


## Article

# Single-Molecule Long-Read Sequencing of Purslane (*Portulaca oleracea*) and Differential Gene Expression Related with Biosynthesis of Unsaturated Fatty Acids

Hongmei Du <sup>1</sup>, Shah Zaman <sup>2</sup> , Shuiqingqing Hu <sup>1</sup> and Shengquan Che <sup>1,\*</sup>

<sup>1</sup> School of Design, Shanghai Jiao Tong University, Shanghai 200240, China; hmdu@sjtu.edu.cn (H.D.); hushuiqingqing@sjtu.edu.cn (S.H.)

<sup>2</sup> School of Agricultural and Biology, Shanghai Jiao Tong University, Shanghai 200240, China; Shah\_Zaman@sjtu.edu.cn

\* Correspondence: chsq@sjtu.edu.cn; Tel.: +86-21-3420-5730

**Abstract:** This study aimed to obtain the full-length transcriptome of purslane (*Portulaca oleracea*); assorted plant samples were used for single-molecule real-time (SMRT) sequencing. Based on SMRT, functional annotation of transcripts, transcript factors (TFs) analysis, simple sequence repeat analysis and long non-coding RNAs (LncRNAs) prediction were accomplished. Total 15.33-GB reads were produced; with 9,350,222 subreads and the average length of subreads, 1640 bp was counted. With 99.99% accuracy, after clustering, 132,536 transcripts and 78,559 genes were detected. All unique SMART transcripts were annotated in seven functional databases. 4180 TFs (including transcript regulators) and 7289 LncRNAs were predicted. The results of RNA-seq were confirmed with qRT-PCR analysis. Illumina sequencing of leaves and roots of two purslane genotypes was carried out. Amounts of differential expression genes and related KEGG pathways were found. The expression profiles of related genes in the biosynthesis of unsaturated fatty acids pathway in leaves and roots of two genotypes of purslane were analyzed. Differential expression of genes in this pathway built the foundation of  $\omega$ -3 fatty acid accumulation in different organs and genotypes of purslane. The aforementioned results provide sequence information and may be a valuable resource for whole-genome sequencing of purslane in the future.

**Keywords:** purslane; RNA-seq; gene; unsaturated fatty acid



**Citation:** Du, H.; Zaman, S.; Hu, S.; Che, S. Single-Molecule Long-Read Sequencing of Purslane (*Portulaca oleracea*) and Differential Gene Expression Related with Biosynthesis of Unsaturated Fatty Acids. *Plants* **2021**, *10*, 655. <https://doi.org/10.3390/plants10040655>

Academic Editor: Vagner Benedito

Received: 29 January 2021

Accepted: 24 March 2021

Published: 30 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Purslane (*Portulaca oleracea*) first came from India and Iran and has spread around the world. It is a warm-climate, juicy annual plant spread around the world, belonging to the Portulacaceae family. Purslane is one of the most abundant terrestrial vegetables in spite of its genetic assortment [1–3] and generally recognized as purslane (the USA and Australia), Ma-Chi-Xian (China) [4] and kurfa in (Pakistan) [5,6]. It has been used as an important traditional medicinal plant and itemized by the World Health Organization as “Global Panacea” [7].

Apart from medicinal and nutritive value, purslane is a high-quality halophyte [8]. Some results have been reported about its salt tolerance mechanism [9–12]. Since, the development of high-throughput sequencing technology, transcripts sequencing has become an important means to study gene expression regulation, after the whole-genome sequencing of human were completed in 2004 [13]. However, due to the read length limitation of the second-generation sequencing in different organisms [14–18], the full-length transcript obtained by splicing is not complete. The third-generation sequencing technology signified by Pacific Biosciences (PacBio, Menlo Park, CA, USA) effectively overcome this problem [19,20]. Single-molecule real-time (SMRT) sequencing can directly obtain full-length splice isoforms without the need for assembly [21,22], improve the draft genome annotation in species with reference genome and facilitate comparative transcriptome studies

and gene functional annotation [22]. Up to date, this technology has been successfully utilized in some species, such as perennial ryegrass (*Lolium perenne*) [22], *Rhododendron lapponicum* [23], strawberry (*Fragaria × ananassa*) [24], *Gnetum luofuense* [25], and maize (*Zea mays*) [21]. However, up to date, no researcher studied the full-length transcript of purslane using single-molecule long-read sequencing. In this study, SMRT sequencing was achieved in purslane. After detecting transcripts, we completed functional annotations of transcripts, transcript factors (TFs) and simple sequence repeat (SSR) analysis, long non-coding RNAs (LncRNA) prediction. This study established a high-quality reference transcriptome for purslane, which provides valuable resources for further investigation of related molecular mechanisms, especially biosynthesis of unsaturated fatty acids pathway in purslane.

## 2. Results

### 2.1. Single-Molecule Real-Time Sequencing of Purslane

Leaves and roots from “Pakistan local” (“PL”-North American origin) and a wild variety “Liaoning, China local” (“LCL”) were used for RNA extraction and cDNA library construction. After removing adaptor sequences, low-quality sequences, and short sequences less than 50 bp, a total of 9,350,222 subread (15.33 Gb) were obtained, through normal subread length of 1640 bp, and N50, of 3093 bp (Figure 1A). After self-correction of subread sequences (with min passes = 2, min predicted accuracy = 0.8), a total of 375,102 circular consensus sequence (CCS) were obtained. After sequencing, 259,265 full-length and 254,692 full-length non-chimeric (FLNC) picks were identified. The average FLNC read length was 2808 bp (Figure 1B). The FLNC sequences of the same transcript were clustered using the iterative isoform-clustering (ICE) algorithm, and 132,536 consensus reads were obtained after clustering (Figure 1C).

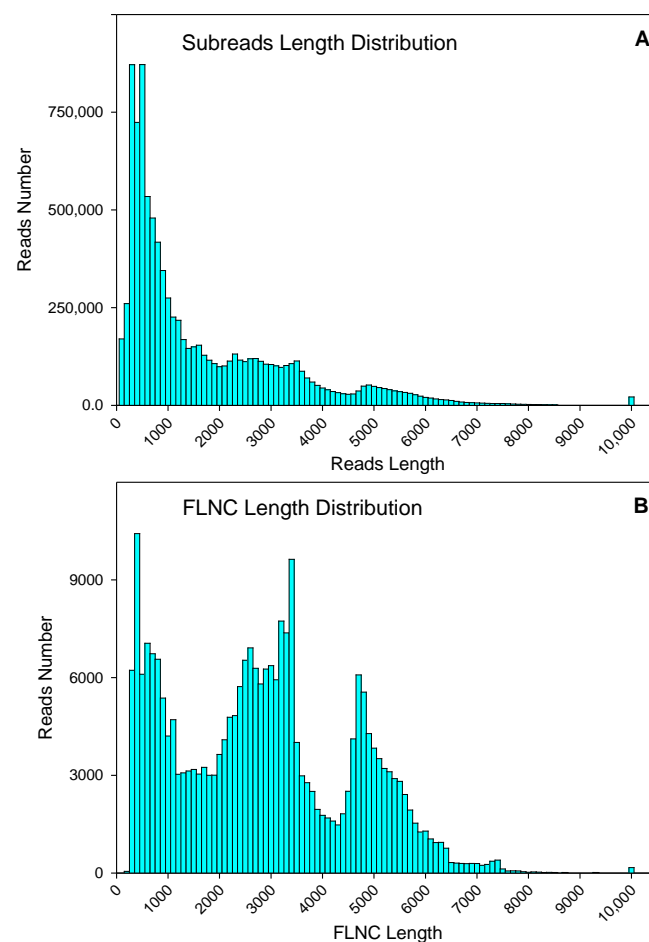
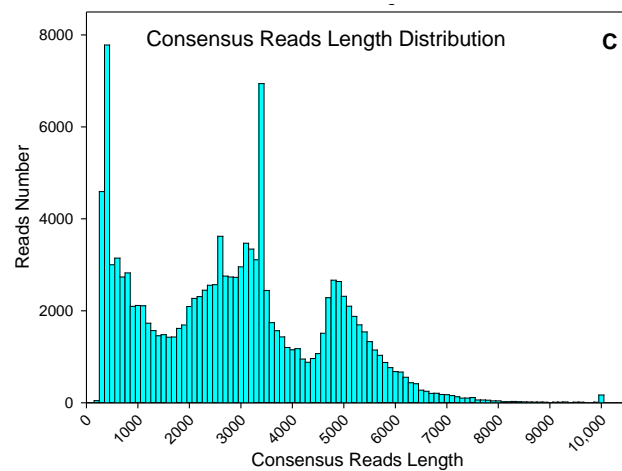
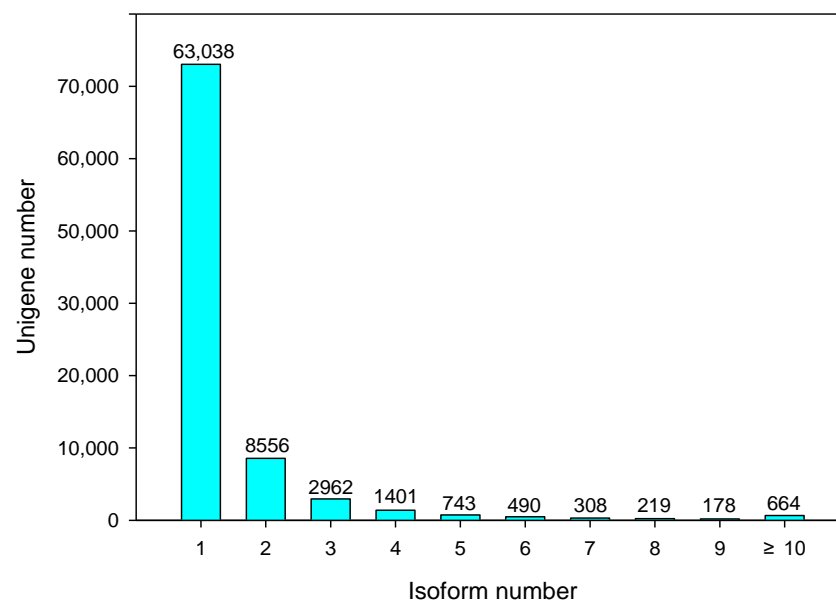


Figure 1. Cont.



**Figure 1.** Numbers and length distributions of 9,350,222 subreads (A), 254,692 full-length non-chimeric (FLNC) sequences (B), and 132,536 consensus reads (C) with PacBio single-molecule real-time sequencing method of purslane (*Portulaca oleracea*).

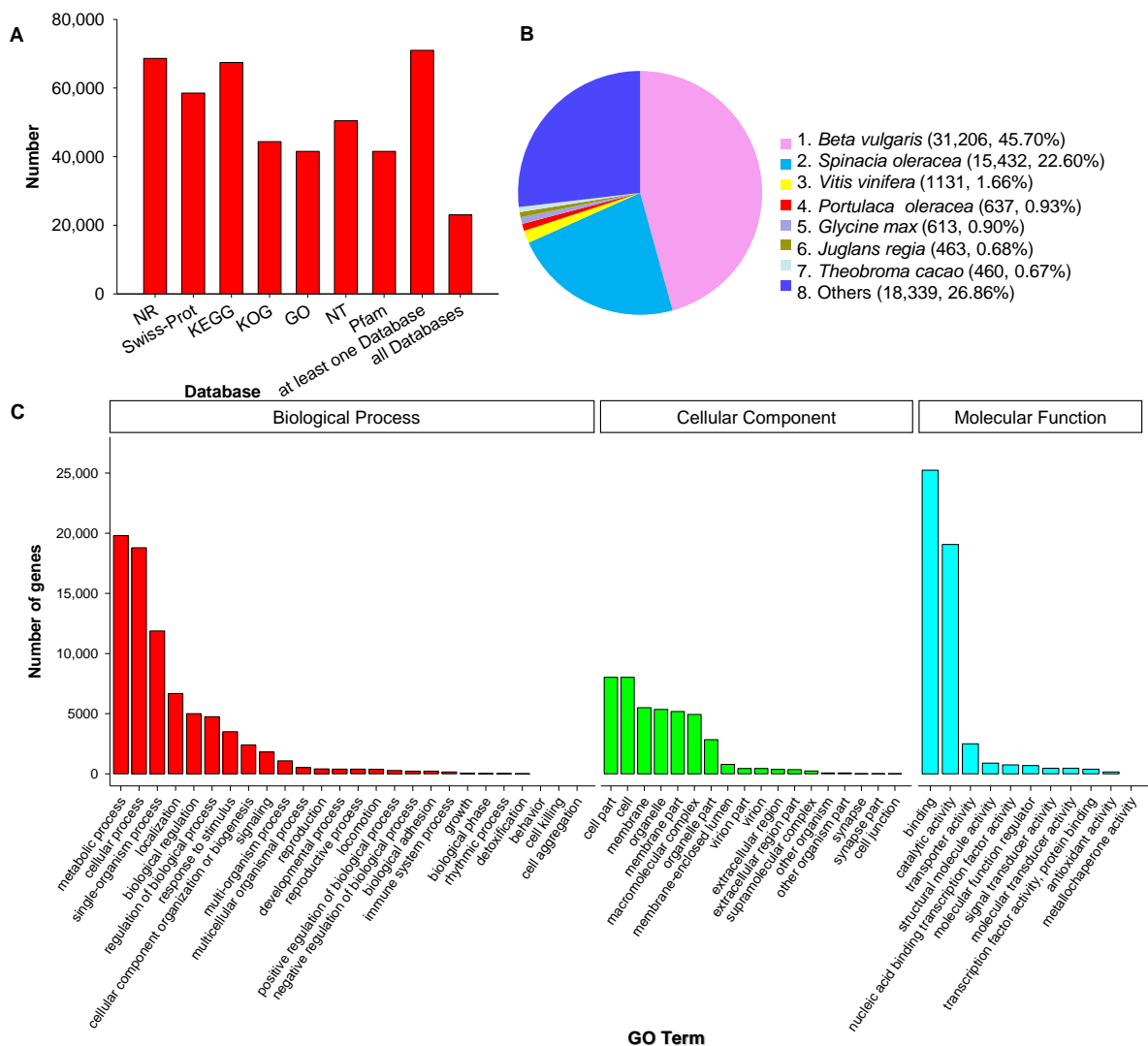
The number of genes with no isoform remained 53,977, and 78,559 transcriptions were found between 1 and 10 isoforms (Figure 2). The transcript length extended from 170 bp to 14,287 bp, with an average length of 3061 bp.



**Figure 2.** Number of isoforms in purslane (*Portulaca oleracea*).

## 2.2. Functional Annotation of Full-Length Transcripts

All 78,559 unique SMRT transcripts were functionally annotated by searching seven data storage, including Gene Ontology (GO), eukaryotic ortholog groups (KOG), protein family (Pfam), NCBI nucleotide sequences (NT), NCBI nonredundant protein sequences (NR), Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG). 90.39% transcript, a total of 71,008, was annotated at one database at least, and 23,013 transcripts were annotated at all seven databases (Figure 3A). We also identified matches to our unique transcripts in clusters of orthologous groups of proteins (COG) (44,376, 56.49%), Pfam database (41,535, 52.87%) and Swiss-Port (58,535, 74.51%). The functional annotations of all 78,559 unique transcripts were listed in Supplementary Table S1.

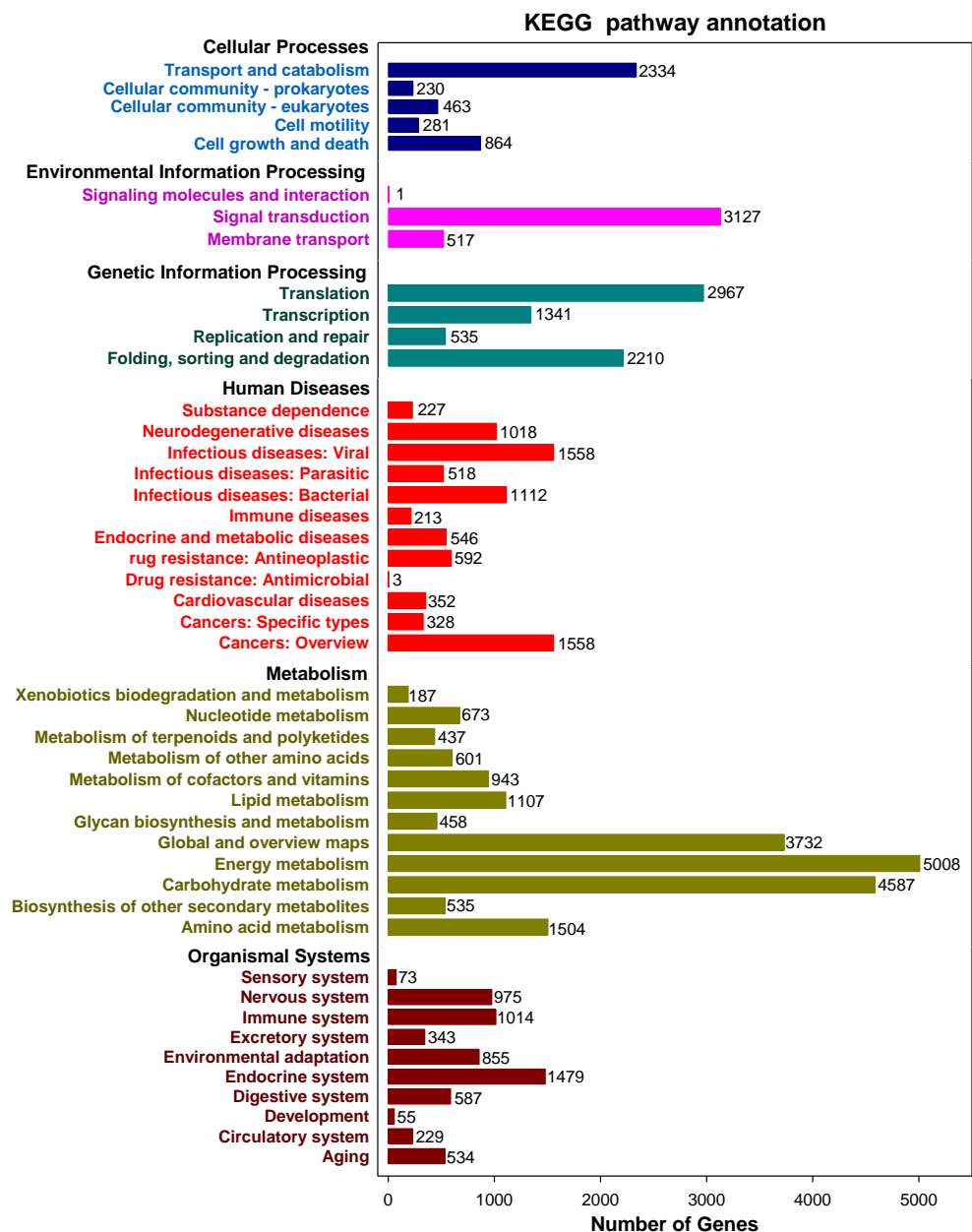


**Figure 3.** Function annotation of corrected isoforms in seven databases (A) (NR, NCBI Nonredundant Protein Database; KEGG, Kyoto Encyclopedia of Genes and Genomes; KOG, cluster of eukaryotic ortholog groups of proteins; GO, Gene Ontology; NT, NCBI nucleotide sequences; Pfam, Protein family), homologous species distribution diagram of transcripts in NCBI nonredundant protein sequences (NR) (B), Gene Ontology (GO) classification of unique transcripts (C) of purslane (*Portulaca oleracea*).

We compared the transcript sequences to NR by homologous species analysis. 68,630 genes were annotated, among them, *Beta vulgaris* (31,206; 45.70%), *Spinacia oleracea* (15,432; 22.60%), *Vitis vinifera* (1131; 1.66%), and *P. oleracea* (637; 0.93%) were the top four species of transcripts distributed (Figure 3B). GO analysis demonstrated that 41,535 unique genes were enriched significantly in three major categories: molecular function (MF), cellular component (CC), biological process (BP). In these three categories, the most abundant GO terms were cellular process and metabolic process in BP, catalytic activity and binding in MF, cell and cell part in CC (Figure 3C).

A total of 67,426 sequences were interpreted by the KEGG data storage and plotted to 366 operative catalogs in purslane. “Metabolism” was the largest transcript category. The first three transcripts-related pathways were carbon metabolism (3007, 4.46%), carbon fixation in photosynthetic organisms (2304, 3.42%) and pyruvate metabolism (2153, 3.19%) (Figure 4 and Supplementary Table S1). A large number of genes, especially interrelated in salt-tolerance and the fatty acid component of purslane, were annotated, such as oxidative phosphorylation (1073), plant hormone signal transduction (506), fatty acid biosynthesis

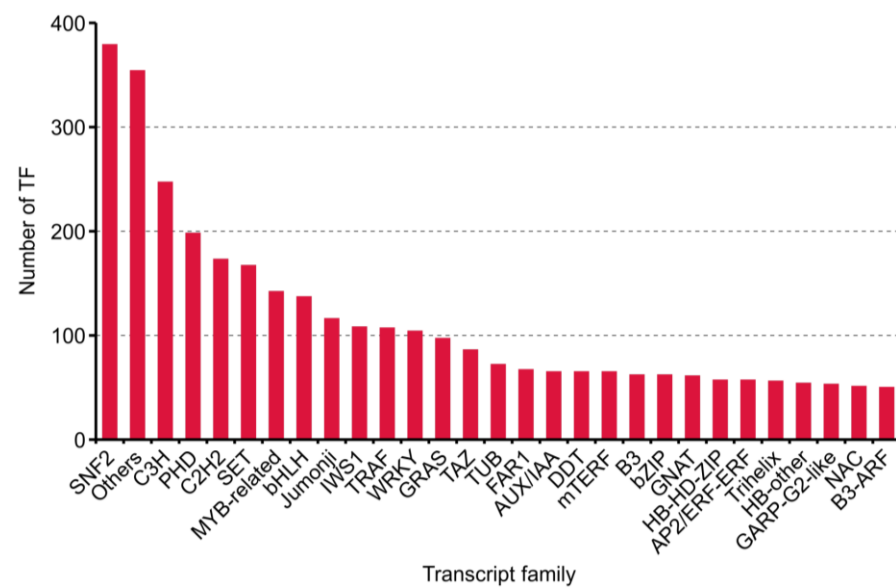
(246), biosynthesis of unsaturated fatty acids (94) and  $\alpha$ -linolenic acid metabolism (199) (Supplementary Table S1).



**Figure 4.** Kyoto Encyclopedia of Genes and Genomes (KEGG) annotated pathways and numbers of the gene in purslane (*Portulaca oleracea*).

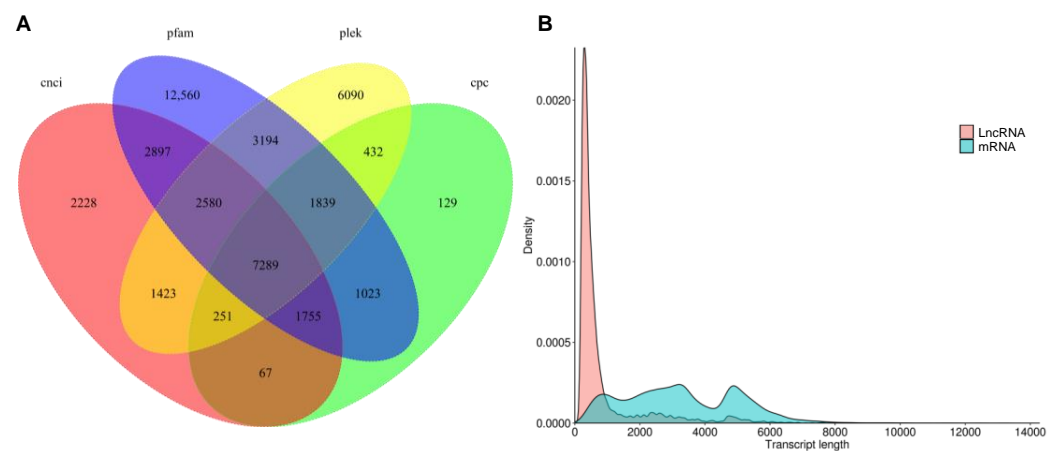
### 2.3. Results of Transcript Factors, Long Non-Coding RNAs and Simple Sequence Repeat

During plant growth and development, TFs and transcription regulation (TRs) acting as dominant characters. Four thousand one hundred eighty transcripts, including 2211 putative TF and 1969 TR from 86 families, were predicted with iTAK software [26] (Supplementary Table S2). The top 30 families annotated are shown in Figure 5.



**Figure 5.** Numbers and families of the top 30 transcript factors (TFs) predicted by iTAK software in purslane (*Portulaca oleracea*).

LncRNA were predicted by PLEK, Pfam, coding potential calculator (CPC) and coding-non-coding index (CNCI) (Figure 6A). Overall, 7289 LncRNA were predicted with a mean length of 848.84 bp with most LncRNAs length ranging from 300 bp to 1000 bp (Figure 6B and Supplementary Table S3).



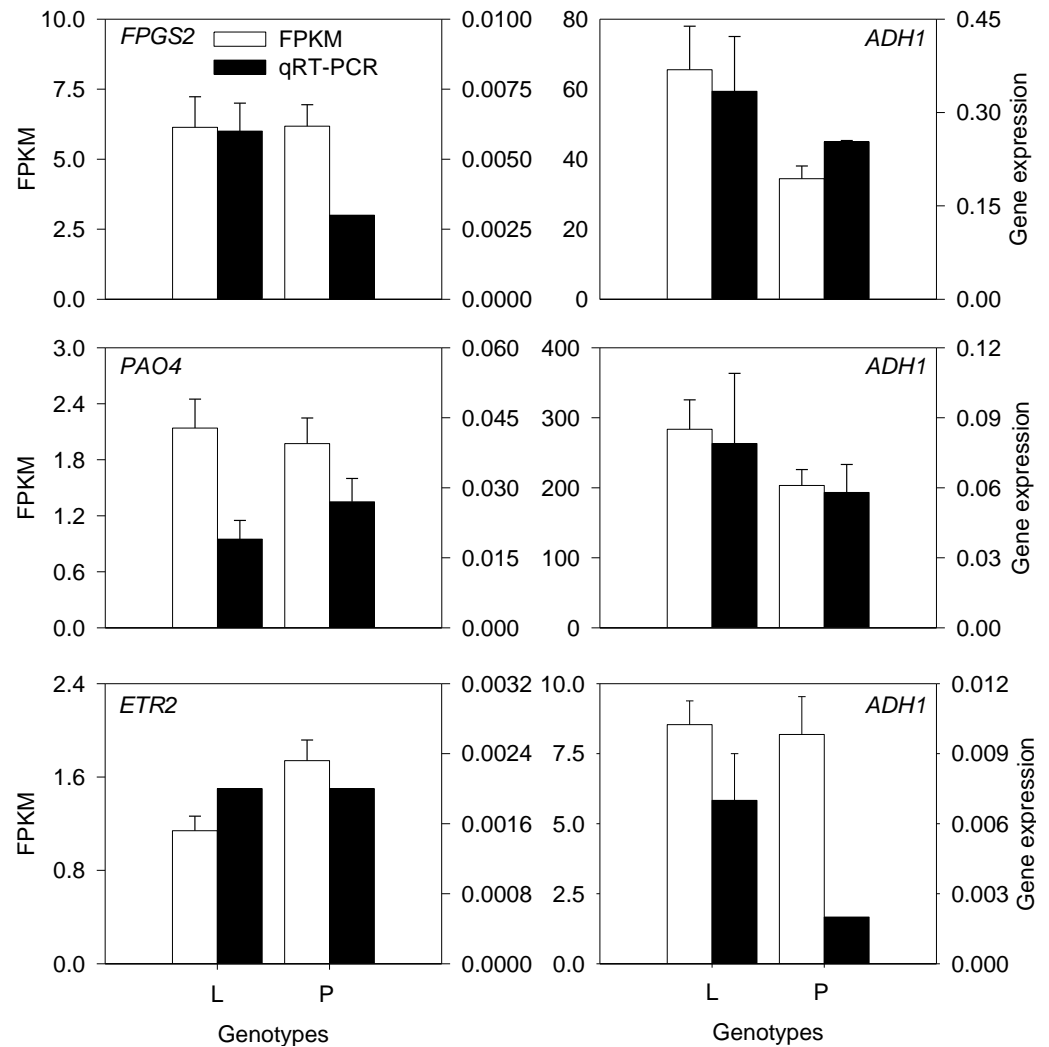
**Figure 6.** Identification of long non-coding RNAs (LncRNAs). (A). Venn diagram of LncRNAs predicted by PLEK, coding-non-coding index (CNCI), coding potential calculator (CPC) and protein family (Pfam) methods. (B). Density and length distributions of LncRNAs and mRNAs in purslane (*Portulaca oleracea*).

SSR is also known as short tandem repeats or microsatellite markers. A total of 58,622 sequences were subjected to SSR analysis. Most of the SSRs identified were one-nucleotide repeats (50.81%), followed by compound nucleotide repeats (17.11%), two-nucleotide repeats (15.20%), three-nucleotide repeats (14.76%), four-nucleotide repeats (1.16%), six-nucleotide repeats (0.73%), and five-nucleotide repeats (0.40%) (Table S4).

#### 2.4. qRT-PCR Validation of Selected Genes

To confirm the results of RNA-seq, we executed qRT-PCR with 6 genes in 4 KEGG pathways, including 1 *FPGS2* in folate biosynthesis, 1 *PAO4* in arginine and proline metabolism, 1 *ETR2* in plant hormone signal transduction and 3 *ADH1* in  $\alpha$ -linolenic

acid metabolism. Gene expression results with qRT-PCR verification showed a very similar tendency with Fragments per kilobase of transcript per million mapped reads (FPKM) values from sequencing at the same conditions (Figure 7), which indicated that the RNA-seq data were reliable.



**Figure 7.** Validation of the mRNA expression levels of a gene by RNA-seq and qRT-PCR analyses.

### 2.5. Illumina Sequencing Results and Differential Gene Expression in Leaves and Roots of Two Purslane Genotypes

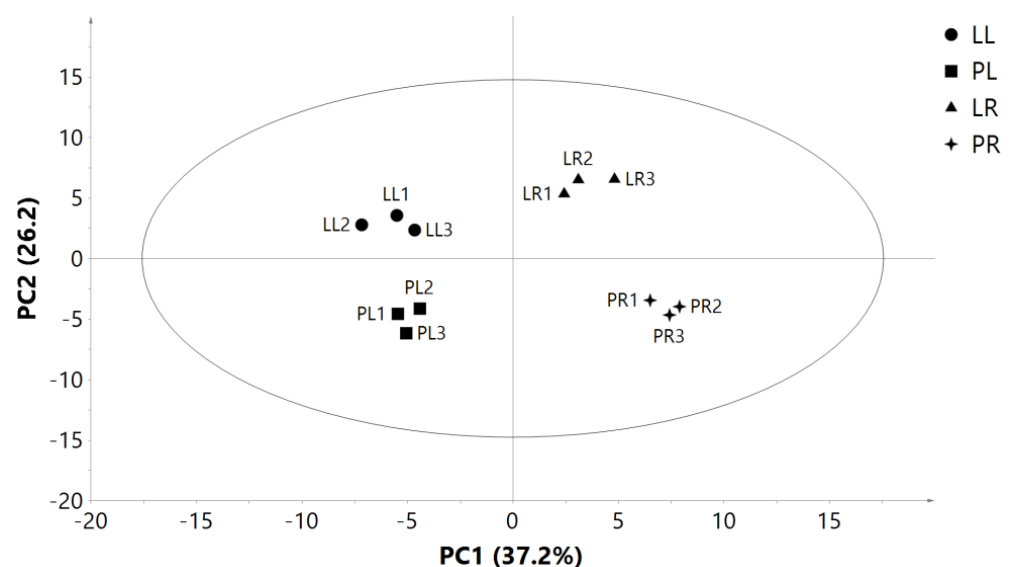
Illumina sequencing was performed on leaves or roots of two purslane genotypes, separately, with three independent biological replicates. Clean reads of leaf and root samples of two genotypes are shown in Table 1. These sequences were mapped to reference sequence (REF) of purslane for annotation of all unigenes. The mapping rate was 57.59%–76.84% for all samples in two purslane genotypes (Table 1). Compared with “LCL”, 9099 and 7408 differentially expressed genes (DEGs) were found upregulated, 9550 and 6802 DEGs downregulated in leaves and roots of “PL” with the criteria of absolute of  $\log_2$ -fold change > 0 and  $p$  value < 0.05 (Supplementary Table S5). Functions of these genes were related to 119 KEGG pathways (Supplementary Table S6).

**Table 1.** Statistics of annotation results for unigenes of two purslanes (*Portulaca oleracea*) genotypes.

Sample	Raw Reads	Clean Reads	Mapping Rate (%)
<b>Leaves</b>			
"LCL", 1	62,777,144	47,789,582	76.13
"LCL", 2	60,625,062	45,707,454	75.39
"LCL", 3	54,712,526	40,774,382	74.52
"PL", 1	54,782,840	41,921,110	76.52
"PL", 2	61,500,192	47,256,666	76.84
"PL", 3	59,127,304	44,841,760	75.84
<b>Roots</b>			
"LCL", 1	58,010,278	34,114,332	58.81%
"LCL", 2	56,153,952	34,729,076	61.85%
"LCL", 3	59,146,128	37,448,562	63.32%
"PL", 1	63,212,838	38,289,822	60.57%
"PL", 2	48,132,318	28,489,398	59.19%
"PL", 3	42,508,608	24,481,842	57.59%

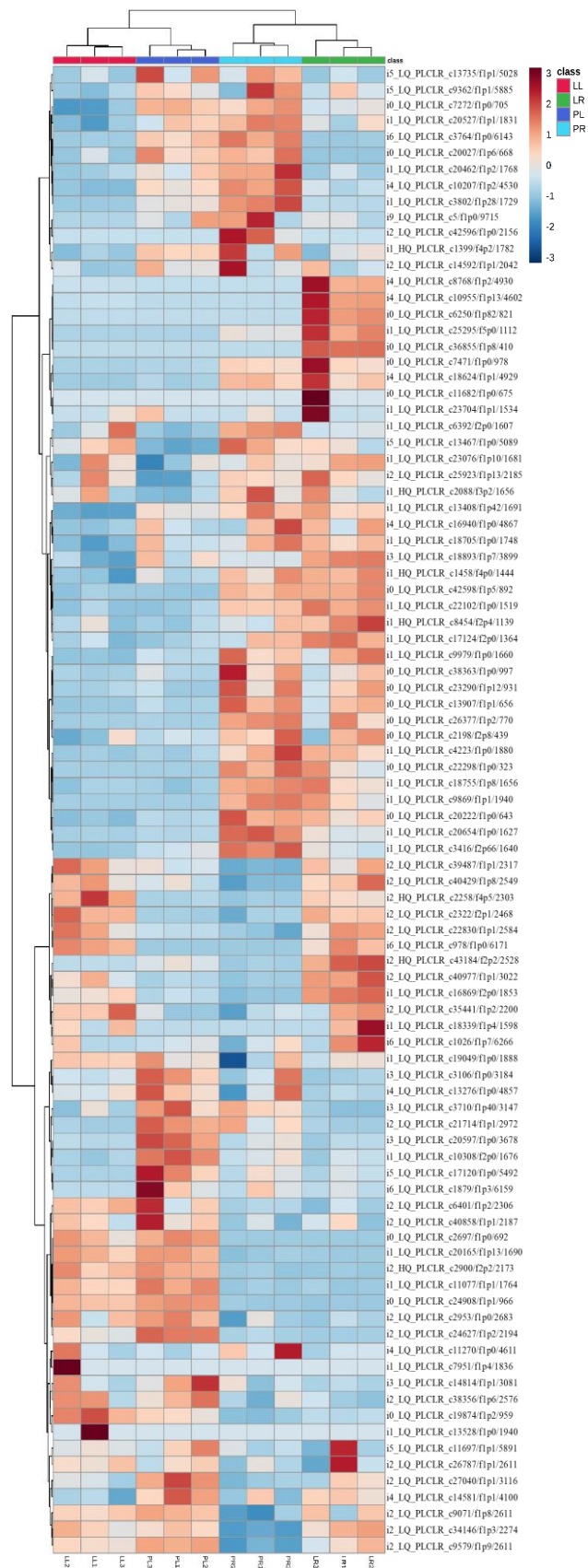
## 2.6. Gene Expression Analysis in the Pathway of Biosynthesis of Unsaturated Fatty Acids

Ninety-four genes were identified to participate in the pathway of "biosynthesis of unsaturated fatty acids" (Table S1). Principle component analysis (PCA) separated these genes expressed in leaves and roots of two purslane genotypes into different principal components. The first and second principal components accounted for 37.25% and 26.2% of the variation, which distinctly separated related gene expressions in leaves and roots of two genotypes (Figure 8). The differentially expressed genes between different organs and genotypes in a heat map (Figure 9) supported our results in PCA analysis. We identified 41 genes with significantly differential expression between "LCL" and "PL" in leaves, of which 30 genes were upregulated and 11 downregulated in "PL" compared with "LCL". Moreover, 32 genes had a significantly differential expression in roots, of which 13 genes upregulated and 19 downregulated in "PL" compared with "LCL" (Supplementary Table S7).



**Figure 8.** Principal component plot of gene expression participated in the pathway of "biosynthesis of unsaturated fatty acids" from leaves (L) and roots (R) of two purslanes (*Portulaca oleracea*) genotypes ("Pakistan" local ("PL") and "Liaoning, China" local ("LCL")). LL: leaves of "LCL"; PL: leaves of "PL"; LR: roots of "LCL"; PR: roots of "LCL".





**Figure 9.** A hierarchically clustered heat map showing the log<sub>10</sub> transformed gene expression values in the pathway of “biosynthesis of unsaturated fatty acids” from leaves (L) and roots (R) two purslane (*Portulaca oleracea*) genotypes (“Pakistan” local (“P”) and “Liaoning, China” local (“L”). Red indicates higher expression, and blue indicates lower expression.

These 94 genes took parts in acyl-(acyl-carrier-protein) desaturase (K00059, 8 genes), acyl-CoA oxidase (K00232, 24 genes), acyl-coenzyme A thioesterase 1/2/4 (K01068, 5 genes), acyl-(acyl-carrier-protein) desaturase (K03921, 16 genes), acetyl-CoA acyltransferase 1 (K07513, 14 genes), 17 beta-estradiol 17-dehydrogenase/very-long-chain 3-oxoacyl-CoA reductase (K10251, 5 genes),  $\omega$ -6 fatty acid desaturase/acyl-lipid  $\omega$ -6 desaturase (K10256, 15 genes), very-long-chain enoyl-CoA reductase (K10258, 6 genes) and very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase (K10703, 1 gene). Among them, genes, which played roles in acyl-(acyl-carrier-protein) desaturase (stearoyl- acyl-carrier-protein desaturase, *SAD*) and  $\omega$ -6 fatty acid desaturase (*FAD2*), participated in oleic acid (C18:1 $\Delta^9$ ) and linoleic acid (C18:2 $\Delta^9, 12$ ) biosynthesis separately, in the  $\omega$ -3 fatty acid synthesis and shown differential expression in leaves or roots in two genotypes.

16 *SAD* related genes were identified, of which 7 genes had not significantly changed, 6 genes upregulated and 3 downregulated in the leaves, 9 genes had not significantly changed, 5 genes upregulated and 2 downregulated in the roots, of "PL" compared with "LCL". 4 genes (i1\_LQ\_PLCLR\_c20654/f1p0/1627, i4\_LQ\_PLCLR\_c10207/f1p2/4530, i1\_LQ\_PLCLR\_c3802/f1p28/1729 and i1\_LQ\_PLCLR\_c3416/f2p66/1640) only upregulated in leaves and roots of "PL" and 2 genes (i4\_LQ\_PLCLR\_c10955/f1p13/4602 and i4\_LQ\_PLCLR\_c8768/f1p2/4930) only upregulated in leaves and roots of "LCL". 14 *FAD2* related genes were identified, of which 10 genes had not significantly changed, 3 genes upregulated and 1 downregulated in the leaves, 11 genes had not significantly changed, 2 genes upregulated and 1 downregulated in the roots, of "PL" compared with "LCL". 1 *FAD2* (i6\_LQ\_PLCLR\_c3764/f1p0/6143) only upregulated in leaves and roots of "PL" and another *FAD2* (i6\_LQ\_PLCLR\_c978/f1p0/6171) only upregulated in leaves and roots of "LCL" (Supplementary Table S7).

### 3. Discussion

Using the PacBio Iso-Seq platform, 15.33 GB subread base and 9,350,222 subreads with 1640, average reads were generated, the total number of transcripts were 132,536, and total genes were 78,559 in purslane. The same with other reports [27–30], our results have also shown that SMRT technology was an efficient tool for full-length cDNA sequencing and provided a rich resource for further functional genomics analysis in purslane.

All high-quality and unique SMRT transcripts were functionally annotated by seven databases. With GO analysis, transcripts associated through the metabolic process and cellular process in BP, cell and cell part in CC, binding and catalytic activity in MF were enriched in subcategories. Metabolism was also the most important transcript group with KEGG annotation, especially for carbon metabolism, carbon fixation in photosynthetic organisms and pyruvate metabolism. This might provide clues for further study for gene expression and regulation in some specific stress conditions in purslane. LncRNAs are regulated with gene expression, either transcriptional or post-transcriptional levels, during plant growth, development and abiotic and biotic stress [21,31–33]. 7289 LncRNA transcripts were identified in purslane with four analytical methods. Their function in purslane needed to be investigated further.

With the gene function interpretation of SMRT transcripts, the second-generation sequencing of purslane was carried out. A large number of differentially expressed genes and related KEGG pathways were found. The biosynthesis of unsaturated fatty acid, especially  $\omega$ -3 fatty acid, is one kind of nutrient substance and important for the abiotic tolerance of plants. *SAD* and *FAD2* were important genes in the pathway of  $\omega$ -3 fatty acid biosynthesis. Tissue-specific expressions of *SAD* and *FAD2* were found in some species. Stearoyl-ACP desaturase (SACPD) catalyzes the conversion of stearic acid (18:0) to oleic acid (18:1) during de novo fatty acid biosynthesis [34,35]. Differential expressions of *SAD* were found in seeds, leaves and roots of sacha inchi (*Plukenetia volubilis*) [36]. Eight putative *SAD* isoforms were found in the cacao (*Theobroma cacao*) genome. These eight isoforms displayed diverse expression patterns in various cacao tissues, and these genes exhibited distinct functions in seed and flower development and fatty acid synthesis [37]. Moreover,

differential expression of *FAD2* was found in different tissues of sunflower (*Helianthus annuus*) [38], cotton (*Gossypium hirsutum*) [39] and purslane [40]. *SAD* and *FAD2* were highly expressed in leaves and poorly in the roots and the stems of lima bean (*Phaseolus lunatus*) [41]. Our previous results have found that compared with “LCL”, “PL” has a significantly higher content of unsaturated fatty acids [42]. Differential expression of these two genes in leaves and roots of two genotypes of purslane laid the foundation for the further study of gene expression and manipulation in the synthesis of  $\omega$ -3 fatty acid in purslane.

The composition of fatty acids in plants may also affect plant morphology [43] and the growth rate of a cell population [44]. The *Arabidopsis thaliana* genome had 7 *SAD* isoforms [34]. *SSI2* was a *SAD* isoform, and *SSI2* mutation caused severe growth defects [45]. “PL” has an upright growth habit and “LCL” with a prostrate growth habit. At the same environmental condition, they have different growth rates [46]. It is needed to be verified that if the differential expressions of *SAD* and *FAD2* influenced the growth and development of the purslane genotype.

Abiotic stress regulated the expression of these two genes [41,47–51]. Cold acclimation upregulated *SAD* expression in the cold acclimating species, *Solanum commersonii*, compared with the cold non-acclimating species, *S. tuberosum* [47]. *FAD2* was required for normal growth of *Arabidopsis* at low-temperature [48,49]. High salinity induced *SAD* upregulation in *Hematococcus pluvialis* [50]. Upregulation of *FAD2* was necessary for salt tolerance during seed germination and early seedling growth of *Arabidopsis* under salinity stress [51]. Purslane is a salt-tolerant plant [8]. According to our previous results, “PL” was salt-tolerance and “LCL” was salt-sensitive purslane genotype, respectively [46,52]. Moreover, under 200 mM NaCl stress,  $\omega$ -3 fatty acid contents in the leaves of “PL” was significantly higher than that in “LCL”. However, the related molecular regulatory mechanism was still unknown. Further studies on elucidating the relationship between *SAD* and *FAD2* expression,  $\omega$ -3 fatty acid content and salt tolerance will build the foundation for selecting or breeding new genotypes with high  $\omega$ -3 fatty acid contents and better salt tolerance.

## 4. Materials and Methods

### 4.1. Plants Samples and Treatment

In order to get the full-length transcriptome of purslane and be used for salt-tolerance-related gene function analysis in the future, two purslane genotypes, “PL” and “LCL,” were chosen. After 14 days of sowing, seedlings were transplanted into plastic hydroponic boxes in a greenhouse on 23 March 2018 at the School of Agriculture and Biology, Shanghai Jiao Tong University, China.

Hydroponic system (absenteeism of  $O_2$ ) was chosen from our previous experiments [46]. For decreasing the environmental influence, 12 plants of selected genotypes remained in the same box as per replicate. The experiment was repeated three times. 15 L quarter-strength of Hoagland’s solution with an electrical conductivity of  $4.0 \text{ dS m}^{-1}$  and a pH of 5.8 were put in each plastic box, and quarter-strength of Hoagland’s solutions was replaced 2 times per week. The plantlets were fully mature in a greenhouse with a day temperature of  $28 \pm 2 \text{ }^\circ\text{C}$  and night temperature of  $16 \pm 2 \text{ }^\circ\text{C}$ , 70–80% relative humidity and  $400 \text{ } \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  photosynthetically active radiation. Plant tissue was collected from biological replicates in each box after two weeks of transplant. The composed tissues were directly ice-covered by liquid nitrogen, reserved at  $-80 \text{ }^\circ\text{C}$  until RNA isolation.

### 4.2. RNA Extraction, Library Construction and SMART Sequencing

Complete RNA samples (0.2 g leaves or roots from “PL” and “LCL” were mixed together) were extracted via RNeasy Plus mini kit (Qiagen, Valencia, CA, USA). After monitoring the deprivation and contamination percentage using agarose gel, RNA integrity (OD260/280) was checked with Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Rockland, DE, USA). RNA level and RNA authenticity were restrained dis-

tinctly by Qubit<sup>®</sup> RNA Assay Kit and RNA Nano 6000 Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA).

Disinfected RNA was secluded from total RNA using oligo (dT) enriches mRNA containing poly-(A) beads. The SMART PCR cDNA synthesis kit (Clontech, CA, USA) was used for cDNA synthesis. For the selection of differently sized full-length cDNA and for the construction of cDNA libraries, the BluePippin<sup>®</sup> (SageScience, Beverly, MA, USA) was used. After screened by Blue Pippin, the fragments were subjected to large-scale PCR to obtain sufficient total cDNA by using a Qubit fluorometer (Life Technologies, Carlsbad, CA, USA). Library uniqueness was kept constant by using Agilent Bioanalyzer 2100 system. SMRT sequencing was achieved via PacBio's real-time sequencer using C2 sequencing reagents. Quality filtering and error correction followed the methods of the literature [53].

#### 4.3. Functional Annotation of Transcripts

We identified functional annotations matching each unique transcript searching NR [54], NT, Pfam (<http://pfam.xfam.org>, accessed on 20 January 2021), KOG (<http://www.ncbi.nlm.nih.gov/COG/>, accessed on 20 January 2021) [55], Swiss-Prot (<http://www.ebi.ac.uk/uniprot/>, accessed on 20 January 2021) [56], KEGG (<http://www.genome.jp/kegg/>, accessed on 20 January 2021) [57] and GO [58]. We used the software of BLAST and NT data storage analysis, software of Diamond BLASTX v2.7.1 and set e-value "1e-10" in KEGG, Swiss-Prot, KOG, NR. and the software of Hmmscan in Pfam database analysis.

#### 4.4. Identification of Transcript Factors, Long Non-Coding RNAs and Simple Sequence Repeat

Plant transcription factors were predicted using iTAK v1.7a (<https://github.com/kentnf/iTAK/>, accessed on 20 January 2021) [26]. Four tools, CNCIv2 (<https://github.com/www-bioinfo-org/CNCI>, accessed on 20 January 2021) [59], CPCvcpc-0.9-r2 (<http://cpc.cbi.pku.edu.cn/>, accessed on 20 January 2021) with e-value '1e-10' [60], Pfam-scan (E 0.001-domE 0.001) [61] and PLEKv1.2 (<https://sourceforge.net/projects/plek/>, accessed on 20 January 2021) with min length 200 [62] were chosen to predict candidate LncRNAs. Transcripts forecasted without coding potential were candidate set of LncRNAs. SSRs were identified by MISA v1.0 (<http://pgrc.ipk-gatersleben.de/misa/>, accessed on 20 January 2021) [63] with default parameters. MISA can identify seven kinds of SSR types.

#### 4.5. Illumina cDNA Library Construction and Second-Generation Sequencing Analysis

Total RNA was extracted from leaves or roots with three independent biological replicates obtained from each genotype. Next, RNA purity was checked using the NanoPhotometer<sup>®</sup> spectrophotometer (IMPLEN, Westlake Village, CA, USA), and RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). A total amount of 1.5 µg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext<sup>®</sup> Ultra<sup>™</sup> RNA Library Prep Kit for Illumina<sup>®</sup> (NEB, Ipswich, MA, USA) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. The clustering of the index-coded samples was performed on a cBot cluster generation system using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq XTEN platform (San Diego, CA, USA), and paired-end reads were generated.

CD-HIT software was used to remove redundant and similar sequences [64], a non-redundant transcriptome generated by SMRT sequencing was used as a REF, then clean reads of each sample obtained from Illumina sequencing were compared to the REF. RSEM software was used in this process [65]. FPKM conversion was performed to analyze the gene expression level. Differential expression analyses of genes in leaves or roots between "PL" and "LCL" were performed using the DESeq R package (1.10.1). The resulting *p* values were adjusted using Benjamini–Hochberg's approach [66] for controlling the false

discovery rate. DEGs were selected with the criteria of absolute of  $\log_2$ -fold change  $> 0$  and  $p$ -value  $< 0.05$  by comparing differences between two genotypes of leaves or roots. GO enrichment analysis and KEGG pathway enrichment analysis were implemented to test the statistical enrichment of DEGs in KEGG pathways, respectively.

#### 4.6. Quantitative RT-PCR Analysis

qRT-PCR was carried out on a Mastercycler ep realplex, real-time PCR system (Eppendorf, Hamburg, Germany) using Bestar SYBR Green qPCR Mastermix (DBI, Bioscience Inc., Hamburg, Germany) [67]. Reactions were performed at 94 °C for 40 s, 30 cycles of 94 °C for 10 s, 54 °C for 30 s and 72 °C for 90 s. Three biological replicates were applied. The specific primers were designed using the Primer Premier 5.0 software. The primer sequences are listed in Supplementary Table S8. Actin was used as the reference genes. The target genes' relative expression levels were calculated using the  $2^{-\Delta\Delta CT}$  approach [68].

## 5. Conclusions

In this scenario, we finished RNA-seq in purslane with SMRT technology. Related results provided significant information for enlightening the whole-genome annotation and transcriptomic characterization that might be useful for genetic manipulation of purslane genotypes under abiotic stresses. With the function of gene function interpretation of RNA-seq technology, we finished Illumina sequencing of leaves and roots of two purslane genotypes and found differential expression of some genes, especially *FAD2* and *SAD*, may be related with  $\omega$ -fatty acid contents of special genotypes of purslane.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/plants10040655/s1>, Table S1: Function annotation of all corrected isoforms in seven databases by SMRT. Table S2: Transcription factor prediction of Iso-Seq reads. Table S3: LncRNAs and mRNAs from Iso-Seq. Table S4: SSR and corresponding primer. Table S5: Differentially expressed transcripts in four comparisons (Leaves or root, "LCL" vs. "PL", up or down). Table S6: KEGG pathway enrichment results of upregulated and downregulated pathways in leaves and roots of two purslane genotypes ("LCL" and "PL"). Table S7: Gene expressions of participated in the biosynthesis of unsaturated fatty acids in leaves and roots of two genotypes ("LCL" and "PL") of purslane (*Portulaca oleracea*). Table S8: Primers used for qRT-PCR analysis of differential genes in purslane (*Portulaca oleracea*).

**Author Contributions:** Conceptualization, H.D. and S.C.; methodology, H.D., S.Z. and S.H.; software, H.D.; validation: H.D.; formal analysis: H.D. and S.Z.; investigation, H.D. and S.H.; resources, H.D. and S.H.; data curation, H.D. and S.H.; writing—original draft preparation, H.D. and S.Z.; writing—review and editing, H.D. and S.C.; visualization, H.D. and S.Z.; supervision, H.D. and S.C.; project administration, S.C.; funding acquisition, S.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Key Technology R&D Program, grant number 2015BAL02B01.

**Data Availability Statement:** Raw CCS reads are available at the SRA database of NCBI GenBank under the accession PRJNA522036.

**Acknowledgments:** We thank Novo Gene Co., China, for technical support in SMRT sequencing and data analysis.

**Conflicts of Interest:** The authors declare that they have no competing interests.

## References

1. Ezekwe, M.O.; Omara-Alwala, T.R.; Membrahtu, T. Nutritive characterization of purslane accessions as influenced by planting date. *Plant Foods Hum. Nutr.* **1999**, *54*, 183–191. [[CrossRef](#)]
2. Simopoulos, A.P.; Norman, H.A.; Gillaspay, J.E. Purslane in human nutrition and its potential for world agriculture. *Plant Foods Hum. Nutr.* **1995**, *77*, 47–74.
3. Uddin, M.; Juraimi, A.S.; Hossain, M.S.; Nahar, M.; Ali, M.; Rahman, M. Purslane weed (*Portulaca oleracea*): A prospective plant source of nutrition, omega-3 fatty acid, and antioxidant attributes. *Sci. World J.* **2014**, *2014*, 951019. [[CrossRef](#)] [[PubMed](#)]

4. Lee, A.S.; Kim, J.S.; Lee, Y.J.; Kang, D.G.; Lee, H.S. Anti-TNF- $\alpha$  activity of *Portulaca oleracea* in vascular endothelial cells. *Int. J. Mol. Sci.* **2012**, *13*, 5628–5644. [[CrossRef](#)]
5. Hayat, M.Q.; Khan, M.A.; Ahmad, M.; Shaheen, N.; Yasmin, G.; Akhter, S. Ethnotaxonomical approach in the identification of useful medicinal flora of tehsil Pindigheb (District Attock) Pakistan. *Ethnobot. Res. Appl.* **2008**, *6*, 35–62. [[CrossRef](#)]
6. Hussain, K.; Nisar, M.F.; Majeed, A.; Nawaz, K.; Bhatti, K.H. Ethnomedicinal survey for important plants of Jalalpur Jattan, district Gujrat, Punjab, Pakistan. *Ethnobot. Leaflet.* **2010**, *14*, 807–825.
7. Ercisli, S.; Coruh, I.; Gormez, A.; Sengul, M. Antioxidant and antibacterial activities of *Portulaca oleracea* L. Grown wild in Turkey. *Ital. J. Food Sci.* **2008**, *20*, 533–542.
8. Grieve, C.M.; Suarez, D.L. Purslane (*Portulaca oleracea* L.): A halophytic crop for drainage water reuse systems. *Plant Soil* **1997**, *192*, 277–283. [[CrossRef](#)]
9. Alam, M.A.; Juraimi, A.S.; Rafii, M.Y.; Hamid, A.A. Effect of salinity on biomass yield and physiological and stem-root anatomical characteristics of purslane (*Portulaca oleracea* L.) accessions. *Biomed. Res. Int.* **2015**, *2015*, 105695. [[CrossRef](#)] [[PubMed](#)]
10. Kafi, M.; Rahimi, Z. Effect of salinity and silicon on root characteristics, growth, water status, proline content and ion accumulation of purslane (*Portulaca oleracea* L.). *Soil Sci. Plant Nutr.* **2011**, *57*, 341–347. [[CrossRef](#)]
11. Mulry, K.R.; Hanson, B.A.; Dudle, D.A. Alternative strategies in response to saline stress in two varieties of *Portulaca oleracea* (Purslane). *PLoS ONE* **2015**, *10*, e0138723.
12. Yazici, I.; Türkan, I.; Sekmen, A.H.; Demiral, T. Salinity tolerance of purslane (*Portulaca oleracea* L.) is achieved by enhanced antioxidative system, lower level of lipid peroxidation and proline accumulation. *Environ. Exp. Bot.* **2007**, *61*, 49–57. [[CrossRef](#)]
13. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **2004**, *431*, 931–945. [[CrossRef](#)] [[PubMed](#)]
14. Dharshini, S.; Chakravarthi, M.; Manoj, V.; Naveenarani, M.; Kumar, R.; Meena, M.; Ram, B.; Appunu, C. De novo sequencing and transcriptome analysis of a low temperature tolerant *Saccharum spontaneum* clone IND 00–1037. *J. Biotechnol.* **2016**, *231*, 280–294. [[CrossRef](#)]
15. Long, W.; Zou, X.; Zhang, X. Transcriptome analysis of canola (*Brassica napus*) under salt stress at the germination stage. *PLoS ONE* **2015**, *10*, e0116217. [[CrossRef](#)] [[PubMed](#)]
16. Li, M.; Liang, Z.; Zeng, Y.; Jing, Y.; Wu, K.; Liang, J.; He, S.; Wang, G.; Mo, Z.; Tan, F. De novo analysis of transcriptome reveals genes associated with leaf abscission in sugarcane (*Saccharum officinarum* L.). *BMC Genom.* **2016**, *17*, 195. [[CrossRef](#)] [[PubMed](#)]
17. Wang, J.; Li, B.; Meng, Y.; Ma, X.; Lai, Y.; Si, E.; Yang, K.; Ren, P.; Shang, X.; Wang, H. Transcriptomic profiling of the salt-stress response in the halophyte *Halostylon glomeratus*. *BMC Genom.* **2015**, *16*, 169. [[CrossRef](#)]
18. Xing, J.C.; Zhao, B.Q.; Dong, J.; Liu, C.; Wen, Z.G.; Zhu, X.M.; Ding, H.R.; He, T.T.; Yang, H.; Wang, M.W.; et al. Transcriptome and metabolome profiles revealed molecular mechanisms underlying tolerance of *Portulaca oleracea* to saline stress. *Russ. J. Plant Physiol.* **2020**, *67*, 146–152. [[CrossRef](#)]
19. Jiao, W.B.; Schneeberger, K. The impact of third generation genomic technologies on plant genome assembly. *Curr. Opin. Plant Biol.* **2017**, *36*, 64–70. [[CrossRef](#)]
20. Slatko, B.E.; Gardner, A.F.; Ausubel, F.M. Overview of next-generation sequencing technologies. *Curr. Protoc. Mol. Biol.* **2018**, *122*, e59. [[CrossRef](#)]
21. Wang, B.; Tseng, E.; Regulski, M.; Clark, T.A.; Hon, T.; Jiao, Y.; Lu, Z.; Olson, A.; Stein, J.C.; Ware, D. Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing. *Nat. Commun.* **2016**, *7*, 11708. [[CrossRef](#)]
22. Xie, L.; Teng, K.; Tan, P.; Chao, Y.; Li, Y.; Guo, W.; Han, L. PacBio single-molecule long-read sequencing shed new light on the transcripts and splice isoforms of the perennial ryegrass. *Mol. Genet. Genom.* **2020**, *295*, 475–489. [[CrossRef](#)]
23. Jia, X.; Tang, L.; Mei, X.; Liu, H.; Luo, H.; Deng, Y.; Su, J. Single-molecule long-read sequencing of the full-length transcriptome of *Rhododendron lapponicum* L. *Sci. Rep.* **2020**, *10*, 6755. [[CrossRef](#)] [[PubMed](#)]
24. Yuan, H.; Yu, H.; Huang, T.; Shen, X.; Xia, J.; Pang, F.; Wang, J.; Zhao, M. The complexity of the *Fragaria*  $\times$  *ananassa* (octoploid) transcriptome by single-molecule long-read sequencing. *Hortic. Res.* **2019**, *6*, 46. [[CrossRef](#)]
25. Deng, N.; Hou, C.; Ma, F.F.; Liu, C.X.; Tian, Y.X. Molecule long-read sequencing reveals the diversity of full-length transcripts in leaves of *Gnetum* (Gnetales). *Int. J. Mol. Sci.* **2019**, *20*, 6350. [[CrossRef](#)]
26. Zheng, Y.; Jiao, C.; Sun, H.; Rosli, H.G.; Pombo, M.A.; Zhang, P.; Banf, M.; Dai, X.; Martin, G.B.; Giovannoni, J.J.; et al. iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* **2016**, *9*, 1667–1670. [[CrossRef](#)]
27. Chao, Y.; Yuan, J.; Li, S.; Jia, S.; Han, L.; Xu, L. Analysis of transcripts and splice isoforms in red clover (*Trifolium pratense* L.) by single-molecule long-read sequencing. *BMC Plant Bio.* **2018**, *18*, 300. [[CrossRef](#)]
28. Workman, R.E.; Myrka, A.M.; Wong, G.W.; Tseng, E.; Welch, J.K.C.; Timp, W. Single-molecule, full-length transcript sequencing provides insight into the extreme metabolism of the ruby-throated hummingbird *Archilochus colubris*. *GigaScience* **2018**, *7*, giy009. [[CrossRef](#)]
29. Zeng, D.; Chen, X.; Peng, J.; Yang, C.; Peng, M.; Zhu, W.; Xie, D.; He, P.; Wei, P.; Lin, Y.; et al. Single-molecule long-read sequencing facilitates shrimp transcriptome research. *Sci. Rep.* **2018**, *8*, 16920. [[CrossRef](#)]
30. Zuo, C.; Blow, M.; Sreedasyam, A.; Kuo, R.C.; Ramamoorthy, G.K.; Torres-Jerez, I.; Li, G.; Wang, M.; Dilworth, D.; Barry, K.; et al. Revealing the transcriptomic complexity of switchgrass by PacBio long-read sequencing. *Biotechnol. Biofuels* **2018**, *11*, 170. [[CrossRef](#)]

31. Chekanova, J.A. Long non-coding RNAs and their functions in plants. *Curr. Opin. Plant Biol.* **2015**, *27*, 207–216. [[CrossRef](#)]
32. Karlik, E.; Marakli, S.; Gozukirmizi, N. Two lncRNAs expression profiles in salt stressed barley (*Hordeum vulgare* L.) roots. *Cytologia* **2018**, *83*, 37–43. [[CrossRef](#)]
33. Wang, J.; Lin, J.; Kan, J.; Wang, H.; Li, X.; Yang, Q.; Li, H.; Chang, Y. Genome-wide identification and functional prediction of novel drought-responsive lncRNAs in *Pyrus betulifolia*. *Genes* **2018**, *9*, 311. [[CrossRef](#)]
34. Kachroo, A.; Shanklin, J.; Whittle, E.; Lapchuk, L.; Hildebrand, D.; Kachroo, P. The Arabidopsis stearyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. *Plant Mol. Biol.* **2007**, *63*, 257–271. [[CrossRef](#)] [[PubMed](#)]
35. Shanklin, J.; Cahoon, E.B. Desaturation and related modifications of fatty acids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 611–641. [[CrossRef](#)]
36. Wang, X.; Liu, A. Expression of genes controlling unsaturated fatty acids biosynthesis and oil deposition in developing seeds of sacha inchi (*Plukenetia volubilis* L.). *Lipids* **2014**, *49*, 1019–1031. [[CrossRef](#)]
37. Zhang, Y.; Maximova, S.N.; Guiltinan, M.J. Characterization of a stearyl-acyl carrier protein desaturase gene family from chocolate tree, *Theobroma cacao* L. *Front. Plant Sci.* **2015**, *6*, 239. [[CrossRef](#)]
38. Martínez-Rivas, J.M.; Martínez, R. Spatial and temporal regulation of three different microsomal oleate desaturase genes (*FAD2*) from normal-type and high-oleic varieties of sunflower (*Helianthus annuus* L.). *Mol. Breed.* **2001**, *8*, 159–168. [[CrossRef](#)]
39. Zhang, D.; Pirtle, I.L.; Park, S.J.; Nampaisansuk, M.; Neogi, P.; Wanjie, S.W.; Pirtle, R.M.; Chapman, K.D. Identification and expression of a new delta-12 fatty acid desaturase (*FAD2-4*) gene in upland cotton and its functional expression in yeast and *Arabidopsis thaliana* plants. *Plant Physiol. Biochem.* **2009**, *47*, 462–471. [[CrossRef](#)]
40. Teixeira, M.C.; Coelho, N.; Olsson, M.E.; Brodelius, P.E.; Carvalho, I.S.; Brodelius, M. Molecular cloning and expression analysis of three omega-6 desaturase genes from purslane (*Portulaca oleracea* L.). *Biotechnol. Lett.* **2009**, *31*, 1089–1101. [[CrossRef](#)]
41. Zhang, Y.; Wang, C.; Hu, H.; Yang, L. Cloning and expression of three fatty acid desaturase genes from cold-sensitive lima bean (*Phaseolus lunatus* L.). *Biotechnol. Lett.* **2011**, *33*, 395–401. [[CrossRef](#)]
42. Hu, S.Q.Q.; Du, H.M. Content analysis of fatty acids and oxalic acid in ten different types of purslane (*Portulaca oleracea*). *Guihaia* **2019**, *39*, 1550–1557, (In Chinese with English abstract).
43. Raheja, R.K.; Batta, S.K.; Ahuja, K.L.; Labana, K.S.; Singh, M. Comparison of oil content and fatty acid composition of peanut genotypes differing in growth habit. *Plant Foods Hum. Nutr.* **1987**, *37*, 103–108. [[CrossRef](#)]
44. Mei, C.; Michaud, M.; Cussac, M.; Albrieux, C.; Gros, V.; Maréchal, E.; Block, M.A.; Jouhet, J.; Rébeillé, F. Levels of polyunsaturated fatty acids correlate with growth rate in plant cell cultures. *Sci. Rep.* **2015**, *5*, 15207. [[CrossRef](#)]
45. Yang, W.; Dong, R.; Liu, L.; Hu, Z.; Li, J.; Wang, Y.; Ding, X.; Chu, Z. A novel mutant allele of *SSI2* confers a better balance between disease resistance and plant growth inhibition on *Arabidopsis thaliana*. *BMC Plant Biol.* **2016**, *16*, 208. [[CrossRef](#)]
46. Zaman, S.; Hu, S.; Alam, M.A.; Du, H.; Che, S. The accumulation of fatty acids in different organs of purslane under salt stress. *Sci. Hortic.* **2019**, *250*, 236–242. [[CrossRef](#)]
47. Vega, S.E.; del Rio, A.H.; Bamberg, J.B.; Palta, J.P. Evidence for the up-regulation of stearyl-ACP ( $\Delta 9$ ) desaturase gene expression during cold acclimation. *Am. J. Potato Res.* **2004**, *81*, 125–135. [[CrossRef](#)]
48. Miquel, M.; James, D.; Dooner, H. Arabidopsis requires polyunsaturated lipids for low-temperature survival. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6208–6212. [[CrossRef](#)]
49. Wallis, J.G.; Browse, J. Mutants of Arabidopsis reveal many roles for membrane lipids. *Prog. Lipid Res.* **2002**, *41*, 254–278. [[CrossRef](#)]
50. Lei, A.; Chen, H.; Shen, G.; Hu, Z.; Chen, L.; Wang, J. Expression of fatty acid synthesis genes and fatty acid accumulation in *Haematococcus pluvialis* under different stressors. *Biotechnol. Biofuels* **2012**, *5*, 18. [[CrossRef](#)]
51. Zhang, J.; Liu, H.; Sun, J.; Li, B.; Zhu, Q.; Chen, S.; Zhang, H. Arabidopsis fatty acid desaturase *FAD2* is required for salt tolerance during seed germination and early seedling growth. *PLoS ONE* **2012**, *7*, e30355. [[CrossRef](#)]
52. Hu, S.Q.Q.; Du, H.M. The effects of NaCl treatment on growth of purslane (*Portulaca oleracea*). *J. Shanghai Jiaotong Univ. (Agri. Sci.)* **2018**, *36*, 67–72, (In Chinese with English abstract).
53. Song, H.; Yang, M.; Yu, Z.; Zhang, T. Characterization of the whole transcriptome of whelk *Rapana venosa* by single-molecule mRNA sequencing. *Mar. Genom.* **2019**, *44*, 74–77. [[CrossRef](#)]
54. Li, W.; Jaroszewski, L.; Godzik, A. Tolerating some redundancy significantly speeds up clustering of large protein databases. *Bioinformatics* **2002**, *18*, 77–82. [[CrossRef](#)] [[PubMed](#)]
55. Tatusov, R.L.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Kiryutin, B.; Koonin, E.V.; Krylov, D.M.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N. The COG database: An updated version includes eukaryotes. *BMC Bioinform.* **2003**, *4*, 41. [[CrossRef](#)] [[PubMed](#)]
56. Bairoch, A.; Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **2000**, *28*, 45–48. [[CrossRef](#)]
57. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [[CrossRef](#)]
58. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T. Gene Ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)]
59. Zhang, C.; Bu, D.; Zhao, G.; Luo, H.; Yu, K.; Sun, L.; Chen, R.; Liu, Y.; Zhao, Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* **2013**, *41*, e166.

60. Kong, L.; Zhang, Y.; Ye, Z.Q.; Liu, X.Q.; Zhao, S.Q.; Wei, L.; Gao, G. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **2007**, *35*, W345–W349. [[CrossRef](#)]
61. Finn, R.D.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2015**, *44*, D279–D285. [[CrossRef](#)]
62. Li, A.; Zhang, J.; Zhou, Z. PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved *k*-mer scheme. *BMC Bioinform.* **2014**, *15*, 311. [[CrossRef](#)]
63. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
64. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. [[CrossRef](#)] [[PubMed](#)]
65. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [[CrossRef](#)]
66. Ferreira, J.; Zwinderman, A. On the Benjamini-Hochberg method. *Ann. Stat.* **2006**, *34*, 1827–1849. [[CrossRef](#)]
67. Zhang, X.; Su, N.; Jia, L.; Tian, J.; Li, H.; Huang, L.; Shen, Z.; Cui, J. Transcriptome analysis of radish sprouts hypocotyls reveals the regulatory role of hydrogen-rich water in anthocyanin biosynthesis under UV-A. *BMC Plant Biol.* **2018**, *18*, 227. [[CrossRef](#)]
68. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)]