



Prediction of Transcription Factors and Their Involvement in Regulating Rifamycin Production in *Amycolatopsis mediterranei* S699

Nirjara Singhvi¹ · Vipin Gupta¹ · Priya Singh^{1,2} · Om Prakash³ · Andreas Bechthold⁴ · Yogendra Singh¹ · Rup Lal⁵

Received: 17 February 2020 / Accepted: 3 April 2020 / Published online: 18 April 2020
© Association of Microbiologists of India 2020

Abstract *Amycolatopsis mediterranei* S699 produces rifamycin B and successors of this strain are in use for the industrial production of rifamycin B. Semisynthetic derivatives of rifamycin B are used against *Mycobacterium tuberculosis* that causes tuberculosis. Although the rifamycin biosynthetic gene cluster was characterized two decades ago, the regulation of rifamycin B biosynthesis in *Amycolatopsis mediterranei* S699 is poorly understood. In this study, we analysed the genome and proteome of *Amycolatopsis mediterranei* S699 and identified 1102 transcription factors which comprise about 10% of the total genome. Using interactomics approaches we delineated 30 unique transcription factors directly involved in secondary

metabolism that regulate rifamycin B biosynthesis. We also predict the role of RifN as hub in controlling the regulation of other genes involved in rifamycin biosynthesis. RifN is important for maintaining the integrity of the rifamycin-network. Thus, these transcription factor can be exploited to improve rifamycin B production in *Amycolatopsis mediterranei* S699.

Keywords Transcription factors · *Amycolatopsis mediterranei* · Rifamycin · Regulation · Interactomics

Introduction

In prokaryotic cells, transcription factors (TF) are proteins that play important role in controlling the rate of transcription of a gene. TF's bind to specific DNA sequences enabling RNA-polymerases to perform the transcription process. There are a wide range of transcription factors. Some harbour DNA binding regions that bind directly to promoters. Others bind to the enhancer region of a gene. Proteins binding directly to the promoter region initiates the transcription process. TFs binding to the enhancer regions are indirectly stimulating or repressing the transcription process [1]. It is known that regulation of transcription factor is the most important form to control a gene expression. TF have been identified in bacterial family actinobacteria like genus *Streptomyces* have been reported to control gene expression and their regulation within the organized network [1, 2].

Amycolatopsis mediterranei S699 is a well-known actinobacterium that produces rifamycin B. It has been developed as industrial strains for the production of rifamycin B primarily by using classical strain improvement methods [3]. Semisynthetic derivatives of rifamycin B are

Nirjara Singhvi and Vipin Gupta have contributed equally.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12088-020-00868-5>) contains supplementary material, which is available to authorized users.

✉ Yogendra Singh
ysinghdu@gmail.com

✉ Rup Lal
ruplal@gmail.com

- ¹ Department of Zoology, University of Delhi, Delhi 110007, India
- ² Present Address: Acharya Narendra Dev College, University of Delhi, New Delhi 110019, India
- ³ National Centre for Cell Sciences, Pune, Maharashtra 411007, India
- ⁴ Pharmaceutical Biology and Biotechnology, Institute of Pharmaceutical Sciences, Albert-Ludwigs University, 79104 Freiburg, Germany
- ⁵ The Energy and Resources Institute, Darbari Seth Block, IHC Complex, New Delhi 110033, India

in use against *Mycobacterium tuberculosis* that causes tuberculosis (TB) [4–8]. With the development of Multi-Drug Resistant (MDR) strains of *Mycobacterium tuberculosis*, there has been an urgent need to develop new drugs against MDR-TB [9–12]. Characterization of the rifamycin biosynthetic gene cluster and a better understanding of the genome of *Amycolatopsis mediterranei* [13, 14] opened up the possibilities to manipulate rifamycin cluster to produce analogs of rifamycin B. An analog of rifamycin B; 24-desmethylrifamycin B was produced by manipulation of the *rifPKS* gene cluster and semisynthetic derivatives of this analog were found to be effective against rifampicin resistant strains of *Mycobacterium tuberculosis* [15, 16]. In order to commercially exploit the mutant strain that produces 24-desmethylrifamycin B, understanding of the regulatory mechanisms of rifamycin B and 24-desmethylrifamycin B production in the wild type and in the mutant strain respectively, is very important. Many studies on rifamycin B production have been performed in classical strain improvement programs and in programs used to understand the biosynthetic network leading to rifamycin B production. In addition, there are scanty reports on the regulation of antibiotic production in species of *Streptomyces* that produce a wide array of antibiotics [17–19]. However, little has been done to understand the role of transcription factors (TF) of *Amycolatopsis mediterranei* S699. As soon as the genome sequence of *Amycolatopsis mediterranei* S699 was available, we analysed its genome to explore different transcriptional regulators playing role in controlling antibiotic production. Cross-regulation among the cluster and regulatory proteins results in a highly complex regulatory network. In addition, this study also revealed that one gene involved in 3-amino 5-hydroxybenzoic acid (ABHA) production, *rifN* [20], acts as the major regulatory hub to control the regulation of the rifamycin biosynthetic gene cluster (RifBGC). This study now reports the crucial transcription factors and important genes that can be exploited to upscale the regulation of rifamycin production.

Methods

Complete Genome Sequencing of *Amycolatopsis mediterranei* S699

Whole genome complete sequencing was performed using Roche 454 system (GS20 version) and Sanger shotgun sequencing by Verma and group (2011) [21]. We reassembled, validated, annotated the genome again (accession number CP002896). The sequence was reanalysed for its genomic features.

Rifamycin BGC Gene Extraction

In order to identify transcription factors involved in rifamycin production it is important to extract genes linked with the antibiotic production. The rifamycin biosynthetic gene cluster harbours 42 genes arranged in four regions. The list of genes and their respective functions are given in Supplementary Table 1 [13, 14]. The 42 gene were identified, and protein sequence and identifier were extracted from UniPROT database [22]. The 42 rifamycin biosynthesis cluster genes were identified from the complete genome of S699 and their interaction map was created using STRING database v10.0 [23].

Transcription Factor Profiling

The genome harbours genes that code for transcription factors that are involved in regulation of gene expression by controlling the process of conversion of DNA to RNA during the process of transcription. Total TFs were annotated from the whole genome of *A. mediterranei* S699 using BLASTn program against P2TF database V2.9 [24] at default parameters. P2TF is an integrated and comprehensive database that hold the compilation of TF genes along with the annotation, sequence features, functional domain. The ID obtained for S699 was converted to ID for strain U32 due to the availability of interaction data and 99.9% similarity of the genome as annotated and shown in Fig. 1. These proteins were then subjected to STRING database to predict the interaction map of TF with that of rifamycin biosynthetic gene cluster to understand the mechanism and role of TF in regulating the production of rifamycin B by the cell. The interactome was then visualized using Cytoscape v3.7.1 [25].

Network Analysis

The network was analyzed for the first neighbor of each gene from rifamycin BGC. The interactome map was created for the selected proteins along with the 42 rifamycin BGC genes. The network was subjected to GeneMania force directed layout for better understanding of the interactions. The network was analyzed using NetworkAnalyzer, the plug-in of Cytoscape v 3.7.1 and Perl programming version 5.18.2.2. Network was visualized based on node degree values to identify key proteins in the cluster [26].

Functional Profiling

The functional profiling of TFs was analyzed using STRINGApp plugin that allow to retrieve functional enrichment for Gene Ontology terms. This can be achieved

Table 1 List of transcription factor found to be involved in rifamycin B production

Sr. no.	UniPROT ID	Family	Type	Interacting RifBGC protein	References
1	RifQ/ AMED_0634	TetR family	1 TetR_N,1 TetR_C	RifP, RifO	[28]
2	AMED_7826	Rok family	1 MarR,1 ROK	RifN	This study
3	AMED_6925	Rok family	1 MarR,1 ROK	RifN	This study
4	AMED_6775	Rok family	1 HTH_CRP,1 ROK	RifN	This study
5	AMED_2585	Rok family	1 MarR,1 ROK	RifN	This study
6	AMED_3749	Rok family	1 MarR,1 ROK	RifN	This study
7	AMED_3418	Rok family	1 HTH_IcIR,1 ROK	RifN	This study
8	AMED_2985	Rok family	1 HTH_5,1 ROK	RifN	This study
9	AMED_1709	TetR family	1 TetR_N	Rif-Orf3	This study
10	AMED_1153	Rok family	1 HTH_5,1 ROK	RifN	This study
11	AMED_4310	Rok family	1 MarR,1 ROK	RifN	This study
12	AMED_5065	Rok family	1 Fe_dep_repress,1 ROK	RifN	This study
13	AMED_1248	TetR family	1 TetR_N,1 TetR_C	RifQ	This study
14	AMED_4179	DNA-binding protein, Hns family	1 Histone_HNS	RifR, RifC, RifD	This study
15	AMED_2021	TetR family	1 TetR_N	Rif-Orf3	This study
16	AMED_8119	LuxR family	1 Trans_reg_C,1 BTAD,1 HTH_LUXR	RifL, RifB	This study
17	AMED_5173	LuxR family	1 HTH_LUXR	Rif-Orf3	This study
18	RifZ/ AMED_0655	LuxR family	1 HTH_LUXR	Rif-Orf19	[29]
19	AMED_6797	TetR family	1 TetR_N	Rif-Orf11	This study
20	AMED_5356	TetR family	1 TetR_N	Rif-Orf19	This study
21	AMED_2205	AsnC family	1 AsnC_trans_reg	RifO	This study
22	AMED_5377	TetR family	1 TetR_N	Rif-Orf3	This study
23	AMED_5358	TetR family	1 TetR_N	Rif-Orf19	This study
24	AMED_4412	IcIR family	1 HTH_IcIR,1 IcIR	RifZ	This study
25	AMED_8118	Sarp family	1 Trans_reg_C,1 BTAD	RifL	This study
26	AMED_2829	Unclassified	1 HTH_11	RifO	This study
27	AMED_1347	TetR family	1 TetR_N	RifB	This study
28	AMED_4442	Sarp family	1 Trans_reg_C,1 BTAD	RifZ	This study
29	AMED_0904	Xre family 1	HTH_3	Rif-Orf20	This study
30	AMED_4014	Sarp family	1 Trans_reg_C,1 BTAD	RifM	This study

This table shows the uniport ID, family and type of transcription factor and their respective references. Also, the interacting rifamycin biosynthetic cluster genes have been enlisted

by converting network to STRING network using STRINGfy and then the network is subjected to STRINGEnrichment.

Results and Discussion

Amycolatopsis mediterranei S699 Genome

The genome of S699 was reanalysed and circular map was generated using CGView Server [27]. The complete

genome size of S699 was found to be 10.23 Mb, with 9575 predicted coding sequences as reported earlier [21]. A 90 kb DNA fragment containing the rifamycin biosynthetic gene cluster harbouring 42 genes was also mapped (Fig. 1). The coding density as reported earlier to be 90%, with an average CDS length of 954 bp was confirmed. A total of 52 tRNA and 4 rRNA operons were present. The genome was found to harbour ~ 10% of the genes that coded for transcription factors involved in regulation of gene expression. High number of TFs are embedded in the genomes of bacterial community that are involved in

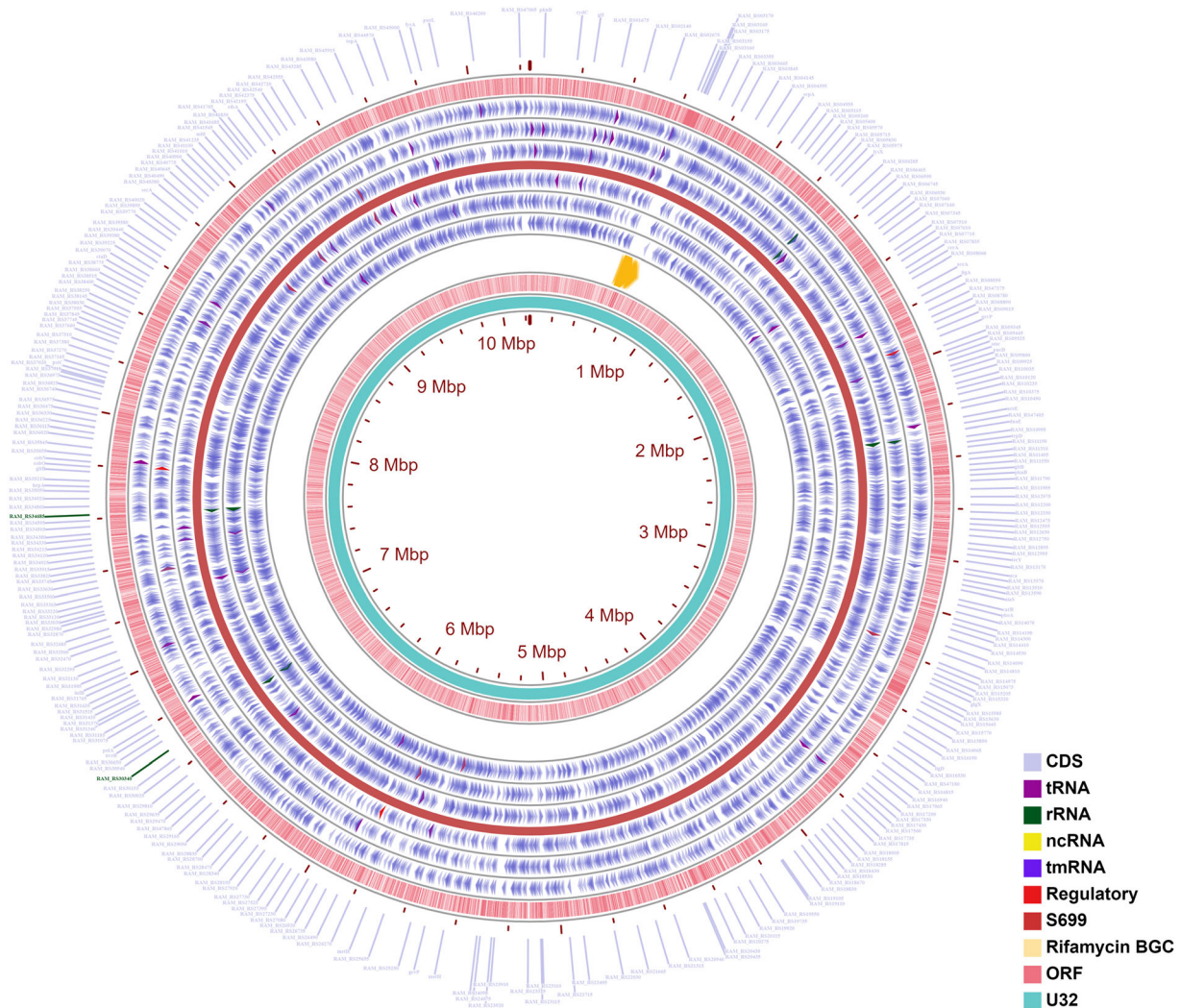


Fig. 1 Circular genome map representation of *Amycolatopsis mediterranei* S699 (red) with genome size of 10.23 Mb generated using CGView. The following features are shown (moving from the innermost track inward, origin of replication at 0 Mbp) (1) The blue track represents the genome of U32, its phylogenetically similar neighbour with ANI similarity of 99.9%. (2) ORF on the forward strand (3) Rifamycin biosynthetic gene cluster have been mapped on

to the genome as orange arrows (4–6) CDS (purple) has been marked along with the genomic features like tRNA, rRNA, ncRNA, Regulatory genes on forward strands (7) backbone ring of the genome S699 (red) (8–10) CDS (purple) has been marked along with the genomic features like tRNA, rRNA, ncRNA, Regulatory genes on reverse strands (11) ORF on the reverse strand (Genome sequence by Verma et al. [21])

secondary metabolite synthesis, residing in tough environments etc. [1].

Identification of Transcription Factor

In order to explore the role of transcription factor involved in controlling rifamycin biosynthetic gene, 1102 TFs were extracted from the reference genome of *Amycolatopsis mediterranei* S699 using P2TF database. TFs along with the 42 genes were then mapped onto STRING database to analyse the interaction among these genes/ proteins they encode. Cytoscape v3.7.1 [25] was used to visualize the final interactome (Fig. 2). Out of 1102, 921 TF's were

found to be interacting with each other and RifBGC genes. We then extracted TFs that were directly interacting with RifBGC using the first neighbouring interacting proteins. A total of 30 TFs were found to be directly involved in the regulation of RifBGC (Table 1).

Diversity of TFs Regulators

Out of 1102 TFs identified in the genome of *A. mediterranei* S699, primarily TFs belonging to MarR-, TetR-, Rok-, PadR-, GntR-, HrcA-, IclR-, LuxR-, LysR-, MerR-, AraC-, ArgR-, ArsR-, AsnC-, CsoR-, DeoR-, and WhiB-families were identified. Proteins belonging to the TetR-

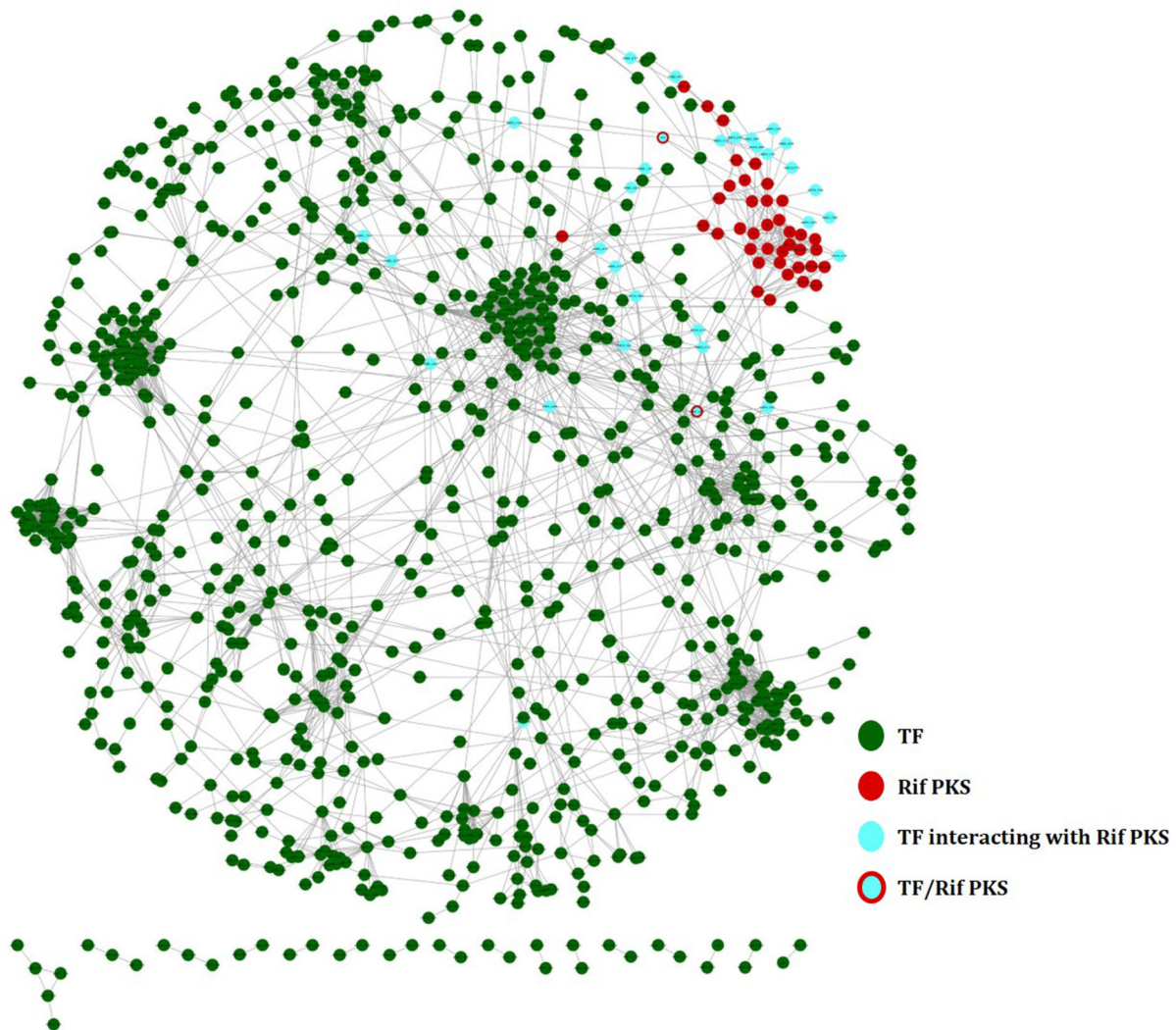


Fig. 2 The whole interaction map of total transcription factors (green) and the rifamycin biosynthetic gene cluster (red) [Node:963 and Edges:3202]. The spherical nodes represent the proteins and grey

lines represents their interaction with each other. The transcription factors interacting directly with rifamycin cluster is marker in blue

(232), MarR- (74) and Xre-family (74) are most abundant. The 30 TFs interacting with RifBGC were found to belong to the Rok-, TetR-, MarR-, LuxR-, Xre- and AsnC-family. It was noted that 2 of these TFs were proteins from RifBGC, namely, RifQ (STRING ID: AMED_0634) and RifZ (STRING ID: AMED_0655) (Fig. 2; Marked blue with red boundary).

RifQ is a TetR type repressor that plays a crucial role in conferring antibiotic resistance to the bacteria [28]. The regulators belonging to this family are associated with the regulation of genes involved in export of small-molecules. RifQ is reported to negatively regulate expression of *rifP*, a membrane protein that is known to export rifamycin B outside the cell and thus balances the extracellular transport [28]. RifZ (a regulator belonging to the LuxR family) is involved in regulating the transcription of all six operons of

the RifBGC and it acts as on/off switch for the synthesis of rifamycin B [29]. Out of the remaining 28 TFs, the highest number of TF ($n = 10$) were found belonging to the ROK family that act as multiple antibiotic resistance regulators and cAMP regulatory proteins (AMED_7826, AMED_6925, AMED_2585, AMED_3749, AMED_2985, AMED_1153, AMED_6775, AMED_3418, AMED_4310, AMED_5065).

Many regulators (Table 1) belonging to other families (e.g. ROK family) are often involved in primary metabolism like glycolysis or fatty acid metabolism but may also play a role in antibiotic production in bacteria like *Streptomyces* [30]. The association and interaction of these TFs with RifBGC correlates with rifamycin biosynthesis. Other five regulatory proteins belong to the MarR family. Proteins of this family are known to regulate the activity of

genes involved in stress response, virulence, and degradation or export of harmful chemicals such as phenolic compounds and antibiotics [31]. MarR related TFs were also found to interact with RifO. This protein play an important role in export of rifamycin interacting with RifP. Further, eight TFs belongs to the TetR family (AMED_1709, AMED_1248, AMED_2021, AMED_6797, AMED_5356, AMED_5377, AMED_5358, AMED_1347), two to the LuxR family (AMED_8119, AMED_5173), two to the IclR family (AMED_3418, AMED_4412) one to the Xre family and one to the Sarp family (AMED_8118, AMED_4442, AMED_4014) (Fig. 3a). Members of the Xre-family are also involved in stress related regulation processes in a cell. Another local regulator, AsnC (AMED_2205) was found to interact specifically with RifQ.

Network Analysis

In order to understand and unveil the role of the TFs in rifamycin production, interaction map was created for 70 proteins including rifamycin BGC and interacting TFs and visualized using Cytoscape v3.7.1. The network constructed obeyed power law, $P(k) \sim k^{-\gamma}$ indicating the

hierarchical nature of the network (Fig. 3b). Degree analysis was carried out in order to check for maximum interaction of the RifBGC node with TF. The predicted network was of tight topology having 70 nodes and 215 edges. Network analysis predicted the order of degree (interaction capacities) of all the TFs (n = 30) with interacting RifBGC proteins (n = 42) for their direct interactions. Two types of hubs were identified, namely, Rif level and TF level hubs Out of 42 RifBGC proteins RifN and Rif-ORF3 were found to interact with highest degree of 10 and 4 TFs respectively at Rif level (Fig. 3c). RifN is known to be involved in converting the amino analogues to aromatic starter unit, 3-amino-5-hydroxybenzoic acid (ABHA). Thus, it acts as a hub to control regulation of other genes in the cluster. Looking into the TF level degree analysis, the highest interacting TFs were AMED_4179 that is a DNA binding protein and found to interacting with RifR, RifC and RifD. Hub proteins are reported to be important for cell growth as well as regulation [26, 32–34]. Absence of the expression of these proteins may disrupt the cluster completely. Thus, these proteins may have very crucial role in regulating rifamycin cluster as a whole and also can be explored during better analog productions of rifamycin B.

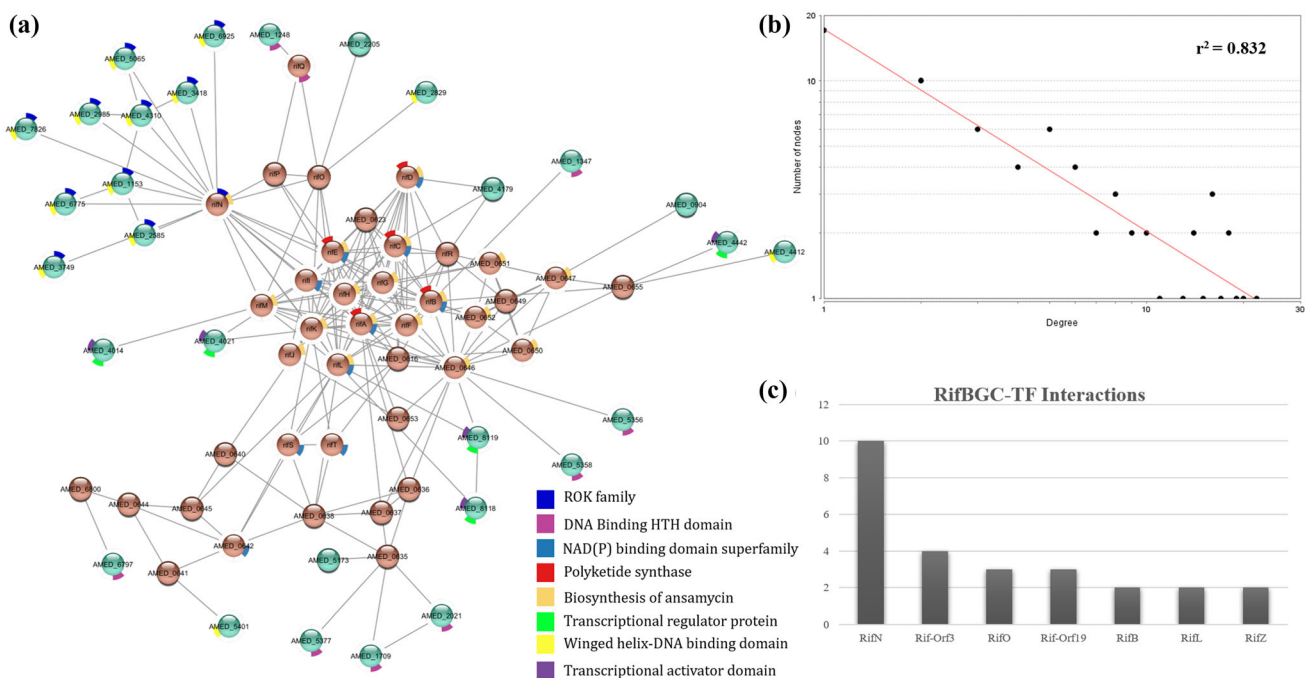


Fig. 3 a The interaction map of rifamycin BGC (brown) along with their interacting transcription factors (green) [Node:70 and Edges: 215]. The ring represents the functional enrichment of each protein description of which are mentioned in the legend. **b** The topological properties of probability of degree distribution $P(k)$ of the network depicting with correlation coefficient values (r^2). The network follows

the power law distribution and show the nature of scale-free network, suggesting a hierarchical organization in the network. **c** Bar graph showing the degree values for rifamycin genes interacting with number of transcription factors, where X-axis denote the rifamycin BGC genes and Y-axis shows number of TF's they are interacting

Conclusions

10% of the genome of the rifamycin producer *Amycolatopsis mediterranei* S699 consist of genes encoding TFs. 30 unique transcription factors are directly involved in regulating rifamycin production and its extra-cellular export. These transcription factors may be involved in feedback mechanisms or act as complex allosteric regulators but need further investigation. RifN is the hub protein regulating the RifBGC-TF network. These findings are very important and open the possibility to improve the production of rifamycin B by *Amycolatopsis mediterranei* S699 and its analog, 24-desmethylrifamycin B by exploiting and manipulating the expression levels of genes encoding these regulators proteins.

Acknowledgements RL acknowledges The National Academy of Sciences, India, for support under the NASI-Senior Scientist Platinum Jubilee Fellowship Scheme. NS acknowledges Council of Scientific and Industrial Research, New Delhi (CSIR) for doctoral fellowship.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Romero-Rodríguez A, Robledo-Casados I, Sánchez S (2015) An overview on transcriptional regulators in *Streptomyces*. *Biochim Biophys Acta* 1849:1017–1039. <https://doi.org/10.1016/j.bbagr.2015.06.007>
- Seshasayee AS, Sivaraman K, Luscombe NM (2011) An overview of prokaryotic transcription factors: a summary of function and occurrence in bacterial genomes. *Subcell Biochem* 52:7–23. https://doi.org/10.1007/978-90-481-9069-0_2
- 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](https://doi.org/10.1016/s0167-7799(00)01443-8)
- Aristoff PA, Garcia GA, Kirchoff PD, Showalter HH (2010) Rifamycins—obstacles and opportunities. *Tuberculosis* 90:94–118. <https://doi.org/10.1016/j.tube.2010.02.001>
- Kumari R, Singh P, Lal R (2016) Genetics and Genomics of the Genus *Amycolatopsis*. *Indian J Microbiol* 56:233–246. <https://doi.org/10.1007/s12088-016-0590-8>
- Gatta L, Scarpignato C (2017) Systematic review with meta-analysis: rifaximin is effective and safe for the treatment of small intestine bacterial overgrowth. *Aliment Pharmacol Ther* 45:604–616. <https://doi.org/10.1111/apt.13928>
- Kumar A, Chettiar S, Parish T (2017) Current challenges in drug discovery for tuberculosis. *Expert Opin Drug Discov* 12:1–4. <https://doi.org/10.1080/17460441.2017.1255604>
- Schön T, Chrystanthou E (2017) Minimal inhibitory concentration distributions for *Mycobacterium avium* complex—towards evidence based susceptibility breakpoints. *Int J Infect Dis* 55:122–124. <https://doi.org/10.1016/j.ijid.2016.12.027>
- Prasad R, Singh A, Balasubramanian V, Gupta N (2017) Extensively drug-resistant tuberculosis in India: current evidence on diagnosis & management. *Indian J Med Res* 145:271–293. https://doi.org/10.4103/ijmr.IJMR_177_16
- Matteelli A, Roggi A, Carvalho AC (2014) Extensively drug-resistant tuberculosis: epidemiology and management. *Clin Epidemiol* 6:111–118. <https://doi.org/10.2147/CLEPS35839>
- Haydel SE (2010) Extensively drug-resistant tuberculosis: a sign of the times and an impetus for antimicrobial discovery. *Pharmaceuticals (Basel)* 3:2268–2290. <https://doi.org/10.3390/ph3072268>
- Shah NS, Wright A, Bai GH, Barrera L, Boulahbal F, Martín-Casabona N, Drobniowski F, Gilpin C, Havelková M, Lepe R, Lumb R, Metchock B, Portaels F, Rodrigues MF, Rüsck-Gerdes S, Van Deun A, Vincent V, Laserson K, Wells C, Cegielski JP (2007) Worldwide emergence of extensively drug-resistant tuberculosis. *Emerg Infect Dis* 13:380–387. <https://doi.org/10.3201/eid1303.061400>
- August PR, Tang L, Yoon YJ, Ning S, Muller R, Yu TW, Taylor M, Hoffman D, Kim CG, Zhang X, Hutchinson CR, Floss HG (1998) Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem Biol* 5:69–79. [https://doi.org/10.1016/S1074-5521\(98\)90141-7](https://doi.org/10.1016/S1074-5521(98)90141-7)
- Xu J, Wan E, Kim CJ, Floss HG, Mahmud T (2005) Identification of tailoring genes involved in the modification of the polyketide backbone of rifamycin B by *Amycolatopsis mediterranei* S699. *Microbiol* 151:2515–2528. <https://doi.org/10.1099/mic.0.28138-0>
- 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 in *Amycolatopsis mediterranei* leads to the production of a new rifamycin B analog with an improved drug activity against rifampicin-resistant *Mycobacterium tuberculosis* strains. *J Biol Chem* 289:21142–21152. <https://doi.org/10.1074/jbc.M114.572636>
- Singh P, Kumari R, Lal R (2017) Bedaquiline: fallible hope against drug resistant tuberculosis. *Indian J Microbiol* 57:371–377. <https://doi.org/10.1007/s12088-017-0674-0>
- McArthur M, Bibb MJ (2008) Manipulating and understanding antibiotic production in *Streptomyces coelicolor* A3(2) with decoy oligonucleotides. *Proc Natl Acad Sci USA* 105:1020–1025. <https://doi.org/10.1073/pnas.0710724105>
- Wang J, Xu J, Luo S, Ma Z, Bechthold A, Yu X (2018) AdpA_{sd}, a positive regulator for morphological development and toyocamycin biosynthesis in *Streptomyces diastatochromogenes* 1628. *Curr Microbiol* 75:1345–1351. <https://doi.org/10.1007/s00284-018-1529-6>
- Xu X, Wang J, Bechthold A, Ma Z, Yu X (2017) Selection of an efficient promoter and its application in toyocamycin production improvement in *Streptomyces diastatochromogenes* 1628. *World J Microbiol Biotechnol* 33:30–38. <https://doi.org/10.1007/s11274-016-2194-1>
- Guo J, Frost JW (2002) Biosynthesis of 1-deoxy-1-imino-D-erythrose 4-phosphate: a defining metabolite in the aminoshikimate pathway. *J Am Chem Soc* 124:528–529. <https://doi.org/10.1021/ja016963v>
- 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>
- UniProt Consortium (2007) The universal protein resource (UniProt). *Nucleic Acids Res* 36:D190–195. [https://doi.org/10.1093/nar/gkm895\(Database issue\)](https://doi.org/10.1093/nar/gkm895(Database issue))
- Mering CV, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B (2003) STRING: a database of predicted functional associations

- between proteins. *Nucleic Acids Res* 31:258–261. <https://doi.org/10.1093/nar/gkg034>
24. Ortet P, De Luca G, Whitworth DE, Barakat M (2012) P2TF: a comprehensive resource for analysis of prokaryotic transcription factors. *BMC Genomics* 13:628–636. <https://doi.org/10.1186/1471-2164-13-628>
 25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
 26. 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* 6:29043–29055. <https://doi.org/10.1038/srep29043>
 27. Grant JR, Stothard P (2008) The CGView server: a comparative genomics tool for circular genomes. *Nucleic Acids Res* 36:W181–W184. <https://doi.org/10.1093/nar/gkn179>
 28. Lei C, Wang J, Liu Y, Liu X, Zhao G, Wang J (2017) A feedback regulatory model for RifQ-mediated repression of rifamycin export in *Amycolatopsis mediterranei*. *Microb Cell Fact* 17:14–23. <https://doi.org/10.1186/s12934-018-0863-5>
 29. Li C, Liu X, Lei C, Yan H, Shao Z, Wang Y, Zhao G, Wang J, Ding X (2017) RifZ (AMED_0655) is a pathway-specific regulator for rifamycin biosynthesis in *Amycolatopsis mediterranei*. *Appl Environ Microbiol* 83:e03201–e3216. <https://doi.org/10.1128/AEM.03201-16>
 30. Świątek MA, Gubbens J, Bucca G, Song E, Yang YH, Laing E, Kim BG, Smith CP, van Wezel GP (2013) The ROK family regulator Rok7B7 pleiotropically affects xylose utilization, carbon catabolite repression, and antibiotic production in *Streptomyces coelicolor*. *J Bacteriol* 195:1236–1248. <https://doi.org/10.1128/JB02191-12>
 31. Grove A (2013) MarR family transcription factors. *Curr Biol* 23:R142–143. <https://doi.org/10.1016/j.cub.2013.01.013>
 32. Verma H, Dhingra GG, Sharma M, Gupta V, Negi RK, Singh Y, Lal R (2020) Comparative genomics of *Sphingopyxis* spp. unravelled functional attributes. *Genomics* 112:1956–1969. <https://doi.org/10.1016/j.ygeno.2019.11.008>
 33. Sood U, Hira P, Kumar R, Bajaj A, Rao DLN, Lal R, Shakarad M (2019) Comparative genomic analyses reveal core-genome-wide genes under positive selection and major regulatory hubs in outlier strains of *Pseudomonas aeruginosa*. *Front Microbiol* 10:53–76. <https://doi.org/10.3389/fmicb.2019.00053>
 34. Kumar R, Verma H, Haider S, Bajaj A, Sood U, Ponnusamy K, Nagar S, Shakarad M, Negi RK, Singh Y, Khurana JP, Gilbert JA, Lal R (2017) Comparative genomic analysis reveals habitat-specific genes and regulatory hubs within the genus *Novosphingobium*. *mSystems* 2:e00020-17. <https://doi.org/10.1128/mSystems.00020-17>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.