

Article

# Genome-Wide Identification, Evolutionary Patterns, and Expression Analysis of *bZIP* Gene Family in Olive (*Olea europaea* L.)

Siyu Rong, Zhiyang Wu, Zizhang Cheng, Shan Zhang, Huan Liu and Qianming Huang \*

Department of Bioresource Chemistry, College of Science, Sichuan Agricultural University, Ya'an 625000, China; 18227591075@163.com (S.R.); wuzhiyang119@163.com (Z.W.); chengxiaoyi\_01@sina.com (Z.C.); zhangshanwts@163.com (S.Z.); liuhuan@163.com (H.L.)

\* Correspondence: hqming@sicau.edu.cn; Tel.: +86-0835-288-6189

Received: 6 March 2020; Accepted: 1 May 2020; Published: 5 May 2020



**Abstract:** Olive (*Olea europaea* L.) is an economically important oleaginous crop and its fruit cold-pressed oil is used for edible oil all over the world. The basic region-leucine zipper (*bZIP*) family is one of the largest transcription factors families among eukaryotic organisms; its members play vital roles in environmental signaling, stress response, plant growth, seed maturation, and fruit development. However, a comprehensive report on the *bZIP* gene family in olive is lacking. In this study, 103 *OebZIP* genes from the olive genome were identified and divided into 12 subfamilies according to their genetic relationship with 78 *bZIP*s of *A. thaliana*. Most *OebZIP* genes are clustered in the subgroup that has a similar gene structure and conserved motif distribution. According to the characteristics of the leucine zipper region, the dimerization characteristics of 103 *OebZIP* proteins were predicted. Gene duplication analyses revealed that 22 *OebZIP* genes were involved in the expansion of the *bZIP* family. To evaluate the expression patterns of *OebZIP* genes, RNA-seq data available in public databases were analyzed. The highly expressed *OebZIP* genes and several lipid synthesis genes (*LPGs*) in fruits of two varieties with different oil contents during the fast oil accumulation stage were examined via qRT-PCR. By comparing the dynamic changes of oil accumulation, *OebZIP1*, *OebZIP7*, *OebZIP22*, and *OebZIP99* were shown to have a close relationship with fruit development and lipid synthesis. Additionally, some *OebZIP* had a significant positive correlation with various *LPG* genes. This study gives insights into the structural features, evolutionary patterns, and expression analysis, laying a foundation to further reveal the function of the 103 *OebZIP* genes in olive.

**Keywords:** olive; *bZIP*; evolution; expression analysis; fruit development; lipid synthesis

## 1. Introduction

The olive (*Olea europaea* L.) is a socioeconomically important oleaginous crop that is planted widely in the Mediterranean basin, symbolizing peace, health, and longevity. The olive is the only member of the Oleaceae family with edible fruit. Olive oil is rich in unsaturated fatty acids and other important secondary metabolites. It also contains more than 30 different phenolic compounds [1] that are strong antioxidants and free radical scavengers [2]. In fact, olive oil has great benefits for human health [3], especially for cardiovascular disease, obesity, and diabetes [4]. It has also been recognized as having potential value in reducing the incidence of Alzheimer's [5]. With the development of China's economy, the market demand for olive oil is increasing dramatically; nevertheless, the domestic planting area is small and the production is low; therefore, consumption depends on imports. As such, it is necessary for us to reduce the cost, increase the planting yield, and improve the quality of olive oil.

The synthesis and accumulation of lipids are regulated by complex and cooperative mechanisms [6]. Transcription factors (TFs) are considered to be important factors in regulating plant development via self-regulation and regulation of downstream target gene expression. In a model plant, *Arabidopsis*, 8% of protein-encoding capacity is attributed to TFs [7], and more than 100 TF families have been found according to the special amino acid sequences and conserved DNA-binding domains. TFs specifically interact with cis-acting elements in the promoter or other TFs through functional domains to activate or inhibit transcription [8]. Studies have found that *MYB* [9,10], *WR11* [11–14], *Dof* [15–17], and *bZIP* [18–21] play key roles in regulating lipid biosynthesis. Among them, we will focus on basic leucine zipper (*bZIP*). The plant *bZIP* TF contains two functional domains: a basic region and an adjacent leucine zipper. The basic region generally consists of 18 amino acid residues (aa), which is very conserved in eukaryotes [22]. The basic regions all have the same aa sequence model like N-X<sub>7</sub>-R/K-X<sub>9</sub>, which is a DNA-binding domain. The *bZIP* target gene sequences consist of an ACGT-core, showing palindromic-hexamers like G-boxes (CACGTG), C-boxes (GACGTC), and A-boxes (TACGTA). However, the leucine zipper region is more variable and is composed of a repeat aa sequence, L-X<sub>6</sub>-L-X<sub>6</sub>-L. This special aa arrangement makes hydrophobic amino acids all appear on one side of the  $\alpha$ -helix. The result is that the leucine zipper forms an amphiphilic helical structure which is involved in the dimerization of the *bZIP* protein before it binds to DNA. So, *bZIP* TFs generally act as dimers to form a ‘scissor-grip’ model and bind dsDNA by the basic region to regulate flexibly.

With respect to function, *bZIP* genes are involved in numerous biological and physiological processes in plant growth and development [23], as well as resistance to stress [24,25]. The functional relevance of *bZIP* in numerous developmental processes is well-established, such as systemic acquired resistance [7], chloroplast development [26], pollen germination [19], stress response [27], and seed maturation, as well as oil accumulation. Research on *bZIP* in seed maturation and oil accumulation is particularly abundant. *AtbZIP10*, *AtbZIP25*, and *AtbZIP53* interacts with *ABI3* [28]. The expression of *AtbZIP53* precedes and overlaps the seed maturation gene (*MAT*). Gain- and loss-of-function approaches indicate a correlation between the amount of the *bZIP53* protein and *MAT* gene expression. *Arabidopsis* *LEC1* and *L1L* activate transcription by interacting with seed-specific ABRE-binding factor, *bZIP67* [18]. Another study found that *AtbZIP67* regulates the  $\omega$ -3 fatty acid content of *Arabidopsis* seed oil by activating fatty acid desaturase 3 [20]. Soybean *GmbZIP123* participates in lipid accumulation in seeds, and the overexpression of *GmbZIP123* can increase the lipid content in the seeds of transgenic *Arabidopsis* [21].

Although the *bZIP* family has been widely studied in many plants, including *Arabidopsis thaliana* (*Arabidopsis*) [22], *Oryza sativa* L. (rice) [29], *Sesamum indicum* (sesame) [30], *Glycine max* (soybean) [31], *Vitis vinifera* L. (grape) [32], *Ricinus communis* L. (castor bean), *Manihot esculenta* (cassava) [33], *Fragaria vesca* (strawberry) [34], *Malus domestica* (apple) [35], and *Brassica napus* (rape) [36], there is still not enough information about it in olive. The genome of the olive was sequenced recently [37]; genome-wide data are available in the Olive Database (<http://olivegenome.org>). Analyzing the gene family at the genome level, we can understand the evolutionary process of the *bZIP* gene family and the differentiation of its function, so it is necessary to explore the structure and evolutionary pattern of *bZIP* at the gene family level. In this study, 103 *OebZIP* genes were identified and divided into 12 subgroups, referencing to the classification of *Arabidopsis* [38]. In addition, *OebZIP* genes were systematically investigated and analyzed, including chromosome distribution, gene structure, motif distribution, dimerization, gene duplication events, and evolutionary analyses with other different plant species. Furthermore, their expression analysis in different tissues (fruits, leaves, stems, pedicels) were examined using RNA-seq data which are available in public databases. To further study this matter, we selected 15 candidate *OebZIP* genes that are highly expressed in fruits based on transcriptome data. Then, we analyzed the expression profiles at different fruit development stages (July to November) of two varieties via qRT-PCR, due to the fact that only a small number of *bZIPs* have been functionally characterized in plants. Oil in olives is produced mainly in the mesocarp. The oil accumulation in the mesocarp follows a typical S-shaped curve [39]. Generally, olive oil is

accumulated mostly between August and September, and reaches a plateau in November. However, due to environmental conditions, different agricultural habits, and olive varieties, the manner of lipid accumulation may be different [40]. Therefore, we preliminarily evaluate the expressive relationship with lipid synthesis genes (*LPGs*) in two varieties. We analyzed the expressive relation between *LPGs* and *OebZIP*, combined with the dynamic diversification of oil content, predicted the upstream cis-element of *LPGs* that contains A-boxes, G-Boxes, and C-Boxes, and performed a correlation analysis. Our study was the first to perform genome-wide identification and an expression analysis of *bZIP* genes. The goal is to provide a foundation to further study the functions of *OebZIP* on lipid synthesis.

## 2. Materials and Methods

### 2.1. Genome Data Sources

The olive (*Olea europaea* L.) genome (v1.0) was downloaded from the olive database (<http://olivegenome.org/>) [41], while the rice (*Oryza sativa* L.) (IRGSP-1.0) [42], Arabidopsis (*Arabidopsis thaliana*) (TAIR10) [43], soybean (*Glycine max*) (v2.1) [44], cacao (*Theobroma cacao*) (v2.0) [45] and grape (*Vitis vinifera* L.) genomes (12X) [46] were downloaded from the Ensembl Plants database (<http://plants.ensembl.org/index.html>). Sesame (*Sesamum indicum*) genome (v1.0) data were downloaded from the *Sesamum indicum* genome database (<http://ocri-genomics.org>) [47].

The bZIP proteins of other species were downloaded from Plant Transcription Factor Database v5.0 (<http://planttfdb.cbi.pku.edu.cn/index.php>).

### 2.2. Identification and Filtration

To retrieve all members of the *bZIP* gene family, we used the HMMER v3.1b2 (<http://hmm.janelia.org/>) Hidden Markov model (PF00170) as a probe and to screen all the candidate proteins with E values of less than 1e-10 in order to build a new, olive-specific Hidden Markov model for the secondary search. Then, the secondary search results of the protein sequences were further confirmed using SMART (<http://smart.embl-heidelberg.de/>) [48] and Pfam (<http://pfam.xfam.org/search>) [49] to ensure their reliability. Finally, all putative bZIP gene models were utilized to analyze the amino acid length, PI, MV by ProtParam (<https://web.expasy.org/protparam/>) and named by their chromosomal location.

### 2.3. Phylogenetic Analysis and Classification

In order to detect the classification and evolutionary profile of all *OebZIP*, the amino acid sequences of the Arabidopsis and olive were determined by ClustalW. The tree was constructed using MEGAX with the Maximum Likelihood (ML) method: A phylogeny test was performed with the bootstrap method of 1000 replicates, substitution with the Poisson model, and gaps Data treatment with pairwise deletion.

### 2.4. Gene Structure and Conserved Motif Analysis

We used the online tool Gene Structure Display Server (GSDS v2.0) (<http://gsds.cbi.pku.edu.cn/index.php>) to analyze the structure of the *OebZIP* gene containing the constituents of the exons/introns. Next, we employed MEME v 5.0.2 (<http://meme-suite.org/tools/meme>) to explore the conservative motifs as well as the bZIP domain, with the following optimized parameters: the maximum number of motifs was set to 10; the width of each motif was 6–50 residues.

### 2.5. Chromosomal Location, Gene Duplication, and Evolutionary Analysis

MG2C v2.0 ([http://mg2c.iask.in/mg2c\\_v2.0/](http://mg2c.iask.in/mg2c_v2.0/)) was used to locate the *OebZIP* gene on the chromosomes of *Olea europaea*.L. Firstly, according to multiple sequence alignments, we got a number of a pairwise genes that had similar lengths and sequences by Circos (<http://circos.ca/>). Afterward, the MCScanX was used to find duplication genes in the olive genome and to analyze collinearity between olive and six other species (Arabidopsis, rice, sesame, soybean, cacao, and grape) [50].

## 2.6. Dimerization Properties of OebZIP Proteins

The leucine zipper region of OebZIP proteins domain, which might be related to the dimerization stability and specificity, was partitioned into different boundaries of N and C terminal [51]. To describe and speculate on the dimerization foundation of 103 OebZIP members, we set seven special amino acid residues that covered four amino acid residues before the appearance of the first leucine in the bZIP domain and the following two amino acid residues, as an initial heptad named L<sub>0</sub>. Next, the amino acid sequences were manually arranged to the end or the tenth heptad, named L<sub>9</sub>. The nomenclature of the position within each heptad was *g, a, b, c, d, e, and f*.

## 2.7. Spatiotemporal Expression Analysis of OebZIP Gene Using Transcriptome Data

To investigate the changes in expression of the *OebZIP* gene across different stages (July, November) and tissues (stem, leaf, pedicel, and fruit), we acquired olive transcriptome data [37] from the NCBI SRA database (<https://www.ncbi.nlm.nih.gov/sra>) (SRR4473639, SRR4473641, SRR44742, SRR4473643, SRR4473644, SRR4473645, SRR4473646, SRR4473647). After downloading the data, we used FastQC v0.11.8 to check the quality of the fastq files and trim adapters using fastp v0.19.4. Next, we mapping reads by HISAT2 v2.1.0. Counts were normalized and standardized using HTSeq v0.9.1. FPKM was calculated by Cufflinks v2.0.2 to obtain gene expression values. Finally, the heat maps were drawn using R v3.5.3, and the data normalization parameter was “scale = row”.

## 2.8. Plant Material

We got fresh samples from July to November in 2019 from Chengdu City, Sichuan province, China (30°45' N, 104°32' E, altitude of 500 ~ 600 m). Olive trees (5 years old) having a vase shape with 5 × 8 distances were randomly cultivated. Different varieties of olive, i.e., ‘Arbequina’, ‘Koroneiki’, ‘Ezhi-8’, ‘Picholine’, and ‘Grossa’, were used in this study. Each experiment was repeated at least three times.

## 2.9. Plant RNA Extraction and qRT-PCR

Total RNA was extracted using a kit (Tiangen, Beijing, China), the reverse-transcribed process was performed with a PrimeScript RT reagent kit (TaKaRa, Dalian, China), and primers were designed by Primer5. The ‘Arbequina’ and ‘Grossa’, due to the enormous difference in their oil contents, were chosen for the qRT-PCR experiment to explore the expression quantity of *OebZIPs* in leaves and fruits at fruit development stages. Then, *LPG* genes were tested.  $\beta$ -*Actin* was used as an internal reference gene, and the expression data were calculated by 2-( $\Delta\Delta$ CT) method [52].

# 3. Results

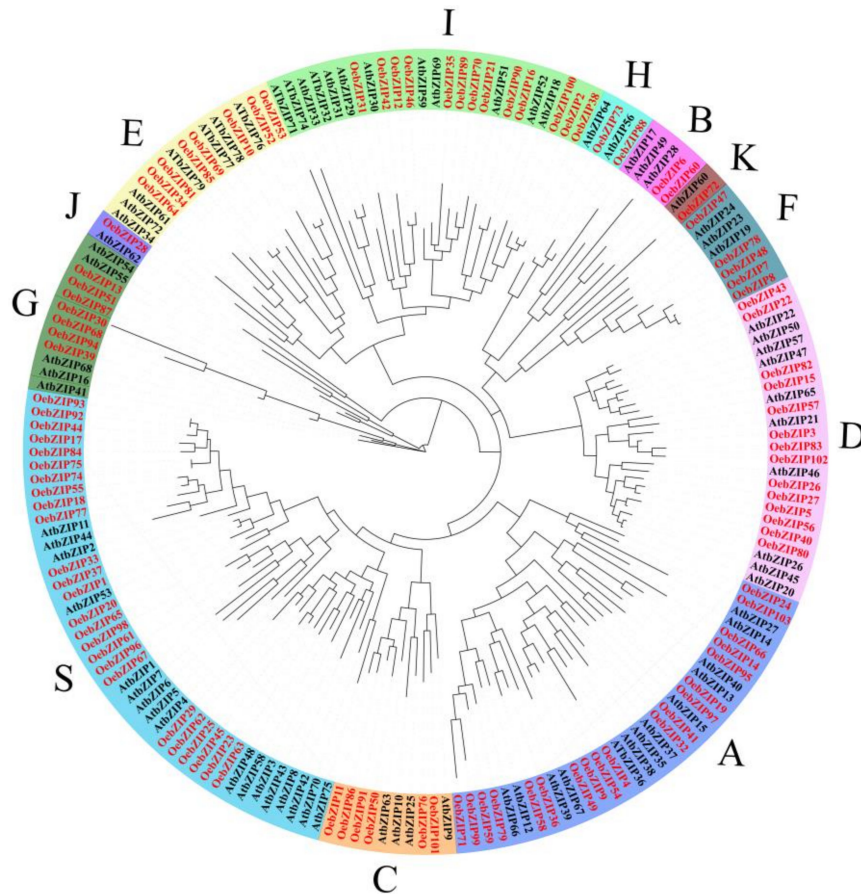
## 3.1. Identification and Naming of OebZIP

Based on a HMMER search analysis, we manually screened the candidates in two databases and got 103 *bZIP* genes in the olive genome; we named these genes according to their distribution on the chromosomes (*OebZIP1*~*OebZIP103*, Table S1). The coding sequence length, the isoelectric points (pIs), and molecular weights (MWs) of these genes were then analyzed, as shown in Table S1. The lengths of the all 103 *OebZIP* proteins were between 96 aa and 772 aa, with an average of 315 aa. The MWs of the proteins ranged from 11.22 (*OebZIP99*) to 84.31 (*OebZIP6*), with an average of 34.96. The isoelectric point (pIs) was between 4.85 (*OebZIP97*) and 10.99 (*OebZIP96*), with an average of 7.20.

## 3.2. Phylogenetic Analysis and Classification

To classify the subgroups of *OebZIPs* and explore the evolutionary relationship between the Arabidopsis and olive, we constructed an unroot phylogenetic tree based on the Maximum Likelihood (ML) method with protein sequences of 75 *AtbZIP* genes and the 103 *OebZIP* genes (Figure 1). The 75 *AtbZIP* genes were divided into 12 subgroups, with 103 *OebZIPs* being distributed in each. Similar to

Arabidopsis, most of the *OebZIP* genes belong to subgroups S and A, with each containing 25 and 19 *OebZIP* genes, respectively. Interestingly, this was also observed in rice, sesame, cacao, and olive, but not grape (Figure S1). Subgroup J and K were the smallest classifications, with each including 1 *OebZIP* gene.



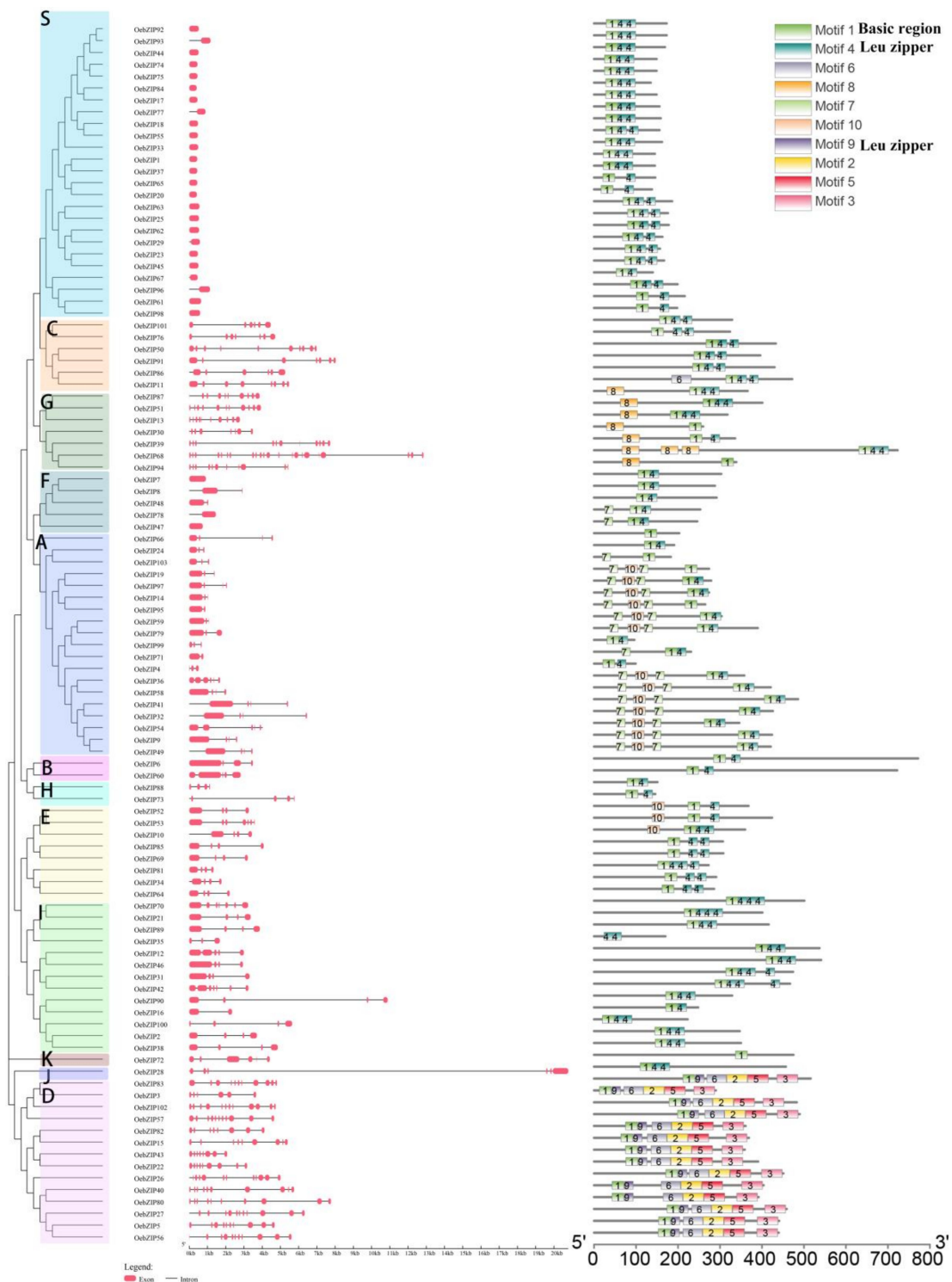
**Figure 1.** Phylogenetic tree of bZIP proteins from Arabidopsis and olive were built by the ML method. The *bZIP* genes from olive and Arabidopsis are marked in red and black, respectively. Plant bZIP members were divided into 12 subgroups (A~K, S). Different color blocks represent different subgroups. The bootstrap was set to 1000.

### 3.3. Gene Structure and Motif Composition of Olive bZIP

As shown in Figure 2, the intron/exon structure was detected using the Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/index.php/>). The number of exons of *OebZIP* genes varied from 1 to 22, of which 22 (21.4%) of the 103 *OebZIP* genes had only one exon, and most of them were fastened to subgroup S. Subgroups C, D, and G had multiple exons: 5 genes had a number of exons ranging from 6 to 10, 14 had between 6 to 12 exons, and 7 had between 8 to 22.

MEME 7.0 was used to detect the conserved motifs in 103 *OebZIP* proteins, and 10 motifs containing the bZIP domain were identified. As shown in Figure 2, most *OebZIP* proteins contained several motifs, while *OebZIP72*, *OebZIP94*, *OebZIP66*, *OebZIP103*, *OebZIP19*, and *OebZIP95* only contain motif 1, which means that only the basic region has no obvious Leu zipper structure. *OebZIP35* only contains motif 4, which only has a Leu zipper region and lacks the Basic region. Motif 8 is highly conserved and only exists in subgroup G, indicating that it is a conserved protein domain which may perform a special function. In subgroup A, except for *OebZIP4*, *OebZIP24*, and *OebZIP99*, all the others contained motifs 7 and 10. The distribution of the protein sequence was motif 7–motif 10–motif 7. *OebZIP* in Subgroup D contained the most motifs among the 13 subgroups, and was arranged in the order of motif 1–motif 9–motif 6–motif 2–motif 5–motif 3; among these, motifs 9, 2, 5 and 3 are unique

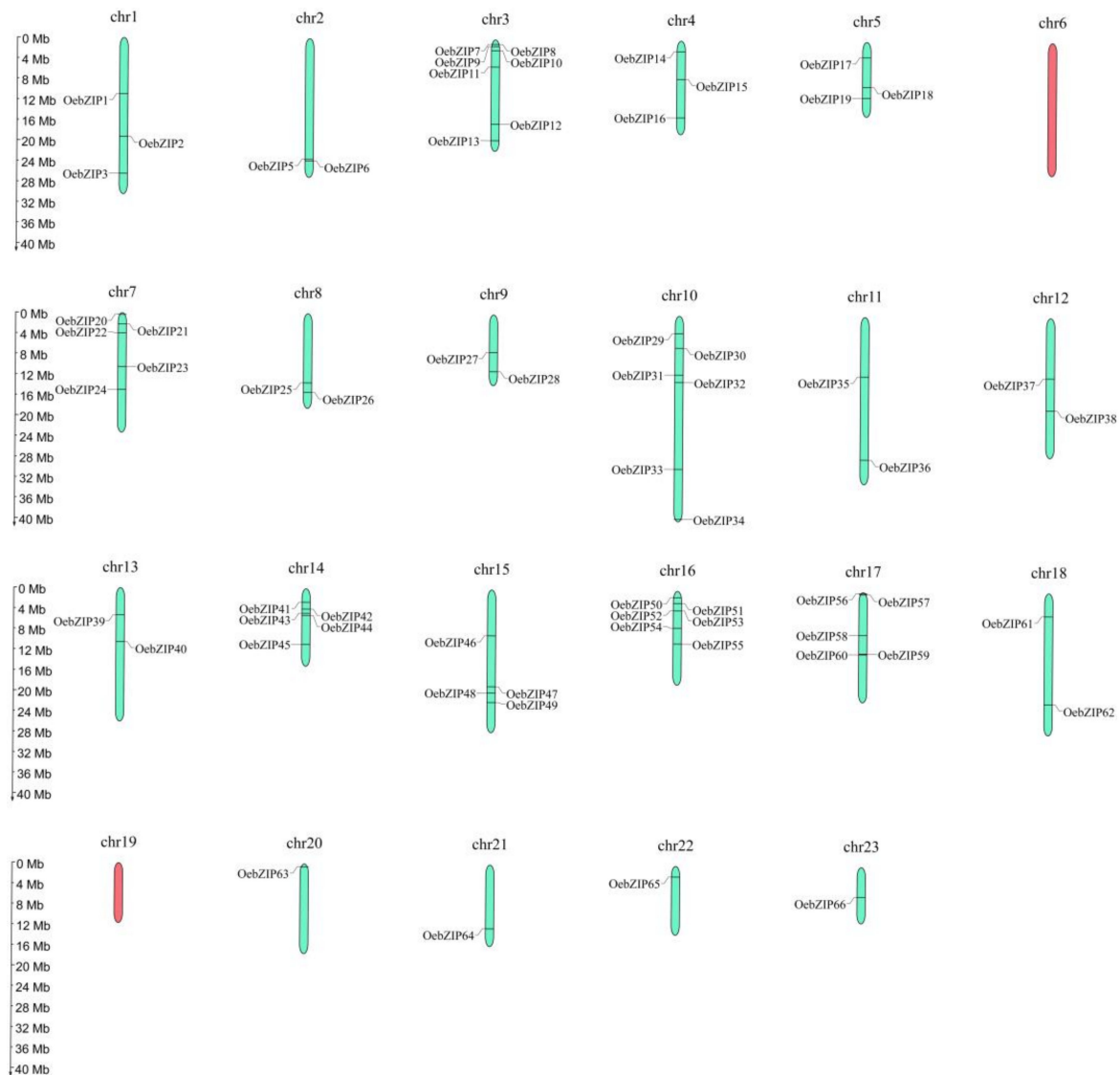
in subgroup D. In addition, the evolutionary tree clustering results are consistent with the intron/exon distribution and motif distribution. This indicated that most of *OebZIP* genes in the same subgroup showed low structural diversity, close evolutionary relationships and high degrees of conservation.



**Figure 2.** Phylogenetic tree. The distribution of the conserved protein motifs and gene structures of *OebZIP* genes. Different color blocks represent different subgroups. Red boxes indicate exon region, and black lines indicate intron. The motifs are numbered 1–10 and displayed in boxes of different colors. Motif 1 indicates the basic region of bZIP domain and motifs 4 and 9 indicate a Leu zipper of the bZIP domain. Detailed sequence information for each motif is shown in Figure S2.

### 3.4. Chromosomal Locations and Gene Collinearity Analysis of *OebZIP*

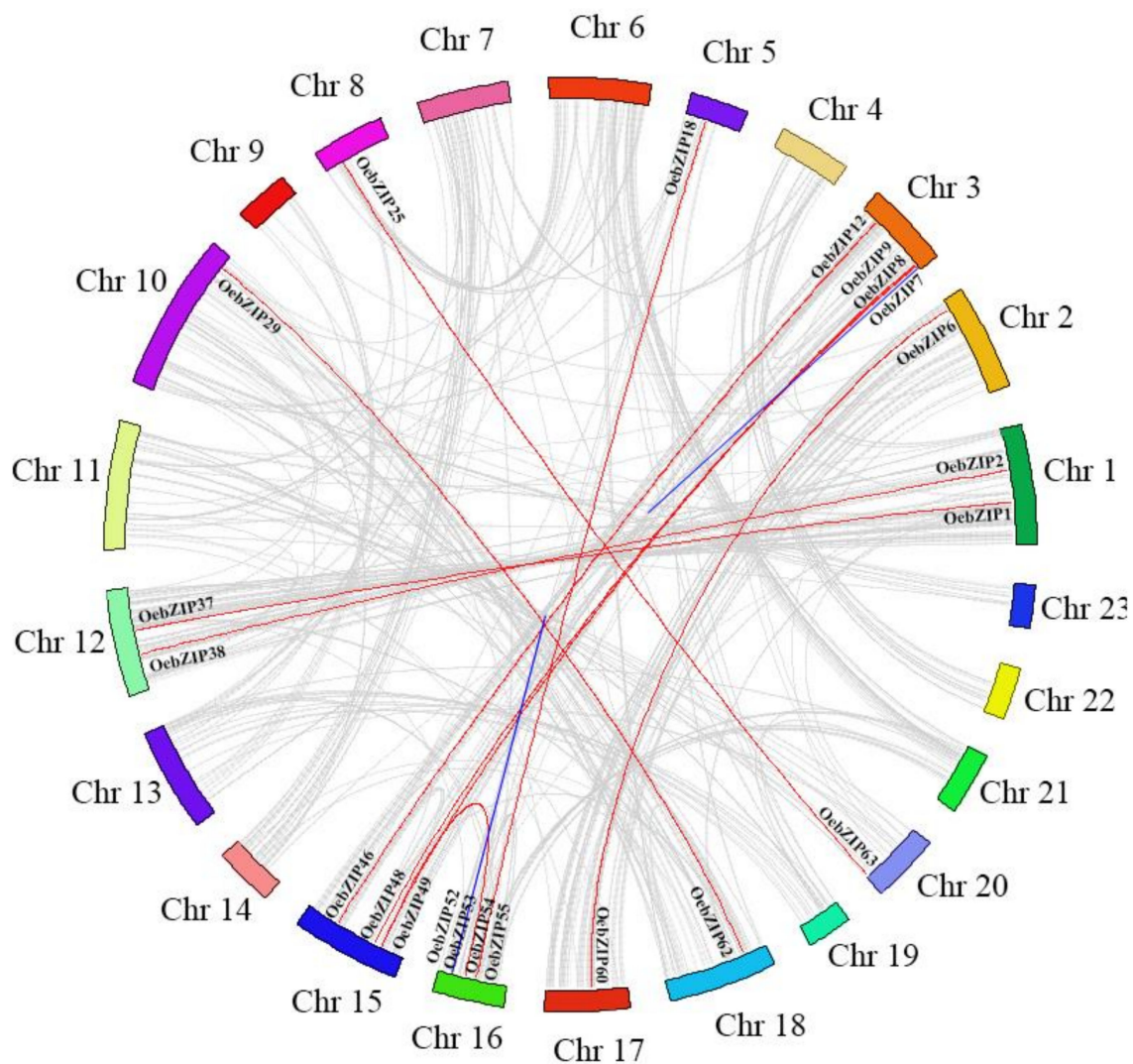
Sixty-seven *OebZIP* gene locations were mapped on the chromosome (Figure 3). Thirty-six *OebZIP* locations were not distributed on the chromosome but rather, on scaffolds, and no *OebZIP* genes were distributed on chr6 and chr19. In some specific regions of chr3 and chr16, *OebZIP* genes were densely distributed.



**Figure 3.** The distribution of *bZIP* genes on olive chromosomes. The red color represents chromosomes without *OebZIP* genes and the green is the chromosome that has *OebZIP* genes. The relative length of the chromosomes is millions of base pairs (Mb).

Gene replication is considered to be one of the main drivers of the evolution of the genome and genetic systems. Segment repeats and tandem repeats are the two main reasons for the expansion of plant gene families. In the olive genome, there were four pairs of tandem repeat events involved in 8 *OebZIP* genes; *OebZIP7* and *OebZIP8*, and *OebZIP52* and *OebZIP53* were distributed on chr3 and chr16, respectively, while *OebZIP74* and *OebZIP75*, and *OebZIP92* and *OebZIP93* were not distributed on the chromosome. As such, they were not represented graphically in Figure 4. Twenty-seven *OebZIP* genes were involved in 14 fragment repeats. A pair of genes that formed one gene duplication event was found to come from the same subgroup. As *OebZIP7* and *OebZIP8* are tandem genes, they were

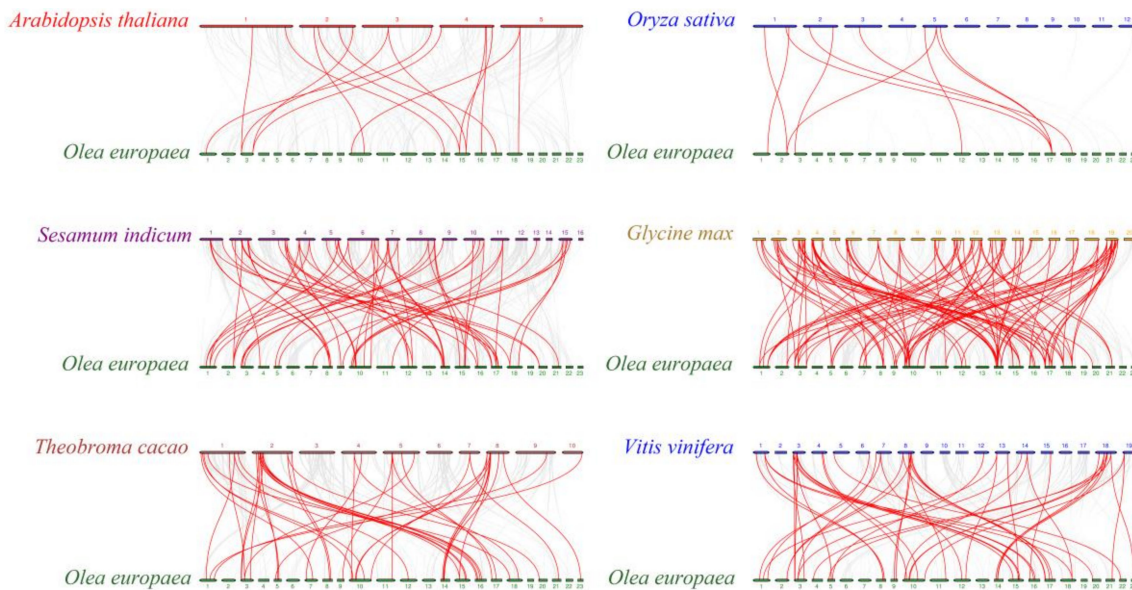
both assigned to subgroup F. Thus, 33.98% of the *OebZIP* genes were shown to be involved in gene duplication, which may be why so many *OebZIP* genes are caused by gene duplication.



**Figure 4.** Gene duplication events in the *OebZIP* gene family. The *OebZIP* genes linked by the red line are the segment repeats, while the *OebZIP* genes linked by the blue line are the tandem repeats. Different color blocks represent different chromosomes.

To explore the evolutionary relationship of *bZIP* genes in different species, we used McScanX to perform a gene collinearity analysis on five dicotyledonous plants: *Arabidopsis thaliana*, *Sesamum indicum*, *Glycine max*, *Vitis vinifera* L., and *Theobroma cacao*; and on a monocotyledon: *Oryza sativa* L. (Figure 5). Many reference plant *bZIP* genes had a colinear relationship with some *OebZIP* genes, i.e., 14 (*Arabidopsis*), 9 (rice), 38 (grape), 35 (*cacao*), 81 (soybean) and 51 (sesame). Soybean had the most *bZIP* genes that were colinear with olive *bZIP* genes, followed by sesame; rice had the fewest. In addition, 5 *OebZIP* genes (*OebZIP1*, *OebZIP29*, *OebZIP46*, *OebZIP33*, *OebZIP44*) were found to be colinear with five dicotyledons, and except for *OebZIP46*, the rest were all in subgroup S, indicating that these genes are highly conserved during evolution and play an important role in the amplification of the *bZIP* gene family in olive.





**Figure 5.** Synteny analysis of *OebZIP* with other species (*Arabidopsis thaliana*, *Sesamum indicum*, *Glycine max*, *Vitis vinifera* L., *Theobroma cacao*, and *Oryza sativa* L.). Gray lines represent collinear blocks within olive and other species, while red lines represent syntenic *bZIP* gene pairs.

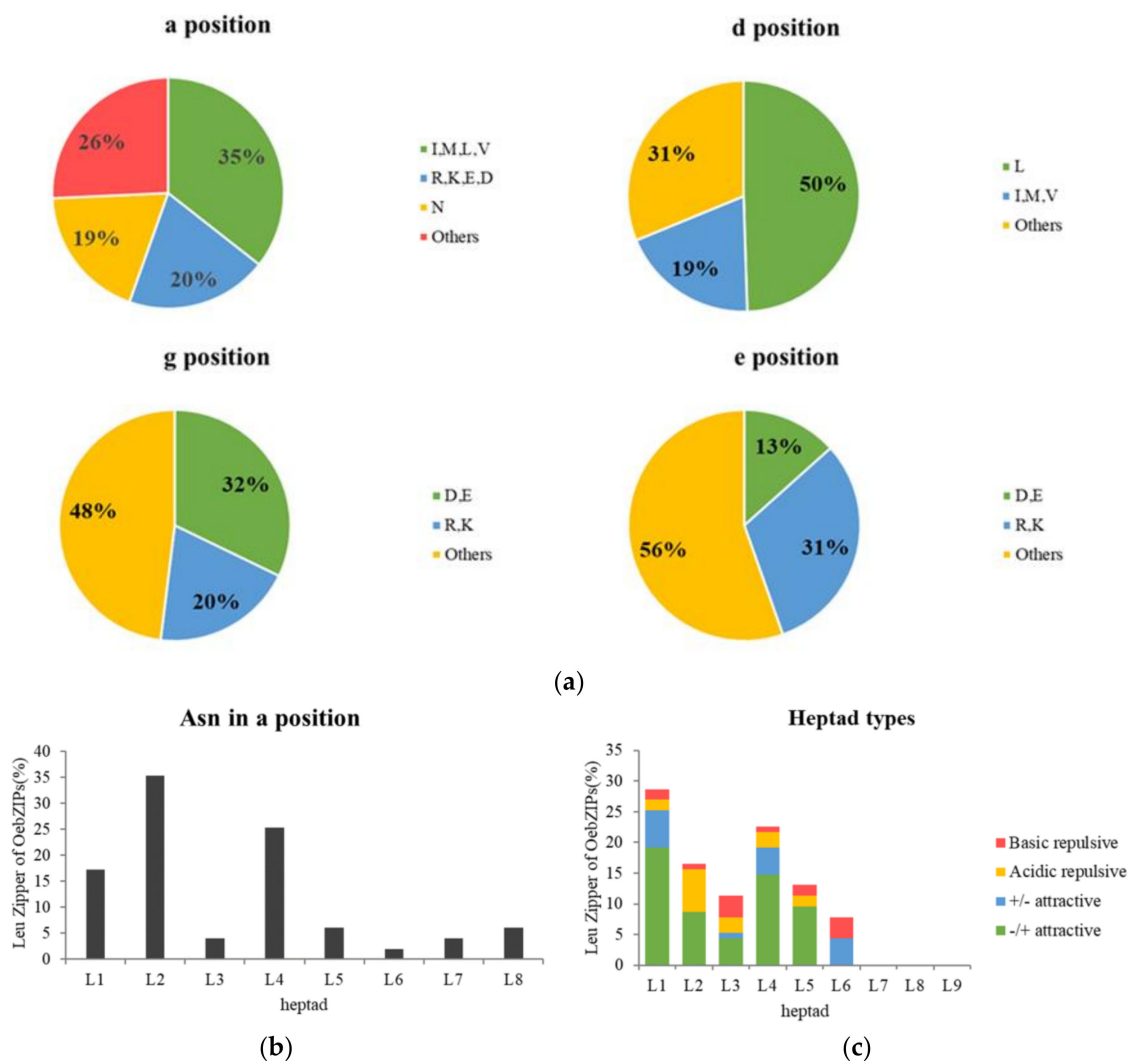
### 3.5. Prediction of Dimerization Properties

The structural characteristics of the Leu zipper region is L- $\chi_6$ -L- $\chi_6$ -L. Hydrophobic and electrostatic interactions mediate  $\alpha$ -helical oligomerization to form homo- or heterologous leucine zipper structures which are involved in the *bZIP* protein dimerization before binding to DNA. Studies have demonstrated that every seven amino acid residues of the leucine zipper region are set as a heptad, and that the arrangement order is *g, a, b, c, d, e, and f*. The possibility, stability, and specificity of the leucine zipper dimerization structure are mainly determined by four sites: *a, d, e, and g*. The *a* and *d* positions are usually hydrophobic amino acids. The two monomer  $\alpha$ -helical structures interact with each other through the hydrophobic amino acids at the *a* and *d* positions to form a hydrophobic inner core, enabling the dimer to exist stably. Asparagine at the *a* position can form a polar pocket on the hydrophobic interface. When *a'* is asparagine, it can form a very stable N-N structure ( $a \leftrightarrow a'$ ), so it tends not to interact with other amino acids, which limits the possibility of interactions between heterologous helices. The *d* position is generally a highly conserved leucine, which is one of the important factors affecting the dimerization structure. However, the *e* and *g* positions that flank the dimerization interface frequently contain charged amino acids including acidic amino acids glutamic acid (E) and aspartic acid (D), and the basic amino acids arginine (R) and lysine (K). As a result, salt bridges are formed to maintain stability between the spirals.

We performed a detailed dimerization analysis to characterize the amino acids present at positions *a, d, e, and g* of the *OebZIP* proteins (Figure S3). The length of the Leu zipper in the *OebZIP* family varied from two to nine heptads. At the *a* position, asparagine (Asn/N) accounted for 19%; meanwhile, at position *a*, the asparagine appeared most frequently in the second heptad ( $L_2$ ), followed by  $L_4$ , accounting for 35.35% and 25.25%, respectively. In addition, at position *a*, the frequency of hydrophobic amino acids (I, V, L, M) was 35%, and charged amino acids (R, K, E, D) was 20%. At the *d* position, the frequency of hydrophobic amino acids was 81%, of which 50% was leucine and 31% was other hydrophobic amino acids (I, V, M).

Four types of  $g \leftrightarrow e'$  interactions in each heptad were then analyzed, including attractive basic–acidic pairs (+/– attractive), attractive acidic–basic pairs (–/+ attractive), repulsive basic pairs, and repulsive acidic pairs (Figure 6). As shown in Figure 6c, the  $L_1$  heptad had a total of 32 of  $g \leftrightarrow e'$  pairs, of which 87.88% were attractive pairs. The  $L_4$  and  $L_5$  heptad had 26 and 15  $g \leftrightarrow e'$  pairs respectively. The

L<sub>6</sub> heptad only had attractive basic–acidic pairs (+/-). The L<sub>7</sub>, L<sub>8</sub>, and L<sub>9</sub> heptad had no complete  $g \leftrightarrow e'$  pair.



**Figure 6.** Prediction of the dimerization properties of OebZIP proteins. (a) Pie charts depicting Figure 3 for the amino acid positions within the Leu zipper regions. (b) Histogram of the frequency of Asn residues present at the position of each heptad within the Leu zippers for the OebZIP proteins. (c) Histogram of the frequency of attractive or repulsive  $g \leftrightarrow e'$  pairs per heptad within Leu zippers for OebZIP proteins.

Based on the dimerization features, OebZIP proteins were divided into three categories: (I) Those that tend to form homodimerization. OebZIP proteins in this category contained the  $g \leftrightarrow e'$  pairs and the asparagine in the *a* position. The members of subgroup A basically belong to this category; (II) Those with both homo- and heterodimerization. Most OebZIP belong to this category, such as those of subgroup C, which contained repulsive  $g \leftrightarrow e'$  pairs and attractive  $g \leftrightarrow e'$  pairs, or contain repulsive  $g \leftrightarrow e'$  pairs. (III) Those that tend to form heterodimerization. OebZIP proteins in this category contained repulsive  $g \leftrightarrow e'$  pairs only; for example, OebZIP30 contained three adjacent repulsive  $g \leftrightarrow e'$  pairs, but there were fewer OebZIPs in this category.

### 3.6. Gene Expression profiles of *OebZIPs*

#### 3.6.1. Gene Expression Profiles of *OebZIPs* on Transcription Data

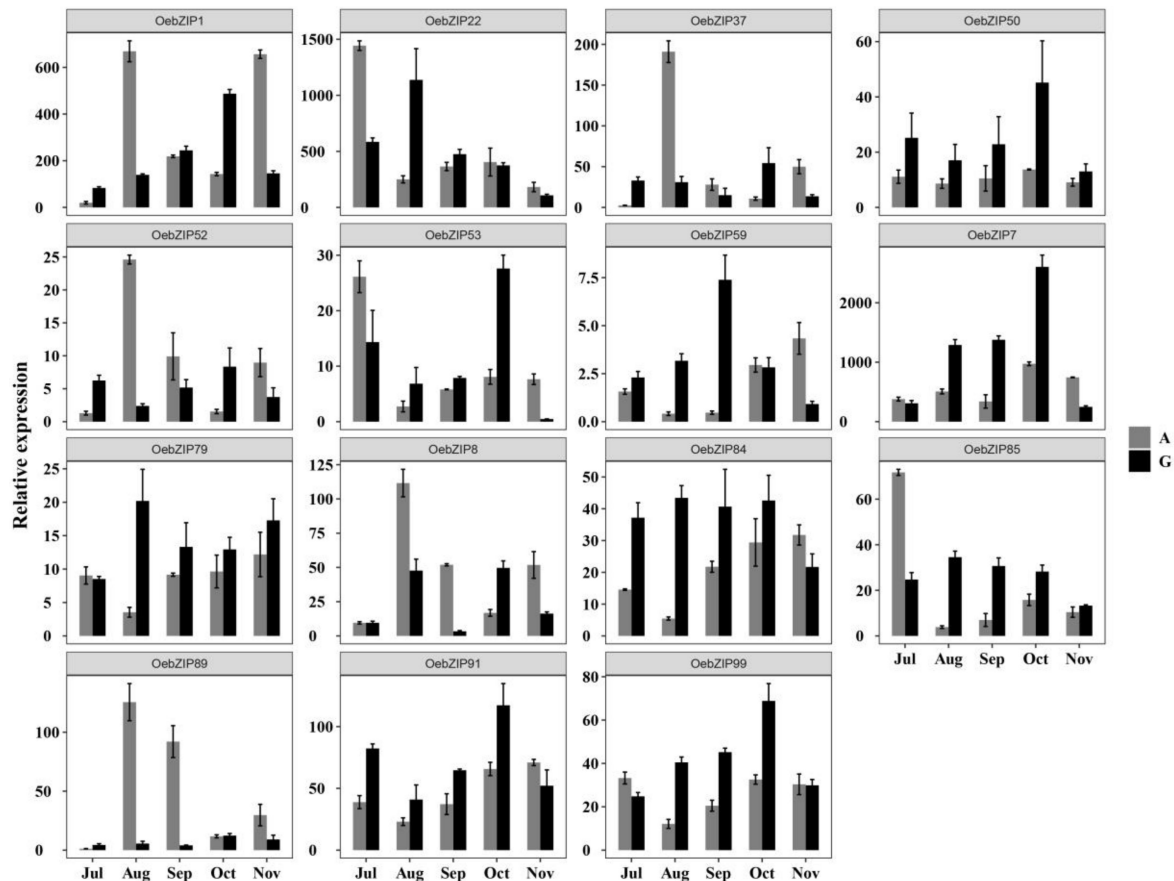
We analyzed the expression profiles based on transcriptome data including different periods and tissues. In July, 92, 95, 97, and 94 *OebZIP* genes were expressed, respectively, in fruits, pedicels, stems, and leaves; In November, 89, 96, 95, and 93 *OebZIP* genes were expressed, respectively, in fruits, pedicels, stems, and leaves. Among them, only one *OebZIP* gene (*OebZIP33*) was not expressed in four tissues in two periods, while the rest were expressed in at least one tissue. However, the gene expression abundance was significantly different. As shown in Figure S4, expression of most *OebZIP* genes was tissue-specific. For example, the *OebZIP84* gene was expressed in fruits only, while the *OebZIP62* gene was expressed in all tissues except for fruits. Moreover, most *OebZIP* genes were significantly up- or down- regulated in different periods. The highly expressed *OebZIP* genes in fruits may be involved in fruit oil accumulation, and there were 17 and 16 *bZIP* genes with higher expression in fruit in July (*OebZIP1*, 7, 8, 10, 12, 22, 27, 37, 50, 52, 53, 78, 84, 85, 91, 94, and 99) and November (*OebZIP22*, 39, 46, 48, 50, 52, 53, 59, 65, 70, 73, 79, 84, 92, 93, and 99), respectively. A total of 5 *OebZIP* genes (*OebZIP53*, 99, 22, 50, and 52) were found to be highly expressed in both periods; 12 *OebZIP* genes were upregulated and 15 *OebZIP* genes were downregulated. Therefore, 15 candidate *OebZIP* genes that possess higher expression in fruits were selected for the next analysis.

#### 3.6.2. Expression Patterns of *OebZIP* Genes in Different Tissues

To give an insight into the functional role of the *OebZIP* gene, we used qRT-PCR to detect the expression profiles of candidate *OebZIP* genes throughout the development of fruit, and the expression of 15 candidate *OebZIP* genes in the leaf and fruit of 'Arbequina' were analyzed (Figure S5). All candidate genes were expressed in both fruits and leaves, but the expression levels were significantly different between different tissues at various development periods. Similar to the result of the transcriptome data analysis, most of the *OebZIP* genes had high expression in fruits. The *OebZIP22*, *OebZIP53*, *OebZIP79*, *OebZIP84*, *OebZIP91*, and *OebZIP99* had higher expression abundance in fruits than leaves in all periods (July to November).

#### 3.6.3. Expression of *OebZIP* Genes in Different Varieties During Fruit Development

We found that different varieties of olive displayed a big difference in the speed of oil accumulation during the fast oil accumulation stages. The fruit development and oil accumulation period of some varieties were significantly earlier than those of others. Therefore, 'Arbequina' ('A') and 'Grossa' ('G') were selected for the real-time PCR, and to explore the candidate *OebZIP* genes among two varieties. As shown in Figure 7, the expression of candidate genes between the two varieties ('A' and 'G') was quite different. We found that some *OebZIP* genes (*OebZIP1*, *OebZIP22*, *OebZIP37*, *OebZIP52*, *OebZIP53*, *OebZIP8*, *OebZIP85*, and *OebZIP89*) in 'A' were concentrated in the early stages of oil accumulation (July to August). Seven *OebZIP* genes (*OebZIP1*, *OebZIP7*, *OebZIP50*, *OebZIP53*, *OebZIP59*, *OebZIP91*, and *OebZIP99*) in variety 'G' were concentrated in the middle stages (September to October) of oil accumulation. Furthermore, the expression levels of five *OebZIP* genes (*OebZIP1*, *OebZIP7*, *OebZIP59*, *OebZIP84*, *OebZIP85*, and *OebZIP85*) in variety 'G' increased dramatically from July to October, reached a peak in September or October, and then decreased rapidly. This observation is similar to the rate of oil accumulation in 'G', varieties indicating that these *OebZIP* genes have tight functions in olive fruit development and lipid synthesis. Interestingly, the expression level of the *OebZIP22* gene in 'A' in July was the same as that of the *OebZIP22* gene in 'G' in August, the expression levels and dynamic changes of the two varieties were again similar, indicating that *OebZIP22* gene may play an important role in fruit development, given the different fruit development periods of the two varieties.



**Figure 7.** The expression level of 15 *OebZIP* genes in different varieties during fruit development. The expression level of 15 *OebZIP* genes in fruit of different varieties from July to November were examined by qRT-PCR. The A indicates 'Arbequina', and the G indicates 'Grossa'. The error bars were obtained from the mean values of three replicates  $\pm$  standard deviation (SD).

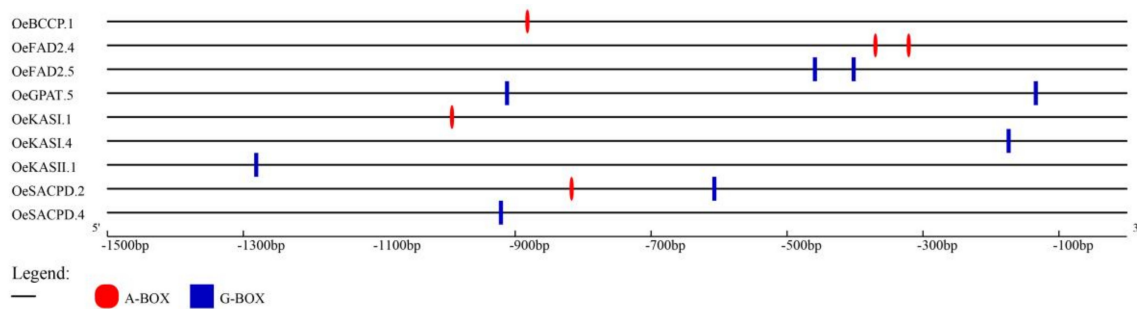
In addition, although two pairs of tandem genes (*OebZIP7* & *OebZIP8*, *OebZIP52*, and *OebZIP53*) and one pair of fragment genes (*OebZIP1* and *OebZIP37*) had sequence similarities, their expression patterns and abundances were disparate, indicating that the repeated genes in olive achieved functional differentiation during their evolution.

### 3.7. Predicting Interaction with Lipid Synthesis Genes

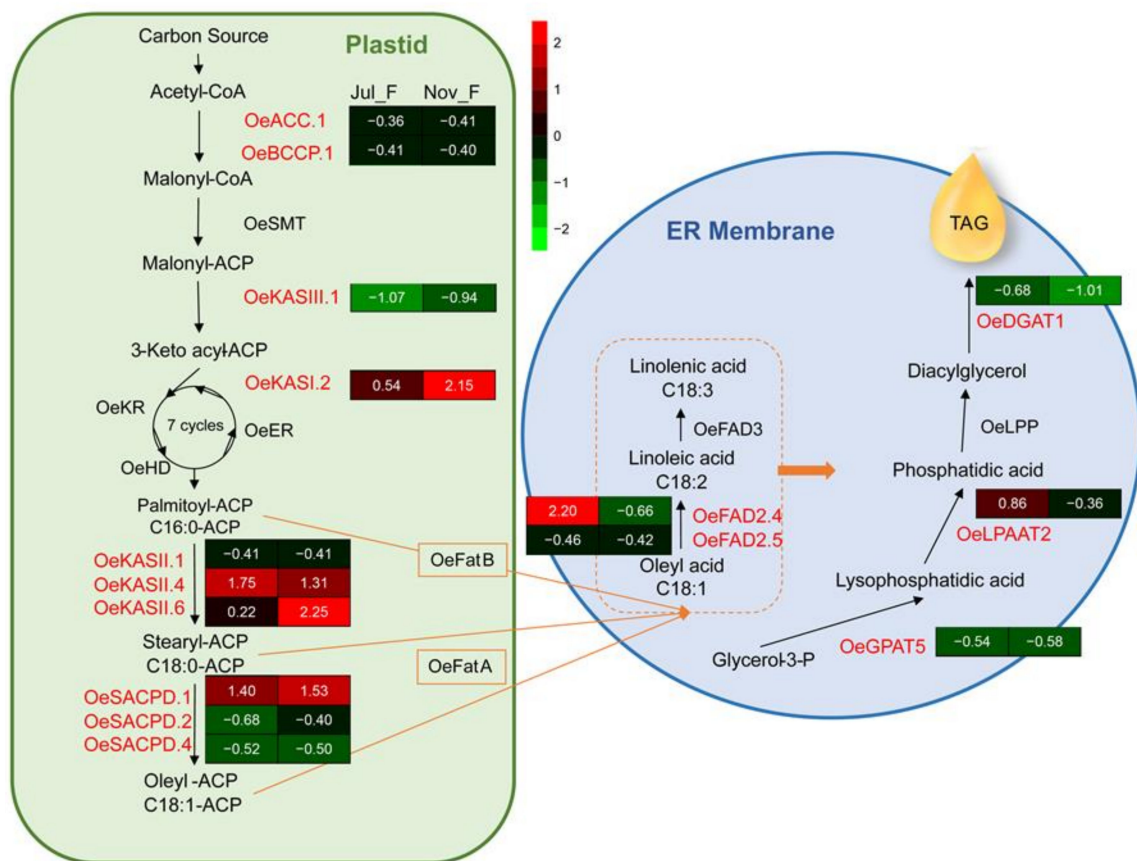
Since the expression of some *OebZIP* genes was found to be significantly different between the two varieties, these genes might correlate with oil accumulation. To preliminarily investigate this interaction, we referred to previous studies and scanned 44 lipid synthesis genes (*LPGs*) named according to their chromosomal location from the olive genome (Table S2).

#### 3.7.1. Expression Patterns of Lipid Synthesis Genes

Based on the abundance of genes expression (Figure S7) and the prediction of upstream cis-acting elements (Figure 8a), 15 genes (*OeACC1*, *OeBCCP2*, *OeDGAT1*, *OeFAD2.4*, *OeFAD2.4*, *OeGPAT5*, *OeKASI.2*, *OeKASII.1*, *OeKASII.4*, *OeKASII.6*, *OeKASIII.1*, *OeLPPAT2*, *OeSACPD1*, *OeSACPD2*, and *OeSADCP4*) were selected. These basically covered most of the genes involved in the lipid synthesis pathway (Figure 8b); therefore, these genes were selected for detailed qRT-PCR analysis.



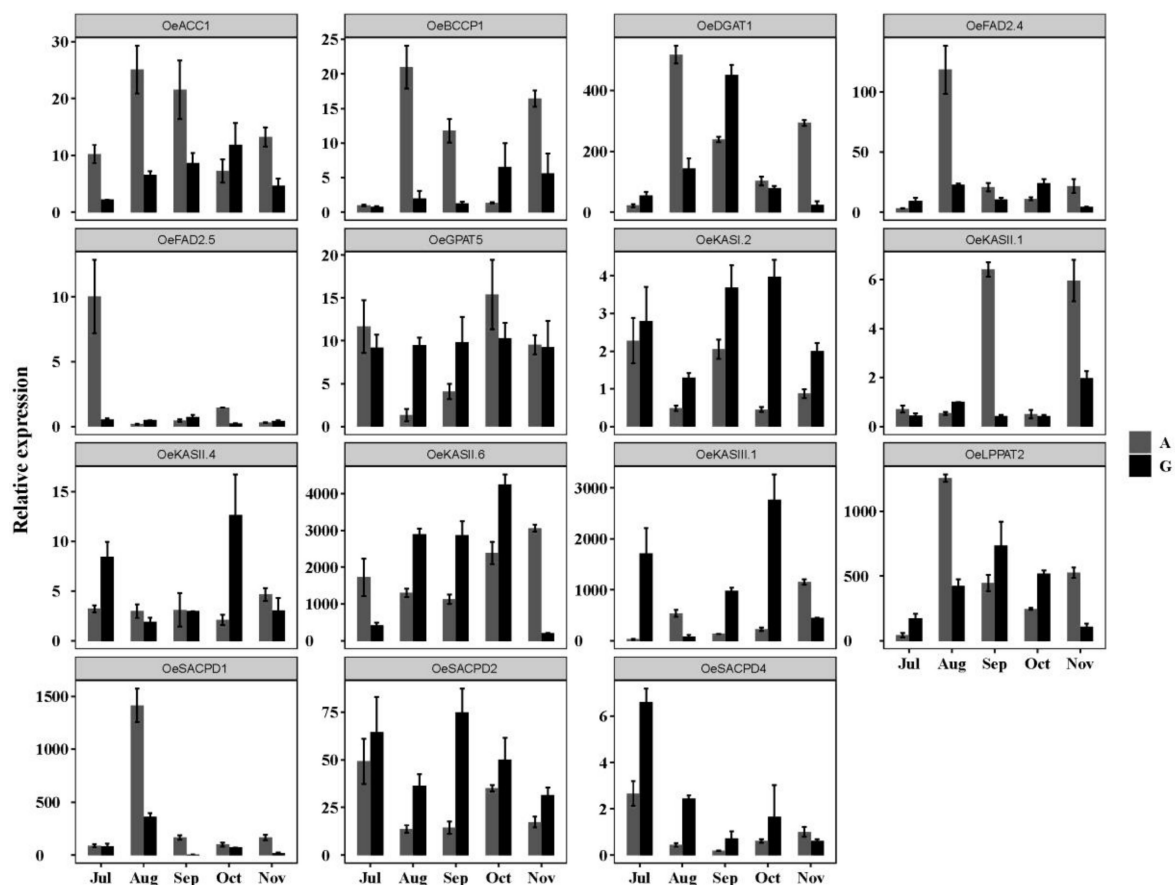
(a)



(b)

**Figure 8.** The prediction of upstream cis-acting elements of *LPG* genes and the triacylglycerol (TAG) biosynthesis pathway in olive. (a) The promoter sequences (−1500 bp) of 9 *LPG* genes were analyzed. The elements were identified as follows: red for the A-box element (TACGTA) and blue for the G-box (CACGTG) element. The scale at the bottom indicates the number of the nucleotides to the translation initiation codon. (b) Red text represents the key enzymes involved in the study, while black represents those not involved. ACC: acetyl-CoA carboxylase; BCCP: biotin carboxyl carrier protein; SMT: S-malonyl-transferase; KAS: β-ketoacyl-ACP synthase; KR: 3-ketobutyryl-ACP reductase; HD: 3-hydroxybutyryl-ACP dehydratase; ER: enoyl-ACP reductase; FatA: fatty acid-ACP thioesterase A; FatB: fatty acid-ACP thioesterase B; SACPD: stearyl-ACP desaturase; FAD: fatty-acid desaturase; GPAT: glycerol-3-phosphateacyltransferase; LPAAT: lysophosphatidate acyltransferase; LPP: phosphatidate phosphohydrolase; DGAT: diacylglycerol acyltransferase. Green indicates low expression, dark indicates intermediate expression, and red indicates high expression. The numbers in the boxes indicate normalized scale.

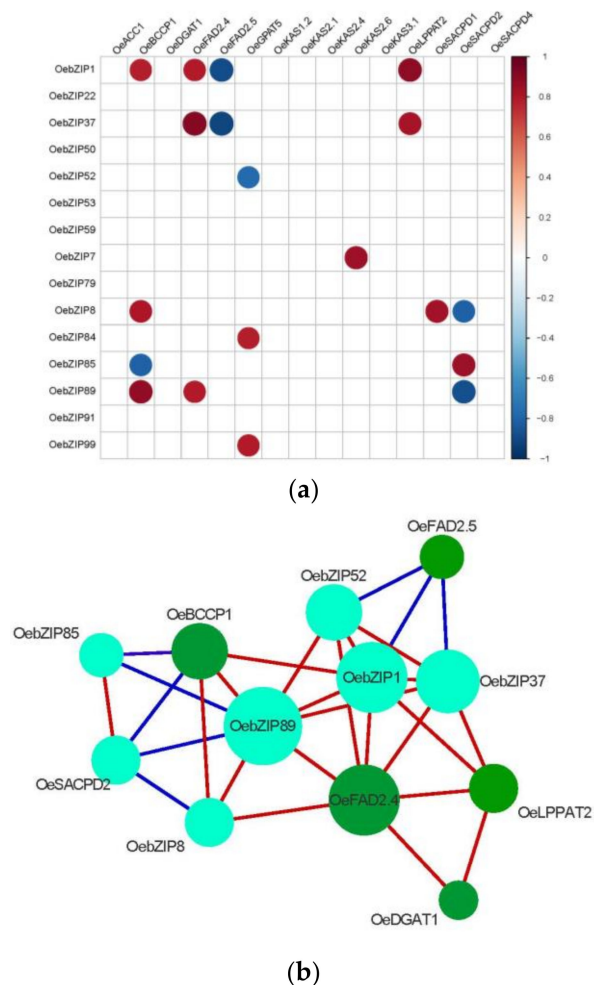
The expression patterns between *LPGs* were significantly different (Figure 9). *OeKASII.6* ( $\beta$ -ketoacyl-ACP synthetase II, C16~C18), *OeKASIII.1* ( $\beta$ -ketoacyl-ACP synthetase III, C2~C4), *OeLPAAT2* (lysophosphatidate acyltransferase) and *OeDGAT1* (diacylglycerol acyltransferase) had higher expression abundance than the other *LPG* genes in both varieties. Moreover, the expression patterns of different members in the gene family were also significantly different. For example, the *OeKASII* and *OeSACPD* families had different expression patterns. Some genes (*OeFAD2.5*, *OeSACPD4*, and *OeKASII.4*) maintained an extremely low level of expression in two varieties, indicating that they may be less correlated with the regulation of fruit development and lipid synthesis. In addition, several *LPGs* had different expression patterns between the two varieties, which might be caused by the difference in the period of oil accumulation, as described above. There were 7 genes in 'A' (*OeACC1*, *OeBCCP1*, *OeDGAT1*, *OeFAD2.4*, *OeFAD2.5*, *OeLPPAT2*, and *OeSACPD1*) that were significantly expressed in the early stage (July to August). In contrast, in 'G' there were 6 genes (*OeACC1*, *OeBCCP1*, *OeDGAT1*, *OeGPAT5*, *OeKASII.6*, *OeLPPAT2*) that were upregulated progressively throughout the oil accumulation until they reached a peak in the middle stages of oil accumulation (September to October), and then downregulated rapidly. Compared with the expression patterns of *OebZIPs* and *LPGs*, the result indicates that some *OebZIPs* (*OebZIP1*, *OebZIP7*, *OebZIP91*, and *OebZIP99*) and *LPGs* (*OeACC1*, *OeDGAT1*, *OeLPPAT2*, and *OeKASII.6*) showed similar expression patterns in different varieties.



**Figure 9.** Expression levels of 15 *LPG* genes during fruit development in different varieties. qRT-PCR was used to detect the expression level of the *LPG* gene in fruits of different varieties from July to November. A represents 'Arbequina', and G represents 'Grossa'. The error bars indicate the mean  $\pm$  standard deviation (SD) of three replicates.

### 3.7.2. Correlation Analysis of *OebZIP* with Lipid Synthesis Genes

Based on qRT-PCR experiment data of the two varieties, a correlation analysis by *R* is summarized in Figure 10. We determined that 7 *OebZIP*s had a significant positive correlation with various *LPG* genes (*OebZIP1* and *OeBCCP1/OeFAD2.4/OeLPPAT2*; *OebZIP7* and *OeKASII.6*; *OebZIP8* and *OeBCCP1/OeSACPD1*; *OebZIP37* and *OeFAD2.4/OeLPPAT2*; *OebZIP84* and *OeGPAT5*; *OebZIP85* and *OeSAD2*; *OebZIP89* and *OeBCCP1/OeFAD2.4*; *OebZIP99* and *OeGPAT5*). In contrast, 6 *OebZIP* genes had a significant negative correlation with *LPG* genes (*OebZIP1* and *OeFAD2.5*; *OebZIP37* and *OeFAD2.5*; *OebZIP52* and *OeGPAT5*; *OebZIP8* and *OeSAD2*; *OebZIP85* and *OeBCCP1*; *OebZIP89* and *OeSAD2*).



**Figure 10.** Correlation analysis of *OebZIP* with lipid synthesis genes (a) Red dot indicates a positive correlation; blue dot indicates a negative correlation, while the different colors represent the *p*-values of the correlation. (b) The predicted interaction network of *OeLPGs* and *OebZIPs*. Red lines are positive regulations, blue lines are negative.

## 4. Discussion

### 4.1. Evolutionary Analysis of the *OebZIP* Gene Family

Olive is a very important oil crop around the world. With the completion of olive genome sequencing and the rapid development of bioinformatics, the mining and identification of gene families based on genome-wide searches have become more and more effective. Therefore, it is necessary to explore the structural and evolutionary characteristics of *OebZIP* genes from the level of gene families. Although the *bZIP* gene family has been studied in many plants, this is not the case for olive. In this study, a total of 103 *OebZIP* genes were identified and screened based on the olive genome. The

average length of *OebZIP* proteins is 315 aa. With respect to numbers and aa length, compared to other plants, Arabidopsis (78, 321 aa) [38], rice (89, 311 aa) [29], sesame (63, 339 aa) [30], soybean (160, 343 aa) [31], grape (47, 336 aa) [32], and apple (114, 411 aa) [35], the number of *bZIP* family members in olive is larger and similar to apple, while the coding sequence length of the genes is clearly different. Several studies [53,54] have found that there is no positive correlation between the numbers of *bZIP* genes and the genome size. The differences in the number of *bZIP* members in plants might be due to whole-genome duplication (WGD) events [37]. Phylogenetic tree analysis with the Arabidopsis *bZIP* protein, which was divided into 12 subgroups, showed that the *AtbZIP* genes have a close kinship with the *OebZIP* genes. The number of identified *OebZIP* genes in the subgroups were similar to other plants (Arabidopsis, rice, sesame, soybean, and cocoa) except grape. Additionally, we found that subgroup H contains 2 *OebZIP* genes, similar to Arabidopsis *bZIP*. This indicates strong homology, and that subgroup H (*OebZIP73*, *OebZIP88*) may play an important role in photomorphogenesis.

Gene replication is considered to be one of the main drivers of genetic system evolution. The two main reasons for the expansion of plant gene families are tandem repeats and segment repeats. Twenty-two olive, 119 soybean, 8 tartary buckwheat, and 28 grape *bZIP* genes are involved in gene duplication events (Figure 4), accounting for around 21.4% (olive), 86.2% (soybean) [31], 8.3% (tartary buckwheat) [54], 59.6% (grape) [32]. The contribution of gene duplication events to the amplification of gene families was significantly different among different species. In addition, by comparing the quantity of the tandem repeats, i.e., 4 pairs of olive, 1 pair of soybean [31], 2 pairs of tartary buckwheat [54], 0 pairs of grape, we concluded that the contribution of tandem repeats during the evolution of *OebZIP* genes was limited, which is consistent with reports on grape, soybean, and tartary buckwheat. Therefore, segment gene duplication plays an important role in *bZIP* gene family expansion in olive. Soybean had the most *bZIP* genes that were colinear with olive *bZIP* genes (Figure 5) and tartary buckwheat *bZIP* genes [54], which may be due to the large genome of soybean and the large number of *GmbZIP* genes.

Most *OebZIP* genes in the same subgroup possess the same features (Figure 2), as also reported for apple [35], melon [55], and cassava [33], i.e., the *OebZIP* genes in each S subgroup have one exon and the motif distribution is simple, just containing motif 1 (basic area) and motif 4 (Leu area). Some specific motifs (motif7, motif10) only exist in a specific subgroup (subgroup A). Meanwhile, the order of motif arrangement in the same subgroup is consistent. Furthermore, motif 8 was found to be quite conserved and focused on subgroup G, indicating this motif might have some specific function.

The function of *bZIP* proteins requires dimerization. We investigated the dimerization patterns of *OebZIP* proteins, which are characterized mainly by the four amino acids presenting at the *a*, *d*, *g*, and *e* positions (Figure 6). At the *d* position, the frequency of Leu aa accounts for 50%, which is close to *AtbZIPs* (56%), but significantly less than in *OsbZIPs* (71%) [29], *ZmbZIPs* (70%) [56], and *GmbZIPs* (68%) [31]. The frequency of Asn at position *a* was 19%, which was lower than that found in *AtbZIP* (40%) [51] and *GmbZIP* (25%) [31]. Another difference is that in the olive, the frequency of Asn at the *a* position was the highest in the L<sub>2</sub> heptad, followed by the L<sub>4</sub> heptad, while in [31] Arabidopsis [51] and strawberry [34], the frequency of Asn in the *a* position was highest in the L<sub>2</sub> heptad, followed by the L<sub>5</sub> heptad. Therefore, it was shown that the *bZIPs* in olive are more prone to heterodimerization, and the length of Leu zipper is shorter compared to soybean [31], Arabidopsis [51], and strawberry [34].

#### 4.2. The Expression Pattern of *OebZIP* was Related to Fruit Development and Lipid Synthesis

Plants have a large number of *bZIP* genes and diverse functions, but investigations of *bZIP* functions have been mainly focused on some model plants [38]. Many study already found that the *bZIP* gene family plays a role in all stages of plant growth and response to biotic/abiotic stresses [18–20,24–26,28,57–63]. However, very few *bZIP* genes involved in fruit development and lipid synthesis have been identified in plants, especially in olive. We explored the possibility of the involvement of candidate *OebZIP* genes in regulating lipid synthesis and gene expression levels of *LPG* genes during fruit development. In a previous study, the oil content of different olive varieties was significantly



differentiated in the fast oil accumulation stage (Figure S6). Several *LPGs* have different expression patterns between the two varieties; this result is similar to that of a study on the gene expression levels of *OeFAD2* [64], *OeFAD3* [65], and *OeSACPD* [66] during fruit development between two olive varieties ('Picual' and 'Arbequina'). The expression levels of some *LPGs*, such as *OeACCI1*, *OeDGAT1*, *OeKASII.6*, and *OeLPAAT2*, were consistent with the oil accumulation trend. The oil in 'A' accumulated mainly in early stage fruit development, while the oil 'G' accumulated mainly in the late stage. The gene expression pattern of *OeDGAT1* in 'G' was consistent with a previous report of *OeDGAT1* in 'Koroneiki' [67], presenting a typical normal curve, i.e., following maximal mRNA levels in September and then declining substantially. Although most TAGs in olive accumulated before they matured, it has been demonstrated that some oil cultivars, like 'A' in our study, have much shorter oil-filling periods than 'Koroneiki' and 'G' [68–70] (Figure S6). Olive oil is known for its high oleic acid content. Several researchers have reported that the composition of olive oil fatty acids, particularly oleic acid, fluctuated according to the variety [68,71,72]. The variation of *OeFAD2* [64,73], *OeFAD3* [65], and *OeSACPD* [66,74] gene expression is the main factor that affects the composition of fatty acids [37]. The gene expression levels of *OeSACPD1* and -2 in 'A' in our study were similar to those in a previous report in the same variety, peaking in the early stages (July to August) and then decreasing [66]. Moreover, the expression level of *OeFAD2.2*, which is considered to be mainly responsible for the linoleic acid content in olive oil, as reported by Hernandez [64] and Bruno [73], was higher than that of *OeFAD2.5* in our study, indicating that *OeFAD2.5* may not be a significant factor affecting the oil content of mesocarp.

The goal of this study was to obtain more insight into the expression patterns and functions of *OebZIP* genes during fruit development and lipid synthesis. An expression profile analysis indicated that *OebZIP* genes expression showed significant variation among different tissues, which implied that they might have diverse functions. It is therefore hypothesized that highly expressed genes in fruits are closely related to fruit development and lipid synthesis. In previous studies, some candidate *OebZIPs* expressed in fruits showed different expressive patterns in the two varieties. In addition, paired gene duplication events showed different expression patterns, suggesting that functional differences had occurred during evolution, as similarly demonstrated in previous reports on rice and grape.

Only four genes, i.e., *OebZIP1*, *OebZIP7*, *OebZIP91*, and *OebZIP99*, were significantly upregulated in the fast oil accumulation stage between the two varieties. In Arabidopsis, *AtbZIP53* was reported to be involved in regulating seed development, and seed maturation was affected when *AtbZIP53* [28] was overexpressed or knocked out. In olive, *OebZIP1* was homologous to *AtbZIP* genes, and *OebZIP1* was clustered into the same group with *AtbZIP53* in the phylogenetic tree (Figure 1). Therefore, we speculate that the *OebZIP1* may also regulate olive fruit development and oil accumulation. In addition, *OebZIP22* was found to have a similar expression pattern between the two varieties, but the expression of *OebZIP22* in 'A' took precedence over that of 'G'. *OebZIP22* is a homolog of *TGA3* (*AtbZIP22*). In Arabidopsis, *TGA3* was involved in mediating hormonal cross-talk between salicylic acid and cytokinin [75]. Therefore, *OebZIP22* may be involved in hormone regulation in fruit development and lipid synthesis. In addition, through a correlation analysis, we found that the expression patterns of some *OebZIP* genes in fruits of different varieties were significantly correlated with oil synthesis genes (Figure 10). For example, there was a significant positive correlation between the expression of *OebZIP1*, *OebZIP37*, *OebZIP 89*, and *OeFAD2.4*. In addition, there were two a-boxes upstream of the *OeFAD2.4* promoter region, and the possible interactions between *OeFAD2.4* and *OebZIP1*, 37, and 89 transcription factors might be the focus of future research. The functions of these *OebZIP* genes need to be verified by experiments in the future.

## 5. Conclusions

In this study, for the first time, we identified 103 *bZIP* genes based on the olive genome, which were divided into 12 subgroups. We analyzed the genetic structure, conserved domain distribution, the dimerization feature, the evolutionary relationship and the gene replication events. Meanwhile, some important candidates of *OebZIP* genes for olive fruit development and lipid synthesis were identified.

This study laid the foundation for the study of the structure and function of the *bZIP* gene family in olive. It provides important clues for increasing the oil content of olive from the direction of *bZIP* TF regulation of lipid synthesis.

**Supplementary Materials:** Supplementary Materials can be found at <http://www.mdpi.com/2073-4425/11/5/510/s1>, Figure S1: Classification of *bZIP* genes in six plants, Figure S2: A detailed motif introduction for all *OebZIP* proteins, Figure S3: Amino acid sequences of the Leu zipper region of *OebZIP* proteins, Figure S4: Heatmap of 103 *OebZIP* genes expressed among four tissues based on transcription data, Figure S5: Relative expression of *OebZIP* genes in fruit and leaf, Figure S6: Dried fruit oil content during the fast oil accumulation stages (July to November), Figure S7: Heatmap of 44 *LPG* genes expressed among four tissues based on transcription data. Table S1: The basic characteristics of olive *bZIP* gene family, Table S2: The chromosomal location of olive *LPG* genes.

**Author Contributions:** Conceptualization, Q.H.; Data curation, S.R. and Z.C.; Investigation, S.R., Z.W. and S.Z.; Methodology, S.R., Z.W., Z.C., S.Z. and H.L.; Project administration, Q.H.; Software, Z.C.; Supervision, Q.H.; Writing—original draft, S.R.; Writing—review & editing, Z.W., Z.C., S.Z., H.L. and Q.H. All authors have read and agree to the published version of the manuscript.

**Funding:** This work was funded by Research Fellowships from the Sichuan Science and Technology Department (12ZC2220).

**Acknowledgments:** We thank to all those who provided useful discussion and technical support for our article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Gouvinhas, I.; Machado, N.; Sobreira, C.; Dominguez-Perles, R.; Gomes, S.; Rosa, E.; Barros, A. Critical Review on the Significance of Olive Phytochemicals in Plant Physiology and Human Health. *Molecules* **2017**, *22*. [[CrossRef](#)] [[PubMed](#)]
- Ghanbari, R.; Anwar, F.; Alkharfy, K.M.; Gilani, A.H.; Saari, N. Valuable nutrients and functional bioactives in different parts of olive (*Olea europaea* L.)—A review. *Int. J. Mol. Sci.* **2012**, *13*, 3291–3340. [[CrossRef](#)] [[PubMed](#)]
- El, S.N.; Karakaya, S. Olive tree (*Olea europaea*) leaves: Potential beneficial effects on human health. *Nutr. Rev.* **2009**, *67*, 632–638. [[CrossRef](#)] [[PubMed](#)]
- Rezaei, S.; Akhlaghi, M.; Sasani, M.R.; Boldaji, R.B. Olive oil improved fatty liver severity independent of cardiometabolic correction in patients with non-alcoholic fatty liver disease, a randomized clinical trial. *Nutrition* **2019**, *57*, 154–161. [[CrossRef](#)] [[PubMed](#)]
- Chiu, C.C.; Su, K.P.; Cheng, T.C.; Liu, H.C.; Chang, C.J.; Dewey, M.E.; Stewart, R.; Huang, S.Y. The effects of omega-3 fatty acids monotherapy in Alzheimer's disease and mild cognitive impairment: A preliminary randomized double-blind placebo-controlled study. *Prog. Neuro Psychopharmacol. Biol. Psychiatr.* **2008**, *32*, 1538–1544. [[CrossRef](#)]
- Manan, S.; Chen, B.; She, G.; Wan, X.; Zhao, J. Transport and transcriptional regulation of oil production in plants. *Crit. Rev. Biotechnol.* **2017**, *37*, 641–655. [[CrossRef](#)]
- Pruneda-Paz, J.L.; Breton, G.; Nagel, D.H.; Kang, S.E.; Bonaldi, K.; Doherty, C.J.; Ravelo, S.; Galli, M.; Ecker, J.R.; Kay, S.A. A genome-scale resource for the functional characterization of Arabidopsis transcription factors. *Cell Rep.* **2014**, *8*, 622–632. [[CrossRef](#)]
- Yamasaki, K.; Kigawa, T.; Seki, M.; Shinozaki, K.; Yokoyama, S. DNA-binding domains of plant-specific transcription factors: Structure, function, and evolution. *Trends Plant Sci.* **2013**, *18*, 267–276. [[CrossRef](#)]
- Liu, Y.F.; Li, Q.T.; Lu, X.; Song, Q.X.; Lam, S.M.; Zhang, W.K.; Ma, B.; Lin, Q.; Man, W.Q.; Du, W.G.; et al. Soybean GmMYB73 promotes lipid accumulation in transgenic plants. *BMC Plant Biol.* **2014**, *14*, 23–31. [[CrossRef](#)]
- Raffaele, S.; Vaillieu, F.; Leger, A.; Joubes, J.; Miersch, O.; Huard, C.; Blee, E.; Mongrand, S.; Domergue, F.; Roby, D. 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*, 752–767. [[CrossRef](#)]
- Kim, H.U.; Jung, S.J.; Lee, K.R.; Kim, E.H.; Lee, S.M.; Roh, K.H.; Kim, J.B. Ectopic overexpression of castor bean LEAFY COTYLEDON2 (LEC2) in Arabidopsis triggers the expression of genes that encode regulators of seed maturation and oil body proteins in vegetative tissues. *FEBS Open Bio.* **2013**, *4*, 25–32. [[CrossRef](#)] [[PubMed](#)]

12. Pouvreau, B.; Baud, S.; Vernoud, V.; Morin, V.; Py, C.; Gendrot, G.; Pichon, J.P.; Rouster, J.; Paul, W.; Rogowsky, P.M. Duplicate maize Wrinkled1 transcription factors activate target genes involved in seed oil biosynthesis. *Plant Physiol.* **2011**, *156*, 674–686. [[CrossRef](#)] [[PubMed](#)]
13. Shen, B.; Allen, W.B.; Zheng, P.; Li, C.; Glassman, K.; Ranch, J.; Nubel, D.; Tarczynski, M.C. Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize. *Plant Physiol.* **2010**, *153*, 980–987. [[CrossRef](#)]
14. Yang, Y.; Munz, J.; Cass, C.; Zienkiewicz, A.; Kong, Q.; Ma, W.; Sedbrook, J.; Benning, C. Ectopic Expression of WRINKLED1 Affects Fatty Acid Homeostasis in *Brachypodium distachyon* Vegetative Tissues. *Plant Physiol.* **2015**, *169*, 1836–1847. [[CrossRef](#)] [[PubMed](#)]
15. Ibanez-Salazar, A.; Rosales-Mendoza, S.; Rocha-Uribe, A.; Ramirez-Alonso, J.I.; Lara-Hernandez, I.; Hernandez-Torres, A.; Paz-Maldonado, L.M.; Silva-Ramirez, A.S.; Banuelos-Hernandez, B.; Martinez-Salgado, J.L.; et al. Over-expression of Dof-type transcription factor increases lipid production in *Chlamydomonas reinhardtii*. *J. Biotechnol.* **2014**, *184*, 27–38. [[CrossRef](#)] [[PubMed](#)]
16. Zhang, J.; Hao, Q.; Bai, L.; Xu, J.; Yin, W.; Song, L.; Xu, L.; Guo, X.; Fan, C.; Chen, Y.; et al. Overexpression of the soybean transcription factor GmDof4 significantly enhances the lipid content of *Chlorella ellipsoidea*. *Biotechnol. Biofuels* **2014**, *7*, 128–135. [[CrossRef](#)]
17. Song, A.; Gao, T.; Li, P.; Chen, S.; Guan, Z.; Wu, D.; Xin, J.; Fan, Q.; Zhao, K.; Chen, F. Transcriptome-Wide Identification and Expression Profiling of the DOF Transcription Factor Gene Family in *Chrysanthemum morifolium*. *Front. Plant Sci.* **2016**, *7*, 199. [[CrossRef](#)]
18. Yamamoto, A.; Kagaya, Y.; Toyoshima, R.; Kagaya, M.; Takeda, S.; Hattori, T. Arabidopsis NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. *Plant J. Cell Mol. Biol.* **2009**, *58*, 843–856. [[CrossRef](#)]
19. Gibalova, A.; Renak, D.; Matczuk, K.; Dupl'akova, N.; Chab, D.; Twell, D.; Honys, D. AtbZIP34 is required for Arabidopsis pollen wall patterning and the control of several metabolic pathways in developing pollen. *Plant Mol. Biol.* **2009**, *70*, 581–601. [[CrossRef](#)]
20. Mendes, A.; Kelly, A.A.; van Erp, H.; Shaw, E.; Powers, S.J.; Kurup, S.; Eastmond, P.J. bZIP67 regulates the omega-3 fatty acid content of Arabidopsis seed oil by activating fatty acid desaturase3. *Plant Cell* **2013**, *25*, 3104–3116. [[CrossRef](#)]
21. Song, Q.X.; Li, Q.T.; Liu, Y.F.; Zhang, F.X.; Ma, B.; Zhang, W.K.; Man, W.Q.; Du, W.G.; Wang, G.D.; Chen, S.Y.; et al. Soybean GmbZIP123 gene enhances lipid content in the seeds of transgenic Arabidopsis plants. *J. Exp. Bot.* **2013**, *64*, 4329–4341. [[CrossRef](#)] [[PubMed](#)]
22. Jakoby, M.; Weisshaar, B.; Droge-Laser, W.; Vicente-Carbajosa, J.; Tiedemann, J.; Kroj, T.; Parcy, F. bZIP transcription factors in Arabidopsis. *Trends Plant Sci.* **2002**, *7*, 106–111. [[CrossRef](#)]
23. Li, S.; Lauri, A.; Ziemann, M.; Busch, A.; Bhave, M.; Zachgo, S. Nuclear activity of ROXY1, a glutaredoxin interacting with TGA factors, is required for petal development in *Arabidopsis thaliana*. *Plant Cell* **2009**, *21*, 429–441. [[CrossRef](#)] [[PubMed](#)]
24. Banerjee, A.; Roychoudhury, A. Abscisic-acid-dependent basic leucine zipper (bZIP) transcription factors in plant abiotic stress. *Protoplasma* **2017**, *254*, 3–16. [[CrossRef](#)] [[PubMed](#)]
25. Pan, Y.; Hu, X.; Li, C.; Xu, X.; Su, C.; Li, J.; Song, H.; Zhang, X.; Pan, Y. SlbZIP38, a Tomato bZIP Family Gene Downregulated by Abscisic Acid, Is a Negative Regulator of Drought and Salt Stress Tolerance. *Genes* **2017**, *8*. [[CrossRef](#)]
26. Gangappa, S.N.; Botto, J.F. The Multifaceted Roles of HY5 in Plant Growth and Development. *Mol. Plant* **2016**, *9*, 1353–1365. [[CrossRef](#)]
27. Finkelstein, R.R.; Lynch, T.J. The Arabidopsis abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. *Plant Cell* **2000**, *12*, 599–609. [[CrossRef](#)]
28. Alonso, R.; Oclate-Selncchez, L.; Weltmeier, F.; Ehlert, A.; Diaz, I.; Dietrich, K.; Vicente-Carbajosa, J.; Droge-Laser, W.J.P.C. A Pivotal Role of the Basic Leucine Zipper Transcription Factor bZIP53 in the Regulation of Arabidopsis Seed Maturation Gene Expression Based on Heterodimerization and Protein Complex Formation. *Plant Cell* **2009**, *21*, 1747–1761. [[CrossRef](#)]
29. Nijhawan, A.; Jain, M.; Tyagi, A.K.; Khurana, J.P. Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. *Plant Physiol.* **2008**, *146*, 333–350. [[CrossRef](#)]
30. Wang, Y.; Zhang, Y.; Zhou, R.; Dossa, K.; Yu, J.; Li, D.; Liu, A.; Mmadi, M.A.; Zhang, X.; You, J. Identification and characterization of the bZIP transcription factor family and its expression in response to abiotic stresses in sesame. *PLoS ONE* **2018**, *13*, e0200850. [[CrossRef](#)]

31. Wang, Z.; Cheng, K.; Wan, L.; Yan, L.; Jiang, H.; Liu, S.; Lei, Y.; Liao, B. Genome-wide analysis of the basic leucine zipper (bZIP) transcription factor gene family in six legume genomes. *BMC Genom.* **2015**, *16*, 1053. [[CrossRef](#)] [[PubMed](#)]
32. Liu, J.; Chen, N.; Chen, F.; Cai, B.; Dal Santo, S.; Tornielli, G.B.; Pezzotti, M.; Cheng, Z.M. Genome-wide analysis and expression profile of the bZIP transcription factor gene family in grapevine (*Vitis vinifera*). *BMC Genom.* **2014**, *15*, 281. [[CrossRef](#)]
33. Hu, W.; Yang, H.; Yan, Y.; Wei, Y.; Tie, W.; Ding, Z.; Zuo, J.; Peng, M.; Li, K. Genome-wide characterization and analysis of bZIP transcription factor gene family related to abiotic stress in cassava. *Sci. Rep.* **2016**, *6*, 22783. [[CrossRef](#)] [[PubMed](#)]
34. Wang, X.L.; Chen, X.; Yang, T.B.; Cheng, Q.; Cheng, Z.M. Genome-Wide Identification of bZIP Family Genes Involved in Drought and Heat Stresses in Strawberry (*Fragaria vesca*). *Int. J. Genom.* **2017**, *2017*, 3981031. [[CrossRef](#)]
35. Li, Y.-Y.; Meng, D.; Li, M.; Cheng, L. Genomes. Genome-wide identification and expression analysis of the bZIP gene family in apple (*Malus domestica*). *Tree Genet.* **2016**, *12*, 82. [[CrossRef](#)]
36. Zhou, Y.; Xu, D.; Jia, L.; Huang, X.; Ma, G.; Wang, S.; Zhu, M.; Zhang, A.; Guan, M.; Lu, K.; et al. Genome-Wide Identification and Structural Analysis of bZIP Transcription Factor Genes in *Brassica napus*. *Genes* **2017**, *8*. [[CrossRef](#)]
37. Unver, T.; Wu, Z.; Sterck, L.; Turktas, M.; Lohaus, R.; Li, Z.; Yang, M.; He, L.; Deng, T.; Escalante, F.J.; et al. Genome of wild olive and the evolution of oil biosynthesis. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 9413–9422. [[CrossRef](#)] [[PubMed](#)]
38. Droge-Laser, W.; Weiste, C. The C/S1 bZIP Network: A Regulatory Hub Orchestrating Plant Energy Homeostasis. *Trends Plant Sci.* **2018**, *23*, 422–433. [[CrossRef](#)]
39. Sanchez, J. Lipid photosynthesis in olive fruit. *Prog. Lipid Res.* **1994**, *33*, 97–104. [[CrossRef](#)]
40. Connor, D.; Fereres, E. The Physiology of Adaptation and Yield Expression in Olive. *Hortic. Rev.* **2010**, *31*, 155–229. [[CrossRef](#)]
41. Olive Genome Consortium. Available online: <http://olivegenome.org/> (accessed on 4 May 2020).
42. National Institute of Agrobiological Sciences. Available online: <https://www.nig.ac.jp/nig/ja/> (accessed on 4 May 2020).
43. Arabidopsis Information Resource (TAIR). Available online: <https://www.arabidopsis.org/index.jsp> (accessed on 4 May 2020).
44. Joint Genome Institute. Available online: <https://jgi.doe.gov/our-projects/> (accessed on 4 May 2020).
45. CIRAD. Available online: <https://www.cirad.fr/en/> (accessed on 4 May 2020).
46. Jaillon, O.; Aury, J.M.; Noel, B.; Policriti, A.; Clepet, C.; Casagrande, A.; Choisne, N.; Aubourg, S.; Vitulo, N.; Jubin, C.; et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **2007**, *449*, 463–467. [[CrossRef](#)]
47. Wang, L.; Yu, J.; Li, D.; Zhang, X. Sinbase: An integrated database to study genomics, genetics and comparative genomics in *Sesamum indicum*. *Plant Cell Physiol.* **2015**, *56*, e2. [[CrossRef](#)] [[PubMed](#)]
48. Letunic, I.; Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* **2018**, *46*, D493–D496. [[CrossRef](#)] [[PubMed](#)]
49. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A.; et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427–D432. [[CrossRef](#)] [[PubMed](#)]
50. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.; Marler, B.; Guo, H.; et al. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [[CrossRef](#)] [[PubMed](#)]
51. Deppmann, C.D.; Acharya, A.; Rishi, V.; Wobbles, B.; Smeekens, S.; Tapparowsky, E.J.; Vinson, C. Dimerization specificity of all 67 B-ZIP motifs in *Arabidopsis thaliana*: A comparison to Homo sapiens B-ZIP motifs. *Nucleic Acids Res.* **2004**, *32*, 3435–3445. [[CrossRef](#)] [[PubMed](#)]
52. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]

53. Sun, W.; Ma, Z.; Chen, H.; Liu, M. MYB Gene Family in Potato (*Solanum tuberosum* L.): Genome-Wide Identification of Hormone-Responsive Reveals Their Potential Functions in Growth and Development. *Int. J. Mol. Sci.* **2019**, *20*. [[CrossRef](#)]
54. Liu, M.; Wen, Y.; Sun, W.; Ma, Z.; Huang, L.; Wu, Q.; Tang, Z.; Bu, T.; Li, C.; Chen, H. Genome-wide identification, phylogeny, evolutionary expansion and expression analyses of bZIP transcription factor family in tartary buckwheat. *BMC Genom.* **2019**, *20*, 483. [[CrossRef](#)]
55. Ünel, N.; Cetin, F.; Karaca, Y.; Celik, A.Y.; Baloglu, M. Comparative identification, characterization, and expression analysis of bZIP gene family members in watermelon and melon genomes. *Plant Growth Regul.* **2018**, *87*. [[CrossRef](#)]
56. Wei, K.; Chen, J.; Wang, Y.; Chen, Y.; Chen, S.; Lin, Y.; Pan, S.; Zhong, X.; Xie, D. Genome-wide analysis of bZIP-encoding genes in maize. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genom.* **2012**, *19*, 463–476. [[CrossRef](#)] [[PubMed](#)]
57. Yang, O.; Popova, O.V.; Suthoff, U.; Luking, I.; Dietz, K.J.; Gollmack, D. The Arabidopsis basic leucine zipper transcription factor AtbZIP24 regulates complex transcriptional networks involved in abiotic stress resistance. *Gene* **2009**, *436*, 45–55. [[CrossRef](#)] [[PubMed](#)]
58. Tsugama, D.; Liu, S.; Takano, T. A bZIP protein, VIP1, is a regulator of osmosensory signaling in Arabidopsis. *Plant Physiol.* **2012**, *159*, 144–155. [[CrossRef](#)] [[PubMed](#)]
59. Gibalova, A.; Steinbachova, L.; Hafidh, S.; Blahova, V.; Gadiou, Z.; Michailidis, C.; Muller, K.; Pleskot, R.; Duplakova, N.; Honys, D. Characterization of pollen-expressed bZIP protein interactions and the role of ATbZIP18 in the male gametophyte. *Plant Reprod.* **2017**, *30*, 1–17. [[CrossRef](#)]
60. Van Leene, J.; Blomme, J.; Kulkarni, S.R.; Cannoot, B.; De Winne, N.; Eeckhout, D.; Persiau, G.; Van De Slijke, E.; Vercruyse, L.; Vanden Bossche, R.; et al. Functional characterization of the Arabidopsis transcription factor bZIP29 reveals its role in leaf and root development. *J. Exp. Bot.* **2016**, *67*, 5825–5840. [[CrossRef](#)]
61. Maurya, J.P.; Sethi, V.; Gangappa, S.N.; Gupta, N.; Chattopadhyay, S. Interaction of MYC2 and GBF1 results in functional antagonism in blue light-mediated Arabidopsis seedling development. *Plant J. Cell Mol. Biol.* **2015**, *83*, 439–450. [[CrossRef](#)]
62. Chuang, C.F.; Running, M.P.; Williams, R.W.; Meyerowitz, E.M. The PERANTHIA gene encodes a bZIP protein involved in the determination of floral organ number in *Arabidopsis thaliana*. *Genes Dev.* **1999**, *13*, 334–344. [[CrossRef](#)]
63. Lara, P.; Onate-Sanchez, L.; Abraham, Z.; Ferrandiz, C.; Diaz, I.; Carbonero, P.; Vicente-Carbajosa, J. Synergistic activation of seed storage protein gene expression in Arabidopsis by ABI3 and two bZIPs related to OPAQUE2. *J. Biol. Chem.* **2003**, *278*, 21003–21011. [[CrossRef](#)]
64. Hernandez, M.L.; Padilla, M.N.; Mancha, M.; Martinez-Rivas, J.M. Expression analysis identifies FAD2-2 as the olive oleate desaturase gene mainly responsible for the linoleic acid content in virgin olive oil. *J. Agric. Food Chem.* **2009**, *57*, 6199–6206. [[CrossRef](#)]
65. Hernández, M.L.; Sicardo, M.D.; Martínez-Rivas, J.M. Differential Contribution of Endoplasmic Reticulum and Chloroplast  $\omega$ -3 Fatty Acid Desaturase Genes to the Linolenic Acid Content of Olive (*Olea europaea*) Fruit. *Plant Cell Physiol.* **2015**, *57*, 138–151. [[CrossRef](#)]
66. Parvini, F.; Sicardo, M.D.; Hosseini-Mazinani, M.; Martinez-Rivas, J.M.; Hernandez, M.L. Transcriptional Analysis of Stearoyl-Acyl Carrier Protein Desaturase Genes from Olive (*Olea europaea*) in Relation to the Oleic Acid Content of the Virgin Olive Oil. *J. Agric. Food Chem.* **2016**, *64*, 7770–7781. [[CrossRef](#)] [[PubMed](#)]
67. Banilas, G.; Karampelias, M.; Makariti, I.; Kourti, A.; Hatzopoulos, P. The olive DGAT2 gene is developmentally regulated and shares overlapping but distinct expression patterns with DGAT1. *J. Exp. Bot.* **2011**, *62*, 521–532. [[CrossRef](#)] [[PubMed](#)]
68. Cheng, Z.; Zhan, M.; Yang, Z.; Zumstein, K.; Chen, H.; Huang, Q. The Major Qualitative Characteristics of Olive (*Olea europaea* L.) Cultivated in Southwest China. *Front. Plant Sci.* **2017**, *8*, 559–571. [[CrossRef](#)] [[PubMed](#)]
69. Zhu, S.; Niu, E.; Shi, A.; Mou, B. Genetic Diversity Analysis of Olive Germplasm (*Olea europaea* L.) With Genotyping-by-Sequencing Technology. *Front. Genet.* **2019**, *10*, 53–61. [[CrossRef](#)]
70. Reboredo-Rodríguez, P.; Gonzalez-Barreiro, C.; Cancho-Grande, B.; Fregapane, G.; Salvador, M.D.; Simal-Gandara, J. Characterisation of extra virgin olive oils from Galician autochthonous varieties and their co-crushings with Arbequina and Picual cv. *Food Chem.* **2015**, *176*, 493–503. [[CrossRef](#)]

71. Ben Ayed, R.; Ennouri, K.; Ercisli, S.; Ben Hlima, H.; Hanana, M.; Smaoui, S.; Rebai, A.; Moreau, F. First study of correlation between oleic acid content and SAD gene polymorphism in olive oil samples through statistical and bayesian modeling analyses. *Lipids Health Dis.* **2018**, *17*, 74. [[CrossRef](#)]
72. Lodolini, E.M.; Polverigiani, S.; Ali, S.; Mutawea, M.; Qutub, M.; Arabasi, T.; Pierini, F.; Abed, M.; Neri, D. Oil Characteristics of Four Palestinian Olive Varieties. *J. Oleo Sci.* **2017**, *66*, 435–441. [[CrossRef](#)]
73. Bruno, L.; Picardi, E.; Pacenza, M.; Chiappetta, A.; Muto, A.; Gagliardi, O.; Muzzalupo, I.; Pesole, G.; Bitonti, M.B. Changes in gene expression and metabolic profile of drupes of *Olea europaea* L. cv Carolea in relation to maturation stage and cultivation area. *BMC Plant Biol.* **2019**, *19*, 428. [[CrossRef](#)]
74. Hernandez, M.L.; Sicardo, M.D.; Alfonso, M.; Martinez-Rivas, J.M. Transcriptional Regulation of Stearoyl-Acyl Carrier Protein Desaturase Genes in Response to Abiotic Stresses Leads to Changes in the Unsaturated Fatty Acids Composition of Olive Mesocarp. *Front. Plant Sci.* **2019**, *10*, 251. [[CrossRef](#)]
75. Choi, J.; Huh, S.U.; Kojima, M.; Sakakibara, H.; Paek, K.H.; Hwang, I. The cytokinin-activated transcription factor ARR2 promotes plant immunity via TGA3/NPR1-dependent salicylic acid signaling in Arabidopsis. *Dev. Cell* **2010**, *19*, 284–295. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).