Chapter 2
Evolutionary Relationship of Penicillin-Binding Protein 2 Coding *penA* Gene and Understanding the Role in Drug-Resistance Mechanism Using Gene Interaction Network Analysis

Sravan Kumar Miryala, Anand Anbarasu and Sudha Ramaiah

Abstract The class A β-lactamase *penA* gene codes for penicillin-binding protein 2 (PBP2) which plays an important role in assembling the peptidoglycans on the outer side of the plasma membrane. The alteration in the structure of PBP2 protein makes the pathogen to gain resistance against penicillin. Thus, it is important to understand the role of drug-resistant mechanism by *penA* gene to develop potent drugs against penicillin-resistant pathogenic strains. In our study, we have used gene interaction network analysis of *penA* gene in various bacteria to understand its role in drug-resistant mechanisms. We have collected a total of 1039 interactions from 28 organisms available from STRING database. The *penA* gene interaction network was constructed using Cytoscape 3.6.1. The network analysis has shown that, along with *penA* gene, the genes *murG, ftsW, murC, ftsA,* and *ftsQ* are observed to have more number of interactors and they may be considered as the key candidates to understand the *penA* drug-resistant mechanism. Functional enrichment analysis has shown the important GO terms and pathways where *penA* gene plays an important role. We have also elucidated the evolutionary relationship of *penA* gene in various Gram-positive and Gram-negative bacteria. Our study helps in understanding the drug-resistant patterns of *penA* gene in various bacteria and also their evolutionary relationships.

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-981-13-7968-0_2) contains supplementary material, which is available to authorized users.

S. K. Miryala · A. Anbarasu · S. Ramaiah (✉)
Medical and Biological Computing Laboratory, School of Biosciences and Technology, Vellore Institute of Technology (VIT), Vellore 632014, Tamil Nadu, India
e-mail: sudhaanand@vit.ac.in

S. K. Miryala
e-mail: miryalasravankumar@vit.ac.in

A. Anbarasu
e-mail: aanand@vit.ac.in

B. Subramanian et al. (eds.), *Emerging Technologies for Agriculture and Environment*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-13-7968-0_2

Keywords Penicillin-binding protein 2 · *penA* gene · Antimicrobial resistance · Gene interaction network · Functional enrichment analysis

2.1 Introduction

Multidrug resistance (MDR) in pathogenic bacteria is a serious problem and becomes a public health threat worldwide. MDR in bacteria occurs mainly by the accumulation of antimicrobial-resistant (AMR) genes on resistance plasmid or transposons. In general, AMR genes code for resistance to a specific agent, but when the genes accumulate in bacterial plasmid, the bacteria show resistance to multiple drugs and make the treatment more critical [1]. Penicillin-binding proteins (PBPs) are one among such a class of proteins which shows reduced susceptibility to penicillin and other beta lactams. Penicillin-binding proteins (PBPs) are the enzymes that assemble the peptidoglycans, which are the main constituents of the cell wall on the outer side of the plasma membrane and help the bacteria to resist the intercellular pressure caused by any external factors such as antimicrobial drugs. Because of the important role in cell wall maintenance, PBP proteins are considered as the major molecular targets for β-lactam antibiotics. The β-lactams such as penicillin inhibit the transpeptidase activity and thus inhibit the peptidoglycan cross-links in the bacterial cell wall [2, 3]. The pathogens acquires resistance to β-lactam drugs by developing more strategies and makes the treatment difficult and cause life threatening disease. Several studies have reported that the alterations in the structure of PBP proteins are the reason for the resistance to penicillin [4, 5]. The β-lactam resistance in Gram-positive bacteria is via two main mechanisms: One mechanism is by enzymatic degradation through the production of β-lactamases, and other mechanism is by the decreasing the affinity of the antibiotics for its target [2], whereas in the Gram-negative bacteria, the β-lactamase-mediated resistance is due to either acquisition of new genes or the mutations affecting the expression of its chromosomal β-lactamases [6].

In our present study, we have analyzed the gene interaction network of the extended spectrum beta lactamase (ESBL), *penA* gene. *penA* gene belongs to the class A of beta lactamases and codes for the protein, penicillin-binding protein 2 (PBP2). Various studies have confirmed that penicillin resistance due to *penA* alleles has been arisen from either recruitment of sequence blocks from natural resistance species or by point mutations such as codon insertion or substitution [3]. The gene interaction network (GIN) approaches are becoming famous and drawing the scientific community attention and are considered to be the more reliable approach to study the antimicrobial-resistant mechanisms in pathogenic bacteria. The GIN analysis in AMR genes provides new insights to various drug-resistant mechanisms by analyzing the AMR genes along with their functional partners [7–11]. We have also collected the *penA* protein sequences from various bacterial strains and constructed the phylogenetic tree to understand the phylogenetic relationship between the bacterial *penA* gene among the bacteria. We have collected functional interactions

of *penA* gene from both Gram-positive and Gram-negative bacteria and constructed gene interaction network to understand the molecular-level interactions of *penA* gene with the functional partners. We further analyzed the functional enrichment of genes in the network to understand the role of *penA* gene along with its functional partners. We have also constructed a phylogenetic tree of *penA* gene from different bacteria to understand the evolutionary relationship between various bacterial species. Our results will be helpful in better understanding the role of *penA* gene in drug-resistant mechanism in various pathogenic bacteria. The genes which play an important role in various biological pathways were reported, and they can be useful as potent drug targets in new drug discovery.

2.2 Materials and Methods

2.2.1 Gene Interaction Data Curation from STRING Database

We have collected the *penA* gene interactions from STRING database from different Gram-negative and Gram-positive bacteria. STRING databases consist of protein–protein interactions curated from various sources such as high-throughput experimental data, data mining, literature survey, and co-expression analysis studies. All the interactions are broadly classified as direct (physical interactions curated from laboratory techniques) and indirect (functional associations extracted from computational prediction). Each interaction is provided with a confidence score that lies between 0 and 1. Each interaction further classified based on the confidence scores as highest (≥0.90–1.0), high (≥0.70–0.89), medium (≥0.40–0.69), and low (≥0.15–0.39) [12].

2.2.2 Interaction Network Construction

We have used Cytoscape 3.6.1 for the gene interaction network construction and analysis. Cytoscape is an open-source visualization tool used for constructing molecular interaction networks and biological pathways. Cytoscape tool comes with the core distribution with basic features for data integration, visualization, and analysis. The additional features such as network and molecular profiling analysis, new layouts, additional file format support, and cross-reference with databases can be available as Cytoscape plugins or apps [13].

2.2.3 Clustering Analysis

Clustering analysis of interacting network was carried out by using Cytoscape plugin MCODE. MCODE is based on molecular complex detection (MCODE) algorithm. The algorithm operates mainly in three different stages such as vertex weighing, complex prediction, and optionally postprocessing. In an interaction network, every vertex is a molecule and edge is an interaction between the molecules. Based on the type of data used, the graphs are divided into direct (known cell signaling and known pathways) and indirect graph [14].

2.2.4 Shortest Path Length and Closeness Centrality Analysis

The interaction network of *penA* gene was analyzed by using Cytoscape app NetworkAnalyzer. It is used to compute and display topological parameters such as a number of nodes, connecting edges, the network diameter, density, radius, centralization, heterogeneity, clustering coefficient, the characteristic path length, the distribution of node degrees, neighborhood connectivity, average clustering coefficients, and the shortest path lengths. It helps in analyzing both the types of networks (directed and undirected) and also allows to construct the intersection or union of two networks [15].

2.2.5 Phylogenetic Tree Construction

For the phylogenetic tree construction, we have used MEGA7. The phylogenetic tree construction involves in multiple sequence alignment (MSA) using MEGA inbuilt tools ClustalW or MUSCLE and followed by the constriction of phylogenetic tree from the aligned sequences. The gaps and the missing data from all the positions were removed from phylogenetic analysis. For phylogenetic tree construction, there are various methods used such as neighbor-joining method, maximum likelihood method, evolutionary distance method, and maximum parsimony method [16].

2.3 Results

2.3.1 penA *Gene Interaction Network Analysis*

Network analysis of *penA* gene along with the functional partners was done using STRING database. A total of 144 *penA* gene interactors with 1039 functional interactions from 28 bacterial strains were curated with medium (>0.4) confidence

scores. Out of 28 bacterial species used in this study, 20 belong to Gram-negative (*Achromobacter piechaudii, Ahrensia* sp. R2A130*, Bordetella petrii, Burkholderia cenocepacia, Burkholderia mallei, Burkholderia pseudomallei, Campylobacter coli, Campylobacter jejuni* 414*, C. jejuni* 81176*, C. jejuni* NCTC11168*, Collimonas fungivorans, Kingella kingae, Legionella pneumophila* Paris, *Neisseria lactamica, Neisseria meningitidis, N.* C102*, Neisseria* sp. F0314*, Oligotropha carboxidovorans, Oxalobacter formigenes* OXCC13*, Rhodopseudomonas palustris* CGA009) and 8 belong to Gram-positive (*Brevibacillus laterosporus, Lactobacillus antri, Saccharopolyspora erythraea, Streptococcus oralis* ATCC35037*, Streptococcus pneumoniae* D39*, S. pneumoniae* R6*, S. pneumoniae* TIGR4). A list of interactions from each bacterial species is provided in Table 2.1. There are no significant interactions with medium and above confidence scores available in STRING for the bacterial species *A. piechaudii*. To obtain the maximum number of *penA* gene interactors from all the possible bacteria, we have curated the low confidence score interactions in the case of *A. piechaudii* (Fig. 2.1). Out of 1039 functional interactions, there are 522, 269, 229, and 19 belong to highest, high, medium, and lowest confidence scores, respectively. Out of 144 functional interactors, 130 directly interact with *penA* gene and are highlighted in red color in the network (Fig. 2.2).

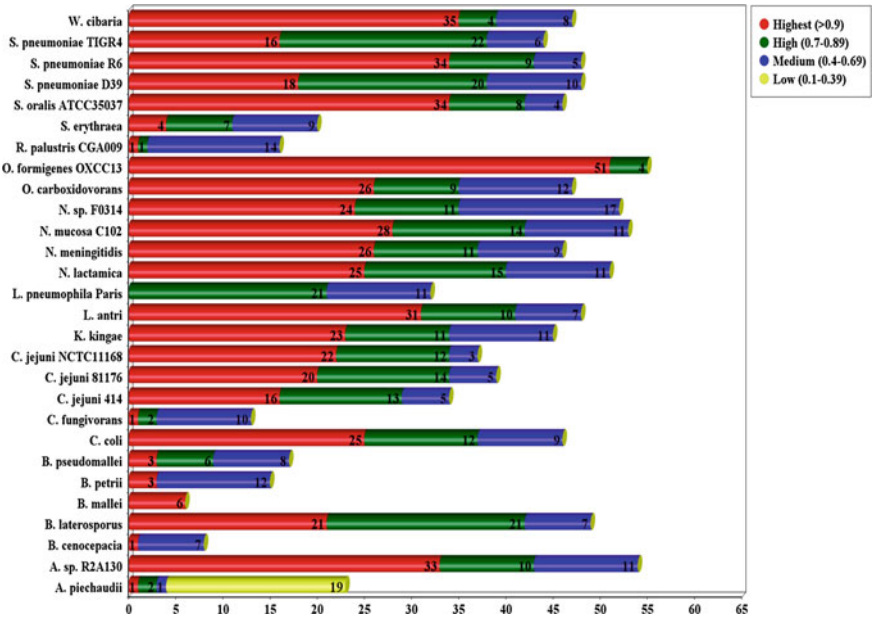

Fig. 2.1 List of *penA* gene interactions collected from STRING database for different organisms. The interactions are given scores as highest (≥0.90–1.0), high (≥0.70–0.89), medium (≥0.40–0.69), and low (≥0.15–0.39) confidence scores. Out of 28 bacterial species, 27 have interactions with medium and above medium confidence scores. In *A. piechaudii*, we have collected interactions with low confidence scores as there are not many interactions above the medium confidence scores

Table 2.1 List of bacteria along with the number of interactions extracted from STRING database

Organism	Combined score			Total number of interactions
	Highest (0.9–1)	High (0.7–0.89)	Medium (0.4–0.69)	
Oxalobacter formigenes OXCC13	51	4	0	55
Ahrensia sp. R2A130	33	10	11	54
Neisseria mucosa C102	28	14	11	53
Neisseria sp. F0314	24	11	17	52
Neisseria lactamica	25	15	11	51
Brevibacillus laterosporus	21	21	7	49
Lactobacillus antri	31	10	7	48
Streptococcus pneumoniae D39	18	20	10	48
Streptococcus pneumoniae R6	34	9	5	48
Oligotropha carboxidovorans	26	9	12	47
Weissella cibaria	35	4	8	47
Campylobacter coli	25	12	9	46
Neisseria meningitidis	26	11	9	46
Streptococcus oralis ATCC35037	34	8	4	46
Kingella kingae	23	11	11	45
Streptococcus pneumoniae TIGR4	16	22	6	44
Campylobacter jejuni 81176	20	14	5	39
Campylobacter jejuni NCTC11168	22	12	3	37
Campylobacter jejuni 414	16	13	5	34

(continued)

Table 2.1 (continued)

Organism	Combined score			Total number of interactions
	Highest (0.9–1)	High (0.7–0.89)	Medium (0.4–0.69)	
Legionella pneumophila Paris	0	21	11	32
Achromobacter piechaudii	1	2	1	4
Saccharopolyspora erythraea	4	7	9	20
Burkholderia pseudomallei	3	6	8	17
Rhodopseudomonas palustris CGA009	1	1	14	16
Bordetella petrii	3	0	12	15
Collimonas fungivorans	1	2	10	13
Burkholderia cenocepacia	1	0	7	8
Burkholderia mallei	6	0	0	6

The interactions are classified as highest, high, and medium based on the confidence scores. Out of 29 bacterial species, 28 have interactions with medium and above medium confidence scores. But there are no interactions with desired confidence scores available in the bacterial species *A. piechaudii*

2.3.2 Network Analysis Using NetworkAnalyzer

NetworkAnalyzer is used for network analysis of *penA* gene. For each node, individual centrality scores, along with the average shortest path length and degree, were given (Supplementary File S1). Among the 145 genes used in the study the top 20 genes with direct interactions (degree), average shortest path length score, closeness centrality, and between centrality scores have been provided in Table 2.2. All the 145 nodes are arranged in 3 layers based on the number of individual direct interactions. There are 37 nodes with more than 10 interactions, 53 nodes with interactions in between 05–10, and 45 nodes with less than 05 direct interactions (Fig. 2.3).

2.3.3 Clustering Analysis Using MCODE

MCODE has been resulted in 12 highly interconnected clusters. Out of 145 genes used in the interaction network, 90 genes were included in 12 clusters and the clusters are named as C1–C12. We have used default MCODE parameters for filtration of

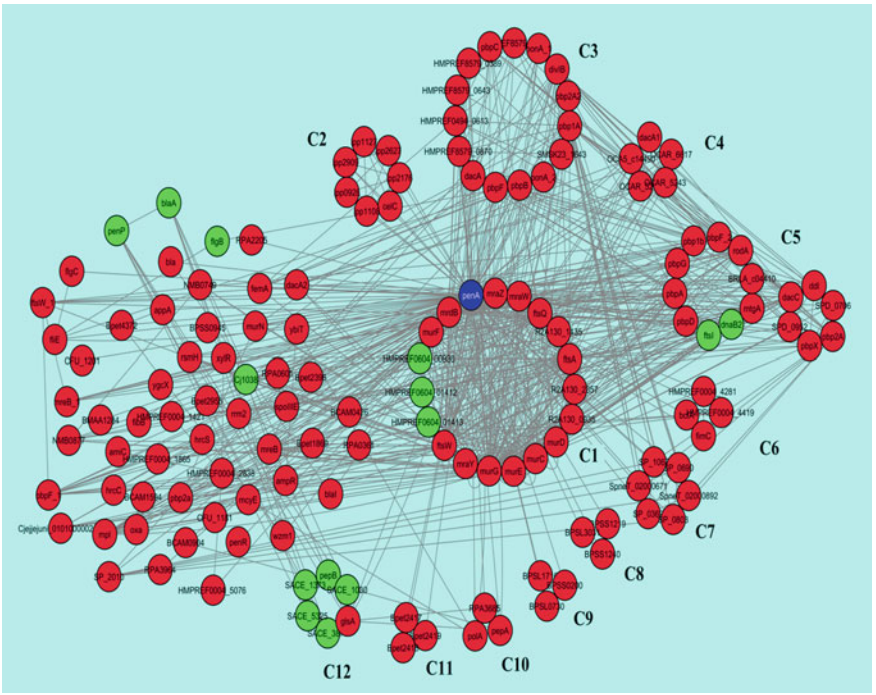

Fig. 2.2 ***penA* gene interaction network along with the functional interactors**. Genes are clustered into 12 densely interconnected clusters (C1–C12) using Cytoscape MCODE. For easy identification, *penA* gene is given in blue color, red-colored nodes are the direct interactors of the *penA* gene, and those nodes which are not direct interactors have given green color

nodes and edges. Among the clusters, there are only C1, C3, and C5 clusters having more than 15 nodes. The *penA* gene was included in cluster C1, and it consists of 19 nodes and 171 edges with 13.44 MCODE score (Table 2.3).

2.3.4 Functional Enrichment Analysis Using STRING Database

All the genes from 12 clusters are analyzed for the functional enrichment of genes using STRING database. Various GO terms such as biological processes (BP), molecular functions (MF), and cellular components are enriched along with the KEGG pathway-related genes, PFAM, and InterPro domain-related genes. Out of 145 genes, 49 genes are enriched in BP, MF, CC, KEGG, PFAM, and InterPro domains (Fig. 2.4). The clusters C2, C6, and C12 have been observed with no significant enrichment results (Supplementary File S2). Cluster C1 consists of genes related to mur ligase family-related genes and peptidoglycan synthase-related genes. The enriched GO

Table 2.2 Network analysis using NetworkAnalyzer tool

Genes	Degree	Avg. shortest path length	Closeness centrality	Betweenness centrality
penA	142	1.125926	0.888158	0.767209
murG	75	1.62963	0.613636	0.056708
ftsW	68	1.681481	0.594714	0.036728
murC	52	1.792593	0.557851	0.018124
ftsA	48	1.807407	0.553279	0.01702
ftsQ	36	1.881481	0.531496	0.006286
mrdB	35	1.859259	0.537849	0.00867
murE	34	1.874074	0.533597	0.007444
pbp1A	29	1.874074	0.533597	0.00944
rodA	29	1.881481	0.531496	0.006594
dacA	27	1.888889	0.529412	0.008643
divIB	27	1.888889	0.529412	0.007284
mtgA	24	1.940741	0.515267	0.005793
murF	24	1.962963	0.509434	0.002863
mraW	23	1.962963	0.509434	0.001896
murD	23	1.925926	0.519231	0.002166
mraY	22	1.948148	0.513308	0.000697
pbp2A	19	1.933333	0.517241	0.003368
pbpB	19	1.925926	0.519231	0.013447
mpl	18	1.985185	0.503731	0.002688

The top 20 genes with the degree, average shortest path length, closeness centrality, and betweenness centrality scores are given. The average shortest path length gives the expected distance between the two connected nodes, and genes with shortest path length and high closeness centrality are considered as the controlling points of molecular communication. Smaller edge betweenness values indicate the stronger interactions

terms include the BP such as regulation of cell shape (GO.0008360), peptidoglycan biosynthetic process (GO.0009252), cell wall organization (GO.0071555), cellular component biogenesis (GO.0044085), nitrogen compound metabolic process (GO.0006807), MF acid-amino acid ligase activity (GO.0016881) and various CC's related to cell (GO.0005623), cell part GO.0044464), intracellular (GO.0005622), cytoplasm (GO.0005737). PFAM and InterPro domains such as mur ligase family, glutamate ligase domain cell cycle protein related are enriched in cluster C1. The genes *mraY, mraZ, murC, murD, murG,* and *rsmH* are found in multiple GO terms and protein domain-related enriched entities.

In cluster C3, there are no significant GO terms enriched, but the genes related to PFAM, InterPro domains related to transglycosylase (PF00912), penicillin-binding protein transpeptidase domain (PF00905), D-alanyl-D-alanine carboxypeptidase, penicillin-binding protein 5-C terminal domain (PF07943),

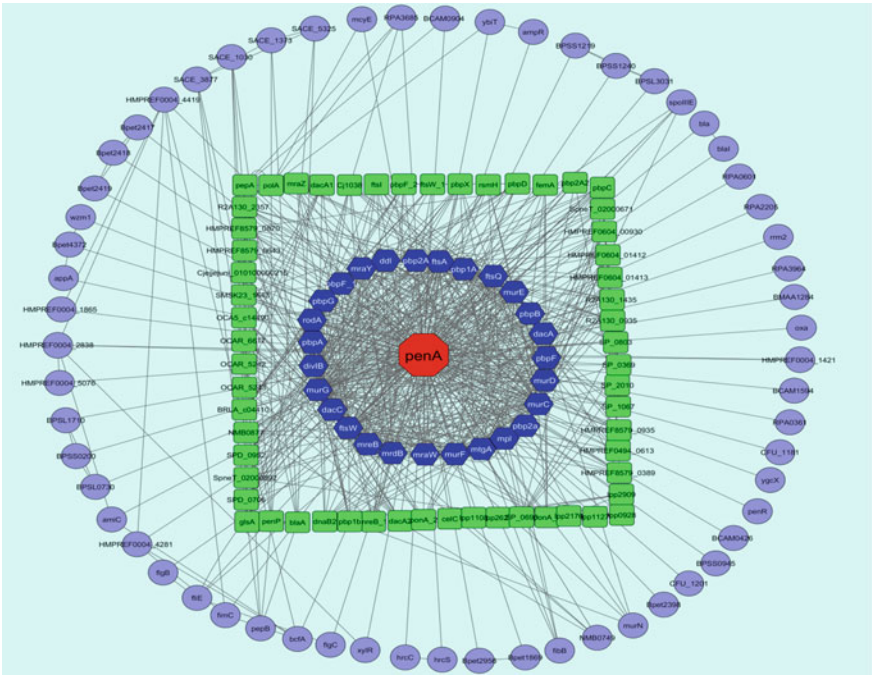

Fig. 2.3 ***penA*** **gene interactions**: All the 145 nodes are arranged in three layers based on the number of individual direct interactions. *penA* gene is denoted with octagonal red color node followed by 37 nodes with more than 10 interactions in hexagonal blue color node followed by 53 nodes with interaction in between 05 and 10 in rounded edge rectangular green color nodes and 45 nodes with less than 5 direct interactions

beta lactamase/transpeptidase-like (IPR012338), glycosyl transferase family 51 (IPR001264), and penicillin-binding protein transpeptidase (IPR001460) were enriched. The cluster C4 genes are enriched with the KEGG pathway peptidoglycan biosynthesis (KEGG ID: 550) along with PFAM and InterPro domains related to penicillin-binding protein transpeptidase domain and transglycosylase related domains. In cluster C5, the BPs related to regulation of cell shape (GO.008360) and peptidoglycan biosynthetic process (GO.009252) and KEGG pathways such as peptidoglycan biosynthesis (KEGG ID: 550) and beta lactam resistance (KEGG ID: 312) were enriched. Along with protein domains related to penicillin-binding protein dimerization, penicillin-binding protein transpeptidase was also enriched. Whereas in cluster C7, cell cycle protein-related PFAM and InterPro domains are enriched.

In cluster C8, KEGG pathways beta lactam resistance (KEGG ID: 312), peptidoglycan biosynthesis (KEGG ID: 550), and protein domains related to penicillin-binding protein dimerization and penicillin-binding protein transpeptidase domains were enriched. In cluster C9, the KEGG pathways penicillin and cephalosporin biosynthesis (KEGG ID: 311) and biosynthesis of secondary metabolites were enriched. The protein domains penicillin amidase, penicillin acylase, and nucleophile

Table 2.3 Clustering analysis of *penA* gene interactions

Cluster	Score	Nodes	Edges	Node IDs
C1	13.444	19	171	*ftsA, ftsQ, mraW, mraZ, R2A130_2357, murE, murF, penA, ftsW, mraY, mrdB, murG, murC, murD, HMPREF0604_01413, HMPREF0604_01412, R2A130_1435, R2A130_0935, HMPREF0604_00930*
C2	7	7	21	*lpp0928, lpp2176, lpp2627, lpp1127, lpp2909, lpp1108, celC*
C3	5.143	15	37	*divIB, HMPREF8579_0389, HMPREF8579_0643, ponA_1, pbp1A, pbp2A2, pbpB, HMPREF8579_0870, ponA_2, dacA, pbpF, HMPREF8579_0935, SMSK23_1643, pbpC, HMPREF0494_0613*
C4	4.5	5	9	*OCAR_5242, OCAR_6617, OCA5_c14490, OCAR_5243, dacA1*
C5	4	16	30	*mtgA, ddl, dnaB2, ftsI, pbpG, pbpF_2, BRLA_c04410, rodA, dacC, pbp1b, SPD_0952, pbpX, pbpD, pbpA, pbp2A, SPD_0706*
C6	4	4	6	*HMPREF0004_4281, bcfA, fimC, HMPREF0004_4419*
C7	3.6	6	9	*SP_0369, SP_0690, SpneT_02000671, SP_0803, SP_1067, SpneT_02000892*
C8	3	3	3	*BPSL3031, BPSS1240, BPSS1219*
C9	3	3	3	*BPSS0200, BPSL1710, BPSL0730*
C10	3	3	4	*pepA, RPA3685, polA*
C11	3	3	3	*Bpet2419, Bpet2417, Bpet2418*
C12	2.8	6	7	*glsA, SACE_3877, SACE_5325, pepB, SACE_1373, SACE_1030*

MCODE has resulted in 12 densely interconnected clusters. Each cluster along with the MCODE scores, nodes, and edges is provided. Clusters are named as C1−C12 for convenience

aminohydrolases related PFAM and InterPro domains were enriched. In cluster C10 there is only one KEGG pathway glutathione metabolism (KEGG ID: 480) was enriched. In cluster C12, the KEGG pathways penicillin and cephalosporin biosynthesis (KEGG ID: 311), beta lactam resistance (KEGG ID: 312), and glutathione metabolism (KEGG ID: 480) were enriched along with cytosol aminopeptidase family domains and beta lactam transpeptidase-like domains.

2.3.5 Phylogenetic Tree Construction and Analysis

We have collected the protein sequences for *penA* gene for 28 bacterial strains. Out of 28 bacterial strains, we have collected available 23 protein sequences from NCBI GenBank database. By using MEGA-muscle MSA tool, we have aligned the 23 amino acid sequences. We have used the aligned sequences for constructing phylogenetic

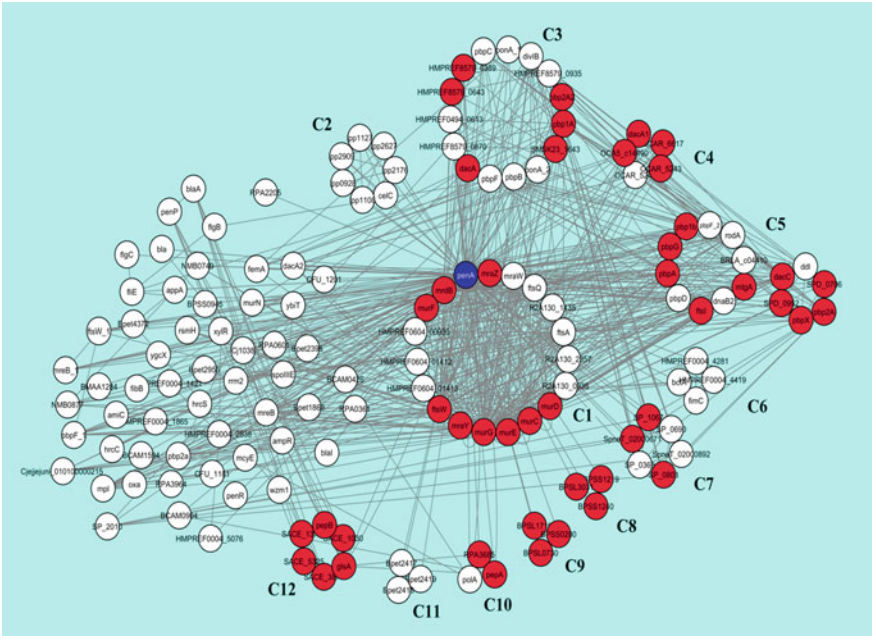

Fig. 2.4 ***penA* gene interaction network functional enrichment analysis**: Out of 145 genes in the network, 49 genes are functionally enriched with biological processes, molecular functions, cellular compartments, KEGG, PFAM, and InterPro domain-related genes. All the 49 enriched genes are highlighted in red color

tree (Fig. 2.5). We have used neighbor-joining method with the bootstrap consensus tree inferred from 1000 replicates. The evolutionary distances were computed using the poisson correction method and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. There were a total of 207 positions in the final dataset.

2.4 Discussion

Multidrug resistance exerted by the pathogenic bacteria has become a global threat for treatment. PBP proteins assemble at the cell wall of the bacteria and help bacteria to resistant the intracellular pressure caused by external agents such as antimicrobials. In our present study, we have used PBP protein-coding gene *penA* to understand the molecular-level interactions of PBP in various bacteria. The mutation in *penA* gene is associated with the resistance to tetracycline/azithromycin, spectinomycin, and ceftriaxone. In *N. gonorrhoeae,* the mutation (insertion of aspartate at 345 position) in *penA* gene along with the *ponA, gyrA,* and *parC* genes determines the susceptibility

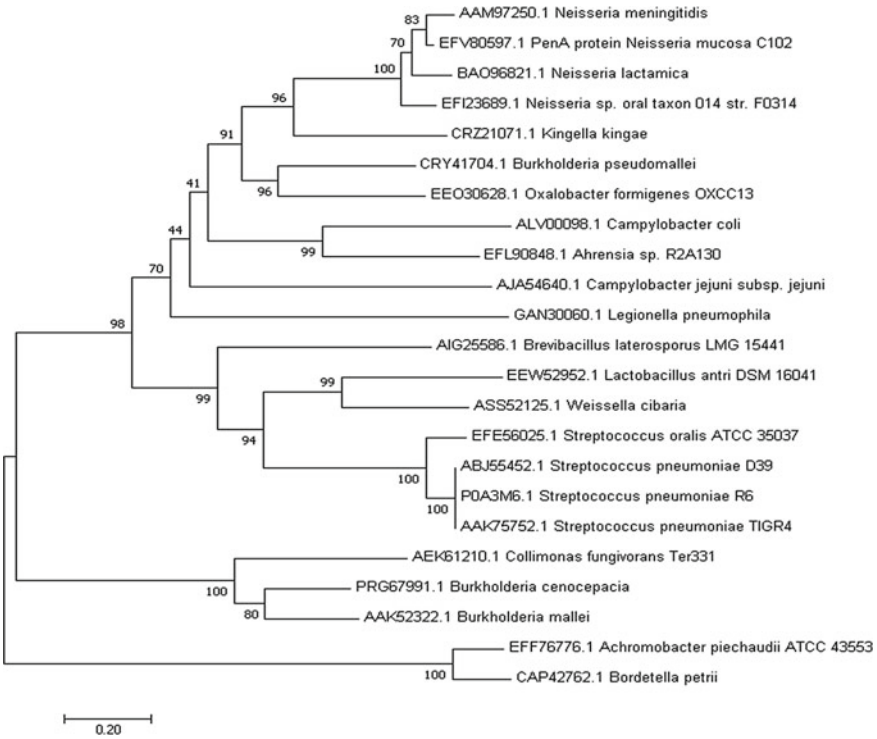

Fig. 2.5 Phylogenetic analysis of *penA* related proteins: Out of 28 bacteria, we have curated 23 available protein sequences and multiple sequence alignment is carried out by using MEGA7-muscle MSA program. The evolutionary history was inferred using the neighbor-joining method. The optimal tree with the sum of branch length = 9.64624886. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches

to penicillins, tetracyclines, and fluoroquinolones [17, 18]. In our results, we have observed that the *penA* gene with 142 total interactions constitutes 13.66% of total interactions in the whole interaction network followed by *murG* (75 interactions; 7.22%), *ftsW* (68 interactions; 6.54%), *murC* (52 interactions; 5%), *ftsA* (48 interactions; 4.62%), and *ftsQ* (36 interactions (3.46%). Along with the *penA* gene, these five genes constitute 40% of total interactions from the whole interaction network. With the maximum number of direct interactions, these nodes may be considered to have a high impact on the network; thus, these can be considered as functional hub nodes and further used as potent drug targets.

The functional enrichment analysis of genes in the interaction network has been shown that the *penA* gene along with the functional interactors is mainly responsible for peptidoglycan biosynthesis, mur ligase synthesis, β-lactam resistance, and glutathione metabolism. It is well known that Gram-positive and Gram-positive bacterial peptidoglycan composition is similar, while Gram-positive bacteria consist of

more thick and cross-linked peptidoglycan layer. Peptidoglycan is the main component of the bacterial cell wall, and it plays a key role in cell shape maintenance, facilitates the attachment for surface-exposed virulence factors, and avoids the modification in internal osmotic pressure. Thus, peptidoglycan biosynthesis mechanism-related genes are the preferred targets in the discovery of new antibiotics for many decades. The antibiotic-resistant mechanisms developed by the pathogenic bacteria against the antibiotics that targets the precursors of bacterial cell wall and biosynthesis machinery-related genes draw the attention of research community in overcoming the resistance strategies exhibited by the pathogens. There is an immediate necessity to study the resistance strategies and identify the new potent drug targets [19]. In our present study, we have observed the genes *OCAR_5243, dacC, BPSL3031, OCAR_6617, ddlB, BPSS1219, ftsI, BPSS1240, mtgA,* and *pbpA* are related to peptidoglycan biosynthesis, and these genes are found to have more dense interactions within the network. There are many clinically used antibiotics, especially various β-lactams, glycopeptides, fosfomycin, and cycloserine, which targets the genes involved in peptidoglycan biosynthesis pathway. Although the amide ligases (*murC, murD, murE,* and *murF*) play an important role in peptidoglycan biosynthesis by catalyzing the non-ribosomal peptide bond formations for the addition of peptide moiety on the peptidoglycan blocks, there are no antibacterial agent targets these amide ligases [20, 21]. In our results, the amide ligase-coded genes *murC, murD, murE,* and *murF* are enriched with Pfam and InterPro domains. It is also noticed that the MUR family genes *murG, murC,* and *murE* genes were in the top 10 genes with more number of direct interactions, which show the importance of amide genes in resistance caused by the *penA* gene along with its interactors. The other notable result is the enrichment of KEGG pathway glutathione metabolism (KEGG ID: 480). The genes *pepA, pepB, RPA3685,* and *SACE_1030* were involved in glutathione metabolism. Various studies have shown that the antibiotics used against the pathogenic Gram-positive and Gram-negative bacteria results in the formation of reactive oxygen species (ROS). When the cell is exposed to an antibiotic, glutathione plays a key role in maintaining the cellular redox poise by the detoxification the xenobiotics [22]. Another interesting observation in our results is the enrichment of KEGG pathway, β-lactam resistance (KEGG ID: 312), which includes the genes *pbp1B, BPSL3031, SACE_1373, pbp2A, BPSS1219, SACE_3877, pbpX,* and *BPSS1240.* The gene *pbp1B* is a peptidoglycan glycosyltransferase or murein synthase, and it plays an essential role in synthesizing peptidoglycan in the absence of a pre-existing template [23]. Whereas the gene *pbp2A* is found in *staphylococcal* cell wall biosynthesis along with the known PBP family proteins *PBP2* and *PBP4*, *PBP2A* which is an acquired transpeptidase plays a crucial role in susceptibility to antimicrobial agents [24]. *pbpX* gene is a homologous to the pneumococcal *PBP2x* gene, and it is isolated from penicillin-sensitive *S. oralis* strain [25]. The genes *SACE_1373* and *SACE_3877* belong to *S. erythraea* strain and coded for beta lactamase. *The genes BPSL3031, BPSS1219, and BPSS1240* belong *to B. pseudomallei* strain. BPSL3031 gene codes for peptidoglycan synthase, and BPSS1219 and BPSS1240 code for penicillin-binding protein.

Out of 28 bacterial *penA*-related proteins, there are only 23 sequences available in NCBI GenBank database. We have collected the 23 *penA*-related protein sequences and used for the construction of phylogenetic tree. Among the bacterial species used for the study, few are non-pathogenic, and most of them are pathogenic. The phylogenetic analysis of *penA* genes from various bacterial species has confirmed the lineage of different bacteria with respect to *penA* gene. We have used neighbor-joining method for the phylogenetic tree construction with 1000 bootstraps. In the phylogenetic tree, the bootstrap values more than 70% show more confidence and less than that show poor confidence of the phylogeny with respect to the entry. In our results, there are only two bootstrap values less than 70%. The constructed phylogenetic tree shows that the bacteria *C. jejuni* showed boot strap values 44% with the other Gram-negative bacteria such as *Neisseria* genus and *K. kingae.* The bacteria belonging to *Neisseria* genus are closely related (bootstrap values 96%) to *K. kingae*. The bacteria *B. pseudomallei* and *O. formigenes* OXCC13 have bootstrap value 96%. The Gram-positive bacteria used in the study have shown the bootstrap values more than 90%, which shows the confidence of the *penA* gene lineage among the Gram-positive bacteria. All the *streptococcus*-related *penA* proteins have shown bootstrap values 100%, and they are highly similar to other Gram-positive bacterial *penA* protein sequences. We have also observed that the bacteria belonging to the same genus showed less than 100% bootstrap values. The bacteria *B. mallei* and *B. cenocepacia* show the bootstrap values 80%; *N. meningitidis* and *Neisseria mucosa* C102 show 83%; and with respect to *N. lactamica* shows, the bootstrap value is 70%. Our phylogenetic studies have provided a detailed lineage of *penA* gene in various pathogenic and non-pathogenic bacteria.

2.5 Conclusion

The gene interaction analysis of PBP2 coding *penA* gene provides a comprehensive evidence on *penA* gene and their functional partners in antibiotic resistance in various bacteria. In our present study, we have identified functional interactors of the *penA* gene from different bacterial species, and by using functional enrichment analysis, we have analyzed the role of these genes in peptidoglycan biosynthesis, mur ligase synthesis, β-lactam resistance, and glutathione metabolism. The constructed *penA* gene network helps in understanding the functional relationship of these interactors in biological pathways. Our results give critical information on various biological processes such as gene functions and complex cellular mechanisms. The phylogenetic tree of *penA* related amino acids from various bacteria gives a glance of lineage of *penA* gene in various Gram-positive and Gram-negative bacteria. To conclude, from our study, we have observed the *penA* gene along with the interactors plays a major role in peptidoglycan biosynthesis, amide ligase biosynthesis, β-lactam resistance, and glutathione metabolism. Our results will help in better understanding the functional role of β-lactamase *penA* gene in β-lactam induction. The molecular interactions of *penA* along with the functional partners will be useful for researchers

exploring the β-lactam-mediated antibiotic resistance in pathogenic bacteria, and the identified resistance genes play major roles in various biological processes, and these can be considered as potent drug targets for developing new drugs.

Acknowledgements The authors gratefully acknowledge the Indian Council of Medical Research (ICMR), Government of India agency for the research grant (IRIS ID: 2014-0099). MSK thanks ICMR for the research fellowship. The authors would like to thank the management of VIT for providing the necessary facilities to carry out this research work.

Conflict of Interest Statement None declared.

References

1. Nikaido, H.: Multidrug resistance in bacteria. Annu. Rev. Biochem. 119–146 (2009). https://doi.org/10.1146/annurev.biochem.78.082907.145923.multidrug
2. Munita, J.M., Bayer, A.S., Arias, C.A.: Evolving resistance among gram-positive pathogens. Clin. Infect. Dis. **61**, S48–S57 (2015). https://doi.org/10.1093/cid/civ523
3. Zapun, A., Morlot, C., Taha, M.-K.: Resistance to β-Lactams in *Neisseria ssp* due to chromosomally encoded penicillin-binding proteins. Antibiotics **5**, 35 (2016). https://doi.org/10.3390/antibiotics5040035
4. Thulin, S., Olcén, P., Fredlund, H., Unemo, M.: Total variation in the penA gene of *Neisseria meningitidis*: correlation between susceptibility to β-lactam antibiotics and penA gene heterogeneity. Antimicrob. Agents Chemother. **50**, 3317–3324 (2006). https://doi.org/10.1128/AAC.00353-06
5. Powell, A.J., Tomberg, J., Deacon, A.M., et al.: Crystal structures of penicillin-binding protein 2 from penicillin-susceptible and -resistant strains of *Neisseria gonorrhoeae* reveal an unexpectedly subtle mechanism for antibiotic resistance. J. Biol. Chem. **284**, 1202–1212 (2009). https://doi.org/10.1074/jbc.M805761200
6. Sanders, C.C., Sanders, W.E.: β-Lactam resistance in gram-negative bacteria: global trends and clinical impact. Clin. Infect. Dis. **15**, 824–839 (1992). https://doi.org/10.1093/clind/15.5.824
7. Anitha, P., Anbarasu, A., Ramaiah, S.: Computational gene network study on antibiotic resistance genes of *Acinetobacter baumannii*. Comput. Biol. Med. **48**, 17–27 (2014). https://doi.org/10.1016/j.compbiomed.2014.02.009
8. Anitha, P., Anbarasu, A., Ramaiah, S.: Gene network analysis reveals the association of important functional partners involved in antibiotic resistance: a report on an important pathogenic bacterium *Staphylococcus aureus*. Gene **575**, 253–263 (2016). https://doi.org/10.1016/j.gene.2015.08.068
9. Parimelzaghan, A., Anbarasu, A., Ramaiah, S.: Gene network analysis of metallo beta lactamase family proteins indicates the role of gene partners in antibiotic resistance and reveals important drug targets. J. Cell. Biochem. **117**, 1330–1339 (2016). https://doi.org/10.1002/jcb.25422
10. Miryala, S.K., Ramaiah, S.: Exploring the multi-drug resistance in *Escherichia coli* O157 : H7 by gene interaction network : a systems biology approach. Genomics (2018) https://doi.org/10.1016/j.ygeno.2018.06.002
11. Miryala, S.K., Anbarasu, A., Ramaiah, S.: Discerning molecular interactions: a comprehensive review on biomolecular interaction databases and network analysis tools. Gene **642** (2018). https://doi.org/10.1016/j.gene.2017.11.028
12. Szklarczyk, D., Morris, J.H., Cook, H., et al.: The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. **45**, D362–D368 (2017). https://doi.org/10.1093/nar/gkw937

13. Shannon, P., Markiel, A., Ozier, O., et al.: Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2498–2504 (2003). https://doi.org/10.1101/gr.1239303.metabolite
14. Bader, G.D., Hogue, C.W.V.: An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics **4**, 1–27 (2003). https://doi.org/10.1186/1471-2105-4-2
15. Assenov, Y., Ramírez, F., Schelhorn, S.E.S.E., et al.: Computing topological parameters of biological networks. Bioinformatics **24**, 282–284 (2008). https://doi.org/10.1093/bioinformatics/btm554
16. Kumar, S., Stecher, G., Tamura, K.: MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. **33**, 1870–1874 (2016). https://doi.org/10.1093/molbev/msw054
17. Török, M.E., Chantratita, N., Peacock, S.J.: Bacterial gene loss as a mechanism for gain of antimicrobial resistance. Curr. Opin. Microbiol. **15**, 583–587 (2012). https://doi.org/10.1016/j.mib.2012.07.008
18. Ohnishi, M., Watanabe, Y., Ono, E., et al.: Spread of a chromosomal cefixime-resistant *penA* gene among different *Neisseria gonorrhoeae* lineages. Antimicrob. Agents Chemother. **54**, 1060–1067 (2010). https://doi.org/10.1128/AAC.01010-09
19. Nikolaidis, I., Favini-Stabile, S., Dessen, A.: Resistance to antibiotics targeted to the bacterial cell wall. Protein Sci. **23**, 243–259 (2014). https://doi.org/10.1002/pro.2414
20. Kouidmi, I., Levesque, R.C., Paradis-Bleau, C.: The biology of Mur ligases as an antibacterial target. Mol. Microbiol. **94**, 242–253 (2014). https://doi.org/10.1111/mmi.12758
21. Vollmer, W., Blanot, D., De Pedro, M.A.: Peptidoglycan structure and architecture. FEMS Microbiol. Rev. **32**, 149–167 (2008). https://doi.org/10.1111/j.1574-6976.2007.00094.x
22. Cameron, J.C., Pakrasi, H.B.: Glutathione facilitates antibiotic resistance and photosystem I stability during exposure to gentamicin in cyanobacteria. Appl. Environ. Microbiol. **77**, 3547–3550 (2011). https://doi.org/10.1128/AEM.02542-10
23. Ranjit, D.K., Jorgenson, M.A., Young, K.D.: PBP1B glycosyltransferase and transpeptidase activities play different essential roles during the de novo regeneration of rod morphology in *Escherichia coli*. J. Bacteriol. **199** (2017). https://doi.org/10.1128/jb.00612-16
24. Leski, T.A., Tomasz, A.: Role of penicillin-binding protein 2 (PBP2) in the antibiotic susceptibility and cell wall cross-linking of *Staphylococcus aureus*. J. Bacteriol. **2**, 1815–1824 (2005). https://doi.org/10.1128/jb.187.5.1815
25. Sibold, C., Henrichsen, J., König, A., et al.: Mosaic pbpX genes of major clones of penicillin-resistant *Streptococcus pneumoniae* have evolved from pbpX genes of a penicillin-sensitive *Streptococcus oralis*. Mol. Microbiol. **12**, 1013–1023 (1994). https://doi.org/10.1111/j.1365-2958.1994.tb01089.x