Research paper

Complete genome sequencing and comparative analyses of broad-spectrum antimicrobial-producing *Micromonospora* sp. HK10

Madhumita Talukdar\(^a,\,1\), Dhrubajyoti Das\(^a,\,1\), Chiranjeea Borah\(^a\), Tarun Chandra Bora\(^a\), Hari Prasanna Deka Boruah\(^a\), Anil Kumar Singh\(^a,\,b,\,*\)

\(^a\) Department of Biotechnology, CSIR-North East Institute of Science and Technology, Jorhat 785006, India
\(^b\) Academy of Scientific and Innovative Research, Rafi Marg, New Delhi 110001, India

**ARTICLE INFO**

**Article history:**
Received 29 December 2015
Received in revised form 29 August 2016
Accepted 2 September 2016
Available online xxx

**Keywords:**
*Micromonospora*
Anti-microbial producing
Bacterial genomics
Secondary metabolite

**ABSTRACT**

*Micromonospora* genus produces > 700 bioactive compounds of medical relevance. In spite of its ability to produce high number of bioactive compounds, no genome sequence is available with comprehensive secondary metabolite gene clusters analysis for anti-microbial producing *Micromonospora* strains. Thus, here we contribute the full genome sequence of *Micromonospora* sp. HK10 strain, which has high antibacterial activity against several important human pathogens like, *Mycobacterium abcessus*, *Mycobacterium smegmatis*, *Bacillus subtilis*, *Staphylococcus aureus*, *Proteus vulgaris*, *Pseudomonas aeruginosa*, *Salmonella* and *Escherichia coli*. We have generated whole genome sequence data of *Micromonospora* sp. HK10 strain using Illumina NexSeq 500 sequencing platform (2 × 150 bp paired end library) and assembled it de novo. The sequencing of HK10 genome enables identification of various genetic clusters associated with known- and probably unknown- antimicrobial compounds, which can pave the way for new antimicrobial scaffolds.

© 2016 Published by Elsevier Ltd.

**1. Introduction**

*Micromonospora* is a genus of Gram-positive bacteria that taxonomically belong to the family of *Micromonosporaceae* with 63 species (Genilloud and Genilloud, 2012; Hirsch and Valdés, 2009; Genilloud, 2015). The members of *Micromonospora* genus are widely distributed in diverse geographical habitats e.g., soil, fresh water, marine sediment, rocks, nitrogen fixing nodules from leguminous and actinorhizal plants (Das et al., 2008; Charan et al., 2004; Trujillo et al., 2010; Valdés et al., 2005; Carro et al., 2013; Trujillo et al., 2006; Trujillo et al., 2007) etc. *Micromonospora* belongs to actinomycetes group, which are prolific producers of > 700 bioactive compounds including gentamicin and calicheamycin (Charan et al., 2004; Bérdy, 2005; Igarashi et al., 2011a; Igarashi et al., 2007; Igarashi et al., 2011b; Furumai et al., 2000; Furumai et al., 2002). Bioactive compounds produced by *Micromonospora* have been studied considerably for their critical roles in human health and medicines (Bérdy, 2005; Igarashi et al., 2011a; Igarashi et al., 2007; Igarashi et al., 2011b; Furumai et al., 2000; Furumai et al., 2002; Shimotohno et al., 1993; Tomita and Tamaoki, 1980; Lam et al., 1996). Advancement of new generation sequencing and improved accuracy of gene annotation, have opened up a new avenue for the discovery of novel secondary metabolite encoding gene clusters and their activities. Although *Micromonospora* genus belong to the order *Actinomycetales* along with *Sporomycetes* genus, which are well known for production of secondary metabolites, a limited numbers of studies have been focused on the secondary metabolites and the biosynthetic gene clusters of this genus. Despite the production of various useful antibiotics by *Micromonospora* species, the full genomes of only limited number of strains have been available to date. The genomes of *Micromonospora aurantiaca* ATCC 27029\(^1\) (accession no. NC_014391.1) and *Micromonospora* sp. ATCC 39149 (accession no. GG657738) isolated from soil and *Micromonospora* sp. L5 and *Micromonospora lupini* Lupac 08, isolated from nitrogen fixing root nodules of *Casuarina equisetifolia* (Valdés et al., 2005; Trujillo et al., 2014) and *Lupinus angustifolius* (Trujillo et al., 2010; Trujillo et al., 2014) respectively, are available in public database. Intrigued by its potential bioactivity, we contribute here a genome sequence of *Micromonospora* sp. strain HK10, a soil isolate from Kaziranga National Park, Assam, India (Talukdar et al., 2016). HK10 strain shows high antibacterial activity against both Gram-negative and Gram-positive bacteria (Talukdar et al., 2012). Since secondary metabolites play crucial roles in several ecological functions for example, chemical communication, nutrient attainment, defense mechanisms etc., we decided to analyze the genome sequences of HK10 strain and present these results in this manuscript.

**Abbreviation:** COG, clusters of orthologous groups; KEGG, Kyoto encyclopedia of genes and genomes; HGT, horizontal gene transfer; GI, genetic island; IS, insertion sequence; CRISPRs, clustered regularly interspaced short palindromic repeats; SRP, signal recognition particle; Sec, sec secretion system; PKS, polypeptide synthase; NRPS, non-ribosomal peptide synthetase

* Corresponding author at: CSIR-North East Institute of Science & Technology, Jorhat 785006, Assam, India.

Email address: 1010am@gmail.com (A.K. Singh)

* Co-first authors.

http://dx.doi.org/10.1016/j.gene.2016.09.005

0378-1119/© 2016 Published by Elsevier Ltd.
netic information related to its unknown active secondary metabolites and respective genetic determinants. HK10 genome was compared within other Micromonospora (Micromonospora sp. L5 and M. aurantiaca ATCC20029) and also with other actinobacteria such as, Salinispora arenicola CNS-205, Streptomyces coelicolor A3 (Hirsch and Valdés, 2009), Nocardioides brasiliensis ATCC 700358 and Rhodococcus pyridinivorans SB3094, which are known to produce bioactive compounds. The comparative genome analysis with other strains revealed some significant genetic differences. Furthermore, we compared various critical systems for their function and distributions of genes like two component regulatory systems, secretion systems and transporters mechanisms, which help the microbes to survive and adapt with their surrounding environments.

2. Material and methods

2.1. Strains, media and antagonism test

The antibacterial assay was performed against various test pathogens: Salmonella sp. (IDH377), enterotoxigenic Escherichia coli E9 (serotype O146) (Pazhani et al., 2011), Mycobacterium smegmatis ATCC700804 and Mycobacterium abscessus ATCC19977. The antibacterial activity of Micromonospora sp. HK10 was tested using agar diffusion assay method (Graner et al., 1976; Forbes et al., 1990) by observing presence of inhibition zone (in mm) around the well. For extraction of crude compound, the pure isolate was cultured into 500 ml Erlenmeyer flasks containing 100 ml of CSYP liquid medium (K-HPO4 0.5 g, casein 3 g, maize starch 10 g, peptone 1 g, malt extract 10 g, yeast extract 1 g, distilled water 1 l. pH 7.5) for different time intervals such as (7, 10, 12 and 14 days) at 28 ± 2 °C at 200 rpm. The culture broth was solvent extracted with equal volume of chloroform (1:1) (v/v) in separating funnels by shaking vigorously for 15 min and the solvent part was dried to yield the crude extract. The dried crude extract was dissolved in 20% dimethyl sulfoxide (DMSO) and antimicrobial bioassay was carried out against the test pathogens. 70 μL of fresh inoculum from each culture was spread on the plates. Wells were punctured on freshly spread bacterial culture on Muller Hinton Agar (HiMedia, India) using sterile cork borer and the bored wells were filled with the 100 μg/ml extract of Micromonospora sp. HK10. All the plates were incubated at 37 °C for 48 h. Growth inhibition of test pathogens was measured with standard scale; the mean values of inhibition zones were taken and compared with 50 μg/ml neomycin sulfate (HiMedia, India).

2.2. HK10 whole genome sequencing, assembly and annotation

The Micromonospora sp. HK10 genome was sequenced with the Illumina NextSeq 500 system using a paired-end 2×150 bp library. The raw data was trimmed using Trimmomatic version 0.30 (Bolger et al., 2014) for high-quality read length (cutoff quality score 20). A total of 8,896,194 high-quality, vector-filtered reads were used for assembly with CLC Genomics Workbench version 6 (CLC bio, Denmark) and the draft genome thus obtained was annotated for protein coding genes with the help of Glimmer (Delcher et al., 2007) and Prodigal v2.60 (Hyatt et al., 2010). This whole-genome shotgun project has been deposited at GenBank.

2.3. Complete genome analysis of Micromonospora sp. HK10

The genome sequence was submitted to the Integrated Microbial Genomes (IMG) server (http://img.jgi.doe.gov) of the Joint Genome Institute (JGI) for more detail analysis and genome comparison (Markowitz et al., 2014a; Markowitz et al., 2014b). A one-sample t-test was carried out by using OriginPro 8 software to evaluate the statistically significant differences of gene abundance in each COG category between Micromonospora sp. HK10 and other five genomes of IMG database: M. aurantiaca ATCC 27029 (accession no. NC_014391.1), Micromonospora sp. L5 (accession no. NC_014815.1), Micromonospora sp. ATCC 39149 (accession no. GG657738.1), M. lupini Lupac 08 (accession no. HF570108), and Micromonospora sp. CNB394 (Biosample no. SAMN02441007). Genomic alignment of the Micromonospora sp. HK10 was performed with Micromonospora sp. L5 and M. aurantiaca ATCC20029 and with four secondary metabolite producing actinobacteria viz., Salinispora arenicola CNS-205, Streptomyces coelicolor A3 (Hirsch and Valdés, 2009), Nocardioides brasiliensis ATCC 700358 and Rhodococcus pyridinivorans (SB3094) strains using MAUVE software (Darling et al., 2010) and average nucleotide sequence identity (ANI) score was validated by NCBI-Blastn. Complete genome sequence was analyzed using IMG server (http://img.jgi.doe.gov) and antiSMASH (Medema et al., 2011; Blin et al., 2013; Weber et al., 2015) was used to detect the secondary metabolite gene clusters. A Venn diagram of unique and shared genes present in the respective genomes was calculated using DrawVenn (http://bioinformatics.psb.ugent.be/webtools/Venn/). HK10 genome was examined for horizontal gene transfer (HGT) which plays critical roles in shaping the genome assisted by transposons and other origins. We performed two analyses: (i) GIs prediction by IslandViewer (using Dimodulicity-mobility analysis (DIMOB) and IslandPath tools) (Dhillon et al., 2015) and (ii) detection of insertion sequences from HK10 genome using IS Finder database (https://www.is.biotoul.fr/). The DIMOB analyses is based on dimodulite bias and mobility genes to more accurately predicted regions, termed DIMOB islands (DIMOB-Is) that may have been acquired by HGT. IslandPath combine these parameters using additional information of tRNAs and G + C percentage. All the predicted GI contents were compared by BLAST to identify the kind of genes acquired by HGT in HK10 strain from other bacteria. The draft genome of Micromonospora sp. HK10 was examined by CRT, PILERCR (Bland et al., 2007; Edgar, 2007), and CRISPR Finder tool (Grissa et al., 2007) to detect the presence of CRISPR elements, which confer immunity against incoming DNA, including bacteriophages. Transport system analyses of HK10 genome was performed by comparing each predicted protein against transporter classification database (TCDB) (http://www.tcdb.org/).

2.4. Phylogenetic analysis of Micromonospora sp. HK10 strain

Total of 67 almost complete 16S rRNA sequences were acquired from public databases NCBI Genbank and aligned by ClustalW using 1.6 DNA matrix (Thompson et al., 1994). The evolutionary history was inferred by using the Maximum Likelihood method (Kimura, 1980; Felsenstein, 1985) based on 1000 bootstrap replicates, Streptomyces coelicolor A3 (Hirsch and Valdés, 2009) was used as an outgroup. Evolutionary analyses were conducted in MEGA7 software (Kumar et al., 2016).

2.5. Nucleotide sequence accession number

HK10 Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number NZ_JT-GL000000001.
3. Results and discussion

3.1. Genome characteristics and comparative analysis of HK10 strain with other Micromonospora strains

The final de novo genome assembly of Micromonospora sp. HK10 contains 294 contigs with a total size of 6,911,179 bp (50% of the entire assembly of contigs i.e., “N50” equal or larger than 61,860 bp and the longest contig assembled measures 352,445 bp) with average G + C content of 73.39% (Table 1). The genome contains a total of 6282 predicted protein coding genes/CDs, having 3 tRNAs and 58 rRNAs (Table 1). From total predicted protein coding genes, 4644 genes (73.93%) were allocated with a biological function while 1552 genes (24.71%) were categorized as of unknown function. Further, out of these 6282 CDS, 3598 genes (57.27%) were observed under the clusters of orthologous groups (COGs) and assigned to 25 different categories (Fig. 1, Table S1 & S2). A one-sample t-test was used to evaluate if there were statistically significant differences in the gene abundances of each COG categories between the genomes of HK10 and the other five Micromonospora strains present in IMG database: M. aurantiaca ATCC 27029\textsuperscript{3}, Micromonospora sp. L5, Micromonospora sp. ATCC 39149, M. lupini Lupac 08, and Micromonospora sp. CNB394 (Fig. 1, Table S1).

The results showed that the abundances of the proteins responsible for carbohydrate transport and metabolism (B, 9.11%), cell motility (D, 0.37%), posttranslational modification, protein turnover, chaperones (S, 4%), signal transduction mechanisms (W, 5.6%) are significantly ($P < 0.005$) higher than the average levels, which are 8.71%, 0.265%, 3.79% and 5.38%, respectively.

The abundances of the proteins responsible for chromatin structure and dynamics (F, 0.02%), coenzyme transport and metabolism (G, 5.99%), defense mechanisms (H, 0.02%), cytoskeleton (I, 2.92%), inorganic ion transport and metabolism (N, 4.49%), lipid transport and metabolism (P, 5.03%), monolipids: propanolipids, transposons (Q, 0.37%), nucleotide transport and metabolism (R, 2.36%), RNA processing and modification (T, 0.02%) and secondary metabolite biosynthesis, transport and catabolism (V, 3.88%) are significantly ($P < 0.005$) lower than the average levels, 0.03%, 6.26%, 0.033%, 2.95%, 4.88%, 5.43%, 0.51%, 2.44%, 0.023%, and 4.26%, respectively (Fig. 1, Table S1).

The genomic characteristics of HK10 strain and five other Micromonospora genomes available in the public databases including M. aurantiaca ATCC 27029\textsuperscript{3}, Micromonospora sp. L5, Micromonospora sp. ATCC 39149, M. lupini Lupac 08, and Micromonospora sp. CNB394 were compared (Table S3). The important differences observed between the six strains were GC percentage and number of tRNA genes. HK10 genome has particularly highest GC content (73.39%) and second highest number of tRNAs (Nikaido and Takatsuka, 1794) compared to other five strains (Table S3). Several reports in bacteria indicate that increase in GC content correlates with a higher temperature optimum and a broader tolerance range for a species (Nishio et al., 2003; Musto et al., 2006). Thus high GC content in HK10 might give more thermo stability for adaptation to the immediate environment and high numbers of tRNA genes can be seen as a potential for high protein synthesis and good growth rate (Rocha, 2004) in Micromonospora strains but it is matter of future investigation.

A complete genome sequence makes the fundamental hereditary content of an organism finite and is the most suitable state for any genome sequence with less error bars. Therefore, we have further compared HK10 genome sequence with two fully sequenced genomes available in the database i.e., M. aurantiaca ATCC 27029\textsuperscript{3} and Micromonospora sp. L5. Our analysis disclosed that the highest number (6360) of annotated genes present in M. aurantiaca ATCC 27029\textsuperscript{3} followed by Micromonospora sp. L5 and HK10 strain (6326 and 6282 respectively) (Table S3). There were highest 60 predicted gene clusters having 861 genes (13.6% of total genome) in Micromonospora sp. L5, followed by M. aurantiaca ATCC 27029\textsuperscript{3} with 893 genes (14.04% of total genome) and HK10 strain detected with lowest 49 gene clusters with 721 genes (11.48% of the total genome). Functional annotation using the Kyoto encyclopedia of genes and genomes (KEGG) exposed the total protein coding genes connected to KEGG pathways: lowest 1213 (19.31%) in Micromonospora sp. HK10, followed by 1247 (19.61%) of M. aurantiaca ATCC 27029\textsuperscript{3} and highest 1277 (20.19%) of Micromonospora sp. L5. Similarly analysis with KEGG Orthology suggested, lowest number (2118) of genes (33.72%) found in Micromonospora sp. HK10, followed by 2237 (35.36%) and 2216 (34.84%) number of genes in Micromonospora sp. L5 and M. aurantiaca ATCC 27029\textsuperscript{3}, respectively.

Pair-wise genome synteny studies of these strains also revealed a high degree of positional conservation between the three genomes (Fig. S1). Although the three genomes have a considerable amount of genetic similarity, they have encountered with many inversion and translocations. HK10 genome possesses large number of non-conserved regions and unique genomic regions as compared to M. aurantiaca ATCC 27029\textsuperscript{3} and Micromonospora sp. L5. The genomic comparison by average nucleotide sequence identity confirmed that HK10 strain has the highest similarity to Micromonospora sp. L5 (with 95% identity and 48% coverage), followed by M. aurantiaca ATCC 27029\textsuperscript{3} (with 95% identity and 47.9% coverage). To evaluate genetic similarity of strains, the comparative Venn diagrams of the total protein coding genes with predicted function was analyzed. Interestingly, HK10 strain contains the highest number (4644, 73.93%) of protein coding genes with predicted function in comparison to Micromonospora sp. L5 (4358, 68.1%) and M. aurantiaca ATCC 27029\textsuperscript{3} (4259, 66.97%) (Fig. S2). Our analysis also found that HK10 genome shared highest common genes (992 genes, 11.1%) with M. aurantiaca ATCC 27029\textsuperscript{3} and lowest common genes (925, 10.2%) with Micromonospora sp. L5 (Fig. S2). Unique genes comparison of HK10 strain has shown highest 3643 genes (78.4%) with Micromonospora sp. L5 followed by 3252 genes (76.3%) with M. aurantiaca ATCC 27029\textsuperscript{3}. Similarly, HK10 strain contains highest 3710 number of genes (79.8%), compared to Micromonospora sp. L5 3337 numbers of genes (76.5%) (Fig. S2). These genetic similarities revealed the closeness of HK10 strain with both strains.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA, total number of bases</td>
<td>6911179</td>
</tr>
<tr>
<td>DNA coding number of bases</td>
<td>6129805</td>
</tr>
<tr>
<td>DNA G + C content (%)</td>
<td>73.39</td>
</tr>
<tr>
<td>Genes total number</td>
<td>6282</td>
</tr>
<tr>
<td>Protein coding genes</td>
<td>6196</td>
</tr>
<tr>
<td>Protein coding genes with function prediction</td>
<td>4644</td>
</tr>
<tr>
<td>Protein coding genes without function prediction</td>
<td>1552</td>
</tr>
<tr>
<td>RNA genes</td>
<td>86</td>
</tr>
<tr>
<td>tRNA genes</td>
<td>3</td>
</tr>
<tr>
<td>5S rRNA</td>
<td>1</td>
</tr>
<tr>
<td>16S rRNA</td>
<td>1</td>
</tr>
<tr>
<td>23S rRNA</td>
<td>1</td>
</tr>
<tr>
<td>tRNA genes</td>
<td>58</td>
</tr>
<tr>
<td>Other RNA genes</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 1
General statistics of Micromonospora sp. HK10 genome.
Fig. 1. (A) Comparison of COG functional categories between Micromonospora sp. HK10 and other five Micromonospora genomes- Micromonospora sp. HK10, Micromonospora sp. L5, M. aurantiaca ATCC 27029T, Micromonospora sp. ATEC 39149, M. lupini Lupac 08 and Micromonospora sp. CNIB394 (from the center to the outer side). (B) Comparison of COG functional categories between Micromonospora sp. HK10 and four actinobacteria genomes- Micromonospora sp. HK10, S. coelicolor A3 (Hirsch and Valdés, 2009), S. arenicola CNS-205, R. pyridinivorans SB3094, N. brasiliensis ATCC 700358) from IMG database, (from the center to outer side).

3.2. Genomic comparison of Micromonospora sp. HK10 with other actinobacteria

Since Micromonosporaceae family is well known for production of a number of medically relevant bioactive compounds, it is interesting to understand the genetic composition of other secondary metabolite producing actinobacteria, viz. Salinispora arenicola CNS-205 (Accession no.: NC_009953), Streptomyces coelicolor A3 (Hirsch and Valdés, 2009) (Accession no.: AL645882), Nocardioides brasiliensis ATCC700358 (Accession no.: NC_018681) and Rhodococcus pyridinivorans SB3094 (Accession no.: NC_023150). One-sample t-test was carried out between Micromonospora sp. HK10 and above four actinobacteria (Fig. 1, Table S2) suggests that the HK10 strain has significantly higher abundance of genes of carbohydrate transport and metabolism (B, 9.11%), cell motility (D, 0.37%), chaperons (S, 4%), and signal transduction mechanisms (W, 5.6%) compared to other actinobacteria strains (Fig. 1, Table S2). On contrary, abundance of genes categorized under cell cycle control, cell division, chromosome partitioning (C, 0.81%), coenzyme transport and metabolism (G, 5.99%), defense mechanisms (H, 0.02%), cytoskeleton (I, 2.92%), energy production and conversion (J, 5.89%), inorganic ion transport and metabolism (N, 4.49%), lipid transport and metabolism (P, 5.03%), mobiolome: prophages, transposons (Q, 0.37%), nucleotide transport and metabolism (R, 2.36%), RNA processing and modification (T, 0.02%), secondary metabolite biosynthesis, transport and catabolism (V, 3.88%), transcription (X, 10.31%) were significantly lower (Table S2). These results revealed a common feature in the secondary metabolite producing bacteria, which is the low abundance of
genes in secondary metabolites biosynthesis and high abundance of genes in carbohydrate metabolism.

Further we carried out a pair wise comparative genome analysis of HK10 strain between S. aurecicola, S. coelicolor, N. brasiliensis and R. pyridinivorans strains using progressive Mauve aligner (Darling et al., 2010) (Fig. S3). Although the five genomes share a notable amount of genetic characteristics, they have encountered various inversions and translocations and Micromonomospora sp. HK10 contains the highest number of non-conserved regions. The genomic comparison by average nucleotide sequence identity confirmed that HK10 strain has the highest similarity to S. aurecicola. The order of similarity was S. aurecicola (identity score of 92%, coverage of 33%) > S. coelicolor (identity score of 85%, coverage of 10%) > N. brasiliensis (similarity score of 85%, coverage of 5%) > R. pyridinivorans (similarity score of 86%, coverage of 4%). These results are consistent with their phylogenetic relation of HK10 strain and S. aurecicola, as both genera belong to Micromonomosporaaceae family.

Total protein coding genes with- and without- predicted function is shown as: (464/1552), (6485/2009), (5364/2846), (3703/1204) and (3316/1790) in Micromonomospora sp. HK10, N. brasiliensis, S. coelicolor, R. pyridinivorans, S. aurecicola, respectively. Here, we have compared ‘total protein coding genes with predicted function’ to validate the similarity between HK10 and the other actinobacteria genomes. Our analysis revealed the similarity in following order, R. pyridinivorans (1953 genes, 42.13% similarity) > N. brasiliensis (1405 genes, 30.31% similarity) > S. aurecicola (836 genes, 18.46% similarity) > S. coelicolor (819 genes, 17.66% similarity) (Fig. S4). Surprisingly, this analysis brought the highest sharing (1953 protein coding genes) between HK10 and R. pyridinivorans strains, but above results (Fig. S3) was found just opposite. These variations might be arising because former results appeared after total genome analysis (Fig. S2) however, latter one utilizing only part of the genome for analysis. It is also noted that actinobacteria genomes have high number of genes categorized under ‘without predicted function’ category, shown as S. coelicolor, 2846 > N. brasiliensis, 2009 > S. aurecicola, 1790 > R. pyridinivorans, 1204, which is also accountable for these differences.

3.3. Genomic Island and horizontal gene transfer

A total of 18 genomic islands (GIs) were predicted by Island Viewer method (Dhillon et al., 2015) and their precise location is shown in Fig. S5 and Table S4. These 18 GIs encode a total of 113 (1.7%) genes in 1,344,243 bp from various sources viz. eukaryote, virus and other bacteria. All the identified genes from these Gls are summarized in Table S4, including genes encoding for numerous modifying genes i.e., hydrolases, acetyltransferases, dehydrogenase, methyl transferase, phosphotransferase, glycosyltransferase, helicase, glycosyl hydrolase. Thus, it shows that HK10 strain evolved by receiving Gls, which make them genetically richer and equipped with secondary metabolite modifying enzymes and eventually making them potential producer of several bioactive compounds. As per IslandPath-DIMOB GI-1 (4,549,755 bp to 4,571,433 bp) was found to be the prominent island representing 16 different genes while GI-3 (7,115,968 bp to 7,130,234 bp) was having 8 genes, which also includes transposes IS 116/IS 110/IS902 family genes (Table S4). Presence of three putative transposase (from GI-3, Ga0063173_03998, Ga0063173_6177 and Ga0063173_0269) and one phase integrase family gene (from GI-4, Ga0063173_04552) or their fragments were scattered in the genome that could support potential active HGT in the strain (Table S4). Interestingly only three Gls (GI-1, GI-11 and GI-16) had the similar DNA composition with the Micromonomospora

genus in general while the rest of the GIs showed similarity with various phyla of bacteria such as, Actinobacteria: Actinoplanes (GI-2, GI-13, GI-15 and GI-17), Salinispora (GI-3 and GI-18), Streptomyces (GI-4), Verrucosispora (GI-5, GI-7, GI-12), Frankia (GI-8), Streptosporangium (GI-14); Chlorobi: Chlororubib tepidum (GI-6); Planctomyces: Planctomyces limnophilus (GI-9); and Proteobacteria: Thioctis violacens (GI-10) (Table S4). IMG server analysis revealed that HK10 strain has acquired lesser genes (48 genes) by recent horizontal transfer than Micromonomospora sp. LS (61 genes) and M. aurantiaca ATCC 27029T (67 genes) strains (Fig. 2 and Table S5). Overall all horizontal gene transfer from phylum Proteobacteria is predominantly observed in all Micromonomospora genome sequences. Interestingly HK10 genome has some unique gene transfer from Armamotoadetes phylum, which was absent in other available genome sequences (Table S5). Moreover, HK10 genome contains 23 mobile genetic elements from 11 different IS families i.e., IS 110, IS 1380, IS 21, IS 256, IS 3, IS 4, IS 5, IS 630, IS As1, IS L3 and IS NCY (Fig. S6 and Table S6). IS 110 family transposases are predominantly present in HK10 genome.

Clustered regularly interspaced short palindromic repeats (CRISPRs) are structures, which conceal complex biological mechanisms to account for their transfer, evolution and behavior. They have probably played an important role in the evolution of archaea and bacteria by providing a defense mechanism against foreign DNA (Horvath and Barrangou, 2004). Using CRT and PILCR tools (Bland et al., 2007; Edgar, 2007) a total of 10 and with CRISPR database (Grissa et al., 2007) a total of 14 CRISPR sequences in HK10 genome were identified (Table S7). Notably no CRISPR sequences were found in Micromonomospora sp. LS, M. aurantiaca ATCC 27029T and M. lupini Lupac 08 strains. The presence of number of CRISPR sequences might be responsible for resistance of foreign DNA sequences in HK10 strain compared to other available genomes of Micromonomospora where no CRISPR sequences have been observed. Number of possible roles for CRISPR sequences have been suggested such as, chromosomal rearrangement, modulation of expression of neighboring genes, target for DNA binding proteins, replicon partitioning, and DNA repair (Makarova et al., 2002). Numerous in silico studies have been reported homology between spacer sequences and extra chromosomal elements (like viruses and plasmids), which led to the hypothesis that CRISPR may provide adaptive immunity against foreign genetic elements (Horvath and Barrangou, 1994; Makarova et al., 2006). These studies lead us to speculate that presence of CRISPR in HK10 strain might help in adaptation under variable environmental conditions.

3.4. Antimicrobial activity of Micromonomospora sp. HK10 and complete secondary metabolism encoding genes

HK10 strain has already been reported to possess a broad range of antibacterial activity against various Gram-positive and Gram-negative bacteria (Talukdar et al., 2012). To reaffirm this broad spectrum phenotype, we have evaluated antibacterial activity of HK10 strain against different clinical pathogens such as, Mycobacterium smegmati s ATCC 700084, Mycobacterium abscessus ATCC 19977T, Salmo nella sp. IDH377, and enterotoxigenic Escherichia coli E9. The zone of inhibition was determined as 27 mm, 19 mm, 18 mm and 18 mm for M. smegmatis, M. abscessus, Salmonella sp. and E. coli respectively, compared to 19 mm, 17 mm, 13 mm and 15 mm against M. smegmatis, M. abscessus, Salmonella sp. IDH377 and E. coli E9, respectively using as positive control neomycin sulfate (Fig. 3). Although the concentration of HK10 extract used was double (100 µg/
Fig. 2. Comparative analysis of horizontally transferred genes of Micromonospora HK10 with other five Micromonospora genomes from IMG database.

Fig. 3. Antimicrobial activity of Micromonospora sp. HK10-chloroform cell extracts (HK10) tested against (i) M. smegmatis, (ii) M. abscessus, (iii) Salmonella sp. IDH 377 (iv) Enterotoxigenic E. coli E9 strains, 50 μg/ml neomycin sulfate (NS) used as positive control and only DMSO used as a control (C).

ml) than the neomycin sulfate (50 μg/ml), reasonably higher inhibition zones of HK10 extracts against all test pathogens clearly indicates the effectiveness of secondary metabolite produced by HK10. Secondary metabolite production of HK10 strain was optimized using different media at various time interval (data not shown) and 10 days old culture of HK10 strain grown on CSPY media emerged as the best for antibacterial activity. The antimicrobial activity of HK10 was absent when other media such as Mueller Hinton
broth (Himedia), Luria Bertani broth (Himedia), Nutrient broth (Himedia) and Starch Casein broth (Himedia) was used.

To investigate the secondary metabolite synthesis potential of HK10 strain, secondary metabolite encoding gene clusters was predicted using IMG server. HK10 genome possesses 49 secondary metabolite gene clusters, which are encoded by 721 genes with 11.48% of genome coverage (Fig. 4, Table S8). Total 49 gene clusters include 13 polyketide synthase (PKS) genes, 2 non-ribosomal peptide synthetase (NRPS)-PKS genes and 34 encoding other secondary metabolites including siderophores (Fig. 4, Table S8). Interestingly, the length of ten gene clusters exceeds > 25 Kb. The presence of large gene clusters in HK10 genome indicates that these clustering confer an ecological advantage and each gene of the cluster contributes to an enzymatic step in their respective cellular metabolic process. These clusters also imply transcriptional co-regulation of genes, which can be controlled at different levels of expression. HK10 genome is exceptionally rich in pks gene cluster (total 13) among sequenced genomes, providing evidence that HK10 strain harbors an extensive secondary metabolite metabolism (Fig. 4, Table S8). To understand the overall secondary metabolite synthesis potential of HK10 strain, these clusters were compared with Micromonospora sp. L5 and M. aurantiaca ATCC 27029\(^1\) (Table S9). Compared to M. aurantiaca ATCC 27029\(^1\), HK10 genome had 8 unique gene clusters, 30 conserved gene clusters (similarity > 60%) and 11 clusters were partially conserved (similarity > 30%) (Table S9). Similarly 9 gene clusters (Cluster 7: Putative, Cluster 19: Putative, cluster 22: Lantipeptide, Cluster 29: Putative, Cluster 44: Lantipeptide, Cluster 45: Putative, Cluster 47: Lantipeptide, NRPS, T1 PKS, Cluster 48: Putative, Cluster 49: Putative) were found to be unique when HK10 genome was compared with Micromonospora sp. L5, along with 29 conserved gene clusters and 11 partially conserved clusters (Table S9, Fig. 5, Table S10).

Overall seven gene clusters have been found unique in HK10 genome (Fig. 5). Our HGT analysis of HK10 genome has confirmed that cluster 7 (present in GI-17), cluster 22 (present in GI-8) and cluster 44 (present in GI-3) have been horizontally transferred from Actinoplanes sp., Frankia sp. and Salinispora arenicola respectively (Table S4 & S10).

To explore the biosynthetic cluster closeness of HK10 strain with other secondary metabolite producing strains, predicted gene clusters have been compared with other actinobacteria. Our results revealed the similarity between HK10 strain and other actinobacteria in the following order Salinispora arenicola (16 conserved, 13 partially conserved, 20 absent) > Streptomyces coelicolor (3 conserved, 15 partially conserved, 31 absent) > Nocardioides brasiliensis (4 partially conserved and 45 absent) > Rhodococcus pyridinivorans (4 partially conserved and 45 absent) (Table S9). Interestingly seven gene clusters (cluster 7, cluster 19, cluster 22, cluster 44, cluster 45, cluster 47 and cluster 48) of HK10 strain, did not show any similarity with other actinobacteria used in this study. These clusters might be interesting to explore in future for their secondary metabolite products. These secondary metabolite encoding gene clusters were again revalidated by antiSMASH cluster predicting tool (Medema et al., 2011; Blin et al., 2013; Weber et al., 2015). After analysis of HK10 genome, total twenty-two secondary metabolite encoding gene clusters have been discovered. Interestingly, antiSMASH analysis of HK10 genome emerged with one more lantipeptide gene cluster (Cluster 1) that was absent in our earlier IMG prediction (Table S8). There could be two possible reasons that cause the differences: (Genilowd and Genus, 2012) in IMG the results of biosynthetic cluster prediction is affiliated with antiSMASH v2 but in our prediction with online versions of antiSMASH, we have used the online version of antiSMASH, i.e., version 3, which now annotates more classes of clusters including lantipeptides, (Hirsch and Valdés, 2009) the number of results also depends on the parameters of running the antiSMASH, for example, IMG allow antiSMASH to predict putative clusters based on an implementation of Clusterfinder and this is not available by default on antiSMASH. Our analysis revealed five annotated biosynthetic gene clusters (Fig. S7): lantipeptide (cluster 1), PKS-hybrid (cluster 4), terpene (cluster 8), PKS III (cluster 13) and lantipeptide-siderophore (cluster 18), showing high level of similarity with SapB biosynthetic gene cluster of S. coelicolor (100%), Chlorothricin biosynthetic gene clus-

Fig. 4. (A) Comparative analysis of total genes distributed under categories of: present in secondary metabolite clusters/not present in secondary metabolite clusters in selected Micromonospora spp. (B) Comparative analysis of total predicted secondary metabolite gene clusters types (PKS, NRPS, PKS-NRPS, OTHER) present in selected Micromonospora spp.
Fig. 5. Seven unique biosynthetic gene clusters present in Micromonospora sp. HK10: (A) Cluster-7 (B) Cluster-19 (C) Cluster-22 (D) Cluster-44 (E) Cluster-45 (F) Cluster-47 (G) Cluster-48.

The six unique biosynthetic gene clusters of Streptomyces antibioticus (44%), Suxanthon biosynthetic gene cluster of Salinispora tropica (80%), Alkyl-O-Dihydrogeranyl-Methoxyhydroquinones biosynthetic gene cluster of M. aurantiaca ATCC 27029\(^7\) (71%), and Desferrioxamine B biosynthetic gene cluster of Streptomyces griseus (80%) (Table S11). These five clusters further aligned with other secondary metabolite producing strains of the database and all the results have been shown in Fig. S7. Lantipeptide encoding cluster 1, PKS-hybrid encoding cluster 4 and lantipeptide-siderophore encoding cluster 18 have been showing highest gene similarity with Streptomyces davawensis (19%), Streptomyces antibioticus (58%) and Streptomyces fradiae (30%) strains respectively, whereas terpene encoded cluster 8 and PKS III encoding cluster 13 have more similar to Micromonospora sp. ATCC 39149 (29%) and M. aurantiaca ATCC 27029\(^7\) (82%) strains (Fig. S7). Overall our study provides critical genetic insights into the HK10 strain, which has antibacterial activity against a wide range of bacteria, including Mycobacteria and can produce several bioactive secondary metabolites.
The antibacterial activity of HK10 strain and its link with predicted secondary metabolite gene clusters is our undergoing research work.

3.5. Secretion systems of *Micromonaspora* sp. HK10

Secretion systems are essential for translocation of various proteins across the cell membrane which plays crucial roles like host-pathogen interactions and colonization (Driessen and Nouwen, 2008). Proteins in Gram-positive bacteria predominantly translocated by conserved Sec translocaze system or twin-arginine translocation (Tat) system. Sometimes the type VII or ESX secretion system has been also reported in some Gram positive bacteria. Similar to them HK10 genome encodes all the Sec components (Table S12), viz: the essential proteins SecY (Ga0063173_02940) and SecE (Ga0063173_02979), along with SecD (Ga0063173_02596) and SecF (Ga0063173_02597). Both co- and post-translational routes of targeting secretory proteins to the Sec translocon are present in HK10 strain. HK10 genome also encodes SecA protein (Ga0063173_02270, Ga0063173_06178) like *Streptomyces lividans*, which has been shown in vitro to be required for the secretion of the Sec-dependent α-amylase (Blanco et al., 1998). HK10 genome also contains *fhh* gene (Ga0063173_03995), encodes for signal recognition particle subunit SRP54 and *fshY* (Ga0063173_03984), encoding the SRP receptor. Mutations in *fshY* gene affect sporation and antibiotic production in *S. coelicolor* (Shen et al., 2008).

HK10 genome showed presence of various Tat pathways homologs. Similar to other actinobacteria, TatA (Ga0063173_05927), and TatC (Ga0063173_05926) are found adjacent to each other while TatB (Ga0063173_05707) was located distantly. Unlike *M. lupini Lupac* 08 genome, which has no TatB homologue (Trujillo et al., 2014), HK10 Tat secretion pathway is very similar with *Micromonaspora* sp. L5 and *M. aurantiaca* ATCC 27029. The Tat pathway is responsible for export of number of pre-folded proteins across bacterial membranes (Lee et al., 2006). Tat proteins are also involved in different cellular activities including cell envelope biogenesis, metal acquisition, anaerobic metabolism and detoxification and virulence (Lee et al., 2006). Hence, presence of Tat pathway in *Micromonaspora* strains might useful in dealing with the changing environment around them.

HK10 genome has total 6 genes (Ga0063173_00846, Ga0063173_00894, Ga0063173_00895, Ga0063173_03797, Ga0063173_03798, Ga0063173_03804), which were categorized under RD1 or ESX-1/Snm protein secretion system family (Table S13 & S14). Like different *Streptomyces*, the function of these genes in *Micromonaspora* genus is currently unknown, but a significant physiological role is suggested by analogy to similar systems in other organisms.

T4SS have been mainly studied in Gram-negative bacteria but in Gram-positive bacteria the information is very limited so far. However, Type IV (Conjugal DNA-Protein transfer or VirB) secretary pathway (IVSP) family gene (Ga0063173_02432) was observed in HK10 genome (Table S14) and similar type of gene (virB4) was also reported in *M. lupini Lupac* 08, an endophytic actinobacterium (Trujillo et al., 2014), but was interestingly absent in most of the sequenced *Micromonaspora* species (data not shown), including *Micromonaspora* sp. L5 and *M. aurantiaca* ATCC 27029 strains. The precise role of IVSP family genes in *Micromonaspora* genus is matter of future research.

3.6. Transporters of HK10 strain

Various active transporter systems are involved in transport of nutrient elements, detoxification and defense mechanisms. HK10 genome contains total 74 families of transporters covered by total 587 genes (9.3% of total CDS), which are involved in the transport systems for carbohydrates, lipid, amino acids, nucleotides, aromatic compounds, metals and inorganic ions (Tables S13 &S14). HK10 genome has emerged with the lowest 587 (9.3%) number of transporter genes with 49 predicted metabolite clusters whereas *M. aurantiaca* ATCC 27029 and *Micromonaspora* sp. L5 contain 630 (9.9%) genes with 59 predicted secondary metabolite gene clusters and 658 (10.4%) with 60 clusters, respectively. The soil actinomycetes, *S. coelicolor* has also total 55 predicted secondary metabolite gene clusters with total 834 transporter genes (10.1% genome coverage) (Bentley et al., 2002). These transporters of HK10 strain are potentially involved in metabolite export and are consistent with such extensive secondary metabolism. HK10 genome contains protein-dependent oligopeptide transporter (POT) (Ga0063173_06054), responsible for peptide transport (Table S14), which regulates the expression of secondary metabolite genes (Lorenzana et al., 2004). Major facilitator superfamily (MFS) transporters show specificity for many vital components like nutrients, peptides, nucleotides, drugs, Krebs cycle metabolites, osmorhites, siderophores (efflux & uptake) organic and inorganic anions, etc. (Law et al., 2008). MFS transporter PenT regulates penicillin production by enhancing the translocation of penicillin precursors across fungal cellular membrane (Yang et al., 2012). RND (Resistance-Nodulation-Division) family transporters are widespread especially among Gram-negative bacteria, and catalyze the active efflux of many antibiotics and chemotherapeutic agents and might help to survive against different toxin compounds (Nakai and Takatsuka, 1794). Interestingly HK10 genome contains highest number of MFS (Zhao et al., 2004) and RND (Trujillo et al., 2007) transporters when compared with *Micromonaspora* sp. L5 (61 MFS & 9 RND) and *M. aurantiaca* ATCC 27029 (58 MFS and 8 RND). Moreover, membrane fusion protein (MFP) transporter (Li and Nakai, 2009) (Ga0063173_03875) presents only in HK10, which was not found in both *Micromonaspora* sp. L5 and *M. aurantiaca* ATCC 27029 strains (Table S14).

HK10 genome contains some other transporters, viz: Major intrinsic protein (MIP) family (TC:1.2.8), Lactate permease (LetP) family (TC:2.14.19 and Zinc (Zn\(^2+\))-Iron (Fe\(^3+\)) permease (ZIP) family (TC:2.1.5), which was absent in both *Micromonaspora* sp. L5 and *M. aurantiaca* ATCC 27029 strains (Table S12). In contrast, several transporters were absent in HK10 strain e.g., Urea transporter (UT) family (TC:1.2.18), Betaine/Carnitine/Choline transporter (BCCT) family (TC:2.1.19), Ca\(^2+\)-Cation Antiporter (CaCA) family (TC:2.1.19), Arsenical resistance-3 (ACR3) family (TC:2.1.5), Monovalent cation (K\(^+\) or Na\(^+\))-Proton antiporter-3 (CPA3) family (TC:2.1.63), Auxin efflux carrier (AEC) family (TC:2.1.69), l-Lysine exporter (LysE) family (TC:2.1.75), GlucosamineH\(^+\) Symporter (GntP) family (TC:2.1.8), the Tricarboxylate Transporter (TTT) family (TC:2.1.8), putative cobalt transporter (ChtAB) family (TC:9.6.69) (Table S13).

3.7. Salt tolerant characteristic of HK10 strain

Genomic background of *Salinispora* suggests its evolution from a terrestrial environment (Penn and Jensen, 2012) which raises the question whether *Micromonaspora* has some common genes related to marine environment. There are *Micromonaspora* species now
which have been reported from marine environment (Das et al., 2008; Phongsopitanun et al., 2015; Supong et al., 2013; Maldonado et al., 2009; Zhao et al., 2004; Mincer et al., 2002) and produce various antimicrobials of medical importance (Charan et al., 2004; Thi et al., 2016; Carlson et al., 2013; Kyeremeh et al., 2014). We have confirmed the salt tolerance ability of HK10 strain by growing them on higher concentration (0.5–3.0%) of sodium chloride and best antibacterial activity was observed at 2.5% sodium chloride (data not shown). This characteristic suggests that certain genes may be adapted from marine microbes. Putative potassium uptake protein (Trk) (Ga0063173_05704) which might be useful in marine adaptation (Schlosser et al., 1993; Nakamura et al., 1994) and NDH protein (Ga0063173_00883) constitutes a part of proton pumping NADH dehydrogenase and creates a proton-motive force for generation of ATP (Schlosser et al., 1993). HK10 genome also contains gene encoding mechanosensitive channels MscL (Ga0063173_00971) and MscS (Ga0063173_02549, Ga0063173_04307, Ga0063173_06285) (Table S15 & S16), which enable growth at low osmotic pressure (Bucarey et al., 2012). These genes are also present in Micromonospora sp. L5 and M. aurantiaca ATCC 27029 strains. Presence of these genes in HK10 strain could suggest an adaptation from marine bacteria and might be helpful in salt tolerance in adverse environmental condition.

3.8. Signal transduction and two-component system

The full list of genes involved in signal transduction and two-component systems (TCS) in HK10 strain are given in Table S15 & S16. Various studies have shown that biosynthesis of antibiotic and many other secondary metabolites are strongly down regulated by inorganic phosphate (Sola-Landa et al., 2003; Ishige et al., 2003; Lefevere et al., 1997; Braissant and Content, 2001; Glover et al., 2007). In HK10 strain, the inorganic phosphate deprivation condition might be controlled by senX3-regX3 two-component system. Although the two-component system senX3-regX3 (Ga0063173_04526 and Ga0063173_04527) was found in HK10 strain, the two-component system phoR-phoP was absent. The alkaline phosphatase, phoA gene (Ga0063173_03810) and ABC phosphate transporter operon pstSCAB (Ga0063173_04482, Ga0063173_04481, Ga0063173_04480, Ga0063173_04479) was also present in HK10 genome. We have also observed a few more genes of phosphate transporter for example, phosphate starvation-inducible protein, PhoH (Ga0063173_04785) and phosphate transport system protein, PhoU (Ga0063173_04525). Instead of alkaline phosphatase, phoA gene, alkaline phosphatase, phoD gene (Micau_4701 and Micau_4898) was present in M. aurantiaca.

K+ intracellular cation is required for various physiological processes like turgor homeostasis (Epstein, 1986), pH regulation (Booth, 1985), gene expression (Prince and Villarejo, 1990) etc. For regulation of K+ ion and adaptation under various stress conditions, HK10 genome has KdpD-KdpE (Ga0063173_03480, Ga0063173_03479) two-component system and K+ transporting ATPase genes kdpABC (Ga0063173_03483, Ga0063173_03482, Ga0063173_03481), whereas KdpF gene was absent in HK10 genome like E. coli (Prince and Villarejo, 1990).

HK10 genome has temperature regulator mechanism DesK-DesR two-component system, which works when the ambient temperature drops below 30 °C (Aguilar et al., 2001). Similar DesK-DesR system is also found in both Micromonospora sp. L5 and M. aurantiaca ATCC 27029. Beside this, tricarboxylic transporter Teta-TetB-TetC operon (ML5_4359, ML5_4358, ML5_4357) is only present in Micromonospora sp. L5 and not found in both HK10 strain and M. aurantiaca ATCC 27029. We have also observed gene related to fatty acid metabolism fumarate reductase, fda (Ga0063173_02826, Ga0063173_03575) present in HK10 genome, which was absent in both Micromonospora sp. L5 and M. aurantiaca ATCC 27029 strains.

3.9. Phylogenetic analysis of Micromonospora

The phylogenetic position based on 16S rRNA gene sequence analysis of strain HK10 indicated that it was placed in a monophyletic clade with closest relatives, M. aurantiinga TT1-111 (99% of identity with 97% coverage), M. peucetia DSM 43633 (98% of identity with 97% coverage) and M. chalubaphumensis MC5-1 (99% of identity with 94% coverage).

On habitat basis HK10 strain was clearly positioned within the genus Micromonospora and form a sub group together with other soil isolates (M. aurantiinga TT1-111, M. peucetia DSM 43633 and M. chalubaphumensis MC5-1) (Fig. S8).

4. Conclusions

Despite the ability to produce several useful antibiotics, very few genome sequences of Micromonospora are available in the database. We contribute with a 6,911,179 bp genome of a soil isolate Micromonospora sp. HK10, with broad spectrum antibacterial activity against Gram-negative and Gram-positive bacteria including enterotoxicigenic E. coli, Salmonella sp., M. abscessus and M. smegmatis. HK10 genome analysis emerged with total 49 secondary metabolite gene clusters (encoding mainly PKS, NRPS, siderophore, terpene etc) including comparatively high (Igarashi et al., 2007) PKS and 7 unique gene clusters, which can be exploited for conceiving the new bioactive molecules. Further, this study also provides a new genetic insight into horizontal gene transfer (18-genetic islands with 3 new secondary metabolite genetic cluster), secretion system, CRISPRs, two-component system, salt tolerant behavior and transport mechanism in Micromonospora. This study will certainly enhance our molecular understanding about secondary metabolite engineering and environmental adaptations of Micromonospora and other related microbes.

Acknowledgement

This work was funded by the Council of Scientific and Industrial Research (CSIR), India (CSC131, BSC109, and HCP0001). M.T., D.D., and C.B. are supported by a research fellowship from the CSIR, India. We thank the Director, CSIR-North East Institute of Science and Technology (CSIR-NEIST), Jorhat, India, for his consistent encouragement and support. We are grateful to Dr. T. Ramamurthy, National Institute of Cholera and Enteric Diseases, Kolkata, India, for his kind gift of Salmonella sp. (ID377) and Escherichia coli E9 (serotype 0146) strains used in this study. Authors are also very thankful to Prof. Paul Jensen for his kind permission to use the Micromonospora sp. CNB394 genome sequence in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2016.09.005.

References


