Evolutionary Analysis of Beta-Lactamases in WHO Priority Pathogens Identifies Conserved Motifs and Positions VIM as an Evolutionary Intermediate Between IMP and NDM

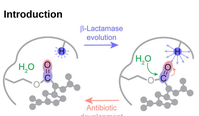
Ananya Anurag Anand, Vidushi Yadav, Kshitiz K, Sreedevi PC, Sarfraz Anwar, Sintu Kumar Samanta*
Department of Applied Sciences, Indian Institute of Information Technology Allahabad, Prayagraj 211012, Uttar Pradesh, India

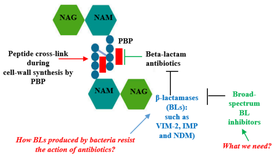
ISAAR 2026 Hong Kong

Abstract

This study presents the first integrated evolutionary analysis of chromosomal AmpC β-lactamases (ADC, ACT, PDC) and plasmid-encoded metallo-beta-lactamases (MBLs) (NDM, VIM, IMP) across three priority pathogens. PPI analysis identified NagZ as the central BL-interacting partner across Enterobacteriaceae (E) and *Pseudomonas aeruginosa* (P), highlighting its potential as a drug target. *A. baumannii* (A) was found to carry highest diversity ratio in PPI. Patristic distance analysis revealed progressive erosion of phylogenetic signal among MBLs correlating with evolutionary age — NDM retaining the strongest structure, VIM intermediate, and IMP the least — while chromosomally encoded proteins consistently showed a conserved inter-organism proximity pattern. Discrete trait migration analysis showed inter-host gene flow is greatest between phylogenetically distant lineages, suggesting selection-driven dissemination. Motif analysis identified a universally conserved FYPG(P/A)(G/A)H(T/S) element across all MBL families studied and the PGFGAVASNGLI motif showing cross-family similarity across all three MBLs. Targetting these motifs can inhibit all three MBLs using a single drug. FADE analysis revealed sites surrounding the active site are under high mutational pressure, providing actionable insights for broad-spectrum BL inhibitor design. VIM consistently showed intermediate characteristics between NDM and IMP across GC percentage, motif, PFAM and PCA analyses. Therefore, VIM could be understood as an evolutionary intermediate or bridge between NDM and IMP. As a result, most inhibitors targetting VIM could also target NDM and IMP.

Introduction





- *First motif-guided broad-spectrum MBL inhibitor design framework*
- *First quantitative co-evolutionary evidence for the AmpC and interacting partners*
- *First proposal of VIM as evolutionary bridge between NDM and IMP*
- *First MBL HGT directionality analysis across WHO Priority Pathogens*

Methodology

1. PPI profiles of BLs across P. aeruginosa, A. baumannii and Enterobacteriaceae retrieved from STRING database; quantitative diversity metrics computed
2. Organism-specific sequences retrieved via NCBI BLASTP; variant sequences of IMP (n=103), VIM (n=94), NDM (n=88), ACT (n=203), ADC (n=372) and PDC (n=676) retrieved from BLDB
3. MSA performed using MUSCLE algorithm in MEGA v12; maximum likelihood phylogenetic trees constructed with 100 bootstrap replicates
4. Patristic distances computed using ape and phangorn packages in R; inter-group proximity assessed via Wilcoxon rank-sum tests
5. Bayesian structured coalescent analysis (MASCOT, BEAST2 v2.7.7) performed to estimate preferential HGT directionality across three demes
6. GC content calculated using Biopython; descriptive statistics and ANOVA with Bonferroni post-hoc tests performed
7. AAC and PCP descriptors computed using Pfeature; PCA performed using scikit-learn
8. PFAM domain annotation performed using HMMER suite; co-occurrence visualised via UpSet plots
9. Conserved motifs identified using MEME; cross-family similarity assessed using TOMTOM
10. Catalytic motif conservation analysed via HMM profiling and Skylign conservation logos
11. Directional amino acid evolution identified using HyPhy FADE; significant sites mapped to VIM-2 crystal structure (PDB: 1KO3)

Results and Discussion

A. baumannii was found to carry highest diversity ratio in PPI.

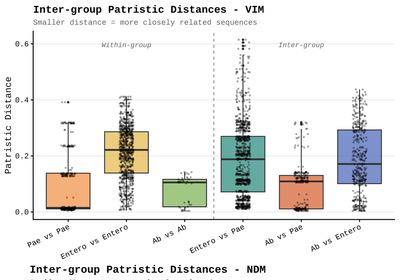
Table 1. Quantitative metrics of PPI interaction diversity analysis

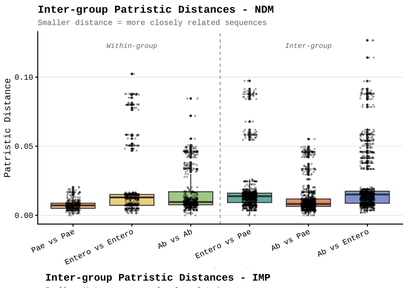
Metric	*P. aeruginosa*	*A. baumannii*	Enterobacteriaceae
No. of BLs	87	18	350
Total interactions	4,455	544	14,596
Mean interactions/BL	51.2	30.2	41.7
Total unique interactors	468	112	2,037
Mean unique interactors/BL	5.38	6.22	5.82
Interactors shared across ≥2 BLs (n)	115	15	347
Interactors shared across ≥2 BLs (%)	24.57%	13.39%	17.03%
Exclusive to 1 BL (n)	353	97	1690
Exclusive to 1 BL (%)	75.43%	86.61%	82.97%
Diversity Ratio (Total unique interactors / Total interactions)	0.105	0.206	0.14
Redundancy Index (Shared across ≥2 BLs / Total unique interactors)	0.246	0.134	0.17

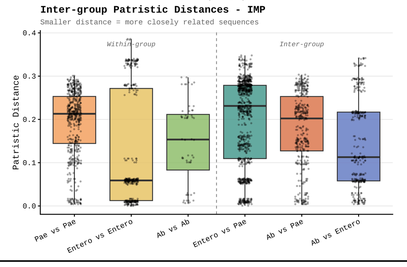
VIM and NDM follow evolutionary patterns of ancestral MBL fold proteins. IMP has diverged.
The co-evolution of AmpC-AmpR-AmpD and the repetition of species-level phylogeny by NagZ, indicates the highly targetable nature of NagZAmpR-AmpC- axis, with the conserved NagZ at the top.

Table 2. Inter-organism Evolutionary Proximity based on Patristic distance analysis

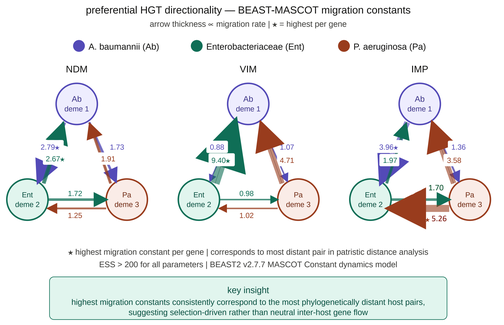
Protein	Age/Type	Closest inter-group	Most distant inter-group
None	**Species-level phylogeny**	**Ab-Pae**	**Ab–Entero / Pae–Entero**
AmpC	Chromosomal, ancient	Entero–Pae	**Ab–Entero**
AmpR	Regulatory, ancient	Entero–Pae	**Ab–Entero**
AmpD	Housekeeping, ancient	Entero–Pae	**Ab–Entero**
NagZ	Housekeeping, ancient	Ab–Pae	**Ab–Entero**
MBL-fold proteins*	Chromosomal	Ab-Pae/Entero-Pae	**Ab–Entero**
NDM	Mobile, **recent MBL**	Ab–Pae	**Ab–Entero**
VIM	Mobile, **intermediate MBL**	Ab–Pae	**Ab–Entero**
IMP	Mobile, **oldest MBL**	Ab–Entero	**Entero–Pae**







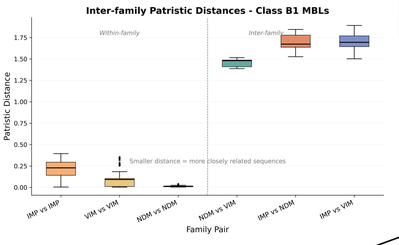
Gene flow is the highest between phylogenetically distant bacterial groups: driven by strong selective pressures rather than neutral diffusion

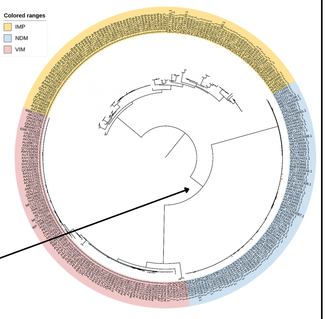


★ highest migration constant per gene | corresponds to most distant pair in patristic distance analysis
ESS > 200 for all parameters | BEAST2 v2.7.7 MASCOT Constant dynamics model

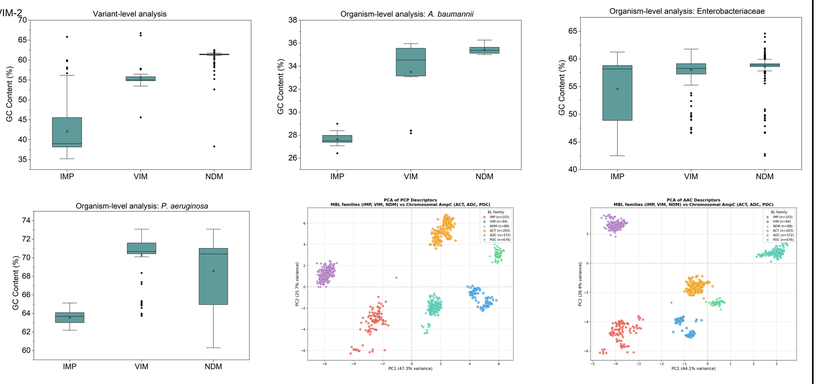
key insight
highest migration constants consistently correspond to the most phylogenetically distant host pairs, suggesting selection-driven rather than neutral inter-host gene flow

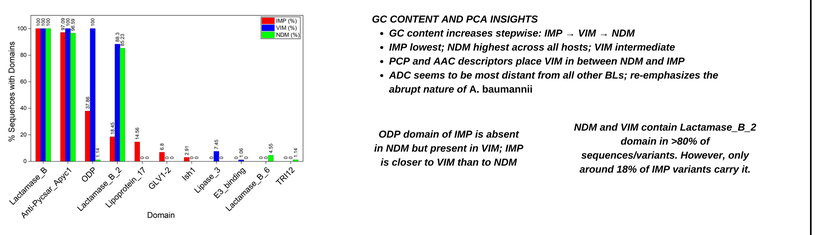
Within-family distances show that IMP is most divergent, while NDM is most conserved.
Interfamily distances show that NDM is closer to VIM than to IMP.

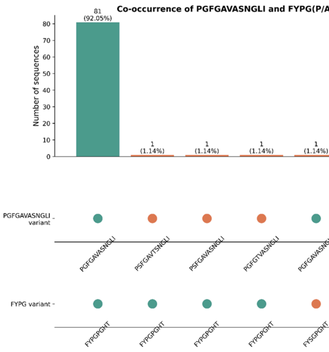




NDM and VIM are sister clades (branch from a common node) in the midpoint-rooted maximum likelihood tree; NDM is closer to VIM than to IMP.





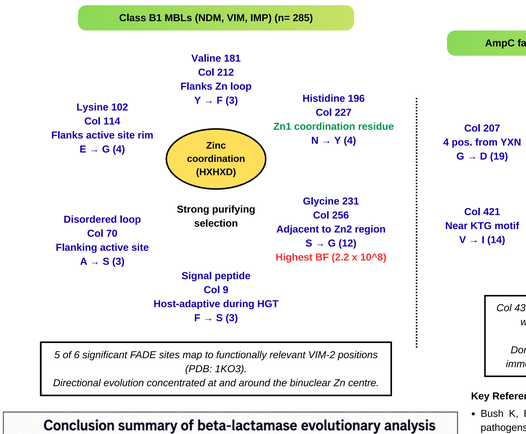


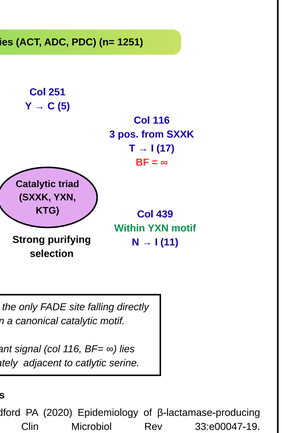
GC CONTENT AND PCA INSIGHTS

- ***GC content increases stepwise: IMP → VIM → NDM***
- ***IMP lowest; NDM highest across all hosts; VIM intermediate***
- ***PCP and AAC descriptors place VIM in between NDM and IMP***
- ***ADC seems to be most distant from all other BLs; re-emphasizes the abrupt nature of* A. baumannii**

ODP domain of IMP is absent in NDM but present in VIM; IMP is closer to VIM than to NDM

NDM and VIM contain Lactamase_B_2 domain in >80% of sequences/variants. However, only around 18% of IMP variants carry it.





Class B1 MBLs (NDM, VIM, IMP) (n= 285)

Valine 181
Col 212
Flanks Zn loop
Y → F (3)

Lysine 102
Col 114
Flanks active site rim
E → G (4)

Zinc coordination (HXHXD)

Histidine 196
Col 227
Zn1 coordination residue
N → Y (4)

Strong purifying selection

Glycine 231
Col 256
Adjacent to Zn2 region
S → G (12)
Highest BF (2.2 x 10^8)

Disordered loop
Col 70
Flanking active site
A → S (3)

Signal peptide
Col 9
Host-adaptive during HGT
F → S (3)

5 of 6 significant FADE sites map to functionally relevant VIM-2 positions (PDB: 1KO3).
Directional evolution concentrated at and around the binuclear Zn centre.

AmpC families (ACT, ADC, PDC) (n= 1251)

Col 251
Y → C (5)

Col 116
3 pos. from SXXK
T → I (17)
BF = ∞

Col 207
4 pos. from YXN
G → D (19)

Catalytic triad (SXXK, YXN, KTG)

Strong purifying selection

Col 421
Near KTG motif
V → I (14)

Col 439
Within YXN motif
N → I (11)

Col 439 is the only FADE site falling directly within a canonical catalytic motif.

Dominant signal (col 116, BF= ∞) lies immediately adjacent to catlytic serine.

Conclusion summary of beta-lactamase evolutionary analysis

Evolutionary Analysis of Beta-Lactamases in WHO Priority Pathogens

CONCLUSION

PPI network
NagZ is the most conserved BL-interacting partner across *P. aeruginosa* and Enterobacteriaceae — a priority broad-spectrum drug target.

phylogenetic signal
MBLs show progressive erosion of phylogenetic signal correlating with evolutionary age — NDM strongest, VIM intermediate, IMP least. Chromosomal proteins show consistent inter-organism proximity.

HGT directionality
Inter-host MBL gene flow is greatest between phylogenetically distant lineages — selection-driven rather than neutral dissemination.

VIM as bridge
VIM occupies an intermediate position between NDM and IMP across GC content, PCA, motif and PFAM analyses — inhibitors targeting VIM may also target NDM and IMP.

FADE analysis
Directional evolutionary pressure is concentrated at scaffold positions surrounding — not within — the active site. Catalytic residues are maintained under strong purifying selection.

NagZ-AmpR-AmpC
First proposal of the NagZ-AmpR-AmpC regulatory cascade as a co-evolutionary unit — a highly targetable axis for overcoming chromosomal AmpC-mediated resistance.

Conserved sequence motifs
FYPG(P/A)(G/A)H(T/S) — universally conserved across all three MBL families
PGFGAVASNGLI — cross-family similarity across IMP, VIM and NDM

KEY TAKEAWAY
This study provides a quantitative evolutionary framework with actionable insights for next-generation broad-spectrum beta-lactamase inhibitor design — from interaction networks to directional gene flow.

Key References

- Bush K, Bradford PA (2020) Epidemiology of β-lactamase-producing pathogens. Clin Microbiol Rev 33:e00047-19. https://doi.org/10.1128/CMR.00047-19
- Jacoby GA (2009) AmpC beta-lactamases. Clin Microbiol Rev 22:161–182. https://doi.org/10.1128/CMR.00036-08
- Naas T, Oueslati S, Bonnin RA, et al (2017) Beta-lactamase database (BLDB) — structure and function. J Enzyme Inhib Med Chem 32:917–919. https://doi.org/10.1080/14756366.2017.1344235
- Bouckaert R, Vaughan TG, Barido-Sottani J, et al (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput Biol 15:e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Acknowledgements

All the authors are thankful to ISAAR 2026 organizing team for the opportunity to present the poster. Ananya Anurag Anand, Kshitiz K and Sreedevi PC are thankful to Ministry of Education (MoE), GOI, for their fellowship. Sarfraz Anwar is thankful to CSIR-UGC (Council of Scientific and Industrial Research - University Grants Commission) for his fellowship. Vidushi Yadav is thankful to CST-UP (Council of Science and Technology, Uttar Pradesh) for her fellowship. All the authors are thankful to IIIT Allahabad for providing the research facility.