

## OVERVIEW OF COFFEE GENOME (*Coffea canephora* L.) AND ITS FUNCTION IN STRESS RESPONSE AND CAFFEINE BIOSYNTHESIS

Nguyen Dinh Sy<sup>1</sup>, Nguyen Van Tinh<sup>2</sup>, Tran Van Cuong<sup>3</sup>, Nguyen Ngoc Huu<sup>3</sup>

Received Date: 22/8/2022; Revised Date: 22/9/2022; Accepted for Publication: 23/9/2022

### SUMMARY

*Coffea canephora*, a coffee species belonging to the Rubiaceae family, is one of the most popular cultivated coffee worldwide. In this study, we overview some important aspects of the genetic genome in *C. canephora* and also review the candidate genes that involve biotic and abiotic stress response and caffeine biosynthesis in the *Coffea* genome. It is reported that the *C. canephora* genome consists of 25,574 protein-coding genes, 2,573 organellar-to-nuclear genome transfers, 6,812 predicted transposable elements (TEs), and 92 microRNA precursors. Coffee-caffeine is synthesized from xanthosine via three methylation steps by NMT genes in *C. canephora* genome: NMT2: cc02\_g09350; DXMT: cc01\_g00720; XMT: cc09\_g06970; MXMT: cc00\_g24720; NMT3: cc09\_g06960; MTL: cc09\_g06950. *C. canephora* genome sequence has recently become a new tool for investigating and analyzing coffee resistance and quality at the molecular level.

**Keywords:** Caffeine, *C. canephora*, chromosome, miRNA, stress, transposable elements.

### 1. INTRODUCTION

There are more than 100 species under the *Coffea*, but most coffee bean productions are *C. canephora* and *C. arabica*, with 30% and 70% in total production worldwide (ICO, 2021). Among them, *C. arabica* is preferred to *C. canephora* due to its low-caffeine content and less-bitter taste. It is estimated that an average of 2.25 billion cups of coffee are consumed worldwide, bringing approximately 12 billion US dollars in income each year (ICO, 2021). Therefore, the *Coffea* sp. genome is chosen to sequence, which aids the scientists in deciphering the molecular mechanisms of metabolites biosynthesis pathway and stress response of coffee plant.

There are several published articles about genome sequencing, the coffee transcriptomes, and expressed sequence tags (ESTs) from both robusta and arabica coffee plants (Pallavicini et al., 2004; Vieira et al., 2006). An oligo-based microarray containing 15,721 coffee genes was constructed to analyze potential genes involved in bean maturation, pathogen resistance, and stress responses (Privat et al., 2011). EST sequences of *C. arabica* were recently released (Mishra and Slater 2012).

*C. canephora* is the diploid species ( $2n = 2x = 22$  chromosomes) and is self-incompatible. *C. canephora* reference genome sequence was completed and analyzed in collaboration with Genoscope, IRD, and Cirad (UMRs AGAP,

DIADÉ, and RPB) funded by ANR. Whole genome sequencing strategy for coffee genome sequencing by method following: Illumina GAIx and Roche/454 GSFLX next generation sequencing platforms were used to generate data. They extracted DNA from coffee leaves. Also, they extracted RNA for transcriptome sequencing. RNA sequences read by Illumina GAI. BWA algorithm technology was used for alignment of RNA reads to produce contigs. *C. canephora* genome was assembled using Solexa/Illumina platform, and GapCloser Software was used for filling Gaps between the contigs. From these data, some specific databases were generated for coffee: Coffee Genome Hub (<http://coffee-genome.org>), an integrated web-based database providing centralized access to coffee community genomics, genetics and breeding data and analysis tools to facilitate basic, translational and applied research in coffee. Data available are the complete genome sequence of *C. canephora* along with gene structure, gene product information, metabolism, gene families, transcriptomics (ESTs, RNA-Seq), genetic markers and genetic maps. The hub also provides tools for easy querying, visualizing, and downloading research data. Also, some information is included in familiar databases, such as NCBI Map Viewer (<http://www.ncbi.nlm.nih.gov/mapview/>). The sequence was published in: “The Coffee Genome Provides Insight into Convergent Evolution of Caffeine Biosynthesis” by Denoeud et al., 2014.

<sup>1</sup>Faculty of Natural Science and Technology, Tay Nguyen University;

<sup>2</sup>Department of Basic Science, Buon Ma Thuot Medical University

<sup>3</sup>Faculty of Agriculture and Forestry, Tay Nguyen University;

Corresponding author: Nguyen Dinh Sy; Tel: 0961367958; E-mail: ndsy@ttn.edu.vn.