de NPs sur les trois régions d'intérêt.	84
Figure 5.12	Densités de NPs sur les trois régions d'intérêt - Sein uniquement. . .	85
Figure 5.13	Densité moyenne par score IHC et par tissu.	85
Figure 5.14	Nombre moyen de NPs par cellule en fonction du type cellulaire. . . .	87
Figure 5.15	Nombre moyen de NPs par cellule - Sein seulement.	87
Figure 5.16	Moyenne du nombre moyen de NPs par cellule épithéliale néoplasique par score IHC et par tissu.	88
Figure 5.17	Indice d'affinité membranaire pour les cellules épithéliales néoplasiques.	89

Figure 5.18	Marquage membranaire autour des noyaux des cellules épithéliales néoplasiques.	89
Figure 5.19	Évaluation de l'attachement global sur les lames analysées.	90
Figure 5.20	Taux d'agrégation pour toutes les lames analysées.	92
Figure 5.21	Résumé des densités de NPs sur différentes zones.	93
Figure 5.22	Densité moyenne sur la zone membranaire des cellules épithéliales néoplasiques, calculée pour tous les scores IHC de chaque lame.	95
Figure 5.23	Nombre moyen de NPs par cellule épithéliale néoplasique, calculé pour tous les scores IHC de chaque lame.	97
Figure 5.24	Classification des carottes équivoques en IHC grâce au marquage IP. .	98
Figure 5.25	Lame 9 du CMA imagée en BF et en SI. Objectif 40x à air NA=0,6 + zoom 1,5x.	99
Figure 5.26	Comparaison des résultats de la segmentation des cellules par CellPose (A), et par le modèle de détection du TissuDetection (B). 1 : CellPose performe bien pour détecter et séparer les cellules adjacentes, TissuDetection ne fait pas la distinction entre les différentes cellules. 2 : CellPose restreint certaines cellules à leur noyau. 3 : CellPose manque la détection de cellules au complet.	100
Figure 5.27	Comparaison des densités obtenues en utilisant <i>TissuDetection</i> ou <i>CellPose</i> pour détecter la surface occupée par les cellules.	101
Figure 5.28	Comparaison de la densité de NPs sur les cellules en fonction de la quantité d'anticorps utilisés pour la bioconjugaison.	102
Figure 6.1	Trois types de NPs fonctionnalisées avec EPR ciblant les billes de contrôle.	104
Figure 6.2	Billes de contrôles utilisées comme contrôle de routine sur les lames de TMA.	107
Figure 6.3	Densité moyenne d'attachement sur la zone membranaire des cellules épithéliales néoplasiques, par score IHC EPR pour toutes les carottes des différentes expériences.	107
Figure 6.4	Quantification des NPs fixées sur une bille de contrôle : de la pile d'images acquise en microscopie à réflexion à une représentation 3D de la bille recouverte de NPs.	109
Figure 6.5	Contrôle quantitatif de l'attachement des NPs sur les billes de contrôle au cours des expériences réalisées.	110
Figure 6.6	Billes de contrôle du calibrateur colorées par IHC.	111
Figure 6.7	Réponse dynamique de la coloration IHC pour des concentrations croissantes de la protéine HER2.	112

Figure 7.1	Variabilité des images acquises et des résultats de segmentation en fonction des systèmes d'acquisition.	115
Figure 7.2	Variabilité des images acquises et des résultats de segmentation en fonction des lames.	116
Figure 8.1	Schéma représentant la stratégie développée pour l'analyse des biopsies marquées par nanoparticules plasmoniques.	121

LISTE DES SIGLES ET ABRÉVIATIONS

ADC	<i>Antibody-Drug Conjugate</i> , conjugué anticorps-médicament
ADN	Acide désoxyribonucléique
ASCO/CAP	<i>American Society of Clinical Oncology / College of American Pathologists</i> , Société Américaine d'Oncologie Clinique et Collège Américain des Pathologistes
BCS	<i>Boston Cell Standards</i>
BF	<i>Bright Field</i> , champ clair
BSA	<i>Bovine Serum Albumin</i> , sérum d'albumine bovine
CEP17	Centromère du chromosome 17
CISH	<i>Chromogenic In Situ Hybridization</i> , hybridation <i>in situ</i> par chromogène
CMA	<i>Cells Microarray</i> , matrice de cellules
CNN	<i>Convolutional Neural Network</i> , réseau neuronal convolutionnel
CRCHUM	Centre de Recherche du Centre Hospitalier de l'Université de Montréal
DAB	Diaminobenzidine
DEL	Diode Électro-Luminescente
DF	<i>Dark Field</i> , champ sombre
DL	<i>Deep Learning</i> , apprentissage profond
EDC	1-éthyl-3-(3-diméthylaminopropyl)carbodiimide
EGFR	<i>Epidermal Growth Factor Receptor</i> , récepteur du facteur de croissance épidermique
ER	<i>Estrogen Receptor</i> , récepteur à œstrogène
FFPE	<i>Formalin-Fixed, Paraffin-Embedded</i> , fixé au formol et inclus en paraffine
FDA	<i>Food and Drug Administration</i>
FISH	<i>Fluorescence In Situ Hybridization</i> , hybridation <i>in situ</i> par fluorescence
FP	Faux positif
GT	<i>Ground Truth</i> , vérité terrain
H&E	Hématoxyline et Éosine
HER2	<i>Human Epidermal Growth Factor Receptor 2</i> , récepteur du facteur de croissance épidermique humain 2
HWHM	<i>Half Width at Half Maximum</i> , demi-largeur à mi-hauteur
HRP	<i>Horseradish Peroxidase</i> , peroxydase de raifort
IA	Intelligence Artificielle
IDC	<i>Invasive Ductal Carcinoma</i> , carcinome canalaire invasif

IF	Immunofluorescence
IHC	Immunohistochimie
ILC	<i>Invasive Lobular Carcinoma</i> , carcinome lobulaire invasif
IP	Immunoplasmonique
IRIC	Institut de Recherche en Immunologie et en Cancérologie
IRM	Imagerie par Résonance Magnétique
LFA	<i>Lateral Flow Assay</i> , test à flux latéral
LOD	<i>Limit of Detection</i> , limite de détection
LP2L	Laboratoire de Plasmonique et de Procédés par Laser
LSPR	<i>Localized Surface Plasmon Resonance</i> , résonance localisée des plasmons de surface
ML	<i>Machine Learning</i> , apprentissage automatique
NA	<i>Numerical Aperture</i> , ouverture numérique
NPs	Nanoparticules
NST	<i>No Special Type</i> , sans type particulier
OMS	Organisation Mondiale de la Santé
PBS	<i>Phosphate Buffered Saline</i> , solution saline tamponnée au phosphate
PBST	<i>Phosphate Buffered Saline with Tween</i> , solution saline tamponnée au phosphate avec Tween
PD-1	<i>Programmed Death Protein 1</i>
PD-L1	<i>Programmed Death - Ligand 1</i>
PEG	Polyéthylène glycol
PR	<i>Progesterone Receptor</i> , récepteur à progestérone
QBCF	<i>Quebec Breast Cancer Foundation</i> , Fondation du cancer du sein du Québec
RGB	<i>Red Green Blue</i> , rouge vert bleu
SAM	<i>Segment Anything Model</i>
SI	<i>Side Illumination</i> , illumination latérale
sulfo-NHS	<i>N-Hydroxysulfosuccinimide</i>
TMA	<i>Tissue Microarray</i> , matrice de tissus
VP	Vrai positif
VN	Vrai négatif
WSI	<i>Whole Slide Imaging</i> , imagerie de lames entières

CHAPITRE 1 INTRODUCTION

Selon les estimations de la Société canadienne du cancer, le cancer du sein représente à lui seul 12 % des cancers diagnostiqués [1]. Le cancer du sein est une maladie particulièrement hétérogène dont le diagnostic précis permet d’orienter le choix thérapeutique vers le traitement offrant le meilleur pronostic. Ainsi, toutes les patientes atteintes d’un cancer du sein doivent effectuer un prélèvement tissulaire, appelé biopsie, qui permet d’observer la tumeur au microscope et de la caractériser. La méthode la plus couramment utilisée pour la caractérisation d’une tumeur est l’immunohistochimie (IHC). Cette méthode permet de colorer l’échantillon tissulaire afin de révéler l’expression de marqueurs d’intérêt, notamment le récepteur du facteur de croissance épidermique humain 2 (HER2, *Human Epidermal growth factor Receptor 2*) et les récepteurs hormonaux (œstrogène ER et progestérone PR) dans le cadre du cancer du sein. Cette évaluation est cruciale, car elle permet de déterminer le profil moléculaire de la tumeur, élément clé pour choisir le traitement le plus approprié. Cependant, l’IHC est souvent critiquée pour son manque de standardisation et la subjectivité de son interprétation, qui entraînent des taux élevés de discordance inter-pathologistes, en particulier dans les cas de faible expression des marqueurs. Or, l’émergence de nouveaux traitements efficaces chez les patientes présentant une faible surexpression de HER2 rend plus que jamais nécessaire l’évaluation précise et fiable de l’expression de ce marqueur.

Pour relever ce défi, le Laboratoire de Plasmonique et de Procédés Laser (LP2L) dirigé par le Professeur Meunier propose d’utiliser des nanoparticules (NPs) plasmoniques fonctionnalisées pour détecter une protéine d’intérêt. Cette méthode, intitulée immunoplasmonique (IP) est basée sur une réaction immunologique couplée aux propriétés de diffusion optique des NPs qui permettent de les visualiser comme des points brillants en microscopie optique. Avec la méthode IP, l’évaluation des marqueurs d’intérêt du cancer du sein devient quantitative : il est possible de compter les NPs une par une et d’évaluer l’expression du marqueur en fonction du nombre de NPs qui se sont fixées sur l’échantillon.

Des preuves de concept réalisées sur des tissus ont déjà permis d’observer une corrélation entre les marquages IHC et IP, révélant le lien entre expression protéique et attachement des NPs. La Figure 1.1 montre un tissu de sein dont l’expression du marqueur HER2 est hétérogène. L’attachement des NPs corrèle avec la coloration brune de l’IHC. Là où les NPs se fixent, la coloration brune de l’IHC est présente et indique une surexpression du marqueur HER2. Au contraire, dans les zones où le tissu n’est pas coloré en brun, les NPs sont absentes.

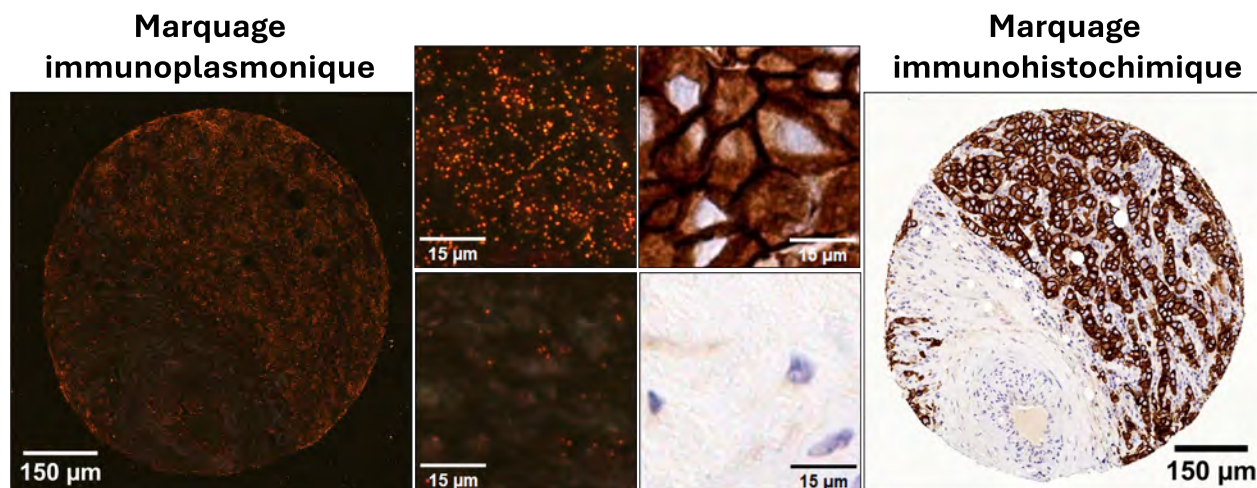


FIGURE 1.1 Corrélation entre le marquage immunoplasmonique et immunohistochimique pour un tissu de sein. Les deux images proviennent de coupes successives d'une même carotte de tissu appartenant à un bloc de tissu microarray (TMA).

Cependant, pour répondre à l'ambition initiale d'une analyse quantitative qui permettrait l'évaluation objective et précise du niveau d'expression protéique, il est essentiel de développer une méthodologie d'analyse complète des biopsies marquées par IP. Cette analyse doit permettre non seulement de quantifier l'attachement des NPs, mais aussi d'interpréter ces résultats dans le contexte biologique du tissu observé. Les travaux présentés dans ce mémoire visent à **établir une chaîne d'analyse complète, depuis l'acquisition des images jusqu'à la description quantitative de l'expression du marqueur HER2**. La stratégie développée est spécifiquement conçue pour l'analyse de carottes de tissu organisées en matrices de tissus (*tissue microarray*, TMA), mais peut aisément être adaptée à d'autres types d'échantillons biologiques.

Le développement d'une telle chaîne d'analyse est organisé autour de trois objectifs principaux :

Le premier objectif vise à **établir un protocole standardisé pour l'acquisition des images**, garantissant la capture des informations pertinentes pour l'analyse dans des conditions optimales. Chaque échantillon est imagé selon deux modalités d'illumination complémentaires. L'acquisition en champ sombre, sous illumination latérale (*Side Illumination*, SI) permet de capturer le signal des NPs. L'acquisition en champ clair (*Bright Field*, BF) permet quant à elle de capturer le contexte biologique de l'échantillon visible grâce à une contre coloration hématoxyline et éosine (H&E). Dans le cas de l'acquisition des carottes de TMA, un balayage dans le plan est nécessaire pour imager la carotte dans son intégralité.

Le second objectif consiste à **développer des outils de segmentation performants**, capables d'extraire automatiquement les informations pertinentes des images acquises dans les deux modalités. L'algorithme Find Maxima d'ImageJ, couramment utilisé pour la détection de maxima locaux dans des images, présente rapidement des limites en termes de performances et d'adaptabilité à la variabilité des images. L'approche proposée pour assurer la détection des NPs repose donc sur le développement de modèles d'apprentissage profond (Deep Learning, DL). Un premier modèle permet de détecter les NPs dans les images SI. Un second modèle permet d'isoler le tissu dans les images BF. Afin d'enrichir l'analyse et d'exploiter pleinement le contexte biologique des images BF, un modèle pré-existant est également utilisé afin de détecter et de classer les noyaux présents dans le tissu selon leur type cellulaire.

Enfin, le dernier objectif est de **générer des métriques permettant de décrire quantitativement les échantillons** à partir des résultats de segmentation. Les outils développés dans le cadre du second objectif sont appliqués aux images acquises de sorte à associer, pour chaque échantillon, la position des NPs au contexte biologique correspondant. Ces informations servent à calculer différents indicateurs quantitatifs reflétant l'expression protéique, conformément à l'objectif initial de la méthode IP. Par ailleurs, d'autres métriques permettent d'évaluer la qualité du marquage IP en décrivant notamment le niveau d'agrégation des NPs et l'attachement non spécifique. Ces métriques sont précieuses pour assurer un contrôle de la qualité des marquages effectués et pour optimiser la méthode IP.

Après cette Introduction, le mémoire est constitué de sept Chapitres.

Le Chapitre 2 propose une revue de littérature présentant les concepts clés sur lesquels repose ce projet, depuis le contexte clinique jusqu'aux outils d'intelligence artificielle utilisés en pathologie digitale, en passant par les principes fondamentaux de la méthode IP.

Le Chapitre 3 décrit les échantillons analysés ainsi que le protocole employé pour l'acquisition des images.

Le Chapitre 4 est consacré au développement des outils de segmentation. Il présente d'abord l'évaluation des performances de l'algorithme Find Maxima, puis le développement des modèles de DL.

Le Chapitre 5 expose le flux de travail complet mis en place pour l'analyse des biopsies et la génération des métriques. Il comprend notamment l'analyse des sept lames étudiées dans le cadre de ce travail, ainsi qu'une section dédiée à l'adaptation de la méthodologie à l'analyse d'échantillons cellulaires.

Le Chapitre 6 constitue une exploration parallèle réalisée à partir des standards de la compagnie Boston Cell Standards (BCS). Les objectifs de ce chapitre sont d'appliquer la méthodologie d'analyse dans un contexte tridimensionnel, de mettre en place des méthodes de contrôle pour la méthode IP et de déterminer sa plage dynamique.

Enfin, le Chapitre 7 propose une discussion critique sur le travail effectué et ses limites, tandis que le Chapitre 8 vient conclure ce mémoire.

CHAPITRE 2 REVUE DE LITTÉRATURE

2.1 Cancer du sein

Au Canada, on estime qu'en 2024, 30 500 femmes ont reçu un diagnostic de cancer du sein. Cela représente 25 % des cancers diagnostiqués chez les femmes, ce qui fait du cancer du sein le cancer le plus diagnostiqué chez les femmes [2]. Par conséquent, le cancer du sein a un impact majeur en santé publique et constitue l'objet de nombreuses recherches, tant pour mieux comprendre les mécanismes de la maladie que pour développer de nouveaux outils diagnostiques et thérapeutiques.

Dans cette première section, nous présenterons le cancer du sein et ses différentes formes. Nous verrons comment le diagnostic permet de caractériser la maladie et d'orienter vers la stratégie thérapeutique la plus adaptée. Nous détaillerons ensuite le statut HER2, central dans le cadre de ce travail.

2.1.1 Hétérogénéité des cancers du sein

Le cancer du sein est une maladie particulièrement hétérogène, qui se manifeste à travers un vaste spectre de variations morphologiques et moléculaires à l'origine d'une grande diversité de manifestations cliniques et de réponses au traitement [3–6]. Au cours des dernières décennies, d'importants efforts de recherche ont été consacrés à mieux comprendre et caractériser ces différents types de cancers du sein, afin de pouvoir les classer plus précisément et développer des approches thérapeutiques personnalisées, offrant ainsi de meilleures perspectives de traitement et de survie.

Anatomie macroscopique

La Figure 2.1 montre l'anatomie macroscopique du sein. Les lobules sont les unités sécrétoires productrices de lait. Les canaux galactophores relient ces glandes entre elles et permettent d'acheminer le lait jusqu'au mamelon. Lobules et canaux sont entourés de tissus conjonctifs et graisseux [7, 8]. La plupart des cancers du sein proviennent des cellules épithéliales qui tapissent les lobules et les canaux. Lorsque ces dernières se multiplient de manière incontrôlée, elles forment des tumeurs. On utilise le terme **carcinome** pour désigner les cancers originaires des cellules épithéliales, en opposition au sarcome par exemple qui prend naissance dans les tissus conjonctifs ou dans les tissus de soutien (os, muscles, graisse, etc.). Les tumeurs peuvent rester cloisonnées dans leur emplacement initial, ou infiltrer les tissus environnants.

Si la tumeur continue à se propager, elle peut atteindre le système sanguin ou lymphatique et migrer vers d'autres parties du corps où elle sera appelée métastase [9].

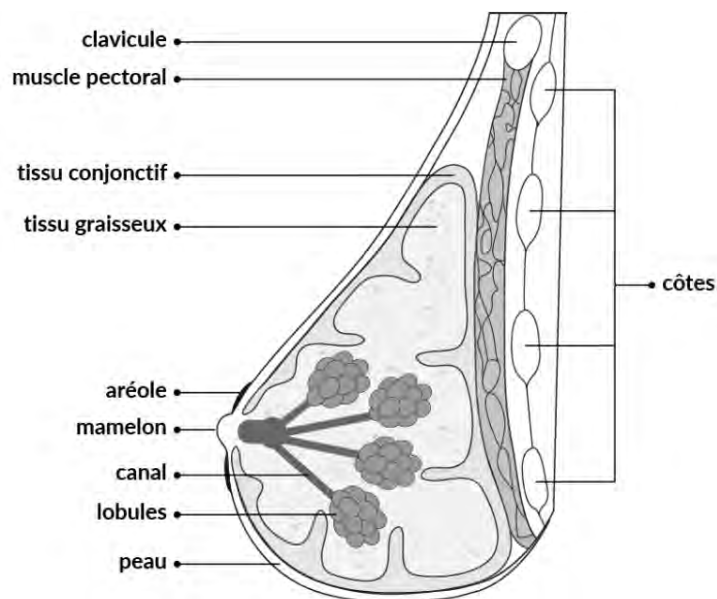


FIGURE 2.1 Schéma de l'anatomie macroscopique du sein. Tiré du site internet de la Société canadienne du cancer [7].

Types histologiques

Le type histologique est une classification morphologique basée sur l'observation microscopique de l'organisation et de la localisation des cellules. Si les cellules tumorales sont confinées dans l'épithélium, on parle de carcinome **in situ** [3]. Si les cellules tumorales ont envahi le stroma (ensemble formé du tissu conjonctif, des vaisseaux sanguins, de cellules immunitaires) environnant, on parle alors de carcinome **invasif** ou **infiltrant** [3]. Selon l'Organisation Mondiale de la Santé (OMS), on compte plus de 21 sous-types de carcinomes invasifs [4, 10]. En fonction de leur localisation (lobule ou canal), on distingue deux grands types de carcinomes invasifs. Les carcinomes canauxiaux invasifs (*invasive ductal carcinoma*, IDC) aussi appelés "sans type particulier" (*no special type*, NST) représentent 40 à 75 % des cas [4, 5]. Les carcinomes restants sont des carcinomes lobulaires invasifs (*invasive lobular carcinoma*, ILC). Il existe ensuite une très grande variété de types plus rares, dits spéciaux, caractérisés par des comportements cellulaires et moléculaires particuliers (carcinome médullaire, carcinome métaplasique, carcinome apocrine, carcinome mucineux, etc. [5, 6]). La variété de types histologiques explique la diversité des structures microscopiques que l'on observera dans les TMA présentés dans le Chapitre 3.

Types moléculaires

Le type moléculaire est une classification basée sur le profil d'expression des gènes et des récepteurs présents dans les cellules tumorales. Les marqueurs systématiquement analysés dans le cadre du cancer du sein sont les récepteurs hormonaux (récepteurs à l'œstrogène ER et à la progestérone PR), le récepteur du facteur de croissance épidermique humain 2 (HER2, *Human Epidermal growth factor Receptor 2*). Le marqueur de prolifération cellulaire Ki-67, bien que non obligatoire, est aussi souvent investigué [3, 4]. Selon la combinaison de l'expression de ces marqueurs, on distingue quatre types moléculaires principaux [5] :

- **Luminal A** : 40 à 50 % des cas, ER positif et/ou PR positif, HER2 négatif, prolifération faible (Ki-67 faible). Ce type est associé à un pronostic favorable.
- **Luminal B** : 20 à 30 % des cas, ER positif et/ou PR positif, HER2 positif ou Ki-67 élevé. Le pronostic est légèrement moins favorable que Luminal A.
- **HER2-positif** : 15 à 20 % des cas, ER et PR négatifs, HER2 positif. Ces cancers ont un pronostic plus sévère, mais peuvent bénéficier de thérapies ciblant HER2.
- **Triple négatif / basal-like** : 10 à 20 % des cas, ER négatif, PR négatif, HER2 négatif. Ce type est souvent plus agressif et présente moins d'options thérapeutiques ciblées.

Cette classification moléculaire est essentielle pour guider la stratégie thérapeutique et personnaliser le traitement.

2.1.2 Procédure de diagnostic

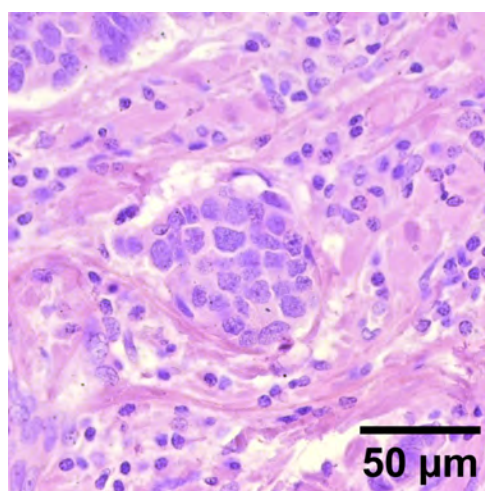
Le diagnostic du cancer du sein débute souvent par une suspicion, généralement lors de la découverte d'une masse à la palpation ou au cours d'une mammographie de dépistage. La détection précoce constitue un enjeu majeur, car elle améliore considérablement le pronostic [11, 12].

Afin de confirmer la présence d'une lésion, des examens d'imagerie médicale (mammographie, échographie, IRM, etc. [11]) sont d'abord réalisés. Ils permettent de visualiser la structure du sein et de détecter d'éventuelles masses ou anomalies. Toutefois, le diagnostic définitif ne peut être posé qu'après l'analyse d'une biopsie réalisée par un pathologiste. Le prélèvement tissulaire est généralement effectué à l'aide d'une fine aiguille, afin de limiter l'invasivité de la procédure. Les échantillons sont ensuite fixés dans du formol et inclus dans de la paraffine pour préserver la morphologie et assurer une conservation longue durée (on parle d'échantillons FFPE pour *Formalin-Fixed, Paraffin-Embedded*). Pour analyser la biopsie, des fines sections sont découpées dans le bloc FFPE et montées sur des lames de microscopie. Chaque lame doit subir une série de traitements (chauffage, déparaffinage, réhydratation, récupéra-

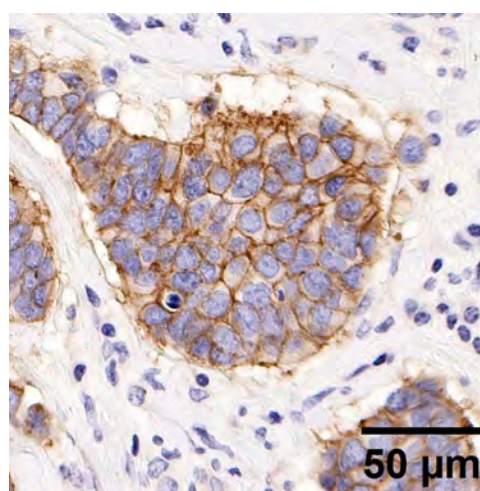
tion antigénique) avant d'être utilisable. Grâce à différentes colorations, le pathologiste peut analyser le tissu, confirmer le diagnostic, et le préciser en indiquant le type histologique et moléculaire du cancer étudié [11, 12].

Contre-coloration à l'hématoxyline-éosine (H&E)

Parmi les colorations de routine, on retrouve la contre-coloration à l'hématoxyline-éosine (H&E). L'hématoxyline est un colorant basique qui colore en bleu violacé les structures acides, telles que l'ADN contenu dans les noyaux cellulaires. L'éosine, au contraire, est un colorant acide qui colore en rose les structures basiques, comme les protéines cytoplasmiques. La coloration H&E permet ainsi de mettre en évidence la position et la morphologie des noyaux dans leur cytoplasme et le stroma environnant [13]. La Figure 2.2 introduit l'un des enjeux de la pathologie digitale : toutes les méthodes de coloration reposent sur la perception d'une couleur, laquelle peut varier de manière importante selon les protocoles employés, le système d'imagerie utilisé, les paramètres d'acquisition ou encore le traitement appliqué aux images. Ici, on voit que les noyaux colorés à l'hématoxyline apparaissent très violets dans la sous-figure (a), et plus bleutés dans la sous-figure (b).



(a) Tissu de cancer du sein coloré par H&E.



(b) Tissu de cancer du sein coloré par IHC, avec une contre-coloration à l'hématoxyline.

FIGURE 2.2 Coupes successives de biopsie de cancer du sein, colorées avec deux techniques différentes.

Immunohistochimie (IHC)

L'immunohistochimie est couramment utilisée pour évaluer l'expression des biomarqueurs d'intérêt [4–6]. Cette technique fonctionne en trois étapes, comme illustré dans la Figure 2.3. Dans un premier temps, un anticorps primaire est incubé afin de cibler spécifiquement le marqueur d'intérêt. Ensuite, un anticorps secondaire est utilisé pour reconnaître la région Fc de l'anticorps primaire, cette région étant spécifique à l'espèce dans laquelle l'anticorps primaire a été produit. L'anticorps secondaire est préalablement couplé à une enzyme, telle que la peroxydase de raifort (horseradish peroxidase, HRP). Lors de l'ajout d'un substrat chromogénique comme la diaminobenzidine (DAB), l'enzyme catalyse sa conversion en un précipité brun, visible en microscopie optique [14, 15].

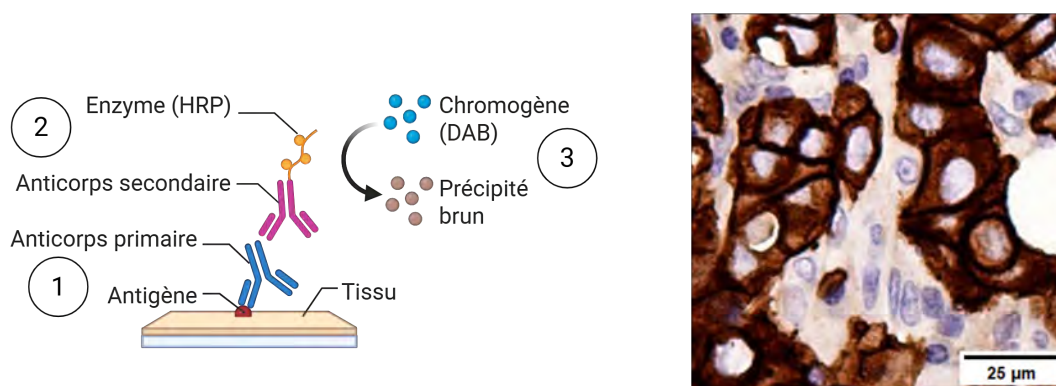


FIGURE 2.3 Schéma du fonctionnement de l'IHC et coloration visible en microscopie optique.

L'IHC est considérée comme la référence clinique pour l'évaluation de l'expression de différents marqueurs. Elle est compatible avec la coloration H&E et permet donc de localiser les protéines dans l'échantillon en connaissance de la morphologie tissulaire. L'IHC est peu coûteuse, et son analyse nécessite uniquement un microscope optique classique et l'expertise d'un pathologiste [16].

En fonction du marqueur étudié, l'évaluation doit être réalisée en suivant les recommandations de la Société Américaine d'Oncologie Clinique et du Collège Américain des Pathologistes (American Society of Clinical Oncology et College of American Pathologists, ASCO/CAP). Pour l'évaluation des récepteurs hormonaux ER et PR, on considère l'échantillon positif si 1 % ou plus des noyaux tumoraux sont colorés [17]. Pour HER2, les échantillons sont classés avec des scores allant de 0 à 3+ en fonction de la complétude de la coloration membranaire des cellules tumorales ainsi que du pourcentage de cellules positives [18].

Ces deux métriques révèlent la limite de l'IHC. En effet, c'est la subjectivité du pathologiste qui tranche pour déterminer si un noyau est positif ou si une membrane est entièrement colorée. On observe donc un taux de discordance inter-observateurs important, particulièrement pour les cas de faible expression. De plus, chaque laboratoire peut suivre son propre protocole d'IHC. On observe donc aussi un important taux de discordance inter-laboratoires causé par le manque de standardisation de la technique. Ces discordances peuvent aller de 1,6% à 76 % selon des études qui s'intéressent à l'évaluation des marqueurs HER2, ER, PR et Ki67 [19].

Bien que très largement utilisée, l'IHC reste une technique semi-quantitative qui ne permet pas toujours de conclure sur l'expression d'un marqueur et doit parfois être complétée par un test en hybridation in situ.

Hybridation in situ par fluorescence (FISH)

L'hybridation in situ par fluorescence (FISH) permet d'étudier l'amplification d'un gène. Cette technique fait appel à une sonde nucléique, dont la séquence est complémentaire à celle du gène ciblé, couplée à un fluorophore. Comme pour l'IHC, le tissu monté sur une lame subit les étapes de chauffage, déparaffinage, réhydratation et récupération antigénique. Une étape de digestion enzymatique est ajoutée pour rendre l'ADN accessible. Ensuite, la sonde se lie spécifiquement à la séquence cible. Après un lavage qui permet d'éliminer les sondes qui ne se sont pas liées à l'échantillon, il est possible de visualiser les copies du gène en microscopie à fluorescence. L'évaluation de l'amplification d'un gène est quantitative puisque l'on mesure le ratio entre le nombre d'un certain chromosome qui porte le gène d'intérêt, et le nombre de copies de ce gène [20].

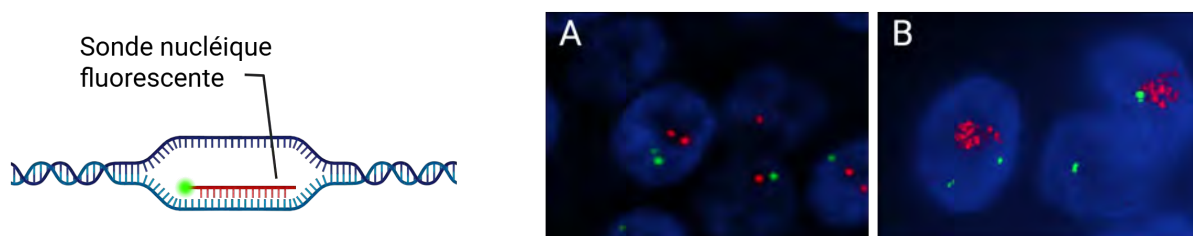


FIGURE 2.4 Représentation schématique d'une sonde utilisée pour le FISH et visualisation des sondes en microscopie à fluorescence. Les images A et B sont tirées de [20]. En vert, le signal de CEP17, en rouge le signal de HER2, en bleu, la contre-coloration au DAPI pour marquer les noyaux. A : le gène HER2 n'est pas amplifié. B : le gène HER2 est amplifié.

La synthèse de la protéine HER2 est contrôlée par le gène ERBB2, également appelé plus

simplement HER2. Dans une cellule normale, on retrouve deux copies de ce gène sur le chromosome 17. Pour déterminer si le gène est amplifié, on s'intéresse au rapport HER2/CEP17, en utilisant deux sondes fluorescentes à des longueurs d'onde différentes : l'une complémentaire au gène HER2 et l'autre ciblant le centromère du chromosome 17. La Figure 2.4 illustre les signaux obtenus pour CEP17 et HER2 dans le cas d'une amplification et d'une expression normale du gène.

Le FISH est une méthode quantitative, donc plus objective que l'IHC. Il permet d'analyser directement l'ADN de la cellule, qui est plus stable et moins susceptible de se dégrader que les protéines membranaires, sensibles aux étapes de prétraitement du tissu. Cette technique est particulièrement utile pour trancher les cas équivoques et réduire la discordance inter-observateurs fréquemment observée avec l'IHC. Pour éviter l'utilisation de la microscopie à fluorescence qui peut s'avérer complexe, il existe une alternative nommée hybridation in situ par chromogène (CISH). Cette technique exploite l'ajout d'un chromogène plutôt que d'un fluorophore sur la sonde nucléique, permettant la visualisation en champ clair.

Toutefois, les méthodes par hybridation restent plus complexes et coûteuses que l'IHC, ce qui explique qu'elles ne fassent que rarement partie des procédures de routine pour le diagnostic du cancer du sein.

Thérapies

Le traitement du cancer du sein repose sur une approche multimodale, combinant chirurgie, radiothérapie, chimiothérapie [12] et parfois hormonothérapie et thérapies ciblées, en fonction du type de la tumeur [4, 5, 12]. Cette stratification thérapeutique illustre l'importance du diagnostic moléculaire présenté précédemment.

L'hormonothérapie permet de traiter les cancers exprimant les récepteurs hormonaux (ER et/ou PR, luminal A et luminal B). Elle vise à bloquer la stimulation hormonale des cellules tumorales, soit par antagonisme des récepteurs (tamoxifène), soit par inhibition de la synthèse d'œstrogènes à l'aide d'inhibiteurs de l'aromatase (letrozole, anastrozole) [3, 21].

Les cancers surexprimant le récepteur HER2 peuvent être traités avec une thérapie ciblant spécifiquement ce récepteur et inhibant la prolifération cellulaire [4, 5].

Les cancers dits triple négatifs (HER2-, ER-, PR-) ne présentent pas de cible thérapeutique connue. De nouvelles approches telles que l'immunothérapie (anticorps anti-PD-1/PD-L1) ou la nanomédecine, sont actuellement explorées afin d'améliorer leur pronostic [12].

2.1.3 Statut HER2 : intérêt clinique et défis à relever

Parmi les récepteurs systématiquement évalués dans le cadre du diagnostic du cancer du sein, le récepteur HER2 occupe une place centrale. Cette protéine, membre de la famille des récepteurs du facteur de croissance épidermique (epidermal growth factor receptor, EGFR), régule la croissance et la prolifération cellulaires. HER2 est surexprimée dans 15 à 20 % des carcinomes invasifs. Les cancers HER2-positifs sont particulièrement agressifs et associés à un pronostic défavorable [22–24].

Cancers HER2-positifs

Les cancers HER2-positifs peuvent être traités par des thérapies ciblées, telles que le trastuzumab (Herceptin, approuvé par la *Food and Drug Administration*, FDA en 1998 [25]), ou le pertuzumab (Perjeta, approuvé par la FDA en 2012). Le trastuzumab se fixe sur le domaine extracellulaire de la protéine HER2 et inhibe les voies de signalisation intracellulaires, ce qui permet de bloquer le cycle cellulaire et de freiner l'évolution de la tumeur [24]. Cependant, l'efficacité de ces traitements repose sur une surexpression réelle de HER2. Il est donc crucial que le diagnostic permette de distinguer clairement les cancers HER2-positifs des cancers HER2-négatifs. Pour cela, les pathologistes suivent les recommandations de l'ASCO/CAP, qui préconisent d'utiliser l'IHC pour évaluer le statut HER2 avec un score de 0 à 3+. Les cas équivoques (score 2+) sont ensuite analysés par FISH afin de détecter l'amplification du gène HER2. Seuls les cancers 3+ ou les cancers 2+ avec amplification génétique confirmée sont considérés comme HER2-positifs et éligibles aux thérapies ciblées anti-HER2, comme le montre la Figure 2.5.

Émergence des cancers HER2-low et HER2-ultralow

Récemment, la compréhension du rôle de HER2 dans le cancer du sein s'est affinée avec l'identification de tumeurs exprimant HER2 à des niveaux plus faibles : les cancers dits HER2-low. Ces cancers, qui ne remplissent pas les critères classiques de HER2-positivité, peuvent toutefois répondre à certains traitements ciblés, notamment aux anticorps conjugués à un agent cytotoxique (*Antibody-Drug Conjugates*, ADC) comme le trastuzumab-deruxtecan. Ces ADC combinent la spécificité d'un anticorps dirigé contre HER2 avec un médicament cytotoxique, permettant de délivrer directement la molécule tueuse dans les cellules tumorales tout en épargnant les cellules saines. Le trastuzumab-deruxtecan a très récemment été approuvé par la FDA pour les cancers HER2-low et HER2-ultralow (27 janvier 2025, commercialisé sous le nom d'Enhertu par Daiichi Sankyo, Inc [26, 27]). L'émergence de ces nouvelles catégories

HER2-low et HER2-ultralow soulève de nouveaux défis pour la stratification du statut HER2.

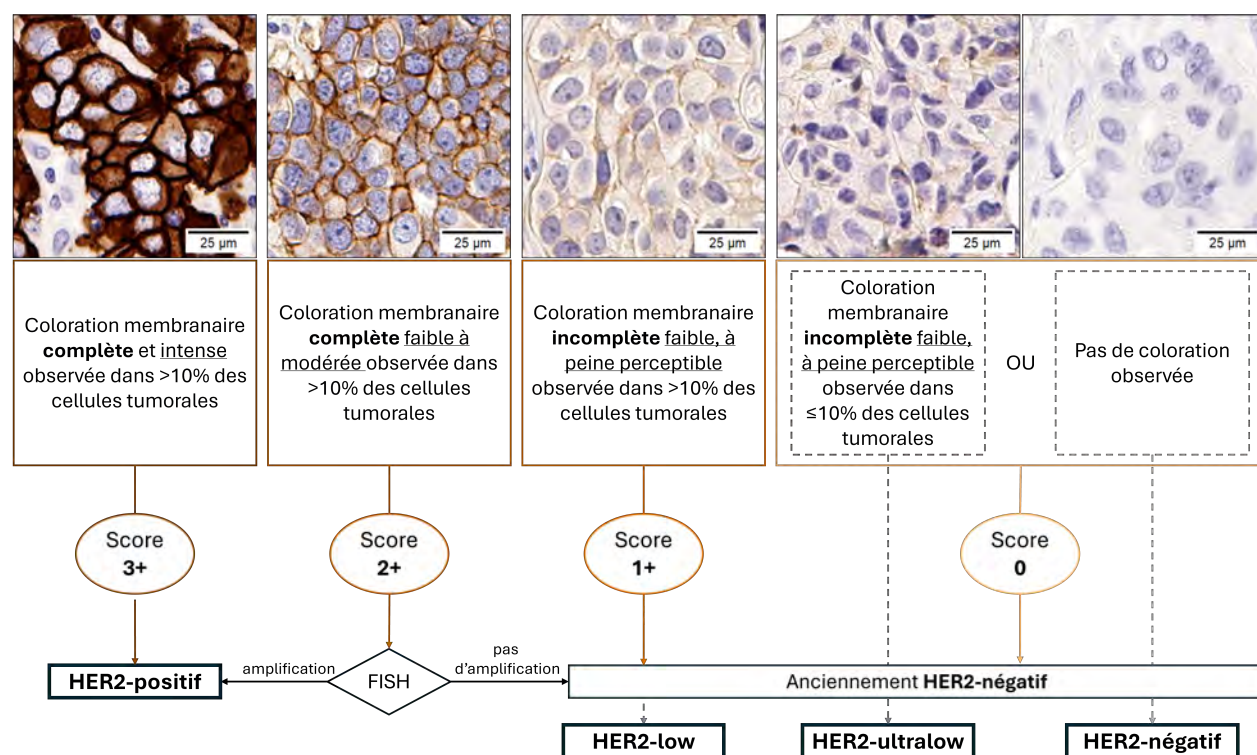


FIGURE 2.5 Consignes d'évaluation du statut HER2 de l'ASCO/CAP, avec ajout des catégories émergentes HER2-low et HER2-ultralow.

Nécessité de nouvelles techniques de diagnostic

De nombreuses études se sont penchées sur la reproductibilité inter-observateurs de l'évaluation du statut HER2, en particulier pour les cas de faible expression (HER2-low et HER2-ultralow). Plusieurs travaux ont mis en évidence une discordance significative entre pathologistes, soulignant la difficulté d'appliquer de manière uniforme les recommandations actuelles de l'ASCO/CAP. Certaines études suggèrent même qu'il serait nécessaire de revoir les critères d'évaluation, jugés parfois trop subjectifs [28]. En effet, l'IHC, bien qu'établie comme méthode de référence pour le statut HER2, demeure une approche intrinsèquement subjective, influencée par la variabilité pré-analytique, la qualité de la coloration et l'interprétation individuelle. L'émergence des catégories HER2-low et HER2-ultralow renforce ainsi la nécessité de développer de nouvelles approches diagnostiques plus objectives et quantitatives. Ces méthodes pourraient permettre une évaluation plus précise et standardisée de l'expression de HER2, optimisant ainsi l'identification des patientes susceptibles de bénéficier des nouvelles thérapies ciblées, notamment les ADC présentés précédemment.

2.2 Immunoplasmonique

Depuis des années, le laboratoire du Professeur Meunier investit l'utilisation de nanoparticules (NPs) plasmoniques dans des contextes biologiques. Elles peuvent notamment être utilisées dans le cadre d'une application diagnostique, pour marquer un échantillon et détecter une protéine d'intérêt, on parle alors d'immunoplasmonique (IP). Ce nom fait écho aux techniques classiques d'immunofluorescence (IF) ou d'immunohistochimie (IHC) qui exploitent des réactions immunologiques couplées respectivement à un fluorophore et un chromogène pour visualiser la présence d'une protéine. La technique IP repose quant à elle sur l'exploitation d'une réaction immunologique couplée aux propriétés de diffusion optique des NPs qui permettent de visualiser les NPs comme des points brillants en microscopie optique. Cette section permet donc dans un premier temps de décrire les propriétés optiques des NPs utilisées, puis de présenter les modifications de surface qu'elles peuvent subir afin de devenir des sondes ciblant une protéine d'intérêt et finalement, de discuter des stratégies de visualisation des NPs permettant leur quantification.

2.2.1 Propriétés optiques des nanoparticules plasmoniques

Lorsqu'une NP métallique (Au, Ag, Pt...) est soumise à une lumière incidente, on observe une oscillation collective des électrons de surface qui se déplacent sous l'effet de l'onde électromagnétique incidente. Cette oscillation collective des charges porte le nom de plasmon, et entraîne un phénomène appelé résonance localisée des plasmons de surface (*Localized Surface Plasmon Resonance*, LSPR). Cette résonance est à l'origine d'une forte absorption, d'une forte diffusion et d'une amplification du champ proche qui peuvent être exploitées dans diverses applications allant du domaine du biomédical à celui des énergies ou de la communication [29]. Une NP peut présenter des propriétés plasmoniques sans être métallique (par exemple dans le cas d'un nanocristal semiconducteur [30]), mais la suite de ce travail se concentrera spécifiquement sur les NPs métalliques plasmoniques (principalement Au, parfois alliage AuAg) utilisées pour réaliser le marquage IP.

La position du pic de résonance plasmonique varie en fonction de la morphologie (taille et forme), de la composition, et du milieu environnant, ce qui permet de concevoir des NPs sur mesure en fonction des besoins de l'application. Les NPs plasmoniques métalliques sont connues pour leur pic de résonance plasmonique dans le domaine du visible, voire du proche-infrarouge, ouvrant les portes à un grand nombre d'applications en microscopie optique.

Lorsque la NP plasmonique métallique est soumise à un champ électromagnétique, les électrons libres se déplacent sur la surface de la NP dans la direction opposée au champ incident.

La résonance localisée des plasmons de surface entraîne une section efficace d'extinction σ_{ext} englobant la section efficace d'absorption σ_{abs} et de diffusion σ_{scat} .

$$\sigma_{ext} = \sigma_{scat} + \sigma_{abs} \quad (2.1)$$

La section efficace d'extinction représente la perte totale de l'onde incidente qui ne sera pas transmise mais absorbée ou diffusée par la NP. La section efficace de diffusion correspond à la lumière diffusée dans toutes les directions par la NP, tandis que la section efficace d'absorption fait état de la résistance électrique de la NP responsable de la dissipation de l'énergie sous forme de chaleur.

Pour calculer la contribution de chaque phénomène, deux approches présentées dans la Figure 2.6 sont envisageables. Pour une longueur d'onde fixe, le champ appliqué peut être considéré invariant spatialement si le diamètre de la NP est bien plus petit que la longueur d'onde étudiée. Il est alors acceptable de considérer la NP comme un dipôle. Cependant, pour des NPs de diamètre plus important, le champ appliqué varie spatialement. L'approximation dipolaire n'est donc plus valide et c'est le modèle de Mie qui doit être utilisé [31].

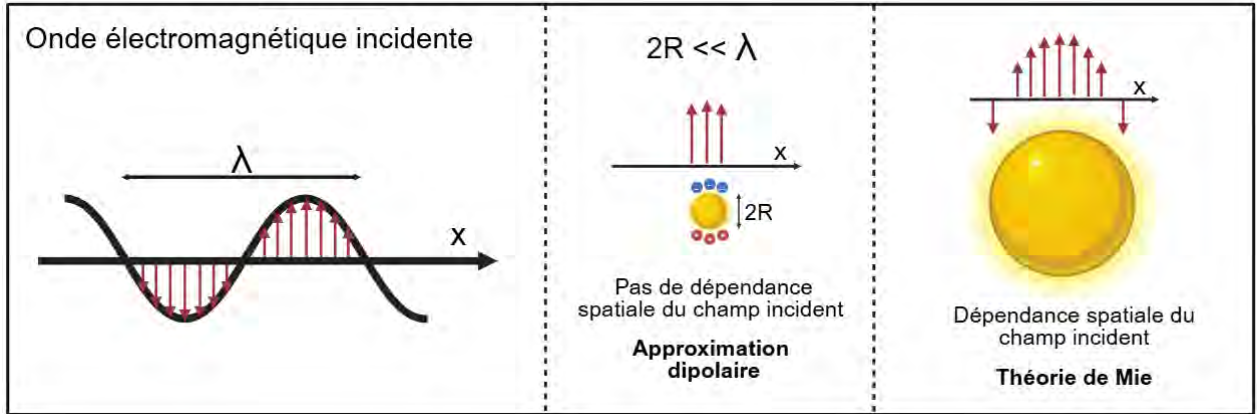


FIGURE 2.6 Effet du diamètre de la NP sur la dépendance spatiale du champ incident.

En sélectionnant soigneusement la taille et la composition des NPs utilisées, il est possible d'atteindre une section efficace de diffusion particulièrement importante. Ainsi, les NPs excitées par une onde électromagnétique diffuseront fortement la lumière à leur longueur d'onde de résonance plasmonique. En les observant en microscopie optique classique, sous certaines conditions d'illumination, discutées dans la section 2.2.3, il devient donc possible de visualiser des NPs sphériques plus petites que la limite de résolution optique en observant la lumière qu'elles diffusent.

2.2.2 Fonctionnalisation des nanoparticules plasmoniques

Le terme fonctionnalisation désigne les modifications de surface que peuvent subir les NPs dans le but de leur conférer de nouvelles propriétés exploitables dans une application précise. Dans notre cas, la fonctionnalisation des NPs permet d'introduire des groupes carboxyles (-COOH) à leur surface, qui serviront de point d'ancrage pour le greffage d'anticorps dirigés contre la protéine cible. La fonctionnalisation doit aussi permettre de prévenir l'aggrégation des NPs que l'on souhaite éviter au maximum afin de conserver les propriétés optiques des NPs. Cette section présente brièvement certaines modifications de surface que peuvent subir les NPs, avec un focus particulier sur la PEGylation et sur la fonctionnalisation par calix[n]arènes particulièrement utilisés dans le cadre de ce travail.

Méthodes de fonctionnalisation

Les anticorps peuvent être greffés à la surface des NPs grâce à des liaisons covalentes, non covalentes ou encore à l'aide de molécules adaptatrices [32].

Les méthodes impliquant des liaisons non covalentes consistent à fixer l'anticorps à la surface de la NP soit par adsorption physique via de faibles interactions (interactions électrostatiques, liaisons hydrogène, interactions hydrophobes et forces de van der Waals), soit par l'intermédiaire de liaisons ioniques entre les NPs dont la surface est chargée positivement et les anticorps porteurs de charges négatives. Ces approches présentent l'avantage d'être réversibles mais par conséquent, très peu stables et difficilement reproductibles.

Au contraire, les méthodes impliquant des liaisons covalentes garantissent une meilleure stabilité, mais nécessitent généralement une étape de pré-activation pour introduire à la surface de la NP un groupe fonctionnel qui sera exploité pour greffer l'anticorps. Parmi ces méthodes, on peut citer la chimie maleimide, la chimie click ou encore la chimie carbodiimide, utilisée par le groupe du Professeur Meunier. Cette dernière méthode exploite un carbodiimide, le 1-éthyl-3-(3-diméthylaminopropyl) carbodiimide, plus connu sous le nom de EDC pour réagir avec un groupe carboxyle et former un intermédiaire instable appelé O-acylisourée. En présence d'une amine (-NH₂), cet intermédiaire réagit pour former une liaison stable. Cependant, il s'hydrolyse très rapidement, ce qui diminue considérablement le rendement de la réaction souhaitée. Il est donc très recommandé de le stabiliser avec un N-hydroxysulfosuccinimide, plus couramment appelé sulfo-NHS, qui permet de former un ester toujours réactif aux amines mais stable dans l'eau.

L'utilisation de la chimie EDC/sulfo-NHS nécessite une étape de pré-activation permettant d'introduire des groupes carboxyles à la surface de la NP. Pour ce faire deux méthodes sont utilisées par le groupe du Pr Meunier : la PEGylation et la fonctionnalisation par calix[n]arènes.

PEGylation

La PEGylation désigne l'utilisation de polyéthylène glycol (PEG) pour recouvrir la surface d'une NP afin d'y introduire des groupes fonctionnels et d'améliorer sa stabilité. Le PEG est un polymère dont le motif $\text{O}-\text{CH}_2-\text{CH}_2$, appelé oxyéthylène, se répète linéairement. L'ajout de ces chaînes à la surface de la NP permet d'introduire une gêne stérique qui diminue le risque d'agrégation et limite l'adsorption non spécifique.

Les deux terminaisons du PEG peuvent être choisies parmi une large variété de groupes fonctionnels permettant de couvrir une grande diversité d'applications [33]. Dans notre cas, on choisit un PEG porteur d'un groupe carboxyle nécessaire au greffage des anticorps par la chimie EDC/sulfo-NHS. L'attache sur la NP se fait généralement grâce à un groupement thiol ($-\text{SH}$) qui permet de faire une liaison covalente Au-S . La Figure 2.7 représente le PEG couramment utilisé. Le soufre nécessaire à la liaison Au-S peut aussi être introduit par un acide lipoïque dont le pont disulfure a été préalablement réduit grâce à l'ajout de TCEP [34].

La diversité des PEG réside aussi dans l'éventail de masses moléculaires disponibles se traduisant par des chaînes plus ou moins longues. Plus la masse moléculaire est grande, plus le motif oxyéthylène se répète et donc plus le PEG est long. Un PEG ayant une masse moléculaire élevée (20 kDa par exemple), sera donc en mesure de s'étendre plus loin de la surface de la NP et donc de maximiser la gêne stérique. Cependant, sa capacité à se déplier, c'est à dire à adopter une certaine conformation dépend de sa densité sur la surface de la NP. En effet, à forte densité le PEG s'organise dans une conformation brosse (*brush conformation*) caractérisée par le déploiement des chaînes : la gêne stérique est maximisée et les groupes fonctionnels sont plus probablement accessibles et orientés vers l'extérieur de la NP. Au contraire, à faible densité, l'espace libre lui permet de se replier sur lui-même (*mushroom conformation*), limitant la gêne stérique et augmentant le risque que les groupes fonctionnels se retrouvent proches de la surface de la NP, orientés vers l'intérieur et donc inexploitable [35, 36]. Un long PEG aura donc tendance à se replier sur lui-même si sa densité n'est pas assez élevée. Pour assurer une fonctionnalisation efficace, il est donc essentiel de choisir des conditions qui favorisent l'adoption d'une conformation brosse.

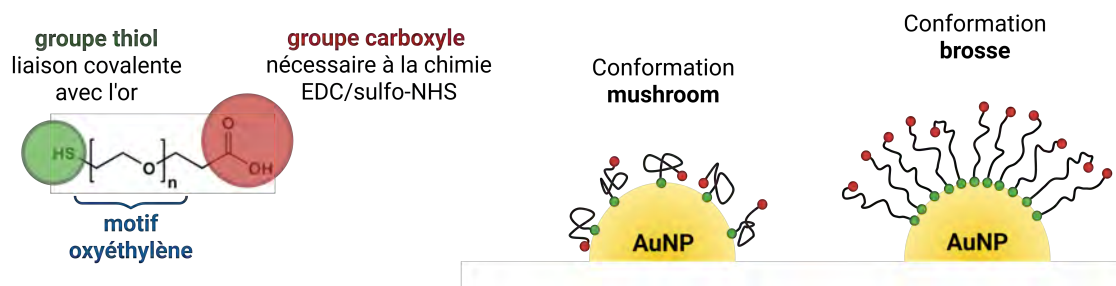


FIGURE 2.7 Illustration d'un SH-PEG-COOH dans les conformations champignon et brosse.

La PEGylation permet aussi d'améliorer la biocompatibilité des NPs d'or. Cette caractéristique n'est pas indispensable pour notre application qui se déroule *in vitro*, mais elle est particulièrement décrite dans la littérature, notamment pour les applications de livraison de médicaments pour lesquelles le PEG permet d'augmenter le temps de circulation des NPs dans le corps [33], [37], [36].

La PEGylation est un mécanisme très versatile, dont de nombreux paramètres peuvent être contrôlés pour s'adapter au mieux à l'application souhaitée. Il est largement décrit dans la littérature comme une référence de la fonctionnalisation des NPs d'or. Pourtant, la PEGylation présente une importante lacune de reproductibilité quand il s'agit de transposer le protocole à la fonctionnalisation de NPs différentes. Dans le Chapitre 3, on verra que deux PEG différents ont été utilisés pour fonctionnaliser des NPs d'or et d'alliage or/argent. Par ailleurs, Charline Courbon a montré dans son mémoire de maîtrise que la PEGylation ne représente pas une solution robuste pour notre utilisation, puisque la variabilité inter-lots pour un même type de NPs, achetées chez la même compagnie a entraîné l'arrêt de la bonne fonctionnalisation avec le protocole de routine de PEGylation [38]. Pour ces raisons, le groupe du Professeur Meunier a investigué une nouvelle alternative permettant d'introduire des groupes carboxyles à la surface des NPs.

Fonctionnalisation par calix[n]arènes

La fonctionnalisation des NPs peut être réalisée par l'intermédiaire de calix[n]arènes. Cette méthode a été adoptée dans le cadre d'une collaboration avec l'Université Libre de Bruxelles. Parmi leurs nombreuses applications, ces plateformes moléculaires macrocycliques peuvent être utilisées pour modifier la surface de NPs d'or [39–42]. Pour notre application, nous utilisons un calix[4]arène portant des groupements diazonium (N_2^+) et des groupements carboxyles [43]. Dans la suite de ce mémoire, on emploiera le terme calix[4]arène pour faire

référence à ce type de calix[4]arène tétra-diazonium portant des groupements carboxyles. Lors de l'ajout de borohydrure de sodium (NaBH_4), la réduction des groupements diazonium génère des radicaux aryles capables de former des liaisons covalentes avec l'or (Au-C). Ces liaisons Au-C ont une énergie supérieure de $14,2 \text{ kJ mol}^{-1}$ à celle des liaisons Au-S [39]. Chaque calix[4]arène est lié à la NP d'or au moyen de quatre liaisons covalentes, ce qui assure une fixation irréversible de la plateforme moléculaire à la surface de la NP [41]. De plus, le calix[4]arène présente une géométrie rigide et compacte qui assure une organisation spatiale contrôlée à la surface des NPs, contrairement aux PEG plus flexibles qui, à faible densité, peuvent se replier sur eux-même et rendre inaccessibles leurs groupes fonctionnels [35]. Enfin, les NPs fonctionnalisées avec des calix[4]arènes ont une excellente stabilité colloïdale puisque dans une solution de PBS à pH physiologique, ($\sim 7,4$) les groupements carboxyles se déprotonent en acide carboxylate (COO^-) permettant de limiter l'agrégation des NPs par répulsion électrostatique [41].

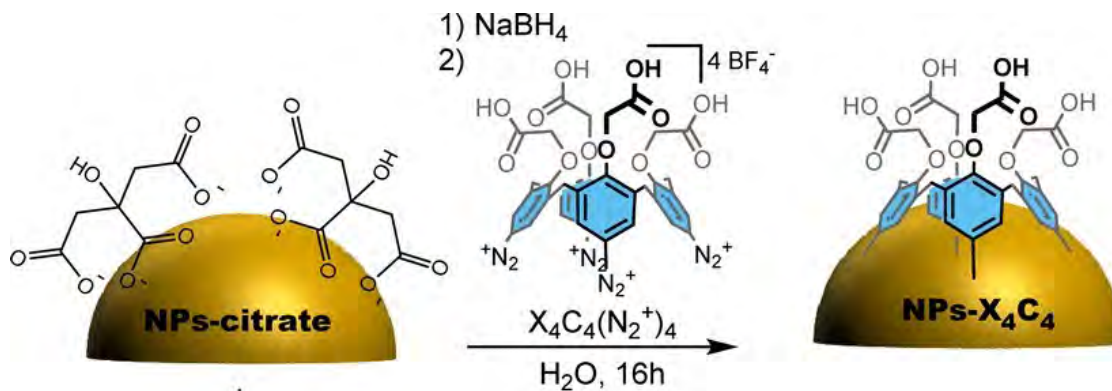


FIGURE 2.8 Structure du calixarène utilisé pour fonctionnaliser les NPs d'or. Tiré de [43].

2.2.3 Techniques d'imagerie et de quantification des nanoparticules plasmoniques

Le potentiel de la méthode IP repose sur les propriétés optiques de diffusion des NPs décrites précédemment. Sous l'incidence d'une onde électromagnétique, les NPs diffusent une lumière correspondant à leur pic plasmonique. Toutefois, l'intensité de la lumière diffusée reste faible en comparaison à celle de la source utilisée pour exciter la NP. Ainsi, la visualisation des NPs n'est possible qu'en utilisant une configuration en champ sombre, qui permet de ne collecter que la lumière diffusée par les NPs, et non celle provenant directement de la source.

Pour atteindre cet objectif, différentes configurations optiques peuvent être utilisées.

Condenseur champ sombre

La première consiste à employer un condenseur champ sombre, associé à un objectif dont l'ouverture numérique est inférieure à celle du condenseur. Dans ce cas, l'angle d'illumination de l'échantillon est trop grand pour que la lumière incidente soit collectée par l'objectif, ce qui permet de ne collecter finalement que la lumière diffusée par l'échantillon. Cependant, dans cette configuration, la lumière diffusée vers l'avant par les cellules est également collectée, ce qui réduit le contraste des NPs et complique leur visualisation [44]. Cette configuration n'a pas été utilisée dans le cadre de ce travail.

Illumination latérale

La seconde option repose sur l'utilisation d'un dispositif d'illumination latérale (*Side Illumination*, SI), développé dans le laboratoire du Professeur Meunier [45]. Ce dispositif, adaptable sur un microscope optique classique, est composé de DEL rouges, vertes et bleues dont l'intensité peut être ajustée individuellement. La lumière incidente se propage dans la lame de microscopie, qui sert de guide d'onde pour atteindre l'échantillon à imager. Comme la diffusion des NPs est isotrope, il est possible de collecter une partie de la lumière diffusée sans être gêné par la diffusion vers l'avant des cellules. Cette méthode de visualisation présente l'avantage d'être simple d'utilisation et compatible avec de nombreux objectifs. Elle permet également de passer aisément en illumination en champ clair (*Bright Field*, BF) par un simple clic, sans modifier la configuration optique. Ainsi, un même échantillon peut être imagé en SI pour visualiser les NPs, puis en BF pour observer le contexte biologique, simplement en illuminant l'échantillon avec une source en transmission.

Épi-illumination

Enfin, la troisième configuration repose sur une épi-illumination de l'échantillon [46]. Dans ce cas, la lumière collectée par l'objectif correspond uniquement à la lumière réfléchiée et diffusée vers l'arrière. La diffusion vers l'avant des cellules n'est donc pas collectée, ce qui permet de récupérer principalement la lumière diffusée par les NPs. Cependant, cette technique nécessite l'utilisation d'un objectif à immersion dans l'huile afin de limiter les réflexions à l'interface air/lame de microscopie. L'huile permet d'employer des objectifs à ouverture numérique élevée, mais rend la procédure d'imagerie plus complexe.

La Figure 2.9 permet de visualiser le chemin d'illumination de chaque modalité utilisée, ainsi que le type d'images acquises.

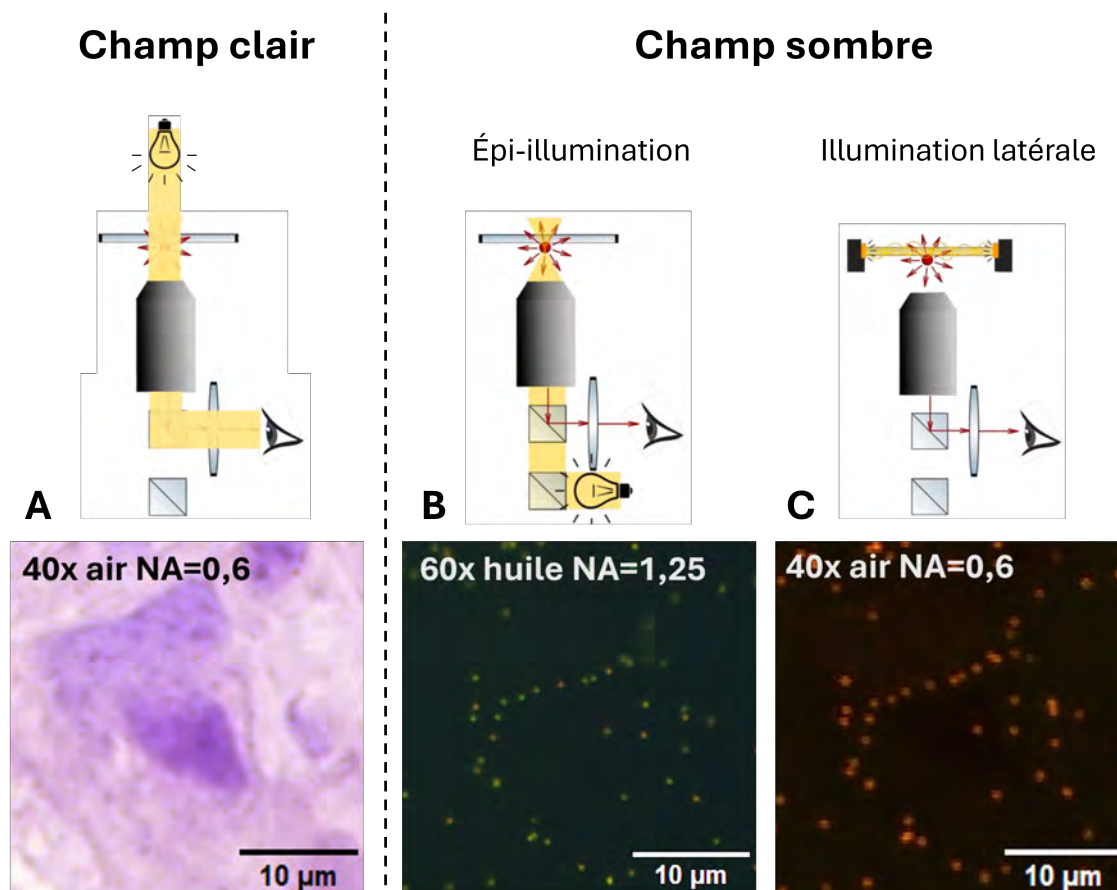


FIGURE 2.9 Modalités d'illumination permettant la visualisation des structures biologiques (champ clair A) et des NPs (épi-illumination B et illumination latérale C). Chaque image montre la même région d'un tissu coloré par H&E et marqué par IP. La différence de couleur entre les images B et C s'explique principalement par la différence dans les spectres des sources d'illumination utilisées par chaque technique.

Une fois les NPs visibles dans les images acquises, la solution la plus simple pour les quantifier consiste à utiliser l'algorithme Find Maxima du logiciel ImageJ. Cet algorithme permet de détecter des maxima locaux dans une image en appliquant un seuil de tolérance au bruit, appelé *prominence*. Le nombre de NPs détectées dans l'image dépend directement de ce seuil, que l'utilisateur doit ajuster manuellement. L'algorithme Find Maxima présente donc des limites d'adaptabilité lorsqu'il est appliqué à des images variées. Pourtant, le développement d'une stratégie d'analyse d'image automatisée nécessite une méthode de détection robuste qui permette de détecter les NPs présentes dans des images variées présentant une densité variable de NPs, et donc une intensité moyenne plus ou moins élevée. Pour garantir une quantification plus robuste des NPs, la solution proposée repose sur l'utilisation d'un modèle d'Intelligence Artificielle (IA) capable de reconnaître automatiquement les NPs dans les images SI.

2.3 Intelligence Artificielle en pathologie digitale

L'émergence des scanners permettant d'imager des lames entières (*Whole Slide Imaging*, WSI) dans les années 1990 a profondément affecté le domaine de la pathologie en rendant possible la numérisation systématique et à haute résolution des lames histologiques [47]. Cette transition technologique a ouvert la voie à l'intégration de l'IA dans les flux d'analyse histopathologique. Les volumes massifs d'images acquises peuvent désormais servir à l'entraînement de modèles permettant d'assister les pathologistes dans leurs tâches de diagnostic. Cette section permet dans un premier temps de présenter les concepts fondamentaux de l'IA, avant de s'intéresser plus en détails à certains modèles de Deep Learning (DL) permettant la détection des cellules et de leur noyaux dans les images histopathologiques.

2.3.1 Concepts fondamentaux de l'Intelligence Artificielle

Le terme Intelligence Artificielle (IA) décrit l'ensemble des méthodes et algorithmes permettant à un ordinateur d'exécuter des tâches nécessitant une forme d'intelligence, telles que la reconnaissance d'images, la prise de décision ou la compréhension du langage. Dans sa forme moderne, l'IA repose principalement sur l'apprentissage machine (*Machine Learning*, ML), où un modèle est entraîné à partir de données afin d'acquérir la capacité de généraliser et de prédire lorsqu'il est confronté à de nouvelles situations.

L'apprentissage peut être supervisé, lorsque le modèle dispose d'une vérité terrain (*Ground Truth*, GT) lui indiquant la sortie correcte pour chaque entrée, ou non supervisé, lorsqu'aucune GT n'est fournie et que le modèle doit apprendre par lui-même à identifier des structures ou des motifs dans les données. Des approches intermédiaires, dites semi-supervisées ou par renforcement, existent également.

Dans le cadre de l'apprentissage supervisé, sur lequel nous allons nous concentrer, l'entraînement d'un modèle se déroule itérativement. À chaque itération, le modèle reçoit une entrée, génère une sortie, puis compare cette sortie à la GT correspondante. Une fonction de perte (aussi appelée fonction d'erreur, *Loss function*) mesure l'écart entre la prédiction et le résultat attendu. Cet écart sert ensuite à ajuster les paramètres internes du modèle à l'aide d'un algorithme d'optimisation (*optimizer*) principalement basé sur le gradient de la fonction de perte. Ainsi, la valeur de la fonction de perte est réduite itérativement et le modèle apprend à produire des sorties de plus en plus proches de la vérité terrain : il améliore progressivement ses performances pour la tâche qui lui est confiée.

Apprentissage Machine

L'apprentissage machine (*Machine Learning*, ML) repose sur la recherche de relations ou de motifs dans les données afin de construire un modèle capable de généraliser à de nouveaux exemples.

Le processus d'apprentissage repose sur trois éléments principaux :

- un **jeu de données d'entraînement**, contenant des exemples représentatifs du problème à résoudre ;
- un **modèle**, défini par une architecture et un ensemble de paramètres ajustables (régression, arbres de décision, forêts aléatoires, méthodes de clustering) ;
- une **fonction de perte**, qui évalue la différence entre les prédictions du modèle et la vérité terrain.

Selon la nature du problème, le ML peut s'appliquer à des tâches de classification (prédire une catégorie), de régression (prédire une valeur continue), de clustering (regrouper des données similaires) ou encore de réduction de dimensionnalité (identifier les composantes principales d'un ensemble de données).

Apprentissage Profond

L'apprentissage profond (*Deep Learning*, DL) constitue une branche spécifique du ML, fondée sur l'utilisation de réseaux de neurones artificiels comportant plusieurs couches dites profondes. Dans le cas de l'analyse d'image, ces couches successives permettent au modèle d'extraire des représentations hiérarchiques des données : les premières couches apprennent des caractéristiques simples (bords, motifs élémentaires), tandis que les couches plus profondes apprennent des concepts plus abstraits (formes, objets, structures complexes). Le Chapitre 4 fait appel à différents modèles de DL pour réaliser des tâches de segmentation permettant d'extraire les informations pertinentes à l'analyse des biopsies marquées par IP (NPs, tissus, noyaux des cellules).

Les étapes de calcul intermédiaires d'un réseau sont généralement appelés neurones. La valeur prise par ces neurones est le résultat d'une combinaison linéaire de ses entrées pondérées par les poids (*weights*) du réseau, à laquelle est appliquée une **fonction d'activation** non linéaire. Une fois que l'entrée a parcouru toutes les couches du réseau, il est possible de calculer le gradient utilisé pour ajuster itérativement les poids du réseau grâce à une étape de rétropropagation (*backpropagation*).

Les architectures utilisées varient selon le type de données, voici ici quelques exemples :

- les **réseaux de neurones convolutionnels** (*Convolutional Neural Networks*, CNN),

- utilisés pour le traitement d’images et de données spatiales ;
- les **réseaux récurrents** (*Recurrent Neural Networks*, RNN) et leurs variantes, comme les *Long Short-Term Memory* (LSTM), adaptés aux données séquentielles ou temporelles ;
- les **transformers**, plus récents, se basant sur des mécanismes d’attention et s’imposant dans le traitement du langage naturel ainsi qu’en vision par ordinateur.

Le DL a permis des avancées majeures dans de nombreux domaines, notamment la reconnaissance d’images, la segmentation de tissus en histologie, la traduction automatique et la génération d’images. Sa puissance repose sur la combinaison de grandes quantités de données, de capacités de calcul élevées et d’architectures de modèles complexes généralement issus d’assemblages des architectures de base.

2.3.2 Modèles de Deep Learning

U-Net

Le modèle U-Net, introduit en 2015 par Ronneberger et al. [48], constitue une architecture de référence en segmentation d’images biomédicales. Sa conception repose sur une structure en forme de « U » composée de deux parties principales : un encodeur et un décodeur.

L’encodeur, ou chemin contractant, est constitué d’une succession de couches de convolutions et de sous-échantillonnages (*max pooling*) qui extraient des caractéristiques locales tout en réduisant la résolution spatiale de l’image. Le décodeur, ou chemin expansif, reconstruit progressivement la carte de segmentation en combinant les informations contextuelles extraites par l’encodeur avec les détails spatiaux issus des couches correspondantes via des connexions de saut (*skip connections*). Ces connexions directes entre les niveaux symétriques de l’encodeur et du décodeur permettent au réseau de mieux localiser les frontières des objets, un atout essentiel pour la segmentation précise des structures biologiques (noyaux, cellules, tissus, etc.). Le U-Net a ainsi démontré d’excellentes performances sur des tâches nécessitant la détection d’objets de petite taille ou la séparation d’instances proches, et constitue aujourd’hui la base de nombreuses architectures dérivées, comme HoVer-Net [49], StarDist [50, 51], CellPose [52, 53] ou CellViT [49, 54] qui seront utilisées dans les Chapitres 4 et 5.

Des variantes ont ensuite été proposées pour améliorer la robustesse ou étendre ses capacités :

- **U-Net++**, introduisant des connexions denses entre les niveaux du décodeur pour une meilleure fusion des caractéristiques [55] ;
- **Attention U-Net**, intégrant des mécanismes d’attention permettant au modèle de se concentrer sur les régions d’intérêt pertinentes [56] ;

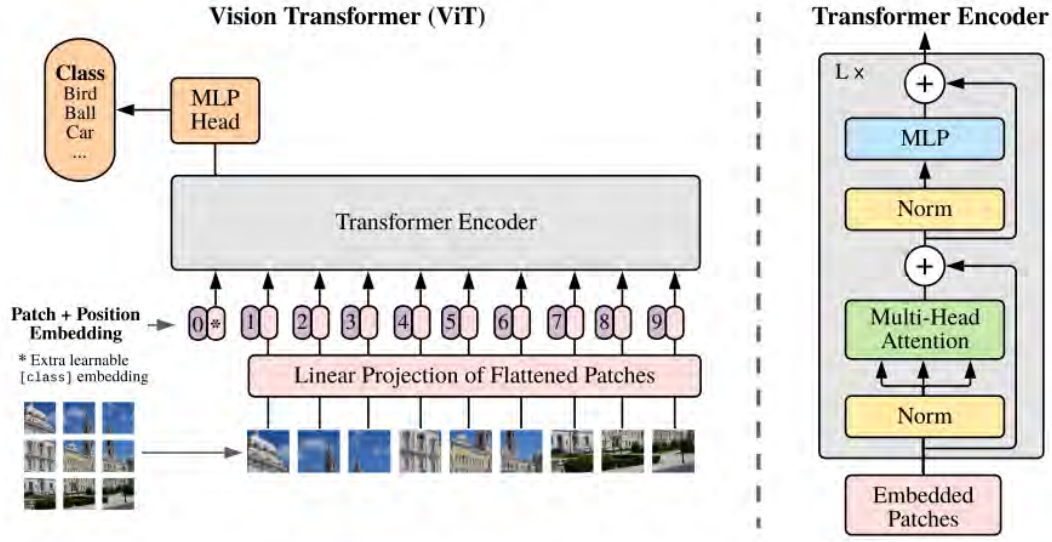


FIGURE 2.11 Architecture du ViT. Tiré de [59].

Comparé au CNN, le ViT ne repose pas sur des convolutions locales mais sur une modélisation explicite des dépendances globales. Cette capacité à capturer simultanément les contextes locaux et globaux a conduit à d'importants progrès en vision médicale, notamment pour la classification et la segmentation d'images histologiques, comme en témoignent les modèles récents tels que Cellpose-SAM [53] et CellViT [54].

2.3.3 Outils de segmentation et de classification cellulaire

Avec l'émergence de la pathologie digitale, de nouveaux besoins en segmentation d'images histologiques sont apparus. En particulier, la détection des cellules à partir de la coloration des noyaux par la contre-coloration H&E est devenue une tâche centrale. L'objectif est d'identifier chaque cellule (ou noyau) comme un objet individuel, on parle alors de segmentation par instance. Cette tâche peut se révéler complexe lorsqu'elle implique des cellules denses qui se touchent. De plus, les cellules peuvent avoir des morphologies hétérogènes, notamment dans le cadre de tissus cancéreux composés de cellules tumorales plus ou moins différenciées. Les principaux défis de la segmentation cellulaire sont donc d'être capable de détecter et de séparer des objets morphologiquement hétérogènes. De nombreuses approches ont été développées pour accomplir cette tâche. La suite de cette revue de littérature se concentre sur trois outils qui seront utilisés aux Chapitres 4 et 5 de ce mémoire, à savoir StarDist [50,51], CellPose [52,53] et CellViT [49,54].

StarDist 2018 et sa version histopathologique de 2022

StarDist a été initialement développé pour la segmentation des noyaux dans les images de microscopie à fluorescence [50], puis adapté en 2022 aux images d’histopathologie colorées à l’H&E [51]. Plutôt que d’utiliser des boîtes englobantes pour délimiter chaque instance, StarDist représente chaque objet par un polygone star-convexe. Concrètement, un CNN prend l’image en entrée et, pour chaque pixel, calcule deux informations : la probabilité que le pixel appartienne à un objet (un noyau), et la distance radiale entre ce pixel et le contour de l’objet correspondant. Tous les pixels dont la probabilité d’appartenance dépasse un seuil sont associés à un polygone représentant l’objet, puis une procédure de suppression des maximums non désirés (*non-maximum suppression*) est appliquée afin de fusionner les polygones redondants.

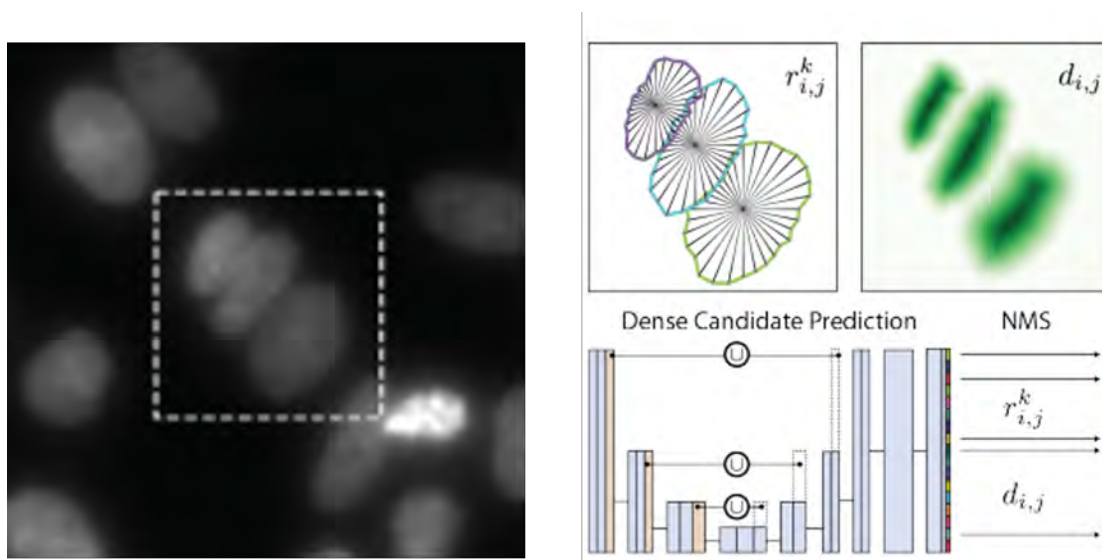


FIGURE 2.12 Méthode utilisée par StarDist pour segmenter les noyaux grâce au calcul de la probabilité d’appartenance et des distances radiales. Tiré de [50].

La version adaptée aux images histopathologiques permet également la classification cellulaire, grâce à l’ajout d’une branche CNN supplémentaire dédiée à la prédiction du type de cellule. Le modèle a été entraîné sur le jeu de données Lizard, composé d’images H&E de tissu colique issues du challenge CoNiC. Ce jeu de données contient 4 981 images RGB de 256×256 pixels, dont les noyaux ont été annotés et répartis en six classes : neutrophiles, cellules épithéliales, lymphocytes, plasmocytes, éosinophiles et cellules du tissu conjonctif. La distribution y est fortement déséquilibrée, avec environ 60 % de cellules épithéliales. L’architecture utilisée repose sur un U-Net de profondeur 4. Les fonctions de perte combinent l’entropie croisée binaire pour la probabilité d’appartenance à un objet, l’erreur absolue moyenne pour les dis-

tances radiales, et, pour la classification, la somme de l'entropie croisée catégorielle et de la perte de Tversky. L'optimiseur employé est Adam, avec un entraînement sur 1000 époques (*epochs*) et une taille de lot (*batch size*) de 4.

Le modèle est mis à disposition par des plugins existant dans les logiciels ImageJ et QuPath [60], bien que la version incluant la classification cellulaire nécessite un réentraînement préalable.

Cellpose 2021 et son évolution en Cellpose-SAM en 2025

CellPose permet de segmenter des cellules grâce à un réseau de neurones capable de prédire des champs de flux vectoriels horizontaux et verticaux, ainsi qu'un masque binaire indiquant si chaque pixel appartient à une cellule ou non [52]. Cette approche permet d'identifier le centre de chaque cellule et de regrouper tous les pixels convergeant vers le même centre pour former une instance. Initialement, CellPose prend en entrée une image composée de deux canaux : le premier canal comporte le marquage cytoplasmique, et le deuxième canal, optionnel, contient le marquage nucléaire, souvent DAPI.

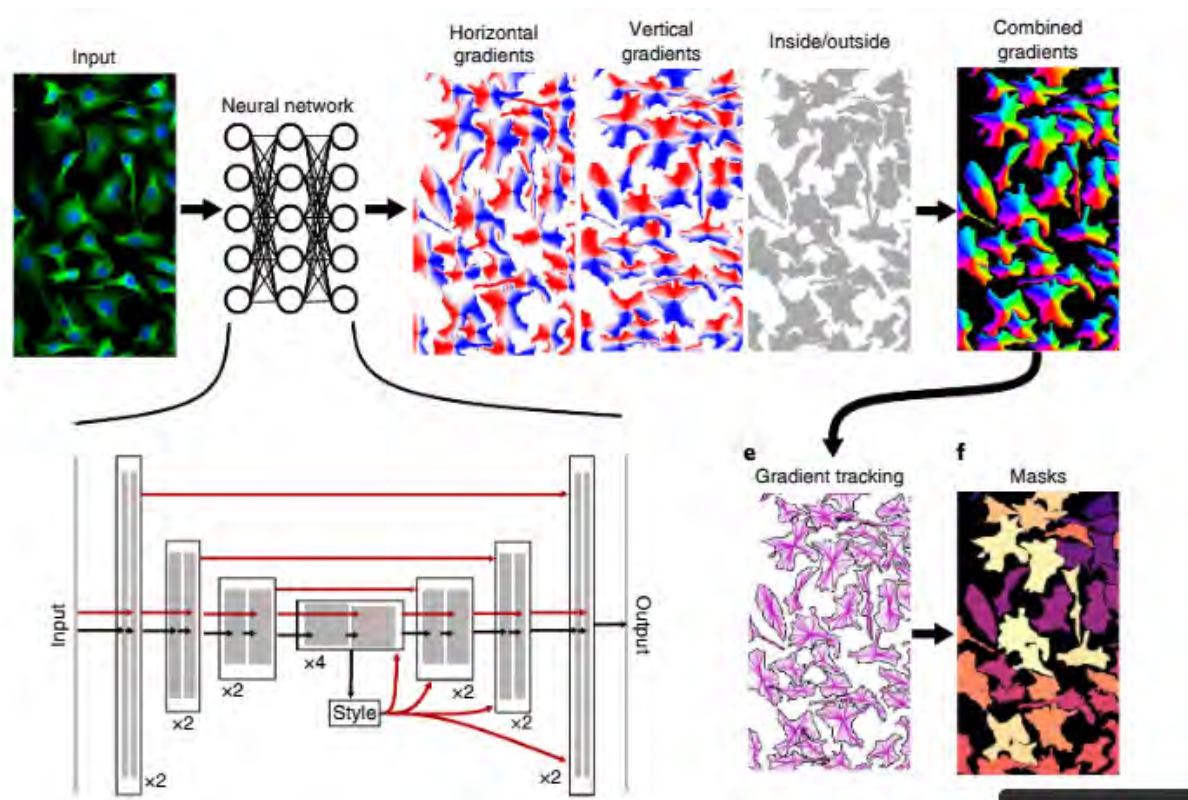


FIGURE 2.13 Méthode utilisée par CellPose pour segmenter les cellules grâce à la prédiction de champs vectoriels horizontaux et verticaux. Tiré de [52]

L’architecture utilisée est basée sur un U-Net. Le modèle a été entraîné sur plusieurs jeux de données. Le premier est déjà segmenté et comporte 1136 images de différentes sources (BBBC038v1, BBBC039v1, MoNuSeg, ISBI 200). Un second jeu de données plus généralisé est composé de 608 images qui peuvent aussi bien être des cellules imagées en microscopie à fluorescence que des images plus générales comportant des motifs répétitifs comme des fruits ou des coquillages. L’optimiseur utilisé est la descente de gradient stochastique, avec un entraînement sur 500 époques et une taille de lot de 8.

Une version récente, SAM-CellPose [53], combine l’encodeur du Segment Anything Model (SAM) [61] et le décodeur de CellPose. L’encodeur Vision Transformer de SAM, entraîné sur des milliards d’images, fournit une représentation très robuste des structures de l’image, tandis que le décodeur CellPose applique la logique des flux vectoriels pour la segmentation cellulaire. Cela améliore la généralisation sur différents types d’images et de conditions de microscopie. Ce modèle est utilisable dans Dragonfly qui sera présenté dans le Chapitre 4. Bien qu’adaptable aux images RGB H&E, CellPose segmente parfois les noyaux plutôt que les cellules à cause du fort marquage nucléaire de la contre-coloration H&E.

CellViT 2023 et son évolution en CellViT++ en 2025

CellViT est un modèle de segmentation et de classification cellulaire basé sur une architecture de type Vision Transformer (ViT) [49]. Le modèle permet à la fois la segmentation des noyaux et leur classification en cinq types cellulaires : cellules épithéliales néoplasiques, cellules épithéliales non néoplasiques, cellules inflammatoires, cellules conjonctives et cellules mortes. L’architecture du modèle repose sur le principe de U-Net encodeur-décodeur, similaire à HoVer-Net [62], mais avec un encodeur remplacé par un Vision Transformer. Ce mécanisme permet de capturer simultanément les informations locales et globales de l’image.

Le décodeur est composé de trois branches :

- Une branche prédisant un masque binaire noyau / pas noyau.
- Une branche calculant des cartes de distances horizontales et verticales, facilitant la séparation des noyaux contigus.
- Une branche dédiée à la classification cellulaire.

Une étape de post-traitement, similaire à celle de HoVer-Net, regroupe ces informations pour produire une segmentation finale des noyaux avec leur type cellulaire associé.

Pour l’encodeur, différents ViT pré-entraînés peuvent être utilisés, tels que ViT256, pré-entraîné sur 104 millions d’images histologiques, ou l’encodeur du Segment Anything Model (SAM).

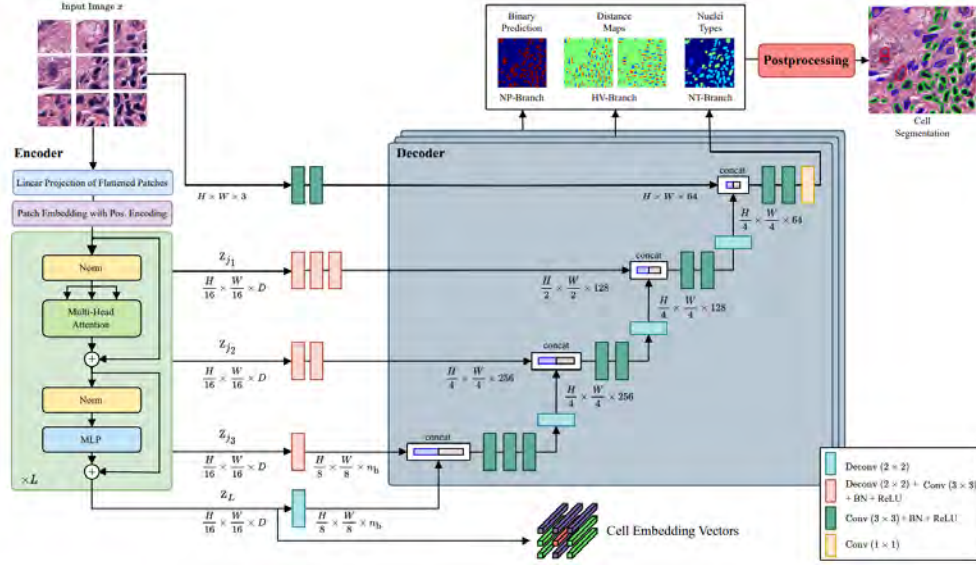


FIGURE 2.14 Méthode utilisée par CellViT pour segmenter et classer les noyaux. Tiré de [49]

CellViT a été entraîné sur plusieurs jeux de données annotés en pathologie digitale, notamment PanNuke (189 744 cellules provenant de 19 types de tissus), CoNSeP et MoNuSeg.

La version mise à jour, CellViT++ [54], améliore la généralisation du modèle et son adaptabilité à de nouvelles données et contextes d'utilisation, tout en conservant les avantages de l'approche Transformer pour la segmentation et la classification simultanées.

CHAPITRE 3 MATÉRIEL ET MÉTHODES POUR LA PRÉPARATION DES ÉCHANTILLONS ET L'ACQUISITION DES IMAGES À ANALYSER

Le développement d'une stratégie d'analyse d'images serait impensable sans disposer d'échantillons à étudier. Ce chapitre est donc consacré à la préparation et à la description des différents échantillons qui ont été analysés, ainsi qu'aux conditions choisies pour l'acquisition des images. Bien que l'objectif final soit l'analyse d'images, il est indispensable de comprendre en profondeur la nature des échantillons étudiés. Cette connaissance détermine les informations pertinentes à extraire des images et oriente les choix méthodologiques pour leur analyse.

La majorité des échantillons ont été préparés et fournis par Jennyfer Zapata Farfan, PhD, qui sera créditée régulièrement.

3.1 Préparation des échantillons marqués par immunoplasmonique

3.1.1 Nanoparticules utilisées

Le marquage IP nécessite dans un premier temps des NPs fonctionnalisées.

Dans le cadre de ce travail, les NPs utilisées ont toutes été fonctionnalisées avec l'anticorps EPR19547-12 (Abcam, réf. ab222482, Anti-ErbB2 / HER2 antibody [EPR19547-12] - BSA and Azide free) qui sera simplement désigné par EPR dans la suite de ce mémoire. Il s'agit d'un anticorps monoclonal recombinant de lapin qui cible la protéine HER2. À l'avenir, il serait pertinent de fonctionnaliser les NPs avec un clone utilisé en clinique, comme par exemple le clone PATHWAY 4B5 de Roche Diagnostics. Cependant, certains composés aminés présents dans les tampons de conservation des anticorps cliniques entrent en compétition avec les anticorps eux-mêmes lors de la fonctionnalisation par EDC/sulfo-NHS, ce qui compromet le succès de la fonctionnalisation des NPs. Un travail d'adaptation des protocoles est en cours pour y parvenir.

Pour chaque lot de NPs fonctionnalisées, le succès de la fonctionnalisation et la qualité des NPs obtenues sont contrôlés grâce à plusieurs tests :

- **Mesure du spectre d'absorbance par UV-visible** : réalisée à partir d'un volume de 2 μ L de NPs déposé sur une plaque Take3 compatible avec le spectrophotomètre Epoch de BioTek. Le spectre d'absorbance est obtenu en illuminant l'échantillon avec un spectre continu (lampe flash au xénon). Il est le résultat de la contribution du très grand nombre de NPs contenues dans le volume utilisé. La position du pic plasmone

nique permet de donner une estimation de la couleur qui sera diffusée par la NP. Un élargissement de ce pic au cours du protocole de fonctionnalisation témoignerait de l'hétérogénéité des NPs contenues dans la solution et par conséquent, de leur possible agrégation. Les valeurs d'absorbance au pic et sur la moyenne du plateau sont utilisées pour calculer la concentration des NPs.

- **Mesure de l'indice de polydispersité (PDI) par Diffusion Dynamique de la Lumière (DLS)** : réalisée grâce à un zétamètre (Zetasizer - Malvern Panalytical). Plus l'indice de polydispersité est faible, plus la solution est monodisperse. Au-delà de 0,08, on considère que la solution présente une polydispersité non acceptable [44].
- **Test à flux latéral (*Lateral Flow Assay*, LFA)** : permet de confirmer la présence d'anticorps à la surface de la NP. Ce test utilise une bandelette faite d'une membrane de nitrocellulose sur laquelle a été déposée une ligne de protéine A/G. Lorsque la bandelette est déposée dans un puits contenant 2 μ L de NPs approximativement concentrées à 1×10^{10} NPs/mL, diluées dans 40 μ L de PBST avec 1 % de BSA, les NPs migrent le long de la membrane. L'affinité entre la portion Fc [63] des anticorps et la protéine A/G [64] interrompt la migration des NPs qui s'accumulent sur la ligne de dépôt jusqu'à l'apparition d'une ligne colorée. Cette ligne témoigne de la présence d'anticorps à la surface des NPs, et donc d'une fonctionnalisation réussie. Ce test n'apporte toutefois aucune information quant à l'activité fonctionnelle des anticorps (l'utilisation d'un test ELISA pourrait remédier à cette lacune). Pour fonctionner, la portion Fc d'une partie des anticorps doit être accessible pour interagir avec la protéine A/G. Comme la chimie EDC/sulfo-NHS utilisée n'est pas orientée, les LFA peuvent être utilisés.

Au90NPs-NanoBrand-X4C4-EPR

Les NPs les plus utilisées dans ce travail sont des NPs d'or sphériques, de 90 nm de diamètre. Elles sont fournies par NanoBrand (réf.NB-GSP-90-CIT-1 90 nm Gold Nano-Spheres, citrate-coated OD=1). Les NPs sont fonctionnalisées par calix[4]arène, suivi du greffage des anticorps EPR par chimie EDC/sulfo-NHS. Le spectre UV-visible montre un pic plasmonique bien défini dont le maximum est atteint à 556 nm. En présence de Permout (indice de réfraction de 1,52) qui sert de milieu de montage, le pic plasmonique subit un léger décalage vers la droite (*red shift*), conférant à la NP une teinte orangée visible en microscopie optique, sous illumination latérale. Ces NPs ont été particulièrement étudiées dans notre groupe, notamment dans le cadre d'une étude de vieillissement menée par Charline Courbon, M. Sc. A [38]. Elle souligne notamment leur impressionnante monodispersité (indice de polydispersité après fonctionnalisation mesuré à 0,02 [38]).

Au90NPs-NanoBrand-X4C4-EPR

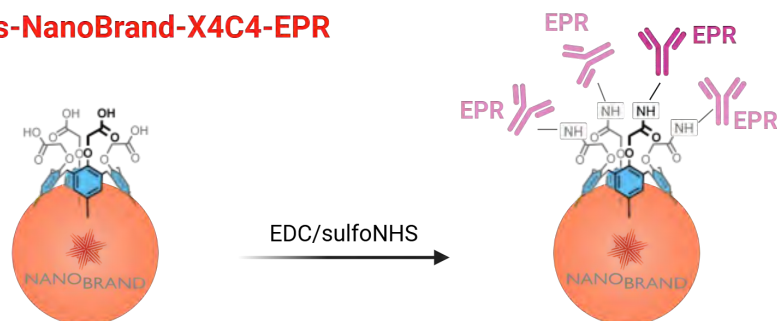


FIGURE 3.1 Représentation schématique d'une Au90NP-NanoBrand-X4C4-EPR. Par souci de simplicité, un seul calix[4]arène est représenté à la surface de la NP. L'orientation des anticorps est aléatoire. Le schéma n'est pas à l'échelle.

Les Au90NPs-X4C4-EPR sont incubées sur les TMA, sur les billes de contrôle BCS ajoutées aux lames de TMA, et sur les CMA. La fonctionnalisation et l'incubation sont réalisés par Jennyfer Zapata Farfan, PhD.

Au90NPs-Nanopartz-LAPEG-EPR

D'autres NPs d'or sphériques, de même diamètre 90 nm, mais cette fois fournies par Nanopartz (réf. A11-90-CIT-DIH-1-50) ont aussi été utilisées. Leur utilisation est antérieure à l'adoption de la fonctionnalisation par calix[4]arène. Elles sont donc PEGylées grâce à un LA-PEG_{1 kDa}-COOH (réf. PG2-CALA-1k, Nanocs). Leur pic plasmonique se situe à 561 nm, Elles apparaissent donc sensiblement de la même couleur que les NPs précédentes lorsqu'on les observe en microscopie optique, sous la même illumination. L'incubation de ces NPs fonctionnalisées a été réalisée sur les billes de contrôle BCS et a permis d'observer un attachement spécifique qui sera présenté dans le Chapitre 6.

Au90NPs-Nanopartz-LAPEG-EPR

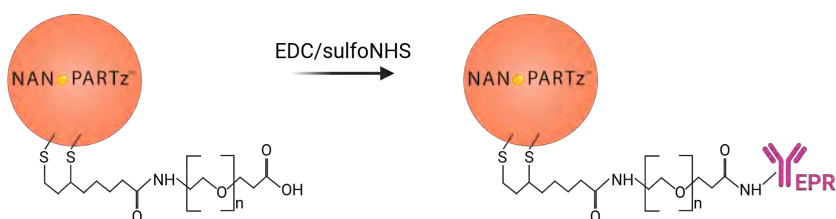


FIGURE 3.2 Représentation schématique d'une Au90NP-Nanopartz-LAPEG-EPR. Par souci de simplicité, une seule chaîne de PEG a été représentée à la surface de la NP. L'orientation de l'anticorps est aléatoire. Le schéma n'est pas à l'échelle.

Toutefois, ces NPs disposent d'une tendance à l'agrégation plus importante que les NPs présentées précédemment. En effet, la Figure 3.3, montre que le pic plasmonique des Au90NPs-Nanopartz-LAPEG-EPR est plus large que celui des Au90NPs-NanoBrand-X4C4-EPR. De plus, l'indice de polydispersité, mesuré en DLS à 0,07 pour les Au90NPs-Nanopartz-LAPEG-EPR confirme la présence de NPs de tailles variées, suggérant la présence d'agrégats. Ce phénomène peut s'expliquer par deux facteurs : d'une part, la qualité des NPs nues (PDI de 0,01 pour NanoBrand, contre 0,05 pour Nanopartz), et d'autre part, l'efficacité de la méthode de fonctionnalisation pour limiter l'agrégation (le PEG court instaure une faible gêne stérique qui peut diminuer encore plus si la densité de PEG n'est pas bien maîtrisée, tandis que le calix[4]arène assure un recouvrement de surface plus homogène, reconnu pour limiter l'agrégation).

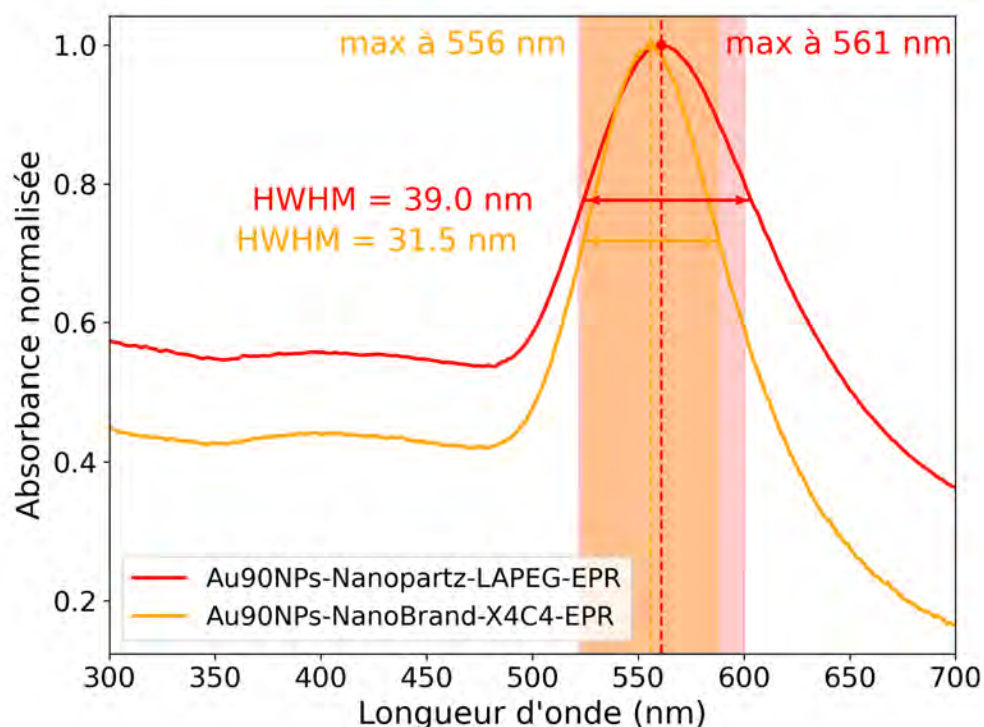


FIGURE 3.3 Absorbances normalisées des Au90NPs-Nanopartz-LAPEG-EPR et des Au90NPs-NanoBrand-X4C4-EPR mesurées par spectroscopie UV-visible. La demi-largeur à mi-hauteur (*Half Width at Half Maximum*, HWHM) a été calculée en utilisant le plateau situé entre 350 et 450 nm comme base pour le calcul de la prominence.

La faible qualité des Au90NPs-Nanopartz-LAPEG-EPR a conduit à leur abandon dès que la fonctionnalisation par calix[4]arène a permis d'utiliser avec succès les NPs de NanoBrand, préférées pour leur monodispersité.

AuAg-HSPEG

Enfin, les alliages or/argent ont fait l'objet d'une courte investigation décrite dans le Chapitre 6. Deux types d'alliages ont été utilisés. Ils ont été synthétisés au laboratoire du Professeur Meunier en utilisant les méthodes de Turkevich et de Seeded Growth [43,65]. La fonctionnalisation a été assurée par PEGylation avec un HS-PEG_{1kDa}-COOH (réf. PG2-CATH-1k, Nanocs). L'alliage 10:90 Au:Ag de 50 nm de diamètre présente un pic plasmonique à 433 nm, il apparaît donc bleu lorsqu'on l'observe au microscope optique, sous illumination latérale. Le second alliage 50:50 Au:Ag de 73 nm de diamètre diffuse quant à lui dans le vert en raison de son pic plasmonique à 491 nm. Les indices de polydispersité mesurés sont supérieurs à 0,08, traduisant une polydispersité non négligeable qui aura pour conséquence d'introduire une variabilité de la couleur perçue en imagerie.

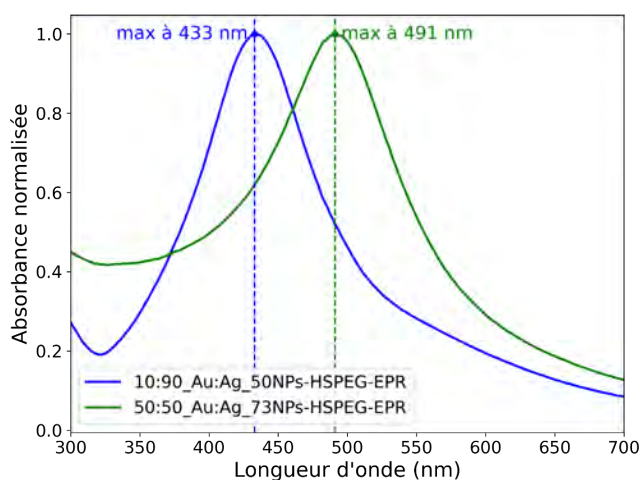
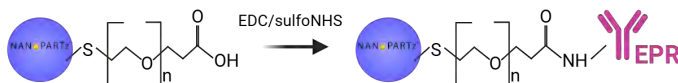


FIGURE 3.4 Absorbances normalisées des alliages mesurées par spectroscopie UV-visible.

10:90 Au:Ag 50NPs-HSPEG-EPR



50:50 Au:Ag 73NPs-HSPEG-EPR

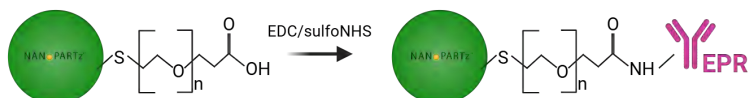


FIGURE 3.5 Représentation schématique des alliages. Par souci de simplicité, une seule chaîne de PEG a été représentée à la surface de chaque NP. L'orientation de l'anticorps est aléatoire. Le schéma n'est pas à l'échelle.

3.1.2 Échantillons utilisés

Une fois les NPs fonctionnalisées, elles sont incubées sur divers échantillons biologiques afin de quantifier leur expression protéique. Cette section décrit les trois types d'échantillons utilisés.

Matrices de tissus (TMA)

L'essentiel du travail a été réalisé sur des matrices de tissus (*tissue microarrays*, TMA). Les TMA sont des assemblages de carottes de tissus, conservées sous la forme d'un bloc de paraffine. Chaque carotte de tissu provient d'un prélèvement tissulaire effectué sur un patiente lors du diagnostic d'un cancer (en favorisant les biopsies à l'aiguille, moins invasives que les prélèvements par chirurgie). Le bloc de TMA est ensuite découpé au microtome en sections d'environ 4 µm d'épaisseur. Ces sections sont collées sur des lames de microscopie et doivent être traitées (chauffage, déparaffinage, réhydratation, récupération antigénique) avant d'être utilisées pour les marquages souhaités (IHC, IF, IP). Le TMA utilisé a été fourni par le Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CR-CHUM) et provient de la biobanque de la fondation cancer du sein du Québec (*Quebec Breast Cancer Foundation*, QBCF). Il contient 28 carottes, mesurant environ 700 µm de diamètre et provenant de cinq types tissulaires, organisées comme le montre la Figure 3.6.

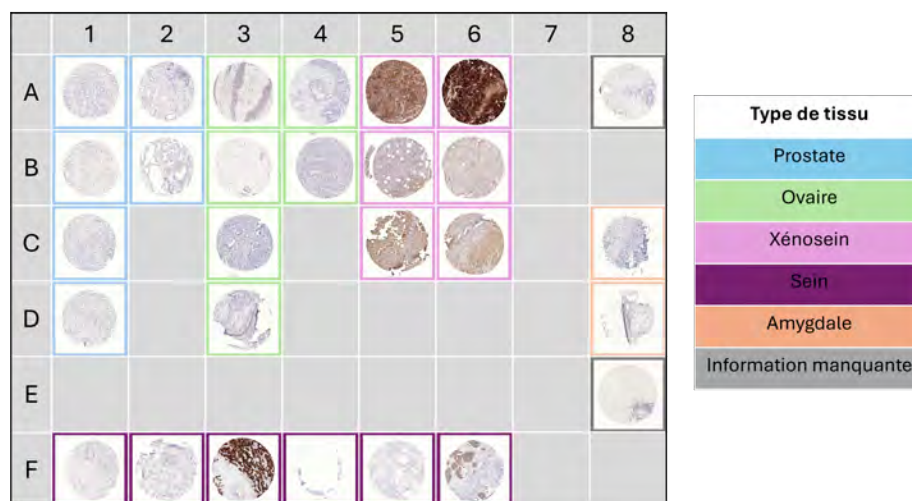


FIGURE 3.6 Carte du TMA utilisé. Les images utilisées pour illustrer chaque carotte sont les images de l'IHC réalisée avec le clone EPR.

L'utilisation d'un tel TMA permet de disposer sur une même lame de différents tissus exprimant la même protéine d'intérêt HER2 à différents niveaux. Cela devrait nous permettre d'évaluer les performances de la méthode IP à quantifier la protéine HER2 et à reconnaître les cas de faible expression HER2-low.

Pour connaître le niveau d’expression de la protéine HER2 de chaque carotte, une IHC avec le clone EPR a été performée sur une coupe du TMA. La procédure a été réalisée au CR-CHUM à l’aide de l’automate Discovery Auto-stainer (Ventana). Le Tableau 3.1 présente les scores IHC établis par deux pathologistes du CR-CHUM en utilisant les directives de l’ASCO/CAP.

TABLEAU 3.1 Scores IHC attribués par deux pathologistes. ND : Non Déterminé. NE : Non Évalué. Les carottes encadrées en rouge sont celles pour lesquelles une discordance inter-pathologistes a été observée. La carotte de sein F4 n’a pas pu être scorée car elle a été arrachée lors du processus de découpe au microtome et de montage sur la lame. Les autres carottes non scorées ont été considérées comme de moindre intérêt par les pathologistes sollicitées.

	Prostate						Ovaire						Xénossein						Sein						Amygdale		ND	
	A1	B1	C1	D1	A2	B2	A3	B3	C3	D3	A4	B4	A5	B5	C5	A6	B6	C6	F1	F2	F3	F4	F5	F6	C8	D8	A8	E8
Pathologiste 1	0	0	0	0	0	0	2	2	0	0	0	1	3	0	0	3	1	3	0	0	3	NE	0	2	0	0	NE	NE
Pathologiste 2	NE	NE	NE	NE	NE	NE	1	NE	NE	NE	NE	NE	3	1	1	3	1	3	1	1	3	NE	0	2	NE	NE	NE	NE
SCORE MOYEN	0	0	0	0	0	0	1.5	2	0	0	0	1	3	0.5	0.5	3	1	3	0.5	0.5	3	NE	0	2	0	0	NE	NE

Parmi les 12 carottes évaluées par les deux pathologistes, on recense 5 cas de discordance. Cette observation confirme la faible concordance inter-pathologistes lors de l’évaluation IHC, en particulier pour les cas de faible expression. Ces carottes équivoques se sont vues attribuer des scores intermédiaires (0,5 et 1,5) et feront l’objet d’une attention particulière lors de l’évaluation par IP.

L’organisation d’un tissu varie énormément en fonction du type tissu étudié mais aussi en fonction de la zone du tissu observée et du type de cancer. La Figure 3.7 montre la diversité de morphologies que l’on peut observer dans un seul et même type de tissu, ici le sein.

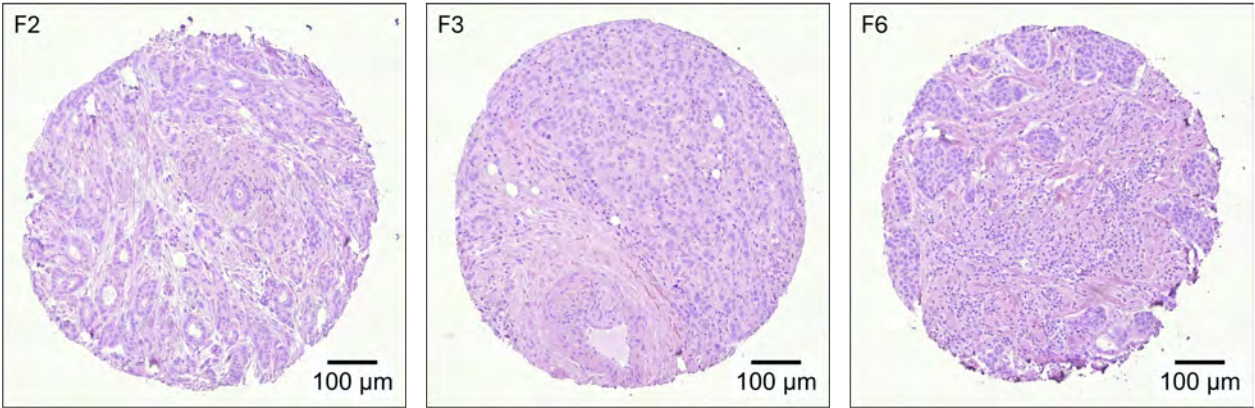


FIGURE 3.7 Carottes de cancer du sein colorées par H&E montrant la diversité de morphologies observables.

Toutefois, on peut simplifier l'architecture d'un tissu en considérant des cellules parenchymateuses (cellules principales responsables de la fonction de l'organe), entourées de tissus conjonctifs (charpente qui soutient les cellules). Les tissus contiennent aussi des vaisseaux sanguins et des nerfs qui permettent le bon fonctionnement des organes. Dans le cas des tissus mammaires, les cellules principales sont les cellules épithéliales. Le tissu conjonctif environnant appelé stroma est composé de matrice extracellulaire, de fibres structurales (collagène, élastine), de cellules conjonctives (fibroblastes, adipocytes) et de cellules inflammatoires (lymphocytes, macrophages, neutrophiles, mastocytes, etc.).

La Figure 3.8 montre la carotte F3 en précisant le type cellulaire auquel appartiennent quelques noyaux. On peut y voir des cellules épithéliales néoplasiques qui ont envahi le stroma de manière désorganisée, il s'agit d'un carcinome invasif.

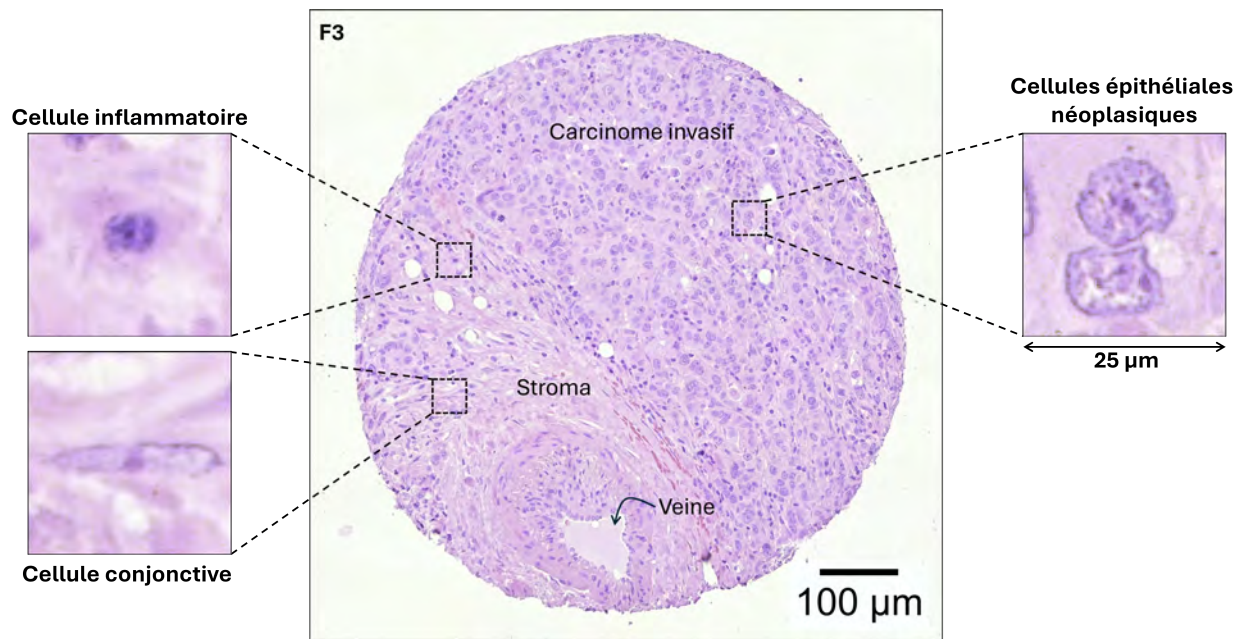


FIGURE 3.8 Carotte de cancer du sein (F3) colorée par H&E. Les régions agrandies de taille 25 µm par 25 µm montrent différents types cellulaires observables.

L'incubation des NPs sur TMA a été réalisée par Jennyfer Zapata Farfan, PhD. Le protocole utilisé inclut des étapes de déparaffinage, récupération antigénique, blocage, incubation des NPs et contre-coloration H&E. Les TMA ont été utilisés pour développer la méthodologie présentée dans le Chapitre 4, et les analyses réalisées dans le Chapitre 5.

Matrices de cellules (CMA)

Les matrices de cellules (*cell microarrays*, CMA) sont des assemblages de carottes similaires aux TMA décrits précédemment, mais constitués de cultures cellulaires plutôt que de tissus. La préparation d'un bloc de CMA nécessite généralement de cultiver les cellules d'intérêt jusqu'à l'obtention d'une confluence élevée. Les cellules sont ensuite fixées à la formaline, détachées, centrifugées puis resuspendues dans une solution d'agarose. Les cylindres contenant ce mélange cellules/agarose sont ensuite inclus dans un bloc de paraffine afin de construire le bloc de CMA. Des sections d'environ 4 μm d'épaisseur sont découpées au microtome et collées sur des lames de microscopie. Celles-ci doivent être traitées comme les coupes de TMA (chauffage, déparaffinage, réhydratation, récupération antigénique) avant d'être utilisées pour les marquages souhaités (IHC, IF, IP) [66,67]. Les CMA représentent un outil précieux pour la recherche clinique. Ils permettent notamment de comparer le marquage d'un même biomarqueur sur différentes lignées cellulaires présentant des niveaux d'expression protéiques distincts. Cela permet de réaliser des preuves de concepts indispensables au développement et à la validation de nouveaux biomarqueurs [66] ou de nouvelles méthodes d'évaluation de l'expression d'un marqueur [68].

Dans le cadre du développement de la méthode IP, les CMA sont utilisés pour étudier l'attachement des NPs à différentes lignées cellulaires dont le profil moléculaire est connu. Contrairement aux TMA, ils restreignent l'objet biologique étudié à un type cellulaire unique par carotte.

Le CMA utilisé dans ce travail a été acquis chez Array Science (réf. BC-7-1-3-2, Breast Marker Control - HER2Low). Il est composé de sept carottes de sept lignées cellulaires (Jurkat, DLD-1, BT-20, T-47D, BT-483, ZR-75-1, BT-474), chacune présentant un niveau distinct d'expression de la protéine HER2, allant du score 0 avec la carotte 1 (Jurkat) au score 3+ avec la carotte 7 (BT-474). Chaque carotte a été initialement évaluée par IHC en utilisant le clone 4B5. Les images et scores fournis par Array Science sont présentés dans la Figure 3.9. Une seconde évaluation par IHC, en utilisant le clone EPR, a été réalisée à l'Institut de Recherche en Immunologie et en Cancérologie (IRIC). Cette évaluation permet la comparaison directe de l'IHC avec la méthode IP qui emploie des NPs fonctionnalisées avec le clone EPR. Le scoring des carottes a été effectué par une pathologiste du CR-CHUM.

Les résultats présentés dans la Figure 3.9 mettent en évidence une forte discordance entre les scores obtenus avec le clone 4B5 et ceux obtenus avec le clone EPR. De telles divergences, déjà rapportées par le groupe du Professeur Meunier sur des TMA [44], sont courantes lors

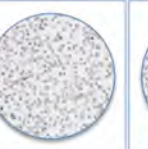
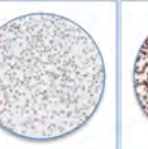
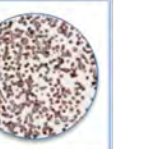
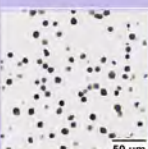
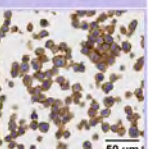
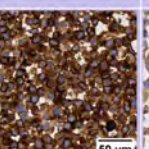
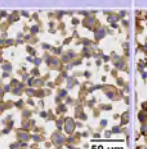
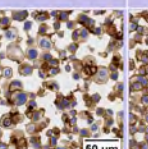
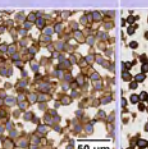
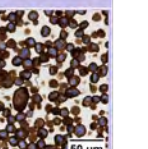
	C1 Jurkat	C2 DLD-1	C3 BT-20	C4 T-47D	C5 BT-483	C6 ZR-75-1	C7 BT-474
Clone 4B5 Images et scores fournis par Array Science							
	0	0 / 1+ (Ultra low)	0 / 1+	1+	1+ / 2+	1+ / 2+	3+
Clone EPR Performé et scanné à l'IRIC	Marquage cytoplasmique	3+	3+	3+	3+	3+	3+
							

FIGURE 3.9 Images des IHC et scores HER2 des sept carottes du CMA HER2Low Array Science. Les clones 4B5 et EPR 19547-12 montrent une très grande discordance de marquage, entraînant une très grande discordance des scores.

de l'évaluation d'un même marqueur avec différents clones [69–73]. Ces variations peuvent s'expliquer par les différences d'accessibilité et d'affinité des épitopes ciblés par chaque clone. Les clones 4B5 et EPR ciblent clairement des épitopes différents puisque EPR cible le domaine intracellulaire de la protéine HER2 tandis que 4B5 cible le domaine extracellulaire de cette même protéine. Ainsi, les épitopes ciblés ont des localisations différentes et correspondent à des séquences d'acides aminés distinctes dont la conformation peut varier. Par conséquent, l'accessibilité d'un épitope et son affinité avec le clone qui le cible peut varier, induisant une différence dans le marquage observé en fonction du clone utilisé. Aussi, les conditions expérimentales retenues pour réaliser l'IHC peuvent contribuer à un marquage plus ou moins intense.

Ces résultats compromettent l'objectif initial d'utiliser ce CMA et le clone EPR pour statuer sur les cas de faible expression. Néanmoins, le marquage étant positif, ce support est tout de même exploité pour le développement de la méthode IP (optimisation des protocoles, multiplexage, adaptation à de nouveaux clones, etc.).

L'incubation des NPs sur CMA a été réalisée par Jennyfer Zapata Farfan, PhD. Le protocole utilisé inclut des étapes de déparaffinage, récupération antigénique, blocage, incubation des NPs et contre-coloration H&E. La section 5.3 présentera l'adaptation de la méthodologie d'analyse développée pour les TMA aux CMA.

Produits Boston Cell Standards - BCS

La compagnie Boston Cell Standards (BCS) souhaite proposer des solutions afin de réduire le taux d'erreur élevé observé en IHC, tel que discuté dans la section 2.1.2. Une de leurs études a permis de conclure que la principale cause de ces erreurs réside dans l'absence de mesures traçables, de standards internationaux, ainsi que de calibration et de contrôle quantitatif [19]. BCS commercialise donc des outils visant à instaurer des procédures de calibration et des contrôles de qualité de routine pour l'IHC. L'entreprise propose deux produits fabriqués à partir de microbilles de verre d'un diamètre de 7 à 8 μm . Ces billes sont recouvertes des protéines d'intérêt (soit de la protéine entière, soit d'un peptide seulement) à une concentration connue et peuvent être colorées par IHC.

IHControls®

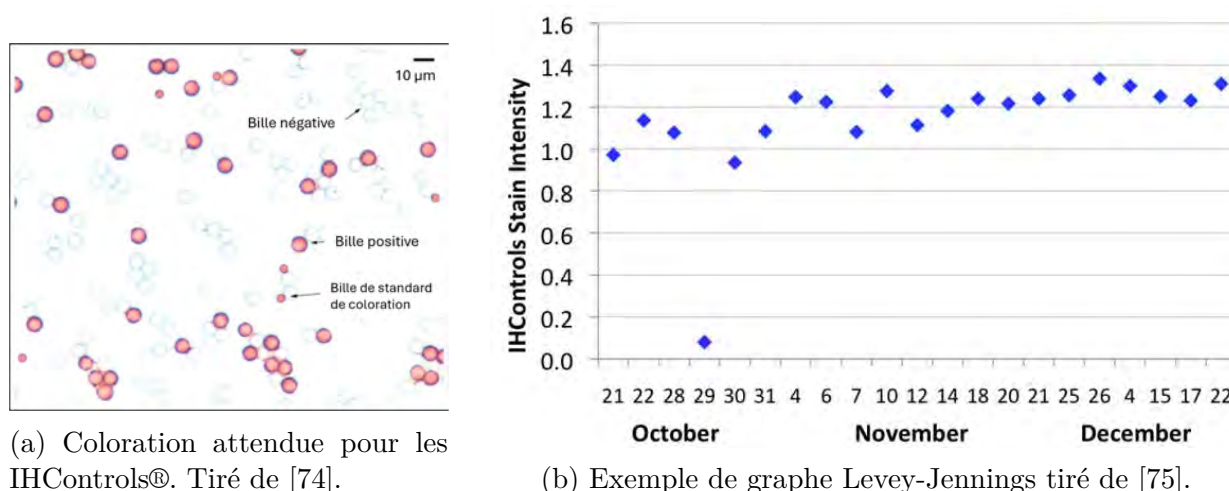


FIGURE 3.10 Résultats attendus et exemple d'analyse pour les IHControls®.

Les contrôles sont des solutions contenant les billes en suspension dans une matrice. Ils sont destinés à être utilisés comme contrôles de routine : lors de chaque procédure d'IHC, une goutte de la solution de contrôle doit être déposée à proximité du tissu à analyser. Les microbilles sont conçues pour subir les mêmes étapes de traitement que l'échantillon tissulaire, incluant le chauffage, le déparaffinage et la récupération antigénique. Après la coloration par IHC, elles peuvent être observées en champ clair au microscope optique. Une coloration brune doit être visible sur certaines microbilles, correspondant aux contrôles positifs, tandis que les contrôles négatifs restent incolores. Enfin, des microbilles de plus petit diamètre présentent une coloration orangée permanente et servent de standard d'intensité pour

la calibration et l'analyse quantitative des images. Les résultats peuvent être analysés pour tracer un graphique de Levey-Jennings permettant de suivre l'intensité de coloration au fil des expériences [75]. Le contrôle utilisé est un panel pour le cancer du sein, qui contient des billes présentant les protéines HER2, ER et PR (réf. BRLS11).

IHCalibrators®

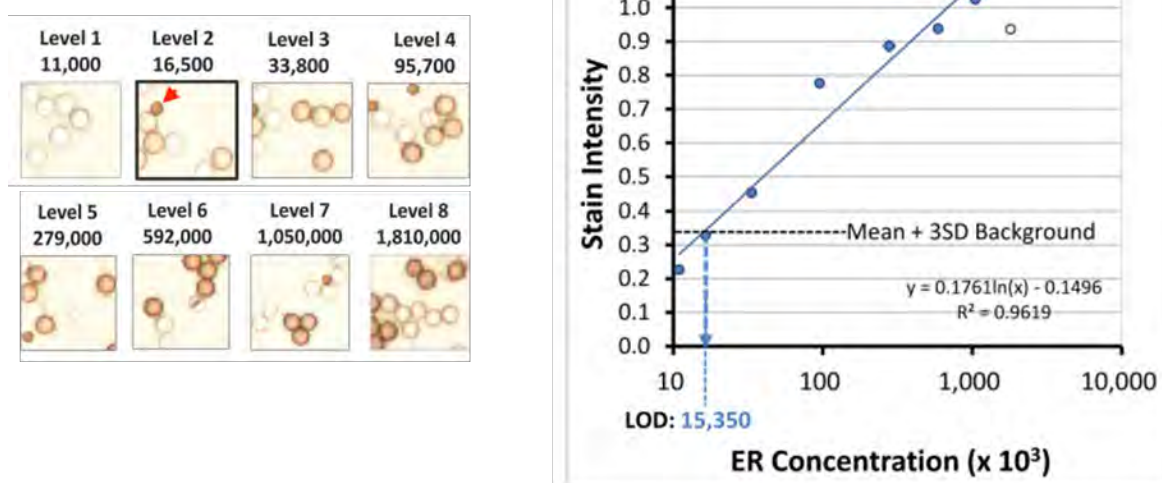


FIGURE 3.11 Exemple de résultats attendus et d'analyse pour les IHCalibrators®. Tiré de [76]

Les calibrateurs sont des lames sur lesquelles des dépôts de billes ont été préalablement réalisés. Chaque lame contient dix dépôts positifs et cinq dépôts négatifs servant de contrôles. Les dépôts positifs contiennent des billes dont la concentration en protéine est connue et croît d'un dépôt à l'autre. Ces lames permettent d'évaluer la coloration en fonction de la concentration protéique afin de déterminer la plage dynamique du test étudié. Les courbes obtenues permettent de déterminer la limite de détection (*Limit of Detection*, LOD), correspondant à la concentration minimale pour laquelle une coloration est observable. Elles permettent également d'identifier la concentration à partir de laquelle la coloration reste stable, indiquant un plateau de saturation. Ces courbes sont particulièrement utiles pour optimiser un protocole, notamment pour sélectionner la dilution optimale de l'anticorps utilisé en IHC, ou pour comparer les performances de deux tests utilisant des protocoles différents, où l'anticorps utilisé peut varier [77]. Le calibrateur utilisé présente dix niveaux croissants de la protéine HER2 (réf. BRRCB).

3.2 Acquisition des images

L'acquisition des images est réalisée grâce au microscope inversé Ti-Eclipse de Nikon. Les mouvements du porte-échantillon en X et Y sont assurés grâce au contrôleur Prior III.

Les images doivent être acquises dans deux modalités :

- En **champ sombre** pour la visualisation des NPs. L'illumination de l'échantillon est assurée par le dispositif d'illumination latérale (SI) décrit précédemment. Les trois DEL RGB sont allumées à leur intensité maximale.
- En **champ clair** pour la visualisation de la contre-coloration H&E. Un condenseur est aligné avec la source pour atteindre l'illumination de Köhler.

Pour les deux modalités d'illumination, les images sont acquises avec une caméra couleur PCO (pco.panda 4.2 sCMOS). Cette caméra est composée de 2048 pixels par 2048 pixels. Les pixels mesurent 6,5 μm par 6,5 μm .

Le Tableau 3.2 présente les paramètres optiques en fonction des différents objectifs utilisés.

TABLEAU 3.2 Paramètres optiques en fonction des objectifs utilisés.

	40x air NA=0,6		60x air NA=0,7		100x huile NA=1,25
Zoom	1x	1,5x	1x	1,5x	1,5x
Taille d'un pixel	0,1625 μm x 0,1625 μm	0,1083 μm x 0,1083 μm	0,1083 μm x 0,1083 μm	0,0722 μm x 0,0722 μm	0,0433 μm x 0,0433 μm
Taille du champ de vue imagé	332,8 μm x 332,8 μm	221,8 μm x 221,8 μm	221,8 μm x 221,8 μm	147,67 μm x 147,9 μm	88,68 μm x 88,68 μm
Limite de résolution latérale	0,52 μm	0,52 μm	0,44 μm	0,44 μm	0,24 μm
Profondeur de champ (DoF)	0,61 μm	0,61 μm	0,43 μm	0,43 μm	0,19 μm
	TMA	CMA	BCS Contrôles qualitatifs	BCS (Figure 6.1)	BCS Contrôles quantitatifs

3.2.1 Imagerie des TMA

La découpe au microtome des blocs de TMA permet d'obtenir des sections de 4 μm d'épaisseur relativement planes. Les NPs sont visibles uniquement à la surface de la coupe, ce qui semble indiquer qu'elles ne pénètrent pas à l'intérieur de la section de tissu. Les sites accessibles au sommet de la coupe semblent être suffisants pour effectuer le marquage IP, comme l'indiquent les preuves de concept réalisées préalablement. En utilisant l'objectif 40x à air (NA=0,6), il est possible de parcourir une carotte d'environ 700 μm de diamètre au complet, sans perdre le focus. La Figure 3.12 illustre schématiquement une coupe de TMA.

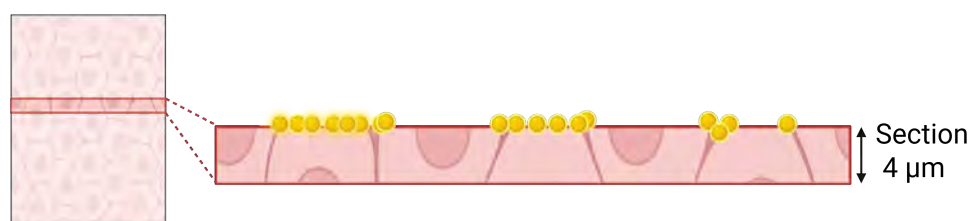


FIGURE 3.12 Représentation schématique d'une section de 4 μm dans un bloc de TMA.

L'acquisition d'une carotte au complet est donc réalisée en balayant l'échantillon dans le plan (X, Y), dans chaque modalité d'illumination. D'abord, le focus en Z est choisi manuellement de sorte à ce que les NPs restent visibles à la surface du tissu lorsque l'on se déplace d'un côté à l'autre de la carotte. Un premier scan est lancé pour acquérir l'image en SI. Puis, on change d'illumination et le focus en Z est réajusté pour garantir une bonne visibilité des structures biologiques et de la coloration H&E. Le deuxième scan est lancé pour acquérir l'image en BF. Les paramètres d'acquisition utilisés pour les deux modalités sont présentés dans le Tableau 3.3.

TABLEAU 3.3 Paramètres d'acquisition utilisés pour imager les TMA. Magnification 40x.

	Champ clair	Illumination latérale
Intensité de la source		Red: max Green: max Blue: max
Temps d'exposition	36 ms	500 ms
Saturation	10	10
Vibrance	-4	-4
Couleur de température	2034	4472
Teinte	16%	-65%
Contraste	0%	0%
Gamma	1,00	1,00
Plage dynamique	0 - 46617	0 - 46617

Les champs de vue acquis successivement sont utilisés pour reconstruire la carotte au complet en utilisant le plugin ImageJ Grid/Collection Stitching [78]. La Figure 3.13 montre une carotte entière après reconstruction dans les deux modalités d'illumination.

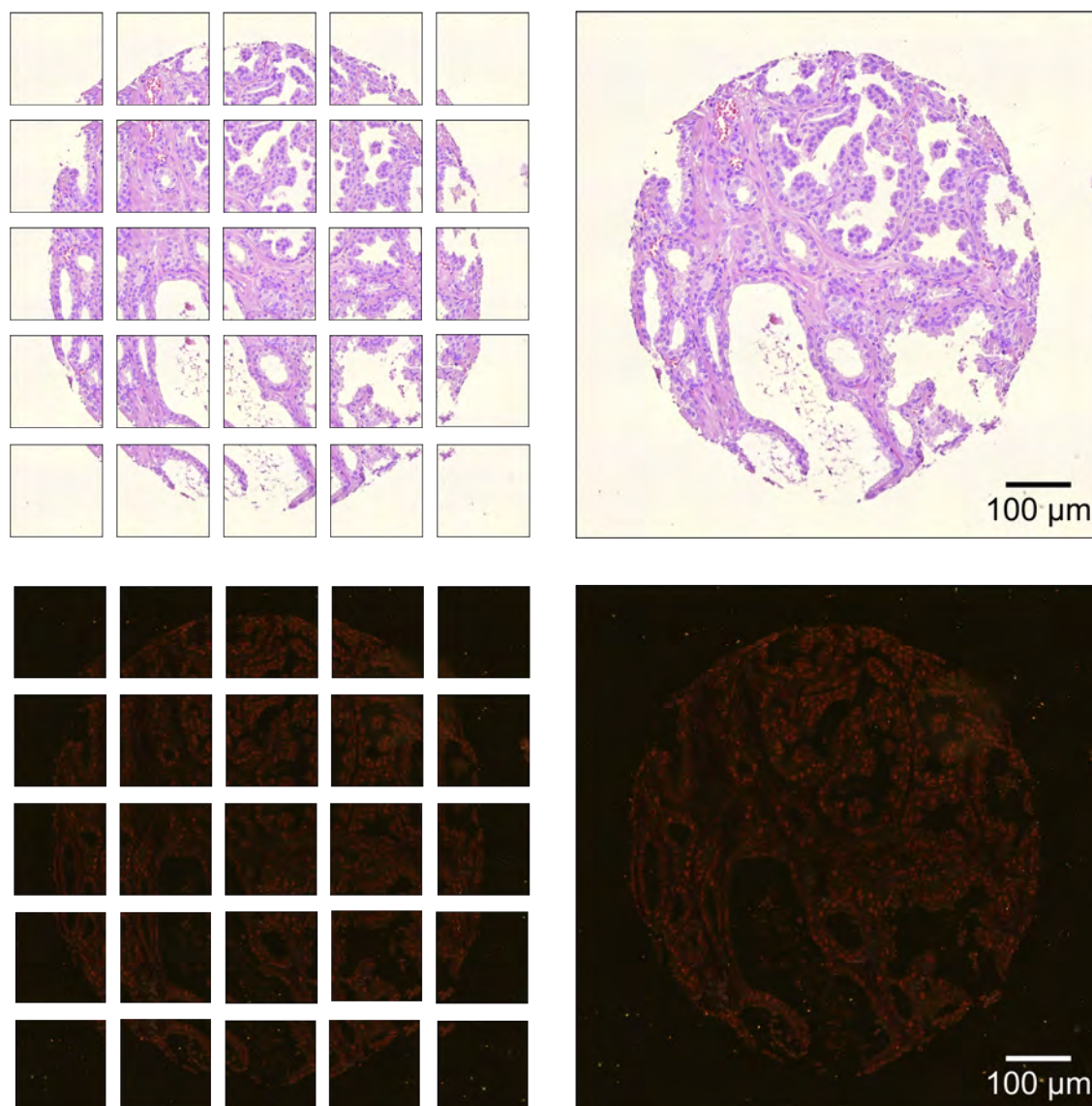


FIGURE 3.13 Reconstructions de la carotte B2 (Prostate) acquise en BF et en SI.

Après reconstruction, les images obtenues doivent être alignées pour que les deux modalités puissent se superposer parfaitement. Cette étape est cruciale afin que la position des NPs détectées dans l'image SI puisse être projetée sur l'image BF contenant la morphologie tissulaire. Pour cela, le plugin ImageJ Linear Alignment with SIFT est utilisé pour un alignement automatique [79]. Avant d'appliquer le plugin, l'image BF est inversée, de sorte à ce que les deux images à aligner soient plus proches visuellement (l'arrière plan de l'image BF passe de

blanc à noir au cours de l'inversion). Malheureusement, cette méthode n'est pas robuste et il est très courant que l'alignement ne soit pas satisfaisant. Dans ce cas, il est nécessaire de réaliser un alignement manuel avec le plugin ImageJ Big Warp [80]. Ce plugin fonctionne en sélectionnant manuellement dans les deux images au moins quatre points identiques.

Une fois alignées, les images sont redimensionnées pour qu'elles mesurent 800 μm par 800 μm et soient centrées autour de la carotte. La valeur moyenne des pixels de la périphérie de l'image est ensuite attribuée aux pixels vides de l'image BF qui peuvent être présents dans la périphérie de l'image après la reconstruction (script Python).

En utilisant cette méthodologie, 7 lames contenant 28 carottes ont été imagées. Comme le focus doit être réajusté pour chaque carotte, le scan est automatisé pour une acquisition carotte par carotte uniquement. Il faut donc compter 3h d'imagerie pour imager une lame au complet, dans les deux modalités. Même si cette durée peut sembler importante pour l'acquisition d'une lame, l'automatisation du scan grâce à un programme LabView a permis de réduire considérablement le temps d'imagerie, qui pouvait durer plusieurs jours. La Figure 3.14 montre un exemple des images obtenues après le scan d'une lame.

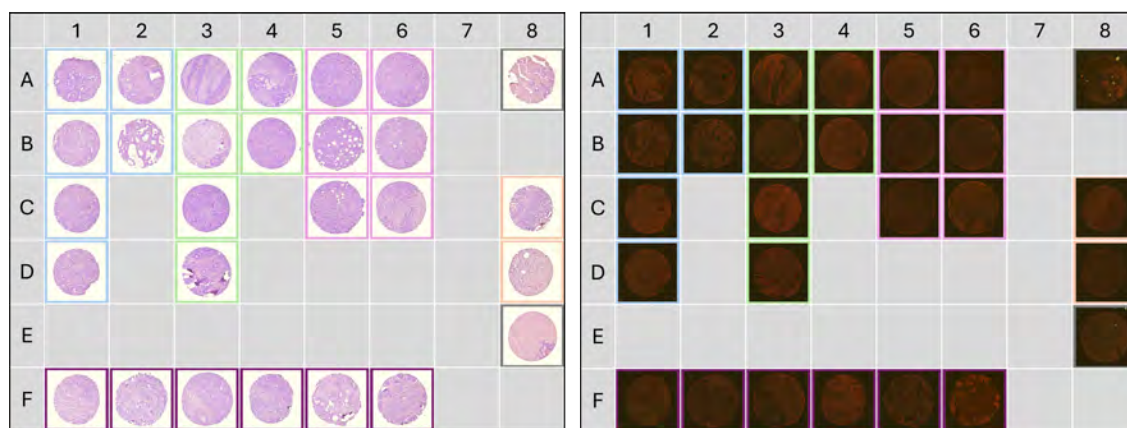


FIGURE 3.14 Lame 14 acquise dans les deux modalités d'illumination.

L'acquisition de ces 196 carottes dans les deux modalités d'illumination a permis de créer une banque d'images considérable qui été utilisée pour l'entraînement et l'évaluation des modèles de DL présentés dans le Chapitre 4.

3.2.2 Imagerie des CMA

Contrairement aux TMA décrits précédemment comme principalement plats, les coupes de CMA révèlent les membranes cellulaires à des hauteurs différentes dans une même section. La Figure 3.15 illustre schématiquement l'exposition variable du noyau, du cytoplasme et de la membrane de chaque cellule après une coupe au microtome. Comme les NPs ciblent la protéine transmembranaire HER2, elles peuvent se fixer à plusieurs hauteurs de la section. Ce paramètre complexifie l'acquisition des images puisque la profondeur de champ théorique de $0,61\text{ }\mu\text{m}$ conférée par l'objectif 40x ne permet pas de voir toutes les NPs qui peuvent s'étaler sur une section de hauteur $4\text{ }\mu\text{m}$. Toutefois, il semble qu'en pratique, il soit possible de visualiser des NPs à différentes hauteurs d'une coupe, même si le focus n'est pas parfait, comme le montre la Figure 3.16.

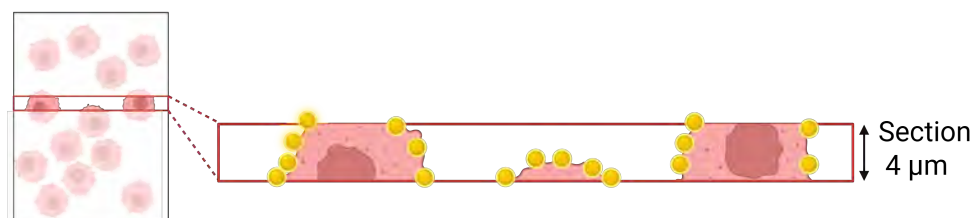


FIGURE 3.15 Représentation schématique d'une section de $4\text{ }\mu\text{m}$ dans un bloc de CMA. Les culots préparés en suivant la méthodologie décrite dans la Section 3.1.2 contiennent des cellules individuelles suspendues dans de l'agarose.

Par ailleurs, les tentatives de scan des carottes de diamètre $1000\text{ }\mu\text{m}$ se sont révélées particulièrement difficiles : le focus se perd facilement à mesure que l'on parcourt l'échantillon. Ce problème peut être attribué à une légère inclinaison de la lame sur la porte-échantillon du microscope, ou à une irrégularité de la coupe. Quelle qu'en soit l'origine, il semble compromis de scanner les carottes dans de telles conditions : le protocole d'acquisition des images ne serait pas suffisamment robuste pour garantir des images de qualité pour réaliser l'analyse.

Comme chaque carotte est composée d'un unique type cellulaire, il a été jugé raisonnable de n'imager qu'un seul champ de vue par carotte, afin de mieux contrôler la mise au point de chaque image acquise. Pour réduire davantage le champ de vue et assurer un contrôle encore plus important sur l'image acquise, le zoom x1.5 du microscope est aussi utilisé.

L'objectif utilisé est le même que pour l'acquisition des TMA (40x à air, $NA=0,6$). L'image acquise pour chaque carotte est un champ de vue de $221,8\text{ }\mu\text{m}$ par $221,8\text{ }\mu\text{m}$. Chaque pixel mesure $0,1075\text{ }\mu\text{m}$. En moyenne, on compte une centaine de cellules par image.

Les paramètres d'acquisition utilisés pour les deux modalités sont présentés dans le Tableau 3.4.

TABLEAU 3.4 Paramètres d'acquisition utilisés pour imager les CMA. Magnification 40x + zoom, champ de vue réduit de 221,8 μm par 221,8 μm .

	Champ clair	Illumination latérale
Intensité de la source		Red: max Green: max Blue: max
Temps d'exposition	36 ms	500 ms
Saturation	10	10
Vibrance	-4	-4
Couleur de température	2034	4472
Teinte	16%	-65%
Contraste	0%	0%
Gamma	1,00	1,00
Plage dynamique	0 - 46617	0 - 46617

La Figure 3.16 montre les images acquises pour une carotte (C7 - BT-474). Les images B1-B4 montrent des régions d'intérêt de 25 μm par 25 μm . On peut voir que toutes les NPs ne sont pas parfaitement nettes, mais le focus choisi permet de s'assurer de la visibilité du plus grand nombre de NPs aussi bien sur le substrat que sur le sommet des cellules. On peut observer un léger changement de couleur des NPs en fonction de leur position par rapport au plan focal.

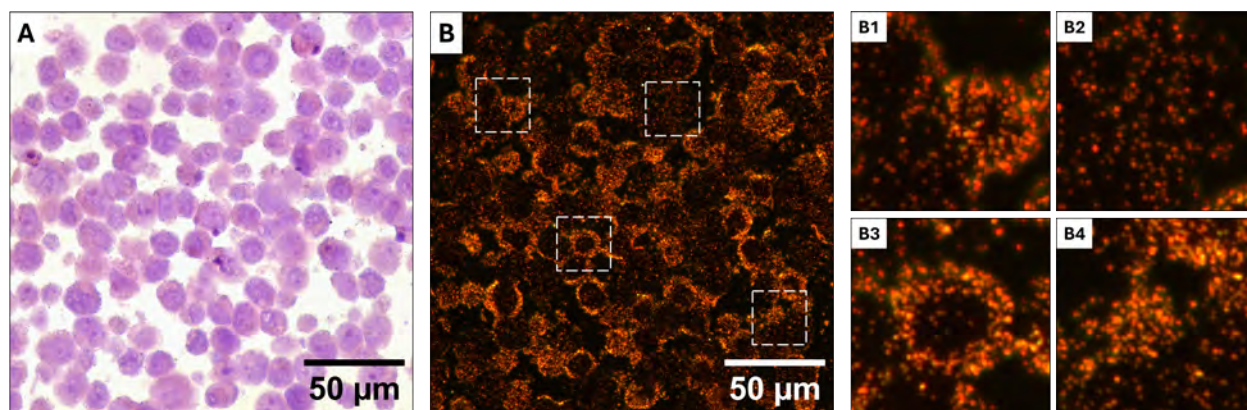


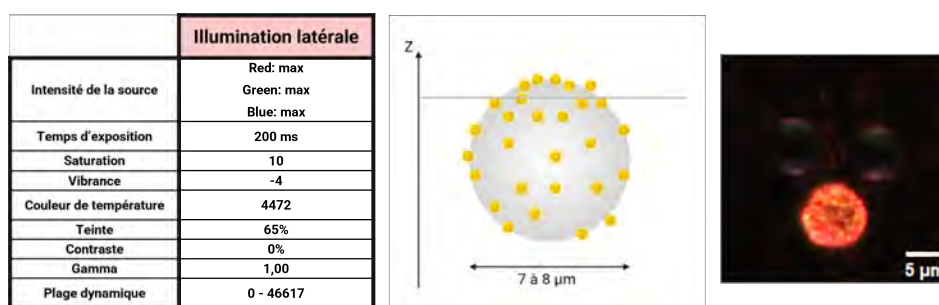
FIGURE 3.16 Exemple d'images BF (A) et SI (B) pour une carotte de CMA (C7 - lignée BT-474). Les régions d'intérêt B1-B4 mesurent 25 μm par 25 μm et montrent que les NPs sont visibles dans tout le champ de vue même si le focus n'est pas optimal pour chacune d'entre elles.

3.2.3 Imagerie des produits BCS

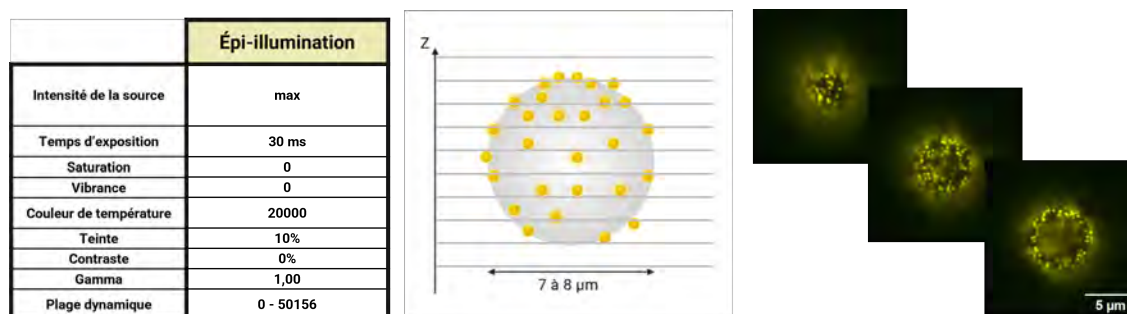
Les microbilles BCS mesurent entre 7 et 8 μm de diamètre.

Pour observer l'attachement des NPs sur les billes avec un aspect qualitatif seulement, il est possible d'acquérir une unique image en SI en s'assurant que le focus soit sur le sommet des billes. Ainsi, il est possible de visualiser le contraste entre les billes recouvertes de la protéine cible, sur lesquelles les billes s'attachent, et les autres billes sur lesquelles les NPs ne s'attachent pas. La Figure 3.17a présente les paramètres utilisés pour l'acquisition des images qualitatives des billes BCS, ainsi que le type d'image obtenu.

Pour imager l'intégralité des NPs attachées à la surface de ces billes couvertes de protéine, il est indispensable de réaliser un balayage selon l'axe Z. De plus, il est intéressant d'améliorer la résolution des images acquises en utilisant l'épi-illumination et un objectif à huile. Le balayage en Z est réalisé manuellement avec un pas d'environ 1 μm permettant d'acquérir 10 images par bille. La Figure 3.17b présente les paramètres utilisés pour l'acquisition des images permettant l'analyse quantitative des billes BCS, ainsi que le type d'image obtenu.



(a) Acquisition des billes BCS pour une analyse qualitative uniquement. Objectif 60x à air NA=0,7.



(b) Acquisition des billes BCS pour une analyse quantitative de l'attachement des NPs à la surface des billes. Objectif 100x à huile NA=1,25.

FIGURE 3.17 Paramètres d'acquisition pour l'imagerie des billes BCS et images obtenues.

CHAPITRE 4 OUTILS DE SEGMENTATION DES IMAGES

Les images acquises dans les deux modalités d’illumination contiennent des informations précieuses et complémentaires. Pour mener une analyse complète des échantillons, il est nécessaire de disposer d’outils de segmentation robustes. Dans les images en SI, l’objectif est de détecter la position des NPs. Les images en BF, quant à elles, fournissent des informations sur la morphologie biologique de l’échantillon. Elles serviront dans un premier temps à extraire le tissu, puis à détecter les noyaux des cellules, voire à les classer selon leur type.

4.1 Génération des masques de vérité terrain

Les 196 carottes de TMA, acquises dans les deux modalités d’illumination constituent une banque d’images considérable. Ces images ont été exploitées pour générer des jeux de données composés d’images brutes et de masques de référence, appelés « vérité terrain » (*Ground Truth*, GT). Un masque correspond à un tableau à deux dimensions de même taille que l’image à laquelle il correspond. Chaque valeur du masque correspond à la valeur attribuée au pixel de l’image ayant les mêmes coordonnées et permet de déterminer à quelle classe chaque pixel appartient. Les masques GT correspondent au résultat de segmentation attendu pour chaque image et sont utiles à l’entraînement et à l’évaluation des outils de segmentation. La Figure 4.1.B permet de visualiser le masque d’une image de NPs imagées en SI.

4.1.1 Méthodologie pour la génération des masques de vérité terrain

Les masques de vérité terrain ont été générés grâce au logiciel Dragonfly [81]. Il s’agit d’un logiciel d’analyse d’images dont l’interface interactive permet à la fois la visualisation d’images scientifiques et leur annotation intuitive. Dragonfly permet également d’utiliser ces annotations pour entraîner des modèles de DL, fonctionnalité qui sera détaillée dans la section 4.3. Les images SI et BF acquises au format RGB sont importées dans Dragonfly en utilisant un unique canal de luminance ($Luminance = 0.2126 R + 0.7152 G + 0.0722 B$). Les outils ROI Painter (*Region Of Interest Painter*) offrent la possibilité de « peindre » directement sur l’image afin d’associer une classe donnée à chaque pixel. Une option de seuillage local selon Otsu est disponible et permet d’annoter facilement les pixels en fonction de leur intensité locale. Il est aussi possible d’utiliser un modèle de DL entraîné à segmenter l’image selon les classes souhaitées pour obtenir une vérité terrain approximative. Cette approche hybride permet de gagner un temps considérable. La correction manuelle reste toutefois indispensable

afin de limiter les biais introduits par le modèle et d’assurer la fiabilité des masques de vérité terrain. Une fois l’annotation terminée, il est possible d’exporter une image en niveaux de gris, où chaque valeur correspond à une classe spécifique.

4.1.2 Vérité terrain pour la segmentation des NPs

Les masques de vérité terrain pour la détection des NPs distinguent deux classes : *nanoparticule* et *arrière-plan*. Les coordonnées correspondant aux NPs prennent la valeur 255 dans le masque en niveaux de gris, tandis que ceux de l’arrière-plan ont la valeur 0. La position de chaque NP est ensuite extraite en séparant chaque NP comme une particule individuelle puis en calculant son centre de masse. La séparation en particules individuelles est réalisée à l’aide de l’outil “*Separate 4 Connected Components*”, qui identifie les composantes connexes en considérant uniquement les pixels voisins par leurs 4 côtés. Pour l’évaluation quantitative de la détection des NPs, les positions vérité terrain des NPs sont finalement exportées dans un fichier Excel.

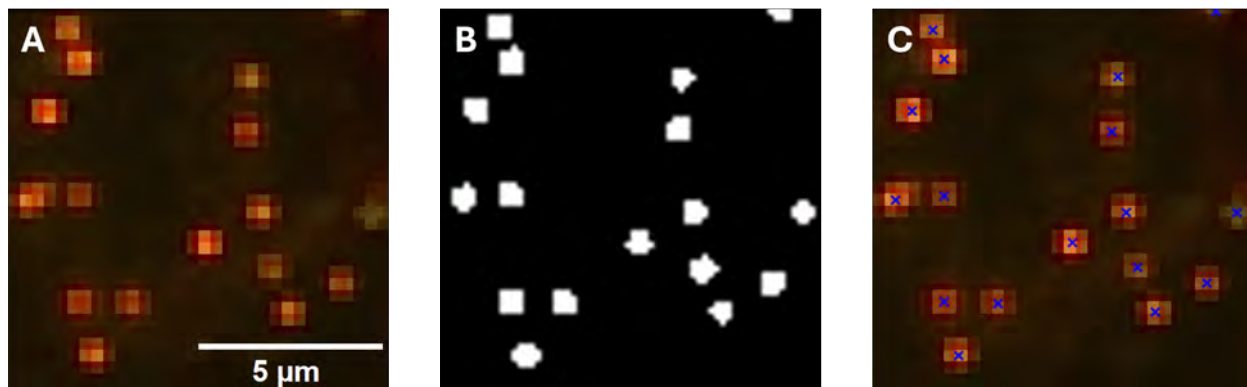


FIGURE 4.1 Génération du masque de vérité terrain pour la segmentation des NPs. A : image SI, B : masque généré grâce aux annotations en deux classes, C : image SI avec la superposition des positions du centre de masse de chaque NPs.

4.1.3 Vérité terrain pour la segmentation du tissu

Pour la détection du tissu, les deux classes des masques de vérité terrain sont *tissu* et *substrat*. Les pixels correspondant au tissu prennent la valeur 255, et ceux du substrat la valeur 0. Les masques en niveaux de gris sont exportés sous la forme d’images .tif.

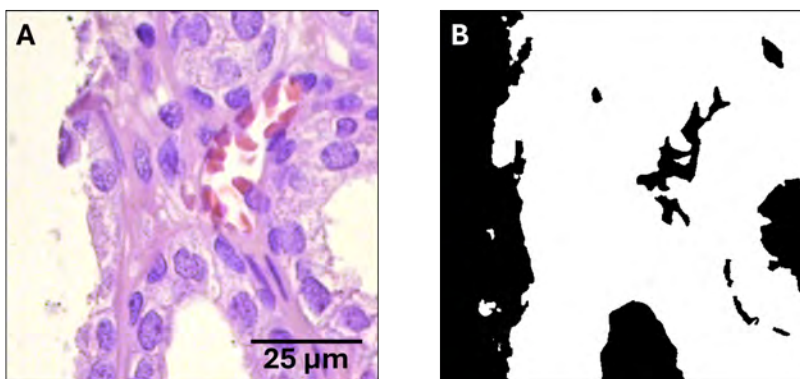


FIGURE 4.2 Génération du masque de vérité terrain pour la détection du tissu. A : image BF, B : masque de vérité terrain.

4.1.4 Jeux de données générés

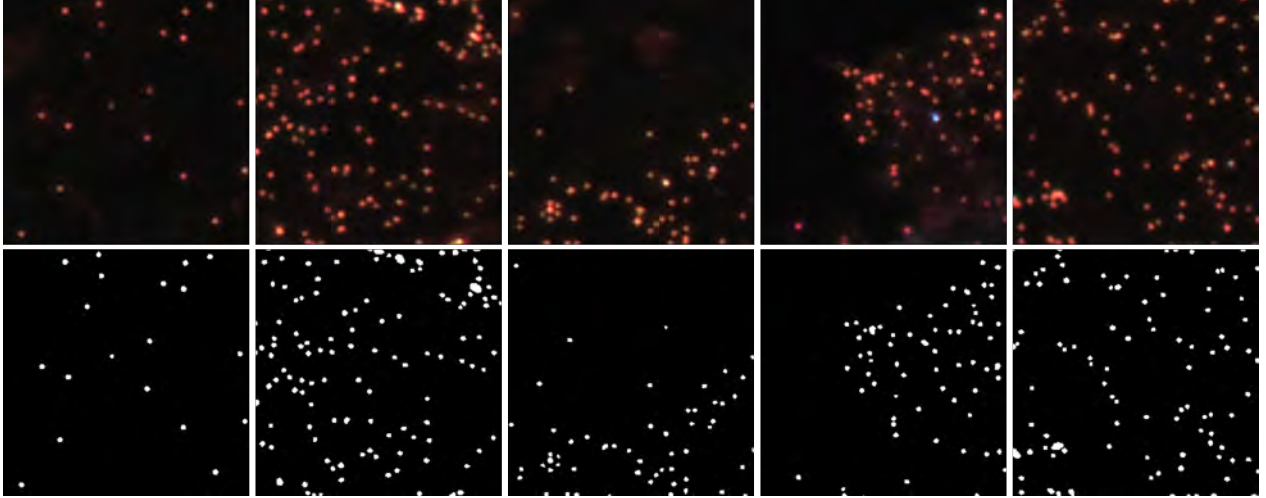
Les paires image / vérité terrain permettent de constituer des banques de données destinées à l'entraînement et à l'évaluation des outils de segmentation. Ainsi, trois jeux de données, présentés dans le Tableau 4.1 ont été générés.

TABLEAU 4.1 Description des jeux de données générés.

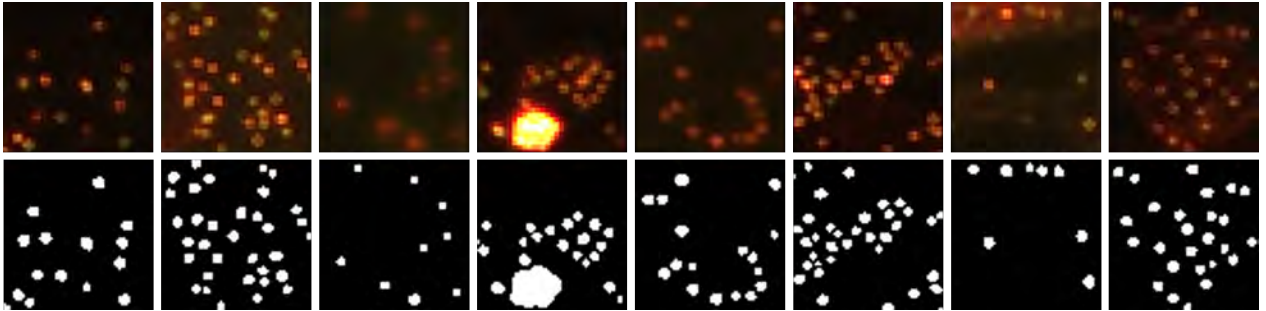
	Segmentation des NPs sur les images SI		Segmentation du tissu sur les images BF
	Jeu de données A	Jeu de données B	Jeu de données C
Conditions d'acquisition des images	SI Objectif 40x air NA=0,6 + zoom 1,5x	SI Objectif 40x air NA=0,6	BF Objectif 40x air NA=0,6
Taille des parcelles	256 pixels x 256 pixels 27,72 µm x 27,72 µm	64 pixels x 64 pixels 10,4 µm x 10,4 µm	512 pixels x 512 pixels 83,2 x 83,2 µm
Origine des parcelles	1 seule carotte d'une seule lame (Lame 7 – A6)	28 carottes de 7 lames différentes	28 carottes de 7 lames différentes
Nombre de parcelles utilisées pour l' entraînement	100 (71%)	300 (75%)	112 (67%)
Nombre de parcelles utilisées pour l' évaluation	40 (29%)	100 (25%)*	56 (33%)

*Images utilisées pour l'évaluation des performances de *Find Maxima*, section 4.2.

Chaque jeu de données est composé de parcelles découpées dans les images acquises. Ces parcelles permettent de témoigner au mieux de la diversité des images, sans avoir à segmenter manuellement et entièrement toutes les images. En effet, il serait extrêmement long de segmenter manuellement les images entières et cela ne serait pas pertinent pour l'entraînement des modèles de DL. Il est préférable de disposer de petites parcelles en grand nombre représentant la diversité des images plutôt que de grandes parcelles moins nombreuses qui ne couvriraient pas l'ensemble de la diversité observée.



(a) Exemples de parcelles issues du **jeu de données A** pour la segmentation des NPs.



(b) Exemples de parcelles issues du **jeu de données B** pour la segmentation des NPs.

FIGURE 4.3 Exemples de parcelles issues des jeux de données A (a) et B (b) pour la segmentation des NPs.

En particulier, le jeu de données B, constitué de parcelles de 64 pixels par 64 pixels, offre une bonne représentation de la diversité des images acquises en SI. On y retrouve des NPs plus ou moins nettes, avec des signaux de différentes intensités, des variations de couleur et des arrière-plans plus ou moins sombres. Les parcelles denses en NPs sont indispensables pour

représenter le principal défi de la segmentation des NPs, à savoir la détection et séparation des NPs qui se trouvent très proches les unes des autres.

Pour les images acquises en BF, les parcelles sélectionnées permettent de couvrir la diversité des morphologies des différents types tissulaires du TMA, ainsi que les variations de contre-coloration H&E observables d’une lame à l’autre.

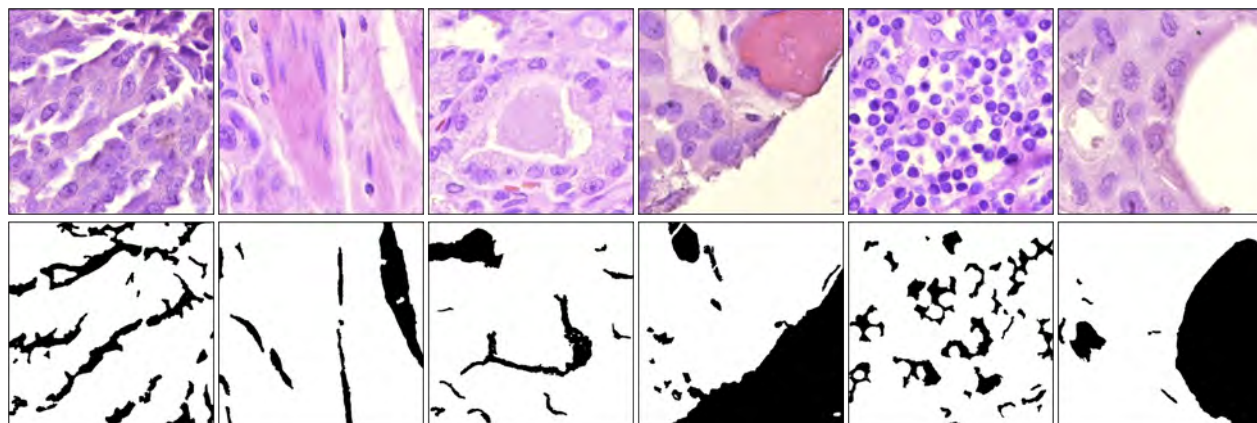


FIGURE 4.4 Exemple de parcelles issues du **jeu de données C** pour la segmentation du tissu.

4.1.5 Méthodologie pour l’évaluation des outils de segmentation

L’évaluation des outils de segmentation permet de s’assurer de leur robustesse, qualité indispensable pour instaurer la confiance nécessaire au développement d’une stratégie d’analyse basée sur ces outils.

L’évaluation des performances est réalisée en comparant la sortie générée par l’outil de segmentation au masque de vérité terrain correspondant, pour des images n’ayant pas été utilisées lors de l’entraînement. La fiabilité de l’évaluation dépend donc directement de la qualité des masques générés précédemment.

Métriques pour l’évaluation de la détection des NPs

Le résultat d’une détection de NPs est évalué en comparant les positions des NPs détectées avec les positions des NPs de la vérité terrain. Chaque détection est considérée comme un Vrai Positif (VP) ou un Faux Positif (FP) en fonction de sa distance par rapport à la NP de la vérité terrain la plus proche, comme présenté dans la Figure 4.5. Le cercle de tolérance utilisé est suffisamment large pour que les NPs détectées par ImageJ, parfois légèrement décallées du

centre de masse, soient bien comptabilisées comme des VP. Les NPs de la vérité terrain qui ne sont pas associées à une NP détectée sont considérées comme des Faux Négatifs (FN).

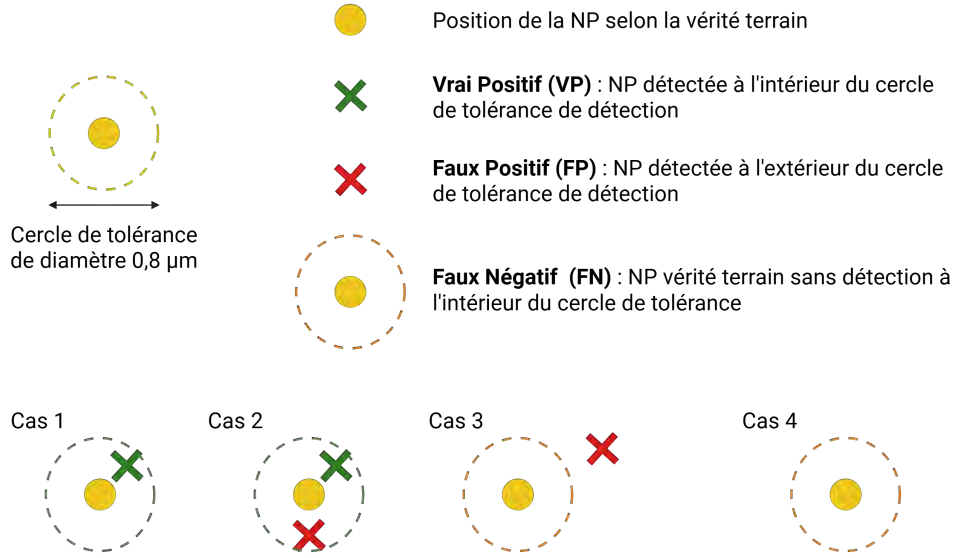


FIGURE 4.5 Représentation schématique de la classification en VP, FP et FN.

Une fois le nombre de VP, FP et FN établi, on peut calculer les trois métriques suivantes :

- **Précision (*Precision*)** : capacité de l'outil de détection à éviter les fausses détection (FP). Plus l'outil de détection surestime le nombre de NPs (FP), plus la précision diminue.

$$\text{Précision} = \frac{VP}{VP + FP} \quad (4.1)$$

- **Sensibilité (*Recall* ou *Sensitivity*)** : capacité de l'outil de détection à ne manquer aucune NP (FN). Plus l'outil de détection manque des NPs (FN), plus la sensibilité diminue.

$$\text{Sensibilité} = \frac{VP}{VP + FN} \quad (4.2)$$

- **F1-score** : moyenne harmonique de la précision et de la sensibilité. Il permet de combiner précision et sensibilité en une seule métrique pour résumer la performance

globale de l'outil de détection.

$$F1 - score = 2 \times \frac{\text{Précision} \times \text{Sensibilité}}{\text{Précision} + \text{Sensibilité}} \quad (4.3)$$

Métriques pour l'évaluation de la détection du tissu

La détection du tissu est évaluée en comparant le résultat de la segmentation avec le masque de vérité terrain comme illustré dans la Figure 4.6. Contrairement à la détection de NPs, l'évaluation se fait sur des régions continues plutôt que sur des positions ponctuelles. La métrique utilisée doit donc permettre de mesurer le chevauchement entre le résultat de la segmentation et le masque de vérité terrain. Pour cela, on considère les masques obtenus de manière binaire : le substrat prend la valeur 0 et le tissu prend la valeur 1. Ainsi, on peut raisonner avec une approche logique où l'intersection des deux masques donne 1 si le pixel étudié est classé comme tissu par les deux masques, ou 0 si la classe de la détection n'est pas la même que celle de la vérité terrain. On note S le résultat donné par le modèle de segmentation, et GT le masque de vérité terrain.

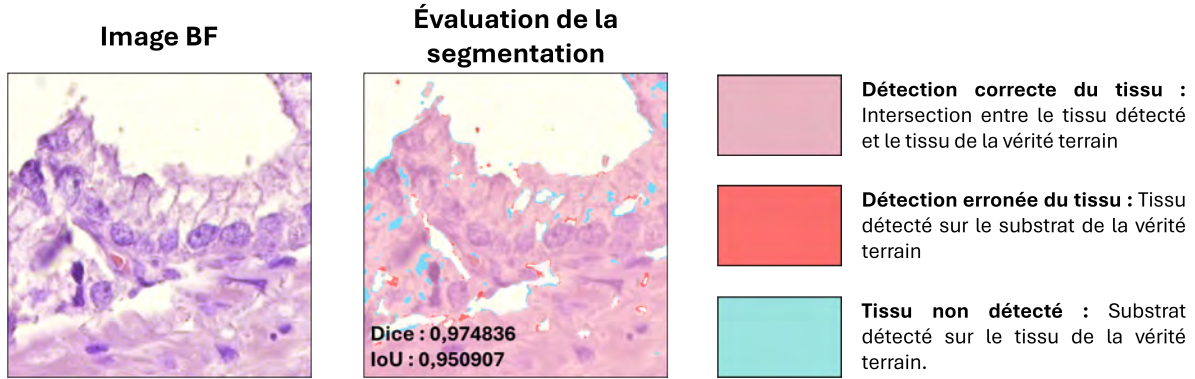


FIGURE 4.6 Évaluation de la détection du tissu.

Deux métriques sont envisageables :

- Le **score Dice** est utilisé pour minimiser l'impact des pixels isolés mal classés et se concentrer sur la qualité de la segmentation à l'échelle globale. En effet, il est souvent difficile de déterminer avec précision si chaque pixel individuel appartient réellement au tissu. Par conséquent, la vérité terrain doit être considérée de manière plus souple, afin de refléter la segmentation globale plutôt que les détails ponctuels. Le score Dice est la métrique la plus pertinente pour notre application.

$$\text{Dice} = \frac{2|S \cap GT|}{|S| + |GT|} \quad (4.4)$$

- L'**Intersection over Union (IoU, ou Jaccard index)** pénalise davantage les chevauchements partiels, ce qui n'est pas particulièrement souhaitable dans un contexte de détection du tissu qui ne se joue pas au pixel près.

$$\text{IoU} = \frac{|S \cap GT|}{|S \cup GT|} \quad (4.5)$$

4.2 Limites de l'algorithme naïf ImageJ

La détection des NPs par l'algorithme naïf *Find Maxima* de Fiji ImageJ est connue pour présenter certaines limites, notamment en termes d'adaptabilité à des images variées. Pour quantifier ces limites, une évaluation des performances de *Find Maxima* a été réalisée. Cette évaluation a aussi permis d'identifier la performance maximale dont est capable *Find Maxima*, qu'il faudra surpasser lors du développement de modèles de DL pour la détection des NPs.

L'évaluation a été menée grâce aux 100 images d'évaluation du jeu de données B. Ces images sont organisées en 4 groupes en fonction de la densité de NPs dans l'image :

- Faible densité : moins de 10 NPs par 100 μm^2 (29 images) ;
- Densité moyenne : entre 10 et 20 NPs par 100 μm^2 (23 images) ;
- Forte densité : entre 20 et 30 NPs par 100 μm^2 (28 images) ;
- Très densité : plus de 30 NPs par 100 μm^2 (20 images).

Les images proviennent de 5 lames différentes et de 10 carottes différentes afin d'introduire de la diversité dans les données d'évaluation. Elles contiennent 2 237 NPs.

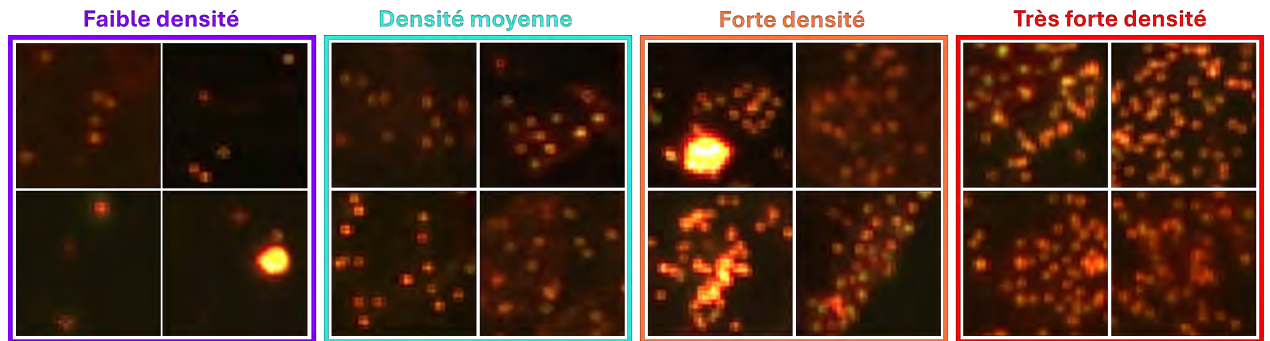


FIGURE 4.7 Exemple d'images pour les 4 groupes de densité des données d'évaluation.

Pour chacune des images d'évaluation, on dispose du masque de vérité terrain. Le centre de masse de chaque NP segmentée est calculé pour être utilisé comme les coordonnées de chaque NP. L'algorithme *Find Maxima* a été appliqué à chaque image en faisant varier la prominance de 1 à 30. Les positions détectées ont été comparées avec celles de la vérité terrain en suivant

la méthodologie présentée à la section 4.1.5 pour calculer la précision, la sensibilité (recall) et le F1-score.

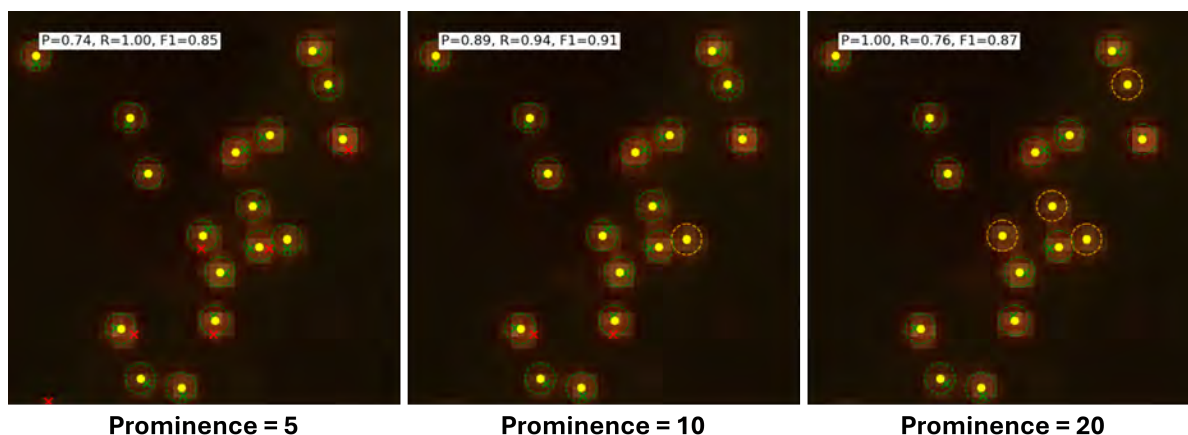


FIGURE 4.8 Détections de l’algorithme *Find Maxima* comparées aux positions vérité terrain. Les positions vérité terrain sont marquées par des points jaunes, tandis que les détections de *Find Maxima* sont en vert s’il s’agit de vrais positifs et en rouge s’il s’agit de faux positifs. Les scores pour chaque images sont affichés dans le coin supérieur gauche : P=précision, R=recall, F1=F1-score.

La Figure 4.9 montre la moyenne de la sensibilité, de la précision et du F1-score pour chaque groupe de densité, en fonction de la prominence. Sans surprise, **la prominence a un effet direct sur les métriques calculées pour chaque image**. Une prominence faible multiplie le nombre de détections, assurant que toutes les NPs soient détectées (sensibilité élevée), mais rapportant aussi un grand nombre de faux positifs (faible précision). Au contraire, une prominence élevée réduit le nombre de détection, limitant le nombre de faux positifs (précision élevée), mais augmentant le risque de ne pas détecter une NP (faible sensibilité). Le F1-score (moyenne harmonique de la précision et de la sensibilité) est quant-à-lui tiré vers le bas par la métrique la moins performante.

La Figure 4.9 suggère aussi que la prominence doit être ajustée en fonction de la densité de NPs dans l’image puisque par exemple, le F1-score maximal n’est pas atteint pour la même valeur de prominence pour les 4 groupes de densités. Cette dépendance entre prominence optimale et densité est visible dans la Figure 4.10.

Même si la dispersion des valeurs est grande, **les courbes de tendance montrent bien que la prominence optimale dépend de la densité de NPs dans l’image analysée**. Cette dépendance est particulièrement observée pour le calcul de la sensibilité, et par conséquent pour le calcul du F1-score. Pour une image avec peu de NPs, la prominence doit être élevée pour limiter les faux positifs. Pour une image avec beaucoup de NPs, la prominence

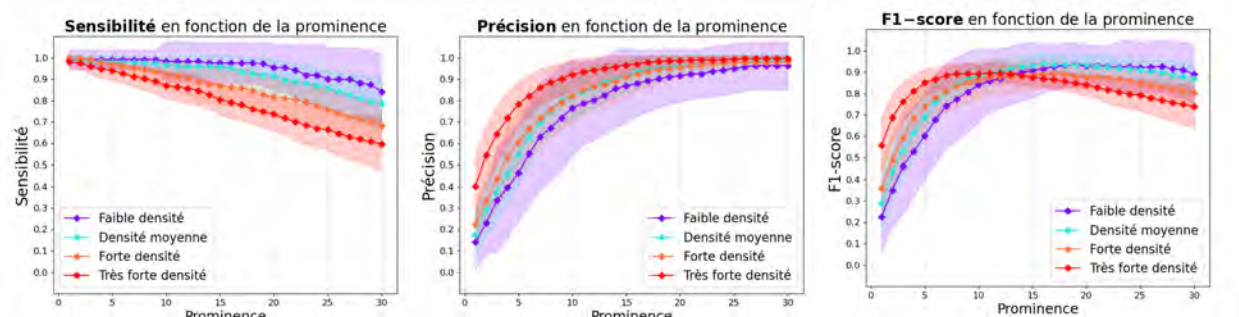


FIGURE 4.9 Sensibilité, précision et F1-score évalués pour chaque groupe de densité en fonction de la prominence. Le halo autour de chaque courbe correspond à l'écart-type des valeurs dans le groupe de densité étudié.

doit être ajustée à la baisse pour garantir un maximum de vrais positifs.

Dans des conditions optimales qui correspondent à l'ajustement de la prominence en fonction de la densité, voire à un ajustement itératif avec vérifications visuelles, l'algorithme *Find Maxima* peut être performant sur des régions d'intérêt de densité uniforme. Cependant, les images acquises sont grandes (800 μm par 800 μm) et la densité des NPs n'est pas uniforme. En effet, les tissus ont souvent des niveaux d'expression protéique hétérogènes, qui se traduisent par un attachement hétérogène des NPs sur l'échantillon. Il est donc impossible d'ajuster *Find Maxima* pour qu'il soit performant sur une image entière de 800 μm par 800 μm . Dans une perspective d'automatisation, l'ajustement manuel serait aussi particulièrement problématique.

La performance maximale de *Find Maxima* a tout de même été déterminée en identifiant la valeur de prominence permettant d'obtenir le meilleur F1-score pour les 100 images d'évaluation. La Figure 4.11 montre la moyenne sur les 100 images de la précision, de la sensibilité et du F-score en fonction de la prominence .

Le F1-score maximum 0,905 est atteint pour une prominence de 17. Ce F1-score sera la performance à surpasser lors du développement du modèle de DL assurant la détection des NPs.

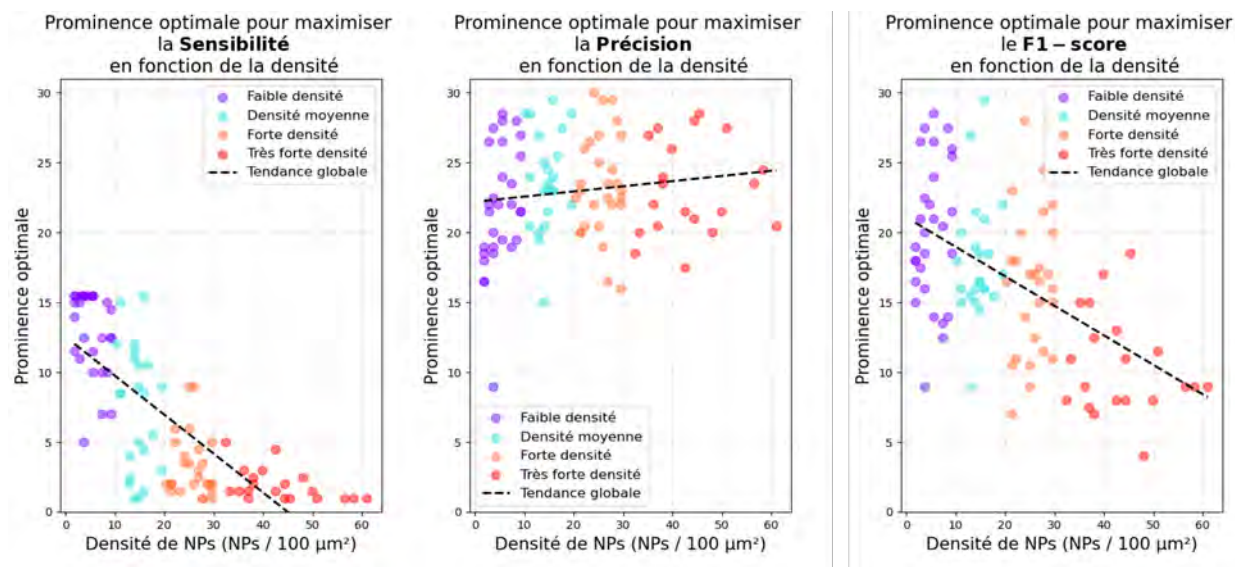


FIGURE 4.10 Prominence optimale permettant de maximiser la sensibilité, la précision et le F1-score pour chaque densité. Lorsque le même résultat peut être atteint avec plusieurs prominences, la prominence optimale a été calculée en faisant la moyenne de toutes les prominences permettant d'atteindre le maximum.

Précision, Sensibilité et F1-score en fonction de la prominence

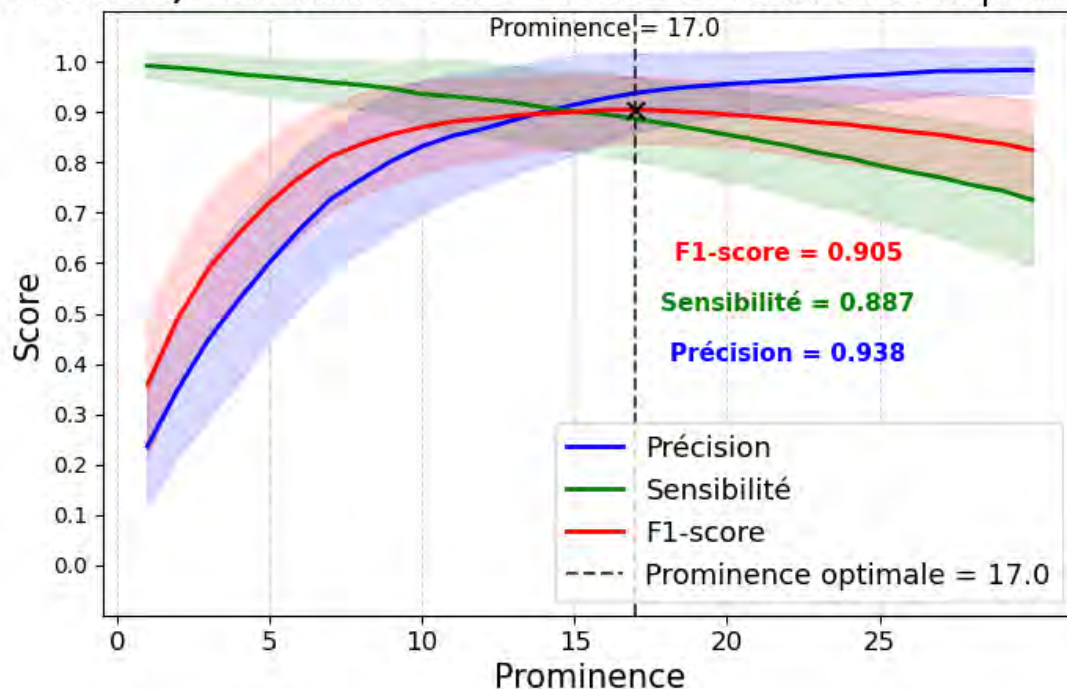


FIGURE 4.11 Sensibilité, Précision et F1-score moyens de la détection *Find Maxima* sur les 100 images d'évaluation, en fonction de la prominence utilisée.

4.3 Modèles de Deep Learning développés avec Dragonfly

Pour pallier les limites de la détection des NPs avec *Find Maxima*, l'approche proposée consiste à développer un modèle de DL capable de détecter les NPs visibles dans les images SI. Dans la même lancée, un autre modèle a été développé pour isoler le tissu dans les images BF.

4.3.1 Entraînement des modèles

L'entraînement des modèles a été réalisé grâce au logiciel Dragonfly présenté précédemment, et plus particulièrement à l'aide de l'interface *Segmentation Wizard*. Cette interface permet de charger des images et leur masque de vérité terrain pour entraîner différents modèles à des tâches de segmentation sémantique. Tous les modèles sont entraînés en utilisant les mêmes hyperparamètres sélectionnés par défaut dans Dragonfly. L'entraînement comporte 100 époques. L'algorithme d'optimisation est Adadelta. La fonction de perte utilisée est l'erreur d'entropie croisée catégorielle (*Categorical Cross Entropy*). Si la fonction de perte ne diminue pas durant 15 époques, on considère une convergence de la performance et l'entraînement prend fin prématurément. Pour rendre l'entraînement plus robuste, les données d'entraînement sont augmentées via des transformations d'images : renversements horizontal et vertical, rotations, déformations et modifications de l'échelle. Comme les différents modèles ont présenté des performances conformes à nos attentes ($F1_score > 0,905$), les hyperparamètres initiaux n'ont pas été modifiés.

Modèle pour la segmentation des NPs - SegNPs

Le modèle **SegNPs** doit permettre de détecter la position des NPs dans les images SI. Pour développer un tel outil, trois architectures différentes ont été investiguées :

- L'architecture **U-Net**, référence dans le domaine de la segmentation d'images médicales [48].
- L'architecture **TransUNet**, qui combine un encodeur s'inspirant des Transformers avec un décodeur de type U-Net [82].
- L'architecture **Attention-UNet**, qui permet d'ajouter des portes d'attention (*Attention Gates*, AGs) à l'architecture U-Net classique, permettant d'apprendre au modèle à focaliser son attention sur les zones importantes dans l'image à segmenter [56].

Un premier modèle a été développé en utilisant l'architecture U-Net uniquement et les 100 images d'entraînement du jeu de données A. Le modèle obtenu est appelé **PreTrained100_UNet**.

L'entraînement de ce modèle a ensuite été approfondi en utilisant les 300 images d'entraînement du jeu de données B. En parallèle, trois autres modèles ont été entraînés avec les mêmes 300 images, sans avoir été pré-entraînés avec les 100 images du jeu de données A. Ces trois modèles sont simplement appelés **TransUNet**, **UNet** et **AttUNet** en référence aux architectures utilisées.

Très vite, l'étude du dernier modèle **AttUNet** a été écartée, car le modèle ne prédisait que des pixels d'arrière plan, même lorsque plusieurs NPs étaient clairement visibles dans l'image.

L'évaluation des performances des trois modèles restants a permis d'identifier le modèle le plus performant et d'analyser l'influence de l'architecture ainsi que du pré-entraînement sur les performances obtenues.

Modèle pour la segmentation du tissu - **TissuDetection**

Le modèle **TissuDetection** doit permettre d'isoler le tissu du substrat dans les images BF. Deux des architectures présentées précédemment ont été utilisées pour développer cet outil :

- L'architecture **U-Net** .
- L'architecture **Attention-UNet**.

Deux modèles basés sur ces deux architectures ont été entraînés avec les 112 images d'entraînement du jeu de données C. L'évolution de leurs performances a ensuite permis d'identifier le modèle le plus performant et d'étudier l'effet de l'architecture sur les performances de segmentation du tissu.

4.3.2 Evaluation des modèles

Segmentation des NPs avec SegNPs

Les performances des modèles sont évaluées en comparant les positions des NPs connues grâce aux masques de vérité terrain avec les positions prédites par les différents modèles. La précision, la sensibilité et le F1-score sont calculés en suivant la méthodologie décrite à la section 4.1.5.

Pour commencer, le modèle **PreTrained100_Unet**, entraîné avec les 100 images d'entraînement du jeu de données A, est évalué avec les 40 images d'évaluation du jeu de données A, et les 100 images d'évaluation du jeu de données B. L'évaluation avec le jeu de données B

permet de comparer les performances du modèle avec les performances de *Find Maxima*. Les résultats sont présentés dans le Tableau 4.2.

TABLEAU 4.2 Résultat de l'évaluation du modèle PreTrained100_UNet obtenus avec les deux jeux de d'évaluation, comparé aux performances de *Find Maxima*.

	Précision	Sensibilité	F1-score	Images d'entraînement	Images d'évaluation
PreTrained100_UNet	0,995	0,914	0,953	100 images A	40 images A
PreTrained100_UNet	0,997	0,875	0,929	100 images A	100 images B
<i>Find Maxima</i> performances maximales (prominence=17)	0,938	0,887	0,905	/	100 images B

Sans surprise, le modèle performe mieux sur les images d'évaluation provenant du jeu de données sur lequel il a été entraîné (F1-score 0,953 > 0,929). Toutefois, les performances sur le second jeu de données dépassent déjà les performances maximales de *Find Maxima* pour deux des trois métriques utilisées. Seule la sensibilité doit encore être améliorée pour surpasser les performances de l'algorithme naïf.

Ensuite, le second jeu de données est utilisé pour continuer l'entraînement du modèle PreTrained100_Unet et débiter l'entraînement des deux autres modèles Unet et TransUnet. Les performances des modèles sont évalués au cours de l'entraînement en utilisant le jeu de données B. Les résultats obtenus sont présentés dans la Figure 4.12.

Avec 10 images d'entraînement seulement, on peut voir que le modèle TransUNet atteint une performance similaire, voire supérieure au modèle PreTrained100_UNet. Au contraire, le modèle UNet n'ayant pas été pré-entraîné est peu sensible et doit être exposé à un plus grand nombre d'images pour rattraper les performances des deux autres modèles.

TABLEAU 4.3 Performances des trois modèles après un entraînement sur 10 images du jeu de données B uniquement.

	Précision	Sensibilité	F1-score	Images d'entraînement	Images d'évaluation
PreTrained100_UNet	0,997	0,875	0,928	100 images A + 10 images B	100 images B
UNet	0,912	0,541	0,663	10 images B	100 images B
TransUNet	0,993	0,892	0,937	10 images B	100 images B

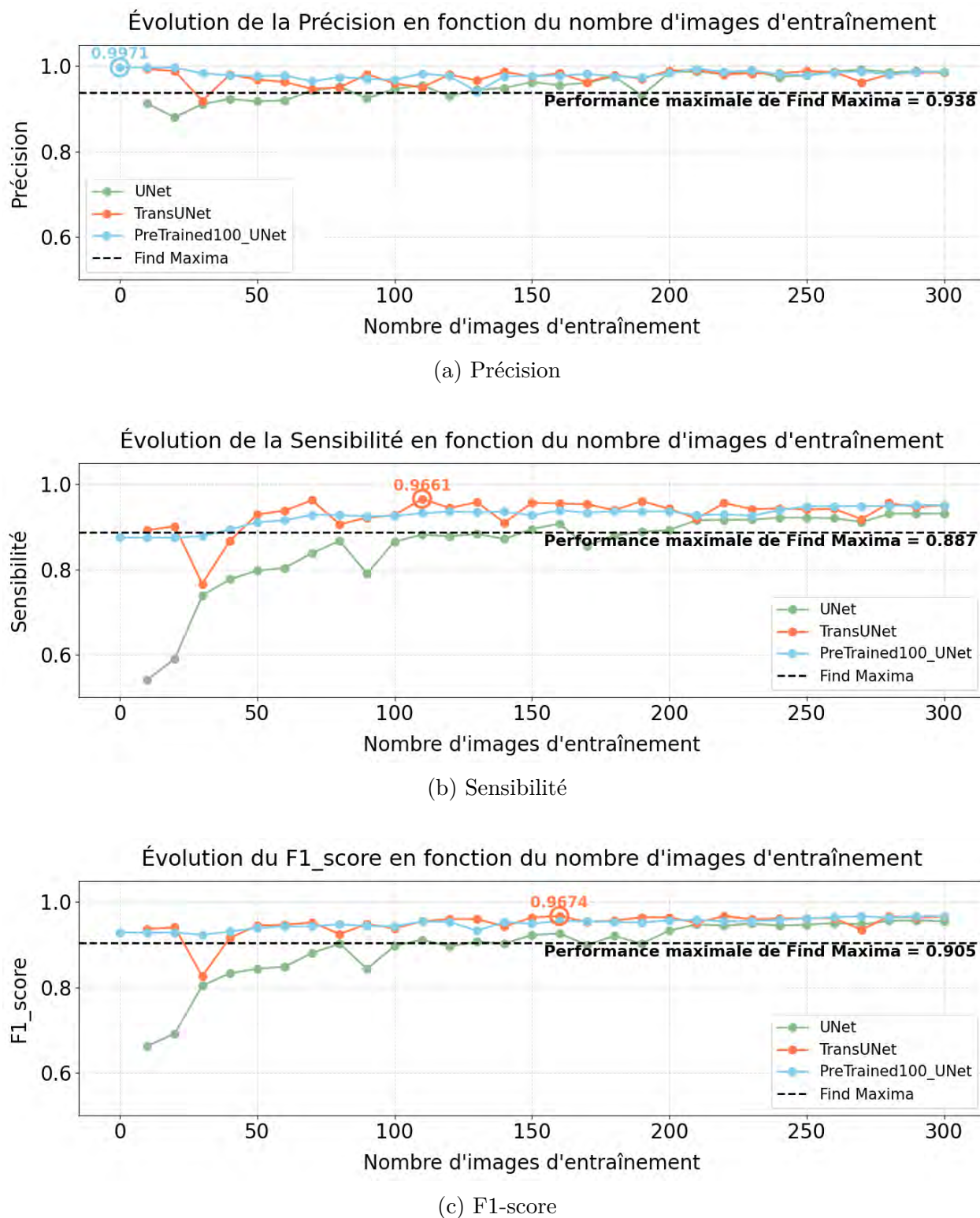


FIGURE 4.12 Évolution des métriques de performance pour les différents modèles. Les points entourés correspondent au score maximal atteint pour chaque métrique.

Après environ 200 images d'entraînement, l'architecture du modèle ne semble plus avoir une grande influence sur les performances. Les trois modèles atteignent des résultats excellents et surpassent largement *Find Maxima*.

D'après la Figure 4.12, le F1-score maximal **0,9674** est atteint avec le modèle TransUnet entraîné avec 160 images. Toutefois, ce maximum n'est pas un pic net, le F1-score maximal reste ensuite constant jusqu'à la fin de l'entraînement. Les modèles PreTrained100_Unet et TransUnet atteignent un F1-score proche du maximum tandis que le modèle UNet est légèrement moins performant avec un score de 0,9657.

TABLEAU 4.4 Performances des trois modèles à la fin de l'entraînement sur les 300 images du jeu de données B.

	Précision	Sensibilité	F1-score	Images d'entraînement	Images d'évaluation
PreTrained100_Unet	0,9872	0,9507	0,9672	100 images A + 300 images B	100 images B
UNet	0,9848	0,9312	0,9551	300 images B	100 images B
TransUnet	0,9849	0,9506	0,9657	300 images B	100 images B

Le maximum atteint avec 160 images n'étant pas très marqué, il semble judicieux de conserver comme **meilleur modèle** le **PreTrained100_Unet** entraîné avec toutes les images d'entraînement disponibles.

L'architecture TransUnet offre de très bonnes performances avec un entraînement limité, mais un entraînement plus long n'apporte pas nécessairement d'amélioration significative par rapport au modèle UNet pré-entraîné.

Bien que plus efficace qu'avec *Find Maxima*, la détection des NPs à haute densité reste difficile. En effet, certaines NPs très proches sont souvent considérées comme une seule et même NP par l'algorithme. Toutefois, dans le cadre de la détection des cas de faible expression de la protéine HER2, cette limite n'est que peu problématique.

Pour l'analyse des TMA, le modèle utilisé a été le PreTrained100_Unet, entraîné uniquement avec les images du jeu de données A. L'utilisation de ce modèle a permis de conserver la même méthodologie pour toutes les analyses, qui ont débuté avant la fin de l'entraînement des modèles. Il pourrait être pertinent de recommencer l'analyse avec le modèle le plus performant pour voir comment les résultats seraient affectés.

L'utilisation du modèle SegNPs, a permis de segmenter les images SI pour obtenir les positions des NPs, comme le montre la Figure 4.13.

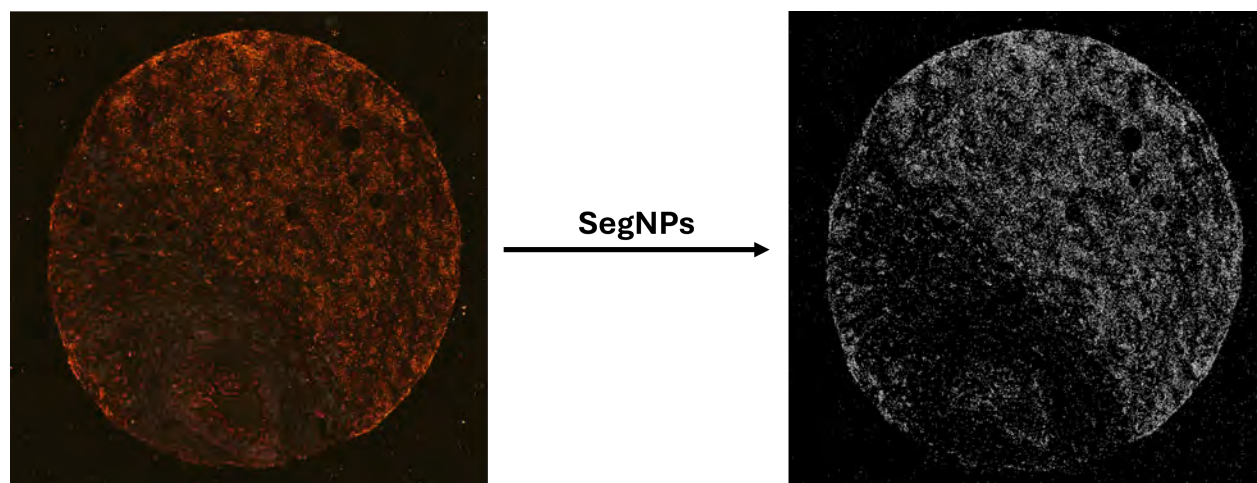


FIGURE 4.13 Détection des NPs dans l'image SI grâce au modèle SegNPs. Lame 16 - Carotte Sein F3.

Détection du tissu avec TissuDetection

La performance des modèles a été évaluée en utilisant les 56 images du jeu de données C destinées à l'évaluation.

Les modèles ont été évalués en calculant les scores Dice et IoU présentés dans la section 4.1.5. Comme discuté précédemment, il est particulièrement compliqué de déterminer au pixel près s'il s'agit du tissu ou du substrat. En effet, le tissu tranché fin peut être relativement transparent, même avec la coloration H&E. La limite entre tissu et substrat est rarement définie très nettement. On considère donc que le modèle performe bien s'il parvient à prédire la forme générale du tissu.

La Figure 4.14 montre l'évolution du score Dice au cours de l'entraînement, tandis que la Figure 4.15 montre l'évolution du score IoU.

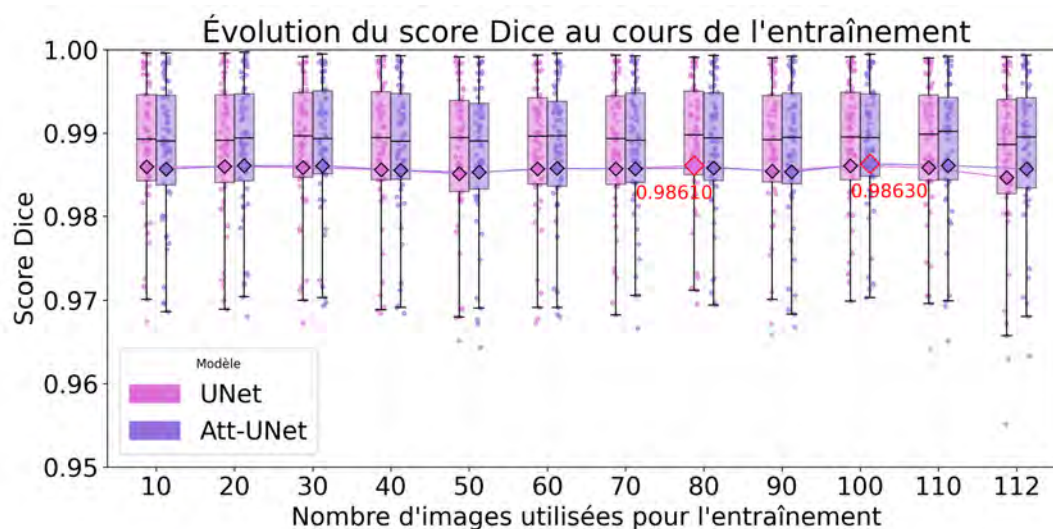


FIGURE 4.14 Évolution du score **Dice** au cours de l'entraînement pour les deux modèles.

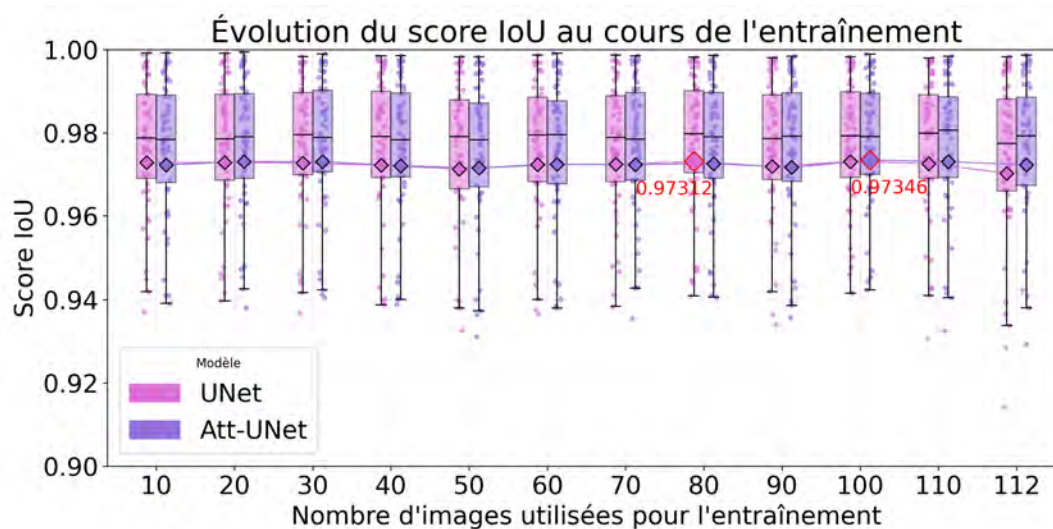


FIGURE 4.15 Évolution du score **IoU** au cours de l'entraînement pour les deux modèles.

Avec uniquement 10 images d’entraînement, le modèle **TissuDetection** performe déjà particulièrement bien avec un score Dice supérieur à 0,98, et ce pour les deux architectures utilisées. Le score Dice maximal 0,98630 est atteint pour 100 parcelles d’entraînement, avec le modèle **Attention-UNet**. Le score IoU moyen est maximal dans les mêmes conditions. L’architecture **UNet**, bien que légèrement moins performante, donne des résultats particulièrement proches de l’autre architecture. Il semble donc que l’architecture choisie pour la segmentation du tissu ait peu d’importance.

Les deux architectures ont montré des performances largement suffisantes à toutes les étapes de l’entraînement et pourraient donc être utilisées pour réaliser la tâche de segmentation du tissu lors de l’analyse des TMA. Cependant, on conserve uniquement le modèle **Att-UNet_TissuDetection_T100** qui est celui qui a permis d’observer le score Dice le plus élevé. Dans la suite de ce mémoire, ce modèle sera simplement appelé **TissuDetection**.

Ce modèle sera utilisé lors de l’analyse des TMA pour obtenir un masque donnant la position du tissu à partir de l’image BF, comme le montre la Figure 4.16.

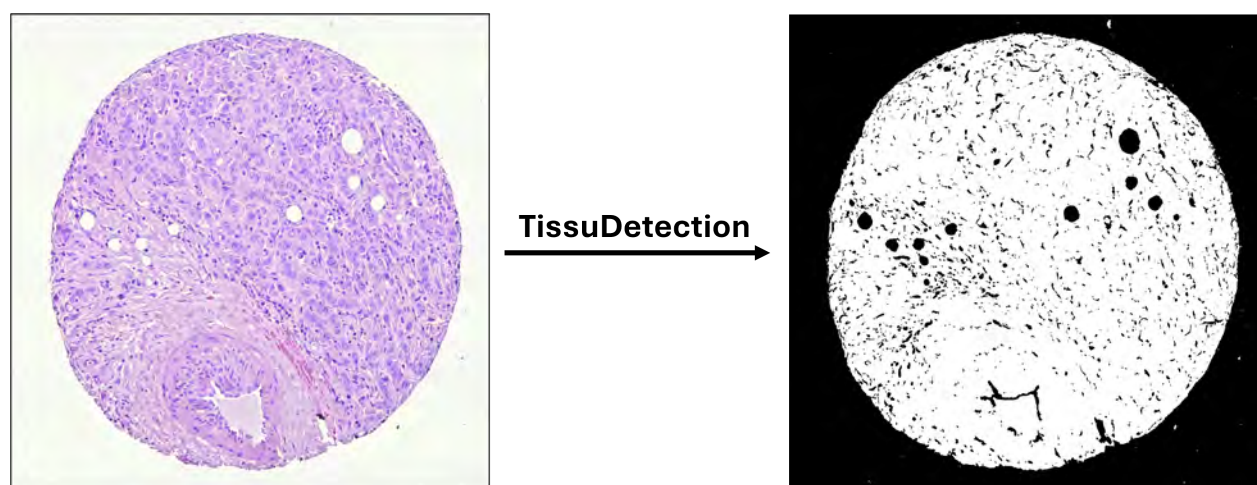


FIGURE 4.16 Génération du masque de tissu grâce au modèle **TissuDetection** appliqué sur l’image BF. Lame 16 - Carotte Sein F3.

4.4 Segmentation et classification des noyaux cellulaires

Comme mentionné dans le Chapitre 2, la segmentation et la classification des noyaux sont des problématiques récurrentes en pathologie digitale. De nombreuses solutions sont développées et commercialisées pour permettre aux pathologistes de travailler avec des outils innovants

basés sur l’usage de l’IA. C’est par exemple le cas du logiciel Aperio Clinical Solution de Leica Biosystems ou encore de Visiopharm. Il existe aussi des modèles développés par différents groupes de recherche utilisables gratuitement dans un cadre académique. Il n’est donc pas indispensable de développer de nouveaux modèles qui auraient nécessité de nouveaux jeux de données basés sur l’expertise d’un pathologiste. Les différents modèles présentés dans la Section à venir ont été choisis par le groupe pour leur performance dans les tâches à réaliser et leur facilité d’implémentation dans la méthodologie complète.

4.4.1 Segmentation des noyaux

La segmentation des noyaux peut facilement être réalisée en utilisant le plugin Stardist disponible dans ImageJ ou encore le plugin CellPose-SAM disponible dans Dragonfly. La Figure 4.17 présente un exemple de segmentation des noyaux présent dans une région de tissu. Pour les deux modèles, la segmentation est bien réalisée sans avoir besoin de modifier le moindre paramètre. Les noyaux de toutes formes sont bien segmentés en instances individuelles, même lorsque les noyaux se touchent. Stardist semble performer légèrement moins bien que CellPose-SAM puisque l’on peut remarquer que certains noyaux se chevauchent avec la segmentation de Stardist.

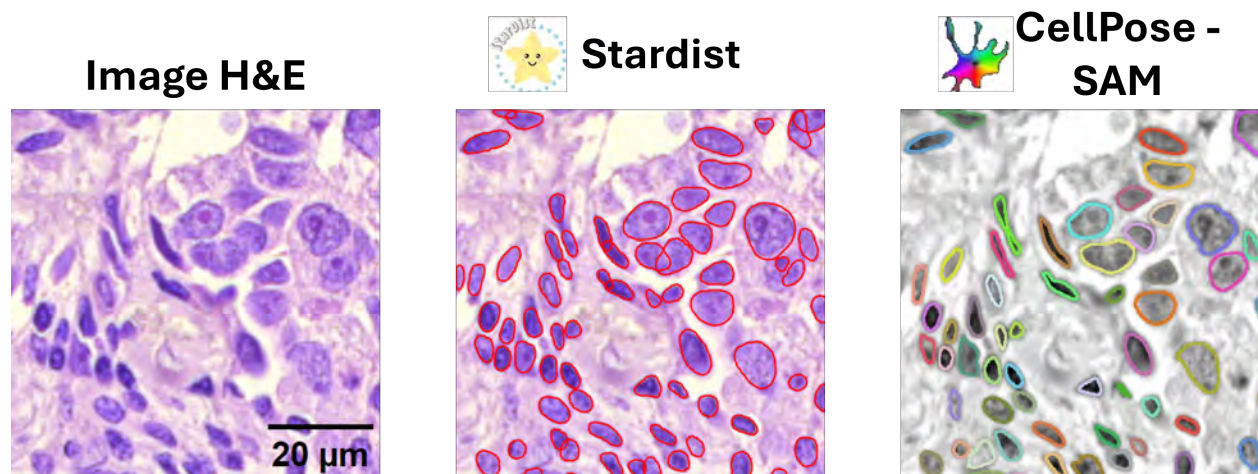


FIGURE 4.17 Résultat de la segmentation des noyaux par les modèles Stardist et CellPose-SAM.

Pour segmenter l’image avec CellPose-SAM dans Dragonfly, l’image doit être chargée avec ses trois canaux RGB. Dans la Figure 4.17, la segmentation apparaît sur une image en niveaux de gris car Dragonfly n’est pas capable de gérer l’affichage des images RGB.

Stardist et CellPose-SAM, bien qu’ayant montré des résultats préliminaires intéressants ne

permettent pas de classer les noyaux en fonction de leur type cellulaire, contrairement à l'outil qui sera présenté à la section suivante. Par conséquent, leur performance n'a pas fait l'objet d'une étude approfondie.

4.4.2 Classification des noyaux

Les cellules qui composent le tissu peuvent appartenir à divers types cellulaires. Pour l'évaluation du statut HER2, on rappelle que les directives de l'ASCO/CAP préconisent de se concentrer sur les cellules tumorales, à savoir les **cellules épithéliales néoplasiques** dans le cadre du cancer du sein. Les autres types de cellules devraient donc être identifiés et écartés du reste de l'analyse (ou du moins, ne pas être utilisés pour quantifier l'expression de HER2).

Bien que CellPose et Stardist aient déjà été utilisés dans la littérature pour réaliser des tâches de classification cellulaire, leur utilisation dans ce contexte n'est pas possible sans ré-entraînement préalable.

C'est donc le modèle CellViT et en particulier sa version CellViT++ qui a été choisie pour intégrer l'analyse des TMA.

Le modèle CellViT++ est appliqué à une image BF, qui a préalablement été convertie en image pyramidale grâce à un script Python. Les images pyramidales sont des images particulièrement courantes en pathologie digitale dont la structure multi-couches permet de stocker une même image à des magnifications différentes dans un même fichier. L'inférence du modèle CellViT++ sur l'image pyramidale permet de générer deux fichiers .json qui contiennent les coordonnées des contours des noyaux détectés, et une valeur correspondant au type cellulaire prédit parmi cinq classes possibles : cellule épithéliale néoplasique, cellule inflammatoire, cellule conjonctive, cellule épithéliale non néoplasique et cellule morte. La Figure 4.18 montre la région du tissu précédemment segmentée par Stardist et CellPose, maintenant segmentée par CellViT++ qui classe aussi les noyaux par classe cellulaire.

4.4.3 Évaluation de la classification des noyaux

L'évaluation de la classification des noyaux n'a pas pu être réalisée en utilisant des métriques quantitatives comme pour les modèles **SegNPs** et **TissuDetection**. Une telle étude aurait requis des masques de vérité terrain annotés par un pathologiste expérimenté pour comparer le type cellulaire prédit avec le type défini par le pathologiste. La génération de tels masques représente un travail colossal pour un pathologiste et n'a pas pu être réalisé. L'évaluation des performances de CellViT++ sur des jeux de données libres d'accès, déjà annotés par des pathologistes, pourrait être intéressante, mais ne garantirai pas des performances équivalentes

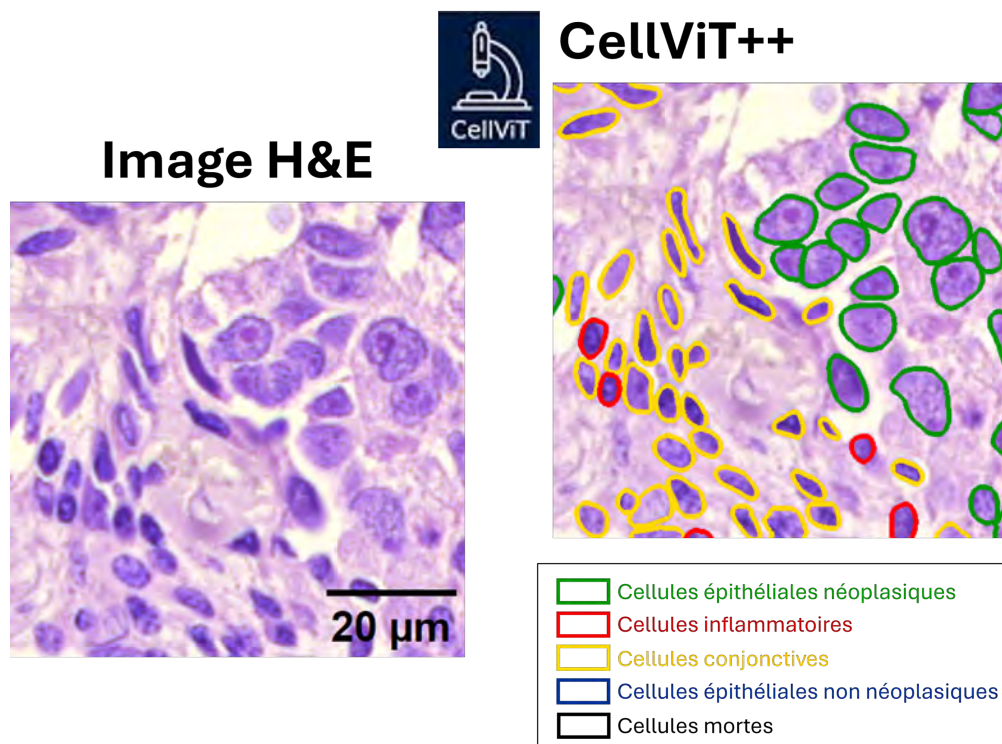


FIGURE 4.18 Résultat de la segmentation des noyaux et de leur classification selon leur type cellulaire par le modèle CellViT++.

sur nos jeux de données acquis expérimentalement. À terme, une évaluation quantitative de CellViT++ en collaboration avec des pathologistes semble inévitable.

Toutefois, une évaluation qualitative des résultats de CellViT++ a été menée avec l'expertise d'une pathologiste du CRCHUM. Les carottes de sein de la lame 16 annotées avec les prédictions de CellViT++ ont été inspectées en présence de la pathologiste pour établir des remarques qualitatives qui ont permis de conclure que les résultats de détection et de classification, bien qu'imparfaits sont satisfaisants et peuvent être utilisés pour la suite de l'analyse.

Le modèle CellViT++ oublie peu de noyaux. Parfois, certains noyaux sont scindés en deux. Outre ce défaut, la détection des noyaux est considérée comme très satisfaisante.

En ce qui concerne la classification par type cellulaire, CellViT++ performe bien à l'échelle globale puisque les annotations par type cellulaire permettent de révéler l'architecture du tissu. La Figure 4.19 montre la corrélation de la classification avec le marquage IHC : on retrouve des cellules épithéliales néoplasiques là où les membranes sont colorées en brun, région qui correspond bien aux cellules tumorales.

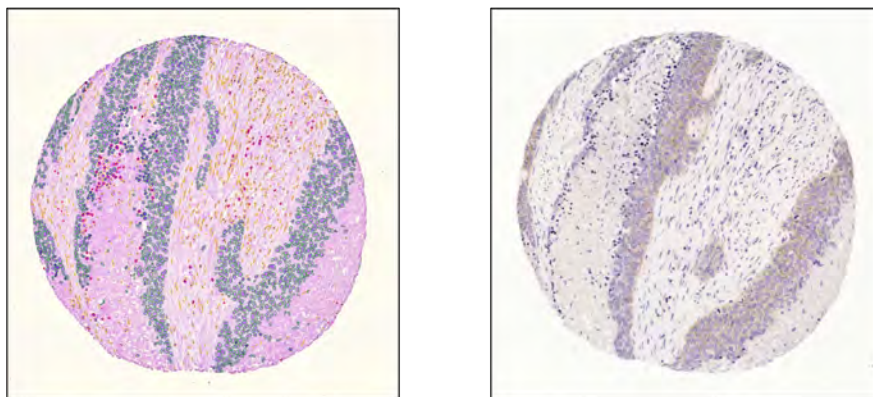
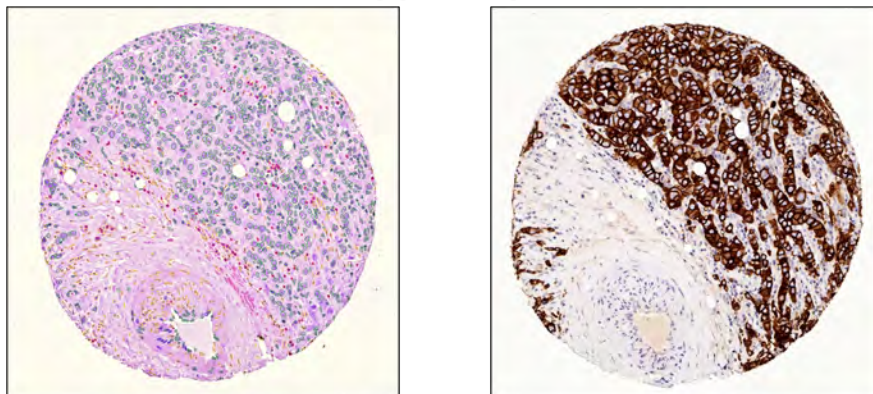
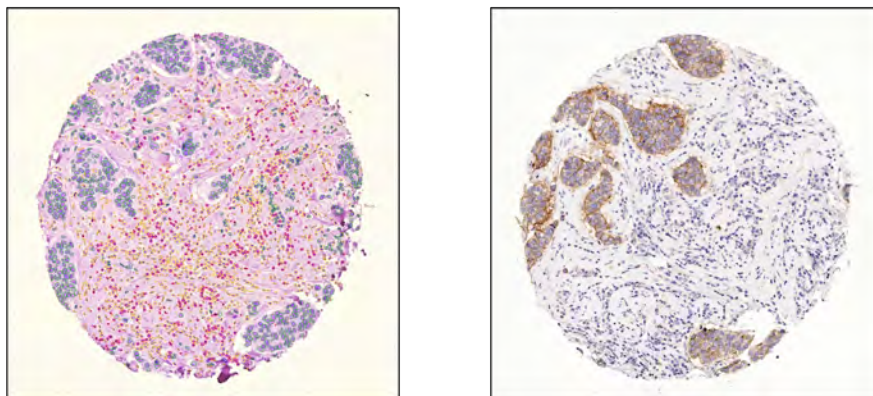
Ovaire – A3**Sein – F3****Sein – F6**

FIGURE 4.19 Corrélation entre la tumeur surexprimant HER2 visible en IHC (images de droite, coloration brune) et les noyaux des cellules épithéliales néoplasiques prédits par Cell-ViT++ (images de gauche, noyaux en vert).

Cependant, à l'échelle individuelle, les erreurs sont plus fréquentes. La classification en cinq types cellulaires ne permet pas de couvrir toute la diversité des cellules observables dans un tissu. Il n'existe par exemple aucune classe pour les cellules endothéliales qui sont souvent considérées à tort comme des cellules épithéliales néoplasiques. Les débris cellulaires issus de cellules mortes sont souvent considérés comme des noyaux de cellules inflammatoires, à cause de leur petite taille et de leur hyperchromasie (intensité de coloration anormalement importante dans les noyaux dû à un excès de matériel génétique). Parfois, l'architecture du tissu n'est pas suffisamment prise en compte par le modèle qui place un noyau d'un certain type cellulaire au milieu d'une structure dans laquelle ce type cellulaire ne devrait pas se trouver. Certains noyaux allongés devraient être classés comme conjonctif mais sont associées à la classe épithéliale néoplasique. Il en est de même pour certains noyaux petits et ronds qui correspondent à des noyaux de cellules inflammatoires mais prennent la classe épithéliale néoplasique. De manière générale, il semble que la classe épithéliale néoplasique soit surestimée.

Ces erreurs de classification pourraient être corrigées en appliquant des filtres (par exemple pour s'assurer qu'un noyau allongé ne soit pas considéré comme épithélial néoplasique, ou encore pour s'assurer que les noyaux épithéliaux néoplasiques aient une taille supérieure à un certain seuil). Cependant, cette proposition présente aussi ses limites. En effet, il est compliqué de discriminer une cellule uniquement en fonction de sa morphologie. Le noyau d'un même type cellulaire peut prendre des formes très différentes, d'une part à cause de l'hétérogénéité intrinsèque des noyaux, et d'autre part à cause de la découpe au microtome des blocs de TMA qui révèle les noyaux à des hauteurs différentes. La classification d'un noyau en fonction de son type cellulaire n'est pas une tâche facile que le pathologiste réalise uniquement en regardant la morphologie du noyau en question, mais en considérant ce noyau au sein de l'architecture dans laquelle il s'inscrit.

Malgré ces erreurs identifiées avec l'aide de la pathologiste, la segmentation et la classification prédites par CellViT++ restent particulièrement pertinentes et intéressantes et seront utilisées pour la suite de l'analyse des TMA.

CHAPITRE 5 ANALYSE DES MATRICES DE TISSUS

Grâce aux outils de segmentation présentés au Chapitre précédent, il est possible d’extraire les informations contenues dans les images issues des deux modalités. Il devient alors possible d’analyser les TMA à l’aide de métriques plus complexes qu’un simple décompte, en tenant compte de la structure biologique à laquelle chaque NP est fixée.

5.1 Méthodologie sur TMA

5.1.1 Flux de travail

Les TMA présentés dans la section 3.1.2 et produits par Jennyfer Zapata Farfan ont été analysés grâce au flux de travail suivant.

Étape 1 : Imagerie

Pour commencer, chaque lame de TMA est imagée en suivant la méthodologie présentée dans la section 3.2.1. Les 28 carottes sont acquises dans les deux modalités d’illumination (SI et BF). Ensuite, les images sont alignées et les pixels vides sont éventuellement remplis. Une fois que l’on dispose de paires d’images alignées de 800 μm par 800 μm , le processus d’analyse peut commencer.

Étape 2 : Extraction des objets d’intérêt

Les images d’une même lame sont chargées dans Dragonfly sous la forme de deux piles d’images, une par modalité d’illumination.

- Le modèle *SegNPs* est appliqué à la pile SI pour détecter les NPs. La liste des NPs est exportée sous la forme d’un tableur `.csv` contenant la liste des NPs détectées avec leur centre de masse X, Y, Z (qui permet de savoir de quelle carotte il s’agit), et leur diamètre circulaire équivalent.
- Le modèle *TissuDetection* est appliqué à la pile BF pour détecter le tissu. Un masque par carotte est exporté sous la forme d’une image `.tif` en niveaux de gris.

CellViT++ est appliqué à la pile d’images BF converties en images pyramidales par un script Python. Pour chaque image traitée, deux fichiers `.json` sont générés. Ils contiennent pour chaque noyau, les positions spatiales des contours détectés, ainsi qu’une valeur de 1 à 5 qui correspond au type cellulaire prédit par le modèle.

Comme la contre-coloration H&E des images BF ne donne aucune information sur la position exacte des membranes, la sortie CellViT++ est utilisée pour extrapoler une zone autour de chaque noyau qui correspondrait au cytoplasme de chaque cellule. La méthode de partage des eaux (*watershed*) est utilisée pour définir des frontières entre les différentes cellules (bibliothèque scikit-image). Un critère de limite d'expansion est appliqué pour éviter que le cytoplasme des cellules ne s'étende trop loin des noyaux, ce qui ne serait pas réaliste biologiquement. Le masque de tissu est aussi utilisé pour restreindre l'expansion des cytoplasmes à l'intérieur du tissu détecté, et ainsi éviter qu'un cytoplasme ne dépasse sur le substrat.

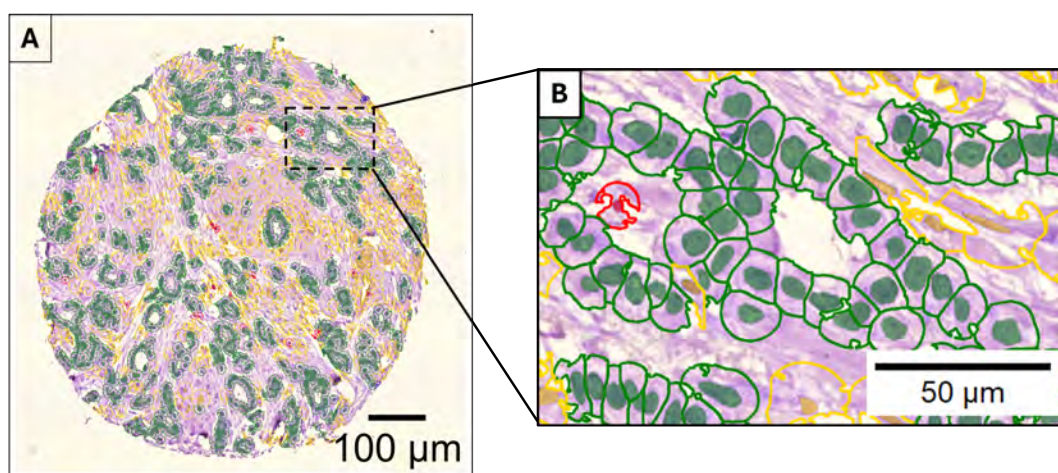


FIGURE 5.1 Carotte imagée en champ clair avec la superposition des noyaux détectés par CellViT++, et l'expansion des cytoplasmes extrapolée par watershed. En vert, les cellules épithéliales néoplasiques, en rouge une cellule inflammatoire, et en jaune les cellules conjonctives dont l'épaisseur du cytoplasme ne peut pas dépasser respectivement 5 μm , 5 μm et 10 μm . Les cytoplasmes sont confinés dans le tissu. Lame 16 - Carotte de Sein F2.

Étape 3 : Classification des NPs

Chaque NP détectée est classée en fonction de deux critères :

- Le **diamètre circulaire équivalent** d'une NP permet de déterminer s'il s'agit d'un agrégat ou non. Les 4 596 958 NPs détectées dans les 7 lames analysées ont permis de tracer l'histogramme de la Figure 5.2 montrant la distribution du diamètre des NPs détectées. Le mode du diamètre des NPs détectées est de 0,686 μm , ce qui est proche de la limite de résolution pour les conditions d'imagerie utilisées. L'écart type des diamètres est de 0,140 μm . Un seuil correspondant au mode augmenté de deux fois l'écart type, soit environ 1 μm , est utilisé pour discriminer les agrégats. Les NPs dont le diamètre est supérieur à 1 μm seront considérées comme des agrégats et écartées du reste de l'analyse. Ce seuil peut être ajusté en fonction de la tolérance aux agrégats souhaitée pour l'analyse.

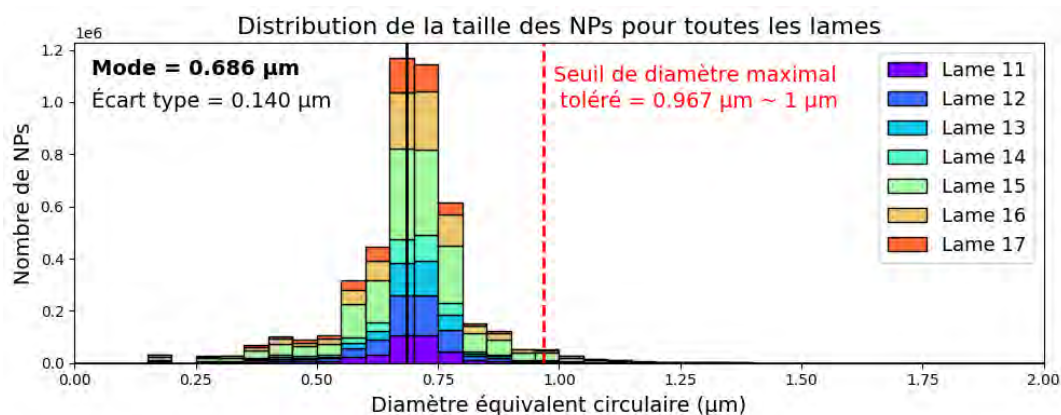


FIGURE 5.2 Histogramme montrant la distribution du diamètre des NPs détectées dans les 7 lames analysées.

- La **position** d'une NP permet de déterminer sur quelle structure biologique elle s'est fixée. Plusieurs échelles peuvent être considérées :

- Une NP peut se trouver sur le substrat ou sur le tissu.
- Une NP sur le tissu peut se trouver sur une cellule ou à l'extérieur d'une cellule.
- Une NP sur une cellule peut se trouver sur le noyau de la cellule ou sur son cytoplasme, comme le montre la Figure 5.3.

Il est possible de décliner les catégories cellulaires avec les cinq classes prédites par Cell-ViT++, à savoir cellules épithéliales néoplasiques, cellules épithéliales non néoplasiques, cellules conjonctives, cellules inflammatoires et cellules mortes.

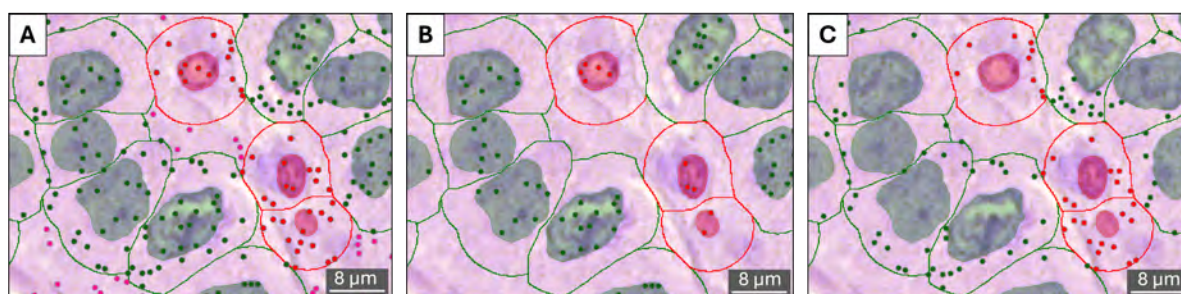


FIGURE 5.3 Positions des NPs détectées sur les cellules. A : les NPs fixées respectivement sur des cellules épithéliales néoplasiques, inflammatoires ou hors des cellules sont en vert, rouge et rose. B : NPs sur les noyaux des cellules uniquement. C : NPs sur le cytoplasme uniquement.

5.1.2 Métriques générées

Les NPs classées sont ensuite utilisées pour calculer des métriques qui permettent de générer des statistiques pour chaque lame. Les métriques doivent permettre d'évaluer quantitativement différents paramètres pertinents lors de l'évaluation d'une carotte marquée par IP. L'objectif premier est bien sûr de développer une métrique qui permette de **mesurer l'expression protéique** de l'échantillon. Cependant, comme la technique IP est encore en développement, d'autres métriques sont indispensables et permettent de **mesurer la qualité du marquage** qui doit respecter certains critères. Pour illustrer cette section, des résultats de l'analyse de la lame 14 ont été utilisés.

Qualité du marquage

Certaines métriques permettent d'évaluer la qualité du marquage IP. Bien qu'elles ne renseignent pas sur le niveau d'expression protéique des échantillons, elles s'assurent que le marquage respecte des critères de qualité. Un marquage réussi est un marquage pour lequel on observe un faible taux d'agrégation, et un faible attachement non spécifique.

i. Taux d'agrégation

Bien que la limite de résolution optique ne permette pas d'affirmer que chaque point brillant corresponde à une NP individuelle, la présence d'agrégats est symptomatique d'une faible stabilité des NPs fonctionnalisées. On considère donc qu'un marquage est réussi si les NPs fixées sur le tissu sont visibles comme des points brillants individuels. L'utilisation du calixarène pour la fonctionnalisation assure une haute stabilité des NPs qui ne devraient pas s'agréger au-delà d'un seuil établi à 5 %.

Le nombre d'agrégats permet de calculer le **taux d'agrégation** observé dans une carotte :

$$\text{Taux d'agrégation (\%)} = \frac{N_{\text{agrégats}}}{N_{\text{NPs détectées, toutes tailles confondues}}} \times 100 \quad (5.1)$$

Pour chaque lame analysée, un graphe permet de visualiser le taux d'agrégation pour chaque carotte. La valeur présente au-dessus de chaque point indique le score IHC attribué à la carotte.

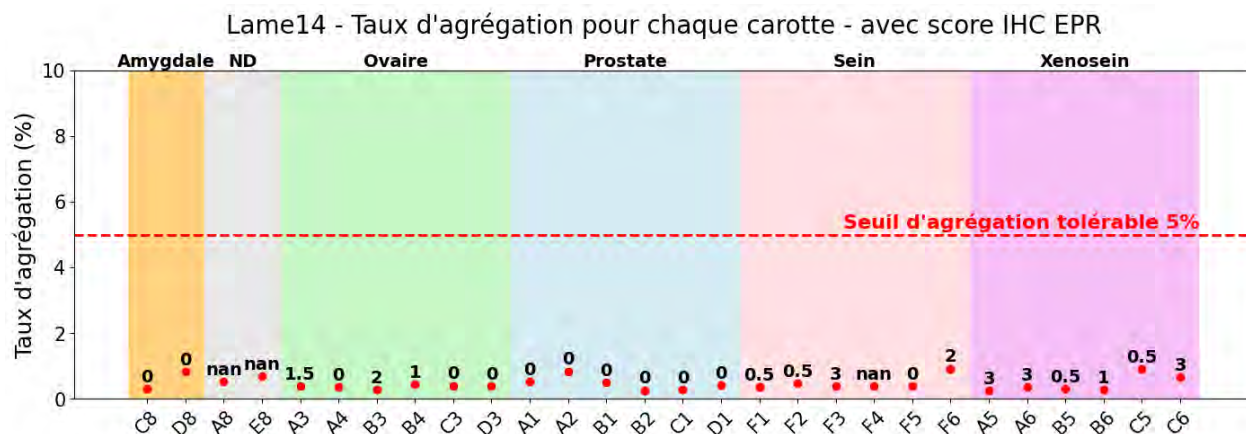


FIGURE 5.4 Graphe montrant le taux d'agrégation pour toutes les carottes d'une lame.

ii. Attachement non spécifique

Un marquage est réussi si l'attachement non spécifique est faible, c'est-à-dire si les NPs se fixent peu sur les structures qui n'expriment pas la protéine cible. Dans les TMA, trois zones sur lesquelles les NPs ne doivent pas se fixer ont été déterminées :

- Le substrat,
- La zone non cancéreuse, qui regroupe tout le tissu qui n'est pas tumoral.
- Les noyaux des cellules, puisque la protéine HER2 est transmembranaire.

Bien qu'un certain niveau d'attachement non spécifique soit inévitable, cet attachement doit rester le plus faible possible afin de pouvoir affirmer avec confiance que l'attachement observé sur les structures exprimant la protéine cible est bien la conséquence de la présence de cette protéine.

La **densité de NPs sur le substrat** permet de quantifier l'attachement des NPs en dehors du tissu, sur la lame de microscopie :

$$\text{Densité de NPs sur le substrat (en NPs / } 100 \mu\text{m}^2) = \frac{N_{\text{NPs sur le substrat}}}{S_{\text{substrat}}} \times 100 \quad (5.2)$$

La densité de NPs sur le substrat est en général bien maîtrisée, et dépasse rarement 1 NP / $100 \mu\text{m}^2$. Si cette densité est trop élevée, il est important de s'interroger sur les conditions d'incubation utilisées et d'ajuster certains paramètres, tels que la concentration en NPs ou la durée d'incubation. En effet, ces paramètres peuvent favoriser l'attachement global, augmentant ainsi l'attachement non spécifique observé sur le substrat.

Pour chaque carotte analysée, une image montrant les NPs détectées sur le tissu et sur le substrat est générée.

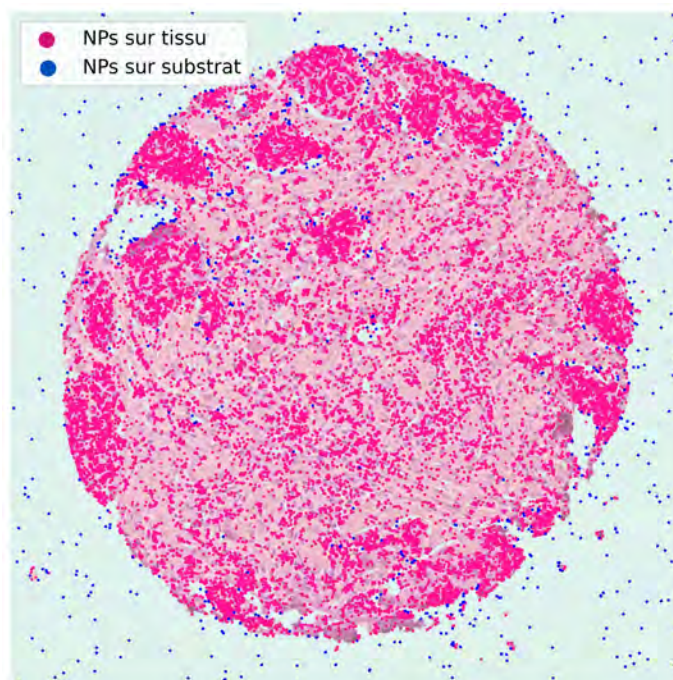


FIGURE 5.5 Image montrant en rose les NPs sur le tissu et en bleu les NPs sur le substrat.

Un graphe permet ensuite de visualiser la densité de NPs sur le substrat pour toutes les carottes d'une lame.

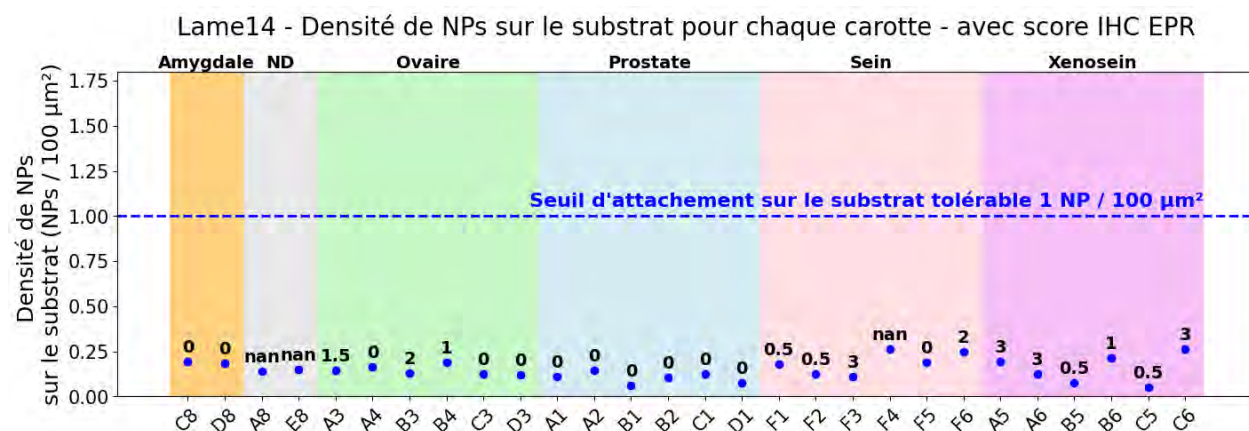
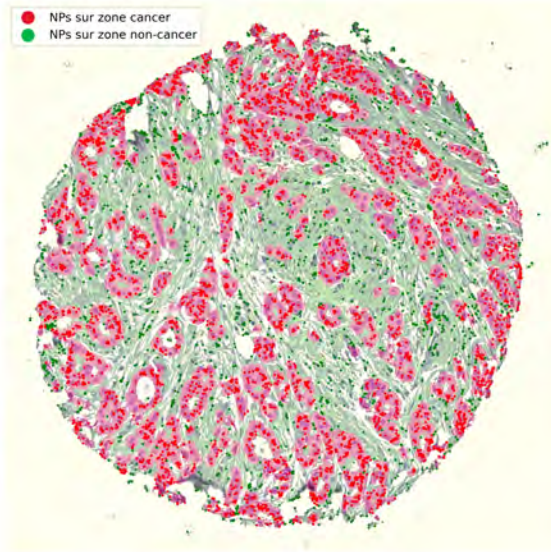


FIGURE 5.6 Graphe montrant la densité de NPs sur le substrat pour toutes les carottes d'une lame.

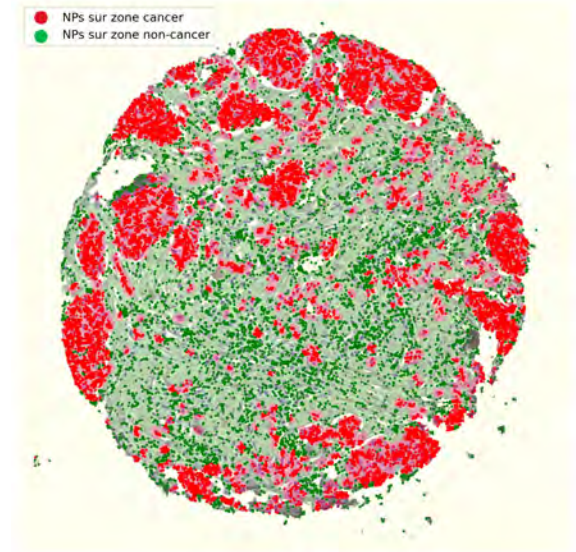
La **densité de NPs sur la zone non cancéreuse** permet de quantifier l'attachement des NPs sur les régions du tissu qui n'expriment pas la protéine HER2. La zone non cancéreuse est composée de tout le tissu qui n'est pas tumoral, c'est-à-dire qui n'est pas classé comme cellule épithéliale néoplasique. Cela englobe tous les autres types cellulaires mais aussi le tissu conjonctif qui se trouve entre les différentes cellules. La densité calculée dépend directement de la qualité de la classification cellulaire réalisée par CellViT++.

$$\text{Densité de NPs sur la zone non cancéreuse (en NPs / } 100 \mu\text{m}^2) = \frac{N_{\text{NPs sur la zone non cancéreuse}}}{S_{\text{zone cancéreuse}}} \times 100 \quad (5.3)$$

Pour chaque carotte analysée, une image montrant les NPs détectées sur la zone non cancéreuse et sur la zone cancéreuse est générée.



(a) Carotte Sein F2 - densité de NPs sur la zone non cancéreuse = 0,85 NP / 100 μm^2



(b) Carotte Sein F6 - densité de NPs sur la zone non cancéreuse = 3,08 NPs / 100 μm^2

FIGURE 5.7 Images montrant en rouge les NPs sur la zone cancéreuse et en vert les NPs sur la zone non cancéreuse. La forte densité parfois mesurée dans une zone cancéreuse peut être causée par un attachement non spécifique anormalement haut, ou par une erreur de classification de CellViT++. Il est important de garder un regard critique sur les résultats obtenus.

Un graphe permet ensuite de visualiser la densité de NPs sur la zone non cancéreuse pour toutes les carottes d'une lame. La densité sur la zone non cancéreuse est très souvent supérieure au seuil de 1 NP / 100 μm^2 qui semblait raisonnable pour étudier l'attachement non spécifique sur le substrat. Les NPs s'attachent plus sur le tissu que sur le substrat, même

si ces deux régions d'intérêt n'expriment pas la protéine. L'optimisation de la méthode IP devrait avoir pour objectif de limiter au maximum l'attachement non spécifique sur les zones non cancéreuses.

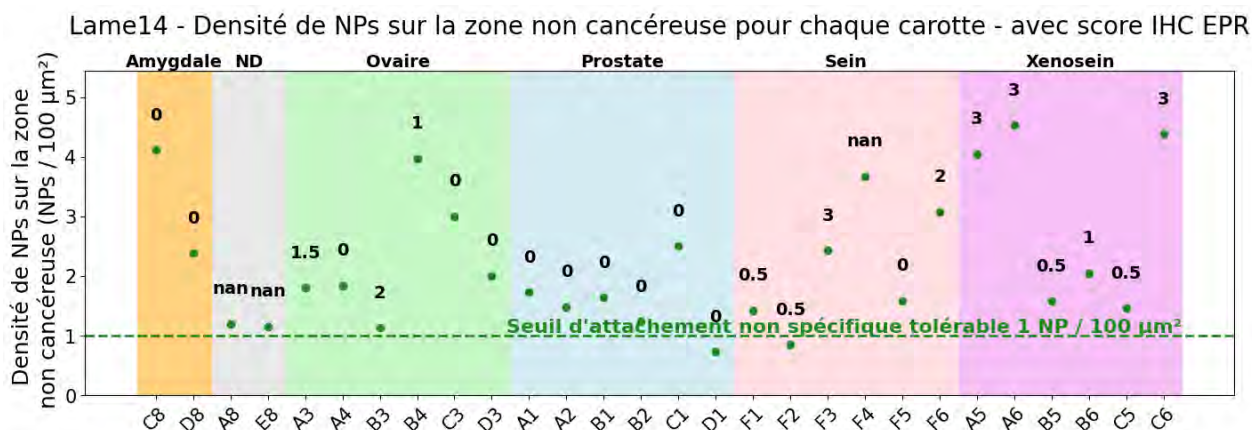


FIGURE 5.8 Graphe montrant la densité de NPs sur la zone non cancéreuse pour toutes les carottes d'une lame.

La **densité de NPs sur les noyaux** permet de quantifier l'attachement non spécifique sur les noyaux :

$$\text{Densité de NPs sur les noyaux (en NPs / 100 } \mu\text{m}^2) = \frac{N_{\text{NPs sur les noyaux}}}{S_{\text{noyaux}}} \times 100 \quad (5.4)$$

Un graphe permet de visualiser la densité de NPs sur les noyaux pour toutes les carottes d'une lame.

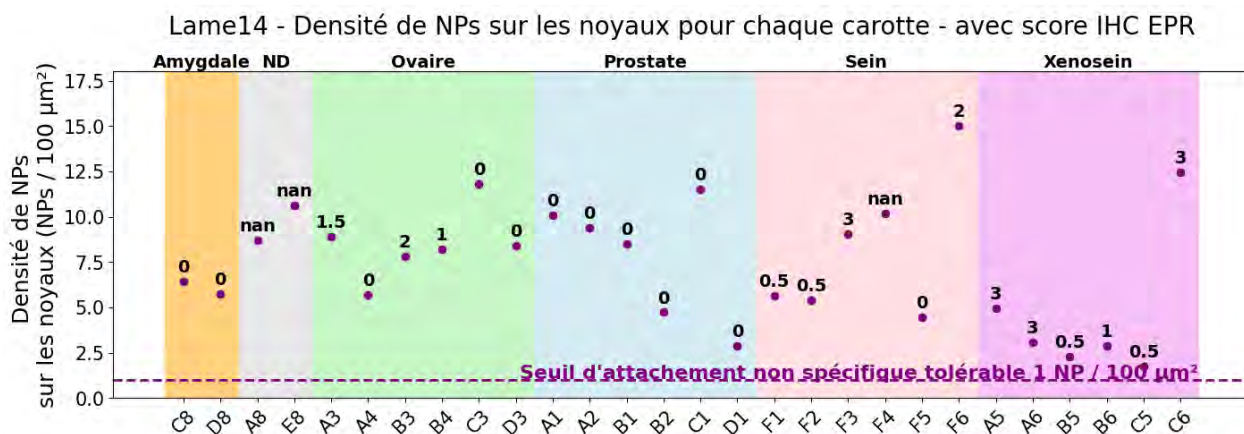


FIGURE 5.9 Graphe montrant la densité de NPs sur les noyaux pour toutes les carottes d'une lame.

La densité de NPs sur les noyaux est particulièrement importante. Elle dépasse systématiquement le seuil d'attachement non spécifique tolérable qui semblait raisonnable pour l'attachement non spécifique sur le substrat.

Cette métrique permet de révéler une affinité particulière des NPs pour les noyaux des cellules. Cette affinité non souhaitée reste pour le moment inexpliquée et constitue un axe majeur d'amélioration de la méthode IP.

L' **indice d'affinité nucléaire** permet de quantifier cette affinité particulière que semblent avoir les NPs pour les noyaux :

$$\text{Indice d'affinité nucléaire} = \frac{\text{Densité de NPs sur les noyaux (tous types cellulaires confondus)}}{\text{Densité de NPs sur le cytoplasme (tous types cellulaires confondus)}}$$

Un graphe permet de visualiser cet indice pour toutes les carottes d'une lame.

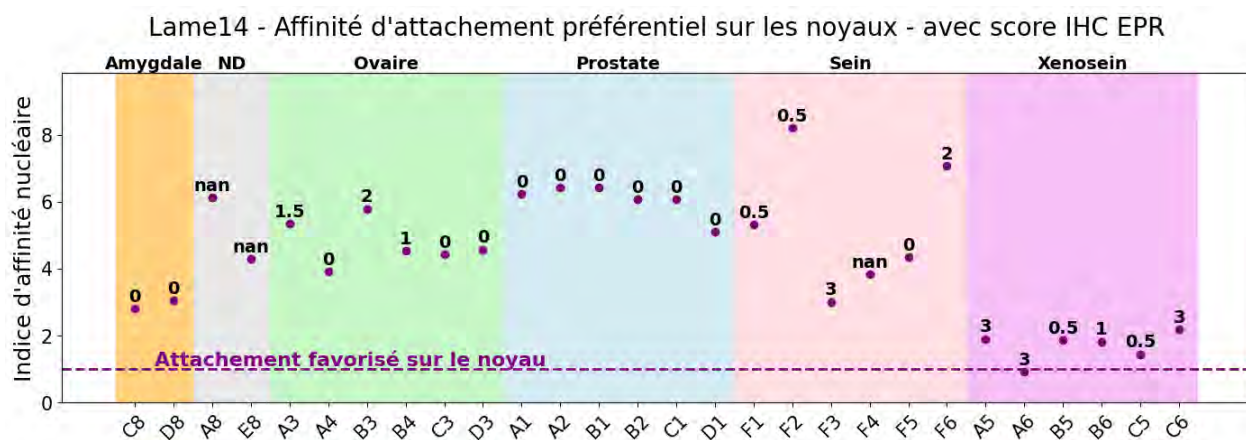


FIGURE 5.10 Graphe montrant l'indice d'affinité nucléaire calculé pour toutes les carottes d'une lame.

L'indice d'affinité nucléaire permet de mesurer à quel point la densité sur le noyau est plus importante que la densité dans le cytoplasme. Pour la lame 14, l'indice d'affinité nucléaire est en moyenne de 4,4 pour toutes les carottes, ce qui signifie que l'attachement dans le noyau est plus de 4 fois plus important dans les noyaux que dans le reste des cellules.

Expression protéique de chaque carotte

Certaines métriques doivent permettre d'évaluer l'expression protéique de chaque carotte. Pour cela, on s'intéresse à l'attachement des NPs sur différentes zones d'intérêt. Cet attachement peut être quantifié à l'aide d'une mesure de densité, ou de nombre par cellule.

i. Densités de NPs

L'évolution de la méthodologie et des outils de segmentation utilisés a permis de mesurer la densité des NPs sur trois régions d'intérêt :

- sur le **tissu** :

$$\text{Densité de NPs sur le tissu (NPs / 100 } \mu\text{m}^2) = \frac{N_{\text{NPs sur le tissu}}}{A_{\text{tissu}}} \quad (5.5)$$

Cependant, le tissu comporte aussi bien des zones cancéreuses composées de cellules épithéliales néoplasiques pouvant surexprimer la protéine HER2 que des zones non cancéreuses composées de cellules épithéliales saines, ou de toute autre composante présente normalement dans un tissu et ne surexprimant pas la protéine HER2. La densité sur le tissu est donc tirée vers le bas par la densité de NPs sur les zones saines pour lesquelles l'attachement observé ne devrait correspondre qu'à de l'attachement non spécifique.

- sur la **zone cancéreuse**, composée de toutes les cellules épithéliales néoplasiques détectées dans le tissu :

$$\text{Densité de NPs sur la zone cancéreuse (NPs / 100 } \mu\text{m}^2) = \frac{N_{\text{NPs sur la zone cancéreuse}}}{A_{\text{zone cancéreuse}}} \quad (5.6)$$

Avec cette seconde métrique, on ne s'intéresse plus à tout le tissu mais seulement à la zone cancéreuse d'intérêt. Cela permet de réhausser la densité calculée en se concentrant uniquement sur les cellules tumorales pour lesquelles l'expression de la protéine HER2 doit être quantifiée.

- sur la **zone membranaire des cellules épithéliales néoplasiques**, donc la zone cancéreuse sans les noyaux.

$$\text{Densité}_{\text{zone membranaire, épi. néo.}} (\text{NPs / 100 } \mu\text{m}^2) = \frac{N_{\text{NPs membranes épi. néo.}}}{A_{\text{membranes épi. néo.}}} \quad (5.7)$$

Cette zone correspond à la zone définie par expansion des cytoplasmes des cellules épithéliales

néoplasiques. Toutefois, le terme cytoplasme n'est pas conservé. En effet, pour l'évaluation de HER2, c'est l'attachement membranaire et non cytoplasmique qui intéresse les pathologistes. Comme la position de la membrane n'est pas précisément connue, on considère qu'elle se trouve quelque part à l'intérieur de la zone d'expansion des cytoplasmes. Isoler cette zone permet donc d'isoler les NPs qui se seraient fixées sur la membrane, et d'exclure les NPs qui se sont fixés sur les noyaux de la zone cancéreuse. Comme mentionné précédemment, les noyaux n'expriment pas la protéine HER2 et correspondent donc à des zones d'attachement non spécifique.

Pour une expérience qui a bien fonctionné, les densités calculées devraient croître avec le score IHC : plus une carotte possède un score élevé, plus elle surexprime la protéine HER2 et plus la densité en NPs devrait être élevée.

Lors de l'analyse de chaque lame, différents graphes sont générés pour représenter les données disponibles, puis les organiser afin de révéler le lien entre densité et expression protéique :

La Figure 5.11 permet de visualiser toutes les densités pour toutes les carottes d'une lame.

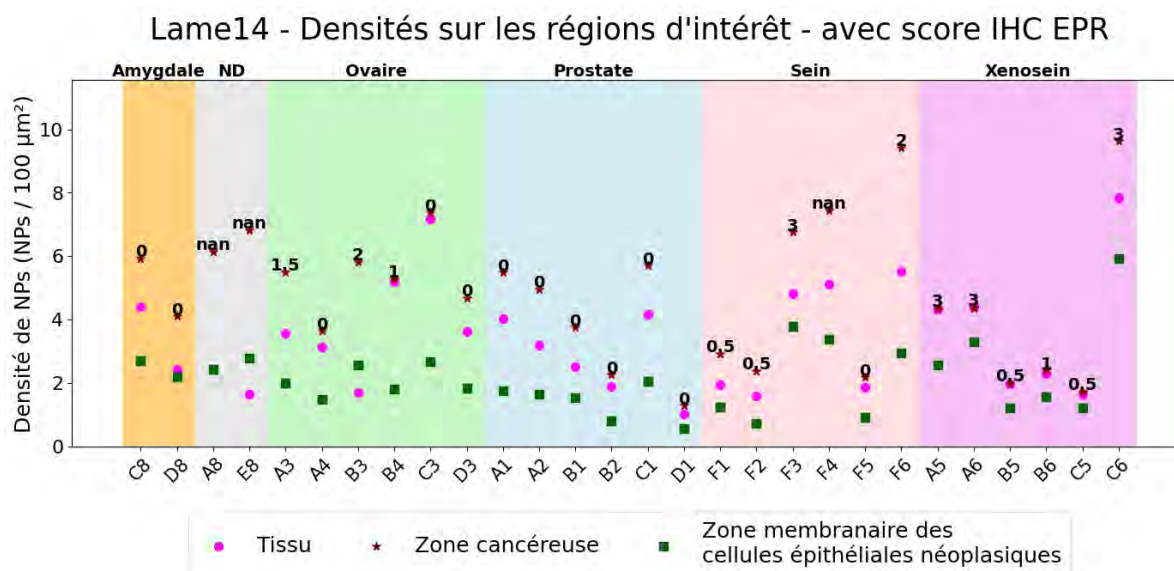


FIGURE 5.11 Densités de NPs sur les trois régions d'intérêt.

La Figure 5.12 permet de s'intéresser à un tissu en particulier pour révéler le lien entre densités et expression protéique.

Lame14 - Densité sur les tissus de Sein

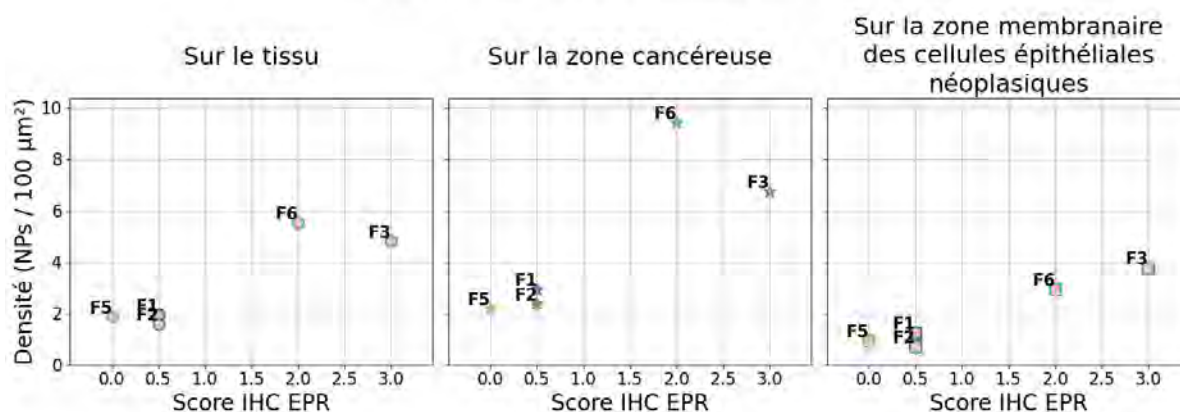


FIGURE 5.12 Densités de NPs sur les trois régions d'intérêt - Sein uniquement.

Comme mentionné précédemment, on peut voir que la densité sur la zone cancéreuse permet de réhausser la densité sur le tissu. Au contraire, la densité sur la zone membranaire des cellules épithéliales néoplasiques est basse comparée à la densité sur la zone cancéreuse, ce qui décrit un fort attachement sur les noyaux.

La Figure 5.13 montre la moyenne de densité calculée pour chaque score IHC de chaque tissu et permet de visualiser le lien entre densité et score IHC pour tous les types de tissus simultanément. La densité utilisée dans ce grahe est la **densité de NPs sur la zone membranaire des cellules épithéliales néoplasiques**. Il s'agit de la métrique la plus pertinente pour évaluer l'attachement des NPs sur la zone inspectée par les pathologistes lors de l'évaluation du statut HER2.

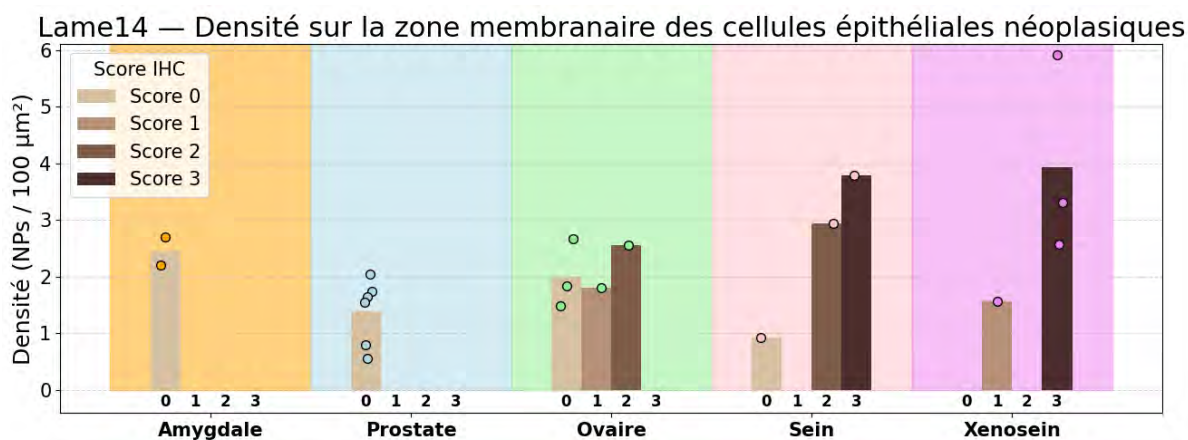


FIGURE 5.13 Densité moyenne par score IHC et par tissu.

ii. Nombre moyen de NPs par cellule

L'attachement des NPs peut aussi être exprimé en nombre moyen de NPs attachées sur une cellule. Cette métrique peut être calculée en comptant toutes les cellules du tissu sans considérer les types cellulaires, ou au contraire en choisissant un type cellulaire d'intérêt (les cellules épithéliales néoplasiques par exemple) :

- Nombre moyen de NPs par cellule, **tous types confondus** :

$$N_{\text{moyen de NPs par cellule}}(\text{NPs} / \text{cellule}) = \frac{N_{\text{NPs sur toutes les cellules}}}{N_{\text{toutes les cellules}}} \quad (5.8)$$

- Nombre moyen de NPs par cellule, pour les **cellules épithéliales néoplasiques uniquement** :

$$N_{\text{moyen de NPs par cellule épi. néo.}}(\text{NPs} / \text{cellule}) = \frac{N_{\text{NPs sur les cellules épi. néo.}}}{N_{\text{cellules épi. néo.}}} \quad (5.9)$$

Comme pour les densités, on s'attend à ce que pour une expérience qui a fonctionné, le nombre moyen de NPs fixées sur une cellule augmente avec le score IHC. Cependant, cette tendance ne doit être observée que pour les cellules épithéliales, cellules qui expriment la protéine HER2. Les cellules conjonctives et inflammatoires n'expriment pas la protéine HER2, le nombre moyen de NPs attachées sur ces cellules devrait donc rester faible et constant puisqu'il s'agit d'un attachement non spécifique.

- Pour ne pas comptabiliser les NPs fixées sur les noyaux des cellules et ainsi exclure l'attachement non spécifique, on peut calculer le nombre moyen de NPs par **membrane de cellule épithéliale néoplasique** :

$$N_{\text{moyen de NPs par membrane de cellule épi. néo.}}(\text{NPs} / \text{cellule}) = \frac{N_{\text{NPs sur les membranes de cellules épi. néo.}}}{N_{\text{cellules épi. néo.}}} \quad (5.10)$$

Lors de l'analyse de chaque lame, différents graphes sont générés pour représenter les données disponibles, puis les organiser afin de révéler le lien entre nombre moyen de NPs par cellule et expression protéique :

La Figure 5.14 permet de visualiser le nombre moyen de NPs par cellule pour tous les types cellulaires, pour toutes les carottes d'une lame.

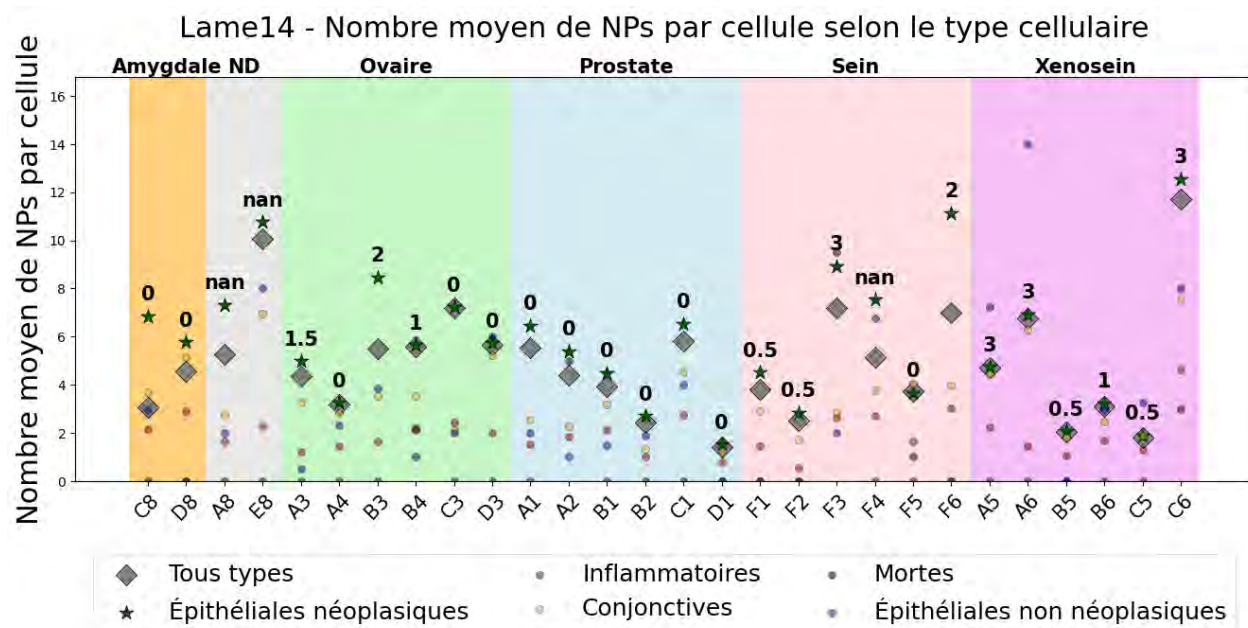


FIGURE 5.14 Nombre moyen de NPs par cellule en fonction du type cellulaire.

La Figure 5.15 permet de s'intéresser à un tissu en particulier pour révéler le lien entre nombre moyen de NPs par cellule et expression protéique.

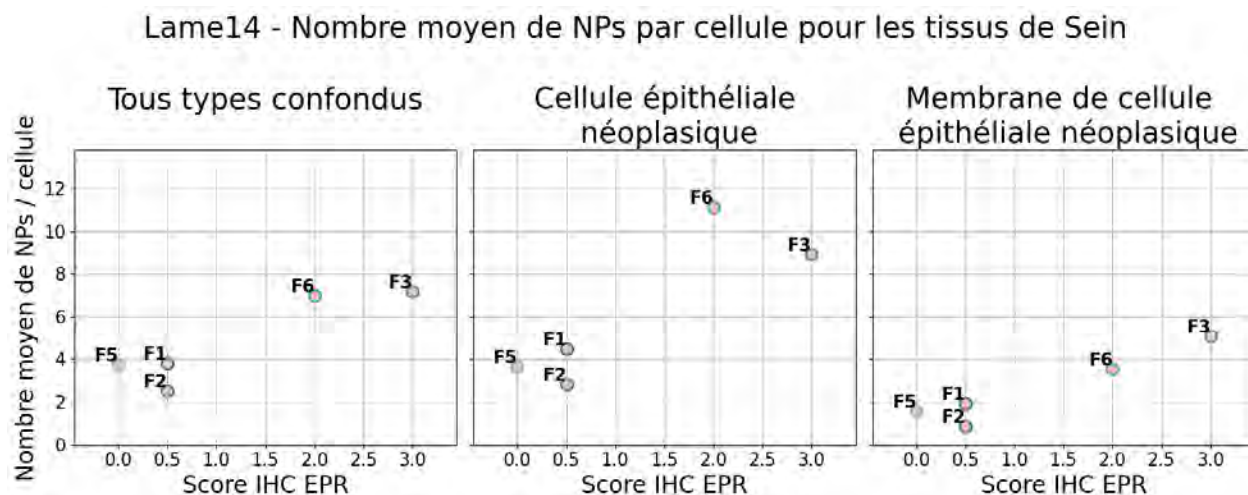


FIGURE 5.15 Nombre moyen de NPs par cellule - Sein seulement.

En se concentrant uniquement sur les cellules épithéliales néoplasiques, le nombre moyen de NPs par cellule est réhaussé. Les cellules n'exprimant pas HER2 pour lesquelles l'attachement observé est faible sont écartées du décompte. Si ce phénomène n'est pas observé, il peut y avoir un problème d'attachement non spécifique trop important sur les autres types

cellulaires. Cependant, lorsque l'on ne comptabilise que les NPs présentes sur les membranes, le nombre moyen de NPs par cellule chute, témoignant là encore d'un fort attachement aux noyaux.

La Figure 5.16 montre la moyenne du nombre moyen de NPs par cellule épithéliale néoplasique calculée pour chaque score IHC de chaque tissu. Elle permet de visualiser le lien entre le nombre moyen de NPs par cellule tumorale et l'expression protéique pour tous les types de tissus simultanément.

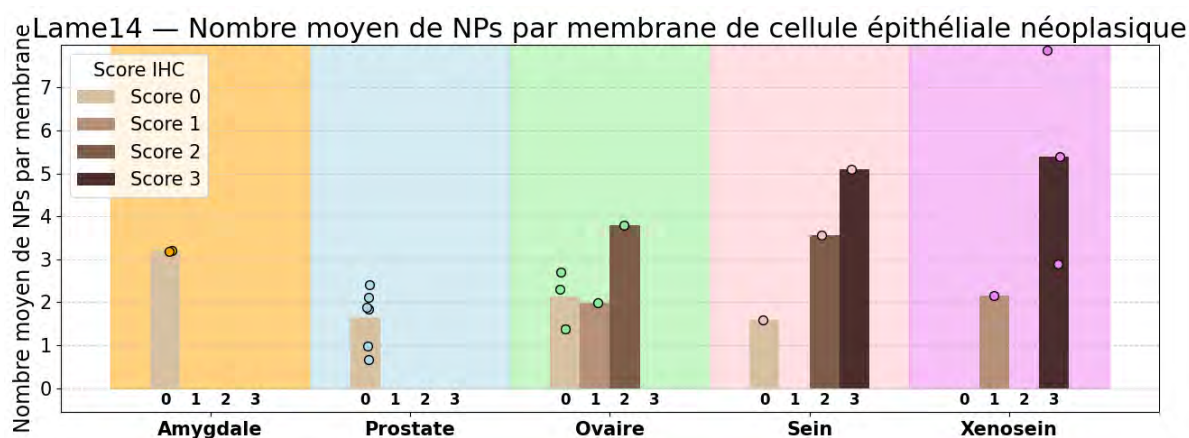


FIGURE 5.16 Moyenne du nombre moyen de NPs par cellule épithéliale néoplasique par score IHC et par tissu.

Les métriques utilisant la densité et le nombre par cellule permettent de rapporter sensiblement les mêmes informations.

iii. Métrique exploratoire

Les discussions avec les pathologistes ont révélé l'importance d'observer des NPs dans la zone membranaire des cellules épithéliales néoplasiques. En effet, c'est cette zone que les pathologistes ont l'habitude d'évaluer lorsqu'ils examinent des biopsies marquées par IHC. En IP comme en IHC, la protéine ciblée est transmembranaire, le marquage doit donc cibler les membranes des cellules tumorales. Malheureusement, la méthode IP ne permet pas toujours d'observer un beau marquage membranaire, et il n'est pas rare d'observer des NPs attachées sur les noyaux. L'**indice d'affinité membranaire épithéliale néoplasique** permet de mettre en avant les carottes pour lesquelles l'attachement membranaire observé pour les cellules épithéliales néoplasiques est plus important que l'attachement nucléaire :

$$\begin{aligned} \text{Indice d'affinité membranaire épi. néo.} &= \frac{1}{\text{Indice d'affinité nucléaire épi. néo.}} \quad (5.11) \\ &= \frac{\text{Densité de NPs dans la zone membranaire (épi. néo.)}}{\text{Densité de NPs dans le noyau (épi. néo.)}} \quad (5.12) \end{aligned}$$

La Figure 5.17 montre cet indice calculé pour toutes les carottes de la lame 14.

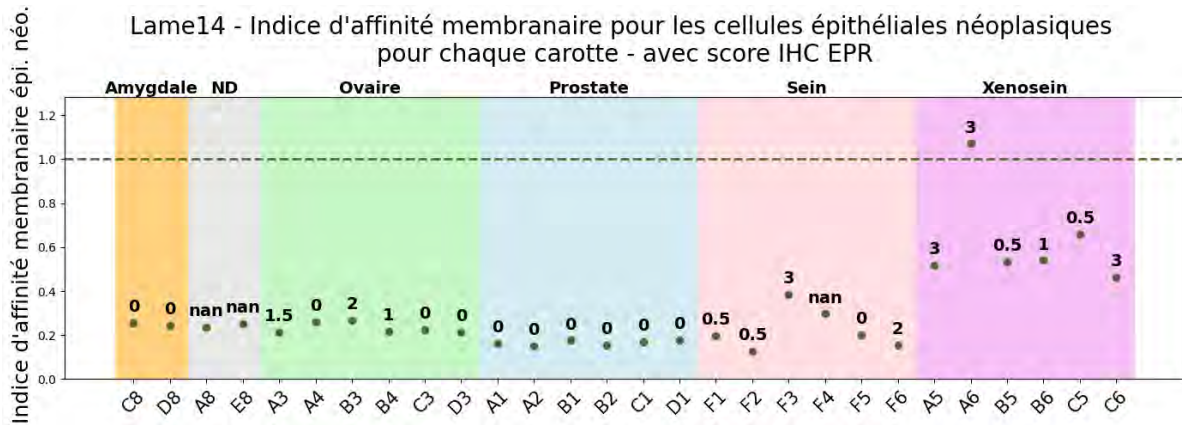


FIGURE 5.17 Indice d'affinité membranaire pour les cellules épithéliales néoplasiques.

Pour cette lame, on remarque que la carotte Xénosein A6 est la seule pour laquelle on trouve plus de NPs dans la zone membranaire que dans les noyaux. Pour toutes les autres carottes, l'attachement ne semble donc pas être spécifique à la membrane sur laquelle on peut retrouver la protéine ciblée HER2. Ces observations corréleront avec les observations qualitatives qui ont pu être réalisées lors de l'imagerie : un marquage membranaire est bien visible sur la carotte A6, contrairement aux autres carottes.

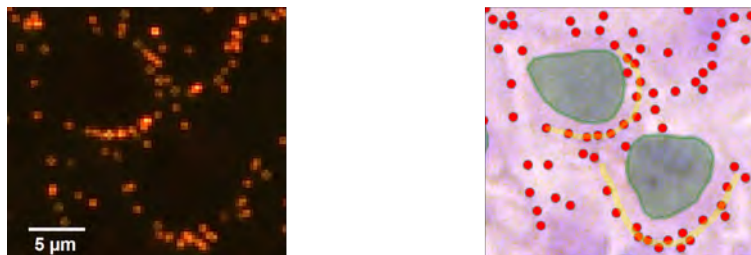


FIGURE 5.18 Marquage membranaire autour des noyaux des cellules épithéliales néoplasiques. Lame 14 - carotte Xénosein A6.

5.2 Résultats sur TMA

La méthodologie présentée à la section précédente a été utilisée pour analyser 7 lames de TMA. Pour toutes ces lames, les différentes métriques ont été calculées afin de générer les différents graphes.

5.2.1 Résultats généraux des 7 lames analysées

Dans cette section, les résultats généraux des 7 lames analysées sont présentés afin de donner un aperçu du travail réalisé et de formuler certaines conclusions globales sur les différentes expériences.

Les 7 lames analysées ont été préparées à des moments différents, en utilisant des NPs fonctionnalisées avec du calixarène provenant de différents lots et en utilisant différents paramètres d'incubation. En un mot, plusieurs paramètres expérimentaux peuvent varier simultanément entre deux expériences. Pour cette raison, les résultats donnés dans cette section permettent uniquement de soulever des tendances globales.

La Figure 5.19 montre le nombre de NPs détectées dans chaque carotte, pour toutes les lames analysées et permet donc d'avoir une idée de la tendance globale de l'attachement des NPs entre les différentes lames.

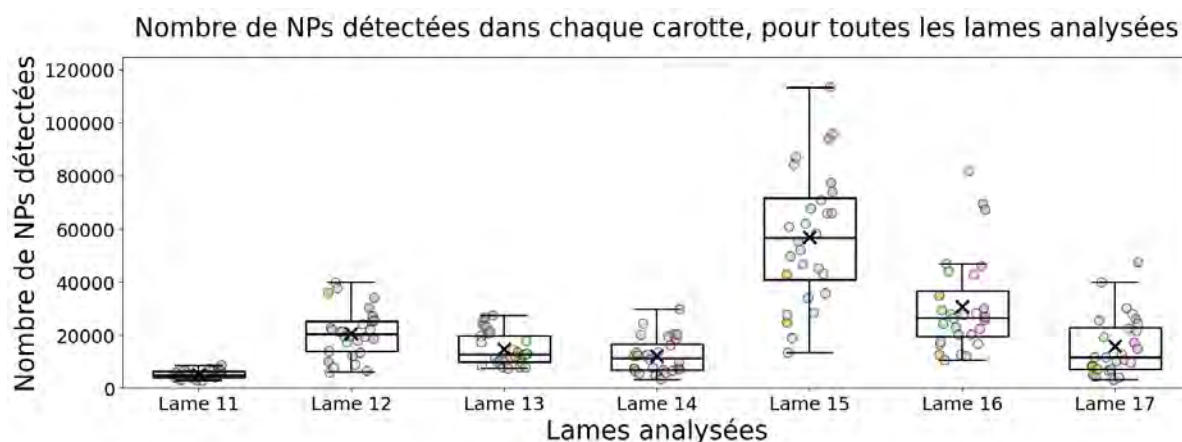


FIGURE 5.19 Évaluation de l'attachement global sur les lames analysées. Chaque point représente le nombre de NPs détectées dans une carotte, la couleur du point indique le type tissulaire.

Deux lames sortent de la tendance d'attachement globale :

- La lame 11 présente un attachement très faible et très uniforme. Il semble déjà compromis de faire la différence entre des carottes exprimant plus ou moins la protéine HER2, tant la dispersion du nombre de NPs détectées par carotte est limitée.
- La lame 15 au contraire, présente une grande dispersion du nombre de NPs détectées par carotte. Les carottes les moins marquées possèdent 20 000 NPs, tandis que les plus marquées comptent plus de 100 000 NPs. Si l'attachement des NPs corrèle avec le score IHC, il pourrait s'agir d'une expérience très réussie. La lame 15 a été préparée en diluant les NPs dans du TBST. L'utilisation de ce tampon dans le passé a déjà permis de constater qu'il semble favoriser l'attachement des NPs. Les analyses présentées ici, appuient cette observation, bien qu'aucune conclusion ne puisse être formulée du fait de la variabilité des conditions expérimentales.

TABLEAU 5.1 Moyenne du nombre de NPs détectées par carotte pour chaque lame.

Lame	Nombre de NPs détectées (moyenne $\pm$ écart-type)	Concentration (NPs/mL)	Tampon de dilution
11	5 121,89 $\pm$ 1 488,82	5×10^9	Diluant vert
12	20 604,50 $\pm$ 9 056,13	5×10^9	PBST
13	14 746,32 $\pm$ 6 133,39	7.5×10^9	Diluant vert
14	12 107,14 $\pm$ 6 534,47	5×10^9	PBST
15	56 869,39 $\pm$ 24 430,19	5×10^9	TBST
16	30 739,32 $\pm$ 17 673,92	7.5×10^9	PBST
17	15 709,82 $\pm$ 11 089,71	7.5×10^9	PBST

Qualité du marquage

Taux d'agrégation

Pour toutes les lames analysées, le taux d'agrégation est particulièrement faible. Il reste toujours inférieur au seuil de 5 %, garantissant une stabilité satisfaisante des NPs. Le taux d'agrégation semble suivre la tendance de l'attachement global : l'agrégation est plus marquée sur la lame 15, où l'attachement des NPs est particulièrement élevé. La forte densité des NPs peut contribuer à surestimer le nombre d'agrégats, car les NPs sont si proches les unes des autres qu'il devient difficile de les séparer. Les NPs détectées peuvent ainsi regrouper plusieurs particules voisines, comptabilisées comme un seul agrégat.

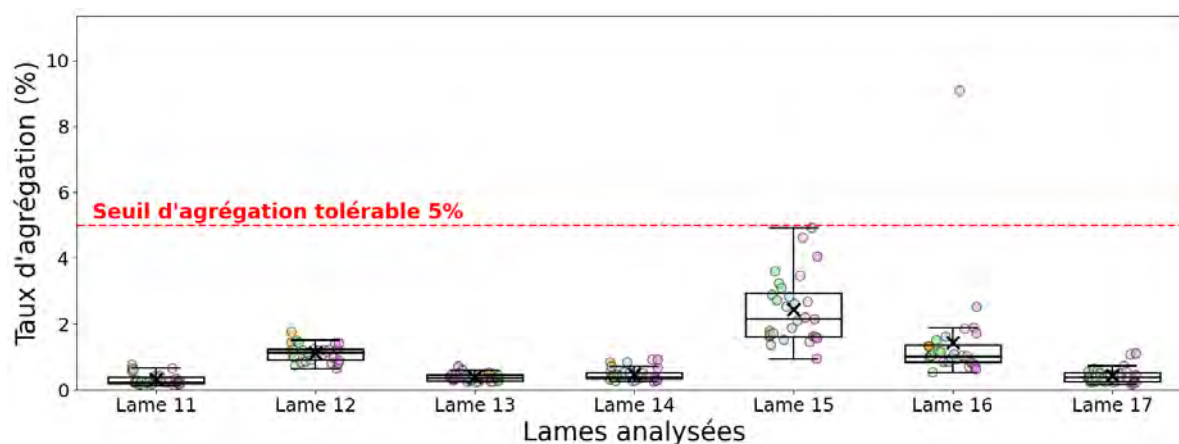


FIGURE 5.20 Taux d'agrégation pour toutes les lames analysées.

Il apparaît donc essentiel de choisir des conditions expérimentales permettant d'obtenir une densité d'attachement raisonnable sur les carottes, afin que les NPs puissent toujours être détectées individuellement. Une densité trop élevée correspondrait à une limite supérieure de la plage dynamique de la méthode IP.

Attachement non spécifique

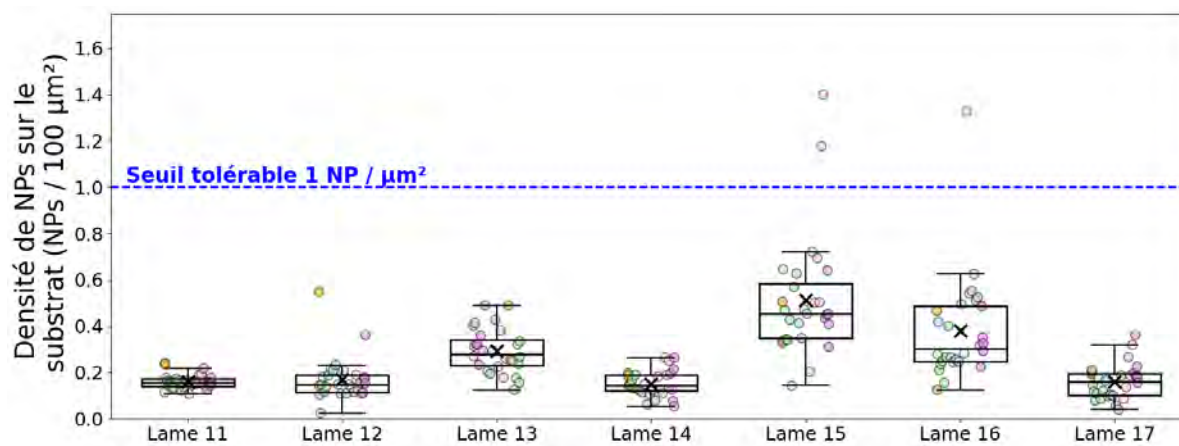
La Figure 5.21 présente les densités d'attachement non spécifique pour les 7 lames analysées.

Densité sur le substrat

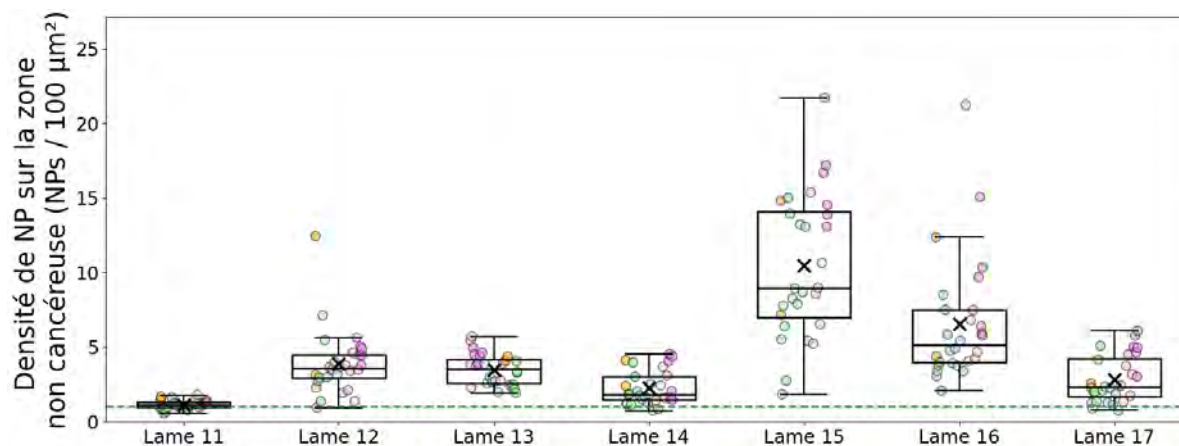
Pour toutes les lames analysées, la densité sur le substrat reste en dessous du seuil établi à 1 NP / 100 μm^2 . On peut reconnaître la tendance de l'attachement global, mais même pour la lame 15, très peu de NPs se fixent sur le substrat.

Densité sur la zone non cancéreuse

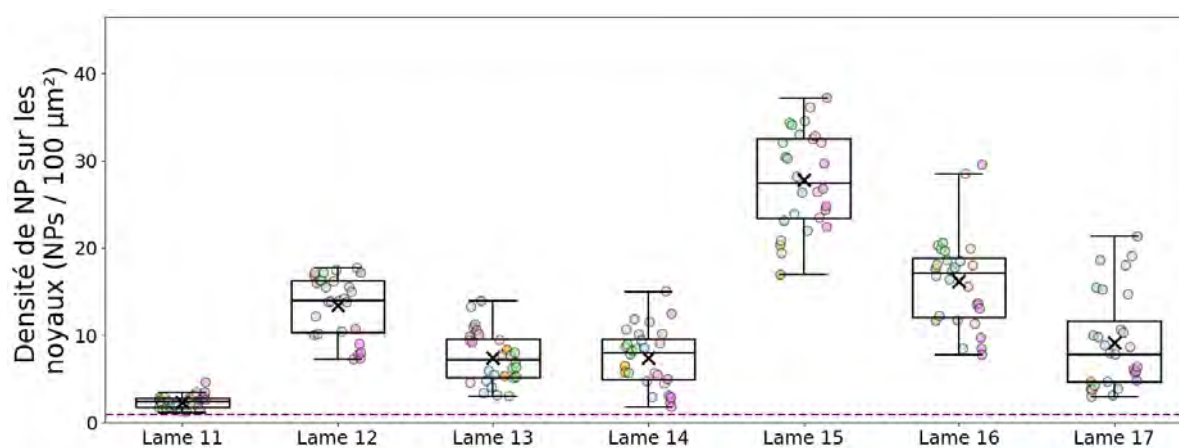
La densité sur la zone non cancéreuse est globalement élevée, hormis sur pour la lame 11. Seules certaines carottes présentent une densité en NPs inférieure au seuil établi à 1 NP / 100 μm^2 . Pour les lames 15 et 16 en particulier, l'attachement non spécifique sur la zone non cancéreuse est important.



(a) Densité de NPs sur le substrat.



(b) Densité de NPs sur la zone non cancéreuse.



(c) Densité de NPs sur les noyaux.

FIGURE 5.21 Résumé des densités de NPs sur différentes zones.

Densité sur les noyaux

La densité sur les noyaux est particulièrement élevée pour toutes les lames, sauf pour la lame 11. Les valeurs de densité, bien plus élevées que pour la densité sur les zones cancéreuse permettent de témoigner de l'ampleur du phénomène d'attachement non spécifique sur les noyaux. On peut encore reconnaître la tendance de l'attachement global, avec un maximum de la densité moyenne sur les noyaux atteint pour la lame 15 avec environ 30 NPs / 100 μm^2 sur les noyaux.

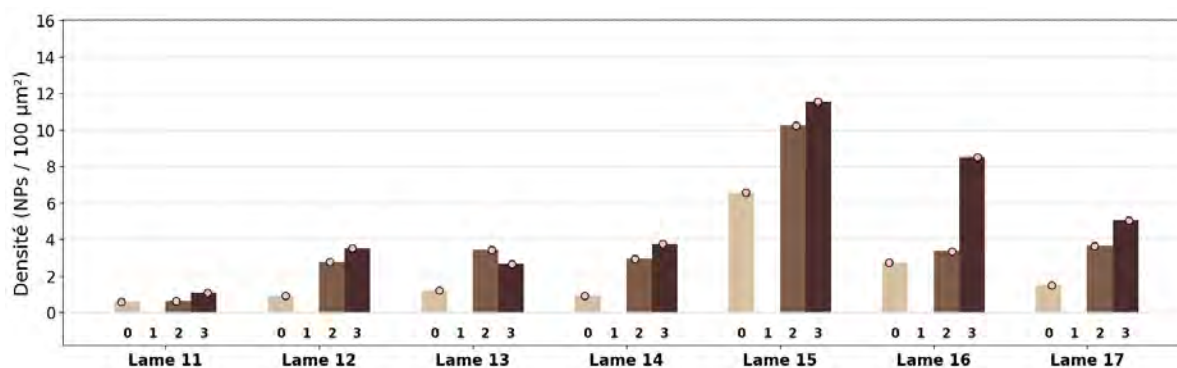
Les résultats de l'analyse de l'attachement non spécifique pour les 7 lames analysées peuvent être résumés comm suit : Même si la densité sur le substrat est correcte pour toutes les lames analysées, l'attachement non spécifique reste trop important. Seule la lame 11, pour laquelle l'attachement semble être uniformément faible, donc semble ne pas avoir fonctionné, présente des densités suffisamment faibles sur les structures n'exprimant pas la protéine HER2. Améliorer le protocole d'incubation pour limiter l'attachement non spécifique semble indispensable pour être en mesure d'assurer que la technique est spécifique. Détecter de faibles niveaux d'expression de la protéine HER2 ne sera possible que si l'absence de protéine se traduit par un attachement des NPs quasi nul.

Evaluation du statut HER2

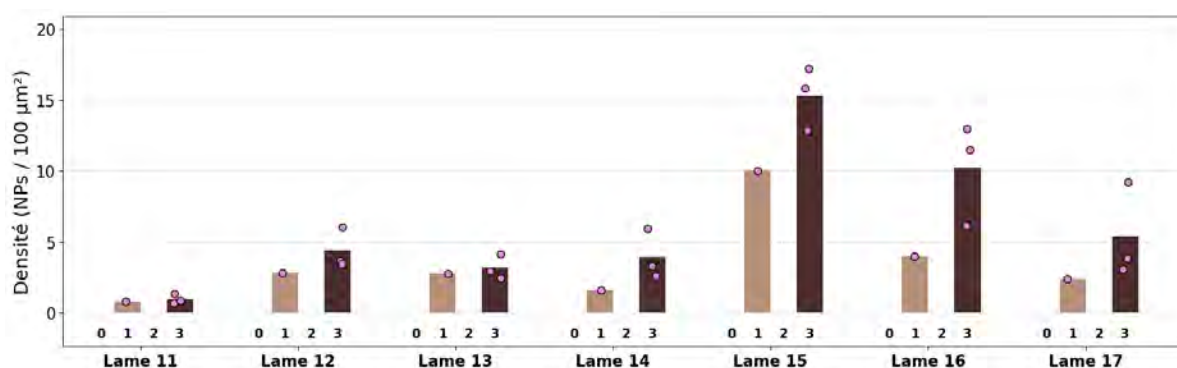
Les Figures 5.22 et 5.23 présentent l'attachement sur les membranes des cellules épithéliales néoplasiques, sous la forme de densité et de nombre par cellule. Les sous-figures permettent d'isoler les types tissulaires Sein et Xénosein particulièrement intéressants dans le contexte de ce travail. L'objectif est d'identifier les lames pour lesquelles l'attachement corrèle avec le score IHC. De plus, il est souhaitable que l'écart entre les scores soit le plus grand possible. Cela signifierait que la plage dynamique est grande et qu'il est possible de distinguer facilement les carottes présentant des expressions protéiques différentes.

Quand on s'intéresse à la **densité sur les membranes épithéliales néoplasiques**, il est possible de faire les observations suivantes. Pour la lame 11, l'attachement est uniformément faible pour tous les scores. Pour les lames 12, 13, 14, l'attachement croît avec le score, mais l'écart entre chaque catégorie est particulièrement faible (moins d'1 NP d'écart par score). Pour la lame 15, l'écart entre chaque score est plus grand, mais les scores 0 et 1 sont difficilement différenciables. Pour la lame 16, seule la catégorie 3 se détache. La lame 17 est celle qui présente les meilleurs résultats. Bien que l'écart entre les différentes catégories reste faible, chaque niveau d'expression est identifiable.

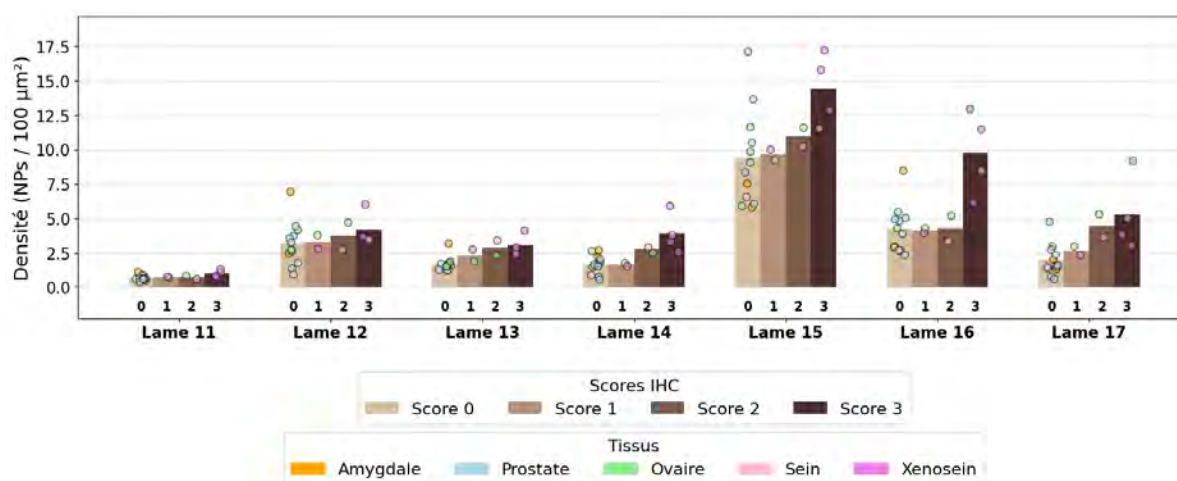
Densité sur la zone membranaire des cellules épithéliales néoplasiques



(a) Tissu de Sein uniquement.



(b) Tissu de Xénosein uniquement.



(c) Tous tissus confondus.

FIGURE 5.22 Densité moyenne sur la zone membranaire des cellules épithéliales néoplasiques, calculée pour tous les scores IHC de chaque lame.

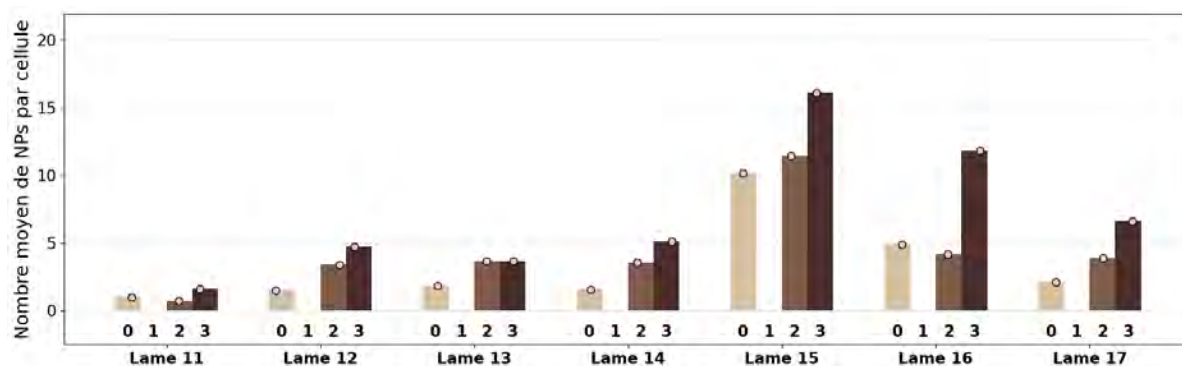
Pour toutes les carottes, les scores faibles 0 et 1 sont réhaussés à cause des carottes d'Ovaire et de Prostate, sur lesquelles un attachement important est souvent observé malgré des niveaux d'expression de HER2 faible (la Figure 5.16 permet d'observer cette tendance). C'est pourquoi il peut être intéressant de regarder les sous-figures qui ne présentent que les types tissulaires d'intérêt Sein et Xénosein. Cependant, le manque de diversité de carottes pour ces tissus limite les données disponibles et la représentation de tous les scores.

La Figure 5.23 qui présente le nombre de NPs par membrane de cellule épithéliale néoplasique permet de faire sensiblement les mêmes observations malgré de petites variations. En particulier, les scores 2 et 3 de la lame 17 sont plus proches qu'auparavant à cause d'une carotte d'Ovaire scorée 2 pour laquelle le nombre moyen de NPs par cellule est élevé. Pour la lame 15 au contraire, la distinction entre le score 0 et 1 est plus marquée.

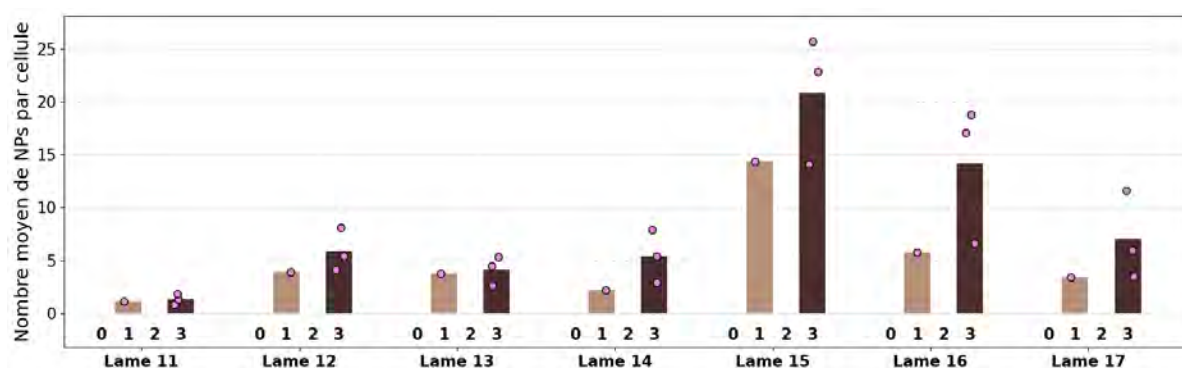
Les résultats de l'analyse du statut HER2 pour les 7 lames analysées peuvent être résumés tel que suit :

La densité de NPs par cellule épithéliale néoplasique, tout comme le nombre de NPs par membrane de cellule épithéliale néoplasique suit bien l'expression protéique, puisque la métrique suit globalement la même tendance que le score IHC. Cependant, il est encore difficile d'avoir un écart important entre les différents scores, qui permettrait facilement de retrouver le score d'une carotte à partir de sa densité en NPs. L'optimisation de la méthode doit continuer afin de déterminer les paramètres expérimentaux qui permettrait d'étendre la plage dynamique suffisamment pour être capable de faire la distinction entre les différentes expressions protéiques. Aussi, l'attachement semble beaucoup dépendre du type de tissu. Il n'est peut être pas pertinent de calculer des moyennes impliquant des échantillons provenant de différents types cellulaires. Il pourrait donc être intéressant d'avoir accès à un TMA présentant une plus grande diversité de carottes pour pouvoir mener des expériences plus significatives.

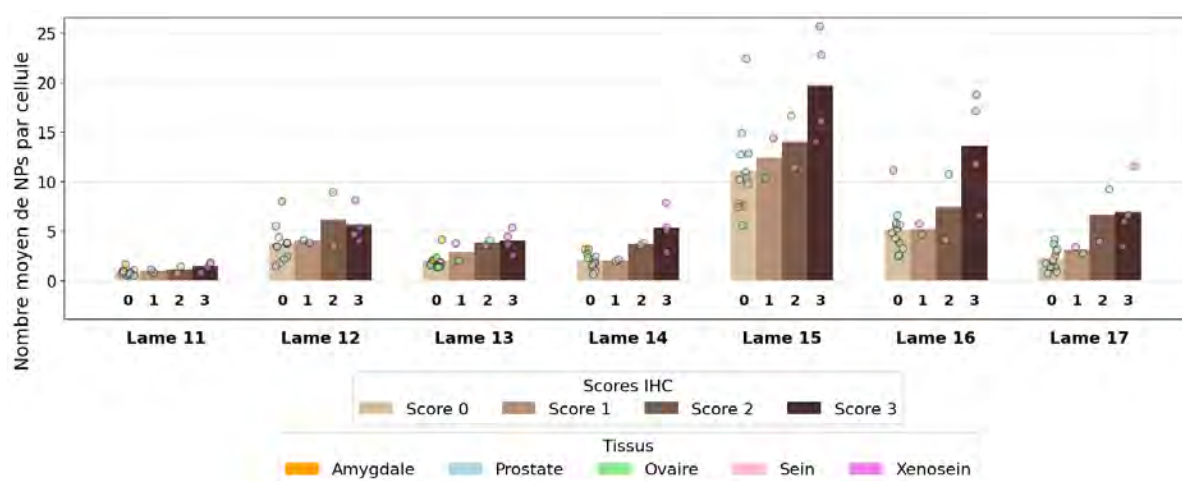
Nombre de NPs par membrane de cellule épithéliale néoplasique



(a) Tissu de Sein uniquement.



(b) Tissu de Xénosein uniquement.



(c) Tous tissus confondus.

FIGURE 5.23 Nombre moyen de NPs par cellule épithéliale néoplasique, calculé pour tous les scores IHC de chaque lame.

5.2.2 Utilisation du marquage IP pour affiner le score IHC

Nous venons de confirmer que le marquage IP a bien fonctionné sur la lame 17, et que la densité de NPs sur la membrane des cellules épithéliales néoplasiques correspond bien au niveau d'expression de HER2.

La Figure 5.24 permet de visualiser la densité de NPs sur les membranes des cellules épithéliales néoplasiques pour toutes les carottes de Sein et de Xénosein. En utilisant ces deux types de tissu, il est possible de disposer d'au moins une carotte par score IHC, tout en excluant les tissus qui semblaient avoir des comportements trop éloignés des tissus de Sein. La moyenne de densité calculée pour chaque score apparaît avec une ligne en pointillés.

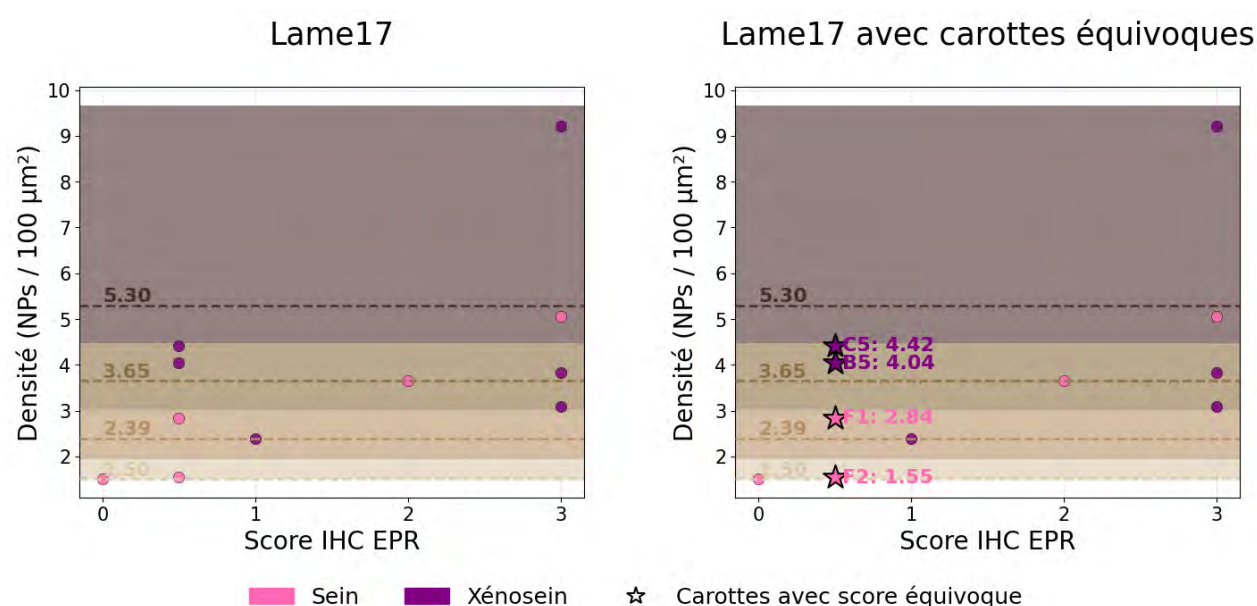


FIGURE 5.24 Classification des carottes équivoques en IHC grâce au marquage IP.

Sur la deuxième sous-figure, les carottes équivoques F1, F2, B5 et B6 ont été ajoutées. Ces carottes se sont vues attribuer le score de 0,5 après que l'évaluation par deux pathologistes ait donné des scores discordants 0 et 1. Grâce aux quatre zones tracées autour de chaque moyenne, il est possible d'utiliser le marquage IP pour préciser le score IHC :

- La carotte F2, avec une densité de 1,55 NPs / 100 μm^2 devrait être scorée 0.
- La carotte F1, avec une densité de 2,84 NPS / 100 μm^2 est plus proche d'un score 1+.
- Les carottes B5 et C5, bien que scorées provisoirement 0,5 (scores 0 et 1+ donnés par les pathologistes interrogés). ont des densités respectives de 4,04 et 4,42 NPs / 100 μm^2 , ce qui devrait leur valloir le score 2+. Ces deux carottes seraient donc définitivement plus des scores 1+ que 0.

5.3 Translation préliminaire aux CMA

Le développement de la méthode IP implique souvent des CMA. En effet, il est souvent plus simple de commencer avec des CMA dont l'analyse peut être plus intuitive, comme chaque carotte ne contient qu'un seul type cellulaire. Il n'est pas nécessaire de comprendre l'hétérogénéité de la morphologie des tissus pour conclure sur les expériences réalisées avec un CMA. Il semble donc pertinent de proposer une adaptation de la méthodologie développée pour les CMA, afin que les différents membres du laboratoire du Professeur Meunier puissent intégrer des analyses quantitatives à leurs résultats. Cela permettrait un meilleur suivi des expériences réalisées et contribuerait à optimiser les protocoles de manière plus efficace.

5.3.1 Méthodologie

Deux lames du CMA présenté dans la section 3.1.2 et préparées par Jennyfer Zapata Farfan ont été imagées selon la méthodologie présentée dans la section 3.2.3.

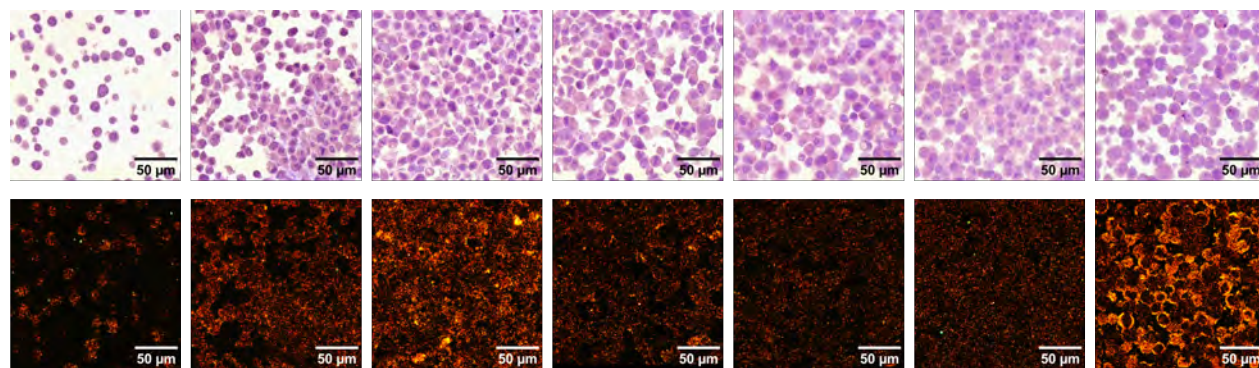


FIGURE 5.25 Lame 9 du CMA imagée en BF et en SI. Objectif 40x à air NA=0,6 + zoom 1,5x.

L'objectif est ensuite d'extraire le contexte biologique de chaque image BF, et de détecter les NPs visibles dans l'image SI. Pour cela, les images d'un même CMA sont chargées dans **Dragonfly** sous la forme de deux piles d'images, une par modalité d'illumination.

Détection des NPs : Comme pour les TMA, le modèle *SegNPs* est appliqué à la pile SI pour détecter les NPs. La liste des NPs est exportée sous la forme d'un tableur `.csv` contenant la liste des NPs détectées avec leur centre de masse X, Y, Z (qui permet de savoir de quelle carotte il s'agit), et leur diamètre circulaire équivalent.

Extraction du contexte biologique : Pour l'analyse des CMA, CellViT n'est plus indispensable. En effet, comme chaque carotte contient un unique type cellulaire, la classification cellulaire n'est plus pertinente. L'utilisation de l'image BF se résume donc uniquement à la détection des cellules, qui peut être réalisée dans Dragonfly. Pour cela, deux approches sont envisageables :

- Si le nombre de cellules n'est pas considéré comme important, le modèle *TissuDetection* peut être appliqué pour récupérer un masque où le tissu détecté correspond aux cellules.
- Si le nombre de cellules est important, le modèle *CellPose*, directement accessible dans Dragonfly peut être appliqué. Il permet de récupérer un masque en niveaux de gris où chaque intensité est propre à une cellule.

Comme pour les TMA, les NPs détectées peuvent ensuite être classées pour discriminer les agrégats, et pour déterminer sur quelle structure chaque NP s'est fixée. Ensuite, des métriques similaires à celles présentées précédemment peuvent être employées (sans décliner en fonction des types cellulaires). La Figure 5.26 présente par exemple les densités de NPs sur le substrat, et sur les cellules calculées avec *TissuDetection* et avec *CellPose*.

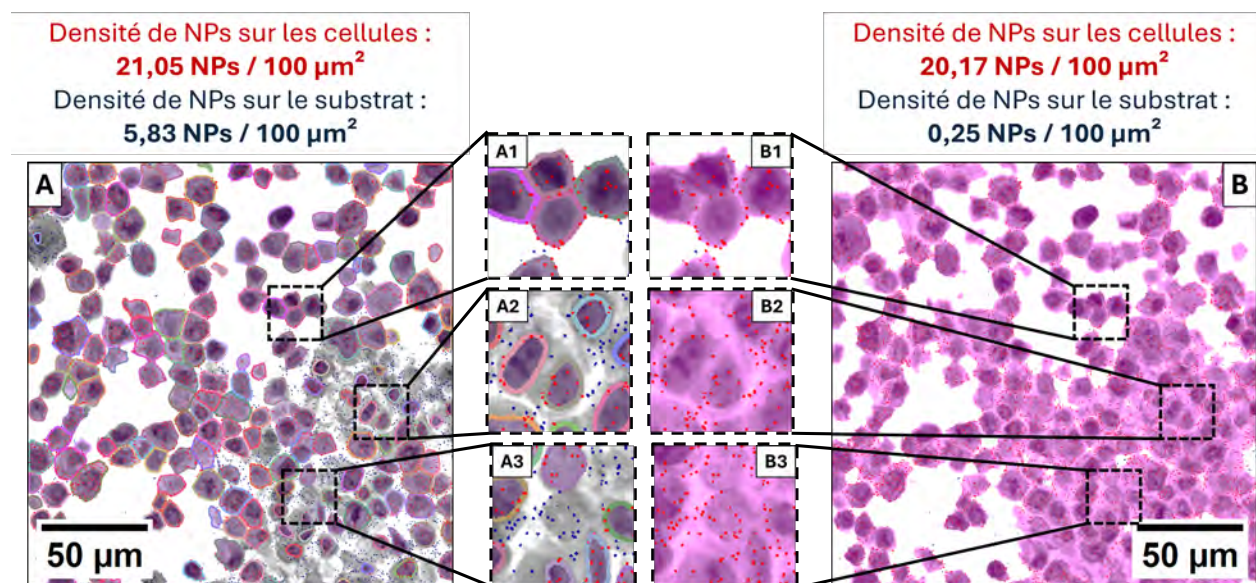


FIGURE 5.26 Comparaison des résultats de la segmentation des cellules par CellPose (A), et par le modèle de détection du TissueDetection (B). 1 : CellPose performe bien pour détecter et séparer les cellules adjacentes, TissueDetection ne fait pas la distinction entre les différentes cellules. 2 : CellPose restreint certaines cellules à leur noyau. 3 : CellPose manque la détection de cellules au complet.

Très vite, les limites de CellPose apparaissent : certaines cellules sont manquées et parfois, seul le noyau est détecté. Par conséquent, la densité de NPs sur le substrat est largement surestimée puisque certaines NPs sont détectées sur le substrat à cause des oublis de CellPose. Toutefois, la densité de NPs sur les cellules n'est que très peu affectée par ces oublis, comme le montre la Figure 5.27. Cela signifie que la surface de cellules oubliée par CellPose reste négligeable devant toute la surface bien détectée. Pour travailler avec la densité de NPs sur les cellules, les deux approches semblent donc pouvoir être utilisées. Pour contrôler l'attachement non spécifique sur le substrat, l'utilisation de *TissuDetection* est préférable.

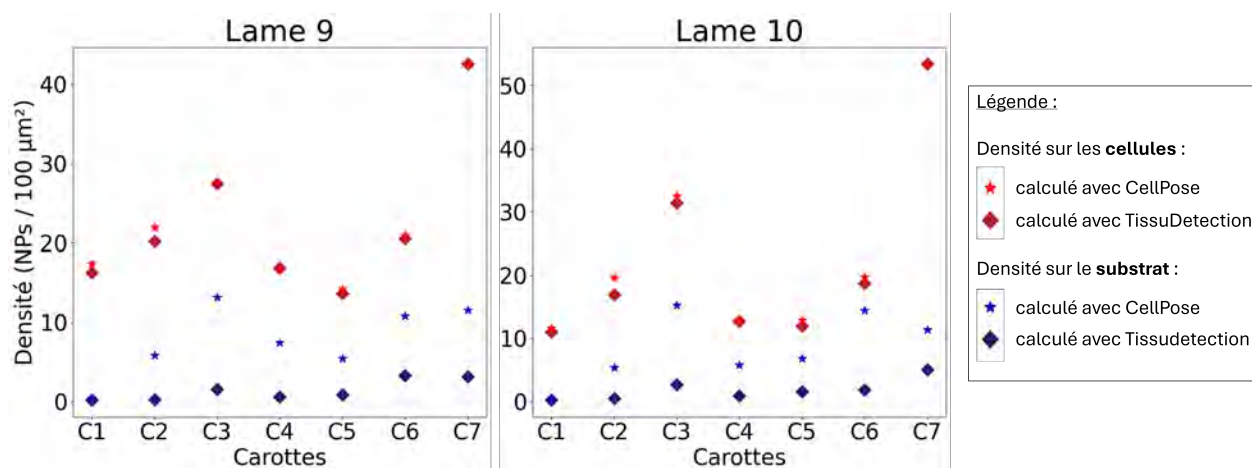


FIGURE 5.27 Comparaison des densités obtenues en utilisant *TissuDetection* ou *CellPose* pour détecter la surface occupée par les cellules.

5.3.2 Exemple de résultats

La méthodologie décrite a été utilisée pour comparer l'attachement observé avec deux lots de NPs incubés dans les mêmes conditions, mais fonctionnalisées différemment. Les NPs de la lame 9 ont été bioconjuguées en utilisant 2 μ L d'anticorps (la quantité préconisée jusqu'à présent), alors que les NPs de la lame 10 ont été bioconjuguées en utilisant 4 μ L. Les densités mesurées dans chaque carotte sont présentées dans la Figure 5.28.

La densité de NPs sur les cellules mesurée pour chaque carotte suit la même tendance pour les deux bioconjugaisons, ce qui montre une forme de reproductibilité. Toutefois, on peut noter que pour les carottes plus denses en NPs (C3 et C7), l'utilisation de 4 μ L d'anticorps a permis d'augmenter la densité. Au contraire, pour les carottes moins denses (C1, C2, C4, C5, C6), l'utilisation d'une plus grande quantité d'anticorps a engendré une diminution de

la densité mesurée. Par conséquence, l'augmentation de la quantité semble augmenter l'écart entre la densité des carottes positives et négatives. Dans un contexte où l'on souhaite s'assurer que cet écart soit suffisamment important pour permettre de classer efficacement les carottes, l'utilisation d'une plus grande quantité d'anticorps pourrait être intéressante. Cette observation serait bien sûr à confirmer avec des répliquats.

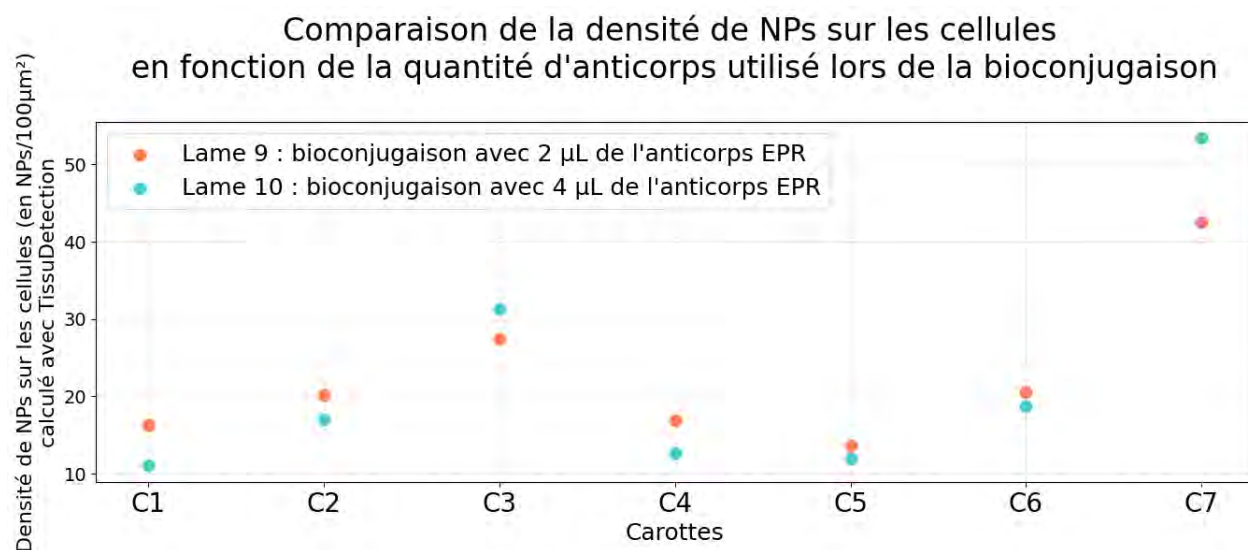


FIGURE 5.28 Comparaison de la densité de NPs sur les cellules en fonction de la quantité d'anticorps utilisés pour la bioconjugaison.

CHAPITRE 6 IMMUNOPLASMONIQUE SUR LES STANDARDS DE LA COMPAGNIE BCS

Un travail parallèle a été mené afin d'appliquer la méthode IP aux standards de la compagnie BCS présentés dans la section 3.1.2. L'objectif était, dans un premier temps, de réaliser une preuve de concept démontrant qu'il est possible de détourner ces produits de leur usage initial pour les utiliser avec un marquage IP plutôt qu'IHC. Dans un second temps, les contrôles ont été mis en place afin de vérifier la bonne fonctionnalisation des NPs utilisées pour marquer les TMA. Enfin, les calibrateurs ont été employés pour tenter de comparer la plage dynamique de l'IHC à celle de l'IP.

Ce chapitre présente donc un travail exploratoire, distinct de l'analyse d'image. Néanmoins, il a permis de mieux comprendre les enjeux associés aux différentes étapes du protocole IP, de s'interroger sur la chimie des réactions immunologiques, d'ajouter des contrôles essentiels à une technique encore en cours d'optimisation, et surtout d'appliquer le modèle de détection des NPs dans un autre contexte, cette fois tridimensionnel, ouvrant ainsi la voie à des analyses plus complexes susceptibles d'être pertinentes pour les CMA.

6.1 Preuve de concept - Attachement des NPs aux contrôles BCS

Dans un premier temps, le protocole IP a été appliqué aux billes de contrôle IHControls®. Un travail préliminaire a été effectué afin de s'assurer du succès de la fonctionnalisation de trois types de NPs. Plusieurs combinaisons NPs / PEG ont été testées avant d'obtenir des résultats positifs au test LFA. Ces combinaisons sont présentées dans le Tableau 6.1.

TABLEAU 6.1 Combinaisons de NPs et de PEG investiguées pour la fonctionnalisation des NPs avec EPR.

	Au90NPs Nanopartz	Au90NPs NanoBrand	Au90NPs synthétisées laboratoire	50:50 Au:Ag 73NPs synthétisées laboratoire	10:90 Au:Ag 50NPs synthétisées laboratoire
HS-PEG _{1kDa}	✗	✗	✗	✓	✓
LA-PEG _{1kDa}	✓	✗	✗	✗	✗

Finalement, ce sont les trois types de NPs présentées à la section 3.1.1 (Au90NPs-Nanopartz-LAPEG-EPR, 50 50 Au :Ag 73NPs-HSPEG-EPR et 10 90 Au :Ag 50NPs-HSPEG-EPR) qui ont permis d'obtenir des tests LFA positifs et un attachement aux billes de contrôle.

Une goutte de la solution de contrôle est déposée sur une lame de microscopie chargée positivement. Après environ cinq minutes de séchage à l'air libre, la matrice se fige et la lame peut être utilisée. Une zone est ensuite délimitée autour du dépôt à l'aide de cire afin de créer une barrière hydrophobe. Une première étape de blocage est réalisée avec 100 μ L d'une solution de PBS contenant 5 % de BSA. Il a été observé qu'augmenter la concentration de BSA de 1 à 5 % permet de limiter l'attachement non spécifique sur le substrat. Après 30 minutes d'incubation, la solution de blocage est retirée et la zone est lavée trois fois avec du PBST. Les NPs sont ensuite concentrées à 5×10^9 NPs/mL dans la même solution de blocage que précédemment, puis 100 μ L sont incubés sur le dépôt de billes. Après une heure d'incubation, les NPs sont retirées et trois nouveaux lavages au PBST sont effectués. Finalement, la lame est déshydratée à l'aide de bains successifs d'éthanol et de xylène, puis montée au Permount. Après 24 heures de séchage, la lame est prête à être observée en microscopie.

L'incubation des trois types de NPs fonctionnalisées avec EPR a permis d'observer un attachement spécifique uniquement sur certaines billes, en accord avec les résultats attendus d'après les spécifications de BCS pour le produit utilisé (présence de billes recouvertes de la protéine HER2, mais aussi de billes recouvertes d'ER et de PR, ainsi que de billes contrôles négatives sans protéine). Les images acquises en SI avec l'objectif 60 $\times$ à air (NA = 0,7) et le zoom de 1,5 $\times$ sont présentées dans la Figure 6.2. On peut clairement y voir des billes positives recouvertes de NPs, et des billes négatives sur lesquelles les NPs ne se sont pas fixées.

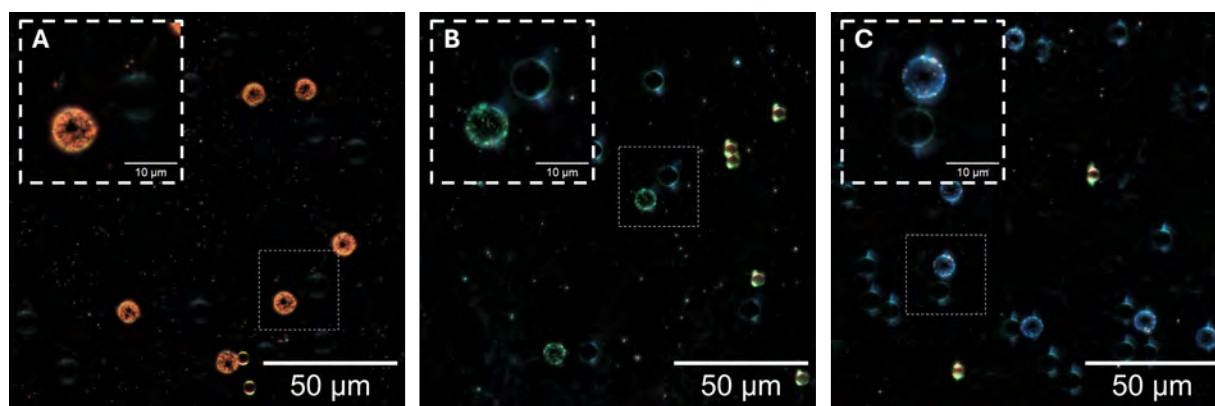


FIGURE 6.1 Trois types de NPs fonctionnalisées avec EPR ciblant les billes de contrôle. SI - Objectif 60x air NA=0,7.

Comme la solution IHControls® utilisée contient des billes exprimant HER2, ER et PR, il serait intéressant de réaliser une preuve de concept de multiplexage en utilisant les trois types de NPs présentés, chacun fonctionnalisé avec un anticorps différent ciblant l'une des trois protéines. Des essais ont donc été effectués afin d'observer un attachement spécifique similaire, mais en utilisant des anticorps différents. Les anticorps employés sont présentés dans le Tableau 6.2.

TABLEAU 6.2 Clones utilisés pour tenter de fonctionnaliser les trois types de NPs, et cibler les protéines HER2, ER et PR à la surface des billes de contrôle.

Cible de l'anticorps	HER2				PR	ER β
Clône	EPR19547-12 Abcam	ICR12 Abcam	4D5-8 (Trastuzumab) Invitrogen	PATHWAY 4B5 * Roche Diagnostics	R.380.2 * Invitrogen	14C8 ** Genetex
Test LFA	✓	✓	✓	✗	✗	✓
Attachement spécifique sur les billes de contrôle	✓	✗	✗	✗	✗	✗

* Clônes contenant du sodium azide comme conservateur, non compatible avec la chimie EDC/sulfo-NHS.

** Clône ciblant l'isoforme β du récepteur ER. L'isoforme d'intérêt pour le cancer du sein est l'isoforme α . Le choix de cet anticorps n'était donc pas pertinent.

Les clones conservés avec du sodium azide ou de la BSA sont incompatibles avec la chimie EDC/sulfo-NHS. En effet, ces composés aminés peuvent entrer en compétition avec les anticorps lors de la fonctionnalisation, entraînant des tests LFA négatifs (synonyme d'échec de la fonctionnalisation). Une étape de filtration serait envisageable mais s'avère délicate à cause du risque d'endommagement et de perte des anticorps durant le procédé.

Pour les clones sans conservateurs, bien que les NPs aient été correctement bioconjuguées, aucun attachement spécifique n'a été observé avec les clones autres que EPR. Chaque clone d'un anticorps reconnaît un épitope particulier de la protéine cible. Les billes utilisées ne sont pas recouvertes de la protéine complète, mais seulement d'une portion de celle-ci, appelée épitope. Par conséquent, la réaction immunologique ne peut se produire que si la NP est fonctionnalisée avec un anticorps ciblant spécifiquement l'épitope exposé à la surface de la bille. Il est souvent difficile de connaître la séquence peptidique exacte des produits commerciaux. Ainsi, la séquence reconnue par les anticorps commerciaux est inconnue, tout comme

la séquence présente à la surface des billes. La compagnie BCS préconise seulement d'utiliser certains clones cliniques, qui ne peuvent pour le moment pas être utilisés en IP, comme discuté précédemment.

Malheureusement, il n'a donc pas été possible de réaliser un multiplexage à l'aide des billes de contrôle. Il faudrait continuer le travail d'adaptation des clones cliniques à la méthode IP, ou bien continuer les expérimentations par essai-erreur pour parvenir à identifier des clones commerciaux permettant de cibler l'épitope présent à la surface des billes. Cependant, l'attachement spécifique observé avec le clone EPR constitue une preuve de concept du marquage des standards BCS par la méthode IP et ouvre la voie à leur utilisation.

6.2 Utilisation des contrôles BCS comme contrôle qualité

Les billes de contrôle ont été adoptées comme contrôle de routine afin de confirmer la bonne fonctionnalisation des NPs utilisées. Lors de la production de chaque nouveau TMA, un dépôt de billes est réalisé sur la lame, à côté du TMA, après les étapes de prétraitement du tissu (chauffage, déparaffinage, récupération antigénique). Il a été observé que des résultats similaires étaient obtenus avec des billes ayant subi ou non ces étapes de prétraitement. Une fois la matrice figée, une zone est délimitée à la cire, puis le même protocole de blocage et d'incubation que celui appliqué au tissu est simultanément appliqué aux billes de contrôle.

6.2.1 Contrôle qualitatif

L'observation des TMA au microscope commence généralement par une vérification qualitative rapide des billes de contrôle, qui permettent de témoigner du succès de la fonctionnalisation. Lorsque certaines billes apparaissent recouvertes de NPs tandis que d'autres restent entièrement vides, la fonctionnalisation est considérée comme réussie. En revanche, si toutes les billes apparaissent vides, cela suggère une anomalie. La Figure 6.2 présente les images des billes de contrôle pour chacune des lames analysées.

L'attachement spécifique observé pour les lames 14, 15, 16 et 17 est de bonne augure. Au contraire, pour les lames 11 et 13, l'absence d'attachement sur les billes semble suggérer une anomalie. Pour la lame 12, les billes de contrôle n'ont pas été utilisées puisque l'on sait déjà que l'anticorps Trastuzumab ne permet pas de cibler l'épitope présent à la surface des billes exprimant HER2 (Tableau 6.2).

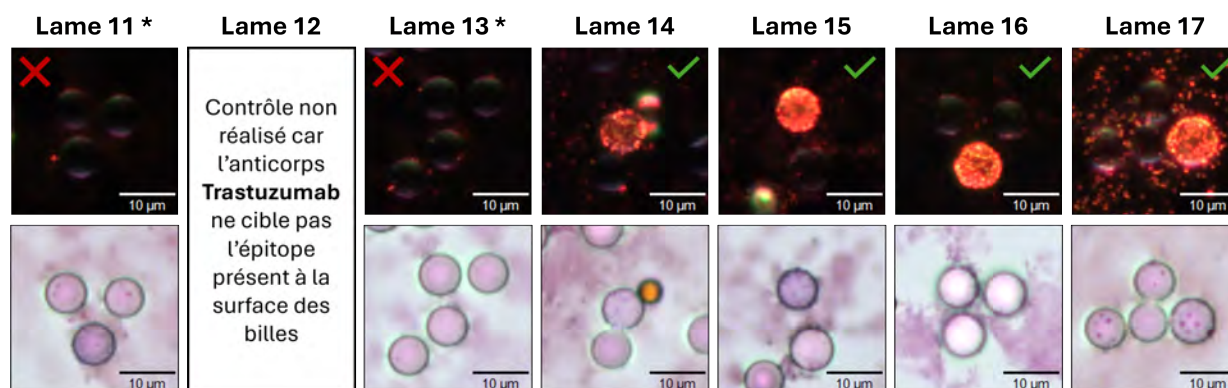


FIGURE 6.2 Billes de contrôles utilisées comme contrôle de routine sur les lames de TMA. Les images en BF témoignent de la présence de billes dans l'image malgré l'absence de NPs sur certaines billes. SI - Objectif 60x air NA=0,7.

La Figure 6.3 reprend les analyses du Chapitre 5. Les carottes du TMA ont été triées en fonction de leur score IHC (ici, on considère uniquement les scores 0, 1+, 2+ et 3+, et non les demi scores des cas équivoques). La Figure 6.3 montre pour chaque lame, la moyenne de la densité d'attachement sur la zone membranaire des cellules épithéliales néoplasiques calculée en prenant toutes les carottes d'un même score IHC.

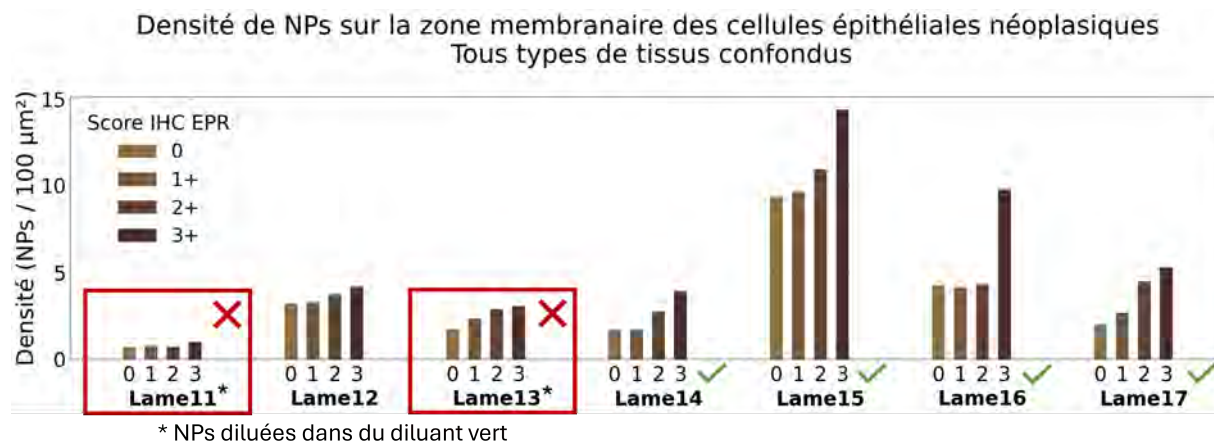


FIGURE 6.3 Densité moyenne d'attachement sur la zone membranaire des cellules épithéliales néoplasiques, par score IHC EPR pour toutes les carottes des différentes expériences. Les lames encadrées en rouges sont celles pour lesquelles aucun attachement n'a été observé sur les billes de contrôles.

Pour les lames 14, 15, 16 et 17, pour lesquelles un attachement spécifique aux billes de contrôle a été observé, on relève une forte densité pour les carottes très positives (3+) et une densité plus faible pour les carottes négatives (0). Pour les lames 15 et 17, on observe même une augmentation progressive de la densité à mesure que le score IHC augmente. Cette forte densité, accompagnée d'un contraste net entre les carottes positives et négatives, témoigne d'une fonctionnalisation réussie : les NPs se fixent bien sur le TMA, et de préférence sur les carottes positives.

En revanche, la lame 11 présente des densités très faibles, même pour les carottes positives. L'absence d'attachement sur les billes de contrôle semble donc corrélée à un échec de la fonctionnalisation, ne permettant pas la fixation des NPs sur l'échantillon.

Cependant, la lame 13 montre des valeurs de densité similaires à celles de la lame 14, et l'on peut également y observer une augmentation de la densité en fonction du score IHC. Il est donc difficile d'affirmer que la fonctionnalisation des NPs de la lame 13 est un échec, si l'on considère que les densités observées pour la lame 14 sont caractéristiques d'une fonctionnalisation réussie.

Les lames 11 et 13 ont toutes deux été préparées en utilisant le diluant vert de Roche Diagnostics (Ventana Antibody Diluent with Casein, réf.06440002001) comme tampon de dilution des NPs avant incubation. Il est donc possible que l'absence d'attachement sur les billes de contrôle résulte d'une incompatibilité entre les billes et le diluant vert, plutôt que d'un problème de fonctionnalisation des NPs. Sur les TMA, l'utilisation du diluant vert semble limiter l'attachement des NPs. Il faut alors utiliser des concentrations plus élevées pour observer un attachement plus fort sur les carottes positives.

Les billes de contrôles semblent être un outil de contrôle qualité intéressant pour évaluer rapidement si la fonctionnalisation des NPs est réussie. Dans le contexte du développement d'une technique innovante dont les résultats peuvent être influencés par de nombreuses variables, il est particulièrement utile de disposer de moyens de contrôle permettant de détecter d'éventuels dysfonctionnements. Toutefois, il est indispensable de rester critique avec ce nouvel outil encore peu connu et dont le comportement semble varier en fonction des conditions d'incubation.

6.2.2 Vers un contrôle quantitatif

Il est possible de pousser ce contrôle de qualité plus loin en le rendant quantitatif. Pour chaque bille à analyser, une pile d'images est acquise en microscopie par réflexion, en suivant

la méthodologie décrite à la section 3.2.3. Le modèle de détection des NPs est ensuite appliqué à chaque tranche de l'image dans Dragonfly, et le centre de masse tridimensionnel de chaque NP est extrait. Cela permet de pouvoir visualiser les NPs fixées sur une bille, comme le montre la Figure 6.4.

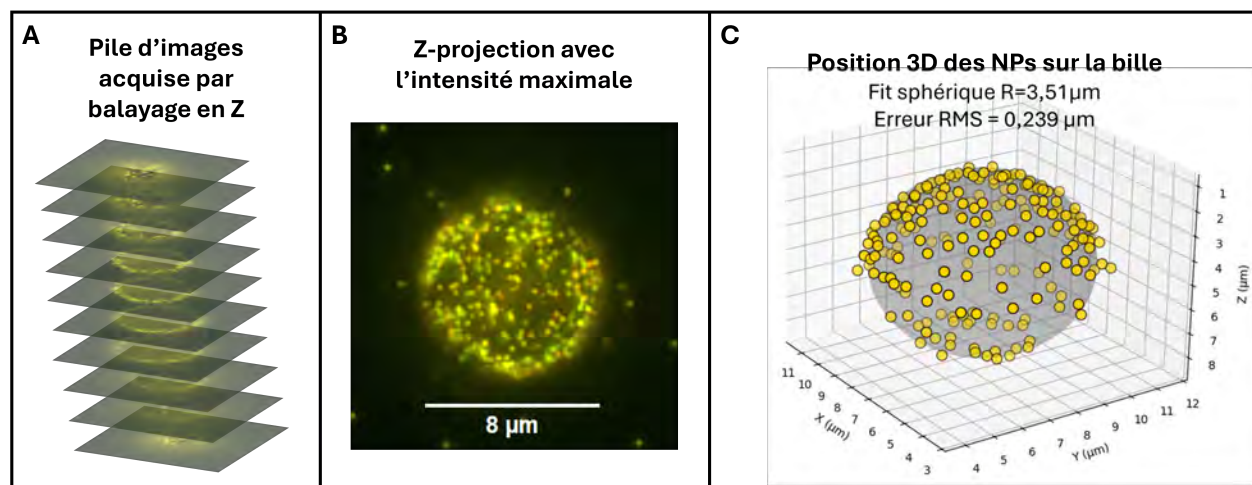


FIGURE 6.4 Quantification des NPs fixées sur une bille de contrôle : de la pile d'images acquise en microscopie à réflexion à une représentation 3D de la bille recouverte de NPs. La sphère grise représente la position de la bille ajustée par interpolation à partir des positions des NPs présentes à sa surface. La bille peut sembler aplatie à cause d'un contrôle imparfait du pas de balayage manuel en Z.

Pour chaque lame, 5 billes positives ont été imagées afin de quantifier le nombre de NPs fixées à leur surface. L'analyse de ces billes a permis de tracer un graphique de Levey-Jennings présenté dans la Figure 6.5.

Le travail effectué ici est peu représentatif en raison du faible nombre de billes analysées pour chaque lame. De plus, les billes imagées ayant été sélectionnées par l'opérateur, un biais évident peut être déploré. Les résultats préliminaires présentés illustrent néanmoins comment le modèle de détection des NPs pourrait être utilisé pour réaliser une analyse tridimensionnelle permettant de contrôler quantitativement la qualité des expériences réalisées au cours du temps. Dans la Figure 6.5, la moyenne du nombre de NPs détectées sur les lames 14, 15, 16 et 17 (lames pour lesquelles l'attachement sur les billes de contrôle s'est montré satisfaisant), est utilisée comme cible d'attachement souhaité. On peut voir que le nombre de NPs détectées par bille reste dans un intervalle de cette cible $\pm 2\sigma$. Au contraire, le nombre de NPs détectées sur les billes des lames 11 et 13 (pour lesquelles aucun attachement n'a été observé) est hors de cet intervalle. Idéalement, cette cible et son intervalle devraient être établis en

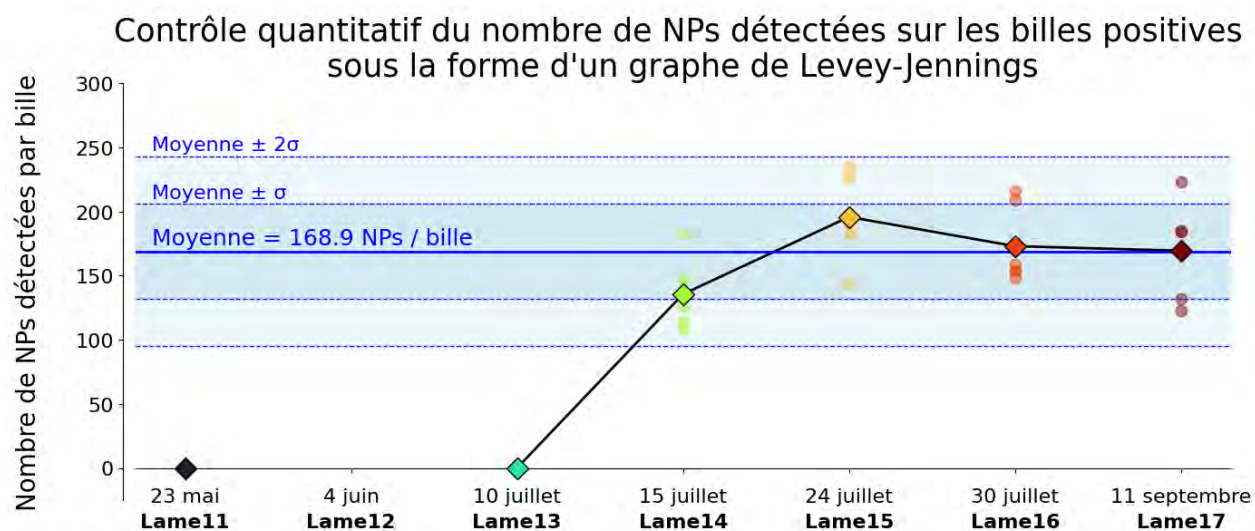


FIGURE 6.5 Contrôle quantitatif de l'attachement des NPs sur les billes de contrôle au cours des expériences réalisées.

utilisant plus d'expériences et de billes analysées, que l'on considère comme le résultat à obtenir.

Avec un système d'acquisition d'images automatisé capable de scanner l'échantillon en Z, il serait possible de généraliser la méthodologie présentée et de produire des résultats basés sur un nombre de billes analysées plus conséquent.

6.3 Vers la caractérisation de la plage dynamique de la méthode IP avec les calibrateurs

Après avoir réussi à obtenir un marquage spécifique des billes de contrôle, un travail a été effectué avec les calibrateurs. L'objectif est de parvenir à tracer la courbe montrant la réponse dynamique de la méthode IP. Ensuite, le protocole de la méthode IP pourrait être optimisé pour améliorer les performances de la technique pour un cas spécifique de détection. Par exemple, pour la détection de HER2-low, il serait souhaitable d'obtenir une limite de détection la plus basse possible et de s'assurer que les faibles variations de concentration en protéine soient perceptibles.

6.3.1 Réponse dynamique de la coloration IHC

Dans un premier temps, les calibrateurs ont été colorés par IHC. La procédure a été réalisée avec l'aide de la plateforme d'histologie de l'IRIC. La lame a ensuite été scannée en champ

clair avec un scanner Hamamatsu NanoZoomer. Les images obtenues sont présentées dans la Figure 6.6. On peut voir que l'intensité de la coloration brune croît avec les niveaux de concentration en protéine. Visuellement, la limite de détection semble se trouver proche du niveau 5. Certaines billes exposent une faible coloration, mais la distinction entre positif et négatif est subjective.

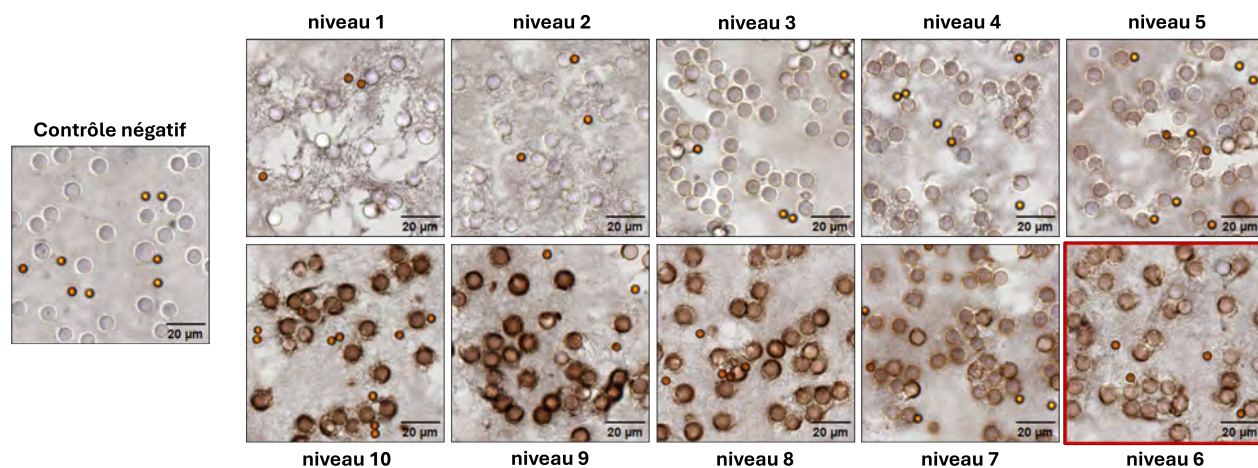


FIGURE 6.6 Billes de contrôle du calibrateur colorées par IHC. La concentration de protéine à la surface des billes est croissante du niveau 1 vers le niveau 10. Le contrôle négatif est composé de billes sans protéine.

Les images acquises ont été converties en niveaux de gris, puis segmentées manuellement sur Dragonfly afin d'isoler les billes. L'intensité moyenne de chaque bille a ensuite été mesurée. Pour chaque niveau de concentration, une trentaine de billes ont été analysées. Les résultats présentés dans la Figure 6.7 montrent que l'intensité en fonction de la concentration suit la tendance logarithmique attendue.

La limite de détection (LOD) a été calculée selon la méthodologie décrite par E. Torlakovic et al. [76]. La moyenne et l'écart-type de l'intensité des billes de contrôle négatives sont mesurés, puis un seuil correspondant à la moyenne à laquelle on soustrait trois fois l'écart-type est appliqué. En deçà de ce seuil, la coloration n'est pas suffisante pour être considérée comme positive. Seul le sixième niveau dépasse ce seuil. Visuellement, c'est en effet à partir de ce niveau de concentration que la coloration devient suffisamment marquée pour être considérée comme positive. Les niveaux 4 et 5 présentent une coloration perceptible, mais trop faible pour être jugée positive, bien qu'elle se situe très près de la limite de détection.

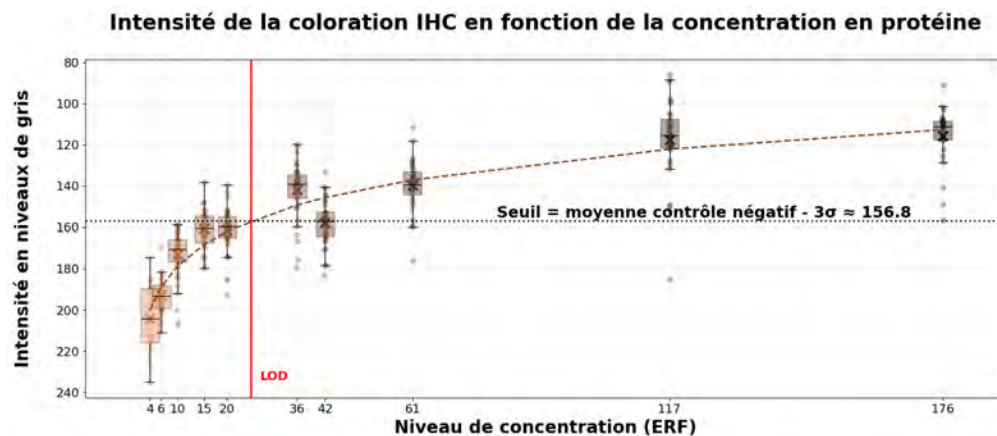


FIGURE 6.7 Réponse dynamique de la coloration IHC pour des concentrations croissantes de la protéine HER2.

6.3.2 Vers la réponse dynamique du marquage IP

Dans un second temps, il faut tracer une courbe semblable montrant le nombre de NPs par bille en fonction du niveau de concentration. Cependant, il n'a pas été possible d'observer un attachement spécifique des NPs sur les billes de calibrateurs. Différentes conditions de prétraitement des lames ont été testées afin de favoriser cet attachement :

- aucune étape de prétraitement ;
- même prétraitement que celui appliqué pour la lame colorée par IHC, réalisé à l'IRIC ;
- prétraitement de routine réalisé au laboratoire et utilisé pour les TMA.

Aucune de ces conditions n'a permis d'observer un attachement des NPs sur les billes qui restaient systématiquement vides. La piste privilégiée pour expliquer cette absence d'attachement est un déparaffinage incomplet. En effet, même sur les images issues de l'IHC, un substrat résiduel est observable. Il pourrait s'agir de restes de la matrice de support des billes, non totalement éliminée lors du déparaffinage. Cela ne semble pas poser de problème pour l'IHC, qui nécessite des sites de fixation de petite taille, les anticorps étant les seuls à devoir se lier à la protéine cible. En revanche, pour l'IP, les NPs, de taille plus importante, pourraient nécessiter des sites de fixation plus larges, ce qui limiterait leur ancrage dans les conditions actuelles.

Des efforts continuent d'être réalisés pour parvenir à obtenir un attachement spécifique sur les calibrateurs et tracer la courbe de réponse dynamique de l'IP qui pourrait avoir un impact important dans le développement de cette technique.

CHAPITRE 7 DISCUSSION

7.1 Outils de segmentation

SegNPs

Le modèle de détection des NPs développé a montré d'excellentes performances, largement supérieures aux limites de l'algorithme FindMaxima. Il convient toutefois de souligner que la qualité de la segmentation dépend fortement de celle des images fournies. Le soin apporté à l'acquisition des images et à la conduite des sessions d'imagerie a permis d'obtenir des images de bonne qualité, propices à une segmentation fiable et à une analyse précise.

Pour renforcer encore la rigueur et améliorer les performances du modèle, il serait intéressant d'augmenter la quantité de données d'entraînement et, surtout, de varier les hyperparamètres afin d'optimiser le modèle.

TissuDetection

Le modèle de détection du tissu requiert peu de ressources puisqu'un entraînement sur seulement dix images pourrait être suffisant. Pourtant, son impact sur la méthodologie est particulièrement significatif. Le masque de tissu constitue en effet la base de l'approche et permet notamment d'éviter facilement que le modèle de détection des cellules identifie des noyaux situés en dehors du tissu, par exemple dans les trous qui peuvent y être présents.

CellViT

La détection et la classification des cellules constituent un aspect central de la méthodologie développée. Toutefois, la classification n'a pas pu être évaluée quantitativement. Bien que les résultats globaux soient satisfaisants, la fiabilité de la classification cellulaire est cruciale. Il serait donc pertinent d'y consacrer davantage d'efforts afin de s'assurer de la robustesse de CellViT. Une piste d'amélioration serait d'impliquer plus étroitement les pathologistes et de réentraîner le modèle sur les images issues de notre propre jeu de données.

Dans le cadre de l'évaluation de CellViT, certaines observations ont été faites concernant l'impact de la variabilité des images acquises sur les résultats de segmentation.

Variabilité entre systèmes d’acquisition

Le premier point étudié concerne l’imagerie d’une même lame à l’aide de systèmes d’acquisition différents. La lame 16 a été scannée selon trois modalités :

- au laboratoire, avec le protocole d’imagerie présenté dans ce mémoire,
- à l’IRIC, avec deux scanners distincts : le VS200 de Evident Scientific et le NanoZoomer de Hamamatsu.

La Figure 7.1 illustre une même carotte imagée avec ces trois systèmes, puis segmentée à l’aide de CellViT. La première observation évidente concerne la différence marquée de balance des blancs entre les systèmes, entraînant des rendus visuels très différents. En conséquence, les résultats de segmentation varient également d’un système à l’autre.

Variabilité inter-expériences pour un même système d’acquisition

Le second point examiné concerne la variabilité entre des images acquises avec un même système d’imagerie, mais lors d’expériences différentes. La Figure 7.2 présente des images d’une même carotte sur les trois lames analysées et imagées au laboratoire. Même si les acquisitions ont été réalisées dans des conditions similaires (même système optique, mêmes paramètres d’acquisition), l’aspect visuel des images varie. Ces différences proviennent certainement de variations expérimentales minimales, inhérentes aux manipulations. Par exemple, une lame laissée une seconde de plus dans le bain d’éosine présentera une coloration plus intense. L’application de CellViT à ces différentes images conduit à des résultats de segmentation variables. Bien que les coupes ne contiennent pas exactement les mêmes cellules, on observe des différences notables.

Cette étude met en évidence la variabilité des sorties de CellViT. Il est difficile d’identifier précisément les paramètres les plus importants pour CellViT qui mènent à cette variabilité de segmentation, mais la cause la plus probable semble être l’absence de normalisation des images. La mise en place d’une étape de normalisation apparaît donc indispensable. Une approche méthodique consisterait à tester différentes méthodes de normalisation, puis à valider, avec l’aide d’un pathologiste, celle qui permet d’obtenir les résultats de segmentation les plus satisfaisants avec CellViT.

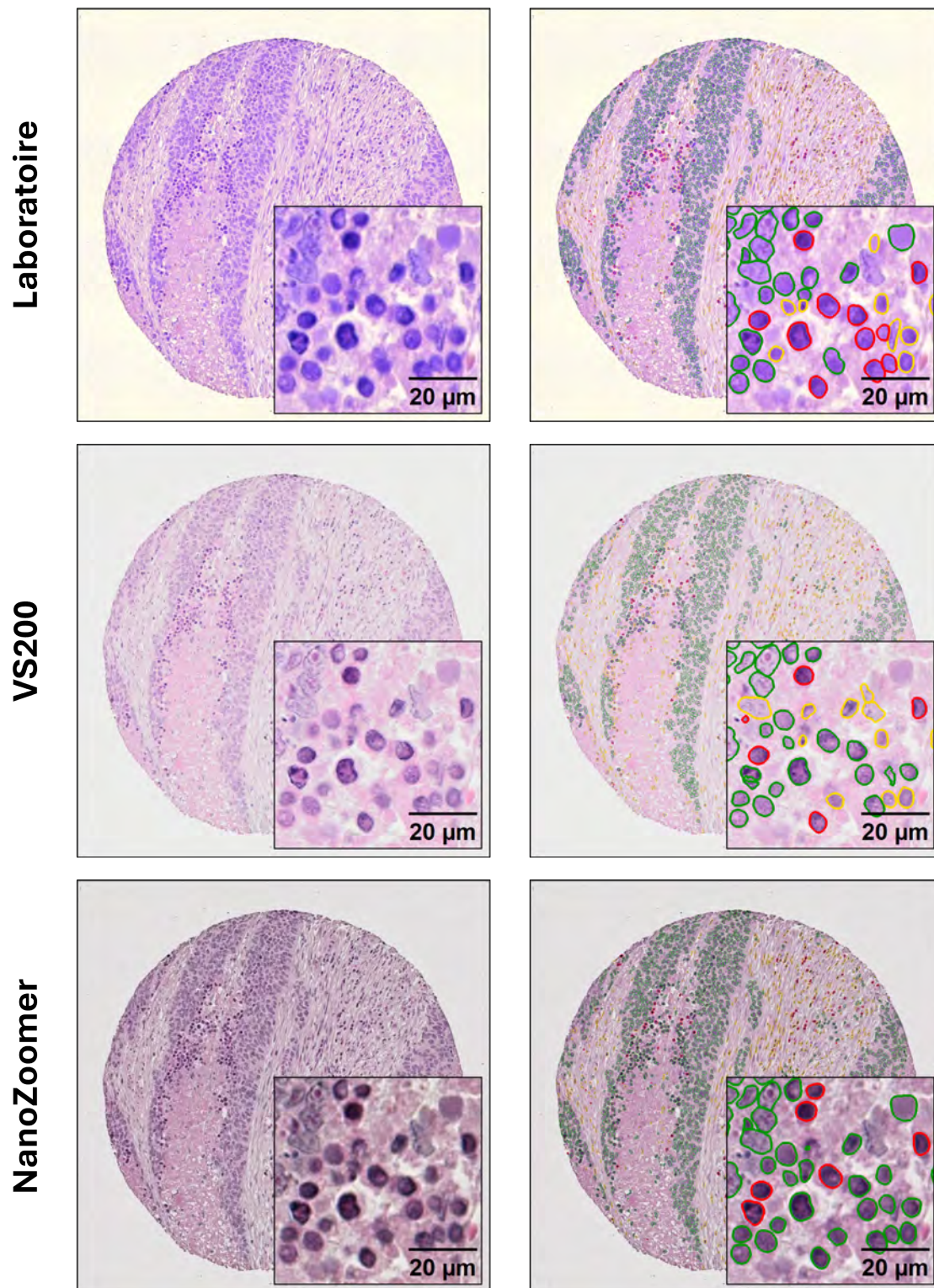
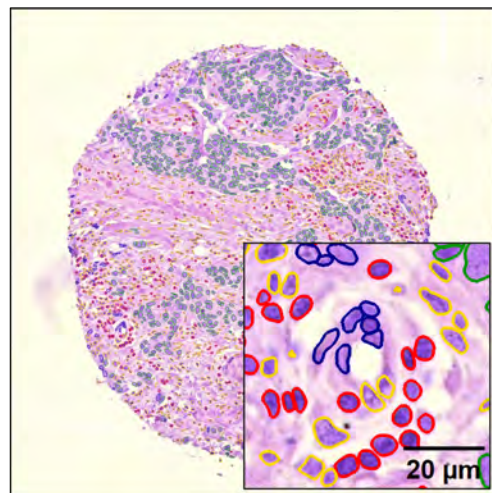
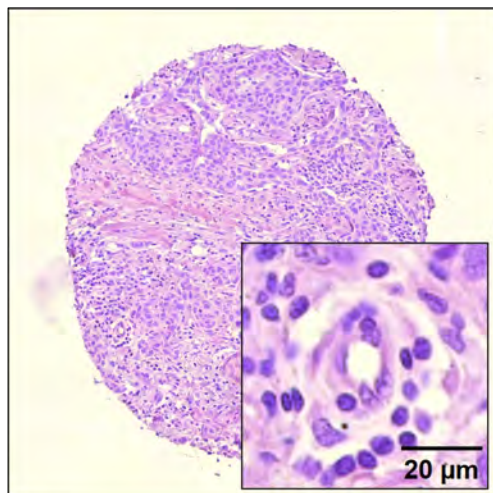
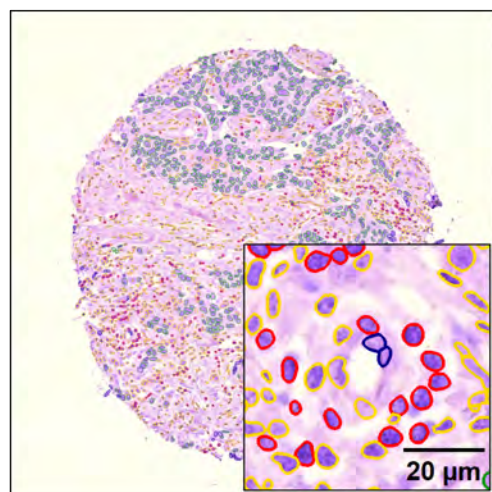
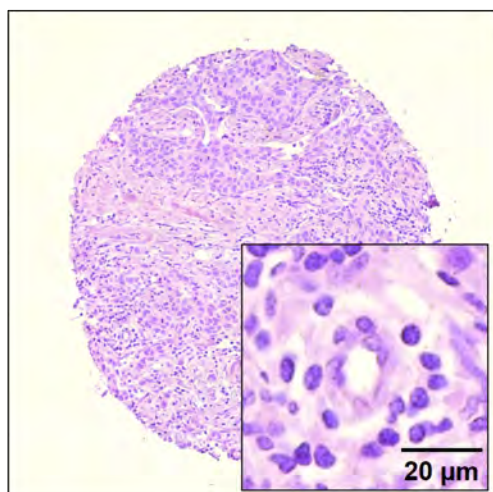


FIGURE 7.1 Variabilité des images acquises (à gauche) et des résultats de segmentation (à droite) en fonction des systèmes d'acquisition.

Lame 12



Lame 13



Lame 14

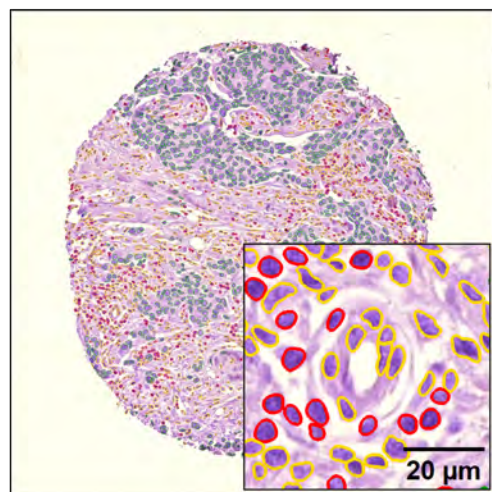
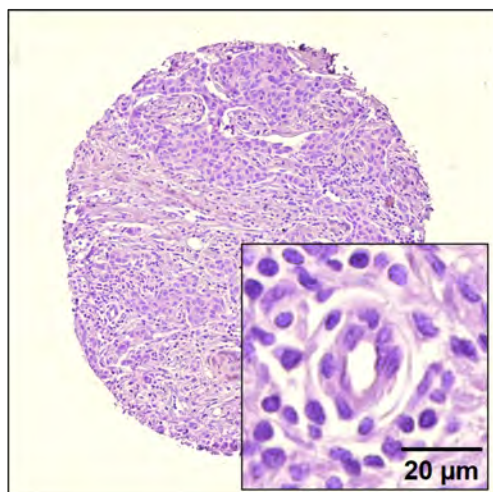


FIGURE 7.2 Variabilité des images acquises (à gauche) et des résultats de segmentation (à droite) en fonction des lames.

7.2 Gestion des agrégats

Dans le flux de travail proposé, les agrégats sont systématiquement exclus de l'analyse après avoir été identifiés et quantifiés par le taux d'agrégation. Ce choix a été fait dans un contexte où l'agrégation constituait un problème majeur du marquage IP. En raison de l'utilisation de la PEGylation comme méthode de fonctionnalisation, les premiers échantillons étudiés présentaient un grand nombre d'agrégats, considérés comme un signe de mauvaise fonctionnalisation.

Aujourd'hui, l'utilisation du calixarène a permis de réduire considérablement ce problème, voire de le faire disparaître complètement. Les agrégats détectés sont généralement causés par une très forte densité de NPs qui se touchent les unes les autres, et que le modèle SegNPs n'est donc pas en mesure de séparer. Ce phénomène a été particulièrement observé pour les CMA, où l'attachement est souvent plus important que pour les TMA et pour lesquels le marquage membranaire devient donc une ligne lumineuse continue.

Dans ce contexte particulier, il serait pertinent de compter les NPs détectées comme des agrégats en tant que NPs individuelles. Pour cela, une approche intéressante consisterait à considérer qu'au-delà d'un certain seuil de taille, si une NP détectée présente une forme allongée, signe d'une ligne de NPs alignées plutôt que d'un agrégat circulaire, on pourrait estimer le nombre de NPs contribuant à cette ligne en divisant l'aire de l'agrégat par l'aire moyenne d'une NP.

7.3 Réflexion sur les métriques

7.3.1 Importance de l'aspect qualitatif

Bien que ce travail repose principalement sur l'aspect quantitatif du marquage IP, notamment grâce à un travail approfondi mené autour des métriques, les discussions avec les pathologistes ont souvent remis l'aspect qualitatif du marquage au centre de l'attention. En effet, il est essentiel pour eux de retrouver, dans le marquage IP, le motif membranaire caractéristique observé en IHC. La protéine HER2 étant une protéine membranaire, il est donc indispensable que le marquage obtenu reflète cette localisation.

Par ailleurs, pour susciter l'intérêt des pathologistes, il semble important de rapprocher le marquage IP du marquage IHC dans un premier temps, afin de démontrer que les deux permettent d'observer et d'évaluer les mêmes structures. Ce n'est qu'ensuite que l'on pourra mettre en avant la valeur ajoutée du marquage IP, en montrant par exemple qu'il permet de

distinguer des cas de faible expression protéique.

De plus, l'aspect qualitatif est resté un critère primordial tout au long de ce projet. Plus qu'une densité élevée de NPs, c'est la présence d'un marquage membranaire sur les tissus qui a été considérée comme un signe de réussite. Il serait donc pertinent de développer une métrique quantitative capable de rendre compte de la présence de motifs membranaires dans l'image.

7.3.2 Non exhaustivité des métriques

Les métriques présentées dans ce travail constituent avant tout des propositions. Elles permettent de caractériser les images avec précision en s'intéressant à différents aspects du marquage et sont indispensables à l'optimisation de la méthode IP. C'est notamment grâce aux métriques décrivant l'attachement sur les noyaux que l'on a pu prendre conscience du phénomène récurrent d'attachement sur les noyaux, et que l'on peut maintenant s'interroger sur l'affinité entre les NPs et les noyaux. Toutefois, le cœur du travail réalisé réside davantage dans la méthodologie employée pour extraire les données que dans la manière de les manipuler.

De nombreuses approches pourraient encore être envisagées pour exploiter ces données : par exemple, calculer différents rapports à partir des densités, ou encore compter le nombre de NPs par cellule de manière individuelle plutôt que d'en faire la moyenne, afin d'analyser la distribution statistique des NPs au sein des cellules.

Il reste encore de nombreuses pistes à explorer, notamment en utilisant des données issues de biobanques, une fois le marquage IP pleinement optimisé. Cela permettra d'étudier les corrélations possibles entre les différentes métriques et des paramètres cliniques plus avancés que le score IHC, tels que la réponse au traitement. Cela permettrait de déterminer si le marquage IP présente une réelle valeur ajoutée par rapport à l'IHC.

7.4 Réflexion sur les produits BCS

7.4.1 Limites liées à l'utilisation d'un produit commercial

Les produits BCS ont représenté une excellente alternative pour travailler avec la méthode IP sans avoir recours à des échantillons biologiques, souvent difficiles à obtenir ou impliquant certaines contraintes liées à la culture cellulaire. Toutefois, le fait qu'il s'agisse de produits

commerciaux n'a pas permis de lever toutes les incertitudes. En particulier, l'absence d'information sur la séquence peptidique présente à la surface des billes a compliqué le choix du clone approprié.

De plus, il reste inexpliqué qu'un attachement ait pu être observé sur les billes de contrôle, même sans étape de déparaffinage, alors que pour les billes des calibrateurs, les résultats sont restés non concluants, que le déparaffinage soit réalisé ou non, et ce malgré l'essai de différentes méthodes. Le manque de transparence sur la composition exacte des produits utilisés rend le travail d'interprétation plus difficile.

7.4.2 Approche future pour l'évaluation de la plage dynamique

Néanmoins, l'idée d'utiliser ce type d'échantillon, présentant plusieurs niveaux de protéines connus, demeure très intéressante pour établir la plage dynamique de la méthode IP. Étant donné que cette approche n'a pas donné de résultats concluants et que, l'aspect tridimensionnel des billes complexifie l'analyse sans réelle valeur ajoutée, il serait pertinent d'envisager une autre stratégie. Une piste prometteuse serait de développer une sorte de test ELISA optique, consistant à déposer la protéine d'intérêt de manière contrôlée sur une lame de verre compatible avec la microscopie. Ainsi, on pourrait imager et quantifier les NPs se liant à des puits recouverts de protéine à différentes densités.

CHAPITRE 8 CONCLUSION

8.1 Synthèse des travaux

Ce projet a permis de développer avec succès une méthodologie complète schématisée dans la Figure 8.1. Cette méthodologie permet l'analyse des biopsies marquées par immunoplasmonique. Un protocole d'acquisition d'images dans les deux modalités d'illumination a été mis en place afin d'assurer la qualité et la standardisation des images obtenues. Par la suite, différents outils de segmentation ont été développés et utilisés pour proposer une détection des NPs plus robuste que celle offerte par la fonction *Find Maxima*, tout en exploitant au mieux le contexte biologique. Ces outils ont été appliqués aux images acquises afin de détecter et de classer les NPs selon la structure sur laquelle elles se sont fixées. Cette approche a permis de générer des métriques décrivant chaque échantillon, à la fois du point de vue qualitatif, en s'intéressant notamment à l'agrégation et à l'attachement non spécifique, et du point de vue de l'expression protéique, à travers l'attachement des NPs aux régions tumorales.

Ce travail a permis de proposer une méthodologie adaptée à l'évaluation des tissus marqués par immunoplasmonique, tout en mettant à disposition des outils destinés aux utilisateurs de la méthode souhaitant approfondir leurs analyses à l'aide d'approches quantitatives. De plus, les différents jeux de données générés restent accessibles pour continuer le développement des outils de segmentation nécessaires à l'analyse des échantillons marqués par immunoplasmonique.

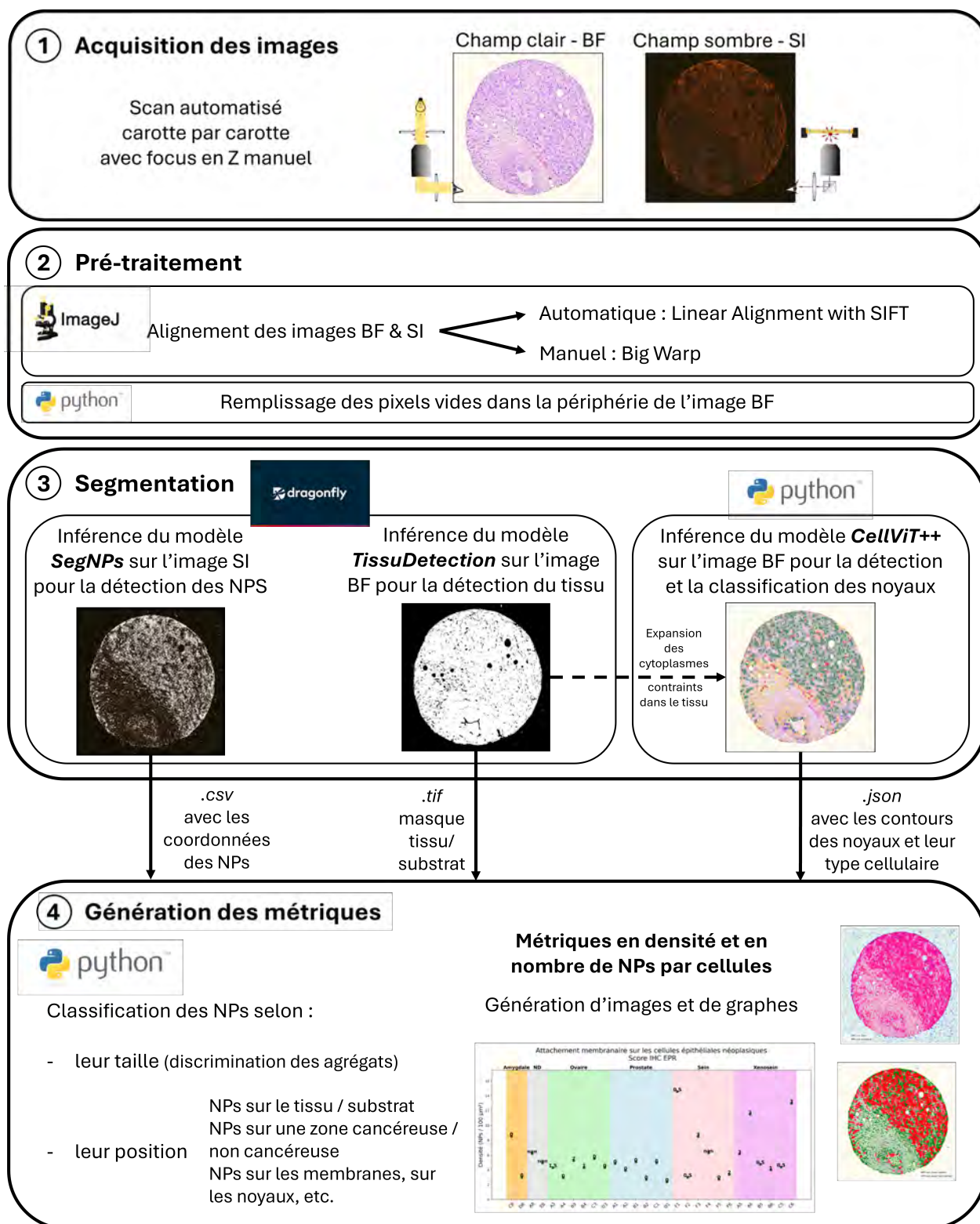


FIGURE 8.1 Schéma représentant la stratégie développée pour l'analyse des biopsies marquées par nanoparticules plasmoniques.

8.2 Limitations

8.2.1 Imagerie avec focus manuel

Tout d’abord, bien que l’acquisition des images soit passée d’un mode entièrement manuel au début du projet à une acquisition automatisée, effectuée carotte par carotte, certains aspects restent perfectibles. En particulier, la mise au point en Z doit encore être choisie manuellement par l’opérateur, qui ne peut sélectionner qu’un seul plan focal par carotte. Bien que les tissus soient globalement plats, l’intégration d’un autofocus capable d’ajuster la mise au point à partir de plusieurs points par carotte permettrait d’améliorer la qualité des images acquises. Cela faciliterait notamment la mise au point précise sur le substrat, actuellement difficile à obtenir avec la configuration existante. De plus, l’utilisation d’un autofocus permettrait de finaliser l’automatisation du processus d’acquisition, en supprimant la nécessité pour l’opérateur d’effectuer manuellement la mise au point en Z pour chaque carotte imagée.

8.2.2 Manque d’automatisation

Dans la continuité du point précédent, la méthodologie présentée souffre encore d’un manque d’automatisation. Outre l’acquisition d’images semi-automatique, le prétraitement des images (qui inclue l’alignement des deux modalités d’illumination ainsi que le redimensionnement autour de chaque carotte) est souvent réalisé manuellement, faute d’outils automatiques suffisamment robustes. Un travail préalable visant à augmenter le contraste des images pourrait favoriser un meilleur alignement à l’aide du plugin *Linear Alignment with SIFT*, mais une solution pleinement satisfaisante reste encore à identifier.

8.2.3 Faible intégration

Le principal défaut de la méthodologie proposée réside dans son manque d’intégration. Les images sont d’abord acquises, puis le prétraitement est effectué à l’aide d’ImageJ et d’un code Python. Par la suite, les images doivent être importées dans Dragonfly pour la segmentation, ainsi que dans un autre code Python pour l’inférence avec CellViT. Enfin, les données sont exportées sous forme de fichiers .tif, .csv et .json, qui doivent être chargés dans un dernier script Python afin d’extraire les métriques finales. Bien que ce flux de travail soit fonctionnel, il manque de cohérence et d’intuitivité. Il serait donc pertinent de regrouper l’ensemble des étapes au sein d’un seul outil Python, afin d’en faciliter l’utilisation par les différents utilisateurs.

8.3 Améliorations futures

8.3.1 Proposer une solution entièrement automatisée et intégrée

Naturellement, la première amélioration proposée consisterait à répondre aux limites précédemment évoquées en finalisant l'automatisation de la méthode au sein d'un outil Python unique. Une telle intégration faciliterait considérablement l'analyse d'images et permettrait d'obtenir plus aisément les résultats des différentes étapes d'analyse. Actuellement, l'analyse des données constitue une étape à part entière pour les utilisateurs de la méthode IP, souvent réduite à une interprétation qualitative faute d'expertise technique. Le développement d'un tel outil représenterait donc une avancée significative, à la fois utile pour les utilisateurs et impactante en raison des résultats générés.

8.3.2 Ajouter la classification des NPs selon leur couleur

La méthodologie d'analyse permet la détection des NPs, quelle que soit leur couleur. Avec le développement des projets de multiplexage proposant d'utiliser plusieurs types de NPs fonctionnalisées avec différents anticorps pour cibler plusieurs protéines au sein d'un même échantillon, il serait pertinent d'ajouter une étape de classification des NPs selon leur couleur. Cela permettrait de prendre aussi en charge l'analyse des expériences de multiplexage.

8.3.3 Intégrer une métrique pour décrire la présence de motifs membranaires

Comme discuté précédemment, l'aspect qualitatif du marquage membranaire est important. Le développement de nouvelles métriques permettant de quantifier la présence et la complétude du marquage membranaire serait particulièrement intéressant. Ces métriques pourraient, par exemple, établir une relation entre la longueur de chaque membrane et le nombre de NPs alignées le long de celle-ci, afin d'évaluer dans quelle mesure le marquage membranaire est complet. Une telle approche permettrait d'analyser la complétude du marquage, conformément aux recommandations de l'ASCO/CAP.

RÉFÉRENCES

- [1] Société canadienne du cancer. (2024) Tableau de bord des statistiques canadiennes sur le cancer (tbscc). Consulté le 1 novembre 2025. [En ligne]. Disponible : <https://cancerstats.ca/?la=fr>
- [2] —, “Statistiques sur le cancer du sein,” Site web – Cancer.ca, 2024, dernière révision médicale : mai 2024. Consulté le 01 octobre 2025. [En ligne]. Disponible : <https://cancer.ca/fr/cancer-information/cancer-types/breast/statistics>
- [3] G. K. Malhotra *et al.*, “Histological, molecular and functional subtypes of breast cancers,” *Cancer Biology & Therapy*, vol. 10, n°. 10, p. 955–960, nov 2010.
- [4] D. Vuong *et al.*, “Molecular classification of breast cancer,” *Virchows Archiv*, vol. 465, n°. 1, p. 1–14, juill. 2014.
- [5] R. G. D. Nascimento et K. M. Otoni, “Histological and molecular classification of breast cancer : what do we know ?” *Mastology*, vol. 30, p. e20200024, 2020.
- [6] B. Weigelt, F. C. Geyer et J. S. Reis-Filho, “Histological types of breast cancer : How special are they ?” *Molecular Oncology*, vol. 4, n°. 3, p. 192–208, juin 2010.
- [7] Société canadienne du cancer, “Qu’est-ce que le cancer du sein ?” Site web – Cancer.ca, 2025, consulté le 01 octobre 2025. [En ligne]. Disponible : <https://cancer.ca/fr/cancer-information/cancer-types/breast/what-is-breast-cancer>
- [8] G. Menon, F. M. Alkabban et T. Ferguson, “Breast cancer,” dans *StatPearls [Internet]*. Treasure Island (FL) : StatPearls Publishing, 2025, [Updated 2024 Feb 25]. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/books/NBK482286/>
- [9] Société canadienne du cancer, “Types de cancer,” Site web – Cancer.ca, 2025, consulté le 1er octobre 2025. [En ligne]. Disponible : <https://cancer.ca/fr/cancer-information/what-is-cancer/types-of-cancer>
- [10] O. mondiale de la santé, édit., *Breast tumours*, 5^e éd., ser. World Health Organization Classification of Tumours. Geneva : OMS, 2019, vol. 2.
- [11] A. Bhushan, A. Gonsalves et J. U. Menon, “Current state of breast cancer diagnosis, treatment, and theranostics,” *Pharmaceutics*, vol. 13, n°. 5, p. 723, mai 2021.
- [12] E. S. McDonald *et al.*, “Clinical diagnosis and management of breast cancer,” *J Nucl Med*, vol. 57, n°. Supplement 1, p. 9S–16S, févr. 2016.
- [13] U. o. L. The Histology Guide, Faculty of Biological Sciences, “What is h&e ?” https://www.histology.leeds.ac.uk/what-is-histology/H_and_E.php, —, consulté le 01 octobre 2025.

- [14] S.-W. Kim, J. Roh et C.-S. Park, “Immunohistochemistry for pathologists : Protocols, pitfalls, and tips,” *Journal of Pathology and Translational Medicine*, vol. 50, n°. 6, p. 411–418, nov. 2016.
- [15] S. Magaki *et al.*, “An introduction to the performance of immunohistochemistry,” dans *Biobanking*, ser. Methods in Molecular Biology, W. H. Yong, édit. New York, NY : Springer New York, 2019, vol. 1897, p. 289–298.
- [16] D. Y. Mebratie et G. G. Dagnaw, “Review of immunohistochemistry techniques : Applications, current status, and future perspectives,” *Seminars in Diagnostic Pathology*, vol. 41, n°. 3, p. 154–160, mai 2024.
- [17] K. H. Allison et et al., “Estrogen and progesterone receptor testing in breast cancer : Asco/cap guideline update,” *Journal of Clinical Oncology*, vol. 38, n°. 12, p. 1346–1366, avr. 2020.
- [18] A. C. Wolff et et al., “Human epidermal growth factor receptor 2 testing in breast cancer : Asco–college of american pathologists guideline update,” *Journal of Clinical Oncology*, vol. 41, n°. 22, p. 3867–3872, août 2023.
- [19] S. A. Bogen, “A root cause analysis into the high error rate in clinical immunohistochemistry,” *Applied Immunohistochemistry & Molecular Morphology*, vol. 27, n°. 5, p. 329–338, mai 2019.
- [20] D. G. Hicks et R. R. Tubbs, “Assessment of the her2 status in breast cancer by fluorescence in situ hybridization : a technical review with interpretive guidelines,” *Human Pathology*, vol. 36, n°. 3, p. 250–261, mars 2005.
- [21] M. Drăgănescu et C. Carmocan, “Hormone therapy in breast cancer,” *Chirurgia*, vol. 112, n°. 4, p. 413, 2017.
- [22] S. Ahn *et al.*, “Her2 status in breast cancer : changes in guidelines and complicating factors for interpretation,” *J Pathol Transl Med*, vol. 54, n°. 1, p. 34–44, jan 2020.
- [23] S. Ahuja, A. A. Khan et S. Zaheer, “Understanding the spectrum of her2 status in breast cancer : From her2-positive to ultra-low her2,” *Pathology - Research and Practice*, vol. 262, p. 155550, oct 2024.
- [24] S. M. Swain, M. Shastri et E. Hamilton, “Targeting her2-positive breast cancer : advances and future directions,” *Nat Rev Drug Discov*, vol. 22, n°. 2, p. 101–126, feb 2023.
- [25] J. Albanell et J. Baselga, “Trastuzumab, a humanized anti-her2 monoclonal antibody, for the treatment of breast cancer,” *Drugs Today (Barc)*, vol. 35, n°. 12, p. 931–946, dec 1999.
- [26] S. Modi *et al.*, “Trastuzumab deruxtecan in previously treated her2-low advanced breast cancer,” *N Engl J Med*, vol. 387, n°. 1, p. 9–20, jul 2022.

- [27] A. Dilawari *et al.*, “Us food and drug administration approval summary : Trastuzumab deruxtecan for the treatment of adult patients with hormone receptor–positive, unresectable or metastatic human epidermal growth factor receptor 2–low or human epidermal growth factor receptor 2–ultralow breast cancer,” *JCO*, vol. 43, n°. 26, p. 2942–2951, sep 2025.
- [28] X. Baez-Navarro *et al.*, “Interobserver variation in the assessment of immunohistochemistry expression levels in her2-negative breast cancer : Can we improve the identification of low levels of her2 expression by adjusting the criteria ? an international interobserver study,” *Modern Pathology*, vol. 36, n°. 1, p. 100009, jan 2023.
- [29] L. Wang, M. Hasanzadeh Kafshgari et M. Meunier, “Optical properties and applications of plasmonic-metal nanoparticles,” *Advanced Functional Materials*, vol. 30, n°. 51, p. 2005400, déc. 2020.
- [30] A. Agrawal *et al.*, “Localized surface plasmon resonance in semiconductor nanocrystals,” *Chemical Reviews*, vol. 118, n°. 6, p. 3121–3207, mars 2018.
- [31] G. Mie, “Beiträge zur optik trüber medien, speziell kolloidaler metallösungen,” *Annalen der Physik*, vol. 330, n°. 3, p. 377–445, janv. 1908.
- [32] A. C. Marques *et al.*, “Functionalizing nanoparticles with cancer-targeting antibodies : A comparison of strategies,” *Journal of Controlled Release*, vol. 320, p. 180–200, avr. 2020.
- [33] L. Shi *et al.*, “Effects of polyethylene glycol on the surface of nanoparticles for targeted drug delivery,” *Nanoscale*, vol. 13, n°. 24, p. 10 748–10 764, 2021.
- [34] A. Doi *et al.*, “Role of lipoic acid and gold nanomaterials and nanoclusters for biomedical and environmental applications,” *Materials Research Bulletin*, vol. 193, p. 113712, janv. 2026.
- [35] P. G. de Gennes, “Conformations of polymers attached to an interface,” *Macromolecules*, vol. 13, n°. 5, p. 1069–1075, sept. 1980.
- [36] J. V. Jokerst *et al.*, “Nanoparticle pegylation for imaging and therapy,” *Nanomedicine*, vol. 6, n°. 4, p. 715–728, juin 2011.
- [37] J. S. Suk *et al.*, “Pegylation as a strategy for improving nanoparticle-based drug and gene delivery,” *Advanced Drug Delivery Reviews*, vol. 99, p. 28–51, avr. 2016.
- [38] C. Courbon, “Optimisation et étude de vieillissement des nanoparticules d’or fonctionnalisées pour la détection du cancer du sein,” Mémoire de maîtrise, Polytechnique Montréal, 2025.
- [39] L. Troian-Gautier *et al.*, “Extremely robust and post-functionalizable gold nanoparticles coated with calix[4]arenes via metal–carbon bonds,” *Chemical Communications*, vol. 52,

- n°. 69, p. 10 493–10 496, 2016.
- [40] A. Mattiuzzi *et al.*, “Electrografting of calix[4]arene diazonium salts to form versatile robust platforms for spatially controlled surface functionalization,” *Nature Communications*, vol. 3, p. 1130, 2012.
 - [41] M. Retout *et al.*, “Ultrastable pegylated calixarene-coated gold nanoparticles with a tunable bioconjugation density for biosensing applications,” *Bioconjugate Chemistry*, vol. 32, n°. 2, p. 290–300, feb 2021.
 - [42] —, “Bifunctional calix[4]arene-coated gold nanoparticles for orthogonal conjugation,” *Langmuir*, vol. 38, n°. 29, p. 9301–9309, 2022.
 - [43] C. Darvriot *et al.*, “Multiplexed immunolabelling of cancer using bioconjugated plasmonic gold–silver alloy nanoparticles,” *Nanoscale Advances*, vol. 6, n°. 17, p. 4385–4393, 2024.
 - [44] C. Darvriot, “Quantitative and multiplexed protein detection using plasmonic nanoparticles : from bioconjugation to hyperspectral detection,” Thèse de doctorat en génie physique, Polytechnique Montréal, juin 2024.
 - [45] M. Qi *et al.*, “Cost-effective side-illumination darkfield nanoplasmonic marker microscopy,” *Analyst*, vol. 144, n°. 4, p. 1303–1308, 2019.
 - [46] S. Patskovsky et M. Meunier, “Reflected light microspectroscopy for single-nanoparticle biosensing,” *Journal of Biomedical Optics*, vol. 20, n°. 9, p. 097001, sept. 2015.
 - [47] S. Shafi et A. V. Parwani, “Artificial intelligence in diagnostic pathology,” *Diagnostic Pathology*, vol. 18, n°. 1, p. 109.
 - [48] O. Ronneberger, P. Fischer et T. Brox, “U-net : Convolutional networks for biomedical image segmentation,” *arXiv preprint arXiv :1505.04597*, 2015. [En ligne]. Disponible : <https://arxiv.org/abs/1505.04597>
 - [49] F. Hörst *et al.*, “Cellvit : Vision transformers for precise cell segmentation and classification,” *arXiv*, 2023.
 - [50] U. Schmidt *et al.*, “Cell detection with star-convex polygons,” dans *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, ser. Lecture Notes in Computer Science, A. F. Frangi *et al.*, édit., vol. 11071. Cham : Springer International Publishing, 2018, p. 265–273.
 - [51] M. Weigert et U. Schmidt, “Nuclei instance segmentation and classification in histopathology images with stardist,” dans *2022 IEEE International Symposium on Biomedical Imaging Challenges (ISBIC)*. Kolkata, India : IEEE, mars 2022, p. 1–4.
 - [52] C. Stringer *et al.*, “Cellpose : a generalist algorithm for cellular segmentation,” *Nature Methods*, vol. 18, n°. 1, p. 100–106, janv. 2021.

- [53] M. Pachitariu, M. Rariden et C. Stringer, “Cellpose-sam : superhuman generalization for cellular segmentation,” *bioRxiv*, mai 2025, [Preprint].
- [54] F. Hörst *et al.*, “Cellvit++ : Energy-efficient and adaptive cell segmentation and classification using foundation models,” *arXiv*, 2025.
- [55] Z. Zhou *et al.*, “Unet++ : A nested u-net architecture for medical image segmentation,” 2018. [En ligne]. Disponible : <https://arxiv.org/abs/1807.10165>
- [56] O. Oktay *et al.*, “Attention u-net : Learning where to look for the pancreas,” 2018. [En ligne]. Disponible : <https://arxiv.org/abs/1804.03999>
- [57] A. Hatamizadeh *et al.*, “Unetr : Transformers for 3d medical image segmentation,” 2021. [En ligne]. Disponible : <https://arxiv.org/abs/2103.10504>
- [58] A. Vaswani *et al.*, “Attention is all you need,” *arXiv preprint arXiv :1706.03762*, 2017. [En ligne]. Disponible : <https://arxiv.org/abs/1706.03762>
- [59] A. Dosovitskiy *et al.*, “An image is worth 16x16 words : Transformers for image recognition at scale,” *arXiv preprint arXiv :2010.11929*, 2020. [En ligne]. Disponible : <https://arxiv.org/abs/2010.11929>
- [60] P. Bankhead *et al.*, “Qupath : Open source software for digital pathology image analysis,” *Scientific Reports*, vol. 7, n^o. 1, p. 16878, déc. 2017.
- [61] A. Kirillov *et al.*, “Segment anything,” *arXiv*, 2023.
- [62] S. Graham *et al.*, “Hover-net : Simultaneous segmentation and classification of nuclei in multi-tissue histology images,” *Medical Image Analysis*, vol. 58, p. 101563, déc. 2019.
- [63] M. L. Chiu *et al.*, “Antibody structure and function : The basis for engineering therapeutics,” *Antibodies*, vol. 8, n^o. 4, p. 55, déc. 2019.
- [64] W. Choe, T. Durgannavar et S. Chung, “Fc-binding ligands of immunoglobulin g : An overview of high affinity proteins and peptides,” *Materials*, vol. 9, n^o. 12, p. 994, déc. 2016.
- [65] D. Rioux et M. Meunier, “Seeded growth synthesis of composition and size-controlled gold–silver alloy nanoparticles,” *The Journal of Physical Chemistry C*, vol. 119, n^o. 23, p. 13 160–13 168, juin 2015.
- [66] X. Wu *et al.*, “A breast cancer cell microarray (cma) as a rapid method to characterize candidate biomarkers,” *Cancer Biology & Therapy*, vol. 15, n^o. 12, p. 1593–1599, déc. 2014.
- [67] H. L. Fedor et A. M. De Marzo, “Practical methods for tissue microarray construction,” dans *Pancreatic Cancer*. New Jersey : Humana Press, 2004, vol. 103, p. 89–102.

- [68] M. Moutafi *et al.*, “Quantitative measurement of HER2 expression to subclassify ERBB2 unamplified breast cancer,” *Laboratory Investigation*, vol. 102, n°. 10, p. 1101–1108, oct. 2022.
- [69] E. Manion *et al.*, “A comparison of equivocal immunohistochemical results with anti-her2/neu antibodies a0485 and sp3 with corresponding fish results in routine clinical practice,” *American Journal of Clinical Pathology*, vol. 135, n°. 6, p. 845–851, juin 2011.
- [70] E. Lucas *et al.*, “Comparison of dako herceptest and ventana pathway anti-her2 (4b5) tests and their correlation with fluorescent in situ hybridization in breast carcinoma,” *Applied Immunohistochemistry & Molecular Morphology*, vol. 27, n°. 6, p. 403–409, juill. 2019.
- [71] A. Glazyrin *et al.*, “Direct detection of herceptin/trastuzumab binding on breast tissue sections,” *Journal of Histochemistry & Cytochemistry*, vol. 55, n°. 1, p. 25–33, janv. 2007.
- [72] D. E. Carvajal-Hausdorf *et al.*, “Measurement of domain-specific her2 (erbb2) expression may classify benefit from trastuzumab in breast cancer,” *JNCI : Journal of the National Cancer Institute*, vol. 107, n°. 8, août 2015.
- [73] M. A. Hempenius *et al.*, “Quantitative comparison of immunohistochemical her2-low detection in an interlaboratory study,” *Histopathology*, vol. 85, n°. 6, p. 920–928, déc. 2024.
- [74] S. R. Sompuram *et al.*, “Standardizing immunohistochemistry : A new reference control for detecting staining problems,” *Journal of Histochemistry & Cytochemistry*, vol. 63, n°. 9, p. 681–690, sept. 2015.
- [75] K. Vani *et al.*, “Levey-jennings analysis uncovers unsuspected causes of immunohistochemistry stain variability,” *Applied Immunohistochemistry & Molecular Morphology*, vol. 24, n°. 10, p. 688–694, nov. 2016.
- [76] E. E. Torlakovic *et al.*, “Development and validation of measurement traceability for in situ immunoassays,” *Clinical Chemistry*, vol. 67, n°. 5, p. 763–771, avr. 2021.
- [77] S. R. Sompuram *et al.*, “Quantitative comparison of pd-l1 ihc assays against nist standard reference material 1934,” *Modern Pathology*, vol. 35, n°. 3, p. 326–332, mars 2022.
- [78] S. Preibisch, S. Saalfeld et P. Tomancak, “Globally optimal stitching of tiled 3d microscopic image acquisitions,” *Bioinformatics*, vol. 25, n°. 11, p. 1463–1465, juin 2009.
- [79] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, n°. 2, p. 91–110, nov. 2004.
- [80] J. A. Bogovic *et al.*, “Robust registration of calcium images by learned contrast synthesis,” dans *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*.

IEEE, avr. 2016, p. 1123–1126.

- [81] C. T. C. Inc., “Dragonfly 3d world,” computer software. [En ligne]. Disponible : <https://dragonfly.comet.tech/>
- [82] J. Chen *et al.*, “Transunet : Transformers make strong encoders for medical image segmentation,” 2021. [En ligne]. Disponible : <https://arxiv.org/abs/2102.04306>