Letters

Are fungi-derived genomic regions related to antagonism towards fungi in mosses?

Introduction

Land plants have been intimately associated with fungi over the course of their evolution. Because of their lack of sophisticated protective structures, early land plants would conceivably have required additional defense strategies against microbial pathogens, including various fungi. On the other hand, a symbiotic association between plants and fungi is instrumental to plant adaptation to terrestrial environments (Selosse & Le Tacon, 1998; Bidartondo *et al.*, 2011; Martin *et al.*, 2017). The crucial role of this partnership is further evidenced by the widespread occurrence of mycorrhiza (root–fungi association) in vascular plants, as well as mycorrhiza-like fungal associations (MFAs hereafter) in nonvascular plants, such as liverworts and hornworts (Wang & Qiu, 2006; Pressel *et al.*, 2014). Surprisingly, although fungal symbiosis is commonly considered to be an ancestral trait for land plants (Wang & Qiu, 2006; Delaux *et al.*, 2015), with the possible exception of the genus *Takakia* (Boullard, 1988; Grosche *et al.*, 2018), no MFAs have been confirmed in other mosses (Pressel *et al.*, 2014; Field *et al.*, 2015), the most diverse group of nonvascular land plants.

Here, we report two genomic regions in the nuclear genome of the moss *Physcomitrium patens*, previously *Physcomitrella patens* (Medina *et al.*, 2019; Rensing *et al.*, 2020), that contain mostly fungi-specific genes and mobile genetic elements. These two regions were identified in our genome screening for horizontally acquired genes in *P. patens*. Available evidence indicates that these fungi-specific genes are probably involved in the interaction between mosses and fungi. We discuss how these fungi-specific genes might have contributed to the defense against fungal and other microbial pathogens, as well as the loss of MFAs in mosses.

Two genomic regions in *P. patens* include mostly fungi-specific genes

The two genomic regions reported here are 513 and 111 kbp long, respectively, and they are referred to as fungal region 1 (FR1) and fungal region 2 (FR2). FR1 is located on chromosome 13, from base pair position 14 225–527 422 bp, and contains 19 genes based on the annotation of *P. patens* v3.3 (Fig. 1a; Supporting Information Table S1) (Lang *et al.*, 2018). With these genes as query, we performed BLAST searches of NCBI nonredundant (*nr*) protein sequences, JGI Fungal Genomics (MycoCosm), OneKP (One Thousand Plant Transcriptomes, 2019), recently sequenced charophytes and hornworts, as well as our own genome and transcriptome data (Table S2; Methods S1). For eight of the 19 annotated genes, hits were mostly found in fungi and nonseed plants, including mosses, liverworts and seedless vascular plants (*E*-value threshold = 0.01); hits were also occasionally found in bacteria and microbial eukaryotes (e.g. haptophytes and stramenopiles), but not in charophyte algae, the closest relatives of land plants (Delwiche & Cooper, 2015), or in hornworts (Table S3). Based on their sequence similarities, these genes could be roughly classified into three families encoding: a heterokaryon incompatibility protein (HET) domain that reportedly functions in the self/nonself recognition system of filamentous fungi (PF06985; loci *Pp3c13_190*, *Pp3c13_650*/*Pp3c13_670*, and *Pp3c13_930*); a HopQ1-like protein (HLP) that may act as a virulence factor in bacteria (*Pp3c13_910*); and a conserved hypothetical protein with unknown function (*Pp3c13_120*, *Pp3c13_300*, and *Pp3c13_690*; termed *PpCF* here for *P. patens* conserved fungal gene (*CF*)) (Fig. 1a; Table S1). Another eight genes are either specific to *P. patens* or restricted to mosses and seedless vascular plants. The remaining three genes in FR1 are truncated but actively transcribed Ty1/Copia retrotransposons. FR2 contains 11 genes positioned from 5180 386 to 5291 530 bp on chromosome 27 (Fig. 1b; Table S1). Among them, seven genes from five loci contain a HET domain (*Pp3c27_8610*, *Pp3c27_8690*/*Pp3c27_8695*, *Pp3c27_8710*, *Pp3c27_8740*, *Pp3c27_8810*/ *Pp3c27_8820*), and one encodes a hypothetical protein of unknown function (*Pp3c27_8730*) that was only found in *P. patens* and two ascomycete fungi (*Aspergillus turcosus*, *Fusarium euwallaceae*) (*E*-value threshold = 0.01). Each of the remaining three annotated genes either encodes a short peptide (< 53 amino acids) or appears to be specific to *P. patens*. Both FR1 and FR2 contain additional degenerated transposable elements (TEs) (Fig. S1). The possibility that FR1 and FR2 represent a contamination could be excluded by the following evidence (Notes S1): CG, CHG and CHH DNA methylation (Fig. S1); homolog distribution of TEs from FR1 and FR2 in other chromosomes of *P. patens*, and in liverworts and lycophytes (Figs S2–S4); genetic linkage map and bacterial artificial chromosome/fosmid paired-end sequence data (Fig. S5); PCR amplification and sequencing of FR1, FR2, and adjacent regions (Fig. S6; Table S4); *de novo* sequencing and assembly of 12 other mosses, including one high-quality genome from an unspecified congeneric species (i.e. *Physcomitrium* sp.; Figs S7, S8).

Horizontal gene transfer of fungi-specific genes

Other than a fraction of lineage-specific or orphan genes, all other genes in FR1 and FR2 were probably subjected to lateral transmission. Retrotransposons have long been known to transfer to, and amplify frequently in, plants and other eukaryotes (El

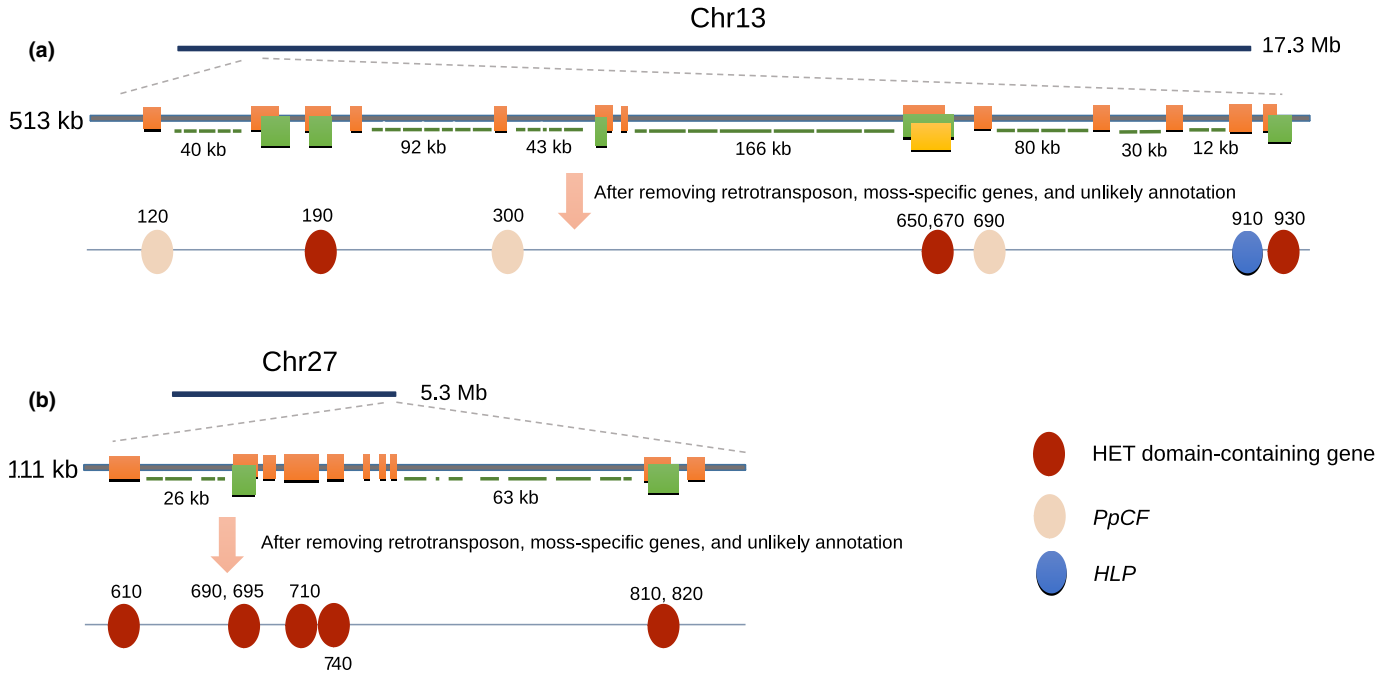

Fig. 1 Schematic illustration of the two large DNA fragments on chromosome 13 (FR1, 513 kb) (a) and chromosome 27 (FR2, 111 kb) (b) harboring genes of potential fungal origin and mobile elements in *Physcomitrella patens*. The distribution of protein-coding genes in each fragment is shown at three levels: the chromosome level, the scaffold level with the original annotation information of *P. patens* v3.3, and the scaffold level after removing the annotated retrotransposons, moss-specific genes and incorrectly annotated genes. Colored boxes indicate protein-coding genes in each region with overlapping genes shown in different colors. Numbers above the genes of fungal origin show the original gene identifiers according to *P. patens* genome annotation v3.3 in Phytozome (e.g. 120 = *Pp3c13_120*). Mobile elements are illustrated by solid dark green lines with length information. Also see Supporting Information Fig. S1 for JBrowse screenshots of the two genomic regions.

Baidouri *et al.*, 2014), and their presence in FR1 and FR2 is not surprising, particularly given that TEs account for 57% of the *P. patens* genome (Lang *et al.*, 2018). The more noteworthy finding is the existence of three gene families with detectable homologs mostly restricted to nonseed plants and fungi (i.e. *PpCF*, *HLP* and HET domain-containing genes). Consistent with sequence similarity comparisons, phylogenetic analyses show that these genes in mosses and other land plants are most closely related to homologs from different fungal lineages (Figs 2, S9–S12; Methods S1; Notes S2). Such phyletic distributions and relationships are typically indicative of historical horizontal gene transfer (HGT) events between fungi and early land plants, although lineage-specific gene loss remains a possible, but less parsimonious, scenario (Huang & Gogarten, 2006) (also see Notes S2). Given their distribution in both Ascomycota and Basidiomycota (and Mucoromycota for *CF*), which originated earlier than land plants (Berbee *et al.*, 2017; Lutzoni *et al.*, 2018), we reason that these genes were probably transferred from fungi to the most recent common ancestor of land plants. In view of their affinities with different fungal lineages, the three gene families, and possibly members of HET domain-containing genes, might have been acquired independently and later evolved into gene clusters in *P. patens* through genome rearrangements; alternatively, they might have been derived from a single HGT event, followed by lineage-specific duplication and losses in both land plants and fungi. We note here that although genes acquired from various sources exist in land plants (Hoang *et al.*, 2009; Yue *et al.*, 2012; Maumus *et al.*, 2014; Zhang *et al.*, 2020), the vast majority of reported HGT events involve mitochondrial genes and/or parasitic plants (Davis & Xi, 2015; Yang *et al.*, 2016); HGT from fungi to land plants was once considered to be extremely rare (Richards *et al.*, 2009), though they have been documented in multiple recent studies (Bowman *et al.*, 2017; Guan *et al.*, 2018; Li *et al.*, 2018; H. Wang *et al.*, 2020; S. Wang *et al.*, 2020). Our finding suggests that HGT of nuclear genes from fungi to plants might occur either more frequently than currently realized or in large fragments.

Some fungi-derived genes are highly responsive to fungal and chitin treatments

Strikingly, all three gene families (i.e. *CF*, *HLP* and HET domain-containing genes) appear to be functionally related to fungal interactions with other organisms, either different fungi or host plants. HLPs contain an inosine-uridine nucleoside hydrolase domain and are homologous to the type III secretion system effector HopQ1 of bacterial pathogens (Li *et al.*, 2013). The HLP in *P. patens* only shares 24–30% sequence identity with bacterial HopQ1, but it was shown to induce plant immune responses when ectopically expressed in the bacterium *Pseudomonas syringae* (Piechocki *et al.*, 2018). This evidence suggests the *HLP* genes in fungi are probably involved in interactions with host plants. Indeed, the moss HLP protein sequences share the highest

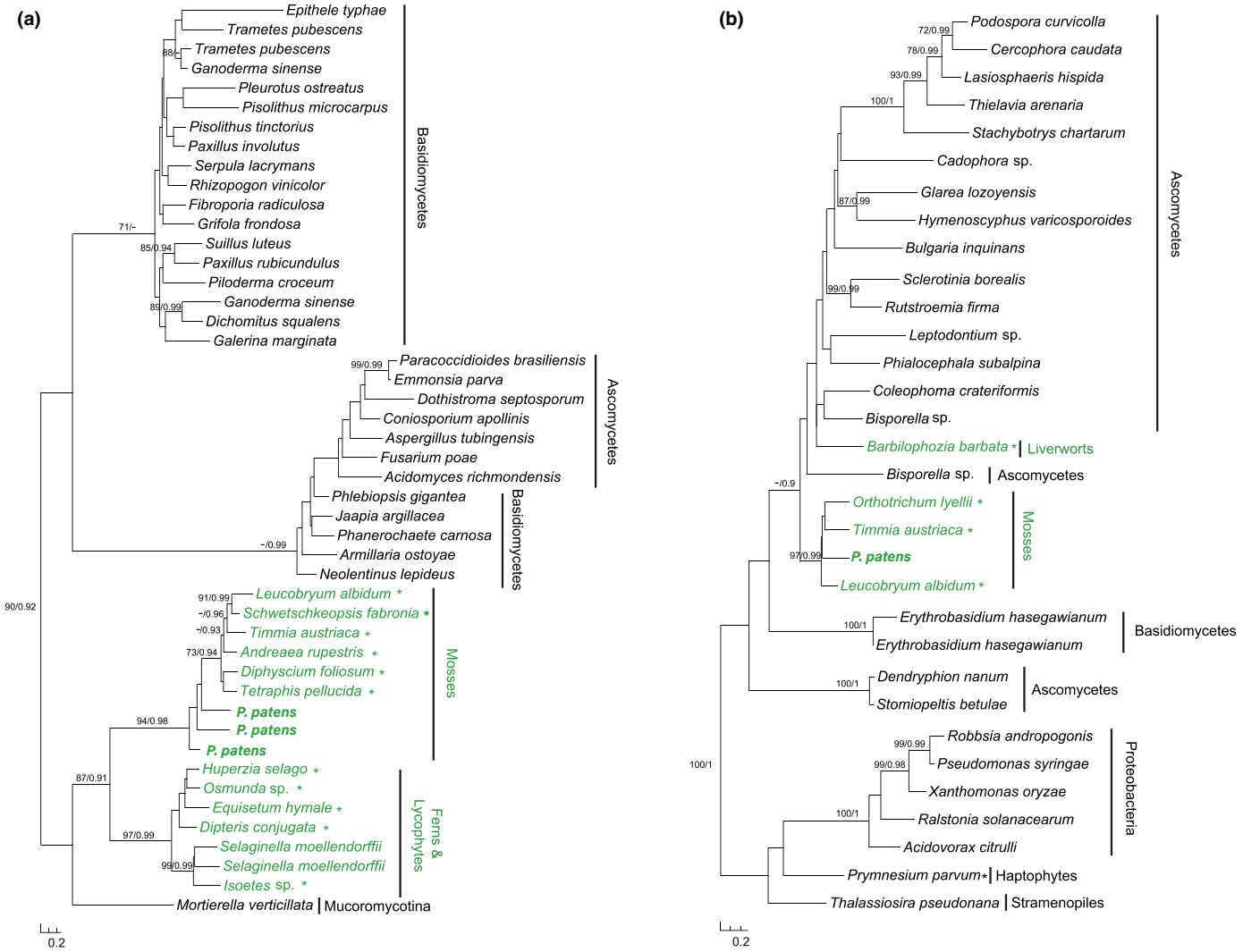

Fig. 2 Molecular phylogenies of CF (a) and HopQ1-like protein (HLP) (b) proteins. The alignment length used for tree construction is 260 amino acids for CFs and 277 amino acids for HLPs. Numbers above branches show bootstrap percentages and posterior probabilities from maximum-likelihood and Bayesian analyses, respectively. Values < 70% in maximum-likelihood analyses and < 0.9 in Bayesian analyses are not shown or indicated by dashes. Asterisks following species names indicate sequences from transcriptomic data. Lineage information is indicated after vertical bars. The two CFs from *Physcomitrium* sp. are 100% identical to each other and share 95% identity with Pp3c13_300v3.2 from *P. patens*. These sequences form a group with 92% bootstrap support in Supporting Information Fig. S9.

percentage identities (up to 41%) with plant-associated fungi such as the saprotroph *Amorphotheca resinae*, the leaf pathogen *Marssonina coronaria*, and the ericoid mycorrhizal fungus *Meliniomyces variabilis*. This pattern of gene distribution, with the highest similarities to sequences from fungal pathogens and symbionts, has also been observed for *PpCF*. Thus far, functional information for *CF* is not available, but gene fusion with an N-terminal HET domain was found in at least four homologous sequences from mycorrhizal fungi *Pisolithus tinctorius* and *Pisolithus microcarpus* (Fig. S13). As gene fusion often suggests functional linkage (Yanai *et al.*, 2001), it is likely that CFs are functionally related to the HET domain in fungi. The HET domain is the most common component of heterokaryon incompatibility genes (*het*) in the vegetative self/nonself recognition system of filamentous fungi (Hall *et al.*, 2010; Van der Nest *et al.*, 2014). Such a system allows hyphal fusion for individuals identical at *het* loci, but induces programmed cell death (PCD) when hyphae with different *het* alleles encounter (Paoletti & Clave, 2007). Currently, the HET domain is believed to encode an effector for PCD in fungal heterokaryon incompatibility (Paoletti & Clave, 2007). In *P. patens*, HET domain-containing genes are also commonly distributed on other chromosomes; our analyses identified at least 27 HET domain-containing genes from 22 gene loci (Table S5; Fig. S14), most of which are combined with a C-terminal sequence region of unknown function that is also present in fungal homologs.

The HET domain-containing genes, along with *PpCF* and *HLP*, are expressed in different tissues under various experimental conditions in *P. patens* (Figs S1, S14–S17). The majority of HET domain-containing genes have low expression levels. However, *Pp3c14_10* and *Pp3c27_8580* are moderately to strongly expressed in protonema, rhizoids, and gametophores (Figs S14, S15).

Interestingly, *Pp3c27_8710* and *Pp3c12_12000* show more than two-fold stronger expression under exudate treatment with *Gigaspora* (Fig. S14), a genus of arbuscular mycorrhizal fungi. These genes are also found to be more highly expressed in quantitative reverse transcription polymerase chain reaction (qRT-PCR) under the treatment of chitooctaose, a chitin derivative (Galotto *et al.*, 2020) (Fig. S16; Table S6). For the 16 genes for which data are present under both exudate and chitin treatments, we observe that 15 are more highly expressed under *Rhizophagus* exudate than in the control, while some are either less or more highly expressed under *Gigaspora* or *Botrytis* treatment (Fig. S18). The chitin treatment leads to genes either more highly expressed at 1 and 6 h, or showing higher expression after 1 h and lower expression after 6 h. Eight out of 16 genes show a congruent pattern of apparent expression stimulation under fungal treatment, while for the other eight genes there are differences between the treatments. This pattern of expression suggests a fine-grained response to different fungal treatments. We note that those loci encoding two genes (marked with blue boxes in Figs S14, S15) show a generally lower expression (Figs S14, S15), but they show detectable expression under treatment with both symbiotic and pathogenic fungal exudate (Fig. S14). The low level of detected expression might be a result of the challenge of representing conflicting gene models, sometimes on opposite strands. However, these regions encoding potential antisense RNAs might also constitute a defense system that makes use of small RNAs. The *PpCF* and *HLP* genes (Fig. S17) generally show stronger expression than the HET domain-containing genes. Similar to the HET domain-containing genes, three of the four *PpCF* and *HLP* genes show pronounced expression under fungal exudate (Fig. S17).

Are fungi-derived genes related to antagonism towards fungi in mosses?

The presence of the HET domain, *PpCF,* and other fungi-derived genes in mosses raises a fascinating question: why do mosses maintain genes that otherwise are specific to fungal activities? To date, most fungi-derived genes in mosses have not been investigated intensively, but it is worthwhile to speculate on their functional roles. It is possible that some of these genes have been coopted by mosses for activities not directly related to fungi. This is evidenced by the expression of *Pp3c14_10* and *Pp3c27_8580* during the various developmental stages of *P. patens* (Figs S14, S15). On the other hand, some other fungi-derived genes might have been recruited by mosses to regulate or counteract the activities of fungi or other microbes, in addition to their roles in other processes. This second possibility is in line with the role of the *HLP* gene in plant antimicrobial responses (Piechocki *et al.*, 2018) and supported by similar cases documented in the literature (Chou *et al.*, 2015; Di Lelio *et al.*, 2019). For example, endogenous viral DNAs are known to contribute significantly to the immunity of host animals against viral infection (Aswad & Katzourakis, 2012). In *P. patens*, remnant sequences of giant viruses are thought to provide protection from nucleocytoplasmic viruses (Lang *et al.*, 2018). Likewise, the *Fusarium* head blight resistance gene *Fhb7* in wheatgrasses of the genus *Thinopyrum* was acquired from fungal endophytes (H. Wang *et al.*, 2020).

Other than the ectopic induction of plant immune response by HLP and the potential functional link between CF and the HET domain, additional lines of evidence are consistent with the possible role of these fungi-derived genes in moss defense. Importantly, *het* genes are known to transfer between species, and acquired *het* genes may indeed induce PCD in filamentous fungi (Paoletti *et al.*, 2006; Wichmann *et al.*, 2008). The *het-c* gene, which encodes a transmembrane protein, is one of the best-studied *het* genes in fungi. In *Neurospora crassa*, *het-c* forms the heterokaryon incompatibility system through nonallelic interactions with the closely linked *pin-c* gene, which contains the HET domain (Kaneko *et al.*, 2006). Homologs of *het-c*, however, are also found in certain strains of the phytopathogenic bacterium *Pseudomonas syringae*, probably resulting from past HGT (Wichmann *et al.*, 2008). Ectopic expression of bacterial *het-c* homologs in *N. crassa* induces PCD of hyphae, suggesting that *P. syringae* might have acquired *het-c* to kill fungal cells to obtain nutrients (Wichmann *et al.*, 2008). We note here the other similarities of these genes in fungi and mosses. For instance, HET domain-containing genes are highly amplified and diverse in both groups. The HET domain has been proposed to evolve in fungi initially for pathogen recognition and host defense (Paoletti & Saupe, 2009). Interestingly, several HET domain-containing loci in *P. patens* (e.g. *Pp3c13_930*, *Pp3c13_970*) are located adjacent to loci encoding leucine-rich repeat receptor-like kinases (e.g., *Pp3c13_960* and *Pp3c13_990*), major players in pathogen-associated molecular pattern recognition of plant immunity (Tena *et al.*, 2011). Furthermore, several HET domain-containing genes are highly upregulated under the treatment of chitin, salicylic acid, and jasmonic acid, common signals of plant defense against pathogens and herbivores (Fig. S16).

Conclusions and perspectives

The acquisition of *PpCF*, *HLP* and HET domain-containing genes, which are seemingly specific to fungal activities, is surprising. The available evidence suggests that some of these genes might be involved in the interactions between mosses and fungi, probably as a defense mechanism in mosses. Such evidence is consistent with the idea that mosses and other bryophytes possess alternative defense strategies contributed by HGT from fungi and bacteria (Ponce de Leon & Montesano, 2017). These genes might have been lost secondarily from certain mosses but retained in other bryophytes or nonseed vascular plants, as often observed for resistance-related genes in plants (Zhao *et al.*, 2018; Zhang *et al.*, 2019). Given the likely antagonistic nature of these fungi-related genes, it is tempting to speculate whether they have possibly contributed to the loss of MFAs in mosses, particularly in light of the role of the HET domain in PCD of hyphae in filamentous fungi, which include mycorrhiza-like fungi. The loss of MFAs in mosses might also involve lineage-specific structural innovations and physiological processes (Field *et al.*, 2015). Nonetheless, many questions remain to be answered. For instance, the HET domain in fungi contains three motifs of 15–30 amino acids with unknown functions (Paoletti & Clave, 2007). If some HET domain-

containing genes are indeed involved in defense in mosses, further investigations are needed on whether all three motifs are required and whether these genes are intended for certain fungal groups or specific types of pathogens. Such information may help us to understand the existence of moss genes with a truncated HET domain (Table S5) and the presence of certain fungal pathogens and endophytes in mosses (Davey & Currah, 2006; U'Ren *et al.*, 2010; Chen *et al.*, 2018). Ultimately, gene knockout and other detailed functional investigations are required to uncover the role of these fungi-related genes in mosses. Such detailed functional investigations should also provide a better understanding of the importance of HGT in plant evolution, as well as in the interplay among genetic integration, organismal interaction and counteraction.

Acknowledgements

We are grateful to John Stiller for reading the manuscript draft and for his constructive comments. Thanks are also due to Jianchao Ma, Guanxiao Chang, and Jinjie Zhao for assistance in data analysis and experimental verification of the studied genes. This work is supported in part by the National Science Foundation of China (NSFC) grants to JH (31970248), GS (31771414) and QW (31800190), the Department of Science and Technology of Henan Province to GS (18HASTIT041), the Chinese Academy of Sciences 'Lights of Western China' grant to JH, the Second Tibetan Plateau Scientific Expedition and Research (STEP) Program to HS (2019QZKK0502), and the Key Projects of the Joint Fund of the National Natural Science Foundation of China to HS (U1802232). The authors declare no competing financial interests.

Author contributions

GS, YG and JH conceived and designed the study. GS, SB, YG, QW, JH, FBH, NF-P, AC and SAR performed analyses. SW performed experiments to confirm the existence and expression of identified genes. YL, HL and BG generated additional genomic and transcriptomic data and contributed to data analyses. MP, YZ, XH and HS participated in data analyses. GS, SAR and JH wrote the manuscript. GS and SB contributed equally to this work.

Data availability

All sequences, alignments, phylogenetic trees, and scripts are available on the GitHub website (https://github.com/slbai01/2019-het-article).

ORCID

Shenglong Bai https://orcid.org/0000-0003-2481-0457
Alia Czyrt https://orcid.org/0000-0002-8897-6969
Noe Fernandez-Pozo https://orcid.org/0000-0002-6489-5566
Fabian B. Haas https://orcid.org/0000-0002-7711-5282
Jinling Huang https://orcid.org/0000-0003-4893-4171
Stefan A. Rensing https://orcid.org/0000-0002-0225-873X
Guiling Sun https://orcid.org/0000-0001-7580-8301

Guiling Sun[1]†, Shenglong Bai[1]†, Yanlong Guan[2], Shuanghua Wang[2], Qia Wang[2], Yang Liu[3,4], Huan Liu[4,5], Bernard Goffinet[6], Yun Zhou[1], Mathieu Paoletti[7], Xiangyang Hu[8], Fabian B. Haas[9], Noe Fernandez-Pozo[9], Alia Czyrt[9], Hang Sun[2]*, Stefan A. Rensing[9]* and Jinling Huang[1,2,10]*

[1]Key Laboratory of Plant Stress Biology, State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng 475004, China;
[2]Key Laboratory for Plant Diversity and Biogeography of East Asia, Yunnan Key Laboratory for Fungal Diversity and Green Development, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China;
[3]Fairy Lake Botanical Garden, Chinese Academy of Sciences, Shenzhen 518004, China;
[4]BGI-Shenzhen, Shenzhen 518083, China;
[5]State Key Laboratory of Agricultural Genomics, BGI-Shenzhen, Shenzhen 518083, China;
[6]Ecology and Evolutionary Biology, University of Connecticut, 75N Eagleville Rd, Storrs, CT 06269-3043, USA;
[7]Laboratoire de Génétique Moléculaire des Champignons, Institut de Biochimie et de Génétique Cellulaires, UMR 5095 CNRS-Université de Bordeaux 2, 1 rue Camille St Saëns, Bordeaux Cedex 33077, France;
[8]College of Life Sciences, Shanghai University, Shanghai 200444, China;
[9]Plant Cell Biology, Department of Biology, University of Marburg, Marburg 35043, Germany;
[10]Department of Biology, East Carolina University, Greenville, NC 28590, USA
(*Authors for correspondence: tel +1 252 3285623, email huangj@ecu.edu (JH); tel +49 6421 28 21940, email stefan.rensing@biologie.uni-marburg.de (SAR); tel +86 871 65215002, email sunhang@mail.kib.ac.cn (HS))
†These authors contributed equally to this work.

References

Aswad A, Katzourakis A. 2012. Paleovirology and virally derived immunity. *Trends in Ecology & Evolution* **27**: 627–636.

Berbee ML, James TY, Strullu-Derrien C. 2017. Early diverging fungi: diversity and impact at the dawn of terrestrial life. *Annual Review of Microbiology* **71**: 41–60.

Bidartondo MI, Read DJ, Trappe JM, Merckx V, Ligrone R, Duckett JG. 2011. The dawn of symbiosis between plants and fungi. *Biology Letters* **7**: 574–577.

Boullard B. 1988. Observations on the coevolution of fungi and hepatics. In: Pirozynski KA, Hawkesworth DL, eds. *Coevolution of fungi with plants and animals*. London, UK: Academic Press, 107–124.

Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, Yamaoka S, Nishihama R, Nakamura Y, Berger F *et al.* 2017. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* **171**: 287–304, e215.

Chen KH, Liao HL, Arnold AE, Bonito G, Lutzoni F. 2018. RNA-based analyses reveal fungal communities structured by a senescence gradient in the moss *Dicranum scoparium* and the presence of putative multi-trophic fungi. *New Phytologist* **218**: 1597–1611.

Chou S, Daugherty MD, Peterson SB, Biboy J, Yang Y, Jutras BL, Fritz-Laylin LK, Ferrin MA, Harding BN, Jacobs-Wagner C *et al.* 2015. Transferred interbacterial antagonism genes augment eukaryotic innate immune function. *Nature* **518**: 98–101.

Davey ML, Currah RS. 2006. Interactions between mosses (Bryophyta) and fungi. *Canadian Journal of Botany* **84**: 1509–1519.

Davis CC, Xi Z. 2015. Horizontal gene transfer in parasitic plants. *Current Opinion in Plant Biology* **26**: 14–19.

Delaux P-M, Radhakrishnan GV, Jayaraman D, Cheema J, Malbreil M, Volkening JD, Sekimoto H, Nishiyama T, Melkonian M, Pokorny L *et al.* 2015. Algal ancestor of land plants was preadapted for symbiosis. *Proceedings of the National Academy of Sciences, USA* **112**: 13390.

Delwiche Charles F, Cooper ED. 2015. The evolutionary origin of a terrestrial flora. *Current Biology* **25**: R899–R910.

Di Lelio I, Illiano A, Astarita F, Gianfranceschi L, Horner D, Varricchio P, Amoresano A, Pucci P, Pennacchio F, Caccia S. 2019. Evolution of an insect immune barrier through horizontal gene transfer mediated by a parasitic wasp. *PLoS Genetics* **15**: e1007998.

El Baidouri M, Carpentier M-C, Cooke R, Gao D, Lasserre E, Llauro C, Mirouze M, Picault N, Jackson SA, Panaud O. 2014. Widespread and frequent horizontal transfers of transposable elements in plants. *Genome Research* **24**: 831–838.

Field KJ, Pressel S, Duckett JG, Rimington WR, Bidartondo MI. 2015. Symbiotic options for the conquest of land. *Trends in Ecology & Evolution* **30**: 477–486.

Galotto G, Abreu I, Sherman CA, Liu B, Gonzalez-Guerrero M, Vidali L. 2020. Chitin triggers calcium-mediated immune response in the plant model *Physcomitrella patens*. *Molecular Plant–Microbe Interactions* **33**: 911–920.

Grosche C, Genau AC, Rensing SA. 2018. Evolution of the symbiosis-specific GRAS regulatory network in bryophytes. *Frontiers in Plant Science* **9**: 1621.

Guan Y, Liu L, Wang Q, Zhao J, Li P, Hu J, Yang Z, Running MP, Sun H, Huang J. 2018. Gene refashioning through innovative shifting of reading frames in mosses. *Nature Communications* **9**: 1555.

Hall C, Welch J, Kowbel DJ, Glass NL. 2010. Evolution and diversity of a fungal self/nonself recognition locus. *PLoS ONE* **5**: e14055.

Hoang QT, Cho SH, McDaniel SF, Ok SH, Quatrano RS, Shin JS. 2009. An actinoporin plays a key role in water stress in the moss *Physcomitrella patens*. *New Phytologist* **184**: 502–510.

Huang J, Gogarten JP. 2006. Ancient horizontal gene transfer can benefit phylogenetic reconstruction. *Trends in Genetics* **22**: 361–366.

Kaneko I, Dementhon K, Xiang Q, Glass NL. 2006. Nonallelic interactions between *het-c* and a polymorphic locus, *pin-c*, are essential for nonself recognition and programmed cell death in *Neurospora crassa*. *Genetics* **172**: 1545.

Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, Haas FB, Piednoel M, Gundlach H, Van Bel M, Meyberg R *et al.* 2018. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal* **93**: 515–533.

Li M, Zhao J, Tang N, Sun H, Huang J. 2018. Horizontal gene transfer from bacteria and plants to the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. *Frontiers in Plant Science* **9**: 701.

Li W, Chiang Y-H, Coaker G. 2013. The HopQ1 effector's nucleoside hydrolase-like domain is required for bacterial virulence in *Arabidopsis* and tomato, but not host recognition in tobacco. *PLoS ONE* **8**: e59684.

Lutzoni F, Nowak MD, Alfaro ME, Reeb B, Miadlikowska J, Krug M, Arnold AE, Lewis LA, Swofford D, Hibbett D *et al.* 2018. Contemporaneous radiations of fungi and plants linked to symbiosis. *Nature Communications* **91**: 1–11.

Martin FM, Uroz S, Barker DG. 2017. Ancestral alliances: plant mutualistic symbioses with fungi and bacteria. *Science* **356**: eaad4501.

Maumus F, Epert A, Nogué F, Blanc G. 2014. Plant genomes enclose footprints of past infections by giant virus relatives. *Nature Communications* **5**: 4268.

Medina R, Johnson MG, Liu Y, Wickett NJ, Shaw AJ, Goffinet B. 2019. Phylogenomic delineation of *Physcomitrium* (Bryophyta: Funariaceae) based on targeted sequencing of nuclear exons and their flanking regions rejects the retention of *Physcomitrella*, *Physcomitridium* and *Aphanorrhegma*. *Journal of Systematics and Evolution* **57**: 404–417.

Paoletti M, Buck KW, Brasier CM. 2006. Selective acquisition of novel mating type and vegetative incompatibility genes via interspecies gene transfer in the globally invading eukaryote *Ophiostoma novo-ulmi*. *Molecular Ecology* **15**: 249–262.

Paoletti M, Clave C. 2007. The fungus-specific HET domain mediates programmed cell death in *Podospora anserina*. *Eukaryotic Cell* **6**: 2001–2008.

Paoletti M, Saupe SJ. 2009. Fungal incompatibility: evolutionary origin in pathogen defense? *BioEssays* **31**: 1201–1210.

Piechocki M, Giska F, Koczyk G, Grynberg M, Krzymowska M. 2018. An engineered distant homolog of *Pseudomonas syringae* TTSS effector from *Physcomitrella patens* can act as a bacterial virulence factor. *Frontiers in Microbiology* **9**: 1060.

Ponce de Leon I, Montesano M. 2017. Adaptation mechanisms in the evolution of moss defenses to microbes. *Frontiers in Plant Science* **8**: 366.

Pressel S, Bidartondo MI, Ligrone R, Duckett JG. 2014. Fungal symbioses in bryophytes: new insights in the twenty first century. *Phytotaxa* **9**: 238–253.

Rensing SA, Goffinet B, Meyberg R, Wu SZ, Bezanilla M. 2020. The moss *Physcomitrium* (*Physcomitrella*) *patens*: a model organism for non-seed plants. *Plant Cell* **32**: 1361–1376.

Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Talbot NJ. 2009. Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *Plant Cell* **21**: 1897.

Selosse MA, Le Tacon F. 1998. The land flora: a phototroph-fungus partnership? *Trends in Ecology & Evolution* **13**: 15–20.

Tena G, Boudsocq M, Sheen J. 2011. Protein kinase signaling networks in plant innate immunity. *Current Opinion in Plant Biology* **14**: 519–529.

One Thousand Plant Transcriptomes I. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679–685.

U'Ren JM, Lutzoni F, Miadlikowska J, Arnold AE. 2010. Community analysis reveals close affinities between endophytic and endolichenic fungi in mosses and lichens. *Microbial Ecology* **60**: 340–353.

Van der Nest MA, Olson A, Lind M, Velez H, Dalman K, Brandstrom Durling M, Karlsson M, Stenlid J. 2014. Distribution and evolution of *het* gene homologs in the basidiomycota. *Fungal Genetics and Biology* **64**: 45–57.

Wang B, Qiu YL. 2006. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**: 299–363.

Wang H, Sun S, Ge W, Zhao L, Hou B, Wang K, Lyu Z, Chen L, Xu S, Guo J *et al.* 2020. Horizontal gene transfer of *Fhb7* from fungus underlies *Fusarium* head blight resistance in wheat. *Science* **368**: eaba5435.

Wang S, Guan Y, Wang Q, Zhao J, Sun G, Hu X, Running MP, Sun H, Huang J. 2020. A mycorrhizae-like gene regulates stem cell and gametophore development in mosses. *Nature Communications* **11**: 2030.

Wichmann G, Sun J, Dementhon K, Glass NL, Lindow SE. 2008. A novel gene, *phcA* from *Pseudomonas syringae* induces programmed cell death in the filamentous fungus *Neurospora crassa*. *Molecular Microbiology* **68**: 672–689.

Yanai I, Derti A, DeLisi C. 2001. Genes linked by fusion events are generally of the same functional category: a systematic analysis of 30 microbial genomes. *Proceedings of the National Academy of Sciences, USA* **98**: 7940.

Yang Z, Zhang Y, Wafula EK, Honaas LA, Ralph PE, Jones S, Clarke CR, Liu S, Su C, Zhang H *et al.* 2016. Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. *Proceedings of the National Academy of Sciences, USA* **113**: E7010.

Yue J, Hu X, Sun H, Yang Y, Huang J. 2012. Widespread impact of horizontal gene transfer on plant colonization of land. *Nature Communications* **3**: 1152.

Zhang J, Fu XX, Li RQ, Zhao X, Liu Y, Li MH, Zwaenepoel A, Ma H, Goffinet B, Guan YL *et al.* 2020. The hornwort genome and early land plant evolution. *Nature Plants* **6**: 107–118.

Zhang T, Qiao Q, Novikova PY, Wang Q, Yue J, Guan Y, Ming S, Liu T, De J, Liu Y *et al.* 2019. Genome of *Crucihimalaya himalaica*, a close relative of *Arabidopsis*, shows ecological adaptation to high altitude. *Proceedings of the National Academy of Sciences, USA* **116**: 7137–7146.

Zhao Q, Feng Q, Lu H, Li Y, Wang A, Tian Q, Zhan Q, Lu Y, Zhang L, Huang T *et al.* 2018. Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. *Nature Genetics* **50**: 278–284.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 JBrowse screenshots of the two regions covering FR1/2 in *P. patens* genome.

Fig. S2 Molecular phylogeny of retroviral integrase core domain RVE (PF00665) type TEs in *P. patens*, *Marchantia polymorpha* and *Selaginella moellendorffii* and their chromosomal distribution in *P. patens*.

Fig. S3 Molecular phylogeny of retroviral aspartyl protease RVP_2 domain (PF08284) in *P. patens*, *M. polymorpha* and *S. moellendorffii* and their chromosomal distribution in *P. patens*.

Fig. S4 Molecular phylogeny of reverse transcriptase RVT_1 domain (PF00078) in *P. patens, M. polymorpha* and *S. moellendorffii* and their chromosomal distribution in *P. patens*.

Fig. S5 Evidence of FR1 and FR2 in the genome assembly of *P. patens*.

Fig. S6 Verification of regions in FR1 and FR2 by PCR amplification and sequencing.

Fig. S7 Distribution of the target genomic regions (i.e. those mostly consisting of *CF*, *HLP* and HET domain-containing genes) in *Physcomitrium* sp., *Ulota hutchinsiae* and *Anomodon attenuatus*.

Fig. S8 Mapping view and depth statistics of the contig that contains *CF*, *HLP* and HET domain-containing genes in *Physcomitrium* sp. (Contig00000005).

Fig. S9 Maximum-likelihood tree of CFs in mosses and other eukaryotes.

Fig. S10 Overview of the phylogenetic relationships of HET domain-containing genes in *P. patens* and other eukaryotes.

Fig. S11 Molecular phylogeny of a subset of HET domain-containing proteins in *P. patens* and the homologs from other eukaryotes.

Fig. S12 Molecular phylogenies of five HET domain-containing genes in FR1 and FR2.

Fig. S13 Gene fusion of HET domain and *CF* homologs in fungi *Pisolithus tinctorius* and *P. microcarpus*.

Fig. S14 Phylogenetic tree, expression profiles and domain structures of 27 HET domain-containing genes in *P. patens*.

Fig. S15 Phylogenetic tree and expression profiles of 27 HET domain-containing genes in *P. patens*.

Fig. S16 Relative expression levels of HET domain-containing genes in *P. patens* after exogenous application of SA, methyl jasmonate, and chitooctaose (a fungal elicitor).

Fig. S17 Expression profiles of *PpCF* and *HLP* genes in *P. patens*.

Fig. S18 Comparison of expression profiles of the HET domain-containing genes in *P. patens*.

Methods S1 Identification of fungal genes in FR1, FR2, and their expression profiles.

Notes S1 Evidence of the presence of FR1 and FR2 in *P. patens*.

Notes S2 On the origin of HET domain-containing genes, *HLP* and *CF* in mosses.

Table S1 Annotated genes in FR1 and FR2 of *P. patens*.

Table S2 List of additional genomes and transcriptomes used in this study and identification of the three gene families.

Table S3 Distribution of the HET domain-containing genes in nonseed land plants with genomic or transcriptomic data available in *nr* and OneKP databases.

Table S4 Sequence-specific primers used for PCR verification of FR1, FR2, and their adjacent regions in *P. patens*.

Table S5 Distribution of all the HET domain-containing genes in the genome of *P. patens*.

Table S6 Sequence-specific primers used for qRT-PCR.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Key words: fungal symbionts, horizontal gene transfer, HET domain, HopQ1, land plant evolution, mycorrhizas, organismal interaction, symbiosis.

Received, 14 April 2020; accepted, 19 June 2020.