

RESEARCH ARTICLE

Open Access



# Identification of candidate genes involved in wax deposition in *Poa pratensis* by RNA-seq

Yu Ni, Na Guo, Qiuling Zhao and Yanjun Guo\*

## Abstract

**Background:** The cuticular wax plays important roles in plant resistance to various biotic and abiotic stresses. Understanding the synthesis and secretion of cuticular waxes is necessary in utilizing cuticular waxes to improve crop productivity and plant ecological adaptation. Due to the lack of genomic resources, little genetic research on cuticular wax deposition has been focused on *Poa pratensis*, a perennial forage and turf grass species that is widely distributed under various habitats. In this study, we performed *de novo* transcriptome sequencing to explore differentially expressed genes between the leaf non-elongation zone (NEZm) and the emerged blade zone (EBZ) and to identify genes related to cuticular wax deposition.

**Results:** A total of 77,707,414 high quality reads were obtained from Illumina HiSeq 2500 platform, which were then assembled into 106,766 unigenes. Among them, 6019 unigenes showed significant differences in expression between NEZm and EBZ. In our assembled sequences, 3087 SSRs molecular markers were discovered. All the unigenes were searched against the NR, Swissprot, GO, COG, and KEGG databases using BLAST program for functional annotation. From 3156 unigenes with more expression in NEZm compared to EBZ, a number of unigenes involved in very long chain fatty acids (VLCFAs) and cuticular wax biosynthesis, transportation and regulation were identified. Several unigenes related to defense response and epidermal patterning were also found. Twelve putative genes involved in VLCFAs and cuticular wax biosynthesis were further analyzed for their expressions using qRT-PCR.

**Conclusions:** The transcriptome of *P. pratensis* leaf was deep sequenced, *de novo* assembled and annotated, and the candidate genes potentially involved in VLCFAs and cuticular wax biosynthesis, secretion and regulation in *P. pratensis* were identified. This provides fundamental genetic resources in improving plant adaptation to abiotic and biotic stresses.

**Keywords:** Cuticular wax, *Poa pratensis*, Transcriptome, Unigenes, Very long chain fatty acids

## Background

Cuticular wax, as the major chemical component of plant cuticles, plays important physiological and ecological roles in the interactions between plants and their abiotic and biotic environments [1]. It has become clear that cuticular wax deposition could prevent water loss through leaf epidermis [2]; reduce water retention on leaf surface [3]; protect plant against enhanced ultraviolet radiation [4]; defend plant against pathogen attack [5]; and protect plant against air pollution and weathering [6]. The cuticular

waxes are very long chain fatty acids (VLCFAs), consisting of hydrocarbons such as *n*-alkanes, primary alcohols, aliphatic ketones, *n*-alkanoic acids, as well as esters, aldehydes, and secondary alcohols [7]. Studies have shown that wax compositions differ in their functions and responses to abiotic and biotic stresses [8, 9]. Therefore, understanding the synthesis and transportation of each cuticular wax composition is necessary in utilizing cuticular waxes to improve crop productivity and plant ecological adaptation.

Wax biosynthesis begins with *de novo* C16 or C18 fatty acid biosynthesis in the plastid of epidermal cells. Then using the C16 or C18 acyl-CoA and malonylcoenzyme A as substrates, fatty acid elongase (FAE) complex

\* Correspondence: qhgyj@126.com  
College of Agronomy and Biotechnology, Southwest University, Chongqing 400716, China

performs a reiterative cycle of four reactions catalyzed by a  $\beta$ -ketoacyl-CoA synthase (KCS), a  $\beta$ -ketoacyl-CoA reductase (KCR), a  $\beta$ -hydroxyacyl-CoA dehydratase (HCD), and an enoyl-CoA reductase (ECR), to produce saturated VLCFAs with 24–36 carbon atoms. These VLCFAs are further modified to various wax molecules through two major pathways, the acyl-reduction and the decarbonylation pathways [10]. Using wax-deficient mutants, some genes encoding enzymes involved in these wax biosynthesis pathways have been cloned and characterized. The mutant loci in *Arabidopsis thaliana* are termed *eceriferum* (*cer*), and 21 independent *cer* loci have been identified in this plant model [11]. For example, *CER6/KCS6* and *CER10* have been identified for encoding KCS and ECR involved in VLCFAs biosynthesis [12, 13]. *CER8* encodes Long-chain Acyl-CoA Synthetase (LACS) catalyzing free fatty acids to CoA [14]. *CER3*, *CER1* and *MAH1* encode enzymes involved in the decarbonylation pathway which produces aldehydes, alkanes, secondary alcohols, and ketones [15, 16]. *CER4* and *WSD1* encode enzymes involved in the acyl reduction which forms primary alcohol and wax esters [17, 18]. *CER5/ABCG12* [19] and *ABCG11* [20], which belong to ATP-binding cassette (ABC) transporters, were reported to be required for *Arabidopsis* wax export. The identification of these wax-related genes helps understanding the production of cuticular wax and their functions.

Understanding the development of wax depositions during plant growth also provides an alternative way in screening cuticular wax genes. A microarray study on *Arabidopsis* by Costaglioli et al. (2005) showed that comparison of gene expression between (younger) wax synthesizing and (older) not wax-synthesizing shoot tissue was better suited for screening wax-related genes than comparison of expression between wild-type and *cer* mutant plants [21]. For monocotyledonous plants, Rhee et al. (1998) analyzed the wax deposition along the length of expanding leek (*Allium porrum* L.) and found that the level of total cuticular wax increased along the length of the leaf and the microsomal fatty acid elongation activities were induced within a defined and identifiable region of the expanding leek leaf [22]. In grasses, Richardson et al. (2007a) reported that cutin deposition in barley leaf epidermis occurred in parallel with cell elongation, whereas deposition of significant amounts of wax commenced as cells ceased to elongate [23]. Based on this finding, Richardson et al. (2007b) compared expression of candidate contig-sequences between leaf Elongation Zone, Non-Elongation Zone and Emerged Blade Zone, and identified candidate genes involved in wax deposition on barley leaves through a microarray approach [24].

*Poa pratensis* L., also known as Kentucky bluegrass, is a perennial forage and turf grass species that is well

adapted to a wide range of mesic to moist habitats, and temperate to alpine conditions [25]. Pertierra et al. (2013) also reported that it was the longest surviving non-native vascular plant colony known in Antarctica from 1954 to 2012 [26]. The proportion of total cuticular wax present as alkanes was the highest in the alpine *Poa* species compared to low land *Poa* species [27], implying that the wide distribution of *P. pratensis* might be attributed to its ability in adjusting cuticular wax depositions. The RNA-Seq is a powerful, recently developed, high-throughput sequencing method, and provides new approaches to explore functional genes. In this study, we performed the pair end transcriptome sequencing of leaf non-elongation zone (NEZm) and the emerged blade zone (EBZ) of *P. pratensis* using the Illumina HiSeq platform. A total of 102,878,869 reads were obtained and assembled into 106,766 Unigenes. Thousands of potential simple sequence repeats (SSRs) molecular markers were discovered. About 6019 unigenes were differentially expressed between two leaf zones and a number of unigenes involved in cuticular wax deposition were identified. Twelve putative differentially expressed unigenes involved in cuticular wax biosynthesis were analyzed for their relative expression by further quantitative real-time PCR. The objective of this study was to provide a comprehensive molecular biology insight into the cuticular wax deposition in *P. pratensis* and to identify genes which are likely to be involved in wax deposition in *P. pratensis* leaves.

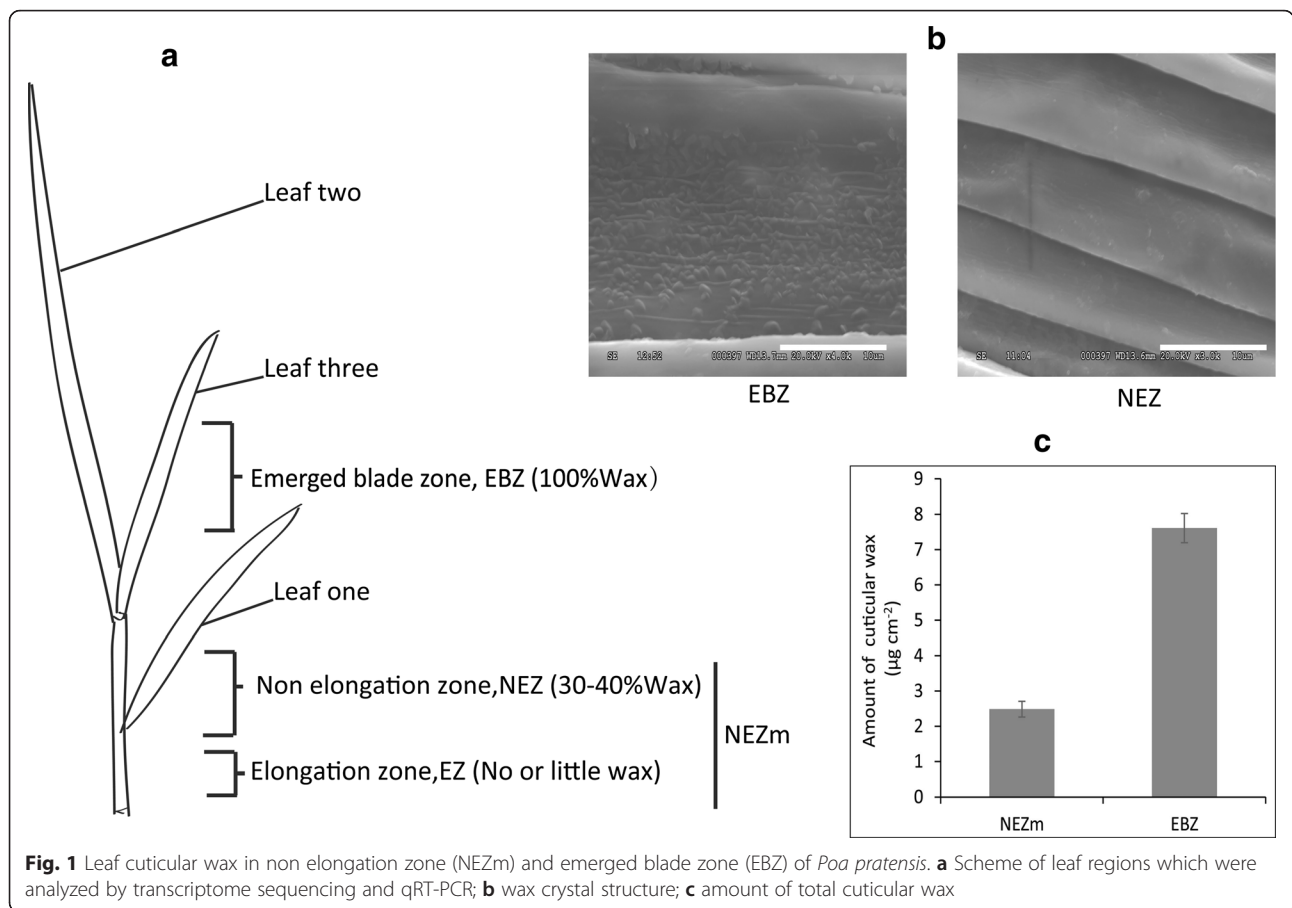
## Results

### Leaf cuticular wax

Leaf three of *P. pratensis* was divided into mixed sample of non-elongation zone and elongation zone (NEZm) and emerged blade zone (EBZ) according to their positions (Fig. 1a). The NEZm was covered by sheath whereas the EBZ was fully expanded and exposed under air. Cuticular wax differentially deposited on the leaf zones of *P. pratensis*. Scanning electron microscope analysis showed that there was less wax crystalloids in NEZm than EBZ (Fig. 1b). The amount of total cuticular wax on EBZ was about three times than that in NEZm (Fig. 1c). In grasses, leaf cells divide and expand within the sheaths of older leaves, where the micro-environment differs from the open atmosphere. By the time epidermal cells are displaced into the atmosphere, they must have a functional cuticle to minimize uncontrolled water loss.

### *Poa Pratensis* transcriptome sequencing and *de novo* assembly

RNA samples from leaf NEZm and EBZ were subjected to Illumina HiSeq 2500 paired-end sequencing. In total, 102,878,869 reads with a total of 20,574,028,763 base-pairs (20.57 Gb) were generated. After stringent quality



assessment and data filtering, 77,707,414 reads (15.54 G) with 93.17 % Q30 bases (those with a base quality greater than 30) were selected as high quality reads for further analysis. An overview of the sequencing was presented in Table 1. All reads of sequencing have been deposited in the NCBI's Short Read Archive (accession number SRX1512962, SRX1512973, SRX1512974, and SRX1512975).

**Table 1** Summary of the sequencing reads and reads after preprocessing

	<i>Poa pratensis</i>
Sequencing reads before preprocessing	
Total reads	102,878,869
Total bases	20,574,028,763
GC%	52.18
Q30%	80.39
Reads after trimming and preprocessing	
Total reads	77,707,414
Total bases	15,540,103,604
GC%	50.81
Q30%	93.17

Using the Trinity *de novo* assembly program [28], short read sequences were assembled into 8,032,387 contigs, of which there were 12,335 contigs coding for transcripts longer than 1 kb and 5435 contigs coding for transcripts longer than 2 kb. The contigs were subsequently subjected to cluster and assembly analysis. A total of 106,766 unigenes with N50 length of 1042 bp and mean length of 640.83 bp were obtained, among which 18,559 genes (17.38 %) were greater than 1 kb. An overview of the assembled contigs, transcripts and unigenes was presented in Table 2. These results demonstrated the effectiveness of Illumina pyrosequencing in rapidly capturing a large portion of the transcriptome.

Since there was no reference genome sequence for *P. Pratensis*, the *de novo* assembled transcriptome sequence by Trinity was regarded as a reference sequence. Bowtie software [29] was used to map all the clean reads to the *de novo* assembly transcriptome reference sequences and to qualify transcriptome by assigning to unigenes with the RSEM (RNA-Seq by Expectation Maximization) software [30]. The 59,429,338 clean reads (76.47 %) were successfully realigned to the reference sequence, showing that the quality of these assembled unigenes was sufficient to

**Table 2** Summary of Illumina transcriptome assembly for *Poa pratensis*

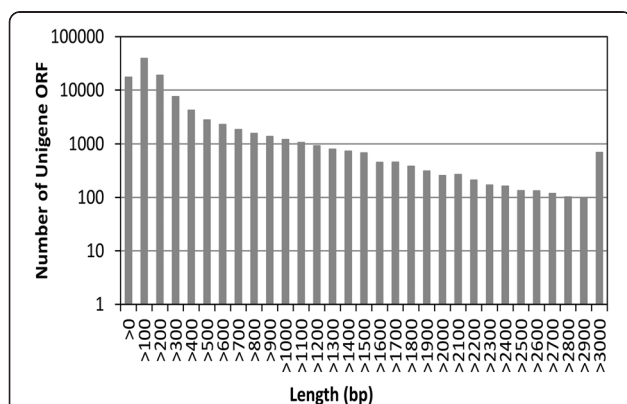
Length range	Total Length (Percentage)		
	Contigs	Transcripts	Unigenes
200–300	7,951,914(99.00 %)	65,336(23.44 %)	41,403(38.78 %)
300–500	40,282(0.50 %)	60,304(21.63 %)	29,234(27.38 %)
500–1000	22,421(0.28 %)	68,871(24.71 %)	17,570(16.46 %)
1000–2000	12,335(0.15 %)	60,743(21.79 %)	12,683(11.88 %)
≥2000	5435(0.07 %)	23,481(8.42 %)	5876(5.50 %)
Total number	8,032,387	278,735	106,766
Total length	466,581,232	239,460,551	68,419,199
N50 length	53	1313	1042
Mean length	58.09	859.10	640.83

conduct the following analysis. Sequencing randomness and sequencing saturation assessment also demonstrated RNA integrity at transcript level (in Additional file 1: Figure S1).

Furthermore, 106,000 unigenes were found to have predicted ORF by using the software Getorf (<http://emboss.sourceforge.net/apps/cvs/emboss/apps/getorf.html>). The predicted ORFs ranged from 57 bp to 10,376 bp in length and N50 is 753 bp. The length distribution of the predicted ORFs was shown in Fig. 2.

**Frequency and distribution of SSRs in the *P. pratensis* leaf transcripts**

SSRs markers are the most important molecular markers in plants and have been proven to be a valuable tool for various applications in genetics and plant breeding. In this study, 18,559 unigenes (more than 1 kb) were used to



**Fig. 2** Length distribution of *Poa pratensis* unigene ORF. The y-axis indicates the number of unigene ORF, and the x-axis indicates the length of unigene ORF. Of the 106,766 unigenes, 106,000 unigenes were found to have predicted ORF by using the software Getorf. The minimum length was 57 bp and the maximum length was 10,376 bp. The largest number of ORF was in the range of 0–300 bp

determine potential microsatellite motifs using MicroSatellite (MISA) software (<http://pgrc.ipk-gatersleben.de/misa>). In total 3087 SSRs were identified in leaf samples (Table 3). Tri-nucleotide repeats (1432, 46.39 %) were the most abundant SSR motif in leaf tissues followed by mono-nucleotide (1032, 33.43 %), di-nucleotide (412, 13.35 %), tetra-nucleotide (29, 0.94 %), penta-nucleotide (8, 0.26 %), and hexa-nucleotide (5, 0.16 %) motifs. The number of compound SSRs and uncertain compound SSRs were 157 (5.09 %) and 12 (0.39 %), respectively.

**Analysis of differentially expressed genes (DEGs)**

The transcript abundance of each unigene was estimated by reads per kilobase of exon per million mapped reads (RPKM). Using software DESeq [31] and FDR ≤ 0.001 and log2FC ≥ 8 as the criteria, 6019 unigenes showed significant differences in expression between the NEZm and the EBZ (Fig. 3), including both 2863 up-regulated unigenes and 3156 down-regulated unigenes in the EBZ library compared to those in the NEZm library.

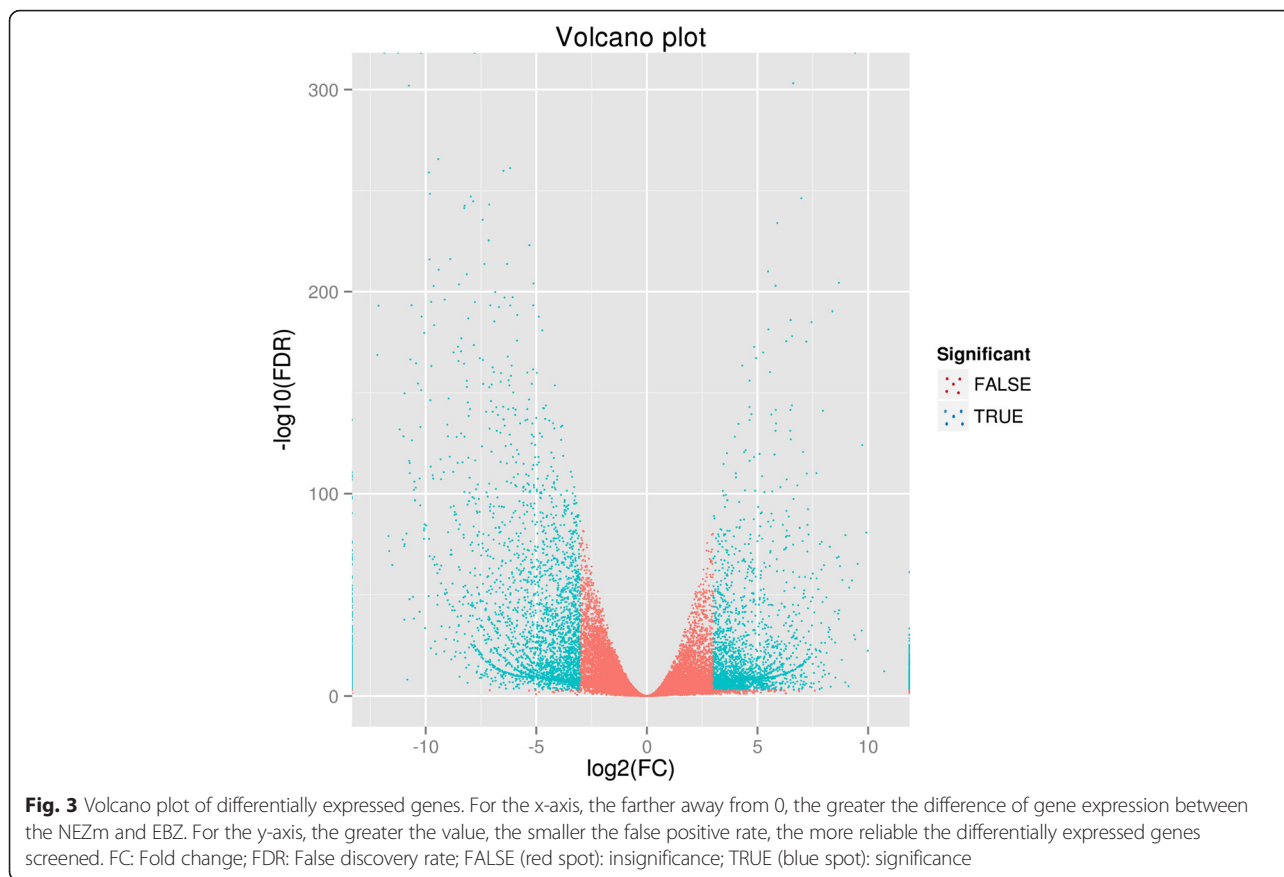
**Functional annotation**

For the validation and annotation of the assembled unigenes, all the assembled unigenes were searched against the NCBI non-redundant (NR), Swissprot, Gene ontology (GO), Clusters of Orthologous Groups (COG), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases using BLAST program with an E-value threshold of 10<sup>-5</sup>. Among 106,766 unigenes, 53,851 (50.44 %) and 32,710 (30.64 %) unigenes had significant matches in the NR and Swiss-Prot database, respectively (Table 4). Furthermore, GO, COG and KEGG annotations were applicable for 33.36, 12.45, and 7.91 % of unigenes, respectively.

**Table 3** Frequency of SSRs in *Poa pratensis*

Motif length	Repeat numbers							Total (%)
	5	6	7	8	9	10	>10	
perfect_SSR_1	-	-	-	-	-	588	444	1032 (33.43 %)
perfect_SSR_2	-	239	82	46	22	10	13	412 (13.35 %)
perfect_SSR_3	1042	280	97	13	-	-	-	1432 (46.39 %)
perfect_SSR_4	25	4	-	-	-	-	-	29 (0.94 %)
perfect_SSR_5	7	1	-	-	-	-	-	8 (0.26 %)
perfect_SSR_6	5	-	-	-	-	-	-	5 (0.16 %)
compound_SSR	-	-	-	-	-	-	-	157 (5.09 %)
compound_SSR*	-	-	-	-	-	-	-	12 (0.39 %)
Total SSR	1079	524	179	59	22	598	457	3087

Perfect\_SSR\_1, SSR\_2, SSR\_3, SSR\_4, SSR\_5, and SSR\_6 represent mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide repeat, respectively. Compound\_SSR represents compound of two or more motifs. \*uncertain SSR resulted from the compound



There were 4782 DEGs annotated in NR database, among which 3442 DEGs were assigned with one or more GO terms (Fig. 4). In the biological process category, 2430 DEGs were classified into 24 GO classes, and “Metabolic process” and “Cellular process” were the most represented GO terms. In the cellular component category and the molecular function category, 1992 and 2545 DEGs fell into 16 GO classes, respectively. “Cell part”, “Cell” and “Organelle” were the most common terms in the cellular component category. Regarding molecular function, unigenes with binding and catalytic activity were highly represented (Fig. 4). Different

enrichment trends between all unigenes and DEG unigenes were mainly observed in “Extracellular matrix”, “Extracellular region” and “Membrane-enclosed lumen” in the cellular component category, “Nutrient reservoir activity”, “antioxidant activity” and “protein binding transcription factor activity” in the molecular function category, and “Cell proliferation” and “Pigmentation” in the biological process category, respectively (Fig. 4).

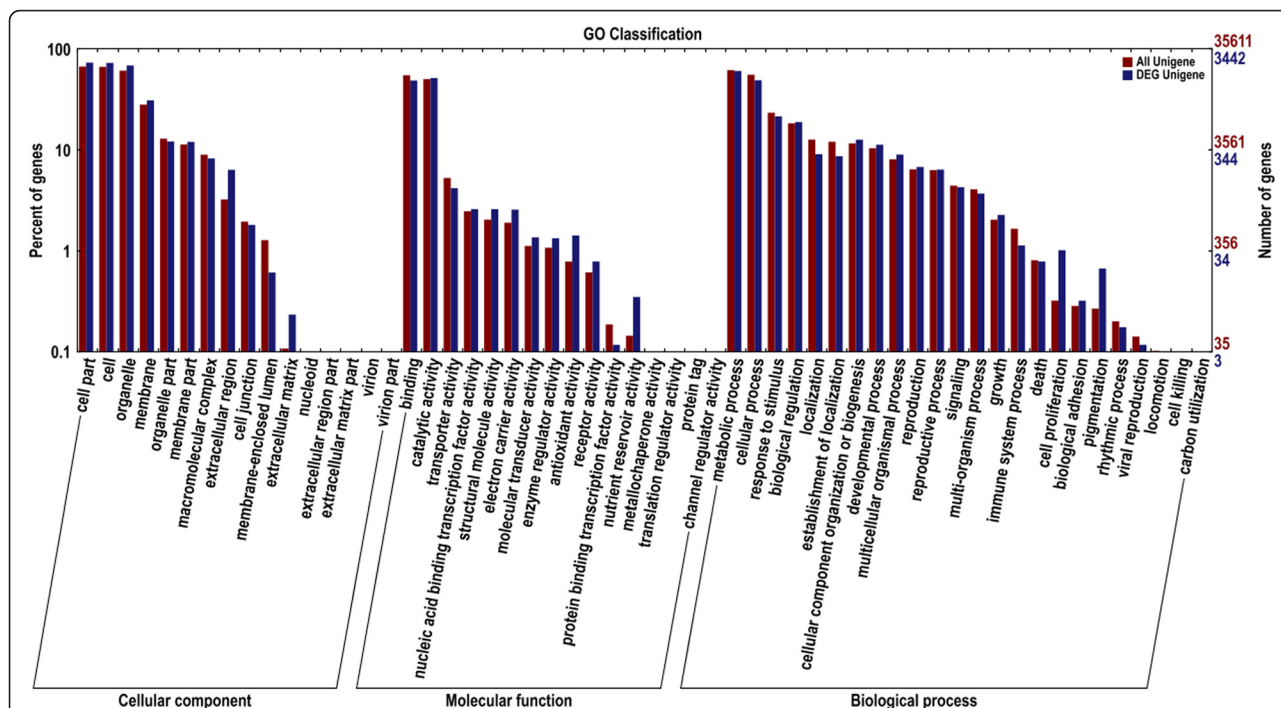
DEG unigenes were also subjected to a search against the COG database for functional prediction and classification. Finally, 1427 DEGs could be assigned to COG classifications (Fig. 5). COG annotated putative proteins were functionally classified into at least 25 protein families involved in lipid transport and metabolism, transcription, signal transduction mechanisms, and so on. The cluster for general function prediction (403, 28.24 %) represented the largest group, followed by replication, recombination and repair (280, 19.62 %), transcription (239, 16.75 %), and signal transduction mechanisms (226, 15.84 %).

To further analyze DEGs between leaf NEZm and the EBZ, all the DEG unigenes were analyzed in KEGG pathway database. As a result, 615 DEGs were found to have significant matches in the database and were

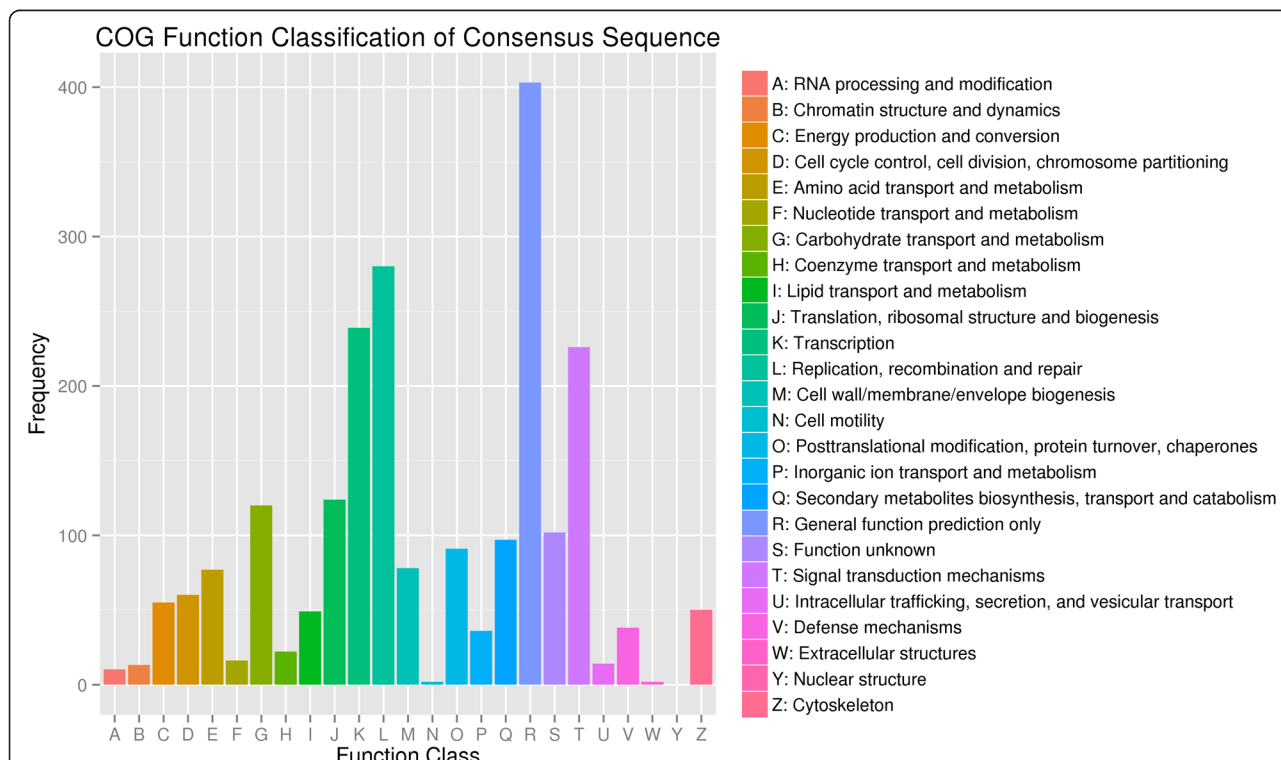
**Table 4** Functional annotation of the unigenes from *Poa pratensis*

Annotated databases	All sequence	>= 300 nt	>= 1000 nt
COG	13,296	11,188	6491
GO	35,619	26,932	12,428
KEGG	8449	6514	3311
Swiss-Prot	32,710	26,122	12,686
NR	53,851	39,903	16,597
All	54,219	40,083	16,623

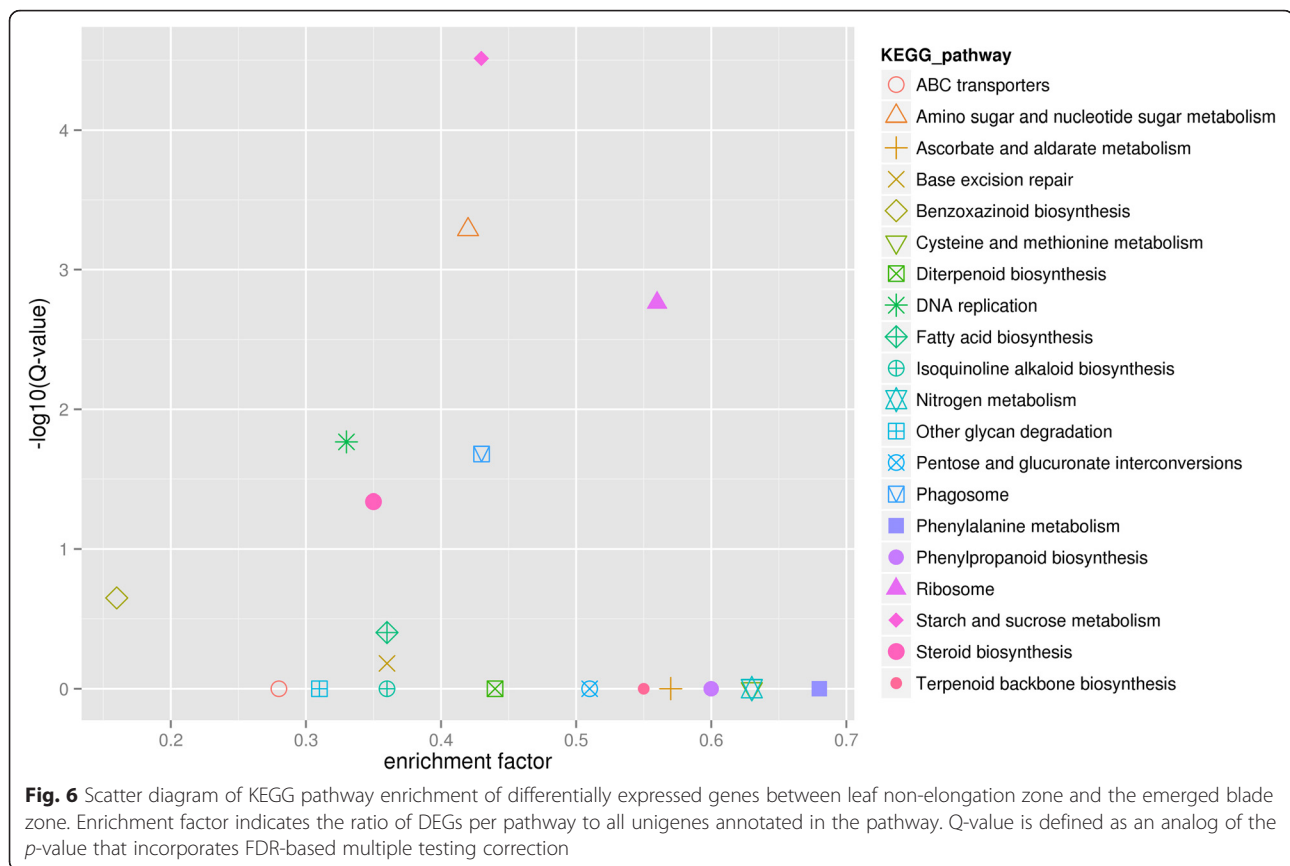




**Fig. 4** Histogram presentation of Gene Ontology classification. The results are summarized in three main categories, Cellular component, Molecular function, and Biological process. The left y-axis indicates the percentage of genes in a category, and the right y-axis indicates the number of genes in a category



**Fig. 5** Histogram presentation of the COG classification of the differentially expression genes between leaf non-elongation zone and the emerged blade zone. In total, 1427 DEGs showing significant homology to the COGs database were grouped into at least 25 categories. The capital letters in x-axis indicate the COG categories as listed on the right and the y-axis indicates the number of DEGs in each category



assigned to 100 KEGG pathways (in Additional file 2: Table S1). The top 20 KEGG pathways with the most significant enrichment were shown in Fig. 6. It was also observed that there were 9 DEGs encoding enzymes that were involved in fatty acid biosynthesis.

#### Differentially expressed genes involved in wax biosynthesis in *P. pratensis*

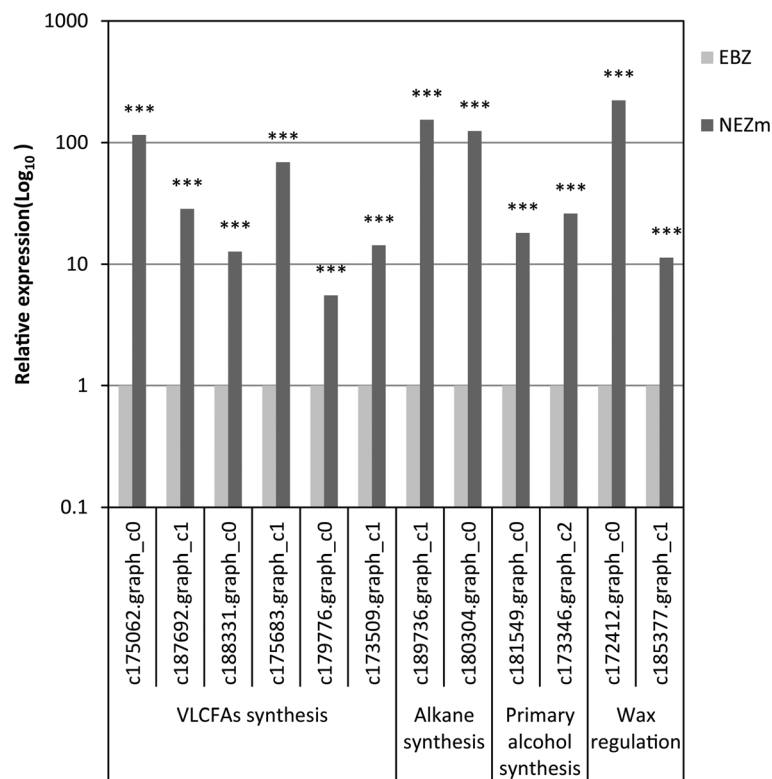
To identify the genes likely involved in wax biosynthesis in *P. pratensis*, we constructed a candidate fatty acid and wax deposition related gene set from down-regulated unigenes in the EBZ (non-active wax synthesizing) compared to the NEZm (active wax synthesizing) based on the GO, COG, KEGG, Swissprot, and NR annotation (in Additional file 2: Table S2). The selected unigene ID and annotation were listed in Additional file 2: Table S2. Three enzymes related to very-long-chain-fatty acid biosynthesis and six enzymes and proteins related to cuticular wax biosynthesis, secretion and regulation were identified. Putative proteins involved in defense response were also found (in Additional file 2: Table S3). In the most cases, more than one unigenes were assigned to the same enzyme or protein. Such unigenes might represent different fragments of a single transcript, different members of a gene family, or both.

#### Quantitative real-time-PCR validation of the candidate DEGs involved in wax biosynthesis

To validate the responsible genes in the candidate gene set involved in wax deposition, twelve unigenes from the above set were selected and detected their expression profile between the NEZm and the EBZ by qRT-PCR. The results showed that twelve unigenes significantly up-regulated in the NEZm compared to the EBZ (Fig. 7), which were in consistent with the transcriptome data.

#### Discussion

Transcriptome sequencing is widely applied in transcriptional and post-transcriptional regulation analysis of genes and global expression pattern analysis of complex genomes. In this study, a comparative *de novo* transcriptome analysis of leaf NEZm versus EBZ was conducted and 102,878,869 reads with a total of 20,574,028,763 bp (20.57 Gb) were generated by Illumina HiSeq 2500 paired-end sequencing. For plants without a genomic sequence, such short reads can be effectively assembled through the improvement of read length by the paired-end sequencing and the using of Trinity [28]. In total, 8,032,387 bp contigs, 278,735 bp transcripts and 106,766 bp unigenes were obtained in *P. pratensis*, respectively.



**Fig. 7** Validation of selected twelve up-regulated transcripts in the non-elongation zone as compared to the emerged blade zone involved in wax biosynthesis by qRT-PCR. EBZ and NEZm represent the emerged blade zone and the non-elongation zone, respectively. The relative gene expression levels as expressed by  $2^{-\Delta\Delta Ct}$  were determined separately for each treatment as the mean  $\pm$  SE. Asterisks indicate levels of significance of differential expression ( $P < 0.001$ )

It was previously reported that wax deposition in barley started to be synthesized from the leaf NEZ wrapped by sheath [23]. In this study, transcript levels from the leaf NEZm and the EBZ were compared to obtain a list of significantly differentially expressed unigenes. Using DESeq software, which is suitable for screening differentially expressed genes in samples with biological replicates, and  $FDR \leq 0.01$  and  $|\log_2FC| \geq 2$  as the criteria, 16,626 unigenes showed significant differences in expression between the NEZm and the EBZ. This number of DEGs was far beyond normal differential range. Considering the potential big difference between the two tissues and to increase the accuracy of the screening results,  $FDR \leq 0.001$  and  $|\log_2FC| \geq 8$  was used as adjusted criteria and finally 6019 DEGs were obtained. The functions of the DEGs by GO annotations were mainly related to “Extracellular matrix” and “Extracellular region” in the cellular components, and “Nutrient reservoir activity” in the molecular function category. These functions were in consistent with the vital importance of plant surface wax in protecting tissue from environmental stresses [9]. During cuticle deposition, a massive flux of lipids occurs from the sites of lipid synthesis in the plastid and the endoplasmic reticulum to the plant surface. The endoplasmic reticulum integrated

with the membrane and plasmodesma in cellular components may play an important role in wax synthesis. Furthermore, DEGs were also mainly annotated with “Cell proliferation”, attributing to different surface area expansion rates in NEZm and EBZ. This supports the phenomenon that leaf cells divide and expand within the sheaths of older leaves in grasses.

Based on Swissprot and NR annotation, several candidate genes involved in VLCFAs and cuticular wax biosynthesis pathway were obtained, such as *LACS*, *KCS*, *KCR*, *FAR*, and *CER1*. VLCFAs, as both wax component and wax precursor, was catalyzed by the elongase complexes, including *KCS*, *KCR*, *HCD*, and *ECR*. *KCS* is the first and committing step in VLCFAs biosynthesis and has broad substrate specificity [32]. Alcohol-forming fatty acyl-CoA reductase (*FAR*) and aldehyde carbonylase *CER1* were identified catalyzing the production of primary alcohols and alkanes, respectively [15, 17]. *LACS* catalyzed free long-chain fatty acids to CoA [14]. In this study, *LACS*, *KCS*, *KCR*, *FAR* and *CER1* were expressed predominantly in the NEZm, the zone where significant wax deposition commences, supporting the previous results that younger shoot tissue was the active wax synthesis site [21]. In the developing barley leaf,



wax deposition commences within the portion of the blade that is enclosed by sheaths of older leaves [33]. Changes in the micro-environment of the leaf from the NEZm to EBZ might alter the expression of the unigenes. However, five unigenes (c183078.graph\_c0, c175984.graph\_c0, c171931.graph\_c1, c175102.graph\_c0 and c171931.graph\_c0) annotated as KCS and c185031.graph\_c1 annotated as KCR were expressed higher in the EBZ compared to NEZm, suggesting the functional differentiation of KCS and KCR gene family members. Besides involving in cuticular wax synthesis, VLCFAs were also related to other physiological function such as involving in membrane lipids and sphingolipids [34]. Therefore, it is possible that only some KCS and KCR gene family members participate in wax production.

ABC transporter family was ubiquitously associated with transport across membranes of a broad range of molecules in prokaryotes and eukaryotes [35]. In this study, putative ABC transporter G, B and C families, and Lipid Transfer Proteins (LTPs) were significantly up-regulated in the NEZm compared to EBZ, indicating that these proteins might play an important role in *P. pratensis* wax secretion. It has been commonly proposed that ABC transporter G family member and LTPs were involved in wax trafficking from the endoplasmic reticulum to the plasma membrane and wax extracellular transportation, respectively [19, 36, 37].

The wax metabolic pathway is regulated by various transcription factors. One of these to be identified and characterized for its role in cuticle metabolism was SHN1/WIN1, a member of the *A. thaliana* ethylene-responsive AP2-domain transcription factor super family [38–40]. In this study, putative ethylene-responsive transcription factor, including *WIN1* and *WR11*, were expressed mainly in the NEZm, which might activate wax deposition in younger tissue. The expression of *WIN1* was also related to defense response of plant. The overexpression of *SISHN3*, an ortholog of the *Arabidopsis* *WIN/SHN3*, increased tomato resistance to *Botrytis cinerea*, whereas *SISHN3*-RNAi plants were more sensitive to *B. cinerea* compared to wild-type plants [41]. The altered defense responses in *SISHN3*-overexpressing plants were correlated with the cuticle permeability and the activation of pathogenesis-related gene *PR1a* and AOS [41]. Several defense response related genes, including *PR1*, were more expressed in NEZm compared to EBZ in this study, indicating that wax deposition might be linked with defense response. Garbay et al. [2007] reported that the amount of *PR1*-mRNA was not directly correlated to the amount of leaf epicuticular wax, but was rather correlated to the presence or absence of some particular lipid-constituents in the epicuticular wax layer [42]. Other transcription factors characterized for its role in cuticle metabolism belonged

to MYB family. It is reported that MYB96 could promote drought tolerance and trigger disease resistance response by inducing *Arabidopsis* cuticular wax accumulation [43, 44]. The overexpression of *MYB30* in *Arabidopsis* stimulated the production of VLCFAs and waxes [45]. In this study, although no MYB96 and MYB30 were identified, fourteen unigenes annotated with Myb protein were found to have more expression in the NEZm compared to the EBZ, suggesting the potential role of MYB family in wax metabolism regulation.

Besides their primary role in stress response, cuticular waxes were also found to be involved in developmental processes, notably through tight connections with the epidermis morphology [46]. The more expression of epidermal patterning factor (EPF) in NEZm than that in EBZ supported the link between cuticular wax metabolism and the process of epidermal cell patterning (in Additional file 2: Table S4). In addition, putative bHLH transcription factors were also up-regulated in the NEZm (in Additional file 2: Table S5). This was in agreement with the function of bHLH, which was important in development or cell activity.

## Conclusions

In this study, a comparative transcriptome sequencing between leaf NEZm and EBZ in *P. pratensis* was performed using Illumina platform. In total, 77,707,414 clean reads were *de novo* assembled into 106,766 unigenes. All unigenes were then evaluated and functionally annotated by comparing with the existing protein databases, including NR, Swissprot, GO, COG, and KEGG database. Several candidate genes potentially involved in cuticular wax biosynthesis, transportation, regulation, development, and defense response, were identified. About 3087 SSR molecular markers were developed. The database will improve our understanding of the molecular mechanism of cuticular wax deposition in *P. pratensis* leaf and will provide the fundamental genetic resources in improving plant adaptation to abiotic and biotic stresses.

## Methods

### Plant material

*Poa pratensis* cv Nuglade plants were grown in pot (10 cm x 15 cm) filled with turfy soil in a growth chamber (15 °C /25 °C, RH 75 %). To make sure the plants grow well, 1/4 Hoagland solution were applied every 5 days. There were 70 pots with 15 plants in each pot. Forty 5 days after germination, most plants went into three leaf stage. According to the study of Richardson et al. (2007a), the leaf of barley was divided into three parts, elongation, non-elongation and emerged blade [23]. According to the protocol from Rymen et al. [47], for leaf three of *P. pratensis* in this study, the length of

emerged blade, non-elongation and elongation parts reached about 10 cm, 1.5 cm and 0.5 cm, respectively. Since it was difficult to separate elongation zone and non-elongation zone clearly for RNA extraction and cuticular wax analysis separately, non-elongation and elongation zone (mainly non-elongation zone) were mixed together as one sample (NEZm) and were used for later analysis as well as emerged blade zone (EBZ) in this study. About 1 cm samples were collected from NEZm and EBZ of leaf three, separately, immersed in liquid nitrogen immediately, and then stored at  $-80^{\circ}\text{C}$ . About 200 plants from 20 pots were collected and bulked into one sample, approximately 0.1 g. In total two bulk samples were collected for RNA extraction from NEZm and EBZ, separately, as two biological replicates.

#### Cuticular wax extraction and analysis

Samples from remaining 30 pots were used for wax extraction, with 10 pots as one replicate. The samples from NEZm and EBZ were separately dipped in 50 ml of chloroform containing 0.25  $\mu\text{g}$  of hexadecane (Sigma) as internal standard for 30s at room temperature. The wax extractions were dried under nitrogen stream, derivatized with 100  $\mu\text{l}$  of BSTFA (N,O-Bis (trimethylsilyl) trifluoroacetamide) for 1 h at  $80^{\circ}\text{C}$ , and the surplus BSTFA was evaporated under nitrogen. The extract was re-dissolved in 1 mL of hexane for wax analysis using GCMS-2010 (Shimadzu Technologies Co, Japan) equipped with a flame ionization detector (FID) as described by Kim et al. (2007) [48]. Peaks were assigned by comparing of their mass spectra with the mass spectral library, GCMS solution Software (Shimadzu, Japan). Amounts of cuticular waxes were expressed in  $\mu\text{g/g}$ .

#### Scanning electron microscopy (SEM) analysis

Samples from NEZm and EBZ were collected, air dried, and then affixed to an aluminum stub with double sided adhesive tape. Stub was coated with gold and placed in the low-vacuum, variable-pressure chamber of the Hitachi S3500 Scanning electron microscopy and photographed with a digital camera with a digital camera at approximately 5000 magnification. Each sample replicated three times.

#### RNA extraction, cDNA preparation and transcriptome sequencing

Total RNA was extracted from the EBZ and NEZm of *P. pratensis* leaf using TransZol kit (TransGen, China). DNA contamination was removed with RNase-free DNase I (Takara, China). RNA quality and quantity were assessed by absorption at 260 nm/280 nm, and gel electrophoresis. Briefly 2.5  $\mu\text{g}$  of total RNA was enriched for Poly-A using NEBNext Poly (A) mRNA Magnetic

Isolation Module (NEB, E7490). Transcriptome library for sequencing was constructed according to NEBNext mRNA Library Prep Master Mix Set for Illumina (NEB, E6110) and NEBNext Multiplex Oligos for Illumina (NEB, E7500). The prepared library was quantified using Library Quantification Kit-Illumina GA Universal (Kapa, KK4824) and validated for quality by running 1.8 % agarose gel electrophoresis. The library products were sequenced via Illumina HiSeq™2500 sequencer.

#### Sequence data processing and *de novo* assembly

The raw reads generated by HiSeq2500 were cleaned by removing adaptor and primer sequences, reads in which the percentage of unknown bases (N) is greater than 5 % and low quality reads in which the percentage of the low quality bases was more than 20 %. Trinity [28] was used in *de novo* sequence assembly. First, Trinity combined the reads with a certain overlap length to form longer fragments, which were called contigs. Next, these reads were mapped back to contigs; with paired-end reads, Trinity was able to detect contigs from the same transcript and determine the distances between these contigs. Finally, Trinity connected these contigs into sequences that could not be extended on their end. Such sequences were defined as unigenes. Sequence saturation and distribution of reads on reference genes were analyzed to assess the overall sequencing quality. The ORFs were identified as the nucleotide sequence or as the protein translation provided by the “Getorf” software (<http://emboss.sourceforge.net/apps/cvs/emboss/apps/getorf.html>). The longest ORF was extracted for each unigene. The gene expression level was calculated using the RPKM method [49].

#### Sequence annotation and functional characterization

The assembled sequences were annotated using BLASTX program against NCBI database and all unigenes were utilized for homology searches against protein databases such as NCBI Nr (<http://www.ncbi.nlm.nih.gov/>) and Swissprot (<http://www.expasy.ch/sprot/>). To further annotate the unigenes in this research, the Blast2GO program was used to get GO annotation according to molecular function, biological process and cellular component ontologies (<http://www.geneontology.org/>). Each annotated sequence may have more than one GO term, either assigned in the different GO categories or in the same category. Secondary metabolic Pathway assignments were performed according to the KEGG pathway database [50]. The unigenes sequences were also aligned to the COG database [51] to predict and classify functions. DESeq [31] was used to identify differentially expressed genes (DEGs) between the NEZm and EBZ. FDR (false discovery rate)  $\leq 0.001$  and Fold Change  $\geq 8$  were used as the threshold to judge the significance of gene expression difference. The

annotation information from all the unigenes was extracted for the DEGs.

### Simple sequence repeats (SSRs) identification

All the unigenes with length more than 1 kb were used in a microsatellite program (MISA) (<http://pgrc.ipk-gatersleben.de/misa/misa/>) for identification of potent SSR motifs. We searched for microsatellites from mono-nucleotide to hexa-nucleotide. Dinucleotide repeats of more than six times and tri-, tetra, penta and hexa-nucleotide repeats of more than five times were considered as the search criteria for SSRs in MISA script. Both perfect (i.e. contain a single repeat motif) and compound repeats (i.e. composed of two or more motifs separated by 100 bases) were identified.

### Quantitative real-time reverse transcription-PCR (qRT-PCR)

Twelve up-regulated unigenes in the NEZm with potential roles in cuticular wax deposition were chosen for validation using qRT-PCR. Total RNA was extracted from the EBZ and the NEZm of *P. pratensis* leaf using TransZol kit (TransGen, China). DNase-treated RNA was used to synthesize first strand cDNA by using SuperScript II reverse transcriptase (Invitrogen, China). The gene names and primers used for qRT-PCR are listed in Additional file 2: Table S6. The quantitative reaction was performed on the CFX96 Real-Time PCR Detection System (Bio-Rad) using the SYBR Premix Ex TaqII (Takara, China). qRT-PCRs were performed as follows: 95 °C for 30 s, 40 cycles of 95 °C 5 s, 59 °C 30 s, and 72 °C 15 s. CFX Manager software (Bio-Rad) was used for data analysis. Expression levels of the selected unigenes were normalized to that of Elongation factor 1 (eEF1-a), an internal reference gene [52]. The relative expression levels of target genes were calculated with the  $2^{-\Delta\Delta C_t}$  method [53]. All the experiments were repeated using three biological and three technical replicates and the data were analyzed statistically.

### Availability of data and material

The data sets supporting the results of this article are included within the article and its additional files.

### Additional files

**Additional file 1: Figure S1.** A, Randomness test of cDNA fragments; B, Sequencing saturation analysis. T1 and T2 represent NEZm; T3 and T4 represent EBZ. (PDF 290 kb)

**Additional file 2: Table S1.** List of DEG unigenes between NEZm and EBZ related to KEGG pathway. **Table S2.** Candidate genes involved in fatty acids and cuticular wax biosynthesis in *P. pratensis* leaf. **Table S3.** List of unigenes involved in defense response that is differentially expressed in EBZ versus NEZm. **Table S4.** Putative epidermal patterning factor (EPF) differentially expressed between NEZm and EBZ. **Table S5.** Putative bHLH transcription factor differentially expressed between NEZm

and EBZ. **Table S6.** The primers used in quantitative real-time PCR analysis. (XLS 118 kb)

### Abbreviations

ABC: ATP-binding cassette; COG: Clusters of Orthologous Groups; EBZ: emerged blade zone; ECR: enoyl-CoA reductase; FAE: fatty acid elongase; FAR: fatty acyl-CoA reductase; GO: Gene ontology; HCD:  $\beta$ -hydroxyacyl-CoA dehydratase; KCR:  $\beta$ -ketoacyl-CoA reductase; KCS:  $\beta$ -ketoacyl-CoA synthase; KEGG: Kyoto Encyclopedia of Genes and Genomes; MISA: MicroSAteLLite; NEZm: mixed sample of non-elongation zone and elongation zone; NR: non-redundant; qRT-PCR: quantitative real-time polymerase chain reaction; RSEM: RNA-Seq by Expectation Maximization; SSRs: simple sequence repeats; VLCFAs: very long chain fatty acids.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

Yu Ni contributed to data analysis and manuscript preparation. Yanjun Guo contributed to the experimental design and manuscript revision. Na Guo and Qiuling Zhao contributed to experiment management, RNA extraction and cuticular wax analysis. All authors have read and approved the final manuscript.

### Acknowledgements

We would like to thank Jianhua Gao and Yuji He of Southwest University for help with RNA extraction and cuticular wax extraction.

### Open access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

### Funding

The work was supported by National Natural Science Foundation of China (31270450), the National Key Basic Research Program of China (2014CB138806) and the Fundamental Research Funds for the Central Universities (XDJK2014B037).

Received: 7 January 2016 Accepted: 22 April 2016

Published online: 29 April 2016

### References

- Shepherd T, Griffiths DW. The effects of stress on plant cuticular waxes. *New Phytol.* 2006;171:469–99.
- Gonzalez A, Ayerbe L. Effect of terminal water stress on leaf epicuticular wax load, residual transpiration and grain yield in barley. *Euphytica.* 2010;172:341–9.
- Koch K, Hartmann KD, Schreiber L, Barthlott W, Neinhuis C. Influences of air humidity during the cultivation of plants on wax chemical composition, morphology and leaf surface wettability. *Environ Exp Bot.* 2006;56:1–9.
- Ni Y, Xia R, Li J. Changes of epicuticular wax induced by enhanced UV-B radiation impact on gas exchange in *Brassica napus*. *Acta Physiol Plant.* 2014;36:2481–90.
- Uppalapati SR, Ishiga Y, Doraiswamy V, Bedair M, Mittal S, Chen JH, et al. Loss of abaxial leaf epicuticular wax in *Medicago truncatula* irg1/palm1 mutants results in reduced spore differentiation of antheracnose and nonhost rust pathogens. *Plant Cell.* 2012;24:353–70.
- Hoad SP, Jeffree CE, Grace J. Effects of wind and abrasion on cuticular integrity in *Fagus sylvatica* and consequences for transfer of pollutants through leaf surfaces. *Agri Ecos Environ.* 1992;42:275–89.
- Riederer M, Müller C. *Biology of the Plant Cuticle. The fine structure of the plant cuticle.* Oxford: Blackwell Publishing Ltd; 2006.
- Gefen E, Talal S, Brendzel O, Dror A, Fishman A. Variation in quantity and composition of cuticular hydrocarbons in the scorpion *Buthus occitanus*

- (Buthidae) in response to acute exposure to desiccation stress. *Comp Biochem Physiol A Mol Integr Physiol.* 2015;182:58–63.
9. Yeats TH, Rose JKC. The formation and function of plant cuticles. *Plant Physiol.* 2013;163:5–20.
  10. Post-Beittenmiller D. Biochemistry and molecular biology of wax production in plants. *Annu Rev Plant Physiol Plant Mol Biol.* 1996;47:405–30.
  11. Koorneef M, Hanhart CJ, Thiel F. A genetic and phenotypic description of eceriferum (*cer*) mutants in *Arabidopsis thaliana*. *J Hered.* 1989;80:118–22.
  12. Fiebig A, Mayfield JA, Miley NL, Chau S, Fischer RL, Preuss D. Alterations in CER6, a gene identical to CUT1, differentially affect long chain lipid content on the surface of pollen and stems. *Plant Cell.* 2000;12:2001–8.
  13. Zheng H, Rowland O, Kunst L. Disruptions of the *Arabidopsis* enoyl-CoA reductase gene reveal an essential role for very-long-chain fatty acid synthesis in cell expansion during plant morphogenesis. *Plant Cell.* 2005;17:1467–81.
  14. Lü S, Song T, Kosma DK, Parsons EP, Rowland O, Jenks MA. *Arabidopsis* CER8 encodes LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. *Plant J.* 2009;59:553–64.
  15. Bernard A, Domergue F, Pascal S, Jetter R, Renne C, Faure JD, et al. Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *Plant Cell.* 2012;24:3106–18.
  16. Greer S, Wen M, Bird D, Wu X, Samuels L, Kunst L, Jetter R. The cytochrome P450 enzyme CYP96A15 is the midchain alkane hydroxylase responsible for formation of secondary alcohols and ketones in stem cuticular wax of *Arabidopsis*. *Plant Physiol.* 2007;145:653–67.
  17. Rowland O, Zheng H, Hepworth SR, Lam P, Jetter R, Kunst L. CER4 encodes an alcohol-forming fatty acyl-coA reductase involved in cuticular wax production in *Arabidopsis*. *Plant Physiol.* 2006;142:866–77.
  18. Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, et al. Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in *Arabidopsis*. *Plant Physiol.* 2008;148:97–107.
  19. Pighin JA, Zheng H, Balakshin LJ, Goodman IP, Western TL, Jetter R, et al. Plant cuticular lipid export requires an ABC transporter. *Science.* 2004;306:702–4.
  20. Bird D, Beisson F, Brigham A, Shin J, Greer S, Jetter R, et al. Characterization of *Arabidopsis* ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. *Plant J.* 2007;52:485–98.
  21. Costaglioli P, Joubes J, Garcia C, Stef M, Arveiler B, Lessire R, et al. Profiling candidate genes involved in wax biosynthesis in *Arabidopsis thaliana* by microarray analysis. *Biochim Biophys Acta.* 2005;1734:247–58.
  22. Rhee Y, Hlousek-Radojicic A, Ponsamuel J, Liu DH, Post-Beittenmiller D. Epicuticular wax accumulation and fatty acid elongation activities are induced during leaf development of leeks. *Plant Physiol.* 1998;116:901–11.
  23. Richardson A, Wojciechowski T, Franke R, Schreiber L, Kerstiens G, Jarvis M, et al. Cuticular permeance in relation to wax and cutin development along the growing barley (*Hordeum vulgare*) leaf. *Planta.* 2007a; 225: 1471–1481.
  24. Richardson A, Boscarri A, Schreiber L, Kerstiens G, Jarvis M, Herzyk P, et al. Cloning and expression analysis of candidate genes involved in wax deposition along the growing barley (*Hordeum vulgare*) leaf. *Planta.* 2007b; 226: 1459–1473.
  25. Reader RJ, Wilson SD, Belcher JW, Wisheu I, Keddy PA, Tilman D, et al. Plant competition in relation to neighbor biomass - an intercontinental study with *Poa pratensis*. *Ecology.* 1994;75:1753–60.
  26. Pertierra LR, Lara F, Benayas J, Hughes KA. *Poa pratensis* L, current status of the longest-established non-native vascular plant in the Antarctic. *Polar Biol.* 2013;36:1473–81.
  27. Pilon JJ, Lambers H, Baas W, Tosserams M, Rozema J, Atkin OK. Leaf waxes of slow-growing alpine and fast-growing lowland *Poa* species: inherent differences and responses to UV-B radiation. *Phytochemistry.* 1999;50:571–80.
  28. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29:644–52.
  29. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10:R25.
  30. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12:323.
  31. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010;11:1–12.
  32. Millar AA, Kunst L. Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. *Plant J.* 1997;12:121–31.
  33. Richardson A, Franke R, Kerstiens G, Jarvis M, Schreiber L, Fricke W. Cuticular wax deposition in growing barley (*Hordeum vulgare*) leaves commences in relation to the point of emergence of epidermal cells from the sheaths of older leaves. *Planta.* 2005;222:472–83.
  34. Abe K, Ohno Y, Sassa T, Taguchi R, Çalıřkan M, Ober C, et al. Mutation for nonsyndromic mental retardation in the trans-2-Enoyl-CoA Reductase TER gene involved in fatty acid elongation impairs the enzyme activity and stability, leading to change in sphingolipid profile. *J Biol Chem.* 2013;288: 36741–9.
  35. Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, et al. Plant ABC proteins – a unified nomenclature and updated inventory. *Trends Plant Sci.* 2008;13:151–9.
  36. Panikashvili D, Savaldi-Goldstein S, Mandel T, Yifhar T, Franke RB, Höfer R, et al. The *Arabidopsis* DESPERADO/AtWBC11 transporter is required for cutin and wax secretion. *Plant Physiol.* 2007;145:1345–60.
  37. Kader JC. Lipid-transfer proteins in plants. *Annu Rev Plant Physiol Plant Mol Biol.* 1996;47:627–54.
  38. Aharoni A, Dixit S, Jetter R, Thoenes E, Arkel G, Pereira A. The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in *Arabidopsis* W inside a box sign. *Plant Cell.* 2004;16:2463–80.
  39. Broun P, Poindexter P, Osborne E, Jiang CZ, Riechmann JL. WIN1, a transcriptional activator of epidermal wax accumulation in *Arabidopsis*. *PNAS, USA.* 2004;101:4706–11.
  40. Kannangara R, Branigan C, Liu Y, Penfield T, Rao V, Mouille G, et al. The transcription factor WIN1/SHN1 regulates cutin biosynthesis in *Arabidopsis thaliana*. *Plant Cell.* 2007;19:1278–94.
  41. Buxdorf K, Rubinsky G, Barda O, Burdman S, Aharoni A, Levy M. The transcription factor SISHINE3 modulates defense responses in tomato plants. *Plant Mol Biol.* 2014;84:37–47.
  42. Garbay B, Tautu MT, Costaglioli P. Low level of pathogenesis-related protein 1 mRNA expression in 15-day-old *Arabidopsis cer6-2* and *cer2* eceriferum mutants. *Plant Sci.* 2007;172:299–305.
  43. Seo PJ, Lee SB, Suh MC, Park MJ, Go YS, Park CM. The MYB96 transcription factor regulates cuticular wax biosynthesis under drought conditions in *Arabidopsis*. *Plant Cell.* 2011;23:1138–52.
  44. Seo PJ, Park CM. MYB96-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in *Arabidopsis*. *New Phytol.* 2010;186:471–83.
  45. Raffaele S, Vaillieu F, Léger A, Joubès J, Miersch O, Huard C, et al. A MYB transcription factor regulates very-long-chain fatty acid biosynthesis for activation of the hypersensitive cell death response in *Arabidopsis*. *Plant Cell.* 2008;20:752–67.
  46. Javelle M, Vernoud V, Rogowsky PM, Ingram GC. Epidermis: the formation and functions of a fundamental plant tissue. *New Phytol.* 2011;189:17–39.
  47. Rymen B, Coppens F, Dhondt S, Fiorani F, Beemster GTS. Kinematic analysis of cell division and expansion. *Methods Mol Biol.* 2010;655:203–27.
  48. Kim KS, Park SH, Jenks MA. Changes in leaf cuticular waxes of sesame (*Sesamum indicum* L.) plants exposed to water deficit. *J Plant Physiol.* 2007; 164:1134–43.
  49. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5:621–8.
  50. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 2004;32 suppl 1:D277–80.
  51. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000;28:33–6.
  52. Dombrowski JE, Martin RC. Evaluation of reference genes for quantitative RT-PCR in *Lolium temulentum* under abiotic stress. *Plant Sci.* 2009;176:390–6.
  53. Wong ML, Medrano JF. Real-time PCR for mRNA quantitation. *Biotechniques.* 2005;39:75–85.