

Gene discovery in the developing xylem tissue of a tropical timber tree species: *Neolamarckia cadamba* (Roxb.) Bosser (*kelampayan*)

Shek Ling Pang¹ · Wei Seng Ho² · M. N. Mat-Isa³ · Julaihi Abdullah¹Received: 8 January 2014 / Revised: 27 March 2015 / Accepted: 15 April 2015 / Published online: 30 April 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract A complementary DNA (cDNA) library was constructed from the developing xylem tissues of *Neolamarckia cadamba*. A total of 10,368 single-pass sequences was generated through high-throughput 5'-expressed sequence tag (EST) sequencing of the cDNA clones, and 6622 high-quality ESTs were obtained after removing the low-quality sequences; this gave approximately 3.17 Mb of data. Clustering of the high-quality ESTs revealed 4728 unigenes, consisting of 2100 consensus and 2628 singletons. A total of 2405 ESTs were successfully annotated with 7753 gene ontology (GO) terms that distributed among three main GO categories, which were biological processes (2333), molecular function (3056) and cellular component (2364). Simple sequence repeat (SSR) mining revealed that the frequency of SSR in the *N. cadamba* EST database (NcbdEST) was 3.3 %, with the GCT/AGC motif being the most abundant

repeat motif. The most abundant transcript with known function found in this database was 60S ribosomal protein followed by 40S ribosomal protein. Some of the important genes involved in xylogenesis and lignin biosynthesis were found in NcbdEST; these include tubulin genes, *cellulose synthase* (*CesA*), *xyloglucan endotransglycosylase* (*XET*), arabinogalactan, *cinnamate 4-hydroxylase* (*C4H*), *caffeoyl-coenzyme A O-methyltransferase* (*CCoAOMT*) and peroxidase. The data obtained from this study will provide a powerful means for identifying mechanisms controlling wood formation pathways of *kelampayan* and supply many new cloned genes for future endeavours to modify wood and fibre properties.

Keywords Expressed sequence tags (ESTs) · Forest plantation · Lignin biosynthesis · *Neolamarckia cadamba* · Wood formation

Communicated by W. Ratnam

This article is part of the Topical Collection on *Genome Biology*

Electronic supplementary material The online version of this article (doi:10.1007/s11295-015-0873-y) contains supplementary material, which is available to authorized users.

✉ Shek Ling Pang
shekling@gmail.com; slpang@sarawakforestry.com

¹ Applied Forest Science and Industry Development, Tree Breeding Unit, SARAWAK FORESTRY, 93250 Kuching, Sarawak, Malaysia

² Forest Genomics and Informatics Laboratory, Department of Molecular Biology, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

³ Malaysia Genome Institute, 43600 Bangi, Selangor Darul Ehsan, Malaysia

Introduction

In Malaysia, the State Government of Sarawak has set a target for 1 million ha of forest land degraded by shifting cultivation to be planted with fast-growing timber tree species by year 2020 in order to meet the increasing demand for wood and wood fibre. To meet this target, an estimation of 42 million quality seedlings is needed for annual planting or reforestation. *Neolamarckia cadamba* (Roxb.) Bosser, locally known as *kelampayan*, has been identified as one of the potential indigenous fast-growing timber tree species for forest plantation in Sarawak (Tchin et al. 2012; Lai et al. 2013; Tiong et al. 2014). It is a large, deciduous and fast-growing tree that gives early economic returns within 8–10 years. Under normal conditions, it attains a height of 17.67 m and diameter of 25.3 cm at breast height within 9 years. *Kelampayan* not only serves as