

Appearance and elaboration of the ethylene receptor family during land plant evolution

Daniel R. Gallie

Received: 9 September 2014 / Accepted: 11 February 2015 / Published online: 15 February 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Ethylene is perceived following binding to endoplasmic reticulum-localized receptors, which in *Arabidopsis thaliana*, include ETR1, ERS1, EIN4, ETR2, and ERS2. These receptors fall into two subfamilies based on conservation of features within their histidine kinase domain. Subfamily 1 contains ETR1 and ERS1 whereas subfamily 2 contains EIN4, ETR2, and ERS2. Because ethylene receptors are found only in plants, this raises questions of when each receptor evolved. Here it is shown that subfamily 1 receptors encoded by a multigene family are present in all charophytes examined, these being most homologous to ETR1 based on their evolutionary relationship as well as containing histidine kinase and receiver domains. In charophytes and *Physcomitrella patens*, one or more gene family members contain the intron characteristic of subfamily 2 genes, indicating the first step in subfamily 2 receptor evolution. ERS1 homologs appear in basal angiosperm species after *Amborella trichopoda* and, in some early and basal angiosperm species and monocots in general, it is the only subfamily 1 receptor present. Distinct EIN4 and ETR2 homologs appear only in core eudicots and ERS2 homologs appear only in the Brassicaceae, suggesting it is the most recent receptor to evolve. These findings show that a subfamily 1 receptor had evolved and a subfamily 2 receptor had begun to evolve in plants prior to the colonization of land and only these two existed up to the appearance of the first basal angiosperm. The appearance of ERS2 in the Brassicaceae suggests ongoing evolution of the ethylene receptor family.

Keywords Ethylene · Ethylene receptor · Gene duplication · Hormone perception · Signal transduction

Abbreviations

ACO	ACC oxidase
ACS	ACC synthase
CC	Coiled coil
CTR1	Constitutive triple response1
Cys	Cysteine
EST	Expressed sequence tag
GAF	cGMP-specific phosphodiesterases, adenylyl cyclases and Fh1A
Tyr	Tyrosine

Introduction

Ethylene serves as a hormone in plants involved in regulating diverse aspects of growth and development and in responses to adverse growth conditions (Mattoo and Suttle 1991; Ables et al. 1992; Bleecker and Kende 2000; Klee 2004; Lin et al. 2009; Schaller 2012; Binder et al. 2012; Shakeel et al. 2013). Ethylene is perceived through its binding to receptors which are localized to the endoplasmic reticulum (Chang and Shockey 1999; Chang and Stadler 2001; Wang et al. 2002; Chen et al. 2002; Stepanova and Alonso 2005; Lin et al. 2009; Schaller 2012; Binder et al. 2012; Shakeel et al. 2013). As negative regulators, ethylene receptors in conjunction with the constitutive triple response1 (CTR1) Raf-like kinase, signal to repress ethylene responses in the absence of the hormone (Kieber et al. 1993; Hua and Meyerowitz 1998; Clark et al. 1998). Consequently, receptor loss-of-function mutants exhibit increased sensitivity to ethylene and/or constitutive ethylene responses in *A. thaliana* and *Solanum lycopersicum* (Hua and Meyerowitz 1998; Tieman et al. 2000).

D. R. Gallie (✉)
Department of Biochemistry, University of California, Riverside,
CA 92521-0129, USA
e-mail: drgallie@citrus.ucr.edu

Five different types of ethylene receptors (i.e., ETR1, ERS1, EIN4, ETR2, and ERS2) are expressed in *A. thaliana* and these belong to one of two subfamilies: subfamily 1 is composed of ETR1 and ERS1 which have functional histidine kinase domains (Gamble et al. 1998; Moussatche and Klee 2004) while subfamily 2 is composed of ETR2, ERS2, and EIN4 which possess Ser/Thr kinase activity (Moussatche and Klee 2004). A C-terminal receiver domain is present downstream of the histidine kinase domain in *A. thaliana* ETR1, ETR2, and EIN4 whereas ERS1 and ERS2 lack a C-terminal receiver domain. Ethylene receptors likely evolved from bacterial and yeast two-component regulators as these also have similar domains for signal input and output and have histidine kinase activity (Schaller 1997; Chang and Stewart 1998; Chang and Stadler 2001; Lohrmann and Harter 2002). Loss of subfamily 1 receptor expression results in a more severe constitutive ethylene response than does loss of subfamily 2 receptors in *A. thaliana* (Hall and Bleecker 2003; Wang et al. 2003; Xie et al. 2006; Qu et al. 2007). Moreover, loss of expression of the subfamily 1 receptors cannot be rescued by the ectopic expression of subfamily 2 receptors (Wang et al. 2003). Ethylene receptors such as ETR1 form homodimers through a disulfide bond between Cys-4 and Cys-6 (Schaller et al. 1995; Rodriguez et al. 1999) and ETR1 can also interact with subfamily 2 receptors, indicating that ethylene receptors are present as clusters which may facilitate their function (Schaller et al. 1995; O'Malley et al. 2005; Grefen et al. 2008; Gao et al. 2008).

Ethylene binding occurs within an N-terminal transmembrane domain that spans the endoplasmic reticulum membrane and a cysteine and histidine residue in transmembrane domain III are responsible for binding copper to which ethylene binds (Rodriguez et al. 1999). Mutation of this cysteine residue to a tyrosine in *A. thaliana* receptors results in ethylene insensitivity in a dominant negative manner as the receptor is unable to bind ethylene and therefore retains its ability to repress ethylene responses in the presence of ethylene (Bleecker et al. 1988; Guzman and Ecker 1990; Chang et al. 1993; Chen and Bleecker 1995). Ethylene binding causes loss of signaling from ethylene receptors and CTR1 to release repression of downstream components of the ethylene signaling pathway which allows the induction of expression of those genes involved in ethylene responses (Chao et al. 1997; Solano et al. 1998; Alonso et al. 1999).

Isolation of ethylene receptors from *Zea mays* revealed that this species expresses just two types: subfamily 1 receptors that lack a receiver domain (S1–R or ERS1-like) and subfamily 2 receptors that contain a receiver domain (S2+R or EIN4/ETR2-like) (Gallie and Young 2004). Two genes were reported encoding S1–R receptors (i.e., *ZmERS1a* and *ZmERS1b*) and two encoding S2+R

receptors (i.e., *ZmETR2a* and *ZmETR2b*) in *Z. mays* (Gallie and Young 2004). A Cys to Tyr mutation introduced in transmembrane domain III of *ZmERS1b* and *ZmETR2b* conferred a state of ethylene insensitivity in a subfamily 1-dependent, dominant manner when the mutant receptors were expressed in *A. thaliana* and resulted in the characteristic ethylene insensitive phenotypes such as increased leaf size and delayed leaf senescence (Chen and Gallie 2010). Expression of just the N-terminal transmembrane domain of mutant *Zmers1b* was sufficient to confer dominance over endogenous *A. thaliana* ethylene receptors whereas expression of the mutant *Zmetr2b* N-terminal domain did not (Chen and Gallie 2010). *Oryza sativa* also expresses S1–R receptors, OsERS1 and OsERS2, and S2+R receptors, OsETR2, OsETR3, and OsETR4 (Yau et al. 2004). This suggests monocots may differ substantially from eudicots in the type of receptors present.

The fact that ethylene receptors are found only in plants raises questions of which receptor represents the foundational member of the gene family, when the other family members evolved, and whether these diverse receptors are shared among modern plants. In this report, the evolution of ethylene receptors in plants was examined to provide insight into these aspects of the gene family that may yield clues into the function of each receptor. Although ethylene receptors seem to be absent from marine algae, charophytes do express subfamily 1 receptors (S1+R or ETR1-like) from a multigene family, suggesting that ethylene receptors as they exist in modern plants may have first appeared in fresh water algae. The first indication of the evolution of subfamily 2 receptor genes was observed in the intron structure of one ethylene receptor gene from the charophyte *Klebsorbidium flaccidum*, which likely arose from a gene duplication event of the progenitor subfamily 1 (ETR1-like) receptor gene, and in four of the eight receptors genes of *P. patens*. An S1–R (ERS1-like) receptor likely did not appear until after *Amborella trichopoda*, the basal most angiosperm species, had evolved. Like *Z. mays* and *O. sativa*, all monocots examined only possess S1–R and S2+R receptors. The differentiation of S2+R receptors into distinct EIN4-like and ETR2-like receptors occurred following the appearance of core eudicots and ERS2 evolved specifically in the Brassicaceae. These results show how the members of the ethylene receptor gene family appeared during plant evolution and how the gene family continues to evolve dynamically.

Materials and methods

Sequence alignment and phylogenetic analysis

The ethylene receptor amino acid sequences of *A. thaliana* used in this study are available in the National Center

for Biotechnology Information (NCBI) sequence database (<http://www.ncbi.nlm.nih.gov/>) and were used as queries to perform BLAST searches of the Phytozome database (v9.1) (<http://www.phytozome.com>) (Goodstein et al. 2012) for orthologs in most of the species used in the study. A BLAST search of the Amborella Genome Database (v1.0) (<http://amborella.org/>) (Amborella Genome Database 2013) was performed to identify ethylene receptors in *A. trichopoda*. A BLAST search of the Spruce Genome Project (v1.0) (<http://congenie.org/blastsearch>) (Nystedt et al. 2013) was performed to identify ethylene receptors in *Picea abies*. A BLAST search of the *K. flaccidum* Genome Project (v1.0) (http://genome.microbedb.jp/blast/blast_search/ klebsormidium/genes) was performed to identify ethylene receptors in *K. flaccidum*. For four species in which genome sequence was not available, EST sequences were obtained from the Onekp project in May 2014. Iterative searches of species were performed using ethylene receptor amino acid sequences from that species. Predicted protein sequences of ESTs were obtained using the ExpASY Translate tool (Gasteiger et al. 2003). With the exception of the sequences from *Chlamydomonas reinhardtii*, only sequences containing the conserved cysteine and histidine residues in the transmembrane domain involved in ethylene binding were included for subsequent analysis. Amino acid sequence alignments were performed by ClustalW2 with the following parameters: pairwise gap opening penalty 10, pairwise gap extension penalty 0.1, multiple gap opening penalty 10, multiple gap extension penalty 0.2, Gonnet protein weight matrix, and no end gap separation.

The evolutionary history was inferred by using the Maximum Likelihood method based on the Le Gascuel model (Le and Gascuel 2008) and the bootstrap consensus tree was inferred from 500 replicates and a LG amino acid replacement matrix to represent the phylogenetic relationship among sequences of the proteins analyzed. The tree with the highest log likelihood (−102975.4201) is shown. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using a JTT model. Similar trees were seen using the Whalen and Goldman substitution model and did not alter the findings. The tree is drawn to scale. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test is shown next to the branches in the figures with branch lengths measured in the number of substitutions per site. No branches were collapsed regardless of bootstrap values. The analysis involved 198 amino acid sequences. There were a total of 1569 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamur et al. 2013).

Gene sequences used the analyses were from *A. thaliana* (AT1G66340; AT2G40940; AT3G23150; AT1G04310;

AT3G04580); *Arabidopsis lyrata* (Al_934910, Al_475737, Al_905365, Al_477715, Al_910099); *Capsella rubella* (Cr_10025204, Cr_10019861, Cr_10008561, Cr_10013026, Cr_10013030); *T. halophila* (Th_10001362, Th_10018186, Th_10007018, Th_10020114, Th_10020124); *Brassica rapa* (Br_004449, Br_004160, Br_030564, Br_015303, Br_023756, Br_040134); *Carica papaya* (Cp_151.32, Cp_84.52, Cp_2388.2); *Gossypium raimondii* (Gr_007G126300, Gr_008G143100, Gr_002G038300, Gr_009G358200, Gr_005G047900, Gr_006G247800, Gr_007G280100); *Theobroma cacao* (Tc_1EG004726, Tc_1EG032339, Tc_1EG026253, Tc_1EG020821); *Populus trichocarpa* (Pt_002G201500, Pt_003G032300, Pt_001G204200, Pt_010G074300, Pt_008G164400, Pt_013G044100, Pt_019G014300); *Eucalyptus grandis* (Eg_K03513, Eg_J02086, Eg_H03145, Eg_H04259); *Citrus sinensis* (Csoorange_1.1g005591, Csoorange_1.1g004636, Csoorange_1.1g004510, Csoorange_1.1g006508); *Cucumis sativus* (Cs_205330, Cs_178860, Cs_255140); *Mimulus guttatus* (Mg_003242, Mg_001910, Mg_001671, Mg_002003); *Manihot esculenta* (Me_4.1_003444, Me_4.1_002394, Me_4.1_002375, Me_4.1_002152, Me_4.1_027924, Me_4.1_002165); *Ricinus communis* (Rc_29986, Rc_28802, Rc_29603, Rc_29680); *Medicago truncatula* (Mt_7g109150, Mt_4g031150, Mt_1g079790, Mt_7g116330); *Prunus persica* (Ppe_002692, Ppe_001917, Ppe_001786, Ppe_001846); *Malus domestica* (Md_257135, Md_242413, Md_300556, Md_557234, Md_267951, Md_920189, Md_195916, Md_219737, Md_393617, Md_231172); *Fragaria vesca* (Fv_11090, Fv_21106, Fv_e32532, Fv_16612); *Phaseolus vulgaris* (Pv_001G210200, Pv_011G216400, Pv_007G271700, Pv_007G129500, Pv_006G106400, Pv_006G106300); *Glycine max* (Gm_03g37470, Gm_19g40090, Gm_12g37050, Gm_10g33240, Gm_20g36440, Gm_10g31040, Gm_19g43840, Gm_03g41220); *Vitis vinifera* (Vv_GSVIVG01028053001, Vv_GSVIVG01038085001, Vv_GSVIVG01024904001, Vv_GSVIVG01027723001, Vv_GSVIVG01036213001); *S. lycopersicum* (Sl_09g075440, Sl_07g056580, Sl_12g011330, Sl_06g053710, Sl_09g089610, Sl_11g006180, Sl_05g055070); *Sorghum bicolor* (Sb_09g004300, Sb_01g010930, Sb_02g035430, Sb_04g007500, Sb_0169s002030, Sb_06g001740); *Brachypodium distachyon* (Bi_2g35080, Bi_4g00200, Bi_1g11540, Bi_3g56550, Bi_3g57807, Bi_3g55730, Bi_5g00700); *Z. mays* (Zm_AC194965, Zm_2G102601, Zm_2G073668, Zm_2G077008, Zm_2G318689, Zm_2G075368, Zm_2G420801, Zm_2G089010); *Oryza sativa* (Os05g06320, Os_03g49500, Os_07g15540, Os_02g57530, Os_04g08740); *Panicum virgatum* (Pv_00024255, Pv_00036840, Pv_00069616, Pv_0003739, Pv_00046085, Pv_00051650, Pv_00003352, Pv_00050400); *Setaria italica* (Si_009598, Si_034660, Si_032462, Si_016376, Si_019927, Si_009431);

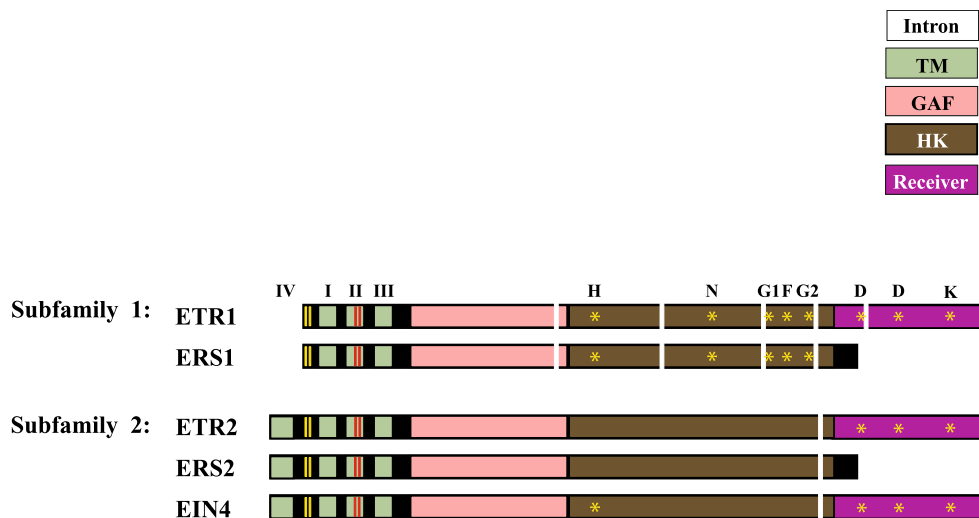


Fig. 1 Domain organization of the ethylene receptor gene family in *A. thaliana*. Comparison of subfamily 1 receptors, i.e., ETR1 and ERS1, with subfamily 2 receptors, i.e., ETR2, EIN4, and ERS2. The N-terminal, hydrophobic, transmembrane domains are indicated by green boxes with the N-proximal cysteine residues involved in receptor dimerization indicated by vertical yellow bars and the cysteine and histidine residues in transmembrane domain II involved in eth-

ylene binding indicated by vertical red bars. The GAF, histidine protein kinase domain, and receiver domains are indicated according to the key at the upper right. The five consensus motifs (H, N, G1, F, and G2) within the histidine kinase domain are indicated with yellow asterisks as are the conserved aspartate and lysine residues in the receiver domain. The position of each intron is indicated by vertical white bars

Magnolia grandiflora (Mgr_8225, Mgr_24121); *Aquilegia coerulea* (Ac_006_00311, Ac_003_00648, Ac_042_00042, Ac_017_00676, Ac_034_00407); *Aristolochia elegans* (Ae_48333, Ae_48454); *Illicium floridanum* (If_106789, If_106923); *Austrobaileya scandens* (As_15022, As_4770); *A. trichopoda* (ATr_00032.33, ATr_00009.122); *P. abies* (Pa_10428397, Pa_67531, Pa_10048); *Selaginella moellendorffii* (Sm_267662, Sm_110685, Sm_230881, Sm_84824); *P. patens* (Pp_003G006300, Pp_013G004000, Pp_027G017000, Pp_016G055700, Pp_010G058200, Pp_001G117000, Pp_004G103700, Pp_004G103600), *K. flaccidum* (EST Kf_HO437642, genomic Kf100524, Kf100196, Kf100708, Kf100794, Kf100385); *Spirogyra pratensis* (EST Sg_GW599340; Sg_GW596207); *Penium margaritaceum* (EST Pm_HO601064); and *C. reinhardtii* (Creinhardtii_27574900, Creinhardtii_27564446).

Results

The appearance of ETR1-like receptors likely predates land plant evolution

The division of the five ethylene receptor genes in *A. thaliana* into subfamily 1 and 2 is based on gene organization, intron structure, the presence or absence of conserved elements in the histidine kinase domain, and their evolutionary relationship (Fig. 1) (Schaller 2012; Binder et al. 2012;

Shakeel et al. 2013). The subfamily 1 receptors ETR1 and ERS1 are S1+R and S1–R type receptors, respectively, whereas the subfamily 2 receptors EIN4 and ETR2 are S2+R type receptors and ERS2 is an S2–R type receptor.

Subfamily 1 members in *A. thaliana* contain multiple introns present in the GAF and histidine kinase domains as well as in the receiver domain when present and contain the conserved motifs within the histidine kinase domain required for the histidine kinase activity they possess (Fig. 1). In contrast, subfamily 2 members contain a single intron at the end of the histidine kinase domain and lack the conserved motifs in the histidine kinase domain with the only exception being the presence of a histidine residue within the H motif of EIN4. The subfamily 2 members, however, possess Ser-Thr kinase activity which ERS1 also possesses (Chang et al. 1993; Hua et al. 1995, 1998; Gamble et al. 1998; Sakai et al. 1998; Moussatche and Klee 2004). ETR1-like or ERS1-like receptors are defined as those homologs that are phylogenetically related to ETR1 or ERS1, respectively, and contain the domain structure and conserved motifs of each. EIN4-like, ETR2-like, or ERS2-like receptors are defined as those homologs that are phylogenetically related to EIN4, ETR2, or ERS2, respectively, and contain the domain structure of each while lacking most or all of the conserved motifs of the histidine kinase domain.

Genes encoding ethylene receptors that are similar in sequence and structure to those in higher plants have been

identified as far back in land plant evolution as bryophytes, e.g., *P. patens* (Binder et al. 2012). Two genes identified in the cyanobacterium *Anabaena* sp. strain PCC 7120 and a gene from *Synechocystis* strain 6803 (slr1212, NP_440714) share limited similarity to ETR1 (Mount and Chang 2002; Rodríguez et al. 1999). *Synechocystis* slr1212, but not the two *Anabaena* proteins, was shown to bind ethylene, data supporting the conclusion that it is an ethylene-binding protein and the possible bacterial origin for ethylene receptors (Rodríguez et al. 1999; Bleecker 1999). However, slr1212 lacks a receiver domain and some of the conserved sequence motifs of the histidine kinase domain (Mount and Chang 2002). Moreover, a coiled coil (CC) region is present between the N-terminal transmembrane and GAF domains of slr1212 and in the two *Anabaena* proteins which is absent in higher plant ethylene receptors. Two of these proteins also contain one or two PAS/PAC [PAS (Per, ARNT, Sim) followed by PAC] domains proximal to the CC domain not present in ethylene receptors of higher plants and neither *Anabaena* protein contains a GAF domain. Therefore, these proteins differ structurally from plant ethylene receptors and no study has reported how these proteins evolved in early plants into modern ethylene receptors. In addition, no other functional component for ethylene signaling, e.g., CTR1, has been demonstrated for either species. Interestingly, *Synechocystis* produces no detectable ethylene nor responds to the hormone while these have not been examined for *Anabaena* (Bleecker 1999).

No obvious ethylene receptor homologs are present in the genome of salt-water algal species such as *C. reinhardtii* nor is ethylene binding activity detected (Wang et al. 2006). Land plants, however, did not evolve from the Chlamydomonadales but rather from aquatic ancestors that are sister groups to charophycean algae (i.e., fresh water algae). Therefore, it is possible that ethylene receptors evolved prior to the appearance of land plants. Supporting this is the presence of ESTs encoding possible homologs to ethylene receptors as well as CTR1, EIN2, EIN3/ERF1, and EBF1 in *Coleochaete orbicularis* (Coleochaetales) and *S. pratensis* (Zygnematales) (Timme and Delwiche 2010).

The genome sequence of the charophyte *K. flaccidum* representing the Klebsormidiales was recently reported and the presence of an ETR1 receptor noted (Hori et al. 2014). A search of the *K. flaccidum* revealed the presence of at least five genes encoding subfamily 1 (ETR1-like) receptors and a search of charophyte EST databases identified one EST from *K. flaccidum* as well as two ESTs from *S. pratensis* and one EST from *P. margaritaceum* (Zygnematales) that encoded subfamily 1 (ETR1-like) receptors (Fig. 2). The predicted encoded proteins exhibit substantial sequence conservation with ETR1 and contain the cysteine and histidine residues involved in ethylene binding in the

transmembrane domain, the H, N, G1, F, and G2 motifs present in the histidine kinase domain of subfamily 1 receptors as well as the conserved aspartic acid and lysine residues in the receiver domain (Fig. 2). Moreover, homologs for CTR1, EIN3, and EBF1 are present in the *K. flaccidum* genome, indicating that critical factors necessary for ethylene signaling in addition to ethylene receptors had evolved in charophytes (Hori et al. 2014). A recent report using shotgun transcriptome assemblies from five charophytes demonstrated the presence of ethylene receptor homologs in *S. pratensis*, *C. orbicularis*, *Nitella mirabilis* but not in *Mesostigma viride*, which likely represents the earliest charophyte lineage (Ju et al. 2015).

Proto-subfamily 2 ethylene receptor genes appear in non-vascular land plants

To examine phylogenetically the relationship of putative charophycean ethylene receptors to those of land plants and to determine when other ethylene receptors may have arisen during plant evolution, ethylene receptor gene sequences were obtained from species throughout plant evolution and phylogenetic analysis of the proteins performed. The tree was rooted with two gene sequences from *C. reinhardtii* that were identified from searches using the ethylene receptor sequences identified in the fresh water algae species, *K. flaccidum*, *P. margaritaceum*, and *S. pratensis*. The two *C. reinhardtii* proteins do not contain predicted N-terminal transmembrane regions or the cysteine and histidine residues involved in ethylene binding but do contain elements of the histidine kinase domain, suggesting they are not ethylene receptors but function as response regulators with some similarity to ethylene receptors of lower plants.

Although previously suggested to contain seven ethylene receptor genes (Binder et al. 2012), the ethylene receptor gene family in *P. patens* had expanded to eight members which form two gene subfamilies (Figs. 3, 4). Of these, four gene members were more closely related to the ETR1 clade than the other four (Fig. 3). To determine whether the subfamily 1 receptors of the first group in *P. patens* contain the conserved motifs of the histidine kinase and receiver domains characteristic of ETR1, the ethylene receptor sequences were examined in detail. Two of the four *P. patens* subfamily 1 receptors contain the conserved motifs of the histidine kinase and receiver domains while a third contains all but the histidine residue within the H motif, and the fourth lacks the conserved lysine residue of the receiver domain (Fig. 5). This fourth member also has two deletions within the histidine kinase and receiver domains (Fig. 5). Analysis of the intron positions within these four genes revealed that they contain all five of the introns present in *A. thaliana* ETR1 (Fig. 5), demonstrating that the gene structure of subfamily 1 receptor genes was established at least

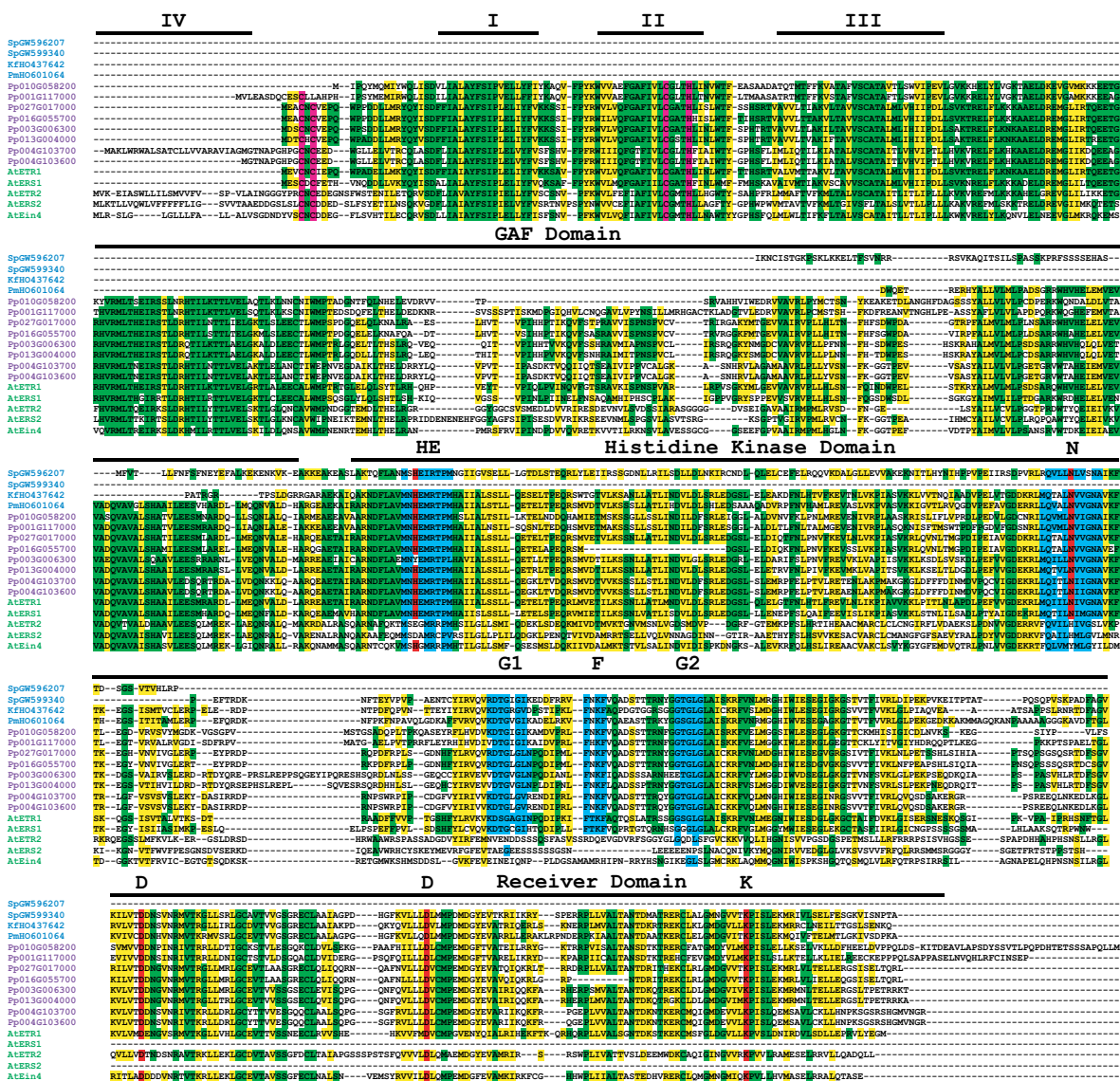


Fig. 2 Sequence alignment of ethylene receptor homologs in charophytes. Alignment of amino acid sequence of ethylene receptor homologs is shown with amino acid identity relative to *A. thaliana* ETR1 highlighted in green amino acid and amino acid similarity highlighted in yellow. The N-proximal cysteine residues involved in receptor dimerization are indicated in purple as are the cysteine and histidine residues in transmembrane domain II that are involved in ethylene binding. The transmembrane, GAF, histidine kinase, and receiver domains are indicated above the pertinent sequence. The sequences of the conserved motifs (H, N, G1, F, and G2) of the

histidine protein kinase domains are indicated in blue with the conserved histidine of the H motif and the conserved asparagine of the N motif indicated in red. The conserved aspartate and lysine residues in the receiver domain are indicated in purple. Protein sequences used were: *S. pratensis* (Sp) GW596207 and GW599340; *K. flaccidum* (Kf) Kf100708; *P. margaritaceum* (Pm) HO61064; *P. patens* (Pp) 01G058200, 01G117000, 02G017000, 01G055700, 03G006300, 01G3004000, 04G103700, and 04G103600; and *A. thaliana* (At) ETR1, ERS1, ETR2, ERS2, EIN4. Only one of the five *K. flaccidum* ETR1-like receptor sequences is shown

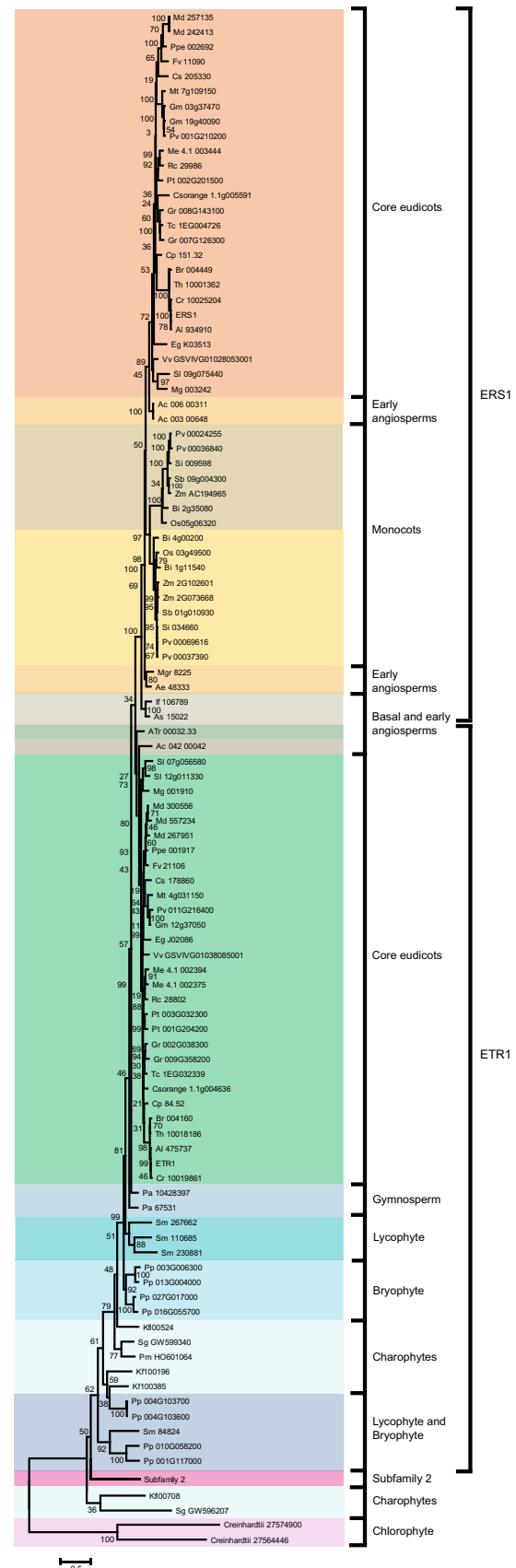
by the appearance of *P. patens*. Of the five genes encoding subfamily 1 receptors in *K. flaccidum*, Kf100524 (named from the scaffold sequence in which it is present) clusters closest to these *P. patens* genes. The *K. flaccidum* EST HO437642 corresponds to Kf100524 based on sequence identity. One of the *S. pratensis* ESTs (GW599340) and the *P. margaritaceum* EST (HO61064) also cluster near to these *P. patens* genes (Fig. 3). *K. flaccidum* Kf100524,

Kf100196, and Kf100708 contain the conserved motifs of the histidine kinase and receiver domains and Kf100524 contains four of the five introns characteristic of subfamily 1 receptors in higher plants as confirmed by comparison with *K. flaccidum* in higher plants as confirmed by comparison with *K. flaccidum* Kf100196 but predicted intron positions include three of the five characteristic subfamily 1 receptor introns (Fig. 5). *K. flaccidum*

Fig. 3 Phylogenetic analysis of ethylene receptors. A phylogenetic tree of subfamily 1 ethylene receptors was constructed using the maximum-likelihood method. The tree with the highest log likelihood is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Numbers on each branch denote percentages of bootstrap support. Only partial sequences were available for *S. pratensis* and *P. margaritaceum* and two of the *P. abies* genes

Kf100708 is unique in that it contains no introns (Fig. 5). Only sequence for the N-terminal transmembrane domain of *K. flaccidum* Kf100794 is available but it shares most similarity with Kf100708 (data not shown).

The remaining four *P. patens* receptors contain multiple introns, ranging from 5 to 9 introns (Fig. 6) that are unlike the other four receptors in this species, suggesting that these four genes are distinct. Of these four receptors, two contain five introns present in the transmembrane, GAF, histidine kinase and receiver domains while the other two contain 8–9 introns in these same domains, seven of which are conserved between the two genes (Fig. 6). In the latter pair, one gene has a deletion within the GAF domain. Interestingly, one intron present at the end of the histidine kinase domain in all four distinct receptor genes is identical in position to the single intron present in subfamily 2 receptors such as EIN4, ETR2, and ERS2 of *A. thaliana* (Fig. 6). As this intron is not present in subfamily 1 receptors (Fig. 5), this suggests that the intron characteristic of subfamily 2 receptors had appeared by the evolution of *P. patens*. However, these four distinct receptor genes also share one intron that is present at the end of the GAF domain of subfamily 1 receptors although they lack the other subfamily 1 introns. All four encoded proteins contain the conserved motifs of the histidine kinase and receiver domains, unlike subfamily 2 receptors, and one of these four distinct *P. patens* receptors contains a predicted fourth transmembrane domain characteristic of subfamily 2 receptors (Fig. 6). As these four distinct receptors contain the ancestral intron conserved among subfamily 2 receptor genes but also retain an intron and the conserved motifs characteristic of subfamily 1 receptors, it is likely that these *P. patens* receptors represent a transitional state from subfamily 1 to subfamily 2 receptors where the intron structure has diverged substantially from subfamily 1 receptors to a new structure that will eventually become that of subfamily 2 receptors while the amino acid sequence retains the defining features of subfamily 1 receptors. The intron structures of these four proto-subfamily 2 receptors also form two subfamilies, suggesting that each gene pair resulted from gene duplication. One of the predicted introns of *K. flaccidum* Kf100385 also corresponds to this intron position characteristic of subfamily 2 receptor genes, suggesting the intron had appeared in charophytes (Fig. 6). An assembled



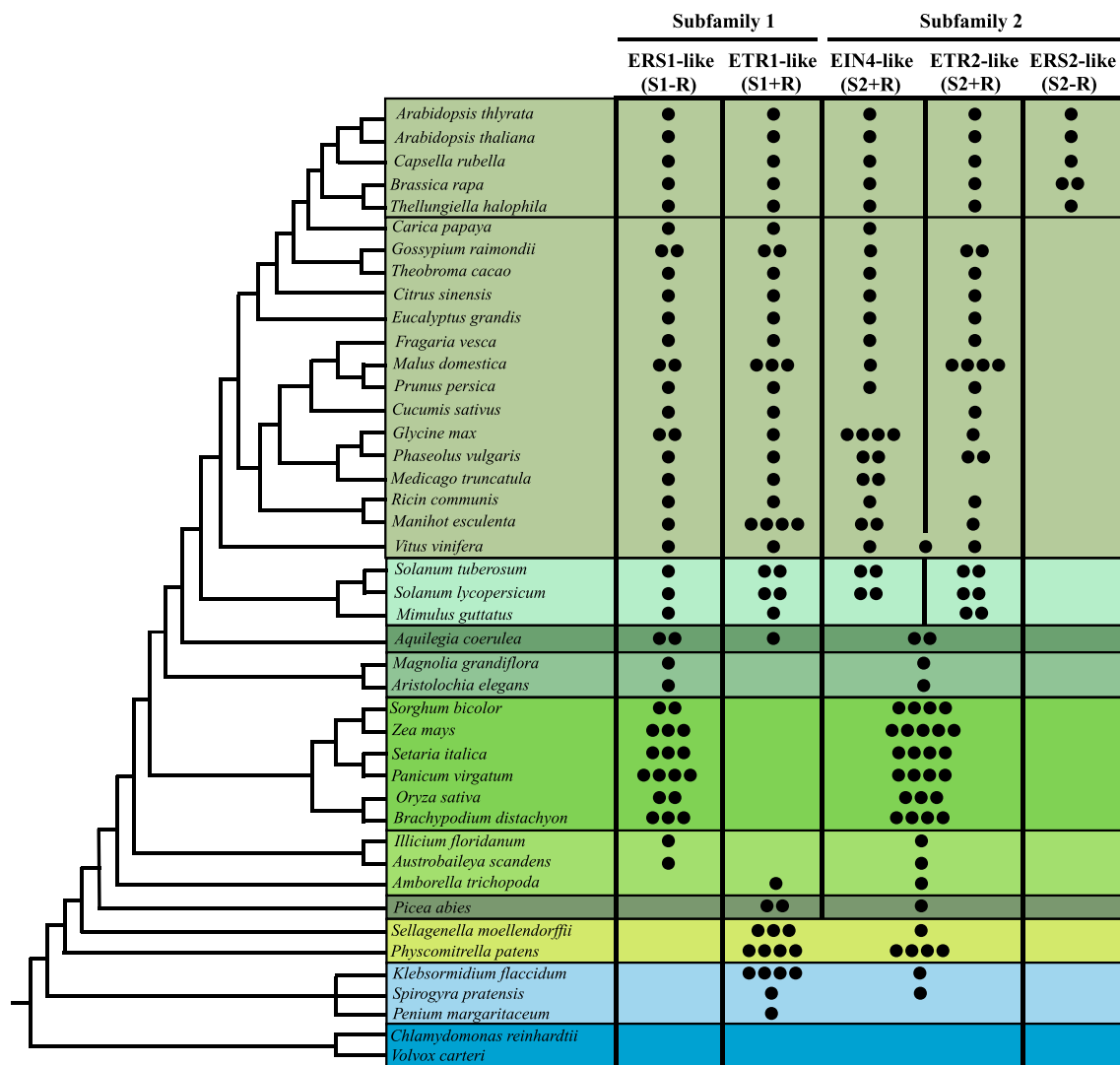


Fig. 4 Appearance and expansion of the ethylene receptor gene family during plant evolution. The presence of the types of ethylene receptor genes in a species is indicated by a *black dot* and the number of genes for each type is indicated by the number of *black dots*. The evolutionary relationship of the species shown is indicated to the *left* of the table. S2+R (EIN4/ETR2-like) receptor genes appeared first as an undifferentiated type of receptor as indicated by the single column in early plants which later differentiated into distinct EIN4 and ETR2 receptor genes. One gene in *V. vinefera* is EIN4/ETR2-like as indicated by the *black dot* straddling the two columns. Note that the subfamily 2 ethylene receptor genes indicated for *P. patens*, *S. moellendorffii*, *K. flaccidum*, and *S. pratensis* should be considered proto-subfamily 2 ethylene receptors as they retain the conserved motifs of the histidine kinase domain characteristic of subfamily 1 receptors. However, in the case of *P. patens*, *S. moellendorffii*, and *K. flaccidum*, the proto-subfamily 2 ethylene receptor genes contain an intron characteristic of subfamily 2 receptors, whereas Ju et al. (2015) concluded that one of the two assembled *S. pratensis* transcripts encoded a protein more similar to subfamily 2 receptors. The two *S. pratensis* gene members are based on transcript assemblies (Ju et al. 2015) and may not reflect the true size of the gene family

transcript from *S. pratensis* shows greater similarity to subfamily 2 receptors than to subfamily 1 receptors (Ju et al. 2015), supporting the notion that proto-subfamily 2 receptors began to evolve in charophytes. *K. flaccidum* Kf100385 contains an intron just upstream of the first conserved aspartic acid residue of the receiver domain that is conserved in *P. patens* genes 001G117000 and 010G058200 (but not in *P. patens* 004G103600 or 004G103700) and in the lycophyte *S. moellendorffii* gene 84824 (Fig. 6),

indicating that these *P. patens* and *S. moellendorffii* genes evolved from the same ancestral gene from which *K. flaccidum* Kf100385 arose. These data demonstrate that ethylene receptors arose prior to the appearance of land plants and the gene family had expanded and diverged by the appearance of charophytes to express subfamily 1 and proto-subfamily 2 receptors.

Although the ethylene receptor gene family is surprisingly large in *P. patens*, by the appearance of *S.*

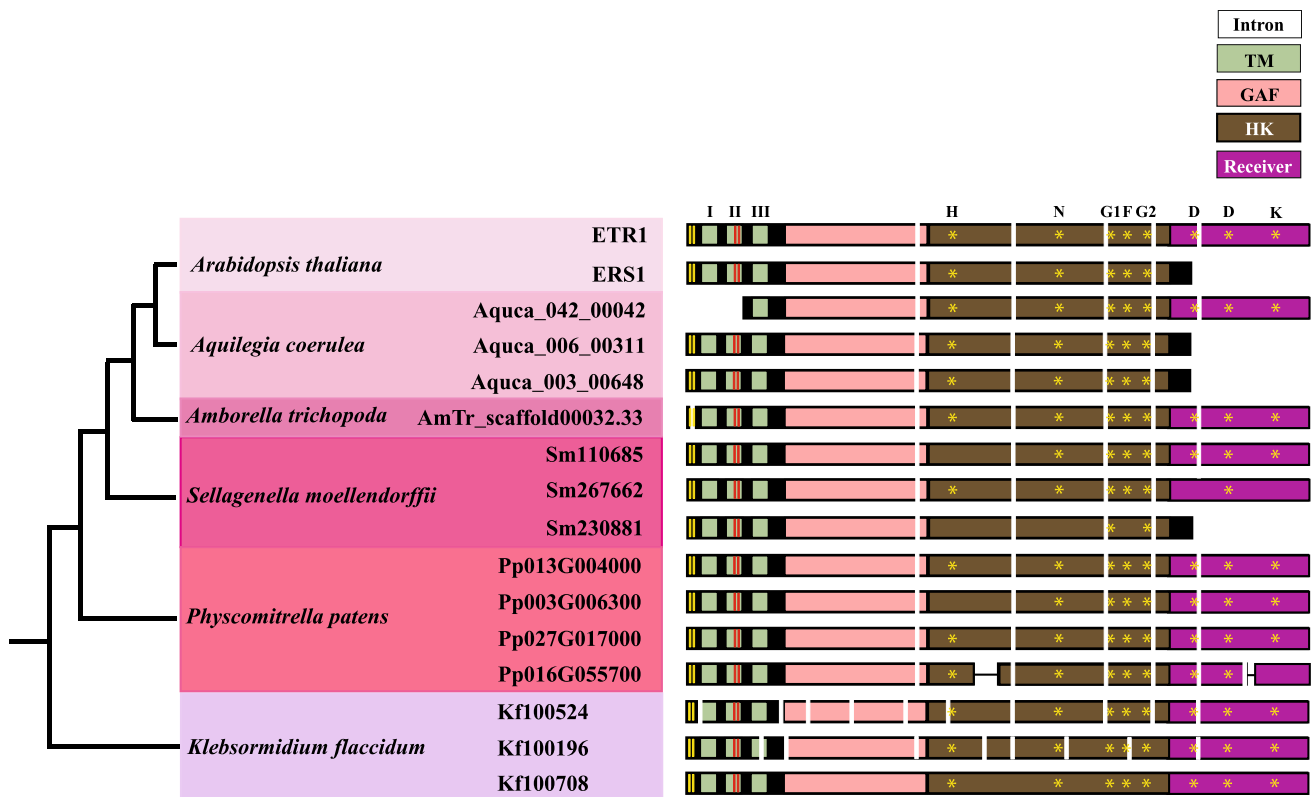


Fig. 5 Early evolution of subfamily 1 ethylene receptor genes. Comparison of the domains and intron positions of subfamily 1 ethylene receptor genes with *A. thaliana* ETR1 and ERS1. The transmembrane, GAF, histidine kinase, and receiver domains are indicated according to the key at the upper right. The N-proximal cysteine residues involved in receptor dimerization are indicated by vertical yellow bars and the cysteine and histidine residues in transmembrane domain II involved in ethylene binding are indicated by vertical red

bars. The consensus motifs of the histidine kinase domain are indicated with yellow asterisks as are the conserved aspartate and lysine residues in the receiver domain. The position of each intron is indicated by vertical white bars. The evolutionary relationship of the species shown is indicated to the left. The sequence for *A. coerulea* gene 042_00042 may be misannotated as the predicted sequence lacks the ethylene binding site and initiates just downstream of the second transmembrane domain

moellendorffii, the gene family was reduced to just four members composed of three subfamily 1 genes and a single proto-subfamily 2 receptor (Figs. 3, 4, 7). Of the three subfamily 1 receptors, one lacked the histidine residue within the H motif, the second lacked one of the two conserved aspartic acid residues and the lysine residues of the receiver domain, and the third lacked the histidine residue of the H motif, the asparagine residue of the N motif, the F motif, and the receiver domain (Fig. 5). Although the lack of a receiver domain might suggest that this establishes the first appearance of an S1–R receptor similar to ERS1, the lack of the aforementioned conserved motifs is inconsistent with an ERS1 receptor. Moreover, this gene is not present in the subsequent proximal evolved species (see below), suggesting it was lost during further evolution. Only the first of the *S. moellendorffii* subfamily 1 genes contains the same introns present in *P. patens* and *A. thaliana* ETR1 genes (Fig. 5). The other two *S. moellendorffii* subfamily 1 genes are missing the intron in the receiver domain. The single *S. moellendorffii* proto-subfamily 2 receptor (84824) contains

the conserved motifs of the histidine kinase and receiver domains (Fig. 6) like the proto-subfamily 2 receptors of *P. patens* with which it clusters in the phylogenetic analysis (Fig. 3). This *S. moellendorffii* proto-subfamily 2 receptor also shares seven of the introns present in two of the *P. patens* proto-subfamily 2 receptor genes (001G117000 and 010G058200), including the intron present at the end of the histidine kinase domain characteristic of subfamily 2 receptors. This suggests that this receptor gene descended from one of these *P. patens* proto-subfamily 2 receptor genes and that the other two-member *P. patens* gene subfamily (004G103600 or 004G103700) was lost or may have appeared during subsequent *P. patens* evolution.

Subfamily 1 and 2 ethylene receptors appear in gymnosperms

Three ethylene receptor genes are present in the gymnosperm species *P. abies* for which the genome sequence was recently reported (Nystedt et al. 2013). The three genes

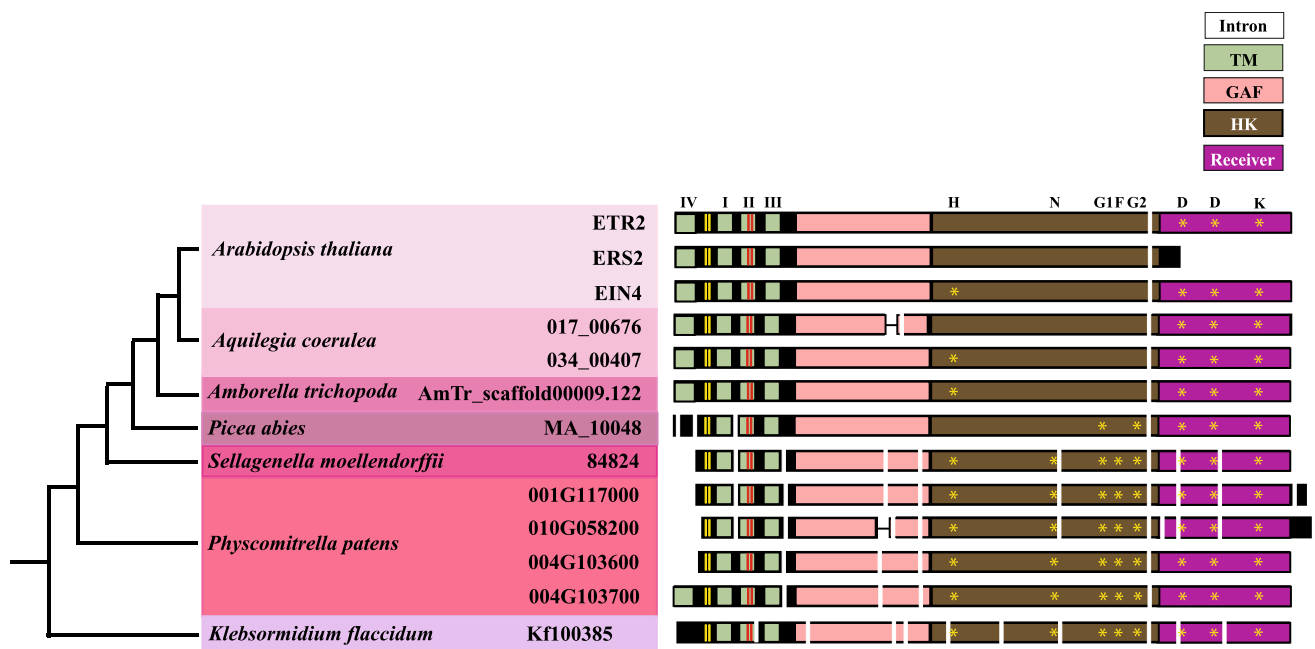


Fig. 6 Early evolution of subfamily 2 ethylene receptor genes. Comparison of the domains and intron positions of subfamily 2 ethylene receptor genes with *A. thaliana* ETR2, ERS2, and EIN4. The transmembrane, GAF, histidine kinase, and receiver domains are indicated according to the key at the upper right. The N-proximal cysteine residues involved in receptor dimerization are indicated by vertical yellow bars and the cysteine and histidine residues in transmembrane

domain II involved in ethylene binding are indicated by vertical red bars. The consensus motifs of the histidine kinase domain are indicated with yellow asterisks as are the conserved aspartate and lysine residues in the receiver domain. The position of each intron is indicated by vertical white bars. The evolutionary relationship of the species shown is indicated to the left

include two subfamily 1 (S1+R) receptors and one subfamily 2 (S2+R) receptor (Figs. 3, 4, 7). The sequence of the two subfamily 1 receptor genes is incomplete but both genes contain the conserved motifs of the histidine kinase and receiver domains and the introns present are conserved with *A. thaliana* ETR1 receptor genes (data not shown). That both are likely ETR1 homologs is supported by the presence of a receiver domain in one case (MA_67531p0010) and the observation that no S1–R receptor gene is present in the genome of *A. trichopoda*, the common ancestor of all extant flowering plants (Amborella Genome Project 2013).

The *P. abies* subfamily 2 receptor gene (MA_10048), for which the sequence is complete, lacks the conserved motifs of the histidine kinase domain but contains the conserved aspartic acid and lysine residues in its receiver domain (Fig. 6). It also lacks most of the introns present in the *P. patens* and *S. moellendorffii* proto-subfamily 2 receptor genes but retains the intron between transmembrane region 1 and II that is present in the *P. patens* proto-subfamily 2 receptor gene subfamily (001G117000 and 010G058200) and the single *S. moellendorffii* proto-subfamily 2 receptor gene (84,824) as well as the intron at the end of the histidine kinase domain that is characteristic of subfamily 2 genes (Fig. 6). The *P. abies* subfamily 2 receptor gene has

also acquired two additional introns on either side of what is the region corresponding to transmembrane region IV, although this sequence may not function as a transmembrane region as it contains several basic residues. The two additional introns are not present in subfamily 2 genes of other species, suggesting these are gymnosperm-specific or unique to *P. abies*. The retention of the intron between transmembrane region 1 and II in the *P. abies* subfamily 2 receptor gene suggests that it evolved from the single *S. moellendorffii* proto-subfamily 2 receptor gene (84824) which in turn likely evolved from the *P. patens* proto-subfamily 2 receptor gene subfamily (001G117000 and 010G058200) which contains this same intron. During its evolution, however, the *P. abies* subfamily 2 receptor gene lost several of the introns present in the homologs of non-seed plants and it has lost three of the five conserved motifs in the histidine kinase domain (Fig. 6), suggesting that it is in transition to the final protein and gene structure observed for EIN4 and ETR2 receptor genes in higher plants.

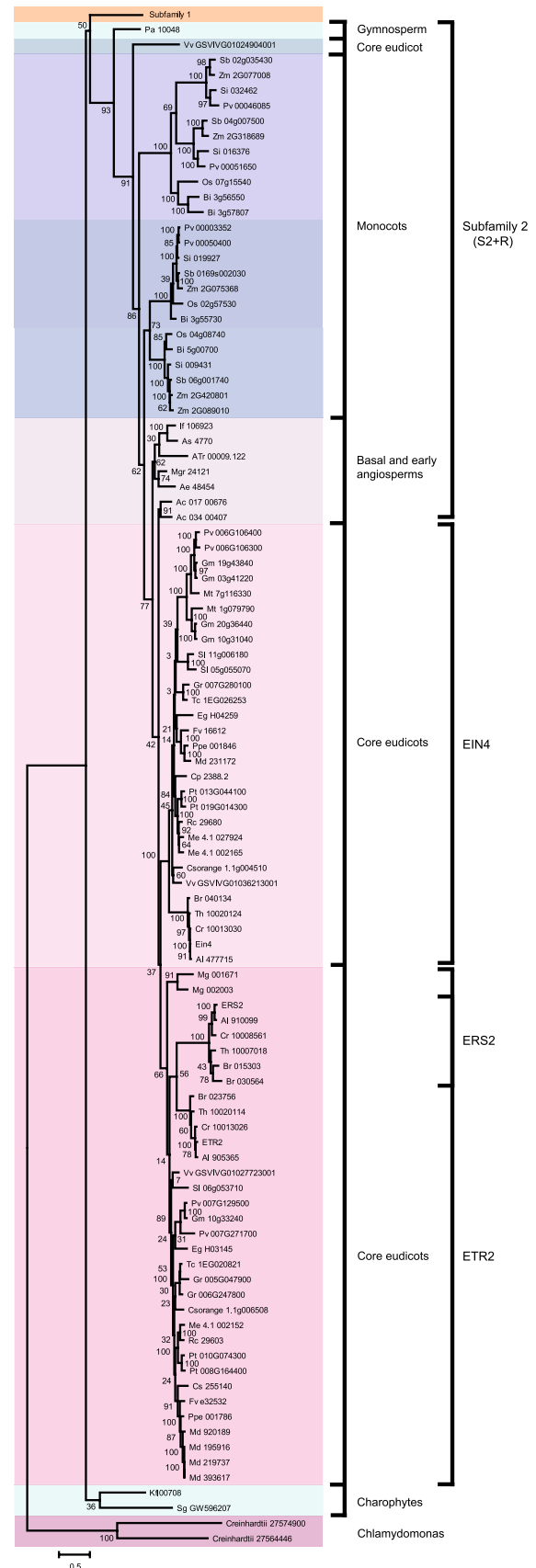
The genome sequence of *A. trichopoda*, which predates angiosperm diversification and therefore is the most basal angiosperm, was recently reported (Amborella Genome Project 2013). The *A. trichopoda* receptor gene family is composed of one gene encoding a subfamily 1 (S1+R) receptor and a second gene encoding a subfamily 2 (S2+R)

Fig. 7 A phylogenetic tree of subfamily 2 ethylene receptors was constructed using the maximum-likelihood method. The sequence representing the receiver domain was omitted for the analysis. The tree with the highest log likelihood is shown. The tree is drawn to scale with branch lengths measured in the number of substitutions per site. Numbers on each branch denote percentages of bootstrap support

receptor (Figs. 3, 4, 7). The intron structure of the *A. trichopoda* subfamily 1 receptor is identical to *A. thaliana* ETR1 (Fig. 5) while the intron structure of the *A. trichopoda* subfamily 2 receptor is identical to *A. thaliana* EIN4 and ETR2 (Fig. 6). Although the phylogenetic analysis places the *A. trichopoda* subfamily 2 receptor outside the EIN4 and ETR2 clades suggesting it is the likely progenitor to both receptor types (Fig. 7), it retains the histidine in the H motif in the histidine kinase domain which suggests that it is more EIN4-like than ETR2-like in this respect.

The appearance of ERS1-like receptors accompanies the loss of ETR1-like receptors in some early angiosperms and in monocots

Despite the fact that one of the three subfamily 1 receptors in *S. moellendorffii* lacks a receiver domain which would render it ERS1-like, the absence of an ERS1 homolog in *A. trichopoda*, and presumably in *P. abies*, suggests that an S1–R receptor had not yet evolved in a stable manner by the appearance of the basal most angiosperm species. Although the genome sequence of subsequent basal angiosperm species, such as *A. scandens* and *I. floridanum*, is not available, ESTs from these species encoding an S1–R receptor in addition to an S2+R receptor indicate that an S1–R receptor had appeared at this point and it remained present throughout subsequent angiosperm evolution (Figs. 3, 4). In general, the S1–R receptor C-terminal sequence is poorly conserved in species that express this receptor and is shorter in the S1–R receptors of *A. scandens* and *I. floridanum* than in species that evolved later (Fig. 8). No EST encoding an S1+R (ETR1-like) receptor was identified in *A. scandens* and *I. floridanu* although no definitive conclusion can be made until the genome sequence is available for these species. If these species do lack an ETR1 homolog, it is possible that the S1–R receptors of these species resulted from a deletion of the C-terminal receiver domain of the S1+R receptor which would account for the simultaneous disappearance of an S1+R receptor gene and the appearance of an S1–R receptor gene. The fact that ETR1 receptor genes are present in later angiosperm evolution suggests that an S1+R receptor gene was maintained in the common ancestral line and that



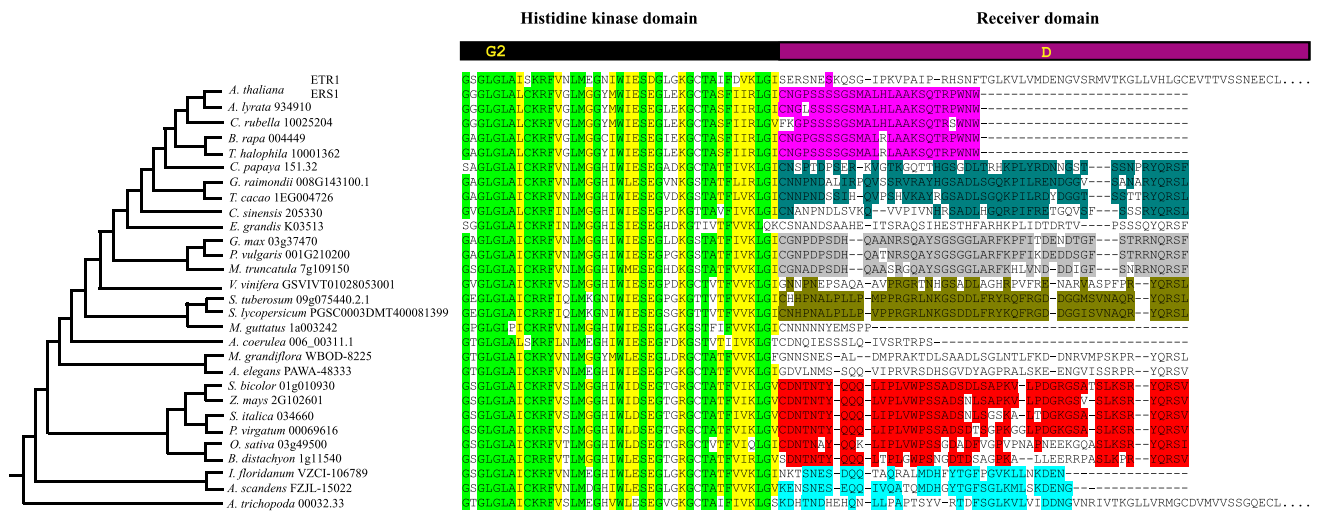


Fig. 8 C-Terminal sequence of ERS1 receptors homologs during plant evolution. The aligned C-terminal ERS1 sequence of the species indicated is shown. Sequence of the histidine kinase domain C-terminal region is included with amino acid identity relative to *A. thaliana* ERS1 highlighted in green and amino acid similarity high-

lighted in yellow. Homology of sequence downstream of the histidine kinase domain is colored according to plant group. The corresponding sequences of *A. thaliana* and *A. trichopoda* ETR1 proteins were included for comparison. The evolutionary relationship of the species shown is indicated to the left

deletion of the C-terminal receiver domain from the S1+R receptor gene may have occurred during the subsequent evolution of *A. scandens* and *I. floridanum*. Such divergence in sequence and length may indicate that this region is not important to S1–R receptor function.

Similar to *A. scandens* and *I. floridanum*, monocot genomes lack an S1+R receptor and contain only S1–R and S2+R receptor genes (Figs. 3, 4, 7). Thus, monocots as a group appear to be unique in that they lack an ETR1 homolog. Expansion of the gene family encoding each receptor type is observed in most monocot species. Phylogenetic analysis of the monocot subfamily 1 receptors indicates that they form two subclades (Fig. 3), suggesting gene duplication early in monocot evolution. The monocot subfamily 1 receptors cluster with eudicot ERS1 receptors.

To establish that the subfamily 1 receptors in monocots are structurally more similar to ERS1 than to ETR1, the monocot subfamily 1 receptors were examined to determine whether they lacked a receiver domain. All monocot subfamily 1 receptors are similar to Arabidopsis ERS1 in that they contain an N-terminal domain composed of three transmembrane spanning regions, followed by a GAF domain and a histidine kinase domain that possesses the amino acid sequences and motifs required for histidine kinase activity, but lack a C-terminal receiver domain although there was some variability in the length of the C-terminal end (Fig. 9). One receptor in *O. sativa* (Os05g06320) is truncated but this may be a misannotation as the predicted protein terminates at the splice site of the conserved intron that lies just upstream of the G1 motif in the histidine kinase domain. For some monocot subfamily

1 receptors, the first aspartic acid residue of the receiver domain is retained whereas in others, it too is absent. All monocot subfamily 1 receptor genes (except Os05g06320) contain the first four introns present in *A. thaliana* ERS1 and ETR1 but do not contain the intron present in the *A. thaliana* ETR1 receiver domain (Fig. 9). Two N-terminal cysteines required for homodimerization and the cysteine and histidine residues in transmembrane domain III that bind Cu (I) needed for ethylene binding are also present in all monocot subfamily 1 receptors. Based on their gene structure and the phylogenetic analysis, monocot subfamily 1 receptors are homologs to ERS1.

Although the subfamily 2 receptors of the basal and early angiosperms *A. trichopoda*, *A. scandens*, and *I. floridanum* lie outside the EIN4 and ETR2 clades, they contain a histidine in the H motif in their histidine kinase domain, suggesting that the early subfamily 2 receptor may be more EIN4-like. Monocot subfamily 2 receptors also lie outside the EIN4 and ETR2 clades consistent with the observation that they predate the appearance of distinct EIN4 and ETR2 receptors. Monocot subfamily 2 receptors are grouped into subclades (Fig. 7), suggesting gene duplication prior to speciation. Two subclades contain a fourth transmembrane domain characteristic of subfamily 2 receptors while the third group lacks this domain (Fig. 10). One of the two subclades containing a fourth transmembrane domain lacks the second aspartic acid residue in the receiver domain, supporting the notion that gene duplication occurred prior to speciation. Almost all contain the two N-terminal cysteines required for homodimerization and the cysteine and histidine residues in transmembrane

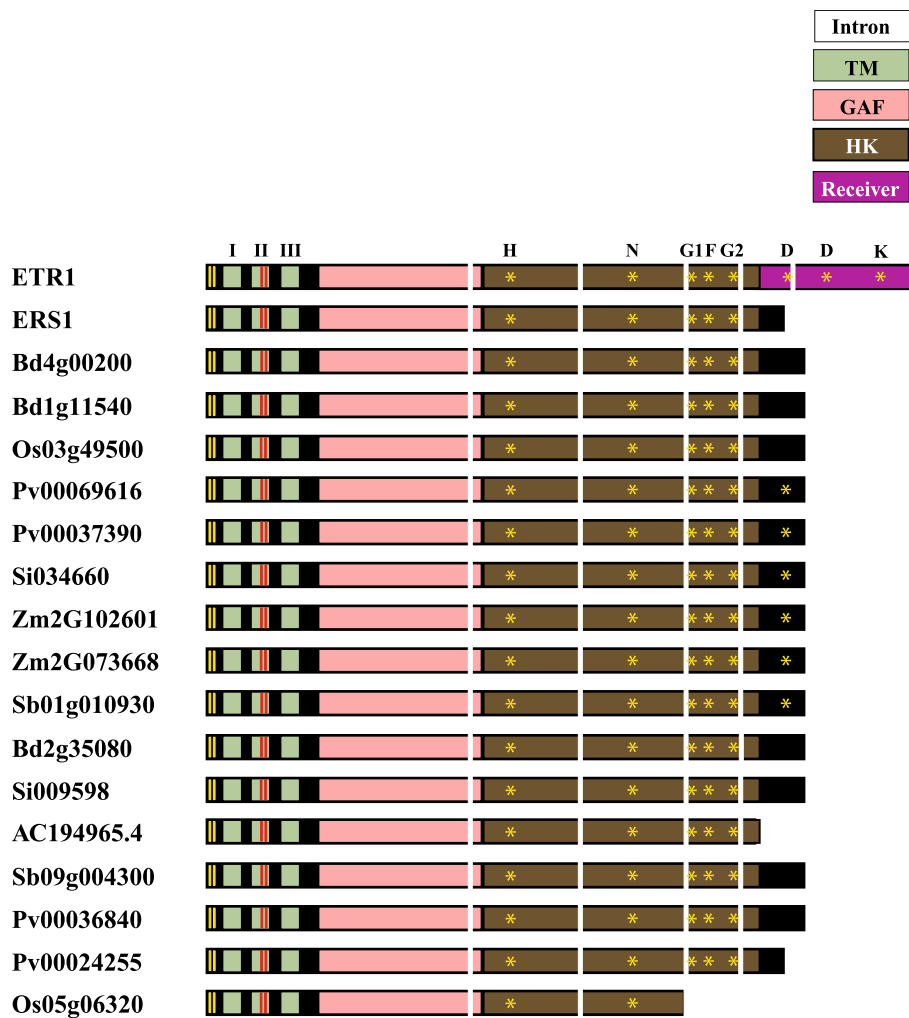


Fig. 9 Monocot subfamily 1 ethylene receptors lack a receiver domain. Comparison of the domains and introns positions of monocot subfamily 1 ethylene receptors with *A. thaliana* ETR1 and ERS1. The transmembrane, GAF, histidine kinase, and receiver domains are indicated according to the key at the upper right. The N-proximal cysteine residues involved in receptor dimerization are indicated by vertical yellow bars and the cysteine and histidine residues in transmembrane domain II involved in ethylene binding are indicated by vertical red bars. The consensus motifs of the histidine kinase

domain are indicated with yellow asterisks as are the conserved aspartate and lysine residues in the receiver domain. The position of each intron is indicated by vertical white bars. Monocot subfamily 1 ethylene receptor genes included are: *S. bicolor* (Sb_09g004300; Sb_01g010930); *B. distachyon* (Bi_2g35080; Bi_4g00200; Bi_1g11540); *Z. mays* (Zm_AC194965; Zm_2G102601; Zm_2G073668); *Oryza sativa* (Os05g06320; Os_03g49500); *P. virgatum* (Pv_00024255; Pv_00036840; Pv_00069616; Pv_00037390); and *S. italica* (Si_009598; Si_034660)

domain III for ethylene binding. With only one exception, all monocot subfamily 2 receptors lack a histidine in the H motif of the histidine kinase domain that is present in EIN4 but not ERS2 of *A. thaliana* (Fig. 10). The presence of the intron characteristic of subfamily 2 receptor genes, the presence of a receiver domain, and the absence of most or all of the conserved motifs of the histidine kinase domain confirmed that they are S2+R receptors. Four of the monocot subfamily 2 receptors have deletions although in the case of Zm2G077008, the absence of a receiver domain may be a misannotation at the splice site of the conserved intron normally near this position. Only one other gene

(Sb0169s002030) lacks the single intron characteristic of subfamily 2 receptors and this gene harbors two deletions. Whether any of these deletions affect receptor function is unknown.

From the available ESTs, the early angiosperm species such as *A. elegans* and *M. grandiflora* appear to be similar to *A. scandens* and *I. floridanum* in that they lack an S1+R receptor gene but, as with *A. scandens* and *I. floridanum*, no definitive conclusion can be made until the genome sequence is available for these species. *A. elegans* and *M. grandiflora*, however, do contain S1–R and S2+R receptor genes (Figs. 3, 4, 7). The presence of an S1–R receptor

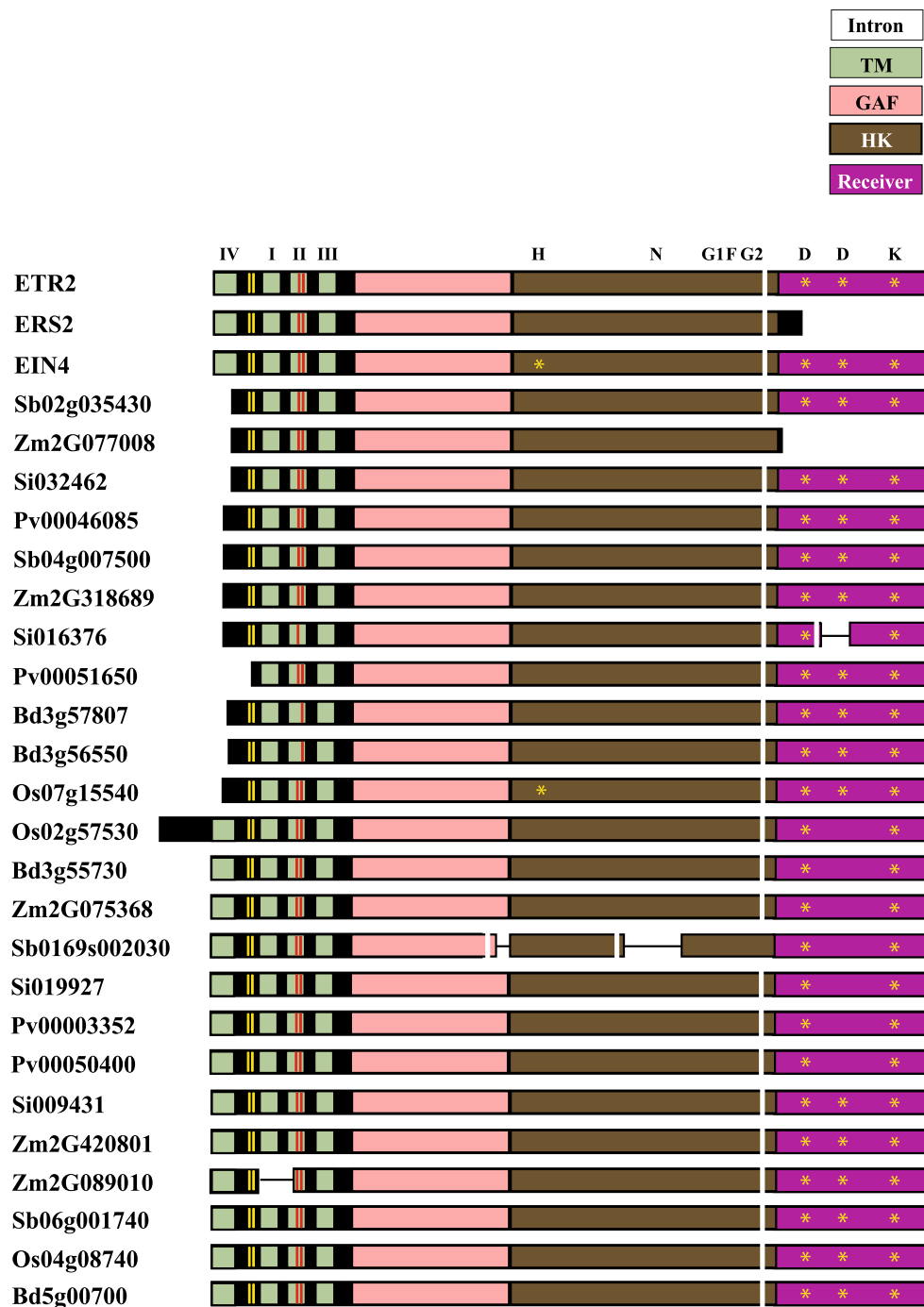


Fig. 10 Monocot subfamily 2 ethylene receptors contain a receiver domain. Comparison of the domains and introns positions of monocot subfamily 2 ethylene receptors with *A. thaliana* ETR2, ERS2, and EIN4. The transmembrane, GAF, histidine kinase, and receiver domains are indicated according to the key at the upper right. The N-proximal cysteine residues involved in receptor dimerization are indicated by vertical yellow bars and the cysteine and histidine residues in transmembrane domain II involved in ethylene binding are indicated by vertical red bars. The consensus motifs of the histidine kinase domain are indicated with yellow asterisks as are the conserved aspartate and lysine residues in the receiver domain. The

position of each intron is indicated by vertical white bars. Deletions in a receptor are indicated with a line. Monocot subfamily 1 ethylene receptor genes included are: *S. bicolor* (Sb_02g035430; Sb_04g007500; Sb_0169s002030; Sb_06g001740); *B. distachyon* (Bi_3g56550; Bi_3g57807; Bi_3g55730; Bi_5g00700); *Z. mays* (Zm_2G077008; Zm_2G318689; Zm_2G075368; Zm_2G420801; Zm_2G089010); *Oryza sativa* (Os_07g15540; Os_02g57530; Os_04g08740); *P. virgatum* (Pv_00046085; Pv_00051650; Pv_00003352; Pv_00050400); and *S. italica* (Si_032462; Si_016376; Si_019927; Si_009431)

was confirmed by the lack of a receiver domain but the C-terminal sequence is not conserved with S1–R receptors in other angiosperms (Fig. 8). This could indicate that, as in *A. scandens* and *I. floridanum*, the S1–R receptors in *A. elegans* and *M. grandiflora* resulted from a deletion of the C-terminal receiver domain of the S1+R receptor during their speciation. This possibility is supported by the fact that the *Aquilegia caerulea* genome contains an S1+R receptor gene as well as two S1–R receptor genes and two S2+R receptor genes (Figs. 3, 4, 7). Although it is possible that the S1+R receptor gene was lost from species following the appearance of *A. trichopoda* only to be regained following the appearance of *A. elegans* and *M. grandiflora*, it is more likely that an S1+R receptor gene was present in the common ancestral line and that S1–R receptor genes arose independently in *S. moellendorffii*, in some basal angiosperms such as *A. scandens* and *I. floridanum*, in the monocot progenitor, and in early angiosperm species such as *A. elegans* and *M. grandiflora*. The conservation of sequence and variable length of the C-terminus of the S1–R receptors in these species supports such a possibility (Fig. 8). The conservation of sequence and the more uniform length of the C-terminus of S1–R receptors in core eudicots suggest that the S1–R receptor gene was stably maintained during their evolution. A notable exception to this is the Brassicaceae in which the length of the C-terminus of the ERS1 receptor is shorter than in other core eudicots (Fig. 8). The presence of introns characteristic of subfamily 1 receptor genes, the absence of a receiver domain, and the presence of the conserved motifs of the histidine kinase domain confirmed that the members included in this clade are ERS1 homologs.

EIN4 and ETR2 homologs appear as distinct receptors in core eudicots

By the appearance of asterids such as *Mimulus guttatus* and *Solanum* species, EIN4-like and ETR2-like receptors had diverged sufficiently to cluster in separate clades and each had undergone gene duplication (Figs. 4, 7). Similarly in rosids, EIN4-like and ETR2-like receptors cluster in separate clades. The genome of *Vitis vinifera* contains an EIN4 homolog and an ETR2 homolog and a third that is equally related to both. The remaining rosid species examined contain genes encoding ETR1, ERS1, EIN4, and ETR2 homologs, although *C. papaya* and *M. truncatula* lack an ETR2 homolog and *Cucumis sativa* lacks an EIN4 homolog (Fig. 4). The number of genes encoding each receptor type varies among these species, suggesting that gene duplication (or loss) likely occurred following speciation. The presence of the intron characteristic of subfamily 2 receptor genes, the presence of a receiver domain, and the absence of most or all of the conserved motifs of the

histidine kinase domain confirmed that they are EIN4 and ETR2 homologs.

ERS2 homologs appear during evolution of the Brassicaceae

ERS2 is the fifth type of ethylene receptor present in *A. thaliana* which is similar to ETR2 in that it lacks the histidine residue in the H motif in the histidine kinase domain but differs in that it lacks a receiver domain. ERS2 homologs are confined to the Brassicaceae and therefore is the most recent receptor type to have evolved. Genes for ERS2 are present in *A. thaliana*, *A. lyrata*, *C. rubella*, *T. halophila* and *B. rapa* (Figs. 4, 7). Phylogenetic analysis suggests that ERS2 is most related to ETR2 as it clusters with the ETR2 clade of Brassicaceae. The presence of the intron characteristic of subfamily 2 receptor genes, the absence of a receiver domain, and the absence of the conserved motifs of the histidine kinase domain confirmed that they are ERS2 homologs.

Discussion

The results presented here demonstrate that an ETR1-like receptor was the first ethylene receptor type to have evolved during plant evolution. The apparent absence of an ethylene receptor homolog in the Chlamydomonadales but their presence in charophycean algae suggests that the ancestral ETR1-like ethylene receptor had evolved by the appearance of fresh water algal species and may indicate that ethylene signaling arose in response to its particular environment. As some fresh water algal species can grow in moist soils, ethylene receptors may have evolved to respond to the abiotic stresses associated with a terrestrial environment such as desiccation or greater exposure to UV radiation.

At least five ETR1-like receptor genes were identified in the *K. flaccidum* genome which includes two genes containing introns characteristic of subfamily 1 receptors, one gene containing the intron characteristic of subfamily 2 receptors, and one gene without introns. The presence of homologs for CTR1, EIN3, and EBF1 in *K. flaccidum* (Hori et al. 2014) also supports the notion that these ethylene receptors are functional and that ethylene signaling likely occurs in this species.

Homologs for ethylene biosynthetic genes encoding ACS and ACO as well as homologs for ethylene signaling genes encoding ethylene receptors, CTR1, EIN2, EIN3, and ERF1 are present in *S. pratensis* (Ju et al. 2015). The demonstration that this species produces and responds to ethylene and that *S. pratensis* ETR1-like and EIN3-like homologs can complement ethylene receptor and *ein3* mutants, respectively, in *A. thaliana* (Ju et al. 2015)

indicates that the ethylene biosynthesis and signaling pathways are functional in this species. That *K. flaccidum* (representing an earlier plant lineage than *S. pratensis*) lacks an ERF1 homolog and its EIN2 homolog may lack the C-terminal signaling domain while *M. viride* (representing the earliest charophyte lineage) appears to lack ethylene receptors altogether (Ju et al. 2015), may indicate that the components of the ethylene signaling pathway were actively evolving during charophycean evolution.

The ethylene receptor gene family had expanded to eight ETR1-like members by the appearance of the non-vascular plant *P. patens*, which includes four genes with introns characteristic of subfamily 1 receptors and four genes whose intron structure is unlike subfamily 1 genes but contains the intron characteristic of subfamily 2 receptors. This gene group may have evolved from a duplication of the canonical subfamily 1 receptor gene because, although they possess an intron structure quite distinct from the subfamily 1 receptor genes, they share the last intron in the GAF domain. They also may have evolved from *K. flaccidum* Kf100385 as they share the intron characteristic of subfamily 2 receptor genes. Because they retain the protein motifs of a subfamily 1 (ETR1-like) receptor, they may exhibit subfamily 1 receptor kinase activity but the phylogenetic analysis indicates they lie more between the subfamily 1 and subfamily 2 receptor clades than do the other four *P. patens* receptor genes. Thus, *K. flaccidum* Kf100385 and this *P. patens* group of genes may have begun to diverge from subfamily 1 genes as the first step in the evolution of subfamily 2 receptors and as such are referred to as proto-subfamily 2 receptor genes. The presence of four proto-subfamily 2 receptor genes that form two subfamilies in *P. patens* supports the notion that this gene group is functionally important in early land plants although whether they differ functionally remains unknown. However, these proto-subfamily 2 receptors retain the conserved motifs of the histidine kinase domain characteristic of subfamily 1 receptors, suggesting that they retain more subfamily 1-like function than subfamily 2-like function. Based on limited phylogenetic analysis, one of the two ethylene receptors identified recently in *S. pratensis* was reported to be a subfamily 2 receptor (Ju et al. 2015). However, it retains the conserved motifs of the histidine kinase domain characteristic of subfamily 1 receptors, so that it, like *K. flaccidum* Kf100385, should be considered a proto-subfamily 2 receptor until proven to possess *bona fide* subfamily 2 receptor function.

The features of the proto-subfamily 2 receptor genes in *P. patens*, i.e., their distinct intron structure while retaining the conserved motifs of the histidine kinase domain characteristic of subfamily 1 receptors, are retained in one of the four ethylene receptor genes of *S. moellendorffii*, suggesting that one of the four proto-subfamily 2 receptor genes of

P. patens was retained during the evolution of *S. moellendorffii* without substantial changes in its gene structure or amino acid sequence which is supported by the phylogenetic analysis. The intron structure of the proto-subfamily 2 receptor genes of *P. patens* and *S. moellendorffii* had begun to disappear by the evolution of gymnosperms, such as *P. abies*, in which its subfamily 2 receptor gene retains only an intron between the transmembrane domains I and II that is present in the *S. moellendorffii* and *P. patens* proto-subfamily 2 receptor genes, and retains only the G1 and G2 motifs of the histidine kinase domain. Even these residual features were lost in the subfamily 2 receptor gene of *A. trichopoda*, the basal most angiosperm species. Subfamily 2 (S2+R) receptors continued as an undifferentiated group throughout monocots and through the evolution of early angiosperms species. EIN4 and ETR2 only diverged into distinct receptors with the evolution of the core eudicots. The presence of one undifferentiated subfamily 2 member in *V. vinifera* in addition to distinct EIN4 and ETR2 homologs may indicate when this divergence was occurring.

ERS2 homologs are found only in the Brassicaceae and as such represents the newest evolved member of the ethylene receptor gene family. It likely arose early in the evolution of the Brassicaceae as all species examined contain an ERS2 homolog. The recent evolution of a fifth receptor in the Brassicaceae demonstrates that the ethylene receptor gene family continues to evolve and that the appearance of this unique ethylene receptor in the Brassicaceae must confer some advantage to species in this family.

The ERS1 receptor likely evolved from duplication of an ETR1-like gene as ERS1 is more similar to ETR1 than it is to subfamily 2 receptors and shares the same introns as the ETR1 gene except in the receiver domain which is missing from all ERS1 receptors. Although one of the three subfamily 1 receptors in *S. moellendorffii* lacks a receiver domain, which might technically make this the first appearance of an S1–R receptor, it does not contain all of the motifs of the histidine kinase domain and this gene was not maintained in subsequent evolution as no S1–R receptor is present in the *A. trichopoda* genome. Supporting this is the absence of an ERS1 homolog in *P. abies* although this represents a single gymnosperm species whose genome sequence is available. However, no ERS1 homolog was observed in the available EST databases for other gymnosperms, including *Cycas micholitzii*, *Sundacarpus amarus*, and *Gnetum montanum*. Nevertheless, just as the *S. moellendorffii* genome appears to contain an S1–R-like receptor, it is formally possible that some gymnosperm species may contain such a receptor as well, e.g., through loss of part or all of the receiver domain from an ETR1 gene. The genome sequence of many additional gymnosperm species will be required to determine whether this is the case.

An ERS1 homolog appears in *A. scandens* and *I. floridanum* and is the only subfamily 1 receptor present as it is in monocots. In the Brassicaceae, the C-terminus of the ERS1 receptor is shorter than in other rosoid species and the sequence is highly conserved but distinct from other rosoids, suggesting that the ERS1 C-terminus continues to undergo dynamic evolution. Whether these changes in the C-terminal sequence and length has consequences on ERS1 function is unknown.

Although ETR1 is considered the dominant receptor in *A. thaliana* (Hua and Meyerowitz 1998), as a group, monocot species differ substantially given the absence of an ETR1 homolog. This might suggest that ethylene signaling is less important in monocots than in those species containing an ETR1 homolog. However, ethylene is known to be involved in monocot growth and development as well as in stress responses. For example, ethylene influences the onset of the programmed cell death of the endosperm during maize kernel development (Young et al. 1997), regulates root growth in response to soil conditions such as mechanical impedance (Whalen and Feldman 1988; Gallie et al. 2009) or low oxygen that occurs during flooding (Drew et al. 1979; Drew 1997; Drew et al. 2000); and promotes maize leaf senescence, inhibits photosynthetic function, and reduces drought tolerance (Young et al. 2004).

Previous work identified four receptor genes from maize (Gallie and Young 2004). The two subfamily 1 receptor genes identified were ZmERS1a (ERS1-14, GenBank accession AY359577) which is GRMZM2G102601 and ZmERS1b (ERS1-25, GenBank accession AY359578) which is GRMZM2G073668. The two subfamily 2 receptor genes identified were ZmETR2a (ETR2-9, GenBank accession AY359580) which is GRMZM2G420801 and ZmETR2b (ETR2-40, GenBank accession AY359581) which is GRMZM2G089010. Mutation of the cysteine residue in transmembrane domain III involved in ethylene binding in ZmERS1b or ZmETR2b to a tyrosine resulted in dominant negative mutant receptors that conferred ethylene insensitivity when expressed in *A. thaliana* (Chen and Gallie 2010) just as the same mutation does for *A. thaliana* ETR1 (Bleecker et al. 1988), demonstrating that these monocot receptors are functionally similar to their eudicot homologs. Collectively, these findings suggest that receptors in monocots regulate ethylene responses and that ethylene signaling in monocots is important despite fewer diverse receptors being present. The presence of subfamily 1 and 2 receptors throughout higher plant species, however, may indicate that both types are needed for optimal ethylene responses.

Acknowledgments This work was funded by the University of California Agricultural Experiment Station.

Conflict of interests The author declares no conflict of interests.

References

- Abeles FB, Morgan PW, Saltveit ME (1992) Ethylene in plant biology, 2nd edn. Academic Press Inc, San Diego
- Alonso JM, Hirayama T, Roman G, Nourizadeh S, Ecker JR (1999) EIN2, a bifunctional transducer of ethylene and stress responses in *Arabidopsis*. *Science* 284:2148–2152
- Amborella Genome Project (2013) The Amborella genome and the evolution of flowering plants. *Science* 342:1241089
- Binder BM, Chang C, Schaller GE (2012) Perception of ethylene by plants-ethylene receptors. *Annual plant reviews, the plant hormone ethylene* 44. Wiley, Hoboken, pp 117–145
- Bleecker AB (1999) Ethylene signaling: an evolutionary perspective. *Trends Plant Sci* 4:269–274
- Bleecker AB, Kende H (2000) Ethylene: a gaseous signal molecule in plants. *Ann Rev Cell Dev Biol* 16:1–18
- Bleecker AB, Estelle MA, Somerville C, Kende H (1988) Insensitivity to ethylene conferred by a dominant mutation in *Arabidopsis thaliana*. *Science* 241:1086–1089
- Bleecker AB, Esch JJ, Hall AE, Rodríguez FI, Binder BM (1998) The ethylene-receptor family from *Arabidopsis*: structure and function. *Philos Trans R Soc Lond B Biol Sci* 353:1405–1412
- Chang C, Shockey JA (1999) The ethylene-response pathway: signal perception to gene regulation. *Curr Opin Plant Biol* 2:352–358
- Chang C, Stadler R (2001) Ethylene hormone receptor action in *Arabidopsis*. *Bioessays* 23:619–627
- Chang C, Stewart RC (1998) The two-component system. Regulation of diverse signaling pathways in prokaryotes and eukaryotes. *Plant Physiol* 117:723–731
- Chang C, Kwok SF, Bleecker AB, Meyerowitz EM (1993) *Arabidopsis* ethylene-response gene ETR1: similarity of product to two-component regulators. *Science* 262:539–544
- Chao Q, Rothenberg M, Solano R, Roman G, Terzaghi W, Ecker JR (1997) Activation of the ethylene gas response pathway in *Arabidopsis* by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell* 89:1133–1144
- Chen QG, Bleecker AB (1995) Analysis of ethylene signal-transduction kinetics associated with seedling-growth response and chitinase induction in wild-type and mutant *Arabidopsis*. *Plant Physiol* 108:597–607
- Chen JF, Gallie DR (2010) Analysis of the functional conservation of ethylene receptors between maize and *Arabidopsis*. *Plant Mol Biol* 74:405–421
- Chen YF, Randlett MD, Findell JL, Schaller GE (2002) Localization of the ethylene receptor ETR1 to the endoplasmic reticulum of *Arabidopsis*. *J Biol Chem* 277:19861–19866
- Clark KL, Larsen PB, Wang X, Chang C (1998) Association of the *Arabidopsis* CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. *Proc Natl Acad Sci USA* 95:5401–5406
- Drew MC (1997) Oxygen deficiency and root metabolism: injury and acclimation under hypoxia and anoxia. *Annu Rev Plant Physiol Mol Biol* 48:223–250
- Drew MC, Jackson MB, Giffard S (1979) Ethylene-promoted adventitious rooting and development of cortical air spaces (aerenchyma) in roots may be adaptive responses to flooding in *Zea mays* L. *Planta* 147:83–88
- Drew MC, He C-J, Morgan PW (2000) Programmed cell death and aerenchyma formation in roots. *Trends Plant Sci* 5:123–127
- Gallie DR, Young TE (2004) The ethylene biosynthetic and perception machinery is differentially expressed during endosperm and embryo development in maize. *Mol Gen Genomics* 271:267–281

- Gallie DR, Geisler-Lee J, Chen J, Jolley B (2009) Tissue-specific expression of the ethylene biosynthetic machinery regulates root growth in maize. *Plant Mol Biol* 69:195–211
- Gamble RL, Coonfield ML, Schaller GE (1998) Histidine kinase activity of the ETR1 ethylene receptor from *Arabidopsis*. *Proc Natl Acad Sci USA* 95:7825–7829
- Gao Z, Wen CK, Binder BM, Chen YF, Chang J, Chiang YH, Kerris RJ 3rd, Chang C, Schaller GE (2008) Heteromeric interactions among ethylene receptors mediate signaling in *Arabidopsis*. *J Biol Chem* 283:23801–23810
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucl Acid Res* 31:3784–3788
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. *Nucl Acid Res* 40:D1178–D1186
- Grefen C, Städele K, Růzicka K, Obrdlík P, Harter K, Horák J (2008) Subcellular localization and in vivo interaction of the *Arabidopsis thaliana* ethylene receptor family members. *Mol Plant* 1:308–320
- Guzman P, Ecker JR (1990) Exploiting the triple response of *Arabidopsis* to identify ethylene-related mutants. *Plant Cell* 2:513–523
- Hall AE, Bleecker AB (2003) Analysis of combinatorial loss-of-function mutants in the *Arabidopsis* ethylene receptors reveals that the *ers1 etr1* double mutant has severe developmental defects that are EIN2 dependent. *Plant Cell* 15:2032–2041
- Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, Moriyama T, Ikeuchi M, Watanabe M, Wada H, Kobayashi K, Saito M, Masuda T, Sasaki-Sekimoto Y, Mashiguchi K, Awai K, Shimojima M, Masuda S, Iwai M, Nobusawa T, Narise T, Kondo S, Saito H, Sato R, Murakawa M, Ihara Y, Oshima-Yamada Y, Ohtaka K, Satoh M, Sonobe K, Ishii M, Ohtani R, Kanamori-Sato M, Honoki R, Miyazaki D, Mochizuki H, Umetsu J, Higashi K, Shibata D, Kamiya Y, Sato N, Nakamura Y, Tabata S, Ida S, Kurokawa K, Ohta H (2014) *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat Commun* 5:3978
- Hua J, Meyerowitz EM (1998) Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana*. *Cell* 94:261–271
- Hua J, Chang C, Sun Q, Meyerowitz EM (1995) Ethylene insensitivity conferred by *Arabidopsis* ERS gene. *Science* 269:1712–1714
- Hua J, Sakai H, Nourizadeh S, Chen QG, Bleecker AB, Ecker JR, Meyerowitz EM (1998) EIN4 and ERS2 are members of the putative ethylene receptor gene family in *Arabidopsis*. *Plant Cell* 10:1321–1332
- Ju C, Van de Poel B, Cooper ED, Thierer JH, Gibbons TR, Delwiche CF, Chang C (2015) Conservation of ethylene as a plant hormone over 450 million years of evolution. *Nat Plants* 1:14004
- Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR (1993) CTR1, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the raf family of protein kinases. *Cell* 72:427–441
- Klee HJ (2004) Ethylene signal transduction. Moving beyond *Arabidopsis*. *Plant Physiol* 135:660–667
- Le SQ, Gascuel O (2008) LG: an improved, general amino-acid replacement matrix. *Mol Biol Evol* 25:1307–1320
- Lin Z, Zhong S, Grierson D (2009) Recent advances in ethylene research. *J Exp Bot* 60:3311–3336
- Lohrmann J, Harter K (2002) Plant two-component signaling systems and the role of response regulators. *Plant Physiol* 128:363–369
- Mattoo AK, Suttle JC (1991) *The Plant Hormone Ethylene*. FL CRC Press, Boca Raton
- Mount SM, Chang C (2002) Evidence for a plastid origin of plant ethylene receptor genes. *Plant Physiol* 130:10–14
- Moussatche P, Klee HJ (2004) Autophosphorylation activity of the *Arabidopsis* ethylene receptor multigene family. *J Biol Chem* 279:48734–48741
- Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, Vicedomini R, Sahlin K, Sherwood E, Elfstrand M, Gramzow L, Holmberg K, Hällman J, Keech O, Klasson L, Koriabine M, Kucukoglu M, Käller M, Luthman J, Lysholm F, Niittylä T, Olson A, Rilakovic N, Ritland C, Rosselló JA, Sena J, Svensson T, Talavera-López C, Theißen G, Tuominen H, Vanneste K, Wu ZQ, Zhang B, Zerbe P, Arvestad L, Bhalerao R, Bohlmann J, Bousquet J, Garcia Gil R, Hvidsten TR, de Jong P, MacKay J, Morgante M, Ritland K, Sundberg B, Thompson SL, Van de Peer Y, Andersson B, Nilsson O, Ingvarsson PK, Lundeberg J, Jansson S (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature* 497:579–584
- O'Malley RC, Rodriguez FI, Esch JJ, Binder BM, O'Donnell P, Klee HJ, Bleecker AB (2005) Ethylene-binding activity, gene expression levels, and receptor system output for ethylene receptor family members from *Arabidopsis* and tomato. *Plant J* 41:651–659
- Qu X, Hall B, Gao Z, Schaller GE (2007) A strong constitutive ethylene-response phenotype conferred on *Arabidopsis* plants containing null mutations in the ethylene receptors ETR1 and ERS1. *BMC Plant Biol* 7:3
- Rodriguez FI, Esch JJ, Hall AE, Binder BM, Schaller GE, Bleecker AB (1999) A copper co-factor for the ethylene receptor ETR1 from *Arabidopsis*. *Science* 283:996–998
- Sakai H, Hua J, Chen QG, Chang C, Medrano LJ, Bleecker AB, Meyerowitz EM (1998) ETR2 is an ETR1-like gene involved in ethylene signaling in *Arabidopsis*. *Proc Natl Acad Sci USA* 95:5812–5817
- Schaller GE (1997) Ethylene and cytokinin signalling in plants: the role of two-component systems. *Essays Biochem* 32:101–111
- Schaller GE (2012) Ethylene and the regulation of plant development. *BMC Biol* 10:9. doi:10.1186/1741-7007-10-9
- Schaller GE, Ladd AN, Lanahan MB, Spanbauer JM, Bleecker AB (1995) The ethylene response mediator ETR1 from *Arabidopsis* forms a disulfide-linked dimer. *J Biol Chem* 270:12526–12530
- Shakeel SN, Wang X, Binder BM, Schaller GE (2013) Mechanisms of signal transduction by ethylene: overlapping and non-overlapping signalling roles in a receptor family. *AoB Plants* 5:plt010
- Solano R, Stepanova A, Chao Q, Ecker JR (1998) Genes Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Development* 12:3703–3714
- Stepanova AN, Alonso JM (2005) *Arabidopsis* ethylene signaling pathway. *Sci STKE* 276:cm4
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Tieman DM, Taylor MG, Ciardi JA, Klee HJ (2000) The tomato ethylene receptors NR and LeETR4 are negative regulators of ethylene response and exhibit functional compensation within a multigene family. *Proc Natl Acad Sci USA* 97:5663–5668
- Timme RE, Delwiche CF (2010) Uncovering the evolutionary origin of plant molecular processes: comparison of *Coleochaete* (Coleochaetales) and *Spirogyra* (Zygnematales) transcriptomes. *BMC Plant Biol* 10:96
- Wang KL, Li H, Ecker JR (2002) Ethylene biosynthesis and signaling networks. *Plant Cell* 14:S131–S151
- Wang W, Hall AE, O'Malley R, Bleecker AB (2003) Canonical histidine kinase activity of the transmitter domain of the ETR1 ethylene receptor from *Arabidopsis* is not required for signal transmission. *Proc Natl Acad Sci USA* 100:352–357

- Wang W, Esch JJ, Shiu SH, Agula H, Binder BM, Chang C, Patterson SE, Bleecker AB (2006) Identification of important regions for ethylene binding and signaling in the transmembrane domain of the ETR1 ethylene receptor of *Arabidopsis*. *Plant Cell* 18:3429–3442
- Whalen MC, Feldman LJ (1988) The effect of ethylene on root growth of *Zea mays* seedlings. *Can J Bot* 66:719–723
- Xie F, Liu Q, Wen C-K (2006) Receptor signal output mediated by the ETR1 N-terminus is primarily subfamily I receptor dependent. *Plant Physiol* 142:492–508
- Yau CP, Wang L, Yu M, Zee SY, Yip WK (2004) Differential expression of three genes encoding an ethylene receptor in rice during development, and in response to indole-3-acetic acid and silver ions. *J Exp Bot* 55:547–556
- Young TE, Gallie DR, DeMason DA (1997) Ethylene-mediated programmed cell death during maize endosperm development of wild-type and *shrunk2* genotypes. *Plant Physiol* 115:737–751
- Young TE, Meeley RB, Gallie DR (2004) ACC synthase expression regulates leaf performance and drought tolerance in maize. *Plant J* 40:813–825