

NGUYỄN THỊ DIỄM

SINH HỌC THỰC NGHIỆM

NĂM 2026

**BỘ GIÁO DỤC
VÀ ĐÀO TẠO**

**VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM**

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Nguyễn Thị Diễm

**NGHIÊN CỨU VAI TRÒ VÀ BIỂU HIỆN CỦA RNA DÀI
KHÔNG MÃ HÓA *CRNDE* TRONG UNG THƯ ĐẠI TRỰC
TRÀNG BẰNG PHÂN TÍCH DỮ LIỆU VÀ
REAL-TIME PCR**

LUẬN VĂN THẠC SĨ SINH HỌC

Hà Nội- Năm 2026

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Nguyễn Thị Diễm

**NGHIÊN CỨU VAI TRÒ VÀ BIỂU HIỆN CỦA RNA DÀI KHÔNG
MÃ HÓA *CRNDE* TRONG UNG THƯ ĐẠI TRỰC TRÀNG BẰNG
PHÂN TÍCH DỮ LIỆU VÀ
REAL-TIME PCR**

LUẬN VĂN THẠC SĨ SINH HỌC

Ngành: Sinh học thực nghiệm

Mã số: 8420114

NGƯỜI HƯỚNG DẪN KHOA HỌC:

TS. Hà Thị Thu Vân

PGS.TS. Võ Thị Bích Thủy

Hà Nội- Năm 2026

LỜI CAM ĐOAN

Tôi xin cam đoan đề tài nghiên cứu trong luận văn này là công trình nghiên cứu của tôi dựa trên những tài liệu, số liệu do chính tôi tự tìm hiểu và nghiên cứu. Chính vì vậy, các kết quả nghiên cứu đảm bảo trung thực và khách quan nhất. Đồng thời, kết quả này chưa từng xuất hiện trong bất cứ một nghiên cứu nào. Các số liệu, kết quả nêu trong luận văn là trung thực nếu sai tôi hoàn chịu trách nhiệm trước pháp luật.

Tác giả luận văn ký và ghi rõ họ tên

LỜI CẢM ƠN

Lời đầu tiên, tôi xin bày tỏ lòng biết ơn chân thành tới Học viện Khoa học và Công nghệ – Viện Hàn lâm Khoa học và Công nghệ Việt Nam, nơi đã tạo điều kiện học tập, nghiên cứu và nuôi dưỡng niềm đam mê khoa học trong tôi. Sự tận tâm của quý thầy cô cùng môi trường học thuật nghiêm túc, giàu cảm hứng tại Học viện đã giúp tôi xây dựng nền tảng kiến thức vững chắc và định hướng rõ ràng trên con đường nghiên cứu khoa học.

Tôi xin được gửi lời cảm ơn sâu sắc tới PGS.TS Võ Thị Bích Thủy – Phòng Công nghệ sinh học môi trường, Viện Sinh học, người đã tận tình hướng dẫn, định hướng và luôn đồng hành cùng tôi trong suốt quá trình thực hiện đề tài. Với kiến thức chuyên môn sâu rộng và sự tận tâm trong công việc, cô đã giúp tôi tiếp cận vấn đề một cách khoa học, hoàn thiện tư duy nghiên cứu và vượt qua những khó khăn trong quá trình thực hiện luận văn.

Tôi cũng xin chân thành cảm ơn TS Hà Thị Thu Vân – Khoa Sinh học phân tử, Bệnh viện Quân y 103, Học viện Quân y, đã nhiệt tình hỗ trợ, tạo điều kiện thuận lợi và trực tiếp hướng dẫn tôi trong quá trình thực nghiệm. Những góp ý và hỗ trợ quý báu của cô đã góp phần quan trọng giúp tôi hoàn thành phần nghiên cứu một cách hiệu quả.

Cuối cùng, tôi xin dành lời cảm ơn sâu sắc tới gia đình, bạn bè và đồng nghiệp, những người đã luôn bên cạnh động viên, hỗ trợ cả về tinh thần lẫn vật chất, là nguồn động lực to lớn giúp tôi vượt qua những khó khăn và hoàn thành luận văn này.

Xin trân trọng cảm ơn!

Tác giả luận văn

MỤC LỤC

LỜI CAM ĐOAN	<i>i</i>
LỜI CẢM ƠN	<i>ii</i>
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ CÁI VIẾT TẮT	<i>v</i>
DANH MỤC CÁC BẢNG	<i>vi</i>
MỞ ĐẦU	<i>1</i>
1. Lý do chọn đề tài	<i>1</i>
2. Mục tiêu nghiên cứu	<i>1</i>
3. Nội dung nghiên cứu	<i>2</i>
4. Cơ sở khoa học và tính thực tiễn của đề tài	<i>2</i>
5. Những đóng góp của luận văn	<i>2</i>
CHƯƠNG 1. TỔNG QUAN NGHIÊN CỨU	<i>3</i>
1.1. TỔNG QUAN VỀ UNG THƯ ĐẠI TRỰC TRÀNG	<i>3</i>
1.1.1 Dịch tễ học UTĐTT	<i>3</i>
1.1.2 Cơ chế bệnh sinh của UTĐTT	<i>4</i>
1.1.3. Phương pháp chẩn đoán UTĐTT	<i>6</i>
1.2. RNA DÀI KHÔNG MÃ HÓA	<i>7</i>
1.2.1. Khái niệm và phân loại lncRNA	<i>7</i>
1.2.2. Cơ chế điều hòa biểu hiện gen của lncRNA	<i>8</i>
1.2.3. Vai trò của lncRNA trong ung thư	<i>9</i>
1.3. LNCRNA CRNDE	<i>11</i>
1.3.1. Đặc điểm cấu trúc và biểu hiện của lnc CRNDE	<i>11</i>
1.3.2. Vai trò của lnc CRNDE trong ung thư đại trực tràng	<i>12</i>
CHƯƠNG 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU	<i>13</i>
2.1. ĐỐI TƯỢNG NGHIÊN CỨU	<i>13</i>
2.1.1. Dữ liệu công khai (TCGA, GEO)	<i>13</i>
2.1.2. Mẫu máu ngoại vi.....	<i>13</i>
2.2. Phạm vi và thiết kế nghiên cứu	<i>14</i>
2.3 Phương pháp phân tích dữ liệu công khai	<i>14</i>
2.3.1 Thu thập và tiền xử lý dữ liệu biểu hiện gen	<i>14</i>
2.3.2 Phân tích sự khác biệt biểu hiện của lnc CRNDE trong UTĐTT	<i>14</i>
2.3.3 Phân tích gen đồng biểu hiện	<i>15</i>
2.3.4 Phân tích chức năng và con đường tín hiệu	<i>15</i>
2.3.5. Phân tích sống sót.....	<i>15</i>

2.4 Phương pháp thực nghiệm	16
2.4.1. Tách chiết RNA tổng số	16
2.4.2 Tổng hợp cDNA.....	16
2.4.3 Đánh giá mức độ biểu hiện lncRNA <i>CRNDE</i> thông qua phản ứng real-time PCR	17
2.4.4 Phân tích dữ liệu real-time PCR	17
2.5 Phương pháp xử lý và phân tích thống kê.....	18
2.6 Đạo đức nghiên cứu.....	18
Sơ đồ nghiên cứu	18
<i>CHƯƠNG 3: KẾT QUẢ VÀ THẢO LUẬN.....</i>	<i>19</i>
3.1. KẾT QUẢ PHÂN TÍCH DỮ LIỆU TIN SINH HỌC	19
3.1.1 Mức độ biểu hiện của <i>CRNDE</i> trong UTĐTT.....	19
3.1.2 Xác định các gen đồng biểu hiện với <i>CRNDE</i>	22
3.1.3 Phân tích tăng cường chức năng (GO).....	24
3.1.4 Phân tích con đường tín hiệu (KEGG)	26
3.1.5 Phân tích sống sót.....	28
3.2 KẾT QUẢ ĐÁNH GIÁ TRÊN MẪU LÂM SÀNG	30
3.2.1 Tách chiết mẫu RNA tổng số, kiểm tra chất lượng và tổng hợp cDNA .	30
3.2.2 Mức độ biểu hiện của <i>CRNDE</i>	31
3.2.3 Đánh giá hiệu quả phân biệt <i>CRNDE</i> giữa UTĐTT và nhóm đối chứng bằng ROC	32
THẢO LUẬN.....	35
<i>KẾT LUẬN VÀ KIẾN NGHỊ</i>	<i>40</i>
KẾT LUẬN.....	40
KIẾN NGHỊ.....	40
<i>DANH MỤC TÀI LIỆU THAM KHẢO</i>	<i>41</i>
<i>PHỤ LỤC 1.....</i>	<i>47</i>
<i>PHỤ LỤC 2.....</i>	<i>50</i>
<i>PHỤ LỤC 3.....</i>	<i>51</i>

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ CÁI VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
ADR	Adenoma Detection Rate	Tỷ lệ phát hiện polyp tuyến
AI	Artificial Intelligence	Trí tuệ nhân tạo
ceRNA	Competing Endogenous RNA	RNA nội sinh cạnh tranh
CIMP	CpG Island Methylator Phenotype	Hiện tượng methyl hóa đảo CpG
CIN	Chromosomal Instability	Bất ổn nhiễm sắc thể
circRNA	Circular RNA	RNA vòng
CRNDE	Colorectal Neoplasia Differentially Expressed	RNA liên quan đến tân sinh đại trực tràng
GEO	Gene Expression Omnibus	Cơ sở dữ liệu biểu hiện gen
GO	Gene Ontology	Phân loại chức năng gen
HGNC	HUGO Gene Nomenclature Committee	Ủy ban Danh pháp Gen Người
KEGG	Kyoto Encyclopedia of Genes and Genomes	Cơ sở dữ liệu con đường sinh học
LINC	Long Intergenic Non-Protein Coding RNA	RNA dài không mã hóa nằm giữa các gen
lncRNA	Long non-coding RNA	RNA dài không mã hóa
miRNA	microRNA	RNA vi mô
MRI	Magnetic Resonance Imaging	Cộng hưởng từ
MSI	Microsatellite Instability	Bất ổn vi vệ tinh
PRC2	Polycomb Repressive Complex 2	Phức hợp ức chế Polycomb 2
TCGA	The Cancer Genome Atlas	Bộ bản đồ hệ gen ung thư
UTĐT	Colorectal cancer	Ung thư đại trực tràng

DANH MỤC CÁC BẢNG

<i>Bảng 2.1 Thành Phần phản ứng tổng hợp cDNA</i>	<i>16</i>
<i>Bảng 2.2 Trình tự môi khuếch đại gen đích và gen nội chuẩn</i>	<i>17</i>
<i>Bảng 3.1 Mức độ biểu hiện của lncRNA CRNDE.....</i>	<i>31</i>
<i>Bảng 3.2 Điểm cắt, AUC, độ nhạy, độ đặc hiệu và giá trị tiên đoán của lncRNA CRNDE</i>	<i>33</i>
<i>Sơ đồ 3.1 Mô hình cơ chế giả định của lncRNA CRNDE trong UTĐTT</i>	<i>31</i>

DANH MỤC CÁC HÌNH VẼ, ĐỒ THỊ

<i>Hình 1.1 Hình ảnh đại trực tràng</i>	3
<i>Hình 1.2 Tổng số ca mắc ung thư trên toàn thế giới theo thống kê của GLOBOCAN năm 2022</i>	4
<i>Hình 1.3 Quá trình bệnh sinh và các biến đổi phân tử của UTĐTT [8]</i>	5
<i>Hình 1.4 Tổng quan về phân loại và chức năng của lncRNA</i>	8
<i>Hình 1.5 Vị trí gen của CRNDE [45]</i>	11
<i>Hình 3.1 Mức độ biểu hiện của lncRNA CRNDE trong dữ liệu TCGA</i>	11
<i>Hình 3.2 Mức độ biểu hiện của lncRNA CRNDE theo từng giai đoạn bệnh trong dữ liệu TCGA</i>	20
<i>Hình 3.3 Mức độ biểu hiện của lncRNA CRNDE trong dữ liệu GEO</i>	21
<i>Hình 3.4 Mức độ biểu hiện của lncRNA CRNDE theo từng giai đoạn bệnh trong dữ liệu GEO</i>	22
<i>Hình 3.5 Mạng lưới các gen có liên quan đến lncRNA CRNDE</i>	23
<i>Hình 3.6 Kết quả phân tích GO từ dữ liệu TCGA</i>	24
<i>Hình 3.7 Kết quả phân tích GO từ dữ liệu GEO</i>	25
<i>Hình 3.8 Kết quả phân tích KEGG pathway từ dữ liệu TCGA</i>	26
<i>Hình 3.9 Kết quả phân tích KEGG pathway từ dữ liệu GEO</i>	28
<i>Hình 3.10 Mức độ biểu hiện lncRNA CRNDE liên quan đến sống sót từ dữ liệu TCGA</i>	29
<i>Hình 3.11 Mức độ biểu hiện lncRNA CRNDE liên quan đến sống sót từ dữ liệu GEO</i>	30
<i>Hình 3.12 Điện di đồ sản phẩm RNA tổng số</i>	31
<i>Hình 3.13 Phân bố mức độ biểu hiện của lncRNA CRNDE trên mẫu lâm sàng</i>	32
<i>Hình 3.14 ROC curve của lncRNA CRNDE cho phân biệt CRC và nhóm đối chứng</i>	33

MỞ ĐẦU

1. Lý do chọn đề tài

Ung thư đại trực tràng (UTĐTT) là một trong những bệnh ung thư phổ biến và gây tử vong cao trên thế giới, xếp thứ ba về tỷ lệ mắc và thứ hai về tỷ lệ tử vong theo GLOBOCAN 2022. Xu hướng này không chỉ thể hiện ở các quốc gia phát triển mà còn ngày càng rõ rệt tại các nước đang phát triển, trong đó có Việt Nam. Đây thực sự là một gánh nặng y tế – xã hội lớn khi UTĐTT thường tiến triển âm thầm, ít biểu hiện ở giai đoạn đầu và chỉ được phát hiện khi bệnh đã ở giai đoạn muộn.

Mặc dù đã có những tiến bộ trong phẫu thuật, hóa trị, xạ trị và các liệu pháp nhắm đích làm tăng hiệu quả điều trị và tỷ lệ sống sót dài hạn của bệnh nhân, việc phát hiện và điều trị UTĐTT vẫn chưa được cải thiện như kỳ vọng. Một trong những nguyên nhân then chốt là thiếu các dấu ấn sinh học tin cậy giúp phát hiện sớm và hỗ trợ tiên lượng. Nếu khoảng trống này không được giải quyết, bệnh nhân sẽ tiếp tục đối mặt với nguy cơ phát hiện muộn, làm tăng tỷ lệ tử vong; đồng thời hệ thống y tế sẽ phải chịu gánh nặng chi phí điều trị ngày càng cao.

Trong bối cảnh đó, việc nghiên cứu và xác định các dấu ấn sinh học mới là một yêu cầu cấp thiết. Trong những năm gần đây, nhóm RNA dài không mã hóa (long non-coding RNA- lncRNA) được xem là nguồn ứng viên tiềm năng cho nghiên cứu, chẩn đoán, và tiên lượng UTĐTT. Tuy nhiên, tại Việt Nam, các nghiên cứu về vai trò và giá trị ứng dụng của lncRNA trong UTĐTT, đặc biệt là lncRNA *CRNDE* (*Colorectal Neoplasia Differentially Expressed*) còn rất hạn chế.

Xuất phát từ thực tiễn trên, đề tài “Nghiên cứu vai trò và biểu hiện của RNA dài không mã hóa *CRNDE* trong ung thư đại trực tràng bằng phân tích dữ liệu và real-time PCR” được triển khai với mục tiêu khai thác dữ liệu lncRNA *CRNDE* công khai kết hợp với thực nghiệm trên mẫu lâm sàng.

2. Mục tiêu nghiên cứu

- Xác định vai trò và mức độ biểu hiện của lncRNA *CRNDE* thông qua phân tích dữ liệu công khai trên TCGA, GEO.
- Xác thực mức độ biểu hiện của lncRNA *CRNDE* bằng real-timePCR trên mẫu lâm sàng (bao gồm mẫu bệnh nhân UTĐTT, polyp và người khỏe mạnh)

3. Nội dung nghiên cứu

Nội dung 1: Khai thác và phân tích dữ liệu công khai (TCGA, GEO ...) đánh giá mức độ biểu hiện của lncRNA *CRNDE* trong ung thư đại trực tràng. (Mục tiêu 1)

Nội dung 2: Phân tích tăng cường chức năng (Mục tiêu 1)

Nội dung 3: Tách chiết RNA tổng số từ mẫu máu ngoại vi của ba nhóm đối tượng (UTĐTT, polyp và khỏe mạnh), tổng hợp cDNA và định lượng biểu hiện của lncRNA *CRNDE* bằng kỹ thuật real-timePCR, so sánh giữa các nhóm mẫu. (Mục tiêu 2)

Nội dung 4: Đối chiếu và phân tích so sánh giữa dữ liệu *in silico* và kết quả thực nghiệm nhằm đánh giá độ tương đồng và tính tin cậy của biểu hiện lncRNA *CRNDE* trong UTĐTT. (Mục tiêu 2)

4. Cơ sở khoa học và tính thực tiễn của đề tài

Đề tài này có ý nghĩa thực tiễn ở chỗ không chỉ làm rõ hơn vai trò sinh học của lncRNA *CRNDE* thông qua phân tích dữ liệu công khai, mà còn bước đầu đánh giá khả năng ứng dụng của phân tử này trong phân biệt UTĐTT.

Ngoài ra, trong bối cảnh tại Việt Nam các nghiên cứu về lncRNA trong UTĐTT còn hạn chế, đặc biệt là các nghiên cứu tích hợp giữa tin sinh học và thực nghiệm, đề tài góp phần bổ sung dữ liệu ban đầu, tạo tiền đề cho các nghiên cứu sâu hơn trong tương lai.

5. Những đóng góp của luận văn

Thứ nhất, nghiên cứu góp phần làm rõ vai trò tiềm năng và đặc điểm biểu hiện của lncRNA *CRNDE* trong UTĐTT.

Thứ hai, luận văn bước đầu xác nhận sự thay đổi biểu hiện của lncRNA *CRNDE* trên mẫu máu ngoại vi của bệnh nhân UTĐTT so với các nhóm đối chứng bằng phương pháp qPCR.

Thứ ba, nghiên cứu góp phần dự đoán các cơ chế phân tử và con đường sinh học liên quan đến lncRNA *CRNDE* thông qua các phân tích tăng cường chức năng và mạng lưới tương tác, từ đó bổ sung hiểu biết về vai trò của lncRNA này trong quá trình sinh ung.

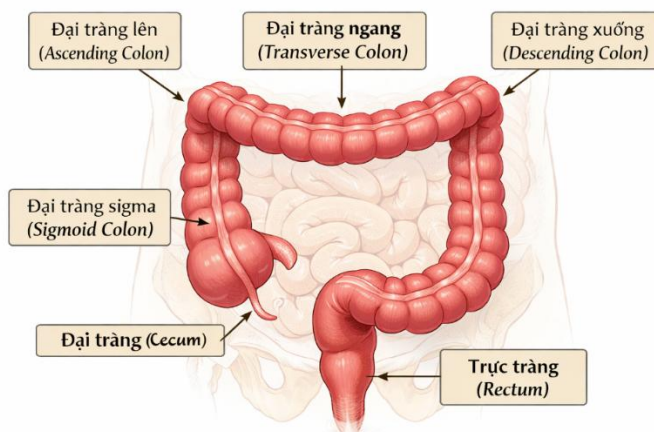
Thứ tư, kết quả của luận văn cung cấp dữ liệu tham khảo có giá trị trong bối cảnh nghiên cứu về lncRNA tại Việt Nam còn hạn chế, đồng thời tạo cơ sở cho các nghiên cứu tiếp theo nhằm xác nhận và mở rộng ứng dụng của lnc *CRNDE* trong phân biệt UTĐTT.

CHƯƠNG 1. TỔNG QUAN NGHIÊN CỨU

1.1. TỔNG QUAN VỀ UNG THƯ ĐẠI TRỰC TRÀNG

1.1.1 Dịch tễ học UTĐTT

Đại trực tràng (ruột già) là đoạn cuối của hệ tiêu hóa, trong đó khoảng 1.8 mét đầu tiên của ruột già được gọi là đại tràng và khoảng 15 cm cuối cùng là trực tràng và hậu môn [1] (Hình 1.1). Chúng có chức năng chính là hấp thu nước, điện giải và hình thành, lưu trữ phân trước khi đào thải ra ngoài cơ thể. Lớp biểu mô niêm mạc dễ xảy ra các biến đổi bất thường dẫn đến hình thành các tổn thương tiền ung thư như polyp và tiến triển thành ung thư.

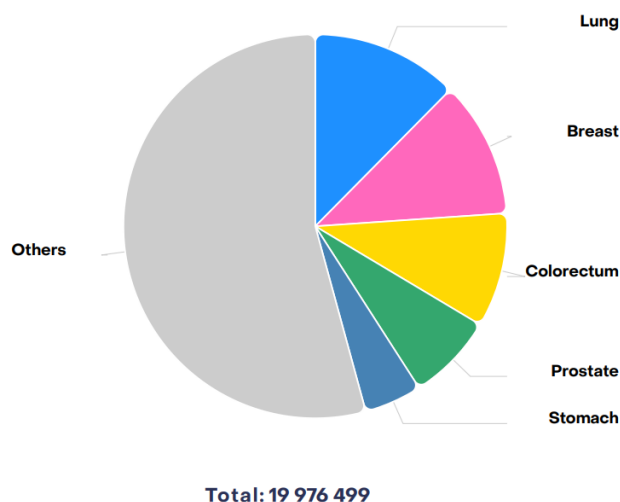


Hình 1.1 Hình ảnh đại trực tràng

Ung thư đại trực tràng là bệnh lý ác tính phát sinh khi các tế bào bất thường tăng sinh không kiểm soát trong mô của đại tràng hoặc trực tràng. Bệnh có khả năng lan rộng sang các mô xung quanh hoặc di chuyển đến các cơ quan xa thông qua hệ bạch huyết và đường máu. Quá trình các tế bào ung thư rời khỏi vị trí ban đầu và hình thành khối u ở cơ quan khác được gọi là di căn. Khi đã xuất hiện di căn, việc điều trị trở nên khó khăn hơn và tiên lượng sống của người bệnh cũng giảm đáng kể [2].

Theo thống kê của GLOBOCAN (2022) trên toàn cầu ghi nhận khoảng 1,9 triệu ca mắc mới UTĐTT và hơn 900.000 trường hợp tử vong, đưa bệnh trở thành nguyên nhân gây tử vong do ung thư đứng hàng thứ hai [3] (Hình 1.2). Các số liệu cập nhật gần đây cũng cho thấy UTĐTT chiếm khoảng 10% tổng số ca ung thư toàn cầu và là gánh nặng y tế đáng kể tại nhiều quốc gia [4]. Tại châu Á, UTĐTT đứng thứ ba về số ca mắc mới (966.399 ca năm 2022) và là nguyên nhân tử vong do ung thư đứng thứ

tur (462.252 ca). Nhật Bản có tỷ lệ mắc và tử vong cao nhất trong khu vực, thậm chí vượt một số nước phương Tây. Ở Philippines, bệnh cũng xếp thứ ba về số ca mắc mới và thứ tư về số ca tử vong trong năm 2022 [4].



Hình 1.2 Tổng số ca mắc ung thư trên toàn thế giới theo thống kê của GLOBOCAN năm 2022

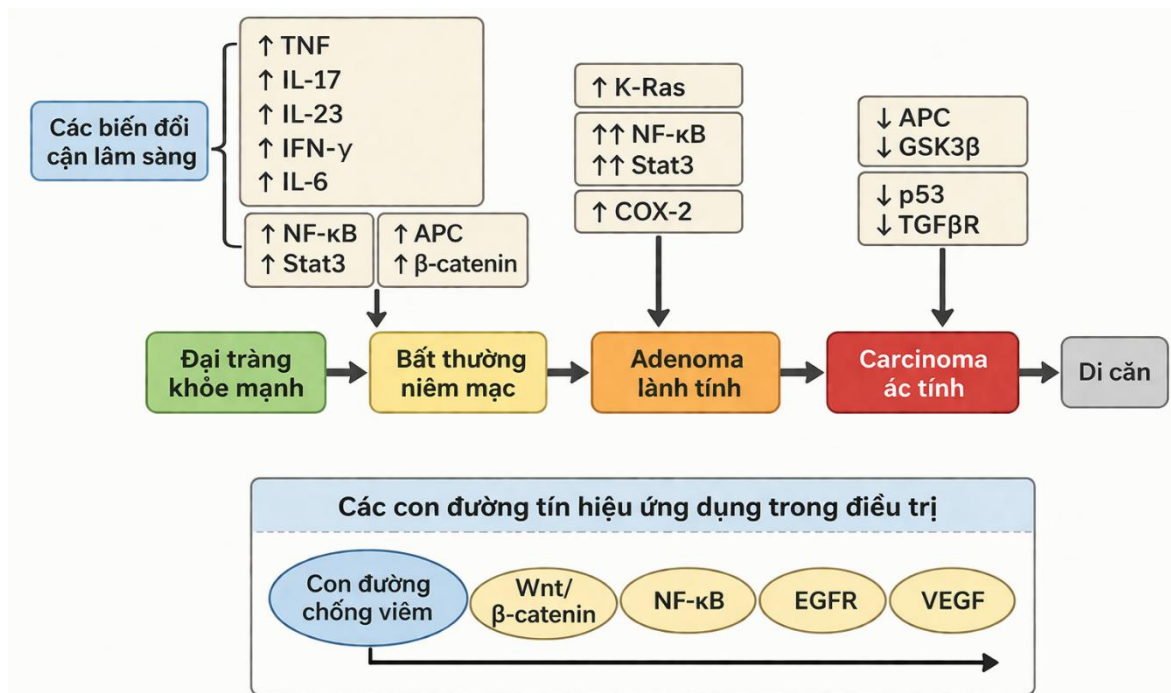
Tỷ lệ mắc UTĐTT có sự khác biệt theo tuổi và giới. Bệnh chủ yếu gặp ở người trên 50 tuổi, tuy nhiên trong những năm gần đây đã ghi nhận xu hướng trẻ hóa ở một số khu vực. Nam giới có nguy cơ mắc bệnh cao hơn nữ giới. Ngoài ra, các yếu tố nguy cơ như chế độ ăn nhiều thịt đỏ, ít chất xơ, béo phì, ít vận động, hút thuốc lá và tiền sử gia đình có vai trò quan trọng trong sự gia tăng tỷ lệ mắc bệnh [4].

Tại Việt Nam, UTĐTT cũng đang trở thành một vấn đề y tế đáng quan tâm. Theo GLOBOCAN 2022, UTĐTT đứng thứ tư về số ca mắc mới với hơn 9.400 trường hợp và có đến 8.454 trường hợp ghi nhận tử vong [5]. Sự gia tăng này được cho là liên quan đến quá trình đô thị hóa, thay đổi lối sống và chế độ ăn sử dụng thức ăn nhanh và đồ chế biến sẵn. Bên cạnh đó, việc tầm soát và phát hiện sớm chưa được triển khai rộng rãi cũng góp phần làm tăng tỷ lệ bệnh nhân được chẩn đoán ở giai đoạn muộn [4].

1.1.2 Cơ chế bệnh sinh của UTĐTT

Ung thư đại trực tràng là kết quả của quá trình tích lũy tuần tự các biến đổi di truyền và biểu sinh, dẫn đến sự chuyển đổi từ niêm mạc bình thường sang tổn thương tiền ung thư và cuối cùng là ung thư xâm lấn (Hình 1.3). Các biến đổi này thường liên quan đến các gen quan trọng như *APC*, *KRAS*, *NRAS*, *BRAF* và *TP53*, làm hoạt

hóa hoặc rối loạn các con đường tín hiệu nội bào như Wnt, EGFR và PI3K/AKT [6]. Quá trình này diễn ra theo nhiều bước, thường được mô tả qua mô hình “adenoma–carcinoma sequence” (chuỗi u tuyến – ung thư biểu mô) [7].



Hình 1.3 Quá trình bệnh sinh và các biến đổi phân tử của UTĐTT [8]

Về cơ chế phân tử, UTĐTT được hình thành chủ yếu thông qua ba con đường chính, bao gồm: bất ổn nhiễm sắc thể (Chromosomal Instability – CIN), bất ổn vi vệ tinh (Microsatellite Instability – MSI) và hiện tượng methyl hóa đảo CpG (CpG Island Methylator Phenotype – CIMP) [6, 9]. Trong đó, con đường CIN là phổ biến nhất, đặc trưng bởi sự tích lũy các đột biến gen và bất thường nhiễm sắc thể trong quá trình tiến triển của khối u [9].

Mô hình kinh điển về sinh UTĐTT được đề xuất bởi Vogelstein cho rằng ung thư phát triển thông qua sự tích lũy tuần tự các đột biến gen. Đột biến sớm thường xảy ra ở gen *APC*, dẫn đến hoạt hóa bất thường con đường tín hiệu Wnt/ β -catenin, làm tăng sinh tế bào biểu mô. Các đột biến tiếp theo ở các gen như *KRAS*, *SMAD4* và đặc biệt là *TP53* thúc đẩy quá trình tiến triển từ u tuyến đến ung thư xâm lấn [10].

Một số nghiên cứu cho thấy tần suất đột biến của các gen này khá cao ở bệnh nhân UTĐTT di căn. Cụ thể, đột biến *TP53* được ghi nhận ở 84,6% trường hợp, *APC* ở 78,0%, *KRAS* ở 49,6% và *SMAD4* ở 22,8% trong nhóm bệnh nhân không có MSI cao [11]. Ngoài ra, sự kết hợp đồng thời các đột biến *KRAS*, *APC* và *TP53* có liên quan đến mức độ tiến triển ác tính cao hơn của khối u [12].

Bên cạnh đó, các cơ chế sinh bệnh khác cũng góp phần quan trọng vào sự hình thành và tiến triển UTĐTT. Trong đó, rối loạn quá trình O-glycosyl hóa kiểu mucin có thể ảnh hưởng đến tăng sinh, di cư và xâm lấn của tế bào ung thư, đồng thời điều hòa đáp ứng miễn dịch của cơ thể [13]. Ngoài ra, sự mất cân bằng hệ vi sinh đường ruột (dysbiosis) cũng được chứng minh có liên quan đến sinh ung, thông qua cơ chế viêm mạn tính, điều hòa miễn dịch và gây tổn thương DNA. Một số vi khuẩn như *Fusobacterium nucleatum*, *Bacteroides fragilis* và *Escherichia coli* có vai trò thúc đẩy quá trình này [14, 15].

1.1.3. Phương pháp chẩn đoán UTĐTT

Chẩn đoán UTĐTT hiện nay, nội soi đại tràng vẫn được xem là tiêu chuẩn vàng nhờ độ nhạy và độ đặc hiệu cao. Các xét nghiệm tìm máu ẩn trong phân cũng được sử dụng rộng rãi trong sàng lọc, tuy nhiên còn hạn chế về độ nhạy và độ đặc hiệu [16-18].

Trong những năm gần đây, nhiều phương pháp chẩn đoán tiên tiến đã được phát triển, đặc biệt là nội soi có hỗ trợ trí tuệ nhân tạo (AI). Kỹ thuật này cho thấy khả năng cải thiện đáng kể tỷ lệ phát hiện polyp tuyến, giảm sự khác biệt giữa các người quan sát và hỗ trợ phân loại tổn thương theo thời gian thực. So với các phương pháp sinh thiết lỏng (liquid biopsy), nội soi có hỗ trợ AI được xem là giải pháp hiệu quả về chi phí, đặc biệt tại các cơ sở y tế có số lượng bệnh nhân lớn [17]. Trong khi đó, sinh thiết lỏng có ưu điểm không xâm lấn và cung cấp thông tin phân tử, nhưng chi phí còn cao [17].

Các dấu ấn sinh học biểu sinh, đặc biệt là các marker methyl hóa DNA, đóng vai trò quan trọng trong chẩn đoán sớm UTĐTT. Một số gen như *SDC2*, *SEPT9* và *SHOX2* cho thấy độ chính xác chẩn đoán cao. Các marker methyl hóa trong phân như *SDC2* và *SHOX2* đạt độ nhạy lên đến 91,35%, trong khi các marker khác như *NDRG4*, *SEPT9* và *BCAT1* có độ nhạy tổng hợp khoảng 82,7% [19-21]. Đối với các marker methyl hóa trong huyết tương như *SEPT9* và *HAND1*, độ nhạy dao động từ 40% đến 75,8% và độ đặc hiệu từ 69% đến 94,7%, tuy nhiên vẫn cần được tối ưu thêm để đáp ứng các tiêu chuẩn lâm sàng [22].

Các phương pháp chẩn đoán hình ảnh như chụp cắt lớp vi tính đại tràng và cộng hưởng từ (MRI) có vai trò quan trọng trong đánh giá giai đoạn bệnh, hỗ trợ lập kế hoạch điều trị và theo dõi bệnh nhân [16].

Gần đây, vai trò của hệ vi sinh vật cũng được quan tâm trong chẩn đoán UTĐTT. Các mô hình chẩn đoán dựa trên hệ vi sinh vật như vi khuẩn *Fusobacterium*, *Porphyromonas* và *Parvimonas* cho thấy độ chính xác cao với AUC từ 0,83 đến 0,94, đặc biệt khi kết hợp mẫu nước bọt và phân [18]. Các xét nghiệm dựa trên DNA và RNA trong phân cũng đang được nghiên cứu nhằm hỗ trợ phát hiện sớm và cải thiện tiên lượng bệnh [23].

Hiện nay, nội soi kết hợp mô bệnh học vẫn là tiêu chuẩn vàng trong chẩn đoán UTĐTT. Tuy nhiên, do tính xâm lấn của các phương pháp này, việc nghiên cứu các dấu ấn sinh học mới có thể phát hiện trong máu ngoại vi đang được quan tâm nhằm hỗ trợ chẩn đoán sớm và cải thiện tiên lượng bệnh.

1.2. RNA DÀI KHÔNG MÃ HÓA

1.2.1. Khái niệm và phân loại lncRNA

Khái niệm lncRNA

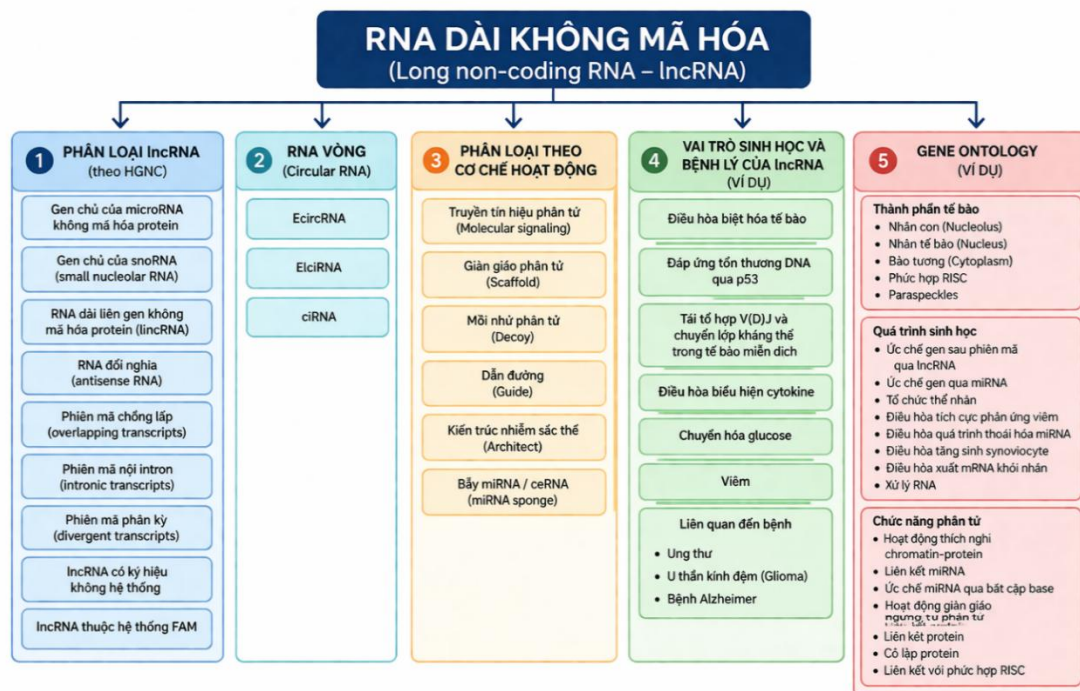
RNA không mã hóa dài là một nhóm lớn và đa dạng của các RNA không tham gia mã hóa protein, nhưng có vai trò quan trọng trong điều hòa nhiều quá trình sinh học ở cả động vật và thực vật [24-26]. Mặc dù phần lớn hệ gen người được phiên mã, chỉ một tỷ lệ nhỏ các bản sao RNA được dịch mã thành protein, điều này cho thấy vai trò thiết yếu của các lncRNA trong điều hòa hoạt động của tế bào [27].

Các lncRNA thường có tính đặc hiệu theo từng loại tế bào và tham gia điều hòa nhiều quá trình sinh lý quan trọng như biệt hóa tế bào, phát triển cơ thể, cũng như các hoạt động nội bào như dịch mã, chuyển hóa và truyền tín hiệu [25, 27]. Về cơ chế tác động, lncRNA có thể tham gia điều hòa biểu hiện gen dưới nhiều hình thức như tín hiệu điều hòa, phân tử đánh lạc hướng, khung liên kết hoặc hấp phụ RNA vi mô. Nhiều lncRNA còn hoạt động như các RNA nội sinh cạnh tranh, qua đó điều hòa biểu hiện gen ở mức sau phiên mã. [26].

Phân loại lncRNA

Theo phân loại của Ủy ban Danh pháp Gen Người (HUGO Gene Nomenclature Committee – HGNC), lncRNA được chia thành chín nhóm chính dựa trên vị trí và

đặc điểm cấu trúc, bao gồm: gen chủ của miRNA không mã hóa, gen chủ của RNA nhân nhỏ không mã hóa, RNA không mã hóa dài nằm giữa các gen (long intergenic non-protein coding RNA – LINC), RNA đối nghĩa (antisense RNA), các bản sao chồng lấp, bản sao nội intron, bản sao phân kỳ, các lncRNA có ký hiệu không hệ thống và các lncRNA thuộc họ FAM [26] (Hình 1.4).



Hình 1.4 Tổng quan về phân loại và chức năng của lncRNA

Ngoài ra, RNA vòng (circular RNA – circRNA) được xem là một nhóm riêng biệt nhưng có nhiều đặc điểm tương đồng với lncRNA. Các circRNA được phân loại thành ba nhóm chính gồm: circRNA có nguồn gốc exon, intron và dạng kết hợp exon–intron [26].

Các nghiên cứu về hệ gen và biểu sinh cho thấy phần lớn các lncRNA nằm giữa các gen (intergenic lncRNA) có nguồn gốc từ các vùng tăng cường (enhancer) hơn là từ vùng khởi đầu phiên mã (promoter), điều này gợi ý vai trò quan trọng của chúng trong điều hòa biểu hiện gen [28].

1.2.2. Cơ chế điều hòa biểu hiện gen của lncRNA

Các lncRNA tham gia điều hòa biểu hiện gen thông qua nhiều cơ chế phức tạp ở các mức độ khác nhau, bao gồm biểu sinh, phiên mã và sau phiên mã [29-31]. Nhờ khả năng tương tác với DNA, RNA và protein, lncRNA đóng vai trò như những phân

tử điều hòa trung tâm trong nhiều quá trình sinh học và bệnh lý, đặc biệt là trong ung thư

Ở mức độ biểu sinh, lncRNA tương tác với các phức hợp điều hòa như PRC2 để kiểm soát methyl hóa DNA và biến đổi histone, từ đó hoạt hóa hoặc ức chế phiên mã gen đích [30, 31]. Chẳng hạn, lncRNA *CRNDE* liên kết với EZH2 thuộc PRC2, thúc đẩy methyl hóa H3K27 và ức chế các gen ức chế khối u như *DUSP5* và *CDKN1A*, góp phần thúc đẩy tiến triển UTĐTT [32].

Ở mức độ phiên mã, lncRNA có thể điều hòa biểu hiện gen bằng nhiều cơ chế như hoạt động như mồi nhử, giá đỡ hoặc dẫn đường cho các yếu tố phiên mã và phức hợp protein [29]. Cụ thể, lncRNA có thể điều hòa phiên mã thông qua tương tác với các yếu tố phiên mã hoặc đóng vai trò giá đỡ cho các phức hợp protein điều hòa gen. Đáng chú ý, lncRNA *CRNDE* có khả năng hoạt hóa con đường tín hiệu Ras/MAPK, góp phần thúc đẩy tăng sinh và sống sót của tế bào. [33]. Bên cạnh đó, lncRNA *CRNDE* còn tham gia điều hòa con đường Wnt/ β -catenin thông qua tương tác với các miRNA, từ đó thúc đẩy tăng sinh và tiến triển tế bào ung thư [32, 34].

Ở mức độ sau phiên mã, lncRNA điều hòa độ ổn định, phân hủy và dịch mã của mRNA, đồng thời ảnh hưởng đến hoạt động của protein và các con đường tín hiệu nội bào [29, 30]. Một cơ chế quan trọng là hoạt động như các RNA hấp phụ miRNA, làm giảm tác dụng ức chế của miRNA lên mRNA đích. Ví dụ, *CRNDE* liên kết với miR-136, làm tăng biểu hiện E2F1, từ đó thúc đẩy di căn và kháng hóa trị trong ung thư đại trực tràng [35]. Ngoài ra, giảm biểu hiện lncRNA *CRNDE* có thể làm tăng độ nhạy của tế bào ung thư với xạ trị thông qua điều hòa miR-384 [36].

1.2.3. Vai trò của lncRNA trong ung thư

Các lncRNA đóng vai trò quan trọng và đa dạng trong quá trình hình thành và tiến triển ung thư, thông qua việc tham gia điều hòa nhiều quá trình sinh học thiết yếu của tế bào.

Điều hòa biểu hiện gen và hành vi tế bào ung thư

Các lncRNA có thể hoạt động như gen sinh ung hoặc gen ức chế khối u tùy thuộc vào bối cảnh sinh học [30]. Thông qua khả năng tương tác với DNA, RNA và protein, lncRNA điều hòa biểu hiện gen ở nhiều mức độ, từ biến đổi chromatin đến điều hòa phiên mã và sau phiên mã. Những cơ chế này ảnh hưởng trực tiếp đến các quá trình tăng sinh tế bào, chết tế bào theo chương trình, xâm lấn và di căn [30, 37].

Một trong những cơ chế quan trọng của lncRNA là tham gia mạng lưới RNA nội sinh cạnh tranh (ceRNA). Trong cơ chế này, lncRNA liên kết với miRNA, làm giảm khả năng gắn kết của miRNA với mRNA đích, từ đó làm tăng biểu hiện của các gen mục tiêu [38, 39]. Thông qua cơ chế ceRNA, lncRNA có thể điều hòa nhiều quá trình sinh học quan trọng như tăng sinh tế bào, apoptosis, xâm lấn và di căn.

Cơ chế tương tác giữa lncRNA và miRNA đã được chứng minh có vai trò trong điều hòa các con đường tín hiệu quan trọng như PI3K/AKT, Wnt/ β -catenin và MAPK [38, 40, 41]. Các con đường tín hiệu này đóng vai trò quan trọng trong điều hòa tăng sinh, sống sót và tiến triển của tế bào ung thư, do đó sự rối loạn mạng lưới ceRNA có thể thúc đẩy quá trình sinh ung. Chẳng hạn, lncRNA *LINC01503* có khả năng hoạt hóa các con đường ERK/MAPK, PI3K/AKT và Wnt/ β -catenin, góp phần thúc đẩy sự phát triển khối [41]. Tương tự, lncRNA *H19* điều hòa con đường PI3K/AKT trong ung thư đại trực tràng thông qua cơ chế ceRNA và có liên quan đến tiên lượng xấu của bệnh nhân [40].

Ảnh hưởng đến vi môi trường khối u và kháng điều trị

Bên cạnh vai trò nội tại trong tế bào ung thư, lncRNA còn tham gia điều hòa vi môi trường khối u thông qua ảnh hưởng đến đáp ứng miễn dịch, sự xâm nhập của tế bào miễn dịch và sản xuất cytokine, từ đó góp phần thúc đẩy sự phát triển của khối u [42].

Ngoài ra, sự biểu hiện bất thường của lncRNA có thể làm thay đổi hoạt động của các con đường tín hiệu quan trọng, dẫn đến giảm hiệu quả điều trị và thúc đẩy hiện tượng kháng thuốc, đặc biệt đối với các liệu pháp nhắm trúng đích[43].

Tiềm năng ứng dụng trong lâm sàng dấu ấn sinh học và mục tiêu điều trị

Nhờ đặc điểm biểu hiện đặc hiệu theo mô và trạng thái bệnh lý, lncRNA được xem là các dấu ấn sinh học tiềm năng trong ung thư. Nhiều lncRNA biểu hiện khác biệt giữa mô ung thư và mô bình thường, đồng thời liên quan đến giai đoạn bệnh, mức độ xâm lấn và tiên lượng sống của bệnh nhân [30, 37]. Đặc biệt, lncRNA có thể được phát hiện trong các dịch sinh học như huyết tương và huyết thanh với độ ổn định cao, tạo điều kiện cho phát triển các phương pháp chẩn đoán không xâm lấn nhằm hỗ trợ phát hiện sớm, theo dõi tiến triển bệnh và đánh giá đáp ứng điều trị[37].

Bên cạnh vai trò là dấu ấn sinh học, lncRNA còn là mục tiêu tiềm năng trong điều trị ung thư. Các chiến lược đang được nghiên cứu bao gồm ức chế lncRNA sinh

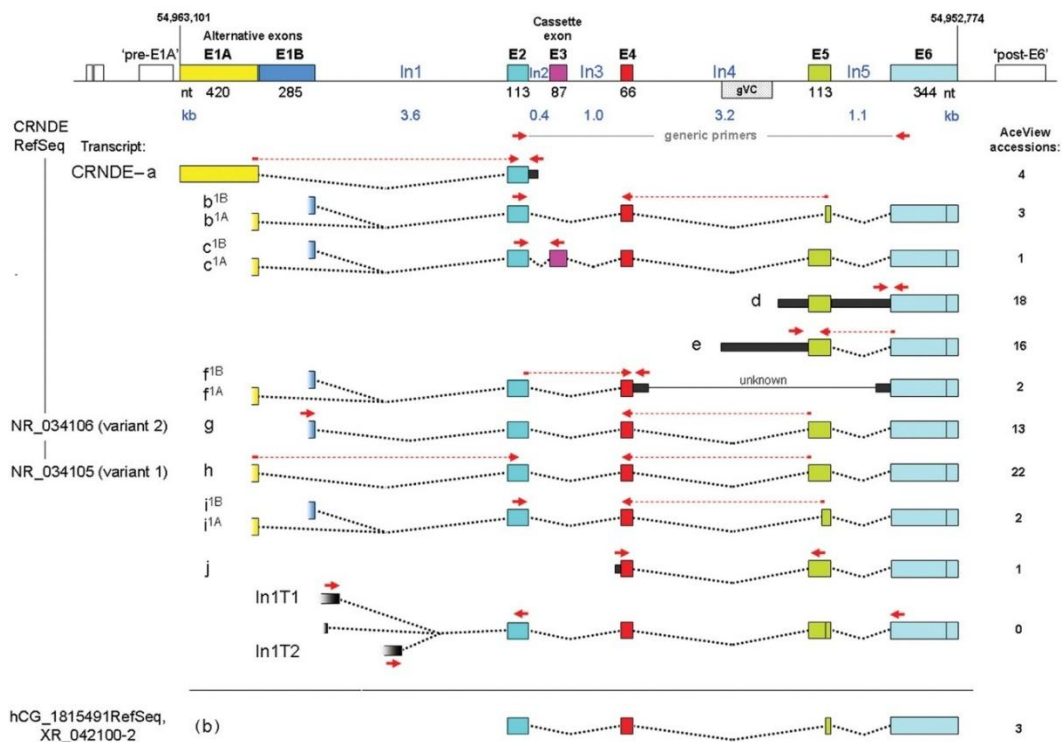
ung hoặc phục hồi chức năng của lncRNA ức chế khối u bằng các công nghệ như antisense oligonucleotide, CRISPR/Cas và các phân tử nhỏ [44]. Ngoài ra, việc can thiệp vào các tương tác giữa lncRNA với miRNA hoặc protein cũng được xem là hướng tiếp cận triển vọng nhằm điều hòa các con đường tín hiệu liên quan đến sinh ung.

Những tiến bộ này cho thấy lncRNA có tiềm năng ứng dụng trong cả chẩn đoán, tiên lượng và điều trị ung thư, góp phần thúc đẩy y học cá thể hóa.

1.3. LNCRNA *CRNDE*

1.3.1. Đặc điểm cấu trúc và biểu hiện của lncRNA *CRNDE*

LncRNA *CRNDE* là một trong những lncRNA được nghiên cứu nhiều nhất trong ung thư đại trực tràng do biểu hiện tăng cao trong nhiều loại ung thư và liên quan mật thiết đến quá trình sinh ung, tiến triển khối u cũng như tiên lượng bệnh.



Hình 1.5 Vị trí gen của *CRNDE* [45]

LncRNA *CRNDE* là một gen mã hóa RNA dài không mã hóa nằm trên nhiễm sắc thể 16q12.2 của người (Hình 1.5). Gen này được phát hiện lần đầu trong mô ung thư đại trực tràng với mức biểu hiện tăng cao rõ rệt so với mô bình thường và được đặt tên dựa trên đặc điểm này [45].

Về cấu trúc, lncRNA *CRNDE* có cấu trúc phức tạp, gồm nhiều exon và tạo ra nhiều biến thể phiên mã thông qua cơ chế cắt nối thay thế. Các isoform như *CRNDE-*

a, *CRNDE-b* và *CRNDE-h* có sự khác biệt về cấu trúc, vị trí phân bố và chức năng sinh học; trong đó *CRNDE-h* được nghiên cứu nhiều do có độ ổn định cao trong huyết thanh và exosome [45, 46].

Tương tự các RNA được phiên mã bởi RNA polymerase II, lncRNA *CRNDE* mang mũ 5', đuôi poly(A) và trải qua quá trình cắt nối exon. Tuy nhiên, lncRNA *CRNDE* không mã hóa protein mà thực hiện chức năng thông qua tương tác với DNA, RNA và protein [47]. Phân tử này có thể hiện diện ở cả trong nhân và bào tương, cho phép tham gia điều hòa biểu hiện gen ở nhiều mức độ, từ điều hòa chromatin và phiên mã đến các cơ chế hậu phiên mã thông qua tương tác với miRNA [46, 47].

Ngoài ra, lncRNA *CRNDE* còn tương tác với các protein điều hòa biểu sinh và tham gia vào các mạng lưới điều hòa RNA, qua đó ảnh hưởng đến nhiều con đường tín hiệu liên quan đến sinh ung và tiến triển khối u. Những đặc điểm này tạo cơ sở cho tiềm năng ứng dụng của lncRNA *CRNDE* như một dấu ấn sinh học trong ung thư đại trực tràng [46, 47].

1.3.2. Vai trò của lncRNA *CRNDE* trong ung thư đại trực tràng

LncRNA *CRNDE* là một trong những lncRNA được nghiên cứu rộng rãi trong UTĐTT do có mức biểu hiện tăng cao đáng kể trong mô ung thư so với mô bình thường. Nhiều nghiên cứu cho thấy lncRNA *CRNDE* đóng vai trò như một lncRNA sinh ung, tham gia điều hòa các quá trình liên quan đến sự hình thành và tiến triển khối u [33, 45].

Về cơ chế phân tử, lncRNA *CRNDE* có khả năng điều hòa các con đường tín hiệu quan trọng như PI3K/AKT, Wnt/ β -catenin và MAPK, qua đó ảnh hưởng đến tăng sinh, sống sót và di căn của tế bào ung thư [48]. Ngoài ra, lncRNA *CRNDE* còn hoạt động theo cơ chế ceRNA, đóng vai trò như một RNA hấp phụ miRNA, làm giảm tác dụng ức chế của miRNA lên các mRNA đích và gián tiếp làm tăng biểu hiện của các gen liên quan đến sinh ung [35].

Bên cạnh vai trò trong cơ chế bệnh sinh, lncRNA *CRNDE* còn có ý nghĩa lâm sàng quan trọng. Mức biểu hiện cao của lncRNA *CRNDE* được ghi nhận có liên quan đến kích thước khối u lớn, giai đoạn bệnh tiến triển, di căn xa và tiên lượng xấu ở bệnh nhân UTĐTT [31, 47]. Đặc biệt, lncRNA *CRNDE* có thể được phát hiện trong huyết tương hoặc huyết thanh, cho thấy tiềm năng ứng dụng như một dấu ấn sinh học không xâm lấn trong chẩn đoán và theo dõi bệnh [46].

CHƯƠNG 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. ĐỐI TƯỢNG NGHIÊN CỨU

2.1.1. Dữ liệu công khai (TCGA, GEO)

Trong nghiên cứu này, dữ liệu biểu hiện gen và thông tin lâm sàng của bệnh nhân ung thư đại trực tràng được thu thập từ các cơ sở dữ liệu công khai, bao gồm The Cancer Genome Atlas (TCGA) (<https://portal.gdc.cancer.gov/>) và Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>). Đây là những nguồn dữ liệu lớn, đáng tin cậy và được sử dụng rộng rãi trong các nghiên cứu tin sinh học nhằm phân tích biểu hiện gen và xác định các dấu ấn sinh học tiềm năng.

Từ TCGA, dữ liệu được khai thác từ hai bộ dữ liệu TCGA-COAD và TCGA-READ, bao gồm dữ liệu biểu hiện RNA (RNA-seq) và các thông tin lâm sàng như tuổi, giới tính, giai đoạn bệnh và tình trạng sống sót. Dữ liệu RNA-seq được chuẩn hóa dưới dạng TPM hoặc FPKM để đảm bảo khả năng so sánh giữa các mẫu.

Từ GEO, các bộ dữ liệu microarray hoặc RNA-seq liên quan đến ung thư đại trực tràng được lựa chọn dựa trên các tiêu chí: có nhóm ung thư và nhóm đối chứng, kích thước mẫu phù hợp và thông tin chú giải đầy đủ. Dữ liệu được tiền xử lý bao gồm chuẩn hóa, loại bỏ dữ liệu thiếu và kiểm tra chất lượng trước khi phân tích.

Việc kết hợp dữ liệu từ TCGA và GEO giúp tăng độ tin cậy thông qua xác nhận chéo giữa các bộ dữ liệu độc lập, đồng thời cho phép đánh giá toàn diện hơn biểu hiện của lncRNA CRNDE trong ung thư đại trực tràng, làm cơ sở cho các phân tích và xác thực thực nghiệm tiếp theo.

2.1.2. Mẫu máu ngoại vi

Mẫu máu ngoại vi được thu thập theo phương pháp thuận tiện, 35 đối tượng tham gia nghiên cứu và chia thành ba nhóm: nhóm ung thư đại trực tràng (n=15) được chẩn đoán xác định bằng mô bệnh học; nhóm polyp đại trực tràng (n=10) được xác định qua nội soi và mô bệnh học; và nhóm chứng khỏe mạnh (n=10) không có tiền sử hoặc biểu hiện lâm sàng liên quan đến bệnh lý đại trực tràng.

Tất cả các mẫu máu được thu thập, xử lý và bảo quản theo quy trình chuẩn nhằm đảm bảo chất lượng RNA cho các phân tích tiếp theo. Việc phân nhóm đối tượng cho phép đánh giá sự khác biệt biểu hiện của lncRNA CRNDE giữa các trạng thái bệnh lý khác nhau và xác định tiềm năng ứng dụng của dấu ấn này trong chẩn đoán ung thư đại trực tràng.

Tiêu chuẩn lựa chọn và loại trừ

Nhóm ung thư đại trực tràng bao gồm các bệnh nhân từ 18 tuổi trở lên, được chẩn đoán xác định bằng mô bệnh học, có đầy đủ thông tin lâm sàng và đồng ý tham gia nghiên cứu. Nhóm polyp đại trực tràng gồm các bệnh nhân từ 18 tuổi trở lên, được xác nhận polyp bằng nội soi và mô bệnh học, không có bằng chứng ung thư. Nhóm chứng khỏe mạnh gồm các đối tượng từ 18 tuổi trở lên, không có tiền sử ung thư hoặc bệnh lý đại trực tràng và đồng ý tham gia nghiên cứu.

Tiêu chuẩn loại trừ áp dụng cho tất cả các nhóm bao gồm: mắc các loại ung thư khác; mắc các bệnh lý viêm đại trực tràng mạn tính như viêm loét đại trực tràng hoặc bệnh Crohn; hoặc mắc các bệnh mạn tính nặng ở gan, thận, tim hoặc phổi.

2.2. Phạm vi và thiết kế nghiên cứu

Nghiên cứu được thiết kế theo dạng mô tả cắt ngang, kết hợp phân tích dữ liệu tin sinh học và xác nhận thực nghiệm.

2.3 Phương pháp phân tích dữ liệu công khai

2.3.1 Thu thập và tiền xử lý dữ liệu biểu hiện gen

Dữ liệu biểu hiện gen và thông tin lâm sàng của bệnh nhân ung thư đại trực tràng được thu thập từ các cơ sở dữ liệu công khai bao gồm TCGA và GEO.

Dữ liệu RNA-seq của ung thư đại tràng được tải từ dự án TCGA-COAD. Quá trình thu thập và xử lý dữ liệu được thực hiện trong môi trường R version 4.4.3.

Các gói R được sử dụng bao gồm:

- TCGAbiolinks: tải dữ liệu biểu hiện gen và dữ liệu lâm sàng từ TCGA [49]
- GEOquery: thu thập dữ liệu biểu hiện gen từ GEO [50]
- dplyr và tidyr: xử lý và tổ chức dữ liệu [51]
- readxl: đọc các tệp dữ liệu dạng bảng [52]

Dữ liệu biểu hiện gen sau khi thu thập được chuẩn hóa nhằm đảm bảo độ tin cậy cho các phân tích tiếp theo.

2.3.2 Phân tích sự khác biệt biểu hiện của *CRNDE* trong UTĐTT

Mức biểu hiện của lncRNA *CRNDE* được so sánh giữa mô ung thư và mô bình thường nhằm đánh giá sự thay đổi biểu hiện của gen này trong ung thư đại trực tràng.

Phân tích biểu hiện sai khác được thực hiện bằng gói edgeR, sử dụng mô hình phân bố âm nhị thức để xác định các gen có sự khác biệt biểu hiện [53].

Kết quả phân tích được trực quan hóa bằng các biểu đồ sử dụng gói ggplot2, bao gồm biểu đồ hộp (boxplot) nhằm minh họa sự khác biệt mức biểu hiện của lncRNA *CRNDE* giữa các nhóm mẫu [54].

2.3.3 Phân tích gen đồng biểu hiện

Để xác định các gen có mối liên quan với lncRNA *CRNDE*, phân tích tương quan biểu hiện gen được thực hiện giữa lncRNA *CRNDE* và toàn bộ các gen khác trong bộ dữ liệu TCGA. Hệ số tương quan Pearson được sử dụng để đánh giá mức độ tương quan giữa mức biểu hiện của lncRNA *CRNDE* và các gen khác [55].

Các gen có hệ số tương quan cao với lncRNA *CRNDE* được lựa chọn làm tập hợp gen đồng biểu hiện cho các phân tích tiếp theo.

Mạng lưới tương tác giữa lncRNA *CRNDE* và các gen đồng biểu hiện được xây dựng nhằm trực quan hóa mối quan hệ giữa các gen, trong đó lncRNA *CRNDE* là nút trung tâm và các gen tương quan được kết nối xung quanh. Việc xây dựng và trực quan hóa mạng lưới được thực hiện bằng gói igraph[56].

2.3.4 Phân tích tăng cường chức năng và con đường tín hiệu

Để tìm hiểu chức năng sinh học của các gen đồng biểu hiện với lncRNA *CRNDE*, phân tích tăng cường chức năng được thực hiện dựa trên cơ sở dữ liệu Gene Ontology (GO) và Kyoto Encyclopedia of Genes and Genomes (KEGG).

Phân tích được thực hiện bằng gói clusterProfiler, cho phép xác định các quá trình sinh học và con đường tín hiệu có sự tham gia đáng kể của các gen nghiên cứu [57].

Kết quả phân tích được trực quan hóa bằng các biểu đồ thanh và biểu đồ bong bóng sử dụng các hàm của clusterProfiler kết hợp với ggplot2 [54, 57].

2.3.5. Phân tích sống sót

Phân tích sống sót được thực hiện nhằm đánh giá giá trị tiên lượng của lncRNA *CRNDE* đối với bệnh nhân ung thư đại trực tràng. Phân tích này được tiến hành trên dữ liệu của dự án TCGA-COAD, do bộ dữ liệu này cung cấp thông tin về thời gian sống và tình trạng sống của bệnh nhân.

Bệnh nhân được chia thành hai nhóm dựa trên mức biểu hiện của lncRNA *CRNDE* (nhóm biểu hiện cao và nhóm biểu hiện thấp) dựa trên giá trị trung vị của mức biểu hiện.

Phân tích sống sót Kaplan–Meier được thực hiện bằng các gói survival và survminer, cho phép xây dựng đường cong sống sót và kiểm định log-rank nhằm đánh giá sự khác biệt giữa các nhóm bệnh nhân [58].

2.4 Phương pháp thực nghiệm

2.4.1. Tách chiết RNA tổng số

Tổng RNA được chiết xuất từ 35 mẫu máu toàn phần, bao gồm các mẫu từ bệnh nhân ung thư đại trực tràng, bệnh nhân polyp và nhóm người khỏe mạnh. Quá trình tách RNA được thực hiện theo hướng dẫn của bộ kit TRIzol® (Thermo Fisher Science, USA). Các mẫu RNA hoàn tất được bảo quản ở -80°C . Nồng độ và độ tinh sạch của RNA được xác định bằng máy đo quang phổ NanoDrop Lite trước khi tiến hành các bước phân tích tiếp theo.

2.4.2 Tổng hợp cDNA

Tổng hợp cDNA được thực hiện dựa trên hướng dẫn của bộ kit Maxime™ RT PreMix (Intron, Hàn Quốc). Trước tiên, RNA khuôn (template RNA) và nước cất xử lý DEPC được thêm vào ống phản ứng chứa Maxime RT PreMix, sao cho tổng thể tích phản ứng đạt $20\ \mu\text{L}$. Thành phần chi tiết của phản ứng được trình bày trong Bảng 2.1. Sau khi thêm đầy đủ các thành phần, hỗn hợp được hòa tan bằng cách hút nhẹ bằng micropipette để đảm bảo đồng nhất.

Bảng 2.1 Thành Phần phản ứng tổng hợp cDNA

Thành phần phản ứng	Nồng độ
Template RNA	0.5 μg
Maxime RT PreMix	
Nước cất (DEPC-treated water)	Thêm đủ đến 20 μl
Tổng thể tích phản ứng	20 μl

Quá trình tổng hợp cDNA được thực hiện trên máy PCR theo chu trình nhiệt sau: tổng hợp cDNA ở 45°C trong 60 phút, sau đó vô hoạt enzyme RTase ở 95°C trong 5 phút và giữ lạnh ở 4°C cho đến khi tiến hành bước tiếp theo. Các chỉ số chất lượng cDNA, bao gồm nồng độ và độ tinh sạch, được đo bằng máy NanoDrop™ Lite (Thermo Fisher Scientific, USA). Cuối cùng, các mẫu cDNA sau khi tổng hợp được điều chỉnh về tổng thể tích $20\ \mu\text{L}$ với nồng độ $500\ \text{ng}/\mu\text{L}$, sẵn sàng cho các thí nghiệm tiếp theo theo yêu cầu nghiên cứu.

2.4.3 Đánh giá mức độ biểu hiện lncRNA *CRNDE* thông qua phản ứng real-time PCR

Phản ứng real-time PCR được thực hiện sử dụng bộ kit SYBR Green qRT-PCR Master Mix (Thermo Fisher Scientific, Hoa Kỳ) trên hệ thống QuantStudio 1 (Thermo, Hoa Kỳ). Mức độ biểu hiện của các gen mục tiêu được đánh giá bằng real-time PCR với các cặp mồi đặc hiệu: nhiệt độ gắn mồi 60°C đối với *CRNDE* và 55°C đối với *GAPDH*. Mỗi phản ứng có thể tích 20 μ L, bao gồm 2x SYBR Green PCR Master Mix, 0,4 μ M mồi xuôi, 0,4 μ M mồi ngược và 2 μ L cDNA khuôn. Điều kiện phản ứng PCR bắt đầu với bước biến tính ban đầu ở 95°C trong 60 giây, tiếp theo là 45 chu kỳ gồm 95°C trong 20 giây, 60°C hoặc 55°C (tùy gen) trong 20 giây và 72°C trong 20 giây. Sau khi kết thúc chu kỳ, phản ứng được thêm giai đoạn high resolution melting (HRM) với các bước: 95°C trong 5 giây, 65°C trong 60 giây và 97°C trong 15 giây. Mỗi mẫu được chạy hai lần lặp để đảm bảo độ chính xác.

Bảng 2.2 Trình tự mồi khuếch đại gen đích và gen nội chuẩn

Tên mồi	Trình tự (5' -> 3')	Kích thước (bp)	T _m (°C)
<i>CRNDE</i> - F	TCAGCCGTTGGTCTTTGAAAT	158	60
<i>CRNDE</i> - R	TTCTGCGTGACAACCTGAGGAT		
<i>GAPDH</i> -F	CCAAGGAGTAAGACCCCTGG	117	55
<i>GAPDH</i> -R	AGGGGAGATTCAGTGTGGTG		

Trình tự mồi trong Bảng 2.2 được thiết kế dựa trên trình tự tham chiếu của gen đích thu thập từ cơ sở dữ liệu GenBank (NCBI), bao gồm lncRNA *CRNDE* (NR_170995) và gen nội chuẩn *GAPDH* (NM_002046.7). Việc thiết kế được thực hiện bằng phần mềm Primer3, đảm bảo các tiêu chí về độ dài mồi, nhiệt độ nóng chảy (T_m), hàm lượng GC và tránh hình thành cấu trúc thứ cấp như hairpin hay primer-dimer. Độ đặc hiệu của mồi được kiểm tra bằng công cụ BLAST nhằm đảm bảo chỉ khuếch đại trình tự mục tiêu. Các cặp mồi được tổng hợp tại Công ty Cổ phần PHÙ SA Genomics (Việt Nam).

2.4.4 Phân tích dữ liệu real-time PCR

Mức độ biểu hiện của gen lncRNA *CRNDE* được đánh giá dựa trên giá trị Ct của mỗi mẫu, với *GAPDH* làm gen nội chuẩn. Các giá trị Ct trung bình của nhóm đối chứng (Healthy) được dùng để chuẩn hóa, và mức biểu hiện tương đối của từng mẫu được tính toán dựa trên phương pháp $2^{-\Delta\Delta Ct}$ [59].

2.5 Phương pháp xử lý và phân tích thống kê

Dữ liệu được xử lý và phân tích bằng R (version 4.4.3). Sự khác biệt biểu hiện lncRNA *CRNDE* giữa hai nhóm được đánh giá bằng kiểm định Student's t-test; giữa nhiều nhóm bằng ANOVA một yếu tố hoặc Kruskal–Wallis, tùy theo phân bố dữ liệu.

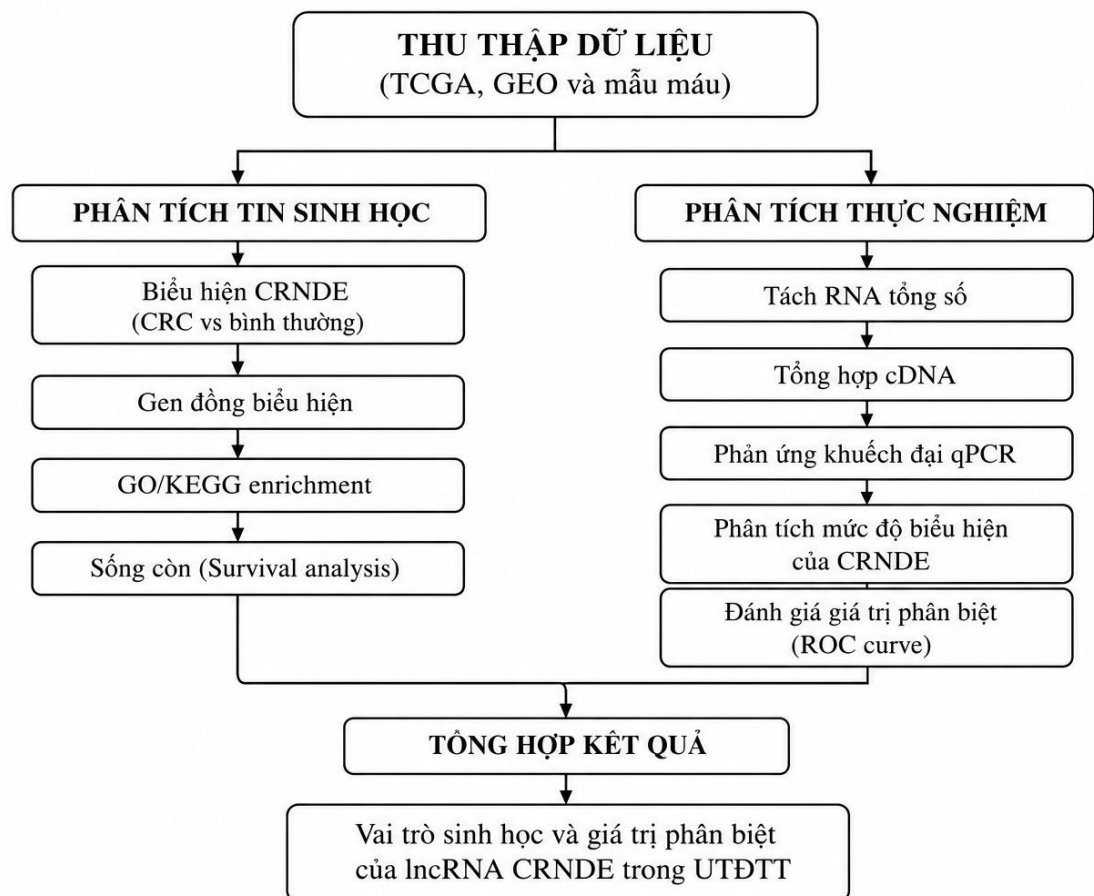
Mối liên quan giữa biểu hiện lncRNA *CRNDE* và các biến số định lượng được đánh giá bằng hệ số tương quan Pearson hoặc Spearman. Giá trị phân biệt của lncRNA *CRNDE* được xác định thông qua đường cong ROC, diện tích dưới đường cong (AUC) và ngưỡng cắt tối ưu theo chỉ số Youden bằng các gói phân tích ROC trong R.

Ngưỡng ý nghĩa thống kê được xác định là $p < 0,05$.

2.6 Đạo đức nghiên cứu

Nghiên cứu được thông qua Hội đồng đạo đức trong nghiên cứu y sinh của Bệnh viện Quân y 103, Học viện Quân y số phê duyệt 341/CNChT-HĐĐĐ, ngày 15 tháng 11 năm 2023.

SƠ ĐỒ NGHIÊN CỨU



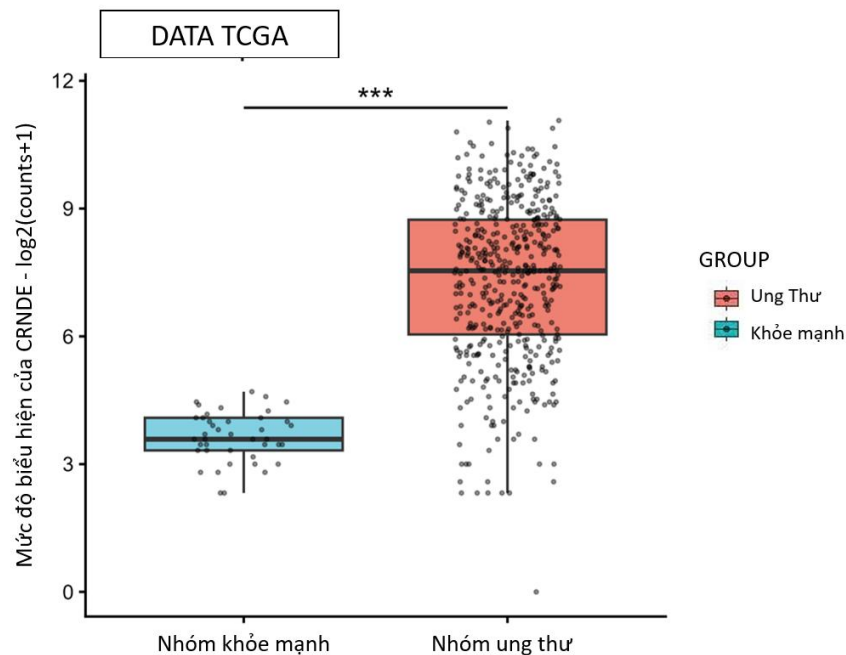
CHƯƠNG 3: KẾT QUẢ VÀ THẢO LUẬN

3.1. KẾT QUẢ PHÂN TÍCH DỮ LIỆU TIN SINH HỌC

3.1.1 Mức độ biểu hiện của lncRNA *CRNDE* trong UTĐTT

Kết quả phân tích data TCGA

Phân tích dữ liệu từ bộ TCGA_COAD bao gồm 483 mẫu ung thư đại trực tràng và 41 mẫu mô đại trực tràng khỏe mạnh cho thấy sự khác biệt rõ rệt về mức độ biểu hiện của lncRNA *CRNDE* giữa hai nhóm. Kết quả hình 3.1 cho thấy lncRNA *CRNDE* được biểu hiện tăng đáng kể trong nhóm ung thư so với nhóm khỏe mạnh ($p < 0,001$).

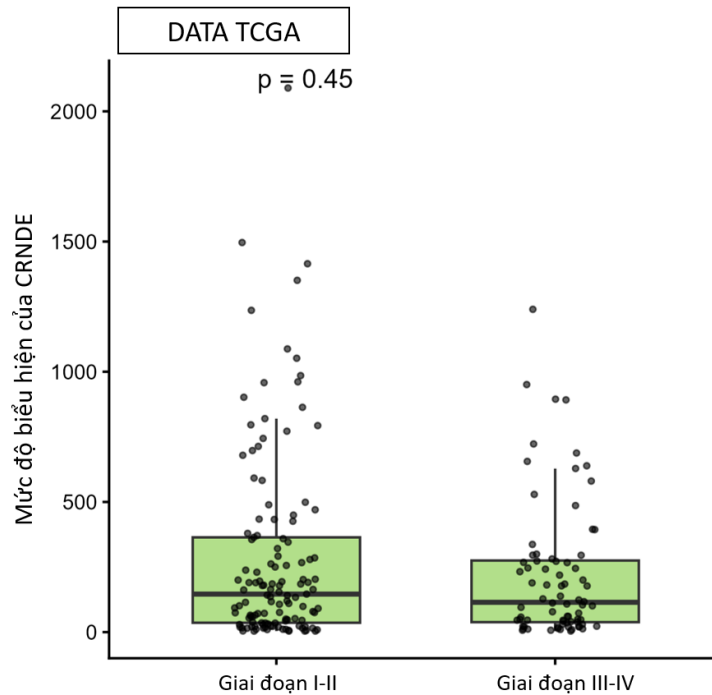


Hình 3.1 Mức độ biểu hiện của lncRNA *CRNDE* trong dữ liệu TCGA

(*** thể hiện $p < 0.001$, sử dụng mô hình phân bố âm nhị thức trong gói edgeR)

Cụ thể, giá trị biểu hiện lncRNA *CRNDE* ($\log_2(\text{counts}+1)$) ở nhóm ung thư có xu hướng cao hơn rõ rệt, với trung vị khoảng 7–8, trong khi nhóm khỏe mạnh chỉ dao động quanh mức 3–4. Ngoài ra, độ phân tán giá trị trong nhóm ung thư cũng lớn hơn, phản ánh tính không đồng nhất sinh học của khối u.

Đánh giá sự khác biệt về mức độ biểu hiện của lncRNA *CRNDE* giữa các giai đoạn bệnh cho thấy không có sự khác biệt có ý nghĩa thống kê giữa hai nhóm giai đoạn sớm (I–II) và giai đoạn muộn (III–IV) ($p = 0,45$) (Hình 3.2).



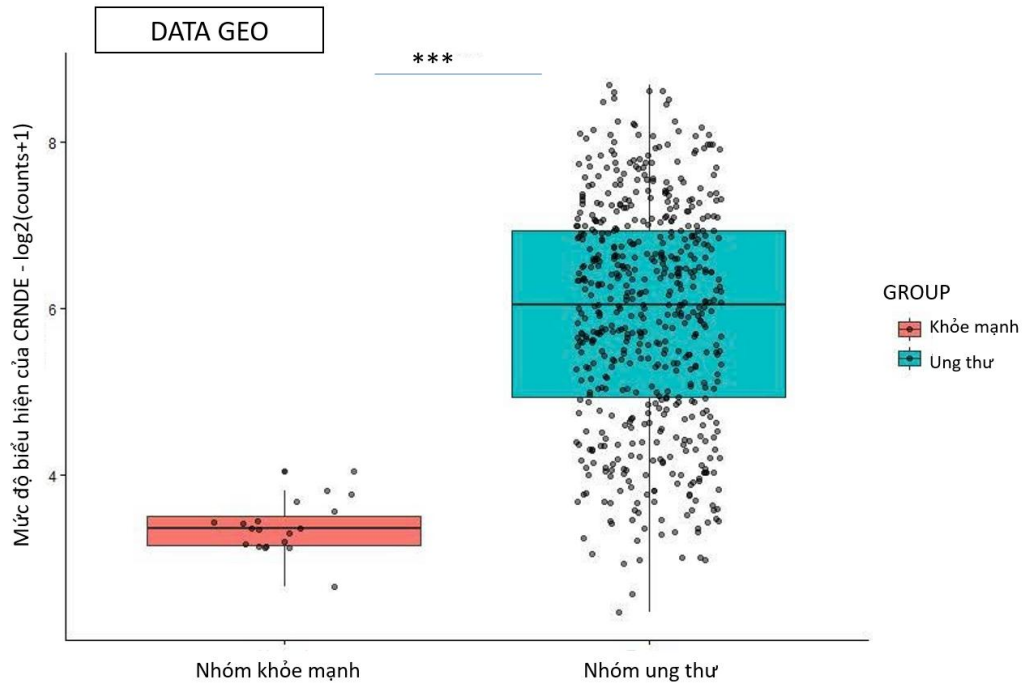
Hình 3.2 Mức độ biểu hiện của lncRNA CRNDE theo từng giai đoạn bệnh trong dữ liệu TCGA

(Mann-Whitney U test)

Mức độ biểu hiện lncRNA CRNDE ở cả hai nhóm giai đoạn có xu hướng tương đương nhau, với trung vị dao động ở mức tương tự. Mặc dù có sự biến thiên lớn về giá trị biểu hiện trong từng nhóm, đặc biệt ở giai đoạn I–II, nhưng sự chồng lấp đáng kể giữa hai nhóm cho thấy không có xu hướng tăng hoặc giảm rõ rệt theo tiến triển giai đoạn bệnh.

Kết quả phân tích data GEO

Phân tích dữ liệu từ tập GSE39582 (GEO), bao gồm 566 mẫu ung thư đại trực tràng và 19 mẫu mô niêm mạc đại trực tràng bình thường, cho thấy sự khác biệt rõ rệt về mức độ biểu hiện của lncRNA CRNDE giữa hai nhóm. Kết quả hình 3.3 cho thấy lncRNA CRNDE được biểu hiện tăng đáng kể trong nhóm ung thư so với nhóm khỏe mạnh ($p < 0,001$). Cụ thể, giá trị biểu hiện lncRNA CRNDE ($\log_2(\text{counts}+1)$) ở nhóm ung thư có trung vị khoảng 6, cao hơn rõ rệt so với nhóm bình thường (khoảng 3–3,5).

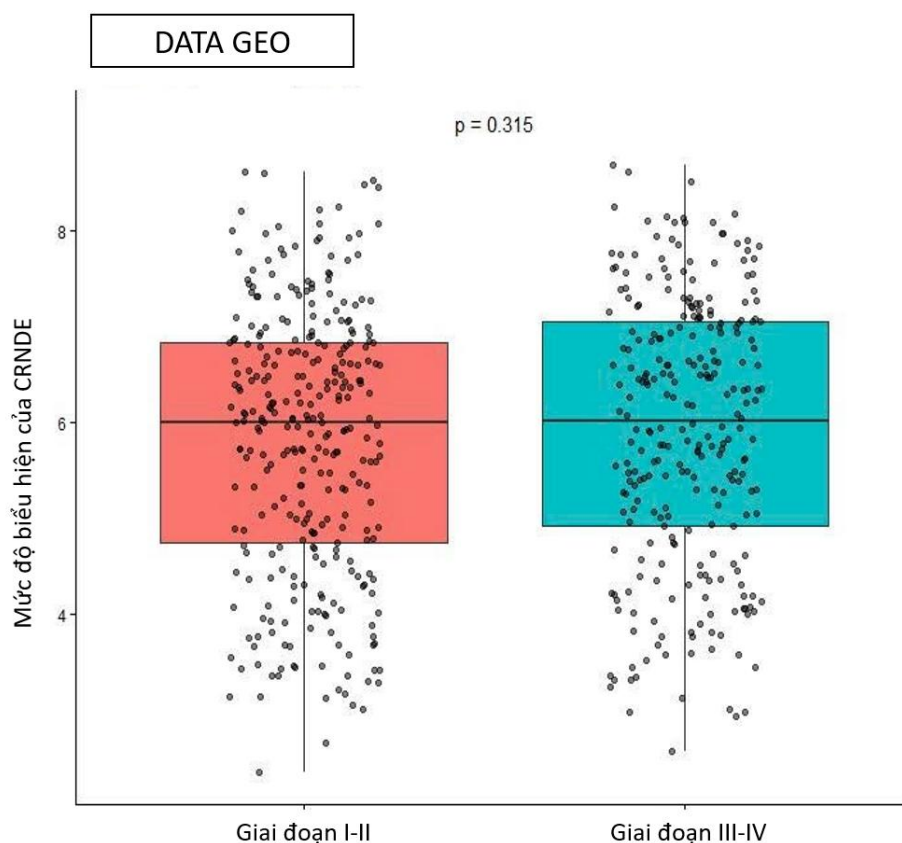


Hình 3.3 Mức độ biểu hiện của lncRNA CRNDE trong dữ liệu GEO

(*** thể hiện $p < 0.001$, sử dụng mô hình phân bố âm nhị thức trong gói edgeR)

Ngoài ra, tương tự như dữ liệu TCGA, mức độ phân tán biểu hiện lncRNA CRNDE trong nhóm ung thư lớn hơn đáng kể so với nhóm khỏe mạnh, phản ánh tính không đồng nhất của quần thể tế bào khối u.

Mối liên quan giữa mức độ biểu hiện của lncRNA CRNDE và giai đoạn bệnh cho thấy không có sự khác biệt có ý nghĩa thống kê giữa hai nhóm giai đoạn sớm (I–II) và giai đoạn muộn (III–IV) ($p = 0,315$) từ dữ liệu GSE39582 (Hình 3.4).



Hình 3.4 Mức độ biểu hiện của lncRNA CRNDE theo từng giai đoạn bệnh trong dữ liệu GEO

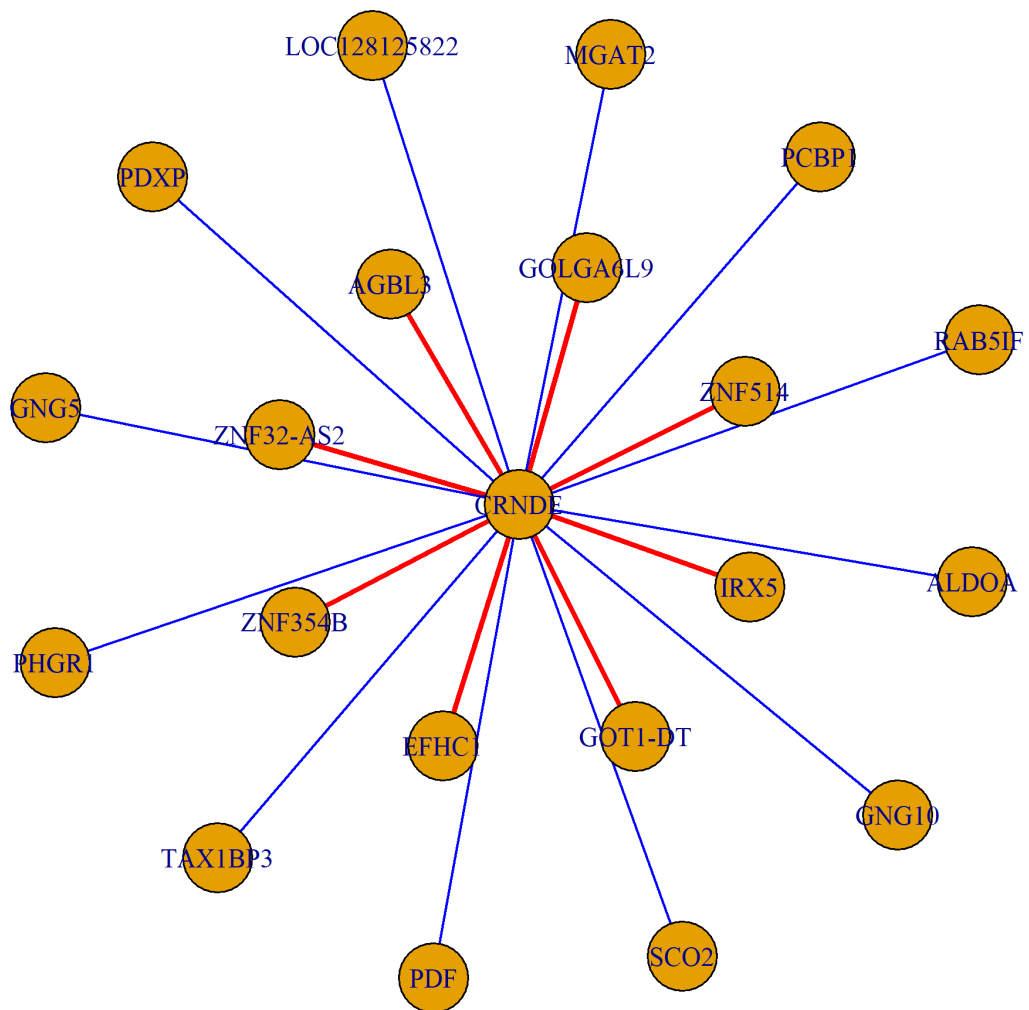
(Mann–Whitney U test)

Kết quả phân tích tương đồng của dữ liệu TCGA và GEO, cho thấy mức độ biểu hiện của gen lncRNA *CRNDE* tăng có ý nghĩa thống kê và không có mối liên quan với giai đoạn tiến triển của bệnh.

3.1.2 Xác định các gen đồng biểu hiện với lncRNA *CRNDE*

Kết quả phân tích cho thấy lncRNA *CRNDE* có mối tương quan biểu hiện với nhiều gen khác nhau, hình thành một mạng lưới đồng biểu hiện (co-expression network). Trong mạng lưới này, lncRNA *CRNDE* đóng vai trò là nút trung tâm, liên kết với nhiều gen thông qua các cạnh thể hiện mối tương quan biểu hiện.

Các gen được xác định có tương quan với *CRNDE* bao gồm: *AGBL3*, *ZNF514*, *ZNF354B*, *ZNF32-AS2*, *GOLGA6L9*, *IRX5*, *EFHC2*, *GOT1-DT*, *ALDOA*, *SCO2*, *RAB5IF*, *PCBP1*, *MGAT2*, *LOC128125822*, *PDGXP*, *GNG5*, *PHGR*, *TAX1BP3*, *GNG10* và *PDF*.



Hình 3.5 Mạng lưới các gen có liên quan đến lncRNA *CRNDE*

Trong mạng lưới:

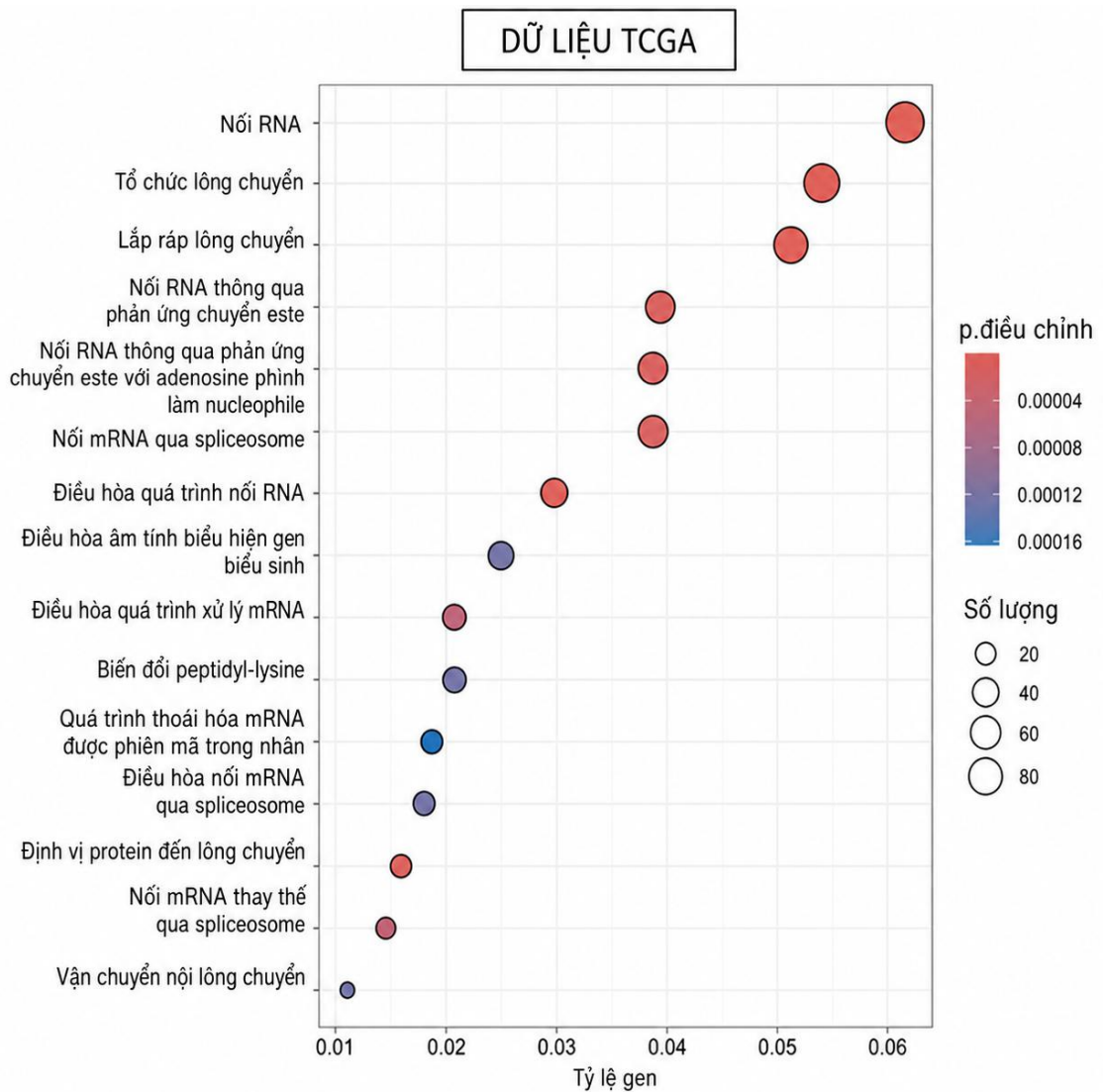
- Các đường màu đỏ biểu thị mối tương quan dương (positive correlation), cho thấy khi mức biểu hiện của lncRNA *CRNDE* tăng thì mức biểu hiện của các gen này cũng có xu hướng tăng.
- Các đường màu xanh biểu thị mối tương quan âm (negative correlation), cho thấy khi lncRNA *CRNDE* tăng thì biểu hiện của các gen này có xu hướng giảm.

Kết quả từ hình 3.5 cho thấy một số gen có mối tương quan dương đáng chú ý với lncRNA *CRNDE* như *AGBL3*, *GOLGA6L9*, *ZNF514*, *ZNF354B* và *EFHC2*, gợi ý khả năng chúng tham gia vào các cơ chế điều hòa hoặc các con đường sinh học liên quan đến hoạt động của lncRNA *CRNDE* trong UTĐTT.

Việc xác định các gen đồng biểu hiện với lncRNA *CRNDE* giúp làm rõ hơn mạng lưới điều hòa phân tử liên quan đến lncRNA này, đồng thời cung cấp cơ sở cho các phân tích chức năng tiếp theo như GO enrichment và KEGG pathway analysis, nhằm dự đoán các quá trình sinh học và con đường tín hiệu mà lncRNA *CRNDE* có thể tham gia trong quá trình sinh UTĐTT.

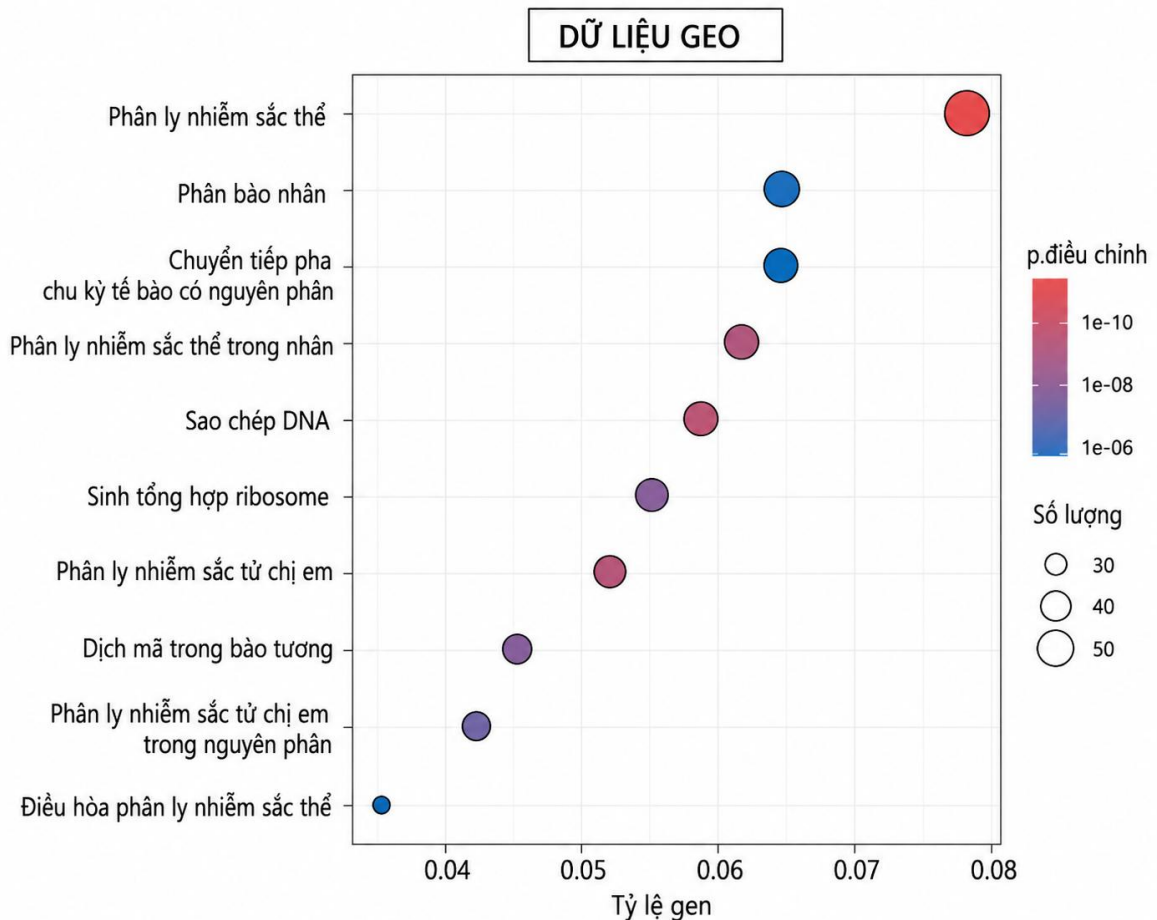
3.1.3 Phân tích tăng cường chức năng (GO)

Để dự đoán vai trò sinh học của các gen đồng biểu hiện với lncRNA *CRNDE*, phân tích tăng cường GO đã được thực hiện trên hai nguồn dữ liệu độc lập gồm TCGA-COAD và các bộ dữ liệu GEO. Các thuật ngữ GO có ý nghĩa thống kê được lựa chọn dựa trên giá trị *p.adjust* nhằm xác định các quá trình sinh học có liên quan đến mạng lưới gen đồng biểu hiện với lncRNA *CRNDE*.



Hình 3.6 Kết quả phân tích GO từ dữ liệu TCGA
(Adjusted p-value (FDR))

Kết quả phân tích từ TCGA-COAD cho thấy gen đồng biểu hiện với lncRNA *CRNDE* chủ yếu liên quan đến xử lý RNA và điều hòa splicing, những quá trình này đóng vai trò quan trọng trong điều hòa biểu hiện gen và sự phát triển của ung thư (Hình 3.6).



Hình 3.7 Kết quả phân tích GO từ dữ liệu GEO

(Adjusted p-value (FDR))

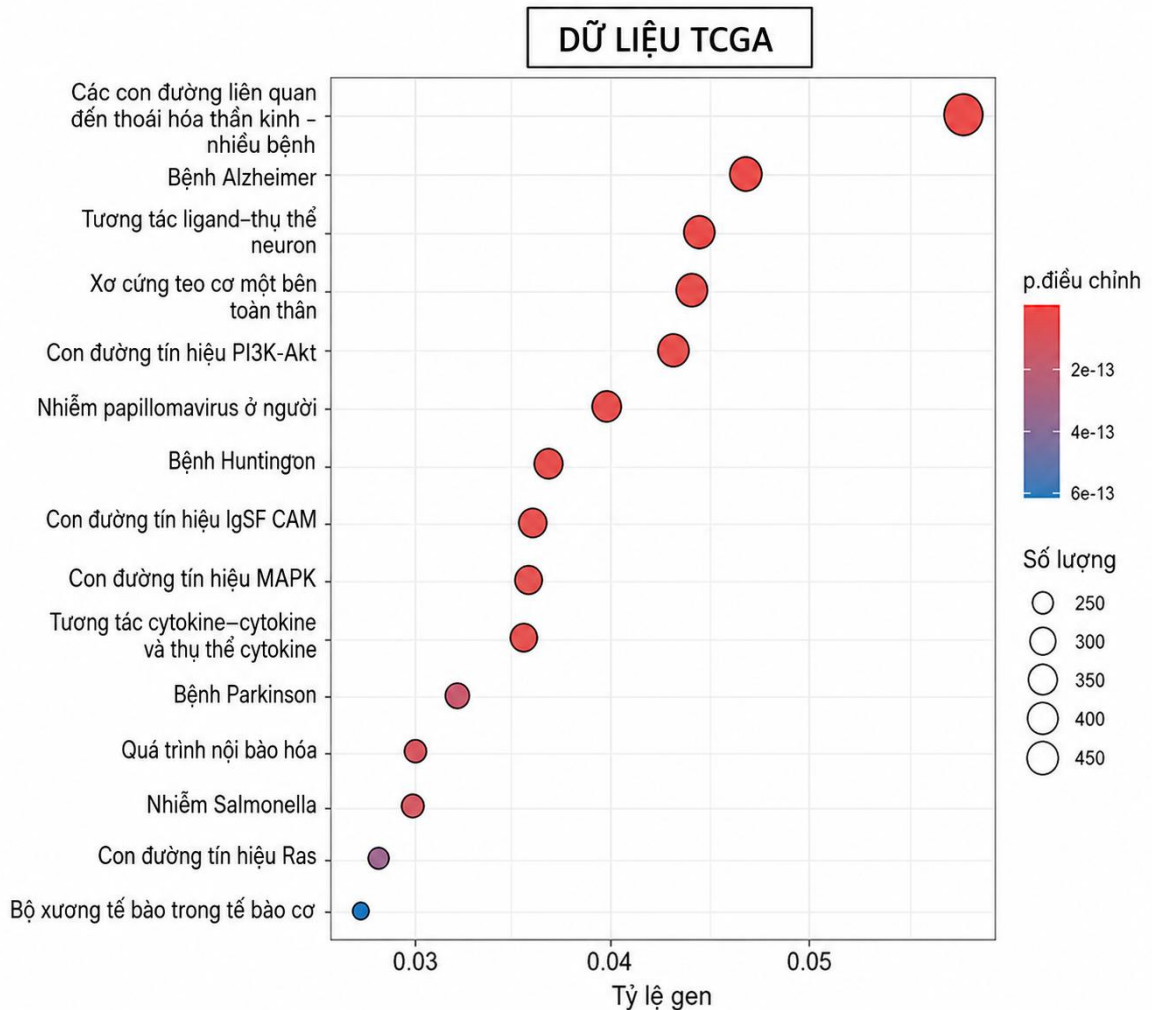
Đối với các bộ dữ liệu GEO, phân tích GO cho thấy các gen liên quan đến *CRNDE* chủ yếu tăng cường trong các quá trình liên quan đến chu kỳ tế bào và tăng sinh tế bào bao gồm từ quá trình sao chép DNA, phân chia nhiễm sắc thể đến khi phân chia tế bào (Hình 3.7).

Mặc dù không có sự đồng nhất 100% giữa kết quả phân tích ở 2 nguồn dữ liệu, tuy nhiên chúng thể hiện sự nhất quán trong xu hướng tham gia vào hai nhóm chức năng sinh học chính là điều hòa xử lý RNA và ghép nối RNA (từ dữ liệu TCGA) và

điều hòa chu kỳ tế bào và phân chia nhiễm sắc thể (từ dữ liệu GEO). Minh chứng cho việc lncRNA *CRNDE* góp phần vào cơ chế sinh bệnh của UTĐTT.

3.1.4 Phân tích con đường tín hiệu (KEGG)

Phân tích KEGG pathway đã được thực hiện trên tập hợp các gen đồng biểu hiện với lncRNA *CRNDE* từ hai nguồn dữ liệu TCGA-COAD và GEO. Các con đường tín hiệu có ý nghĩa thống kê được xác định dựa trên giá trị *p.adjust*.



Hình 3.8 Kết quả phân tích KEGG pathway từ dữ liệu TCGA

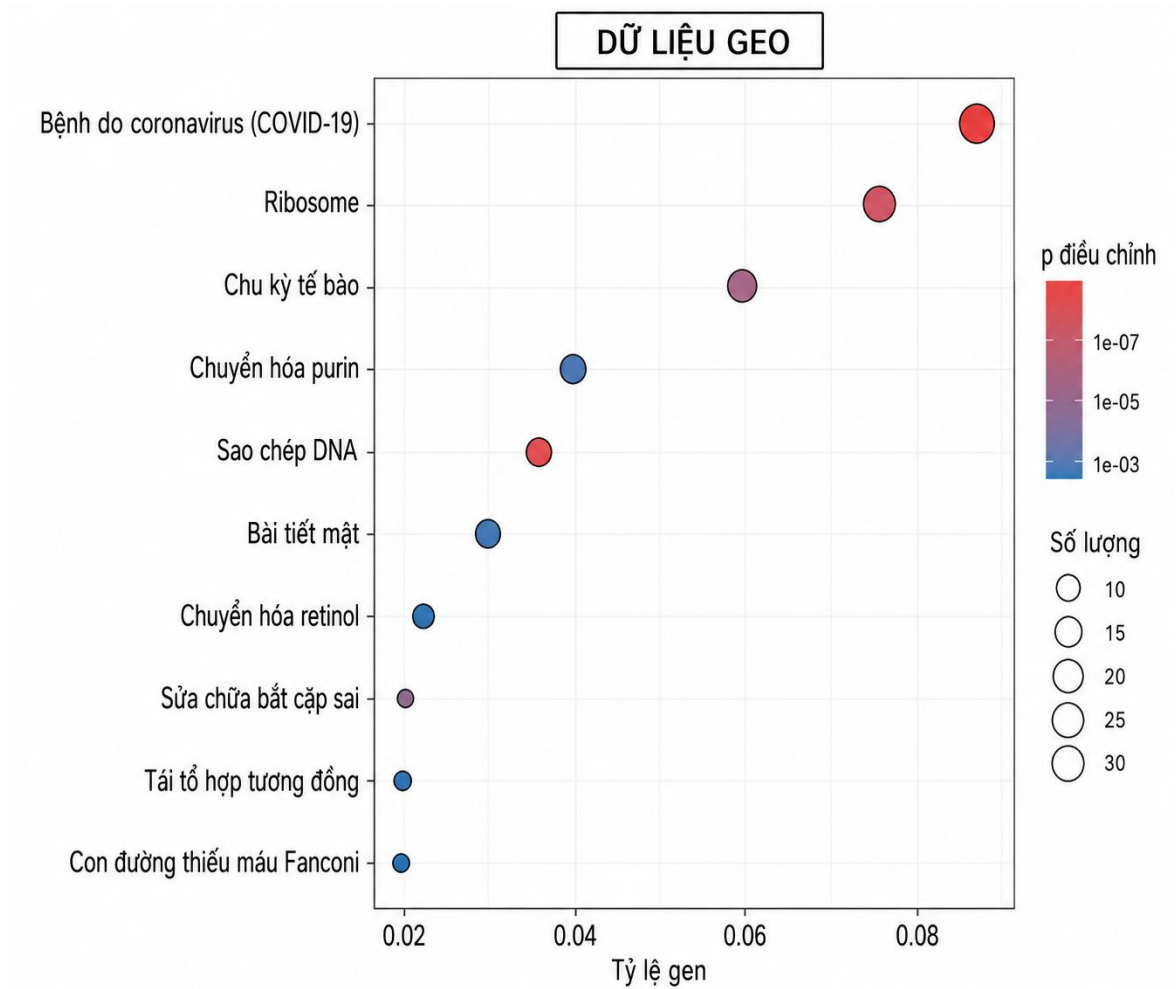
(Adjusted *p*-value (FDR))

Kết quả hình 3.8, phân tích từ dữ liệu TCGA cho thấy các gen đồng biểu hiện với lncRNA *CRNDE* được tăng cường trong nhiều con đường tín hiệu quan trọng liên quan đến ung thư và điều hòa tín hiệu tế bào. Các con đường tín hiệu nổi bật được ghi nhận bao gồm con đường PI3K-Akt, con đường MAPK, con đường Ras, tương tác cytokine-thụ thể cytokine và tương tác ligand-thụ thể thần kinh. Đây đều là những con đường có vai trò quan trọng trong điều hòa tăng sinh, sống sót, biệt hóa và truyền

tín hiệu của tế bào, đồng thời có liên quan chặt chẽ đến quá trình hình thành và tiến triển của ung thư.

Sự tăng cường của các pathway này cho thấy lncRNA *CRNDE* có thể liên quan đến các mạng lưới tín hiệu điều hòa tăng sinh, biệt hóa và sống sót của tế bào, đặc biệt thông qua các con đường PI3K-Akt, MAPK và Ras, vốn được biết là đóng vai trò quan trọng trong sự phát triển và tiến triển của UTĐTT.

Đối với dữ liệu GEO, các gen liên quan đến lncRNA *CRNDE* được tăng cường chủ yếu trong các con đường liên quan đến chu kỳ tế bào, sao chép và sửa chữa DNA. Các pathway nổi bật bao gồm con đường chu kỳ tế bào, sao chép DNA (DNA replication), sửa chữa bắt cặp sai, tái tổ hợp tương đồng và con đường thiếu máu Fanconi. Đây là những con đường có vai trò quan trọng trong kiểm soát tăng sinh tế bào, duy trì tính ổn định của bộ gen và bảo vệ tế bào trước các tổn thương DNA. Sự tăng cường của các pathway này cho thấy *CRNDE* có thể tham gia vào quá trình hình thành và tiến triển của ung thư thông qua điều hòa tăng sinh tế bào và các cơ chế sửa chữa DNA (Hình 3.9).



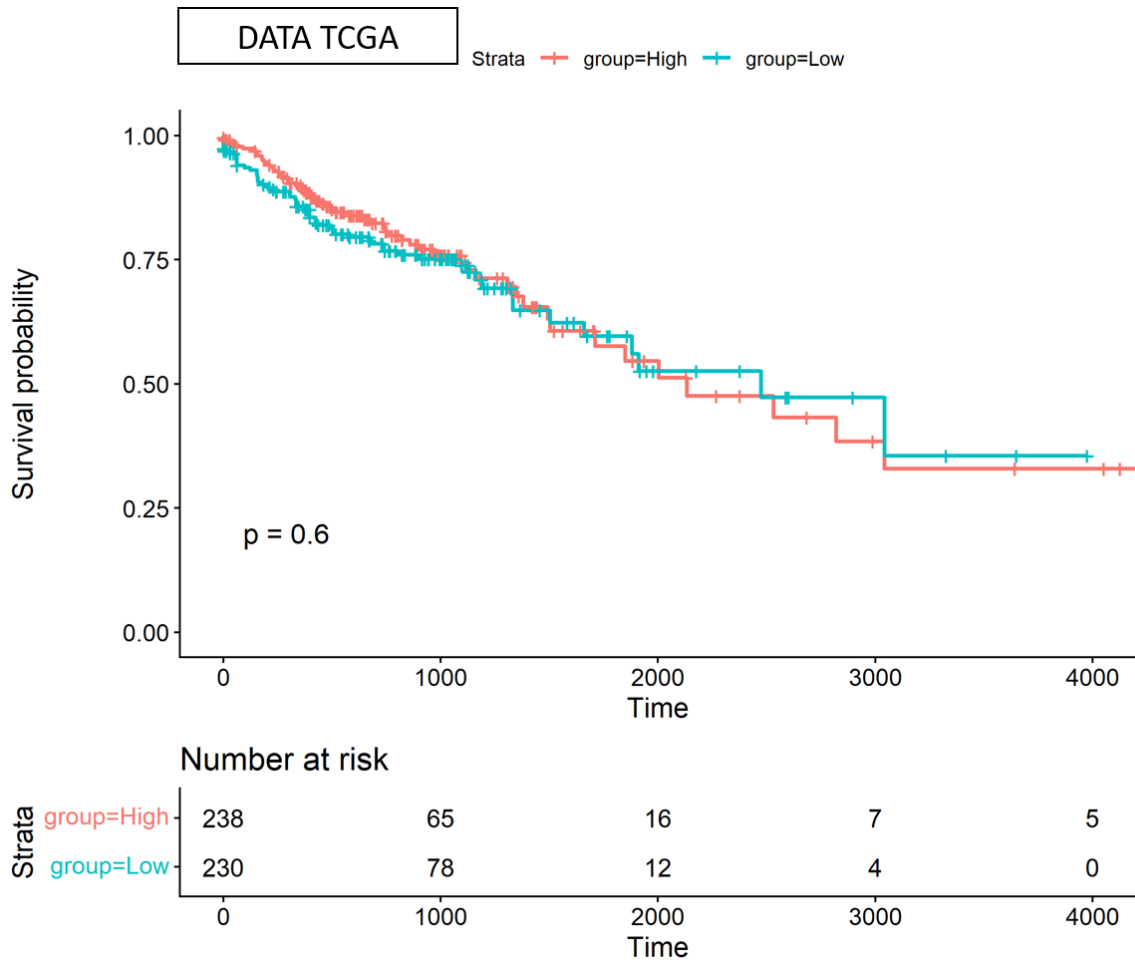
Hình 3.9 Kết quả phân tích KEGG pathway từ dữ liệu GEO
(Adjusted p -value (FDR))

Tổng hợp từ hai nguồn dữ liệu cho thấy các gen đồng biểu hiện với lncRNA *CRNDE* có xu hướng tham gia vào hai nhóm con đường tín hiệu chính: Các con đường tín hiệu liên quan đến ung thư như PI3K-Akt, MAPK và con đường tín hiệu Ras (dữ liệu TCGA). Các con đường liên quan đến chu kỳ tế bào và sao chép DNA như chu kỳ tế bào và nhân đôi DNA (dữ liệu GEO). Những kết quả này gợi ý rằng lncRNA *CRNDE* có thể đóng vai trò trong việc điều hòa các mạng lưới tín hiệu kiểm soát tăng sinh tế bào, chu kỳ tế bào và sửa chữa DNA, từ đó góp phần vào quá trình sinh UTĐTT.

3.1.5 Phân tích sống sót

Đường cong Kaplan–Meier cho thấy xác suất sống của hai nhóm High và Low gần như chồng lấp trong suốt thời gian theo dõi. Kiểm định log-rank cho kết quả $p = 0,6$, cho thấy không có sự khác biệt có ý nghĩa thống kê về thời gian sống giữa hai

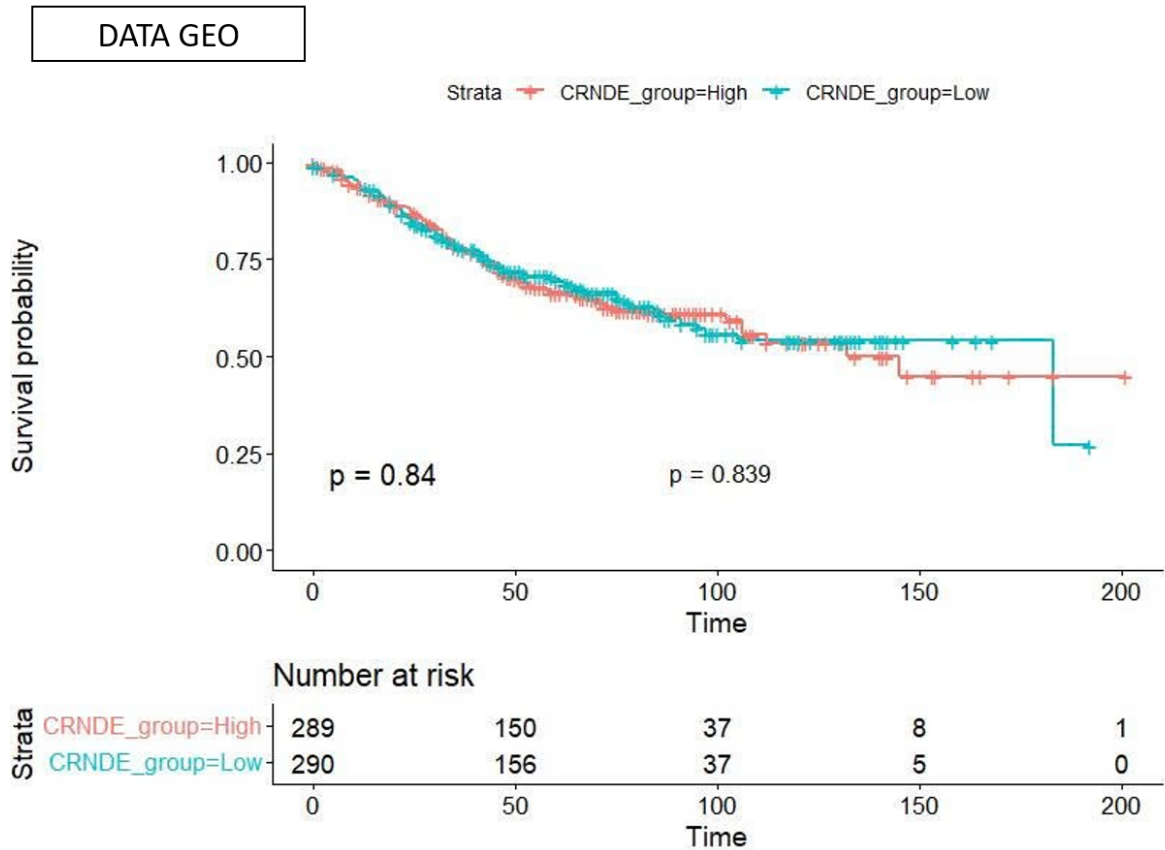
nhóm (Hình 3.10). Điều này gợi ý rằng trong bộ dữ liệu TCGA, mức biểu hiện gen chưa thể hiện vai trò tiên lượng rõ rệt.



Hình 3.10 Mức độ biểu hiện CRNDE liên quan đến sống sót từ dữ liệu TCGA

(Log-rank test)

Tương tự, ở bộ dữ liệu GEO, hai đường cong sống sót giữa nhóm lncRNA CRNDE biểu hiện cao và thấp cũng không có sự tách biệt rõ ràng. Giá trị $p = 0,84$ (và $\sim 0,839$ trong phân tích chi tiết) tiếp tục cho thấy không có sự khác biệt có ý nghĩa thống kê giữa hai nhóm (Hình 3.11).



Hình 3.11 Mức độ biểu hiện lncRNA *CRNDE* liên quan đến sống sót từ dữ liệu GEO

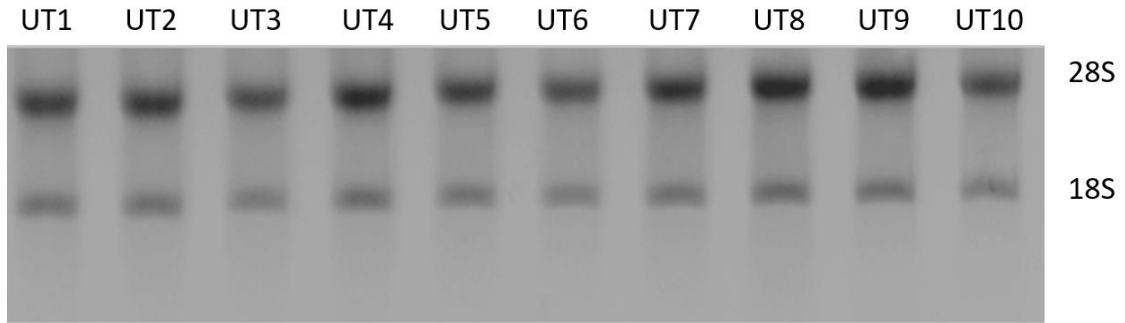
(Log-rank test)

Kết quả từ cả hai bộ dữ liệu đều nhất quán, cho thấy không có mối liên quan đáng kể giữa mức biểu hiện gen lncRNA *CRNDE* và thời gian sống tổng thể và gen này chưa chứng minh được giá trị tiên lượng độc lập trong bối cảnh nghiên cứu hiện tại.

3.2 KẾT QUẢ ĐÁNH GIÁ TRÊN MẪU LÂM SÀNG

3.2.1 Tách chiết mẫu RNA tổng số, kiểm tra chất lượng và tổng hợp cDNA

Các mẫu được tiến hành tách chiết RNA tổng số bằng bộ kit Trizol® (Thermo Fisher Scientific) theo quy trình do nhà sản xuất cung cấp. Sau khi thu nhận, RNA được đánh giá độ tinh sạch và nồng độ bằng máy đo quang phổ NanoDrop, đồng thời kiểm tra tính toàn vẹn bằng điện di trên gel agarose 1%. Kết quả được thể hiện ở hình 3.12.



Hình 3.12 Điện di đồ sản phẩm RNA tổng số

Các mẫu RNA tổng số được xác định nồng độ bằng máy NanoDrop và đánh giá chất lượng thông qua điện di trên gel agarose 1%, nhuộm ethidium bromide và ghi nhận kết quả bằng phần mềm Quantity One (Gel Doc EQ; Bio-Rad, Hercules, CA). Kết quả kiểm tra cho thấy tất cả các mẫu đều xuất hiện rõ hai băng 28S và 18S, đảm bảo tính nguyên vẹn của RNA.

Mẫu RNA đạt yêu cầu được chuẩn hóa về cùng nồng độ 0.5 $\mu\text{g}/\mu\text{L}$ với thể tích 20 μL , sau đó tổng hợp cDNA.

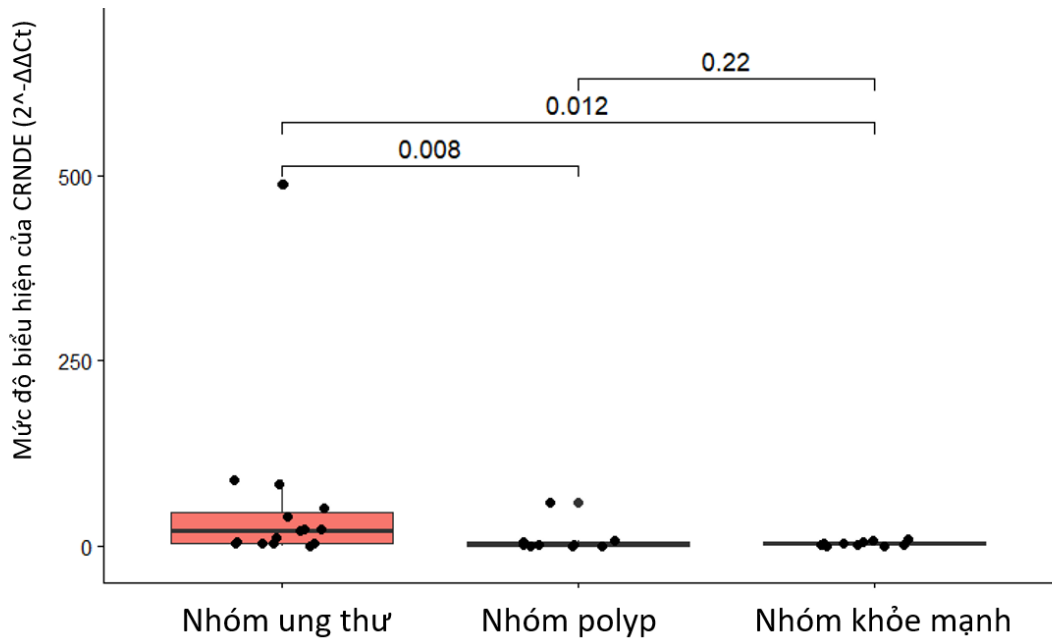
3.2.2 Mức độ biểu hiện của lncRNA *CRNDE*

Mức biểu hiện tương đối của gen *CRNDE* được đánh giá bằng phương pháp $2^{-\Delta\Delta\text{Ct}}$, sử dụng *GAPDH* làm gen nội chuẩn. Kết quả bảng 3.1 và hình 3.13 cho thấy mức biểu hiện của lncRNA *CRNDE* cao hơn ở nhóm bệnh nhân UTĐTT so với nhóm polyp và nhóm khỏe mạnh.

Bảng 3.1 Mức độ biểu hiện của lncRNA *CRNDE*

Nhóm	n	Mức độ biểu hiện (IQR)	p
UTĐTT	15	20.15 (3.68–45.5)	0.06
Polyp	10	2.69 (2.21–4.76)	
Khỏe mạnh	10	0.99 (0.33–4.71)	

Dunn post-hoc test

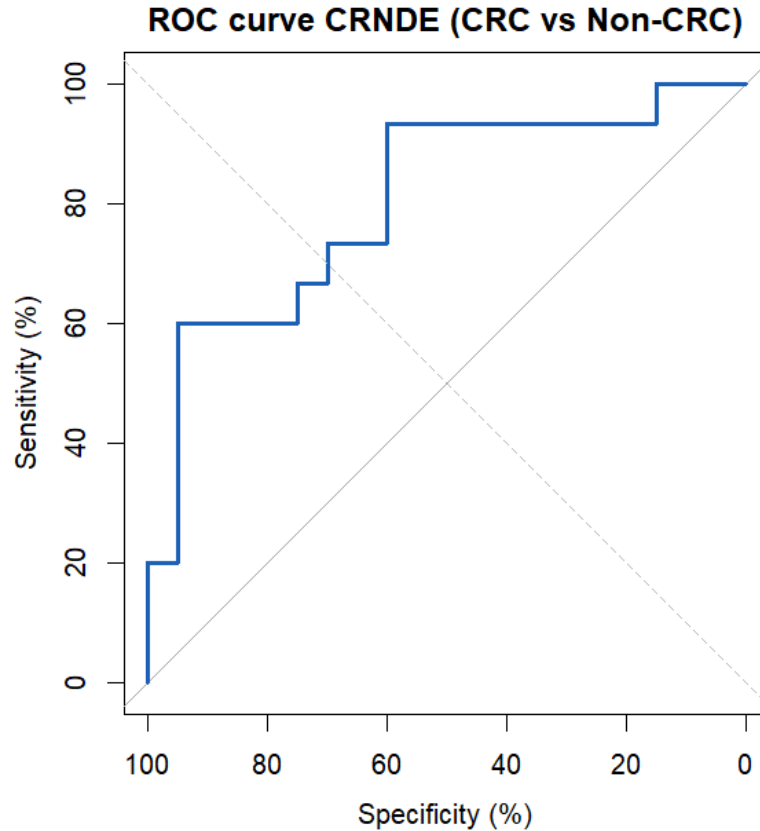


Hình 3.13 Phân bố mức độ biểu hiện của lncRNA CRNDE trên mẫu lâm sàng
(Kruskal-Wallis test)

So sánh giữa các nhóm bằng Kruskal-Wallis test kết hợp Dunn post-hoc test cho thấy, mức biểu hiện của lncRNA CRNDE ở nhóm UTĐTT cao hơn đáng kể so với nhóm polyp ($p = 0,008$) và so với nhóm khỏe mạnh ($p = 0,012$). Ngược lại, sự khác biệt giữa nhóm polyp và nhóm khỏe mạnh không có ý nghĩa thống kê ($p = 0,22$). Những kết quả này cho thấy lncRNA CRNDE có xu hướng tăng biểu hiện ở bệnh nhân UTĐTT, có khả năng phân biệt nhóm UTĐTT với nhóm polyp và nhóm khỏe mạnh, trong khi giữa polyp và nhóm khỏe mạnh mức biểu hiện không khác biệt rõ rệt.

3.2.3 Đánh giá hiệu quả phân biệt lncRNA CRNDE giữa UTĐTT và nhóm đối chứng bằng ROC

Phân tích đường cong ROC được thực hiện nhằm đánh giá khả năng phân biệt giữa nhóm ung thư đại trực tràng và nhóm đối chứng dựa trên mức độ biểu hiện của lncRNA CRNDE.



Hình 3.14 ROC curve của lncRNA CRNDE cho phân biệt CRC và nhóm đối chứng
 Bảng 3.2 Điểm cắt, AUC, độ nhạy, độ đặc hiệu và giá trị tiên đoán của
 lncRNA CRNDE

Chỉ số	Điểm cắt	AUC	Se (%)	Sp (%)	PPV (%)	NPV (%)	p
CRNDE	9.45	80.67	60	95	90	76	<0.001

DeLong test

Hình 3.14 trình bày đường cong ROC của lncRNA CRNDE trong việc phân biệt hai nhóm nghiên cứu. Diện tích dưới đường cong (AUC) đạt 0,8067, cho thấy khả năng phân biệt ở mức khá đến tốt. Giá trị AUC lớn hơn 0,8 chứng tỏ lncRNA CRNDE có tiềm năng đáng kể trong vai trò dấu ấn sinh học phân biệt.

Điểm cắt tối ưu được xác định là 9,45, tại đó độ nhạy (Se) đạt 60% và độ đặc hiệu (Sp) đạt 95% (Bảng 3.2). Kết quả này cho thấy lncRNA CRNDE có xu hướng ưu tiên độ đặc hiệu cao hơn độ nhạy, tức là khả năng loại trừ các trường hợp không mắc bệnh rất tốt, trong khi vẫn duy trì mức độ phát hiện bệnh ở mức chấp nhận được. Bên cạnh đó, giá trị tiên đoán dương (PPV) đạt 90% và giá trị tiên đoán âm (NPV) đạt 76%, phản ánh rằng phần lớn các trường hợp được xác định dương tính bởi

lncRNA *CRNDE* thực sự là bệnh nhân ung thư đại trực tràng. Tuy nhiên, giá trị NPV ở mức trung bình cho thấy vẫn tồn tại khả năng bỏ sót một số trường hợp bệnh. Sự khác biệt giữa hai nhóm có ý nghĩa thống kê với $p < 0,001$, khẳng định giá trị phân biệt của *CRNDE* không phải do ngẫu nhiên.

Nhìn chung, kết quả phân tích ROC cho thấy lncRNA *CRNDE* là một dấu ấn sinh học tiềm năng trong hỗ trợ phân biệt ung thư đại trực tràng. Với độ đặc hiệu cao, lncRNA *CRNDE* đặc biệt phù hợp trong các chiến lược phân biệt UTĐTT nhằm giảm tỷ lệ dương tính giả, và có thể hữu ích khi kết hợp với các dấu ấn sinh học khác để cải thiện độ nhạy tổng thể.

THẢO LUẬN

Kết quả của nghiên cứu này cho thấy sự tăng biểu hiện của lncRNA *CRNDE* trong ung thư đại trực tràng một cách nhất quán trên cả dữ liệu công khai (TCGA, GEO) và mẫu máu lâm sàng. Đáng chú ý, sự nhất quán này không chỉ thể hiện giữa các nguồn dữ liệu khác nhau mà còn phù hợp với xu hướng đã được ghi nhận trong nhiều nghiên cứu trước đây. Sự đồng nhất về xu hướng giữa các nguồn dữ liệu độc lập và giữa hai loại mẫu (mô và máu) cho thấy lncRNA *CRNDE* có tính ổn định cao về biểu hiện, đây là một đặc điểm quan trọng của một dấu ấn sinh học tiềm năng.

Cụ thể, điểm tương đồng nổi bật là lncRNA *CRNDE* đều được báo cáo tăng biểu hiện trong mô ung thư so với mô bình thường. Nhiều nghiên cứu trước đã chỉ ra rằng lncRNA *CRNDE* là một trong những lncRNA được hoạt hóa mạnh trong UTĐTT, với mức biểu hiện cao hơn đáng kể ở mô khối u [32, 34, 46, 48]. Kết quả của chúng tôi tái khẳng định xu hướng này khi cả hai bộ dữ liệu độc lập (TCGA và GEO) đều cho thấy sự gia tăng có ý nghĩa thống kê của lncRNA *CRNDE* ở nhóm ung thư. Đồng thời, việc ghi nhận sự tăng biểu hiện này trên mẫu máu lâm sàng càng củng cố tính ổn định và khả năng phát hiện của lncRNA *CRNDE* trong các loại mẫu khác nhau.

Tuy nhiên, cần lưu ý rằng dữ liệu TCGA/GEO được xây dựng chủ yếu từ mô khối u, trong khi nghiên cứu hiện tại sử dụng mẫu máu ngoại vi. Sự tương đồng về xu hướng biểu hiện giữa hai loại mẫu này có thể được giải thích bởi khả năng các lncRNA, bao gồm lncRNA *CRNDE*, được giải phóng vào tuần hoàn thông qua các túi ngoại bào như exosome và tồn tại ổn định trong huyết tương [46]. Điều này cho thấy lncRNA *CRNDE* có tiềm năng ứng dụng trong sinh thiết lỏng.

Sự tương đồng về xu hướng tăng biểu hiện giữa nghiên cứu hiện tại và các công bố trước không chỉ làm tăng độ tin cậy của kết quả, mà còn cho thấy lncRNA *CRNDE* có thể là một đặc điểm sinh học phổ biến và nhất quán trong UTĐTT. Điều này đặc biệt quan trọng trong bối cảnh phát triển biomarker, vì một dấu ấn sinh học chỉ thực sự có giá trị khi nó thể hiện được tính lặp lại và ổn định trên nhiều quần thể và phương pháp nghiên cứu khác nhau. Do đó, việc lncRNA *CRNDE* liên tục được ghi nhận tăng biểu hiện trong ung thư đại trực tràng qua nhiều nghiên cứu độc lập, bao gồm cả nghiên cứu này, đã củng cố thêm bằng chứng về vai trò tiềm năng của nó trong phân biệt bệnh. Đặc biệt, nghiên cứu này cung cấp thêm dữ liệu trên quần thể người Việt

Nam, góp phần bổ sung bằng chứng về tính phổ quát của lncRNA *CRNDE* trong các nhóm dân số khác nhau.

Trong nghiên cứu này không ghi nhận sự khác biệt biểu hiện rõ ràng mức độ biểu hiện của lncRNA *CRNDE* trong các giai đoạn bệnh, cũng như biểu hiện tăng của gen có ý nghĩa đối với tiên lượng sống sót của bệnh nhân. Điều này tương đồng với kết quả nghiên cứu của Ma và cộng sự (2022) khi chỉ ra mối liên hệ giữa biểu hiện lncRNA *CRNDE* cao với tiên lượng không thuận lợi ở sáu loại ung thư: ung thư vòm thượng thận, ung thư tế bào sáng thận, ung thư tuyến giáp, u tế bào thần kinh đệm cấp thấp, u nguyên bào thần kinh đệm đa hình và ung thư tế bào gan và hoàn toàn không bao gồm UTĐTT [60]. Trong khi đó nghiên cứu của Ding và cộng sự (2017) báo cáo mức độ biểu hiện của lncRNA *CRNDE* có sự khác nhau có ý nghĩa thống kê giữa giai đoạn sớm (I-II) và giai đoạn muộn (III-IV) [32]. Sự không nhất quán này có thể do một số nguyên nhân, bao gồm cỡ mẫu còn hạn chế, sự không đồng đều giữa các giai đoạn bệnh, cũng như đặc điểm sinh học dị hợp của UTĐTT. Điều này cho thấy vai trò tiên lượng của lncRNA *CRNDE* trong UTĐTT vẫn chưa thống nhất và cần được nghiên cứu thêm trên các quần thể lớn hơn.

Kết quả phân tích tăng cường chức năng cho thấy các gen đồng biểu hiện với lncRNA *CRNDE* có sự khác biệt đáng kể giữa dữ liệu GEO và TCGA. Tuy nhiên, sự khác biệt này không mang tính mâu thuẫn mà phản ánh các khía cạnh bổ sung về vai trò sinh học của lncRNA *CRNDE* trong UTĐTT. Cụ thể, dữ liệu GEO chủ yếu phản ánh các thay đổi ở mức độ chức năng tế bào, trong khi dữ liệu TCGA phản ánh các cơ chế điều hòa ở mức độ phân tử.

Trong bộ dữ liệu GEO, các gen liên quan đến lncRNA *CRNDE* chủ yếu được tăng cường trong các quá trình liên quan đến chu kỳ tế bào và phân bào, bao gồm phân ly nhiễm sắc thể, phân chia nhân và sao chép DNA. Điều này cho thấy lncRNA *CRNDE* có thể liên quan đến sự tăng sinh của tế bào ung thư. Các con đường KEGG như chu kỳ tế bào, sửa chữa DNA và tái tổ hợp tương đồng cũng củng cố nhận định này, cho thấy lncRNA *CRNDE* có thể góp phần vào việc duy trì sự tăng trưởng và ổn định bộ gen của tế bào khối u. Những kết quả này phù hợp với các nghiên cứu trước đây cho thấy lncRNA *CRNDE* thúc đẩy tăng sinh tế bào thông qua điều hòa chu kỳ tế bào và các gen liên quan đến phân bào.

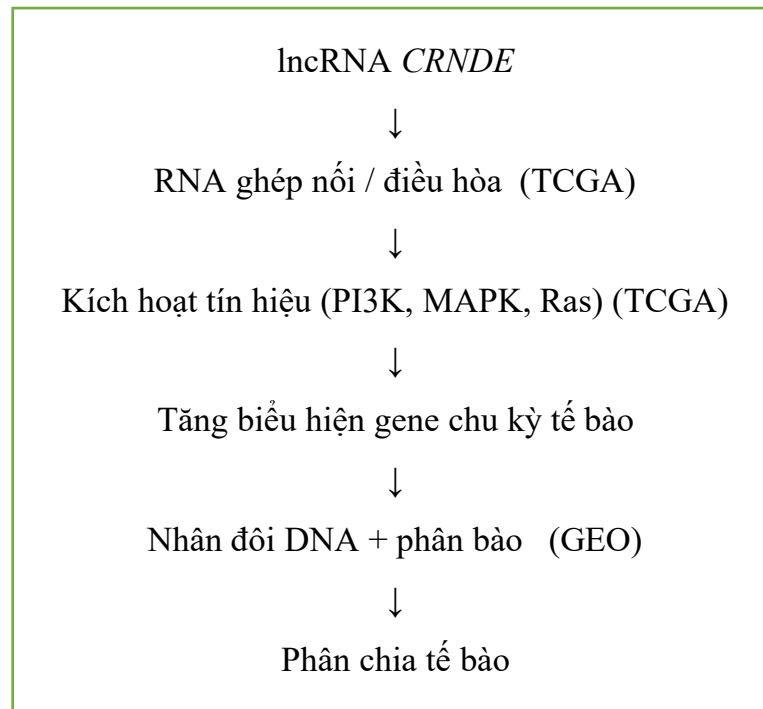
Kết quả phân tích chức năng từ dữ liệu GEO trong nghiên cứu này cho thấy các gen liên quan đến lncRNA *CRNDE* chủ yếu được tăng cường trong các quá trình sinh học như chu kỳ tế bào, phân bào và sao chép DNA. Điều này phù hợp với nhiều nghiên cứu trước đây, trong đó lncRNA *CRNDE* được chứng minh có vai trò thúc đẩy tăng sinh tế bào ung thư. Ví dụ, Han et al. (2017) cho thấy việc ức chế lncRNA *CRNDE* dẫn đến bất giữ chu kỳ tế bào và giảm tăng sinh tế bào trong UTĐTT [61]. Ngoài ra, các nghiên cứu khác như Ding et al. (2017) đã chỉ ra lncRNA *CRNDE* tham gia vào các con đường tín hiệu như Wnt/ β -catenin, góp phần thúc đẩy phân bào và tăng trưởng khối u [32]. Những bằng chứng này củng cố nhận định rằng lncRNA *CRNDE* có thể đóng vai trò quan trọng trong việc điều hòa chu kỳ tế bào và duy trì sự tăng sinh của tế bào UTĐTT. Ngược lại, trong dữ liệu TCGA, các gen đồng biểu hiện với lncRNA *CRNDE* chủ yếu được tăng cường trong các quá trình liên quan đến xử lý RNA như cắt nối RNA và xử lý mRNA. Kết quả này phù hợp với các nghiên cứu trước đây cho thấy lncRNA *CRNDE* có nhiều biến thể splice khác nhau và tham gia vào các cơ chế điều hòa RNA ở mức độ hậu phiên mã [48]. Bên cạnh đó, lncRNA *CRNDE* cũng được ghi nhận có vai trò trong điều hòa biểu hiện gen ở mức độ phiên mã, góp phần điều chỉnh hoạt động của nhiều gene liên quan đến ung thư [45].

Ngoài ra, các phân tích KEGG trong nghiên cứu này cho thấy sự tăng cường của các con đường tín hiệu quan trọng như PI3K-Akt, MAPK và Ras. Điều này tương đồng với các báo cáo trước đó, trong đó lncRNA *CRNDE* được chứng minh có khả năng hoạt hóa tín hiệu PI3K/Akt và điều hòa con đường MAPK, từ đó thúc đẩy tăng sinh, sống sót và tiến triển của tế bào ung thư [48]. Sự khác biệt giữa hai bộ dữ liệu có thể được giải thích bởi sự khác nhau về nền tảng công nghệ và đặc điểm dữ liệu. Cụ thể, dữ liệu GEO sử dụng nền tảng microarray với số lượng probe giới hạn, chủ yếu phản ánh các gen có mức biểu hiện cao, trong khi dữ liệu TCGA dựa trên RNA-seq cho phép phát hiện toàn diện hơn các transcript, bao gồm cả các gen có mức biểu hiện thấp và các RNA không mã hóa. Do đó, TCGA có xu hướng phản ánh tốt hơn các quá trình điều hòa phức tạp như xử lý RNA và tín hiệu nội bào, trong khi GEO nhấn mạnh các quá trình sinh học nổi bật như chu kỳ tế bào.

Một số con đường KEGG liên quan đến bệnh truyền nhiễm như bệnh *Coronavirus* hoặc nhiễm *Salmonella* cũng được ghi nhận trong phân tích. Tuy nhiên, các con đường này thực chất chứa nhiều gen liên quan đến đáp ứng stress, viêm và

tổng hợp protein, là những cơ chế chung giữa các quá trình bệnh lý khác nhau, bao gồm cả ung thư. Do đó, sự xuất hiện của các con đường này không phản ánh mối liên quan trực tiếp giữa lncRNA *CRNDE* và các bệnh truyền nhiễm, mà chủ yếu cho thấy sự chồng lấp về chức năng của các gen tham gia.

Tổng hợp các kết quả, có thể giả định rằng lncRNA *CRNDE* đóng vai trò như một phân tử điều hòa trong ung thư đại trực tràng, tham gia vào các quá trình xử lý RNA và điều hòa tín hiệu nội bào, từ đó gián tiếp thúc đẩy chu kỳ tế bào, sao chép DNA và sự tăng sinh của tế bào ung thư.



Sơ đồ 3.1 Mô hình cơ chế giả định của lncRNA *CRNDE* trong UTĐTT

Sự khác biệt giữa kết quả từ GEO và TCGA phản ánh các mức độ sinh học khác nhau của hoạt động lncRNA *CRNDE*. Trong khi dữ liệu TCGA cho thấy vai trò điều hòa ở mức độ phiên mã và tín hiệu, thì dữ liệu GEO phản ánh các hậu quả sinh học ở mức độ tế bào như tăng sinh và phân bào. Hai kết quả này mang tính bổ sung, góp phần làm rõ cơ chế hoạt động của lncRNA *CRNDE* trong ung thư đại trực tràng.

Nghiên cứu này vẫn còn một số hạn chế. Thứ nhất, đây là nghiên cứu đơn trung tâm và chưa có sự xác nhận trên các quần thể độc lập hoặc đa trung tâm, do đó khả năng khái quát hóa kết quả còn hạn chế. Thứ hai, các phân tích chức năng như GO và KEGG chủ yếu dựa trên phương pháp tin sinh học, do đó chỉ mang tính suy luận và chưa được kiểm chứng bằng các thí nghiệm chức năng *in vitro* hoặc *in vivo*. Ngoài

ra, cỡ mẫu lâm sàng còn tương đối hạn chế và việc sử dụng mẫu máu có thể chưa phản ánh đầy đủ đặc điểm sinh học của mô khối u. Do đó, các nghiên cứu tiếp theo với thiết kế đa trung tâm và kết hợp thực nghiệm chức năng là cần thiết để xác nhận và làm rõ vai trò của lncRNA *CRNDE* trong UTĐTT.

KẾT LUẬN VÀ KIẾN NGHỊ

KẾT LUẬN

Nghiên cứu này đã đạt được các mục tiêu đề ra với các kết quả chính như sau:

(1) Đối với mục tiêu 1, phân tích dữ liệu từ các cơ sở dữ liệu công khai (TCGA, GEO) cho thấy lncRNA *CRNDE* tăng biểu hiện có ý nghĩa thống kê trong UTĐTT so với mô bình thường. Đồng thời, phân tích tăng cường chức năng cho thấy lncRNA *CRNDE* liên quan đến các quá trình sinh học quan trọng như chu kỳ tế bào, phân bào, sao chép DNA và xử lý RNA, cũng như các con đường tín hiệu ung thư như PI3K-Akt, MAPK và Ras.

(2) Đối với mục tiêu 2, kết quả thực nghiệm trên mẫu máu ngoại vi cho thấy lncRNA *CRNDE* tăng biểu hiện rõ rệt ở nhóm ung thư so với nhóm polyp và nhóm khỏe mạnh. Đặc biệt, khi đối chiếu giữa dữ liệu *in silico* và kết quả thực nghiệm, nghiên cứu ghi nhận sự tương đồng về xu hướng biểu hiện, qua đó củng cố độ tin cậy của lncRNA *CRNDE* như một dấu ấn sinh học tiềm năng.

Tóm lại, lncRNA *CRNDE* là lncRNA có biểu hiện tăng đặc trưng trong ung thư đại trực tràng, có liên quan đến các cơ chế sinh học quan trọng và có tiềm năng ứng dụng trong phân biệt bệnh.

KIẾN NGHỊ

Cần thực hiện các nghiên cứu với cỡ mẫu lớn hơn và thiết kế đa trung tâm nhằm xác nhận tính ổn định và giá trị ứng dụng của lncRNA *CRNDE* trong UTĐTT. Đồng thời, cần tiến hành các nghiên cứu thực nghiệm *in vitro* và *in vivo* để làm rõ cơ chế phân tử của lncRNA *CRNDE*, đặc biệt trong điều hòa chu kỳ tế bào và các con đường tín hiệu liên quan đến ung thư. Bên cạnh đó, nên mở rộng đánh giá lncRNA *CRNDE* trên nhiều loại mẫu lâm sàng và xem xét kết hợp với các dấu ấn sinh học khác nhằm nâng cao hiệu quả phân biệt và theo dõi bệnh.

DANH MỤC TÀI LIỆU THAM KHẢO

1. National Cancer Institute. *Cancer Stat Facts: Colorectal Cancer*. Accessed February 2026.
2. Công thông tin Thành phố Hồ Chí Minh. *UNG THƯ ĐẠI TRỰC TRÀNG: NGUYÊN NHÂN, TRIỆU CHỨNG, CÁCH ĐIỀU TRỊ VÀ PHÒNG NGỪA*. 2023. Truy cập tháng 02 năm 2026
3. Organization, W.H. *Colorectal cancer*. 2026. Accessed February 2026.
4. MIMS Vietnam. *Ung thư đại trực tràng: Tổng quan về bệnh*. 2025 Truy cập tháng 02 năm 2026.
5. International Agency for Research on Cancer, *Viet Nam Fact Sheet 2022*, in *GLOBOCAN 2022*. 2024, International Agency for Research on Cancer (IARC). Accessed February 2026.
6. Duta-Ion, S.G.; Juganaru I.R.; Hotinceanu I.A.; Dan A.; Burtavel L.M.; Coman M.C.; Focsa I.O.; Zaruha A.G.; Codreanu P.C.; Bohiltea L.C. Redefining therapeutic approaches in colorectal cancer: targeting molecular pathways and overcoming resistance. *Int. J. Mol. Sci.* 2024, 25, 12507.
7. Kasi, A.; Handa S.; Bhatti S.; Umar S.; Bansal A.; Sun W. Molecular pathogenesis and classification of colorectal carcinoma. *Curr. Colorectal Cancer Rep.* 2020, 16, 97-106.
8. Nguyễn Thị Quỳnh Hương; Hồ Thị Ngọc; Hà Hải Anh. Các con đường tín hiệu chủ yếu trong nghiên cứu thuốc điều trị nhắm trúng đích cho ung thư đại trực tràng. Truy cập tháng 02 năm 2026.
9. Cao, Y.; Wang X. Effects of molecular markers on the treatment decision and prognosis of colorectal cancer: a narrative review. *J. Gastrointest. Oncol.* 2021, 12, 1191.
10. Yamagishi, H.; Kuroda H.; Imai Y.; Hiraishi H. Molecular pathogenesis of sporadic colorectal cancers. *Chin. J. Cancer* 2016, 35, 4.
11. Huang, Y.-H.; Lin P.-C.; Su W.-C.; Chan R.-H.; Chen P.-C.; Lin B.-W.; Shen M.-R.; Chen S.-H.; Yeh Y.-M. Association between altered Oncogenic Signaling Pathways and overall survival of patients with metastatic Colorectal Cancer. *Diagnostics* 2021, 11, 2308.

12. Telysheva, E.; Shaikhaev E.; Snigireva G. Mutational profile of KRAS-positive colorectal cancer. *Siberian journal of oncology* 2022, 21, 47-56.
13. Jankowski, W.M.; Fichna J.; Tarasiuk-Zawadzka A. Molecular mechanisms and pathophysiological implications of mucin-type O-glycosylation dysregulation in colorectal cancer progression. *Naunyn-Schmiedeberg's Archives of Pharmacology* 2025, 398, 13049-13063.
14. Zhang, C.; Geng H.; Tan Y.; Wang L. Multidimensional regulation of the microbe-TLR4 signaling Axis in colorectal cancer: From molecular mechanisms to microbe-targeted therapies. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 2025, 189397.
15. Paduraru, D.N.; Palcau A.C.; Dinca V.G.; Ciuc D.M.; Constantinescu A. The Role of Gut Microbiota in Colorectal Cancer Pathogenesis: A Comprehensive Literature Review. *Int. J. Mol. Sci.* 2025, 26, 11870.
16. Souza, G.D.d.; Souza L.R.Q.; Cuenca R.M.; Vilela V.M.; Santos B.E.d.M.; Aguiar F.S.d. Pre-and postoperative imaging methods in colorectal cancer. *ABCD. Arquivos Brasileiros de Cirurgia Digestiva (São Paulo)* 2018, 31, e1371.
17. Chitca, D.-D.; Popescu V.; Dumitrescu A.; Botezatu C.; Mastalier B. Advancing colorectal cancer diagnostics from barium enema to AI-assisted colonoscopy. *Diagnostics* 2025, 15, 974.
18. Didrikaitė, G.; Dulskas A. Oral Microbiota: an Innovative Approach to Colorectal Cancer Diagnosis. *Lietuvos chirurgija* 2025, 24, 176-183.
19. Chilimoniuk, Z.; Gładysz K.; Moniczewska N.; Chawrylak K.; Pelc Z.; Mlak R. The Role of Epigenetic Biomarkers as Diagnostic, Predictive and Prognostic Factors in Colorectal Cancer. *Cancers (Basel)* 2025, 17, 2632.
20. Hallom, P.; Naidoo P.; Senzani S.; Kader S.S.; Mkhize-Kwitshana Z.L. Stool- and Blood-Associated Colorectal Cancer Biomarkers: A Systematic Review. *Cancers (Basel)* 2025, 18, 96.
21. Ruiz-Quinga, A.A.; Noriega-Puga V.R. Evaluation of Epigenetic Markers in the Diagnosis and Prognosis of Colorectal Cancer. *SALUD, CIENCIA Y TECNOLOGÍA Учредители: Salud, Ciencia y Tecnologia* 2025, 5, 1139.

22. Olorosisimo, H.E.T.; Sumperos C.G.; Adviento A.L.A.; Lachica A.A.T.; Cabaltega J.L.A.; Ubiña A.G.R.; Valentin J.T.; Bremner P.R. Evaluating plasma and tissue biopsy for DNA methylation markers in early colorectal cancer detection: a systematic review. *Asian Biomed. (Res. Rev. News)* 2025, 19, 266.
23. Igder, S.; Zamani M.; Fakher S.; Siri M.; Ashktorab H.; Brim H.; LOS MAREK J.; Mokarram P. Novel Molecular Prognostic and Diagnostic Techniques in Colorectal Cancer. 2022,
24. Alabd, A.; Ni J.; Wang X.; Bai S.; Teng Y. Long non-coding RNAs: A promising tool to improve horticultural quality traits. *Current Plant Biology* 2024, 40, 100413.
25. Mattick, J.S.; Amaral P.P.; Carninci P.; Carpenter S.; Chang H.Y.; Chen L.-L.; Chen R.; Dean C.; Dinger M.E.; Fitzgerald K.A. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nature reviews Molecular cell biology* 2023, 24, 430-447.
26. Chodurska, B.; Kunej T. Long non-coding RNAs in humans: Classification, genomic organization and function. *Non-coding RNA research* 2025, 11, 313-327.
27. Alessio, E.; Bonadio R.S.; Buson L.; Chemello F.; Cagnin S. A single cell but many different transcripts: a journey into the world of long non-coding RNAs. *Int. J. Mol. Sci.* 2020, 21, 302.
28. Hon, C.-C.; Ramilowski J.A.; Harshbarger J.; Bertin N.; Rackham O.J.; Gough J.; Denisenko E.; Schmeier S.; Poulsen T.M.; Severin J. An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* 2017, 543, 199-204.
29. Zhang, X.-Z.; Liu H.; Chen S.-R. Mechanisms of long non-coding RNAs in cancers and their dynamic regulations. *Cancers (Basel)* 2020, 12, 1245.
30. Statello, L.; Guo C.-J.; Chen L.-L.; Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nature reviews Molecular cell biology* 2021, 22, 96-118.
31. Luo, Z. Regulatory mechanisms of long non-coding RNAs. *Oncology and Translational Medicine* 5(3):p, 147-151

32. Ding, J.; Li J.; Wang H.; Tian Y.; Xie M.; He X.; Ji H.; Ma Z.; Hui B.; Wang K. Long noncoding RNA CRNDE promotes colorectal cancer cell proliferation via epigenetically silencing DUSP5/CDKN1A expression. *Cell Death Dis.* 2017, 8, e2997-e2997.
33. Jiang, H.; Wang Y.; Ai M.; Wang H.; Duan Z.; Wang H.; Zhao L.; Yu J.; Ding Y.; Wang S. Long noncoding RNA CRNDE stabilized by hnRNPUL2 accelerates cell proliferation and migration in colorectal carcinoma via activating Ras/MAPK signaling pathways. *Cell Death Dis.* 2017, 8, e2862-e2862.
34. Yu, B.; Ye X.; Du Q.; Zhu B.; Zhai Q.; Li X.-X. The long non-coding RNA CRNDE promotes colorectal carcinoma progression by competitively binding miR-217 with TCF7L2 and enhancing the Wnt/ β -catenin signaling pathway. *Cell. Physiol. Biochem.* 2017, 41, 2489-2502.
35. Gao, H.; Song X.; Kang T.; Yan B.; Feng L.; Gao L.; Ai L.; Liu X.; Yu J.; Li H. Long noncoding RNA CRNDE functions as a competing endogenous RNA to promote metastasis and oxaliplatin resistance by sponging miR-136 in colorectal cancer. *Oncotargets Ther.* 2017, 205-216.
36. SUN, X.; LIAN Y.; BAI Y.; YANG C.; HU S.; WANG G. Effect of lncRNA CRNDE targeting miR-384 on radiosensitivity of colorectal cancer cells. *Chinese Journal of Radiological Medicine and Protection* 2019, 893-898.
37. Schmitt, A.M.; Chang H.Y. Long noncoding RNAs in cancer pathways. *Cancer Cell* 2016, 29, 452-463.
38. Sun, J.; Yan J.; Yuan X.; Yang R.; Dan T.; Wang X.; Kong G.; Gao S. A computationally constructed ceRNA interaction network based on a comparison of the SHEE and SHEEC cell lines. *Cell. Mol. Biol. Lett.* 2016, 21, 21.
39. Norouzi, R.; Mohamadzade Z.; Norouzi R.; Norouzi R.; Esmaeili R.; Soltani B.M. In-silico and in-vitro evidence suggest LINC01405 as a sponge for miR-29b and miR-497-5p, and a potential regulator of Wnt, PI3K, and TGFB signaling pathways in breast carcinoma. *Cancer Reports* 2024, 7, e1972.

40. Zhong, M.-E.; Chen Y.; Zhang G.; Xu L.; Ge W.; Wu B. LncRNA H19 regulates PI3K–Akt signal pathway by functioning as a ceRNA and predicts poor prognosis in colorectal cancer: integrative analysis of dysregulated ncRNA-associated ceRNA network. *Cancer Cell Int.* 2019, 19, 148.
41. Bao, L. Roles, underlying mechanisms and clinical significances of LINC01503 in human cancers. *Pathology-Research and Practice* 2024, 254, 155125.
42. Arshi, A.; Mahmoudi E.; Raeisi F.; Dehghan Tezerjani M.; Bahramian E.; Ahmed Y.; Peng C. Exploring potential roles of long non-coding RNAs in cancer immunotherapy: a comprehensive review. *Front. Immunol.* 2024, 15, 1446937.
43. Yang, L. Abstract ES3-3: Progress, paradigms, and pathways of long non-coding RNAs in cancer. *Cancer Research* 2018, 78, ES3-3-ES3-3.
44. Elimam, H.; Radwan A.F.; El Said N.H.; Elfar N.; Abd-Elmawla M.A.; Aborehab N.M.; Nassar K.; Mohammed O.A.; Doghish A.S. Long non-coding RNAs and signaling networks in non-small cell lung cancer: mechanistic insights into tumor pathogenesis. *Cancer Gene Ther.* 2025, 32, 1145-1165.
45. Graham, L.D.; Pedersen S.K.; Brown G.S.; Ho T.; Kassir Z.; Moynihan A.T.; Vizgoft E.K.; Dunne R.; Pimlott L.; Young G.P. Colorectal neoplasia differentially expressed (CRNDE), a novel gene with elevated expression in colorectal adenomas and adenocarcinomas. *Genes & cancer* 2011, 2, 829-840.
46. Liu, T.; Zhang X.; Gao S.; Jing F.; Yang Y.; Du L.; Zheng G.; Li P.; Li C.; Wang C. Exosomal long noncoding RNA CRNDE-h as a novel serum-based biomarker for diagnosis and prognosis of colorectal cancer. *Oncotarget* 2016, 7, 85551.
47. Hor, Y.Z.; Salvamani S.; Gunasekaran B.; Yian K.R. CRNDE: a pivotal oncogenic long non-coding RNA in cancers. *The Yale Journal of Biology and Medicine* 2023, 96, 511.
48. Lu, Y.; Sha H.; Sun X.; Zhang Y.; Wu Y.; Zhang J.; Zhang H.; Wu J.; Feng J. CRNDE: an oncogenic long non-coding RNA in cancers. *Cancer Cell Int.* 2020, 20, 162.

49. Colaprico, A.; Silva T.C.; Olsen C.; Garofano L.; Cava C.; Garolini D.; Sabedot T.S.; Malta T.M.; Pagnotta S.M.; Castiglioni I. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016, 44, e71-e71.
50. Davis, S.; Meltzer P.S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 2007, 23, 1846-1847.
51. Siddiqui, N., *Dplyr and TidyR*, in *Cultural Analytics in R: A Tidy Approach*. 2025, Springer. p. 57-95.
52. Wickham, H.; Bryan J.; Kalicinski M.; Valery K.; Leittenne C.; Colbert B.; Hoerl D.; Miller E.; Bryan M.J. Package 'readxl'. *Version* 2019, 13,
53. Robinson, M.D.; McCarthy D.J.; Smyth G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26, 139-140.
54. Wickham, H. ggplot2. *Wiley interdisciplinary reviews: computational statistics* 2011, 3, 180-185.
55. Chattamvelli, R., *Pearson's correlation*, in *Correlation in Engineering and the Applied Sciences: Applications in R*. 2024, Springer. p. 55-76.
56. Csardi, M.G. Package 'igraph'. *Last accessed* 2013, 3, 2013.
57. Yu, G.; Wang L.-G.; Han Y.; He Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012, 16, 284-287.
58. Jiang, F.; Guterman E. Survival analysis. *Statistical Methods in Epilepsy* 2024, 124-142.
59. Livak, K.J.; Schmittgen T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻ ΔΔCT method. *methods* 2001, 25, 402-408.
60. Ma, X.; Jin W.; Zhao C.; Wang X.; Wang K. CRNDE: A valuable long noncoding RNA for diagnosis and therapy of solid and hematological malignancies. *Molecular Therapy Nucleic Acids* 2022, 28, 190-201.
61. Han, P.; Li J.-w.; Zhang B.-m.; Lv J.-c.; Li Y.-m.; Gu X.-y.; Yu Z.-w.; Jia Y.-h.; Bai X.-f.; Li L. The lncRNA CRNDE promotes colorectal cancer cell proliferation and chemoresistance via miR-181a-5p-mediated regulation of Wnt/β-catenin signaling. *Molecular cancer* 2017, 16, 9.

PHỤ LỤC 1**DANH SÁCH MẪU MÁU SỬ DỤNG TRONG NGHIÊN CỨU***(Tên bệnh nhân lấy mẫu được viết tắt để đảm bảo yếu tố đạo đức trong Y học)*

TT	Lab ID-IGR	Họ và tên đệm	Tên	Năm sinh	Giới tính	Chẩn đoán
1	128	Phạm Quang	Đ.	1990	Nam	ung thư trực tràng 1/3 trên gđ cT3N2Mx (td di căn thùy trên phổi phải) đã hoá chất 1 chu kỳ
2	144	Tô Minh	Th.	1952	Nam	ung thư đại tràng sigma pT3N2M0
3	147	Nguyễn Thị	H.	1958	Nữ	ung thư 1/3 dưới trực tràng giai đoạn T2N2M0
4	148	Dương Văn	L.	1956	Nam	ung thư đại tràng sigma gđ cT4NxMx đã PT làm hậu môn nhân tạo
5	149	Nguyễn Bá	Đ.	1958	Nam	ung thư đại tràng sigma pT3N0M0
6	150	Đặng Văn	Ch.	1955	Nam	ung thư đại tràng sigma pT3N2M0
7	151	Trần Thị	S.	1932	Nữ	ung thư đại tràng sigma giai đoạn IV cT4N2M1 (di căn tử cung) đang hoá chất
8	152	Đặng Thị	M.	1942	Nữ	ung thư trực tràng
9	273	Trịnh Thị	H.	1952	Nữ	ung thư biểu mô tuyến đại tràng xuống T4aN0M0
10	282	Đặng Viết	G.	1954	Nam	ung thư biểu mô tuyến trực tràng
11	161	Phạm Thị	H.	1935	Nữ	ung thư đại tràng xuống T2N0M0/ K trực tràng T1N0M0

12	220	Lâm Đức	Ch.	2004	Nam	k trực tràng trên - giữa T3NxM1 - di căn phúc mạc
13	221	Đoàn Thị	L.	1964	Nữ	ung thư đại tràng ngang T3N1M0
14	224	Trịnh Xuân	V.	1984	Nam	tắc ruột do K đại tràng xuống di căn phúc mạc
15	225	Nguyễn Thị	Nh.	1952	Nữ	ung thư đại tràng gđ IV di căn gan đa ổ đã hoá chất
16	191	Đông Bá	C.	1956	Nam	polyp trực tràng
17	222	Nguyễn Quý	Ch.	1961	Nam	polyp đại tràng / viêm dạ dày cấp
18	253	Nguyễn Văn	CH.	1868	Nam	polyp đại tràng
19	308	Mai Nhất	L.	1963	Nam	polyp đại tràng
20	232	Nguyễn Đức	S.	1959	Nam	đa polyp đại tràng/ Gout mạn
21	243	Nguyễn Thị	L.	1957	Nữ	polyp đại tràng
22	252	Ngô Văn	Ch.	1952	Nam	polyp đại tràng
23	247	Đỗ Xuân	H.	1954	Nam	polyp đại tràng xuống.
24	256	Phạm Văn	V.	1955	Nam	polyp đại tràng
25	254	Dư Thị	H.	1991	Nữ	polyp đại tràng sigma
26	226	Lưu Thanh	H.	1986	Nam	Khỏe mạnh
27	238	Nguyễn Thị	H.	1959	Nữ	Khỏe mạnh
28	233	Dương Thị	T.	1963	Nữ	Khỏe mạnh
29	239	Đoàn Thị	T.	1983	Nữ	Khỏe mạnh
30	240	Phan Đình	T.	1995	Nam	Khỏe mạnh
31	235	Trần Thị	T.	1958	Nữ	Khỏe mạnh
32	244	Nguyễn Thị	Th.	1990	Nữ	Khỏe mạnh

33	250	Phùng Văn	K.	1985	Nam	Khỏe mạnh
34	294	Nguyễn Văn	D.	1958	Nam	Khỏe mạnh
35	251	Cù Anh	T.	1983	Nam	Khỏe mạnh

PHỤ LỤC 2

Kết quả realtime PCR bằng hệ thống QuantStudio 1

	<i>CRNDE</i>		<i>GAPDH</i>	
UT1	20.704	20.896	22.715	22.613
UT2	28.735	28.506	33.301	33.605
UT3	27.458	29.49	31.862	32.668
UT4	26.498	26.652	29.637	29.391
UT5	27.1	27.109	30.945	29.197
UT6	26.369	26.557	28.676	29.824
UT7	28.698	28.327	34.101	32.776
UT8	27.258	27.063	30.58	31.972
UT9	27.167	26.947	26.968	27.143
UT10	28.561	28.428	28.559	28.458
UT11	28.504	28.152	25.25	25.299
UT12	26.718	26.459	27.206	27.525
UT13	25.973	26.115	26.768	26.516
UT14	28.359	28.495	28.473	28.406
UT15	24.046	24.401	31.068	32.153
POLYP1	24.039	24.506	25.738	25.816
POLYP2	28.911	28.917	26.619	26.353
POLYP3	31.445	31.283	32.146	31.572
POLYP4	26.641	26.645	26.312	26.147
POLYP5	28.948	28.925	26.686	26.325
POLYP6	26.695	26.791	26.307	26.541
POLYP7	27.641	27.378	27.136	27.15
POLYP8	26.088	25.911	28.218	25.322
POLYP9	24.848	25.919	25.325	25.566
POLYP10	30.9	30.474	32.146	31.572
KM1	28.277	28.098	29.28	29.115
KM2	28.685	28.974	27.191	27.136
KM3	26.402	26.902	27.807	28.256
KM4	29.668	29.186	26.009	25.826
KM5	33.073	32.6	25.462	25.724
KM6	26.83	27.001	25.328	25.562
KM7	27.999	28.07	26.073	25.102
KM8	28.472	28.5	27.367	27.078
KM9	30.306	31.111	26.112	26.127
KM10	22.57	22.684	26.979	26.967

PHỤ LỤC 3

BỆNH VIỆN QUÂN Y 103
HỘI ĐỒNG ĐẠO ĐỨC TRONG
NGHIÊN CỨU Y SINH
Số: 341/CNCHT-HĐDD

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc
Hà Nội, ngày 15 tháng 11 năm 2023

CHỨNG NHẬN

Chấp thuận đề cương nghiên cứu

Căn cứ Quyết định số 3847/QĐ-BVQY103 ngày 26/8/2022 của Giám đốc Bệnh viện về việc kiện toàn Hội đồng, Ban cố vấn Hội đồng đạo đức trong nghiên cứu y sinh học cấp cơ sở Bệnh viện Quân y 103 nhiệm kỳ 2019-2024;

Căn cứ Quy chế tổ chức và hoạt động của Hội đồng đạo đức trong nghiên cứu y sinh học Bệnh viện Quân y 103 nhiệm kỳ 2019-2024;

Căn cứ biên bản họp Hội đồng đạo đức cấp cơ sở tại Bệnh viện Quân y 103 và bản giải trình của nghiên cứu viên (Nghiên cứu sinh);

Xét đề nghị của Thường trực Hội đồng Đạo đức Bệnh viện Quân y 103.

Hội đồng đạo đức trong nghiên cứu y sinh học Bệnh viện Quân y 103 chấp thuận:

A. *Khía cạnh khoa học và đạo đức đối với nghiên cứu:*

- Tên đề tài: Nghiên cứu mức độ biểu hiện và vai trò của một số ARN chuỗi dài không mã hoá máu ngoại vi trong chẩn đoán ung thư đại trực tràng.
- Đối tượng nghiên cứu: Đối tượng nghiên cứu gồm 2 nhóm (nhóm bệnh và nhóm chứng) được thu thập tại Bệnh viện Quân Y 103, Bệnh viện Trung ương Quân đội 108 và Bệnh viện K cơ sở Tân Triều. Nhóm bệnh gồm 100 bệnh nhân ung thư đại trực tràng ở các giai đoạn khác nhau. Nhóm chứng gồm 120 bệnh nhân, trong đó có 90 bệnh nhân u tuyến đại trực tràng và 30 bệnh nhân có rối loạn tiêu hoá nhưng soi đại tràng không phát hiện ung thư đại trực tràng hoặc u tuyến đại trực tràng
- Mã số (chuyên ngành): 9720107
- Nghiên cứu viên (Nghiên cứu sinh): Nguyễn Ngọc Khánh
- Địa điểm triển khai: Bệnh viện Quân y 103, Bệnh viện Trung ương Quân đội 108, Bệnh viện K cơ sở Tân Triều
- Thời gian nghiên cứu: từ tháng 11 năm 2023 đến tháng 11 năm 2026.

B. Cho phép sử dụng các tài liệu sau trong nghiên cứu nêu trên: xem Phụ lục đính kèm.

Ngày chấp thuận: 15/11/2023

Nghiên cứu viên chính có trách nhiệm: tuân thủ nội dung đề cương nghiên cứu, hồ sơ nghiên cứu đã được phê duyệt; các quy trình thực hiện chuẩn, các nguyên tắc GCP và các quy định của pháp luật hiện hành về đạo đức nghiên cứu.

Nơi nhận:

- Nghiên cứu viên (Nghiên cứu sinh);
- Lưu: VT, HĐDD. T03.

CHỦ TỊCH HỘI ĐỒNG ĐẠO ĐỨC
PHÓ GIÁM ĐỐC



PGS. TS. Tạ Bá Thắng

**DANH MỤC TÀI LIỆU NGHIÊN CỨU ĐƯỢC HỘI ĐỒNG ĐẠO ĐỨC
TRONG NGHIÊN CỨU Y SINH BỆNH VIỆN QUÂN Y 103 PHÊ DUYỆT**
(Kèm theo GCN số: **CNCHT-HĐĐĐ** ngày .../11/2023)

STT	Tên tài liệu
1.	Đơn xin đánh giá đạo đức trong nghiên cứu y sinh
2.	Quyết định thành lập tiểu ban chuyên môn xét duyệt đề cương nghiên cứu sinh của ứng viên Phạm Khánh Hưng, Biên bản họp xét tuyển của tiểu ban chuyên môn, Bản cam kết đã hoàn thiện hồ sơ theo ý kiến của Tiểu ban chuyên môn xét tuyển Nghiên cứu viên.
3.	Bản đăng kí chứng nhận chấp thuận Đạo đức trong nghiên cứu y sinh học
4.	Đề cương nghiên cứu đã được phê duyệt
5.	Chiến lược bảo vệ đối tượng nghiên cứu là con người
6.	Lý lịch khoa học của ứng viên Nghiên cứu viên chính/Nghiên cứu sinh
7.	Bản cung cấp thông tin cho đối tượng tham gia nghiên cứu
8.	Mẫu phiếu chấp thuận tình nguyện tham gia nghiên cứu
9.	Bản cam kết thực hiện theo đúng các nguyên tắc về đạo đức trong nghiên cứu của nhà nghiên cứu
10.	Bệnh án nghiên cứu
11.	Thư đồng thuận cho phép triển khai nghiên cứu của cơ quan nơi triển khai nghiên cứu.

[Handwritten Signature]