Journal of Proteomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Complementary transcriptome and proteome profiling in cabbage buds of a recessive male sterile mutant provides new insights into male reproductive development

Jialei Ji, Limei Yang*, Zhiyuan Fang, Mu Zhuang, Yangyong Zhang, Honghao Lv, Yumei Liu, Zhansheng Li

Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture/Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, No. 12 ZhongGuanCun South St., Beijing 100081, People's Republic of China

ARTICLE INFO

Keywords:
Cabbage
Male sterility
Transcriptome
Proteome
Correlation

ABSTRACT

Plant male reproductive development is a very complex biological process that involves multiple metabolic pathways. To reveal novel insights into male reproductive development, we conducted an integrated profiling of gene activity in the developing buds of a cabbage recessive genetic male sterile mutant. Using RNA-Seq and label-free quantitative proteomics, 2881 transcripts and 1245 protein species were identified with significant differential abundance between the male sterile line 83121A and its isogenic maintainer line 83121B. Analyses of function annotations and correlations between transcriptome and proteome and protein interaction networks were also conducted, which suggested that the male sterility involves a complex regulatory pattern. Moreover, several key biological processes, such as fatty acid metabolism, tapetosome biosynthesis, amino acid metabolism and protein synthesis and degradation were identified as being of relevance to male reproductive development. A large number of protein species involved in sporopollenin synthesis, amino acid synthesis, ribosome assembly, protein processing in endoplasmic reticulum and lipid transfer were observed to be significantly down-accumulated in 83121A buds, indicating their potential roles in the regulation of cabbage microspore abortion. In summary, the conjoint analysis of the transcriptome and proteome provided a global picture regarding the molecular dynamics in male sterile buds of 83121A.

Biological significance: Male sterile mutants are excellent materials for the study of plant male reproductive development. This study revealed the molecular dynamics of recessive male sterility in cabbage at the transcriptome and proteome levels, which deepens our understanding of the metabolic pathways involved in male development. Moreover, the male sterility-related genes identified in this study could provide a reference for the artificial regulation of cabbage fertility by using genetic engineering technology, which may result in potential applications in agriculture such as production of hybrid seeds using male sterility.

1. Introduction

Cabbage (*B. oleracea* var. *capitata*), a significant species of the Brassicaceae family, is an economically and nutritionally worldwide cultivated vegetable. The Food and Agriculture Organization of the United Nations reported that global cabbage and other brassicas production in 2014 was 71.78 million tons, with China producing 47.30% of the global total (http://faostat.fao.org/). Cabbage is an important source of vitamin C, vitamin K and other phytochemicals, such as sulforaphane and indole-3-carbinol [1]. Cabbage displays strong heterosis, which has been increasingly applied in cabbage production, with the aim of developing better performance, a higher yield and more vigorous cultivars. Male sterility, as an economical and effective system for pollination control, plays an important role in hybrid cabbage seed production.

We have previously reported a male sterile mutant 83121A that was found in the autumn cabbage line 83121. The male sterility of 83121A was found to be controlled by a single recessive gene [2]. Phenotypic and biochemical analyses demonstrated that the failed formation of pollen exine is the main cause of microspore abortion in 83121A [2]. Furthermore, the male sterile line was found to contain an insertion of a long terminal repeat-retrotransposon (LTR-RT) in the first exon of *BoCYP704B1*, which involves the production of sporopollenin, a primary component of the exine [2]. The insertion of LTR-RT not only inhibited

* Corresponding author.
E-mail address: yanglimei@caas.cn (L. Yang).

https://doi.org/10.1016/j.jprot.2018.03.003
Received 1 January 2018; Received in revised form 22 February 2018; Accepted 1 March 2018

the expression of *BoCYP704B1* but also changed the splice sites on pre-mRNA of *BoCYP704B1* [2]. Markers designed to distinguish between the *BoCYP704B1* alleles in the male sterile line and its isogenic maintainer line co-segregated perfectly with male fertility/sterility in test crosses, confirming that this gene is responsible for the fertility defect [2].

Anther development is a very complex biological process that begins with the emergence of the stamen primordia and concludes with the release of pollen grains at anther dehiscence [3]. Although it has been shown that the LTR-RT insertion may have completely destroyed the biological function of *BoCYP704B1*, the gene regulatory mechanisms during sexual reproduction of 83121A are still unclear. Analysis of the molecular mechanism of male reproductive development could contribute to discovering new male reproduction-related genes. Moreover, the exploration of gene regulation mechanisms during male reproduction of the flowering plant may result in potential applications in agriculture such as production of hybrid seeds by using male sterility [3].

High-throughput profiling of transcripts or protein species is a powerful method for analyzing changes in intricate biological processes, with the advantages of dynamic landscapes, deep coverage and high resolution [4]. In an attempt to further explore the regulation mechanism of the male sterility of 83121A, an integrated profiling of gene activity by RNA-Seq and label-free quantitative proteomics was performed. By analyzing the differentially expressed genes in developing buds comparing 83121A and its maintainer line 83121B, we found several key biological processes related to male sterility, such as fatty acid metabolism in tapetum, tapetosome biosynthesis, amino acid metabolism and protein synthesis and degradation. The conjoint analysis of the transcriptome and proteome provides a complete picture regarding the physiological state in male sterile buds and expands our understanding of the regulation of the metabolic network in cabbage male reproductive development.

2. Materials and methods

2.1. Plant materials

The near isogenic lines, including the male sterile line 83121A and its maintainer line 83121B, were provided by the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China. After vernalization, the two lines were planted in a climate-controlled greenhouse with 20/15 °C (12/12 h) day/night temperature and 70% relative humidity (RH). During the flowering stage, the buds before the bicellular pollen stage (bud length < 4 mm) were collected from 10 cabbage plants, and the buds were sampled for transcriptomic and proteomic analyses. All experiments were repeated three times. The experimental system is shown in Fig. S1.

2.2. Transcriptome sequencing

Total RNA from all of the obtained samples was extracted using the RNAprep Pure Plant Kit (TIANGEN) according to the manufacturer's instructions. The library construction and sequencing of six RNA samples were completed by Annoroad Gene Technology Co. Ltd. (Beijing, China). The RNA-Seq data have been submitted to Sequence Read Archive (SRA) with the accession number SRP091687. After data processing of raw sequences, clean reads were aligned against the *B. oleracea* genome (http://www.ocri-genomics.org/bolbase/) using Tophat (v2.0.12) [5,6]. The gene expression level was estimated by the reads per kilobase per million mapped reads (RPKM) [7]. The differential gene expression analysis was performed using DESeq (v1.16). Genes with $q < 0.05$ and $|\log2 \text{ ratio}| > 1$ were identified as differentially expressed genes.

2.3. Protein extraction and digestion

Briefly, all tissue samples were transferred to a 1.5-mL screw capped tube and centrifuged at 10,000 ×*g* for 30 min at 4 °C. Next, 100 µL of lysis buffer (8 M urea, 100 mM Tris-HCl, 1× protease inhibitors) was added to each sample, and then they were ultrasonically crushed to extract total protein. Protein was precipitated with trichloroacetic acid (TCA) for 30 min on ice and centrifuged at 40,000 ×*g* for 30 min. Protein concentration was determined with a BCA protein assay kit using bovine serum albumin (BSA) as a standard and a wavelength of 562 nm. An appropriate amount of protein was taken for electrophoresis. The gel was stained and destained until the background was clear. Each lane, corresponding to one replicate, was divided into three pieces of proper gel blocks longitudinally, and each gel block was cut into small pieces of 1 mm^3. After completely destaining and washing, the tryptic digestion of gel pieces was performed with an overnight incubation at 37 °C. The in-gel digestion of protein was performed as previously described by Shevchenko et al. [8].

2.4. Mass spectrometry

The digested peptides were separated using a Thermo Scientific EASY-nLC 1000 System. Peptide mixtures were loaded onto a self-made C18 trap column (Acclaim PepMap100 column, 2 cm × 100 µm, C18, 5 µm) in solution A (0.1% formic acid) and then separated with a self-made Capillary C18 column (EASY-Spray column, 12 cm × 75 µm, C18, 3 µm) with gradient solution B (100% acetonitrile and 0.1% formic acid) at a flow rate of 350 nL/min. The separation gradient is as follows: 4% solution B for 5 min, 15% for 35 min, 25% for 25 min, 35% for 5 min, and then 95% solution B for 15 min. The separated peptides were examined in an Orbitrap Fusion mass spectrometer (Thermo Scientific). The spray voltage of the ion source was set to 2.1 kV. Full-scan mass spectra were acquired in the MS over 350–1800 *m/z* with a resolution of 70,000. The HCD spectra resolution was 17,500. The normalization collision energy was set to 29%.

2.5. Data processing for proteomics

The MS raw data were processed using MaxQuant software. The MS data were searched against the cabbage genome database (http://www.ocri-genomics.org/bolbase/). Parameters of MaxQuant searches were set as follows: carbamidomethylation (C) and oxidization (M) were set as fixed and variable modifications, respectively; precursor ion mass tolerance was 15 ppm, and fragment ion mass tolerance was 20 mmu; fully tryptic peptides with ≤2 missed cleavages were permitted. The threshold of global false discovery rates (FDR) for peptide and protein species identification was set to 0.01. Following protein identification, the intensity for each identified protein was calculated using peptide signal intensities. According to MaxLFQ label-free quantification method, retention time alignment, label-free quantification, and MaxLFQ normalization were performed as previously described [9]. The identification transfer protocol ("match-between-runs" feature in MaxQuant) was conducted within the experimental replicates to extract the quantification information across the replicates [9,10].

To obtain the quantitative data for all of the peptides in the samples, Perseus (version 1.4.1.3) was used to compare the peak intensities from the whole set of measurements. The LFQ protein intensities from the MaxQuant analysis were imported and transformed to a logarithmic scale with base two. The missing values were replaced with the value of the lowest intensity to compensate for the low signals of the low-abundant proteins. The protein quantification and statistical significance analysis were performed using two-way Student's *t*-tests. Protein species with fold-change > 1.5 and FDR-adjusted (p-value < 0.05) were identified as differentially abundant protein species (DAPS) between the experimental groups. FDR-adjusted ($p < 0.05$) was corrected using the method of Benjamini–Hochberg.