Indonesian Journal of Biotechnology

VOLUME 22(1), 2017, 43–48 | RESEARCH ARTICLE

Simultaneous clustering analysis with molecular docking in network pharmacology for type 2 antidiabetic compounds

Nur Azizah Komara Rifai[1,*], Farit Mochamad Afendi[1], and I Made Sumertajaya[1]

[1]Department of Statistics, Bogor Agricultural University, Jalan Meranti Wing 22 Level 4, Kampus IPB Darmaga, Bogor 16680, Indonesia
[*]Corresponding author: azizah_kr@yahoo.com

ABSTRACT The database of drug compounds and human proteins plays a very important role in identifying the protein target and the compound in drug discovery. Recently, a network pharmacology approach was established by updating the research paradigm from the current "one disease-one target-one drug" to a new "drug-target-disease network". Ligand-protein interactions can be analyzed quantitatively using simultaneous clustering and molecular docking. The docking method offers the ability to quickly and cheaply predict the ligand-protein binding free energy (ΔG) in structure-based virtual screening. Meanwhile, simultaneous clustering was used to find subgroups of compounds that exhibit a high correlation with subgroups of target proteins. This study is focused on the interaction between the 306 compounds from medicinal plants (brotowali *Tinospora crispa*, ginger *Zingiber officinale*, pare *Momordica charantia*, sembung *Blumea balsamifera*), synthetic drugs (FDA-approved) and the 21 significant human proteins associated with type 2 diabetes. We found that brotowali (B018), sembung (S031), pare (P231), and ginger (J036, J033) were close to the synthetic drugs and can possibly be developed as antidiabetic drug candidates. Likewise, the proteins AKT1, WFS1, APOE, EP300, PTH, GCG, and UBC which assemble each other and which have a high association with INS can be seen as target proteins that play a role in type 2 diabetes.

KEYWORDS molecular docking; network pharmacology; simultaneous clustering analysis; type 2 diabetes

1. Introduction

Type 2 diabetes is a chronic disease characterized by the body being unable to effectively produce insulin. World Health Organization (2016) reported that the number of people with diabetes has risen from 108 million in 1980 to 422 million in 2014. Diabetes is one of the leading causes of death in the world and it has been rising more rapidly in low-middle income countries such as Indonesia. Indonesia was ranked one of the world's top ten countries for about 10 million adults with diabetes in 2015 (International Diabetes Federation 2015) and was estimated to reach 21.3 million by 2030 (World Health Organization 2016). This suggests that genetic risk factors, obesity, unhealthy diet, and physical inactivity have also increased. This case of diabetes now occurs not only in adults but also in children and adolescents. Thus, diabetes problems need to be handled seriously.

From a metabolic perspective, a disease occurs because the protein function is impaired. In order to make the protein function normally, it needs to be treated with drugs that contain several chemical compounds that can influence or inhibit the activity of proteins or target networks. In recent years, a network pharmacology approach was established by updating the research paradigm from the current "one disease-one target-one drug" to a new "drug-target-disease networks" (Yang et al. 2013). Therefore, determination of the compounds that target the proteins associated with a certain disease is crucial to understanding the molecular mechanisms in drug design.

There are several methods to investigate drug candidate at the molecular level. One of the most common uses in computational technique (*in silico*) is molecular docking. Molecular docking offers the ability to predict quickly and cheaply the ligand-protein interaction in structure-based virtual screening which results in the binding free energy (ΔG) scores. The binding free energy (ΔG) scores show the bond strength between molecules or the magnitude of ligand-protein interaction. Lalitha and Sripathi (2011) have shown that Xylitol can be used for antidiabetic using molecular docking. However, the compounds isolated from medicinal plants are safer and more appropriately used for metabolic and degenerative diseases (Greer et al. 1994). Thus, in this study we use the ligand not only from synthetic drugs but also from medicinal plants associated with type 2 diabetes.

The ligand-protein interaction can be analyzed quantitatively by clustering method in order to find subgroups of compounds and subgroups of target proteins that exhibit a high correlation in two-way data analysis. Most of the standard clustering literature focuses on one-sided clustering algorithms, but in bioinformatics, it allows to deal with sparse and high dimensional data matrices. Madeira and Oliveira (2004) have been proposed a survey of simultaneous clustering algorithms on biological data. Simultaneous clustering is a method to avoid some limitations of standard clustering approach. Simultaneous clustering is more robust and more informative than standard clustering, as it involves rows and columns together. Thus, this study is focused on the interaction between the compounds of medicinal plants, synthetic drugs (FDA-approved) and the human proteins associated with type 2 diabetes by using simultaneous clustering with molecular docking approach.

2. Materials and methods

2.1. Dataset preparation

The data of 287 compounds of four medicinal plants (Qomariasih 2015) including 15 from brotowali (*Tinospora crispa*), 173 from ginger (*Zingiber officinale*), 87 from pare (*Momordica charantia*), 12 from sembung (*Blumea balsamifera*) and 19 synthetic drugs were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov) in .sdf format and were modeled using Marvinsketch in .pdb format. In addition, the 21 human protein associated with type 2 diabetes (Usman 2016) sequences were obtained from NCBI (https://www.ncbi.nlm.nih.gov) and were modeled using SWISS-MODEL (http://swissmodel.expasy.org) in .pdb format.

2.2. Molecular docking

Molecular docking is a method that predict the most favorable orientation of a ligand when interacting with a protein to form a stable complex (Nogara et al. 2015). Ligand-protein docking is an important type molecular docking in modern structure-based drug design (Huang and Zou 2010). There are two essential components in ligand-protein docking method, namely the search algorithm and the scoring function. The search algorithm is responsible for searching through different ligand conformations and orientations (poses) within a given target protein. The scoring function is responsible for estimating the binding affinities of the generated poses, ranking them, and identifying the most favorable binding modes of the ligand to the given target. Before beginning the docking, drug-likeness for the ligand was analyzed by Lipinski's Rule parameters including hydrogen bond donors ≤ 5, hydrogen bond acceptors ≤ 10, molecular weight ≤ 500 g/mol, and partition coefficient logP ≤ 5 (Lipinski et al. 2001). Molecules violating more than one of these rules may delete because they have problems with bioavailability. Optimization of ligand geometry was conducted by using wash to improve the structure of the ligand and the position of hydrogen atom. Minimization of ligand energy was conducted by using a modified Merck Molecular Forcefield 94 (MMFF94) and gradient root mean square (RMS) 0.001 kkal/Åmol. Whereas, optimization of protein geometry was conducted by adding polar hydrogens, protonation, and partial charges. Minimization of protein energy was conducted using Merck Molecular Forcefield 94x (MMFF94x) and solvation in gas phase with fixed charges, then minimize with gradient root mean square (RMS) 0.05 kkal/Åmol. Ligand-protein docking process used placement Triangle Matcher with retain 5. Then, scoring function used London dG, refinement is Forcefield with retain 1. Molecular docking produces a binding free energy (ΔG). Thermodynamically, the ligand-protein interaction occurs when it produces $\Delta G < 0$. The smaller values of ΔG, the ligand-protein bonds are more stable or ligand-protein interactions are stronger.

2.3. Simultaneous clustering

Simultaneous clustering performs clustering in the two dimensions simultaneously (Charrad and Ahmed 2011). Simultaneous clustering, usually called by bi-clustering, co-clustering, two-way clustering, or block clustering, is an important technique to find sub-matrices. The sub-matrices are subgroups of rows and subgroups of columns that exhibit a high correlation in two-way data analysis. In this study, singular value decomposition (SVD) approach was applied for simultaneous clustering analysis. If $\mathbf{X}$ is an $n \times p$ matrix with n observations in a row and p variables in column then the SVD is (Jolliffe 2002)

$$\mathbf{X} = \mathbf{ULA}^T \qquad (1)$$

where $\mathbf{U}$ and $\mathbf{A}$ are $n \times r$ and $p \times r$ orthonormal column matrices ($\mathbf{U}^T\mathbf{U} = \mathbf{A}^T\mathbf{A} = \mathbf{I}_r$), $\mathbf{L}$ is $r \times r$ diagonal matrix, and r is the rank of $\mathbf{X}$.

The SVD in the dimension reduction is relevant to PCA in several respects. The columns of the matrices $\mathbf{U}$ and $\mathbf{A}$ are eigenvectors of the matrices $\mathbf{XX}^T$ and $\mathbf{X}^T\mathbf{X}$, respectively, and the decreasing non-negative entries $l_1^{1/2} \geq l_2^{1/2} \geq ... \geq l_r^{1/2}$ in the diagonal matrix $\mathbf{L}$ are square roots of the non-zero eigenvalues of $\mathbf{XX}^T$ and also of $\mathbf{X}^T\mathbf{X}$. We denote the ith columns of the matrices $\mathbf{U}$ and $\mathbf{A}$ by u_i and a_i, respectively. The vectors u_i and a_i are called the left and right singular vectors of $\mathbf{X}$, and the values l_i are called the *singular values*.

The SVD in clustering is a generalization of the algorithm shows a transformation that permit us to get two matrices from one matrix. Now define $\mathbf{L}^\alpha$, for $0 \geq \alpha \geq 1$, as the diagonal matrix whose elements are $l_1^{\alpha/2}, l_2^{\alpha/2}, \ldots, l_r^{\alpha/2}$ with a similar definition for $\mathbf{L}^{1-\alpha}$, and let $\mathbf{G} = \mathbf{UL}^\alpha, \mathbf{H}^T = \mathbf{L}^{1-\alpha}\mathbf{A}^T$. Then

$$\mathbf{GH}^T = \mathbf{UL}^\alpha\mathbf{L}^{1-\alpha}\mathbf{A}^T = \mathbf{ULA}^T = \mathbf{X} \qquad (2)$$

We used $\alpha = \dfrac{1}{2}$ so $\mathbf{G} = \mathbf{UL}^{0.5}$ and $\mathbf{H}^T = \mathbf{L}^{0.5}\mathbf{A}^T$. $\mathbf{G}$

and **H** matrices represent the information of compounds and proteins, respectively. Then, these matrices were analyzed by hierarchical clustering and plotted in heat map with two-dimensional dendrograms simultaneously. **G** matrix produced a column-side dendrogram and **H** matrix produced a row-side dendrogram, or reversed. A heat map is a literal way of visualizing the binding free energy (ΔG) scores with colored cells. We used the Euclidean distance and linkage method both for clustering of compounds and proteins to built a two-dimensional dendrograms.

Euclidean distance: the usual square distance between two vectors, is given by:

$$d(x,y) = (\sum_{j=1}^{d}(x_j - y_j)^2)^{\frac{1}{2}} \tag{3}$$

Single linkage: the distance D_{ij} between two clusters C_i and C_j is the minimum distance between two points x and y, with $x \in C_i$ and $y \in C_j$, is given by:

$$D_{ij} = \min_{x \in C_i, y \in C_j} d(x,y) \tag{4}$$

Complete linkage: the distance D_{ij} between two clusters C_i and C_j is the maximum distance between two points x and y, with $x \in C_i$ and $y \in C_j$, is given by

$$D_{ij} = \max_{x \in C_i, y \in C_j} d(x,y) \tag{5}$$

Average linkage: the distance D_{ij} between two clusters C_i and C_j is the mean of the distance between the pair of points x and y, where $x \in C_i$ and $y \in C_j$, is given by:

$$D_{ij} = \sum_{x \in C_i, y \in C_j} \frac{d(x,y)}{n_i \times n_j} \tag{6}$$

Ward: the total within-cluster sum of square (SSE) is computed to determine the next two groups merged at each step of algorithm, where $\mathbf{x}_j$ is multivariate measurement associated with the jth object and $\bar{\mathbf{x}}$ is the mean of all the object, is given by:

$$SSE = \sum_{j=1}^{N}(\mathbf{x}_j - \bar{\mathbf{x}})'(\mathbf{x}_j - \bar{\mathbf{x}}) \tag{7}$$

Furthermore, we compared the performance of linkage method using heat maps against a number of reasonable benchmarks.

3. Results and discussion

The results of drug likeness test showed that out of 306 compounds, 199 compounds (185 from medicinal plants and 14 from synthetic drugs) satisfied the Lipinski's Rule properties. The boxplot of ΔG scores from medicinal plant and synthetic drug compounds in Figure 1 are not greatly different. Although the mean of ligand-protein interaction from medicinal plants is lower than synthetic drugs, some of medicinal plant compounds produce the lowest ΔG scores < -15 kJ/mol. The most stable binding complex or strongest interaction for medicinal plants is -16.97 kJ/mol (J156 and INS) whereas for synthetic drugs is -14.56 kJ/mol (DB11 and INS). Therefore, some medicinal plants have a better stability than synthetic drugs and it can possible be developed as antidiabetic drug candidates.

The relative position of the compounds based on the target proteins can be described by plot PCs (Figure 2). The plot of first two PCs in Figure 2 shows that compounds tend to gather in quadrant III and outside quadrant III. The synthetic drug compounds (red color) spread in quadrants I, II, and III. In quadrant I, there is one synthetic drug compound that is close to pare (purple color) and ginger (black color). In quadrant II, there are 4 synthetic drug compounds that tend to gather with ginger and sembung (green color). In quadrant III, there are many synthetic drugs that are close to each other with some medicinal plant compounds of brotowali (blue color), ginger, pare, and sembung. In quadrant IV, there is no synthetic drug compounds. The first two PCs explain 87.29% of the total variance. This plot is an exploration of the compound grouping based on PCs of target proteins. Furthermore, the cluster analysis was performed using simultaneous clustering.

The index values of cluster validity show that the best number of clusters for the compounds is two clusters (Table 1) and for the proteins is two clusters (Table 2). The hierarchical method that produce the most two clusters for the compounds is complete linkage method, whereas for the proteins are complete linkage and Ward method. Therefore, the best hierarchical method used for the simultaneous clustering is complete linkage method.

The two-dimensional dendrogram generated by complete linkage method in Figure 3 is trimmed into two clusters both in row (protein) and column (compound). The strong ligand-protein interactions (light yellow colors) are found horizontally on the right-side (quadrant I and IV) and vertically on the up-side (quadrant I and II). In addition, 13 synthetic drug compounds are more commonly

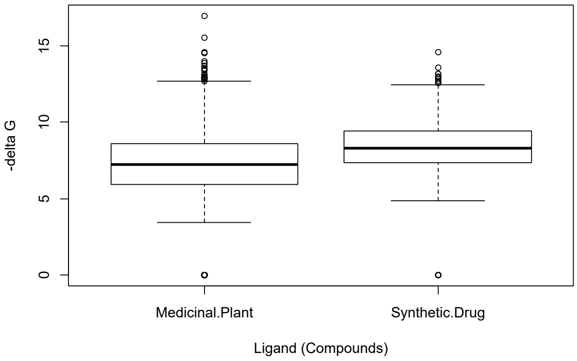


FIGURE 1 Boxplot of ligand-protein interaction (ΔG) scores.

TABLE 1 Cluster validity of ligand.

Validity Index	Linkage Method	Number of clusters (compound)								
		2	3	4	5	6	7	8	9	10
CH	Single	4.75[a]	3.82	4.26	3.87	3.77	3.67	3.53	3.36	3.30
	Complete	15.18[a]	9.59	7.93	8.22	7.35	9.15	9.55	8.90	8.30
	Average	4.75	3.95	3.62	3.38	4.04	3.77	3.57	3.41	3.64
	Ward	26.25[a]	23.06	18.82	16.49	15.16	14.27	13.35	12.68	12.18
Pseudo t^2	Single	0.21[a]	4.98	-0.12	0.47	0.22	0.05	-0.42	0.09	-0.21
	Complete	3.88[a]	-0.69	8.46	3.53	19.42	9.28	3.57	0.00	5.34
	Average	-1.62[a]	0.21	-0.11	6.31	-0.49	0.03	0.37	4.93	6.07
	Ward	18.83	11.30	7.12[a]	8.76	6.82	7.68	5.24	-1.60	5.14
Frey	Single	39.24	15.67	6.38[a]	-0.40	1.23	4.72	8.82	2.33	3.99
	Complete	-1.38	-1.49	0.45	-0.69	-0.21	1.75	-0.14	-0.32	0.34
	Average	30.17	14.95	12.84	6.02	6.33	5.57	5.43	4.84	3.27
	Ward	0.07	0.82	0.35	2.10	0.68	-0.10	0.29	-0.14	0.26
Gap	Single	0.52	0.53[a]	0.03	-1.24	-1.78	-2.17	-2.18	-2.45	-2.50
	Complete	-0.25[a]	-1.05	-1.56	-1.95	-2.38	-2.52	-2.79	-2.96	-3.14
	Average	-0.39[a]	-1.11	-1.74	-2.04	-2.43	-2.87	-3.09	-3.17	-3.34
	Ward	-0.33[a]	-0.85	-1.46	-1.79	-2.20	-2.65	-2.77	-2.91	-3.07

[a] Index values that show the best number of clusters.

TABLE 2 Cluster validity of protein.

Validity Index	Linkage Method	Number of clusters (compound)								
		2	3	4	5	6	7	8	9	10
CH	Single	2.49[a]	1.89	1.74	1.70	1.77	1.72	1.70	1.71	1.64
	Complete	2.49[a]	1.91	1.76	1.96	1.89	1.84	1.79	1.78	1.76
	Average	2.49[a]	1.89	1.74	1.98	1.89	1.81	1.77	1.74	1.69
	Ward	2.49[a]	2.43	2.14	1.98	1.89	1.84	1.80	1.78	1.76
Pseudo t^2	Single	0.00[a]	0.09	0.12	1.74	-0.36	-0.39	-0.38	0.00	-0.19
	Complete	1.32[a]	0.08	2.26	0.00	-0.31	-0.27	-0.40	-0.48	1.21
	Average	0.00[a]	0.09	2.30	-0.35	-0.27	-0.41	-0.50	-0.24	0.13
	Ward	2.26[a]	0.08	0.00	1.39	-0.31	-0.40	-0.27	-0.48	-0.24
Frey	Single	3.44	4.38	1.72	1.46	1.13[a]	0.85	0.53	2.71	0.72
	Complete	22.79[a]	-0.02	0.07	-0.09	0.17	0.14	0.19	0.29	0.58
	Average	3.44	4.38	1.42[a]	0.83	0.69	0.71	0.82	0.73	0.58
	Ward	2.70[a]	0.57	-0.03	4.35	0.17	0.23	0.09	0.29	0.17
Gap	Single	-0.12[a]	-0.27	-1.82	-2.22	-2.64	-2.65	-3.03	-3.70	-4.06
	Complete	-0.30[a]	-1.13	-2.30	-2.47	-2.80	-3.11	-3.48	-3.69	-4.05
	Average	-0.84[a]	-1.60	-2.08	-2.54	-2.91	-3.12	-3.48	-3.70	-4.07
	Ward	-0.84[a]	-1.55	-2.03	-2.54	-2.91	-3.14	-3.48	-3.69	-4.05

[a] Index values that show the best number of clusters.

found on the right-side and only one synthetic drug compound on the left-side. Thus, we only focused on the ligand-protein interactions in quadrant I.

There are 146 medicinal plants clustered with synthetic drugs in quadrant I, those are 7 from brotowali, 95 from ginger, 23 from pare, and 8 from sembung. The medicinal plants which close to synthetic drugs (DB12, DB11, DB08, DB17) and (DB05) are (B018, S031) and (P231), respectively. The synthetic drugs (DB01, DB13, DB02, DB15, DB03, DB18) are close and clustered to each other. Next, the medicinal plants (J036, J033) are close to synthetic drugs (DB06, DB16).

The medicinal plants B018, S031, P231, J036, and J033 which clustered with synthetic drugs in quadrant I,

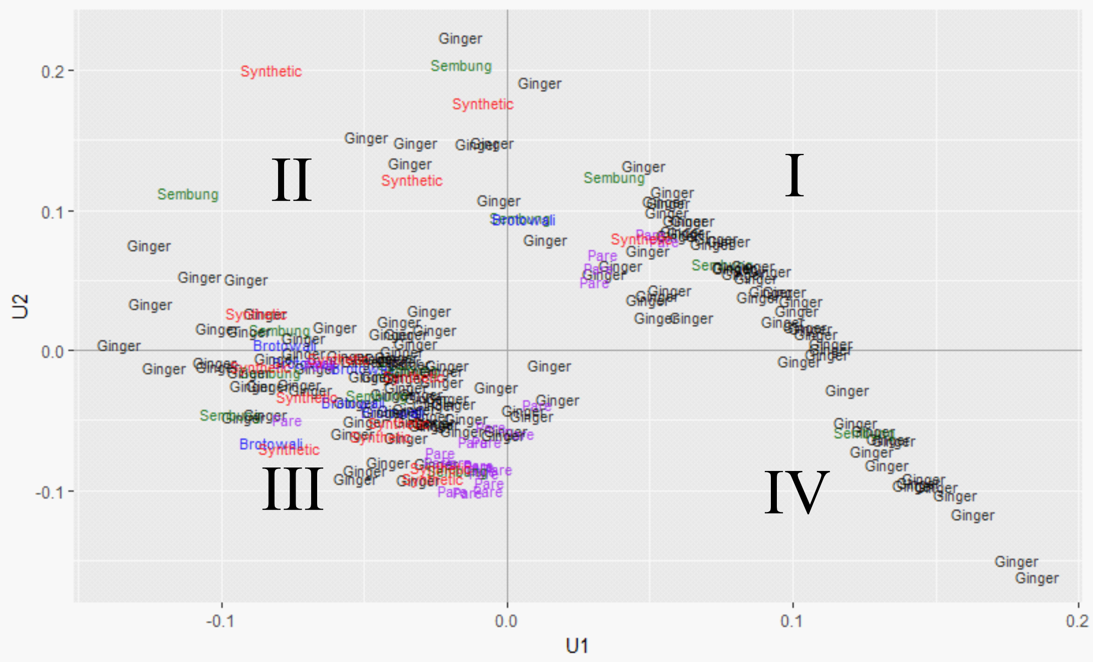
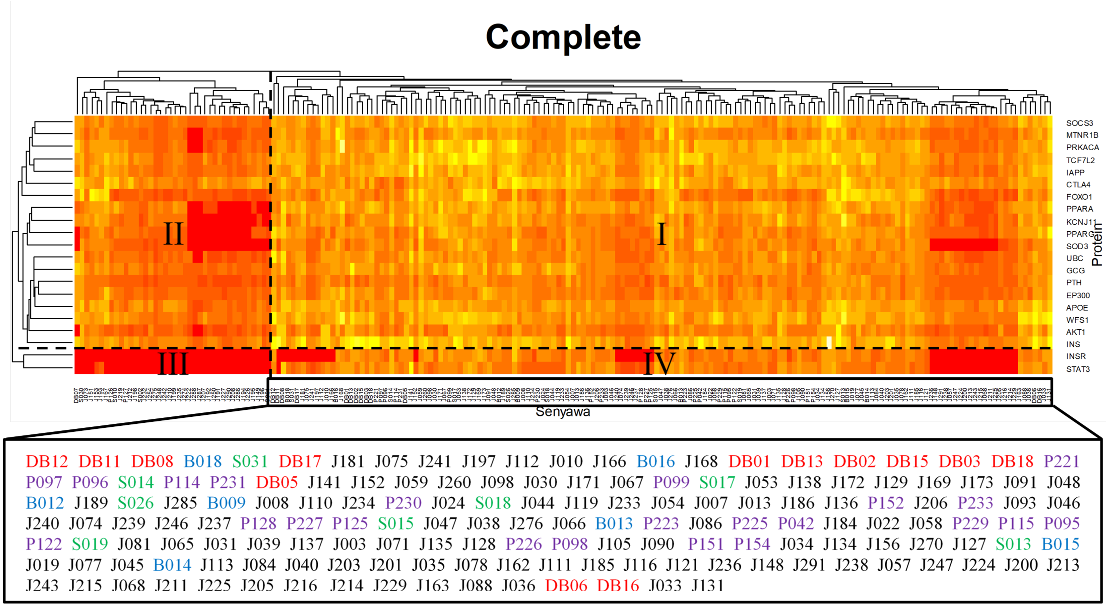


FIGURE 2 Two-dimensional plot of the compounds (medicinal plants and synthetic drugs) based on the first two PCs of target proteins.



DB12 DB11 DB08 B018 S031 DB17 J181 J075 J241 J197 J112 J010 J166 B016 J168 DB01 DB13 DB02 DB15 DB03 DB18 P221
P097 P096 S014 P114 P231 DB05 J141 J152 J059 J260 J098 J030 J171 J067 P099 S017 J053 J138 J172 J129 J169 J173 J091 J048
B012 J189 S026 J285 B009 J008 J110 J234 P230 J024 S018 J044 J119 J233 J054 J007 J013 J186 J136 P152 J206 P233 J093 J046
J240 J074 J239 J246 J237 P128 P227 P125 S015 J047 J038 J276 J066 B013 P223 J086 P225 P042 J184 J022 J058 P229 P115 P095
P122 S019 J081 J065 J031 J039 J137 J003 J071 J135 J128 P226 P098 J105 J090 P151 P154 J034 J134 J156 J270 J127 S013 B015
J019 J077 J045 B014 J113 J084 J040 J203 J201 J035 J078 J162 J111 J185 J116 J121 J236 J148 J291 J238 J057 J247 J224 J200 J213
J243 J215 J068 J211 J225 J205 J216 J214 J229 J163 J088 J036 DB06 DB16 J033 J131

FIGURE 3 Heat map of ligand-protein interaction by complete linkage method.

have the strongest interaction with protein APOE, IAPP, INS, WFS1, and WFS1, respectively. The compound P231 (Kuguacin Q) in pare that clustered with synthetic drug DB05 (Glimepiride) has the lower ΔG scores with protein INS than other proteins that is -10.30 kJ/mol, thus pare can possible be developed as type 2 antidiabetic drug candidates. Moreover, pare is known to have antihyperglycemic effects for both human and animals because it contain alkaloids, flavonoids, saponins, and tannins (Leelaprakash et al. 2011).

Nevertheless, the light yellow-row cells in quadrant I are more commonly found to the protein INS (Insulin). It is also has shown by the highest ΔG scores in quadrant I is for protein INS. That is, the protein INS is more targeted by some compounds especially synthetic drugs than the other proteins. The proteins that closest to protein INS are AKT1, WFS1, APOE, EP300, PTH, GCG, dan UBC.

4. Conclusions

Based on molecular docking and simultaneously clustering analysis was obtained that medicinal plants such as brotowali (B018), sembung (S031), pare (P231), and ginger (J036, J033) are found close to synthetic drugs. Thus, those compounds especially pare (P231) that target the protein insulin (INS) can possible be developed as type 2 antidiabetic drug candidates. Besides, the protein AKT1, WFS1, APOE, EP300, PTH, GCG, dan UBC are found as target proteins that play a role in type 2 diabetes.

Acknowledgments

We would like to thank the Department of Statistics, Bogor Agricultural University for the funding support to this publication and the LPDP scholarship of the Republic of Indonesia Endowment Funding in Education for the funding support to this research project. We would also like to thank Rudi Heryanto, Luthfan Irfana, and Tony Ibnu Sumaryada for helpful discussions about the chemistry and the physics. We are grateful to the anonymous reviewers for their insightful feedback on the content of the manuscript.

Authors' contributions

NAKR and FMA conceived and planned the study. FMA developed the docking idea. IMS developed the clustering. NAKR analyzed the data, perform the computations, and wrote the manuscript. FMA and IMS supervised the findings of this work. All authors discussed the results, read, and approved the final version of the manuscript.

Competing interests

The authors declare that there is no significant competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described in this manuscript.

References

Charrad M, Ahmed MB. 2011. Simultaneous clustering: a survey. In: Pattern recognition and machine intelligence. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. p. 370–375. doi:10.1007/978-3-642-21786-9_60.

Greer J, Erickson JW, Baldwin JJ, Varney MD. 1994. Application of the three-dimensional structures of protein target molecules in structure-based drug design. J Med Chem. 37(8):1035–1054. doi:10.1021/jm00034a001.

Huang SY, Zou X. 2010. Advances and challenges in protein-ligand docking. Int J Mol Sci. 11(8):3016–3034. doi:10.3390/ijms11083016.

International Diabetes Federation. 2015. IDF Diabetes Atlas. 7th edition. Brussels: International Diabetes Federation.

Jolliffe IT. 2002. Principal Component Analysis. 2nd edition. New York: Springer Science & Business Media.

Lalitha P, Sripathi S. 2011. In silico ligand-receptor docking of new cyclitols for type II diabetes using Hex. Pharma Sci Monit. 2(3, Suppl-1):S32–41.

Leelaprakash G, Rose JC, Gowtham BM, Javvaji PK, Prasad SA. 2011. In vitro antimicrobial and antioxidant activity of *Momordica charantia* leaves. Pharmacophore 2(4):244–252.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Delivery Rev. 46(1-3):3–26. doi:10.1016/S0169-409X(00)00129-0.

Madeira SC, Oliveira AL. 2004. Biclustering algorithms for biological data analysis: a survey. IEEE/ACM Trans Comput Biol Bioinf. 1(1):24–45.

Nogara PA, Saraiva RdA, Caeran Bueno D, Lissner LJ, Lenz Dalla Corte C, Braga MM, Rosemberg DB, Rocha JBT. 2015. Virtual screening of acetylcholinesterase inhibitors using the lipinski's rule of five and ZINC databank. doi:10.1155/2015/870389.

Qomariasih N. 2015. Simultaneous clustering analysis and network pharmacology in the determination of active compounds of jamu for diabetes type 2. Thesis. [Bogor, Indonesia]: Bogor Agricultural University.

Usman M. 2016. Identifications of significant proteins associated with diabetes mellitus (DM) type 2 using network topology analysis of protein-protein interactions. Thesis. [Bogor, Indonesia]: Bogor Agricultural University.

World Health Organization. 2016. Global report on diabetes: executive summary. Technical report. World Health Organization. Geneva.

Yang M, Chen JL, Xu LW, Ji G. 2013. Navigating traditional Chinese medicine network pharmacology and computational tools. J Evidence-Based Complementary Altern Med 2013:1–23. doi:10.1155/2013/731969.