Relationship Between Phylogenetic of Apium and Foeniculum Plants from Central Java, Indonesia, and Their Secondary Metabolites Potency against COVID-19 Protease

Hermín Pancasakti Kusumaningrum*, Rejeki Siti Ferniah¹, Siti Nur Jannah¹, Mufida Budi Kurniawati¹, Anis Afifah¹, Yoshua Mario Sumbodo¹, Salma Sefilia Hanif¹, Tia Erfianti¹, Yuriza Eshananda**

¹Biotechnology Study Program, Department of Biology, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia; ²Post Graduate Study Program, Faculty of Biology, Gadjah Mada University, Yogyakarta, Indonesia; ³Biology Study Program, Faculty of Biology, Jenderal Soedirman University, Purwokerto, Indonesia

Abstract

BACKGROUND: The emergence of COVID-19 in the late of 2019 resulted in the massive screening of drugs, including natural products, to support the current vaccines. Apium and Foeniculum vegetables are members of the Apiaceae family that potentially used to be natural immunosuppressant.

AIM: The purpose of this research is to analyze the phylogenetic position between these two plants as well as find out their secondary metabolites potency against COVID-19 main protease (Mpro) and the papain-like protease (PLpro).

METHODS: The phylogenetic analysis of Apium and Foeniculum from Indonesia was carried out based on internal transcribed spacer (ITS) region and the bioactive virtual screening assay was completed through AutoDock Vina software.

CONCLUSION: Overall, Apium and Foeniculum have close relationships among the members of Apiaceae after maximum likelihood analysis. Furthermore, it also has 70 similar bioactive compounds that some of these potentially inhibit both of COVID-19 proteases.

Introduction

The initial case of COVID-19 pandemic was reported first in the seafood market of Wuhan, China, at the end of 2019. The infected people in the surrounding areas suffered from multiple symptoms such as fever, cough, short of breath, headache, and even diarrhea [1]. It was initially called the novel coronavirus (2019-nCoV) until the genomic analysis was conducted to obtain the exact phylogenetic position [2]. The analysis showed that the novel coronavirus was taxonomically included in the SARS-CoV group and it was named SARS-CoV-2 as the naturally taxonomic practice in the SARS group [3], [4].

There are two viral cysteine proteases that are encoded by SARS-CoV-2, namely, the main protease (Mpro) and the papain-like protease (PLpro). All of these enzymes have been validated as drug targets to inhibit viral replication [5]. Several drugs such as amprenavir, darunavir, lopinavir, and ritonavir have been developed specifically as viral protease targeting inhibitors [6]. However, the screening of SARS-CoV-2 protease inhibitors, mainly from herbal medicine, is still needed to find the potential alternative compounds to support the massive ongoing vaccination.

Celery plant (Apium graveolens L.) is the member of Apiaceae family that spreads across Europe and the tropical and subtropical regions of Africa and Asia [7]. In Indonesia, it is usually used as the condiment for salad and soup to strengthen their flavor [8]. Regarding pharmacological activity, several researchers reported that Apium has antimicrobial activity, antifungal activity, antioxidant activity, and antidiabetic activity [9]. However, the screening and study about its antiviral activity has not existed yet.
Fennel plant (*Foeniculum vulgare* Mill.) is the consumable plant with medicinal usage to treat various diseases and has an impact on human hormones. In Indonesia, *Foeniculum* is available to grow at an altitude more than 1500 m above sea level [10]. In addition, it also has other pharmacological abilities such as antioxidant, antitumor, chemopreventive, cytoprotective, hepatoprotective, hypoglycemic, and estrogenic activities [11]. There is also no report about the antiviral ability from this plant.

Both *Apium* and *Foeniculum* are the members of Apiaceae which are the largest family in the order of Apiales and have been well known as the medicinal and consumable plants [12]. Nevertheless, this family is currently under the major taxonomic rearrangements [13]. Therefore, the aim of this research is to reveal the phylogenetic position as well as the connection with the secondary metabolites content between *Apium* and *Foeniculum*. Furthermore, we also evaluated the ability of their secondary metabolites against COVID-19 main protease and papain-like protease through virtual screening.

**Materials and Methods**

**Plant materials and DNA extraction**

The plants used in this study (*Apium* and *Foeniculum*) were obtained from Central Java, Indonesia. The DNA extraction was carried out using Doyle and Doyle methods [14]. The concentration and purity of DNA was measured using Nanodrop2000.

**DNA amplification, analysis, and phylogenetic relationship**

DNA amplification of the internal transcribed spacer (ITS) gene (ITS5: 5′-GGA AGT AAA AGT CGT AAC AAG G-3′ and ITS4: 5′-TCC TCC GCT TAT TGA TAT GC-3′) was carried out by a total reaction of 50 μl using MY TAQ HS REDMIX from Bioline. Polymerase chain reaction for 35 cycles was finished with initial denaturation at 94°C for 3 min; denaturation at 94°C for 25 s, annealing at 55°C for 30 s, and extension at 72°C for 50 s; and final extension at 72°C for 2 min, according to our previous research [15]. The phylogenetic relationship was analyzed through Molecular Evolutionary Genetics Analysis (MEGA) X software [16] using the maximum likelihood model with Kimura 2-parameter method.

**Secondary metabolites analysis**

The secondary metabolites data of *Apium* and *Foeniculum* plant were extracted from two databases comprise of KNAPSAcK: A Comprehensive Species-Metabolite Relationship Database (http://www.knapsackfamily.com/KNAPSAcK/) and Dr. Duke’s Phytochemical and Ethnobotanical Databases (https://phytochem.nal.usda.gov/phytochem/search). The similar secondary metabolites between these two plants then used as ligands for further molecular docking analysis. All of SDF format of these ligands were downloaded from PubChem database (http://pubchem.ncbi.nlm.nih.gov) and ZINC (https://zinc20.docking.org/substances/subsets/natural-products/).

**Protein and ligand preparation**

The structure of SARS coronavirus main protease (COVID-19 3CL\textsuperscript{pro}/M\textsuperscript{pro}) and papain-like protease was retrieved from the structure database (PDB) with ID 6LU7 [17] and 6WX4 [18], respectively. The native ligands from both receptors were separated as positive control and the active site of the grid box is adjusted by position of native ligands in the structure through AutoGrid in Autodock 4.0 software [19]. The grid size was set to 25 × 54 × 32 (6LU7) and 30 × 30 × 40 (6WX4) xyz points. The dimension of grid center for 6LU7 was −9.768 (x), 11.436 (y), and 68.904 (z) while for 6WX4 was 9.507 (x), −27.445 (y), and −37.252 (z). The ligand molecules were minimized using MMFF94 and converted into.pdbqt format using Open Babel program in Linux operating system.

**Molecular docking analysis**

The molecular docking process is carried out using AutoDock Vina (AV) software [20] operated in Windows subsystem Linux throughout the scripting process in Linux terminal. The binding affinity value and RMSD used as the main parameter. In addition, the molecular interaction between ligands and receptors was analyzed using Biovia Discovery Studio software and LIGPLOT.

**Lipinski rule of five test, absorption, and toxicology analysis**

The physicochemical test of the top 10 metabolites with highest binding against COVID-19 main proteases affinity was carried out by entering the sdf 2D file format (obtained from the PubChem database http://pubchem.ncbi.nlm.nih.gov) on the Lipinski rule of five pages (http://www.scbio-litd.res.in/software/drugdesign/lipinski.jsp). Meanwhile, the absorption and toxicology properties analyzed through pkCSM server (http://biosig.unimelb.edu.au/pkcsms/prediction) and Pro-Tox II (http://tox.charite.de/protox_II/), respectively.
Results and Discussion

In this study, the DNA isolation results of *Apium* and *Foeniculum* plants used the Doyle and Doyle method without the addition of polyvinylpyrrolidone (PVP) are exhibited in Table 1. Sambrook and Russell [21] argue that the use of cetyltrimethylammonium bromide (CTAB) which is a non-ionic detergent that can precipitate nucleic acids and polysaccharides from solutions with low ionic strength has a function to help purify nucleic acids from other organisms that produce polysaccharides in large quantities such as plant. Based on research by Saravanaperumal and La Terza [22], the use of DNA isolation procedures with CTAB without additional antioxidants (PVP) for young *Foeniculum* leaf tissue is able to produce superior DNA quality.

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Plant</th>
<th>DNA concentration (ng/μl)</th>
<th>A260/A280</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Apium</td>
<td>366.7</td>
<td>1.75</td>
</tr>
<tr>
<td>2</td>
<td>Foeniculum</td>
<td>300</td>
<td>2.04</td>
</tr>
</tbody>
</table>

The two samples have nearly representative DNA purity which ranged from 1.8 to 2.0. The DNA purity <1.8 indicates the presence either of phenolic compounds or protein. Meanwhile, the value more than 2.0 means the contamination of RNA during the isolation process [23]. Previous research by Chaudary et al. [24] showed that the DNA extraction of *Foeniculum* from India seed bank using CTAB method had the yield of more than 1000 ng/μl with the purity of 1.8. Regarding the *Apium* plant, Pafundo et al. [25] reported that the DNA isolation of *Apium* using various methods yielded <50 ng/μl with purity ranging from 1.6 to 2.0.

The results of DNA amplification showed that the region of internal transcribed spacer (ITS) in the *Apium* and *Foeniculum* is about 700 bp length, as shown in Figure 1. These were confirmed with the DNA alignment in the National Centre of Biotechnological Information (NCBI). The sequence of *Foeniculum* has 634 bp length and is highly similar with *Foeniculum vulgare* voucher 200505069 (99,53%). On the other hand, the *Apium* sequence (711 bp) is closely related with *Apium graveolens* voucher M.V. Chase 2523 (99%). All of these results indicate that all the samples in this research belong to the similar species which has been registered in GenBank.

In a previous study by Zhou et al. [26] using several plants belonging to the same family as *Foeniculum*, namely, *Apiaceae* from India, a PCR amplification process was carried out using ITS 4 and ITS 5 primers. The PCR condition comprises of pre-denaturation at 94°C for 3 min, denaturation at 94°C for 3 min 45 s, the annealing temperature is 55°C for 1 min, extension 72°C for 3 min and final extension 72°C for 7 min. The similar annealing temperature in PCR condition with this study gives indication that optimal annealing temperature of *Apiaceae* family for ITS amplification gene is 55°C.
are in the same branch (monophyletic) or in the same taxa. Based on the phylogenetic analysis above, it can be seen that the scale substitution rate (scale bar) is 0.02, this value means that for every 100 bases, there are two different nucleotide substitutions. The existence of genetic distances affects the level of gene differences (genomic differences) in a population [28]. The result in Figure 2 also gives indication that Apium and Foeniculum are the closely related species among the group of Apiaceae even though separated in two distinct clades.

Besides the data of phylogenetic relationship, we obtained hundreds secondary metabolites contained in Apium and Foeniculum plant, respectively (Supplementary 1) collected from both KNApSAcK and Dr. Duke’s Database Phytochemical and Ethnobotanical Databases. KNApSAcK database was built in the year of 2011 for metabolomics research purpose [29] and currently containing 141,466 of metabolite-species pair. Meanwhile, the Dr. Duke’s database has existed since the year of 1992 and stores thousands of biologically active phytochemicals from higher plants. Until now, there are about 49,000 plant secondary metabolites that can be explored [30]. Interestingly, this research found 87 compounds that exactly match between these two plants and 70 compounds that could be used for further virtual screening against COVID-19 main protease and papain-like protease. Others categorized as elements that could not be used for molecular docking. We selected the 10 out of 70 compounds based on the lower binding energy and root mean square deviation (RMSD) results as can be seen from Table 2 (6LU7) and Table 3 (6WX4). The complete virtual screening results are available in Supplementary 2.

The virtual screening results in Table 2 showed that rutin has the closest binding affinity towards 6LU7 receptor (N3) followed by stigmasterol, campesterol, and beta-carotene (binding energy ≤ 9). The binding between protein and ligand occurs when the required energy is low so that the more negative binding affinity number gives higher possibility regarding the stability interaction between ligands and receptor [31]. However, all of these compounds did not meet the criteria of Lipinski’s rule of five. Regarding the rutin, it only meets the criteria of logP (1.8) while others exceed maximum number (5) (Supplementary 3). It can also be observed that only three compounds in the top 10 of the virtual screening, namely, riboflavin, colombiane, and beta-caryophyllene that fulfill Lipinski’s rule of five. All of these compounds have differences of about 3.0 kcal/mol of binding energy compared to N3.

The RMSD value for all screened ligands showed the value <2 which indicates that the docking method used has been well validated [32]. Interestingly, the virtual screening against 6WX4 papain-like proteases has almost the similar result regarding the top 10 bioactive compounds with lowest binding energy. Nevertheless, almost all of the bioactive compounds had lower or similar energy numbers compared to VIR251 as native ligand. It means that almost all of these ligands can bind spontaneously to protein [31]. Although there are only three out of 10 compounds that fulfill Lipinski’s rule of five, the presence of seselin might become the promising compound to inhibit the 6WX4 papain-like protease. Besides the lower binding affinity compared to the native ligand, previous in vitro research revealed the antiviral ability of seselin compound against nuclear polyhedrosis virus infection in the larvae silkworm [33]. The result of drugs oral availability in Table 4 based on water solubility, Caco-2 cell permeability (Papp), and human intestinal absorption (HIA) indicating

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Bioactive compound</th>
<th>Molecular formula</th>
<th>Binding energy (Kcal/mol)</th>
<th>RMSD upper/lower bound</th>
<th>Lipinski rule of five</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rutin</td>
<td>C_{6}H_{10}O_{2}</td>
<td>-8.3</td>
<td>0.0</td>
<td>Not fulfill</td>
</tr>
<tr>
<td>2</td>
<td>Stigmasterol</td>
<td>C_{6}H_{10}O</td>
<td>-8.6</td>
<td>0.0</td>
<td>Not fulfill</td>
</tr>
<tr>
<td>3</td>
<td>Campesterol</td>
<td>C_{6}H_{10}O</td>
<td>-8.6</td>
<td>0.0</td>
<td>Not fulfill</td>
</tr>
<tr>
<td>4</td>
<td>Beta-carotene</td>
<td>C_{6}H_{10}O</td>
<td>-8.4</td>
<td>0.0</td>
<td>Not fulfill</td>
</tr>
<tr>
<td>5</td>
<td>Beta-sitosterol</td>
<td>C_{6}H_{10}O</td>
<td>-8.3</td>
<td>0.0</td>
<td>Not fulfill</td>
</tr>
<tr>
<td>6</td>
<td>Phytosterol</td>
<td>C_{6}H_{10}O</td>
<td>-8.3</td>
<td>0.0</td>
<td>Not fulfill</td>
</tr>
<tr>
<td>7</td>
<td>Chlorogenic Acid</td>
<td>C_{6}H_{10}O</td>
<td>-8.3</td>
<td>0.0</td>
<td>Fulfill</td>
</tr>
<tr>
<td>8</td>
<td>Riboflavin</td>
<td>C_{6}H_{10}O</td>
<td>-8.3</td>
<td>0.0</td>
<td>Fulfill</td>
</tr>
<tr>
<td>9</td>
<td>Colombianite</td>
<td>C_{6}H_{10}O</td>
<td>-8.3</td>
<td>0.0</td>
<td>Fulfill</td>
</tr>
<tr>
<td>10</td>
<td>Beta-caryophyllene</td>
<td>C_{6}H_{10}O</td>
<td>-8.3</td>
<td>0.0</td>
<td>Fulfill</td>
</tr>
</tbody>
</table>

RMSD: Root mean square deviation.
Table 4: Absorption and toxicity data of selected ligands

<table>
<thead>
<tr>
<th>Bioactive compound</th>
<th>pCKSM</th>
<th>Water solubility (log mol/L)</th>
<th>Caco-2 permeability (cm/s)</th>
<th>HIA (% absorbed)</th>
<th>P-glycoprotein inhibitor/substrate</th>
<th>ProTox II</th>
<th>Acute toxicity class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hepatotoxicity</td>
<td>Carcinogenicity</td>
</tr>
<tr>
<td>Rutin</td>
<td>–2.892</td>
<td>–0.949</td>
<td>23.446</td>
<td>91.732</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Stigmasterol</td>
<td>–6.882</td>
<td>1.213</td>
<td>94.97</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Campesterol</td>
<td>–7.068</td>
<td>1.223</td>
<td>94.53</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Beta-carotene</td>
<td>–7.39</td>
<td>1.262</td>
<td>91.732</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Beta-sitosterol</td>
<td>–6.773</td>
<td>1.201</td>
<td>94.464</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Phytochemicals</td>
<td>–6.773</td>
<td>1.201</td>
<td>94.464</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Chlorogenic acid</td>
<td>–2.449</td>
<td>–0.84</td>
<td>36.377</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Riboflavin</td>
<td>–2.029</td>
<td>–0.679</td>
<td>36.107</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Beta-caryophyllene</td>
<td>–5.555</td>
<td>1.423</td>
<td>96.577</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Colombianite</td>
<td>–3.172</td>
<td>1.18</td>
<td>97.946</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
<tr>
<td>Seselin</td>
<td>–3.401</td>
<td>1.129</td>
<td>97.946</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Inactive</td>
<td>Active</td>
</tr>
</tbody>
</table>

HIA: Human intestinal absorption.
and Pro248 (Supplementary 5) in 6WX4 protein. Only riboflavin has hydrogen interactions through Asp164 that we can also find in VIR251 (Figure 4). Based on Rut et al. [18], a deep hydrophobic pocket in papain-like protease, formed in the residue of Tyr264, Tyr268, Pro247, Pro248, Tyr273, Thr301, and Met208. The previous study [40] found that the hydrogen interaction with Asp164 increased the stability of the complexes. Therefore, riboflavin based on this study forms the most stable interaction with 6WX4.

Figure 3: The comparison of molecular interaction between N3 as native ligands of 6LU7 and riboflavin

Figure 4: The comparison of molecular interaction between VIR251 as native ligands of 6WX4 and riboflavin
Based on the research result, the close phylogenetic analysis between *APIUM* and *FOENICULUM* plants was followed by the similarity of several secondary metabolites which have promising ability to inhibit both of COVID-19 protease and resulted riboflavin as the best compound which meets drug-like properties and toxicity criteria. These results supported the previous study which shows that natural immunosuppressant compounds derived from plant sources are known to inhibit the production and release of pro-inflammatory cytokines and chemokines in COVID-19 [41], [42], [43], [44]. The use of natural immunosuppressants from these two plants was used as an adjuvant to ameliorate cytokine storms, as was other plants from the previous studies. The results show the potential of the two plants which when combined with other currently used antiviral agents and drugs will have the potential as a new, synergistic approach for more effective COVID-19 treatment and cure.

**Conclusion**

Phylogenetic analysis indicates that Indonesian *APIUM* and *FOENICULUM* have close relationships inside of the family *APIACEAE*. Besides, it also shares similar bioactive compounds that potentially bind strongly to the COVID-19 protease receptors. Further *in vitro* assay is needed to validate this study.

**References**

20. Trot O, Olson AJ. AutoDock vina: Improving the speed
Kusumaningrum et al. Apium and Foeniculum Against COVID-19 Protease


