Transcriptome Study During Manufacturing Process of Semi-fermented Tea

INTRODUCION

Tea (*Camellia sinensis* (L.) O. Kuntze) belongs to the family Theaceae and is one of the world's three major beverage crops. Tea is native to China and by now wildly cultivated and consumed worldwide. Harvested tea shoots mainly processed into three different categories of teas including non-fermented green tea, semi-fermented oolong tea and fully fermented black tea. The appealing flavor, aroma, potential health benefits and medical properties of tea make it popular and highly valued.

The flavor of tea is principally determined by complex chemical components, mainly comprising of polyphenols such as catechins and flavonoids that contribute to important physiological activities and the quality of tea. Other constituents are caffeine, theobromine, theophylline and other alkaloids, proteins and free amino acids, carbohydrates, organic acids, volatile compounds, minerals, vitamins, and pigments (Balentine et al., 1997).

General manufacturing process of semi-fermented oolong tea in Taiwan is described as follow. After tea shoots with one tip and two leaves were plucked from the tea plant, they were exposed to sunlight for about one hour solar withering, and then indoor withering at room temperature, accompanied by around four times shaking that turn over tea shoots, with approximately one to two hours intervals for fermentation process. Following steps were blenching with panning, rolling, and finally drying at 90ºC for 2 hours to obtain the raw tea (Fig. 1).

During oolong tea and black tea manufacturing, withering induces the oxidation and fermentation process, and the oxidation of catechins, amino acids and carbohydrates forms different flavor and aroma compounds. Fresh tea shoots are also virtually odorless

or with slightly green smell, suggesting that most floral and fruity aroma compounds of oolong tea are produced during the tea manufacturing process and the action of specific enzymes may also be involved.

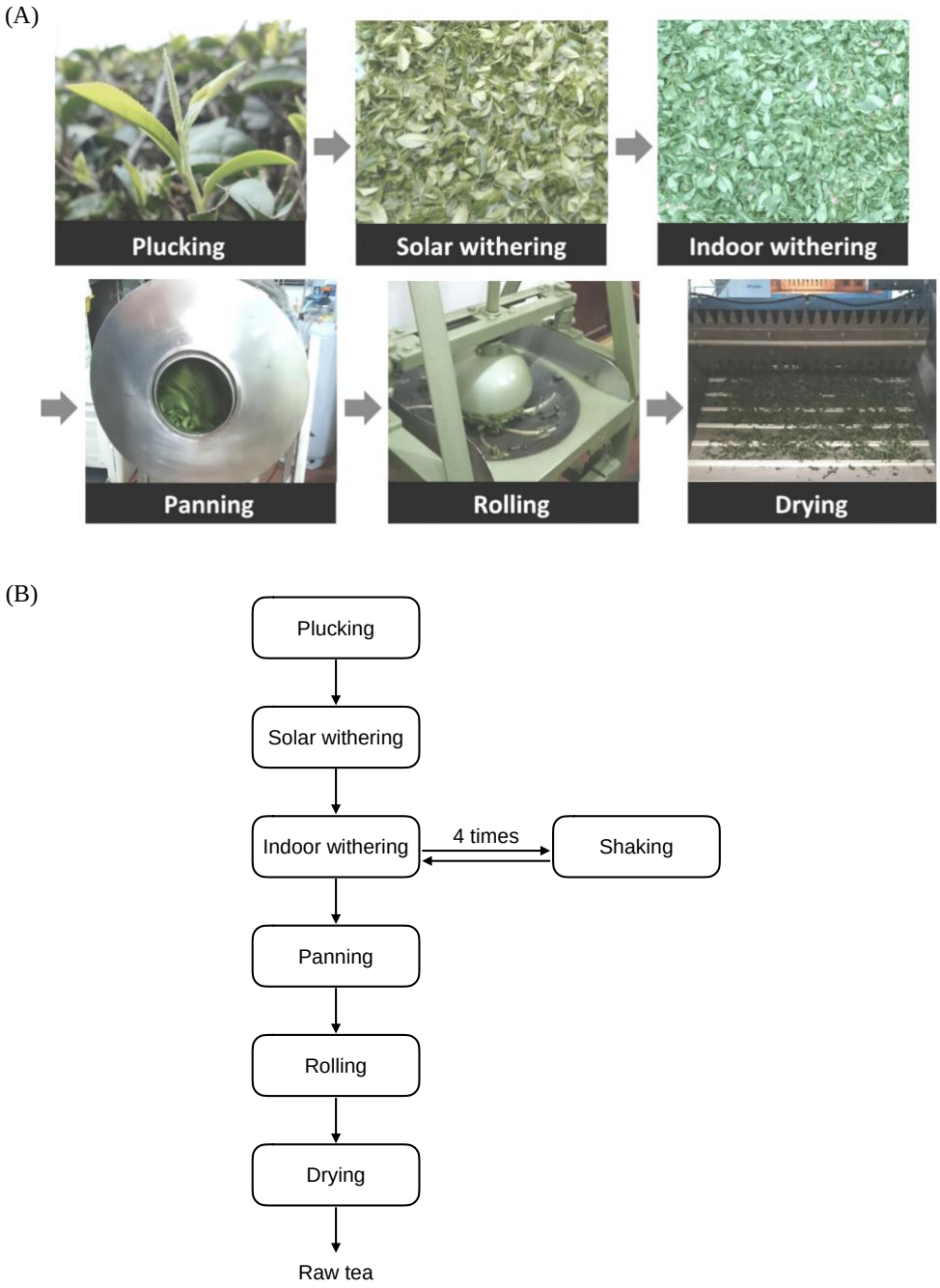

Fig. 1. Semi-fermented oolong tea manufacturing process.

The arrival of NGS technologies had changed the vision of basic, applied and clinical researches (Metzker, 2010). RNA-seq, or massively parallel cDNA sequencing, revolutionized transcriptomics by allowing RNA analysis through cDNA sequencing at massive scale, that provides a progressively fuller knowledge of transcript biology in both quantitative and qualitative aspects. The applications and advances of RNA sequencing included protein coding gene annotation, transcription start sites mapping, strand-specific RNA sequencing, alternative splicing pattern and gene fusion detection, direct RNA sequencing, small RNA and low-quantity RNA sample profiling (Ozsolak and Milos, 2011). In plant biology, RNA-seq can be applied to molecular marker development, hybridization and introgression, polyploid genetics, phylogenetic and ecological studies (Egan et al., 2012). Methods to design RNA-seq projects for non-model plant species and to analyze and interpret the data had been reviewed and proposed (Strickler et al., 2012).

Recent studies involved next generation sequencing techniques in tea transcriptome researches in order to gain more information. Some omics studies about the biosynthesis of biochemical compounds during different tea manufacturing process had also been published (Chen, et al., 2020; Zhang, et al., 2021; Zeng, et al., 2019). *C. sinensis* is a diploid dicot (2n=30) with a large genome. The draft genome of *C. assamica* and *C. sinensis* had also been performed recently, which may provide molecular and genetic insights into the biosynthesis and regulation of important tea compounds (Wei et al., 2018; Xia et al., 2017).

THE AIM OF THIS STUDY

Polyphenols and volatile organic compounds are key chemical compounds that determines the quality of tea. However, previous researches put more emphasis on the analysis of chemical contents and composition of tea. Further researches focusing on the

biosynthesis pathway and the enzymes responsible for the synthesis of these compounds are needed, especially during manufacturing process which produces different flavor and aroma compounds including theaflavins, thearubigins, and most floral and fruity aroma compounds.

In this study, we planned to identify the expression profile of genes related to the biosynthesis of key soluble chemical compounds and volatile organic compounds during tea manufacturing process with the use of RNA-seq and data analysis techniques. The analysis of sequencing data will include reads preprocessing and quality analysis, transcripts assembly, gene annotation, gene expression analysis, gene clutering and functional analysis, and differentially expressed gene analysis (Fig. 2). The results were expected to provide insights into the physiological significance of each manufacturing step, and possible approach to improve the natural flavors of tea.

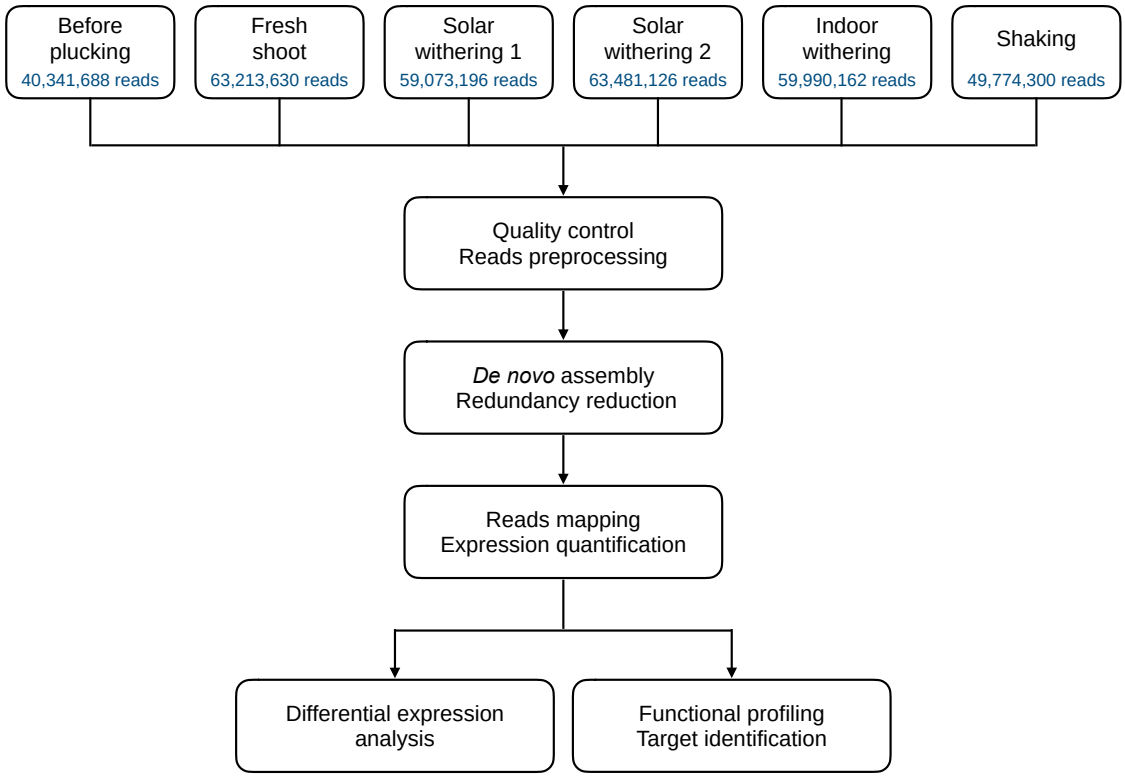

Fig. 2. Pipeline for tea transcriptome analysis.

Tea cultivar Taiwan Tea Experiment Station No.12 (TTES No.12), also called Chinhsuan Oolong, was developed by Taiwan Tea Research and Extension Station (TRES) in 1981, which was suitable for making semi-fermented tea. With high yield, long plucking period and special creamy taste, TTES No.12 was one of the most important tea cultivars in Taiwan, and was chosen to establish Taiwan semi-fermented tea transcriptome database in this study.

PRELIMINARY RESEARCH

Tea plant (*Camellia sinensis* var. *sinensis* cv TTES No.12) was cultivated in the Tea Research and Extension Station (TRES) at Taoyuan, Taiwan, and the following manufacture and sampling were carried out at National Taiwan University. Sampling was carried out during early stage of tea manufacturing, and the samples were immediately frozen with liquid nitrogen, and then stored at -80ºC. The six sampling time points for next generation sequencing analysis were ten days before plucking and manufacturing (B), fresh tea shoots after plucking (F), solar withering for 20 minutes (SW1), complete solar withering for 70 minutes (SW2), indoor withering for 90 minutes without shaking (IW), and finally at 45 minutes after first shaking (S) (Table 1).

Table 1. Sampling points for RNA sequencing

Sampling point	Manufacturing Process
0	Immature tea shoots before plucking (B)
1	Fresh tea shoots with one bud and two leaves (F)
2	Solar withering for 20 minutes (SW1)
3	Solar withering for 70 minutes (SW2)
4	Indoor withering for 90 minutes (IW)
5	45 minutes after first shaking (S)

To obtain a comprehensive *C. sinensis* gene expression profile during early stage of tea manufacturing process, total RNA was extracted from six tea samples including immature and mature tea fresh shoots and samples taken at different manufacturing stages. The RNA concentrations of all samples were above 500 ng/μL, the ratios of absorbance at 260 nm and 280 nm (A260/280) were between 2.0 and 2.1, and the RNA integrity number (RIN) values were above 7.8. For each sample, approximately 40.3 to 63.5 million pair-end clean reads with the length of 90 bp were generated using Illumina HiSeq2000 sequencing system. The Q20 value, indicating 1% of sequencing error rate, was above 98.3%, and the GC percentage was near 44% (Table 2).

Table 2. The result of *Camellia sinensis* RNA sequencing

Sample	Read length (bp)	Clean reads	Clean bases	Q20 (%)	GC (%)
Before plucking	90	40,341,688	3,630,751,920	98.59;98.43	44.43
Fresh	90	63,213,630	5,689,226,700	98.62;98.40	44.71
Solar withering 1	90	59,073,196	5,316,587,640	98.61;98.34	44.16
Solar withering 2	90	63,481,126	5,713,301,340	98.64;98.36	44.15
Indoor withering	90	59,990,162	5,399,114,580	98.65;98.41	44.00
Shaking	90	49,774,300	4,479,687,000	98.60;98.40	44.15

REFERENCES

Balentine, D.A., Wiseman, S.A. and Bouwens, L.C. 1997. The chemistry of tea flavonoids. *Crit. Rev. Food Sci. Nutr.* 37:693-704.

Chen, S., Liu, H., Zhao, X., Li, X., Shan, W., Wang, X., Wang, S., Yu, W., Yang, Z. and Yu, X. 2020. Non-targeted metabolomics analysis reveals dynamic changes of volatile and non-volatile metabolites during oolong tea manufacture. *Food Res. Int.*

128:108778.

Egan, A.N., Schlueter, J. and Spooner, D.M. 2012. Applications of next-generation sequencing in plant biology. *Amer. J. Bot*. 99:175-185.

Metzker, M.L. 2010. Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11:31-46.

Ozsolak, F. and Milos, P.M. 2011. RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet*. 12:87-98.

Strickler, S.R., Bombarely, A. and Mueller, L.A. 2012. Designing a transcriptome next-generation sequencing project for a non-model plant species. *Amer. J. Bot*. 99: 257-266.

Wei, C., Yang, H., Wang, S., Zhao, J., Liu, C., Gao, L., and Wan, X. 2018. Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. *PNAS*. 115:E4151-E4158.

Xia, E.H., Zhang, H.B., Sheng, J., Li, K., Zhang, Q.J., Kim, C., and Gao, L.Z. 2017. The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. *Mol. Plant*. 10:866-877.

Zhang, J., Jia, H., Zhu, B., Li, J., Yang, T., Zhang, Z.Z. and Deng, W.W. 2021. Molecular and biochemical characterization of jasmonic acid carboxyl methyltransferase involved in aroma compound production of methyl jasmonate during black tea processing. *J. Agric. Food Chem*. 69:3154-3164.

Zeng, L., Wang, X., Liao, Y., Gu, D., Dong, F. and Yang, Z. 2019. Formation of and changes in phytohormone levels in response to stress during the manufacturing process of oolong tea (*Camellia sinensis*). *Postharvest Biol. Technol.* 157:110974.