Identification of Potential Bacteria on Several Lakes in East Java, Indonesia Based on 16S rRNA Sequence Analysis

Achmad Rodiansyah, Ainul Fitria Mahmudah, Mastika Marisahani Ulfah, Uun Rohmawati, Dwi Listyorini, Eko Agus Suyono, Sitoresmi Prabaningtyas

1Department of Biology, Faculty of Mathematics and Natural Science, Universitas Negeri Malang, Malang, Indonesia
2Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, Indonesia

1. Introduction

The previous research was found 96 bacterial isolates from Ranu Pani, Ranu Regulo, Ranu Grati, and Telaga Ngebel which have multiple potencies as follows, 40 isolates have the potency to produce IAA, 31 phosphate solubilizing bacterias, two dominant bacterias from Ranu Pani and Ranu Grati can promote microalgae growth (Prabaningtyas et al. 2018a, 2018b), 47 cellulolytic bacterias (Ardilla 2018), and 53 amylolytic bacterias (Basitoh 2018). Identification of bacteria that have the highest potential for producing IAA and solubilizing phosphate yet not been done. Moreover, the results of identification based on morphology and physiology of highest amylolytic bacteria identified as Bacillus subtilis (Basitoh 2018), highest cellulolytic bacteria identified as Bacillus alvei (Ardilla 2018), and identification of dominant bacteria from Ranu Grati identified as Enterobacter gergoviae and dominant bacteria from Ranu Pani cannot be identified (Prabaningtyas and Witjoro 2017).

The identification of microorganisms important for microbiologists and scientists involved in many fields of research and applied industry (Bisen et al. 2012). This identification of potential bacteria has beneficial for biological conservation, regulation study, biological activity, and especially in a consortium of bacterial in co-culture microalgae. The consortium of bacteria in co-culture microalgae can increase the efficiency of biomass production from microalgae (Fuentes et al. 2016; Ramanan et al. 2016). This consortium can increase the growth of algae because bacteria provide inorganic essential mineral elements (Marañón et al. 2005; Thyssen 2005) like carbon that needed for metabolic algae (Jena et al. 2011; Kropat et al. 2011).

The identification of bacteria based on the phenotype has been carried out, but this identification method needs more time and energy (López-Campos et al. 2012), different environmental conditions may decreases the validity (Tshikhudo et al. 2013), some bacteria difficult to be cultured so that identification...
cannot be done (Gigliandolo et al. 2011), and there are some morphological and physiological similarities that are difficult to be distinguished (Boivin-Jahns et al. 1995). A fast and reliable method is widely used to overcome this problem by analyzing the barcode 16S rRNA gene (Buszewski et al. 2017; Tshikhudo et al. 2013).

DNA barcoding is an efficient, accurate, cheap and standardized method for identifying organisms using short DNA sequences from whole-genome DNA in the sample (Cristescu 2014; Lebonah et al. 2014). This method can be used to deepen the understanding of ecology, biodiversity, evolution, taxonomy, and conservation (Hanner et al. 2011). The 16S rRNA gene is one of the barcode genes which is usually used to identify bacteria and its population. That gene was chosen because has an informative area and enough variation to distinguish between taxa in the taxon (Wang and Qian 2009). 16S rRNA sequences can be found in all prokaryotic cells with 1-15 copies (Werner et al. 2012; Rosselli et al. 2016), and have widely used in microbial community studies (Rosselli et al. 2016). The use of this gene depends on significant differences in base sequence between species and fewer differences in species (Woo et al. 2009) so it can be used to estimate the genetic distance between species.

2. Materials and Methods

2.1. Isolates

Six bacterial isolates were obtained from the previous study has various highest potential (Listed in Table 1), those isolates were already in the Laboratory of Microbiology, Department of Biology, Universitas Negeri Malang. Each isolate was cultured on 5 ml nutrient broth and incubated in incubator for overnight at 37°C. These bacterial cultures were used for genome DNA (gDNA) isolation.

Table 1. Bacterial isolates in this study

| Code | Source | Potential       |
|------|--------|-----------------|
| PIS  | Ranu Pani | IAA-Producer    |
| GPS  | Ranu Grati | Phosphate solubilizer |
| PSS  | Ranu Pani | Cellulolytic     |
| PAS  | Ranu Pani | Amylolytic       |
| PØD  | Ranu Pani | Growth promoter  |
| GØD  | Ranu Grati | Growth promoter  |

2.2. gDNA Isolation and 16S rRNA Amplification

The gDNA from each sample was isolated using the QIAmp DNA Mini Kit (Qiagen, Germany). The result from gDNA isolation was measured using NanoDrop ND-2000 Spectrophotometer (ThermoScientific™) to determine the concentration and purity of gDNA. The results from gDNA isolation were used as a template in the Polymerase Chain Reaction (PCR) to amplification the 16S rRNA sequence. Top Taq Master Mix reagents from Qiagen were used for PCR. Primers used for amplification are 27F: 5’-GAGTTTGATCMTGGCTCAG-3’ and 1492R: 5’-ACGGYTACCTTGTTACGACTT-3’ (Devereux and Willis 2004; Wang and Qian 2009). The PCR reaction for 30 cycles was performed as follows: initial denaturation 94°C/3 min, denaturation 94°C/1 min, annealing 50°C/30 s, extension 72°C/1 min 30 s, final extension 72°C/10 min, and hold 4°C. The PCR product was checked on 1% agarose gel electrophoresis.

2.3. DNA Sequencing, Phylogenetic Trees Reconstruction and Genetic Distance Analysis

The PCR products were determined by DNA sequencing in 1st BASE Laboratories, Malaysia. Sequencing results were analyzed using DNA baser software to determine sequence consensus and sequence alignment using BLAST program on NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Phylogenetic trees reconstruction and genetic distance were analyzed using MEGA6 software. The phylogenetic trees were constructed using Neighbor-Joining (NJ) method (Saitou and Nei 1987) and Minimum-Evolution (ME) method (Rzhetsky and Nei 1992). The evolutionary distances were calculated using Kimura 2-parameter method (Kimura 1980) with bootstrap test 1,000 replicates (Felsenstein, 1985).

3. Results

3.1. Genome DNA Isolation and PCR

Total gDNA obtained from isolation had the various purity and concentration from each isolate (Table 2). The PCR result showed that band DNA in gel electrophoresis has a target gene with length about 1500 bp (Figure 1) and then verified by DNA sequencing.
3.2. DNA Sequencing and Sequence Analysis

Consensus sequence length from DNA sequencing listed in Table 3. BLAST analysis from all consensus sequences showed that they have high similarity and high query coverage with 16S rRNA sequence on the NCBI database. Several similar sequences from BLAST were downloaded and used for phylogenetic trees reconstruction and genetic distance analysis.

3.3. Phylogenetic Trees Reconstruction and Genetic Distance Analysis

Phylogenetic trees reconstruction using NJ and ME showed that three main clades and one clade as out of the group, these three main clades are Enterobacter clade (red line), Bacillus cereus Group clade (blue line), and Bacillus subtilis Group clade (green line) (Figure 2a and b). The PIS and PSS isolates were located in Bacillus cereus Group clade, with bootstrap value NJ: 100 and ME: 100, the PAS isolate was located in Bacillus subtilis Group clade, with bootstrap value NJ: 82 and ME: 83, the GPS isolate located on Enterobacter clade with bootstrap value NJ: 100 and ME: 100, PØD and GØD isolates identified as Enterobacter cloacae a member of Enterobacter cloacae Complex (ECC) with bootstrap value NJ: 61 and ME: 62.

Phylogenetic trees showed a correlation with genetic distance analysis. Three main clades in Figure 2 have a significant genetic distance value there are indicating a different genus. Group Enterobacter cloacae Complex has genetic distance value 0.270 with Bacillus subtilis Group and 0.258 with Bacillus cereus Group (black box in Figure 3). It means there’s are 73.0-74.2% sequences similarity in the Enterobacter group with both genus Bacillus so that clearly defined as a different genus. Bacillus cereus Group has a genetic distance value 0.060 or 94% similarity with the Bacillus subtilis Group (red box in Figure 3), so as for them separate in sub-genus level.

Analysis of genetic distance in PIS and PSS samples identified in one species to members species of Bacillus cereus Group with similarity 99.9-100% (red square in Figure 4), and these samples located in the Bacillus cereus Group’s clade. PAS sample is located in Bacillus subtilis Group's clade and this sample identified as B. amyloliquefaciens with 100% similarity but it has closely
Figure 2. Phylogenetic trees based on 16S rRNA gene, (a) NJ tree, (b) ME tree. The trees were constructed with 1,000 replicates using Kimura 2-parameter model and *Acinetobacter baumannii* and *Streptomyces* sp. were used as an out of the group.
related to with *B. subtilis*, *B. velezensis*, and *B. vallismortis* with similarity >99.8% (black square in Figure 4). The genetic distance of the GPS sample has a similarity <95% with another species in genus *Enterobacter*, this result indicated that GPS sample cannot be defined as species but defined in the genus level (blue square in Figure 4). The genetic distance of PØD and GØD samples were defined as *Enterobacter cloacae* complex with similarity >99%, but these samples were closely related to *Enterobacter cloacae* with similarity values 99.8% on GØD and 99.2% on PØD (yellow box in Figure 4). Intraspécific genetic distance analysis showed that there’s are no significant variation 16S rRNA sequences within groups, so the sequences variety each group are relatively homogeneous with similarity >99% (Figure 5).

### 4. Discussion

Phylogenetic trees reconstruction revealed that PSS and PIS samples were located in clade with *Bacillus cereus* Group and PAS sample located in *Bacillus subtilis* Group closely to *Bacillus amyloliquefaciens* NBRC 15535, GØD and PØD located in *Enterobacter cloacae* Complex, while GPS is located in *Enterobacter* clade with form a new clade (can be defined as novel *Enterobacter*) (Figure 2). All of the branches from phylogenetic trees show a confident position with trusted bootstrap value of >50 (Gregory 2008). The position of each sample in the phylogenetic tree showed a correlation with genetic distance analysis. Almost all of them were separated in a new clade as a group supported by its significant genetic distance (Figure 3), and within-group genetic distance (>99%)
Molecular signatures for *Bacillus* spp. were categorized on several groups based on 16S rRNA, whereas *Bacillus substilis* Group in clade I and *Bacillus cereus* Group in clade II (Rooney *et al.* 2009; Bhandari *et al.* 2013; Fan *et al.* 2017). *B. amyloliquefaciens, B. siamensis, B. velezensis, B. methylotrophicus* are a member of the new phylogenetic branch of “operational group *Bacillus amyloliquefaciens*” from *Bacillus substilis* Species Complex (Fan *et al.* 2017). Identification based on morphology and physiology in the previous study PAS sample identified as *B. substilis* (Basitoh 2018) but with 16S rRNA, we can identify more deeply to *B. amyloliquefaciens* NBRC 15535. This species can hydrolyze amylopectin using amylase enzymes (Deb *et al.* 2013). In the other group of *Bacillus* spp., *Bacillus cereus* Group or (*B. cereus* sensu lato) is sub-division in genus *Bacillus*, this group include *B. anthracis, B. cereus, B. mycoides, B. pseudomyoides, B. weihenstephanensis, B. thuringiensis, B. parantarhis, B. pacificus, B. tropicus, B. albus, B. mobilis, B. luti, B. proteolyticus, B. nitratireducens, B. paramyoides* (Nakamura and Jackson 1995; Vilas-Bôas *et al.* 2007; Liu *et al.* 2017). Most species in genus *Bacillus* can produce IAA (Patten *et al.* 2013; Liu *et al.* 2017; Ozdal *et al.* 2017; Susilowati *et al.* 2018) via dependent pathways with tryptophan as a precursor (Spaepen *et al.* 2007; Shao *et al.* 2015) and also cellulose hydrolysis (Balasubramanian *et al.* 2012; Gupta *et al.* 2012; Lin *et al.* 2012; Chantarasiri 2015) with celluloses enzyme group such as endoglucanase, exoglucanase, and β-glucosidase (Bayer *et al.* 2007). The little genetic diversity on those group *Bacillus*, identification on species-level mostly cannot be defined with 16S rRNA (Rasko *et al.* 2005; Rooney *et al.* 2009).

In the *Enterobacter* clade, sample GPS identified as novel species in genus *Enterobacter* with similarity below 95%. In the research progress with that isolate, confirmed that Pho regulon includes phoA, phoB, and phoR genes are probably act on solubilizing of complex phosphates (Ongoing research, data not published). Two dominant isolates (GØD and PØD) are located one clade with *Enterobacter cloacae* ATCC 13047 and they have >99% similarity within group ECC, so it makes identification on species-level probably bias. Some species in genus *Enterobacter*, especially ECC has multiple potentials such as solubilizing phosphate, produce IAA, and produce ammonia (Nhu and Diep 2017).

This research has shown the PIS and PSS isolates were identified as *Bacillus cereus* Group closely related with *Bacillus paramyoides*, PAS isolate identified as *Bacillus amyloliquefaciens* which is a member of *Bacillus substilis* Group, GPS isolate identified as novel species in *Enterobacter*, and GØD and PØD isolates identified as *Enterobacter cloacae* which is a member of ECC. The 16S rRNA sequence can be defining samples in sub-genus or group within a genus level. Species can be defined but not trusted. Other barcodes/marker genes are required to identification species-level in this group. Identification based on whole-genome sequencing, plasmid characterization, phenotypes characterization, and their combination are required to determining the further classification of species (Konstantinidis and Tiedje 2005; Woo *et al.* 2009; Tindall *et al.* 2010; Brady *et al.* 2013).

**Acknowledgements**

The authors would like to extend our grateful to the Direktorat Riset dan Pengabdian Masyarakat (DRPM), Direktorat Jenderal Penguatan Riset dan Pengemembangan from Ministry of Research and Technology Republic Indonesia for funding this research.

**References**

Ardilla R. 2018. Isolasi bakteri selulolitik potensial dari Ranu Pan, Ranu Regulo, Ranu Grati, dan Telaga Ngebel [Thesis]. Malang, Indonesia: Universitas Negeri Malang.

Balasubramanian N *et al.* 2012. Purification and biochemical characterization of a novel thermo-stable carboxymethyl cellulase from Azorean isolate *Bacillus mycoides* S122C. *Appl Biochem Biotechnol* 168:2191–2204.

Basitoh YK. 2018. Eksplorasi bakteri amilolitik potensial dari Ranu Pan, Ranu Regulo, Ranu Grati, dan Telaga Ngebel [Thesis]. Malang, Indonesia: Universitas Negeri Malang.

Bayer EA *et al.* 2007. The potential of cellulases and cellulosomes for cellulosic waste management. *Curr Opin Biotechnol* 18:237–245.
Rooney AP et al. 2009. Phylogeny and molecular taxonomy of the *Bacillus subtilis* species complex and description of *Bacillus subtilis* subsp. inaquosorum subsp. nov. *Int J Syst Evol Microbiol* 59:2429–2436.

Rosselli R et al. 2016. Direct 16S rRNA-seq from bacterial communities; a PCR-independent approach to simultaneously assess microbial diversity and functional activity potential of each taxon. *Sci Rep* 6:32165. DOI: 10.1038/srep32165.

Rzhetsky A, Nei M. 1992. A simple method for estimating and testing minimum-evolution trees. *Mol Biol Evol* 9:945–967.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425.

Schloss PD, Handelsman J. 2004. Status of the microbial census. *Microbiol Mol Biol Rev* 68:686–691.

Shao J et al. 2015. Analysis and cloning of the synthetic pathway of the phytohormone indole-3-acetic acid in the plant-beneficial *Bacillus amyloliquefaciens* SQR9. *Microb Cell Fact* 14:130. DOI: 10.1186/s12934-015-0323-4.

Spaepen S et al. 2007. Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiol Rev* 31:425–448.

Stackebrandt E, Goebel BM. 1994. Taxonomic note: a place for DNA-DNA reassocation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* 44:846–849.

Susilowati DN et al. 2018. Indole-3-acetic acid producing bacteria and its application on the growth of rice. In: *Proceedings of the 5th International Conference on Biological Science*. Yogyakarta: AIP Conference Proceedings. pp. 1-9.

Thyssen M. 2005. Spatial distribution of heterotrophic bacteria in the northeast Atlantic (POMME study area) during spring 2001. *J Geophys Res* 110:1–16.

Tindall BJ et al. 2010. Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* 60:249–266.

Truong DT et al. 2017. Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* 27:626–638.

Tshikhudo P et al. 2013. Bacterial species identification getting easier. *Afr J Biotechnol* 12:5975–5982.

Vilas-Bôas GT et al. 2007. Biology and taxonomy of *Bacillus cereus*, *Bacillus anthracis*, and *Bacillus thuringiensis*. *Can J Microbiol* 53:673–687.

Wang Y, Qian PY. 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE* 4:e7401. DOI: 10.1371/journal.pone.0007401.

Werner JJ et al. 2012. Impact of training sets on classification of high-throughput bacterial 16S rRNA gene surveys. *ISME J* 6:94–103.

Woo PCY et al. 2009. Guidelines for interpretation of 16S rRNA gene sequence-based results for identification of medically important aerobic Gram-positive bacteria. *J Med Microbiol* 58:1030–1036.