In silico screening of oil palm early and continuously flowering gene candidates for faster breeding program

A Polosoro\textsuperscript{1,2,*}, W Enggarini\textsuperscript{2}, T Hadiarto\textsuperscript{2}, E D J Supena\textsuperscript{1} and Suharsono\textsuperscript{1}

\textsuperscript{1}Department of Biology, Graduate Program of Plant Biology Study Program, Faculty of Mathematics and Natural Sciences, IPB University, Dramaga Campus, Bogor 16680, Indonesia.
\textsuperscript{2}Indonesian Center for Agricultural Biotechnology and Genetic Resource Research and Development (ICABIOGRD), Jl. Tentara Pelajar No. 3A, Cimanggu, Bogor, Indonesia.

*Email: aqwin@pertanian.go.id

\textbf{Abstract.} Oil palm (\textit{Elaeis guineensis} Jacq.) is a vegetable oil-producing crop used for consumption and biodiesel production. Fastrack breeding is one of breeding techniques to reduce plant vegetative phase significantly by a combination of conventional and transgene technologies. However, the lack information of flowering genes in oil palm become a gap for developing this technology. This paper aimed to discover the candidate genes from oil palm for developing early flowering oil palm lines. Screening potential florigen protein by comparing rice OsHd3a protein sequence in the NCBI database successfully discovered 13 florigen protein candidates with more than 53% similarity and up to 90% query cover. All protein candidates were analyzed by protein structure homology modeling (SWISS-MODEL) based on several parameters, including GMQE score and QMEAN Z-score. The modeling discovered that at least eight proteins predicted a similar structure to \textit{AtFTL1} and four proteins are related to \textit{AtTFL1}, which repress flowering process. Based on GMQE, QMEAN Z-score, and phylogenetic tree, we suggested three proteins (XP_010911427, XP_010925712, and XP_019705101) as the strong candidates for florigen protein. However, to verify their functions, all selected proteins should be inserted into the model plant and evaluate them as morphologically and physiologically.

\section*{1. Introduction}

Oil palm (\textit{Elaeis guineensis} Jacq.) is a vegetable oil-producing crop used for consumption and biodiesel production. As valuable food and bioenergy crop, African oil palm just has only one close relative from South and Central America, which is \textit{Elaeis oleifera}, both of them classified to the Arecaceae family [1,2]. Several methods have been used to differentiate oil palm varieties, including based on the thickness of the endocarp and the fruit color. For example, Dura, Pisifera, and Tenera are varieties that differentiated by endocarp thickness [3], while Nigrescens, Virescens, and Albescens are differentiated by fruit color [4].

The demand increment of palm oil products and its economic value causes the oil palm to become an important crop not only for a big corporation but also for the small farmers in Indonesia. The Ministry of Agriculture [5] has reported the increment of oil palm production from 17.54 million tons...
in 2008 to 23.52 million tons in 2012 or the growth rate around 7.7% in a year. Moreover, the Ministry of Finance [6] informed that 50% of Indonesian Crude Palm Oil production and 85% of Indonesian Crude Palm Kernel Oil (CPKO) were exported. In 2012, this sector earned 8.54 million dollars or 4.43% of total Indonesian export. Based on these achievements, Indonesia becomes the largest producer of palm oil worldwide. However, Indonesian palm oil production is far from optimal condition because of the low productivity rate. The productivity of the national oil palm company only reaches 3.69 tons CPO/ha [7]. Indonesia needs to enhance the seed quality to solve this problem.

Conventional breeding is the primary technique to generate oil palm with high productivity, environmental stress tolerance, and pest resistance. However, the long vegetative stage of oil palm (+/-5 years) is the main problem for the breeding process, because it takes a very long time. Fastrack breeding is the acceleration breeding technique to cut off the plant vegetative phase by using Early and Continuously Flowering (ECF) transgene technology. Fastrack breeding has successfully accelerated the apple (Malus fusca) breeding process by insertion and overexpression of the BpMADS4 gene from silver birch [8]. The discovery of the FLOWERING LOCUS T1 (FT1) gene has been also successfully applied in several breeding processes, such as orange [9], apple [10], and plum [11]. The transition product of this breeding process is categorized as a transgenic plant, but the final product will be selected by segregation to get a non-transgenic plant [12].

Flowering is a critical mechanism for the breeding process. However, the flowering mechanism is also restricted by the maturity process and environmental induction that cause some plants to have an obstacle in the breeding process [13]. Flowering induction mechanism in Arabidopsis is initiated by four main factors, such as photoperiod, vernalization, gibberellic acid (GA), and autonomous pathways [14]. All induction signals, except signal from GA pathways, will be transmitted to two central flowering regulators, CONSTANS (CO) and Flowering Locus C (FLC) [15]. CO accommodates photoperiod signals, while FLC assists vernalization and autonomous pathways. Those signals are used as "flowering integrator," such as SUPPRESSOR OF OVEREXPRESSION OF CO I (SOC1) and FLOWERING LOCUS T (FT), which act in the upstream flowering pathways. Those genes drive flower morphogenesis genes, such as APETALA1 (API) and LEAF (LFY) [15].

The FT and TERMINAL FLOWER 1 (TFL1) are grouped in phosphatidylethanolamine-binding protein (PEBP) genes, which are known as central to plant development. Those genes have essential functions especially for regulating the flowering time and plant architecture. They have similarities in protein sequence despite the antagonistic role in the flowering mechanism. FT induces flowering time while TFL represses it [11]. Induced by the transcription factor CO in leaves, FT protein is transferred from leaves to the shoot apical meristem (SAM) by phloem. In SAM, FT forms a protein complex with FD which induces flowering by inducing API and FUL (FRUITFULL) expression [12]. CO also regulates TFL1, which represses LFY and API expression level in inflorescence meristem (IM) and SAM resulting in flowering delays [13]. Moreover, several flowering studies showed pathways variation between species influenced by environmental conditions, but they remain similar main mechanisms [16,17]. Based on the taxonomy, the closest model plant relative to oil palm is rice, which has the most comprehensive molecular information in monocot [18]. These two factors became the pointer to use OsHd3a (FT homolog protein in rice) to discover oil palm florigen. Using Blastp tools in NCBI and specified it on the oil palm protein sequence, it is possible to screen the candidate gene by using query cover and percent identity.

SWISS-MODEL program is an integrated service Web-based protein modeling computing system [19]. The system is programmed to identify suitable templates based on an experimental protein structure database. Based on the target protein and template sequence alignment, simulation of a three-dimensional target protein model is generated from reference structure. Model quality evaluation tools, such as GMQE (Global Model Quality Estimation) and QMEAN (Qualitative Model Energy Analysis) are used to assess the reliability of the 3D model structure [20]. Generated by the modeled protein target-template alignment, GMQE scores are used to measure protein structural reliability, the highest score is the best model (0-1 scale) [21]. Meanwhile, QMEAN score utilizes mean force as statistical
potentials to produce global and per residue quality estimates [22]. Currently, homology modeling is the most accurate strategy to simulate reliable three-dimensional protein models.

This paper aimed to discover the candidate gene from oil palm for developing early flowering oil palm lines. Bioinformatic tools such as protein blast, multiple alignments, as well as homology protein modeling are potential tools to identify the ECF gene based on protein sequences of several flowering genes such as *TaFT*, *HvFT*, *MdFT1*, *AtFT1* and its homolog protein *AtTFL1*. This research will be followed by function gene validation in the model plant to prove the in-silico process.

2. Materials and Methods

2.1. Sequence Retrieval
National Center for Biotechnology Information (NCBI) database provides *OsHd3a* (XP_015641951) sequence for the primary source of the florigen protein sequence. Moreover, several proved florigen protein sequences were collected to cluster florigen candidate genes, such as *TaFT* (AAW23034), *HvFT* (AAZ38709), *MdFT1* (ACV92037), *AtFT1* (BA77838), and its antagonistic homologs protein *AtTFL1* (At5g03840).

2.2. Sequence Alignment and Phylogenetic Tree Construction
By used several florigens above, the clusterization of the oil palm candidate gene is possible to do by MegAlign (DNASTAR). Briefly, the analysis was done by using the Clustal W algorithm and default multiple alignment parameters such as gap penalty 10, gap length penalty 0.2, delay divergent seq 30%, and protein weight matrix of Gonnet Series [23].

2.3. SWISS-MODEL protein structure homology-modeling
SWISS-MODEL is a protein structure homology-modeling program fully automated. SWISS-MODEL is capable to simulate the 3D structure based on the reference of the PDB database [24]. Each model is evaluated by QMEAN to measure the quality of the model, not only on a global scale by Z-score but also by local quality estimate. Moreover, the GMQE score is also used for quality estimation by combined properties from the target–template alignment and the template search method. There are several main steps in the SWISS-MODEL program, firstly identification of structural template(s), alignment of a target sequence, and template structure(s), followed by model-building and model quality evaluation [21].

3. Result and Discussion

3.1 Protein blast screening
The Blastp of *OsHd3a* in NCBI oil palm transcriptomic database has successfully discovered 18 protein candidates (Table 1), but 5 of those just only has less than 60% query cover. In this case, 13 candidates of protein had more than 90% query cover used for further experiments. However, the percent identity of those protein sequences was varying start from 53% to 83%. Protein *HEADING DATE 3A isoform X2* (XP_010925712) was chosen as the highest score with a 96% query cover and 82.72% percent identity. We also discovered two genes that expressed more than one protein and formed isoform protein. LOC105048191 gene is alternatively spliced and translating into XP_010925712 and XP_019707288 proteins. LOC105040685 was also alternatively spliced into six isoform proteins [25].
### Table 1. The result of OsHd3a (XP_015641951) Blastp in NCBI oil palm protein database

| No | Accession Number | Description | Genes | Query Cover | Percent identity | E Value |
|----|------------------|-------------|-------|-------------|------------------|---------|
| 1  | XP_010925712     | Protein HEADING DATE 3A isoform X2 | LOC105048191 | 96% | 83.72% | 1e-108 |
| 2  | XP_019705101     | Protein HEADING DATE 3A isoform X2 | LOC105040685 | 97% | 79.31% | 2e-108 |
| 3  | XP_019705100     | Protein HEADING DATE 3A isoform X1 | LOC105040685 | 97% | 75.82% | 2e-105 |
| 4  | XP_010911427     | Protein HEADING DATE 3A-like | LOC105037463 | 96% | 78.03% | 4e-104 |
| 5  | XP_010930170     | Protein HEADING DATE 3A-like | LOC105051422 | 96% | 73.84% | 3e-95 |
| 6  | XP_010912811     | Protein FLOWERING LOCUS T-like | LOC105038642 | 95% | 73.10% | 1e-91 |
| 7  | XP_010912140     | Protein HEADING DATE 3A | LOC105038131 | 95% | 67.84% | 3e-87 |
| 8  | XP_010915452     | Protein HEADING DATE 3A-like | LOC105040569 | 93% | 66.67% | 2e-84 |
| 9  | XP_010919262     | CEN-like protein | LOC105043421 | 93% | 63.10% | 1e-75 |
| 10 | XP_010907126     | Protein SELF-PRUNING-like | LOC105033865 | 95% | 62.21% | 1e-73 |
| 11 | XP_010912272     | Protein SELF-PRUNING | LOC105038236 | 95% | 61.63% | 1e-70 |
| 12 | XP_010940015     | Protein MOTHER of FT and TFL1 homolog 1 | LOC105058700 | 93% | 54.39% | 5e-62 |
| 13 | XP_010936814     | Protein MOTHER of FT and TFL1 homolog 1 | LOC105056345 | 93% | 53.80% | 3e-59 |
| 14 | ARO70141         | Heading date 3a isoform X5 | LOC105040685 | 55% | 74.00% | 3e-53 |
| 15 | XP_029119083     | Protein HEADING DATE 3A isoform X3 | LOC105040685 | 55% | 73.00% | 1e-52 |
| 16 | XP_019707288     | Protein HEADING DATE 3A isoform X1 | LOC105048191 | 57% | 75.73% | 3e-51 |
| 17 | ARO70139         | Heading date 3a isoform X3 | LOC105040685 | 55% | 67.59% | 1e-49 |
| 18 | ARO70144         | heading date 3a isoform X8 | LOC105040685 | 36% | 72.31% | 3e-28 |

### 3.2 Clustal W and phylogenetic tree

The progressive alignment algorithm was obtained using Clustal W. It was successfully performed multiple sequence alignment and processed the result into a phylogenetic tree (fig. 1). All florigen genes from the Poaceae family were grouped into one area. There were several candidate florigen proteins which relatively had similar sequence to a group of florigen proteins, such as XP_010911427, XP_010925712, XP_019705101, and XP_019705100. On the other side, the antagonist protein to florigen (AtTFL1) also had close relatives, such as XP_010907126 and XP_010912272.

![Figure 1. Phylogenetic tree of florigen protein (AtFT1, OsHd3a, PtFT, TaFT, HvFT, and MdFT1) compared to the five of florigen protein candidates from oil palm.](image-url)
3.3 Protein modeling analysis

SWISS-MODEL protein structure homology-modeling program showed that eight proteins had a similar structure with AtFT1 protein. However, they had a different quality level based on GMQE and QMEAN scores, these significant differences were believed to affect the protein structure and the protein function. Moreover, we also found five proteins that had a similar structure to AtTFL1, even at different quality levels. All the structure had been chosen by several indicators such as high resolution (<2.2Å) and coverage level. Then the results of the analysis were sorted according to the quality level of the structure (table 2).

Table 2. Protein modeling parameters of florigen protein candidates were showed by SWISS-MODEL program

| Protein ID     | GMQE | QMEAN | Seq Identity | Seq Similarity | Coverage | Found by Method | Resolution | Description (Template) |
|----------------|------|-------|--------------|----------------|----------|-----------------|------------|------------------------|
| XP_010911427   | 0.82 | -0.58 | 73.46        | 0.54           | 0.92     | BLAST X-ray     | 1.33 Å     | FLOWERING LOCUS T (6igi.1.A) |
| XP_010925712   | 0.80 | -0.33 | 73.78        | 0.54           | 0.93     | HHblits X-ray   | 1.01 Å     | FLOWERING LOCUS T (6igh.1.A) |
| XP_019705101   | 0.80 | -0.47 | 73.01        | 0.54           | 0.89     | BLAST X-ray     | 1.50 Å     | FLOWERING LOCUS T (6igi.1.A) |
| XP_010930170   | 0.79 | -0.41 | 69.33        | 0.52           | 0.91     | BLAST X-ray     | 1.01 Å     | FLOWERING LOCUS T (6igh.1.A) |
| XP_010912140   | 0.81 | -1.12 | 66.27        | 0.51           | 0.96     | BLAST X-ray     | 1.01 Å     | FLOWERING LOCUS T (6igh.1.A) |
| XP_010915452   | 0.78 | -0.81 | 64.20        | 0.51           | 0.94     | HHblits X-ray   | 1.01 Å     | FLOWERING LOCUS T (6igh.1.A) |
| XP_010912811   | 0.77 | -0.96 | 70.37        | 0.52           | 0.91     | HHblits X-ray   | 1.33 Å     | FLOWERING LOCUS T (6igh.1.A) |
| XP_019705100   | 0.75 | -2.80 | 73.01        | 0.54           | 0.85     | BLAST X-ray     | 1.33 Å     | FLOWERING LOCUS T (6igh.1.A) |
| XP_010907126   | 0.81 | -0.02 | 74.12        | 0.54           | 0.98     | HHblits X-ray   | 1.80 Å     | TERMINAL FLOWER 1 (1wko.2.A) |
| XP_010912272   | 0.81 | -0.14 | 72.51        | 0.53           | 0.99     | BLAST X-ray     | 1.80 Å     | TERMINAL FLOWER 1 (1wko.2.A) |
| XP_010919262   | 0.80 | -0.67 | 64.33        | 0.50           | 0.99     | BLAST X-ray     | 1.80 Å     | TERMINAL FLOWER 1 (1wko.2.A) |
| XP_010940015   | 0.78 | -0.87 | 56.02        | 0.47           | 0.95     | BLAST X-ray     | 1.80 Å     | TERMINAL FLOWER 1 (1wko.2.A) |
| XP_010936814   | 0.78 | -0.91 | 55.69        | 0.47           | 0.95     | BLAST X-ray     | 1.80 Å     | TERMINAL FLOWER 1 (1wko.2.A) |

XP_010911427 protein (figure 2a) had the highest GMQE score. It also had a good quality structure globally by QMEAN Z-score (|Z-score|>1) and by local quality estimate. At the same time, XP_019705100 (figure 2c) had the lowest GMQE score and showed a low-quality structure by QMEAN Z-score (|Z-Score|>2). The low-quality region was indicated by red mark in the 3D protein structure and less than 0.6 scores in the local quality estimate. On the other hand, XP_010907126 protein (figure 2d) and four other proteins were predicted as different proteins, and all parameters show all those 3D protein structures were closed to AtTFL protein.
Figure 2. Picture of protein structure, the graph of comparison with a non-redundant set of PDB structure, and the local quality estimate of (a). XP_010911427, (b). XP_019705101, (c). XP_019705100, and (d) XP_010907126 proteins. Blue color indicated good quality and red color indicated bad quality scores for the specific feature in the 3D protein structure.

3.4 Protein sequence comparative studies

As a monocotyledon, oil palm by the phylogenetic way is a close relative of rice compared to other model plants. Even though Arabidopsis flowering pathways and mechanism is well mapped compared to rice, both of the model plants are still sharing the main pathways and mechanism [26]. For example, AtFT has a similar role to OsHd3a and OsFLT gene as florigen or AtCO and OsHd1 as a circadian-clock mediator. The similarity is not only sharing the same pathways but also sharing similar protein sequences [27]. In this paper, we used OsHd3a as a preliminary screening by Blastp in NCBI protein database because of rice phylogenetically close to oil palm. However, in the modeling by SWISS-MODEL software, OsHd3a reference structure did not meet the minimum required resolution.
Therefore, ArFT was used as a template protein for protein modeling because of its availability of high-resolution protein structure. High-resolution template protein was needed to get a more detailed and precision model [20,28,29].

Information on a three-dimensional protein structure is essential to study and understand its function and activity. The 3-dimensional protein structure of an unknown protein is possible to be constructed on experimentally determined structures of close relatives protein as templates [20]. However, the similarity of protein sequence level does not guarantee protein behave similar structure and function. In the case of isoform protein from LOC105040685 gene, the addition of 8 amino acids in the 75th aa sequence of XP_019705100 compared XP_019705101 gave many changes in the protein structure. GMQE indicated the changes and QMEAN score in SWISS-MODEL analysis, which completely dropped it to the lowest level in both parameters. It was also physically observable in the local quality estimate and 3D model, which marked by red color. These parameters were used to analyze protein profoundly that impossible to Clustal and phylogenetic analysis do. However, several studies has successfully indicated the similarity of the florigen protein sequence based on family linked that was also important for screening the protein candidates [30–32].

3.5 Protein function based on the PDB structure database

Based on the SWISS-MODEL modeling and its scoring parameters, we successfully identified two antagonistic proteins from all candidates. Flowering Locus T and Terminal Flower were observed as "twins" proteins with having an antagonist role [33–36]. However, the mutation of at least four different residues of FT surface charge successfully changed FT protein to mimics FTL1 role because two proteins are sharing a lot of structural similarities [37]. In this study, Clustal W algorithm and phylogenetic tree successfully split between FT and TFL with the same result to SWISS-MODEL tools. However, SWISS-MODEL analysis could be more selective for insertion and deletion cases than phylogenetic analysis.

Build upon the GMQE and QMEAN parameters also combined with a phylogenetic tree, at least three strong protein candidates were mathematically proved as strong candidates for florigen protein in oil palm. XP_010911427, XP_010925712, and XP_019705101 were the top three with the highest GMQE and QMEAN score. The protein sequences also closed enough to monocot florigen proteins based on phylogenetic analysis.

4. Conclusion

The modeling successfully selected at least eight proteins from thirteen protein candidates. However, by considering to quality of protein modeling structure, protein sequence phylogenetic analysis, and the closeness of taxon, XP_010911427, XP_010925712, and XP_019705101 were chosen as the strong candidate for florigen protein. Moreover, we also found five protein structures resembled Terminal Flowering protein, which was an FT homolog and ruled antagonistic as a flowering repressor in plant. All protein candidates should be verified by overexpression those genes into the model plant and evaluating the changes of flowering time.

References
[1] Nambiappan B, Iail A, Hashim N, Ismail N, Shahari D N, Idris N A N, Omar N, Salleh K M, Hassan N A M and Kushairi A 2018 J. Oil Palm Res. 30 13–25
[2] Verhuy R W 2009 Land Use, Land Cover and Soil Sciences vol I
[3] Odonore V D, Eke C R, Asemota O and Shittu H O 2015 Eur. Int. J. Sci. Technol. 4 155–60
[4] Ying S T, Zaman F Q, Ling H O C, Ithnin M and Rao V 2007 J. Oil Palm Res. 19 381–92
[5] Ministry of Agriculture of Indonesia 2012 Statistik Pertanian (Agricultural Statistics) (Jakarta)
[6] Ministry of Finance of Indonesia 2013 Nota Keuangan dan Rancangan Anggaran dan Pendapatan Negara RI (Financial Notes and Budget Draft and State Revenue of the Republic of Indonesia) (Jakarta)
[7] Purba J H V 2019 IOP Conf. Ser. Earth Environ. Sci. 336
[8] Flachowsky H, Peil A, Sopanen T, Elo A and Hanke V 2007 Plant Breed. 126 137–45
[9] Matsuda N, Ikeda K, Kurosaka M, Takashina T, Isuzugawa K, Endo T and Omura M 2009 J. Japanese Soc. Hortic. Sci. 78 410–6
[10] Wenzel S, Flachowsky H and Hanke M-V V 2013 Plant Cell. Tissue Organ Cult. 115 127–37
[11] Srinivasan C, Dardick C, Callahan A and Scorza R 2012 PLoS One 7 1–11
[12] van Nocker S and Gardiner S E 2014 Hortic. Res. 1 14022
[13] McGarry R C, Klocko A L, Pang M, Strauss S H and Ayre B G 2017 Plant Physiol. 173 47–55
[14] Mouradov A, Cremer F and Coupland G 2002 Plant Cell 14 Suppl S111–30
[15] Fornara F, de Montaigu A and Coupland G 2010 Cell 141 3–5
[16] Izawa T, Takahashi Y and Yano M 2003 Curr. Opin. Plant Biol. 6 113–20
[17] Shrestha R, Gómez-Ariza J, Brambilla V and Fornara F 2014 Ann. Bot. 114 1445–58
[18] Sakata K, Antonio B A, Mukai Y, Nagasaki H, Sakai Y, Makino K and Sasaki T 2000 Nucleic Acids Res. 28 97–101
[19] Arnold K, Bordoli L, Kopp J and Schwede T 2006 Bioinformatics 22 195–201
[20] Bordoli L, Kiefner F, Arnold K, Benkert P, Battey J and Schwede T 2009 Nat. Protoc. 4 1–13
[21] Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefner F, Gallo Cassarino T, Bertoni M, Bordoli L and Schwede T 2014 Nucleic Acids Res. 42 W252–8
[22] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer F T, de Beer T A P, Rempfer C, Bordoli L, Lepore R and Schwede T 2018 Nucleic Acids Res. 46 W296–303
[23] Ravn L S, Aizawa Y, Pollevick G D, Hofman-Bang J, Cordeiro J M, Dixen U, Jensen G, Wu Y, Burashnikov E, Haunso S, Guerichoff A, Hu D, Svendsen J H, Christiansen M and Antzelevitch C 2008 Hear. Rhythm 5 427–35
[24] Studer G, Rempfer C, Waterhouse A M, Gumienny R, Haas J and Schwede T 2020 Bioinformatics 36 1765–71
[25] Whankaew S, RuttaJorn K, Madsen C K, Asp T, Xu L, NakkaeW and Phongdara A 2019 J. Sci. Technol. 41 332–40
[26] Takahashi Y and Shimamoto K 2011 Genes Genet. Syst. 86 175–82
[27] Guo X, Wang Y, Wang Q, Xu Z and Lin X 2015 J. Plant Biochem. Biotechnol. 25 168–178
[28] Fiser A and Šali A 2003 Methods Enzymol. 374 461–91
[29] Kapupara P P, Dholakia S P, Patel V P and Suhagia B N 2011 J. Chem. Pharm. Res. 3 287–94
[30] Zhang J, Song Q, Cregan P B, Nelson R L, Wang X, Wu J and Jiang G L 2015 BMC Genomics 16 1
[31] Xiang L, Li X, Qin D, Guo F, Wu C, Miao L and Sun C 2012 Plant Physiol. Biochem. 58 98–105
[32] Tränkner C, Lehmann S, Hoenicka H, Hanke M V, Fladung M, Lenhardt D, Dunemann F, Gau A, Schlangen K, Malnoy M and Flachowsky H 2010 Planta 232 1309–24
[33] Wickland D P and Hanzawa Y 2015 Mol. Plant 8 983–97
[34] Kotake T, Takada S, Nakahigashi K, Ohto M and Goto K 2003 Plant Cell Physiol. 44 555–64
[35] Flachowsky H, Peil A, Hanke M-V, Tränkner C, Szankowski I and Lehmann S 2012 Acta Hortic. 929 351–356
[36] Pin P A, Benlloch R, Bonnet D, Wremerth-Weich E, Kraft T, Gielen J J L and Nilsson O 2010 Science (80-. ). 330 1397–400
[37] Ho W W H and Weigel D 2014 Plant Cell 26 552–64