

Shot-gun proteome and transcriptome mapping of the jujube floral organ and identification of a pollen-specific S-locus F-box gene

Ruihong Chen¹, Guoliang Chen¹ and Jian Huang²

¹ Shaanxi Province Key Laboratory of Jujube, College of Life Science, Yan'an University, Yan'an, China

² College of Forestry, Northwest A & F University, Yangling, China

ABSTRACT

The flower is a plant reproductive organ that forms part of the fruit produced as the flowering season ends. While the number and identity of proteins expressed in a jujube (*Ziziphus jujuba* Mill.) flower is currently unknown, integrative proteomic and transcriptomic analyses provide a systematic strategy of characterizing the floral biology of plants. We conducted a shotgun proteomic analysis on jujube flowers by using a filter-aided sample preparation tryptic digestion, followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). In addition, transcriptomics analyses were performed on HiSeq2000 sequencers. In total, 7,853 proteins were identified accounting for nearly 30% of the 'Junzao' gene models (27,443). Genes identified in proteome generally showed higher RPKM (reads per kilobase per million mapped reads) values than undetected genes. Gene ontology categories showed that ribosomes and intracellular organelles were the most dominant classes and accounted for 17.0% and 14.0% of the proteome mass, respectively. The top-ranking proteins with iBAQ > 10¹⁰ included non-specific lipid transfer proteins, histones, actin-related proteins, fructose-bisphosphate aldolase, Bet v I type allergens, etc. In addition, we identified one pollen-specificity S-locus F-box-like gene located on the same chromosome as the S-RNase gene. Both of these may activate the behaviour of gametophyte self-incompatibility in jujube. These results reflected the protein profile features of jujube flowers and contributes new information important to the jujube breeding system.

Submitted 23 January 2017

Accepted 27 June 2017

Published 17 July 2017

Corresponding author

Jian Huang, huangj@nwsuaf.edu.cn

Academic editor

Chun Liang

Additional Information and
Declarations can be found on
page 14

DOI 10.7717/peerj.3588

© Copyright
2017 Chen et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Agricultural Science, Genomics, Plant Science

Keywords *Ziziphus jujuba*, Floral biology, Transcriptome, Proteome, Gametophytic self-incompatibility, S-locus F-box gene

INTRODUCTION

Chinese jujube (*Ziziphus jujuba* Mill.) is one of the most popular fruit trees in China. The fruits are valued for their good taste and their high nutritional and medical value (Liu et al., 2014; Huang et al., 2016). The increasing economic value of the Chinese jujube has resulted in an increase in cultivation acreage.

Flowers are a critical reproductive organ and are the origin of these valuable fruits. They begin to develop in early June, originating from the leaf axil with shoot growth, and continue to be produced until the end of October in northern China

(Niu, Zhang & Yuan, 2015). Jujube flowers also produce abundant nectar, and honey made from jujube nectar is one of the most popular types of honey in China. However, jujube flowers are small and drop easily; 87.9–99.9% of the flowers and 68.5% of young fruits usually drop during development, and only 1.1% of the fruits reach maturity on average (Liu *et al.*, 2009). Therefore, understanding jujube floral biology is important and may provide insights into fruit setting and jujube breeding. Although several studies have focused on the biology of flower differentiation (He, Zhang & Jiang, 2009; Song *et al.*, 2013; Niu, Peng & Li, 2011), a global analysis of protein expression in jujube flowers is needed to understand floral biology.

Proteomics provide a comprehensive and quantitative approach to identify proteins that are expressed in a given organ, tissue, or cell line. With the advance in high-resolution mass spectrometry (MS)—based shotgun proteomics (Mann, Kulak & Nagaraj, 2013), several complete proteome maps have been assembled in recent years. However, these studies were restricted to model organisms (Wilhelm *et al.*, 2014; Kim *et al.*, 2014; Baerenfaller *et al.*, 2008). Recently, whole-genome sequencing has been applied to more organisms, and the developed gene models have provided information on the theoretical protein-coding capacity of an organism. These advances have enabled complete proteome analysis for more species. Since the first proteome map of a model plant species, *Arabidopsis thaliana*, was assembled (Baerenfaller *et al.*, 2008), rapid progress in plant proteome has been made in various plant species at different tissue levels, such as fruit, seeds, flowers and leaves roots (Zeng *et al.*, 2011; Girolamo, D'Amato & Righetti, 2012; Dai *et al.*, 2013; Zhu *et al.*, 2015). Undoubtedly, sequencing of the jujube genome will also contribute to the proteomics study of jujube (Liu *et al.*, 2014).

Some genes specifically expressed in floral organs have been shown to contribute to flower development and reproductive function. In a previous study, gametophyte self-incompatibility (GSI) induced in *Z. jujuba* led to high percentage of seedless fruits (Asatryan & Tel-Zur, 2013). Self-incompatibility/compatibility is controlled by the pistil S determinant ribonuclease (S-RNase) and the pollen S determinant F-box (SFB) gene(s). The GSI has been well characterized in some core-eudicots, such as *Solanaceae*, *Plantaginaceae* and *Rosaceae*. We have previously identified an S-RNase candidate gene in the chromosome 1 of the assembled 'Junzao' jujube genome (Huang *et al.*, 2016). However, the SFB candidate genes specifically expressed in pollens have not been identified.

A combination of transcriptome and proteome technologies offers a new comprehensive approach to identify the genes involved in developmental systems and environmental responses. This study is the first to generate a draft proteome map of jujube flowers based on the liquid chromatography-tandem mass spectrometry (LC-MS/MS). To further improve the molecular knowledge of GSI in *Z. jujuba*, we attempt to identify the SFB gene(s) specially expressed in anthers/pollens linked to the S-RNase gene. Information from our study will provide a better understanding of jujube flower development and the breeding system.

MATERIALS AND METHODS

Sample collection

Budding flowers were collected from nine year old jujube trees (*Ziziphus jujuba* 'Junzao') in 2013. We collected two biological replicates of flower samples from each of two trees. The collected flowers were immediately immersed in liquid nitrogen, and stored at -80°C for transcriptome and proteome analysis.

Protein extraction and quality determination

We selected one sample replicate for proteome analysis. The trichloroacetic acid (TCA)/acetone precipitation method was used to extract protein from jujube flowers (Wu *et al.*, 2014). Briefly, frozen flowers were ground in liquid nitrogen and precipitated with TCA/acetone (1:9) overnight at -20°C . Then, the homogenate was centrifuged ($7,000 \times g$, 30 min) and washed three times with acetone. The precipitated proteins were freeze-dried and then re-suspended in 800 μl of SDT buffer (4% SDS, 100 mM DTT, 150 mM Tris-HCl, pH 8.0). After 15 min of incubation in boiling water, the homogenate was fractionated by ultrasound (80 W, 10 s, 10 times) on ice. The homogenate was centrifuged at $14,000 \times g$ at 4°C for 30 min, and the supernatant was collected. Finally, the protein concentration was determined using a bicinchoninic acid protein assay kit (BCA, Beyotime, China) and adjusted to 5 $\mu\text{g}/\mu\text{l}$. Protein was also analysed by electrophoresis on a 12.5% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gel (7 cm \times 20 cm) and stained with Coomassie Brilliant Blue (CBB) staining buffer (GelCode blue; Pierce, Waltham, MA, USA).

In-solution protein digestion

Proteins were digested by following the filter-aided sample preparation (FASP) procedure (Wiśniewski, Zougman & Nagaraj, 2009). Total-protein samples (50 μg) were incorporated into 30 μl of SDT buffer (4% SDS, 100 mM DTT, 150 mM Tris-HCl, pH 8.0) and then heated at 95°C for 5 min. Thereafter, 200 μl of UA buffer (8 M urea, 150 mM Tris-HCl pH 8.0) was added to the protein solution, mixed to remove the detergent, DTT, and passed through an ultrafiltration unit with a nominal molecular weight (MW) cut-off of 30 kDa (Cat No. MRCF0R030, Millipore) to remove other low-molecular-weight components. Then, 100 μl of 0.05 M iodoacetamide in UA buffer was added, and the samples were incubated for 20 min in the dark. The filter was washed three times with 100 μl of UA buffer, followed by two washes with 100 μl of 25 mM NH_4HCO_3 . Finally, 2 μg of trypsin (Promega, Madison, WI, USA) was added, the protein suspension was digested overnight at 37°C and the resulting peptides were collected as the filtrate.

Peptide fractionation with strong cation exchange (SCX) chromatography

Peptides were fractionated by SCX chromatography using an AKTA Purifier system (GE Healthcare). The dried peptide mixture was reconstituted and acidified with 2 ml of buffer A (10 mM KH_2PO_4 in 25% of acetonitrile (ACN), pH 2.7) and loaded onto a Poly SULFOETHYL 4.6 \times 100 mm column (5 μm , 200 \AA ; PolyLC Inc, Maryland, USA). The

peptides were eluted at a flow rate of 1 ml/min using a gradient of 0–10% buffer B (500 mM KCl, 10 mM KH₂PO₄ in 25% of ACN, pH 2.7) for 2 min, 10–20% buffer B for 25 min, 20–45% buffer B for 5 min, and 50%–100% buffer B for 5 min. The eluents were monitored by absorbance at 214 nm and fractions were collected every 1 min. The collected fractions (approximately 30) were combined into 25 pools and desalted on C18 Cartridges (Empore™ SPE Cartridges C18 (standard density), bed I.D. 7 mm, volume 3 ml; Sigma Aldrich, St. Louis, MO, USA). Each fraction was concentrated by vacuum centrifugation and reconstituted in 40 µl of 0.1% (v/v) trifluoroacetic acid. All samples were stored at –80 °C until LC-MS/MS analysis could be conducted.

Liquid Chromatography (LC)—Electrospray ionization (ESI) tandem MS (MS/MS) analysis by Q Exactive

Experiments were performed on a Q Exactive mass spectrometer that was coupled to a nanoflow HPLC (Easy nLC; Proxeon Biosystems/Thermo Fisher Scientific, Waltham, MA, USA). Each fraction (6 µl) was injected for Nano LC-MS/MS analysis. The peptide mixture (5 µg) was loaded onto a C18-reversed phase column (Thermo Scientific Easy Column, 10 cm long, 75 µm inner diameter, 3 µm resin) in buffer A (0.1% formic acid) and separated using a linear gradient of buffer B (80% acetonitrile and 0.1% formic acid) at a flow rate of 250 nl/min controlled by IntelliFlow technology over 60 min.

MS data were acquired dynamically using a data-dependent top10 method by choosing the most abundant precursor ions from the survey scan (300–1,800 m/z) for HCD fragmentation. Target value determination was based on predictive automatic gain control (pAGC). The dynamic exclusion duration was 60 s. Survey scans were acquired at a resolution of 70,000 at 200 m/z, and for HCD spectra, the resolution was set to 17,500 at 200 m/z. The normalized collision energy was 30 eV, and the underfill ratio, which specifies the minimum percentage of the target value likely to be reached at maximum fill time, was defined as 0.1%. The instrument was run in peptide recognition mode.

RNA extraction and Transcriptome analyses

Total RNA of whole budding flowers was isolated using an RNA prep Pure Tissue Kit (Tiangen, Beijing, China) for RNA extraction and then treated with RNase-free DNase I treatment (Promega, Madison, WI, USA). RNA purity was checked using the NanoPhotometer® spectrophotometer (Implen, Westlake Village, CA, USA). RNA concentration was measured using Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). RNA integrity was assessed using an RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Finally, RNA degradation and contamination was visualized on 1% agarose gels. A total of 3 µg of RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, Ipswich, MA, USA) according to manufacturer's recommendation. The libraries were sequenced on an Illumina HiSeq 2000 platform and 100 bp paired-end reads were generated. The raw-sequence reads data are deposited in NCBI Sequence Read Archive ([SRX1518648](#) and [SRX1518646](#)).

After quality control of raw reads, filtered reads were mapped to the 'Junzao' genome using Tophat2 (Kim *et al.*, 2013). Total aligned read numbers were normalized by gene length and sequencing depth for an accurate estimate of expression levels (Robinson & Oshlack, 2010). HTSeq 0.6.1 was used to count the reads mapping to each gene (Anders, Pyl & Huber, 2014). In addition, the RPKM (reads per kilobase per million mapped reads) of each gene was calculated based on the length of the gene to represent the expression level (Mortazavi *et al.*, 2008).

Sequence database searching and data analysis

The mass spectrometric raw data from the top10 method were analysed using MaxQuant 1.3.0.5 (Cox & Mann, 2008). MS/MS data were searched against the gene predicted gene models of *Ziziphus jujuba* 'Junzao' (Accession: LPXJ000000000 and Dryad Digital Repository. doi: 10.5061/dryad.83fr7). An initial search was set at a precursor mass window of 6 ppm. The search followed an enzymatic cleavage rule of Trypsin/P and allowed a maximum of two missed cleavage sites and a mass tolerance of 20 ppm for fragment ions. Enzyme = Trypsin, Missed cleavage = 2, Fixed modification: Carbamidomethyl (C), Variable modification: Oxidation (M), Decoy database pattern = decoy. The cut-off of global false discovery rate (FDR) for peptide and protein identification was set to 0.01. The iBAQ algorithm was used to calculate the approximate abundance of each protein, which normalizes the summed peptide intensities by the number of theoretically observable peptides of the protein (Schwanhäusser *et al.*, 2013).

Gene Ontology (GO) term enrichment analysis of expressed genes and proteins

GO term enrichment analysis was conducted for expressed genes and proteins in the jujube flower using the R package TopGO, and is available from Dryad Digital Repository (doi: 10.5061/dryad.83fr7). GO term enriched results for the transcriptome and proteome were compared for the three aspects of cell components, molecular function and biological progress.

Gene expression validation by quantitative Real-time PCR (qRT-PCR) analysis

We performed a qRT-PCR analysis to validate the relative gene expression levels in the flowers. RNA samples from the same flower samples used for RNA-sequencing analysis were extracted using the same method as described above. cDNA was synthesized using a PrimeScript™ RT reagent kit with gDNA Eraser (TaKaRa, Dalian, China) according to the manufacturer's instruction. We selected 30 genes with different levels of RPKM and iBAQ values for qRT-PCR analyses (Table S1). The primers were designed using Primer-BLAST online software (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Primer specificities were initially verified by PCR amplification of a single product and product size was visualized after electrophoresis on a 1.2% agarose gel. The cDNAs were then amplified in a Bio-Rad CFX Real-Time System (Bio-Rad, Hercules, CA, USA) with SYBR Premix Ex Taq™ II (Takara, Dalian, China). The qRT-PCR amplification protocol was set at 95 °C for 3 min, 95 °C for 10 s, 40 cycles of 95 °C for 10 s and 58 °C for 10 s, and

finally, 72 °C for 30 s. Three replicates were performed for each gene. The UBG gene (Zj.jz021445712) of jujube was chosen as an internal reference gene for normalization (Zhang, Huang & Li, 2015). The average threshold cycle (Ct) was normalized, and the relative expression was calculated using the $2^{-\Delta C_t}$ method.

Identification of S-locus F-box (SFB) genes

We used previously reported S-locus F-box genes as queries to BLAST search 'Junzao' gene models, such as those from *Plantaginaceae* (6 S-locus F-box (SLF) genes), *Solanaceae* (35 SLFs), and *Maleae* (22 S-locus F-box brothers (SFBB)), 8 SLF-like genes and 3 SFB genes of *Prunus*. Because we have identified the S-RNase gene (Zj.jz035833030) located on the linkage group (LG) 1 of the assembled 'Junzao' genome, we treated the newly identified F-box genes located on LG1 as potential SFB genes and then subjected them to phylogenetic analysis with the known reported S-locus F-box genes. All the amino acid sequences were aligned and a neighbour-joining tree was constructed using Mega 6.0.6 (Tamura et al., 2013). Then, RPKM values of those F-box genes were calculated based on the transcriptomic analysis of leaves, flowers, shoot and fruits based on the previous report (Huang et al., 2016). We further performed tissue-specific expression analysis of those F-box genes among different tissues/organs (leaves, shoot, young fruit, ripe fruit, pollen, pistil, petal, flower disc) using RT-PCR. RNA extraction, cDNA synthesis and primer design (Table S3) were performed as described above. PCR was performed in using 2 × GoldStar Best MasterMix (CWBIO, Beijing, China) according to its protocol. PCR products were visualized on a 1.2% agarose gel after electrophoresis.

RESULTS

Transcriptome and Proteome analysis of jujube flowers

RNA sequencing of two independent flower samples (flower 1 and 2) on an Illumina HiSeq 2000 platform generated 52.8 and 59.4 million clean paired-end reads after quality control for flower 1 and 2, respectively (Table S1). In general, 90.1% of the reads from flower 1 and 92.2% of flower 2 were mapped onto the 'Junzao' jujube gene models, and 88.1% and 90.3% of reads were mapped to unique sites, respectively. A total of 18,646 (67.9%) and 17,995 (65.6%) genes had RPKM values ≥ 1 in two samples, respectively (Table S2). The two samples showed a strong correlation (Pearson $r = 0.96$, $P < 0.01$) on the RPKM metrics. In addition, the rank pattern of expression levels of the selected 30 genes generated by qRT-PCR analysis showed a linear correlation with their RPKM values ($r = 0.95$, $P < 0.01$). Taken together, both correlation analyses and independent qRT-PCR evaluation confirmed the reproducibility and validity of the relative expression levels based on the RPKM methods in the transcriptome analysis (Table S2).

In total, 80,802 MS/MS spectra were obtained and evaluated against the 'Junzao' jujube gene prediction database. We identified 7853 proteins with 42,925 unique peptides originating from 80,394 tandem mass spectrometry spectrum assignments with a false-discovery rate (FDR) below 1%, which accounts for nearly 30% of all predicted 'Junzao' gene models (27,443).

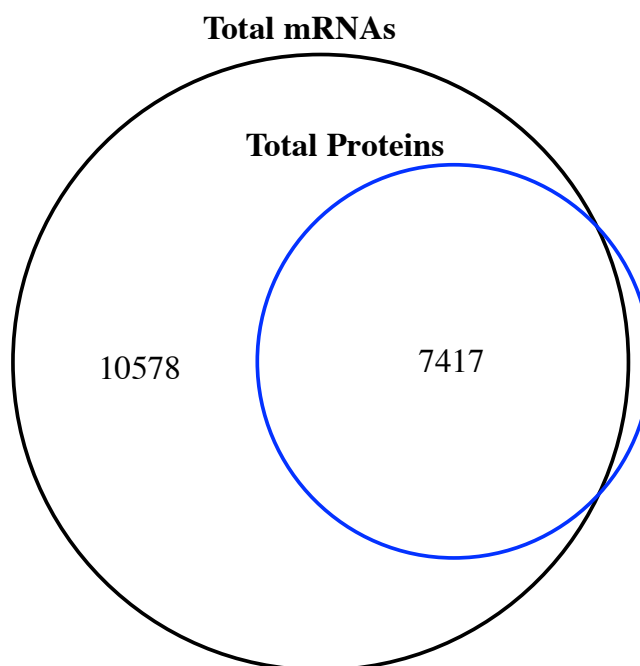


Figure 1 Venn diagram of the number of expressed genes with RPKM ≥ 1 and the identified proteins in the proteome of jujube flowers.

Comparison of the proteome and transcriptome

We compared the transcriptome and proteome data generated from the flower samples based on the 'Junzao' gene models. As shown in the Venn diagram (Fig. 1), 94% of genes identified by MS as protein could also be detected at the mRNA expression level (RPKM ≥ 1), and only 80 genes were missing. The identified proteins accounted for 41.2% of the transcripts with RPKM values ≥ 1 .

The RPKM values of all detected genes show a bimodal distribution (Fig. 2A) in which 81.27% (17,995) and 64.08% (14,182) of genes had signals of RPKM values ≥ 1 and 5, respectively. When the criterion was set at 5, the left shoulder of the bimodal distribution line disappeared and transformed into a Gaussian distribution. By comparison, the identified protein abundance showed a broader distribution range than the filtered mRNA abundance (Figs. 2A and 2B). Among the 7,497 genes identified in the proteomic analysis, only 129 (1.6%) and 473 (6.0%) of genes had RPKM values < 1 and 5, respectively. RPKM values of the genes identified in the proteome or those not identified differ clearly (Fig. 2C). We also found that the RPKM-based transcript abundance values correlated (Spearman's correlation 0.5) well with iBAQ-based protein abundance values in jujube flowers (Fig. 2D). The top10 categories of GO for the proteome and the corresponding transcriptome in the three aspects of cell components, biological progress and molecular functions are shown in Fig. 3. The ranking patterns were generally similar in the transcriptome and proteome in the three aspects, and the highest-ranking categories were consistent among the three aspects.

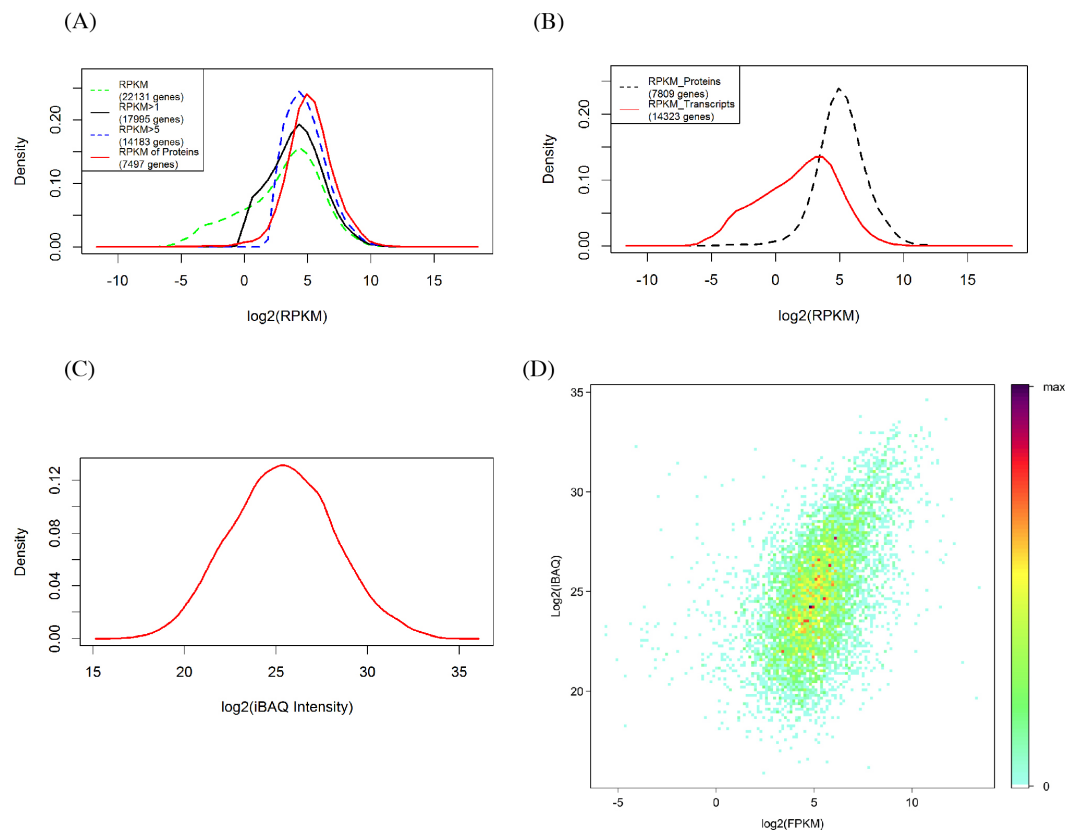


Figure 2 Comparison of proteomics and RNA-Seq data generated from jujube flowers. (A) Distribution of RPKM values in different levels. (B) Distribution of RPKM values of identified and unidentified genes in proteomic. (C) Distribution of proteins abundance (iBAQ intensities, FDR = 1%). (D) A density scatter plot of iBAQ versus RPKM values. The colour indicates the percentage points that were included in a region representing a specific colour.

Protein abundance

The iBAQ values are considered as estimates of the absolute amount of each protein. The most abundant 224 proteins (2.85%) constituted 50% of the jujube flower proteome, and the 691 most abundant proteins, representing 80% abundance, dominated the total protein abundance (Fig. 4). The 41 most abundant proteins comprised 20% of the proteome. The top7 proteins with iBAQ values $>10^{10}$ included two non-specific lipid transfer proteins, histones, actin-related proteins, fructose-bisphosphate aldolase, Bet v I type allergens and unknown proteins. However, ribosomal proteins with six members showed the most abundant protein mass (50, 381, 000, 000) within the top20 proteins (Supplementary Data).

Abundance values of each protein reflect their contribution to the total proteome (Nagaraj et al., 2011). As shown in the GO category, more members of those proteins did not necessarily possess a higher proportion of the protein mass (Fig. S1A). Some protein classes contributed a significantly higher proportion of the total protein mass compared with the corresponding identified protein number. Based on the cellular components, we identified 175 (2.3%) different ribosomes accounting for 17% of the total protein mass

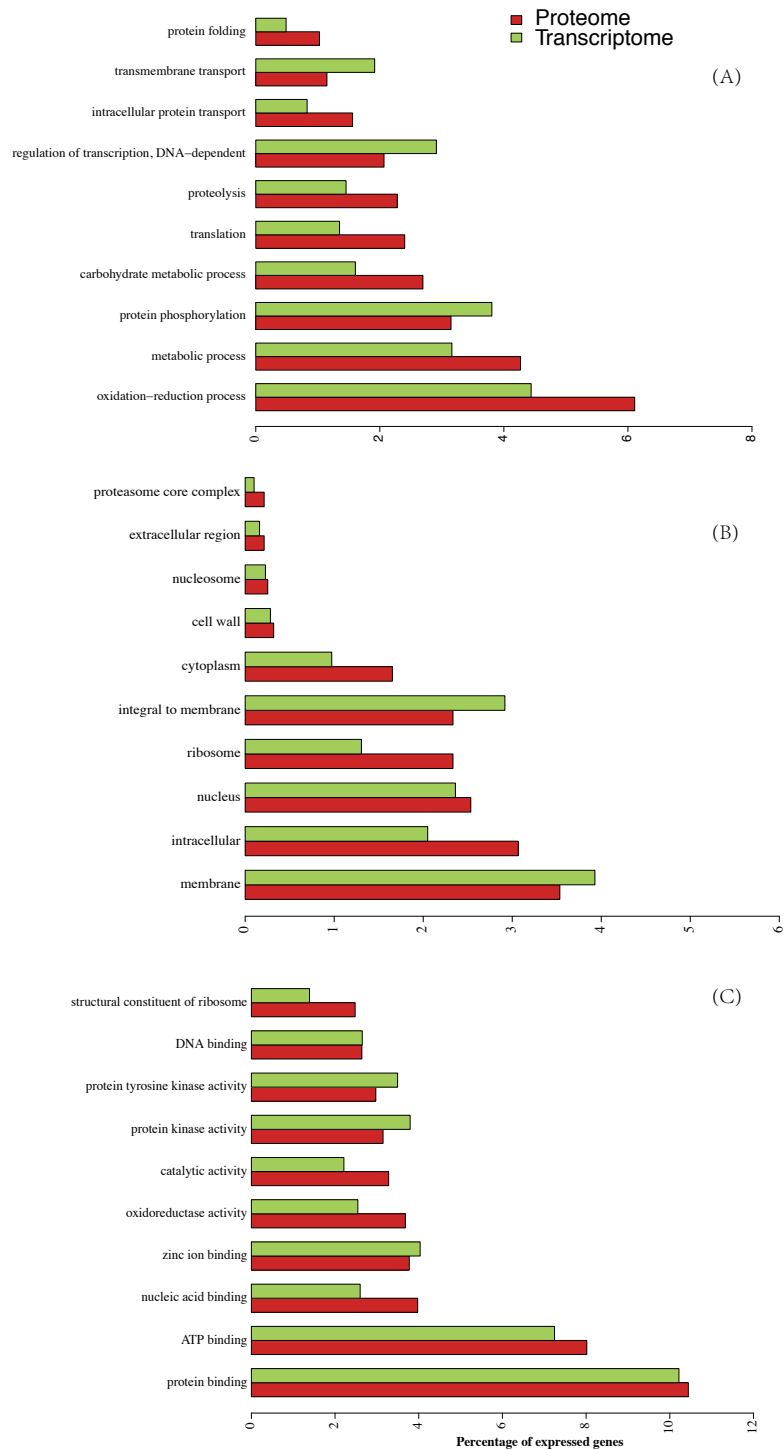


Figure 3 Lists of ten highly ranked GO categories of proteome and their corresponding transcriptome. (A) Biological progress, (B) cellular components, (C) molecular function.

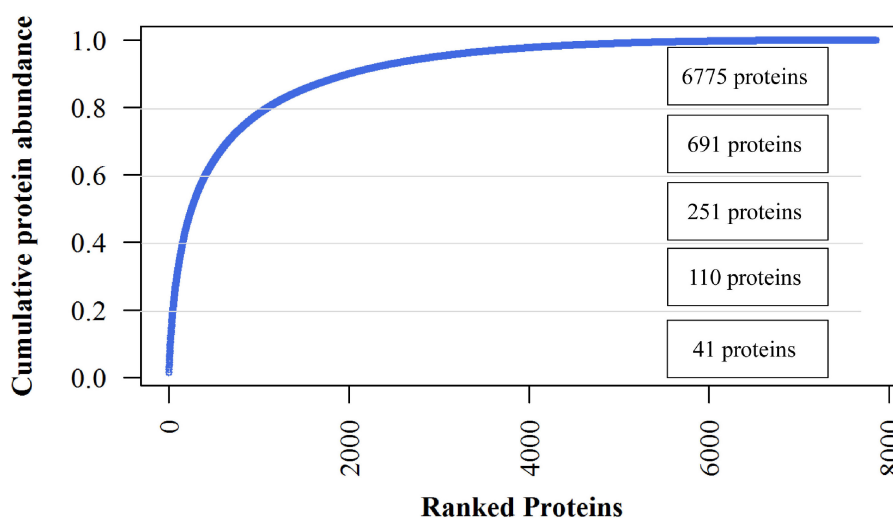


Figure 4 Cumulative protein mass from the highest to the lowest abundance proteins identified in jujube flowers.

in our data (Fig. S1A). Similarly, the intracellular components with 230 genes (3.1%) contributed 13.96% of the proteome mass. This large difference was also observed in the “structural constituents of ribosomes” belonging to ‘molecular function’ and ‘translation’ belonging to ‘biological progress’ (Figs. S1B and S1C). In contrast, integral membrane proteins accounted for 2.3% of the protein mass but contributed much less to the proteome (1.0% of total protein mass).

Identification of SFB gene

We first identified seven candidate SFB genes (Table S3) in LG1 around S-RNase (Zj.jz035833030) by Blastp analysis. Phylogenetic analysis showed that S-locus F-box genes belonging to *Plantaginaceae*, *Solanaceae*, and *Maleae* clustered into one group. Jujube SFB genes were located between *Prunus* SFB genes and SLFL genes (Fig. 5). According to the transcriptomic analysis, we found that No. 5 (Zj.jz026591005) and No. 7 (Zj.jz039541036) were slightly expressed in budding flowers while the other four genes were expressed in several tissues (Fig. 6A). RT-PCR analyses showed that Zj.jz039541036 was strongly expressed in pollens while Zj.jz026591005 was only weakly detected in all flower organs. In contrast to RT-PCR results, Zj.jz027555002 was not detected in any of the tissues (Fig. 6B). Therefore, we consider Zj.jz039541036 as the candidate SFB gene of *Z. jujuba*.

DISCUSSION

In this study, we used an FASP-SCX-LC-MS/MS method to resolve the proteome of jujube flower organs. In total, we identified 7497 proteins that accounted for nearly 30.0% of all predicted ‘Junzao’ gene models (27,443). This depth was similar to the results of the shotgun proteome study of *Arabidopsis* flowers (Baerenfaller et al., 2008), in which 7040 proteins (26.1% of all gene models) were identified and represented nearly 50% of all identified *Arabidopsis* genes in the proteome map. Compared to the improved LC-MS/MS

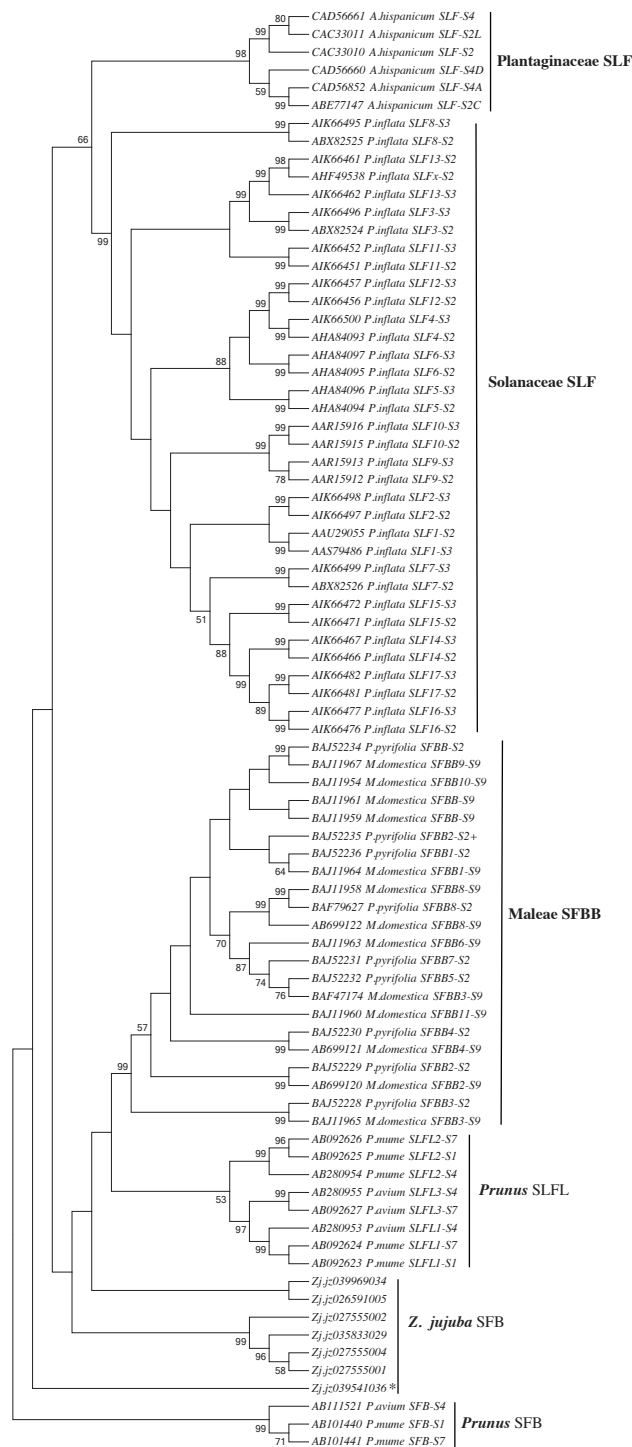


Figure 5 Phylogenetic analysis of the S locus-linked F-box genes from different taxa exhibiting S-RNase-based GSI and seven F-box genes of *Ziziphus jujuba* located in the same chromosome with the S-RNase gene. The neighbor-joining tree was constructed based on aligned amino acid sequences corresponding to those S locus-linked F-box genes from well characterized S haplotypes in each plant species using Mega 6.0 (bootstrap = 1,000). Zj. jz039541036 was the pollen-specificity candidate SFB gene of Jujube.

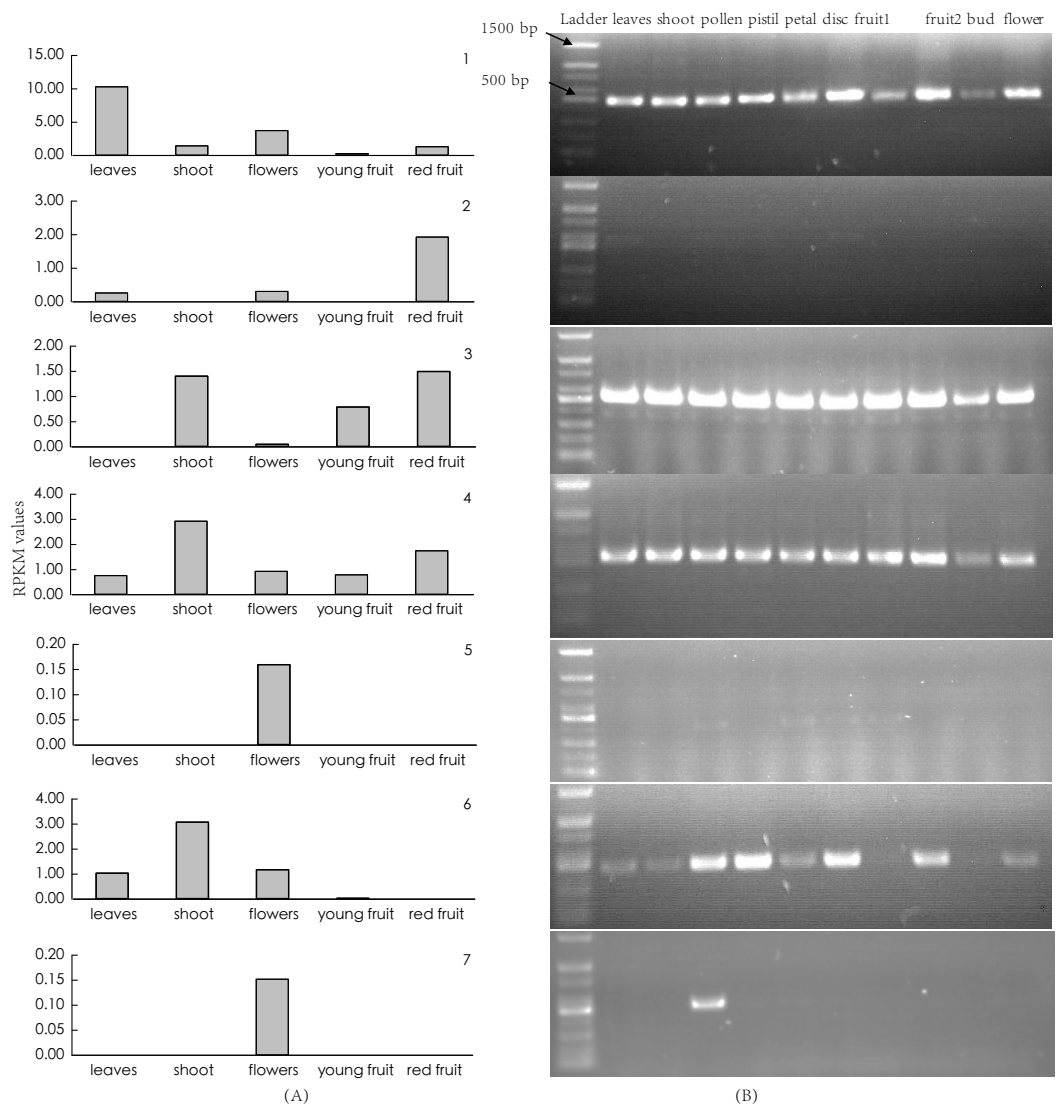


Figure 6 (A) Transcriptomic and (B) RT-PCR analysis of candidate S-locus F-box genes expression in different organs. fruit 1: young fruit, fruit 2: red fruit, flower: opening flower. The number of each photo corresponds to the F-box genes listed in Table S2.

technologies, two-dimensional electrophoresis/mass spectrometry (2-DGE/MS) and multi-dimensional protein identification technology (MudPIT) approaches only identified 2,400 floral proteins in wild *Arabidopsis* flowers (Feng et al., 2009). Lack of genome resources could undermine the efficiency of protein identification. In one study, only 1,488 proteins were obtained from the fruit of *Punica granatum* by nanoliquid chromatography high-resolution tandem mass spectrometry (Capriotti et al., 2013), and 554 of the 2,043 peptides could not be annotated in the proteome of the midgut of *Spodoptera litura* larvae by shotgun ESI-MS (Liu et al., 2010). In the present study, 58% of the genes with RPKM values ≥ 1 were still not covered by the proteome in jujube flowers, suggesting the limitation of the present research strategies. The flower is a complex plant organ that secretes many sugars

that likely interfere with protein separation (Righetti & Boschetti, 2016). When the flower is separated into different organs, such as pistil, petal, stigma, and style, the depth of the flower proteome may be further increased. In general, plant proteomics often identify fewer proteins than animal and human cell or organ proteomics because of higher amount of proteins in animals compared to the protein levels in plant biomass (for example, cell walls contain 90–95% polysaccharides and 5–10% proteins). Therefore, a new strategy for protein extraction should be considered (e.g., CPL; Righetti & Boschetti, 2016).

Recent advances in genome sequencing and proteomics make it possible to study protein abundance and its relationship to corresponding mRNA levels. The distribution pattern of all transcripts in transcriptome showed a bimodal distribution. Transcripts in the left part of bimodal distributions (Fig. 2A) possessed less than one copy and were not expressed as functional protein in tissues (Hebenstreit et al., 2011). When the RPKM value criterion increased to 5, the RPKM values of transcripts showed a Gaussian distribution and were close to the distribution of genes detected in the proteome (Fig. 2A). In addition, the expression levels of those genes that have their corresponding proteins identified in the proteome were generally higher than those of which the proteins were not detected in the proteome (Fig. 2B). Genome-scale proteome analysis of *A. thaliana* preferentially detected proteins that are expressed at higher transcript frequencies (Baerenfaller et al., 2008). Feng et al. (2009) also found that proteins encoded by genes with relatively high levels of expression were detected with greater frequency in the proteome of wild *Arabidopsis* flower investigated by 2-DGE/MS and MudPIT approaches. These results confirm that genes with higher expression levels have a high likelihood of being translated into active protein.

Unlike the observed domination of RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase) in leaf proteins (>50%) and of storage proteins in seeds (>80%), in the jujube flower proteome, ribosomal proteins were the most abundant protein class, accounting for 17% of the total protein mass. This reflects the different biological functions of floral organs in the role of protein composition. Other abundant proteins also contribute to the special features of flowers. Non-specific lipid transfer proteins are abundant and widely distributed in plants, representing as much as 4% of total soluble proteins (Liu et al., 2015). Plant actin-related protein are also an important component and are widely distributed in plant cells. Fructose-bisphosphate aldolase is a key enzyme involved in glycolysis, gluconeogenesis in the cytoplasm, and in the Calvin cycle. Members of fructose-bisphosphate aldolase that are highly expressed in flowers may play an important role in pollen tube growth (Lu et al., 2012; Tang, 2013). In addition, abundant Bet v I type allergens make jujube flowers an allergen source for humans. This finding could also characterize the feature of jujube flower based on proteomics.

GSI is an important feature of reproductive biology in core-eudicots and exhibited in *Z. jujuba*. It has been well characterized that GSI is controlled by one pistil-specificity S-RNase and a single (or multiple) pollen-specificity F-box gene(s). We found there were some differences between the RNA-Seq and RT-PCR. According to the transcriptomic analysis, we were failed to identify a flower-specificity F-box gene linked to S-RNase based on the transcriptomic analysis (Fig. 6A). In contrast, we found a single pollen-specific F-box gene (Zj.jz 039541036) based on RT-PCR analysis after dissecting the flower into its component

organs. Most likely this is because the RNA-Seq of whole flowers is not reflecting the gene expression level in specific organs or tissues. Phylogenetic analysis revealed that the jujube SFB gene was close to that of *Prunus* (Fig. 5), in which GSI exhibited a distinct system of recognition between pollen and pistil involving a single SFB gene. This contrasts with multiple F-box genes that are activated in other plants (Matsumoto & Tao, 2016). With the formerly identified S-RNase gene in *Z. jujuba*, therefore, we concluded that jujube GSI behaviour was like *Prunus* and controlled by the S-RNase and a single SFB gene together. This finding will facilitate further characterization of the underlying molecular mechanism of the reproductive biology of *Z. jujuba*.

Abbreviations

LC-MS/MS	Liquid Chromatography-tandem Mass Spectrometry
MS	Mass Spectrometry
RPKM	Reads Per Kilobase per Million mapped reads
FDR	False discovery rate
CBB	Coomassie Brilliant Blue
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
MW	Molecular Weight
GO	Gene Ontology
qRT-PCR	Quantitative real-time reverse transcription-PCR
iBAQ	Intensity Based Absolute Quantification

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was supported by doctoral research funds from Yan'an University, the Key Laboratory project fund of Shanxi Provincial Education Department (Grant No. 16JS117), and doctoral research funds from Shaanxi Province to Dr. Huang. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Shanxi Provincial Education Department: 16JS117.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Ruihong Chen performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Guoliang Chen performed the experiments, reviewed drafts of the paper.
- Jian Huang conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

Data Availability

The following information was supplied regarding data availability:

The raw data has been supplied as a [Supplementary File](#). and the raw-sequence reads data were deposited in NCBI Sequence Read Archive ([SRX1518648](#) and [SRX1518646](#)).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.3588#supplemental-information>.

REFERENCES

- Anders S, Pyl PT, Huber W. 2014. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169 DOI 10.1093/bioinformatics/btu638.
- Asatryan A, Tel-Zur N. 2013. Pollen tube growth and self-incompatibility in three *Ziziphus* species (Rhamnaceae). *Flora - Morphology, Distribution, Functional Ecology of Plants* 208:390–399 DOI 10.1016/j.flora.2013.04.010.
- Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S, Zimmermann P, Grossniklaus U, Gruissem W, Baginsky S. 2008. Genome-Scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* 320:938–941 DOI 10.1126/science.1157956.
- Capriotti AL, Caruso G, Cavaliere C, Foglia P, Piovesana S, Samperi R, Laganà A. 2013. Proteome investigation of the non-model plant pomegranate (*Punica granatum* L.). *Analytical and Bioanalytical Chemistry* 405:9301–9309 DOI 10.1007/s00216-013-7382-3.
- Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* 26:1367–1372 DOI 10.1038/nbt.1511.
- Dai L, Kang G, Li Y, Nie Z, Duan C, Zeng R. 2013. In-depth proteome analysis of the rubber particle of *Hevea brasiliensis* (para rubber tree). *Plant Molecular Biology* 82:155–168 DOI 10.1007/s11103-013-0047-y.
- Feng B, Li L, Zhou X, Stanley B, Ma H. 2009. Analysis of the *Arabidopsis* floral proteome: detection of over 2000 proteins and evidence for posttranslational modifications. *Journal of Integrative Plant Biology* 51:207–223 DOI 10.1111/j.1744-7909.2008.00787.x.
- Girolamo F, D'Amato A, Righetti PG. 2012. Assessment of the floral origin of honey via proteomic tools. *Journal of Proteomics* 75:3688–3693 DOI 10.1016/j.jprot.2012.04.029.
- He R, Zhang Q, Jiang X. 2009. Observation on flower-bud morph differentiation and flower structure on jujube. *Journal of Tarim University* 21:11–14.
- Hebenstreit D, Fang M, Gu M, Charoensawan V, Van Oudenaarden A, Teichmann SA. 2011. RNA sequencing reveals two major classes of gene expression levels in metazoan cells. *Molecular Systems Biology* 7:497.

- Huang J, Zhang CM, Zhao X, Fei ZJ, Wan KK, Zhang Z, Pang XM, Yin X, Bai Y, Sun XQ, Gao LZ, Li RQ, Zhang JB, Li XG. 2016. The jujube genome provides insights into genome evolution and the domestication of sweetness/acidity taste in fruit trees. *PLOS Genetics* 12(12):e1006433 DOI 10.1371/journal.pgen.1006433.
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14:R36 DOI 10.1186/gb-2013-14-4-r36.
- Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabuddhe NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LD, Patil AH, Nanjappa V, Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK, Marimuthu A, Sathe GJ, Chavan S, Datta KK, Subbannayya Y, Sahu A, Yelamanchi SD, Jayaram S, Rajagopalan P, Sharma J, Murthy KR, Syed N, Goel R, Khan AA, Ahmad S, Dey G, Mudgal K, Chatterjee A, Huang TC, Zhong J, Wu X, Shaw PG, Freed D, Zahari MS, Mukherjee KK, Shankar S, Mahadevan A, Lam H, Mitchell CJ, Shankar SK, Satishchandra P, Schroeder JT, Sirdeshmukh R, Maitra A, Leach SD, Drake CG, Halushka MK, Prasad TS, Hruban RH, Kerr CL, Bader GD, Iacobuzio-Donahue CA, Gowda H, Pandey A. 2014. A draft map of the human proteome. *Nature* 509:575–581 DOI 10.1038/nature13302.
- Liu F, Zhang X, Lu C, Zeng X, Li Y, Fu D, Wu G. 2015. Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. *Journal of Experimental Botany* 66:5663–5681 DOI 10.1093/jxb/erv313.
- Liu J, Zheng S, Liu L, Li L, Feng Q. 2010. Protein Profiles of the Midgut of *Spodoptera litura* larvae at the sixth instar feeding stage by shotgun ESI–MS approach. *Journal of Proteome Research* 9:2117–2147 DOI 10.1021/pr900826f.
- Liu MJ, Zhao J, Cai QL, Liu GC, Wang JR, Zhao ZH, Liu P, Dai L, Yan G, Wang WJ, Li XS, Chen Y, Sun YD, Liu ZG, Lin MJ, Xiao J, Chen YY, Li XF3, Wu B, Ma Y, Jian JB, Yang W, Yuan Z, Sun XC, Wei YL, Yu LL, Zhang C, Liao SG, He RJ, Guang XM, Wang Z, Zhang YY, Luo LH. 2014. The complex jujube genome provides insights into fruit tree biology. *Nature Communications* 55:315.
- Liu P, Liu MJ, Zhao ZH, Liu XY, Wang JR, Yan C. 2009. *Investigation on the characteristics of fruiting and seed development in Chinese jujube (Ziziphus jujuba Mill.)*. Leuven: International Society for Horticultural Science (ISHS), 209–214.
- Lu W, Tang X, Huo Y, Xu R, Qi S, Huang J, Zheng C, Wu CA. 2012. Identification and characterization of fructose 1,6-bisphosphate aldolase genes in Arabidopsis reveal a gene family with diverse responses to abiotic stresses. *Gene* 503:65–74 DOI 10.1016/j.gene.2012.04.042.
- Mann M, Kulak NA, Nagaraj N, Cox J. 2013. The coming age of complete, accurate, and ubiquitous proteomes. *Molecular Cell* 49:583–590 DOI 10.1016/j.molcel.2013.01.029.
- Matsumoto D, Tao R. 2016. Distinct Self-recognition in the *Prunus* S-RNase-based Gametophytic Self-incompatibility System. *The Horticulture Journal* 85:289–305 DOI 10.2503/hortj.MI-IR06.

- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008.** Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5:621–628 DOI [10.1038/nmeth.1226](https://doi.org/10.1038/nmeth.1226).
- Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Pääbo S, Mann M. 2011.** Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology* 7:548.
- Niu HL, Zhang HW, Yuan B, Li XG. 2015.** Flower formation and endogenous hormones dynamic in Chinese jujube. *Acta Horticulturae Sinica* 42:655–664.
- Niu YF, Peng JY, Li L. 2011.** Study on morphology of flower organ and anatomical characteristics of microspore different developmental period in Chinese jujube and wild jujube. *Journal of Plant Genetic Resources* 12:158–162.
- Righetti PG, Boschetti E. 2016.** Global proteome analysis in plants by means of peptide libraries and applications. *Journal of Proteomics* 143:3–14 DOI [10.1016/j.jprot.2016.02.033](https://doi.org/10.1016/j.jprot.2016.02.033).
- Robinson MD, Oshlack A. 2010.** A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* 11:R25 DOI [10.1186/gb-2010-11-3-r25](https://doi.org/10.1186/gb-2010-11-3-r25).
- Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2013.** Corrigendum: global quantification of mammalian gene expression control. *Nature* 495:126–127 DOI [10.1038/nature11848](https://doi.org/10.1038/nature11848).
- Song YQ, Li WQ, Li LL, Liu H, Lu GB. 2013.** Observation on jujube flowers morphology and anther characteristics. *Journal of Shanxi Agricultural Sciences* 41:1190–1193.
- Tang X. 2013.** *New functional studies of two cytosolic Fructose Bisphosphate Aldolase genes in Arabidopsis [D]*. Taian: Shandong Agriculture University.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: Molecular Evolutionary Genetics Analysis version 6. 0. *Molecular Biology and Evolution* 30:2725–2729 DOI [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197).
- Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmair A, Faerber F, Kuster B. 2014.** Mass-spectrometry-based draft of the human proteome. *Nature* 509:582–587 DOI [10.1038/nature13319](https://doi.org/10.1038/nature13319).
- Wisniewski JR, Zougman A, Nagaraj N, Mann M. 2009.** Universal sample preparation method for proteome analysis. *Nature Methods* 6:359–362 DOI [10.1038/nmeth.1322](https://doi.org/10.1038/nmeth.1322).
- Wu X, Xiong E, Wang W, Scali M, Cresti M. 2014.** Universal sample preparation method integrating trichloroacetic acid/acetone precipitation with phenol extraction for crop proteomic analysis. *Nature Protocols* 9:362–374 DOI [10.1038/nprot.2014.022](https://doi.org/10.1038/nprot.2014.022).
- Zeng Y, Pan Z, Ding Y, Zhu A, Cao H, Xu Q, Deng X. 2011.** A proteomic analysis of the chromoplasts isolated from sweet orange fruits (*Citrus sinensis* (L.) Osbeck). *Journal of Experimental Botany* 62:5297–5309 DOI [10.1093/jxb/err140](https://doi.org/10.1093/jxb/err140).
- Zhang C, Huang J, Li X. 2015.** Identification of appropriate reference genes for RT-qPCR analysis in *Ziziphus jujuba* Mill. *Scientia Horticultura* 197:166–169 DOI [10.1016/j.scienta.2015.09.026](https://doi.org/10.1016/j.scienta.2015.09.026).

Zhu W, Hu J, Wang X, Tian J, Komatsu S. 2015. Organ-specific analysis of mahonia using gel-free/label-free proteomic technique. *Journal of Proteome Research* 14:2669–2685 DOI [10.1021/acs.jproteome.5b00208](https://doi.org/10.1021/acs.jproteome.5b00208).