

RESEARCH ARTICLE

Open Access

Evolutionary conservation of *MLO* gene promoter signatures



Giuseppe Andolfo¹, Paolo Iovieno¹, Luigi Ricciardi², Concetta Lotti³, Edgardo Filippone¹, Stefano Pavan^{2,4*} and Maria Raffaella Ercolano^{1*}

Abstract

Background: Powdery mildew (PM) is a widespread fungal disease of plants in temperate climates, causing significant economic losses in agricultural settings. Specific homologs of the *MLO* gene family are PM susceptibility factors, as their loss-of-function results in durable PM resistance (*mlo* resistance) in several plant species. The role of *MLO* susceptibility genes in plant-pathogen interactions is still elusive, however it is known that they are strongly upregulated following PM infection.

Results: In this study, we investigated the structure of 414 Putative Promoter Regions (PPRs) of *MLO* genes and highlighted motif and regulatory element patterns related to genomic relationships among species and phylogenetic distance among homologs. A TC box-like motif and a thymine-rich motif were found to be overrepresented in *MLO* genes transcriptionally upregulated upon infection with PM fungi. As proof of concept, we showed that the expression of a melon (*Cucumis melo* L.) gene enriched for the motifs above mentioned was strongly upregulated upon infection with the PM fungus *Podosphaera xanthii*.

Conclusion: While identifying a candidate *MLO* susceptibility gene in melon, this study provides insight on the transcriptional control of *MLO* genes and indicates diagnostic features useful to identify *MLO* susceptibility genes across species affected by the PM disease.

Keywords: Cis-acting regulatory element, Motif, *MLO*, Powdery mildew resistance, Transcription

Background

The powdery mildew (PM) disease, caused by ascomycete fungi from the order of Erysiphales, can severely affect the yield and the quality of several agricultural species cultivated in temperate regions, including melon (*Cucumis melo* L.). Chemical control of PM is generally costly and raises important environmental issues. Therefore, understanding molecular mechanisms underlying genotypic resistance to PM is of great interest for agricultural genetics and breeding.

Specific homologs of the *Mildew Locus O* (*MLO*) plant-specific gene family are PM susceptibility factors, as their loss-of-function results in a durable and broad-spectrum form of resistance, known as *mlo*

resistance [1]. Originally discovered in barley, *mlo* resistance was later shown to occur in several other species, namely Arabidopsis, tomato, pea, pepper, tobacco and wheat [2–10], and proved to be effective against both epiphytic and endophytic PM species [10, 11]. Therefore, several authors suggested the opportunity of a plant breeding strategy based on the identification and inactivation of *MLO* susceptibility genes [1, 7].

With no exception known so far, *MLO* susceptibility genes belong to two phylogenetic clades, referred to as clade IV and V, which are specific for monocots and eudicots, respectively [12–16]. However, not all clade IV and clade V *MLO* genes are PM susceptibility factors. For example, while silencing of the tomato gene *SIMLO1* results in *mlo* resistance to the powdery mildew pathogen *Oidium neolycopersici*, silencing of the other clade V genes *SIMLO3*, *SIMLO5* and *SIMLO8* is still associated with a susceptible phenotype [17]. Previous investigations showed *MLO* homologs playing a major role in

* Correspondence: stefano.pavan@uniba.it; ercolano@unina.it

²Department of Soil, Plant and Food Science, University of Bari "Aldo Moro", Via Amendola 165/A, 70126 Bari, Italy

¹Department of Agricultural Sciences, University of Naples "Federico II", Via Università 100, 80055 Portici (Naples), Italy

Full list of author information is available at the end of the article



PM susceptibility, including barley *HvMLO*, tomato *SlMLO1* and pepper *CaMLO2*, are transcriptionally up-regulated a few hours after infection with PM fungi [2, 11, 18].

Transcription is a very important biological process regulated at several stages. Understanding how gene regulation is orchestrated is an important challenge for characterizing complex events such as plant-pathogen interactions [19, 20]. Proteins known as transcription factors (TFs) modulate gene expression in specific ways through their binding to DNA regulatory elements [21]. Therefore, a key aim of gene expression analysis is the identification of transcription factors and DNA regulatory elements. Throughout the last decade, empirical data have accumulated suggesting that mutations in regulatory elements could be a major cause of phenotypic divergence [22, 23].

The term ‘promoter’ is used to designate the genomic sequence located upstream of a gene, which tends to contain cis-acting regulatory elements (CREs). Based on the distance from the transcription start site (TSS), the terms of ‘proximal promoter’ (several hundred nucleotides upstream of the TSS) and ‘distal promoter’ (thousands and more nucleotides upstream of the TSS) are also used. Most CREs are composed of 5–20 nucleotides and are localized in the proximal promoter [24]. It is generally thought that genes having similar expression patterns contain common motifs in their promoter regions [25].

The availability of whole genome data and next generation sequencing technology has opened up new avenues for the analysis of gene regulation and expression. The development of specialized databases of CREs in plants [26] and the development of bioinformatics tools to discover specific motifs in DNA sequences [27] greatly facilitate in silico analysis of promoters. Various methodologies can aid the identification of regulatory motifs, including deletion based functional analysis, comparative genomics, analysis of their co-expressed genes, and ChIP-Chip or ChIP-Seq [28–30] of co-expressed genes for overrepresented motifs [31, 32].

Prior to this study, a unique CRE was associated with the promoter region of the Arabidopsis susceptibility gene *AtMLO2* and co-expressed genes, which was found to be responsive to senescence, light stress, wounding and PM inoculation [33]. The aim of our research was to mine motifs and regulatory elements in the Putative Promoter Regions (PPRs) of *MLO* genes from different plant species, and to identify putative regulatory sequences specific for clade V homologs or for *MLO* genes known to be readily upregulated upon PM infection. The conclusions drawn by bioinformatics analysis were substantiated by the transcriptional characterization of

MLO genes that we have recently described in the genome of melon [13].

Results and discussion

Identification and distribution of *MLO* CREs in

Viridiplantae

Putative Promoter Regions (PPRs), corresponding to the 2 Kbp sequence located upstream the predicted TSS, were successfully extracted from 414 out of 447 *MLO* genes previously described in 25 plant genomes [13]. To identify CREs putatively involved in the transcriptional regulation of *MLO*-expression, a search was performed against the PLACE database. Three hundred and sixteen non-redundant CREs, distributed in variable number and order, were detected (Additional file 1: Table S1). In total, 186,060 elements were annotated in the PPRs of our 414 *MLO* gene dataset, and an average number of 449 CREs in forward- and reverse-complement orientation per homolog was detected (Table 1). *Capsicum annuum* and *Arabidopsis thaliana* genomes had the highest (476) number of CREs per *MLO* homolog, whereas the *Brachypodium distachyon* genome had the lowest number (346). The most abundant CREs were DOFCOREZM (PLACE-ID:S000265), CACTFTPPCA1 (ID:S000449) and CAATBOX1 (ID:S000028), which were recorded 10,983, 10,431 and 8533 times, respectively, in the whole dataset (Additional file 1: Table S1). DOFCOREZM is the core site required for binding of DOF (DNA-binding One zinc Finger) proteins, a family of transcription factors involved in many biological processes in higher plants [34]. The tetranucleotide CACT, present in the CACTFTPPCA1 element, is the key component of mesophyll expression in C4 plants [35, 36].

Generally, flowering plants exhibited a family-specific pattern in terms of number of elements and global density distribution (Fig. 1). Indeed, homogenous CRE profiles were observed for Cucurbitaceae (*Cucumis melo* and *Cucumis sativus*), Rosaceae (*Malus domestica* and *Prunus persica*), Poaceae (*Brachypodium distachyon*, *Oryza sativa* and *Sorghum bicolor*), Solanaceae (*Solanum lycopersicum* and *Nicotiana tabacum*) and Brassicaceae (*Arabidopsis thaliana* and *Capsella rubella*). In other cases, the *MLO*-CRE profile appeared to be family-independent, such as in the case of *Brassica rapa* and *Glycine max*. This might be due to missing information on the CRE sequence of some *MLO* family members, or the occurrence of true species-specific patterns. Analysis of individual CREs indicated that nineteen CREs were specifically present in monocots or eudicots. As expected, the *MLO* CRE profile of algae (*Volvox carteri* and *Chlamydomonas reinhardtii*) was clearly distinct from the one of flowering plants (Fig. 1), and included 147 elements which might be associated with

Table 1 Candidate cis-regulatory elements (CREs) in the putative promoter regions (PPRs) of *MLO* genes

Family	Species	N. Analysed Promoters	N. CREs ^a	N. CREs/ <i>MLO</i>
Volvocaceae	<i>Volvox carteri</i>	3 (3) ^b	1047	349
Chlamydomonadaceae	<i>Chlamydomonas reinhardtii</i>	4 (4)	941	235
Brassicaceae	<i>Arabidopsis thaliana</i>	15 (15)	7147	476
	<i>Brassica rapa</i>	23 (23)	10,212	444
	<i>Capsella rubella</i>	16 (16)	7288	456
Cucurbitaceae	<i>Citrullus lanatus</i>	14 (14)	6184	442
	<i>Cucumis melo</i>	15 (16)	6980	465
	<i>Cucumis sativus</i>	13 (13)	6105	470
Euphorbiaceae	<i>Manihot esculenta</i>	19 (19)	8899	468
Fabaceae	<i>Glycine max</i>	41 (41)	19,242	469
	<i>Medicago truncatula</i>	16 (16)	7170	448
	<i>Phaseolus vulgaris</i>	20 (20)	9418	471
Poaceae	<i>Brachypodium distachyon</i>	13 (13)	5157	397
	<i>Oryza sativa</i>	12 (12)	5282	440
	<i>Sorghum bicolor</i>	15 (15)	6145	410
	<i>Triticum aestivum</i>	26 (55)	10,595	408
Rosaceae	<i>Fragaria vesca</i>	17 (17)	7263	427
	<i>Malus domestica</i>	20 (21)	8533	427
	<i>Prunus persica</i>	19 (19)	8869	467
Solanaceae	<i>Solanum lycopersicum</i>	17 (17)	8051	474
	<i>Nicotiana tabacum</i>	14 (14)	6623	473
	<i>Capsicum annuum</i>	15 (15)	7139	476
	<i>Solanum melongena</i>	16 (18)	7489	416
	<i>Solanum tuberosum</i>	12 (12)	5573	464
Vitaceae	<i>Vitis vinifera</i>	19 (19)	8708	458
Total		414 (447)	186,060	10,930

^aReferred to the cis-regulatory elements (CREs) annotated on the 2 Kbp upstream sequences from the putative initiation codon. ^bIn brackets, the number of total annotated *MLO* genes characterized by Iovieno et al. (2016)

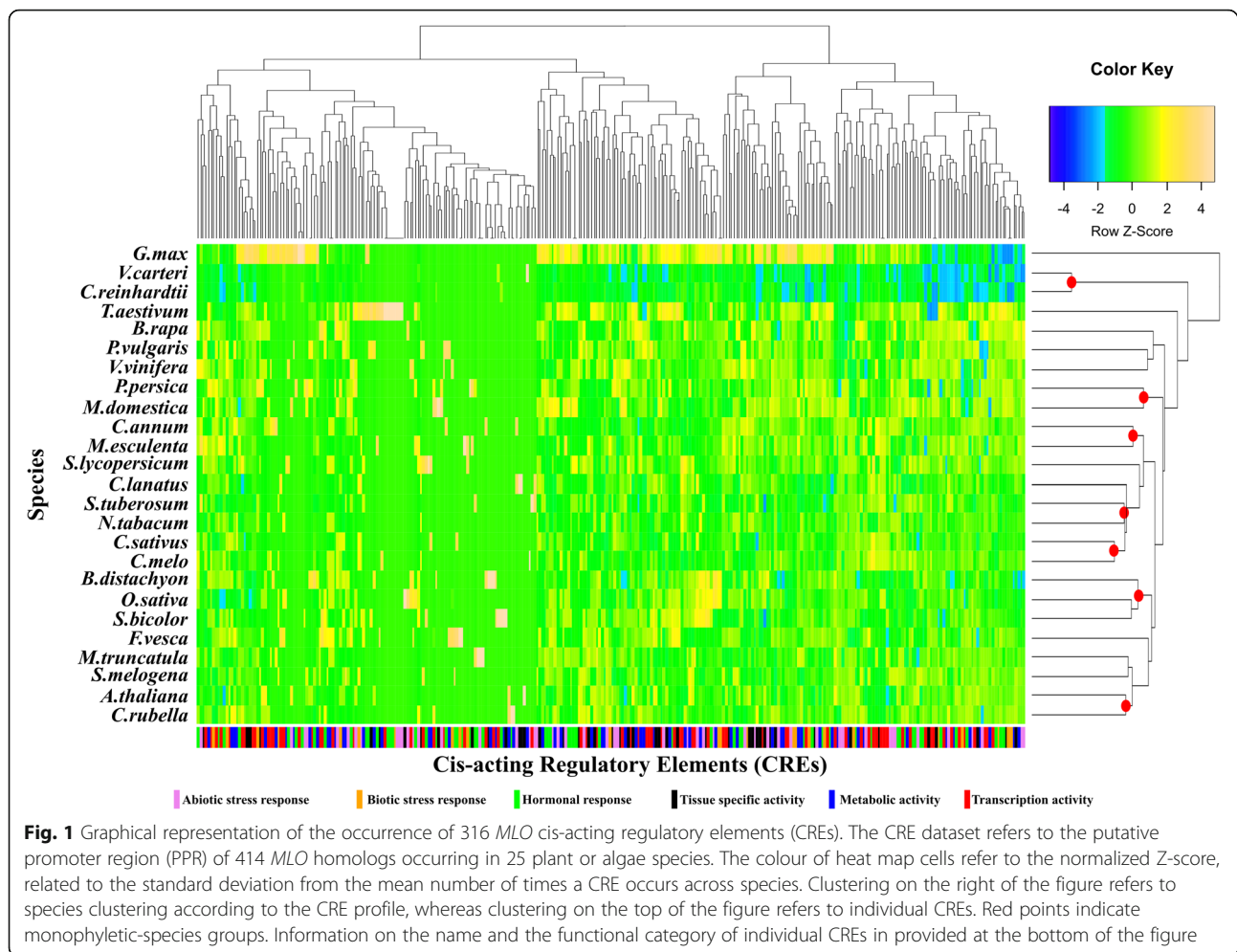
conserved features during the evolution of green plants and involved in the transcription of *MLO* homologs.

Distribution of *MLO* CREs in phylogenetic clades

Identified CREs were subdivided in six functional categories: metabolic activity (MA) (63), transcription activity (TA) (74), tissue specific activity (TSA) (45), hormonal response (HR) (60), biotic stress response (BSR) (24) and abiotic stress response (ASR) (50). A phylogenetic clade-specific CRE profile was obtained by normalizing, for each clade, the abundance of each CRE class on the total of CREs belonging to the same class [$\text{CREs}^{\text{CLADE CLASS}}/\text{CREs}^{\text{CLASS}}$] (Additional file 1: Table S2). Overall, the distribution of the classes was not markedly different among the eight phylogenetic clades under study (Fig. 2). MA, HR and BSR classes were evenly represented in all phylogenetic clades. A higher percentage of CREs associated with TA was observed in the *MLO*-PPRs of clade I, III and V (31%) and IV (33%). The highest percentage in CRE related to TSA and

ASR was observed in the clades IV (26%) and VIII (20%) (Fig. 2). A core of 119 (~40% of total) CRE elements was highly conserved in all regulatory profiles of the eight clades. This subset could help to explain the coordinate up-regulation of *MLO* homologs from different clades in response to the same biotic or abiotic stress [37]. Possibly, a core regulatory structure of PPRs of *MLO*s belonging to different clades promotes a basic gene activation to environmental cues. No specific CRE was identified in either pathogen-responsive or clade V *MLO* genes (Table 2).

In order to identify CREs enriched in specific phylogenetic clades, a motif enrichment analysis was performed. Five clades (I, II, III, IV and V) displayed significantly enriched CREs (e-CREs) (Additional file 1: Table S3). In total, 43 over-represented CREs were identified, of which about 40% were grouped into the TA-class (Additional file 1: Table S3). *MLO* PPRs clustering in clades I and II displayed enrichment of specific transcription factor binding sites (MYCATRD22 for



clade I; RAV1BAT and WBOXNTCHN48 for clade II). Clade III showed five e-CREs (POLASIG1, POLASIG2, GT1GMSCAM4, TATABOX5 and MARTBOX). The element AGCBOXNPGLB, associated with stress signal-response factors (as ER and PR-proteins), was the only CRE significantly enriched in clade IV. Interestingly, more than 50% (24 out of 43) of identified e-CREs, mostly related to the TA class, were associated with PPRs of *MLO* homologs clustering in clade V (adjusted p -value < 0.01).

De novo identification of putative regulatory motifs in pathogen-responsive *MLO*s

Additional regulatory motifs in PPRs of transcriptionally responsive *MLO*-homologs were disclosed. From a first global alignment of the PPR-dataset, the highest sequence identity was observed in the first 200 bp upstream of the predicted TSS. A similarity matrix restricted to this region highlighted the occurrence of a large cluster containing about 65% of the PPRs, including those of 17 (out of 18) *MLO* genes known to be up-regulated upon PM infection (Table 2 and Fig. 3a).

Multiple alignment of the same region revealed higher sequence identity (33%) and AT-content (70%) in the PPR of *MLO* homologs upregulated upon PM infection compared to the rest of the dataset (27 and 64%, respectively) (Fig. 3a and b). Interestingly, a significant association between expression and the AT-content of PPRs was previously reported for genes different from *MLO*s [38].

A MEME D-mode analysis was carried out to identify enriched motifs in the PPR of *MLO* genes upregulated during the interaction with PM fungi or included in clade V (Table 2). Notably, two overrepresented motifs (M1 and M2) were found in *MLO* genes up-regulated upon PM infection (Additional file 1: Table S4). Except for *AtMLO3*, these motifs were present within the 0.4 Kbp sequence upstream the predicted TSS. Three conserved motifs (M3, M4 and M5) were over-represented in clade V homologs (Additional file 1: Table S4). The consensus sequence of M1 (Fig. 3c) is similar to a TC-element sequence. Previous studies suggested that TC-elements might constitute a novel class of regulatory sequences participating to the complex modulation of

Table 2 Features of *MLO* genes upregulated upon powdery mildew infection and/or belonging to the phylogenetic clade V

Species	Upregulated	Clade V	Phylogenetic clade (Iovieno et al., 2016)	Consortium gene ID	Reference for upregulated <i>MLO</i> s
<i>Arabidopsis thaliana</i>	AtMLO2	AtMLO2	V	AT1G11310	[48]
	AtMLO3	–	VI	AT3G45290	
	AtMLO6	AtMLO6	V	AT1G61560	
	AtMLO12	AtMLO12	V	AT2G39200	
<i>Vitis vinifera</i>	VvMLO3	VvMLO3	V	GSVIVG01025653001	[13, 49]
	VvMLO4	VvMLO4	V	GSVIVG01025162001	
	VvMLO9	–	VII	GSVIVG01020617001	
	–	VvMLO13	V	GSVIVG01016302001	
	VvMLO17	VvMLO17	V	GSVIVG01016304001	
<i>Malus domestica</i>	MdMLO11	MdMLO11	V	MDP0000239643	[50]
	–	MdMLO5	V	MDP0000163089	
	–	MdMLO7	V	MDP0000123907	
	MdMLO18	–	VII	MDP0000928368	
	MdMLO19	MdMLO19	V	MDP0000168714	
<i>Capsicum annuum</i>	–	CaMLO1	V	Capana11g000102	[10]
	CaMLO2	CaMLO2	V	Capana06g001935	
<i>Solanum lycopersicum</i>	SIMLO1	SIMLO1	V	Solyc04g049090	[17]
	–	SIMLO3	V	Solyc06g010030	
	SIMLO4	–	III	Solyc00g007200	
	–	SIMLO5	V	Solyc03g095650	
	–	SIMLO8	V	Solyc11g069220	
	SIMLO14	–	I	Solyc07g063260	
<i>Cucumis sativus</i>	CsMLO1	CsMLO1	V	Csa1M085890.	[51]
	–	CsMLO8	V	Csa5M623470	
	–	CsMLO11	V	Csa6M292430	
<i>Citrullus lanatus</i>	–	CIMLO2	V	Clao005044	[7, 13]
	–	CIMLO5	V	Clao008753	
	CIMLO12	CIMLO12	V	Clao020573	
<i>Cucumis melo</i>	–	CmMLO3	V	MELO3C005044	This work
	CmMLO5	CmMLO5	V	MELO3C012438	
	–	CmMLO12	V	MELO3C025761	

represent a general signature of *MLO* genes, regulating expression in response to homeostasis perturbations.

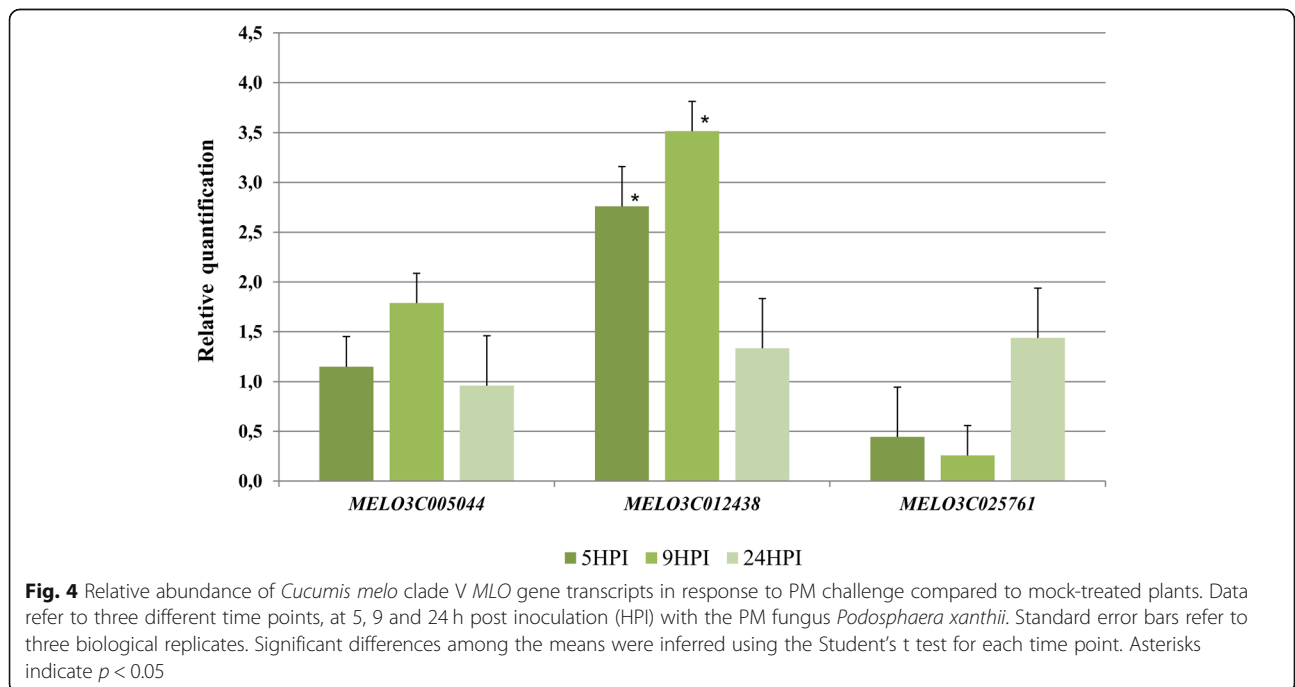
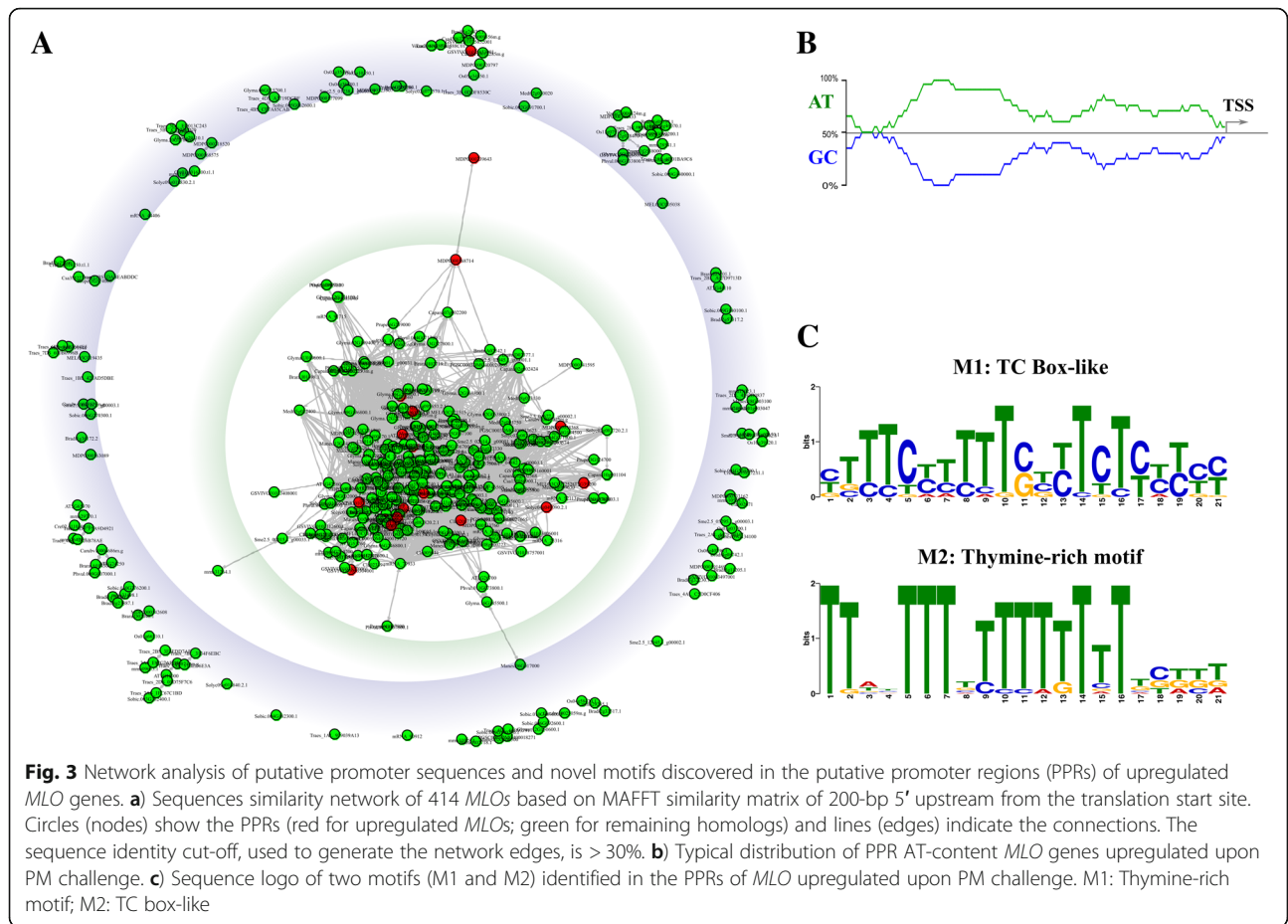
Transcriptional validation of bioinformatics results

In silico analyses highlighted the overrepresentation of two putative PPR regulatory elements, namely a TC box-like and a thymine-rich motif, in *MLO* homologs up-regulated upon PM infection. As proof of concept, we analysed the transcriptional response of three clade V genes previously found in the genome of melon [13] (*MELO3C005044*: *CmMLO3*, *MELO3C012438*: *CmMLO5* and *MELO3C025761*: *CmMLO12*). Among them, *MELO3C012438* shows the best combined match (combined *p*-value <1e-5) for the over-represented motifs in the PPR

region. Following artificial inoculation with the PM fungus *Podosphaera xanthii*, a strong upregulation of *MELO3C012438* was observed at both 5 and 9 h post-inoculation. In contrast, no significant difference was detected for *MELO3C005044*, and *MELO3C025761* (Fig. 4). As eudicot *MLO* susceptibility genes belong to clade V and are upregulated upon PM challenge [2, 11, 18], our results indicate that *MELO3C012438* is a PM susceptibility factor in melon.

Conclusions

Mechanisms controlling the expression of *MLO* homologs are not clear. Here we report the characterization of *MLO* PPRs in a large dataset extracted from 25 plant



and algal species. Distribution patterns of putative *MLO* regulatory elements varied according to genomic relatedness among species and phylogenetic distance among homologs. Two motifs, referred to as M1 and M2, were overrepresented in the PPRs of clade V *MLO* genes upregulated upon PM infection, which are known to act as PM susceptibility genes. In accordance with this finding, we showed that, in melon, a clade V *MLO* gene enriched for M1 and M2 is upregulated following challenge with the melon PM fungus *Podosphaera xanthii*. Overall, this study provides insight on the PPR of the *MLO* gene family, reports the characterization of a candidate *MLO* susceptibility gene in melon, and highlights PPR diagnostic features that could be used to identify *MLO* susceptibility genes across cultivated species. This information is extremely valuable for breeding purposes, as the inactivation of *MLO* susceptibility genes, for example by targeted silencing and genome editing, might lead to the identification of PM resistant genotypes.

Methods

Putative promoter region dataset

447 *MLO* genes from twenty-five different genomes (*Volvox carteri*, *Chlamydomonas reinhardtii*, *Arabidopsis thaliana*, *Brassica rapa*, *Capsella rubella*, *Citrullus lanatus*, *Cucumis melo*, *Cucumis sativus*, *Manihot esculenta*, *Glycine max*, *Medicago truncatula*, *Phaseolus vulgaris*, *Brachypodium distachyon*, *Oryza sativa*, *Sorghum bicolor*, *Triticum aestivum*, *Fragaria vesca*, *Malus domestica*, *Prunus persica*, *Solanum lycopersicum*, *Nicotiana tabacum*, *Capsicum annuum*, *Solanum melongena*, *Solanum tuberosum* and *Vitis vinifera*), described in the work of Iovieno et al. (2016), were considered as initial dataset. Putative promoter regions (PPRs) (2 Kbp upstream the predicted TSS reported in gene annotations) were extracted using BEDTools v2.17.0 (<http://samtools.sourceforge.net>) and 414 PPRs (out of 447) with a PPR length > 1 Kbp were used further.

Detection and distribution of CREs in Viridiplantae *MLO* homologs

PPR-sequences were scanned in forward and reverse complement orientation to search for cis-acting regulatory elements (CREs) present in the Plant cis-acting regulatory DNA elements (PLACE) database (<https://sogo.dna.affrc.go.jp/cgi-bin/sogo.cgi?lang=en&pj=640&action=page&page=newplace>) [26]. The most stringent identification criterion (exact match) was used. The Needleman-Wunsch algorithm (default setting) was used to generate a non-redundant dataset by eliminating duplicated motifs defined by 'identical sequences'. In order to obtain a CRE map of Viridiplantae *MLO* homologs, a hierarchical clustering analysis of annotated CREs was performed. A heat map was generated starting from a

CRE-abundance matrix using the 'GPLOTS' R software package [41].

Characterization of CREs with respect to phylogenetic clades and transcriptional response

Identified CREs (316) were manually assigned to six different functional categories (classes) based on the regulatory functions reported in literature for each element:

- **Metabolic activity** (MA): implicated into biosynthesis of primary and secondary metabolites;
- **Transcription activity** (TA): recognised by transcription factors;
- **Tissue specific activity** (TSA): associated to a specific tissue;
- **Hormonal response** (HR): induced by phytohormones;
- **Biotic stress response** (BSR): induced by biotic stresses;
- **Abiotic stress response** (ASR): induced by abiotic stresses.

A clade-specific CRE profile was obtained by normalizing, for each clade, the abundance of each CRE class on the total of CREs belonging to the same class [$\text{CREs}^{\text{CLADE CLASS}} / \text{CREs}^{\text{CLASS}}$]. In order to identify CREs overrepresented in specific clades, PPRs associated with each clade were compared to those of the remaining dataset, using the Analysis of Motif Enrichment (AME) tool available in the MEME suite and the default cut-off value [42]. Motif enrichment significance was assessed using the Ranksum test, also known as the Mann-Whitney U test, with subsequent multiple testing correction applied to the obtained *P* values (adjusted *p*-value < 0.05). Analysis of motif enrichment and individual CRE annotation were also carried out to compare PPR of homologs upregulated during PM infection or included in clade V with those of other *MLO* homologs.

PPR similarity network

MAFFT version v6.814b was used to align the first 200 bp of PPRs, and to generate a similarity matrix, using the EINS-i algorithm [43]. A weighted correlation network analysis was carried out using the 'WGCNA' R software package [44].

De novo annotation and function prediction of *MLO*-motifs

The Multiple EM for Motif Elicitation (MEME) algorithm [27] was used to identify motifs that are specifically enriched in the PPRs of *MLO* genes included in clade V or previously shown to be upregulated upon PM challenge (Table 2). The analysis was carried out in discriminative-mode (D-mode), using the default cut-off value for statistical confidence.

The best combined match of enriched motifs identified by MEME analysis was detected using Motif Alignment & Search Tool (MAST) (default cut-off: p-value $1e-4$ and E-value < 10) [27]. MAST was also used to confirm the presence of the motif previously described by Humphry et al. (2010) [33] in the PPR dataset, using two different settings (p-value $< 1e-4$ and p-value $< 1e-3$). Enriched motifs were analysed using the Gene Ontology for Motifs (GOMo, <http://meme-suite.org>) tool [29] to determine if any motif was significantly associated with genes linked to one or more Genome Ontology (GO) terms [45].

Transcriptional characterization of melon *MLO* genes

Plants of *Cucumis melo* (cv. Tendral Verde TS258) were inoculated with the PM fungus *Podosphaera xantii* by brushing heavily infected leaves collected in the field. Plants brushed with healthy leaves were used as experimental controls. Leaf samples were collected at 5, 9 and 24 h after inoculation (three biological replicates for each treatment and three technical replicates). Protocols for RNA isolation, cDNA preparation and qRT-PCR were the same as those reported by Iovieno et al. (2015) [7]. Relative transcript abundance of the three clade V melon genes *MELO3C005044*, *MELO3C012438* and *MELO3C025761* was assessed using the $2^{-\Delta\Delta Ct}$ method [46], using the β -actin gene as housekeeping for target gene normalization [47]. The Student's t-test was used to test the null hypothesis. All the primer pairs used for gene expression analysis are listed in Additional file 1: Table S5. For each primer pair, a dissociation kinetics analysis was performed to ensure the specificity of the amplification product and amplification efficiency was assessed by generating a standard curve with five serial 10-fold dilutions.

Additional file

Additional file 1: Table S1. List of 316 candidate CREs in the putative promoter regions of MLOs characterized by Iovieno et al., (2016). **Table S2.** Functional annotation of CREs in *MLO*-PPRs. The classification is based on the molecular function and the biological processes of the genes containing CREs. **Table S3.** List of enriched-CREs identified for each phylogenetic clade. **Table S4.** List of enriched motifs identified in the putative promoter regions (PPRs) of *MLO* genes upregulated during the interaction with PM fungi and/or belonging to the phylogenetic clade V. **Table S5.** Gene-specific primers used in this study for gene-expression analysis and their amplification efficiency. (XLSX 31 kb)

Abbreviations

AME: Analysis of Motif Enrichment; ASR: Abiotic Stress Response; BSR: Biotic Stress Response; CRE: Cis-acting Regulatory Element; e-CREs: Enriched-CRE; GO: Genome Ontology; GOMo: Gene Ontology for Motif; HR: Hormonal Response; MA: Metabolic Activity; MAST: Motif Alignment & Search Tool; MEME D-mode: Multiple EM for Motif Elicitation Discriminative-mode; MLO: Mildew Locus O; PLACE: Plant cis-acting regulatory DNA element; PM: Powdery Mildew; PPR: Putative Promoter Region; TA: Transcription Activity; TSA: Tissue Specific Activity; TSS: Transcription Start Site

Acknowledgements

We are grateful to Prof. Luigi Frusciante for valuable discussions on conducted analysis.

Funding

Financial support for this work was provided by the Italian Ministry of University and Research (GenHort Project).

Availability of data and materials

All the supporting data are included within the article. The Supplementary Material (Additional file 1: Tables S1, S2, S3, S4 and S5) for this article can be found online at: (<https://figshare.com/s/6adb5ce53cd02838ffbd>). Users can download and use the data freely for research purpose only with acknowledgment to us and quoting this paper as reference to the data.

Authors' contributions

GA conceived the idea of study and was centrally involved in bioinformatic analysis, plant inoculation, results interpretation and manuscript writing. PI was involved in data analysis, plant inoculation and qRT-PCR experiments. LR and CL critically revised the manuscript. EF provided a critical reading of the manuscript. SP was involved in experimental planning, results interpretation and manuscript writing. MRE coordinated the project and contributed to results interpretation and manuscript writing. All of the authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable. This manuscript does not report on or involve the use of any animal or human data or tissue.

Consent for publication

Not applicable.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Agricultural Sciences, University of Naples "Federico II", Via Università 100, 80055 Portici (Naples), Italy. ²Department of Soil, Plant and Food Science, University of Bari "Aldo Moro", Via Amendola 165/A, 70126 Bari, Italy. ³Department of Agriculture, Food and Environmental Science, University of Foggia, via Napoli 25, 71100 Foggia, Italy. ⁴Institute of Biomedical Technologies, National Research Council (CNR), Via Amendola 122/D, 70126 Bari, Italy.

Received: 29 March 2018 Accepted: 31 March 2019

Published online: 17 April 2019

References

- Pavan S, Jacobsen E, Visser RG, Bai Y. Loss of susceptibility as a novel breeding strategy for durable and broad-spectrum resistance. *Mol Breed*. 2010;25:1–12. <https://doi.org/10.1007/s11032-009-9323-6>.
- Bai Y, Pavan S, Zheng Z, Zappel NF, Reinstädler A, Lotti C, 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;21:30–9. <https://doi.org/10.1094/MPMI-21-1-0030>.
- Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, et al. The barley Mlo gene: a novel control element of plant pathogen resistance. *Cell*. 1997;88:695–705. [https://doi.org/10.1016/S0092-8674\(00\)81912-1](https://doi.org/10.1016/S0092-8674(00)81912-1).
- Consonni C, Humphry ME, Hartmann HA, Livaja M, Westphal L, Vogel J, et al. Conserved requirement for a plant host cell protein in powdery mildew pathogenesis. *Nat Genet*. 2006;38:716–20. <https://doi.org/10.1038/ng1806>.
- Fujimura T, Sato S, Tajima T, Arai M. Powdery mildew resistance in the Japanese domestic tobacco cultivar Kokubu is associated with aberrant splicing of MLO orthologues. *Plant Pathol*. 2016;(8):1358–65. <https://doi.org/10.1111/ppa.12498>.

6. Humphry ME, Reinstaedler 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;12:866–78. <https://doi.org/10.1111/j.1364-3703.2011.00718.x>.
7. Iovieno P, Andolfo G, Schiavulli A, Catalano D, Ricciardi L, Frusciante et al. structure, evolution and functional inference on the mildew locus O (MLO) gene family in three cultivated Cucurbitaceae. *BMC Genomics*. 2015;16:1–13. <https://doi.org/10.1186/s12864-015-2325-3>.
8. Pavan S, Schiavulli A, Appiano M, Marcotrigiano AR, Cillo F, Visser RGF, et al. Pea powdery mildew *er1* resistance is associated to loss-of-function mutations at a MLO homologous locus. *Theor Appl Genet*. 2011;123:1425–31. <https://doi.org/10.1007/s00122-011-1677-6>.
9. Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol*. 2014;32:947–51. <https://doi.org/10.1038/nbt.2969>.
10. Zheng Z, Nonomura T, Appiano M, Pavan S, Matsuda Y, Toyoda H, et al. 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. <https://doi.org/10.1371/journal.pone.0070723>.
11. Seifi A, Gao D, Zheng Z, Pavan S, Faino L, Visser RGF, et al. Genetics and molecular mechanisms of resistance to powdery mildews in tomato (*Solanum lycopersicum*) and its wild relatives. *Eur J Plant Pathol*. 2014;138:641–65. <https://doi.org/10.1007/s10658-013-0314-4>.
12. Appiano M, Catalano D, Santillán Martínez M, Lotti C, Zheng Z, Visser RGF. 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. <https://doi.org/10.1186/s12870-015-0639-9>.
13. 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;98:587–94. <https://doi.org/10.4454/JPP.V98I3.027>.
14. 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;8:878–95. <https://doi.org/10.1093/gbe/evw036>.
15. Panstruga R. Discovery of novel conserved peptide domains by ortholog comparison within plant multi-protein families. *Plant Mol Biol*. 2005;59:485–500. <https://doi.org/10.1007/s11103-005-0353-0>.
16. Risapañil N, Rubiales D. Genome-wide identification and comparison of legume MLO gene family. *Sci Rep*. 2016;6:32673. <https://doi.org/10.1038/srep32673>.
17. Zheng Z, Appiano M, Pavan S, Bracuto V, Ricciardi L, Visser RGF, et al. Genome-wide study of the tomato SIMLO gene family and its functional characterization in response to the powdery mildew fungus *Oidium neolyopersici*. *Front Plant Sci*. 2016;7:380. <https://doi.org/10.3389/fpls.2016.00380>.
18. Piffanelli P, Zhou FS, Casais C, Orme J, Jarosch B, Schaffrath U, Collins NC, et al. The barley MLO modulator of defense and cell death is responsive to biotic and abiotic stress stimuli. *Plant Physiol*. 2002;129:1076–85. <https://doi.org/10.1104/pp.010954>.
19. Andolfo G, Ercolano MR. Plant innate immunity multicomponent model. *Front Plant Sci*. 2015;6:987. <https://doi.org/10.3389/fpls.2015.00987>.
20. Sanseverino W, Ercolano MR. In silico approach to predict candidate R proteins and to define their domain architecture. *BMC Res Notes*. 2012;5(1):–11. <https://doi.org/10.1186/1756-0500-5-678>.
21. Lenhard B, Sandelin A, Carninci P. Regulatory elements metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat Rev Genet*. 2012;13:233–45. <https://doi.org/10.1038/nrg3163>.
22. Bulyk M. Computational prediction of transcription-factor binding site locations. *Genome Biol*. 2003;5:201–11. <https://doi.org/10.1186/gb-2003-5-1-201>.
23. Zhu Z, Shendure J, Church GM. Discovering functional transcription factor combinations in the human cell cycle. *Genome Res*. 2005;15:848–55. <https://doi.org/10.1101/gr.3394405>.
24. Rombauts S, Florquin K, Lescot M, Marchal K, Rouzé P, Van de Peer Y. Computational approaches to identify promoters and cis-regulatory elements in plant genomes. *Plant Physiol*. 2013;132:1162–76. <https://doi.org/10.1104/pp.102.017715>.
25. Vilo J, Brazma A, Jonassen I, Robinson A, Ukkonen E. Mining for putative regulatory elements in the yeast genome using gene expression data. *Proc Int Conf Intell Syst Mol Biol*. 2000;8:384–94.
26. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res*. 1999;27:297–300. <https://doi.org/10.1093/nar/27.1.297>.
27. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37:W202–8. <https://doi.org/10.1093/nar/gkp335>.
28. Ji H, Jiang H, Ma W, Johnson DS, Myers RM, Wong WH. An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat Biotechnol*. 2008;26:1293–300. <https://doi.org/10.1038/nbt.1505>.
29. Tompa M, Li N, Bailey TL, Church GM, De Moor B, Eskin E, Favorov AV, et al. Assessing computational tools for the discovery of transcription factor binding sites. *Nat Biotechnol*. 2005;23:137–44. <https://doi.org/10.1038/nbt1053>.
30. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, et al. Systematic discovery of regulatory motifs in human promoters and UTRs by comparison of several mammals. *Nature*. 2005;434:338–45. <https://doi.org/10.1038/nature03441>.
31. Maclean D, Jerome CA, Brown AP, Gray JC. Co-regulation of nuclear genes encoding plastid ribosomal proteins by light and plastid signals during seedling development in tobacco and Arabidopsis. *Plant Mol Biol*. 2008;66:475–90. <https://doi.org/10.1007/s11103-007-9279-z>. Epub 2008 Jan 10.
32. Vandepoele K, Quimbaya M, Casneuf T, De Veylder L, Van de Peer Y. Unraveling transcriptional control in Arabidopsis using cis-regulatory elements and coexpression networks. *Plant Physiol*. 2009;150:535–46. <https://doi.org/10.1104/pp.109.136028>.
33. Humphry ME, Bednarek P, Kemmerling B, Koh S, Stein M, Göbel U. A regulon conserved in monocot and dicot plants defines a functional module in antifungal plant immunity. *Proc Natl Acad Sci U S A*. 2010;107:21896–901. <https://doi.org/10.1073/pnas.1003619107>.
34. Noguero M, Atif RM, Ochatt S, Thompson RD. The role of the DNA-binding one zinc finger (DOF) transcription factor family in plants. *Plant Sci*. 2013;209:32–45. <https://doi.org/10.1016/j.plantsci.2013.03.016>.
35. Akyildiz M, Gowik U, Engelmann S, Koczor M, Streubel M, Westhoff P. Evolution and function of a cis-regulatory module for mesophyll-specific gene expression in the C4 dicot *Flaveria trinervia*. *Plant Cell*. 2007;19:3391–402. <https://doi.org/10.1105/tpc.107.053322>.
36. Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, et al. Cis-regulatory elements for mesophyll-specific gene expression in the C4 plant *Flaveria trinervia*, the promoter of the C4 phosphoenolpyruvate carboxylase gene. *Plant Cell*. 2004;16:1077–90. <https://doi.org/10.1105/tpc.019729>.
37. Nguyen VNT, Vo KTX, Park H, Jeon JS, Jung K-H. A systematic view of the MLO family in rice suggests their novel roles in morphological development, diurnal responses, the light-signaling pathway, and various stress responses. *Front Plant Sci*. 2016;7:1413. <https://doi.org/10.3389/fpls.2016.01413>.
38. Kravatskaya GI, Chechetkin VR, Kravatsky YV, Tumanyan VG. Structural attributes of nucleotide sequences in promoter regions of supercoiling-sensitive genes: how to relate microarray expression data with genomic sequences. *Genomics*. 2013;101:1–11. <https://doi.org/10.1016/j.jygeno.2012.10.003>.
39. Bernard V, Brunaud V, Lecharny A. TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. *BMC Genomics*. 2010;11:166. <https://doi.org/10.1186/1471-2164-11-166>.
40. Lascaris RF, Groot E, Hoen PB, Mager WH, Planta RJ. Different roles for *abf1p* and a T-rich promoter element in nucleosome organization of the yeast *RPS28A* gene. *Nucleic Acids Res*. 2000;28:1390–6. <https://doi.org/10.1093/nar/28.6.1390>.
41. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W et al. Package ‘*gplots*’. 2012. <http://cran.r-project.org>.
42. McLeay RC, Bailey TL. Motif enrichment analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics*. 2010;11:165. <https://doi.org/10.1186/1471-2105-11-165>.
43. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30:3059–66. <https://doi.org/10.1093/nar/gk436>.
44. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9(1):–13. <https://doi.org/10.1186/1471-2105-9-559>.
45. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000;25:25–9. <https://doi.org/10.1038/75556>.
46. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using realtime quantitative PCR and the 2^{-ΔΔCT} method. *Methods*. 2001;25:402–8. <https://doi.org/10.1006/meth.2001.1262>.

47. Kong Q, Yuan J, Niu P, Xie J, Jiang W, Huang Y, et al. Screening suitable reference genes for normalization in reverse transcription quantitative real-time PCR analysis in melon. *PLoS One*. 2014;9:e87197. <https://doi.org/10.1371/journal.pone.0087197>.
48. Chen Z, Hartmann H, Wu M, Friedman E, Chen J, Pulley M, et al. Expression analysis of the AtMLO gene family encoding plant-specific seven-transmembrane domain proteins. *Plant Mol Biol*. 2006;60:583–97. <https://doi.org/10.1007/s11103-005-5082-x>.
49. Feechan A, Jermakow AM, Dry IB. Grapevine MLO candidates required for powdery mildew pathogenicity? *Plant Signal Behav*. 2009;4:522–3. <https://doi.org/10.4161/psb.4.6.8575>.
50. Pessina S, Pavan S, Catalano D, Gallotta A, Visser RG, Bai Y, et al. Characterization of the MLO gene family in Rosaceae and gene expression analysis in *Malus domestica*. *BMC Genomics*. 2014;15:618. <https://doi.org/10.1186/1471-2164-15-618>.
51. Schouten H, Krauskopf J, Visser R, Bai Y. Identification of candidate genes required for susceptibility to powdery or downy mildew in cucumber. *Euphytica*. 2014;200:475–86. <https://doi.org/10.1007/s10681-014-1216-z>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

