Knowledge E
Engaging minds

Conference Paper

Biological Network Analysis of Genes Involve in Embryogenesis of Oil Palm using ClueGO

Widyartini Made Sudania, Zulfikar Achmad Tanjung, Nurita Toruan-Mathius, and Tony Liwang

Plant Production and Biotechnology Division, PT. SMART Tbk, Sinar Mas Land Plaza, 2nd Tower, 10th Fl., Jalan M.H. Thamrin No.51, Jakarta 10350, Indonesia

Abstract

The application of DNA sequencing technologies has a major impact on molecular biology, especially in understanding genes interaction in a certain condition. Due to a large number of genes produced by this high-throughput technology, a proper analysis tool is needed for data interpretation. ClueGO is a bioinformatics tool, an easy to use Cytoscape plug-in that strongly improves biological function interpretation of genes. It analyzes a cluster or comparing two clusters and comprehensively visualizes their group functions. This tool is applied to identify biological networks of genes involved in embryogenesis of oil palm, the most critical phase in oil palm tissue culture process. Two ESTs sequencing data from the GenBank database under accession number EY396120-EY413718 and DW247764-DW248770 were used in this study. Fifty-two and one hundred eight groups of genes were identified using biological process in Gene Ontology setting from the database of EY396120-EY413718 and DW247764-DW248770, respectively. Thirty-one groups of genes were consistently occurred in both ESTs. According to the literature, these genes play an important role in cell formations and developments, stresses and stimulus responses, photosynthesis and metabolic processes that indicate the involvement of these groups of genes in oil palm embryogenesis processes. ClueGO is the appropriate tool to analyze a large data set of genes in a specific condition, such as embryogenesis of oil palm.

🔓 **OPEN ACCESS**

1. Introduction

The use of clonal planting materials in large-scale oil palm plantations offers several advantages compared with seedling planting materials. It allows high productivity (increasing yields between 20% to 30%), rapid multiplication and uniform growth of palms with desired traits [1]. As the limitation, it has various critical aspects that require extra precaution for successful inductions and developments of somatic embryos such as genotype, cytoskeleton, arabinogalactan protein, environment and plant hormones [2]. Somatic embryo formation is an important and major constraint in oil palm tissue culture due to low average rate of embryogenesis at 3% to 6% [3]. On the other hand,

How to cite this article: Widyartini Made Sudania, Zulfikar Achmad Tanjung, Nurita Toruan-Mathius, and Tony Liwang, (2017), "Biological Network Analysis of Genes Involve in Embryogenesis of Oil Palm using ClueGO" in *ICBS Conference Proceedings, International Conference on Biological Science (2015)*, KnE Life Sciences, pages 265–274. DOI 10.18502/kls.v3i4.714

Page 265
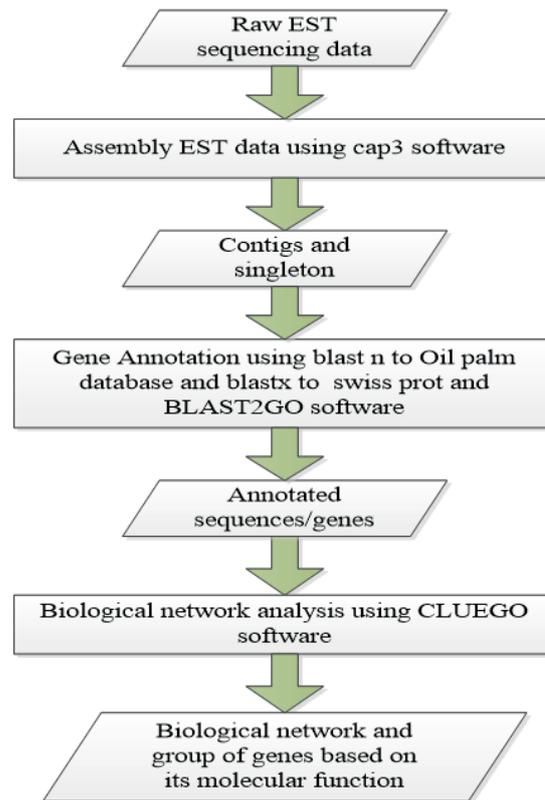
Figure 1: Flowchart of biological network analysis using ClueGO software.

information on molecular mechanisms and genes associated with embryogenesis processes in oil palm tissue culture remains limited.

Numerous approaches have been used to understand the complexity of gene interaction in embryogenesis of oil palm tissue culture, including expressed sequence tags (EST). Some of EST data sequences have been published in GenBank database [4, 5]. These data can be analyzed in advance to get more comprehensive information about biological networks of genes, which are relevant in embryogenesis using bioinformatics tools, such as ClueGO.

ClueGO is a plug-in Cytoscape software that strongly improves biological interpretations of genes. ClueGO integrates Gene Ontology (GO) terms as well as KEGG/BioCarta pathways and creates a functionally organized GO/pathway term networks [6]. As the ClueGO developer [6], has introduced and conducted an experimental trial of ClueGO for biological networks analysis in 2009 to analyze natural (NK) cell genes in human. It showed a good result in grouping either up-regulated or down-regulated genes into some biological networks. Based on this result, this tool was applied to identify biological networks of genes from EST sequencing data during embryogenesis of oil palm callus.
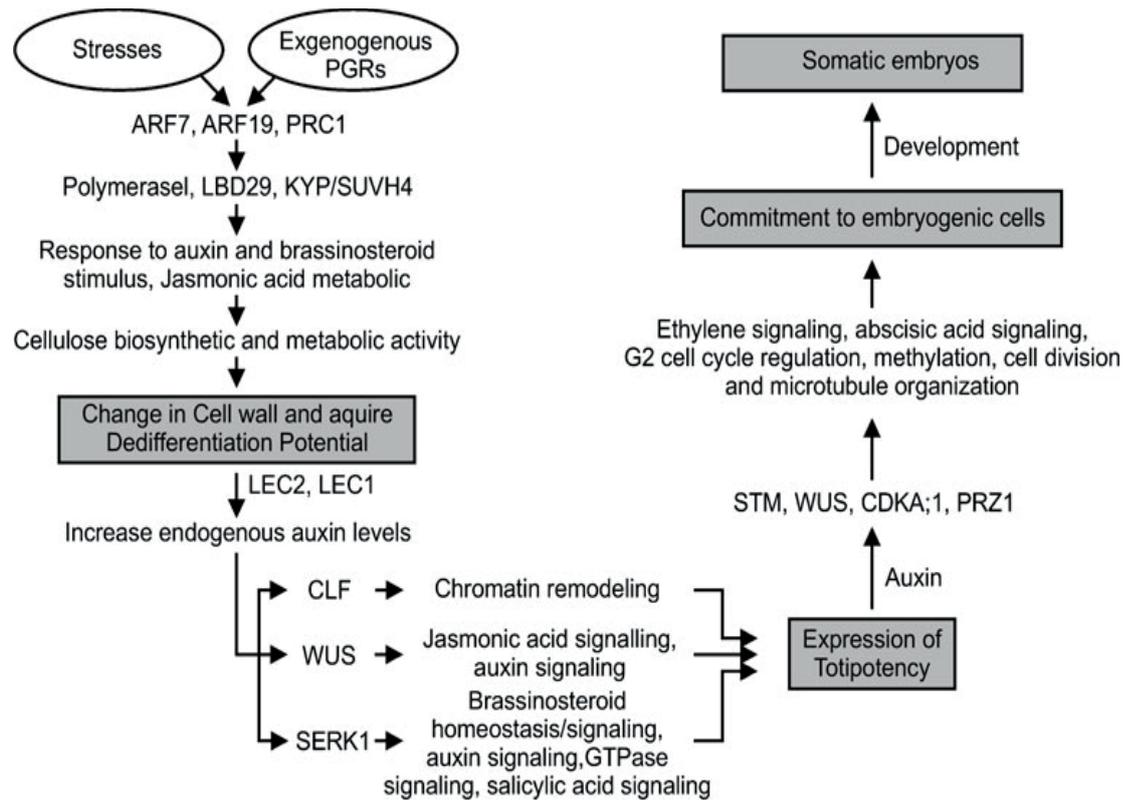
Figure 2: Hypothesized molecular and physiological pathways of somatic embryogenesis. Source: Elhiti et al. [10].

2. Materials and Methods

This study utilized two oil palm EST libraries from GenBank under accession numbers of EY396120-EY413718 and DW247764-DW248770 [4, 5]. There were two libraries under accession of EY396120-EY413718, namely embryogenic callus (EC) and non-embryogenic callus (NEC), and one library under accession numbers EY396120-EY413718, which was embryoid. The analysis comprised of sequence assembly using cap3 [7], sequence blasting (similarity analysis) using BLASTX software ver.2.2.25 [8], gene annotating using BLAST2GO software ver.2.8 [9], and biological networking analysis using ClueGO software ver.1.8 [6] (see Figure 1).

Hypothesized molecular and physiological pathways of plant somatic embryogenesis described by Elhiti *et al.* [10] were used as a reference library to identify all stages in embryogenesis that were affected by these genes product shown in Figure 2.

3. Result and Discussion

There were 53 and 108 biological network groups, which putatively involved in oil palm embryogenesis were identified using ClueGo analysis from EST EY396120-EY413718
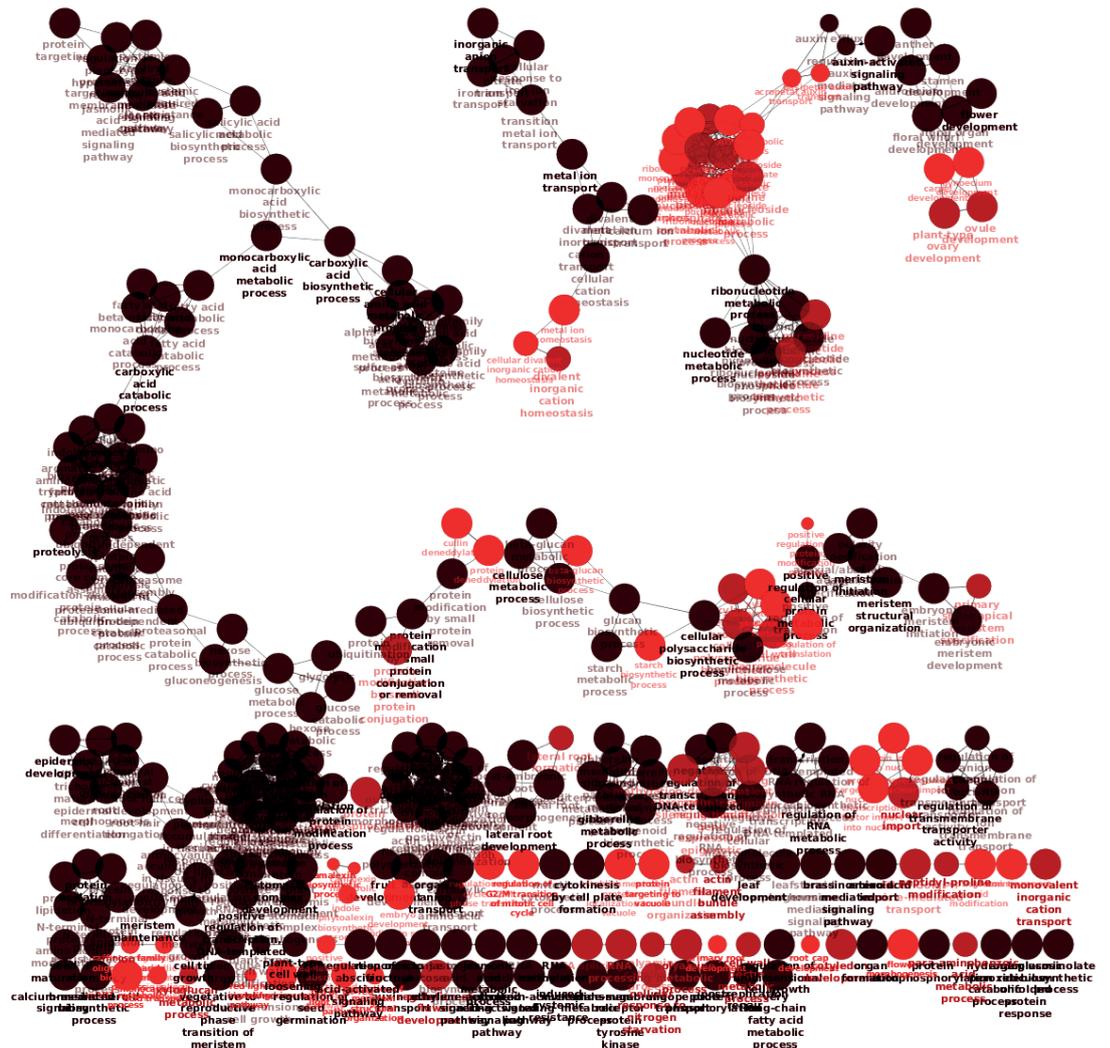
Figure 3: Biological network groups of genes derived from EST EY396120-EY413718.

and DW247764-DW248770, respectively (see Figure 3 and Figure 4). Biological network groups were illustrated as nodes, linked based on their kappa score levels (≥0.5). The node size (small to big) and color gradient (red to dark red) were represented more significance.

Specific genes networking overviews and their distributions in EC from EST EY396120-EY413718 and EST DW247764-DW248770 showed several groups that consisted of a large number of genes (see Figure 5 and Figure 6). These groups of genes were suspected to be up-regulated during embryogenesis process.

There were 31 groups of genes networking, which consistently occurred in the two EST datasets were identified from the data comparison analysis (Table 1). Several groups of genes were up-regulated in the two data sets. These groups of genes may have a function in regulating somatics embryogenesis, especially in embryonic induction. According to the literature in plant somatic embryogenesis, eleven functions in
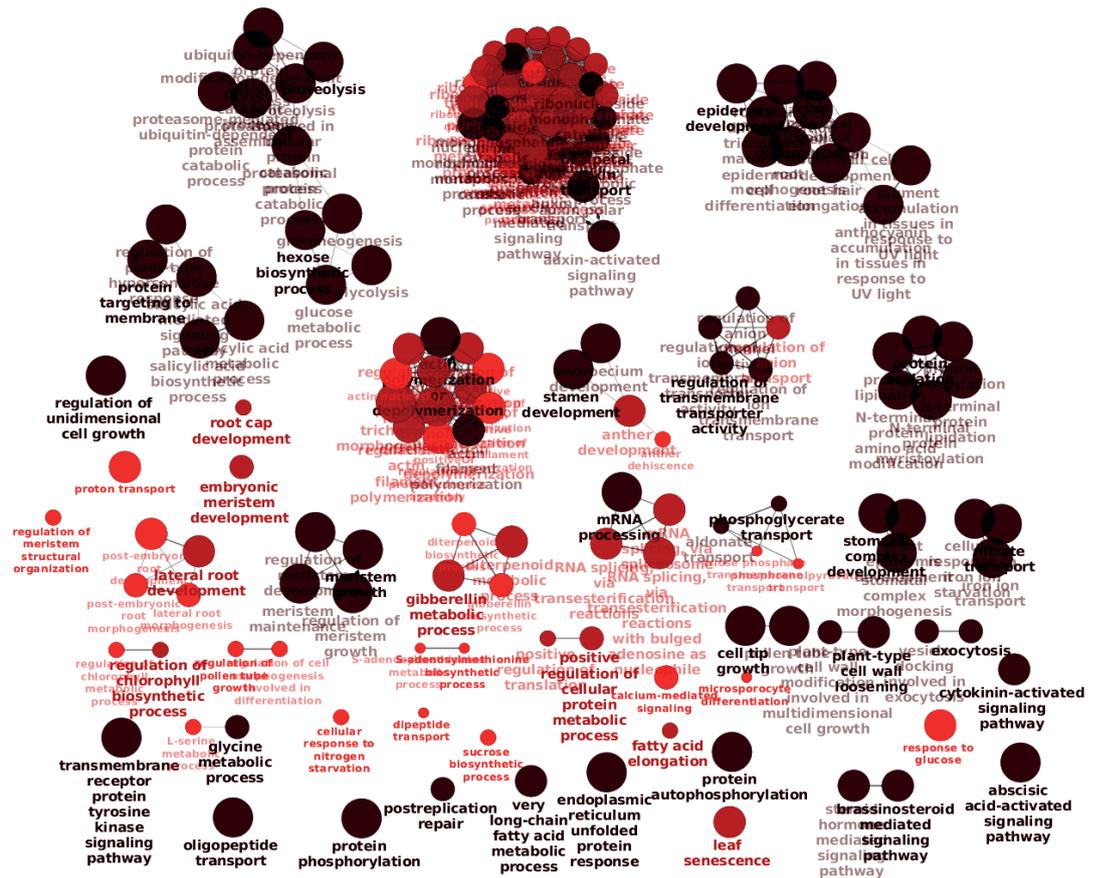
Figure 4: Biological network groups of genes derived from EST DW247764-DW24877.

somatics embryogenesis were affected by these groups of genes. These functions are control cell positions and its shape correctly during proliferations; promotions of cell elongations, cell divisions, and dedifferentiations throughout the plant cell life cycles; ethylene precursor; meristem formation and regulation; modification in cell wall composition (SE induction effect); ethylene synthesis; Gibberellin (GA) precursor; Jasmonic acid (JA) precursor; genetic information to control totipotency; and cell cycles [10–16].
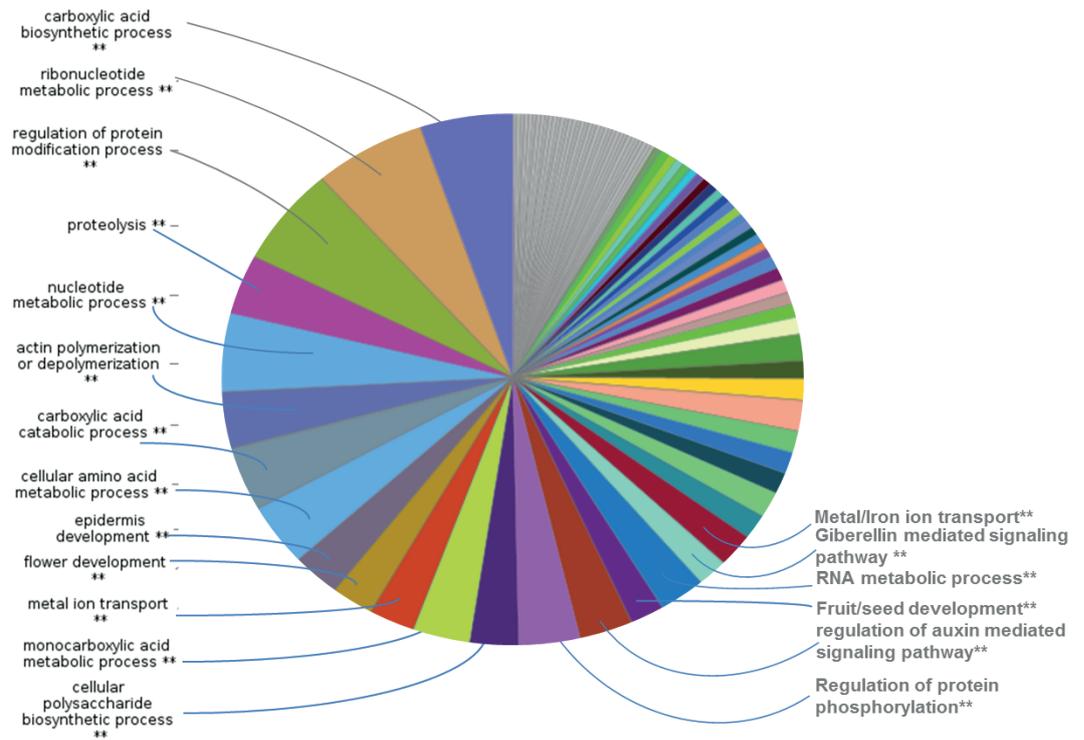
Figure 5: Genes networking overviews from EST EY396120-EY413718 data set.
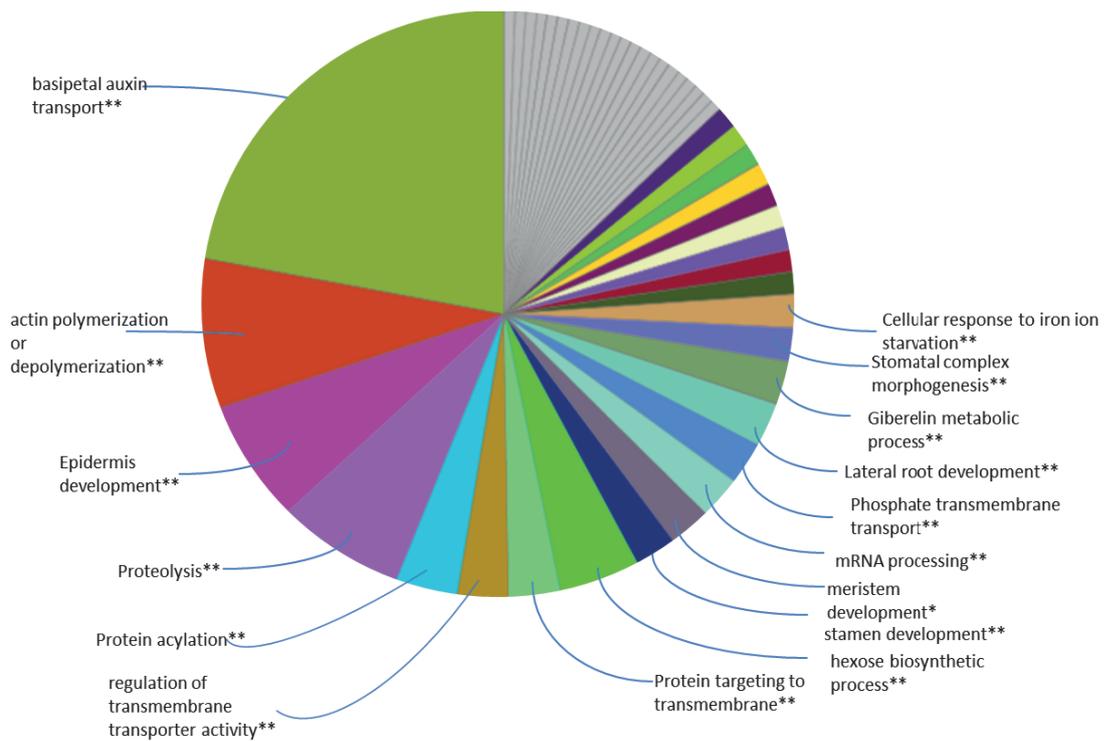
Figure 6: Genes networking overview from EST DW247764-DW248770 data set.

No	Groups of genes	Function in embryogenesis	Phase
1	actin filament bundle assembly [11, 12]	control cell position and its shape correctly during proliferation; needed in developmentalswitch from unpolarized to polarized cells in response to auxin	dedifferentiation
2	actin polymerization or depolymerization*) [11, 12]	control cell position and its shape correctly during proliferation(11); needed in developmentalswitch from unpolarized to polarized cells in response to auxin	dedifferentiation
3	brassinosteroid-mediated signaling pathway [10]	promotion of cell elongation, cell division, and dedifferentiation throughout the plant cell life cycle	dedifferentiation
4	carboxylic acid biosynthetic process* [13]	ethylene precursor (low level inhibited the formation of somatic embryos, whereas high levels elevated the frequency of somatic embryos induction)	dedifferentiation
5	carboxylic acid catabolic process* [13, 16]	ethylene precursor (low level inhibited the formation of somatic embryos, whereas high levels elevated the frequency of somatic embryos induction)	dedifferentiation
6	cell tip growth [10]	meristem formation and regulation	commitment
7	cellular polysaccharide biosynthetic process* [14]	modification in cell wall composition (SE induction effect)	commitment
8	cellulose metabolic process [14]	modification in cell wall composition (SE induction effect)	commitment
9	cytokinesis by cell plate formation [11, 12]	control cell position and its shape correctly during proliferation; needed in developmentalswitch from unpolarized to polarized cells in response to auxin	dedifferentiation
10	epidermis development*) [10]	meristem formation and regulation	commitment
11	fruit development* [15]	involved in ethylene synthesis	dedifferentiation
12	gibberellin metabolic process*) [10]	Gibberellin (GA) precursor (low level inhibited the formation of somatic embryos, whereas high levels elevated the frequency of somatic embryos induction)	dedifferentiation
13	jasmonic acid metabolic process [10, 16]	Jasmonic acid (JA) precursor (low level inhibited the formation of somatic embryos, whereas high levels elevated the frequency of somatic embryos induction), ethylene precursor	dedifferentiation
14	lateral root development) [10]	meristem formation and regulation	commitment
15	leaf development [10]	meristem formation and regulation	commitment
16	meristem structural organization [10]	meristem formation and regulation	commitment
17	monocarboxylic acid metabolic process [14]	modification in cell wall composition (SE induction effect)	commitment
18	negative regulation of transcription, DNA-templated [10]	genetic information required to control totipotency	expression of totipotency
19	plant-type cell wall loosening [10]	involved in cell cycle	commitment

No	Groups of genes	Function in embryogenesis	Phase
20	positive regulation of cellular protein metabolic process [10]	genetic information required to control totipotency	expression of totipotency
21	positive regulation of transcription, DNA-templated [10]	genetic information required to control totipotency	expression of totipotency
22	protein acylation⁾[10]	genetic information required to control totipotency	expression of totipotency
23	protein targeting to vacuole [10]	genetic information required to control totipotency	expression of totipotency
24	proteolysis*⁾[10]	genetic information required to control totipotency	totipotency
25	regulation of abscisic acid-activated signaling pathway [10, 13, 16]	ethylene precursor (low level inhibited the formation of somatic embryos, whereas high levels elevated the frequency of somatic embryos induction)	dedifferentiation
26	regulation of protein phosphorylation*⁾[10]	genetic information required to control totipotency	expression of totipotency
27	regulation of RNA metabolic process [10]	genetic information required to control totipotency	expression of totipotency
28	regulation of transmembrane transporter activity [10]	genetic information required to control totipotency	expression of totipotency
29	response to fructose [14]	modification in cell wall composition (SE induction effect)	commitment
30	seed maturation [10]	meristem formation and regulation	commitment
31	stomatal complex development [10]	meristem formation and regulation	commitment

*: up-regulated in EST EY396120-EY413718 data set; ⁾: up-regulated in EST DW247764-DW248770;

*⁾: up-regulated in the two data sets

TABLE 1: List of genes networking groups consistently occurred in the two EST datasets.

4. Conclusion

Thirty-one groups of genes that may play an important role in embryogenesis of oil palm *in vitro* culture were identified using ClueGO analysis. These groups affect in eleven functions of embryogenesis based on numerous previous publications in plant somatic embryogenesis studies. ClueGO is considered as the appropriate bioinformatic tool to conduct *in silico* analysis and identify candidate genes that may impact in a specific condition, including oil palm embryogenesis from large data set of genes, such as EST data set.

Acknowledgements

References

[1] M. B. Wahid, S. N. A. Abdullah, and I. E. Henson, "Oil palm - Achievements and potential," *Plant Production Science*, vol. 8, no. 3, pp. 288–297, 2005.

[2] J. Šamaj, M. Bobák, A. Blehová, and A. Pret'Ová, "Importance of cytoskeleton and cell wall in somatic embryogenesis," *Plant Cell Monographs*, vol. 2, pp. 35–50, 2005.

[3] PBH. Tinker and RHV. Corley, "The Oil Palm," in *The Oil Palm*, p. 284, Blackwell Science, Oxford, 4th edition, 2003.

[4] E.-T. L. Low, H. Alias, S.-H. Boon et al., "Oil palm (Elaeis guineensis Jacq.) tissue culture ESTs: Identifying genes associated with callogenesis and embryogenesis," *BMC Plant Biology*, vol. 8, article no. 62, 2008.

[5] P.-L. Chan, L.-S. Ma, E.-T. L. Low et al., "Normalized embryoid cDNA library of oil palm (Elaeis guineensis)," *Electronic Journal of Biotechnology*, vol. 13, no. 1, 2010.

[6] G. Bindea, B. Mlecnik, H. Hackl et al., "ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks," *Bioinformatics*, vol. 25, no. 8, pp. 1091–1093, 2009.

[7] X. Huang and A. Madan, "CAP3: A DNA sequence assembly program," *Genome Research*, vol. 9, no. 9, pp. 868–877, 1999.

[8] S. F. Altschul, W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman, "Basic Local Alignment Search Tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.

[9] A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles, "Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research," *Bioinformatics*, vol. 21, no. 18, pp. 3674–3676, 2005.

[10] M. Elhiti, C. Stasolla, and A. Wang, "Molecular regulation of plant somatic embryogenesis," *In Vitro Cellular and Developmental Biology - Plant*, vol. 49, no. 6, pp. 631–642, 2013.

[11] A. P. Smertenko and P. J. Hussey, "Immunolocalization of Proteins in Somatic Embryos," in *Plant Embryogenesis*, vol. 427 of *Methods in Molecular Biology*, pp. 157–171, Humana Press, Totowa, NJ, 2008.

[12] C. Holweg, C. Süßlin, and P. Nick, "Capturing in vivo dynamics of the actin cytoskeleton stimulated by auxin or light," *Plant and Cell Physiology*, vol. 45, no. 7, pp. 855–863, 2004.

[13] B. Bai, Y. H. Su, J. Yuan, and X. S. Zhang, "Induction of somatic embryos in arabidopsis requires local YUCCA expression mediated by the down-regulation of ethylene biosynthesis," *Molecular Plant*, vol. 6, no. 4, pp. 1247–1260, 2013.

[14] L. Portillo, A. Olmedilla, and F. Santacruz-Ruvalcaba, "Cellular and molecular changes associated with somatic embryogenesis induction in Agave tequilana," *Protoplasma*, vol. 249, no. 4, pp. 1101–1107, 2012.

[15] IE. Yates and CC. Reilly, "Somatic embryogenesis and plant development in eight cultivars of pecan," *HortScience*, vol. 25, no. 5, pp. 573–576, 1990.

[16] P. Piyatrakul, R.-A. Putranto, F. Martin et al., "Some ethylene biosynthesis and AP2/ERF genes reveal a specific pattern of expression during somatic embryogenesis in Hevea brasiliensis," *BMC Plant Biology*, vol. 12, article no. 244, 2012.