

Article The First Genome-Wide Mildew Locus O Genes Characterization in the Lamiaceae Plant Family

Andolfo Giuseppe * 🕑 and Ercolano Maria Raffaella * 🕩

Department of Agricultural Sciences, University of Naples "Federico II", Via Università 100, Portici, 80055 Naples, Italy * Correspondence: giuseppe.andolfo@unina.it (A.G.); ercolano@unina.it (E.M.R.)

Abstract: Powdery mildew (PM) is a widespread plant disease that causes significant economic losses in thousands crops of temperate climates, including Lamiaceae species. Multiple scientific studies describe a peculiar form of PM-resistance associated at the inactivation of specific members of the Mildew Locus O (MLO) gene family, referred to as mlo-resistance. The characterization of Lamiaceae MLO genes, at the genomic level, would be a first step toward their potential use in breeding programs. We carried out a genome-wide characterization of the MLO gene family in 11 Lamiaceae species, providing a manual curated catalog of 324 MLO proteins. Evolutionary history and phylogenetic relationships were studied through maximum likelihood analysis and motif patter reconstruction. Our approach highlighted seven different clades diversified starting from an ancestral MLO domain pattern organized in 18 highly conserved motifs. In addition, 74 Lamiaceae putative PM susceptibility genes, clustering in clade V, were identified. Finally, we performed a codon-based evolutionary analysis, revealing a general high level of purifying selection in the eleven Lamiaceae MLO gene families, and the occurrence of few regions under diversifying selection in candidate susceptibility factors. The results of this work may help to address further biological questions concerning MLOs involved in PM susceptibility. In follow-up studies, it could be investigated whether the silencing or loss-of-function mutations in one or more of these candidate genes may lead to PM resistance.

Keywords: medicinal plants; MLO; powdery mildew; functional inference; disease resistance

1. Introduction

The mint family (*Lamiaceae*) is an important group of medicinal plants. It consists of more than 7000 aromatic species, distributed nearly worldwide, and commonly widespread in regions of lands around the Mediterranean area [1,2]. Some species (such as mint, basil, lavender and rosemary) are famous for their use in essential oil (EO) production worldwide. EOs can be extracted from several parts of medicinal plants [3] for foods, cosmetic, cleaning, and pharmaceutical industries applications [4,5]. Generally, EO yield is correlated with the number and distribution of glandular trichomes, in which oil components are synthesized and stored [6]. *Lamiaceae* pathogens can certainly compromise plant growth, directly or indirectly affecting the quality and abundance of EOs. Among them, the powdery mildew (PM) is considered to be one of the main biotic threats for the cultivation of species of the *Lamiaceae* species [7,8]. PM colonizes young stems and leaves producing the characteristic white mycelium; when the attack is very severe, stems can be completely killed. The use of resistant cultivars can significantly reduce the amount of fungicides necessary to control PM.

Several studies revealed that the loss of function of Mildew Locus O (MLO)-specific gene members confers resistance to powdery mildew [9–13]. The numbers of members in the MLO family in the plant kingdom ranged from three in *Volvox carteri* to 55 in *Triticum aestivum* [14]. Previous inference analyses grouped MLO proteins in distinct subfamilies, or phylogenetic clades, with a varying number of clades depending on the



Citation: Giuseppe, A.; Raffaella, E.M. The First Genome-Wide Mildew Locus O Genes Characterization in the *Lamiaceae* Plant Family. *Int. J. Mol. Sci.* 2023, 24, 13627. https://doi.org/ 10.3390/ijms241713627

Academic Editor: Zeng-yan Zhang

Received: 30 June 2023 Revised: 24 August 2023 Accepted: 25 August 2023 Published: 4 September 2023



Copyright: © 2023 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/). analyzed species [14–16]. Available scientific literature indicates that two clades (IV and V in monocots and dicots, respectively) host MLO proteins associated with PM susceptibility [14,15,17,18]. Indeed, the inactivation of these genes by gene silencing, genome editing, or TILLING leads to a peculiar form of resistance (known as mlo-resistance) based on the enhancement of pre-penetrative defense responses [19–21]. The broad-spectrum effectiveness and durability makes mlo-resistance a general breeding strategy against PM of cultivated species.

Although MLO genes have been studied in several species [15], they have not yet been identified in *Lamiaceae* plant family. The growing interest for medicinal properties as well as for economic importance in *Lamiaceae* species has led to the sequencing of several genomes in recent years [22–29], thus providing an opportunity to explore the variability of their MLO gene families.

In this study, we used the genome sequences of *Callicarpa americana*, *Mentha longifolia*, *Ocimum basilicum*, *Origanum majorana*, *Origanum vulgare*, *Rosmarinus officinalis*, *Salvia bowleyana*, *Salvia hispanica*, *Salvia miltiorrhiza*, *Salvia officinalis* and *Salvia splendens* to annotate the corresponding MLO gene families. Through a phylogenetic analysis and the characterization of MLO motifs, we gained information with respect to their structure, evolutionary history and function. Furthermore, genetic diversity studies were performed to improve knowledge on the main MLO selection force driving evolution associated with PM susceptibility. A better understanding of *Lamiaceae* MLOs will promote progress in both basic research and practical breeding activities aimed to select PM-resistant cultivars in these species.

2. Results

2.1. MLO Gene Annotation in Eleven Lamiaceae Species

In order to identify the genes encoding for the MLO proteins in *Lamiaceae* plant family, a protein domain search analysis into the *C. americana*, *M. longifolia*, *O. basilicum*, *O. majorana*, *O. vulgare*, *R. officinalis*, *S. bowleyana*, *S. hispanica*, *S. miltiorrhiza*, *S. officinalis* and *S. splendens* proteomes was performed. Since the official gene annotation of *M. longifolia* and *Salvia officinalis* genome assemblies are not available (Supplementary Table S1), the loci harboring a typical MLO domain were identified by using an in-house-built pipeline described in the Materials and Methods section.

A total of 324 MLO genes were identified and finely characterized into the 11 *Lamiaceae* genomes (Table 1; Supplementary Table S2). MLO gene family size in *Lamiaceae* ranged from 20 to 54 members in *O. vulgare* and *S. splendens*, respectively (Table 1). A conspicuous diversification in the size of the MLO families across species was revealed, while within the same genus (e.g., *Origanum* sp., *Salvia* sp.) narrow differences were observed (Table 1). A positive relationship between the size of a predicted gene set in each genome assembly and the relative number of MLO genes was displayed by correlation analysis (Pearson's r = 0.769).

The size of manually curated MLO genes identified in *Lamiaceae* genomes, composed of 11 to 14 exons, ranged between 3591 bp (*O. majorana*) and 5105 bp (*S. bowleyana*) (Table 2 and Supplementary Table S3), while the number of amino acid residues encoded by these MLO genes ranged from 407 to 520 (Table 2). The average length of MLO domain (pfam03094), archived in "The Protein Family" (Pfam) database, resulted of 367 amino acids (Supplementary Table S3). The MLO genes predicted and characterized in the *M. longifolia* and *S. officinalis* genomes [23,29] were consistently similar to the ones annotated in remaining analyzed species in terms of structural genomic properties (Table 2). The longest MLO gene model belonged to *S. splendens* (KAG6400176.1), with a total length of 15,015 nucleotides (Supplementary Table S3).

Species	Common Name	Genome Size (Mb)	Predicted Gene Set	Number of MLO Loci	Reference Genome
Callicarpa americana	American beautyberry	506	32,164	23	Hamilton et al., 2020 [26]
Mentha longifolia	Silver mint	470	52,520 *	33	Vining et al., 2016 [23]
Ocimum basilicum Origanum majorana Origanum vulgare Rosmarinus officinalis	Basil Marjoram Oregano Rosemary	2068 761 630 1014	78,990 33,929 32,623 51,389	45 21 20 29	Bornowski et al., 2020 [25]
Salvia bowleyana	Nan-dan-shen	462	44,044	22	Zheng et al., 2021 [28]
Salvia hispanica	Chia	348	31,069	28	Wang et al., 2022 [22]
Salvia miltiorrhiza	Dan-shen	595	32,483	21	Song et al., 2020 [24]
Salvia officinalis	ulvia officinalis Sage 480		50,957 *	28	Li et al., 2022 [29]
Salvia splendens	Scarlet sage	808	56,267	54	Jia et al., 2021 [27]
	Tota	al			324

Table 1. Number of MLO genes identified in eleven *Lamiaceae* genomes. Genome size and predicted gene sets were reported for each plant species.

* Number of gene models predicted in this study by Augustus on *Mentha longifolia* [30] and *Salvia officinalis* [29] genome assemblies.

Table 2. Structural properties of MLO gene models identified in eleven genomes. Gene, protein and domain lengths, and number of exons were reported for each *Lamiaceae* species. The average number refers to the arithmetic mean between the members of each MLO gene family.

Species	Genomic Length (bp)	Number of Exons for Locus	Protein Length (aa)	Domain Length (aa) *
Callicarpa americana	5084	12	458	387
Mentha longifolia	4763	12	507	334
Ocimum basilicum	4101	11	407	335
Origanum majorana	3960	11	433	348
Origanum vulgare	4215	13	482	404
Rosmarinus officinalis	3591	11	412	341
Salvia bowleyana	5105	13	436	344
Salvia hispanica	3958	13	494	406
Salvia miltiorrhiza	5008	14	520	449
Salvia officinalis	4167	12	474	344
Salvia splendens	4082	11	454	343

* Length of MLO Pfam domain [31] predicted in this study by InterProScan [32].

2.2. Identification of MLO Genes Associated to PM Susceptibility

The MLO-Pfam domain sequences of *C. americana*, *M. longifolia*, *O. basilicum*, *O. majorana*, *O. vulgare*, *R. officinalis*, *S. bowleyana*, *S. hispanica*, *S. miltiorrhiza*, and *S. splendens* MLO proteins were used to infer phylogenetic distances among them and with respect to functionally validated MLOs in other plant species (Supplementary Table S4). In total, 293 MLO proteins collapsed in seven robust phylogenetic clades (bootstrap index \geq 95) (Figure 1; Supplementary Figure S1 and Table S5). These were designated with the Roman numerals from I to VII, based on the position of Arabidopsis and monocot MLO homologs reported in the previous studies [14,33].

Clade I includes fifty-one MLO proteins, three of which are annotated in Arabidopsis (AtMLO4, AtMLO11 and AtMLO14), 10 in *S. splendens*, five in *M. longifolia*, five *O. basilicum*, five in *S. hispanica*, four in *S. bowleyana*, four in *C. americana*, four in *O. vulgare*, four in *R. officinalis*, four in *S. miltiorrhiza*, and three in *O. majorana*. Clade II groups forty-eight MLO homologs, three of which identified in Arabidopsis (AtMLO1, AtMLO13 and AtMLO15), six in *R. officinalis* and *O. basilicum*, five in *M. longifolia*, and four in all remaining analyzed species (*C. americana*, *O. majorana*, *O. vulgare*, *S. bowleyana*, *S. hispanica*, *S. miltiorrhiza* and *S. splendens*).



Figure 1. Maximum likelihood analysis of *Lamiaceae* MLO proteins. The tree includes 247 *Lamiaceae* MLO homologs and 25 reference MLO proteins already characterized in other species. Clades were enumerated with the Roman numerals from I to VII on the base of position of Arabidopsis and monocot MLO genes. The tree was drawn to scale, with branch lengths proportional to the number of substitutions per site. Bootstrap values of collapsed clades are indicated above the branches. The taxa to which the protein sequences belong are indicated by colored spots. Pie charts illustrate the number of MLO genes belonging to a given clade in each taxonomic group.

A common ancestor (bootstrap support 98/100) in the evolutionary history of clades III, IV, V, VI and VII was found (Figure 1). Clade III includes thirty-six *Lamiaceae* homologs (eight in *S. splendens*; seven in *O. basilicum*; five in *S. hispanica*; three in *C. americana* and *R. officinalis*; and two in *M. longifolia*, *S. bowleyana*, *S. miltiorrhiza*, *O. majorana* and *O. vulgare*) together with five Arabidopsis proteins (AtMLO5, AtMLO7, AtMLO8, AtMLO9 and AtMLO10). Phylogenetic data also provided evidence for a common ancestor originating clades III and IV (Supplementary Figure S1 and Table S5). Clade IV, which is conventionally associated with MLO proteins (TaMLO_A1b, TaMLO_B1, HvMLO1 and OsMLO3) from monocot species, also contains 13 *Lamiaceae* MLO proteins (two in *S. splendens*, *O. basilicum* and *M. longifolia*; one in *C. americana*, *O. majorana*, *O. vulgare*, *R. officinalis*, *S. bowleyana*, *S. hispanica* and *S. miltiorrhiza*).

Nine M. longifolia, O. basilicum and S. splendens, seven S. hispanica and S. miltiorrhiza, six S. bowleyana and O. vulgare, five C. americana and R. officinalis, and four O. majorana

proteins cluster together in the phylogenetic clade V, including all the dicot MLO homologs experimentally shown to act as PM susceptibility factors. Within the clade V, based on a bootstrap index \geq 53, two subclades (Va and Vb) were distinguishable, including eight functionally characterized MLO proteins (AtMLO2, AtMLO6, AtMLO12, SIMLO1, PsMLO1, CaMLO2, LjMLO1 and MtMLO1) in the subclade Va and one (VvMLO3) in the subclade Vb (Figure 1). Clade VI is located on an ancestral position (Supplementary Figure S1) and included AtMLO3 protein and 25 *Lamiaceae* MLOs. In particular, four MLO proteins annotated in *O. basilicum* and three in *R. officinalis, S. splendens* and *S. hispanica*, while only two AtMLO3 homologs were found in the remaining analyzed species (Figure 1; Supplementary Figure S1). Finally, clade VII includes 13 *Lamiaceae* MLO proteins (three in *S. splendens*; two *R. officinalis* and *S. bowleyana*; and one in *C. americana*, *M. longifolia*, *O. basilicum*, *O. vulgare*, *S. hispanica* and *S. miltiorrhiza*).

2.3. Characterization of Clade-Specific Amino Acid Motifs

To explain the evolutionary routes emerging in the phylogenetic tree of *Lamiaceae* MLO gene family, a motif-based sequence analysis was carried out. Towards this goal, the MLO Pfam domain (pfam03094) of our protein dataset was divided into 30 ungapped motifs by MEME (Supplementary Table S6), which were used by MAST for a sequence pattern analysis. The motifs structure of MLO Pfam domains for each collapsed clade was showed in Figure 2.



Figure 2. Identification and distributions of clade-specific amino acid motifs in MLOs. The six phylogenetic clades were collapsed and marked with different colors. For each clade the sequence logos of highly conserved motifs were showed. Black lines and colored boxes denote MLO domain sequences and motif positions, respectively. Red and blue rectangles delimit the conserved motif patterns between clades. Red lines connect motifs conserved between clades IV and V.

MLO Pfam domain was outlined from a total of 43 motifs since some MEME-motifs were detected multiple times along the protein sequences (Figure 2). The number of MLO-related motifs increased from clade VII (27) to clades III and V (36). Our analysis has underlined an ancestral MLO-core structure including 23 motifs highly conserved between the seven clades, of which 78% (18 out of 23) were organized in four blocks (red rectangles in Figure 2). Starting from this ancestral structure, the MLO family has diversified and functionally specialized in the seven phylogenetic clades. The presence of 20 additional motifs contribute to generate the seven clade-specific patterns (Supplementary Table S7).

Clade V MLO homologs showed the most complex profile, beside the red conserved blocks, six clade-specific motifs were identified between the red blocks three and four (Figure 2).

An additional block composed from four motifs (MEME-8, MEME-29, MEME-14 and MEME-16) was present only in clades III, IV, V and VI (blue rectangles in Figure 2). Noteworthily, three motifs (MEME-26, MEME-17 and MEME-25) were simultaneously conserved between clade IV and clade V (red lines in Figure 2). Analyzing in more detail the MLO domain structure of all the dicot MLO homologs experimentally tested so far for PM susceptibility (AtMLO2, AtMLO6, AtMLO12, SIMLO1, PsMLO1, CaMLO2, LjMLO1, MtMLO1 and VvMLO3), 30 conserved motifs were detected (Supplementary Table S7), suggesting that they are functionally related. Only 24% (18 out of 74) of *Lamiaceae* MLO proteins showed the conserved pattern of molecularly validated MLO genes for PM susceptibility (Supplementary Table S8). Diversely, the functionally well-characterized monocot MLOs (TaMLO_A1b, TaMLO_B1, HvMLO1 and OsMLO3) presented four conserved motifs that were absent in all *Lamiaceae* MLOs of clade IV.

2.4. Selection Pressure Acting on Lamiaceae MLO Phylogenetic Clades

In order to discover the direction and magnitude of natural selection acting on MLO protein coding genes in *C. americana*, *M. longifolia*, *O. basilicum*, *O. majorana*, *O. vulgare*, *R. officinalis*, *S. bowleyana*, *S. hispanica*, *S. miltiorrhiza*, *S. officinalis* and *S. splendens*, we used a combination of maximum likelihood and counting approaches implemented into the single-likelihood ancestor counting (SLAC) method [34].

Neutrality tests performed on the MLO gene family of each of the eleven *Lamiaceae* species, yielded π and ω values ranges from 0.342 to 0.431 and from 0.311 to 0.212, respectively (Table 3). On average, ~239 (max. 292; min. 180) negatively selected codon sites are statistically significant in each MLO gene family (Table 3). Comprehensive analyses reveal that negative selection has been acting against extreme polymorphic variants in *Lamiaceae* plant species. Indeed, only eleven positively selected codon sites were statistically supported in five out of eleven analyzed species (Table 3).

Table 3. Estimation of evolutionary divergence (π) and non-synonymous to synonymous substitutions (ω) mean dissimilarity for each MLO gene family and phylogenetic clade.

Aligned Coding Sequences	MLO Loci (n.)	Average Identity (%)	Π*	ω	PSCs	NSCs
<i>C. americana</i> MLO family	23	39.2	0.364	0.223	0	247
M. longifolia MLO family	33	25.7	0.428	0.349	0	192
O. basilicum MLO family	45	29.9	0.398	0.250	0	243
O. majorana MLO family	21	33.2	0.374	0.251	2	292
<i>O. vulgare</i> MLO family	20	37.9	0.41	0.212	0	250
R. officinalis MLO family	29	34.3	0.369	0.249	2	180
S. bowleyana MLO family	22	36.7	0.369	0.259	0	182
S. hispanica MLO family	28	35.9	0.407	0.257	1	253
S. miltiorrhiza MLO family	21	45.1	0.431	0.221	2	264
S. officinalis MLO family	28	33.7	0.409	0.256	0	264
S. splendens MLO family	54	29.9	0.342	0.311	4	267
Lamiaceae MLO clade I	51	48.9	0.314	0.307	0	279

Aligned Coding Sequences	MLO Loci (n.)	Average Identity (%)	Π*	ω	PSCs	NSCs	
Lamiaceae MLO clade II	49	55.6	0.307	0.267	0	311	
Lamiaceae MLO clade III	39	50.8	0.268	0.291	0	255	
Lamiaceae MLO clade IV	14	68.3	0.146	0.236	0	96	
Lamiaceae MLO clade V	73	44.2	0.254	0.375	4	371	
Lamiaceae MLO clade VI	27	53.6	0.252	0.369	0	178	
Lamiaceae MLO clade VII	13	73	0.115	0.216	0	65	

Table 3. Cont.

* All positions with less than 80% site coverage were eliminated. The values of ω (=dN/dS) reported in the table were calculated using SLAC method and the number of positive (PSCs: positively selected codons) and negative (NSCs: negatively selected codons) codons statistically significant (p < 0.1) in MLO protein-coding genes, and are significant.

Similarly, a purifying selection (0.216–0.375 ω values) was observed into the seven phylogenetic clades (Table 3). Indeed, the average sequence identity ranges from 44.2% to 73% in clades V and VII, respectively. Likely, natural selection acts against mutations that have deleterious effects on functionally important amino acid residues. Single-codon analysis highlighted the presence of several negatively selected sites varying from 371 to 65 in the V and VII clades, respectively (Table 3). With respect to other phylogenetic clades, clade V displays a higher ω (0.375) value (Table 3), thus suggesting that it is subjected to a larger number of selective constraints. Indeed, 371 out of 954 negatively selected codon sites were also statistically significant (Table 3; Figure 3). Finally, a total of 407 positively selected codon sites ($\omega > 1$) were identified along the MLO protein-coding-genes collapsed in clade V, of which only four statistically significant (Figure 3). The relaxed purifying selection might be the main driving force during the evolution of MLO genes associated with PM susceptibility.



Figure 3. Natural selection acting on *Lamiaceae* clade V MLO homologs. Direction and magnitude of positive and negative dN–dS values are showed with blue and red lines, respectively. Extracellular, transmembrane and intracellular MLO protein regions are indicated with the violet, grey and green boxes, respectively. The red arrows indicate codons characterized by a significant *p*-value (p < 0.1) for positive selection (in position 955, 1135, 1180 and 1944 of the alignment).

3. Discussion

In the present study, we exploited available genomic resources to characterize the MLO gene family in cultivated *Lamiaceae*. A total of 324 MLO protein encoding genes from eleven medicinal plant genomes were identified. It was observed that a direct correlation between predicted gene set sizes and the number of MLO gene family members. For example, the MLOs annotated in the genome assemblies released by Bornowski et al. [25] were higher in

genome with a greater gene set, and vice versa (Table 1). Diversely, the number of MLO genes of the *S. miltiorrhiza* (32,483 loci) [24] and *C. americana* (32,164 loci) [26] did not change significantly.

In addition, we also noted that, the tetraploid *O. basilicum* cultivar (Genovese) $(2n = 4 \times = 48)$ and *S. splendens* $(2n = 4 \times = 44)$ genomes (Bornowski et al., 2020; Jia et al., 2021), displayed half of MLO orthologs in multiple copies, consistent with their polyploidy [35]. These findings indicate that the whole-genome duplication events contributed to the expansion of MLO gene families [36]. All the MLO proteins identified in *Lamiaceae* genomes have amino acid lengths comparable to those of Arabidopsis AtMLO homologs, ranging from 460 to 593 residues [33]. Similarly, the de novo gene models predicted in *M. longifolia* and *S. officinalis* genomes, obtained in this work, showed structural features in line with the MLO gene annotation reported in other species (Table 2).

Relations of orthology between Lamiaceae MLO members were inferred based on phylogenetic relatedness [15,33,37]. Based on the functionally characterized MLO genes, the gene family is collapsed into six major clades, of which the clade IV was exclusively present in monocots, while the clade V was only present in dicots [33]. Iovieno et al. [14] reported for the first time the occurrence of MLOs belonging to dicotyledons in clade IV. Furthermore, some authors claim that the MLO genes are organized in seven clades [16], although not all plant species harbor members in all clades [38]. An extra clade (VIII) has been proposed, and various subclades have been identified by Iovieno et al. [14], but no consensus has been reached yet. Our maximum likelihood analysis assigned Lamiaceae MLO proteins to the seven evolutionary clades. We found considerable differences in the variation of the number of genes (from one to nine) per clade across species. Although Clade IV is considered exclusively composed of monocot MLO proteins [19], we found 14 Lamiaceae MLOs grouped in it, confirming recent findings [14]. In Lamiaceae clade IV, we distinguished a primary subgroup (bootstrap index = 99), including all the monocot homologs involved in the interaction with PM fungi (namely barley HvMLO, rice OsMLO3 and wheat TaMLO_A1 and TaMLO_B1) separated from the group including the 14 Lamiaceae MLOs. Probably, clade IV Lamiaceae MLO homologs have the same role of MLOs previously identified in [14].

Interestingly, 74 MLO homologs were found to cluster together in the phylogenetic clade V, including also all the dicot reference genes (namely Arabidopsis AtMLO2, AtMLO6 and AtMLO12; tomato SIMLO1; grapevine VvMLO3 and VvMLO4; tobacco NtMLO1; pepper CaMLO2 and barrel clover MtMLO1) functionally associated with PM susceptibility [17,37]. Similar to clade IV, it was determined that clade V is composed of two subclades, named Va and Vb (Figure 1). Subclade Va is composed of all the reference genes, except for VvMLO3, seven homologs belonging to S. miltiorrhiza, five to C. americana, seven to S. officinalis, nine to M. longifolia, four to O. majorana, six to O. vulgare, five to R. officinalis, six to *S. bowleyana*, seven to *S. hispanica*, nine for *S. splendens*, and nine for *O. basilicum*. The subclade Vb, includes VvMLO3, three members belonging to S. splendens, two to R. officinalis and S. bowleyana, and one to all the remaining Lamiaceae species. Our study provides information on 65 and 14 putative novel PM susceptibility factors in subclades Va and Vb, respectively. These MLO gene targets could be used for future breeding activities aimed to introduce mlo-resistance in cultivated species. Noteworthy, transcription activator-like effector nucleases (TALEN) and clustered regularly interspaced short palindromic repeat (CRISPR) technology can be used to inactivate PM susceptibility genes [20,21].

We reanalyzed the multiple alignment of MLO protein dataset to detect amino acid motifs that, being highly conserved, are predicted to play a major role for PM susceptibility function. For the first time, we found that each clade was characterized by a specific conserved motifs patter, implying that specific conserved profiles were likely required for subfamily-specific functions. In the *Lamiaceae* MLO gene family, the seven MLO clades were evolved from the ancient MLO domain structure. Subsequently, short duplication, insertion or deletion of specific motifs occurred in the MLO protein domain (Figure 2). Clade-related evolutionary patters exhibit some degree of local sequence conservation and diversification that could result in fold change or deterioration. Gene sequence divergence, recombination and duplications resulted the main driving forces for the gene diversification [39–41].

As expected, a very high conservation level (pairwise average identity: ~38%) was found in the MLO Pfam domain (Pfam03094) sequences, including the extracellular loops or transmembrane regions, previously found to be essential for MLO functionality and/or stability [42]. In addition, we detected a series of conserved motifs scattered in the clade V *Lamiaceae* MLO homologs. These might be specifically important for the evolution of functional isoforms associated with PM susceptibility. With this respect, our study complements the previous works on conserved amino acid residues and motifs [15,42,43], reporting the identification of a motif pattern specifically conserved in putative PM susceptibility factors.

In line with the identification of several conserved amino acidic motifs, the neutrality test addressed to infer evolutionary forces acting on *Lamiaceae* MLO homologs suggested a general high level of negative selection (Table 3). The negative selection contributes to low gene family variability, which was consistent with the hypothesis that highly conserved genes remain in the genome due to purifying selection. Most likely, the strong negative selection has likely contributed to the maintenance of consistent clade sizes across species. Interestingly, single-codon analysis of clade V *Lamiaceae* MLO homologs highlighted the occurrence of three protein regions that are likely under positive selection pressure (Figure 3), as reported by Ioieno et al. [15] for Cucurbitaceae and Rosaceae clade V MLO homologs. Generally, the gene copies with high variability in specific sites tend to be involved in plant-environment interactions, particularly biotic interactions [44]. Positive selection drives plant/pathogen co-evolution [45,46], in accordance with an "arms race" model [47–49]. Although the molecular interaction between MLO proteins and pathogen effectors is still elusive to date, it might be speculated that positively selected codon sites located in the intracellular C-terminus are implicated in pathogen sensing [15].

4. Conclusions

The comprehensive analysis of the MLO gene families in eleven economically important *Lamiaceae* species allowed us to assess the variability existing in this botanical family. Starting from the MLO core structure reconstructed in this work, we were able to depict the diversification that occurred in seven phylogenetic clades. For the first time, we highlighted the conserved motifs patter of each MLO clade. Our findings are important for the identification of a motif profile specifically associated with PM susceptibility. The conserved pattern of *Lamiaceae* MLOs in clade IV and V could be further investigated to identify putative PM susceptibility factors. Furthermore, the identification of codons under positive selection pressure in clade V members may direct the selection of appropriate candidate genes for further breeding activities aimed at introducing PM resistance.

5. Materials and Methods

5.1. Taxa Data Set

To investigate the evolution of MLO genes along the *Lamiaceae* plant family, we used the genomic data of eleven taxa (*Callicarpa americana, Mentha longifolia, Ocimum basilicum, Origanum majorana, Origanum vulgare, Rosmarinus officinalis, Salvia bowleyana, Salvia hispanica, Salvia miltiorrhiza, Salvia officinalis and Salvia splendens*). The genome and proteome sequences were retrieved from five (DRYAD, GigaDB, CNGBdb, GenBank, NGDC and GWH) public database (Supplementary Table S1). Moreover, for comparative purposes we also added 15 MLO genes identified in *Arabidopsis thaliana* retrieved from TAIRdb (www.arabidopsis.org, accessed on 1 June 2023) and 10 well-characterized reference MLO genes in other species (*A. thaliana, C. annuum, H. vulgare, L. japonicas, M. truncatula, O. sativa, P. sativum, S. lycopersicum, T. aestivum* and *V. vinifera*) (Supplementary Table S4).

5.2. Identification and Manual Curation of Eleven Lamiacaeae MLO Gene Families

In order to retrieve predicted *Lamiaceae* MLO genes, we scanned the *C. americana*, *L. angustifolia*, *O. basilicum*, *O. majorana*, *O. vulgare*, *R. officinalis*, *S. hispanica*, *M. longifolia*,

S. miltiorrhiza, S. officinalis and *S. splendens* proteome data set (Supplementary Table S1) with the hidden Markov model of the MLO domain (Pfam: PF03094) using HMMER v3.0 with default parameters [50]. In addition, a local BLASTp search (e-value $< 1 \times 10^{-5}$) was performed by mapping A. thaliana amino acid sequences as queries (AtMLO1-AtMLO15 (Supplementary Table S4), NCBI accession numbers: NP_192169.1, NP_172598.1, NP_566879.1, NP_563882.1, NP_180923.1, NP_176350.1, NP_179335.3, NP_565416.1, NP_174980.3, NP_201398.1, NP_200187.1, NP_565902.1, NP_567697.1, NP_564257.1 and NP_181939.1) to the *Lamiaceae* protein data set. The protein domain architecture of HMMER and BLAST outputs was annotated using InterProScan [32] and conserved domain search [51] with default parameters.

The official gene annotation of *M. longifolia* and *S. officinalis* were not made available on GenBank (GCA001642375.2) and NGDC (GWHBJVP00000000), respectively. To identify the MLO coding sequences located on *M. longifolia* scaffolds and *S. officinalis* chromosomes, we performed a gene annotation using the Augustus (v2.5.5) pipeline [30] with Arabidopsis parameters. Only the predicted genes encoding proteins of more length, 50 amino acids and a confidence score > 0 were used to perform *M. longifolia* and *S. officinalis* MLO annotations.

5.3. Multiple Sequence Alignments and Phylogenetic Analysis

Sequence similarities were determined performing a ClustalW (v2.0) [52] multiple alignment with default settings, using the conserved MLO domain sequence of the PFAM database (ID: PF03094) as input. Phylogenetic analysis was performed by using newly identified *Lamiaceae* MLO homologs containing at least 50% of the full-length MLO domain. The dataset was completed with the whole Arabidopsis MLO protein family and the following proteins previously shown to act as PM susceptibility factors: pea PsMLO1, barley HvMLO, rice OsMLO2, pepper CaMLO2, tomato SIMLO1, grape VvMLO3, barrel clover MtMLO1, lotus LjMLO1, and wheat TaMLO1_A1b and TaMLO_B1 (Supplementary Table S4). All these sequences were extracted from the NCBI database (http://www.ncbi.nlm.nih.gov, accessed on 1 June 2023). Evolutionary relationships between MLO proteins were inferred using the maximum likelihood method based on the Whelan and Goldman model [53], using the MEGA6 software (http://www.megasoftware.net, accessed on 1 June 2023) [54]. The model with the lowest BIC (Bayesian Information Criterion) score was considered to describe the substitution pattern. The bootstrap consensus tree, obtained from 100 replicates, was taken to represent the MLO family phylogenetic history [55].

5.4. De Novo Prediction of MLO-Encoding Genes Motifs

The Multiple EM for Motif Elicitation (MEME) (https://meme-suite.org/, accessed on 1 June 2023) algorithm was used to decompose in motifs (Supplementary Table S3) the MLO Pfam domain (PF03094) of protein sequence data set. The analysis was carried out using the default cut-off value for statistical confidence. Motif Alignment and Search Tool (MAST) (https://meme-suite.org/, accessed on 1 June 2023) was also used to confirm the presence of MEME-motifs previously identified (Supplementary Table S3), using default setting [56]. The MEME-motifs with less than 70% site coverage were eliminated from the motif structure of phylogenetic clades.

5.5. Evolution Rates at Codon Sites

Evolutionary forces acting on MLO homologs, in *Lamiaceae* plant family, were investigated by determining two parameters based on the number of non-synonymous and synonymous substitutions per non synonymous and synonymous site (dN and dS, respectively), indicated as ω (dN/dS). Tests were conducted to estimate the evolution of each codon: positive ($\omega > 1$); neutral ($\omega = 1$); and negative ($\omega < 1$). All MLO-coding DNA sequences were aligned using ClustalW (v2.0) [52], and positions with less than 80% site coverage were eliminated from the analysis. To clearly depict the proportion of sites under selection, an evolutionary fingerprint analysis was carried out using the implemented SLAC algorithm in the Datamonkey server at the default value [34]. Pressure selection analyses were performed on the same sequence subset of phylogenetic analysis.

The average nucleotide diversity (π) was computed using the Tajima's Test of Neutrality method implemented in the MEGA v6 software [54].

Supplementary Materials: The following supporting information can be download at: https://www.mdpi.com/article/10.3390/ijms241713627/s1.

Author Contributions: A.G. conceived the research idea, performed all the analyses, and was centrally involved in results interpretation and in manuscript writing. E.M.R. contributed to results interpretation and manuscript revision. All authors have read and agreed to the published version of the manuscript.

Funding: This study was carried out within the Agritech National Research Center and received funding from European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RE-SILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1032 17 June 2022, CN00000022). This manuscript reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data generated and presented in this study are available in the Supplementary Materials.

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

Abbreviations

mildew locus O
essential oil
powdery mildew
multiple EM for motif elicitation
single-likelihood ancestor counting
positively selected codon
negatively selected codon
motif alignment and search tool
Bayesian information criterion
protein family database

References

- 1. Tamokou, J.D.D.; Mbaveng, A.T.; Kuete, V. Antimicrobial Activities of African Medicinal Spices and Vegetables. In *Medicinal Spices and Vegetables from Africa*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 207–237; ISBN 9780128094419.
- Medeiros-Neves, B.; Teixeira, H.F.; von Poser, G.L. The Genus *Pterocaulon* (Asteraceae)—A Review on Traditional Medicinal Uses, Chemical Constituents and Biological Properties. *J. Ethnopharmacol.* 2018, 224, 451–464. [CrossRef] [PubMed]
- 3. Tongnuanchan, P.; Benjakul, S. Essential Oils: Extraction, Bioactivities, and Their Uses for Food Preservation. *J. Food Sci.* 2014, 79, R1231–R1249. [CrossRef] [PubMed]
- 4. Patrignani, F. Lamiaceae in the Treatment of Cardiovascular Diseases. Front. Biosci. 2021, 26, 4909. [CrossRef]
- 5. Nieto, G. Biological Activities of Three Essential Oils of the Lamiaceae Family. Medicines 2017, 4, 63. [CrossRef]
- Lange, B.M.; Mahmoud, S.S.; Wildung, M.R.; Turner, G.W.; Davis, E.M.; Lange, I.; Baker, R.C.; Boydston, R.A.; Croteau, R.B. Improving Peppermint Essential Oil Yield and Composition by Metabolic Engineering. *Proc. Natl. Acad. Sci. USA* 2011, 108, 16944–16949. [CrossRef] [PubMed]
- Wichura, A. Golovinomyces orontii and Other Powdery Mildews on Rosmarinus officinalis. Plant Pathol. Quar. 2012, 2, 162–166. [CrossRef]
- 8. Szpyrka, E.; Słowik-Borowiec, M. Consumer Health Risk to Pesticide Residues in *Salvia officinalis* L. and Its Infusions. *J. Environ. Sci. Health Part B Pestic. Food Contam. Agric. Wastes* **2019**, *54*, 14–19. [CrossRef]
- Zheng, Z.; Nonomura, T.; Appiano, M.; Pavan, S.; Matsuda, Y.; Toyoda, H.; Wolters, A.M.A.; Visser, R.G.F.; Bai, Y. Loss of Function in Mlo Orthologs Reduces Susceptibility of Pepper and Tomato to Powdery Mildew Disease Caused by *Leveillula taurica*. *PLoS* ONE 2013, 8, e70723. [CrossRef]
- 10. Nekrasov, V.; Wang, C.; Win, J.; Lanz, C.; Weigel, D.; Kamoun, S. Rapid Generation of a Transgene-Free Powdery Mildew Resistant Tomato by Genome Deletion. *Sci. Rep.* **2017**, *7*, 482. [CrossRef]

- Ingvardsen, C.R.; Massange-Sánchez, J.A.; Borum, F.; Uauy, C.; Gregersen, P.L. Development of Mlo-Based Resistance in Tetraploid Wheat against Wheat Powdery Mildew. *Theor. Appl. Genet.* 2019, 132, 3009–3022. [CrossRef]
- Kim, M.C.; Lee, S.H.; Kim, J.K.; Chun, H.J.; Choi, M.S.; Chung, W.S.; Moon, B.C.; Kang, C.H.; Park, C.Y.; Yoo, J.H.; et al. Mlo, a Modulator of Plant Defense and Cell Death, Is a Novel Calmodulin-Binding Protein. J. Biol. Chem. 2002, 277, 19304–19314. [CrossRef]
- Zhang, Z.; Sun, Y.; Lu, P.; Feng, C.; Niu, Q.; Lin, G.; Kong, D.; Liu, L.; Luan, S.; Li, L.; et al. Arabidopsis MLO4 Functions as a Ca²⁺ Channel Essential for Mechanosensing in Root Tips. *bioRxiv* 2022, bioRxiv:2022.06.05.494847. [CrossRef]
- Iovieno, P.; Bracuto, V.; Pavan, S.; Lotti, C.; Ricciardi, L.; Andolfo, G. Identification and Functional Inference on the MLO-Family in Viridiplantae. J. Plant Pathol. 2016, 98, 587–594. [CrossRef]
- Iovieno, P.; Andolfo, G.; Schiavulli, A.; Catalano, D.; Ricciardi, L.; Frusciante, L.; Ercolano, M.R.; Pavan, S. Structure, Evolution and Functional Inference on the Mildew Locus O (MLO) Gene Family in Three Cultivated Cucurbitaceae spp. *BMC Genom.* 2015, 16, 1112. [CrossRef]
- Pépin, N.; Hebert, F.O.; Joly, D.L. Genome-Wide Characterization of the MLO Gene Family in *Cannabis sativa* Reveals Two Genes as Strong Candidates for Powdery Mildew Susceptibility. *Front. Plant Sci.* 2021, 12, 729261. [CrossRef] [PubMed]
- Bai, Y.; Pavan, S.; Zheng, Z.; Zappel, N.F.; Reinstädler, A.; Lotti, C.; De Giovanni, C.; Ricciardi, L.; Lindhout, P.; Visser, R.; et al. Naturally Occurring Broad-Spectrum Powdery Mildew Resistance in a Central American Tomato Accession Is Caused by Loss of Mlo Function. *Mol. Plant-Microbe Interact.* 2008, 21, 30–39. [CrossRef] [PubMed]
- Humphry, M.; Reinstädler, A.; Ivanov, S.; Bisseling, T.; Panstruga, R. Durable Broad-Spectrum Powdery Mildew Resistance in Pea Er1 Plants Is Conferred by Natural Loss-of-Function Mutations in PsMLO1. *Mol. Plant Pathol.* 2011, 12, 866–878. [CrossRef] [PubMed]
- Acevedo-Garcia, J.; Kusch, S.; Panstruga, R. Magical Mystery Tour: MLO Proteins in Plant Immunity and Beyond. *New Phytol.* 2014, 204, 273–281. [CrossRef] [PubMed]
- 20. Wan, D.-Y.; Guo, Y.; Cheng, Y.; Hu, Y.; Xiao, S.; Wang, Y.; Wen, Y.-Q. CRISPR/Cas9-Mediated Mutagenesis of VvMLO3 Results in Enhanced Resistance to Powdery Mildew in Grapevine (*Vitis vinifera*). *Hortic. Res.* **2020**, *7*, 116. [CrossRef]
- Wang, Y.; Cheng, X.; Shan, Q.; Zhang, Y.; Liu, J.; Gao, C.; Qiu, J.-L. Simultaneous Editing of Three Homoeoalleles in Hexaploid Bread Wheat Confers Heritable Resistance to Powdery Mildew. *Nat. Biotechnol.* 2014, 32, 947–951. [CrossRef]
- 22. Wang, L.; Lee, M.; Sun, F.; Song, Z.; Yang, Z.; Yue, G.H. A Chromosome-Level Genome Assembly of Chia Provides Insights into High Omega-3 Content and Coat Color Variation of Its Seeds. *Plant Commun.* **2022**, *3*, 100326. [CrossRef] [PubMed]
- Vining, K.J.; Pandelova, I.; Lange, I.; Parrish, A.N.; Lefors, A.; Kronmiller, B.; Liachko, I.; Kronenberg, Z.; Srividya, N.; Lange, B.M. Chromosome-Level Genome Assembly of *Mentha longifolia* L. Reveals Gene Organization Underlying Disease Resistance and Essential Oil Traits. *G3 Genes Genomes Genet.* 2022, *12*, jkac112. [CrossRef]
- 24. Song, Z.; Lin, C.; Xing, P.; Fen, Y.; Jin, H.; Zhou, C.; Gu, Y.Q.; Wang, J.; Li, X. A High-quality Reference Genome Sequence of *Salvia miltiorrhiza* Provides Insights into Tanshinone Synthesis in Its Red Rhizomes. *Plant Genome* **2020**, *13*, e20041. [CrossRef]
- 25. Bornowski, N.; Hamilton, J.P.; Liao, P.; Wood, J.C.; Dudareva, N.; Buell, C.R. Genome Sequencing of Four Culinary Herbs Reveals Terpenoid Genes Underlying Chemodiversity in the Nepetoideae. *DNA Res.* **2020**, *27*, dsaa016. [CrossRef] [PubMed]
- Hamilton, J.P.; Godden, G.T.; Lanier, E.; Bhat, W.W.; Kinser, T.J.; Vaillancourt, B.; Wang, H.; Wood, J.C.; Jiang, J.; Soltis, P.S.; et al. Generation of a Chromosome-Scale Genome Assembly of the Insect-Repellent Terpenoid-Producing Lamiaceae Species, *Callicarpa americana*. *Gigascience* 2020, 9, giaa093. [CrossRef] [PubMed]
- 27. Jia, K.-H.; Liu, H.; Zhang, R.-G.; Xu, J.; Zhou, S.-S.; Jiao, S.-Q.; Yan, X.-M.; Tian, X.-C.; Shi, T.-L.; Luo, H.; et al. Chromosome-Scale Assembly and Evolution of the Tetraploid *Salvia splendens* (Lamiaceae) Genome. *Hortic. Res.* **2021**, *8*, 177. [CrossRef]
- Zheng, X.; Chen, D.; Chen, B.; Liang, L.; Huang, Z.; Fan, W.; Chen, J.; He, W.; Chen, H.; Huang, L.; et al. Insights into Salvianolic Acid B Biosynthesis from Chromosome-scale Assembly of the *Salvia bowleyana* Genome. *J. Integr. Plant Biol.* 2021, 63, 1309–1323. [CrossRef]
- Li, C.-Y.; Yang, L.; Liu, Y.; Xu, Z.-G.; Gao, J.; Huang, Y.-B.; Xu, J.-J.; Fan, H.; Kong, Y.; Wei, Y.-K.; et al. The Sage Genome Provides Insight into the Evolutionary Dynamics of Diterpene Biosynthesis Gene Cluster in Plants. *Cell Rep.* 2022, 40, 111236. [CrossRef]
- 30. Keller, O.; Kollmar, M.; Stanke, M.; Waack, S. A Novel Hybrid Gene Prediction Method Employing Protein Multiple Sequence Alignments. *Bioinformatics* **2011**, *27*, 757–763. [CrossRef]
- Devoto, A.; Piffanelli, P.; Nilsson, I.; Wallin, E.; Panstruga, R.; von Heijne, G.; Schulze-Lefert, P. Topology, Subcellular Localization, and Sequence Diversity of the Mlo Family in Plants. J. Biol. Chem. 1999, 274, 34993–35004. [CrossRef]
- 32. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-Scale Protein Function Classification. *Bioinformatics* **2014**, *30*, 1236–1240. [CrossRef] [PubMed]
- Devoto, A.; Hartmann, H.A.; Piffanelli, P.; Elliott, C.; Simmons, C.; Taramino, G.; Goh, C.S.; Cohen, F.E.; Emerson, B.C.; Schulze-Lefert, P.; et al. Molecular Phylogeny and Evolution of the Plant-Specific Seven-Transmembrane MLO Family. *J. Mol. Evol.* 2003, 56, 77–88. [CrossRef] [PubMed]
- Kosakovsky Pond, S.L.; Frost, S.D.W. Datamonkey: Rapid Detection of Selective Pressure on Individual Sites of Codon Alignments. Bioinformatics 2005, 21, 2531–2533. [CrossRef]
- Carović-Stanko, K.; Liber, Z.; Besendorfer, V.; Javornik, B.; Bohanec, B.; Kolak, I.; Satovic, Z. Genetic Relations among Basil Taxa (*Ocimum* L.) Based on Molecular Markers, Nuclear DNA Content, and Chromosome Number. *Plant Syst. Evol.* 2010, 285, 13–22. [CrossRef]

- 36. Panchy, N.; Lehti-Shiu, M.; Shiu, S.H. Evolution of Gene Duplication in Plants. Plant Physiol. 2016, 171, 2294–2316. [CrossRef]
- Appiano, M.; Catalano, D.; Santillán Martínez, M.; Lotti, C.; Zheng, Z.; Visser, R.G.F.; Ricciardi, L.; Bai, Y.; Pavan, S. Monocot and Dicot MLO Powdery Mildew Susceptibility Factors Are Functionally Conserved in Spite of the Evolution of Class-Specific Molecular Features. *BMC Plant Biol.* 2015, 15, 257. [CrossRef]
- Kusch, S.; Pesch, L.; Panstruga, R. Comprehensive Phylogenetic Analysis Sheds Light on the Diversity and Origin of the MLO Family of Integral Membrane Proteins. *Genome Biol. Evol.* 2016, *8*, 878–895. [CrossRef]
- Lin, Y.; Cheng, Y.; Jin, J.; Jin, X.; Jiang, H.; Yan, H.; Cheng, B. Genome Duplication and Gene Loss Affect the Evolution of Heat Shock Transcription Factor Genes in Legumes. *PLoS ONE* 2014, 9, e102825. [CrossRef]
- 40. Andolfo, G.; Dohm, J.C.; Himmelbauer, H. Prediction of NB-LRR Resistance Genes Based on Full-length Sequence Homology. *Plant J.* **2022**, *110*, 1592–1602. [CrossRef]
- 41. Andolfo, G.; Sánchez, C.S.; Cañizares, J.; Pico, M.B.; Ercolano, M.R. Large-Scale Gene Gains and Losses Molded the NLR Defense Arsenal during the Cucurbita Evolution. *Planta* **2021**, 254, 82. [CrossRef]
- Elliott, C.; Müller, J.; Miklis, M.; Bhat, R.A.; Schulze-Lefert, P.; Panstruga, R. Conserved Extracellular Cysteine Residues and Cytoplasmic Loop–Loop Interplay Are Required for Functionality of the Heptahelical MLO Protein. *Biochem. J.* 2005, 385, 243–254. [CrossRef] [PubMed]
- Panstruga, R. Discovery of Novel Conserved Peptide Domains by Ortholog Comparison within Plant Multi-Protein Families. Plant Mol. Biol. 2005, 59, 485–500. [CrossRef] [PubMed]
- 44. Sun, X.; Cao, Y.; Wang, S. Point Mutations with Positive Selection Were a Major Force during the Evolution of a Receptor-Kinase Resistance Gene Family of Rice. *Plant Physiol.* **2006**, *140*, 998–1008. [CrossRef] [PubMed]
- 45. Amoroso, C.G.; D'Esposito, D.; Cigliano, R.A.; Ercolano, M.R. Comparison of Tomato Transcriptomic Profiles Reveals Overlapping Patterns in Abiotic and Biotic Stress Responses. *Int. J. Mol. Sci.* **2023**, *24*, 4061. [CrossRef]
- Andolfo, G.; Di Donato, A.; Darrudi, R.; Errico, A.; Aiese Cigliano, R.; Ercolano, M.R. Draft of Zucchini (*Cucurbita pepo* L.) Proteome: A Resource for Genetic and Genomic Studies. *Front. Genet.* 2017, *8*, 181. [CrossRef]
- Terauchi, R.; Yoshida, K. Towards Population Genomics of Effector–Effector Target Interactions. *New Phytol.* 2010, 187, 929–939.
 [CrossRef]
- 48. Andolfo, G.; Ercolano, M.R. Plant Innate Immunity Multicomponent Model. Front. Plant Sci. 2015, 6, 987. [CrossRef]
- Ercolano, M.R.; D'Esposito, D.; Andolfo, G.; Frusciante, L. Multilevel Evolution Shapes the Function of NB-LRR Encoding Genes in Plant Innate Immunity. *Front. Plant Sci.* 2022, 13, 1007288. [CrossRef]
- 50. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER Web Server: Interactive Sequence Similarity Searching. *Nucleic Acids Res.* 2011, 39, W29–W37. [CrossRef]
- Marchler-Bauer, A.; Bo, Y.; Han, L.; He, J.; Lanczycki, C.J.; Lu, S.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; et al. CDD/SPARCLE: Functional Classification of Proteins via Subfamily Domain Architectures. *Nucleic Acids Res.* 2017, 45, D200–D203. [CrossRef]
- 52. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; Mcgettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X Version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [CrossRef] [PubMed]
- 53. Whelan, S.; Goldman, N. A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. *Mol. Biol. Evol.* **2001**, *18*, 691–699. [CrossRef]
- Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol. Biol. Evol. 2021, 38, 3022–3027. [CrossRef] [PubMed]
- 55. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. Evolution 1985, 39, 783. [CrossRef]
- 56. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. Nucleic Acids Res. 2015, 43, W39–W49. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.