



HAL
open science

Signature pronostique dans le cancer du côlon par l'étude de l'expression ectopique de gènes spécifiques de tissus

Aurélien Spinelli

► To cite this version:

Aurélien Spinelli. Signature pronostique dans le cancer du côlon par l'étude de l'expression ectopique de gènes spécifiques de tissus. Médecine humaine et pathologie. 2024. dumas-04738592

HAL Id: dumas-04738592

<https://dumas.ccsd.cnrs.fr/dumas-04738592v1>

Submitted on 15 Oct 2024

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

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

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance.

La propriété intellectuelle du document reste entièrement celle du ou des auteurs. Les utilisateurs doivent respecter le droit d'auteur selon la législation en vigueur, et sont soumis aux règles habituelles du bon usage, comme pour les publications sur papier : respect des travaux originaux, citation, interdiction du pillage intellectuel, etc.

Il est mis à disposition de toute personne intéressée par l'intermédiaire de [l'archive ouverte DUMAS](#) (Dépôt Universitaire de Mémoires Après Soutenance).

Si vous désirez contacter son ou ses auteurs, nous vous invitons à consulter en ligne les annuaires de l'ordre des médecins, des pharmaciens et des sages-femmes.

Contact à la Bibliothèque universitaire de Médecine Pharmacie de Grenoble :

bump-theses@univ-grenoble-alpes.fr

Année : 2024

**SIGNATURE PRONOSTIQUE DANS LE CANCER DU CÔLON PAR
L'ÉTUDE DE L'EXPRESSION ECTOPIQUE DE GÈNES SPÉCIFIQUES
DE TISSUS**

THÈSE

POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

SPÉCIALITÉ : ANATOMIE ET CYTOLOGIE PATHOLOGIQUES

SOUTENUE PUBLIQUEMENT À LA FACULTÉ DE MÉDECINE DE GRENOBLE

Le 07/10/2024

Par M. Aurélien SPINELLI

[Données à caractère personnel]

DEVANT LE JURY COMPOSÉ DE :

Présidente du jury :

Mme le Pr Séverine VALMARY-DEGANO (co-directrice de thèse)

Membres :

M. le Dr Julio ABBA

Mme le Dr Ekaterina BOUROVA-FLIN (co-directrice de thèse)

M. le Pr Thomas DECAENS

Mme le Dr Marie-Hélène LAVERRIÈRE

L'UFR de Médecine de Grenoble n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions sont considérées comme propres à leurs auteurs.

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
MCF	ADAM Jean-François	Sciences de la rééducation et de réadaptation
PU-PH	ALBALADEJO Pierre	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH	ARVIEUX-BARTHÉLÉMY Catherine	Chirurgie viscérale et digestive
PU-PH	ATALLAH Ihab	Oto-rhyno-laryngologie
PU-PH	BAILLET Athan	Rhumatologie
MCU-PH	BAILLIEUL Sébastien	Physiologie
PU-PH	BARONE-ROCHETTE Gilles	Cardiologie
PR Attaché	BARTH Johannes	Chirurgie de l'épaule et du genou
PU-PH	BAYAT Sam	Physiologie
MCU-PH	BELLIER Alexandre	Anatomie
PR Ass.MG	BENDAMENE Farouk	Médecine Générale
PU-PH	BENHAMOU Pierre-Yves	Endocrinologie, diabète et maladies métaboliques
PU-PH	BERGER François	Biologie cellulaire
PU-PH	BÉTRY Cécile	Nutrition
MCU-PH	BIDART-COUTTON Marie	Biologie cellulaire
PU-PH	BIOULAC-ROGIER Stéphanie	Pédopsychiatrie ; addictologie
PU-PH	BLAISE Sophie	Chirurgie vasculaire ; médecine vasculaire
PU-PH	BOISSET Sandrine	Bactériologie-virologie ; Hygiène hospitalière
PU-PH émérite	BONAZ Bruno	Gastroentérologie ; hépatologie
PU-PH	BONNETERRE Vincent	Médecine et santé au travail
PU-PH	BOREL Anne-Laure	Nutrition
PU-PH	BOSSON Jean-Luc	Biostatistiques, informatique médicale et technologies de communication
MCU-PH émérite	BOTTARI Serge	Biologie cellulaire
PU-PH	BOUDISSA Mehdi	Chirurgie orthopédique et traumatologique
PU-PH	BOUILLET Laurence	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
PU-PH	BOUSSAT Bastien	Épidémiologie, économie de la santé et prévention
PU-PH	BOUZAT Pierre	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH émérite	BRAMBILLA Christian	Pneumologie
PU-PH émérite	BRAMBILLA Elisabeth	Anatomie et cytologie pathologiques
MCU-PH	BRENIER-PINCHART Marie Pierre	Parasitologie et mycologie
PU-PH	BRICAULT Ivan	Radiologie et imagerie médicale
PU-PH émérite	BRICHON Pierre-Yves	Chirurgie thoracique et cardiovasculaire
MCU-PH	BRIOT Raphaël	Thérapeutique-médecine de la douleur ; Addictologie
PU-PH émérite	CAHN Jean-Yves	Hématologie
PU-PH émérite	CARPENTIER Patrick	Chirurgie vasculaire, médecine vasculaire
PR Ass.MG	CARRILLO Yannick	Médecine Générale
PR Ass. Méd.	CASEZ Olivier	Neurologie
MCU-PH	CASPAR Yvan	Bactériologie-virologie ; hygiène hospitalière
PU-PH émérite	CESBRON Jean-Yves	Immunologie

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
PU-PH	CHABARDÈS Stephan	Neurochirurgie
PU-PH émérite	CHABRE Olivier	Endocrinologie, diabète et maladies métaboliques
PU-PH	CHAFFANJON Philippe	Anatomie
MCF Ass.MG	CHAMBOREDON Benoît	Médecine Générale
PU-PH	CHARLES Julie	Dermato-vénérologie
PU-PH	CHAVANON Olivier	Chirurgie thoracique et cardiovasculaire
MCU-PH	CHEVALLIER Marie	Pédiatrie
PU-PH	CHIQUET Christophe	Ophthalmologie
PU-PH	CHIRICA Mircea	Chirurgie viscérale et digestive
PU-PH	CINQUIN Philippe	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	CLAVARINO Giovanna	Immunologie
MCU-PH	CLIN CHERPEC Rita	Nutrition
PU-PH	COHEN Olivier	Histologie, embryologie et cytogénétique
PU-PH	COURVOISIER Aurélien	Chirurgie infantile
PU-PH	COUTTON Charles	Génétique
PU-PH	CRACOWSKI Jean-Luc	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH	DEBATY Guillaume	Médecine d'Urgence
PU-PH	DEBILLON Thierry	Pédiatrie
PU-PH	DECAENS Thomas	Gastroentérologie ; hépatologie ; addictologie
PR Attaché	DEFAYE Pascal	Cardiologie
PU-PH	DEGANO Bruno	Pneumologie ; addictologie
PU-PH	DEMATTEIS Maurice	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH émérite	DEMONGEOT Jacques	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	DERANSART Colin	Physiologie
PU-PH	DESCOTES Jean-Luc	Urologie
PU-PH	DETANTE Olivier	Neurologie
MCU-PH	DIETERICH Klaus	Génétique
PU-PH	DJAILEB Loïc	Biophysique et médecine nucléaire
PU-PH	DONDE-COQUELET Clément	Psychiatrie d'adultes
MCU-PH	DOUTRELEAU Stéphane	Physiologie
MCU-PH	DREVET Sabine	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
PU-PH	DUMAS Guillaume	Médecine intensive-réanimation
PU-PH	DUMESTRE PÉRARD Chantal	Immunologie
PU-PH	ÉPAULARD Olivier	Maladies infectieuses ; Maladies tropicales
MCU-PH	EVAIN Jean-Noël	Anesthésiologie-réanimation et médecine périopératoire
MCU-PH	EYSSERIC Hélène	Médecine légale et droit de la santé
PU-PH émérite	FAGRET Daniel	Biophysique et médecine nucléaire
PU-PH	FAUCHERON Jean-Luc	Chirurgie viscérale et digestive
PU-PH	FAURÉ Julien	Biochimie et biologie moléculaire
PU-PH	FERRETTI Gilbert	Radiologie et imagerie médicale
PU-PH	FIARD Gaëlle	Urologie
PU-PH	FONTAINE Éric	Nutrition
PU-PH émérite	FRANÇOIS Patrice	Épidémiologie, économie de la santé et prévention
PR Ass. Méd.	FREY Gil	Chirurgie thoracique et cardiovasculaire

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
MCU-MG	GABOREAU Yoann	Médecine Générale
PU-PH	GARBAN Frédéric	Hématologie ; Transfusion
PU-PH	GAUDIN Philippe	Rhumatologie
MCU-PH	GAUTIER-VEYRET Elodie	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH	GAVAZZI Gaétan	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
PU-PH	GAY Emmanuel	Neurochirurgie
MCU-PH	GIAI Joris	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	GILLOIS Pierre	Biostatistiques, informatique médicale et technologies de communication
PU-PH	GIOT Jean-Philippe	Chirurgie plastique, reconstructrice et esthétique ; Brûlologie
PU-PH	GIRARD Edouard	Chirurgie viscérale et digestive
MCF Ass.MG	GIRARD Pauline	Médecine Générale
MCU-PH	GRAND Sylvie	Radiologie et imagerie médicale
PU-PH émérite	GRIFFET Jacques	Chirurgie infantile
PU-PH	HAINAUT Pierre	Biochimie et biologie moléculaire
PU-PH	HENNEBICQ Sylviane	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PR Ass. Méd.	HODAJ Hasan	Thérapeutique-médecine de la douleur
PU-PH	HOFFMANN Pascale	Gynécologie-obstétrique ; gynécologie médicale
PU-PH émérite	HOMMEL Marc	Neurologie
PU-PH émérite	JOUK Pierre-Simon	Génétique
PU-PH	JOUVE Thomas	Néphrologie
PU-PH	KAHANE Philippe	Physiologie
PU-PH	KASTLER Adrian	Radiologie et imagerie médicale
MCU-PH	KHERRAF Zine-Eddine	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PU-PH	KRAINIK Alexandre	Radiologie et imagerie médicale
PU-PH	LABARÈRE José	Épidémiologie, économie de la santé et prévention
PU-PH	LABLANCHE (CORNALI) Sandrine	Endocrinologie, diabète et maladies métaboliques
PU-PH	LANDELLE Caroline	Bactériologie – virologie ; Hygiène hospitalière
PU-PH	LANTUEJOUL Sylvie	Anatomie et cytologie pathologiques
PR Ass. Méd.	LARAMAS Mathieu	Cancérologie ; radiothérapie
MCU-PH	LARDY Bernard	Biochimie et biologie moléculaire
PU-PH	LAURENT-COSTENTIN Charlotte	Gastroentérologie ; hépatologie ; addictologie
MCU-PH	LE GOUELLEC LE PISSART Audrey	Biochimie et biologie moléculaire
MCU-PH	LE MARÉCHAL Marion	Maladies infectieuses ; maladies tropicales
PU-PH	LECCIA Marie-Thérèse	Dermato-vénéréologie
PR Ass.MG	LEDOUX Jean-Nicolas	Médecine Générale
PU-PH émérite	LETOUBLON Christian	Chirurgie viscérale et digestive
PU-PH émérite	LEVY Patrick	Physiologie
PU-PH	LONG Jean-Alexandre	Urologie
MCU-PH	LUPO Julien	Bactériologie-virologie ; Hygiène hospitalière
MCU-PH	MARLU Raphaël	Hématologie ; Transfusion
PU-PH	MAURIN Max	Bactériologie-virologie ; Hygiène hospitalière
PR Ass. Méd.	MAURENT-PALOMBI Karine	Ophtalmologie
MCU-PH	MC LEER Anne	Histologie, embryologie et cytogénétique
MCU-PH	MEONI Sara	Neurologie

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
MCU-PH	MEONI Sara	Neurologie
MCU-PH	MEUNIER Mathieu	Hématologie ; Transfusion
PR Ass. Méd.	MICHY Thierry	Gynécologie-obstétrique
MCF Ass.MG	MINIER Pierre	Médecine Générale
MCU-PH	MONDET Julie	Histologie, embryologie et cytogénétique
PU-PH	MORAND Patrice	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	MOREAU-GAUDRY Alexandre	Biostatistiques, informatique médicale et technologies de communication
PU-PH	MORO Elena	Neurologie
PU-PH	MORO-SIBILOT Denis	Pneumologie ; addictologie
PU-PH	MORTAMET Guillaume	Pédiatrie
PU-PH émérite	MOUSSEAU Mireille	Cancérologie ; radiothérapie
PU-PH émérite	MOUTET François	Chirurgie plastique, reconstructrice et esthétique ; brûlologie
MCF Ass.MG	NAVARRE Manon	Médecine Générale
PR Ass. Méd.	NGUON Anne-Sophie	Psychiatrie d'adultes
PR Ass.MG	ODDOU Christel	Médecine Générale
MCU-PH	PACLET Marie-Hélène	Biochimie et biologie moléculaire
PU-PH	PAILHÉ Régis	Chirurgie orthopédique et traumatologie
PU-PH	PALOMBI Olivier	Anatomie
PU-PH	PARK Sophie	Hématologie ; Transfusion
PU-PH émérite	PAYEN DE LA GARANDERIE Jean-François	Anesthésiologie-réanimation et médecine péri-opératoire
MCU-PH	PAYSANT François	Médecine légale et droit de la santé
MCU-PH	PELLETIER Laurent	Biologie cellulaire
PU-PH	PELLOUX Hervé	Parasitologie et mycologie
PU-PH	PÉPIN Jean-Louis	Physiologie
PU-PH	PÉRENNOU Dominique	Médecine physique et de réadaptation
PAST	PICARD Julien	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH	PERNOD Gilles	Chirurgie vasculaire ; Médecine vasculaire
PR	PINSAULT Nicolas	Sciences de la rééducation et de réadaptation
PU-PH	PIOLAT Christian	Chirurgie infantile
PU-PH émérite	PISON Christophe	Pneumologie ; Addictologie
PU-PH émérite	PLANTAZ Dominique	Pédiatrie
PU-PH	POIGNARD Pascal	Bactériologie-virologie ; Hygiène hospitalière
PU-PH émérite	POLACK Benoît	Hématologie ; Transfusion
PU-PH	POLOSAN Mircea	Psychiatrie d'adultes ; Addictologie
MCU-PH	RABATTU Pierre-Yves	Anatomie
PU-PH	RAY Pierre	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
MCU-PH	RENDU John	Biochimie et biologie moléculaire
MCU-PH émérite	RIALLE Vincent	Biostatistiques, informatique médicale et technologies de communication
PU-PH	RIETHMULLER Didier	Gynécologie-obstétrique ; gynécologie médicale
PU-PH	RIGHINI Christian	Oto-rhino-laryngologie
PU-PH émérite	ROMANET Jean Paul	Ophthalmologie
PU-PH	ROSTAING Lionel	Néphrologie
PU-PH	ROUSTIT Matthieu	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie

CORPS	NOM Prénom	DISCIPLINE UNIVERSITAIRE
MCU-PH	ROUX-BUISSON Nathalie	Biochimie et biologie moléculaire
PR Ass.MG	ROYER DE VÉRICOURT Guillaume	Médecine Générale
PU-PH émérite	SARAGAGLIA Dominique	Chirurgie orthopédique et traumatologie
MCU-PH	SATRE Véronique	Génétique
PU-PH	SAUDOU Frédéric	Biologie cellulaire
PU-PH	SCHMERBER Sébastien	Oto-rhino-laryngologie
PU-PH	SCHWEBEL Carole	Médecine intensive-réanimation
PU-PH	SCOLAN Virginie	Médecine légale et droit de la santé
PU-PH	SEIGNEURIN Arnaud	Épidémiologie, économie de la santé et prévention
PU-PH	SPEAR Rafaëlle	Chirurgie vasculaire ; médecine vasculaire
PU-PH émérite	STAHL Jean-Paul	Maladies infectieuses ; Maladies tropicales
PU-PH	STANKE Françoise	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
MCU-PH	STASIA Marie-José	Biochimie et biologie moléculaire
PU-PH	STURM Nathalie	Anatomie et cytologie pathologiques
PU-PH	TAMISIER Renaud	Physiologie
PU-PH	THEVENON Julien	Génétique
PU-PH	TOFFART Anne-Claire	Pneumologie ; Addictologie
PU-PH	TONETTI Jérôme	Chirurgie orthopédique et traumatologie
PU-PH	TOUSSAINT Bertrand	Biochimie et biologie moléculaire
PU-PH	TRILLING Bertrand	Chirurgie viscérale et digestive
PU-PH	VALMARY-DEGANO Séverine	Anatomie et cytologie pathologiques
PU-PH	VANZETTO Gérald	Cardiologie
PU-PH	VERRY Camille	Cancérologie ; Radiothérapie
PU-PH	VIGLINO Damien	Médecine d'urgence
PU-PH émérite	ZARSKI Jean-Pierre	Gastroentérologie ; hépatologie ; addictologie
PU-PH	ZAOUI Philippe	Néphrologie

PU-PH	Professeur des universités - Praticien hospitalier
MCU-PH	Maître de conférences des universités - Praticien hospitalier
PU-PH émérite*	Professeur des universités - Praticien hospitalier émérite
MCU-PH émérite*	Maître de conférences des universités - Praticien hospitalier émérite
PU-MG	Professeur des universités de médecine générale
MCU-MG	Maître de conférences des universités de médecine générale
MCF	Maître de conférences des universités
PR Ass. Med.	Professeur des universités associé (à mi-temps)
PR Ass. MG	Professeur des universités de médecine générale associé (à mi-temps)
MCF Ass.MG	Maître de conférences des universités de médecine générale associé (à mi-temps)
PAST	Professeur associé en service temporaire
PR Attaché	Professeur Attaché

REMERCIEMENTS

Je tiens tout d'abord à remercier les membres du jury :

- Mme le **Pr Séverine Valmary-Degano**, présidente du jury et co-directrice de la thèse. Vous m'avez fait découvrir l'anatomopathologie avec passion et enthousiasme, ce qui m'a encouragé à choisir cette spécialité et dont je ne regrette pas le choix. Vous m'avez également encadré dans ce travail avec bienveillance et en restant toujours présente, tout en m'encourageant à aller plus loin.
- Mme le **Dr Ekaterina Bourova-Flin**, co-directrice de la thèse. Merci énormément pour ton encadrement sans faille et pour toute la passion que tu entretiens et que tu transmets avec la bio-informatique et la recherche. Tu m'as fait découvrir un monde fascinant et inspirant.
- Mme le **Dr Marie-Hélène Laverrière**. Votre bienveillance et votre bonne humeur au quotidien sont contagieuses. Je ne vous remercierai jamais assez des connaissances que vous nous transmettez au microscope ou en macroscopie.
- M. le **Pr Thomas Decaens**. Je vous remercie d'avoir répondu présent afin de lire mon travail et d'être présent en tant que membre de jury, malgré un planning très certainement amplement chargé. Avoir votre avis de clinicien sur mon travail de laboratoire est crucial et indispensable.
- M. le **Dr Julio Abba**. Je vous remercie également de votre présence dans le jury de ma thèse et du point de vue intéressant du chirurgien que vous nous apporterez. J'ai pu passer dans votre service en tant qu'externe et vous avez pu me montrer que la gentillesse et le professionnalisme ne sont pas incompatibles.

Je remercie toute l'équipe « Régulations épigénétiques » de l'IAB du **Dr Saadi Khochbin** avec **Dr Sophie Rousseaux** et **Mme Anne-Laure Vitte**. Outre votre aide précieuse et votre encadrement dans mon travail de thèse, vous m'avez fait découvrir un univers passionnant avec bienveillance et compétence. J'espère pouvoir continuer à travailler avec vous pour la suite du projet.

À mes séniors et collègues de l'anatomopathologie :

Je remercie les médecins pathologistes du CHU Grenoble Alpes : **Pr Nathalie Sturm**, **Dr Eleni Nika**, **Dr Hugo Gil**, **Dr Jean Boutonnat**, **Dr Pierre Dal-Col**, **Dr Catherine Meilhac-Fournier**, **Dr Diane Giovannini** et **Dr Nathalie Chiquet**. Je remercie également les

médecins pathologistes du CH Métropole Savoie : **Dr Jean-Jacques Roux, Dr Geneviève Herve, Dr Anna Gerbe, Dr Adina Petcu, Dr Pierre Pocachard, Dr Claire Arbault et Dr Amandine Bailly**. Vos connaissances, votre pédagogie et vos conseils ont été, sont et seront utiles pour ma carrière professionnelle.

Une attention particulière à mes jeunes cheffes absolument géniales : **Dr Aurélie Cuypers et Dr Romane Siest**.

Un grand merci à **l'équipe technique, aux ingénieurs, aux ARC et au secrétariat** des services d'anatomopathologie de Grenoble et de Chambéry. Vous avez pu supporter mes maladresses et mon humour peu glorieux. Vous m'avez également apporté beaucoup sur le plan professionnel.

À mes séniors et collègues de biologie médicale :

Je remercie l'ensemble des médecins et pharmaciens biologistes du CHU Grenoble Alpes qui m'ont encadré et soutenu pendant mon année d'internat de biologie médicale et de repassage des ECN. En particulier les services d'hématologie biologique (**Pr Pascal Mossuz, Dr Raphaël Marlu, Dr Gautier Szymanski, Dr Claire Vettier, Dr Simon Chevallier, Dr Bénédicte Bulabois, Dr Claire Barro et Dr Sylvain Carras**) ainsi que le service de pharmacologie et de toxicologie (**Pr Françoise Stanke, Dr Hélène Eysseric, Dr Mireille Bartoli, Dr Xavier Foronse, Dr Julia Tonini et Dr Elodie Gautier-Veyret**). Votre compréhension et votre aide dans mon choix de changement de spécialité ont été plus que réconfortant. Je n'ai jamais regretté mon choix de la biologie médicale et la quitter n'a pas été un choix simple. L'ensemble des connaissances et compétences acquises pendant cette année d'internat a été et sera importante dans ma vie professionnelle.

Un grand merci à toute **l'équipe technique, aux ingénieurs et aux secrétaires** des services de biologie, de jour comme de nuit, du CHU de Grenoble, avec qui j'entretiens toujours autant de bonnes relations et qui m'ont tant appris.

À mes co-internes et amis pathologistes :

- **Anouk** (Nouk-nouk) qui me supporte depuis le premier jour en anapath et avec qui je discute comme un daron de nos chiens. Elle trouvera toujours une raison pour râler mais sera toujours présente pour aider. Bientôt l'ouverture de Proxipath. Un grand respect au vaillant orthopédiste Valentin et pleins de bisous à Maki le corgi !
- **Nicolas** qui aurait dû être député et/ou pâtissier sur la Canebière.
- **Maureen** qui est clairement la meilleure d'entre nous. Pédagogue, drôle et bienveillante.

- **Jérémie** avec qui je partage le même humour.
- **Rami** qui aura toujours une question précise à vous poser.
- **Jéromine** qui a l'énergie d'une centrale nucléaire et qui a disparu de l'IBP depuis plus d'un an.
- **Seda** qui sous ses airs timides cache du caractère, de la gentillesse et beaucoup d'humour.
- **Ayman** qui sera toujours présent en soirée (sans lui, Nicolas ne viendrait sûrement jamais).
- **Siwar**, notre influenceuse, qui arrive à rendre l'anapath élégant et moins dégueu (grâce notamment à son masque anti-odeur de macro digestive).
- **Maxence** qui a tout pour devenir un futur chef d'entreprise.
- **Noélie** qui est un saule pleureur toxique, mais très sympathique en vrai.

C'est une chance inouïe d'être accompagné d'une équipe aussi exceptionnelle et compétente.

À mes co-internes et amis biologistes du « club des 5 » :

- **Sébastien** le métabologue de Valence (la petite Arménie), qui ne sait clairement pas se faire d'ennemis et qui brille tant par ses connaissances que par son humour (il s'auto-surnomme « dieu », en toute modestie).
- **Charlotte** l'hématologue nord iséroise, qui ne sait toujours pas dire non, malgré un DU en autorité. Mariée récemment avec le brillant **Thomas** (le camarade de bouffe), elle voue une passion pour les 106 de collection.
- **Jeanne** les embrouilles, qui arrive toujours à se retrouver dans des histoires improbables (elle est où Jeanne ?). La distance entre elle et le CHU est proportionnelle à son nombre de semestre d'internat.
- **Lydie** l'aventurière de Gap, qui est toujours occupée et jamais sur Grenoble. La légende raconte qu'on la croise parfois au labo.

Merci à vous 4 de m'avoir toujours soutenu pour le repassage du concours. Cette année de bio avec vous était incroyable, malgré le COVID, les confinements et les révisions.

À tous les anciens et nouveaux collègues bio disséminés aux quatre coins de la France et de l'IBP : **Capucine, Arthur, Edouard, Laura G, Laura S, Lionel, Giang, Julien A, Julien M, Morgane, Elisa, Nathan, Heidi ...**

À mes amis pharmaciens : **Quentin, Morgane, Manon, Houssein**

À mes amis et co-survivants de l'enfer des études de médecine :

- **Timour**, le très fier nord-isérois, qui aura toujours des histoires folles à vous raconter (j'ai en tête une sombre histoire de chèvre et de rat) et qui est incontestablement un brillant neurologue (ou rappeur)
- **Clémence**, la camarade rebelle, dont la légende raconte que si elle ne se plaint pas, c'est qu'elle est malade. Une médecin généraliste consciencieuse qui sera toujours à l'écoute de ses patients.
- **Rony et Esther**, le couple façon *love story* à l'américaine. Adorable, bienveillant et intelligent sont clairement les adjectifs qui pourraient les définir. Bientôt la famille s'agrandira !
- **Melissa et Alexis**, nos deux médecins généralistes à l'aventure. Avec Melissa, on a tout survécu ensemble, de la P1 à la D4. Et maintenant avec Alexis vous semblez profiter de votre liberté en allant exercer aux quatre coins de la région. Bravo à vous !
- **Gabin**, à mon ami kiné de Saint-Marcellin, avec qui j'ai passé la P1 et qui restera également mon collègue du LADAF. Un gars qui est toujours curieux et passionné par ce qu'il fait.

Merci à vous de votre soutien durant ces études de malade.

À ma **belle-famille Laurent, Fabienne, Leo, Audrey, Axel, Josette, Jean-Claude, Nicole, Sandrine, Richard, Monique, Jean-Pierre et Sabine**. Vous m'avez accepté rapidement comme un membre de la famille et cela me touche sincèrement.

À mes **parents Raphaël et Murielle**, qui ont toujours été là et qui m'ont soutenus de manière sans faille durant toutes ces études. J'en suis là grâce à vous, je vous dédie donc cette thèse.

À ma **sœur Amélie**, qui sera malgré elle une excellente pharmacienne hospitalière à l'accent sudiste prononcé. Même si tu deviendras sûrement une tête plate d'Occitanie, tu resteras à jamais ma sœur et je te dédie également cette thèse (en espérant être invité à la tienne, notamment pour le pot !).

Au reste de ma famille.

À mon **toutou Arya** qui dans tous les problèmes arrivaient à trouver une solution et qui est la seule à me faire sortir lorsque l'envie m'échappe.

À toi, **l'amour de ma vie Marie**. Ce travail t'est entièrement dédié. Tu es ma minidoc pharmacienne toute mignonne avec beaucoup de caractère. Tu as toujours été là pour moi dans les moments les plus difficiles et tu es la source de tous les moments les plus magiques. Tu doutes toujours de toi alors que tu es parfaite. Les patients ont de la chance de t'avoir pour corriger toutes les bêtises des médecins. Cette année 2024 aura été particulière pour nous deux puisqu'on a passé nos thèses, tu as terminé tes longues études et nous nous sommes pacsés. De belles choses nous attendent ensemble dans le futur. *Ti amo e voglio fare la mia vita con te.*

Ces remerciements n'arriveront jamais à la hauteur de toute la gratitude que je vous dois. Vos supports et vos conseils ont su guider ma vie professionnelle et personnelle.

Aurélien Spinelli

RÉSUMÉ

Titre : Signature Pronostique dans le Cancer du Côlon par l'Etude de l'Expression Ectopique de Gènes Spécifiques de Tissus

Résumé :

Introduction : Le cancer du côlon représente, en 2020, le 4^{ème} et 2^{ème} cancer en termes d'incidence et de mortalité respectivement. Malgré de nombreux critères histo-pronostiques, le pronostic et la prise en charge des cancers du côlon en 2024, particulièrement pour les stades II et IV, reste parfois incertains. L'Institut pour l'Avancée des Biosciences a démontré que l'expression ectopique de certains gènes tissus-spécifiques dans le cancer, alors qu'ils sont silencieux dans le tissu sain, est significativement associée à un pronostic défavorable. Ainsi, ces gènes représentent une nouvelle source de biomarqueurs pronostiques. L'objectif principal de cette étude est de créer une signature pronostique dans les cancers du côlon, basé sur les activations ectopiques, pour mieux guider la prise en charge thérapeutique des patients.

Méthode : Nous avons utilisé le pipeline de *machine learning* « ectopy » développé à l'Institut pour l'Avancée des Biosciences qui permet de découvrir des biomarqueurs pronostiques robustes à partir de données publiques d'expression de gènes. Ces biomarqueurs sont combinés dans un outil pronostique GEC qui permet de stratifier les patients en fonction du nombre d'activations dans le panel. Plus grand est le nombre de gènes activés de manière ectopique, moins longue est la probabilité de survie des patients.

Résultat : En appliquant la méthode « ectopy » à trois cohortes publiques du cancer du côlon (effectif total n=1141), nous avons défini une signature pronostique GEC de 4 gènes : ERFE, HOXC6, LAMP5 et ULBP2. En analyses uni- et multivariée, l'activation ectopique d'au moins un de ces 4 gènes est un facteur de pronostic défavorable, tous stades confondus, ainsi que dans chaque stade de cancer du côlon séparément. Suite à ces travaux, nous avons conçu une nouvelle cohorte rétrospective monocentrique du CHUGA, de 2012 à 2014, constituée de 133 patients afin de valider la signature GEC par la technologie RT-qPCR.

Discussion : La signature GEC prédit le pronostic de survie dans le cancer du côlon, indépendamment d'autres critères pronostiques. Cet outil aidera à la prise en charge des patients, en particulier dans les stades II de pronostic incertain et les stades IV où un traitement chirurgical par chimiothérapie hyperthermique est discuté. Les perspectives de l'étude sont une analyse par RT-qPCR sur la cohorte de cancer du côlon du CHUGA ainsi qu'une mise au point d'un test immunohistochimique pour un passage en routine.

Mots-clés : cancer du côlon ; expression ectopique ; analyse de survie ; signature pronostique.

ABSTRACT

Title: Prognostic Signature in Colon Cancer based on Ectopic Expression of Tissue-Specific Genes

Abstract:

Introduction: Colon cancer represents, in 2020, the 4th and 2nd cancer in term of incidence and mortality respectively. Despite numerous histoprognostic factors, prognosis and care of colon cancer in 2024, particularly stage II and IV, remains uncertain. The Institute for Advanced Biosciences showed that ectopic expression of some tissue-specific genes in cancer, instead of remaining silent in normal tissue, is significantly associated with poor prognosis. Thus, those genes represent a new source of prognostics biomarkers. The main objective of this study is to create a prognostic signature in colon cancer, based on ectopic activation, to improve the therapeutic care of patients.

Methods: We used « ectopy »; a machine learning pipeline developed at the Institute for Advanced Biosciences which allows to discover robust prognostic biomarkers from gene expression data. These biomarkers are combined in a GEC prognostic tool that allows patients to be stratified according to the number of activations in the panel. The greater the number of ectopically activated genes, the shorter the probability of patient survival.

Results: By applying the « ectopy » method to three public colon cancer cohorts (effective n=1141), we defined a GEC prognostic signature of 4 genes: ERFE, HOXC6, LAMP5 and ULBP2. In uni- and multivariate analyses, ectopic activation of at least one of these 4 genes is a negative prognostic factor, all stages combined, as well as in each stage of colon cancer separately. Following this work, we designed a new retrospective single-center cohort at the CHUGA, from 2012 to 2014, consisting of 133 patients in order to validate the GEC signature by RT-qPCR technology.

Discussion: The GEC signature predicts the prognosis of survival in colon cancer, independent of other prognostic criteria. This tool will help in the management of patients, especially in stages II of uncertain prognosis and stages IV where surgical treatment with hyperthermic chemotherapy is discussed. The prospects of the study are an RT-qPCR analysis on the CHUGA colon cancer cohort as well as the development of an immunohistochemically test for a routine transition.

Keywords: colon cancer; ectopic expression; survival analysis; prognostic signature.

TABLE DES MATIÈRES

REMERCIEMENTS	7
RÉSUMÉ	12
ABSTRACT	13
TABLE DES MATIÈRES	14
LISTE DES FIGURES	17
LISTE DES TABLEAUX	19
LISTE DES ABRÉVIATIONS	20
INTRODUCTION	22
1. Cancer du côlon.....	22
1.1. Épidémiologie	22
1.2. Étiologies et facteurs de risque.....	22
1.3. Histoire naturelle de l'adénocarcinome colorectal.....	23
1.4. Diagnostic.....	28
1.5. Prise en charge thérapeutique.....	40
1.6. Des résultats contrastés et la nécessité d'une signature pronostique	41
2. Expression ectopique de gènes spécifiques de tissus de tissus en cancérologie	42
2.1. Généralité	42
2.2. Impact sur la cancérogénèse et le pronostic	44
2.3. Intérêt diagnostique	44
2.4. Intérêt thérapeutique.....	45
3. Objectifs du travail de thèse	45
MATÉRIEL ET MÉTHODE	47
1. Étude bio-informatique	47
1.1. Matériel : cohortes publiques	47
1.2. Analyses statistiques des données	50
2. Étude sur une cohorte rétrospective du CHUGA	54

2.1.	Constitution de la cohorte	54
2.2.	Méthodologie de l'analyse par RT-qPCR et préparation des amorces	56
2.3.	Hypothèse et méthodes d'analyse statistique	56
2.4.	Éthique et confidentialité	57
RÉSULTATS		58
1.	Résultats de l'étude bio-informatique	58
1.1.	Identification des gènes spécifiques de tissus et ceux activés de manière aberrante dans le cancer du côlon	58
1.2.	Création de la signature pronostique GEC par la méthode « ectopy »	60
1.3.	Évaluation de l'impact pronostique de la signature GEC par rapport à d'autres facteurs de risque connus	63
2.	Identification des signatures moléculaires des tumeurs agressives GEC+ par l'analyse d'enrichissement GSEA	65
3.	Cohorte du CHU Grenoble Alpes.....	69
3.1.	Caractéristiques démographiques de la cohorte	69
3.2.	Mise au point des amorces pour la RT-qPCR.....	72
4.	Valorisation des résultats	72
DISCUSSION		73
1.	Une nouvelle signature dans le CC basée sur l'expression ectopique de 4 gènes spécifiques de tissu	73
2.	Revue systématique de la littérature des gènes GEC identifiés	74
2.1.	ERFE	74
2.2.	HOXC6.....	75
2.3.	LAMP5.....	76
2.4.	ULBP2.....	77
3.	Connaissances actuelles des expressions ectopiques de gènes spécifiques de tissus dans le cancer du côlon	78
4.	Limites de l'étude.....	79

5. Perspectives du travail de thèse.....	79
5.1. Projet SIROCCO	79
5.2. Recherche d'une signature diagnostique à l'aide « d'ectopy ».....	80
CONCLUSION.....	82
CONCLUSION SIGNÉE PAR LE DOYEN ET LA PRÉSIDENTE DU JURY.....	83
BIBLIOGRAPHIE	84
ANNEXE	98
Annexe 1. Algorithmes décisionnels thérapeutiques proposés par le TNCD.....	98
Annexe 2. Liste non exhaustive des signatures pronostiques proposées dans le cancer du côlon	100
Annexe 3. Séquences d'amorces utilisées dans les RT-qPCR.....	101
Annexe 4. Expression physiologique des gènes ERFE, HOXC6, LAMP5 et ULBP2 selon les tissus	102
Annexe 5. Résultats de l'évaluation de l'interaction protéine-protéine avec IntAct.....	103
Annexe 6. Liste de gènes spécifiques de tissus dont l'expression ectopique aurait un impact sur le pronostic dans le cancer du côlon	104
SERMENT D'HIPPOCRATE.....	108

LISTE DES FIGURES

Figure 1. Classifications moléculaires proposées dans le cancer du côlon et illustrées dans la revue de Dunne P. et Arends M. en 2024	25
Figure 2. Illustration des neuf principales voies moléculaires associées à la carcinogenèse colorectale	27
Figure 3. Différences épidémiologiques, pronostiques et du phénotype moléculaire selon la localisation primitive du cancer du côlon entre le côlon droit et le côlon gauche	27
Figure 4. Images histologiques d'une muqueuse normale et de quelques sous-types morphologiques de l'adénocarcinome colique.....	30
Figure 5. Images histologiques de quelques facteurs histo-pronostiques.....	31
Figure 6. Images histologiques et immunohistochimiques d'une biopsie d'un adénocarcinome colique bien différencié avec perte d'expression de protéines MMR	32
Figure 7. Cascade de réalisation des tests somatiques et indication de consultation d'oncogénétique pour les colorectaux, issue des recommandations INCa 2021	38
Figure 8. Arbre décisionnel des biomarqueurs nécessaires pour le traitement des patients atteints d'adénocarcinomes colorectaux, issu des recommandations de l'INCa de juin 2022 .	39
Figure 9. Voies de cancérogénèse et l'interaction du micro-environnement tumoral pouvant constituer des cibles potentielles pour les biothérapies.....	43
Figure 10. Étapes principales du pipeline « ectopy » qui permet de découvrir des biomarqueurs pronostiques (Gene Expression Classifier, GEC) à partir des données d'expression de gènes	50
Figure 11. Les étapes principales d'entraînement, de validation et de test des modèles prédictifs de survie dans la méthode « ectopy ».....	53
Figure 12. Nombre de gènes spécifiques de tissus présentés par ordre d'abondance.....	59
Figure 13. Exemple de la fréquence d'activation du gène ULBP2.....	59
Figure 14. Fréquences d'activations ectopiques de 3554 gènes spécifiques de tissus dans le CC.	60
Figure 15. Fréquences d'activation ectopique pour les 20 gènes spécifiques de tissus les plus fréquemment activés dans le CC.....	60
Figure 16. Courbes de survie de Kaplan-Meier des quatre gènes au sein des datasets TCGA-COAD et GSE39582 en fonction du statut d'activation du gène (OFF/ON)	62
Figure 17. Courbes de survie de Kaplan-Meier au sein des cohortes publiques du CC en fonction de l'expression de la signature GEC.	62

Figure 18. Courbes de survie de Kaplan-Meier des groupes GEC+ et GEC- issues de l'analyse en sous-groupe des trois cohortes publiques rassemblées.....	64
Figure 19. Analyse multivariée des trois cohortes en tenant compte du statut N, des mutations KRAS, du grade et du stade UICC.....	65
Figure 20. Résultats de l'analyse différentielle du transcriptome total dans le dataset TCGA-COAD entre les groupes GEC+ et GEC-.....	66
Figure 21. Résultats de l'analyse GSEA pour les profils transcriptomiques GEC+ versus GEC- dans une sélection de gene sets significativement enrichis (score d'enrichissement normalisé NES > 0) ou déplétés (NES < 0).....	67
Figure 22. Le score normalisé d'enrichissement NES obtenu par l'analyse GSEA entre les groupes GEC+ versus GEC- pour trois cohortes du CC et pour différents gene sets.....	68
Figure 23. Algorithme décisionnel thérapeutique des CC superficiels (stade 0 et I) proposé par le TNCD.....	98
Figure 24. Algorithme décisionnel thérapeutique des CC de stade II proposé par le TNCD..	98
Figure 25. Algorithme décisionnel thérapeutique des CC de stade III proposé par le TNCD	99
Figure 26. Indications des biothérapies dans les formes métastatiques (stade IV) du CC proposées par le TNCD, selon le phénotype moléculaire et la localisation du primitif.....	99
Figure 27. Histogrammes illustrant les expressions en fonction des tissus des gènes ERFE, HOXC6, LAMP5 et ULBP2.....	102
Figure 28. Représentation des potentielles interactions protéiques retrouvées sur IntAct pour ERFE, HOXC6, LAMP5 et ULBP2 chez l'Homme.....	103

LISTE DES TABLEAUX

Tableau 1. Facteurs histo-pronostiques reconnus par l'OMS 5 ^{ème} édition.....	33
Tableau 2. Classification clinique et pathologique du TNM de l'adénocarcinome du côlon issu de la 8 ^{ème} édition de l'UICC	35
Tableau 3. stades d'évolution de l'adénocarcinome du côlon issu de la 8 ^{ème} édition de l'UICC	36
Tableau 4. Caractéristiques démographiques des cohortes issues de la littérature	48
Tableau 5. Caractéristiques démographiques de la cohorte de 2012 à 2014 du CHU Grenoble-Alpes.....	71
Tableau 6. Liste non exhaustive des signatures pronostiques proposées dans le cancer du côlon	100
Tableau 7. Amorces utilisées pour l'analyse par RT-qPCR	101
Tableau 8. Liste de gènes spécifiques de tissus dont l'expression ectopique aurait un impact sur le pronostic dans le CC.....	107

LISTE DES ABRÉVIATIONS

5-FU	5-fluoro-uracile
ADN	Acide désoxyribonucléique
AJCC	<i>American Joint Committee on Cancer</i>
ARN	Acide ribonucléique
ARNc	Acide ribonucléique complémentaire
ARNm	Acide ribonucléique messenger
ATCD	Antécédent
BAD-LAMP	<i>Brain and Dendritic Cell associated LAMP-like</i>
CC	Cancer du côlon
CCR	Cancer colo-rectal
CHIP	Chimiothérapie intra-péritonéale
CHUGA	Centre-hospitalo universitaire Grenoble Alpes
CIN	Instabilité chromosomique
CIRC	Centre International de Recherche sur le Cancer
CLR	Réaction inflammatoire péritumorale Crohn-like
CMH1	Complexe d'histocompatibilité majeur de classe 1
CMS	<i>Consensus Molecular Subtypes</i>
CMV	Cytomégalovirus
CRCSC	Colorectal Cancer Subtyping Consortium
CRIS	Colorectal intrinsic subtypes
CTA	<i>Cancer/testis antigens</i>
CTRIP	<i>Clq tumor necrosis factor-related protein</i>
DPD	Dihydropyrimidine deshydrogenase
EPO	Erythropoïétine
ERFE	Erythroferrone
FDR	<i>False Discovery Rate</i>
FPKM	<i>Fragments Per Kilobase Million</i>
GEC	<i>Gene Expression Classifier</i>
GEO	<i>Gene Expression Omnibus</i>
GSEA	<i>Gene Set Enrichment Analysis</i>
HOX	Homeobox
HOXC6	Homeobox C6
HR	<i>Hazard ratios</i>
IAB	Institut pour l'Avancée des Biosciences
iCMS	<i>Single Cell Intrinsic Consensus Molecular Subtypes</i>
INCa	Institut national du cancer
IQR	Intervalle inter-quartile
LAMP5	<i>Lysosomal associated membrane protein family member 5</i>

lncRNA	Longues sections non codantes d'ARN
LNK	Lymphocytes <i>natural killer</i>
MMR	Gènes de réparation des mésappariements
MSI	Instabilité des microsatellites
NCBI	<i>National Center for Biotechnology Information</i>
NKG2D	<i>Natural Killer Group 2 Member D</i>
NOS	Non spécifique
OMS	Organisation Mondiale de la Santé
PDS	<i>Pathway Derived Subtypes</i>
PIPAC	Chimiothérapie intra-péritonéale pressurisée par aérosol
POLE	ADN polymérase epsilon
PRAME	<i>Preferentially expressed antigen of melanoma</i>
RCUH	Recto-colite ulcéro-hémorragique
RNIPH	Recherche n'impliquant pas la personne humaine
ROC	Fonction d'efficacité du récepteur
RPKM	<i>Reads per kilo base per million mapped reads</i>
RT-qPCR	<i>Polymerase Chain Reaction (PCR)</i> quantitative après transcription inverse
SNFGE	<i>Société Nationale Française de Gastro-Entérologie</i>
SSA	Adénome festonné sessile
TCGA	The Cancer Genome Atlas
TIL	Réponse immune intratumorale
TME	Micro-environnement tumoral
TNCD	Thésaurus National de Cancérologie Digestive
TRG	Grade de régression tumorale
TSA	Adénome festonné traditionnel
UICC	Union for International Cancer Control
ULBP2	UL16 binding protein 2

INTRODUCTION

1. Cancer du côlon

1.1. Épidémiologie

Le cancer colo-rectal (CCR) est l'un des cancers les plus fréquents en France et dans le monde. Sur l'année 2020, en France, il s'agissait du 4^{ème} cancer en termes d'incidence avec 48.000 personnes atteintes et du 2^{ème} cancer selon la mortalité avec 21.000 décès, avec des chiffres assez similaires entre les hommes et les femmes. Il représente près de 12% de l'ensemble des décès par cancer, notamment chez les patients âgés de 65 ans et plus (1,2). Il est à noter qu'une augmentation de son incidence dans la population âgée de moins de 50 ans a été rapportée (3).

Depuis 2015, un dépistage national est organisé en France par un test immunologique de détection de sang occulte dans les selles. Il s'adresse aux femmes et aux hommes, âgés de 50 à 74 ans, invités tous les deux ans au programme national de dépistage du cancer colorectal, ne présentant ni symptômes, ni antécédents personnels ou familiaux de polype, de cancer ou de maladie touchant le côlon ou le rectum, ni facteurs de risque particuliers. Un test positif conduit à la réalisation d'un examen coloscopique avec biopsie ou exérèse endoscopique des lésions suspectes (2). Cette conduite de dépistage large des populations en Europe a permis une réduction significative de la mortalité (4). Néanmoins, L'Institut National du Cancer (INCa) estime qu'environ un CCR sur cinq est diagnostiqué à un stade local limité, ce qui reste relativement faible et accessible au dépistage, permettant à ces patients un traitement moins lourd et de meilleurs résultats (2).

1.2. Étiologies et facteurs de risque

De nombreux facteurs de risque ont été décrits dans le CCR. Le mode de vie et les habitudes alimentaires en font partis. On peut noter que le régime alimentaire habituellement retrouvé dans les pays à haut revenu combiné à un mode de vie sédentaire est souvent évoqué. Ceci pouvant inclure la consommation de viande rouge, d'alcool et un indice de masse corporel élevé. L'impact du tabac est encore discuté. Au contraire, un régime alimentaire composé de fibres alimentaires et de produits laitiers, ainsi qu'une activité physique régulière seraient des facteurs protecteurs.

Les prédispositions génétiques correspondent au facteur de risque parmi les plus importants du CCR. Le syndrome génétique le plus à risque, avec un risque durant leur vie de 100% de survenue de CCR, est la polypose adénomateuse familiale correspondant à une mutation du gène APC avec une transmission autosomique dominante (5). En termes de fréquence, on retrouve également le syndrome de Lynch correspondant à une mutation, par transmission autosomique dominante, des gènes de réparation des mésappariements (MMR) de l'ADN, incluant notamment les gènes MLH1, MSH2, MSH6 et PMS2, et pouvant conduire à un risque de 10 à 50% de survenue de CCR (6).

Les patients porteurs d'une inflammation chronique de l'intestin, telle que la maladie de Crohn ou la recto-colite ulcéro-hémorragique (RCUH) sont à haut risque de développer des lésions dysplasiques ainsi qu'à une transformation maligne plus rapide de celles-ci en CCR.

Par ailleurs, un antécédent d'irradiation pelvienne (7), de mucoviscidose (8) et d'acromégalie (9) ont également été décrits comme de facteurs de risque de CCR.

La prise en compte de l'ensemble de ces facteurs de risque permet d'identifier les patients à risque et à adapter leur prise en charge. Ainsi, les patients considérés à haut risque (15 à 20% des cas) sont ceux ayant un antécédent personnel de CCR et/ou d'adénomes multiples, un ou des antécédent(s) familial(aux) au premier degré de CCR et/ou d'adénome dont la taille est supérieure à 1 cm, ou ayant une maladie inflammatoire chronique de l'intestin (maladie de Crohn ou recto-colite ulcéreuse hémorragique) évoluant depuis au moins 20 ans. Les patients à très haut risque (1 à 3% des cas) sont ceux avec une prédisposition génétique personnelle connue : polypose adénomateuse familiale (mutations des gènes APC ou MUTYH) ou syndrome de Lynch (mutations des gènes MMR) (2,10).

1.3. Histoire naturelle de l'adénocarcinome colorectal

1.3.1. Classifications moléculaires

Dans le cadre de la compréhension de la cancérogénèse, deux classifications basées sur les examens de biologie moléculaire ont été initialement décrits dans l'adénocarcinome colorectal :

- Une classification basée sur des données génomiques réalisée par le réseau *The Cancer Genome Atlas* (TCGA) en 2012. Celle-ci propose deux groupes selon le taux de mutation dans le génome (basé sur l'ADN tumoral) : hypermuté (15% des CCR) et non hypermuté (85% des CCR). Dans les formes hypermutées, on retrouvera les instabilités des microsatellites (MSI) et les mutations des gènes codant l'ADN polymérase epsilon

(POLE). Dans les formes non hypermutées, il n'existe pas d'instabilité des microsatellites, mais des gènes souvent mutés tels qu'APC, TP53 et KRAS (11).

- Une autre classification publiée en 2015 par le *Colorectal Cancer Subtyping Consortium* (CRCSC) se nommant *Consensus Molecular Subtypes* (CMS) et se basant sur un profil transcriptomique (basé sur l'ARN). Ce CMS est composé de quatre sous-types principaux : les instabilités des microsatellites (CMS1), canonique (CMS2), métabolique (CMS3) et mésenchymateuse (CMS4). Un sous-type comprenant des formes mixtes (12,13).

La revue publiée par Dunne P. et Arends M. en 2024 (14) répertorie d'autres sous-types moléculaires intéressants, notamment *ColoRectal Intrinsic Subtypes* (CRIS) et *Single Cell Intrinsic CMS* (iCMS). Ils sont résumés dans la Figure 1. Le CRIS correspond à une ré-évaluation du CMS après étude des cellules carcinomateuses uniquement, indépendamment du stroma. L'iCMS se base qu'à lui sur un séquençage ARN en cellule unique, ce qui a permis d'identifier deux groupes pronostics du CMS4 (mésenchyme) : iCMS2 et iCMS3 avec respectivement un meilleur et un mauvais pronostic. Il est par ailleurs récemment décrit des sous-types basés sur les voies d'évolution des CC : *Pathway Derived Subtypes* (PDS). Ces derniers sont représentés par trois voies en fonction notamment de la population souche (cellules cylindriques des cryptes LGR5+ ou cellules souches régénératives ANXA1+), du degré de prolifération et du CMS.

L'établissement de ces sous-types moléculaires ou phénotypiques n'est pas réalisé en pratique lors de l'étude anatomopathologique d'un adénocarcinome du côlon, en-dehors de la recherche des instabilités des microsatellites. L'étude des autres gènes décrits nécessite des analyses moléculaires peu réalisées ou coûteuses. Des évaluations simplifiées du CMS par immunohistochimie avec cinq anticorps ou par analyse d'images histologiques numérisées en coloration standard à l'hématoxyline et à l'éosine par *deep learning* ont été proposées, mais avec des résultats encore à vérifier (14).

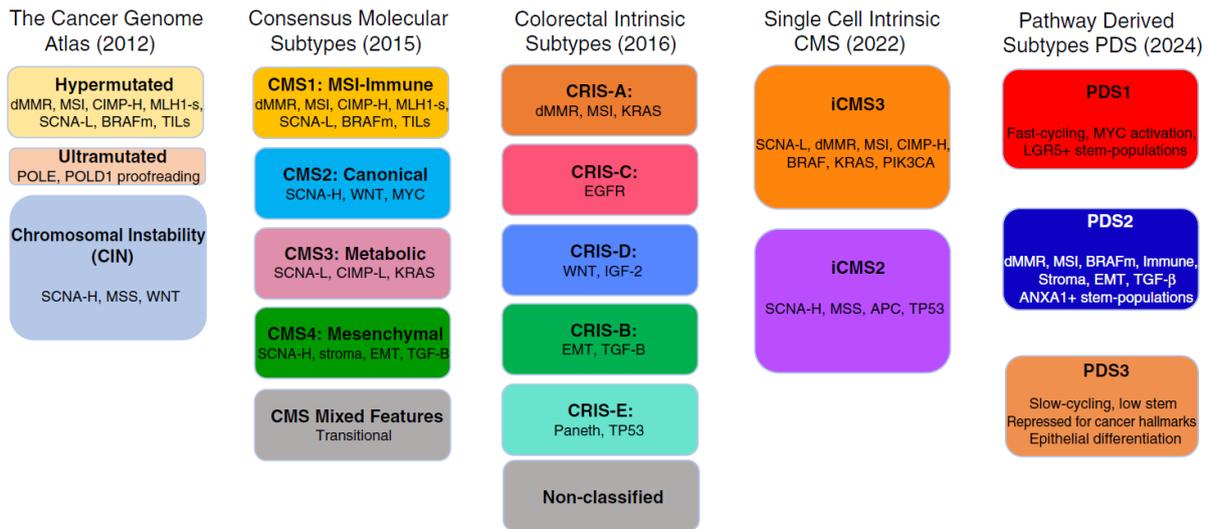


Figure 1. Classifications moléculaires proposées dans le cancer du côlon et illustrées dans la revue de Dunne P. et Arends M. en 2024 (14)

1.3.2. Voies et séquences en lien avec la carcinogénèse colique

L'étude de la carcinogénèse colorectale a permis d'identifier trois principales voies : la voie de l'instabilité chromosomique, la voie de l'instabilité des microsatellites (MSI) et la voie des polypes festonnés (15).

La voie de l'instabilité chromosomique est impliquée dans plus de 70% des CC avec une prédominance dans les lésions du côlon gauche. Elle correspond à des altérations chromosomiques multiples et variées impliquant des mutations somatiques du gène APC, mais également l'activation de KRAS et la perte de p53.

La voie MSI expliquant 10 à 20% des CC, avec une prédominance dans les lésions du côlon droit, se caractérise par un dysfonctionnement des gènes permettant la réparation des mésappariements de l'ADN (MMR), tels que MLH1, MSH2, MSH6 et PMS2. Cette voie implique des mutations de BRAF V600E, de TGFBR2 ou d'APC.

Enfin, la voie des polypes festonnés qui survient dans 10 à 20% des CC peut se retrouver dans l'ensemble des lésions du côlon. Elle présente deux types selon le précurseur précancéreux impliqué : l'adénome festonné traditionnel (TSA) et l'adénome festonné sessile (SSA). Le TSA se caractérise par des mutations KRAS, tandis que le SSA par des mutations BRAF et/ou un profil MSI.

Des séquences sont également décrites et correspondent aux différentes étapes de l'évolution d'un CC à travers les lésions précancéreuses impliquées. La séquence pathogénique classique adénome-adénocarcinome, par instabilité chromosomique, permet d'expliquer entre

70% et 90% adénocarcinomes colorectaux. La cellule à l'origine de ces lésions tumorales est, dans la majorité des cas, une cellule souche présente dans le fond des cryptes glandulaires de Lieberkühn de la muqueuse colique (16,17). Cette cellule conduisant à la formation de cryptes aberrantes, hyperplasiques et aux noyaux dysplasiques. Ceci formant dans un premier temps un adénome (tubuleux, vilieux ou tubulo-vilieux) associé à une dysplasie de bas ou haut grade, constituant une lésion pré-cancéreuse avec un risque de dégénérescence en un véritable adénocarcinome avec un potentiel invasif certain. Les autres séquences décrites contiennent la voie des polypes festonnés et la séquence hamartome-carcinome (15).

Ces différentes séquences et voies de cancérogénèse peuvent être associées à une ou plusieurs des neuf voies de signalisation moléculaire participant le plus fréquemment à la carcinogénèse colorectale (Figure 2) : Wnt/ β -caténine, JAK/STAT, Ras/ERK, PI3K/AKT/mTOR, TGF- β /SMAD, Notch, Hedgehog, TRAIL, TP53. Ces dernières peuvent intervenir dans des mécanismes clés de la cancérisation cellulaire telles que l'apoptose, la prolifération cellulaire, l'évasion immune, la différenciation ou la transition épithélio-mésenchymateuse (15).

Comme indiqué précédemment, la localisation de l'adénocarcinome dans le colon présente des différences majeures sur le plan anatomo-chirurgical, moléculaire et embryonnaire (Figure 3). Le côlon droit est représenté par le caecum, le côlon ascendant, l'angle colique droit et les deux tiers proximaux du côlon transverse. Le côlon gauche est représenté par le tiers distal du côlon transverse, l'angle colique gauche, le côlon descendant et le sigmoïde. Les adénocarcinomes localisés au sein du colon gauche présentent un meilleur pronostic que ceux du colon droit. Le profil mutationnel est différent avec une prédominance des MSI et des SSA au sein du colon droit sur le plan épidémiologique à titre d'exemple (10,18). Embryologiquement, la différence est déjà présente avec un côlon droit issu de l'intestin moyen vascularisé par l'artère mésentérique supérieur et un côlon gauche issu de l'intestin postérieur vascularisé par l'artère mésentérique inférieur (19).

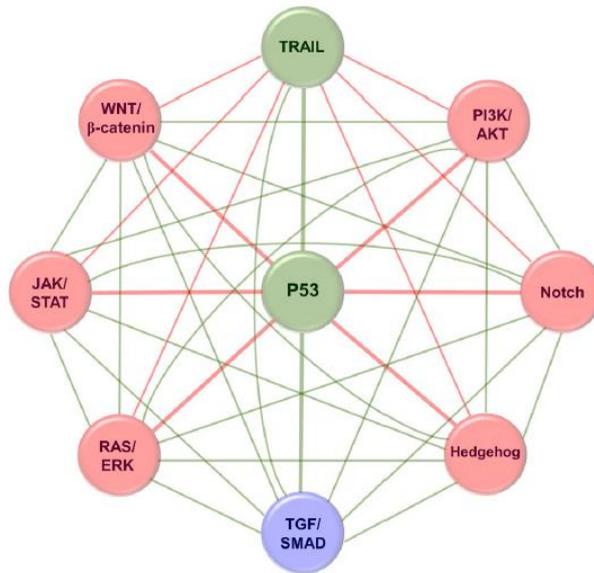


Figure 2. Illustration des neuf principales voies moléculaires associées à la carcinogénèse colorectale (15)
 Dans cette figure, ces voies moléculaires peuvent être suppressives de tumeur (cercle vert), oncogéniques (cercle rouge), et variable (cercle bleue). Les interactions entre ces voies peuvent être synergiques (trait vert) ou inhibitrices (trait rouge).

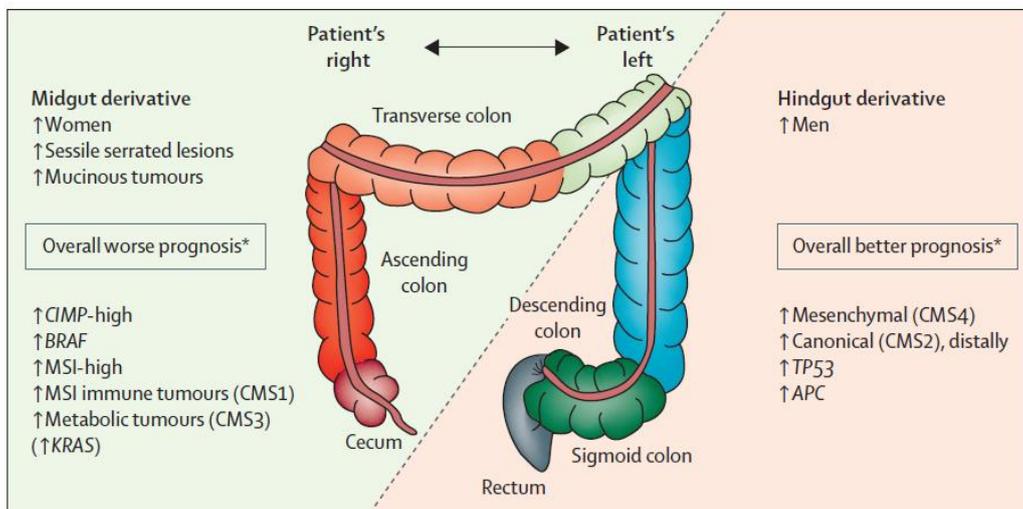


Figure 3. Différences épidémiologiques, pronostiques et du phénotype moléculaire selon la localisation primitive du cancer du côlon entre le côlon droit et le côlon gauche (10)

1.4. Diagnostic

1.4.1. Clinique

En-dehors du dépistage national organisé, des signes cliniques doivent faire évoquer cette hypothèse diagnostique : rectorragies, troubles du transit, douleurs abdominales, masse à la palpation abdominale ou au toucher rectal, anorexie, asthénie, perte de poids, anémie ferriprive et/ou maladies thrombo-emboliques d'étiologie indéterminée. Dans les stades les plus avancées, des adénopathies ou des signes en lien avec une localisation métastatique peuvent également être observées. Selon la situation, un examen complémentaire par endoscopie digestive basse et/ou par imagerie pourra être envisagé (10,20).

1.4.2. Anatomopathologie

D'après la 5^{ème} édition de 2019 des recommandations du Centre Internationale de Recherche sur le Cancer (CIRC) de l'Organisation Mondiale de la Santé (OMS), deux tumeurs épithéliales malignes sont décrites : l'adénocarcinome colorectal et les néoplasies neuroendocrines colorectales. D'autres tumeurs ou néoplasies de nature maligne sont également décrites d'origine mésenchymateuse (GIST, tumeurs vasculaires, etc...) ou hématopoïétique (lymphome de MALT, etc...), ainsi que des localisations métastatiques (20). Cependant, le CCR est très largement représenté par l'adénocarcinome colorectal (90% des CCR) et sera ainsi le sujet de la thèse.

Macroscopiquement, l'adénocarcinome colorectal peut se présenter selon des apparences variables avec des formes exophytiques dites « polypoïde » ou des formes endophytiques. Le niveau de fibrose est également variable. Habituellement, il est observé une plage ulcérée entourée de bords arrondis s'étendant le long de la circonférence colique. A l'examen macroscopique, le degré d'infiltration de la paroi colique sera évalué, ainsi que l'atteinte des tranches de section chirurgicale.

Microscopiquement, il est défini par la présence de cellules carcinomateuses de différenciation glandulaire liberkühnienne infiltrant au minimum la musculaire muqueuse. Bien que la majorité des adénocarcinomes colo-rectaux soient sans spécificité apparente, certains sous-types histologiques d'adénocarcinome existent : adénocarcinome mucineux, adénocarcinome à cellules en bague à chaton, carcinome médullaire, adénocarcinome vilieux, adénocarcinome micropapillaire, adénocarcinome adénome-like, carcinome adéno-squameux, carcinome avec composants sarcomatoïdes, carcinome indifférencié.

L'examen immunohistochimique est peu utilisé au diagnostic de la localisation primitive. Cependant, il peut s'avérer crucial dans les prélèvements des localisations métastatiques afin de déterminer l'origine colique de l'adénocarcinome. Dans ce cas, la présence d'un marquage par les anticorps ciblant CK20, CDX2 et SATB2, associée à l'absence de marquage par l'anticorps anti-CK7 sont évocatrices d'un adénocarcinome d'origine colique. Le statut MMR peut également être évalué par une analyse immunohistochimique avec les anticorps ciblant MLH1, MSH2, MSH6 et PMS2, afin de mettre en évidence une perte d'expression d'une ou de plusieurs protéines du système de réparation, conduisant une instabilité des microsatellites qui sera confirmée par analyse de biologie moléculaire.

Des critères histologiques associés à des pronostics particuliers (critères histopronostiques) sont décrits dans les comptes rendus des pathologistes en pratique quotidienne selon les recommandations de l'OMS 5^{ème} édition (20) (Tableau 1).

A titre d'illustrations, la Figure 3 montre du tissu colique sain et quelques sous-types d'adénocarcinome colique. La Figure 4 s'intéresse à quelques critères histopronostiques. La Figure 5 représente une biopsie d'adénocarcinome colique et quelques analyses immunohistochimies sur celle-ci.

De nouveaux critères basés sur des images numérisées analysées par l'intelligence artificielle ont également été décrits, tels que l'Immunoscore (21). Ce dernier correspond à un algorithme évaluant un score basé sur la densité des lymphocytes T marqués par l'anticorps CD8 par rapport aux lymphocytes T marqués par l'anticorps CD3. Un score élevé est témoin d'un meilleur pronostic (21).

Dans le cadre d'un traitement néo-adjuvant, précédant le geste chirurgical, la réponse thérapeutique histologique (*Tumor Regression Grade*, TRG) est à évaluer, et plusieurs scores ont été décrits par l'OMS (22,23). Cependant, il n'existe pas à l'heure actuelle de Gold Standard dans l'évaluation histologique de la régression tumorale devant une absence de reproductibilité de ces résultats. Le score de Dworak modifié, habituellement utilisé au CHUGA, est composé de quatre grades de régression tumorale histologique : absence de régression (TRG 0), régression minimale (TRG 1), modérée (TRG 2), presque complète (TRG 3) et complète (TRG 4).

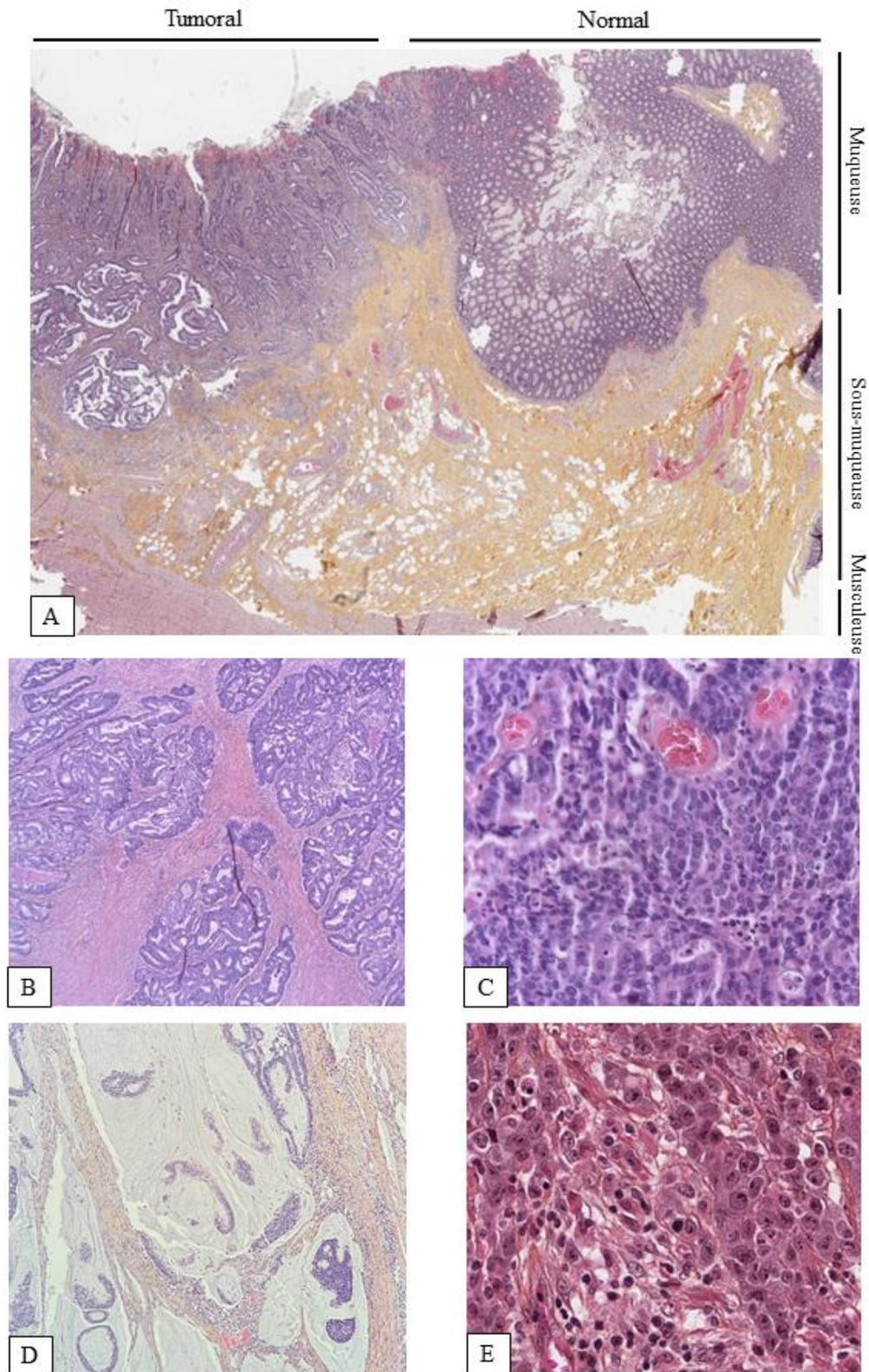


Figure 4. Images histologiques d'une muqueuse normale et de quelques sous-types morphologiques de l'adénocarcinome colique

A) Coupe histologique illustrant un tissu colique infiltré jusqu'à la sous-muqueuse par un adénocarcinome colique. Le tissu normal est à droite et l'adénocarcinome est à gauche (HES faible grossissement) ; B) Adénocarcinome infiltrant bien différencié avec de nombreuses glandes visibles (HES x5) ; C) Adénocarcinome infiltrant peu

différencié formé par une plage solide de cellules cancéreuse, sans formation glandulaire (HES x40) ; D) Adénocarcinome mucineux avec une nette prédominance d'une substance mucineuse faiblement colorée renfermant quelques glandes cancéreuses (HES x20) ; E) Carcinome médullaire dont les cellules cancéreuses ont un cytoplasme éosinophile et un noyau nucléolé accompagnés par un infiltrat lymphocytaire et polynucléaire significatif (HES x40)

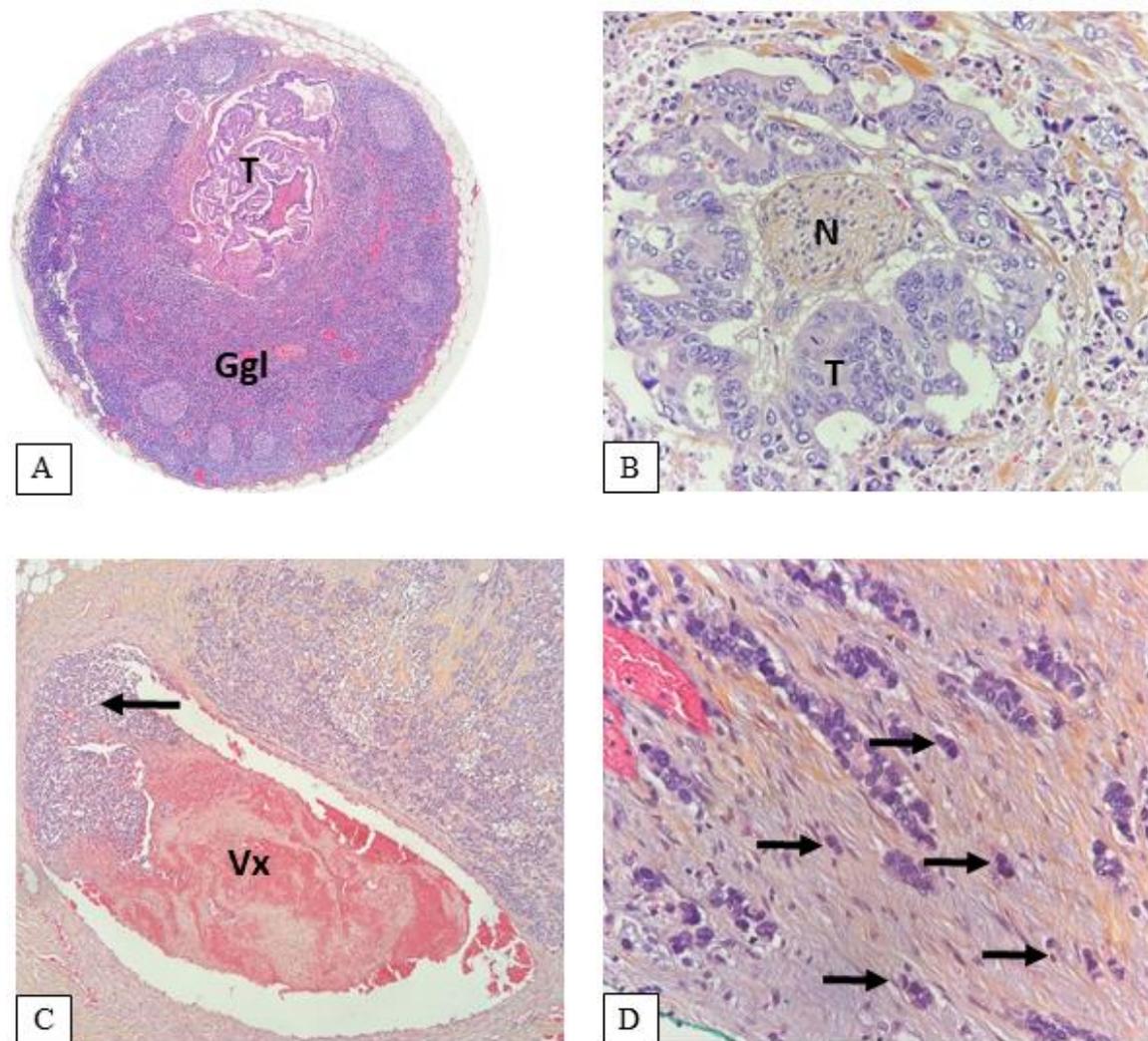


Figure 5. Images histologiques de quelques facteurs histo-pronostiques

A) Métastase d'un adénocarcinome colique bien différencié (T) au sein d'un ganglion lymphatique (Ggl) (HES x5), B) Engainement périnerveux (N : nerf ; T : tumeur) (HES x40), C) Embole veineux (Vx : vaisseaux sanguin) (HES x5), D) Budding (flèches) (HES x40)

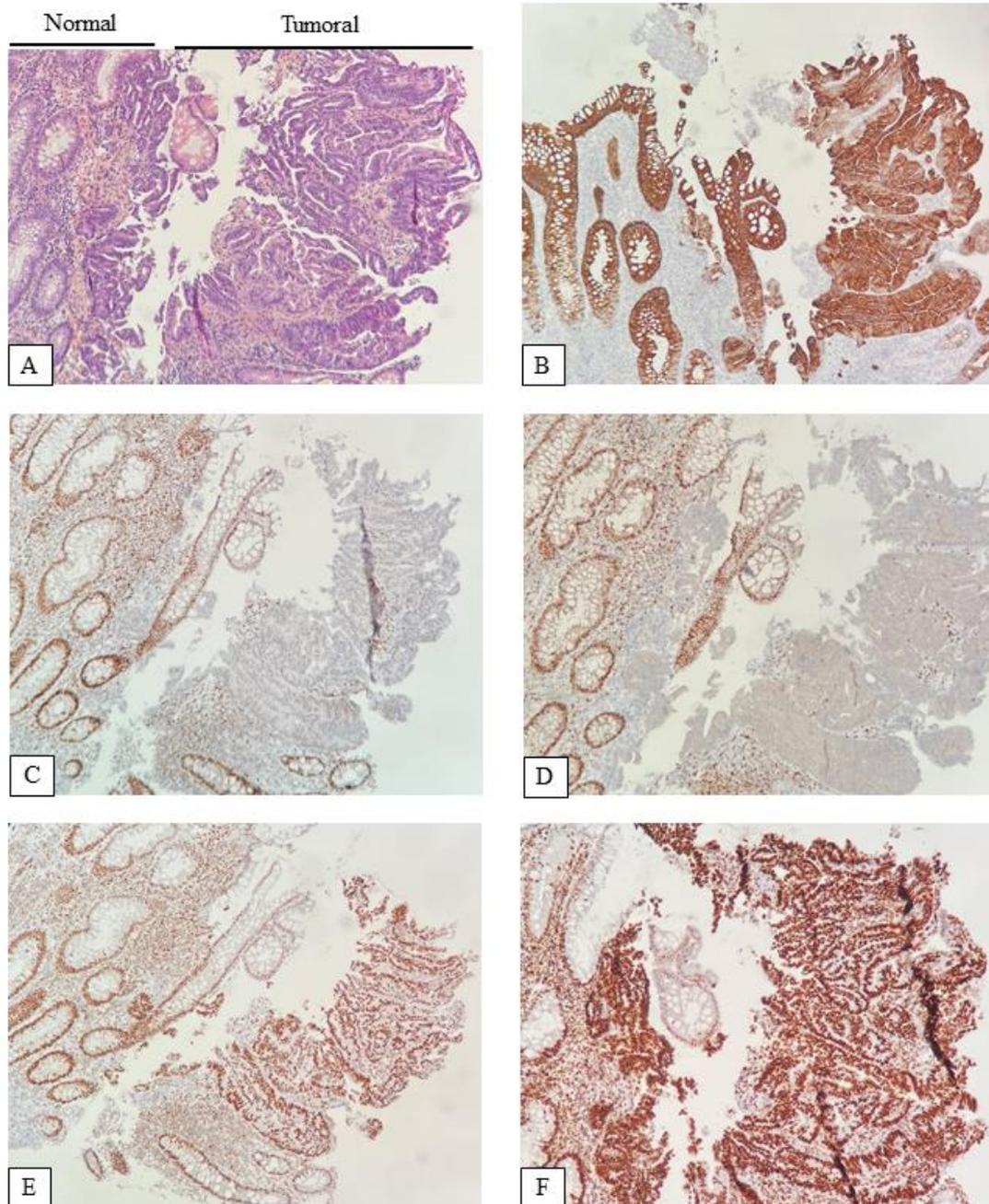


Figure 6. Images histologiques et immunohistochimiques d'une biopsie d'un adénocarcinome colique bien différencié avec perte d'expression de protéines MMR

A) Biopsie colique avec des glandes coliques à gauche et des glandes tumorales bien différenciées à droite (HES x10), B) Marquage immunohistochimique membranaire et cytoplasmique diffus et intense de la CK20, C) Perte de l'expression immunohistochimique nucléaire de MLH1, D) Perte de l'expression PMS2, E) Persistance du marquage nucléaire MSH2, F) Persistance du marquage nucléaire MSH6.

Critères histo-pronostiques	Description
Sous-types histologiques : non spécifique, mucineux, à cellules en bague à chatons, médullaire (20)	<p>Certains adénocarcinomes coliques prennent des aspects architecturaux ou cytologiques particuliers. Ces aspects peuvent être associés à des niveaux de pronostics variables.</p> <p>L'absence de particularité architecturale est ainsi appelée « non spécifique » (NOS) et constitue le principal type architectural de l'adénocarcinome colique.</p> <p>A titre d'exemple, l'adénocarcinome à cellules en bague à chatons présente au moins plus de la moitié des cellules tumorales avec de la mucine intracytoplasmique, et il est associé à un pronostic plus péjoratif par rapport au type non spécifique.</p>
Différenciation et grade (20)	La lésion tumorale peut être bien, moyennement ou peu différenciée. Le grade dépend du degré de différenciation : bas grade (bien ou moyennement différencié) ou haut grade (peu différencié).
Classification pTNM (20)	Cf partie 1.4.3
Embole lymphatique (24)	Présence de cellules tumorales au sein de vaisseaux lymphatiques.
Embole vasculaire (25)	Présence de cellules tumorales au sein de vaisseaux sanguins.
Engainements périnerveux (26)	Présence de cellules tumorales infiltrant un ou des fibres nerveuses.
Nodules tumoraux satellites (27)	Amas de cellules carcinomateuses dans le tissu adipeux péri-colique, sans continuité avec la tumeur, et étant corrélé à un pronostic plus péjoratif.
« Budding » tumoral (28)	Petits amas de moins de cinq cellules carcinomateuses en bordure de la tumeur et étant statistiquement associée à un plus mauvais pronostic.
Clusters peu différenciés (29)	<p>Amas de clusters, d'au moins cinq cellules tumorales, dépourvus d'architecture glandulaire au sein d'un <i>hot spot</i> au grossissement x20.</p> <p>Le résultat est rendu sous la forme : G1 (<5 clusters), G2 (5–9 clusters), G3 (≥10 clusters).</p>
Réponse immune (30)	<p>Infiltration de lymphocytes T au sein de la tumeur (TIL) et/ou autour de la tumeur (réaction inflammatoire péri-tumoral en amas de type Crohn-like ou CLR) qui serait associé à un meilleur pronostic.</p> <p>La CLR est signalée dans le compte rendu par la présence d'au moins trois agrégats lymphoïdes par section.</p> <p>La présence d'une TIL nécessite la présence d'au moins deux lymphocytes intratumoraux sur une moyenne de cinq champs au fort grossissement.</p>

Tableau 1. Facteurs histo-pronostiques reconnus par l'OMS 5^{ème} édition (20)

1.4.3. Classification TNM

La classification pTNM est indispensable en cancérologie et doit apparaître dans le compte rendu du pathologiste. Elle permet de standardiser le stade d'extension locale (identifiée par le « T »), ganglionnaire régionale (« N ») et métastatique (« M ») des cancers, étudié par le pathologiste (« p »). Développée et mise à jour régulièrement par l'UICC (*Union for International Cancer Control*), la 8^{ème} édition de cette classification est la dernière et date de 2017 (31). L'adénocarcinome du côlon et du rectum est ainsi concerné par cette classification et est défini dans le Tableau 2.

Le stade d'atteinte des ganglions lymphatiques (N) est habituellement rendu qu'après étude microscopique d'au moins douze ganglions lymphatiques régionaux. Dans le cadre d'un traitement néoadjuvant, il sera précisé « ypTNM » dans la classification rendue associée à un score de régression tumorale (TRG) de Dworak si celui-ci est réalisable. Il peut également être précisé la présence d'embolie vasculaire (V0/1/2), d'embolie lymphatique (L0/1) ou d'engainement périnerveux (Pn0/1). Dans le cadre des exérèses chirurgicales thérapeutiques, l'état de résection complète, définie en France par l'absence de cellules tumorales au niveau des berges d'exérèse, est également précisé avec « R0 » pour une résection complète et « R1 » pour une résection incomplète.

Par ailleurs, la combinaison des variables du TNM permet de déterminer un stade d'évolution du cancer, corrélé au pronostic, selon le schéma suivant dans le cancer colo-rectal, issu de la classification de la 8^{ème} édition de l'UICC (Tableau 3).

T (extension locale)	N (localisation ganglionnaire régionale)	M (localisation métastatique)
TX : tumeur primitive ne pouvant être évaluée	NX : ganglions lymphatiques ne pouvant être évalués	M0 : Absence de métastase à distance
T0 : absence de cellule carcinomateuse	N0 : absence de métastase ganglionnaire	M1a : métastase confinée à un organe, sans métastase péritonéal
Tis : carcinome <i>in situ</i> : intramuqueux ou invasion de la lamina propria	N1 :	M1b : métastase dans plus d'un organe, sans métastase péritonéal
T1 : envahissement de la sous-muqueuse	- N1a : métastase dans 1 ganglion lymphatique	M1c : métastase dans le péritoine avec ou sans atteinte d'autres organes (carcinose péritonéale)
T2 : envahissement de la musculieuse	- N1b : métastase dans 2 à 3 ganglions lymphatiques	
T3 : envahissement de la sous-séreuse ou dans le tissu péricolique ou périrectal non péritonéalisé	- N1c : dépôts tumoraux (nodules satellites) dans la sous-séreuse ou tissu non péritonisé péricolique ou péri-rectal sans métastase	
T4 :	N2 :	
- T4a : tumeur perforant le péritoine viscéral	- N2a : métastase dans 4 à 6 ganglions lymphatiques	
- T4b : tumeur envahissant directement d'autres organes ou structures	- N2b : métastase dans au moins 7 ganglions lymphatiques	

Tableau 2. Classification clinique et pathologique du TNM de l'adénocarcinome du côlon issu de la 8ème édition de l'UICC (20,31)

Stade	T	N	M
Stade 0			
Stade 0	Tis	N0	M0
Stade I			
Stade I	T1, T2	N0	M0
Stade II			
Stade IIA	T3	N0	M0
Stade IIB	T4a	N0	M0
Stade IIC	T4b	N0	M0
Stade III			
Stade IIIA	T1, T2	N1	M0
	T1	N2a	M0
Stade IIIB	T1, T2	N2b	M0
	T2, T3	N2a	M0
	T3, T4a	N1	M0
Stade IIIC	T3, T4a	N2b	M0
	T4a	N2a	M0
	T4b	N1, N2	M0
Stade IV			
Stade IVA	Tous T	Tous N	M1a
Stade IVB	Tous T	Tous N	M1b
Stade IVC	Tous T	Tous N	M1c

Tableau 3. stades d'évolution de l'adénocarcinome du côlon issu de la 8^{ème} édition de l'UICC (20,31)

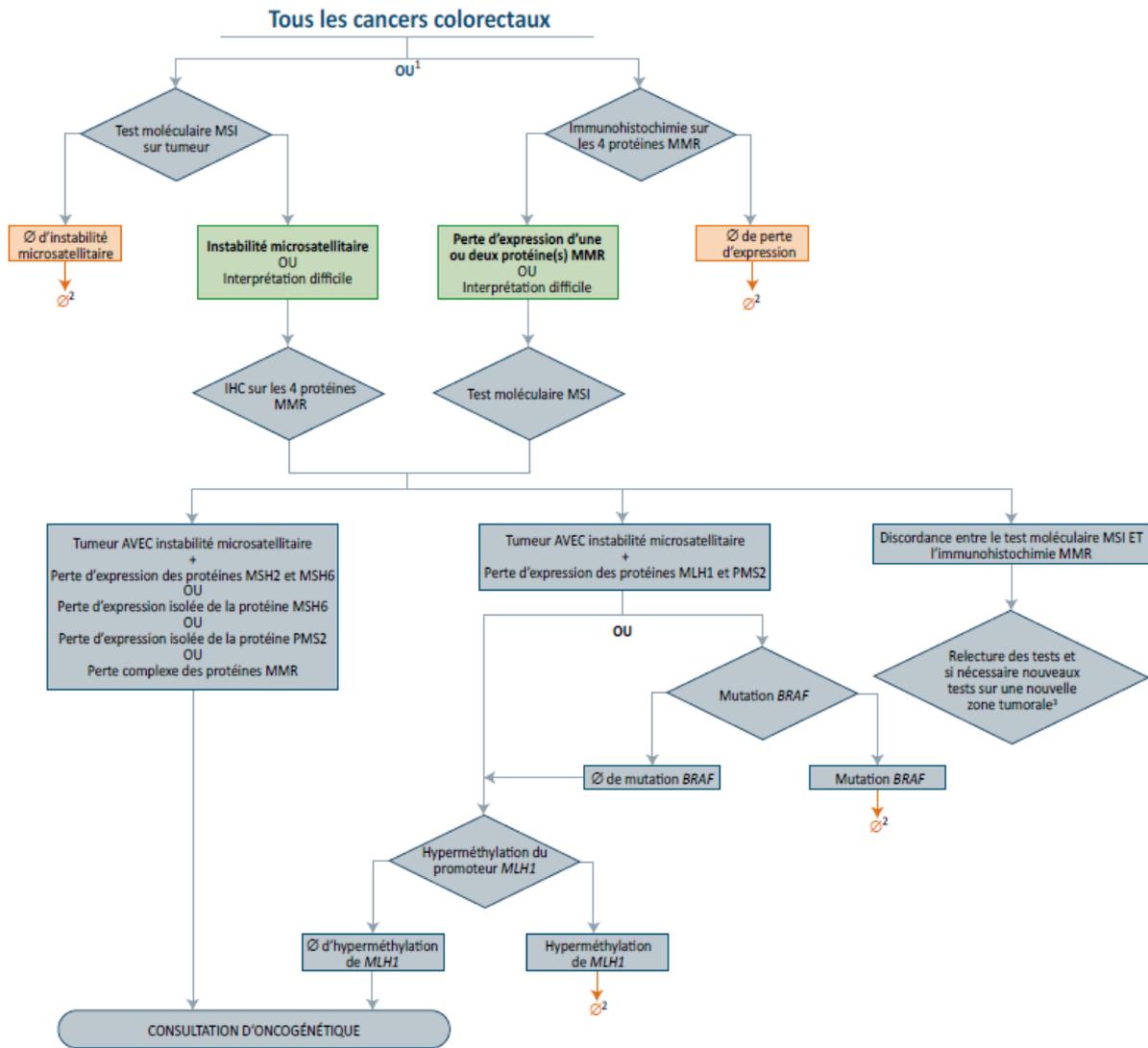
1.4.4. Biomarqueurs recommandés par l'INCa

En 2024, les examens de biologie moléculaire dans le CCR sont limités et leurs indications sont basées sur les recommandations de l'INCa.

Dans le cadre de la recherche systématique du syndrome de Lynch, il est recommandé d'évaluer, indépendamment de l'âge et du contexte familial et personnel du patient, le système de réparation des mésappariements (MMR) par immunohistochimie - pMMR (proficient) ou dMMR (système déficient). Cette recherche pourra être complétée par une étude par biologie moléculaire des instabilités des microsatellites - statut MSS (stable) ou MSI (instable). Elle permet d'orienter les patients et leur famille vers une consultation d'oncogénétique, ainsi que d'organiser une prise en charge thérapeutique personnalisée (notamment pour les stades II et IV) de ces patients (Figure 7) (32).

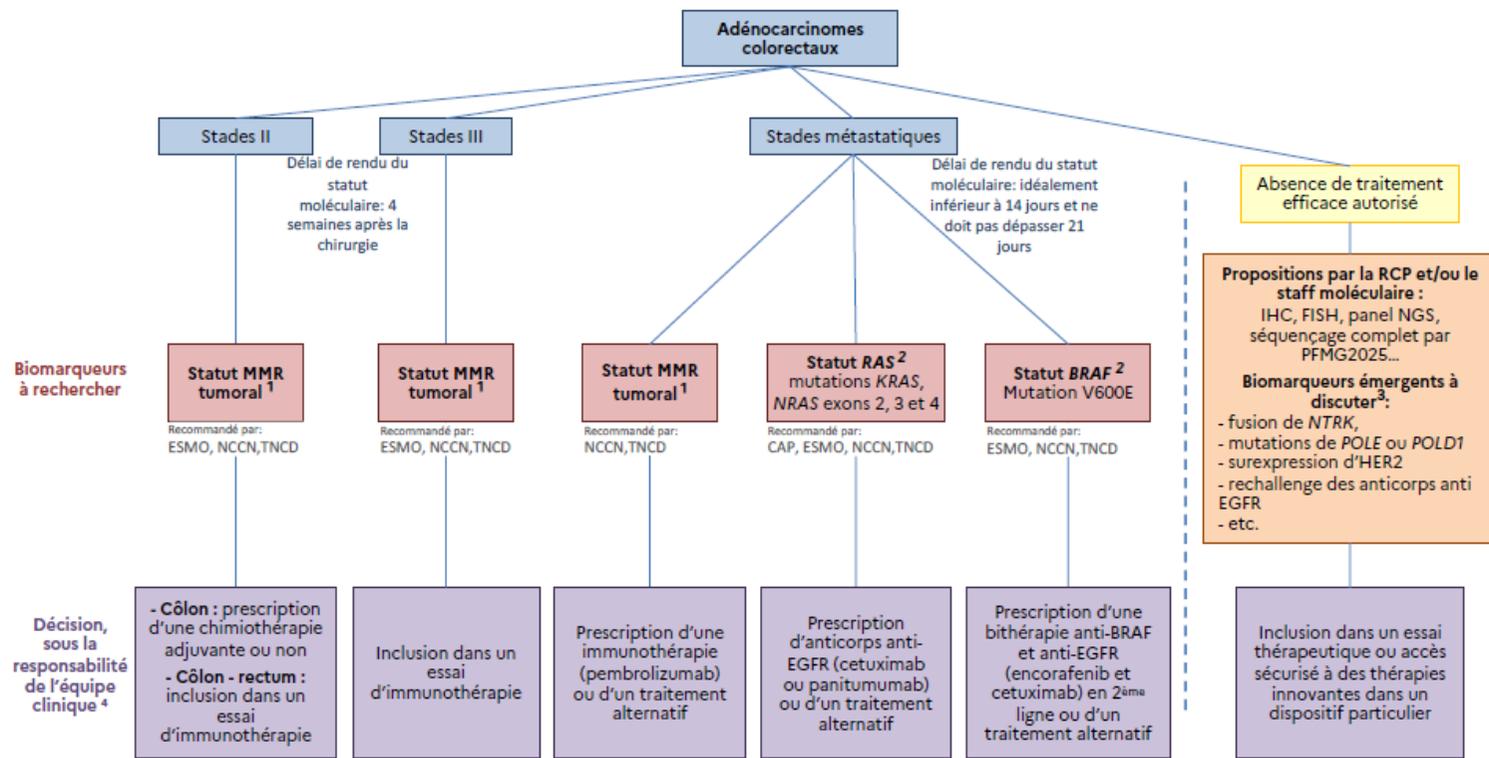
Des biomarqueurs théranostiques sont principalement recommandés et utilisés en situation métastatique (Figure 8) (33) :

- La recherche de mutations des gènes KRAS et NRAS (exons 2, 3 et 4), présentes dans environ la moitié des CCR, permet de prédire l'inefficacité des biothérapies anti-EGFR (HER1), comme le cetuximab (ERBITUX®) ou le panitumumab (VECTIBIX®), dans le cas où ceux-ci sont mutés (34,35) ;
- La recherche de la mutation V600E de BRAF, mutés dans environ 10% des cas, permet d'indiquer la prescription d'une bithérapie anti-BRAF et anti-EGFR à partir de la deuxième ligne de traitement. Cette mutation est également de mauvais pronostic et pourrait également prédire l'inefficacité des biothérapies anti-EGFR, ainsi que la réponse aux chimiothérapies si elle s'associe à des instabilités des gènes microsatellites (34,35) ;
- La recherche d'une instabilité des microsatellites (MSI), présentes dans environ 15% des CCR, permet de prédire la réponse aux biothérapies anti-PDL1 (36).



- ¹ En cas de forte suspicion de syndrome de Lynch ou pour les patients métastatiques potentiellement éligibles à un essai d'immunothérapie, les 2 tests moléculaire et immunohistochimique doivent être réalisés d'emblée. Une stratégie d'analyse combinée des deux techniques d'emblée est également possible.
- ² Un résultat qui n'est pas en faveur du diagnostic de syndrome de Lynch doit néanmoins être interprété en fonction des données cliniques du patient et de ses antécédents familiaux, particulièrement lorsqu'une prédisposition génétique à une pathologie digestive autre que le syndrome de Lynch est suspectée ou envisagée.
- ³ Si une discordance persiste le cas doit être discuté en centre expert entre cliniciens, anatomopathologistes, biologistes moléculaires et/ou oncogénéticiens.

Figure 7. Cascade de réalisation des tests somatiques et indication de consultation d'oncogénétique pour les colorectaux, issue des recommandations INCa 2021 (32)



- 1 - L'évaluation du statut MMR tumoral peut être réalisée par IHC ou biologie moléculaire. Cependant, lorsqu'un traitement avec des inhibiteurs de point de contrôle immunitaire est envisagé, dans un contexte familial évocateur d'un syndrome de Lynch, ou si l'interprétation des analyses est difficile, les deux techniques doivent être réalisées. Si le résultat obtenu avec une technique (IHC ou biologie moléculaire) atteste d'une déficience MMR, il devra être confirmé par l'autre technique. (Recommandations pour l'évaluation du statut MMR tumoral, INCa 2021).
- 2 - En cas d'absence de matériel tumoral ou de délai prévisionnel de rendu de résultats >21 jours, si le patient nécessite une mise sous traitement rapide, la recherche de ce biomarqueur peut être réalisée à partir d'ADNtc. Ce test devra être réalisé dans un centre expert. Si la présence d'ADNtc n'est pas confirmée dans l'échantillon ou s'il n'y a pas de métastases hépatiques, un résultat négatif pour l'identification des mutations recherchées doit être considéré comme non contributif et un nouvel examen à partir d'un échantillon de tissu tumoral doit alors être réalisé.
- 3 - La recherche de ces biomarqueurs émergents ne peut être recommandée de façon indispensable. Elle est à discuter en fonction de la situation du patient et des essais cliniques ouverts. La liste proposée n'est pas exhaustive et est susceptible d'évoluer en fonction des essais cliniques ouverts.
- 4 - Selon les indications des AMM des différentes molécules et les essais cliniques ouverts.

Figure 8. Arbre décisionnel des biomarqueurs nécessaires pour le traitement des patients atteints d'adénocarcinomes colorectaux, issu des recommandations de l'INCa de juin 2022 (33)

En-dehors du cadre des biothérapies, le phénotypage et le génotypage de l'enzyme dihydropyrimidine deshydrogenase (DPD), qui est impliqué dans le métabolisme et dégradation du 5-fluoro-uracile (5-FU), permettent de dépister les patients à risque de surdosage et donc de complications iatrogènes. Ainsi, la pharmacogénétique et la pharmacologie jouent également un rôle clé dans la prise en charge du CC, étant donné que le 5-FU est une des principales chimiothérapies dans celui-ci (37).

L'ensemble de ces informations, utile dans la prise en charge globale et personnalisée du patient, permet de comprendre l'intérêt de la biologie moléculaire dans le CC. Ceci justifiant également la poursuite des recherches de nouveaux biomarqueurs à visée pronostique et théranostique.

1.5. Prise en charge thérapeutique

Les recommandations actuelles du traitement du CC en France proviennent principalement du Thésaurus National de Cancérologie Digestive (TNCD) issu de la Société Nationale Française de Gastro-Entérologie (SNFGE) (38,39). Elles sont différentes entre le colon et le rectum, ce qui nécessite d'être pris en compte dans la réalisation des études de survie. De plus, la prise en charge est différente entre une forme métastatique et une forme non-métastatique au diagnostic initial. Les algorithmes décisionnels recommandés dans la prise en charge du CC sont illustrés en Annexe 1.

La prise en charge thérapeutique prend en compte la présence de métastase (nécessité d'un bilan d'extension exhaustif), le critère d'opérabilité du patient (statut clinique OMS, avis anesthésique, souhait du patient), la localisation de la tumeur, la résécabilité de la lésion (envahissement des organes et des structures de voisinage, du risque d'exérèse incomplète) et du statut moléculaire (MMR, BRAF, RAS).

Le traitement chirurgical est envisagé en première intention dans les formes non-métastatiques et est habituellement suffisant jusqu'au stade II de l'OMS. La chirurgie emportera un segment colique, par voie cœlioscopique idéalement, avec une marge d'exérèse proximale et distale d'au minimum 5 cm, ainsi que du mésocolon et repérage du pédicule vasculaire. Le segment réséqué dépendra de la localisation de la lésion et des recommandations d'expert. Les polypes coliques à risque dégénératif pourront être réséqués par traitement endoscopique avec marges de sécurité latérales et profondes. L'exérèse d'une lésion pT4b emportera les organes et structures envahies. L'exérèse des métastases pourra être envisagée si l'évolution après traitement est efficace et si les lésions sont accessibles par le geste chirurgical. Une

chimiothérapie avant chirurgie (néo-adjuvant) pourra être réalisée en cas de difficulté d'exérèse prévue en première intention et dans les stades avancés.

Le traitement adjuvant sera envisagé à partir du stade II de l'OMS en prenant en compte le pTNM, l'exérèse chirurgicale complète, les facteurs histo-pronostiques et le statut MMR. Les protocoles de chimiothérapies principalement préconisées dans le CC comportent les protocoles FOLFOX (5-fluoro-uracile associée à l'acide folinique et de l'oxaliplatine) et XELOX/CAPOX (association d'oxaliplatine et de capecitabine). Ils seront administrés sur plusieurs mois par cycles de deux ou trois semaines, en fonction de la tolérance du patient.

En cas de carcinose péritonéale, une chimiothérapie intra-péritonéale pourra être envisagée. Basé sur la relation effet-dose, elle consiste à intensifier les concentrations locales des traitements cytotoxiques pour augmenter la destruction cellulaire tumorale et/ou lever des mécanismes de résistance aux médicaments. Deux techniques sont habituellement proposées dans la carcinose péritonéale d'un CC : la chimiothérapie hyperthermique intrapéritonéale (CHIP) ou la chimiothérapie intrapéritonéale pressurisée par aérosol (PIPAC). Une chirurgie complète de cytoréduction consistant en l'exérèse de l'ensemble des lésions péritonéales et hépatiques peut précéder une CHIP. Ces méthodes permettent une amélioration de la survie des patients mais l'efficacité doit être pondérée à ces complications parfois majeures (40).

Les biothérapies sont également utilisées dans le CC, notamment dans les formes métastatiques, en fonction de la localisation de la tumeur, du statut MMR et des mutations RAS et BRAF (38).

1.6. Des résultats contrastés et la nécessité d'une signature pronostique

Grace à une meilleure compréhension de la physiopathologie du cancer, un dépistage national organisé, des traitements standardisés et la présence de biomarqueurs théranostiques, le cancer du côlon connaît une baisse de la mortalité depuis une dizaine d'années avec un meilleur contrôle des cas avancés. Néanmoins, le cancer du côlon reste le 2^{ème} cancer en termes de mortalité et on estime que 15%-30% des patients sont diagnostiqués à un stade métastatique d'emblée et que 20% à 50% des patients initialement diagnostiqué à un stade localisé vont développer des métastases (41).

Les registres américains du réseau SEER de 1992 à 2004 ont permis d'évaluer la survie à 5 ans après chirurgie à visée curative des stades IIA, IIB et IIC à 87%, 80% et 58% respectivement (42). Cette même étude estime que des tumeurs T4 sans envahissement

ganglionnaire (stade II) pouvait avoir une survie à 5 ans inférieure à des tumeurs moins infiltrantes avec un envahissement ganglionnaire (stade III) (42,43). Il est important de remarquer qu'il est recommandé d'évaluer l'indication de la chimiothérapie adjuvante dans le cancer du côlon de stade II à l'aide de critères cliniques et histopronostiques nombreux et à l'appréciation des médecins. Il serait intéressant dans cette situation de trouver une aide aux cliniciens qui permettrait d'aiguiller de façon standardisée l'indication de la chimiothérapie adjuvante (44). Par ailleurs, dans les stades IVC, avec carcinoses péritonéales, un traitement par chimiothérapie intrapéritonéale peut être proposé et l'éligibilité des patients à ce traitement peut se discuter (40).

Ainsi, améliorer la prévision du pronostic des patients à l'aide d'une signature pronostique permettrait d'aider le clinicien dans le choix des traitements et du suivi les plus adaptés au patient. En plus des sous-types moléculaires actuellement définis, des signatures pronostiques ont également été proposées dans le cancer du côlon, et plus particulièrement pour les CC de stade II. Basées sur des analyses et points de vue moléculaires différents, ces signatures peuvent comprendre des longues sections non codantes d'ARN (lncRNA) ou des microARN, pouvant nécessiter des analyses sur puces à ADN, des techniques d'amplification ou FISH sur du matériel congelé ou fixé. Une liste non exhaustive est proposée en Annexe 2. Actuellement, en pratique, l'identification du sous-type moléculaire ou la recherche d'une signature pronostique ne sont pas réalisées, en-dehors de la recherche d'instabilités des microsatellites dans tous les CC et de la recherche de mutations BRAF ou RAS en situation métastatique.

La recherche d'une nouvelle signature pronostique dans le CC, réalisable en pratique anatomopathologique quotidienne et corrélée à des connaissances cliniques, moléculaires et thérapeutiques, permettra une prise en charge adaptée au patient.

2. Expression ectopique de gènes spécifiques de tissus de tissus en cancérologie

2.1. Généralité

L'étude du génome des cancers, notamment de l'expression des gènes dans les lésions cancéreuses ou pré-cancéreuses est le cœur de la recherche en cancérologie. En effet, ces derniers joueraient un rôle crucial en permettant aux cellules tumorales d'acquérir de nouvelles propriétés, qui peuvent participer à l'immortalisation de ces cellules et à l'adaptation à son environnement (Figure 9). Contourner la réaction immunitaire (échappement immunitaire),

favoriser la prolifération vasculaire au sein de la lésion (néovascularisation) ou favoriser l'entrée dans le cycle de prolifération cellulaire sont par exemple des propriétés favorisant les cellules cancéreuses à proliférer et à devenir invasif. Ces nouvelles caractéristiques peuvent être de potentielles cibles thérapeutiques (45).

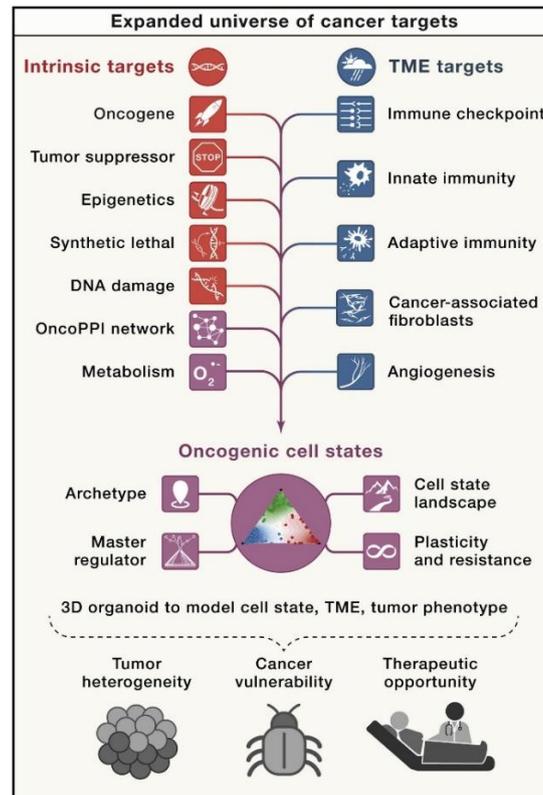


Figure 9. Voies de cancérogénèse et l'interaction du micro-environnement tumoral pouvant constituer des cibles potentielles pour les biothérapies (45)

TME : *tumor microenvironment*

Parmi les gènes exprimés dans une prolifération tumorale, on peut retrouver des gènes dits spécifiques de tissus dont l'expression physiologique est limitée à un ou quelques tissus normaux. L'expression de ces gènes dans une prolifération tumorale au sein d'un tissu n'exprimant pas ou très peu est ainsi considérée comme une expression « ectopique » ou « aberrante » dans la littérature scientifique. Les gènes exprimés physiologiquement et uniquement durant l'embryogénèse, dans le placenta et/ou le testicule, s'appellent « *cancer/testis antigens* » (CTA).

Des banques de données ont été construites pour répertorier les informations et les conséquences des expressions en cancérologie. Par exemple, la base de données CTdatabase contient une liste des principaux CTA (46). Les banques de données publiques telles que Human

Protein Atlas (47) et le National Center for Biotechnology Information (NCBI) (48) permettent d'avoir une vue d'ensemble des connaissances biologiques sur les gènes.

2.2. Impact sur la cancérogénèse et le pronostic

Les travaux de recherche menés à l'Institut pour l'Avancée des Biosciences (IAB) dans l'équipe du Dr. Saadi Khochbin (49) ont établi que l'apparition des expressions ectopiques dans les cancers suite à des dérégulations majeures du système de signalisation épigénétique pourraient avoir des conséquences sur la cancérogénèse et l'évolution tumorale, et par conséquent avoir une influence sur le pronostic du patient.

En se basant sur ce postulat, de nombreux projets de recherche menés en collaboration entre les équipes médicales du CHU Grenoble Alpes (CHUGA) et l'équipe de recherche à l'IAB ont démontré un impact pronostique des expressions ectopiques de gènes spécifiques de tissus dans différents types de cancers, en utilisant un algorithme bio-informatique, développé à l'IAB et basé sur une analyse de données d'expression de gènes. Cette approche a permis d'identifier des gènes dont l'activation « hors contexte » aurait un impact favorable ou défavorable sur le pronostic des cancers du poumon (50), des carcinomes épidermoïdes de la cavité buccale (51), des cancers du sein (52), des lymphomes B diffus à grandes cellules (53) et des leucémies aiguës lymphoblastiques à cellules B et T (54,55). Il est intéressant de noter que l'expression ectopique d'un même gène dans des cancers de localisation différente peut avoir un impact pronostique différent (49,53). L'activation aberrante de certains gènes spécifiques de tissus dans les cancers est significativement associée à des formes particulièrement agressives des tumeurs, de pronostic vital défavorable.

L'expression ectopique de certains gènes permettent aux cellules tumorales d'acquérir de nouvelles propriétés tels que l'expression du CTA NUT dans le carcinome NUT qui modifie complètement l'acétylation chromatinienne CBP/p300-dépendante (56). L'expression de la protéine CYCLON dans les lymphomes B diffus à grandes cellules impacteraient la réponse au Rituximab et la croissance cellulaire (53).

L'algorithme bio-informatique mise au point par l'IAB n'a pour l'instant jamais été utilisé en cancérologie digestive.

2.3. Intérêt diagnostic

L'expression ectopique de certains gènes spécifiques de tissus peut s'avérer utile au diagnostic de cancers et parfois expliquer la physiopathologie de la maladie (57). Par exemple,

l'expression ectopique du gène NUT, physiologiquement limitée au testicule, établit le diagnostic du carcinome NUT et permettrait d'expliquer sa carcinogenèse (56). De même, la translocation du gène SS18 avec un gène spécifique de tissus SSX (SSX1, SSX2, SSX4) est un marqueur moléculaire pathognomonique des synoviosarcomes (58). L'étude des CTA dans les cancers du testicule permet également de préciser des sous-types selon leur expression à travers l'immunohistochimie dans les tumeurs spermatocytiques (59). Enfin, l'utilisation de l'immunohistochimie par l'anticorps anti-PRAME (*PReferentially expressed Antigen of Melanoma*) peut être un outil d'aide au diagnostic du mélanome, du synoviosarcome ou du liposarcome myxoïde (60).

2.4. Intérêt thérapeutique

Les biothérapies constituent depuis quelques années une voie thérapeutique innovante permettant de cibler un ou plusieurs antigènes afin de réduire son expression. En cancérologie, ces traitements permettent d'atteindre des cibles parfois non affectées par la chimiothérapie avec l'avantage de réduire les effets indésirables.

Ainsi, l'expression ectopique des gènes spécifiques de tissus peut constituer une voie intéressante dans les biothérapies. L'expression physiologique des CTA étant limitée au testicule et au placenta (en-dehors de l'embryogénèse), les biothérapies les ciblant auraient ainsi une limitation des effets secondaires avec une spécificité importante contre le tissu tumoral (61,62). A titre d'exemple, la famille MAGE et la famille NY-ESO-1 dans lesquels de nombreuses études assez avancées et prometteuses sont en cours (63,64).

De plus, la synergie avec les chimiothérapies et biothérapies actuelles pourrait être également intéressante, telle que l'exemple de la Decitabine (analogue du désoxynucléotide cytidine) qui permettrait l'expression d'antigènes dont des CTA comme le NY-ESO-1 et ainsi rendre possible l'accès aux biothérapies anti-CTA dans des cancers initialement non expresseurs (65).

3. Objectifs du travail de thèse

Malgré les connaissances actuelles et une prise en charge standardisée par le TNCD et l'INCa en France, la fréquence et la létalité du cancer du côlon restent importantes et nécessitent de plus amples recherches. Le développement d'un marqueur prédictif du pronostic des patients à des stades localement avancés (stade II) ou à des stades métastatiques (stade IV) représentent une avancée considérable dans cette lutte. Par ailleurs, l'étude de l'expression ectopique de

gènes spécifiques de tissu dans les cancers semble être un point de vue intéressant dans l'exploration du pronostic, du diagnostic et probablement du traitement de certains cancers.

Ainsi, le travail de thèse s'intéresse à déterminer et à caractériser dans le cancer du côlon l'activation aberrante des gènes spécifiques de tissu, normalement non exprimés ou faiblement exprimés dans le colon sain, et si cela est le cas, de pouvoir mettre en application l'algorithme bio-informatique développé par l'IAB dans le cancer du côlon afin d'évaluer si cette expression est significativement associée à un pronostic défavorable. L'apport de nouvelles connaissances sur le pronostic en cancérologie digestive a pour objectif d'optimiser la prise en charge des patients et de mettre en évidence de nouvelles cibles thérapeutiques potentielles.

L'objectif principal de cette thèse consiste à identifier la présence d'expressions ectopiques de gènes spécifiques de tissus dans le CC à partir de données publiques d'expression de gènes et à concevoir un outil stratifiant le risque pronostique basé sur le nombre de gènes exprimés de façon aberrante. Les objectifs secondaires sont de consolider les résultats de l'impact pronostique en approche multivariée, en prenant notamment en compte le stade UICC et d'autres facteurs histo-pronostiques connus. Ensuite, la valeur pronostique de l'outil à l'aide d'analyses par RT-qPCR et immunohistochimiques sera confirmée sur une cohorte rétrospective indépendante du CHUGA élaborée spécifiquement.

Il s'agit d'un travail collaboratif entre le département d'Anatomie et de Cytologie Pathologiques, dirigé par le Pr Nathalie STURM du CHU Grenoble Alpes (adresse : Site Santé – Boulevard de la Chantourne, 38700 La Tronche), et de l'unité de recherche INSERM U1209, dans le département Signalisation et chromatine, dans l'équipe « Régulations épigénétiques », dirigée par le Dr Saadi KHOCHBIN de l'IAB, Centre de Recherche UGA, CNRS UMR 5309 (adresse : Site Santé – Allée des Alpes, 38700 La Tronche).

MATÉRIEL ET MÉTHODE

1. Étude bio-informatique

1.1. Matériel : cohortes publiques

1.1.1. Critères de sélection

Différents jeux de données d'expression de gènes issus de la littérature scientifique sont extraits à partir des banques de données internationales telles que le Gene Expression Omnibus (GEO) du National Center for Biotechnology Information (NCBI) et le GDC Data Portal (66). Ces biobanques en libre accès permettent d'obtenir les données de biologie moléculaire des cohortes de patients issues des études scientifiques publiées.

D'après les critères de sélection appliqués dans ce projet, les cohortes retenues doivent contenir au moins 100 échantillons de patients présentant un adénocarcinome du colon (« *colon adenocarcinoma* » ; « *colon cancer* ») avec des données cliniques de survie globale et des données de biologie moléculaire disponibles. Ces dernières sont retenues si elles sont extraites par des techniques de puce à ADN ou de séquençage haut-débit (critère « *study type : expression profiling by array + high throughput sequencing* » sur GEO).

Nous avons ainsi identifié trois cohortes publiques du CC : TCGA-COAD (séquençage de l'ARN, n=439), GSE39582 (puce à ADN, n=585) et GSE17536 (puce à ADN, n=177). Des informations complémentaires telles que le stade UICC sont récupérées. Les caractéristiques démographiques des 3 cohortes sont rapportées présentées dans le Tableau 4.

Un jeu de données transcriptomique supplémentaire du portail public GTEX (<https://gtexportal.org/home>) qui contient environ 3000 échantillons de 48 tissus normaux a également été utilisé pour l'identification des gènes spécifiques de tissus.

Cohorte	TCGA-COAD	GSE39582	GSE17536
Effectif	439	585	177
Tumeur, n (%)	398 (91%)	566 (97%)	177 (100%)
Non tumeur, n (%)	41 (9%)	19 (3%)	0 (0%)
Méthode analyse	Séquençage ARN	Puce à ADN	Puce à ADN
Age, médiane [IQR]	68,2 [57,7-76,4]	68,1 [59,0-76,0]	66,0 [57,0-75,0]
<i>MD</i>	0	1	0
Sexe, n (%)			
Homme	204 (51,3%)	310 (54,8%)	96 (54,2%)
Femme	194 (48,7%)	256 (45,2%)	81 (45,8%)
<i>MD</i>	0	0	0
Localisation, n (%)			
Côlon droit	174 (43,7%)	224 (39,6%) *	-
Côlon gauche	120 (30,2%)	342 (60,4%) *	-
<i>MD</i>	104	0	177
Stade UICC, n (%)			
0	0 (0%)	4 (0,7%)	0 (0%)
I	68 (17,1%)	33 (5,8%)	24 (13,6%)
II	160 (40,2%)	264 (46,6%)	57 (32,2%)
III	108 (27,1%)	205 (36,2%)	57 (32,2%)
IV	52 (13,1%)	60 (10,6%)	39 (22%)
<i>MD</i>	10	0	0
N+, n (%)	156 (39,2%)	238 (44,1%)	96 (54,2%) **
<i>MD</i>	0	26	0
Sous-type, n (%)			
NOS	338 (84,9%)	566 (100%)	177 (100%)
Mucineux	60 (15,1%)	0 (0%)	0 (0%)
<i>MD</i>	0	0	0
Récidive, n (%)	45 (11,3%)	177 (31,6%)	36 (24,8%)
<i>MD</i>	0	6	32
Décès, n (%)	79 (19,8%)	191 (34,0%)	73 (41,2%)
<i>MD</i>	0	4	0

Tableau 4. Caractéristiques démographiques des cohortes issues de la littérature
IQR : écart interquartile ; MD : données manquantes ; * information donnée en localisation « proximale »
et « distale » ; ** information obtenue en associant les stades III et IV

1.1.2. Préparation et normalisation des données transcriptomiques

Les données d'expression de gènes ont été téléchargées à partir des bases de données publiques. Elles ont été ensuite prétraitées et normalisées par le personnel de la plateforme EpiMed à l'IAB.

Les données transcriptomiques des jeux de données publics GSE39582 et GSE17536 ont été obtenues avec la technologie de puces à ADN Affymetrix Human Genome U133 Plus 2.0 Arrays. Les données ont été normalisées en utilisant la méthode Robust Multi-array Average (RMA) (67), puis transformées en pseudo-logarithme en base 2.

Pour le jeu de données TCGA-COAD, obtenu avec la technologie RNA-seq, nous avons utilisé les valeurs de niveaux d'expression normalisées par la méthode FPKM, directement fournies par le portail de données GDC (<https://portal.gdc.cancer.gov>). Les valeurs FPKM ont été transformées en pseudo-logarithme en prenant $\log_2(1+FPKM)$.

Pour le jeu de données de tissus normaux GTEX, nous avons effectué un prétraitement complet à partir des données brutes RNA-seq. Les *reads* ont été alignés sur le génome de référence humain à partir des fichiers FASTQ de séquences brutes, en utilisant le logiciel STAR (68). Les *reads* alignés ont été ensuite comptés en utilisant le logiciel HTSeq (69). À partir des comptages de *reads*, les valeurs RPKM ont été calculées et transformées en pseudo-logarithme en prenant $\log_2(1+RPKM)$.

1.2. Analyses statistiques des données

1.2.1. Déroulé de l'étude

Les étapes principales du projet de thèse sont présentées dans la Figure 10. L'étude commence par une analyse bio-informatique des jeux de données publiques qui consiste à identifier d'abord les gènes spécifiques de tissus, à sélectionner ensuite les gènes fréquemment activés dans le CC, à découvrir des biomarqueurs candidats associés au pronostic vital des patients et à établir finalement une signature pronostique, appelée *Gene Expression Classifier* (GEC). Ce travail bio-informatique correspond aux étapes de 1 à 3 sur le schéma de la Figure 10. Cette approche a été initialement développée à l'IAB pour les données de puces à ADN (50) et ensuite mise à jour et adaptée également aux données de séquençage par RNA-seq (51,52).

La dernière étape du projet (étape 4 dans la Figure 10) vise à construire une nouvelle cohorte rétrospective du CHUGA afin de valider, à terme, la signature GEC dans cette cohorte par la technologie RT-qPCR. Le projet de thèse comprend le travail sur la constitution de la cohorte du CHUGA, l'estimation du nombre d'échantillons nécessaires et des tests biologiques de faisabilité. Les mesures des niveaux d'expression par RT-qPCR pour toute la cohorte du CHUGA et les analyses statistiques de ces données seront réalisées *a posteriori*, dans le cadre du projet de recherche SIROCCO.

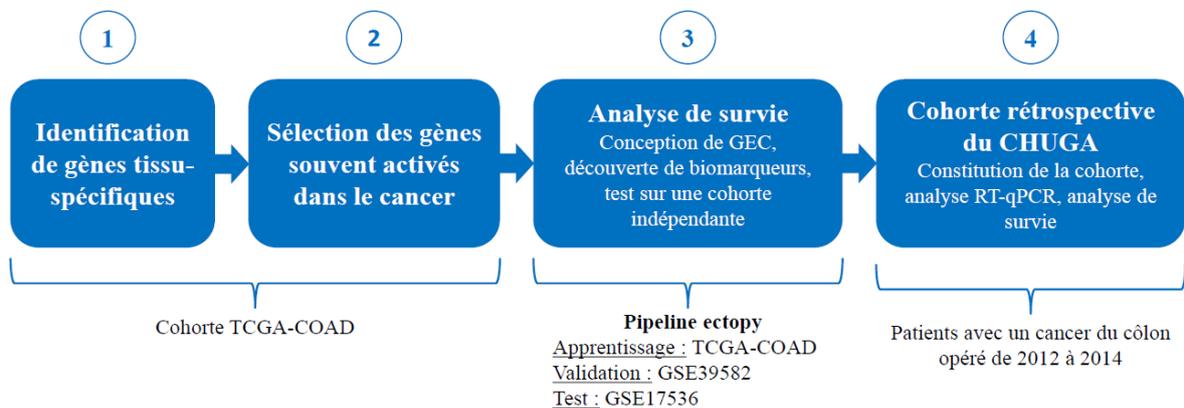


Figure 10. Étapes principales du pipeline « ectopy » qui permet de découvrir des biomarqueurs pronostiques (Gene Expression Classifier, GEC) à partir des données d'expression de gènes

1.2.2. Identification des gènes spécifiques de tissus non exprimés dans le côlon normal

La première étape consiste à identifier des gènes spécifiques de tissus au sein de la cohorte GTEX, en respectant les critères suivants :

- Les gènes doivent être exprimés de façon prédominante dans un tissu autre que le côlon et non exprimés ou faiblement exprimés dans les autres tissus somatiques adultes. La prédominance d'expression est déterminée avec la technique de z-score, calculée pour les niveaux moyens d'expression de chaque gène dans 48 tissus disponibles dans le jeu de données GTEX. Le z-score est égal à la différence entre le niveau d'expression du gène dans chaque tissu et le niveau d'expression moyen, divisée par l'écart-type. Le tissu dans lequel le z-score est supérieur au seuil de 60% par rapport au z-score maximal est identifié comme le tissu de prédominance pour le gène correspondant (52).
- L'expression dans le tissu de prédominance (correspondant au niveau d'expression maximal du gène dans un tissu extra-colique) doit être au moins 4,5 fois plus importante que son expression dans le tissu colique normal.

Au total, cette étape a permis d'identifier 3554 gènes spécifiques de tissus qui ne sont pas exprimés dans le côlon normal.

1.2.3. Identification des gènes spécifiques de tissus fréquemment activés dans les CC

Pour chacun des gènes spécifiques de tissus identifiés, la fréquence d'activation aberrante de ces derniers dans l'adénocarcinome colique est calculée en utilisant les données de la cohorte TCGA-COAD qui contient à la fois des échantillons de tumeur et des échantillons non-tumoraux du colon. Pour chaque gène, un seuil empirique est établi, égal à la moyenne d'expression dans les échantillons non-tumoraux et de 2 écart-types (seuil $m2sd$). La fréquence d'activation aberrante correspond au pourcentage d'échantillons tumoraux dans lesquels le niveau d'expression du gène est supérieur à ce seuil. Les gènes sont considérés comme fréquemment activés si leur fréquence d'activation est supérieure à 10% dans les échantillons tumoraux.

1.2.4. Création d'un outil pronostique *Gene Expression Classifier* (GEC)

Parmi les gènes sélectionnés, il est appliqué une approche originale de *machine learning* pour identifier un panel de biomarqueurs, appelé *Gene Expression Classifier* (GEC), qui prédit de façon robuste et reproductible le pronostic des patients atteints du cancer du côlon. Cette approche a été développée dans l'équipe « Régulations épigénétiques » de l'IAB et

implémentée en langage Python sous le nom de package « ectopy ». Le code *open source* du logiciel « ectopy » est publiquement disponible dans le dépôt Github (70).

L'idée principale de cette approche est basée sur l'hypothèse que l'expression (« ON ») de certains gènes spécifiques de tissus peut permettre aux cellules tumorales d'acquérir de nouvelles propriétés d'agressivité, conduisant à un mauvais pronostic de survie pour les patients.

Le pipeline « ectopy » évalue l'impact pronostique des activations ectopiques des gènes candidats sur la probabilité de survie des patients. En suivant les principes du *machine learning*, les jeux de données de cancer sont séparés en plusieurs *datasets* (Figure 11) :

- Un *dataset d'entraînement* qui sert à créer un modèle de prédiction et à identifier les gènes candidats les plus robustes sur la base d'une activation OFF/ON. L'état OFF correspond à des niveaux d'expression d'un gène inférieurs à la valeur seuil et l'état ON correspond à des niveaux d'expression supérieurs à la valeur seuil. Les gènes candidats marqueurs présentent un état d'expression OFF/ON qui est associé à une différence significative entre les probabilités de survie. Les gènes candidats les plus robustes sont sélectionnés suite à des validations croisées au sein du dataset d'entraînement.
- Un ou plusieurs *datasets de validation* qui servent à confronter les gènes candidats identifiés en phase d'entraînement à une hétérogénéité possible entre différents jeux de données de cancer. Les gènes candidats qui montrent les résultats les plus stables dans tous les jeux de données (entraînement et validation) sont retenus comme biomarqueurs finaux. Les biomarqueurs ainsi identifiés sont combinés dans un panel de gènes prédictif du pronostic, appelé Gene Expression Classifier (GEC).
- Un ou plusieurs *datasets de test* qui vont servir à tester le panel de biomarqueurs GEC dans des jeux de données complètement indépendants, jamais utilisés lors de l'identification des gènes du panel GEC.

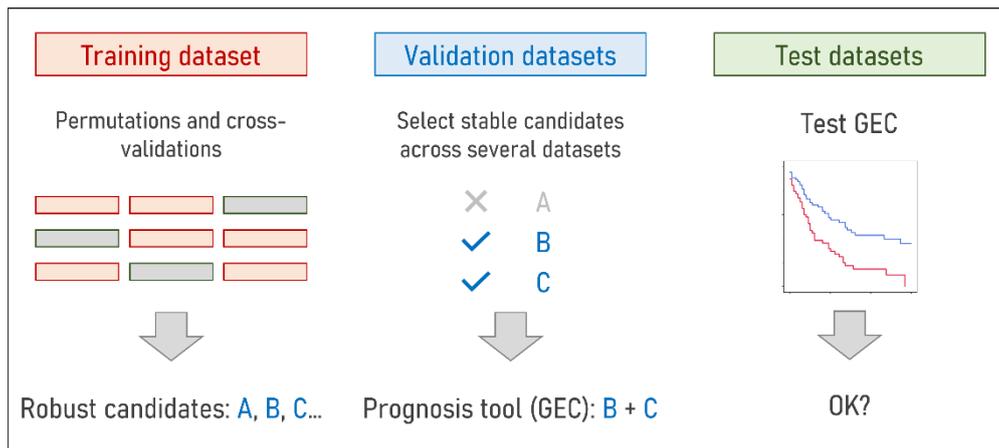


Figure 11. Les étapes principales d'entraînement, de validation et de test des modèles prédictifs de survie dans la méthode « ectopy »

L'introduction des approches du *machine learning* assure une meilleure stabilité des résultats de prédiction dans différents jeux de données de cancer, même quand ces jeux de données présentent une hétérogénéité importante. Cette hétérogénéité est fréquente dans les jeux de données publics et peut être liée, par exemple, à une composition différente en sous-types moléculaires des tumeurs, à des différences dans la prise en charge thérapeutique des patients qui évolue dans le temps, ainsi qu'à d'autres causes possibles lors de la création des *datasets* (provenance de différents pays, de différentes périodes de temps, critères d'inclusion des patients, méthodes diagnostiques disponibles). Le pipeline « ectopy » déploie de nombreux cribles de sélection des gènes candidats lors des phases d'entraînement et de validation. Ceci permet de réduire un surapprentissage éventuel du modèle, l'apparition d'artefacts et de faux positifs dans les résultats.

A l'issue de l'étape de validation, les biomarqueurs retenus sont combinés pour créer un outil de prédiction nommé Gene Expression Classifier (GEC) qui stratifie les patients selon le nombre de gènes activés dans le panel. Chaque gène du panel contribue à la prédiction avec le même poids. Les patients dont la tumeur n'exprime aucun gène du panel GEC ont un pronostic favorable. Les patients dont la tumeur exprime un ou plusieurs gènes du panel auront un pronostic plus sombre. Plus grand est le nombre de gènes du panel activés (ON) de manière ectopique, moins longue est la survie des patients.

A la dernière étape du pipeline, l'outil GEC est testé dans des jeux de données de test indépendants qui n'ont pas été utilisés dans l'identification des gènes du panel GEC. Le résultat de ces tests donne le verdict définitif sur l'efficacité de l'outil et sur son application possible en pratique clinique.

2. Étude sur une cohorte rétrospective du CHUGA

2.1. Constitution de la cohorte

Pour confirmer l'impact pronostic dans le CC, il est nécessaire de mettre à l'épreuve les gènes candidats au sein d'une cohorte de patients issue du CHU Grenoble-Alpes dans le cadre d'une étude de survie évaluant la reproductibilité des résultats par des méthodes d'analyse différentes et couramment utilisées en pratique quotidienne de laboratoire : RT-qPCR et immunohistochimie. Dans ce travail de thèse, seule la réalisation de RT-qPCR sur quelques échantillons a été effectuée pour démontrer la faisabilité de l'approche.

Ainsi, une cohorte de patients dont la pièce opératoire parvenue au laboratoire d'anatomopathologie du CHUGA est conçue. Les critères d'inclusion nécessitent des patients de plus de 18 ans avec un adénocarcinome colique opéré entre 2012 et 2014 inclus et pour lesquels des échantillons tumoraux et des données cliniques sont disponibles. La non-opposition des patients est également un critère recherché systématiquement. Les échantillons tumoraux doivent comporter des blocs de tissus tumoraux et non-tumoraux fixés dans le formol et inclus en paraffine.

L'estimation du nombre de sujet nécessaire est réalisée par l'approche théorique développée par Schoenfeld en 1983 (71) qui est applicable aux analyses de survie utilisant un modèle statistique de risque proportionnel de Cox. Les calculs sont réalisés par un langage informatique Python avec le *package* « *lifelines* ». Les proportions de patients des différents sous-types et les « *hazard ratios* » (HR) sont estimées au travers de résultats préliminaires issus des cohortes extraites des données publiques enregistrées sur GEO.

Ainsi, en prenant un HR de 2,1 pour un risque alpha de 5% et une puissance statistique de 80%, il est nécessaire d'obtenir une cohorte d'au moins 134 patients afin de valider l'outil pronostic GEC. Néanmoins, afin de réaliser une analyse multivariée en prenant en compte le stade UICC dont le critère le plus discriminant étant la présence ou non de métastase ganglionnaire lymphatique (pN0 ou pN1/2), un nombre total de 193 échantillons est nécessaire, avec au moins 110 échantillons pN0 et 83 échantillons pN1/2.

Les variables recueillies sont prédéfinies avant le début de l'étude et regroupent des données cliniques et des données anatomopathologiques. On retrouve ainsi :

- Démographie et métrologie : âge au diagnostic (en année), sexe ;
- Clinique : date du diagnostic (mois et année), date de chirurgie (mois et année) ;

- Facteurs de risque : antécédent de tabagisme (actif ou sevré), antécédent de polypes, de polypose ou de maladie inflammatoire chronique de l'intestin ;
- Antécédents digestifs personnels (adénome ou CC) ;
- Antécédents familiaux (si oui, lesquels et duquel degré familial) ;
- Données chirurgicales : nombre et dates d'interventions, localisation, qualité de l'exérèse (R0 ou R1) ;
- Traitements : traitement de première et seconde intention (néoadjuvant et adjuvant) ;
- Inclusion dans un essai thérapeutique (si oui, lequel) ;
- Données anatomo-pathologiques : le sous-type histologique, le stade pTNM, les facteurs histopronostiques présents dans le compte rendu anatomopathologique : emboles vasculaires, engainements périnerveux, nodules satellites, statut MMR par immunohistochimie ;
- En cas de traitement néoadjuvant, le score de régression tumorale histologique TRG (Dworak modifié) ;
- Données moléculaires : RAS, BRAF, MSS/MSI ;
- Suivi (mois, année) : date des dernières nouvelles, date de récurrence, date de décès ;
- Cause du décès : en lien avec le cancer ou indépendant.

Les variables nouvellement acquises comportent :

- Les paramètres histopronostiques non précisés dans le compte rendu et évalués à la relecture des lames ;
- Résultats de la détection par RT-q-PCR de l'expression ARN des gènes candidats codants et non codants.

L'ensemble des données préexistantes est récupéré à partir des dossiers cliniques du patient au sein du CHUGA et des comptes rendus anatomopathologiques. La date de décès a été vérifiée sur la base de données des personnes décédées en France depuis 1970 répertoriée mensuellement par l'Institut National de la Statistique et des Etudes Economiques (INSEE) (72).

L'Immunoscore n'est pas réalisé car celui-ci nécessite des analyses immunohistochimiques, une numérisation des lames et l'application d'un logiciel propriétaire, non effectuées actuellement en pratique courante au CHUGA.

2.2. Méthodologie de l'analyse par RT-qPCR et préparation des amorces

La méthode d'analyse par *Polymerase Chain Reaction* Quantitative après transcription inverse (RT-qPCR) consiste à détecter et à quantifier de l'ARN. Dans un premier temps, une transcription inverse de l'ARN total ou de l'ARN messager (ARNm) en ADN complémentaire (ADNc) au moyen de l'enzyme transcriptase inverse est réalisée. Ensuite, une amplification et détection de cibles spécifiques de cet ADNc par une PCR quantitative (ou PCR en temps réel) qui permet à chaque cycle d'amplification par PCR de mesure en temps réel la quantité d'ADN produit.

Afin de réaliser cette méthode, des amorces (ou *primers*) spécifiques des séquences étudiées sont nécessaires. A l'aide de ces amorces, les polymérases reconnaissent les séquences des gènes étudiés afin de les amplifier et de les quantifier de manière spécifique. La constitution des amorces spécifiques est réalisée informatiquement par l'outil Primer-BLAST de NCBI (73) avec des séquences FAST de chaque gène récupéré sur l'outil Nucleotide NCBI (74). Les séquences sélectionnées avaient pour condition d'avoir une taille limitée entre 70 à 150 paires de base et une température d'amorçage optimale (T_m) de 60°C (+/- 1°C). Une paire d'amorce est sélectionnée et se compose d'une amorce *Forward* et d'une amorce *Reverse* afin d'augmenter la sensibilité de l'analyse.

Outre les gènes identifiés par la signature GEC, quatre gènes le plus souvent exprimés dans le CC et trois gènes souvent exprimés dans le tissu colique sain sont utilisés afin de contrôler la qualité de l'analyse. Ainsi, chaque prélèvement nécessitera l'analyse de onze gènes par RT-qPCR :

- quatre gènes de la signature GEC avec un impact pronostique péjoratif issus de l'analyse bio-informatique,
- quatre gènes d'expression quasi-spécifique dans le cancer du côlon identifiés dans la cohorte TCGA-COAD (ESM1, CEMIP, WNT2, COL11A1),
- trois gènes « de ménage » (actine, U6, AUP1) exprimé systématiquement dans le tissu colique normal et habituellement utilisé dans l'étude par RT-qPCR du côlon.

Les paires d'amorce des onze gènes sont présentées dans l'Annexe 3.

2.3. Hypothèse et méthodes d'analyse statistique

Dans le cadre des analyses de survie, une méthode de type Kaplan-Meier suivie d'une comparaison des courbes de survie par test du log-rank a été réalisée. L'évènement retenu dans

l'objectif principal est la survenue d'un décès en lien avec le cancer du côlon selon la durée de suivi. L'objectif principal concerne une étude de la survie à 5 ans, soit 60 mois.

2.4. Éthique et confidentialité

Cette recherche correspond à une Recherche N'Impliquant pas la Personne Humaine (RNIPH) et entre dans le cadre de la « Méthodologie de référence » (MR-004) ; en application des dispositions de l'article 54 alinéa 5 de la loi du 6 janvier 1978 modifiée relative à l'information, aux fichiers et aux libertés. Ce changement a été homologué par décision du 3 mai 2018. Le CHUGA a signé un engagement de conformité à cette « Méthodologie de référence ». Ainsi, avant de débiter le travail de recherche, un protocole de RNIPH a été déposé et a été enregistré par la direction de la recherche clinique et de l'innovation (DRCI) dans le registre interne des recherches sur données au CHUGA afin de superviser le caractère éthique de la démarche scientifique de l'étude.

Les personnes ayant un accès direct aux données ont pris toutes les précautions nécessaires en vue d'assurer la confidentialité des informations relatives aux recherches, aux personnes qui s'y prêtent et notamment en ce qui concerne leur identité. Ces personnes, au même titre que les médecins eux-mêmes, sont soumises au secret professionnel. Les données enregistrées à l'occasion de cette recherche ont fait l'objet d'un traitement informatisé au nom de la structure responsable du traitement des données dans le respect de la loi 2018-493 du 20 juin 2018 relative à la protection des données personnelles. Le traitement des données est mis en œuvre dans le cadre de l'article 9 du règlement UE 2016/679. Les documents seront archivés par le nom de l'étude dans les locaux du service d'Anatomie et de Cytologie Pathologiques du CHUGA pour une durée conforme à la réglementation en vigueur.

Pendant la recherche ou à son issue, les données recueillies sur les personnes qui s'y prêtent et transmises au responsable du traitement de données par les intervenants dans la recherche ont été rendues anonymes. Les sujets sont identifiés par un code de 10 caractères : 2 lettres pour les initiales du projet (CC), 4 chiffres pour son numéro d'inclusion unique, 2 lettres pour l'initiale du nom et l'initiale du prénom, 2 chiffres pour le centre d'étude (01). La création de la base de données a été effectuée suite à l'autorisation à démarrer la recherche par la DRCI à l'aide de l'outil informatique RedCap. Les variables extraites et enregistrées ont été clairement définies dans le protocole de recherche avant le début de l'extraction des données.

Aucun financement extérieur n'a été obtenu pour la réalisation du projet de thèse et aucun conflit d'intérêt n'est à déclarer de la part des intervenants au projet.

RÉSULTATS

1. Résultats de l'étude bio-informatique

1.1. Identification des gènes spécifiques de tissus et ceux activés de manière aberrante dans le cancer du côlon

En utilisant les données d'expression de gènes dans la cohorte GTEx, qui contiennent environ 3000 échantillons de 48 tissus normaux, nous avons identifié 3554 gènes spécifiques de tissus. La Figure 12 présente le nombre de gènes spécifiques de tissus en fonction de leur tissu de prédominance. La proportion la plus importante de ces gènes est spécifique à la lignée germinale mâle (1511 gènes sur 3554, 42%). La liste des gènes inclut également quelques centaines de gènes prédominants dans d'autres tissus, par exemple, les gènes spécifiques du cerveau (432 gènes, 12%), des cellules embryonnaires souches (350 gènes, 10%), du foie (288 gènes, 8%) et du placenta (210 gènes, 6%). Certains gènes peuvent être prédominants dans deux tissus à la fois. En particulier, 12 gènes identifiés sont prédominants à la fois dans la lignée germinale mâle et dans le cerveau.

Nous avons ensuite calculé les fréquences d'activation de 3554 gènes spécifiques de tissus dans le cancer du côlon, en utilisant les données de la cohorte TCGA-COAD qui contient 398 échantillons du colon tumoral et 41 échantillons du côlon non-tumoral. La fréquence d'activation correspond au pourcentage d'échantillons tumoraux dans lesquels le gène est exprimé au-dessus du seuil empirique « m2sd » égal à la moyenne d'expression dans les échantillons non-tumoraux plus deux écarts-types. A titre d'exemple, la Figure 13 montre le calcul de la fréquence d'activation pour le gène ULBP2 fréquemment activé dans le cancer du côlon. Les résultats du calcul pour tous les 3554 gènes sont présentés dans la Figure 14.

Nous avons trouvé que 748 gènes spécifiques de tissus sur 3554 (21%) ont été fréquemment activés dans plus de 10% des CC. Nous avons retenu ces gènes en tant que candidats biomarqueurs potentiels pour réaliser une analyse de survie à l'étape suivante avec la méthode « ectopy ». La Figure 15 montre les vingt gènes les plus fréquemment activés dans le cancer du côlon. En particulier, le gène ESM1, exprimé normalement dans les reins et le poumon, et non exprimé dans le colon sain, est activé de façon aberrante dans 100% des CC.

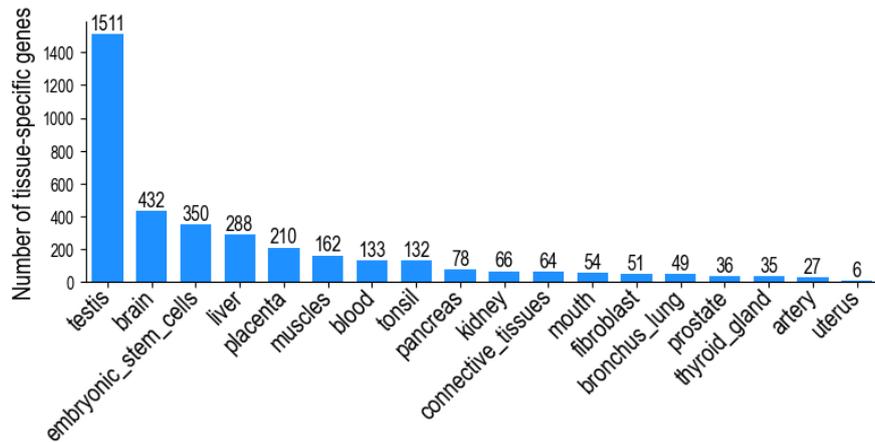


Figure 12. Nombre de gènes spécifiques de tissus présentés par ordre d'abondance

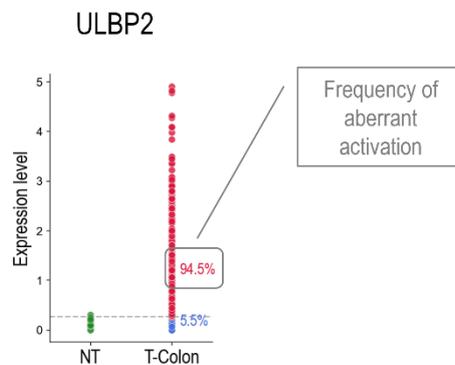


Figure 13. Exemple de la fréquence d'activation du gène ULBP2.

Cette fréquence d'activation est calculée dans les échantillons du cancer du côlon (T-Colon) du jeu de données TCGA-COAD, en utilisant le seuil « m2sd » calculé à partir des échantillons du colon non-tumoral (NT). Chaque cercle correspond à un échantillon dans le dataset TCGA-COAD. Les échantillons non-tumoraux sont marqués en vert, les échantillons tumoraux où le gène ULBP2 n'est pas activé sont présentés en bleu, les échantillons tumoraux où le gène ULBP2 est activé sont en rouge. La position du seuil « m2sd » est représenté par le trait gris pointillé.

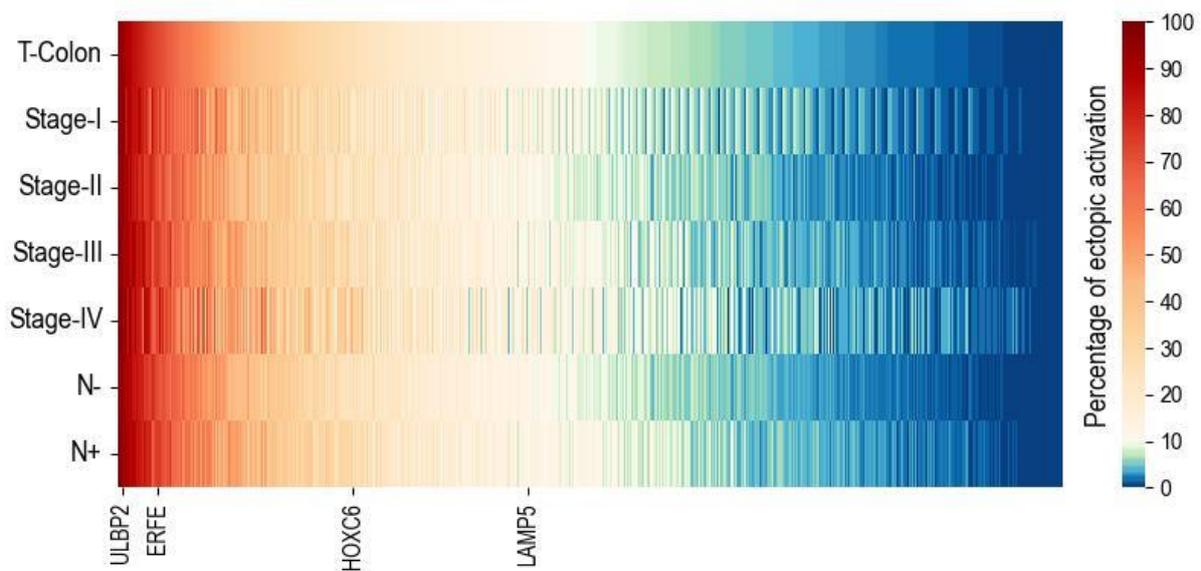


Figure 14. Fréquences d'activations ectopiques de 3554 gènes spécifiques de tissus dans le CC.

Ce graphique illustre également les stades UICC et le statut d'atteinte ganglionnaire N pour le dataset TCGA-COAD. Les gènes sont ordonnés le long de l'axe des abscisses selon le pourcentage d'activation dans la cohorte totale (T-Colon). Les activations fréquentes supérieures à 10% sont présentées en couleur rouge. Les activations peu fréquentes inférieures à 10% sont montrées en bleu.

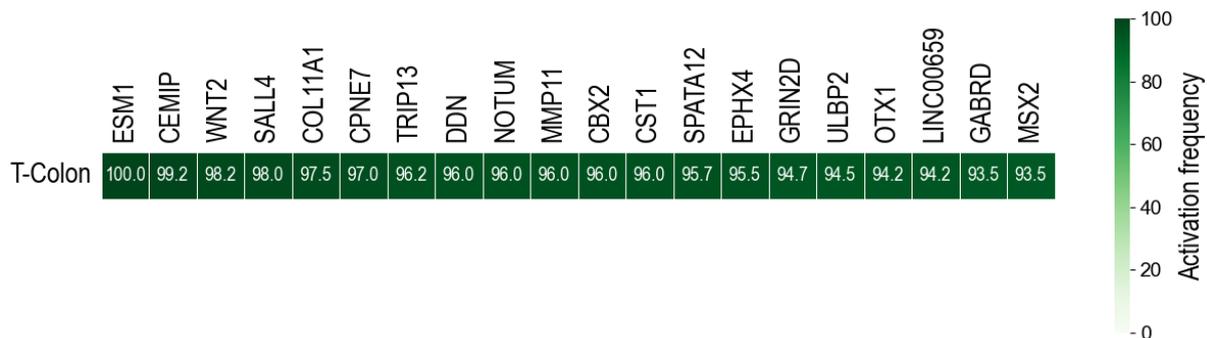


Figure 15. Fréquences d'activation ectopique pour les 20 gènes spécifiques de tissus les plus fréquemment activés dans le CC

1.2. Création de la signature pronostique GEC par la méthode « ectopy »

Pour identifier les gènes biomarqueurs dont l'activation aberrant a un impact significatif sur la survie, nous avons utilisé l'approche « ectopy ». Dans cette approche, toutes les données disponibles doivent être séparées en trois types de datasets : le dataset d'entraînement, le dataset de validation et le dataset de test. Nous avons utilisé la cohorte TCGA-COAD des données RNA-seq en tant que dataset d'entraînement. Les deux cohortes de puces à ADN, GSE39582 et GSE17536, ont été utilisées en tant que datasets de validation et de test respectivement.

Lors de la phase d'apprentissage dans la cohorte TCGA-COAD, parmi les 748 gènes spécifiques de tissus fréquemment activés dans le cancer du côlon, nous avons identifié 53 gènes pour lesquels un seuil d'activation OFF/ON stable a été détecté et dont le statut activé (ON) du gène était significativement associé à une probabilité de survie plus courte ($HR > 1$, $p < 0,05$).

Dans la phase de validation, nous avons comparé l'impact sur la survie de 53 gènes candidats entre la cohorte de validation GSE39582 et la cohorte d'apprentissage TCGA-COAD par le test statistique du logrank en fonction de l'état d'activation de chaque gène (OFF/ON). Cette étape permet d'évaluer la stabilité des biomarqueurs candidats par rapport à une éventuelle hétérogénéité entre les différentes cohortes du cancer du côlon. Pour quatre gènes parmi 53, nous avons obtenu les p-valeurs du test du logrank significatives dans les deux jeux de données à la fois : ERFE (érythroferrone), HOXC6 (*homeobox C6*), LAMP5 (*lysosomal associated membrane protein family member 5*) et ULBP2 (*UL16 binding protein 2*). Les résultats de l'analyse de survie pour ces gènes sont présentés dans la Figure 16. Les profils d'expression des quatre gènes dans les tissus normaux humains sont montrés dans l'Annexe 4. Les autres gènes de la liste de 53 candidats n'ont pas atteint le seuil de significativité dans le dataset de validation. Pour cette raison, ils ont été écartés de la sélection.

Les quatre gènes biomarqueurs retenus, ERFE, HOXC6, LAMP5 et ULBP2, ont finalement été combinés ensemble pour construire l'outil pronostique Gene Expression Classifier (GEC). Le GEC permet de stratifier les patients selon le nombre de gènes activés dans la tumeur parmi les 4 gènes du panel, de 0 à 4. Plus grand est le nombre de gènes activés, moins favorable est le pronostic vital des patients.

Dans la phase de test, l'outil GEC a été utilisé dans une cohorte indépendante GSE17536 qui n'a jamais participé à l'identification des biomarqueurs du panel. Ce test a confirmé avec succès la capacité du score GEC à prédire le pronostic des patients. Les résultats d'application du GEC dans les trois cohortes du cancer du côlon sont présentés dans la Figure 17. Le hazard ratio varie de 1,6 à 3,4 selon les cohortes, et reste statistiquement associé à un pronostic péjoratif du CC en cas d'expression d'un ou plusieurs gènes de la signature GEC.

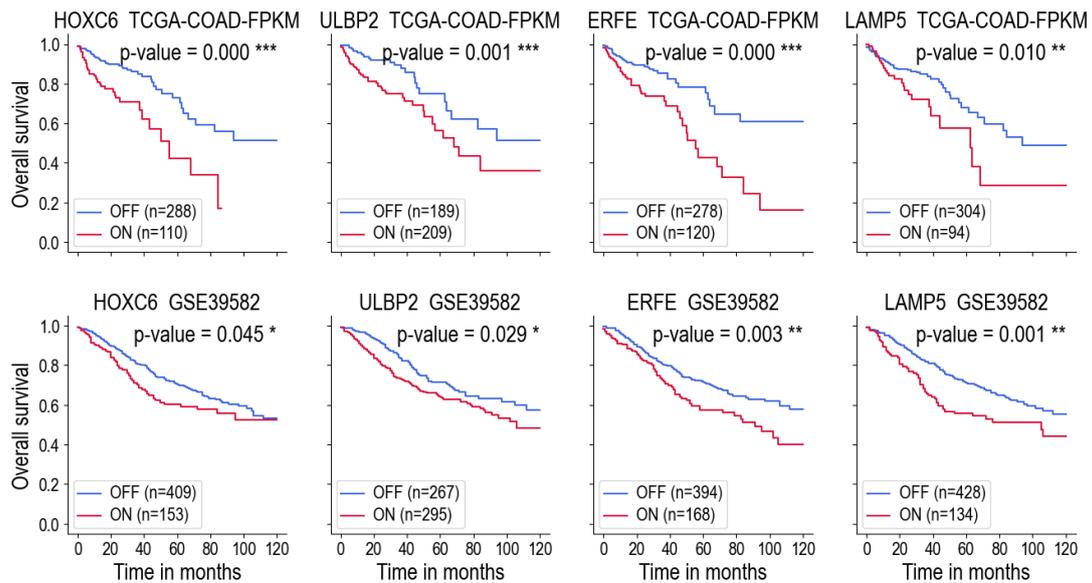


Figure 16. Courbes de survie de Kaplan-Meier des quatre gènes au sein des datasets TCGA-COAD et GSE39582 en fonction du statut d'activation du gène (OFF/ON)

Les p-valeurs correspondent au test du logrank entre les populations OFF et ON.

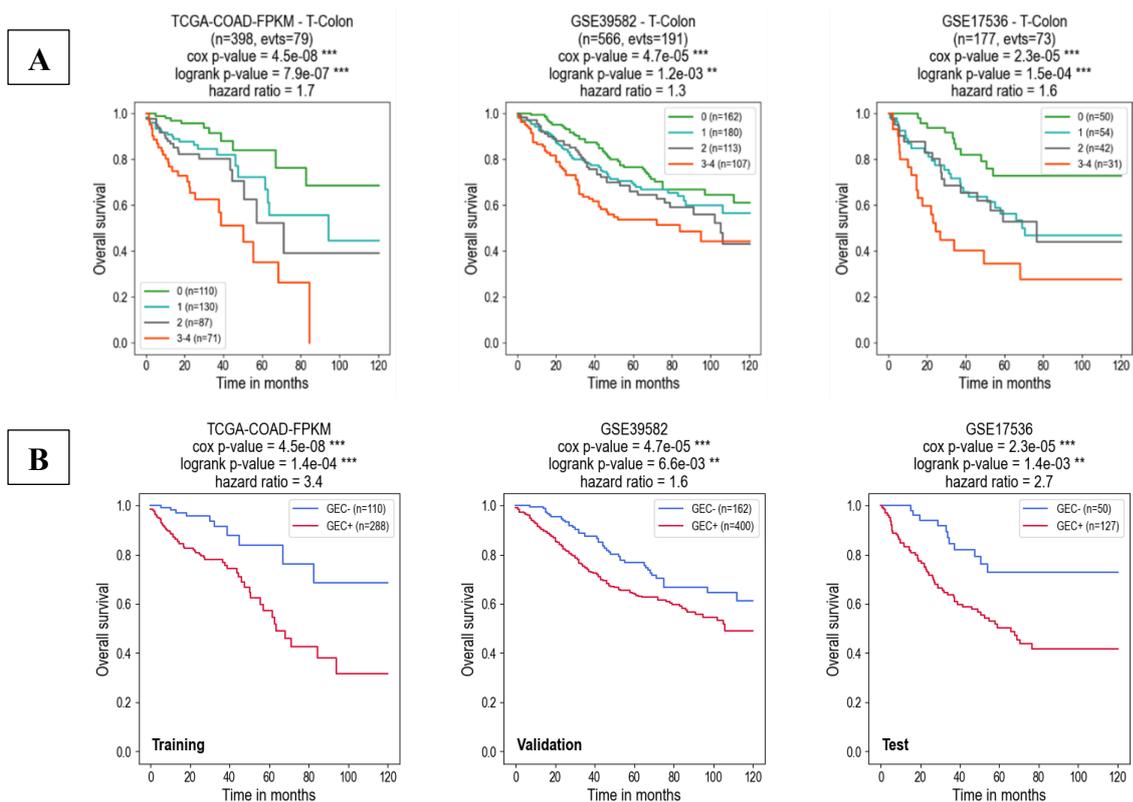


Figure 17. Courbes de survie de Kaplan-Meier au sein des cohortes publiques du CC en fonction de l'expression de la signature GEC.

A) Représentation selon le nombre de gènes activés de la signature GEC (0, 1, 2, 3-4 gènes). B) Représentation selon l'activation d'au moins un gène de la signature GEC. Dans le groupe GEC- aucun gène n'est activé dans le panel GEC. Dans le groupe GEC+ au moins un gène est activé dans le panel GEC.

1.3. Évaluation de l'impact pronostique de la signature GEC par rapport à d'autres facteurs de risque connus

L'analyse de survie de la signature GEC dans des sous-groupes du cancer du côlon des cohortes rassemblées a porté sur le stade UICC, le degré de différenciation, le statut ganglionnaire (N du TNM) et la présence de mutations de KRAS. Les résultats de cette analyse sont présentés dans la Figure 18. En effet, le score GEC permet de prédire la probabilité de survie des patients dans chaque stade du cancer, à l'exception du stade I. Le score GEC permet également d'estimer le pronostic selon le statut ganglionnaire N0 (HR = 2,1, p = 0,0015) et N1 (HR 1,9, p = 0,044), et non dans le sous-groupe N2. Pour le sous-groupe du degré de différenciation, le score GEC semble être plus intéressant dans le sous-groupe moyennement différencié (HR 2,3, p = 0,023). Enfin, le GEC présente un intérêt si KRAS n'est pas muté (HR = 1,6, p = 0,026).

Une analyse multivariée a été réalisée en prenant en compte le statut ganglionnaire, le stade, le grade et la présence de mutations de KRAS au sein des 3 cohortes, dont les résultats sont représentés dans la Figure 19. Ainsi, le GEC est significativement associé à un mauvais pronostic, indépendamment du statut ganglionnaire N, du grade et de la présence de mutations de KRAS.

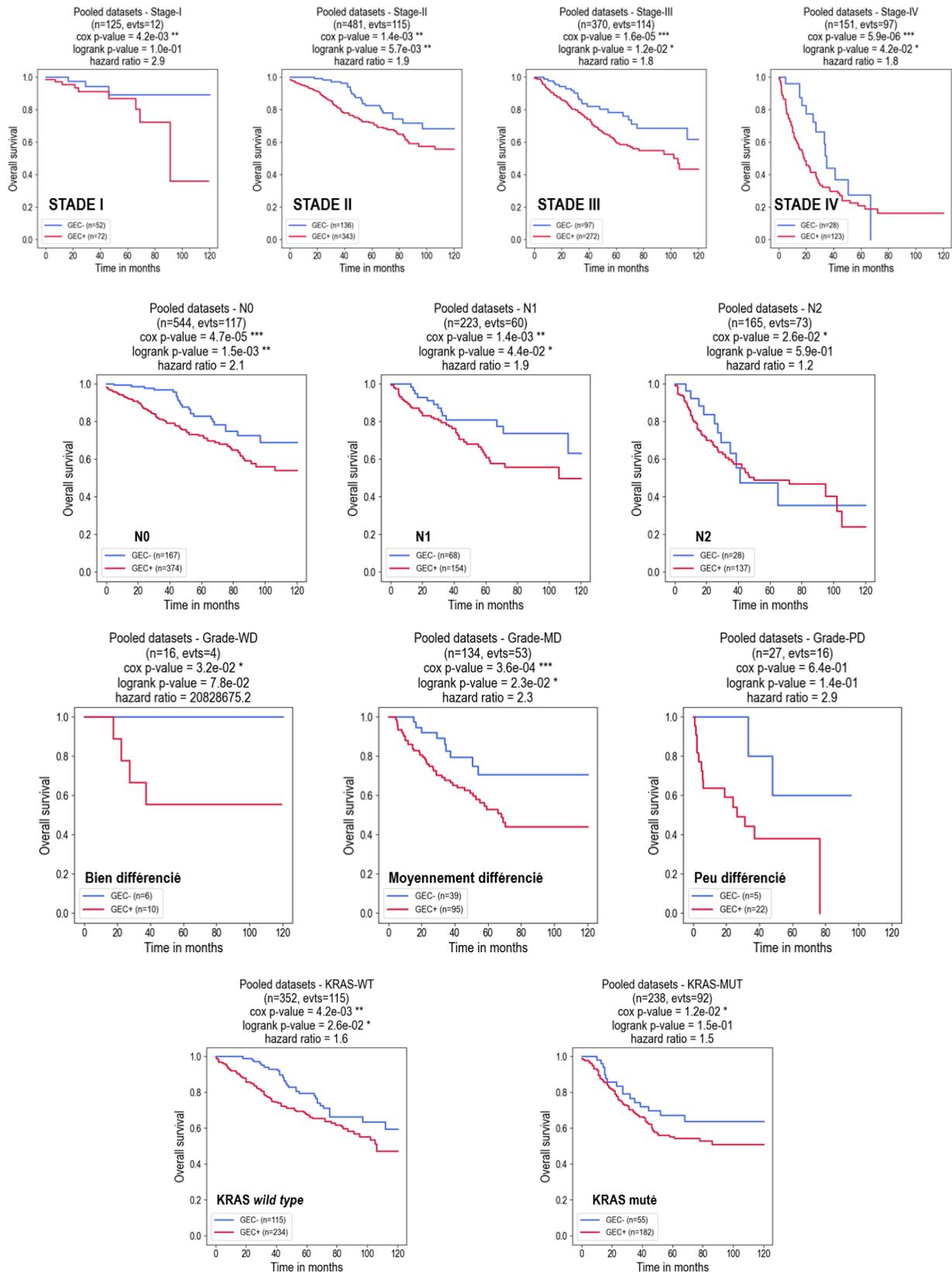


Figure 18. Courbes de survie de Kaplan-Meier des groupes GEC+ et GEC- issues de l'analyse en sous-groupe des trois cohortes publiques rassemblées

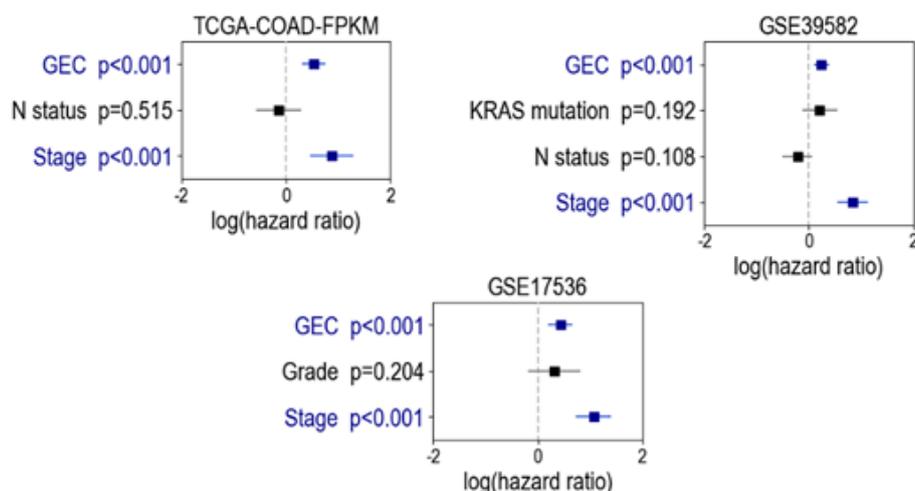


Figure 19. Analyse multivariée des trois cohortes en tenant compte du statut N, des mutations KRAS, du grade et du stade UICC

Les résultats et les incertitudes sont respectivement représentés au sein d'un Forrest-plot par un carré et des barres.

2. Identification des signatures moléculaires des tumeurs agressives GEC+ par l'analyse d'enrichissement GSEA

Le *Gene Set Enrichment Analysis* (GSEA) est une méthode d'analyse de données d'expression de gènes largement utilisée dans la littérature en biologie. Elle vise à comparer un vecteur de données numériques, indexé par l'ensemble des gènes du génome, à un certain ensemble de gènes (*gene set*), connu pour être associé à une signature moléculaire d'un phénomène biologique spécifique. Nous avons utilisé l'analyse GSEA pour identifier les signatures moléculaires qui caractérisent les tumeurs agressives GEC+ dans lesquelles plusieurs gènes du panel GEC sont activés par rapport aux tumeurs GEC- dans lesquelles aucun gène du panel n'est activé. La base de données de différents ensembles de gènes *Molecular Signatures Database* (MsigDB) nécessaires à l'analyse GSEA est publiquement disponible à la communauté scientifique sur le site web <https://www.gsea-msigdb.org/gsea>.

Dans un premier temps, nous avons procédé à l'analyse différentielle du transcriptome total dans chaque dataset disponible entre deux groupes : le groupe sans activation ectopique GEC- et le groupe avec au moins une activation ectopique GEC+. La Figure 20 montre les gènes sur-exprimés et sous-exprimés dans le groupe GEC+ par rapport au groupe GEC- avec la valeur absolue du *fold change* > 1.5 et la p-valeur du test statistique de Mann-Whitney ajustée FDR < 0.05. Respectivement, 150 gènes sous-exprimés et 281 gènes sur-exprimés ont été trouvés dans le dataset TCGA-COAD.

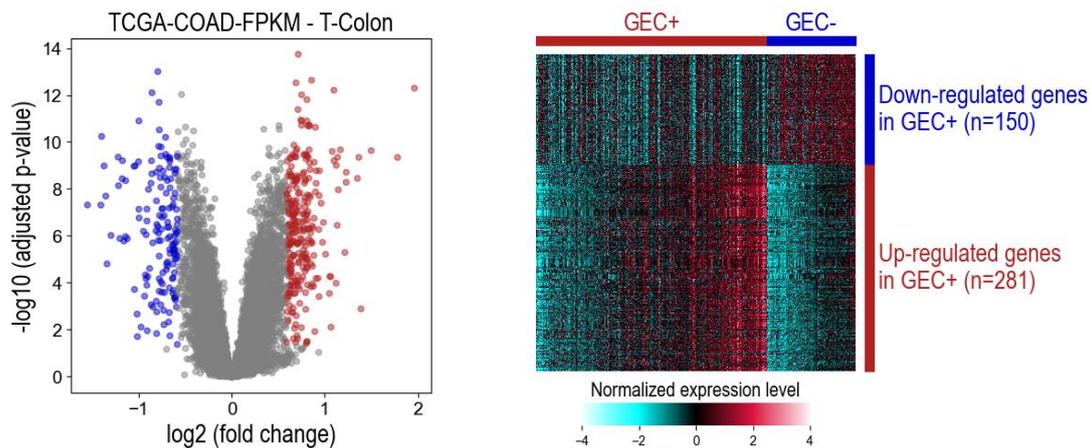


Figure 20. Résultats de l'analyse différentielle du transcriptome total dans le *dataset* TCGA-COAD entre les groupes GEC+ et GEC-

Les gènes significativement sous-exprimés sont présentés en bleu et les gènes significativement surexprimés en rouge. Le graphique de gauche correspond à un *volcano plot* dont chaque point est un gène. Le graphique de droite correspond à un *heat map* dont chaque colonne représente un prélèvement différent et chaque ligne un gène.

Nous avons ordonné les gènes du transcriptome total selon le score \log_2 (*fold change*) obtenu lors de l'analyse différentielle. La recherche des signatures moléculaires a ensuite été effectuée dans trois collections de référence de gene sets dans la base MSigDB : collection H (hallmark gene sets), collection C2 (curated gene sets), collection C5 (ontology gene sets) et collection MitoCarta (voies de signalisation des mitochondries). Un gene set était considéré significativement associé au groupe GEC+ si la p-valeur calculée était inférieure à 0.05 et si la p-valeur ajustée FDR correspondante était inférieure à 0.25.

Les résultats de l'analyse GSEA sont présentés dans les figures 21 et 22. Les profils GSEA des formes agressives du cancer du côlon GEC+ ont démontré un enrichissement significatif dans les gene sets associés à l'instabilité des microsatellites, à la transition épithélio-mésenchymateuse, à la motilité cellulaire, à la rechute du cancer et à la résistance aux traitements. De plus, nous avons trouvé que plusieurs signatures moléculaires liées à la réponse immune et à l'inflammation ont été également significativement enrichis dans les tumeurs GEC+ dans tous les datasets. Enfin, nous avons découvert que les tumeurs agressives GEC+ sont associées à une hyperméthylation aberrante de l'ADN qui cible particulièrement l'inactivation des gènes suppresseurs de tumeurs. Ces signatures sont accompagnées d'une perturbation importante des voies métaboliques, qui concerne notamment le métabolisme des lipides, en favorisant l'accumulation du cholestérol dans les cellules.

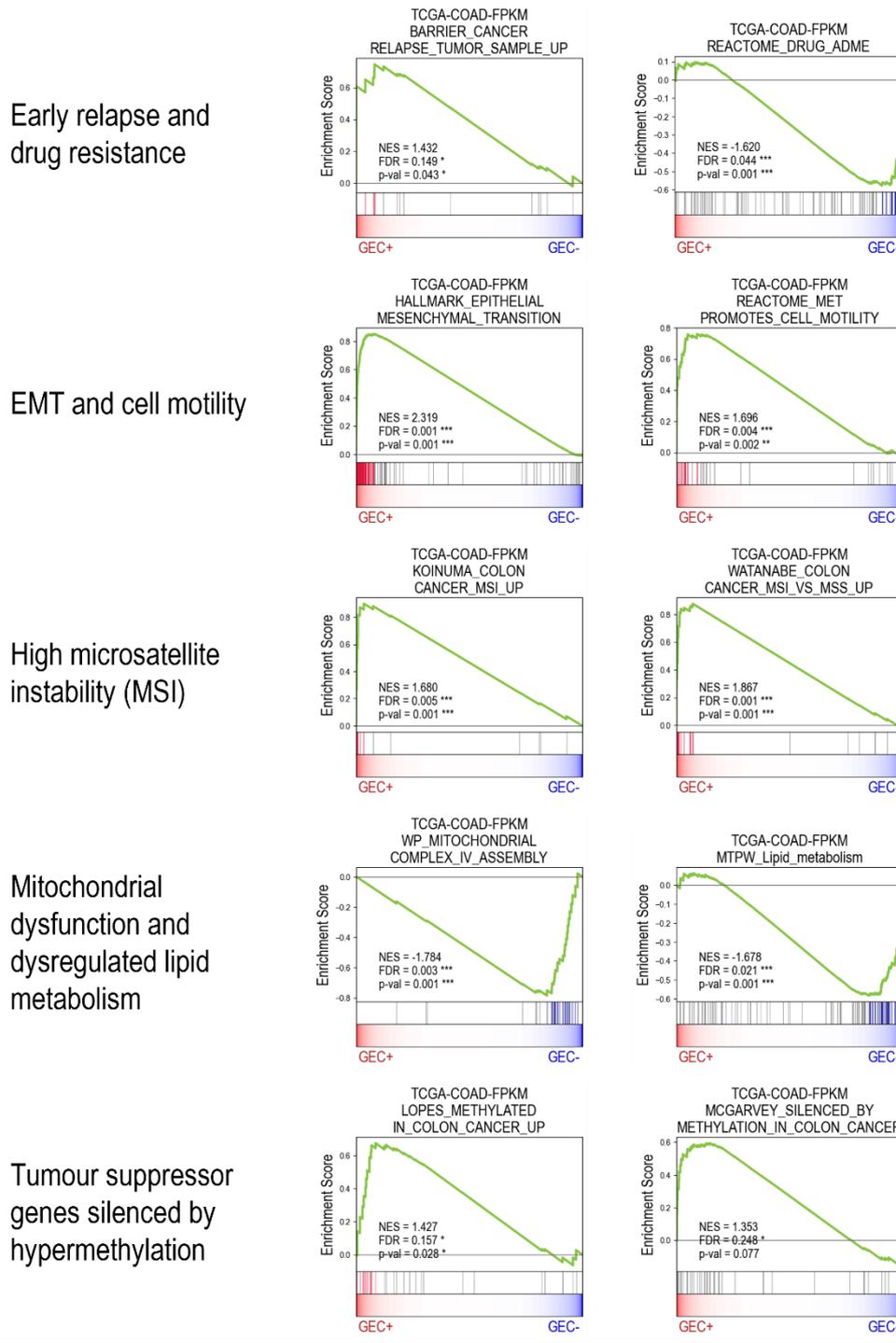


Figure 21. Résultats de l'analyse GSEA pour les profils transcriptomiques GEC+ versus GEC- dans une sélection de *gene sets* significativement enrichis (score d'enrichissement normalisé NES > 0) ou déplétés (NES < 0)

Les résultats se présentent sous la forme d'une courbe dont l'abscisse correspond à chaque gène du panel étudiés et l'ordonnée le score d'enrichissement (différentiel d'expression d'un gène entre GEC+ et GEC-). La ligne sous la courbe représente des gènes en barre verticales dont leur expression est plus ou moins différente selon le GEC ; les plus significatifs sont illustrés en gras et en couleur. Ainsi, un gène surexprimé dans la population GEC+ sera plus à gauche du graphique avec un score d'enrichissement supérieur à 0 ; ce qui sera l'inverse pour un gène déplété.

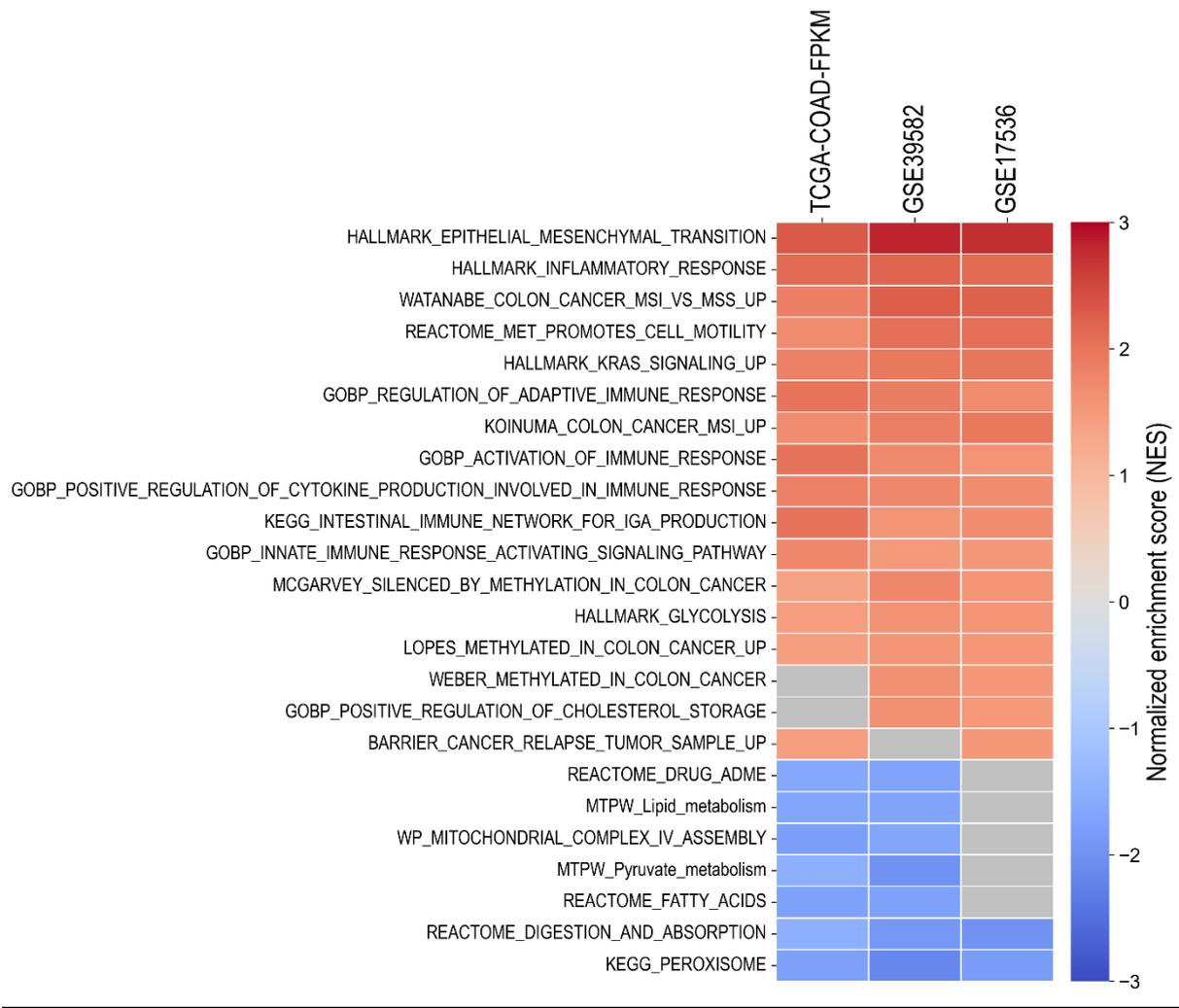


Figure 22. Le score normalisé d'enrichissement NES obtenu par l'analyse GSEA entre les groupes GEC+ versus GEC- pour trois cohortes du CC et pour différents *gene sets*.

Les *gene sets* significativement enrichis sont présentés en rouge. Les *gene sets* significativement déplétés sont montrés en bleu. Les cases grises correspondent au score NES non significatif (la p-valeur ajustée FDR > 0.25).

3. Cohorte du CHU Grenoble Alpes

3.1. Caractéristiques démographiques de la cohorte

De 2012 à 2014 inclus, 133 patients ont été opérés d'un adénocarcinome du côlon au CHUGA dont une analyse anatomopathologique a été réalisée. Les données démographiques et pathologiques de cette nouvelle cohorte sont résumées dans le Tableau 5. Cette cohorte présente une population majoritairement âgée de 61 à 81 ans et d'homme (61,7%). La notion de tabagisme est très peu retrouvée dans les dossiers des patients, mais concernent 22,6% de la cohorte totale, soit 57,7% des patients avec la notion retrouvée dans le dossier. Des antécédents familiaux de cancer (de tout type) ou de polypes colorectaux sont retrouvés chez 9% des patients. Des antécédents personnels de polype (24,8%), de polypose familiale (0,8%) et de cancer du côlon considéré en rémission (2,3%) sont également retrouvés.

Sur le plan pathologique, les patients ont majoritairement été traités par des colectomies partielles, dont 41,4% par colectomie droite et 40,6% par colectomie gauche ou sigmoïdectomie. Les localisations des CC étaient principalement au niveau du côlon droit (45,1%) et de la charnière recto-sigmoïdienne (34,6%). Une discrète surreprésentation des CC de stade II (37,6%) et de statut N0 (58,6%) est à noter. Les sous-types histologiques spécifiques sont limités qu'à 5,3% des patients, avec le carcinome médullaire, l'adénocarcinome mucineux et le carcinome à cellules en bague à chaton. Parmi les autres facteurs histo-pronostiques décrits dans les comptes rendus, il est identifié 18% de haut grade, 38,3% d'embole vasculaire, 24,8% d'engainement péri nerveux et 8,3% de nodule satellite. Un traitement néo adjuvant a été introduit chez 7,5% des patients, avec aucun score TRG proposé dans le compte rendu anatomopathologique.

Au niveau de la biologie moléculaire, peu de données ont pu être retrouvées au sein des comptes rendus cliniques et pathologiques, très certainement lié à un défaut de conservation des données lié au changement des logiciels. Seuls 8 patients avaient une étude immunohistochimique des MMR avec 2 dMMR. Un statut MSI est retrouvé chez 2 patients, des mutations de KRAS chez 10 patients, et des mutations de BRAF chez 1 patient.

	CHUGA 2012-2014	Données manquantes
Effectif	133	
Age au diagnostic, médiane [IQR]	71 [61-81]	0
Sexe, n (%)		0
Femme	51 (38,3%)	
Homme	82 (61,7%)	
Tabagisme (actif ou sevré), n (%)	30 (22,6%)	81
ATCD familial (cancer ou polype), n (%)	12 (9,0%)	0
ATCD personnel, n (%)		0
Polype	33 (24,8%)	
Polypose	1 (0,8%)	
Cancer du côlon	3 (2,3%)	
Année de chirurgie, n (%)		0
2012	37 (27,8%)	
2013	44 (33,1%)	
2014	52 (39,1%)	
Type de chirurgie, n (%)		0
Mucosectomie (EMR)	4 (3,0%)	
Colectomie droite	55 (41,4%)	
Colectomie transverse	3 (2,3%)	
Colectomie gauche ou sigmoïdectomie	54 (40,6%)	
Colectomie totale	17 (12,8%)	
Traitement néoadjuvant, n (%)	10 (7,5%)	0
Taille de la tumeur (en cm), médiane [IQR]	4,5 [3,1-7,0]	3
Localisation de la tumeur, n (%)		0
Caecum/droit	60 (45,1%)	
Transverse	8 (6,0%)	
Gauche	19 (14,3%)	
Charnière recto-sigmoïdienne	46 (34,6%)	

Stade UICC, n (%)		0
0	4 (3,0%)	
I	18 (13,5%)	
II	50 (37,6%)	
III	30 (22,6%)	
IV	31 (23,3%)	
Statut N, n (%)		0
N0	78 (58,6%)	
N1/N2	55 (41,4%)	
Exérèse complète (R0), n (%)	128 (96,2%)	0
Sous-type histologique, n (%)		0
Non spécifique	126 (94,7%)	
Médullaire	1 (0,8%)	
Mucineux	3 (2,3%)	
Bague à chatons	3 (2,3%)	
Grade, n (%)		4
Bas grade	105 (78,9%)	
Haut grade	24 (18,0%)	
Embole vasculaire, n (%)	51 (38,3%)	0
Engainement périnerveux, n (%)	33 (24,8%)	0
Nodule satellite, n (%)	11 (8,3%)	1
Biologie moléculaire, n (%)		
dMMR-IHC	2 (1,5%)	125
MSI	2 (1,5%)	127
Mutation KRAS	10 (7,6%)	109
Mutation BRAF	1 (0,8%)	122

Tableau 5. Caractéristiques démographiques de la cohorte de 2012 à 2014 du CHU Grenoble-Alpes

3.2. Mise au point des amorces pour la RT-qPCR

L'étude de la qualité des amorces des gènes de la signature (ERFE, HOXC6, LAMP5, ULBP2) et de ceux à visée de contrôle de la technique (ESM1, CEMIP, WNT2, COL11A1, actine, U6, AUP1) a été réalisé sur des échantillons de tissus humains normaux au des laboratoires de l'IAB. Chaque paire d'amorce (*Forward* et *Reverse*) a été analysée par qPCR dans des échantillons de tissus coliques sains et de tissus où ils sont normalement exprimés. Deux paires d'amorces ont été étudiées pour chaque gène afin de sélectionner le plus efficace. Deux puits contenant l'ADN-polymérase et un puit n'en contenant pas ont été réalisés pour chaque gène et chaque tissu afin de s'assurer de la fiabilité des résultats.

Ainsi, les amorces des gènes de la signature GEC étaient amplifiées dans les tissus où ils sont normalement exprimés et ne l'étaient pas dans le tissu colique sain. Les amorces des gènes fréquemment exprimés dans le CC sont également fonctionnelles. Cependant, les amorces des gènes fréquemment exprimés dans le tissu colique sain n'ont pas été amplifiées dans les échantillons coliques et nécessitent d'être explorés davantage.

4. Valorisation des résultats

Les résultats présentés dans ce travail de thèse ont été présentés sous la forme de posters à la journée de la recherche médicale à Grenoble (6 juin 2024), à la 36^{ème} édition du congrès européen de pathologie de la société européenne de pathologie (*European Society of Pathology*, ESP) à Florence, Italie (7-11 septembre 2024), ainsi qu'à la 14^{ème} édition des congrès internationaux des pathologies malignes du péritoine (PSOGI-ISSPP) à Lyon (26-28 septembre 2024). Enfin, le travail fait l'objet d'une communication orale au congrès annuel de la Société Française de Pathologie (20-22 novembre 2024). Un article basé sur les données bio-informatiques et complété par des analyses moléculaires est en préparation.

Des demandes de financement auprès d'appels d'offre interne du CHUGA ou de sociétés savantes ont été réalisées. D'autres demandes de financement auprès de futurs appels d'offre du CHUGA ou de la société française de pathologie seront réalisées.

DISCUSSION

1. Une nouvelle signature dans le CC basée sur l'expression ectopique de 4 gènes spécifiques de tissu

Ce travail de thèse a mis en évidence des gènes spécifiques de tissu dans le CC. Après une analyse originale par une approche par *machine learning*, le pipeline « ectopy » a permis de concevoir une signature GEC basée sur 4 gènes dont leur expression ectopique dans CC aurait un impact péjoratif significatif sur la survie des patients : ERFE, HOXC6, LAMP5 et ULBP2. Lorsque ces gènes sont exprimés, le pronostic est dégradé. Le pronostic est aggravé proportionnellement au nombre de gènes exprimés. La signature GEC appliquée à des sous-populations du CC montre un intérêt particulier de cette signature dans les stades avancés (II, III et IV) et dans les cancers dépourvus de mutation de KRAS. Les résultats sont d'autant plus intéressants qu'ils révèlent que la signature permettrait de repérer les patients au pronostic péjoratif à long terme (> 5 ans de survie) dans les stades I ainsi que d'identifier ceux à court terme (< 5 ans de survie) dans les stades IV. Une analyse multivariée a permis de montrer une indépendance de la signature GEC vis-à-vis de mutations KRAS, du grade et du statut ganglionnaire N du TNM.

L'exploration *in silico* de l'enrichissement des gènes GEC par GSEA a mis en évidence des résultats intéressants notamment une corrélation entre l'expression ectopique d'au moins un des gènes de la signature et la survenue d'évènements suivants : la récurrence précoce, la résistance à des médicaments, la transition épithélio-mésenchymateuses, la mobilité cellulaire, l'instabilité des microsatellites, la dysfonction mitochondriale et le métabolisme des lipides dérégulés, l'hyperméthylation de gènes suppresseurs de tumeurs. L'ensemble de ces évènements enrichis peut donner des pistes de mécanismes moléculaires impliqués dans l'expression de la signature GEC et nécessitant d'être exploré.

Cette thèse a également permis de mettre en place une cohorte de 133 patients ayant un CC opéré et dont le prélèvement a eu une étude anatomopathologique complète au sein du CHUGA durant les années de 2012 à 2014. Afin de convenir aux nombre de sujets nécessaires en vue de réaliser une validation indépendante de l'impact du score GEC sur la survie, la cohorte du CHUGA sera complétée à l'avenir par des patients de 2015 à 2017 inclus dans le cadre du projet SIROCCO.

Enfin, il a été identifié et contrôlé des amorces des gènes qui seront utilisables par RT-qPCR dans le cadre du projet SIROCCO. L'analyse par RT-qPCR des prélèvements de la cohorte pourra donc débuter dès que le projet SIROCCO sera financé.

2. Revue systématique de la littérature des gènes GEC identifiés

2.1. ERFE

L'érythroferrone (ERFE, anciennement nommé CTRP15, myonectine ou FAM132B) appartient à la famille des CTRP (*C1q tumor necrosis factor-related protein*). Le gène ERFE est localisé sur le chromosome 2 (2q37.3) avec une expression restreinte aux tissus thyroïdiens, testiculaires et musculaires. Elle n'est pas exprimée physiologiquement dans le côlon. Les CTRP constituent une famille impliquée dans la régulation de la prolifération cellulaire, de l'inflammation, de l'apoptose et du métabolisme glycolipidique (47,75).

ERFE est une hormone dont la principale fonction physiologique est de participer à l'homéostasie du fer en inhibant l'hepcidine, permettant une libération du fer afin de répondre rapidement à une anémie (hémorragie ou altitude). Sa synthèse, supposée par les érythroblastes de la moelle hématopoïétique, est sous le contrôle étroit de l'érythropoïétine (EPO). Des essais cliniques sont en cours pour évaluer l'intérêt de traitement par des agonistes des récepteurs ERFE afin d'agir sur l'anémie en cas d'insuffisance rénale chronique, ou par des anticorps anti-ERFE afin d'éviter des surcharges en fer. ERFE pourrait également être un marqueur d'utilisation de l'EPO en cas de dopage dans le sport de haut niveau (76).

D'autres fonctions ont également été décrites, telles que sa participation dans le métabolisme glycolipidique en agissant comme une adipokine, sous l'influence hypothétique de l'indice de masse corporelle et de l'activité physique, et un lien avec le diabète de type 2. ERFE aurait une action cardioprotectrice, comme démontré au cours d'une étude où des souris surexprimant ERFE limiteraient les dommages tissulaires du cœur contrairement aux souris *wild type*. Une action dans l'embryogénèse a également été évoquée (76).

Sur une étude *in silico* de 2023, son expression ectopique serait associée à un mauvais pronostic dans le CC, ainsi que dans le corticosurréalome, le mésothéliome, l'adénocarcinome pancréatique, le carcinome rénal à cellules claires et le mélanome. Néanmoins, elle serait associée à un meilleur pronostic dans le carcinome rénal papillaire. L'étude a également identifié une association entre la surexpression d'ERFE et des charges mutationnelles tumorales de gènes somatiques, notamment dans le CC ; situation évoquant le sous-type moléculaire d'instabilité chromosomique (CIN). Une surexpression d'ERFE serait associée à des mutations

de TP53 dans d'autres cancers. Cette surexpression semblerait être associée à des modifications du micro-environnement tumoral, avec un remodelage de la matrice extracellulaire, une participation à l'inflammation notamment dans le recrutement des cellules de l'immunité. Une implication d'ERFE dans la voie de signalisation mTOR/PI3K/AKT est évoquée, cette voie est connue pour participer à la carcinogenèse colorectale (15,77).

L'étude de Liang et al., en 2018, a permis d'explorer *in silico* le profil génomique des CC selon leur localisation dans le côlon et a ainsi mis en évidence 164 gènes avec une expression différente entre le côlon droit et le côlon gauche, ainsi qu'une association de l'expression d'ERFE dans le côlon droit avec un mauvais pronostic (78).

En évaluant l'interaction protéine-protéine sur des bases de données à l'aide d'IntAct (Annexe 5), ERFE présenterait une association modérément fiable (score MI : 0,53) avec le gène LRP1 qui serait impliqué dans l'endocytose, la signalisation cellulaire, le métabolisme lipoprotéique et la mobilité cellulaire. LRP1 est également supposé être un gène suppresseur de tumeur, dont son inhibition induirait à une augmentation de l'invasion cancéreuse (79). La précision de l'interaction entre ERFE et LRP1 n'est pas retrouvée.

Néanmoins, il est important de remarquer que la réaction inflammatoire en lien avec un cancer, souvent de stade avancé, peut être responsable d'une anémie, qui elle-même induirait une expression exagérée d'ERFE. Cette association physiopathologique connue pourrait interférer avec l'association statistique entre l'expression ectopique d'ERFE et le mauvais pronostic dans le CC (80). L'anémie étant par ailleurs un signe clinique devant faire évoquer un CC à cause des saignements digestifs occultes en lien avec le polype ou le cancer.

2.2. HOXC6

Le gène HOXC6 appartient à la famille des gènes homeobox (HOX) et code pour un facteur de transcription hautement conservé chez les mammifères qui participe à la morphogenèse et au développement embryonnaire. Il est par ailleurs décrit 4 clusters d'homeobox chez les vertébrés : HOXA, HOXB, HOXC, HOXD. HOXC6 est quant à lui localisé sur le chromosome 12 (12q13.13) avec une expression relativement restreinte aux tissus ovariens, surrénaliens, adipeux, rénaux et placentaires. Elle n'est pas exprimée physiologiquement dans le côlon (47,81). En cancérologie, la surexpression d'HOXC6 a été décrite dans la littérature comme étant de mauvais pronostic dans les cancers du sein, de la prostate, du poumon, ainsi que dans des leucémies, des ostéosarcomes et des médulloblastomes.

Dans le CC, cette surexpression a également été décrite, avec une association un pronostic péjoratif. Elle s'associerait à une baisse de la survie globale, de la survie sans progression, à des stades T et N avancés, à la présence de métastases synchrones et au recours à une chimiothérapie adjuvante (81). Cette surexpression est plutôt retrouvée dans le côlon droit avec un pronostic plus péjoratif dans les CC droit (78,82,83). Plusieurs études l'intègrent dans des propositions de signatures pronostiques du CC (78,82–85).

Sur le plan physiopathologique dans le CC, des recherches ont été menés afin de comprendre les mécanismes pouvant expliquer ce pronostic péjoratif en cas de surexpression. HOXC6 semblerait favoriser la prolifération des cellules du CC et la cancérogénèse à travers l'inhibition de l'autophagie et favoriserait la prolifération cellulaire à l'aide des voies moléculaires PI3K/AKT/mTOR (81). Une activation des voies Wnt et TGF- β /SMAD a également été décrite, ce qui pourrait favoriser la prolifération cellulaire et le potentiel métastatique du CC (85). La surexpression d'HOXC6 pourrait impacter le micro-environnement tumoral à travers des mécanismes d'échappement du système immunitaire (85), de transition épithélio-mésenchymateuse (83,86) et l'hypoxie tissulaire (84). La transition épithélio-mésenchymateuse passerait par l'activation de la voie Wnt/ β -caténine et l'inhibition de DKK1 (83,86). D'autres mécanismes ont été décrits dans d'autres localisations carcinomateuses avec surexpression d'HOXC6, tels que l'activation d'oncogènes comme JunD ou la régulation de Bcl-2 (81).

HOXC6 pourrait également être à l'origine de résistance à de certains traitements du CC, comme les inhibiteurs du check point immunitaire (83) ou l'irinotécan (86). Un mécanisme de résistance des chimiothérapies passant par la régulation de MDR-1 a été évoqué (81).

L'exploration à l'aide d'interactions protéine-protéine sur IntAct n'a pas montré de lien significatif (Annexe 5).

2.3. LAMP5

Le gène LAMP5 (*Lysosome Associated Membrane Protein-5*), aussi nommé BAD-LAMP (*Brain and Dendritic Cell associated LAMP-like*) ou C20orf103, code pour des protéines de la membrane des lysosomes, qui sont des organites intracellulaires permettant la dégradation de molécules à l'aide de nombreuses enzymes. LAMP5 est localisé sur le chromosome 20 (20p12.2). Son expression normale tissulaire est habituellement limitée au cerveau et aux ovaires, en prédominance dans les cellules dendritiques. Les protéines codées par les gènes de la famille LAMP interviennent dans le fonction des lysosomes et de la mort

cellulaire (apoptose, nécrose et autophagie), dont 3 membres de cette famille seraient spécifiques de tissus : LAMP3, LAMP4 (ou CD68) et LAMP5 (87).

Plusieurs études ont montré un lien entre la surexpression de LAMP5 dans le cancer de l'estomac avec un mauvais pronostic et un risque de dissémination métastatique plus important (88,89). Son implication dans des hémopathies malignes, telles que le myélome multiple et les leucémies de phénotype ambiguë, a été rapportée (90,91).

Dans le cancer du côlon, des études ont également identifié un pronostic dégradé lorsque LAMP5 est surexprimé, avec une association à des formes riches en lymphocytes intratumorales (92,93).

Des hypothèses sur le mécanisme physiopathologique de l'implication de LAMP5 en cancérologie ont été proposées. Une action empêchant l'apoptose et/ou l'autophagie des cellules cancéreuses au travers d'une inhibition des caspases, enzymes impliquées dans ces mécanismes, notamment au travers du contrôle de cette expression par une lncRNA LAMP5-AS1 et l'histone méthyltransférase DOTL1, ou bien avec l'activation concomitante de gènes participant à l'autophagie lorsque LAMP5 est surexprimé (88,90). Une régulation du cycle cellulaire au travers de l'activation de la MAP kinase p38 et un contrôle de la phase G2 pourrait faciliter la prolifération des cellules carcinomateuses (88,91). Une interaction avec la transition épithélio-mésenchymateuse et le contrôle de l'activation de ROS ont été suggérés (88).

L'exploration à l'aide d'interactions protéine-protéine sur IntAct n'a pas montré de lien significatif (Annexe 5).

2.4. ULBP2

Le gène ULBP2 (*UL16 binding protein 2*), appartient à la famille des ULBP qui codent pour des molécules en lien avec le complexe d'histocompatibilité majeur de classe I (CMH1) qui se lie aux récepteurs NKG2D (*Natural Killer Group 2 Member D*) des lymphocytes Natural Killer (LNK) et des lymphocytes T CD8⁺. Localisé sur le chromosome 6 (6p25.1), l'expression physiologique d'ULBP2 est limitée principalement à l'œsophage. L'expression normale du CMH1 permet aux lymphocytes de reconnaître les cellules du soi et d'éviter la cytolyse.

Cette famille de gène a été découverte en étudiant l'échappement immunitaire des cellules infectées par le cytomégalo virus humain (CMV) par l'intermédiaire de l'interaction avec la protéine UL16 (94). Ce mécanisme a été décrit récemment dans l'infection par le virus JC chez les patients immunodéprimés atteints de leucoencéphalopathies multifocales (95).

En cancérologie, son expression ectopique dans le CC, ainsi que dans les cancers du pancréas (96) ou du sein (97), serait associée à une baisse de la survie globale et de la survie sans récurrence. Le mécanisme mise en évidence serait l'échappement au système immunitaire, notamment des LNK et des lymphocytes T CD8+, en empêchant la cytolysé des cellules cancéreuses par l'intermédiaire de son expression et de celle du CMH1 (98–101). Outre son impact sur le pronostic, il est également décrit un potentiel rôle diagnostique avec une aire sous la courbe ROC oscillant entre 0,7 et 0,98 (98,101).

L'exploration à l'aide d'interactions protéine-protéine sur IntAct (Annexe 5) montre une association assez forte (Miscore à 0,61) avec le gène KLRK1 qui code pour le récepteur NKG2D. Elle montre également une association modeste (Miscore à 0,5) avec les gènes MED21 (enzyme régulant la transcription de l'ARN polymérase II), YIPF6 (participant au développement des cellules épithéliales intestinales), DERL3 (participant à l'homéostasie du réticulum endoplasmique et évoqué dans certains cancers) et TNFSF8 (ligand du récepteur CD30, décrit dans les réponses immunes antibactériennes et dans certains lymphomes).

3. Connaissances actuelles des expressions ectopiques de gènes spécifiques de tissus dans le cancer du côlon

À l'aide de cohortes publiques telles que TCGA-COAD disponibles sur NCBI Genomics Data Common (102), l'utilisation de données cliniques et moléculaires quasiment exhaustives a permis de mettre en évidence des gènes dont l'expression aurait un impact pronostic significatif (103). À titre d'exemple, The Human Protein Atlas a recensé ces gènes dans le cancer du côlon avec 243 gènes dont leur expression (ectopique ou non) est associée à un impact péjoratif sur le pronostic et 360 gènes associée à un impact mélioratif (47).

Après revue systématique des articles référencés sur NCBI PubMed, il a été identifié 36 gènes spécifiques de tissus dont leur expression ectopique pourrait être associée à un impact sur la survie globale ou sans récurrence, le potentiel métastatique ou une fonction cellulaire pouvant être associée à un impact sur la survie. Ces gènes sont listés en Annexe 6. La majorité des gènes identifiés est associée à un pronostic péjoratif.

L'activation du gène Wnt5A correspond au seul gène dont son expression est associée à un meilleur pronostic dans le CC, démontré par une étude de survie. La famille des gènes Wnt est connue pour être impliquée dans une voie de signalisation Wnt/B-caténine.

Dix de ces gènes sont des CTA, tels que CABYR, NY-ESO-1 (CTAG1A) ou MAGEA3. À noter que PLAC1 est le seul gène uniquement exprimé dans le placenta qui aurait un impact

pronostic péjoratif en cas d'expression ectopique de celui-ci. L'expression aberrante de ce dernier pourrait être induit par le virus Epstein-Barr (104), ainsi l'exposition environnementale doit être explorée afin de mieux comprendre l'origine des expressions ectopiques de gènes spécifiques de tissus.

4. Limites de l'étude

Bien que la signature GEC semble prometteuse en raison des résultats significatifs dans ce travail *in silico*, une étude expérimentale de celle-ci sur une cohorte nouvellement conçue reste nécessaire. Cette dernière permettra de confronter l'expression de ces gènes en fonction des facteurs histo-pronostiques connus et des biomarqueurs étudiés en pratique. Une analyse multivariée plus large permettra aussi d'attester de l'indépendance de la signature avec des données connues sur le pronostic du CC. Ou au contraire, en attestant une association à des critères pronostiques connus, cette signature pourrait permettre d'apporter de nouvelles connaissances sur le pronostic du CC. De plus, il est nécessaire d'évaluer l'application du score GEC par des méthodes d'analyse différentes et accessibles en pratique, comme, par exemple, l'immunohistochimie. Par ailleurs, une meilleure compréhension des gènes de la signature GEC dans la cancérogénèse colique et des voies impliquées dans le pronostic du CC est nécessaire, bien que des pistes soient proposées par l'étude GSEA, puisque de potentiels biais sont possibles tels que la relation entre ERFE et l'anémie ferriprive habituellement présent dans le CC. L'implication de certaines voies moléculaires déjà connues comme le CMS devrait être explorée de manière plus approfondie. La relation ou la comparaison avec d'autres signatures pronostiques des gènes GEC pourrait être nécessaire afin d'objectiver l'apport de l'étude SIROCCO, notamment dans les stades II et IV.

5. Perspectives du travail de thèse

5.1. Projet SIROCCO

Ce travail de thèse s'inscrit dans le projet SIROCCO qui consiste à élaborer puis étudier cette nouvelle signature pronostique à l'aide de données bio-informatiques plus complètes et d'une étude expérimentale au sein de la cohorte du CHUGA mais également dans le cadre d'une étude multicentrique nationale avec l'avis favorable et engageant du réseau RENAPE.

Les données bio-informatiques peuvent être complétées, en s'intéressant particulièrement aux implications des gènes de la signature avec des biomarqueurs utilisés en

routine (BRAF, RAS, MSI). Il serait également intéressant d'approfondir la connaissance des signatures moléculaires des gènes dans les sous-types des CC, notamment dans les cancers avec une instabilité des microsatellites MSI, ce qui permettrait de donner des pistes de compréhension de la relation pronostic et gènes de la signature GEC. Enfin, l'étude de mécanisme moléculaire impliquant l'expression ectopique d'un des gènes GEC permettrait une meilleure compréhension de la cancérogénèse colique et la piste d'une voie thérapeutique innovante. Pour illustrer ce dernier point, l'équipe de l'IAB avait identifié un impact péjoratif sur le pronostic en cas d'expression ectopique de la prolactine dans les cancers du poumon ainsi que son impact dans la méthylation et l'épigénétique des cancers lorsqu'elle est exprimée ou réprimée. La prolactine est une hormone ayant de nombreuses actions, tout comme ERFE (105). Afin d'explorer ces pistes, un travail de recherche orienté statistique et bio-informatique en cancérologie sera proposé à des étudiants de master en Statistiques et Science des Données (SSD).

Comme énoncé précédemment, la cohorte du CHUGA nouvellement créée est en cours de finalisation et sera parachevée en incluant les patients des années 2015 à 2017 inclus. Sur cette cohorte, la valeur pronostique de la signature GEC sera confirmée par une analyse par RT-qPCR à l'IAB afin de l'évaluer sur l'ensemble des stades et des critères histo-pronostiques révisés. Une analyse par immunohistochimie au laboratoire d'anatomopathologie du CHUGA sera également effectuée afin de promouvoir une méthode plus accessible en routine. La comparaison avec leur profil génotypique (RAS, BRAF, MSI) pourra également être envisagée. L'ensemble des études du projet SIROCCO constituera une avancée majeure dans la compréhension du pronostic en cancérologie digestive, tant sur l'évolution naturelle du cancer que sur la prise en charge thérapeutique.

5.2. Recherche d'une signature diagnostique à l'aide « d'ectopy »

L'exploration de l'expression aberrante de gènes spécifiques de tissus pourrait être envisagée à visée diagnostic dans le cadre d'un dépistage précoce du cancer colique. En effet, dans le protocole nous avons identifié des gènes fréquemment activés dans plus de 10% des CC, remplacer ce seuil par un seuil proche de 100% permettrait de rechercher des expressions ectopiques en commun avec l'ensemble des CC. Tel qu'illustré dans la Figure 15, les gènes ESM1 (rein et poumon, 100%), CEMIP (fibroblaste, 99,2%), WNT2 (placenta, 98,2%) pourraient notamment être de potentiels gènes d'intérêt diagnostique à étudier.

Le dépistage organisé en France est basé sur un test immunochimique de recherche de sang occulte dans les selles. Ce dernier est un test sensible mais peu spécifique et si celui-ci se révèle positif, un examen endoscopique est ainsi proposé. Si une signature diagnostique robuste ressort, l'utilisation hypothétique au sein du même prélèvement que le test du sang occulte des selles pourrait être envisagée et représenterait une approche nouvelle et originale dont l'intérêt serait d'optimiser le dépistage et de cibler de façon plus précise les patients nécessitant une coloscopie.

CONCLUSION

Ce travail de thèse a permis de mettre en évidence la présence d'expression ectopique de gènes spécifiques de tissu dans le cancer du côlon, ainsi que d'identifier quatre gènes dont leur activation aberrante dans ce cancer aurait un impact pronostic péjoratif significatif : ERFE, HOXC6, LAMP5 et ULBP2 ; ces derniers constituent une signature pronostique intéressante et robuste. Ce travail a également permis de constituer une cohorte rétrospective monocentrique issue du CHUGA issue des prélèvements chirurgicaux de 2012 à 2014 comprenant 133 patients, qui sera complétée par la suite par des données de 2015 à 2017. Cette dernière a pour intérêt de mettre à l'épreuve cette nouvelle signature pronostique à l'aide d'analyses par RT-qPCR puis par immunohistochimie, afin d'évaluer cette pratique en routine dans un service d'anatomie et cytologie pathologiques. La découverte de ces quatre gènes ouvre des perspectives de recherche pour permettre une meilleure compréhension du pronostic du cancer du côlon ainsi qu'un apport potentiel dans la prise en charge personnalisée des patients.

CONCLUSION SIGNÉE PAR LE DOYEN ET LA PRÉSIDENTE DU JURY

THÈSE SOUTENUE PAR : M. Aurélien SPINELLI

TITRE :

Signature pronostique dans le cancer du côlon par l'étude de l'expression ectopique de gènes spécifiques de tissus

CONCLUSION :

Introduction : Le cancer du côlon représente, en 2020, le 4ème et 2ème cancer en termes d'incidence et de mortalité respectivement. Malgré de nombreux critères histo-pronostiques, le pronostic et la prise en charge des cancers du côlon en 2024, particulièrement pour les stades II et IV, reste parfois incertains. L'Institut pour l'Avancée des Biosciences a démontré que l'expression ectopique de certains gènes tissus-spécifiques dans le cancer, alors qu'ils sont silencieux dans le tissu sain, est significativement associée à un pronostic défavorable. Ainsi, ces gènes représentent une nouvelle source de biomarqueurs pronostiques. L'objectif principal de cette étude est de créer une signature pronostique dans les cancers du côlon, basé sur les activations ectopiques, pour mieux guider la prise en charge thérapeutique des patients.

Méthode : Nous avons utilisé le pipeline de machine learning « ectopy » développé à l'Institut pour l'Avancée des Biosciences qui permet de découvrir des biomarqueurs pronostiques robustes à partir de données publiques d'expression de gènes. Ces biomarqueurs sont combinés dans un outil pronostique GEC qui permet de stratifier les patients en fonction du nombre d'activations dans le panel. Plus grand est le nombre de gènes activés de manière ectopique, moins longue est la probabilité de survie des patients.

Résultat : En appliquant la méthode « ectopy » à trois cohortes publiques du cancer du côlon (effectif total n=1141), nous avons défini une signature pronostique GEC de 4 gènes : ERFE, HOXC6, LAMP5 et ULBP2. En analyses uni- et multivariée, l'activation ectopique d'au moins un de ces 4 gènes est un facteur de pronostic défavorable, tous stades confondus, ainsi que dans chaque stade de cancer du côlon séparément. Suite à ces travaux, nous avons conçu une nouvelle cohorte rétrospective monocentrique du CHUGA, de 2012 à 2014, constituée de 133 patients afin de valider la signature GEC par la technologie RT-qPCR.

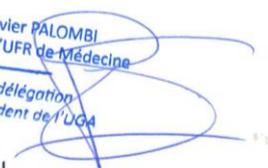
Discussion : La signature GEC prédit le pronostic de survie dans le cancer du côlon, indépendamment d'autres critères pronostiques. Cet outil aidera à la prise en charge des patients, en particulier dans les stades II de pronostic incertain et les stades IV où un traitement chirurgical par chimiothérapie hyperthermique est discuté. Les perspectives de l'étude sont une analyse par RT-qPCR sur la cohorte de cancer du côlon du CHUGA ainsi qu'une mise au point d'un test immunohistochimique pour un passage en routine.

VU ET PERMIS D'IMPRIMER

Grenoble, le : 09/08/2024

LE DOYEN DE LA FACULTÉ DE MÉDECINE

Pr Olivier PALOMBI
Doyen de l'UFR de Médecine
Par délégation
du Président de l'UGA



Pr Olivier PALOMBI

LE PRÉSIDENT / LA PRÉSIDENTE DU JURY



Pr Séverine VALMARY-DEGANO

BIBLIOGRAPHIE

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021;71(3):209-49.
2. Institut National du Cancer [Internet]. 2023 [cité 16 juill 2023]. Le cancer colorectal - Épidémiologie. Disponible sur: <https://www.e-cancer.fr/Professionnels-de-sante/Les-chiffres-du-cancer-en-France/Epidemiologie-des-cancers/Les-cancers-les-plus-frequents/Cancer-colorectal>
3. Siegel RL, Fedewa SA, Anderson WF, Miller KD, Ma J, Rosenberg PS, et al. Colorectal Cancer Incidence Patterns in the United States, 1974–2013. *JNCI J Natl Cancer Inst.* 1 août 2017;109(8):djw322.
4. Cardoso R, Guo F, Heisser T, Hackl M, Ihle P, De Schutter H, et al. Colorectal cancer incidence, mortality, and stage distribution in European countries in the colorectal cancer screening era: an international population-based study. *Lancet Oncol.* 1 juill 2021;22(7):1002-13.
5. Vasen HFA, Möslein G, Alonso A, Aretz S, Bernstein I, Bertario L, et al. Guidelines for the clinical management of familial adenomatous polyposis (FAP). *Gut.* 1 mai 2008;57(5):704-13.
6. Yurgelun MB, Kulke MH, Fuchs CS, Allen BA, Uno H, Hornick JL, et al. Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer. *J Clin Oncol.* avr 2017;35(10):1086-95.
7. Rombouts AJM, Hugén N, Beek JJP van, Poortmans PMP, Wilt JHW de, Nagtegaal ID. Does pelvic radiation increase rectal cancer incidence? – A systematic review and meta-analysis. *Cancer Treat Rev.* 1 juill 2018;68:136-44.
8. Hadjiliadis D, Khoruts A, Zauber AG, Hempstead SE, Maisonneuve P, Lowenfels AB, et al. Cystic Fibrosis Colorectal Cancer Screening Consensus Recommendations. *Gastroenterology.* 1 févr 2018;154(3):736-745.e14.

9. Rokkas T, Pistiolas D, Sechopoulos P, Margantinis G, Koukoulis G. Risk of colorectal neoplasm in patients with acromegaly: A meta-analysis. *World J Gastroenterol*. 14 juin 2008;14(22):3484-9.
10. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *The Lancet*. 19 oct 2019;394(10207):1467-80.
11. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, et al. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. juill 2012;487(7407):330-7.
12. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. nov 2015;21(11):1350-6.
13. Müller MF, Ibrahim AEK, Arends MJ. Molecular pathological classification of colorectal cancer. *Virchows Arch*. 1 août 2016;469(2):125-34.
14. Dunne PD, Arends MJ. Molecular pathological classification of colorectal cancer—an update. *Virchows Arch*. 1 févr 2024;484(2):273-85.
15. Kim JC, Bodmer WF. Genomic landscape of colorectal carcinogenesis. *J Cancer Res Clin Oncol*. 1 mars 2022;148(3):533-45.
16. Nassar D, Blanpain C. Cancer stem cells: Basic concepts and therapeutic implications. *Annual Review of Pathology: Mechanisms of Disease*. 2016;11(1):47–76. doi:10.1146/annurev-pathol-012615-044438.
17. Medema JP. Cancer stem cells: The challenges ahead. *Nat Cell Biol*. avr 2013;15(4):338-44.
18. Loree JM, Pereira AAL, Lam M, Willauer AN, Raghav K, Dasari A, et al. Classifying colorectal cancer by tumor location rather than sidedness highlights a continuum in mutation profiles and consensus molecular subtypes. *Clinical Cancer Research*. 2018;24(5):1062–72. doi:10.1158/1078-0432.ccr-17-2484.
19. Larsen W, Schoenwolf G, Brauer P, Francis-West P. *Embryologie humaine*. 4ème. De Boeck Supérieur; 2017. 560 p.

20. WHO Classification of Tumours Editorial Board. Digestive system tumours [Internet]. Lyon (France): International Agency for Research on Cancer; 2019. (WHO classification of tumours series, 5th ed.; vol. 1). Available from: <https://tumourclassification.iarc.who.int/chapters/31>.
21. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet*. 26 mai 2018;391(10135):2128-39.
22. Kim SH, Chang HJ, Kim DY, Park JW, Baek JY, Kim SY, et al. What Is the Ideal Tumor Regression Grading System in Rectal Cancer Patients after Preoperative Chemoradiotherapy? *Cancer Res Treat*. 22 oct 2015;48(3):998-1009.
23. Trakarnsanga A, Gönen M, Shia J, Nash GM, Temple LK, Guillem JG, et al. Comparison of Tumor Regression Grade Systems for Locally Advanced Rectal Cancer After Multimodality Treatment. *JNCI J Natl Cancer Inst*. 1 oct 2014;106(10):dju248.
24. Yuan H, Dong Q, Zheng B, Hu X, Xu JB, Tu S. Lymphovascular invasion is a high risk factor for stage I/II colorectal cancer: a systematic review and meta-analysis. *Oncotarget*. 16 févr 2017;8(28):46565-79.
25. Knijn N, van Exsel UEM, de Noo ME, Nagtegaal ID. The value of intramural vascular invasion in colorectal cancer – a systematic review and meta-analysis. *Histopathology*. 2018;72(5):721-8.
26. Knijn N, Mogk SC, Teerenstra S, Simmer F, Nagtegaal ID. Perineural Invasion Is a Strong Prognostic Factor in Colorectal Cancer: A Systematic Review. *Am J Surg Pathol*. janv 2016;40(1):103.
27. Nagtegaal ID, Knijn N, Hugén N, Marshall HC, Sugihara K, Tot T, et al. Tumor Deposits in Colorectal Cancer: Improving the Value of Modern Staging—A Systematic Review and Meta-Analysis. *J Clin Oncol*. avr 2017;35(10):1119-27.
28. Lugli A, Kirsch R, Ajioka Y, Bosman F, Cathomas G, Dawson H, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Mod Pathol*. 1 sept 2017;30(9):1299-311.

29. Konishi T, Shimada Y, Lee LH, Cavalcanti MS, Hsu M, Smith JJ, et al. Poorly Differentiated Clusters Predict Colon Cancer Recurrence: An In-Depth Comparative Analysis of Invasive-Front Prognostic Markers. *Am J Surg Pathol*. juin 2018;42(6):705.
30. Rozek LS, Schmit SL, Greenson JK, Tomsho LP, Rennert HS, Rennert G, et al. Tumor-Infiltrating Lymphocytes, Crohn's-Like Lymphoid Reaction, and Survival From Colorectal Cancer. *JNCI J Natl Cancer Inst*. 1 août 2016;108(8):djw027.
31. Brierley J, Gospodarowicz MK, Wittekind C. TNM classification of malignant tumours. 8th ed. Chichester, West Sussex, UK: John Wiley & Sons, Inc.; 2017.
32. Evaluation du statut MMR tumoral, collection Recommandations et référentiels, Institut national du cancer, décembre 2021 [Internet]. [cité 24 mars 2024]. Disponible sur: <https://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Evaluation-du-statut-MMR-tumoral-synthese-2021>
33. Patients atteints d'un adénocarcinome colorectal / Indications des tests moléculaires en vue de la prescription de traitements de précision, collection Recommandations et référentiels, Institut national du cancer, juin 2022 [Internet]. [cité 20 févr 2024]. Disponible sur: <https://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Patients-atteints-d-un-adenocarcinome-colorectal-Indications-des-tests-moleculaires-en-vue-de-la-prescription-de-traitements-de-precision>
34. Thierry AR, Mouliere F, El Messaoudi S, Mollevi C, Lopez-Crapez E, Rolet F, et al. Clinical validation of the detection of KRAS and BRAF mutations from circulating tumor DNA. *Nat Med*. avr 2014;20(4):430-5.
35. Hsu HC, Thiam TK, Lu YJ, Yeh CY, Tsai WS, You JF, et al. Mutations of KRAS/NRAS/BRAF predict cetuximab resistance in metastatic colorectal cancer patients. *Oncotarget*. 12 mars 2016;7(16):22257-70.
36. Dudley JC, Lin MT, Le DT, Eshleman JR. Microsatellite Instability as a Biomarker for PD-1 Blockade. *Clin Cancer Res*. 14 févr 2016;22(4):813-20.
37. Rousseau A, Marquet P, Bastian G, Milano G, Catalin J, Gamelin E, et al. Suivi thérapeutique pharmacologique du 5-fluoro-uracile (5-FU). *EMC - Biologie Médicale*. 2006;1(1):1-5.

38. Phelip JM, Tougeron D, Léonard D, Benhaim L, Desolneux G, Dupré A, Di Fiore F, Michel P, Penna C, Tournigand C, Louvet C, Christou N, Chevallier P, Dohan A, Rousseau B, Bouché O. Metastatic colorectal cancer (mCRC): French intergroup clinical practice guidelines for diagnosis, treatments and follow-up (SNFGE, FFCD, GERCOR, UNICANCER, SFCD, SFED, ACHBT, SFRO, SFR). *Dig Liver Dis.* 2019 Oct;51(10):1357- 1363. doi: 10.1016/j.dld.2019.05.035.
39. Lecomte T, Tougeron D, Chautard R, Bressand D, André T, Bibeau F, Blanc B, Cohen R, Jacques J, Lagasse JP, Laurent-Puig P, Lepage C, Lucidarme O, Martin-Babau J, Panis Y, Portales F, Taïeb J, Aparicio T, Bouché O. Non-metastatic colon cancer - French Intergroup Clinical Practice Guidelines for diagnosis, treatments and follow-up (TNCD, SNFGE, FFCD, GERCOR, UNICANCER, SFCD, SFED, SFRO, ACHBT, SFP, AFEF, SFR). *Dig Liver Dis.* 2023 in press.
40. Goere D. Prise en charge de la carcinose péritonéale d'origine colique. POST'U [Internet]. 2022 [cité 28 mai 2024]; Disponible sur: <https://www.fmcgastro.org/texte-postu/postu-2022/prise-en-charge-de-la-carcinose-peritoneale-dorigine-colique/>
41. Cervantes A, Adam R, Roselló S, Arnold D, Normanno N, Taïeb J, et al. Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol.* 1 janv 2023;34(1):10-32.
42. Gunderson LL, Jessup JM, Sargent DJ, Greene FL, Stewart A. Revised Tumor and Node Categorization for Rectal Cancer Based on Surveillance, Epidemiology, and End Results and Rectal Pooled Analysis Outcomes. *J Clin Oncol.* 10 janv 2010;28(2):256-63.
43. O'Connell JB, Maggard MA, Ko CY. Colon Cancer Survival Rates With the New American Joint Committee on Cancer Sixth Edition Staging. *JNCI J Natl Cancer Inst.* 6 oct 2004;96(19):1420-5.
44. Mitry E. Cancer colique de stade II : quel traitement adjuvant pour quels patients ? POST'U [Internet]. 2014 [cité 2 juin 2024]; Disponible sur: <https://www.fmcgastro.org/textes-postus/postu-2014/cancer-colique-de-stade-ii-quel-traitement-adjuvant-pour-quels-patients/>
45. Hahn WC, Bader JS, Braun TP, Califano A, Clemons PA, Druker BJ, et al. An expanded universe of cancer targets. *Cell.* 4 mars 2021;184(5):1142-55.

46. LNCC. CTpedia. [cité 2 juin 2024]. CTDatabase. Disponible sur: <http://www.cta.lncc.br/>
47. Uhlen M, Fagerberg L, Hallström B, Lindskog C, Oksvold P. The Human Protein Atlas Sweden (2015). Available at: <https://www.proteinatlas.org>.
48. NCBI Gene [Internet]. [cité 2 juin 2024]. Disponible sur: <https://www.ncbi.nlm.nih.gov/gene/>
49. Rousseaux S, Wang J, Khochbin S. Cancer hallmarks sustained by ectopic activations of placenta/male germline genes. *Cell Cycle*. 1 août 2013;12(15):2331-2.
50. Rousseaux S, Debernardi A, Jacquiau B, Vitte AL, Vesin A, Nagy-Mignotte H, et al. Ectopic Activation of Germline and Placental Genes Identifies Aggressive Metastasis-Prone Lung Cancers. *Sci Transl Med*. 22 mai 2013;5(186):186ra66-186ra66.
51. Bourova-Flin E, Derakhshan S, Goudarzi A, Wang T, Vitte AL, Chuffart F, et al. The combined detection of Amphiregulin, Cyclin A1 and DDX20/Gemin3 expression predicts aggressive forms of oral squamous cell carcinoma. *Br J Cancer*. oct 2021;125(8):1122-34.
52. Jacquet E, Chuffart F, Vitte AL, Nika E, Mousseau M, Khochbin S, et al. Aberrant activation of five embryonic stem cell-specific genes robustly predicts a high risk of relapse in breast cancers. *BMC Genomics*. 17 août 2023;24(1):463.
53. Emadali A, Rousseaux S, Bruder-Costa J, Rome C, Duley S, Hamaidia S, et al. Identification of a novel BET bromodomain inhibitor-sensitive, gene regulatory circuit that controls Rituximab response and tumour growth in aggressive lymphoid cancers. *EMBO Mol Med*. août 2013;5(8):1180-95.
54. Wang J, Mi JQ, Debernardi A, Vitte AL, Emadali A, Meyer JA, et al. A six gene expression signature defines aggressive subtypes and predicts outcome in childhood and adult acute lymphoblastic leukemia. *Oncotarget*. 12 mai 2015;6(18):16527-42.
55. Peng LJ, Zhou YB, Geng M, Bourova-Flin E, Chuffart F, Zhang WN, et al. Ectopic expression of a combination of 5 genes detects high risk forms of T-cell acute lymphoblastic leukemia. *BMC Genomics*. 24 juin 2022;23(1):467.
56. Rousseaux S, Reynoird N, Khochbin S. NUT Is a Driver of p300-Mediated Histone Hyperacetylation: From Spermatogenesis to Cancer. *Cancers*. janv 2022;14(9):2234.

57. Naciri I, Laisné M, Ferry L, Bourmaud M, Gupta N, Di Carlo S, et al. Genetic screens reveal mechanisms for the transcriptional regulation of tissue-specific genes in normal cells and tumors. *Nucleic Acids Res.* 23 avr 2019;47(7):3407-21.
58. Gazendam AM, Popovic S, Munir S, Parasu N, Wilson D, Ghert M. Synovial Sarcoma: A Clinical Review. *Curr Oncol.* juin 2021;28(3):1909-20.
59. Bode PK, Thielken A, Brandt S, Barghorn A, Lohe B, Knuth A, et al. Cancer testis antigen expression in testicular germ cell tumorigenesis. *Mod Pathol.* 1 juin 2014;27(6):899-905.
60. Lezcano C, Müller AM, Frosina D, Hernandez E, Geronimo JA, Busam KJ, et al. Immunohistochemical Detection of Cancer-Testis Antigen PRAME. *Int J Surg Pathol.* 1 déc 2021;29(8):826-35.
61. Haen SP, Löffler MW, Rammensee HG, Brossart P. Towards new horizons: characterization, classification and implications of the tumour antigenic repertoire. *Nat Rev Clin Oncol.* oct 2020;17(10):595-610.
62. Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer.* févr 2014;14(2):135-46.
63. Gordeeva O. Cancer-testis antigens: Unique cancer stem cell biomarkers and targets for cancer therapy. *Semin Cancer Biol.* 1 déc 2018;53:75-89.
64. Thomas R, Al-Khadairi G, Roelands J, Hendrickx W, Dermime S, Bedognetti D, et al. NY-ESO-1 Based Immunotherapy of Cancer: Current Perspectives. *Front Immunol.* 2018;9.
65. Taib N, Merhi M, Inchakalody V, Mestiri S, Hydrose S, Makni-Maalej K, et al. Treatment with decitabine induces the expression of stemness markers, PD-L1 and NY-ESO-1 in colorectal cancer: potential for combined chemoimmunotherapy. *J Transl Med.* 31 mars 2023;21(1):235.
66. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 1 janv 2013;41(D1):D991-5.

67. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostat Oxf Engl.* avr 2003;4(2):249-64.
68. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinforma Oxf Engl.* 1 janv 2013;29(1):15-21.
69. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinforma Oxf Engl.* 15 janv 2015;31(2):166-9.
70. EpiMed. *epimed/ectopy* [Internet]. 2024 [cité 2 juin 2024]. Disponible sur: <https://github.com/epimed/ectopy>
71. Schoenfeld DA. Sample-Size Formula for the Proportional-Hazards Regression Model. *Biometrics.* 1983;39(2):499-503.
72. INSEE. Décès en France - Registre des personnes décédées en France depuis 1970 [Internet]. [cité 2 juin 2024]. Disponible sur: <https://www.deces-en-france.fr>
73. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics.* 18 juin 2012;13(1):134.
74. Nucleotide - NCBI [Internet]. [cité 20 févr 2024]. Disponible sur: <https://www.ncbi.nlm.nih.gov/nucleotide/>
75. Kong M, Gao Y, Guo X, Xie Y, Yu Y. Role of the CTRP family in tumor development and progression (Review). *Oncol Lett.* 1 oct 2021;22(4):1-9.
76. Srole DN, Ganz T. Erythroferrone structure, function, and physiology: Iron homeostasis and beyond. *J Cell Physiol.* 2021;236(7):4888-901.
77. Xu Q, Altrock E, Schmitt N, Streuer A, Rapp F, Nowak V, et al. In Silico Pan-Cancer Analysis Reveals Prognostic Role of the Erythroferrone (ERFE) Gene in Human Malignancies. *Int J Mol Sci.* janv 2023;24(2):1725.

78. Liang L, Zeng J hui, Qin X gan, Chen J qiang, Luo D zhong, Chen G. Distinguishable Prognostic Signatures of Left- and Right-Sided Colon Cancer: a Study Based on Sequencing Data. *Cell Physiol Biochem*. 17 juill 2018;48(2):475-90.
79. Masson E. EM-Consulte. [cité 25 juin 2024]. Implications multiples du récepteur LRP-1 dans la progression tumorale. Disponible sur: <https://www.em-consulte.com/article/231561/implications-multiples-du-recepteur-lrp-1-dans-la->
80. Fraenkel PG. Anemia of Inflammation: A Review. *Med Clin North Am*. 1 mars 2017;101(2):285-96.
81. Ji M, Feng Q, He G, Yang L, Tang W, Lao X, et al. Silencing homeobox C6 inhibits colorectal cancer cell proliferation. *Oncotarget*. 12 avr 2016;7(20):29216-27.
82. Alajez NM. Large-Scale Analysis of Gene Expression Data Reveals a Novel Gene Expression Signature Associated with Colorectal Cancer Distant Recurrence. *PLOS ONE*. 9 déc 2016;11(12):e0167455.
83. Wang H, Li Z, Ou S, Song Y, Luo K, Guan Z, et al. Tumor Microenvironment Heterogeneity-Based Score System Predicts Clinical Prognosis and Response to Immune Checkpoint Blockade in Multiple Colorectal Cancer Cohorts. *Front Mol Biosci* [Internet]. 28 juin 2022 [cité 27 juin 2024]; Disponible sur: <https://www.frontiersin.org/journals/molecular-biosciences/articles/10.3389/fmolb.2022.884839/full>
84. Bai S, Chen L, Yan Y, Li R, Zhou Y, Wang X, et al. Exploration of Different Hypoxia Patterns and Construction of a Hypoxia-Related Gene Prognostic Index in Colorectal Cancer. *Front Immunol* [Internet]. 2022 [cité 1 oct 2023];13. Disponible sur: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.853352>
85. Qi L, Ye C, Zhang D, Bai R, Zheng S, Hu W, et al. The Effects of Differentially-Expressed Homeobox Family Genes on the Prognosis and HOXC6 on Immune Microenvironment Orchestration in Colorectal Cancer. *Front Immunol* [Internet]. 2021 [cité 1 oct 2023];12. Disponible sur: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.781221>

86. Qi L, Chen J, Zhou B, Xu K, Wang K, Fang Z, et al. HomeoboxC6 promotes metastasis by orchestrating the DKK1/Wnt/ β -catenin axis in right-sided colon cancer. *Cell Death Dis.* 1 avr 2021;12(4):1-13.
87. Alessandrini F, Pezzè L, Ciribilli Y. LAMPs: Shedding light on cancer biology. *Semin Oncol.* 1 août 2017;44(4):239-53.
88. Umeda S, Kanda M, Shimizu D, Nakamura S, Sawaki K, Inokawa Y, et al. Lysosomal-associated membrane protein family member 5 promotes the metastatic potential of gastric cancer cells. *Gastric Cancer.* 1 mai 2022;25(3):558-72.
89. Lee J, Sohn I, Do Ig, Km K, Park Sh, Park Jo, et al. Nanostring-based multigene assay to predict recurrence for gastric cancer patients after surgery. *PloS One* [Internet]. 3 mai 2014 [cité 4 oct 2023];9(3). Disponible sur: <https://pubmed.ncbi.nlm.nih.gov/24598828/>
90. Chen TQ, Huang HJ, Zhu SX, Chen XT, Pu KJ, Wang D, et al. Blockade of the lncRNA-DOT1L-LAMP5 axis enhances autophagy and promotes degradation of MLL fusion proteins. *Exp Hematol Oncol.* 19 févr 2024;13(1):18.
91. Chen Y, Ma T. LAMP5 may promote MM progression by activating p38. *Pathol Oncol Res.* 22 mars 2023;29:1611083.
92. Martinez-Romero J, Bueno-Fortes S, Martín-Merino M, Ramirez de Molina A, De Las Rivas J. Survival marker genes of colorectal cancer derived from consistent transcriptomic profiling. *BMC Genomics.* 11 déc 2018;19(8):857.
93. Shi J, Jiang D, Yang S, Sun Y, Wang J, Zhang X, et al. Molecular profile reveals immune-associated markers of lymphatic invasion in human colon adenocarcinoma. *Int Immunopharmacol.* 1 juin 2020;83:106402.
94. Sutherland CL, Chalupny NJ, Cosman D. The UL16-binding proteins, a novel family of MHC class I-related ligands for NKG2D, activate natural killer cell functions. *Immunol Rev.* 2001;181(1):185-92.
95. Jost S, Ahn J, Chen S, Yoder T, Gikundiro KE, Lee E, et al. Upregulation of the NKG2D Ligand ULBP2 by JC Polyomavirus Infection Promotes Immune Recognition by Natural Killer Cells. *J Infect Dis.* 15 juin 2024;229(6):1836-44.

96. Kegasawa T, Tatsumi T, Yoshioka T, Suda T, Ikezawa K, Nakabori T, et al. Soluble UL16-binding protein 2 is associated with a poor prognosis in pancreatic cancer patients. *Biochem Biophys Res Commun*. 10 sept 2019;517(1):84-8.
97. Feng R, Xu J, Huang J, Liu J, Wang X, Wang J, et al. An immune-related prognostic gene ULBP2 is correlated with immunosuppressive tumor microenvironment and immunotherapy in breast cancer. *Heliyon*. 15 janv 2024;10(1):e23687.
98. Yang X, Su X, Wang Z, Yu Y, Wu Z, Zhang D. ULBP2 is a biomarker related to prognosis and immunity in colon cancer. *Mol Cell Biochem*. 1 oct 2023;478(10):2207-19.
99. Demirkol S, Gomceli I, Isbilen M, Dayanc BE, Tez M, Bostanci EB, et al. A Combined ULBP2 and SEMA5A Expression Signature as a Prognostic and Predictive Biomarker for Colon Cancer. *J Cancer*. 9 avr 2017;8(7):1113-22.
100. Li C, Shen Z, Zhou Y, Yu W. Independent prognostic genes and mechanism investigation for colon cancer. *Biol Res*. 13 avr 2018;51(1):10.
101. Ruan GT, Wang S, Zhu LC, Liao XW, Wang XK, Liao C, et al. Investigation and verification of the clinical significance and perspective of natural killer group 2 member D ligands in colon adenocarcinoma. *Aging*. 27 avr 2021;13(9):12565-86.
102. NCBI GDC. The Cancer Genome Program, Colorectal Adenocarcinoma [Internet]. [cité 2 juin 2024]. Disponible sur: <https://portal.gdc.cancer.gov/projects/TCGA-COAD>
103. Huo T, Canepa R, Sura A, Modave F, Gong Y. Colorectal cancer stages transcriptome analysis. *PLOS ONE*. 28 nov 2017;12(11):e0188697.
104. Wang X, Baddoo MC, Yin Q. The placental specific gene, PLAC1, is induced by the Epstein-Barr virus and is expressed in human tumor cells. *Virology*. 9 juin 2014;11(1):107.
105. Bescont AL, Vitte AL, Debernardi A, Curtet S, Buchou T, Vayr J, et al. Receptor-Independent Ectopic Activity of Prolactin Predicts Aggressive Lung Tumors and Indicates HDACi-Based Therapeutic Strategies. *Antioxid Redox Signal*. 7 juill 2015;23(1):1.
106. del Toro N, Shrivastava A, Ragueneau E, Meldal B, Combe C, Barrera E, et al. The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res*. 7 janv 2022;50(D1):D648-53.

107. Latini FRM, Hemerly JP, Freitas BCG, Oler G, Riggins GJ, Cerutti JM. ABI3 ectopic expression reduces in vitro and in vivo cell growth properties while inducing senescence. *BMC Cancer*. 11 janv 2011;11:11.
108. Chen J, Ding J, Wang Z, Zhu J, Wang X, Du J. Identification of downstream metastasis-associated target genes regulated by LSD1 in colon cancer cells. *Oncotarget*. 21 mars 2017;8(12):19609-30.
109. Paravani Palacon M, Morandin Ferrisse T, De Oliveira Barbeiro C, Maria Sgavioli Massucato E, Bufalino A. Calreticulin Expression in Human Carcinomas: A Systematic Review and Meta-Analysis. *Asian Pac J Cancer Prev APJCP*. 1 sept 2023;24(9):2929-40.
110. Zhang P, Pan Y, Sun J, Pan G. Aberrant expression of LncRNA CASC2 mediated the cell viability, apoptosis and autophagy of colon cancer cells by sponging miR-19a via NF- κ B signaling pathway. *Int J Exp Pathol*. 2021;102(3):163-71.
111. Xie L, Song X, Lin H, Chen Z, Li Q, Guo T, et al. Aberrant activation of CYR61 enhancers in colorectal cancer development. *J Exp Clin Cancer Res*. 22 mai 2019;38(1):213.
112. Deb M, Sengupta D, Rath SK, Kar S, Parbin S, Shilpi A, et al. Clusterin gene is predominantly regulated by histone modifications in human colon cancer and ectopic expression of the nuclear isoform induces cell death. *Biochim Biophys Acta BBA - Mol Basis Dis*. 1 août 2015;1852(8):1630-45.
113. Almutairi MH, Alrubie TM, Alamri AM, Almutairi BO, Alrefaei AF, Arafah MM, et al. Cancer-Testis Gene Biomarkers Discovered in Colon Cancer Patients. *Genes*. 1 mai 2022;13(5):807.
114. Rashad Y, Olsson L, Israelsson A, Öberg Å, Lindmark G, Hammarström ML, et al. Lymph node CXCL17 messenger RNA: A new prognostic biomarker for colon cancer. *Tumour Biol J Int Soc Oncodevelopmental Biol Med*. sept 2018;40(9):1010428318799251.
115. Tarnowski M, Czerewaty M, Deskur A, Safranow K, Marlicz W, Uraśńska E, et al. Expression of Cancer Testis Antigens in Colorectal Cancer: New Prognostic and Therapeutic Implications. *Dis Markers*. 2016;2016:1987505.

116. Motalebzadeh J, Shabani S, Rezayati S, Shakournia N, Mirzaei R, Mahjoubi B, et al. Prognostic Value of FBXO39 and ETS-1 but not BMI-1 in Iranian Colorectal Cancer Patients. *Asian Pac J Cancer Prev APJCP*. 26 mai 2018;19(5):1357-62.
117. Zhang SD, McCrudden CM, Meng C, Lin Y, Kwok HF. The significance of combining VEGFA, FLT1, and KDR expressions in colon cancer patient prognosis and predicting response to bevacizumab. *OncoTargets Ther*. 2015;8:835-43.
118. Weng W, Okugawa Y, Toden S, Toiyama Y, Kusunoki M, Goel A. FOXM1 and FOXQ1 Are Promising Prognostic Biomarkers and Novel Targets of Tumor-Suppressive miR-342 in Human Colorectal Cancer. *Clin Cancer Res Off J Am Assoc Cancer Res*. 1 oct 2016;22(19):4947-57.
119. Schaks M, Allgoewer K, Nelson N, Ehm P, Heumann A, Ewald F, et al. Ectopic Expression of Hematopoietic SHIP1 in Human Colorectal Cancer. *Biomedicines*. juill 2020;8(7):215.
120. Yan G, Li S, Yue M, Li C, Kang Z. Lysine demethylase 5B suppresses CC chemokine ligand 14 to promote progression of colorectal cancer through the Wnt/ β -catenin pathway. *Life Sci*. 1 janv 2021;264:118726.
121. Liu Z, Guan C, Lu C, Liu Y, Ni R, Xiao M, et al. High NUSAP1 expression predicts poor prognosis in colon cancer. *Pathol - Res Pract*. 1 juill 2018;214(7):968-73.
122. Weber L, Al-Refae K, Ebbert J, Jägers P, Altmüller J, Becker C, et al. Activation of odorant receptor in colorectal cancer cells leads to inhibition of cell proliferation and apoptosis. *PLOS ONE*. 8 mars 2017;12(3):e0172491.
123. Ramadan RA, Morsy H, Samir M, Shamseya M, Shalaby M, El-Attar EA. Significance of cancer testis-associated antigens (SPAG9 and FBXO39) in colon cancer. *Indian J Cancer*. 2022;59(3):394-401.
124. Deng B, Liu XP, Wang X. Prognostic and Immunological Role of THBS2 in Colorectal cancer. *BioMed Res Int*. 2021;2021:1124985.

125. Ling Y, Zhang X, Bai Y, Li P, Wei C, Song T, et al. Overexpression of Mps1 in colon cancer cells attenuates the spindle assembly checkpoint and increases aneuploidy. *Biochem Biophys Res Commun.* 8 août 2014;450(4):1690-5.
126. Lund CM, Dyhl-Polk A, Nielsen DL, Riis LB. Wnt5a expression and prognosis in stage II-III colon cancer. *Transl Oncol.* janv 2021;14(1):100892.

ANNEXE

Annexe 1. Algorithmes décisionnels thérapeutiques proposés par le TNCD

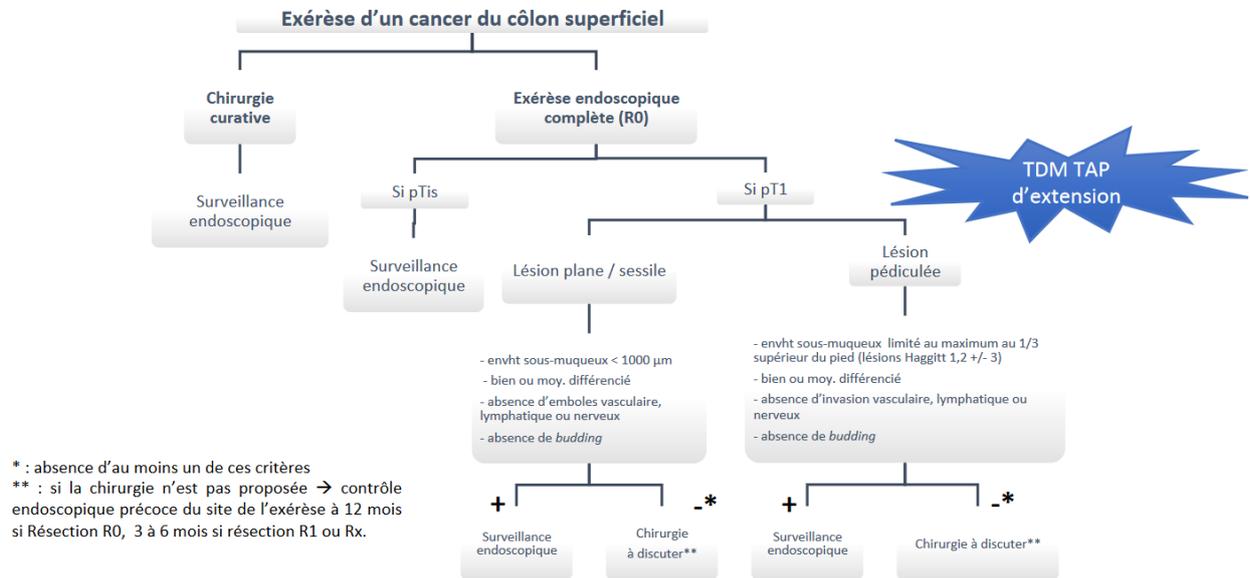


Figure 23. Algorithme décisionnel thérapeutique des CC superficiels (stade 0 et I) proposé par le TNCD (39)

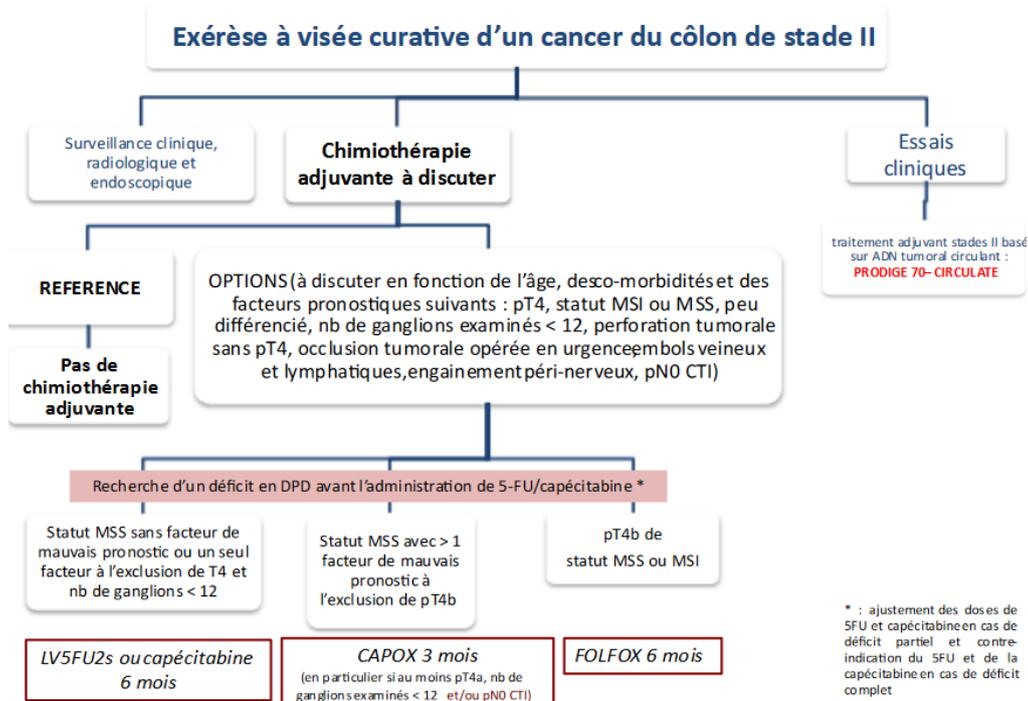
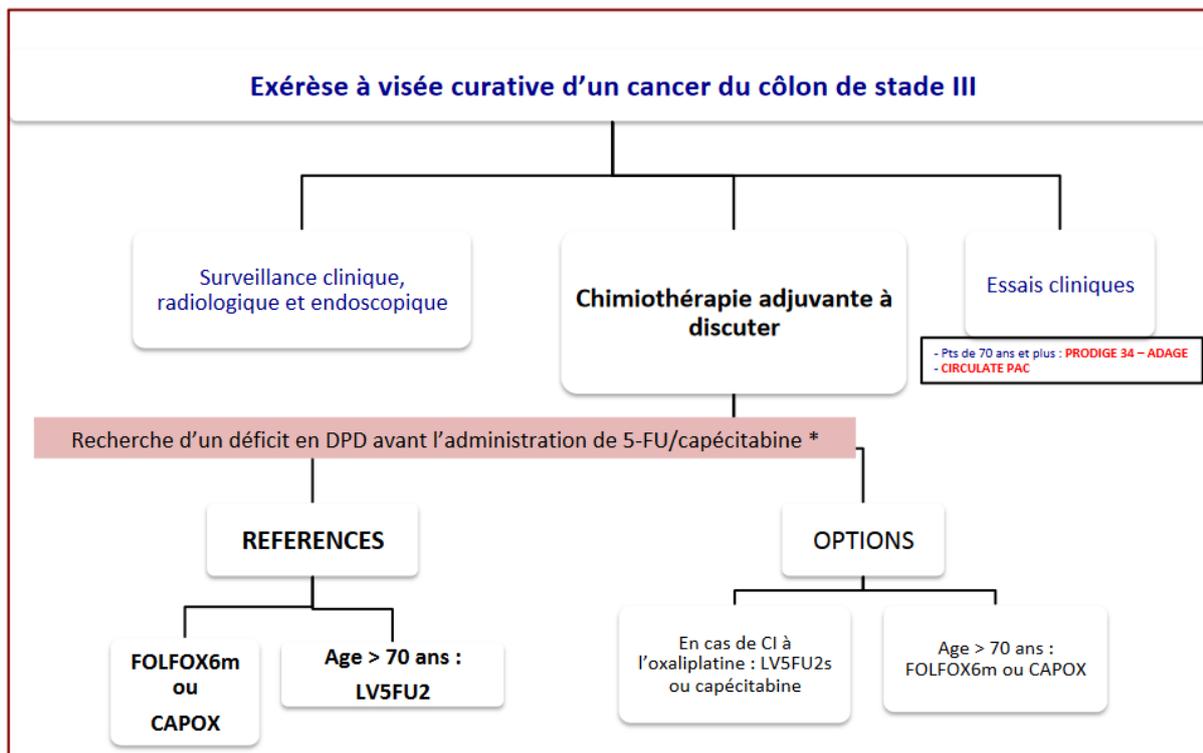


Figure 24. Algorithme décisionnel thérapeutique des CC de stade II proposé par le TNCD (39)



*: ajustement des doses de 5FU et de capécitabine en cas de déficit partiel et contre-indication du 5FU et de la capécitabine en cas de déficit complet; option : association raltitrexed-oxaliplatine

Figure 25. Algorithme décisionnel thérapeutique des CC de stade III proposé par le TNCD (39)

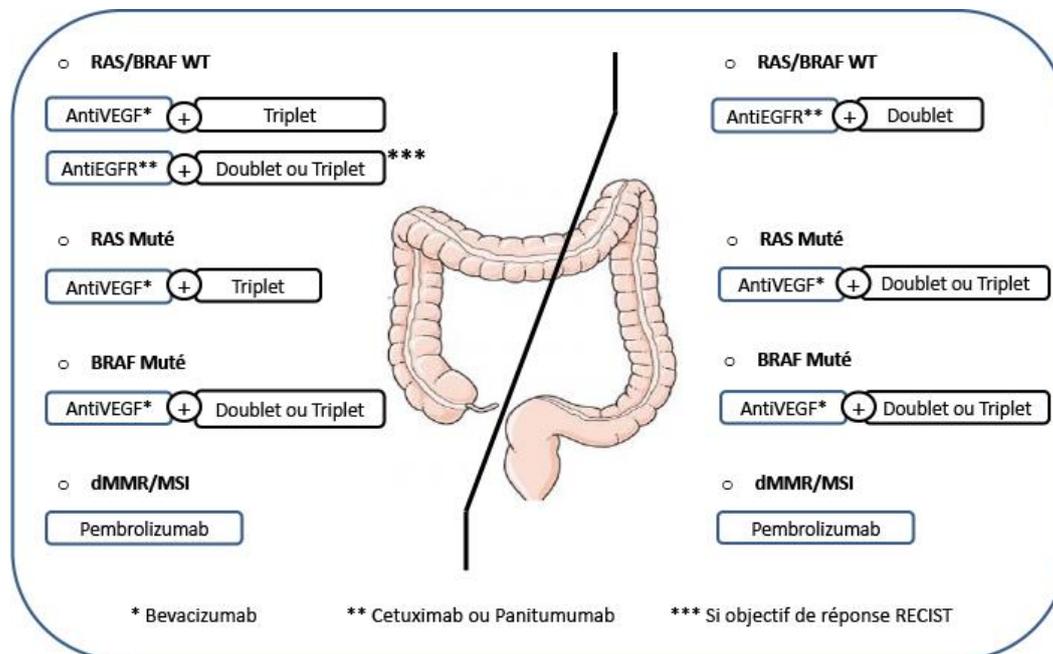


Figure 26. Indications des biothérapies dans les formes métastatiques (stade IV) du CC proposées par le TNCD, selon le phénotype moléculaire et la localisation du primitif (38)

Annexe 2. Liste non exhaustive des signatures pronostiques proposées dans le cancer du côlon

Nom (entreprise/pays)	Stade étudié	Stratification du pronostic	Technique	Prélèvement	Nombre de gènes
Oncotype DX (Genomic Health Inc.)	2	3 groupes	RT-qPCR	FFPE	12
Coloprint (Agendia)	2	2 groupes	FISH	Congélation	18
ColDx (Almac Group Limited Ltd.)	2	2 groupes	Puce Affymétrie	FFPE	634
Colo Guide Ex (Norvège)	2	2 groupes	Puce Affymétrie	Congélation	13
Onco Defender- CRC (Gopath DX)	1 et 2	2 groupes	PCR	FFPE	5
Sun Yat-sen University (Chine)	2	2 groupes	RT-qPCR	FFPE	6

Tableau 6. Liste non exhaustive des signatures pronostiques proposées dans le cancer du côlon

FFPE = Fixé dans le formol et enrobé dans la paraffine

Annexe 3. Séquences d'amorces utilisées dans les RT-qPCR

Gène	Forward (5'-3')	Reverse (5'-3')
Signature GEC		
ERFE	AACTAGCCTCCGGATGAGGT	CTTTCCCAGGGCAGAGTGTT
HOXC6	CTTCCTTATCCTGCCACCTCG	GTAGATCCGGTTCTGGGCAA
LAMP5	GACGTGTCTCATGGCAGAGT	GCGATATCGGCCTGTTCTGT
ULBP2	TCCAGGCTCTCCTTCCATCA	AGAAGGATCTTGGTAGCGGC
Gènes fréquemment activés dans CC		
CEMIP	ACCCATCACTCGGTCTCTGA	CAACTCAAGGCCAGGCAGAT
COL11A1	ACAATAGCACAGACGGAGGC	CCCAGAAACATGCCTAGGAGC
ESM1	TTGCTACCGCACAGTCTCAG	TGCCGTAGGGACAGTCTTTG
WNT2	AAAGAAGATGGGAAGCGCCA	TTCATCAGGGCTCTGGCATC
Gènes fréquemment activés dans tissu colique normal		
Actine	GCTGCACCCTTCAACAACCTG	GGTCAGTCCCTGGATCTCCT
U6	AAGGCCGGGCCATTACAATTA	CTTCATGCTAACACGGGCCT
AUP1	GGGAACTCAGTCCCTACCCA	GGCTTGCGGAGTCATCATCA

Tableau 7. Amorces utilisées pour l'analyse par RT-qPCR

Les amorces utilisées pour l'analyse par RT-qPCR ont été conçues à l'aide de l'outil Primer-BLAST du NCBI, avec comme condition que les séquences nécessitaient une taille limitée entre 70 à 150 paires de base et une température d'amorçage optimale (T_m) de 60°C (+/- 1°C).

Annexe 4. Expression physiologique des gènes ERFE, HOXC6, LAMP5 et ULBP2 selon les tissus

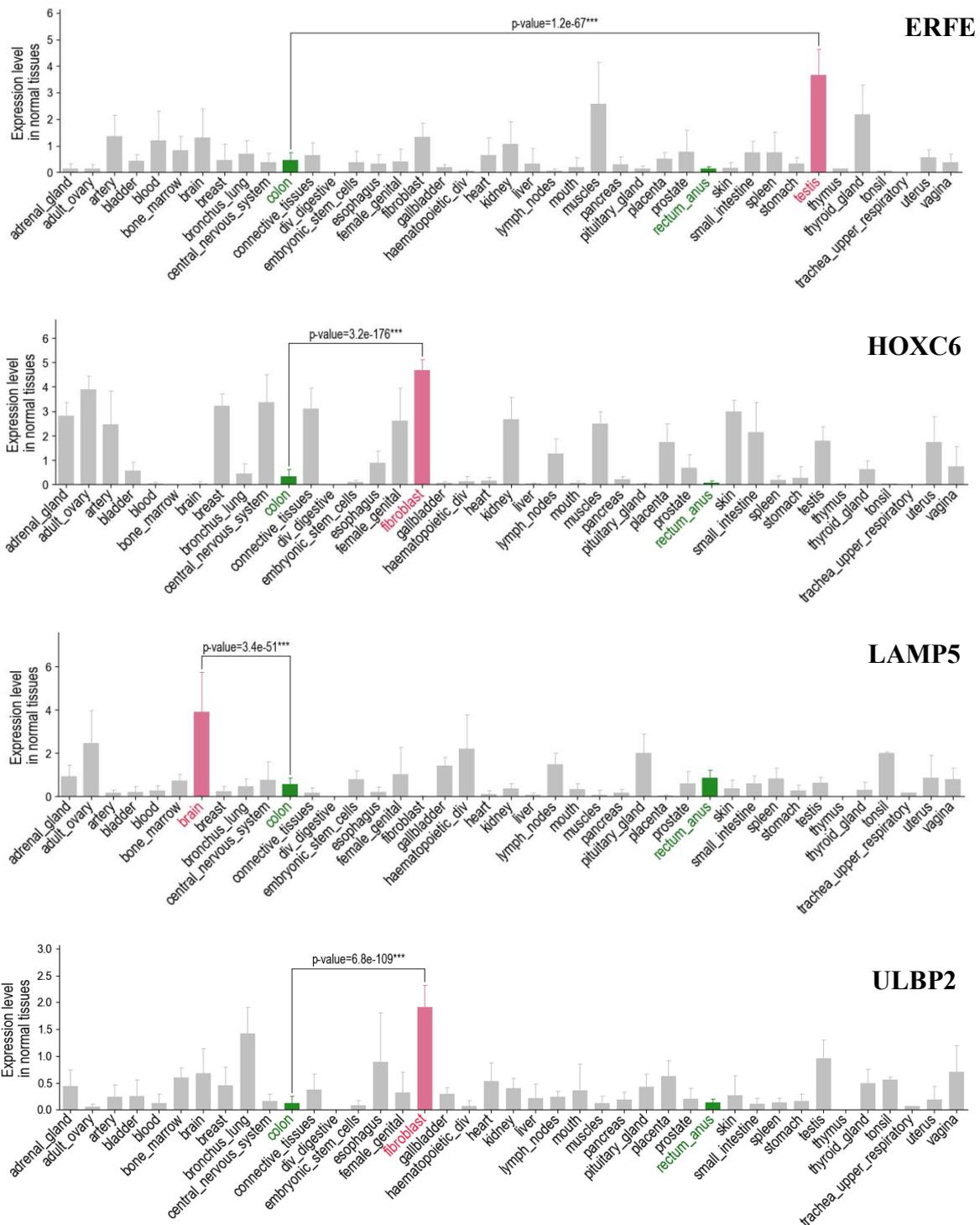


Figure 27. Histogrammes illustrant les expressions en fonction des tissus des gènes ERFE, HOXC6, LAMP5 et ULBP2

On note l'expression quasiment négligeable dans le côlon et la comparaison entre l'expression maximale dans un tissu et le côlon est statistiquement significative.

Annexe 5. Résultats de l'évaluation de l'interaction protéine-protéine avec IntAct

L'interaction protéine-protéine correspond à l'hypothèse que plusieurs protéines peuvent participer à une fonction plus ou moins commune en interagissant entre-elles. Plusieurs méthodes d'analyse permettent d'identifier ces interactions. La base de données IntAct se base sur les interactions protéine-protéine répertoriées dans la littérature scientifique et permet ainsi une recherche facilitée de celle-ci avec une estimation d'un indice de confiance (MI score). Ce dernier se basant notamment sur le nombre de revues publiées et la fiabilité des méthodes d'analyse utilisées pour déterminer cette interaction ; plus cet indice est important, plus l'interaction est fiable (106).

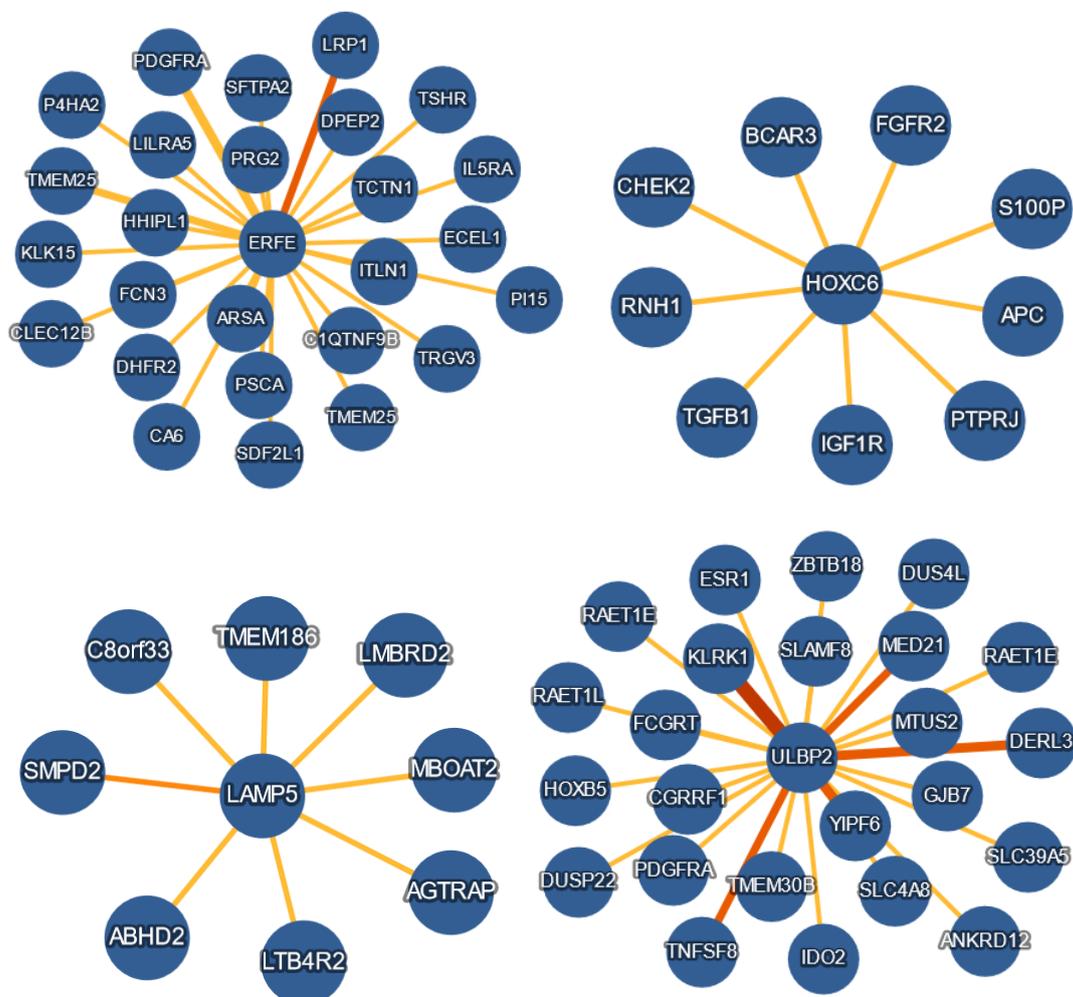


Figure 28. Représentation des potentielles interactions protéiques retrouvées sur IntAct pour ERFE, HOXC6, LAMP5 et ULBP2 chez l'Homme

Annexe 6. Liste de gènes spécifiques de tissus dont l'expression ectopique aurait un impact sur le pronostic dans le cancer du côlon

Gène	Famille	Locus	Fonction physiologique	Tissus normaux	CTA	Pronostic supposé dans le CC	Etude de survie réalisée	Retrouvé dans la liste des expressions ectopiques de notre étude (TCGA)	Source
ABI3	ABI	17q21.32	gène suppresseur de tumeur	rate, ganglion lymphatique, appendice	Non	Non connu (bon ?)	Non	Non	(107)
CABYR	CTA non lié à X	18q11.2	participe à la capacitation des spermatozoïdes	testicule	Oui	Non connu (mauvais ?)	Non	Oui	(108)
CALR	protéine chaperonne	19p13.13	adhésion cellulaire, homéostasie calcique, régulation transcriptomique	thyroïde, placenta, foie	Non	Mauvais	Oui	Oui	(109)
CASC2	ARN long non-codant	10q26.11	gène suppresseur de tumeur	thyroïde, cerveau, rein, testicule	Non	Non connu (bon ?)	Non	Oui	(110)
CCN1	facteur de croissance cellulaire	1p22.3	cycle cellulaire, angiogénèse, apoptose	vésicule biliaire, rate, vessie	Non	Non connu	Non	Non	(111)
CLU	protéine chaperonne	8p21.1	processus de mort cellulaire	cerveau, foie, testicule	Non	Non connu (bon ?)	Non	Non	(112)

CTAG1A	CTA lié à X	Xq28	Non connu	testicule, placenta	Oui	Non connu	Non	Oui	(113)
CXCL17	chimiokine	19q13.2	activation cellulaire et la stimulation de la migration des leucocytes	estomac, foie, œsophage	Non	Mauvais	Oui	Non	(114)
DCBLD2	protéine de la membrane plasmique	3q12.1	inhibition de la croissance cellulaire	testicule, cœur, thyroïde	Non	Mauvais	Oui	Non	(92)
DKKL1	CTA non lié à X	19q13.33	voie de signalisation Wnt	testicule	Oui	Mauvais	Oui	Oui	(115)
ERFE	C1q-related proteins (CTRP)	2q37.3	inhibition de la néoglucogénèse, activation de l'insuline, régulation transport lipides	thyroïde, testicule	Non	Mauvais	Oui	Oui	(77)
FBXO39	CTA non lié à X	17p13.1	participe à l'ubiquitination, régulation des processus cellulaires variables, inhibition de p53	testicule	Oui	Non connu (mauvais ?)	Non	Oui	(116)
FLT1	VEGFR	13q12.3	angiogénèse, vasculogénèse, croissance du placenta	placenta	Non	Mauvais	Oui	Oui	(117)
FOXM1	facteur de transcription de la famille FOX	12p13.33	développement embryonnaire, régulation du cycle cellulaire, signalisation cellulaire	testicule, moelle osseuse	Non	Mauvais	Oui	Non	(118)
FOXQ1	facteur de transcription de la famille FOX	6p25.3	développement embryonnaire, régulation du cycle cellulaire, signalisation cellulaire	estomac, vessie	Non	Mauvais	Oui	Non	(118)
GADD45B	gestion du stress cellulaire	19p13.3	régulation de la croissance et de l'apoptose cellulaire	foie, moelle osseuse	Non	Mauvais	Oui	Non	(92)

GALNAC T2	chondroïtine synthetase	10q11.21	formation de la matrice extracellulaire et signalisation cellulaire	moelle osseuse	Non	Mauvais	Oui	Non	(92)
HOXC6	homeobox	12q13.13	facteur de transcription impliqué dans la morphogenèse	ovaire, surrénal	Non	Mauvais	Oui	Oui	(85)
INPP5D	inositol phosphatase	2q37.1	inhibe la différenciation myéloïde	ganglion lymphatique, rate	Non	Mauvais	Oui	Non	(119)
KDM5B	CTA non lié à X	1q32.1	répression transcriptionnelle de gène suppresseur de tumeur	testicule	Oui	Mauvais	Oui	Non	(120)
LAMP5	protéine glycosylé	20p12.2	membrane des lysosomes	cerveau, ovaire	Non	Mauvais	Oui	Oui	(92)
LCA5	centrosome	6q14.1	potentielle fonction ciliaire primitive	testicule, thyroïde	Non	Mauvais	Oui	Oui	(92)
LEMD1	CTA non lié à X	1q32.1	composant membranaire	testicule	Oui	Mauvais	Oui	Oui	(92)
MAGEA3	CTA lié à X	Xq28	interaction avec p53	testicule	Oui	Non connu (mauvais ?)	Non	Oui	(115)
NPR3	récepteur de peptide natriurétique	5p13.3	natriurèse dans la régulation de la tension artérielle	rein	Non	Mauvais	Oui	Non	(92)
NUSAP1	protéine nucléolaire	15q15.1	organisation des microtubules	moelle osseuse, ganglion lymphatique	Non	Mauvais	Oui	Non	(121)
OR51B4	RCPG	11p15.4	récepteur olfactif	système olfactif	Non	Non connu (bon ?)	Non	Non	(122)

PLAC1	CTA lié à X	Xq26.3	développement du placenta, adhésion du trophoblaste	placenta	Oui	Mauvais	Oui	Oui	(104)
PTPN14	tyrosine phosphatase	1q32.3-q41	croissance cellulaire, différenciation et cycle mitotique	peau, rein	Non	Mauvais	Oui	Non	(92)
SLC2A3	Glucose binding protein	12p13.31	participe au transport intracellulaire du glucose	moelle osseuse	Non	Mauvais	Oui	Non	(92)
SPAG9	CTA non lié à X	17q21.33	participe au cycle mitotique et au transport des microtubules	testicule	Oui	Non connu (mauvais ?)	Non	Non	(123)
THBS2	thrombospondine	6q27	participe à l'adhésion cellulaire et la migration	vésicule biliaire, endomètre	Non	Mauvais	Oui	Non	(124)
TM4SF1	tetraspanin	3q25.1	transduction du signal	vésicule biliaire, poumon, œsophage	Non	Mauvais	Oui	Non	(92)
TTK	CTA non lié à X	6q14.1	régulation de la mitose	testicule, moelle osseuse	Oui	Non connu (mauvais ?)	Non	Oui	(125)
ULBP2	CMH classe I	6q25.1	CMH1, activation lymphocyte NK (ligand des NK)	œsophage, poumon	Non	Mauvais	Oui	Oui	(99)
WNT5A	Wnt	3p14.3	signalisation cellulaire	vessie, endomètre	Non	Bon	Oui	Non	(126)

Tableau 8. Liste de gènes spécifiques de tissus dont l'expression ectopique aurait un impact sur le pronostic dans le CC

SERMENT D'HIPPOCRATE

Le serment d'Hippocrate

Texte revu par l'Ordre des médecins en 2012

“ Au moment d'être admis(e) à exercer la médecine, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité.

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences.

Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences.

Je donnerai mes soins à l'indigent et à quiconque me les demandera. Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis(e) dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçu(e) à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs.

Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité. Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses ; que je sois déshonoré(e) et méprisé(e) si j'y manque.

”