A Combinatorial Approach of High-Throughput Genomics and Mass Proteomics for Understanding the Regulation and Expression of Secondary Metabolite Production in Actinobacteria

Vipin Gupta,[a] Nirjara Singhvi[b]

[a]The Ministry of Environment, Forest and Climate Change, New Delhi, Delhi, India
[b]PhiXgen Pvt. Ltd., Gurugram, India

ABSTRACT Secondary metabolites produced by *Actinobacteria* are an important source of antibiotics, drugs, and antimicrobial peptides. However, the large genome size of actinobacteria with high gene coding density makes it difficult to understand the complex regulation of biosynthesis of such critically and economically important products. In the last few decades, apart from genomics sequences, high-throughput proteomics has proven beneficial to understand the key players regulating the expression pattern of secondary metabolite and antibiotic production in different experimental set-ups. In the past, we have been analyzing the genomics data and mass spectrometry-based proteomics to predict the regulation dynamics and crucial regulatory hubs in *Actinobacteria*. The multidirectional regulation and expression of the biosynthetic gene cluster responsible for the production of important metabolite take their cue from the other primary metabolism pathways with which they show intricate interactions in the interactome. The regulation occurs by not only the action and expression of the biosynthetic gene cluster but also the role of transcription factors and primary metabolic pathways. Using the key players of these interactomes, we can regulate the synthesis/production of these valuable peptides/metabolites. Simultaneously, the multi-omics approach has now opened new gateways in investigation, screening, and identification of naturally occurring antimicrobial peptides from actinobacteria which are beneficial for humans and also provide economic and industrial benefits to humankind.

KEYWORDS combinatorial biosynthesis, proteomics, *Actinobacteria*, secondary metabolite

*A*ctinobacteria represent one of the most varied Gram-positive microbial phyla with a high GC content and a remarkable range of complex morphologies, ranging from unicellular cocci to rods (*Micrococcus* and *Mycobacterium*, *Amycolatopsis*, *Frankia*, and *Streptomyces*) (1–3). Ecologically diversified, they have been reported from many varied habitats like terrestrial, aquatic, and mammalian microbiomes, etc. (4, 5). *Actinobacteria* are well known not only for their ability to produce antibiotics but also as a rich source of promising compounds with herbicidal, antitumor, antifungal, and anthelminthic activities. Common examples of these actinobacteria and their produced antibiotics are *Streptomyces hygroscopicus* (rapamycin [immunosuppressant]), *Amycolatopsis mediterranei* (rifamycin [antimicrobial]), *Saccharopolyspora erythraea* (erythromycin [broad-spectrum antibiotic]), *Amycolatopsis orientalis* (vancomycin [antibiotic]), etc. (6, 7). Due to such clinical properties, they are now acting as a major economic player in the pharmaceutical industry (8).

With the exponential increase in the global multifaceted phenomenon of multidrug resistance strains, there has been an alarmingly low rate of discovery of new antibiotics and their clinical approvals (9, 10). The apparent reason for the low rate of discovery

for new antibiotics is likely due to the lack of deep understanding of the complex regulatory mechanism that is required to activate the expression of the biosynthetic gene clusters (BGCs) responsible for antibiotic synthesis in *Actinobacteria* under laboratory conditions (11). Genome sequencing projects have revealed that most of the natural-product-rich *Actinobacteria* have larger genomes (>8 Mb) that contain an average of ~2,030 biosynthetic gene clusters (BGCs) for secondary metabolism. About half of these BGCs are nonribosomal peptide synthetases (NRPSs)/polyketide synthases (PKSs) (12). NRPSs/PKSs are basically the multidomain and multienzymatic megasynthase units that are involved in secondary metabolite synthesis in actinobacteria (13). In comparison to the other bacterial phyla like *Cyanobacteria*, where the gene coding density is around 52% (14), the gene coding density in *Actinobacteria* like *Corynebacterium* ranges up to 73% (15). This high coding density offers a wide scope of evolutionary dynamics in terms of horizontal gene transfer (16), gene duplication (17), gene decay (18), and genomic arrangements (19); thus, a prolonged genomic heterogeneity is associated with actinobacteria (15). Many groups have attempted to manipulate actinobacteria with a complex set of BGCs like *Amycolatopsis mediterranei* to augment the production of antibiotics but failed because of the lack of a stable cloning vector transformation system until 1991.

Later, the combinatorial biosynthesis approach (genetic engineering of natural biosynthetic clusters) used the information from genomic studies and established the first secondary metabolite databases like antiSMASH, which provides comprehensive information on BGC counts in actinobacteria (20). Combinatorial biosynthesis is considered the most successful approach of genetic engineering to produce new analogs of natural products by the process of modification of biosynthetic clusters and has been used worldwide for generating new antibiotics or their analogs. It was first used back in 1985, when Hopwood and group used this approach, targeting and cloning *Streptomyces coelicolor* BGC genes coding for actinorhodin into the medermycin and dihydrogranaticin antibiotic producers, respectively (21, 22). The new transformant thus generated produced large amounts of a new compound, including a hybrid antibiotic, mederrhodin A (with additional -OH group of actinorhodin). Similarly, a new compound-producing transformant was also generated producing dihydrogranatirhodin (23). Both these new analogs were found to harbor important clinical properties against bacterial infections (21–23).

A similar attempt was made by our group in 1991 with the successful development of the first hybrid plasmid vector construct as a prerequisite for transformation procedures for *Amycolatopsis mediterranei*, which further laid the foundation for its genetic manipulation (24). The cloning vector and transformation system were developed with the aim to study the biochemistry, physiology, and genetics of rifamycin B production in *Amycolatopsis mediterranei*. It was 2011 when the complete genome of *A. mediterranei* S699 was sequenced with 10.2 Mb and 9,575 predicted coding regions by the same group (12). Later, combinatorial biosynthesis was performed where substitution of the acyltransferase domain of module 6 of rifamycin polyketide synthase with that of module 2 of rapamycin generated the 24-desmethylrifamycin-B analog (9). This high-throughput genomics and combinatorial biosynthesis approach created a versatile platform for mining secondary metabolite-associated NRPSs/PKSs and their global regulatory genes in the other actinobacteria like *Streptomyces* and *Actinoplanes*, *Nocardia*, and *Sorangium* spp. for production of bioactive molecules (25, 26). Thus, basic understanding of the biosynthesis of antibiotics like erythromycin, rapamycin, and rifamycin opened up new possibilities for developing new molecules by their genetic manipulations (27–30). This method was found more convenient than all the other chemical and biological modifications of existing antibiotic molecules (29, 31–33). But with time, the number of successful attempts started to drop, and no new analogs were produced. The major problem faced by scientists across the world was the reduction in the amount of secondary metabolite production in the mutants compared to the wild-type strains after their genetic manipulations (9). Apart from the complicated handling conditions for the actinobacteria, their complex genomic architecture also heightens

the failure in genetic manipulations. It was seen that the manipulations may result in changes in the expression profile of the gene involved in the antibiotic production pathway (10). The major reason found was the changed structure of the product, which creates impact on the functioning in downstream processing in the cell or disruption of the closely linked gene during the manipulation process (10), thus resulting in higher probabilities of having low potent bioactivity.

In order to overcome these limitations of the genomic and combinatorial biosynthesis approach, the proteomics approach became popular. Initial proteomic investigations dealing with microbial secondary metabolism were largely based on targeted proteomic profiling formats and included two approaches for NRPS/PKS study: PrISM (PRediction Informatics for Secondary Metabolomes) and OASIS (Orthogonal Active Site Identification System) (34, 35). The two used different methods for the targeting or enrichment of PKS and NRPS. PrISM is a computational biology approach that selects nonribosomal peptides and type I and II polyketides by their large sizes (35). OASIS (experimental approach) chemically reacts with the active sites of NRPS/PKS for affinity enrichment and offers a valuable tool for enzyme discovery, culture condition optimization, and strain comparison (34). However, a gap still remained in genetics-based investigations of polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) biosynthetic pathways and understanding of their regulation, interaction, and activity in living systems. To overcome these hurdles, our group attempted strain-specific high-throughput proteomic mining overlapped with genomics using interactomics (36, 37). This multi-omic approach was employed to improve the production of 24-desmethylrifamycin-B, which was highly effective against rifampin-resistant strains of the tuberculosis bacterium (RR-TB) but had low yield. Where the proteomics approach unveiled the expression profile during the rifamycin production at a different timescale and was a major cause of low yield, the interactomics approach helped to uncover the major regulatory hub proteins that were regulating the secondary metabolite production (10, 36). These hubs were later targeted to enhance the production of the antibiotic.

Thus, applying a similar approach and combining the proteomics and *in silico* interactomics approaches, the unknown mechanism can be understood better and the information can be then used for promoting antibiotic yield. We strongly believe that overlapping high-throughput genomics with proteomics and their integration via interactomics can crucially provide the upper hand for understanding the current gaps in the regulatory structures of actinobacteria.

REFERENCES

1. Atlas R. 1997. Principles of microbiology. WCB McGraw-Hill, New York, NY.
2. Goodfellow M, Kämpfer P, Busse HJ, Trujillo ME, Suzuki K, Ludwig W, Whitman WB (ed). 2012. Bergey's manual of systematic bacteriology, vol 5. The Actinobacteria, part A and B. Springer, New York, NY.
3. Tan GYA, Goodfellow M. 2015. *Amycolatopsis*, p 1–40. *In* Whitman WB (ed). Bergey's manual of systematic bacteriology, John Wiley and Sons, Hoboken, NJ.
4. Qin S, Li WJ, Klenk HP, Hozzein WN, Ahmed I. 2019. Editorial: actinobacteria in special and extreme habitats: diversity, function roles and environmental adaptations, second edition. Front Microbiol 10:944. https://doi.org/10.3389/fmicb.2019.00944.
5. Servin JA, Herbold CW, Skophammer RG, Lake JA. 2008. Evidence excluding the root of the tree of life from the actinobacteria. Mol Biol Evol 25: 1–4. https://doi.org/10.1093/molbev/msm249.
6. Frasch HJ, Kalan L, Kilian R, Martin T, Wright GD, Stegmann E. 2015. Alternative pathway to a glycopeptide-resistant cell wall in the balhimycin producer *Amycolatopsis balhimycina*. ACS Infect Dis 1:243–252. https://doi.org/10.1021/acsinfecdis.5b00011.
7. Xu L, Huang H, Wei W, Zhong Y, Tang B, Yuan H, Zhu L, Huang W, Ge M, Yang S, Zheng H, Jiang W, Chen D, Zhao GP, Zhao W. 2014. Complete genome sequence and comparative analyses of the vancomycin-producing *Amycolatopsis orientalis*. BMC Genomics 15:363–381. https://doi.org/10.1186/1471-2164-15-363.
8. Hussein EI, Jacob JH, Shakhatreh MAK, Al-Razaq MAA, Juhmani AF, Cornelison CT. 2018. Detection of antibiotic-producing *Actinobacteria* in the sediment and water of Ma'in thermal springs (Jordan). Germs 8: 191–198. https://doi.org/10.18683/germs.2018.1146.
9. Nigam A, Almabruk KH, Saxena A, Jongtae Y, Mukherjee U, Kaur H, Kohli P, Kumari R, Singh P, Zakharov LN, Singh Y, Mahmud T, Lal R. 2014. Modification of rifamycin polyketide backbone leads to improved drug activity against rifampicin-resistant *Mycobacterium tuberculosis*. J Biol Chem 289: 21142–21152. https://doi.org/10.1074/jbc.M114.572636.
10. Singhvi N, Singh P, Prakash O, Gupta V, Lal S, Bechthold A, Singh Y, Singh RK, Lal R. 2021. Differential mass spectrometry-based proteome analyses unveil major regulatory hubs in rifamycin B production in *Amycolatopsis mediterranei*. J Proteomics 239:104168. https://doi.org/10.1016/j.jprot.2021.104168.
11. Baral B, Akhgari A, Metsä-Ketelä M. 2018. Activation of microbial secondary metabolic pathways: avenues and challenges. Synth Syst Biotechnol 3:163–178. https://doi.org/10.1016/j.synbio.2018.09.001. (Erratum, 5:328, 2020, https://doi.org/10.1016/j.synbio.2020.10.003.)
12. Verma M, Kaur J, Kumar M, Kumari K, Saxena A, Anand S, Nigam A, Ravi V, Raghuvanshi S, Khurana P, Tyagi AK, Khurana JP, Lal R. 2011. Whole genome

sequence of rifamycin B- producing strain *Amycolatopsis mediterranei* S699. J Bacteriol 193:5562–5563. https://doi.org/10.1128/JB.05819-11.

13. Ansari MZ, Yadav G, Gokhale RS, Mohanty D. 2004. NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. Nucleic Acids Res 32:W405–W413. https://doi.org/10.1093/nar/gkh359.

14. Land M, Hauser L, Jun SR, Nookaew I, Leuze MR, Ahn TH, Karpinets T, Lund O, Kora G, Wassenaar T, Poudel S, Ussery DW. 2015. Insights from 20 years of bacterial genome sequencing. Funct Integr Genomics 15: 141–161. https://doi.org/10.1007/s10142-015-0433-4.

15. Ventura M, Canchaya C, Tauch A, Chandra G, Fitzgerald GF, Chater KF, van Sinderen D. 2007. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiol Mol Biol Rev 71:495–548. https://doi.org/10.1128/MMBR.00005-07.

16. Gogarten JP, Doolittle WF, Lawrence JG. 2002. Prokaryotic evolution in light of gene transfer. Mol Biol Evol 19:2226–2238. https://doi.org/10.1093/oxfordjournals.molbev.a004046.

17. Andersson SG. 2000. The genomics gamble. Nat Genet 26:134–135. https://doi.org/10.1038/79835.

18. Mira A, Ochman H, Moran NA. 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet 17:589–596. https://doi.org/10.1016/S0168-9525(01)02447-7.

19. Wolf YI, Rogozin IB, Kondrashov AS, Koonin EV. 2001. Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. Genome Res 11:356–372. https://doi.org/10.1101/gr.GR-1619R.

20. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res 39(Web Server issue):W339–W346. https://doi.org/10.1093/nar/gkr466.

21. Takano S, Hasuda K, Ito A, Koide Y, Ishii F. 1976. A new antibiotic, medermycin. J Antibiot (Tokyo) 29:765–768. https://doi.org/10.7164/antibiotics.29.765.

22. Corbaz R, Ettlinger L, Gäumann E, Kalvoda J, Keller-Schierlein W, Kradolfer F, Manukian BK, Neipp L, Prelog V, Reusser P, Zähner H. 1957. Stoffwechselprodukte von Aktinomyceten. 9. Mitteilung, Granaticin. Helv Chim Acta 40:1262–1269. https://doi.org/10.1002/hlca.19570400518.

23. Hopwood DA. 1973. Genetics of the Actinomycetales. Soc Appl Bacteriol Symp Ser 2:131–153.

24. Lal R, Lal S, Grund E, Eichenlaub R. 1991. Construction of a hybrid plasmid capable of replication in *Amycolatopsis mediterranei*. Appl Environ Microbiol 57:665–671. https://doi.org/10.1128/aem.57.3.665-671.1991.

25. Hackl S, Bechthold A. 2015. The gene bldA, a regulator of morphological differentiation and antibiotic production in *Streptomyces*. Arch Pharm (Weinheim) 348:455–462. https://doi.org/10.1002/ardp.201500073.

26. Chater KF. 2006. Streptomyces inside-out: a new perspective on the bacteria that provide us with antibiotics. Philos Trans R Soc Lond B Biol Sci 361:761–768. https://doi.org/10.1098/rstb.2005.1758.

27. Cortes J, Haydock SF, Roberts GA, Bevitt DJ, Leadlay PF. 1990. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. Nature 348:176–178. https://doi.org/10.1038/348176a0.

28. Cane DE, Walsh CT. 1999. The parallel and convergent universe of polyketide synthases and non-ribosomal peptide synthases. Chem Biol 6:319–325.

29. Lal R, Kumari R, Kaur H, Khanna R, Dhingra N, Tuteja D. 2000. Regulation and manipulation of the gene clusters encoding type-I PKSs. Trends Biotechnol 18:264–274. https://doi.org/10.1016/S0167-7799(00)01443-8.

30. Lal R, Lal S. 1994. Recent trends in rifamycin research. Bioessays 16: 211–216. https://doi.org/10.1002/bies.950160313.

31. Staunton J, Wilkinson B. 2001. Combinatorial biosynthesis of polyketides and nonribosomal peptides. Curr Opin Chem Biol 5:159–164. https://doi.org/10.1016/S1367-5931(00)00185-X.

32. Walsh CT. 2002. Combinatorial biosynthesis of antibiotics: challenges and opportunities. Chembiochem 3:124–134. https://doi.org/10.1002/1439-7633(20020301)3:2/3<124::AID-CBIC124>3.0.CO;2-J.

33. Rix U, Fischer C, Remsing LL, Rohr J. 2002. Modification of post-PKS tailoring steps through combinatorial biosynthesis. Nat Prod Rep 19:542–580. https://doi.org/10.1039/b103920m.

34. Meier JL, Niessen S, Hoover HS, Foley TL, Cravatt BF, Burkart MD. 2009. An orthogonal active site identification system (OASIS) for proteomic profiling of natural product biosynthesis. ACS Chem Biol 4:948–957. https://doi.org/10.1021/cb9002128.

35. Skinnider MA, Merwin NJ, Johnston CW, Magarvey NA. 2017. PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. Nucleic Acids Res 45:W49–W54. https://doi.org/10.1093/nar/gkx320.

36. Singhvi N, Gupta V, Singh P, Prakash O, Bechthold A, Singh Y, Lal R. 2020. Prediction of transcription factors and their involvement in regulating rifamycin production in *Amycolatopsis mediterranei* S699. Indian J Microbiol 60:310–317. https://doi.org/10.1007/s12088-020-00868-5.

37. Gupta V, Haider S, Sood U, Gilbert JA, Ramjee M, Forbes K, Singh Y, Lopes BS, Lal R. 2016. Comparative genomic analysis of novel *Acinetobacter* symbionts: a combined systems biology and genomics approach. Sci Rep 6: 29043. https://doi.org/10.1038/srep29043.