

CHARACTERIZATION OF THE MUTATIONAL SPECTRUM AND ITS PROGNOSTIC SIGNIFICANCE IN PEDIATRIC ACUTE LEUKEMIA

Giang Bui Thi Huong¹, Loc Thi Pham¹

¹Hai Phong University of Medicine and Pharmacy-No. 72A Nguyen Binh Khiem, Gia Vien Ward, Hai Phong, Vietnam

Received: 05/11/2025

Revised: 06/01/2026; Accepted: 04/06/2026

ABSTRACT

Objective: To determine the frequency of gene mutations and their association with overall survival in pediatric acute leukemia (AL), including T-ALL, B-ALL, and AML.

Materials and Methods: A retrospective secondary analysis of 763 pediatric patients with T-ALL (n=265), B-ALL (n=251), and AML (n=247) was performed using publicly available, de-identified data from St. Jude Cloud (GenomePaint/PeCan). Gene events (SNVs, indels, structural rearrangements, pathogenic fusions) were aggregated. Overall survival was analyzed using Kaplan–Meier and log-rank tests.

Results: High-frequency mutations (>10%) included RUNX1 (46.4%), NRAS (45.9%), NOTCH1 (41.8%), ETV6 (37.9%), KRAS (30.9%), FLT3 (22.0%), MLLT3 (18.2%), WT1 (13.9%), KDM6A (13.0%), BCR (12.6%), TP53 (12.2%), and KIT (11.3%). RUNX1, NRAS, ETV6, and KRAS predominated in B-ALL and AML, whereas NOTCH1 was almost exclusive to T-ALL. Intermediate-frequency mutations (5–10%) included CD56, CBL, EZH2, FOXP1, TPK1, MSH2, MRD5, and KDM4C. NRAS mutation in B-ALL was significantly associated with poorer overall survival ($p = 0.005$); KRAS mutations showed no significant prognostic impact.

Discussion and Conclusion: The mutational landscape varies markedly across AL subtypes and has potential clinical value for risk stratification and personalized treatment.

Keywords: Acute leukemia; gene mutation; prognosis; T-ALL; B-ALL; AML.

*Corresponding author

Email: ptloc@hpmu.edu.vn Phone: (+84) 943839666 DOI: 10.52163/yhc.v67iCD7.5586

NGHIÊN CỨU ĐẶC ĐIỂM CÁC ĐỘT BIẾN GEN VÀ Ý NGHĨA TIỀN LƯỢNG TRONG LƠ XÊ MI CẤP Ở TRẺ EM

Bùi Thị Hương Giang¹, Phạm Thị Lộc¹

¹Trường Đại học Y Dược Hải Phòng-Số 72A Nguyễn Bình Khiêm, phường Gia Viên, Hải Phòng, Việt Nam

Ngày nhận bài: 05/11/2025

Ngày chỉnh sửa: 06/01/2026; Ngày duyệt đăng: 04/06/2026

TÓM TẮT

Mục tiêu: Xác định tần suất đột biến gen và mối liên quan với tiên lượng sống còn ở bệnh nhi lơ xê mi cấp (LXMc), bao gồm T-ALL, B-ALL và AML.

Đối tượng và phương pháp: Phân tích thứ cấp trên 763 bệnh nhi LXMc gồm T-ALL (n=265), B-ALL (n=251) và AML (n=247), khai thác từ bộ dữ liệu công khai đã ẩn danh của St. Jude Cloud (nền tảng GenomePaint/PeCan). Các biến cố di truyền bao gồm SNV, indel, tái sắp xếp cấu trúc và tổ hợp gen có ý nghĩa bệnh sinh. Sống còn toàn bộ được phân tích bằng Kaplan–Meier và kiểm định log-rank.

Kết quả: Các gen tần suất cao (>10%): RUNX1 (46,4%), NRAS (45,9%), NOTCH1 (41,8%), ETV6 (37,9%), KRAS (30,9%), FLT3 (22,0%), MLLT3 (18,2%), WT1 (13,9%), KDM6A (13,0%), BCR (12,6%), TP53 (12,2%) và KIT (11,3%). RUNX1, NRAS, ETV6 và KRAS chủ yếu ở B-ALL và AML; NOTCH1 ưu thế rõ rệt ở T-ALL. Các gen tần suất trung bình (5–10%) gồm CD56, CBL, EZH2, FOXP1, TPK1, MSH2, MRD5 và KDM4C. Đột biến NRAS ở B-ALL liên quan tiên lượng xấu có ý nghĩa thống kê ($p = 0,005$); KRAS không cho thấy ý nghĩa tiên lượng.

Bàn luận và kết luận: Phổ đột biến gen khác biệt rõ giữa các thể LXMc và có giá trị tiên lượng, hỗ trợ phân tầng nguy cơ và điều trị cá thể hóa.

Từ khóa: Lơ xê mi cấp; đột biến gen; tiên lượng; T-ALL; B-ALL; AML.

1. ĐẶT VẤN ĐỀ

Lơ xê mi cấp (LXMc) là nhóm bệnh ác tính của hệ tạo máu, gồm LXMc dòng lympho (T-ALL, B-ALL) và LXMc dòng tủy (AML); đây là bệnh ác tính thường gặp nhất ở trẻ em, có diễn tiến nhanh và tiên lượng phụ thuộc nhiều yếu tố lâm sàng, miễn dịch và di truyền. Sự phát triển của giải trình tự gen thế hệ mới đã giúp xác định nhiều đột biến quan trọng như RUNX1, NRAS, NOTCH1, ETV6, KRAS và FLT3, có vai trò trong sinh bệnh học, tiên lượng và định hướng điều trị cá thể hóa. Tuy nhiên, tần suất và ý nghĩa tiên lượng của các đột biến này khác nhau giữa các thể bệnh và nhiều nghiên cứu hiện nay còn giới hạn ở từng nhóm bệnh hoặc cỡ mẫu nhỏ. Trên cơ sở dữ liệu giải trình tự quy mô lớn của St. Jude Children's Research Hospital, chúng tôi phân tích đa thể LXMc ở trẻ em nhằm đánh giá phổ đột biến gen, tần suất và mối liên quan với tiên lượng sống còn, góp phần hỗ trợ phân tầng nguy cơ và cá thể hóa điều trị.

2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng và thiết kế nghiên cứu

Nghiên cứu hồi cứu, mô tả, phân tích thứ cấp trên 763 bệnh nhi LXMc, gồm T-ALL (n=265), B-ALL (n=251) và AML (n=247), khai thác từ bộ dữ liệu công khai của St. Jude Cloud.

2.2. Xác định đột biến gen và biến số nghiên cứu

Dữ liệu biến thể được khai thác từ nền tảng GenomePaint/PeCan của St. Jude Cloud [1], tổng hợp các kết quả WGS và

WES đã được chú giải bởi nhóm tin sinh học của cơ sở nguồn. Một bệnh nhân được coi là mang “đột biến” của một gen khi có ít nhất một biến thể có ý nghĩa bệnh sinh tại gen đó, bao gồm: biến thể điểm (SNV), biến thể chèn/mất nhỏ (indel), tái sắp xếp cấu trúc hoặc tổ hợp gen pathogenic. Các biến thể dòng tế bào mầm (germline), biến thể đồng nghĩa và biến thể có ý nghĩa lâm sàng chưa rõ (VUS) được loại trừ theo bộ lọc mặc định của nền tảng nguồn. Cần lưu ý đây là các biến cố di truyền ở mức các sự kiện phân tử, không phải mức biểu hiện gen.

Các biến nghiên cứu gồm: loại đột biến gen, tần suất đột biến, phân nhóm bệnh và tình trạng sống còn. Đột biến được phân nhóm theo tần suất: >10% (tần suất cao) và 5–10% (tần suất trung bình).

2.3. Xử lý số liệu và phương pháp thống kê

So sánh tỷ lệ giữa các nhóm sử dụng kiểm định Chi-square hoặc Fisher's exact test khi thích hợp. Phân tích sống còn được thực hiện bằng phương pháp Kaplan–Meier và so sánh bằng kiểm định log-rank. Biến cố sống còn là sống còn toàn bộ (overall survival, OS) với sự kiện được mã hóa là tử vong do mọi nguyên nhân; bệnh nhân còn sống tại lần cập nhật cuối được coi là dữ liệu kiểm duyệt (censored). Thời gian theo dõi được tính từ thời điểm chẩn đoán đến lần cập nhật cuối theo cơ sở dữ liệu nguồn. Đối với từng gen, bệnh nhân được phân thành hai nhóm có và không có đột biến.

*Tác giả liên hệ

Email: ptloc@hpmu.edu.vn Điện thoại: (+84) 943839666 DOI: 10.52163/yhc.v67iCD7.5586

Mức ảnh hưởng tiên lượng của từng gen trên từng thể bệnh (Bảng 3) được phân loại định tính: “xấu” khi $p < 0,05$ và đường cong Kaplan–Meier của nhóm mang đột biến nằm thấp hơn; “tốt” khi $p < 0,05$ và đường cong nhóm mang đột biến nằm cao hơn; “trung bình” khi $p \geq 0,05$ hoặc số sự kiện không đủ để kết luận. Đây là phân loại sơ bộ, cần kiểm chứng bằng mô hình hồi quy Cox đa biến với báo cáo HR và 95% CI. Giá trị $p < 0,05$ được coi là có ý nghĩa thống kê.

2.4. Đạo đức nghiên cứu

Đây là nghiên cứu phân tích thứ cấp, không can thiệp, trên bộ dữ liệu đã ẩn danh hoàn toàn của Saint Jude Children’s Research Hospital (St. Jude Cloud); đồng thuận và phê duyệt đạo đức cho việc tạo lập bộ dữ liệu gốc đã được cơ sở nguồn thực hiện theo điều khoản sử dụng dữ liệu công khai. Việc sử dụng dữ liệu thứ cấp cho nghiên cứu này đã được Hội đồng Đạo đức Trường Đại học Y Dược Hải Phòng phê duyệt với mã số 14 / IRB_HPMU, phạm vi phê duyệt là phân tích dữ liệu thứ cấp đã ẩn danh. Do dữ liệu hoàn toàn ẩn danh và không có sự can thiệp trực tiếp đến bệnh nhân, nghiên cứu không tạo thêm nguy cơ nào đối với đối tượng tham gia.

3. KẾT QUẢ

3.1. Tần suất các đột biến gen

Bảng 1. Đặc điểm các gen đột biến tần suất cao ($>10\%$) trong LXMic (n=763)

STT	Tên gen	T-ALL (n=265)	B-ALL (n=251)	AML (n=247)	Tổng (n=763)	Tỷ lệ (%)
1	RUNX1	33	212	109	354	46,4
2	NRAS	35	206	109	350	45,9
3	NOTCH1	264	47	8	319	41,8
4	ETV6	25	237	27	289	37,9
5	KRAS	12	194	30	236	30,9
6	FLT3	21	94	53	168	22,0
7	MLLT3	24	81	34	139	18,2
8	WT1	49	14	43	106	13,9
9	KDM6A	27	59	13	99	13,0
10	BCR	1	92	3	96	12,6
11	TP53	11	79	3	93	12,2
12	KIT	4	10	72	86	11,3

Nhận xét: Các gen có tần suất cao nhất là RUNX1 (46,4%), NRAS (45,9%) và NOTCH1 (41,8%). RUNX1, NRAS, ETV6 và KRAS phân bố chủ yếu ở B-ALL và AML, trong khi NOTCH1 ưu thế rõ rệt ở T-ALL (264/265, 99,6%). FLT3 (22,0%), MLLT3 (18,2%) và WT1 (13,9%) cũng có tần suất đáng kể, chủ yếu ở B-ALL và AML.

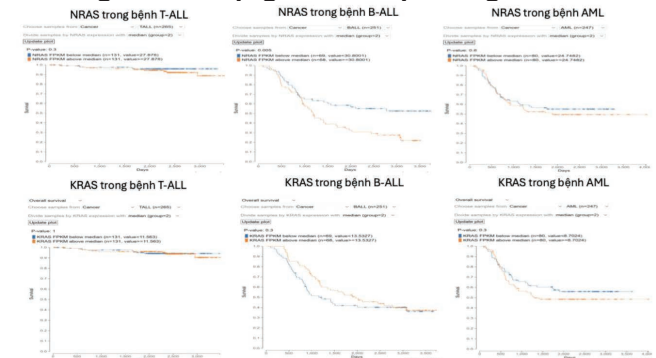
Bảng 2. Các gen đột biến tần suất trung bình (5–10%) trong LXMic (n=763)

STT	Tên gen	T-ALL (n=265)	B-ALL (n=251)	AML (n=247)	Tổng (n)	Tỷ lệ (%)
1	CD56	1	62	12	75	9,8
2	CBL	3	37	25	65	8,5
3	EZH2	24	22	14	60	7,9
4	FOXP1	6	41	13	60	7,9

STT	Tên gen	T-ALL (n=265)	B-ALL (n=251)	AML (n=247)	Tổng (n)	Tỷ lệ (%)
5	TPK1	5	39	7	51	6,7
6	MSH2	3	38	9	50	6,6
7	MRD5	4	33	6	43	5,6
8	KDM4C	7	23	11	41	5,4

Nhận xét: Các gen nhóm tần suất trung bình dao động 5,4–9,8%, trong đó CD56 (9,8%) và CBL (8,5%) cao nhất. Phần lớn các đột biến trong nhóm này phân bố ưu thế ở B-ALL, ngoại trừ EZH2 phân bố tương đối đồng đều giữa ba thể bệnh.

3.2. Ý nghĩa tiên lượng của các đột biến gen



Hình 1. Kết quả tỉ lệ sống còn của đột biến gen NRAS, KRAS trên từng nhóm bệnh

Nhận xét: Hình Kaplan–Meier cho thấy sự khác biệt sống còn theo mức biểu hiện NRAS và KRAS giữa T-ALL, B-ALL và AML. NRAS ở B-ALL có ý nghĩa thống kê ($p = 0,005$) với nhóm biểu hiện cao liên quan tiên lượng xấu hơn rõ rệt, trong khi T-ALL và AML không có khác biệt. Ngược lại, KRAS không ghi nhận sự khác biệt sống còn có ý nghĩa thống kê ở cả ba nhóm bệnh.

Tổng hợp phân loại ảnh hưởng tiên lượng của các đột biến gen được trình bày ở Bảng 3.

Bảng 3. Phân loại ảnh hưởng tiên lượng của các đột biến gen theo thể bệnh

STT	Tên gen	T-ALL	B-ALL	AML
1	RUNX1	Trung bình	Xấu	Xấu
2	NRAS	Trung bình	Xấu	Trung bình
3	NOTCH1	Tốt	Trung bình	Trung bình
4	ETV6	Tốt	Tốt	Trung bình
5	KRAS	Trung bình	Trung bình	Trung bình
6	FLT3	Trung bình	Xấu	Xấu
7	MLLT3	Trung bình	Trung bình	Xấu
8	WT1	Trung bình	Trung bình	Xấu
9	KDM6A	Trung bình	Trung bình	Trung bình
10	BCR	Trung bình	Xấu	Trung bình
11	TP53	Trung bình	Trung bình	Xấu
12	KIT	Trung bình	Trung bình	Xấu
13	EZH2	Trung bình	Trung bình	Trung bình
14	CBL	Trung bình	Trung bình	Trung bình

STT	Tên gen	T-ALL	B-ALL	AML
15	CD56	Trung bình	Trung bình	Xấu
16	FOXP1	Trung bình	Trung bình	Trung bình
17	TPK1	Trung bình	Trung bình	Trung bình
18	MSH2	Trung bình	Trung bình	Trung bình
19	MRD5	Trung bình	Trung bình	Trung bình
20	KDM4C	Trung bình	Trung bình	Trung bình

Nhận xét: Trong AML, các đột biến RUNX1, FLT3, WT1, MLLT3, TP53, KIT và CD56 liên quan tiên lượng xấu. Ở B-ALL, RUNX1, NRAS, FLT3 và BCR liên quan tiên lượng xấu, trong khi ETV6 liên quan tiên lượng tốt. Trong T-ALL, NOTCH1 và ETV6 có xu hướng tiên lượng tốt; các đột biến còn lại chưa cho thấy ý nghĩa tiên lượng rõ rệt.

4. BÀN LUẬN

4.1. Phổ đột biến gen theo thể bệnh

Nghiên cứu ghi nhận phổ đột biến đa dạng trong LXMc trẻ em, với các gen có tần suất cao nhất là RUNX1 (46,4%), NRAS (45,9%) và NOTCH1 (41,8%) – đây là các gen quan trọng trong điều hòa biệt hóa, tăng sinh và chết theo chương trình của tế bào tạo máu. Kết quả phù hợp với các nghiên cứu lớn cho thấy đột biến RAS và yếu tố phiên mã (RUNX1, ETV6) chiếm tỷ lệ cao và có tính đặc hiệu theo thể LXMc [2,3]. Về phân bố theo thể bệnh, RUNX1, NRAS, ETV6 và KRAS chủ yếu gặp ở B-ALL và AML, trong khi NOTCH1 ưu thế rõ rệt ở T-ALL (264/265 ca, 99,6%) – phù hợp với vai trò chủ chốt của NOTCH1 trong T-ALL [4] và vai trò của RUNX1, ETV6 trong rối loạn biệt hóa dòng B và dòng tủy [5]. FLT3 (22,0%), MLLT3 (18,2%) và WT1 (13,9%) có tỷ lệ cao, phù hợp với các nghiên cứu của Kottaridis và cs, Döhner và cs trong đó FLT3 và WT1 được xác định là yếu tố tăng sinh mạnh và tiên lượng xấu trong AML [6,7].

Trong nhóm tần suất trung bình, các gen như CD56 (9,8%), CBL (8,5%), EZH2 (7,9%) và FOXP1 (7,9%) chiếm ưu thế và tập trung ở B-ALL, cho thấy vai trò của các cơ chế ngoại gen và tín hiệu nội bào trong sinh bệnh học LXMc dòng B. EZH2 (thành phần phức hợp PRC2) có vai trò methyl hóa histone, điều hòa biểu hiện gen và liên quan đến tiên lượng không đồng nhất [8]. Tương tự, CBL – một E3 ubiquitin ligase – tham gia điều hòa tín hiệu tyrosine kinase, và đột biến CBL làm tăng hoạt hóa đường dẫn tín hiệu tăng sinh trong LXMc [9].

4.2. Ý nghĩa tiên lượng

Phân tích Kaplan–Meier cho thấy NRAS ở B-ALL liên quan tiên lượng xấu có ý nghĩa thống kê ($p = 0,005$), phù hợp với các nghiên cứu trước về vai trò của đột biến RAS trong kháng hóa trị và tái phát. Ngược lại, KRAS không cho thấy ý nghĩa tiên lượng rõ ràng. Trong AML, các đột biến như RUNX1, FLT3, WT1, TP53 và KIT liên quan tiên lượng xấu, phù hợp với hệ thống phân tầng nguy cơ của European LeukemiaNet [7]. Trong T-ALL, NOTCH1 có xu hướng tiên lượng tốt, mặc dù còn phụ thuộc vào các đồng đột biến. Ở B-ALL, ETV6 liên quan tiên lượng tốt, trong khi BCR liên quan tiên lượng xấu, phù hợp với các phân loại nguy cơ hiện hành.

4.3. Hạn chế của nghiên cứu

Nghiên cứu có một số hạn chế. Thứ nhất, tần suất đột biến của một số gen (RUNX1, NRAS, NOTCH1) cao hơn các báo cáo cổ điển; điều này có thể phản ánh đặc điểm của cơ sở dữ liệu nguồn vốn bao gồm các trường hợp đã giải trình tự sâu, tiêu chí chú giải biến thể rộng (gộp SNV, indel, tái sắp xếp

cấu trúc và tổ hợp gen) và có thể bao gồm cả các bệnh nhân tái phát/kháng trị. Các giá trị tần suất nên được hiểu là tỷ lệ biến cố di truyền tích lũy ở mức bệnh nhân, không trực tiếp so sánh được với tần suất đột biến điểm tại thời điểm chẩn đoán trong các nghiên cứu đơn trung tâm. Thứ hai, do thiết kế phân tích thứ cấp, kết quả phụ thuộc vào quy trình chú giải mặc định của St. Jude Cloud. Thứ ba, phân loại tiên lượng theo “tốt/trung bình/xấu” là định tính, không bao gồm hiệu chỉnh các yếu tố nhiễu (tuổi, nhóm nguy cơ ban đầu, các đồng đột biến); cần kiểm chứng bằng mô hình hồi quy Cox đa biến với báo cáo HR và 95% CI. Thứ tư, với các đột biến hiếm hoặc khi phân tầng theo thể bệnh dẫn tới số sự kiện rất ít, ước lượng sống còn có thể thiếu ổn định.

5. KẾT LUẬN

Phổ đột biến gen trong LXMc trẻ em đa dạng và khác biệt rõ rệt giữa các thể bệnh, phản ánh vai trò đặc hiệu của từng nhóm gen trong cơ chế sinh bệnh. Sự kết hợp đồng thời của các đột biến tín hiệu (RAS), yếu tố phiên mã (RUNX1, ETV6, NOTCH1) và gen điều hòa epigenetic (EZH2, KDM6A) cho thấy LXMc có nền tảng di truyền phức tạp. Đột biến NRAS ở B-ALL liên quan tiên lượng xấu có ý nghĩa thống kê. Kết quả nghiên cứu nhấn mạnh tầm quan trọng của việc tích hợp hồ sơ đột biến gen vào phân tầng nguy cơ, góp phần định hướng điều trị cá thể hóa và cải thiện tiên lượng cho bệnh nhi LXMc.

6. LỜI CẢM ƠN

Chúng tôi xin bày tỏ lòng biết ơn sâu sắc tới Ủy ban Nhân dân Thành phố Hải Phòng, Sở Khoa học và Công nghệ Hải Phòng, Trường Đại học Y Dược Hải Phòng và các đơn vị liên quan đã hỗ trợ và tài trợ kinh phí cho đề tài nghiên cứu mã số 966/HĐ-ĐT.YD.2023.

7. TÀI LIỆU THAM KHẢO

- [1] Zhou X, Wang J, Patel J, et al. Exploration of coding and non-coding variants in cancer using GenomePaint. *Cancer Cell*. 2021;39(1):83–95.
- [2] Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016;374(23):2209–2221.
- [3] Inaba H, Greaves M, Mullighan CG. Acute lymphoblastic leukaemia. *Lancet*. 2013;381(9881):1943–1955.
- [4] Weng AP, Ferrando AA, Lee W, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science*. 2004;306(5694):269–271.
- [5] Osato M. Point mutations in the RUNX1/AML1 gene: another paradigm of how hematopoietic transcription factors are altered in leukemia. *Int J Hematol*. 2004;79(5):380–384.
- [6] Kottaridis PD, Gale RE, Frew ME, et al. The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia adds important prognostic information to cytogenetic risk groups. *Blood*. 2001;98(6):1752–1759.
- [7] Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424–447.
- [8] Ernst T, Chase AJ, Score J, et al. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet*. 2010;42(8):722–726.
- [9] Sanada M, Suzuki T, Shih LY, et al. Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature*. 2009;460(7257):904–908.