agronomy

MDPI

Article

Genome Wide Identification and Analysis of the *R2R3-MYB* Transcription Factor Gene Family in the Mangrove *Avicennia marina*

Seema Pradhan, P Sushree Shyamli, Sandhya Suranjika and Ajay Parida *

Department of Plant and Microbial Biotechnology, Institute of Life Sciences, NALCO Square, Chandrasekharpur, Bhubaneswar-23, Odisha 751023, India; seema@ils.res.in (S.P.); shyamli.p@ils.res.in (P.S.S.); sandhya@ils.res.in (S.S.)

* Correspondence: director@ils.res.in

Abstract: Drought and salinity stress have become the major factors for crop yield loss in recent years. Drastically changing climatic conditions will only add to the adverse effects of such abiotic stresses in the future. Hence, it is necessary to conduct extensive research to elucidate the molecular mechanisms that regulate plants' response to abiotic stress. Halophytes are plants that can grow in conditions of high salinity and are naturally resistant to a number of abiotic stresses. *Avicennia marina* is one such halophyte, which grows in tropical regions of the world in areas of high salinity. In this study, we have analysed the role of *R2R3-MYB* transcription factor gene family in response abiotic stress, as a number of transcription factors have been reported to have a definite role in stress manifestation. We identified 185 *R2R3 MYB* genes at genome-wide level in *A. marina* and classified them based on the presence of conserved motifs in the protein sequences. Cis-regulatory elements (CREs) present in the promoter region of these genes were analysed to identify stress responsive elements. Comparative homology with genes from other plants provided an insight into the evolutionary changes in the *A. marina R2R3 MYB* genes. In silico expression analysis revealed 34 *AmR2R3 MYB* genes that were differentially regulated in the leaves and root tissue of *A. marina* subjected to drought and salinity stress. This study is the first report of the *R2R3 MYB* gene family in the *A. marina* genome and will help in selecting candidates for further functional characterisation.

Keywords: *Avicennia*; *R2R3 MYB*; phylogeny; differential expression; abiotic stress

Citation: Pradhan, S.; Shyamli, PS.; Suranjika, S.; Parida, A. Genome Wide Identification and Analysis of the *R2R3-MYB* Transcription Factor Gene Family in the Mangrove *Avicennia marina*. *Agronomy* **2021**, *11*, 123. https://doi.org10.3390/agronomy11010123

Received: 25 November 2020
Accepted: 30 December 2020
Published: 11 January 2021

1. Introduction

The phenomenon of global warming is responsible for drastic climate changes leading to an increase in the incidences of both biotic and abiotic stress conditions all over the world. In light of this worldwide problem, sustenance and enhancement of agricultural productivity in future is a global concern since such extremes in climatic conditions adversely affect the development, growth and productivity of major agricultural systems. Plants in general, evolve different mechanisms to cope with environmental stresses such as salinity, drought and low/high temperature. Among various strategies used for the development of abiotic stress tolerant plants, the identification of candidate genes and dissecting signaling mechanisms in abiotic stress responses have been widely attempted by researchers [1]. Recent studies have analysed the expression patterns of an array of genes regulating plant growth and development in order to counteract adverse environmental stress conditions [2,3]. Comprehension and analysis of the role of transcription factors is believed to be a more practical approach than focusing on individual gene components. Transcription factors (TFs) have been reported to play an important role in plant growth, development, and stress response through self-regulation [4–7]. They also regulate the expression of downstream target genes [5,6]. These abilities make them suitable candidates for genetic manipulation of complex stress tolerance traits. However, the difficulty often

lies in the selection of a suitable gene for propagation. This problem is now alleviated by contemporary high throughput analyses which have now facilitated screening multiple gene families in silico to select the best candidate for further characterisation.

The MYB transcription factor family is one of the largest and functionally diverse TF families in plants and their role in plant development, abiotic stress response, secondary metabolism regulation, and hormone signal transduction have been well studied in Arabidopsis [4,8], *Setaria italica* [9], *Vitis vinifera* [10], *Zea mays* [11], *Populus trichocarpa* [12], *Gossypium raimondii* [13], and *Oryza sativa* [14]. Over 198 out of the 1600 transcription factors identified in *Arabidopsis thaliana* [15], belong to the MYB family [16]. The conserved MYB DNA-binding domain is essential for functioning of the MYB genes and is comprised of up to four imperfect repeats (R) of 50 to 53 amino acid sequences [17]. Depending on the number of repeats, MYB proteins have been grouped into four classes: 1R (R1/2, R3-MYB), 2R (R2R3-MYB), 3R (R1R2R3-MYB), and 4R (four R1/R2-like-MYB). In plants, R2R3-MYB proteins constitute the largest class of MYB transcription factors and have been further reassigned into 28 subgroups [18]. MYB gene in plants was first identified in *Zea mays* where a putative full set of R2R3-MYBs in the maize genome were identified comprising of a total of 157 typical R2R3-MYB encoding genes [11,19].

R2R3 MYB genes have been reported to act as regulators of abiotic stress response in economically relevant plants like cotton [20], wheat [4,21], apple [22], and rice [23], among many others. Moreover, studies have shown that ectopic expression of MYB genes resulted in enhanced tolerance to freezing, drought, and salt stress in non-halophytic plants [24,25]. In the model halophyte *Thellungiella halophile*, expressed sequence tags (ESTs) related to MYB proteins were reported [26] however, their roles in salt stress tolerance in halophytic plants have not been studied so far. The function of several other transcription regulators, such as the AP2 domain transcription factor and the homeobox-leucine rich protein has been extensively studied, with reference to stress tolerance in plants like *Arabidopsis thaliana* and *Oryza sativa* [27,28], whereas little is known about the function of MYB family transcription factors in the abiotic stress tolerance of halophytes [29].

Halophytes, owing to their capability to thrive under extremely saline conditions, are considered as one of the best source materials for identification of salinity tolerant genes. *Avicennia marina*, a salt-tolerant [30] mangrove species is one such halophyte and has been selected for the present study. Despite having information about a few candidate genes and their role in abiotic stress tolerance [31], information on molecular defense mechanism of halophytes against salt stress is scanty [32,33]. Therefore, the present study emphasizes the importance in assessing the diversity and potential role of MYB gene family in *A. marina*, a small evergreen tree that grows more than 10 m in height and is the most widely distributed mangrove species in the Indo-Western Pacific area belonging to Acanthaceae (Avicenniaceae) family. The plant is commonly known as the gray or white mangrove and is highly resistant to environmental stresses and can flourish under adverse environmental conditions of extreme tides, high salinity, high temperature, strong winds and anaerobic soil [34].

In this study, we have systematically investigated the *R2R3 MYB* gene family in *Avicennia marina* at the whole genome level, their classification into different groups, chromosomal distribution, presence of conserved motifs, phylogenetic relationship, and sequence homology with members of *Arabidopsis thaliana* and *Oryza sativa*. The study will provide us with an insight into the role of *R2R3 MYB* genes of *A. marina* in response to abiotic stresses like drought and salinity.

2. Materials and Methods

2.1. Identification of R2R3-MYB Transcription Factors in A. marina Genome

The CDS and protein sequences for the annotated proteins of *A. marina* were downloaded from DRYAD [35]. The MYB consensus sequence [36] and the HMM profile for MYB DNA-binding domain from Pfam (https://pfam.xfam.org/family/PF00249) were used to identify putative MYB transcription factors in the peptide sequences of anno-

tated CDS of *A. marina*. Thereafter, sequence identifiers provided by [14] were used to download the peptide sequences for R2R3-MYB transcription factors of rice and *Arabidopsis* from Rice Genome Annotation project (http://rice.plantbiology.msu.edu/) and TAIR (https://www.arabidopsis.org/) respectively. These sequences were aligned and the alignment file was used to build the HMM profile of *R2R3-MYB* domain using HMMER v3.3 (http://hmmer.org/) and this profile was used to identify the *R2R3-MYB* domain containing protein sequences amongst the MYB TFs in *A. marina* genome. Chemical properties of the protein sequences were evaluated with the ProtParam tool of ExPASy (https://web.expasy.org/protparam/; [37]) and subcellular localisation was predicted with BUSCA (http://busca.biocomp.unibo.it/; [38])

2.2. Chromosomal Location and Nomenclature

The standalone version of BLAST software (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/) was used to map the corresponding CDS sequences of the MYB TFs of *A. marina* onto the genome using Blastn. The output file was used to locate the positions of all predicted MYB genes on the various scaffolds of *A. marina* genome. The MYB coding genes were named sequentially based on their position on the chromosomes/scaffolds, starting from AmMYB1. Only the *R2R3-MYB* TFs were taken for further analysis and were localised onto the chromosomes using MapChart [39].

2.3. Phylogenetic Analysis

The protein sequences of the R2R3-MYB TFs were used to generate the phylogenetic tree on MEGA X software ([40]; https://www.megasoftware.net/). The sequences were aligned with ClustalW and the tree was constructed with the neighbour-joining method. Bootstrap value was set to 100 and branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The evolutionary distances were computed using the Jones-Taylor-Thornton (JTT) matrix-based method.

2.4. Exon-Intron Structure and Conserved Motif Analysis

The CDS sequences of *R2R3 MYB* TFs were mapped onto the genome of *A. marina* using Blastn and the start and end positions of the mapped gene positions were used to prepare the bed file for extracting the fasta sequences of the genomic regions. The fasta formatted CDS sequences and genomic sequences were uploaded to Gene Structure Display Server (http://gsds.gao-lab.org/) to generate exon-intron structure. Conserved motifs in the protein sequences of R2R3 MYB TFs were identified with MEME tool of MEME-suite (http://meme-suite.org/tools/meme) with the parameters set as Zero to one occurrences per sequence; motifs to be found = 15 and motif width = 6–200.

2.5. Duplication, dN/dS and Homology/Synteny Analysis

Protein sequences of R2R3 MYB TFs in *A. marina* were aligned with clustalW and duplicate genes were identified using MCScanX [41]. The ratios of non-synonymous (dN) and synonymous (dS) substitutions were calculated using PAL2NAL v14 (http://www.bork.embl.de/pal2nal/#Download) and CODEML program of PAML package (http://abacus.gene.ucl.ac.uk/software/paml.html). The *R2R3 AmMYB* genes were compared with the CDS sequences of *R2R3 MYB* genes of *Arabidopsis thaliana* and *Oryza sativa* and their genomic positions were recorded. The Circos tool (circos-0.69-9; http://circos.ca/) was used to generate the figure using the genomic positions and chromosome lengths.

2.6. Conserved Motif Identification in Promoter Region

A stretch of 2000 bp upstream of the transcription start site of *A. marina R2R3 MYB* TFs were extracted from the genome using BEDtools [42]. The motifs in promoter sequences were identified with PlantCARE [43]; http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

2.7. In Silico Expression Analysis

The short reads of *Avicennia marina* control leaf (Accession SRR2029733), control root (Accession SRR2029734), salt stressed leaf (Accession SRR2029735), salt stressed root (Accession SRR2029736), drought stressed leaf (Accession SRR2029738) and drought stressed root (Accession SRR2029739) were obtained from the SRA database of NCBI (https://www.ncbi.nlm.nih.gov/sra). These were aligned onto the CDS sequences of *A. marina R2R3 MYB* TFs and their normalised expression values were calculated using the tools provided in Trinity package v2.11.0 (https://github.com/trinityrnaseq/trinityrnaseq/releases; [44]). Detailed instructions for differential gene expression analysis can be found at: https://github.com/trinityrnaseq/trinityrnaseq/wiki/Trinity-Differential-Expression [45]. The TMM method was used for preparing expression matrix and the FPKM values were normalised by calculating the log2 values. Z score values were calculated for the $\log_2$ normalized FPKM according to the formulae described by [46] values to further remove bias andthese expression values were used to generate the heat map on MeV v4.8.1 (http://mev.tm4.org/; [47]).

3. Results

3.1. R2R3-MYB Transcription Factor Family in Avicennia marina

A total of 284 MYB TF proteins were identified in the genome of *A. marina* based on the presence of the MYB consensus sequence as well as the HMM profile of MYB-DNA binding domain. They were named AmMYB1-AmMYB284 based on their positions on *A. marina* chromosome/scaffold. Of these, 185 members were identified as R2R3 MYB based on their similarity to the R2R3 MYB proteins of rice and Arabidopsis (File S1). Analysis of the chemical properties of *A. marina* R2R3 MYB proteins revealed that their length ranged from 100 bp to 1748 bp, molecular weight ranged from 11.46 kDa to 191.24 kDa, and their isoelectric point ranged from 4.65 to 10.28 (Table 1). A majority of them were predicted to be localised to nucleus (89%) while a few were localised to chloroplast (7%). Four of the AmR2R3 MYB proteins (AmMYB41, AmMYB79, AmMYB130 and AmMYB133) contained a signal peptide implying that these are secretory proteins that localise to extracellular space while one member (AmMYB251) was localised to the mitochondrion (Table 1).

Table 1. The length distribution and chemical properties of Am R2R3 MYB proteins.

Accession	Protein Length (aa)	Protein Molecular Weight (in kDa)	Isoelectric Point (pI)	Subcellular Localisation
AmMYB1	307	34.68	6.20	nucleus
AmMYB2	349	38.13	5.21	nucleus
AmMYB4	146	16.72	9.68	nucleus
AmMYB6	304	34.17	8.90	nucleus
AmMYB7	156	17.61	10.28	chloroplast
AmMYB9	335	37.07	5.93	nucleus
AmMYB10	307	33.42	7.78	nucleus
AmMYB11	373	42.58	6.08	nucleus
AmMYB12	287	33.23	9.58	nucleus
AmMYB13	533	59.68	7.63	endomembrane system
AmMYB14	100	11.46	9.83	nucleus
AmMYB16	295	31.79	8.65	nucleus
AmMYB17	441	48.95	8.55	nucleus
AmMYB19	386	41.64	6.17	nucleus

Table 1. *Cont.*

Accession	Protein Length (aa)	Protein Molecular Weight (in kDa)	Isoelectric Point (pI)	Subcellular Localisation
AmMYB20	216	24.35	9.38	chloroplast
AmMYB21	279	31.89	5.71	nucleus
AmMYB22	173	20.09	5.85	chloroplast
AmMYB24	214	23.67	9.54	nucleus
AmMYB25	470	51.73	5.38	nucleus
AmMYB26	275	31.23	7.69	nucleus
AmMYB29	372	42.06	8.42	chloroplast
AmMYB30	273	30.71	8.97	nucleus
AmMYB32	353	38.86	7.55	nucleus
AmMYB33	321	35.03	8.98	nucleus
AmMYB34	235	25.83	9.30	nucleus
AmMYB36	274	30.40	9.11	nucleus
AmMYB37	275	31.30	5.64	nucleus
AmMYB38	273	30.48	9.06	nucleus
AmMYB39	301	33.26	8.78	nucleus
AmMYB40	295	32.96	8.62	nucleus
AmMYB41	420	46.47	6.13	extracellular space
AmMYB42	330	37.25	7.88	nucleus
AmMYB45	1159	128.76	6.09	nucleus
AmMYB46	307	34.10	9.86	chloroplast
AmMYB47	189	21.43	9.52	nucleus
AmMYB49	331	36.60	6.76	nucleus
AmMYB50	291	31.09	8.55	nucleus
AmMYB51	379	43.13	5.65	nucleus
AmMYB53	195	22.89	9.31	nucleus
AmMYB54	250	28.53	7.67	nucleus
AmMYB55	236	27.15	6.19	nucleus
AmMYB57	311	35.24	5.92	nucleus
AmMYB58	529	59.04	7.61	nucleus
AmMYB60	294	32.99	6.16	nucleus
AmMYB61	411	45.86	6.19	nucleus
AmMYB62	273	29.06	8.15	nucleus
AmMYB63	261	28.93	5.22	nucleus
AmMYB65	358	40.16	9.06	nucleus
AmMYB69	280	30.61	9.05	nucleus
AmMYB71	327	36.19	6.20	nucleus
AmMYB72	304	32.81	6.45	nucleus
AmMYB73	286	32.95	9.27	nucleus
AmMYB74	465	51.35	6.38	nucleus
AmMYB76	959	106.61	5.09	nucleus

Table 1. *Cont.*

Accession	Protein Length (aa)	Protein Molecular Weight (in kDa)	Isoelectric Point (pI)	Subcellular Localisation
AmMYB77	597	63.17	6.22	nucleus
AmMYB79	283	32.45	5.22	extracellular space
AmMYB80	205	22.99	6.10	nucleus
AmMYB82	320	35.69	6.27	nucleus
AmMYB83	314	34.90	9.21	nucleus
AmMYB86	568	61.70	4.88	nucleus
AmMYB87	183	20.58	4.94	nucleus
AmMYB88	308	34.32	8.86	nucleus
AmMYB91	180	20.46	6.01	nucleus
AmMYB98	295	33.90	7.06	nucleus
AmMYB100	295	31.74	6.80	nucleus
AmMYB101	290	32.41	6.72	nucleus
AmMYB102	333	36.08	7.23	chloroplast
AmMYB105	1046	115.14	5.10	nucleus
AmMYB106	575	65.27	5.73	nucleus
AmMYB107	181	20.94	6.36	nucleus
AmMYB108	353	39.25	6.03	nucleus
AmMYB112	367	41.38	5.93	nucleus
AmMYB116	275	31.23	9.46	nucleus
AmMYB117	385	43.53	9.31	nucleus
AmMYB118	320	34.81	9.34	nucleus
AmMYB119	353	39.43	8.04	nucleus
AmMYB120	202	22.88	6.65	nucleus
AmMYB121	482	52.81	7.55	nucleus
AmMYB122	334	37.37	5.62	nucleus
AmMYB127	328	36.67	6.70	nucleus
AmMYB128	431	47.87	6.95	nucleus
AmMYB129	299	32.84	9.75	nucleus
AmMYB130	121	13.72	7.19	extracellular space
AmMYB131	252	28.48	8.84	nucleus
AmMYB132	329	36.00	5.09	nucleus
AmMYB133	321	36.29	8.98	extracellular space
AmMYB134	386	43.40	5.73	nucleus
AmMYB135	232	25.94	6.24	nucleus
AmMYB137	259	29.05	4.65	chloroplast
AmMYB138	551	59.97	5.11	nucleus
AmMYB139	321	36.13	6.21	nucleus
AmMYB140	271	31.28	6.45	nucleus
AmMYB143	251	28.41	6.13	nucleus
AmMYB146	228	25.90	9.73	nucleus
AmMYB147	343	38.22	7.92	nucleus

Table 1. *Cont.*

Accession	Protein Length (aa)	Protein Molecular Weight (in kDa)	Isoelectric Point (pI)	Subcellular Localisation
AmMYB151	307	34.58	7.90	nucleus
AmMYB152	403	43.71	5.25	nucleus
AmMYB153	345	38.36	9.33	chloroplast
AmMYB155	376	41.99	8.99	nucleus
AmMYB157	424	47.74	6.14	nucleus
AmMYB158	261	28.50	9.69	nucleus
AmMYB159	352	39.46	6.45	nucleus
AmMYB161	258	29.06	8.94	nucleus
AmMYB163	302	33.13	9.38	nucleus
AmMYB164	425	46.93	6.89	nucleus
AmMYB166	539	58.42	4.87	nucleus
AmMYB167	338	38.15	7.42	nucleus
AmMYB168	242	28.69	9.55	nucleus
AmMYB172	240	27.00	6.96	nucleus
AmMYB174	405	45.33	8.93	nucleus
AmMYB175	359	40.00	6.01	nucleus
AmMYB177	308	34.25	8.47	nucleus
AmMYB178	331	36.56	8.19	nucleus
AmMYB181	282	30.99	9.48	nucleus
AmMYB182	535	59.41	8.17	nucleus
AmMYB183	361	41.32	9.15	nucleus
AmMYB184	303	34.58	7.27	nucleus
AmMYB185	253	28.50	9.15	nucleus
AmMYB187	416	46.08	6.64	nucleus
AmMYB189	323	35.53	8.94	nucleus
AmMYB190	320	35.91	5.46	nucleus
AmMYB191	266	30.00	9.12	nucleus
AmMYB192	463	50.47	6.40	nucleus
AmMYB194	994	110.87	5.20	nucleus
AmMYB196	1748	191.24	6.20	nucleus
AmMYB198	338	36.80	9.25	nucleus
AmMYB199	482	54.22	5.75	nucleus
AmMYB201	398	43.35	6.03	nucleus
AmMYB202	386	42.74	7.23	nucleus
AmMYB204	314	34.54	7.72	nucleus
AmMYB205	135	15.62	10.20	nucleus
AmMYB206	228	25.45	9.08	nucleus
AmMYB208	317	36.35	9.36	nucleus
AmMYB209	288	31.91	8.84	nucleus
AmMYB211	343	38.51	6.28	nucleus

Table 1. *Cont.*

Accession	Protein Length (aa)	Protein Molecular Weight (in kDa)	Isoelectric Point (pI)	Subcellular Localisation
AmMYB212	269	30.88	5.79	nucleus
AmMYB213	284	31.65	9.05	nucleus
AmMYB214	294	31.42	8.96	nucleus
AmMYB216	287	32.41	5.93	nucleus
AmMYB218	483	54.26	6.23	nucleus
AmMYB220	348	38.90	8.80	endomembrane system
AmMYB221	336	37.14	8.12	nucleus
AmMYB222	315	35.43	4.98	nucleus
AmMYB225	113	12.98	10.13	nucleus
AmMYB226	108	12.31	10.05	nucleus
AmMYB227	211	24.32	7.01	nucleus
AmMYB228	334	37.68	5.34	nucleus
AmMYB229	385	41.70	6.28	nucleus
AmMYB230	265	29.92	9.43	chloroplast
AmMYB231	251	28.28	5.48	nucleus
AmMYB232	325	35.58	7.76	nucleus
AmMYB233	351	39.47	6.55	nucleus
AmMYB234	269	31.22	9.30	chloroplast
AmMYB235	148	16.54	10.15	nucleus
AmMYB236	478	52.71	8.68	nucleus
AmMYB237	302	34.65	6.17	nucleus
AmMYB240	244	27.54	5.62	nucleus
AmMYB241	258	29.42	6.08	nucleus
AmMYB242	354	40.16	7.17	nucleus
AmMYB244	303	32.14	9.21	nucleus
AmMYB245	612	67.24	7.47	nucleus
AmMYB247	321	35.66	9.07	nucleus
AmMYB248	342	38.21	5.63	nucleus
AmMYB250	417	46.73	6.56	nucleus
AmMYB251	313	34.57	8.99	mitochondrion
AmMYB253	252	28.07	9.09	nucleus
AmMYB255	308	34.72	8.42	nucleus
AmMYB256	399	43.21	5.59	nucleus
AmMYB257	312	34.83	7.53	chloroplast
AmMYB259	415	46.14	8.70	nucleus
AmMYB260	217	23.93	9.15	nucleus
AmMYB262	334	37.34	8.15	nucleus
AmMYB263	383	42.63	8.94	nucleus
AmMYB265	351	39.18	5.54	nucleus
AmMYB266	332	38.23	8.40	nucleus

Table 1. *Cont.*

Accession	Protein Length (aa)	Protein Molecular Weight (in kDa)	Isoelectric Point (pI)	Subcellular Localisation
AmMYB267	251	28.39	7.31	nucleus
AmMYB268	519	56.12	7.25	chloroplast
AmMYB271	377	42.75	5.76	nucleus
AmMYB274	332	36.77	6.87	nucleus
AmMYB275	330	37.51	6.73	chloroplast
AmMYB277	370	40.70	7.61	nucleus
AmMYB279	472	52.65	7.11	nucleus
AmMYB280	325	35.50	9.14	nucleus
AmMYB283	219	24.82	7.87	nucleus
AmMYB284	1743	191.10	5.83	nucleus

3.2. Phylogenetic Analysis and Classification of R2R3 MYB Genes

The phylogenetic tree was prepared based on the alignment of the *R2R3 MYB* proteins of *A. marina* amongst themselves. The tree was corelated with the gene structure (intron-exon arrangement) and the motif composition of these proteins. Based on the type of motifs present in the protein sequences, the R2R3 AmMYB were classified into 8 groups (I–VIII) and sub-groups (Figure 1a). Majority of the members belonged to group III and contained motifs 1, 2, 3, and 5 (motif information is provided in Figure S1) while some of the members also contained the motif 7. The similarities were also reflected in the exon-intron structure of the genes. For example, the members of group I were either intronless or consisted of only one intron. Similar results were found for most of the groups with the exception of few members which deviated from the group morphology. Interestingly, AmMYB63 and AmMYB137 did not contain motifs similar to the other AmR2R3 MYB but both had the R2R3 MYB domain and a number of distinct motifs (Figure S2).

3.3. Chromosomal Location and Gene Duplication

The recent *A. marina* genome assembly reports a chromosome level assembly with 32 super scaffolds representing the chromosomes [35]. The *R2R3 MYB* genes of *A. marina* were found to be located exclusively in these 32 chromosomes (Supplementary Table S1, Figure 2). Chromosomes 8 and 11 of *A. marina* contained the loci for highest number of *R2R3 MYB* genes (10 each) while chromosome 7 contained only one. Most of the chromosomes contained 3-7 *R2R3 MYB* genes implying that the distribution of *AmR2R3 MYB* genes is fairly uniform in all chromosomes. Analysis of duplication events revealed seven tandem duplication events and six collinear gene pairs that have resulted in the course of evolution of *AmR2R3 MYB* family (Figure 2). These gene pairs were analysed for the ratio of rates of non-synonymous (dN) to synonymous (dS) substitutions (Table 2). Generally, a value of <1 suggests purifying selection resulting in functional constraint, a value of >1 implies positive selection resulting in accelerated evolution and a value = 1 implies neutral selection. It was observed that the value of dN/dS for all the gene pairs was < 1 with the value for AmMYB225/AmMYB226 almost nearing 1 (Table 2).

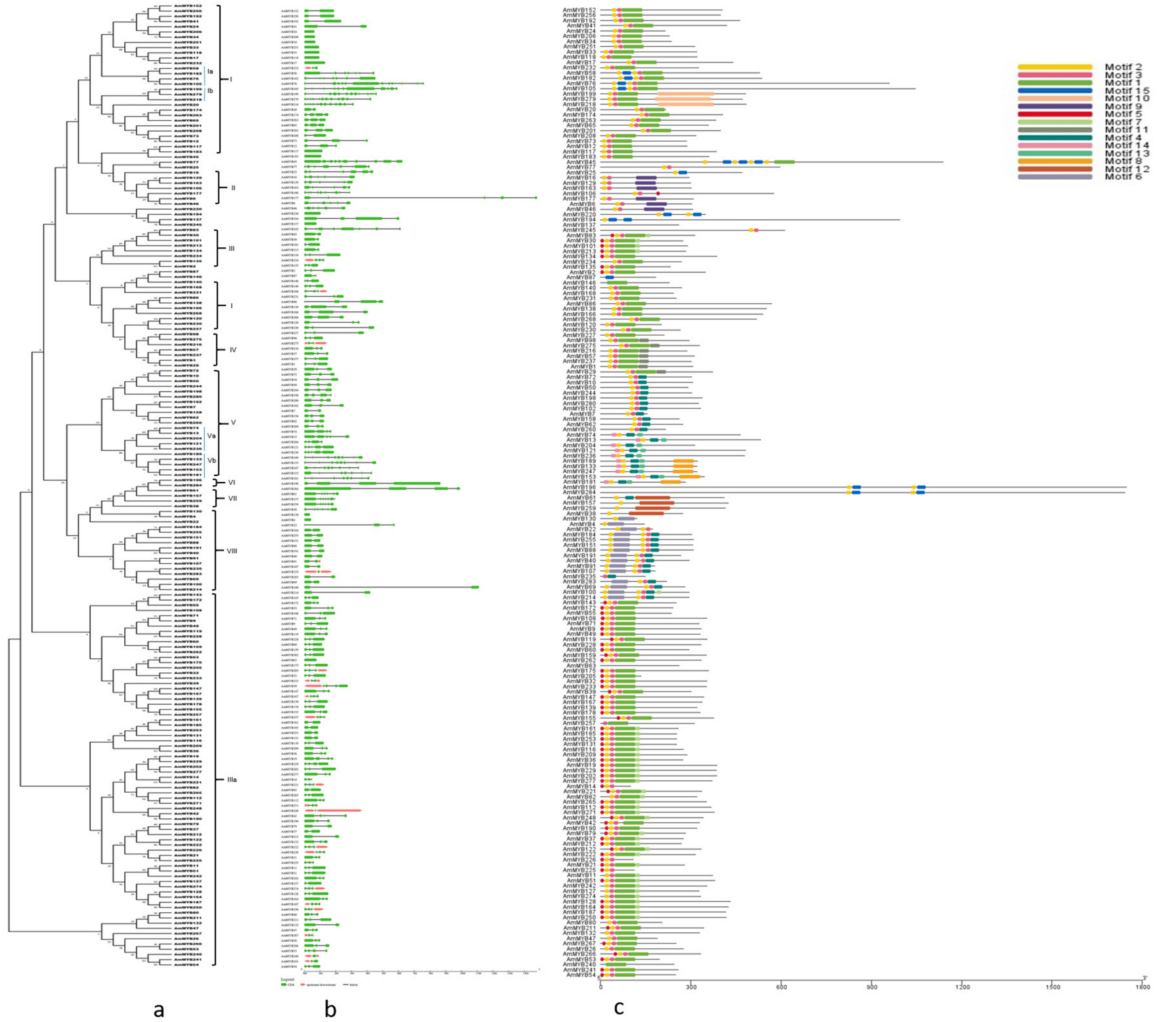

Figure 1. Phylogenetic classification and motif analysis of the *AmR2R3 MYB* genes. (**a**) The phylogenetic tree was constructed using the MEGA X software with Neighbour-Joining algorithm with 100 bootstraps. The exon-intron structures of genes (**b**) and conserved motifs in protein sequences (**c**) are depicted alongside the relative phylogenetic positions of *AmR2R3 MYB* genes.

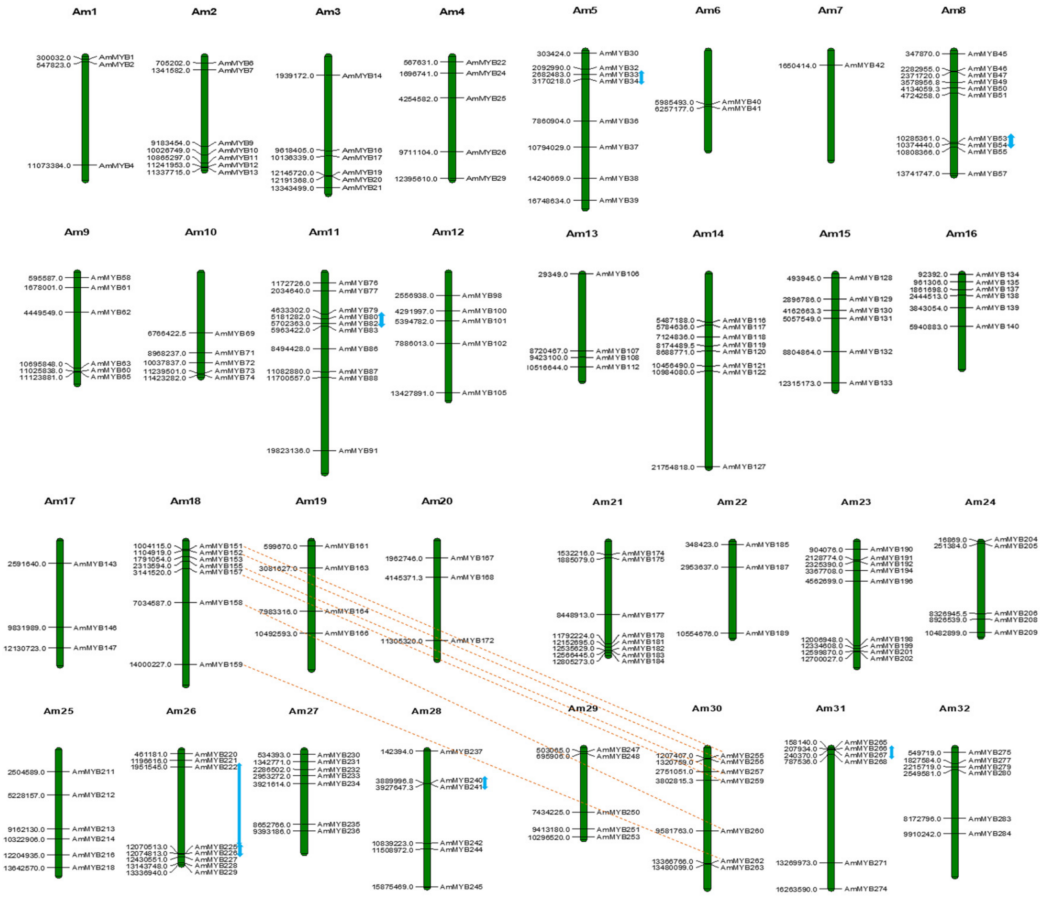

Figure 2. Chromosomal location and duplication events of *AmR2R3 MYB*. The *AmR2R3 MYB* genes were distributed on 22 pseudo-chromosomes of A. marina based on their physical positions. The tandemly duplicated genes are indicated with blue arrows. The red dashed lines between pseudo-chromosomes 18 and 30 indicate the collinear genes.

Table 2. The ratio of non-synonymous to synonymous substitution rates (dN/dS) in the *AmR2R3 MYB* genes.

Tandem Duplication Events				
Gene Pair	dN	dS	dN/dS	Type of selection
AmMYB222/AmMYB225	0.3481	2.708	0.1285	Purifying
AmMYB225/AmMYB226	0.1159	0.1196	0.9684	Purifying/Neutral
AmMYB240/AmMYB241	0.3453	0.9179	0.3761	Purifying
AmMYB266/AmMYB267	0.5621	1.9159	0.2934	Purifying
AmMYB33/AmMYB34	0.4598	7.248	0.0634	Purifying
AmMYB53/AmMYB54	0.2837	0.6626	0.4281	Purifying
AmMYB80/AmMYB82	0.6197	23.7159	0.0261	Purifying
Collinear duplications				
AmMYB151/AmMYB255	0.1009	0.6072	0.1662	Purifying
AmMYB152/AmMYB256	0.1861	0.441	0.4221	Purifying
AmMYB155/AmMYB257	0.2913	0.7077	0.4116	Purifying
AmMYB157/AmMYB259	0.162	0.5626	0.2879	Purifying
AmMYB158/AmMYB260	0.2732	0.8214	0.3326	Purifying
AmMYB159/AmMYB262	0.1536	0.5262	0.292	Purifying

3.4. Promoter Motif Analysis

The promoter regions of genes contain conserved motifs which act as recognition and binding sites for various proteins. These interactions play an integral role in the regulation of gene expression and thereby affect the important biological processes in an organism. Therefore, conserved motifs in the promoter region of *AmR2R3 MYB* genes were analysed to evaluate their role in abiotic stress tolerance. The most frequently found motifs were core elements necessary for transcription such as CAAT-box (6126) and TATA-box (7599). A large number of cis-regulatory elements (CREs) related to stresss tolerance were identified (Figure 3, Table 3). Most abundant of these were the water and drought response elements MYB (872) and ABRE elements (668) followed by MYC (593), which is a drought response element, ARE (327) i.e., anaerobic-responsive elements and STRE (308) which is stress-responsive element. A number of other stress response CREs, such as ARE and ERE elements (oxidative stress responsive elements), W box (binding site for WRKY TFs), LTR (low temperature responsive elements), and DRE core (cold and dehydration responsive element), were also identified in the promoter regions of *AmR2R3 MYB* genes (Table 3).

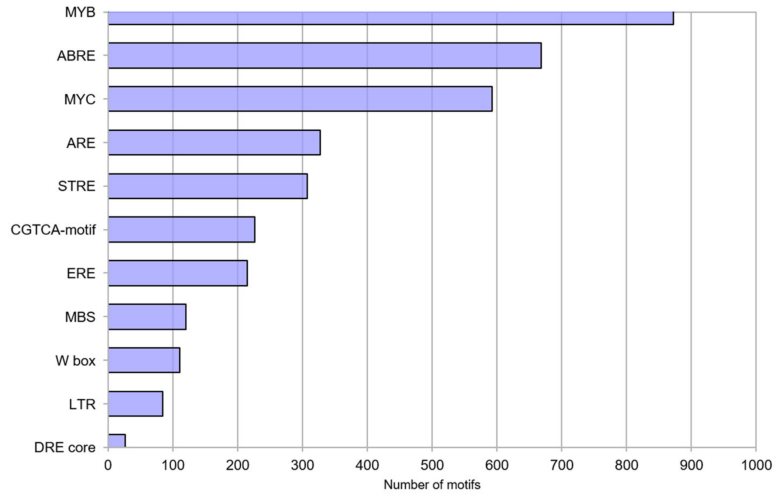

Figure 3. Cis-regulatory elements in the promoter region of *AmR2R3 MYB* genes. The figure represents the number of each type of stress responsive motifs identified in the promoter sequence of *AmR2R3 MYB*.

Table 3. Cis-regulatory motifs in promoter regions of *AmR2R3 MYB* genes.

Name of the Motif	Motif Sequence	Function
ABRE	TACGTG; ACGTG; CACGTA; CACGTG; CGCACGTGTC; CGTACGTGCA; AACCCGG	cis-acting element involved in the abscisic acid responsiveness
ARE	AAACCA	cis-acting regulatory element essential for the anaerobic induction
CGTCA-motif	CGTCA	cis-acting regulatory element involved in the MeJA-responsiveness
DRE core	GCCGAC	cis-acting regulatory element involved in cold and dehydration response
ERE	ATTTTAAA	cis-acting regulatory element involved in oxidative stress response
LTR	CCGAAA	cis-acting element involved in low-temperature responsiveness
MBS	CAACTG	MYB binding site involved in drought-inducibility
MYB	TAACCA; CAACCA; CAACAG; CAACTG; TAACTG; TAACCA	cis-acting element involved in drought responsiveness
MYC	CATGTG; CATTTG; TCTCTTA; CAATTG	cis-acting element involved in drought responsiveness
STRE	AGGGG	cis-regulatory element able to mediate transcriptional induction by different forms of stress
W box	TTGACC	cis-regulatory element that acts as a binding site for WRKY TFs

3.5. Homology with R2R3 MYB Genes from Rice and Arabidopsis

Many transcription factor domains are known to be conserved across plant genera [48]. Probable reason for the high level of conservation could be the essential function of these genes in growth and development of plants. Thus, analysing the homology between different plants can provide an idea about the evolution of the gene family. Therefore, the AmR2R3 MYB proteins were compared with those from Arabidopsis and rice. It was interesting to note that none of the *A. marina* R2R3 MYB proteins grouped with the R2R3 MYB from rice and Arabidopsis (Figure 4a). The Am R2R3 MYB formed a distinctly separate clade while a number of members from Arabidopsis (with locus Ids starting with AT) and rice (with locus Ids starting with LOC_Os) showed similarity to each other. Thereafter, the *AmR2R3 MYB* genes (CDS sequences) were compared to the *R2R3 MYB* genes of rice (*OsR2R3 MYB*) and Arabidopsis (*AtR2R3 MYB*). We observed that 28 out of the 185 *AmR2R3 MYB* genes found homologous counterparts in *AtR2R3 MYB* genes and 27 found homologous counterparts in *OsR2R3 MYB* genes (Supplementary Table S2). Many of the *AmR2R3 MYB* genes had multiple homologs in Arabidopsis and rice (Figure 4b).

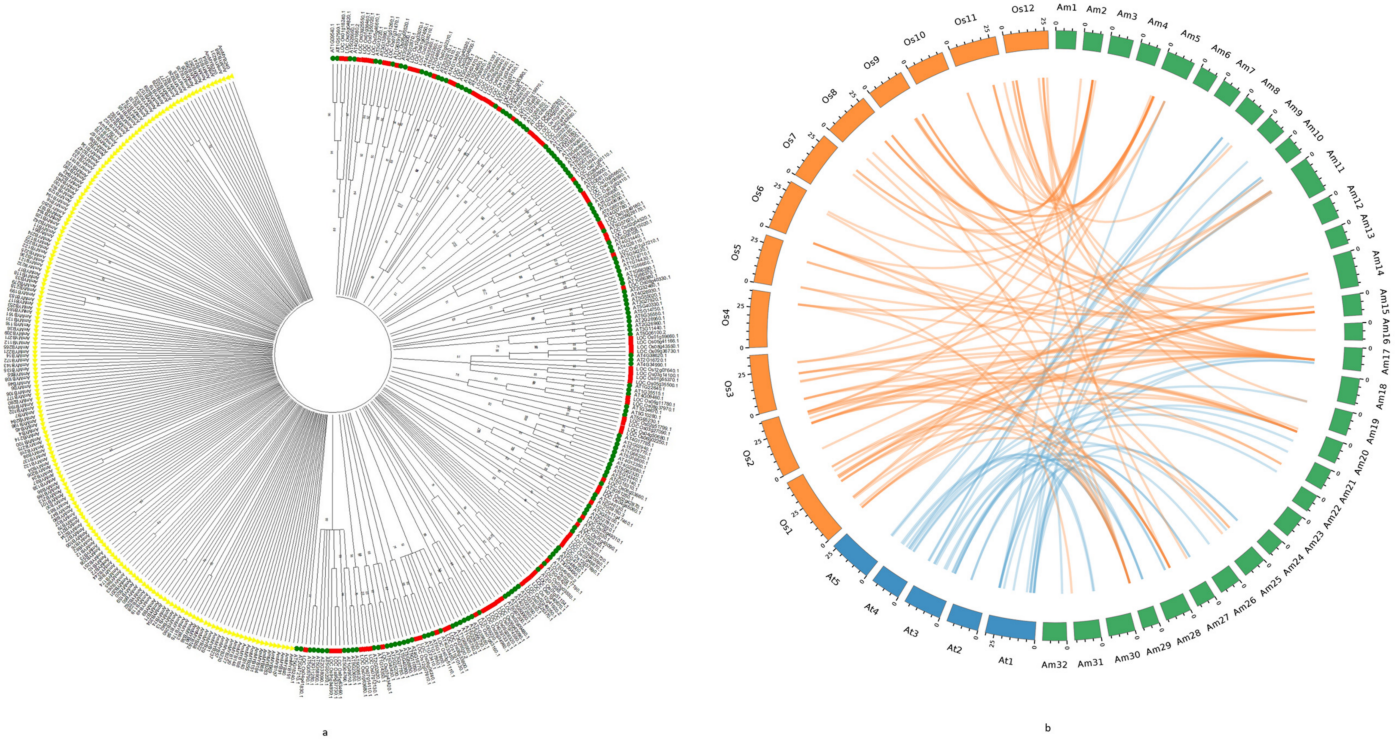

Figure 4. Phylogenetic relationship with R2R3 MYB from *Arabidopsis* and *Oryza*. (**a**) The AmR2R3 MYB proteins were compared to those from Arabidopsis (Ids starting with AT) and rice (Ids starting with Os_LOC) (**b**) Comparison of *AmR2R3 MYB* CDS sequences (Am1-32) homologous with the *R2R3 MYB* from Arabidopsis (At1-5) and rice (Os1-12) based on their chromosomal positions.

3.6. In Silico Gene Expression Analysis

In silico expression analysis was carried out using the RNA-Seq data for control and treated *A. marina* samples exposed to drought and salinity stress available at NCBI. The raw FPKM values were calculated using publicly available perl scripts and normalised by calculating their $\log_2$ values and z scores (Supplementary Table S3). The hierarchical clustering of values based on the pattern of expression clustered the *AmR2R3 MYB* genes into distinct blocks (Figure 5a). It was observed that drought stress had a more noticeable effect on the expression of *AmR2R3 MYB* genes as compared with salt stress. The differential expression was also more pronounced in the leaf tissue compared with root. *AmMYB275*, *61*, *65*, *237*, *135*, *190*, *216*, *1*, *29*, *209*, *257*, *112*, *122*, and *187* were upregulated in leaf and root

tissues in response to drought stress and AmMYB267 was especially upregulated in root tissue (Figure 5b). On the other hand, *AmMYB117*, *158*, *168*, *183*, *26*, *80*, and *140* were down regulated in leaf and root tissue in response to drought stress. *AmMYB146*, *AmMYB164*, and *AmMYB51* were down regulated in leaf tissue while having no significant differential expression in root (Figure 5c). Salt stress had little effect on the expression of *AmR2R3 MYB* genes. Two of the members showed a noticeable change in expression in response to salt stress. *AmMYB87* was upregulated in response to salt stress and downregulated in response to drought in both leaf and root tissue while the opposite was observed for *AmMYB39* (Figure 5d).

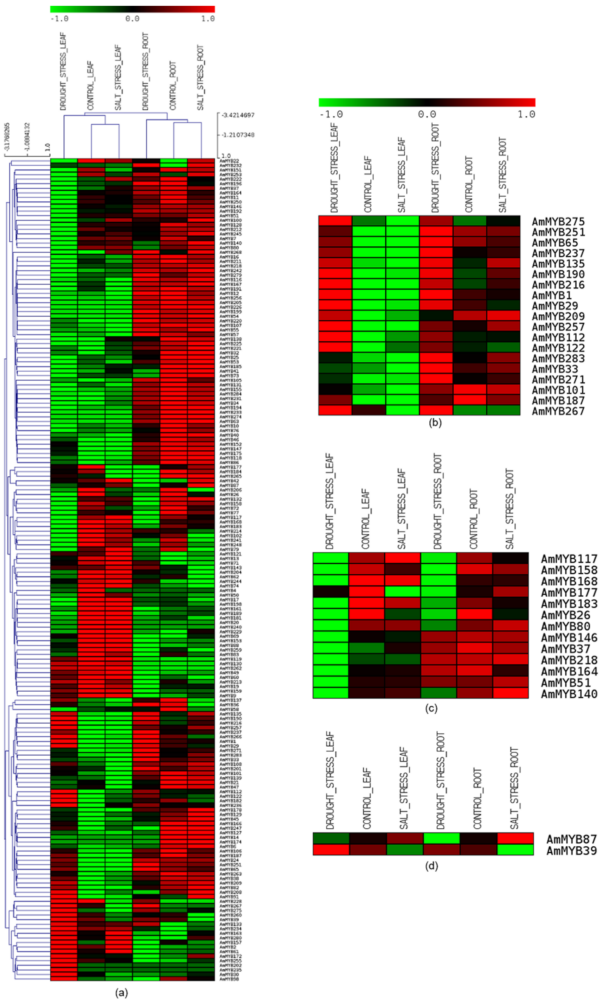

Figure 5. In silico differential expression of *AmR2R3 MYB* genes in response to abiotic stress. (**a**) Hierarchical clustering of the *AmR2R3 MYB* genes based on their expression in leaf and root tissue under salinity and drought stress. (**b**) *AmR2R3 MYB* genes upregulated in *A. marina* tissues in response to drought stress (**c**) *AmR2R3 MYB* genes downregulated in *A. marina* tissues in response to drought stress and (**d**) *AmR2R3 MYB* genes differentially regulated in response to salt stress in *A. marina*.

4. Discussion

Plants response to environmental stress is one of the most extensively studied research area and assumes significance in developing strategies for addressing the growing concern of declining plant productivity in response to adverse climatic impacts. Abiotic stress conditions like drought and salinity are responsible for heavy yield losses in plants [49]. Therefore, understanding the molecular basis of abiotic stress response in plants has been a focal research area. Halophytes such as *Avicennia marina* typically grow in conditions of

high salinity in various regions of the world [50]. Such plants can be a valuable resource for identifying genes which regulate their enhanced response to such stresses. The *R2R3 MYB* gene family is one of the most widespread family of transcription factors and transcription regulators with diverse functions, especially in abiotic stress response in plants [51]. In this study, we have undertaken the identification and analysis of the *R2R3 MYB* gene family in *Avicennia marina* which has led to a number of interesting observations.

4.1. The Evolution of A. marina R2R3 MYB Gene Family

There are numerous parameters to assess the evolutionary history of gene families and phylogenetic relationships form the foundation of such analyses. The AmR2R3 MYB were grouped into eight groups and sub-groups based on the presence of similar motifs in their peptide sequences. A key factor that contributes to the evolution of a gene family is the duplication events. A number of tandem duplications as well as collinear *AmR2R3 MYB* gene pairs were identified and classified into negatively, positively and neutrally selected genes based on their dN/dS ratios. The *AmR2R3 MYB* genes were under negative or purifying selection. This type of expansion of a gene family ensures functional conservation of the genes in the course of evolution and is commonly observed in transcription factor families [52,53].

On one hand, the expansion of *R2R3 MYB* gene family in *A. marina* was found to be under negative selection leading to the conservation of gene function. However, contrasting results were obtained on analysing the phylogenetic association of the members between *A. marina*, *Arabidopsis*, and *O. sativa*. There was some intermixing observed for R2R3 MYB proteins from Arabidopsis and rice, both of which are glycophytes, while the R2R3 MYB proteins of the halophyte *A. marina* formed a clade that was distinctly separate from the counterparts in both Arabidopsis and rice. This outcome is fortified by the fact that *Avicennia* is a monotypic genus which implies that it would have a markedly different genetic composition to model plants like rice and Arabidopsis [50]. This supports the idea that the R2R3 MYB family in *A. marina* has evolved to help the plant to combat abiotic stress and is therefore crucial to its survival and growth in conditions of high salinity. However, a one on one BLASTn based comparison between the CDS sequences of *R2R3 MYB* from *A. marina* with those from *Arabidopsis thaliana* and *Oryza sativa* revealed 28 and 27 members of *AmR2R3 MYB* family to have similarities with *AtR2R3 MYB* and *OsR2R3 MYB* respectively. This suggests that although there may be similarities between the *R2R3 MYB* genes of *A. marina* with the other two plants at the nucleotide level, the results may not be reflected in phylogenetic tree generated using the corresponding proteins. This may be due to the fact that the algorithms for phylogenetic analysis consider only the best alignment scenarios when comparing gene sequences from multiple genera simultaneously. Therefore, AmR2R3 MYB formed a separate clade, as the protein sequences for R2R3 MYB from *A. thaliana* and *O. sativa* were more similar to each other than to *A. marina*.

4.2. AmR2R3 MYB in the Regulation of Gene Expression During Drought and Salinity Stress

The conserved DNA motifs in promoter regions of genes allow various regulators to bind and control gene expression [54]. Analysis of these motifs in *AmR2R3 MYB* gene promoter regions led to identification of a number of motifs involved in abiotic stress response. The "MYB" motif was found to be most abundant and has been reported to be involved in drought response by acting as a favourable binding site for bZIP TFs [55,56]. Similarly, "MYC" CRE is associated with response to drought stress and acts as a binding site for bHLH TFs [57,58]. "ABRE" and "DRE" motifs have been well characterised and known to regulate gene expression in response to drought stress in ABA-dependent and ABA-independent gene expression, respectively [59]. In addition to these, numerous other stress responsive elements, such as "STRE", "LTR", and "W box", were also found which indicate that *AmR2R3 MYB* gene family plays a crucial role in abiotic stress tolerance in *A. marina*.

In silico gene expression analysis revealed a number of members of the *AmR2R3 MYB* family which are differentially expressed during drought and salinity stress. Interestingly, the expression of these genes was affected more by drought stress than salt and was more evident in leaf tissue rather than root. Roots are the first tissues to be affected by salinity [60] and the *R2R3 MYB* genes are known to be differentially expressed in plants like *Arabidopsis* and rice [3] under conditions of high salinity. However, the absence of such marked difference in expression of *AmR2R3 MYB* could be attributed to the natural adaptation of *A. marina* to regions of high salinity which allows it to tolerate higher levels of salinity. Drought stress, on the other hand, had a more noticeable effect on the *AmR2R3 MYB*, especially in the leaf tissue. Only two members, *AmMYB87* and *AmMYB39*, showed considerable differential expression in response to salt stress and could be promising candidates for functional characterisation.

5. Conclusions

This study enumerates the unique characteristics of the *R2R3 MYB* gene family in the monotypic halophyte *Avicennia marina*. It sheds light on the evolution and functional diversification of the AmR2R3 MYB family based on comparisons with model plants such as Arabidopsis and rice. We have also identified important candidates that may be crucial to abiotic stress response in plants. This study can act as a foundation for selecting candidates for further functional characterisation.

Supplementary Materials: The following are available online at https://www.mdpi.com/2073-4395/11/1/123/s1, Figure S1: Logos and sequences of the motifs identified in the AmR2R3 MYB peptide sequences, Figure S2: Motifs identified in the peptide sequences of AmMYB63 and AmMYB137, File S1: File containing the fasta formatted CDS and peptide sequences of AmR2R3 MYB genes predicted in this study, Table S1: Nomenclature of scaffold to designate chromosomes, Table S2: Chromosomal positions of homologous R2R3 MYB CDS sequences in *A. marina*–*A. thalina* and *A. marina*–*O. sativa*, Table S3: Z score values of Log2 normalised FPKM values from in silico expression analysis of *AmR2R3 MYB* genes.

Author Contributions: Conceptualization, A.P.; data curation, P.S.S. and S.S.; formal analysis, S.P., P.S.S., and S.S.; funding acquisition, A.P.; methodology, S.P.; supervision, A.P.; validation, S.P., P.S.S., and S.S.; writing—original draft, S.P.; writing—review and editing, P.S.S. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the institutional grant from Department of Biotechnology, Government of India.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: Whole genome assembly and annotation files for *Avicennia marina* can be found at [35]. RNA-seq data used in gene expression analysis was downloaded from NCBI SRA database: Accession numbers SRR2029733, SRR2029734, SRR2029735, SRR2029736, SRR2029738, SRR2029739.

Acknowledgments: P.S.S. acknowledges the fellowship from DST-INSPIRE, Ministry of Science and Technology, Government of India. The authors acknowledge Department of Biotechnology, Government of India.

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

References

1. Anwar, A.; Kim, J.-K. Transgenic Breeding Approaches for Improving Abiotic Stress Tolerance: Recent Progress and Future Perspectives. *Int. J. Mol. Sci.* **2020**, *21*, 2695. [CrossRef]
2. Lamers, J.; van der Meer, T.; Testerink, C. How Plants Sense and Respond to Stressful Environments. *Plant Physiol.* **2020**, *182*, 1624–1635. [CrossRef]
3. Haak, D.C.; Fukao, T.; Grene, R.; Hua, Z.; Ivanov, R.; Perrella, G.; Li, S. Multilevel Regulation of Abiotic Stress Responses in Plants. *Front. Plant Sci.* **2017**, *8*, 1564. [CrossRef]
4. Li, X.; Tang, Y.; Li, H.; Luo, W.; Zhou, C.; Zhang, L.; Lv, J. A Wheat R2R3 MYB Gene TaMpc1-D4 Negatively Regulates Drought Tolerance in Transgenic Arabidopsis and Wheat. *Plant Sci.* **2020**, *299*, 110613. [CrossRef]

5. Baillo, E.H.; Kimotho, R.N.; Zhang, Z.; Xu, P. Transcription Factors Associated with Abiotic and Biotic Stress Tolerance and Their Potential for Crops Improvement. *Genes* **2019**, *10*, 771. [CrossRef]
6. Wang, P.; Yang, C.; Chen, H.; Luo, L.; Leng, Q.; Li, S.; Han, Z.; Li, X.; Song, C.; Zhang, X.; et al. Exploring Transcription Factors Reveals Crucial Members and Regulatory Networks Involved in Different Abiotic Stresses in Brassica Napus L. *BMC Plant Biol.* **2018**, *18*, 202. [CrossRef]
7. Zuo, C.; Tang, Y.; Fu, H.; Liu, Y.; Zhang, X.; Zhao, B.; Xu, Y. Elucidation and Analyses of the Regulatory Networks of Upland and Lowland Ecotypes of Switchgrass in Response to Drought and Salt Stresses. *PLoS ONE* **2018**, *13*, e0204426. [CrossRef]
8. Mondal, S.K.; Roy, S. Genome-Wide Sequential, Evolutionary, Organizational and Expression Analyses of Phenylpropanoid Biosynthesis Associated MYB Domain Transcription Factors in Arabidopsis. *J. Biomol. Struct. Dyn.* **2018**, *36*, 1577–1601. [CrossRef]
9. Muthamilarasan, M.; Khandelwal, R.; Yadav, C.B.; Bonthala, V.S.; Khan, Y.; Prasad, M. Identification and Molecular Characterization of MYB Transcription Factor Superfamily in C4 Model Plant Foxtail Millet (*Setaria italica* L.). *PLoS ONE* **2014**, *9*, e109920. [CrossRef]
10. Wong, D.C.J.; Schlechter, R.; Vannozzi, A.; Höll, J.; Hmmam, I.; Bogs, J.; Tornielli, G.B.; Castellarin, S.D.; Matus, J.T. A Systems-Oriented Analysis of the Grapevine R2R3-MYB Transcription Factor Family Uncovers New Insights into the Regulation of Stilbene Accumulation. *DNA Res.* **2016**, *23*, 451–466. [CrossRef]
11. Du, H.; Feng, B.-R.; Yang, S.-S.; Huang, Y.-B.; Tang, Y.-X. The R2R3-MYB Transcription Factor Gene Family in Maize. *PLoS ONE* **2012**, *7*, e37463. [CrossRef]
12. Wilkins, O.; Nahal, H.; Foong, J.; Provart, N.J.; Campbell, M.M. Expansion and Diversification of the Populus R2R3-MYB Family of Transcription Factors. *Plant Physiol.* **2009**, *149*, 981–993. [CrossRef]
13. He, Q.; Jones, D.C.; Li, W.; Xie, F.; Ma, J.; Sun, R.; Wang, Q.; Zhu, S.; Zhang, B. Genome-Wide Identification of R2R3-MYB Genes and Expression Analyses During Abiotic Stress in Gossypium Raimondii. *Sci. Rep.* **2016**, *6*, 22980. [CrossRef]
14. Katiyar, A.; Smita, S.; Lenka, S.K.; Rajwanshi, R.; Chinnusamy, V.; Bansal, K.C. Genome-Wide Classification and Expression Analysis of MYB Transcription Factor Families in Rice and Arabidopsis. *BMC Genom.* **2012**, *13*, 544. [CrossRef]
15. Jung, C.; Seo, J.S.; Han, S.W.; Koo, Y.J.; Kim, C.H.; Song, S.I.; Nahm, B.H.; Choi, Y.D.; Cheong, J.-J. Overexpression of AtMYB44 Enhances Stomatal Closure to Confer Abiotic Stress Tolerance in Transgenic Arabidopsis. *Plant Physiol.* **2008**, *146*, 623–635. [CrossRef]
16. He, C.; Teixeira da Silva, J.A.; Wang, H.; Si, C.; Zhang, M.; Zhang, X.; Li, M.; Tan, J.; Duan, J. Mining MYB Transcription Factors from the Genomes of Orchids (Phalaenopsis and Dendrobium) and Characterization of an Orchid R2R3-MYB Gene Involved in Water-Soluble Polysaccharide Biosynthesis. *Sci. Rep.* **2019**, *9*, 13818. [CrossRef]
17. Liu, J.; Osbourn, A.; Ma, P. MYB Transcription Factors as Regulators of Phenylpropanoid Metabolism in Plants. *Mol. Plant* **2015**, *8*, 689–708. [CrossRef]
18. Stracke, R.; Werber, M.; Weisshaar, B. The R2R3-MYB Gene Family in Arabidopsis Thaliana. *Curr. Opin. Plant Biol.* **2001**, *4*, 447–456. [CrossRef]
19. Paz-Ares, J.; Ghosal, D.; Wienand, U.; Peterson, P.A.; Saedler, H. The Regulatory C1 Locus of Zea Mays Encodes a Protein with Homology to Myb Proto-Oncogene Products and with Structural Similarities to Transcriptional Activators. *EMBO J.* **1987**, *6*, 3553–3558. [CrossRef]
20. Zhao, Y.; Yang, Z.; Ding, Y.; Liu, L.; Han, X.; Zhan, J.; Wei, X.; Diao, Y.; Qin, W.; Wang, P.; et al. Over-Expression of an R2R3 MYB Gene, GhMYB73, Increases Tolerance to Salt Stress in Transgenic Arabidopsis. *Plant Sci.* **2019**, *286*, 28–36. [CrossRef]
21. Yu, Y.; Bi, C.; Wang, Q.; Ni, Z. Overexpression of TaSIM Provides Increased Drought Stress Tolerance in Transgenic Arabidopsis. *Biochem. Biophys. Res. Commun.* **2019**, *512*, 66–71. [CrossRef]
22. Zhang, Y.-L.; Zhang, C.-L.; Wang, G.-L.; Wang, Y.-X.; Qi, C.-H.; Zhao, Q.; You, C.-X.; Li, Y.-Y.; Hao, Y.-J. The R2R3 MYB Transcription Factor MdMYB30 Modulates Plant Resistance against Pathogens by Regulating Cuticular Wax Biosynthesis. *BMC Plant Biol.* **2019**, *19*, 362. [CrossRef] [PubMed]
23. Tang, Y.; Bao, X.; Zhi, Y.; Wu, Q.; Guo, Y.; Yin, X.; Zeng, L.; Li, J.; Zhang, J.; He, W.; et al. Overexpression of a MYB Family Gene, OsMYB6, Increases Drought and Salinity Stress Tolerance in Transgenic Rice. *Front. Plant Sci.* **2019**, *10*, 168. [CrossRef]
24. Dai, X.; Xu, Y.; Ma, Q.; Xu, W.; Wang, T.; Xue, Y.; Chong, K. Overexpression of an R1R2R3 MYB Gene, OsMYB3R-2, Increases Tolerance to Freezing, Drought, and Salt Stress in Transgenic Arabidopsis. *Plant Physiol.* **2007**, *143*, 1739–1751. [CrossRef]
25. Pasquali, G.; Biricolti, S.; Locatelli, F.; Baldoni, E.; Mattana, M. Osmyb4 Expression Improves Adaptive Responses to Drought and Cold Stress in Transgenic Apples. *Plant Cell Rep.* **2008**, *27*, 1677–1686. [CrossRef]
26. Wang, Z.L.; Li, P.H.; Fredricksen, M.; Gong, Z.Z.; Kim, C.S.; Zhang, C.; Bohnert, H.J.; Zhu, J.K.; Bressan, R.A.; Hasegawa, P.M.; et al. Expressed Sequence Tags from Thellungiella Halophila, a New Model to Study Plant Salt-Tolerance. *Plant Sci.* **2004**, *166*, 609–616. [CrossRef]
27. Uno, Y.; Furihata, T.; Abe, H.; Yoshida, R.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Arabidopsis Basic Leucine Zipper Transcription Factors Involved in an Abscisic Acid-Dependent Signal Transduction Pathway under Drought and High-Salinity Conditions. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 11632–11637. [CrossRef]
28. Oh, S.J.; Kim, Y.S.; Kwon, C.W.; Park, H.K.; Jeong, J.S.; Kim, J.K. Overexpression of the Transcription Factor AP37 in Rice Improves Grain Yield under Drought Conditions. *Plant Physiol.* **2009**, *150*, 1368–1379. [CrossRef]
29. Ganesan, G.; Sankararamasubramanian, H.M.; Harikrishnan, M.; Ganpudi, A.; Parida, A. A MYB Transcription Factor from the Grey Mangrove Is Induced by Stress and Confers NaCl Tolerance in Tobacco. *J. Exp. Bot.* **2012**, *63*, 4549–4561. [CrossRef]

30. Tomlinson, P.B. *The Botany of Mangroves, 1986 the Botany of Mangroves*, 1st ed.; Cambridge University Press: Cambridge, UK, 2016.
31. Parida, A.; George, S. Sustaining and Enhancing Crop Productivity in an Era of Climate Change. *Curr. Sci.* **2015**, *109*, 462–473.
32. Tuteja, N.; Gill, S.S.; Tiburcio, A.F.; Tuteja, R. *Improving Crop Resistance to Abiotic Stress*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
33. Joshi, R.; Mangu, V.R.; Bedre, R.; Sanchez, L.; Pilcher, W.; Zandkarimi, H.; Baisakh, N. Salt Adaptation Mechanisms of Halophytes: Improvement of Salt Tolerance in Crop Plants. In *Elucidation of Abiotic Stress Signaling in Plants: Functional Genomics Perspectives, Volume 2*; Pandey, G.K., Ed.; Springer: New York, NY, USA, 2015; pp. 243–279. [CrossRef]
34. Das, S.K.; Patra, J.K.; Thatoi, H. Antioxidative Response to Abiotic and Biotic Stresses in Mangrove Plants: A Review. *Int. Rev. Hydrobiol.* **2016**, *101*, 3–19. [CrossRef]
35. Friis, G.; Vizueta, J.; Nelson, D.R.; Khraiwesh, B.; Qudeimat, E.; Salehi-Ashtiani, K.; Ortega, A.; Marshell, A.; Duarte, C.M.; Burt, J.A. A High-Quality Genome Assembly and Annotation of the Gray Mangrove, Avicennia Marina. *bioRxiv* **2020**. [CrossRef]
36. Tombuloglu, H. Genome-Wide Identification and Expression Analysis of R2R3, 3R- and 4R-MYB Transcription Factors during Lignin Biosynthesis in Flax (Linum Usitatissimum). *Genomics* **2020**, *112*, 782–795. [CrossRef] [PubMed]
37. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J.M., Ed.; Springer Protocols Handbooks; Humana Press: Totowa, NJ, USA, 2005; pp. 571–607. [CrossRef]
38. Savojardo, C.; Martelli, P.L.; Fariselli, P.; Profiti, G.; Casadio, R. BUSCA: An Integrative Web Server to Predict Subcellular Localization of Proteins. *Nucleic Acids Res.* **2018**, *46*, W459–W466. [CrossRef] [PubMed]
39. Voorrips, R.E. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J. Hered.* **2002**, *93*, 77–78. [CrossRef]
40. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]
41. Wang, Y.; Tang, H.; DeBarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.; 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]
42. Quinlan, A.R.; Hall, I.M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* **2010**, *26*, 841–842. [CrossRef]
43. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a Database of Plant Cis-Acting Regulatory Elements and a Portal to Tools for in Silico Analysis of Promoter Sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327. [CrossRef]
44. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [CrossRef]
45. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [CrossRef]
46. Howe, E.; Holton, K.; Nair, S.; Schlauch, D.; Sinha, R.; Quackenbush, J. MeV: MultiExperiment Viewer. In *Biomedical Informatics for Cancer Research*; Ochs, M.F., Casagrande, J.T., Davuluri, R.V., Eds.; Springer: Boston, MA, USA, 2010; pp. 267–277. [CrossRef]
47. Yanagisawa, S. Transcription Factors in Plants: Physiological Functions and Regulation of Expression. *J. Plant Res.* **1998**, *111*, 363–371. [CrossRef]
48. Zörb, C.; Geilfus, C.-M.; Dietz, K.-J. Salinity and Crop Yield. *Plant Biol.* **2019**, *21* (Suppl. S1), 31–38. [CrossRef]
49. Kavitha, K.; George, S.; Venkataraman, G.; Parida, A. A Salt-Inducible Chloroplastic Monodehydroascorbate Reductase from Halophyte Avicennia Marina Confers Salt Stress Tolerance on Transgenic Plants. *Biochimie* **2010**, *92*, 1321–1329. [CrossRef] [PubMed]
50. Khan, S.-A.; Li, M.-Z.; Wang, S.-M.; Yin, H.-J. Revisiting the Role of Plant Transcription Factors in the Battle against Abiotic Stress. *Int. J. Mol. Sci.* **2018**, *19*, 1634. [CrossRef] [PubMed]
51. Cao, Y.; Han, Y.; Li, D.; Lin, Y.; Cai, Y. MYB Transcription Factors in Chinese Pear (Pyrus Bretschneideri Rehd.): Genome-Wide Identification, Classification, and Expression Profiling during Fruit Development. *Front. Plant Sci.* **2016**, *7*, 577. [CrossRef] [PubMed]
52. Yan, Q.; Wu, F.; Ma, T.; Zong, X.; Ma, Q.; Li, J.; Zhao, Y.; Wang, Y.; Zhang, J. Comprehensive Analysis of BZIP Transcription Factors Uncovers Their Roles during Dimorphic Floret Differentiation and Stress Response in Cleistogenes Songorica. *BMC Genom.* **2019**, *20*, 760. [CrossRef] [PubMed]
53. Biłas, R.; Szafran, K.; Hnatuszko-Konka, K.; Kononowicz, A.K. Cis-Regulatory Elements Used to Control Gene Expression in Plants. *Plant Cell Tissue Organ Cult.* **2016**, *127*, 269–287. [CrossRef]
54. Heinekamp, T.; Kuhlmann, M.; Lenk, A.; Strathmann, A.; Dröge-Laser, W. The Tobacco BZIP Transcription Factor BZI-1 Binds to G-Box Elements in the Promoters of Phenylpropanoid Pathway Genes in Vitro, but It Is Not Involved in Their Regulation in Vivo. *Mol. Gen. Genom.* **2002**, *267*, 16–26. [CrossRef]
55. Heinekamp, T.; Strathmann, A.; Kuhlmann, M.; Froissard, M.; Müller, A.; Perrot-Rechenmann, C.; Dröge-Laser, W. The Tobacco BZIP Transcription Factor BZI-1 Binds the GH3 Promoter in Vivo and Modulates Auxin-Induced Transcription. *Plant J.* **2004**, *38*, 298–309. [CrossRef]
56. Abe, H.; Yamaguchi-Shinozaki, K.; Urao, T.; Iwasaki, T.; Hosokawa, D.; Shinozaki, K. Role of Arabidopsis MYC and MYB Homologs in Drought- and Abscisic Acid-Regulated Gene Expression. *Plant Cell* **1997**, *9*, 1859–1868. [CrossRef] [PubMed]

57. Klepikova, A.V.; Kasianov, A.S.; Gerasimov, E.S.; Logacheva, M.D.; Penin, A.A. A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *Plant J.* **2016**, *88*, 1058–1070. [CrossRef] [PubMed]
58. Abe, H.; Urao, T.; Ito, T.; Seki, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Arabidopsis AtMYC2 (BHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signaling. *Plant Cell* **2003**, *15*, 63–78. [CrossRef] [PubMed]
59. Yamaguchi-Shinozaki, K.; Shinozaki, K. Organization of Cis-Acting Regulatory Elements in Osmotic- and Cold-Stress-Responsive Promoters. *Trends Plant Sci.* **2005**, *10*, 88–94. [CrossRef] [PubMed]
60. Isayenkov, S.V.; Maathuis, F.J. Plant salinity stress: Many unanswered questions remain. *Front. Plant Sci.* **2019**, *10*, 8.