

## DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **29 septembre 2025**

Nom de famille et prénom de l'auteur. e : **M. Ahmed Abdenour BOURAS**

Titre de la thèse : Explorations génétiques et moléculaires pour optimiser le diagnostic du cancer colorectal héréditaire sur la base de données de séquençage à haut débit

### Résumé



Les syndromes de prédisposition héréditaire au cancer colorectal (CCR) sont le plus souvent associés à des variants pathogènes dans des gènes à haute pénétrance tels que APC (Adenomatous Polyposis Coli) ou dans les gènes du système de réparation des mésappariements de l'ADN (MMR). Toutefois, malgré les avancées du séquençage à haut débit (SHD), un nombre significatif de cas familiaux reste sans explication moléculaire. Cette thèse explore des approches complémentaires visant à améliorer le diagnostic du CCR héréditaire, en identifiant et caractérisant de nouveaux variants et gènes candidats. Le premier axe porte sur l'analyse de la transcription et de l'épissage de l'ARNm comme outil de classification des variants de signification incertaine (VSI) et de résolution des cas en impasse diagnostique. Trente-six variants affectant les gènes MMR, sélectionnés sur la base de prédictions in silico, ont été étudiés par RT-PCR et/ou RNA-seq ciblé. Ces analyses ont permis la reclassification de 80 % des VSI (72 % comme pathogènes ou probablement pathogènes, 8 % comme probablement bénins). Le RNA-seq ciblé, plus sensible que la RT-PCR, a permis d'identifier des événements d'épissage complexes, tels que l'inclusion d'exons alternatifs en aval de la région 3'UTR de MSH2, ainsi que des altérations introniques profondes, notamment des insertions de rétrotransposons. Indétectables par les analyses ADN classiques en raison de leur nature ou de leur localisation, ces anomalies révèlent de nouveaux mécanismes d'inactivation génique et contribuent à expliquer une partie de l'hérédité manquante dans le CCR. Le deuxième axe est centré sur l'analyse des variants du gène PMS2 (Postmeiotic Segregation Increased 2), dont l'étude constitue un défi en SHD du fait de sa forte homologie avec le pseudogène PMS2CL. Dans cette étude, 42 variants détectés par SHD chez 76 patients ont été analysés par PCR long-range spécifique, couplée à une étude des transcrits. Cette approche a permis de reclasser 13 variants comme pathogènes, incluant des réarrangements

complexes (duplications de grande taille, conversions géniques), et d'exclure plusieurs SNPs bénins situés dans PMS2CL, parfois initialement interprétés à tort comme pathogènes. Le troisième axe porte sur l'exploration de gènes candidats dits « recherche », encore non intégrés aux recommandations cliniques en raison de la rareté des mutations rapportées et du manque de données robustes sur leur contribution au risque tumoral. L'analyse de 325 panels NGS a permis d'identifier neuf variants candidats (2,7 %) dans RNF43, AXIN2, NTHL1, TP53 et BUB1, associés à des phénotypes compatibles avec les rares cas décrits dans la littérature. Bien que non actionnables à ce jour, ces gènes illustrent l'importante hétérogénéité génétique du CCR héréditaire, chacun expliquant une part très minoritaire des cas. Ces travaux démontrent que l'intégration d'analyses fonctionnelles, de stratégies ciblant des régions génomiques complexes et d'un élargissement raisonné des panels de gènes permet d'améliorer significativement le diagnostic moléculaire du CCR héréditaire. Le RNA-seq ciblé s'impose comme un outil essentiel pour lever les impasses diagnostiques en oncogénétique.

**Mots-clés :** cancers colorectal héréditaires, séquençage à haut débit, variants de signification indéterminé, Syndrome de Lynch, RNA-seq, analyse d'épissage