

## ABSTRAK

Pemetaan *Quantitative Trait Locus* (QTL) merupakan metode yang paling efisien dalam menentukan hubungan antara fenotip yang kompleks dengan kelimpahan genotip pada genom. Genom organisme yang telah dipetakan dengan baik dan lengkap adalah *Arabidopsis thaliana*. Hingga saat ini hampir 25000 urutan kode genetik yang telah teridentifikasi di seluruh genomnya, namun sekitar 9% yang masih diketahui memiliki asosiasi dengan fungsi fisiologisnya. Untuk mendapatkan hasil analisis QTL yang tepat, dibutuhkan set populasi *Recombinant Inbred Lines* (RILs). Namun, dalam pembuatannya dibutuhkan waktu yang relatif lama dan biaya yang besar. Oleh sebab itu, penelitian ini dilakukan secara *in silico* dengan memanfaatkan publikasi data RILs *Arabidopsis thaliana ecotype Bay-0xSha* yang tersedia pada *Web Service VNAT*. Penelitian ini bertujuan memetakan QTL untuk gen yang berperan dalam pengaturan sifat waktu pembungaan pada *Arabidopsis thaliana* dengan program QTL *Cartographer* dan *WebQTL*. Data yang diperoleh dari *Web Service VNAT* berupa data persilangan 420 turunan RILs dengan 38 penanda mikrosatelit, peta genetik, dan *trait values* untuk tiap fenotip yang telah dipublikasikan. *Trait values* yang digunakan adalah waktu pembungaan pada hari panjang (16 jam) dan hari pendek (8 jam). Data tersebut disimulasi menggunakan program QTL *Cartographer*. Sedangkan program *WebQTL* melakukan simulasi dengan *database*-nya sendiri. Metode analisis QTL yang digunakan adalah *Interval Mapping* dan *Composite Interval Mapping*. Metode ini digunakan untuk menduga posisi QTL dalam kromosom pada sifat waktu pembungaan. Hubungan tersebut dideteksi dengan skor LOD dari grafik pemetaan QTL yang dihasilkan kedua metode. Ambang batas skor LOD untuk metode IM sebesar 2,2 dan CIM sebesar 2,3 untuk program QTL *Cartographer*. Sedangkan ambang batas skor LOD pada *WebQTL* untuk metode IM sebesar 2,31-2,33 dan CIM sebesar 2,19-2,34. Lokus mikrosatelit yang berasosiasi dengan sifat waktu pembungaan diambil dari skor LOD tertinggi di atas ambang batas. Konsensus lokus mikrosatelit yang menentukan posisi gen yang berperan dalam pembentukan sifat kuantitatif waktu pembungaan pada kedua program dengan metode IM adalah NGA249 (kromosom 5) untuk hari panjang serta MSAT4.8 (kromosom 4) yang mengarah ke parental *Shahdara* dan NGA249 (kromosom 5) untuk hari pendek yang mengarah ke parental *Bayreuth*. Untuk metode CIM adalah NGA249 (kromosom 5) pada hari panjang dan pendek. Identifikasi kandidat gen pembungaan *Arabidopsis* dilakukan menggunakan *database Sequence Viewer TAIR* dan *Map Viewer NCBI*. Dari hasil *Sequence Viewer*, kandidat gen yang diduga terdapat pada lokus AT5G08550. Selanjutnya digunakan *Map Viewer* untuk mengidentifikasi lebih spesifik. Kandidat gen yang diidentifikasi untuk lokus NGA249 adalah ILP1 (*Increased Level of Poliploidy1-1D*). Gen ini berperan dalam endoreduplikasi pada *Arabidopsis thaliana*. Hanya lokus NGA249 yang ada pada *database TAIR* dan *NCBI*. Oleh sebab itu, dilakukan *Primer BLAST* untuk *primer MSAT4.8*. Dari hasil *Primer BLAST*, kandidat gen yang diidentifikasi untuk lokus MSAT4.8 adalah MEE47 (*Maternal Effect Embryo Arrest 47*). Gen ini berperan sebagai faktor transkripsi pada *Arabidopsis thaliana*.

Kata kunci : pemetaan QTL, populasi RILs, *Arabidopsis*, penanda mikrosatelit, kandidat gen pembungaan

## ABSTRACT

QTL mapping is the most efficient method in determining the relationship between the complex phenotype and genotype abundance in the genome. An organism's genome that have been correctly and completely mapped is *Arabidopsis thaliana*. Until now, almost 25,000 coding sequences have been identified in the whole genome, but only 9% have known association with their physiological function. To get the results of QTL analysis appropriately, a population set of Recombinant Inbred Lines (RILs) is required. However, making this set will take a relatively long time and have a high cost. Therefore, this research would be conducted *in silico* by using publications of RILs *Arabidopsis thaliana* ecotype Bay-0xSha data that are available on the VNAT Web Service. This study aims to identify candidate flowering genes of *Arabidopsis* through QTL analysis of RILs flowering time with the program QTL Cartographer and WebQTL. Data from VNAT Web Service that consist of 420 derived RILs crossing data with 38 microsatellite markers, the genetic map, and the trait values for each phenotype that have been published. Trait values used in the flowering time is for long day (16 hours) and short day (8 hours). The data was simulated using the program QTL Cartographer. Meanwhile, the simulation program WebQTL uses its own database. QTL analysis method used was the Interval Mapping and Composite Interval Mapping. This method is used to infer the position of QTL in the chromosome in the nature of flowering time. Relationship is detected with LOD scores from QTL mapping graphs generated by both methods. LOD threshold for the IM method is 2.2 and 2.3 for the CIM program of QTL Cartographer. While the threshold LOD score on WebQTL method for IM and CIM is 2,31-2,33 of 2,19-2,34. Microsatellite locus associated with flowering time is taken from the highest LOD score above the threshold limit. Consensus microsatellite locus that determine the position of a role in the establishment of quantitative nature of flowering time in both programs with the IM method is NGA249 (chromosome 5) for long day and the nature MSAT4.8 (chromosome 4) that leads to parental *Shahdara* and NGA249 (chromosome 5) that leads to parental *Bayreuth* for short day. For CIM, NGA249 (chromosome 5) was detected for long day and short day. Identification of candidate flowering genes is done using the *Arabidopsis* database Sequence Viewer TAIR and Map Viewer NCBI. From the results of the Sequence Viewer, the candidate genes were suspected to be on the locus AT5G08550. Next Map Viewer is used to identify more specific. Candidate genes identified for locus NGA249 is ILP1 (Increased Level of Poliploidy1-1D). This gene play a role in endoreduplication process for *Arabidopsis thaliana*. Only locus NGA249 was present in the NCBI and TAIR databases. Therefore, Primer BLAST for MSAT4.8 was used. From the Primer BLAST results, candidate genes identified for locus MSAT4.8 is MEE47 (Maternal Effects Embryo Arrest 47). That gene play a role as transcription factor in *Arabidopsis thaliana*.

Keywords: QTL mapping, RILs population, *Arabidopsis*, microsatellite markers, candidate flowering genes