

Phân tích các biến đổi phân tử của CDKN2A và mối liên quan đến biểu hiện gen trong ung thư phổi không tế bào nhỏ

CDKN2A alterations and their association with gene expression in non-small cell lung cancer

Trần Thanh Bình¹, Nguyễn Văn Khôi²
và Trần Thị Thu Huyền^{2,*}

¹Bệnh viện Quân y 103

²Học viện Quân y

Tóm tắt

CDKN2A là gene ức chế khối u quan trọng thường bị bất thường trong ung thư phổi không tế bào nhỏ (NSCLC) thông qua methyl hóa promoter, đột biến gene và biến đổi số lượng bản sao (CNV). Dữ liệu methyl hóa DNA, biểu hiện mRNA, đột biến gene, CNV và thông tin lâm sàng của 831 bệnh nhân NSCLC (459 trường hợp ung thư biểu mô tuyến phổi và 372 trường hợp ung thư biểu mô vảy phổi) được phân tích tích hợp. Kết quả cho thấy mất đoạn sâu có mối liên quan mạnh nhất với giảm biểu hiện mRNA CDKN2A ($\log_2$ (raw STAR count + 1)), với trung vị giảm từ 9,86 xuống 4,91 và mức giảm rõ rệt hơn ở ung thư biểu mô vảy phổi so với ung thư biểu mô tuyến phổi. Ngược lại, đột biến CDKN2A không liên quan đến giảm biểu hiện mRNA tương ứng. Tóm lại, các biến đổi phân tử của CDKN2A có mối liên quan khác nhau với mức biểu hiện mRNA trong NSCLC, trong đó mất đoạn sâu liên quan mạnh nhất với giảm biểu hiện CDKN2A và rõ rệt hơn ở LUSC so với LUAD.

Từ khóa: Ung thư phổi không tế bào nhỏ, CDKN2A, CNV, methyl hóa, TCGA.

Summary

CDKN2A is an important tumor suppressor gene involved in cell cycle regulation and is frequently altered in non-small cell lung cancer (NSCLC) through multiple mechanisms, including promoter methylation, gene mutation, and copy number variation (CNV). Integrated analyses were performed using DNA methylation, mRNA expression, mutation, CNV, and clinical data from 831 NSCLC patients, including 459 lung adenocarcinoma (LUAD) and 372 lung squamous cell carcinoma (LUSC) cases. The results demonstrated that deep deletion was the molecular alteration most strongly associated with reduced CDKN2A expression, with median expression levels decreasing from 9.86 to 4.91, and the reduction was more pronounced in LUSC than in LUAD. In contrast, CDKN2A mutations were not associated with a corresponding decrease in mRNA expression. These findings indicate that CDKN2A expression in NSCLC is influenced by multiple molecular alterations, among which deep deletion is the alteration most strongly associated with reduced gene expression. In addition, CDKN2A alterations were more prominent in LUSC, suggesting distinct molecular mechanisms between the major histological subtypes of NSCLC.

Keywords: Non-small cell lung cancer, CDKN2A, methylation, copy number variation, TCGA.

Ngày nhận bài: 22/6/2026, ngày chấp nhận đăng: 26/7/2026

* Tác giả liên hệ: tranthuhuyen.cnsh@gmail.com - Học viện Quân y

I. ĐẶT VẤN ĐỀ

Ung thư phổi là một trong những nguyên nhân hàng đầu gây tử vong do ung thư trên toàn thế giới¹. Trong đó, ung thư phổi không tế bào nhỏ (non-small cell lung cancer, NSCLC) chiếm khoảng 85% tổng số trường hợp ung thư phổi². Ung thư phổi không tế bào nhỏ bao gồm hai phân nhóm mô bệnh học chính là ung thư biểu mô tuyến phổi (lung adenocarcinoma, LUAD) và ung thư biểu mô vảy phổi (lung squamous cell carcinoma, LUSC)³.

CDKN2A là một gene ức chế khối u quan trọng nằm trên vùng nhiễm sắc thể 9p21, tham gia điều hòa chu kỳ tế bào thông qua ức chế cyclin-dependent kinases 4 và 6 (CDK4/6)⁴⁻⁷. Sự bất hoạt CDKN2A có thể dẫn đến mất kiểm soát chu kỳ tế bào và thúc đẩy quá trình sinh ung thư. Nhiều nghiên cứu đã ghi nhận CDKN2A thường gặp các biến đổi phân tử như methyl hóa promoter, đột biến gene và biến đổi số lượng bản sao (copy number variation, CNV).^{6,8-10}

Trong những năm gần đây, các cơ sở dữ liệu hệ gene quy mô lớn như The Cancer Genome Atlas (TCGA) cho phép đánh giá đồng thời nhiều lớp dữ liệu phân tử trên cùng bệnh nhân, bao gồm methyl hóa DNA, biểu hiện gene, đột biến và CNV. Cách tiếp cận phân tích tích hợp đa hệ dữ liệu (multi-omics integration) giúp làm rõ vai trò của từng cơ chế biến đổi phân tử trong điều hòa biểu hiện gene¹¹. Mặc dù các cơ chế biến đổi của CDKN2A đã được khảo sát trong các nghiên cứu trước, bao gồm cả đánh giá đồng thời nhiều dạng biến đổi phân tử, các phân tích trên dữ liệu đa hệ quy mô lớn của TCGA nhằm so sánh mối liên quan của từng dạng biến đổi với biểu hiện mRNA, đồng thời phân tầng theo phân nhóm mô học và trạng thái CNV, vẫn còn hạn chế¹².

Xuất phát từ những vấn đề trên, chúng tôi thực hiện nghiên cứu nhằm đánh giá mối liên quan giữa các biến đổi phân tử của CDKN2A, bao gồm methyl hóa promoter, đột biến gene và biến đổi số lượng bản sao, với mức biểu hiện mRNA CDKN2A trên quần thể bệnh nhân NSCLC trong cơ sở dữ liệu TCGA, đồng thời so sánh giữa hai phân nhóm mô học LUAD và LUSC.

II. Đối tượng và phương pháp

2.1. Đối tượng

Nghiên cứu sử dụng dữ liệu công khai từ cơ sở dữ liệu TCGA, bao gồm các trường hợp ung thư phổi không tế bào nhỏ thuộc hai phân nhóm mô học: Ung thư biểu mô tuyến phổi (LUAD) và Ung thư biểu mô vảy phổi (LUSC)

Tiêu chuẩn lựa chọn: Các trường hợp được đưa vào phân tích khi có đồng thời:

Dữ liệu methyl hóa promoter CDKN2A;

Dữ liệu biểu hiện mRNA CDKN2A;

Mã định danh bệnh nhân phù hợp giữa các bộ dữ liệu.

Tiêu chuẩn loại trừ:

Mẫu trùng lặp bệnh nhân

Trường hợp thiếu mã định danh phục vụ tích hợp dữ liệu;

Các mẫu không đạt yêu cầu kiểm soát chất lượng dữ liệu;

Bộ dữ liệu phân tích cuối cùng gồm 831 trường hợp NSCLC.

2.2. Phương pháp

Thiết kế nghiên cứu: Nghiên cứu mô tả cắt ngang, phân tích dữ liệu thứ cấp.

Thời gian và địa điểm nghiên cứu: Nghiên cứu được thực hiện từ tháng 01 năm 2026 đến tháng 6 năm 2026 tại Viện Nghiên cứu y dược học Quân sự, Học viện Quân y.

Thu thập và xử lý dữ liệu: Dữ liệu TCGA-LUAD và TCGA-LUSC được truy xuất bằng gói TCGAbiolinks trong phần mềm R. Dữ liệu biểu hiện gene được lấy từ GDC RNA-seq theo quy trình STAR – Counts và được chuyển đổi theo công thức $\log_2(\text{count} + 1)$ trước khi phân tích. Biểu hiện CDKN2A được phân tích ở mức gene và không phân biệt các transcript/isoform của CDKN2A. Dữ liệu methyl hóa DNA được thu thập từ nền tảng Illumina HumanMethylation450K dưới dạng β -value. Điểm methyl hóa vùng promoter của CDKN2A được tính bằng giá trị trung bình β -value của bảy CpG probes (cg00748581, cg03079681, cg08971670,

cg10848754, cg13601799, cg14430974 và cg26673943) thuộc vùng promoter của gene CDKN2A theo chú giải Illumina HumanMethylation450K. Điểm methyl hóa promoter được phân nhóm thành methyl hóa cao và methyl hóa thấp theo giá trị trung vị của điểm methyl hóa promoter trong quần thể nghiên cứu.

Dữ liệu đột biến gene được xử lý từ Mutation Annotation Format (MAF), sau đó lọc các trường hợp mang biến đổi tại gene CDKN2A. Dữ liệu CNV sử dụng gene-level GISTIC thresholded calls để xác định trạng thái biến đổi số lượng bản sao của CDKN2A. Trạng thái CNV của CDKN2A được phân loại thành mất đoạn sâu, mất đoạn nông (shallow deletion), bình thường (diploid), tăng bản sao (gain) và khuếch đại gene (amplification). Việc phân loại được thực hiện dựa trên gene-level GISTIC thresholded calls của TCGA, với các giá trị tương ứng: mất đoạn sâu (≤ -2), mất đoạn nông (-1), bình thường (0), tăng bản sao (1) và khuếch đại (≥ 2).

Sau khi xử lý từng lớp dữ liệu, các bộ dữ liệu được tích hợp theo mã định danh bệnh nhân nhằm xây dựng bộ dữ liệu đa hệ dữ liệu cuối cùng.

Các biến số nghiên cứu: là các cơ chế biến đổi CDKN2A bao gồm: methyl hóa promoter, đột biến CDKN2A, trạng thái CNV của CDKN2A.

Biểu hiện CDKN2A được đánh giá dựa trên mức biểu hiện mRNA ở mức gene.

2.3. Phân tích thống kê

Các biến định lượng được trình bày dưới dạng trung vị và khoảng tứ phân vị (IQR). So sánh hai nhóm sử dụng kiểm định Wilcoxon rank-sum. So sánh nhiều nhóm sử dụng kiểm định Kruskal-Wallis. Tương quan được đánh giá bằng hệ số Spearman. Các phép kiểm thống kê được xác định trước theo

mục tiêu nghiên cứu nhằm đánh giá mối liên quan giữa các biến đổi phân tử của CDKN2A và mức biểu hiện mRNA CDKN2A. Các phân tích thống kê được thực hiện bằng R phiên bản 4.6.0, sử dụng các gói TCGAbiolinks, dplyr, ggplot2 và các gói Bioconductor. Giá trị $p < 0,05$ được xem là có ý nghĩa thống kê.

2.4. Đạo đức nghiên cứu

Nghiên cứu sử dụng dữ liệu công khai từ cơ sở dữ liệu The Cancer Genome Atlas (TCGA), không bao gồm thông tin định danh cá nhân của bệnh nhân. Toàn bộ dữ liệu được khai thác và phân tích tuân thủ các quy định sử dụng dữ liệu của TCGA. Do nghiên cứu sử dụng dữ liệu thứ cấp công khai và không can thiệp trên người bệnh, nghiên cứu không yêu cầu lấy ý kiến đồng thuận trực tiếp từ đối tượng nghiên cứu.

III. KẾT QUẢ

3.1. Đặc điểm đối tượng tham gia nghiên cứu

Nghiên cứu bao gồm 831 trường hợp NSCLC, trong đó có 459 trường hợp LUAD và 372 trường hợp LUSC. Đột biến CDKN2A được ghi nhận ở 9,4% trường hợp, với tần suất cao hơn ở LUSC so với LUAD (15,1% so với 4,8%). Có 825 trường hợp có dữ liệu CNV/GISTIC của CDKN2A để phân tích, bao gồm 456 trường hợp LUAD và 369 trường hợp LUSC. Trong số các trường hợp này, mất đoạn nông là trạng thái phổ biến nhất của CDKN2A (44,7%), tiếp theo là mất đoạn sâu (21,5%). Cả hai dạng mất đoạn đều xuất hiện thường xuyên hơn ở LUSC, trong khi trạng thái CNV bình thường gặp nhiều hơn ở LUAD. Tỷ lệ methyl hóa promoter mức cao tương đối tương đồng giữa hai phân nhóm mô học (Bảng 1).

Bảng 1. Đặc điểm đối tượng tham gia nghiên cứu (n = 831)

Đặc điểm	Overall (n = 831)	LUAD (n = 459)	LUSC (n = 372)	Giá trị p
Tuổi, trung vị (IQR), năm	67 (60 - 73)	65 (59 - 72)	69 (62 - 74)	<0,001 ^a
Thiếu dữ liệu	28 (3,4)	19 (4,1)	9 (2,4)	
Giới tính, n (%)				<0,001 ^b

Đặc điểm	Overall (n = 831)	LUAD (n = 459)	LUSC (n = 372)	Giá trị p
Nữ	342 (41,2)	246 (53,6)	96 (25,8)	
Nam	489 (58,8)	213 (46,4)	276 (74,2)	
Đột biến CDKN2A, n (%)				
Không	753 (90,6)	437 (95,2)	316 (84,9)	<0,001 ^b
Có	78 (9,4)	22 (4,8)	56 (15,1)	
Trạng thái số bản sao CDKN2A (GISTIC), n (%) [*]				
Mất đoạn sâu	177 (21,3)	74 (16,1)	103 (27,7)	<0,001 ^c
Mất đoạn nông	369 (44,4)	183 (39,9)	186 (50,0)	
Bình thường (Diploid)	202 (24,3)	152 (33,1)	50 (13,4)	
Tăng bản sao	71 (8,5)	44 (9,6)	27 (7,3)	
Khuếch đại gen	6 (0,7)	3 (0,7)	3 (0,8)	
Thiếu dữ liệu	6 (0,7)	3 (0,7)	3 (0,8)	
Tình trạng methyl hóa promoter, n (%)				0,121 ^b
Thấp	418 (50,3)	242 (52,7)	176 (47,3)	
Cao	413 (49,7)	217 (47,3)	196 (52,7)	

Dữ liệu được trình bày dưới dạng trung vị (IQR) đối với biến liên tục và số lượng (tỷ lệ %) đối với biến phân loại. Tỷ lệ phần trăm được tính trên tổng số bệnh nhân của từng nhóm; các giá trị thiếu dữ liệu được báo cáo riêng khi có. ^a Kiểm định Mann-Whitney U. ^b Kiểm định Chi-square của Pearson. ^c Kiểm định chính xác Fisher. ^{*} Phần trăm của các trạng thái CNV được tính trên số bệnh nhân có dữ liệu CNV/GISTIC: Overall n = 825, LUAD n = 456, LUSC n = 369. Sáu bệnh nhân thiếu dữ liệu CNV/GISTIC, bao gồm 3 LUAD và 3 LUSC.

3.2. Mối liên quan giữa các biến đổi phân tử của CDKN2A và biểu hiện gene

Để khảo sát mối liên quan các cơ chế biến đổi phân tử lên biểu hiện CDKN2A, chúng tôi tiến hành so sánh mức biểu hiện gene giữa các nhóm bệnh nhân theo trạng thái CNV, đột biến và methyl hóa promoter (Hình 1).

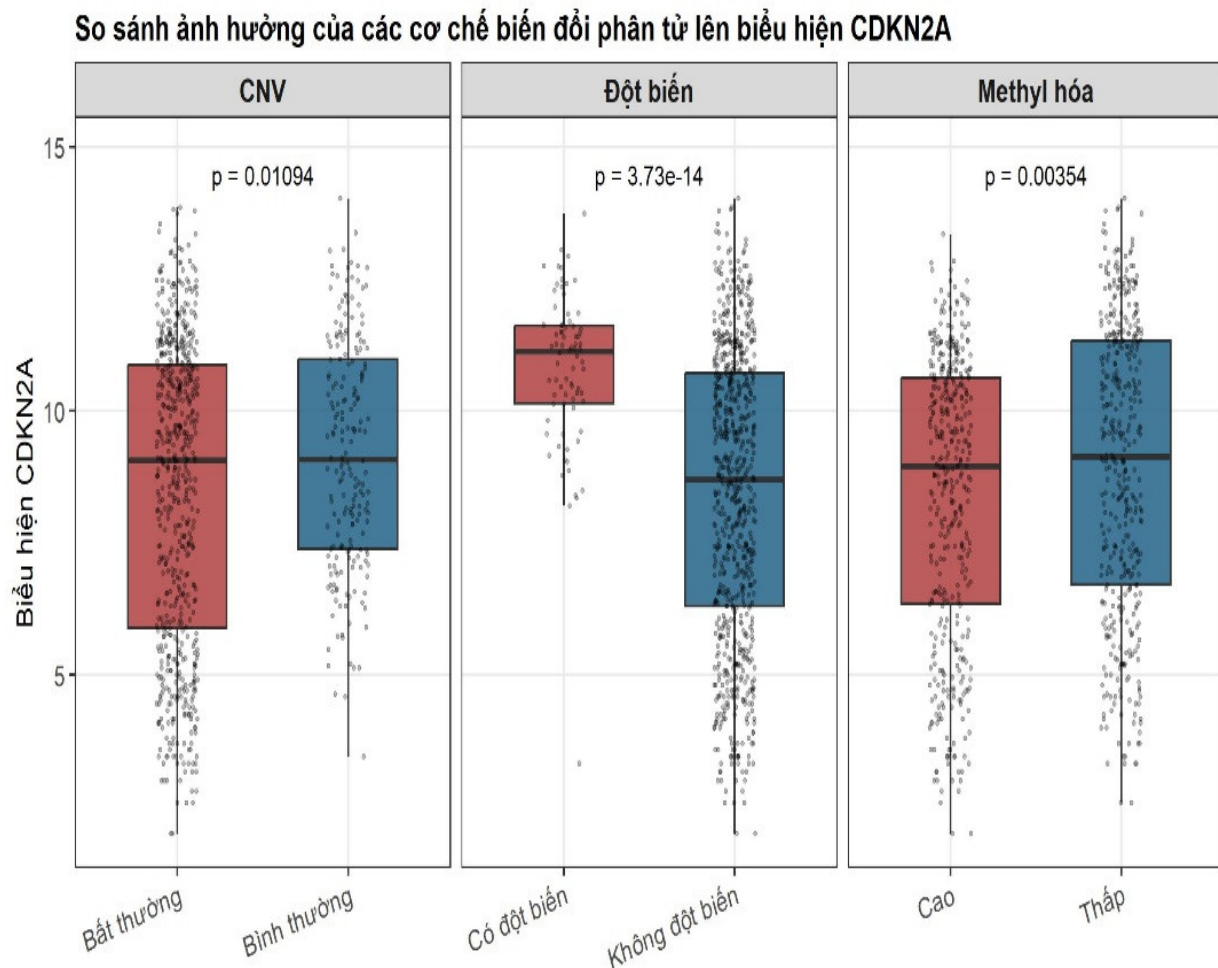
Phân tích mối liên quan giữa các cơ chế biến đổi phân tử với biểu hiện CDKN2A cho thấy cả ba cơ chế đều có mối liên quan với mức biểu hiện CDKN2A, tuy nhiên mức độ ảnh hưởng khác nhau. Nhóm bệnh nhân mang đột biến CDKN2A không ghi nhận

giảm mức biểu hiện mRNA tương ứng. Trái lại, mức biểu hiện mRNA CDKN2A ($\log_2(\text{raw STAR count} + 1)$) ở nhóm có đột biến cao hơn nhóm không có đột biến (trung vị: 11,1 so với 8,70; $p=3,73 \times 10^{-14}$). Tuy nhiên, phân tích mối liên quan giữa đột biến và biểu hiện trong nhóm CNV bình thường cũng như sau khi loại trừ các trường hợp mất đoạn sâu, và hiệu chỉnh theo trạng thái CNV trong mô hình hồi quy, kết quả cho thấy trạng thái CNV gây nhiễu ở mức độ khác nhau giữa các phân nhóm mô học. Trong toàn bộ cohort và ở LUAD, mối liên quan giữa đột biến và biểu hiện CDKN2A suy giảm sau khi kiểm soát CNV nhưng vẫn được duy trì. Ngược lại, ở LUSC, mối liên quan này không còn ý nghĩa thống kê sau khi loại trừ ảnh hưởng của mất đoạn sâu hoặc giới hạn phân tích trong nhóm CNV bình thường.

Trong khi đó, methyl hóa promoter mức cao liên quan đến giảm nhẹ biểu hiện CDKN2A so với nhóm methyl hóa thấp (trung vị: 8,95 so với 9,13; chênh lệch trung vị -0,19; $p=0,00354$). Phân tích methyl hóa dưới dạng biến liên tục không ghi nhận mối tương quan có ý nghĩa giữa điểm methyl hóa promoter và biểu hiện CDKN2A (Spearman $\rho=-0,020$; $p=0,556$).

Đối với biến đổi số lượng bản sao, mức biểu hiện CDKN2A khác biệt giữa các trạng thái CNV (Hình 2). Trong đó, nhóm mất đoạn sâu có mức biểu

hiện thấp nhất, tiếp theo là nhóm mất đoạn nông, trong khi các nhóm tăng bản sao và khuếch đại không ghi nhận xu hướng giảm biểu hiện tương tự.

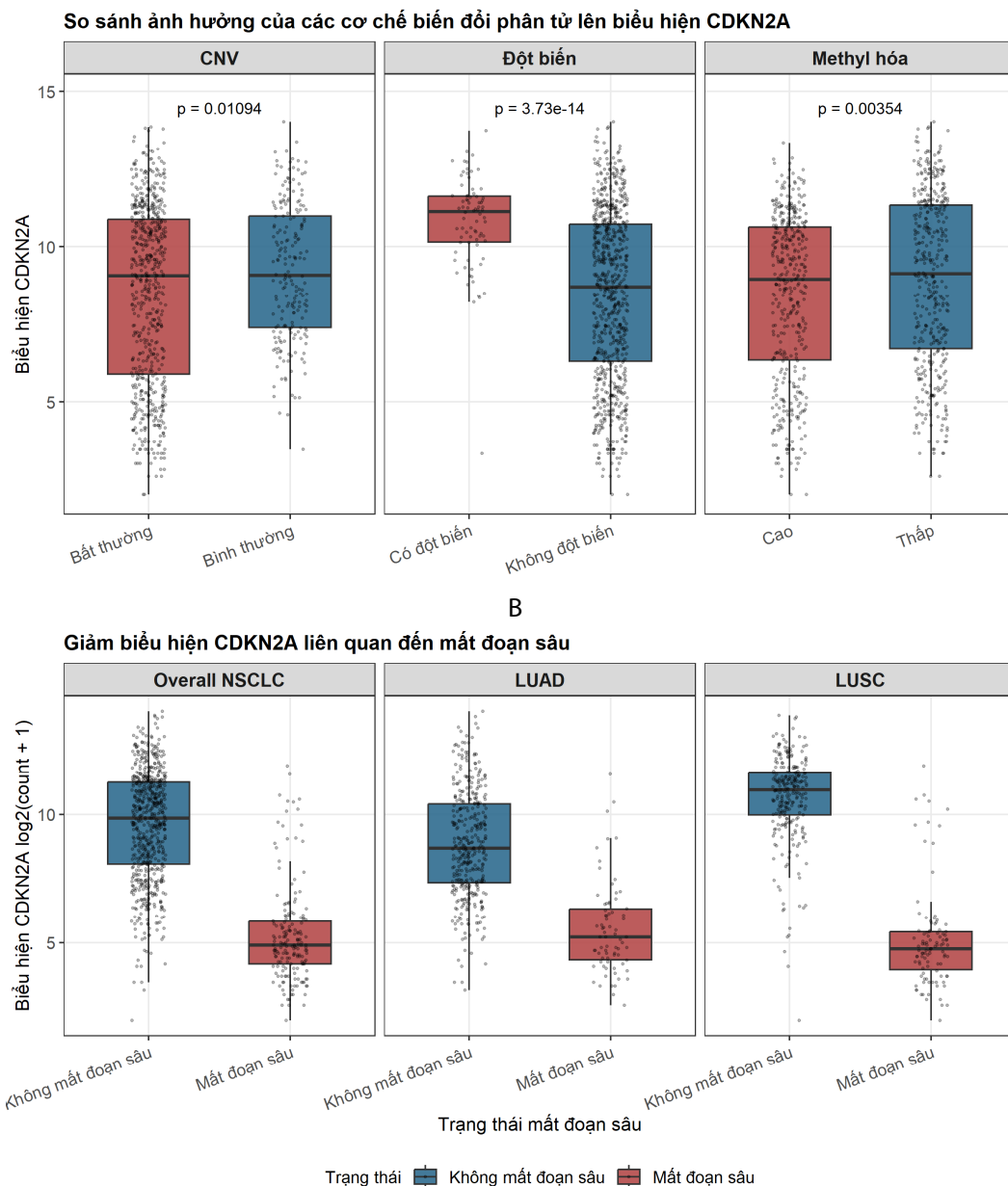


Hình 1. Mối liên quan giữa các biến đổi phân tử của CDKN2A và mức biểu hiện gene trong NSCLC

Biểu đồ hộp so sánh mức biểu hiện mRNA CDKN2A giữa các nhóm theo ba cơ chế biến đổi phân tử. Ở panel CNV, nhóm Bất thường bao gồm các trường hợp có trạng thái CNV khác bình thường (diploid), gồm mất đoạn sâu, mất đoạn nông, tăng bản sao và khuếch đại ($n = 623$); nhóm Bình thường gồm các trường hợp diploid ($n = 202$). Ở panel đột biến, nhóm Có đột biến gồm các trường hợp có ít nhất một biến thể CDKN2A trong dữ liệu MAF ($n = 78$), so với nhóm Không đột biến ($n = 753$). Ở panel methyl hóa, nhóm Cao và Thấp được phân nhóm theo giá trị trung vị của điểm methyl hóa promoter CDKN2A ($n = 413$ và $n = 418$). So sánh giữa hai nhóm được thực hiện bằng kiểm định Wilcoxon rank-sum.

3.3. Phân tích biểu hiện CDKN2A theo từng trạng thái CNV

Trong số 831 mẫu NSCLC nguyên phát của TCGA, 825 mẫu có đầy đủ dữ liệu CNV/GISTIC và biểu hiện mRNA của CDKN2A được đưa vào phân tích. Mức biểu hiện CDKN2A được so sánh giữa các trạng thái CNV khác nhau. Kết quả cho thấy mất đoạn sâu có mối liên quan mạnh nhất với giảm biểu hiện CDKN2A (Hình 2).



Hình 2. Ảnh hưởng của biến đổi số lượng bản sao đến biểu hiện CDKN2A trong NSCLC

A. Biểu hiện CDKN2A theo các trạng thái biến đổi số lượng bản sao. Biểu đồ hộp so sánh biểu hiện mRNA CDKN2A theo năm trạng thái biến đổi số lượng bản sao (CNV) của CDKN2A, bao gồm: mất đoạn sâu (GISTIC = -2; n = 177), mất đoạn nông (GISTIC = -1; n = 369), bình thường (diploid) (GISTIC = 0; n = 202), tăng bản sao (GISTIC = 1; n = 71) và khuếch đại (GISTIC = 2; n = 6). Sự khác biệt về mức biểu hiện giữa các nhóm được đánh giá bằng kiểm định Kruskal-Wallis ($p = 2,29 \times 10^{-66}$).

B. Mối liên quan giữa mất đoạn sâu và biểu hiện CDKN2A. Biểu đồ hộp so sánh mức biểu hiện mRNA CDKN2A giữa nhóm không mất đoạn sâu và nhóm mất đoạn sâu trong toàn bộ quần thể NSCLC và theo từng phân nhóm mô học. Mất đoạn sâu được xác định bởi giá trị GISTIC = -2; nhóm không mất đoạn sâu bao gồm các trường hợp có dữ liệu CNV nhưng không thuộc nhóm mất đoạn sâu. Cỡ mẫu gồm: toàn bộ NSCLC, không mất đoạn sâu (n = 648) và mất đoạn sâu (n = 177); LUAD, không mất đoạn sâu (n = 382) và mất đoạn sâu (n = 74); LUSC, không mất đoạn sâu (n = 266) và mất đoạn sâu (n = 103). So sánh giữa hai nhóm được thực hiện bằng kiểm định Wilcoxon rank-sum.

Trong toàn bộ cohort, nhóm bệnh nhân mang mất đoạn sâu có mức biểu hiện CDKN2A thấp hơn rõ rệt so với nhóm không có mất đoạn sâu (trung vị: 4,91 so với 9,86; $p < 0,001$). Xu hướng này được ghi nhận nhất quán ở cả LUAD và LUSC. Đáng chú ý, LUSC có mức biểu hiện mRNA CDKN2A nền cao hơn đáng kể so với LUAD (trung vị: 10,32 so với 8,29; $p < 0,001$). Tuy nhiên, chênh lệch biểu hiện giữa nhóm có và không có mất đoạn sâu vẫn lớn hơn ở LUSC so với LUAD (-6,20 so với -3,48).

Khi phân tích theo các trạng thái CNV cụ thể, nhóm mất đoạn sâu có mức biểu hiện CDKN2A thấp nhất trong tất cả các nhóm biến đổi số lượng bản sao. Ngược lại, các trường hợp tăng và khuếch đại gen lại có xu hướng biểu hiện CDKN2A cao hơn. Phân tích Kruskal–Wallis cho thấy sự khác biệt mức biểu hiện giữa các nhóm CNV có ý nghĩa thống kê ($p = 2,29 \times 10^{-66}$).

IV. BÀN LUẬN

Trong phân tích tích hợp đa hệ dữ liệu của chúng tôi trên 831 bệnh nhân ung thư phổi không tế bào nhỏ từ cơ sở dữ liệu TCGA, CDKN2A được ghi nhận chịu ảnh hưởng bởi nhiều cơ chế biến đổi phân tử khác nhau. Trong số các biến đổi được khảo sát, mất đoạn sâu cho thấy mối liên quan mạnh nhất với giảm biểu hiện mRNA CDKN2A. Đáng chú ý, mặc dù LUSC có mức biểu hiện mRNA CDKN2A nền cao hơn LUAD, mức giảm biểu hiện liên quan đến mất đoạn sâu vẫn lớn hơn ở LUSC. Phân tích này được bổ sung nhằm loại trừ khả năng sự khác biệt quan sát được chỉ phản ánh khác biệt về mức biểu hiện nền giữa hai phân nhóm mô học. Kết quả này phù hợp với các nghiên cứu trước đây cho thấy mất đoạn tại locus 9p21, đặc biệt mất đoạn đồng hợp tử CDKN2A, là một trong những cơ chế điều hòa quan trọng của gene này trong NSCLC¹³

Về mặt sinh học, CDKN2A nằm trên vùng nhiễm sắc thể 9p21.3, một trong những vùng mất đoạn tái diễn phổ biến ở nhiều loại ung thư, bao gồm cả ung thư phổi không tế bào nhỏ. Các nghiên cứu phân tích biến đổi số lượng bản sao trên quy mô lớn cho thấy những vùng mất đoạn tái diễn này thường chứa các gen ức chế khối u quan trọng và được chọn

lọc trong quá trình hình thành, tiến triển của ung thư^{10,14,15}. Trong bối cảnh đó, mất đoạn sâu làm mất cả hai bản sao của CDKN2A, dẫn đến giảm số lượng bản sao của gen và từ đó có thể làm giảm đáng kể lượng mRNA được phiên mã. Ngược lại, mất đoạn nông chỉ làm mất một bản sao của gen nên thường chỉ gây ảnh hưởng một phần hoặc không đồng nhất đến mức biểu hiện. Cơ chế này phù hợp với kết quả của nghiên cứu hiện tại, trong đó mất đoạn sâu là biến đổi có mối liên quan mạnh nhất với mức biểu hiện mRNA CDKN2A thấp. Kết quả này cũng tương đồng với nghiên cứu của TCGA trên ung thư biểu mô vảy phổi, trong đó biểu hiện CDKN2A được đánh giá trực tiếp theo trạng thái số lượng bản sao của gen và giảm rõ rệt ở nhóm mất đoạn đồng hợp tử¹³.

Một trong những phát hiện đáng chú ý của nghiên cứu là tần suất biến đổi CDKN2A cao hơn rõ rệt ở LUSC so với LUAD. Kết quả này phù hợp với các nghiên cứu hệ gene trước đây cho thấy vùng 9p21.3 chứa CDKN2A là một trong những locus bị mất đoạn thường gặp trong ung thư phổi, đặc biệt ở LUSC. Nhiều nghiên cứu TCGA cũng ghi nhận các biến đổi số lượng bản sao tại vùng này xuất hiện với tần suất cao hơn ở ung thư biểu mô vảy so với ung thư biểu mô tuyến. Những dữ liệu này gợi ý rằng các biến đổi cấu trúc hệ gene, đặc biệt là mất đoạn nhiễm sắc thể, có thể đóng vai trò nổi bật hơn trong cơ chế điều hòa biểu hiện CDKN2A ở LUSC so với LUAD^{13,16,17}.

Trái với nghiên cứu của Tam và cộng sự, nhóm bệnh nhân mang đột biến CDKN2A trong nghiên cứu của chúng tôi không ghi nhận giảm biểu hiện mRNA tương ứng, thậm chí có mức biểu hiện CDKN2A cao hơn nhóm không có đột biến. Để làm rõ hiện tượng này, chúng tôi tiếp tục đánh giá mối liên quan giữa đột biến CDKN2A và trạng thái CNV. Kết quả cho thấy đột biến CDKN2A phân bố không đồng đều giữa các trạng thái CNV, với tần suất thấp nhất ở nhóm mất đoạn sâu. Sau khi kiểm soát ảnh hưởng của CNV bằng các phân tích độ nhạy, mức độ khác biệt về biểu hiện CDKN2A giữa nhóm có và không có đột biến đều giảm. Ở toàn bộ cohort, mối liên quan này vẫn được duy trì sau khi hiệu chỉnh; tuy nhiên, ở LUSC, mối liên quan không còn ý nghĩa

thống kê sau khi kiểm soát ảnh hưởng của CNV. Những kết quả này cho thấy trạng thái CNV là một yếu tố gây nhiễu quan trọng đối với mối liên quan giữa đột biến CDKN2A và biểu hiện mRNA, đặc biệt trong phân nhóm LUSC. Điều này có thể góp phần giải thích sự khác biệt giữa kết quả của chúng tôi và Tam và cộng sự. Trong khi Tam và cộng sự đánh giá đồng thời trạng thái bất hoạt CDKN2A với cả biểu hiện mRNA và protein trên dòng tế bào kết hợp với mô bệnh phẩm¹², nghiên cứu hiện tại sử dụng dữ liệu biểu hiện gene từ RNA-seq ở mức gene trên mô khối u TCGA và không đánh giá biểu hiện protein. Do đó, mức biểu hiện mRNA quan sát được có thể đồng thời chịu ảnh hưởng của trạng thái CNV và các biến đổi phân tử đồng diễn ra.

Trong nghiên cứu của chúng tôi, nhóm có mức methyl hóa promoter cao có biểu hiện CDKN2A thấp hơn so với nhóm methyl hóa thấp, mặc dù mức độ khác biệt tương đối nhỏ. Tuy nhiên, khi methyl hóa promoter được phân tích dưới dạng biến liên tục, không ghi nhận mối tương quan có ý nghĩa với biểu hiện CDKN2A. Kết quả này cho thấy mối liên quan giữa methyl hóa promoter và biểu hiện CDKN2A không đồng nhất trong toàn bộ cohort NSCLC. Nhận định này phù hợp với các nghiên cứu trước đây, trong đó methyl hóa promoter được xác định là một cơ chế bất hoạt biểu sinh của CDKN2A ở ung thư phổi, nhưng đồng thời cũng cho thấy vai trò của methyl hóa phụ thuộc vào bối cảnh phân tử của từng khối u và chỉ là một trong nhiều cơ chế điều hòa biểu hiện CDKN2A^{11-13,18-20}.

Phân tích ban đầu cho thấy biểu hiện CDKN2A có xu hướng giảm theo số lượng biến đổi phân tử đồng thời. Tuy nhiên, phân tích độ nhạy sau khi loại các trường hợp mất đoạn sâu cho thấy xu hướng này không còn ý nghĩa thống kê, cho thấy kết quả ban đầu chủ yếu phản ánh ảnh hưởng nổi trội của mất đoạn sâu.

Nhìn chung, kết quả nghiên cứu cho thấy các biến đổi phân tử của CDKN2A có mối liên quan khác nhau với mức biểu hiện mRNA trong NSCLC, trong đó mất đoạn sâu là biến đổi có mối liên quan mạnh nhất với giảm biểu hiện CDKN2A. Mối liên quan này rõ rệt hơn ở LUSC so với LUAD, đồng thời các biến

đổi của CDKN2A cũng xuất hiện với tần suất cao hơn ở LUSC.

Hạn chế của nghiên cứu

Nghiên cứu còn một số hạn chế. Thứ nhất, do sử dụng dữ liệu TCGA với thiết kế cắt ngang, nghiên cứu chỉ đánh giá mối liên quan giữa các biến đổi phân tử và biểu hiện mRNA CDKN2A mà chưa làm rõ cơ chế chức năng ở mức protein. Thứ hai, phân tích được thực hiện ở mức gene và chưa đánh giá riêng các transcript hoặc isoform của CDKN2A. Ngoài ra, nghiên cứu chưa đánh giá ảnh hưởng của độ tinh khiết khối u cũng như mối liên quan với các đặc điểm lâm sàng và tiên lượng. Những hạn chế này đồng thời là định hướng cho các nghiên cứu tiếp theo của chúng tôi nhằm làm sáng tỏ hơn ý nghĩa sinh học và giá trị ứng dụng của các biến đổi CDKN2A trong ung thư phổi không tế bào nhỏ.

V. KẾT LUẬN

Các biến đổi phân tử của CDKN2A có mối liên quan khác nhau với mức biểu hiện mRNA trong NSCLC. Trong đó, mất đoạn sâu là biến đổi có mối liên quan mạnh nhất với giảm biểu hiện CDKN2A và rõ rệt hơn ở LUSC so với LUAD.

TÀI LIỆU THAM KHẢO

1. Bray F, Laversanne M, Sung H et al (2024) *Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA Cancer J Clin 74(3): 229-263. doi:10.3322/caac.21834
2. Herbst RS, Morgensztern D, Boshoff C (2018) The biology and management of non-small cell lung cancer. Nature 553(7689): 446-454. doi:10.1038/nature25183
3. Travis WD, Brambilla E, Nicholson AG et al (2015) *The 2015 world health organization classification of lung tumors*. J Thorac Oncol 10(9): 1243-1260. doi:10.1097/JTO.0000000000000630
4. Zhao L, Zhou X, Li H, Yin T, Jiang Y (2024) *Prognosis of immunotherapy for non-small cell lung cancer with CDKN2A loss of function*. J Thorac Dis 16(1): 507-515. doi:10.21037/jtd-23-1017

5. Deng C, Li Z xi, Xie C jun et al (2024) *Pan-cancer analysis of CDKN2A alterations identifies a subset of gastric cancer with a cold tumor immune microenvironment*. Hum Genomics 18(1): 55. doi:10.1186/s40246-024-00615-7
6. Feng X, Chang H, Xu Y et al (2025) *CDKN2A alternations in non-small cell lung cancer: An evaluation using next-generation sequencing*. Pathol - Res Pract 274: 156177. doi:10.1016/j.prp.2025.156177
7. Soltan MA, Alhanshani AA, Shati AA et al (2023) *Cyclin dependent kinase inhibitor 2a genetic and epigenetic alterations interfere with several immune components and predict poor clinical outcome*. Biomedicines. 11(8): 2254. doi:10.3390/biomedicines11082254
8. Nobori T, Miura K, Wu DJ, Lois A, Takabayashi K, Carson DA (1994) *Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers*. Nature 368(6473): 753-756. doi:10.1038/368753a0
9. Merlo A, Herman JG, Mao L, et al (1995) *5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers*. Nat Med 1(7): 686-692. doi:10.1038/nm0795-686
10. Beroukhi R, Mermel CH, Porter D et al (2010) *The landscape of somatic copy-number alteration across human cancers*. Nature. 2010;463(7283):899-905. doi:10.1038/nature08822
11. The Cancer Genome Atlas Research Network (2014) *Comprehensive molecular profiling of lung adenocarcinoma*. Nature 511(7511): 543-550. doi:10.1038/nature13385
12. Tam KW, Zhang W, Soh J et al (2013) *CDKN2A/p16 Inactivation mechanisms and their relationship to smoke exposure and molecular features in non-small-cell lung cancer*. J Thorac Oncol 8(11): 1378-1388. doi:10.1097/JTO.0b013e3182a46c0c
13. The Cancer Genome Atlas Research Network (2012) *Comprehensive genomic characterization of squamous cell lung cancers*. Nature 489(7417): 519-525. doi:10.1038/nature11404
14. Zack TI, Schumacher SE, Carter SL et al (2015) *Pan-cancer patterns of somatic copy number alteration*. Nat Genet 45(10): 1134-1140. doi:10.1038/ng.2760
15. Kim TM, Xi R, Luquette LJ, Park RW, Johnson MD, Park PJ (2013) *Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes*. Genome Res 23(2): 217-227. doi:10.1101/gr.140301.112
16. Tuna M, Mills GB, Amos CI (2023) *Common and distinct patterns of acquired uniparental disomy and homozygous deletions between lung squamous cell carcinomas and lung adenocarcinoma*. Neoplasia 45: 100932. doi:10.1016/j.neo.2023.100932
17. Gonçalves PG, Reis RM, Bidinotto LT (2022) *Significance of Chr9p22.1-p21.3 Deletion in Cancer Development: A Pan-cancer In Silico Analysis*. Anticancer Res 42(11): 5291-5304. doi:10.21873/anticancer.16036
18. Merlo A, Herman JG, Mao L et al (1995) *5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers*. Nat Med 1(7): 686-692. doi:10.1038/nm0795-686
19. Belinsky SA, Nikula KJ, Palmisano WA et al (1998) *Aberrant methylation of p16^{INK4a} is an early event in lung cancer and a potential biomarker for early diagnosis*. Proc Natl Acad Sci 95(20): 11891-11896. doi:10.1073/pnas.95.20.11891
20. Zöchbauer-Müller S, Fong KM, Virmani AK, Geradts J, Gazdar AF, Minna JD (2001) *Aberrant promoter methylation of multiple genes in non-small cell lung cancers*. Cancer Res 61(1): 249-255.