Horizontal gene transfer allowed the emergence of broad host range entomopathogens

Qiangqiang Zhang[a,1], Xiaoxuan Chen[a,1], Chuan Xu[a], Hong Zhao[a], Xing Zhang[a], Guohong Zeng[a], Ying Qian[a], Ran Liu[a], Na Guo[a], Wubin Mi[a], Yamin Meng[a], Raymond J. St. Leger[b], and Weiguo Fang[a,2]

[a]MOE Key Laboratory of Biosystems Homeostasis & Protection, College of Life Science, Zhejiang University, Hangzhou 310058, China; and [b]Department of Entomology, University of Maryland, College Park, MD 20742

Edited by Thomas A. Richards, University of Exeter, Exeter, United Kingdom, and accepted by Editorial Board Member W. F. Doolittle March 10, 2019 (received for review September 24, 2018)

The emergence of new pathogenic fungi has profoundly impacted global biota, but the underlying mechanisms behind host shifts remain largely unknown. The endophytic insect pathogen *Metarhizium robertsii* evolved from fungi that were plant associates, and entomopathogenicity is a more recently acquired adaptation. Here we report that the broad host-range entomopathogen *M. robertsii* has 18 genes that are derived via horizontal gene transfer (HGT). The necessity of degrading insect cuticle served as a major selective pressure to retain these genes, as 12 are up-regulated during penetration; 6 were confirmed to have a role in penetration, and their collective actions are indispensable for infection. Two lipid-carrier genes are involved in utilizing epicuticular lipids, and a third (MrNPC2a) facilitates hemocoel colonization. Three proteases degraded the procuticular protein matrix, which facilitated up-regulation of other cuticle-degrading enzymes. The three lipid carriers and one of the proteases are present in all analyzed *Metarhizium* species and are essential for entomopathogenicity. Acquisition of another protease (MAA_01413) in an ancestor of broad host-range lineages contributed to their host-range expansion, as heterologous expression in the locust specialist *Metarhizium acridum* enabled it to kill caterpillars. Our work reveals that HGT was a key mechanism in the emergence of entomopathogenicity in *Metarhizium* from a plant-associated ancestor and in subsequent host-range expansion by some *Metarhizium* lineages.

M [illegible] *m* | entomopathogenic fungi | HGT | pathogenic fungi emergence | fungal virulence evolution

New fungal diseases are constantly emerging in natural ecosystems and are responsible for high-profile declines and extinctions in wildlife as exemplified by the critical situation of bats, frogs, soft corals, and bees (1–4). The emergence of new pathogens frequently involves host switching, but the mechanisms for host-range changes remain largely unknown (5, 6). *Metarhizium* spp. are ubiquitous insect pathogens that differ significantly in host range. The early-diverging *Metarhizium album* and *Metarhizium acridum* have narrow host ranges, whereas the main clade of more recently evolved generalist species have a broad host range (7). *Metarhizium* spp. evolved from beneficial associates of plants, and some generalists, including *Metarhizium robertsii* [an emerging model for investigating fungal evolution in natural communities (8)], remain endophytic, with entomopathogenicity being a more recently acquired adaptation (7, 9, 10).

The insect cuticle is a complex structure composed of an outer epicuticular layer comprising diverse hydrocarbons and lipids and an inner procuticular layer containing chitin fibrils embedded in a composite protein matrix. The intact cuticle is impervious to organisms that do not possess an active mechanism of cuticle penetration (e.g., bacteria, viruses, and nonentomopathogenic fungi) (11). Like other insect pathogenic fungi, an essential infection step by *Metarhizium* spp. is penetration through the host insect's cuticle (9), and the evolutionary processes that transformed *Metarhizium* spp. to insect pathogens must have involved specific adaptations that enable it to breach this barrier.

Systematic studies have shown that horizontal gene transfer (HGT, i.e., the movement of genetic material between distant organisms) is prevalent in prokaryotes, in which it serves as an important mechanism for the emergence of new bacterial pathogens. However, the extent to which HGT contributes to the evolution of eukaryotic pathogens is largely unknown (12), in large measure because of a lack of systematic functional characterization of HGTs (13). In this study, we report that HGT of 18 genes, many involved in cuticle penetration, was a key mechanism in the emergence of entomopathogenicity in *Metarhizium*, and that acquisition and/or retention of several horizontally acquired genes expanded host range in some late-evolving *Metarhizium* species. Our work provides significant insights into the evolution and development of pathogenesis in *Metarhizium* spp., and also represents a genomewide functional characterization of the contributions of HGT to niche adaption in a eukaryote.

Results

Genomewide Identification of HGT Genes in *M. robertsii*. We first used two complementary genomewide screening methods (Darkhorse and a BLASTP-based method; *SI Appendix*, Fig. S1) to identify 118 putative interkingdom and 36 putative intrafungal kingdom HGT genes in *M. robertsii* (*SI Appendix*, Table S1). Among these, only 18 interkingdom HGT events were validated

Significance

Recently emerged fungal diseases are contributing toward global declines in wildlife, but the mechanisms for emergence of new pathogenic fungi remain mysterious. The entomopathogen *Metarhizium robertsii* evolved from plant symbionts, and here we report that this host shift was facilitated by 18 horizontal gene transfers (HGTs). The necessity of breaching cuticular barriers selected for retention of these genes, as 12 are up-regulated during penetration, and their collective actions are indispensable for infection by processes including degradation of procuticular proteins and utilization of epicuticular lipids. Five of the most recently acquired HGT genes contributed to host-range expansion in some late-evolving species. Our work reveals that HGT was a key mechanism for emergence and host-range expansion of fungal entomopathogens.

Author contributions: W.F. designed research; Q.Z., X.C., H.Z., X.Z., G.Z., Y.Q., R.L., N.G., W.M., and Y.M. performed research; Q.Z., X.C., C.X., and W.F. analyzed data; and R.J.S. and W.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. T.A.R. is a guest editor invited by the Editorial Board.

Published under the PNAS license.

[1]Q.Z. and X.C. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: wfang1@zju.edu.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1816430116/-/DCSupplemental.

Published online April 4, 2019.

Downloaded at Shanghai Jiao Tong University on January 5, 2021

by using the gold standard for identifying HGT (phylogenetic incongruence with statistical measures of confidence; *SI Appendix*, Figs. S2–S18 and Table S2) (14). We compared the topologies of the obtained gene trees with constrained trees by using CONSEL. For all 18 HGT genes, the approximately unbiased (AU) test showed that the gene trees obtained were the best supported, and alternative hypotheses were all statistically (P < 0.05) rejected (*SI Appendix*, Figs. S2–S18 and Dataset S1) (14). The results obtained by the seven other tests (np, bp, pp, kh, sh, wkh, and wsh) available in CONSEL were consistent with the AU test (*SI Appendix*, Figs. S2–S18). The two lipid-carrier genes (MAA_05244 and MAA_05652) each clustered among arthropod clades, and alternative hypotheses of vertical inheritance based on the constraint that animal sequences clustered together (unordered) were rejected (P < 0.05; *SI Appendix*, Figs. S2 and S3). Two hypothetical proteins (MAA_01380 and MAA_08605) and acetyltransferase MAA_00129 clustered among bacterial clades with nonbacterial outgroups, but the alternative hypotheses that the three *Metarhizium* genes (clade A in the obtained trees) clustered (unordered) with outgroups were rejected (P < 0.05; *SI Appendix*, Figs. S4, S14, and S16). The remaining 12 genes all clustered among bacterial clades with no nonbacterial outgroups, and alternative hypotheses of vertical inheritance based on the constraint that bacterial sequences clustered together (unordered) were rejected (P < 0.05; *SI Appendix*, Figs. S5–S13, S15, S17, and S18). For the hypothetical gene (MAA_09352), the constraint that Cordycipitaceae (*Cordyceps* and *Beauveria*) sequences clustered together was rejected (P < 0.05), suggesting a possible HGT between the ancestor of Cordycipitaceae and that of *Metarhizium* after an ancient HGT from a bacterium to a fungus (*SI Appendix*, Fig. S8).

Predicted Functions of the HGT Genes. We identified two Gene Ontology (GO) terms that were enriched in the set of 18 HGT genes (P < 0.05). MAA_05652 and MAA_05244 were assigned a "lipid binding function" (GO:0008289). MAA_03817 was not included in GO:0008289, but it also encodes a lipid carrier (14). Despite their predicted function, these three proteins showed no significant similarity ($>1e^{-05}$) to each other, but all three clustered with lipid-carrier genes from arthropods [MAA_05652 and MAA_05244 with mite sequences (*SI Appendix*, Figs. S2 and S3) and MAA_03817 with insect sequences (14)]. Phylogenetic analysis showed that 10 of the remaining 15 sequences clustered with Actinobacteria (*SI Appendix*, Figs. S4–S10 and S16–S18) and five with Proteobacteria (*SI Appendix*, Figs. S11–S15).

MAA_01413 and MAA_00986 were assigned "serine-type endopeptidase activity" (GO:0004252), and clustered with chymotrypsins (*SI Appendix*, Figs. S5 and S11 and Table S2). The *M. robertsii* genome contains 26 chymotrypsins (9). Phylogenetic analysis of these, along with representative fungal and bacterial sequences (*SI Appendix*, Figs. S5 and S11), showed that MAA_01413 and MAA_00986 clustered in different clades of bacterial chymotrypsins that were phylogenetically distant from each other and from other *M. robertsii* chymotrypsins that clustered in their own separate clades (*SI Appendix*, Fig. S19).

MAA_09637 encodes a pyroglutamyl peptidase I-like enzyme that would likely also be involved in proteolysis. Of the remaining 12 genes, six encoded hypothetical proteins and the other six encoded proteins belonging to diverse functional groups (*SI Appendix*, Table S2). Five HGT genes (27.8%), including the three proteolysis-related genes (MAA_01413, MAA_00986, and MAA_09637), encoded proteins with secretion signal peptides (*SI Appendix*, Table S2), compared with 17.6% of *M. robertsii* proteins overall (9).

Distribution and Synteny of the HGT Genes. A phylogeny inferred from genome sequences of seven representative *Metarhizium* species has been reported (7). *M. robertsii*, *Metarhizium brunneum*, and *Metarhizium anisopliae* have broad host ranges; *Metarhizium album* and *M. acridum* have narrow host ranges; and *Metarhizium guizhouense* and *Metarhizium majus* have intermediate host ranges (7). We looked for homologs of the 18 *M. robertsii* HGT genes in the other genomes. *M. brunneum* lacked a homolog for lipid carrier MAA_03817, and *M. anisopliae* lacked a homolog to a hypothetical protein (MAA_08684), but they otherwise retained homologs of the *M. robertsii* HGT genes (Fig. 1). Of the 13 HGT genes absent in *M. album*, seven were also absent in another specialist, *M. acridum*. Four of these seven were lacking in *M. majus*, and two were lacking in *M. guizhouense* (Fig. 1). Except for the absence of lipid carrier MAA_03817 in *M. brunneum*, all seven *Metarhizium* species had homologs to the three lipid carriers, the protease MAA_00986, and a tRNA-ribosyltransferase-isomerase (MAA_05974; Fig. 1).

A polysaccharide lyase (MAA_06640), a β-xylosidase (MAA_07693), and the pyroglutamyl peptidase I-like enzyme MAA_09637 were found in only *Metarhizium* spp. (*SI Appendix*, Figs. S7, S9, and S18), and other genes were also present in a small minority of sequenced ascomycete fungi, mostly hypocrealean genomes (*SI Appendix*, Figs. S2–S6, S8, and S10–S17).

Two genes [glyoxalase MAA_06871 and a hypothetical gene (MAA_06534)] were in conserved microsyntenic regions in *Metarhizium* species and several other hypocrealean fungi (*SI Appendix*, Figs. S12 and S17). There was no synteny between *Metarhizium* and non-*Metarhizium* fungi in the remaining genes (*SI Appendix*, Figs. S2–S11, S13–S16, and S18). Hypothetical genes MAA_08604 and MAA_08605 were physically linked on the *Metarhizium* chromosomes; their homologs are also physically linked in their putative rhizobacterial donors (*SI Appendix*, Figs. S13 and S14), indicating that a bacterial DNA fragment containing MAA_08604 and *MAA_08605* was acquired by an ancestor of *Metarhizium*.

A Majority of HGT Genes Are Up-Regulated During Cuticle Penetration by *M. robertsii*. The expression profile of *Mr-npc2a* (MAA_03871) was previously reported (14). Quantitative RT-PCR (qRT-PCR) analysis was used to compare the expression of the remaining 17 HGT genes during (*i*) saprophytic growth in Sabouraud dextrose broth plus 1% yeast extract (SDY), (*ii*) symbiotic growth in the rhizospheres of *Arabidopsis thaliana*, and (*iii*) pathogenesis (i.e., cuticle penetration and colonization of hemocoel). Compared with saprophytic growth, hemocoel colonization and symbiotic

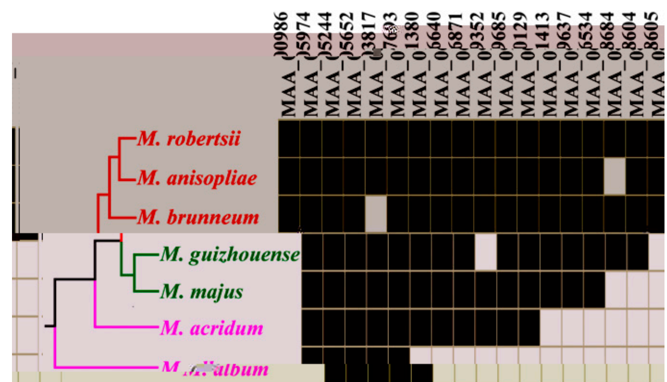

Fig. 1. Distribution of homologs of the *M.* HGT genes in other *M* species. (*L*) A reconstruction of the phylogeny of seven *M* species based on genomic sequences (n = 7). Pink branches indicate specialists; green branches indicate species with intermediate host range; red branches indicate generalists. (*R*) The presence of homologs of the HGT genes found in *M.* in other *M* species. Black box, present; white box, absent. Note: homologs of the 13 HGT candidates found in *M.* were absent in the specialist *M.*, 7 of which were also absent in another specialist, *M.*.

Downloaded at Shanghai Jiao Tong University on January 5, 2021

growth caused up-regulation of two and four genes, respectively (Fig. 2*A*). Two thirds (n = 12) of the HGT genes were up-regulated during cuticle penetration (Fig. 2*A*), in contrast to only 7% of *M. robertsii* genes overall (15).

Cuticle penetration is regulated by several major signaling components, including a PKA (16), three MAPKs (Hog1-, Fus3-, and Slt2-MAPK), the adaptor protein Mr-STE50, and the membrane anchor protein Mr-OPY2 (15). Nine of the 12 genes up-regulated on cuticle were regulated by at least one of these signaling components (Fig. 2*B*).

Collective Action of HGT Lipid Carriers Is Indispensable for Infection by *M. robertsii*. The lipid-carrier gene (MAA_03817) was deleted previously (14). For this study, we deleted the remaining 17 HGT genes, and all deletion mutants were complemented with their respective genomic clones (*SI Appendix*, Fig. S20). As shown in the figures and tables, none of the complemented strains were significantly different from WT in any assays, so they are not discussed further. No differences were observed between the WT and the deletion mutants in colony morphology and growth rate on potato dextrose agar (*SI Appendix*, Fig. S21). Virulence was assayed on *Galleria mellonella* larvae that were infected by topically applying conidia onto the insect cuticle (i.e., natural infection) or by direct injection of conidia into the hemocoel (to bypass the cuticle). Following direct injection, the deletion mutants showed no significant ($P > 0.05$) differences from the WT in speed of kill (Table 1 and *SI Appendix*, Table S3). However, deletion mutants of the three proteases, the lipid carrier MAA_05652, and a hypothetical protein (MAA_08605) that were all up-regulated during cuticle penetration had significantly ($P < 0.05$) reduced virulence when topically applied to insects (Table 1), indicating that their roles were limited to cuticle penetration.

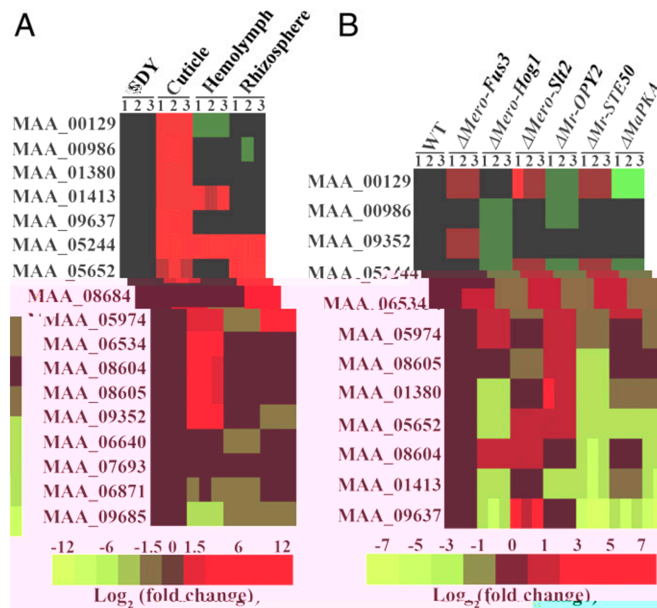

Fig. 2. Expression and regulation of HGT genes in *M.* . (*A*) qRT-PCR analysis of gene expression during saprophytic growth (grown in nutrient-rich SDY medium), two key infection stages (cuticle, cuticle penetration; hemolymph, hemocoel colonization), and plant root colonization (rhizosphere). The expression level of a gene during saprophytic growth (i.e., in SDY) is set to 1; the values represent the log_2-transformed fold changes of differential gene expression during infection or root colonization compared with the saprophytic growth in SDY. (*B*) qRT-PCR analysis of expression of the 12 HGT genes up-regulated during cuticle penetration (shown in *A*) in WT and 6 signaling mutants (impaired in cuticle penetration). These include three MAPK mutants (*ΔM -F 3*, *ΔM -H 1*, and *ΔM -S 2*), the PKA mutant *ΔM PKA*, the membrane protein Mr-OPY2 mutant *ΔM -OPY2*, and the adaptor Mr-STE50 mutant *ΔM -STE50*. The expression level of WT is set to 1; the values represent the log_2-transformed fold changes of differential gene expression of mutants vs. WT. The experiments were repeated three

Deleting the lipid-carrier gene MAA_05244 did not reduce virulence even though, like MAA_05652, it was up-regulated during cuticle penetration (Fig. 2*A*). To look for synergistic interactions between the carriers, we constructed a double gene-deletion mutant (*ΔMAA_05652*::*ΔMAA_05244*; *SI Appendix*, Fig. S20). When topically inoculated, *ΔMAA_05652*::*ΔMAA_05244* took 1.6-fold longer ($P < 0.05$) to kill than *ΔMAA_05652*, whereas both mutants showed WT levels of virulence when conidia were injected (Table 1). Four days after topical application, hyphal bodies were 9.5-fold more abundant in the hemolymph of WT infected insects than in insects infected with *ΔMAA_05652*::*ΔMAA_05244* (19 ± 10 hyphal bodies per milliliter vs. 2 ± 1.8 hyphal bodies per milliliter, respectively).

The previously studied lipid carrier MAA_03817 allows *M. robertsii* to utilize host sterols during hemocoel colonization (14). To investigate the collective contributions of the three HGT lipid carriers, we constructed the triple gene-deletion mutant (*ΔMAA_05652*::*ΔMAA_05244*::*ΔMAA_03871*; *SI Appendix*, Fig. S20). Fifteen days after inoculation, the mortality rate (22.6%) caused by the double gene-deletion mutant (*ΔMAA_05652*::*ΔMAA_05244*) was significantly (*t* test, $P < 0.05$) higher than the 12.8% mortality rate caused by the triple gene-deletion mutant (Table 1 and *SI Appendix*, Fig. S22).

Lipid Carriers MAA_05652 and MAA_05244 Are Involved in Utilization of Epicuticular Lipids. To assay their ligand-binding ability, MAA_05244 and MAA_05652 were expressed in *Escherichia coli* (*SI Appendix*, Figs. S23 and S24), and the recombinant proteins were subjected to ligand-binding assays. MAA_05652 had the greatest affinity to linoleic acid, but also bound palmitic and stearic acids (Fig. 3). MAA_05652 showed low affinity to two other fatty acids (dodecanoic acid and octanoic acid) and hydrocarbons. MAA_05244 showed strong affinity to only stearic acid (Fig. 3).

In half-strength SDY, single gene-deletion mutants of the lipid carriers (*ΔMAA_05652*, *ΔMAA_05244*) and the double gene-deletion mutant (*ΔMAA_05652*::*ΔMAA_05244*) germinated at similar rates ($P > 0.05$), but all germinated significantly faster than the WT ($P < 0.05$, Tukey's test in one-way ANOVA; *SI Appendix*, Fig. S25). To investigate the involvement of MAA_05244 and MAA_05652 in utilizing lipids, we used relative germination rates (*Materials and Methods*) to measure fungal growth on individual fatty acids as sole carbon source. Compared with WT, we found that *ΔMAA_05652*, *ΔMAA_05244*, and *ΔMAA_05652*::*ΔMAA_05244* grew significantly less well with linoleic, palmitic, or stearic acid ($P < 0.05$, Tukey's test in one-way ANOVA; *SI Appendix*, Table S4). In a control medium (containing oleic acid that does not bind to MAA_05652 or MAA_05244), relative germination rates were similar in WT and the three mutants (*SI Appendix*, Table S4).

The single (*ΔMAA_05652*, *ΔMAA_05244*) and double (*ΔMAA_05652*::*ΔMAA_05244*) gene-deletion mutants also grew less well than the WT in fatty acid and hydrocarbon epicuticular fractions ($P < 0.05$, Tukey's test in one-way ANOVA), again without differing significantly ($P > 0.05$) from each other (*SI Appendix*, Table S4). These in vitro results were reproduced in vivo, as relative germination rates of the mutants on locust hindwings were significantly reduced ($P < 0.05$; *SI Appendix*, Table S4).

The Three HGT Proteases Are Important for Procuticle Degradation by *M. robertsii*. In addition to the single gene-deletion mutants, we also constructed a triple gene-deletion mutant (*ΔMAA_00986*::*ΔMAA_09637*::*ΔMAA_01413*) of the two serine endoproteases

Table 1. Virulence (expressed as LT_{50} values) of WT and HGT gene mutants

Strains	Virulence, LT_{50}		Genes deleted or knocked down
	Topical application	Injection	
WT	$9.1 \pm 0.57^{*}$	$4.8 \pm 0.35^{*}$	—
Single gene-deletion mutants			
ΔMAA_01413	$15.2 \pm 1.87^{\dagger}$	$5.2 \pm 0.35^{*}$	Protease
ΔMAA_00986	$15.3 \pm 1.51^{\dagger}$	$5.1 \pm 0.18^{*}$	Protease
ΔMAA_09637	$16.7 \pm 1.27^{\dagger}$	$5.3 \pm 0.19^{*}$	Pyroglutamyl peptidase
ΔMAA_05244	$9.3 \pm 0.28^{*}$	$4.8 \pm 0.26^{*}$	Lipid carrier
ΔMAA_05652	$13.9 \pm 1.07^{\dagger}$	$5.2 \pm 0.12^{*}$	Lipid carrier
ΔMAA_08605	$12.5 \pm 0.93^{\dagger}$	$4.9 \pm 0.38^{*}$	Hypothetical protein
Multiple gene-deletion mutants			
ΔMAA_00986::ΔMAA_09637::ΔMAA_01413	$23.7 \pm 3.28^{\ddagger}$	$4.5 \pm 0.13^{*}$	Three proteases
ΔMAA_00986::ΔMAA_09637::ΔMAA_01413::MAA_05652 #1	NC (8.8%)	$4.8 \pm 0.14^{*}$	Three proteases, one lipid carrier
ΔMAA_05244::ΔMAA_05652	$22.2 \pm 0.91^{\ddagger}$	$4.0 \pm 0.15^{*}$	Two lipid carriers
ΔMAA_05244::ΔMAA_05652::ΔMAA_03817	NC (12.8%)	$5.9 \pm 0.13^{\dagger}$	Three lipid carriers
ΔMAA_05244::ΔMAA_05652::ΔMAA_03817::MAA_00986 #1	NC (0.9%)	$6.1 \pm 0.12^{\dagger}$	Three lipid carriers, one protease

kd, knockdown; LT_{50}, time required to kill 50% of the insects, NC, not calculated because of low mortality rate (rate in parentheses). This table shows only the LT_{50}s of mutants disrupted in multiple genes and single gene-deletion (Δ) mutants significantly different from WT following topical application. Data on other mutants, including complemented deletion mutants, are shown in *SI A* Table S3. For each inoculation method (i.e., each column), values followed by different symbols (*, †, ‡) are significantly different ($P < 0.05$, Tukey's test in one-way ANOVA). All assays were repeated three times. Data are expressed as the mean ± SE.

(*MAA_00986* and *MAA_01413*) and the pyroglutamyl peptidase (MAA_09637) to study their synergy (*SI Appendix*, Fig. S20). The three single and the triple gene-deletion mutants showed WT virulence when injected into the hemocoel (Table 1). However, following topical infection, the triple gene-deletion mutant took 2.5-fold longer to kill than the WT, and also took significantly longer to kill than the single gene-deletion mutants ($P < 0.05$, Tukey's test in one-way ANOVA; Table 1), indicating synergy between the proteases. Four days after topical applications, the triple gene-deletion mutant had produced 9.5-fold fewer hyphal bodies (2 ± 0.8/mL) in infected insects compared with WT (19 ± 5/mL).

To investigate how the three proteases act together to facilitate cuticle penetration, we quantified cuticle-degradation products following 12 h fungal growth on locust cuticle. Although growth rates of the WT and mutants on cuticle were similar ($P > 0.05$; *SI Appendix*, Fig. S26), all four mutants released significantly ($P < 0.05$, Tukey's test in one-way ANOVA) fewer amino acids from the cuticle (Fig. 4*A*).

As the expression of cuticle-degrading enzymes could have synergistic impacts during cuticle penetration (17), we also investigated whether the three HGT proteases impacted production of previously characterized *M. robertsii* cuticle-degrading enzymes. Compared with the WT, Pr1 subtilisin protease activity was not significantly ($P > 0.05$) altered in the three single gene-deletion mutants, but was significantly ($P < 0.05$, Tukey's test in one-way ANOVA) reduced in the triple gene-deletion mutant (Fig. 4*B*). The mutants each secreted significantly less total extracellular protease activity than the WT ($P < 0.05$, Tukey's test in one-way ANOVA; Fig. 4*B*). Likewise, the WT produced significantly more chitinase than *ΔMAA_01413* or *ΔMAA_09637*, which in turn produced significantly more chitinase than *ΔMAA_00986* ($P < 0.05$, Tukey's test in one-way ANOVA). The triple gene-deletion mutant had the lowest chitinase activity, producing 2.5-fold less than the WT (Fig. 4*B*).

Biochemical Characteristics of the Three HGT Proteases. It has long been assumed that *M. robertsii* secretes a large number of proteases that function redundantly to degrade cuticular proteins (9). In this context, it is surprising that the three HGT proteases had such a significant impact on cuticle degradation. To investigate further, we expressed each protease in *E. coli* for biochemical assays (*SI Appendix*, Figs. S27–S29).

Fig. 3. Ligand binding of two HGT-derived lipid carriers (MAA_05244 and MAA_05652). Binding of recombinant MAA_05652 (*L*[illegible]) and MAA_05244 (*R*[illegible]) to hydrocarbons and fatty acids. All assays were repeated three times. Data are expressed as mean ± SE.

MICROBIOLOGY

Downloaded at Shanghai Jiao Tong University on January 5, 2021

As predicted from its pyroglutamyl peptidase annotation in GenBank, the purified recombinant MAA_09637 protein released pyroglutamyl from locust cuticle (*SI Appendix*, Fig. S30). *M. robertsii* proteases produced during growth on locust cuticle solubilized significantly more protein (*t* test, $P = 0.007$) from the MAA_09637 pyroglutamyl peptidase-treated insect cuticle than from untreated insect cuticle, as did *Tritirachium album* subtilisin proteinase K (Sigma; *t* test, $P = 0.01$; *SI Appendix*, Table S5). Therefore, the MAA_09637 pyroglutamyl peptidase increases the susceptibility of insect cuticle to proteolysis. A genomewide search found that *M. robertsii* has an additional pyroglutamyl peptidase (MAA_04392), but qRT-PCR analysis showed that it was not up-regulated during cuticle penetration.

MAA_01413 and MAA_00986 are both annotated as chymotrypsin serine endopeptidases, but, to confirm specificities, we assayed their activities against 16 typical protease substrates (*SI Appendix*, Tables S6 and S7). Among these, MAA_00986 showed the highest activity against the trypsin/thrombin substrate Bz-Phe-Val-Arg-NA. The importance of secondary specificity was shown by relatively low activity against several other trypsin substrates. MAA_01413 also preferentially hydrolyzed Bz-Phe-Val-Arg-NA, but, unlike MAA_00986, it effectively hydrolyzed the other trypsin substrates as well as typical substrates for some elastases (e.g., Suc-Ala-Ala-Val-Ala-NA; *SI Appendix*, Tables S6 and S7). Neither enzyme hydrolyzed chymotrypsin substrates, including Suc-Ala-Ala-Pro-Phe-NA. The activity of proteinase K against the nonspecific protease substrate azocasein (Sigma-Aldrich) is significantly higher than that of MAA_01413, which in turn was significantly higher than that of MAA_00986 (*P* <(5rateolu346..5 (Bz6teolyantly)-401.8 (hi6teinamo9 (b45Phe-)-332.335r1.15
Sageinaa.5203s3991Tc-390.muSe741ta-n-35ntsubs75T1_11Tf26676601Tc-chT_11

mutant caused very low mortality rates, ranging from 0.9% to 5.6% (Table 1 and *SI Appendix*, Table S3).

HGTs Associated with Determination of Host Ranges in *Metarhizium* Fungi. To test whether HGT genes are associated with host-range determination in *Metarhizium* fungi, the seven HGT genes identified in the generalist *M. robertsii*, which lacked homologs in the specialists *M. acridum* and *M. album* (Fig. 1), were individually transferred into *M. acridum*, driven by the constitutive *Ptef* promoter from *Aureobasidium pullulans* (18). RT-PCR confirmed that the seven genes were expressed in their respective transformants during growth in SDY or locust cuticle (*SI Appendix*, Fig. S32*A*). WT *M. acridum* cannot infect *G. mellonella* caterpillars via topical application. However, transformants expressing protease MAA_01413, acetyltransferase MAA_00129, and three hypothetical proteins (MAA_08604, MAA_08605, MAA_08684) were able to infect this nonnatural host, albeit weakly (Fig. 5*A*). We also bioassayed a conidial mixture of the five transformants, but no significant difference in mortality was found between the individual transformants and the mixture (*SI Appendix*, Fig. S33).

To further investigate how expression of the HGT protease (MAA_01413) expands host range, we constructed an *M. acridum* transformant (*T-MAA_01413*) with MAA_01413 driven by its native promoter (*SI Appendix*, Fig. S32*A*). *T-MAA_01413* up-regulated *MAA_01413* expression on *G. mellonella* or locust cuticles (*SI Appendix*, Fig. S32*C*), showing that the *MAA_01413* promoter retains its *M. robertsii* expression profile in the transformed *M. acridum*. *T-MAA_01413* retained virulence to locust hosts (*SI Appendix*, Fig. S32*E*), but, unlike WT *M. acridum*, *T-MAA_01413* infected *G. mellonella* (*SI Appendix*, Fig. S32*D*). We compared the ability of WT *M. acridum* and *T-MAA_01413* to digest locust and *G. mellonella* cuticles. After 12 h growth on locust cuticle, the biomass of WT and *T-MAA_01413* was similar (*SI Appendix*, Fig. S34*A*), as was production of cuticle degradation products (peptides and amino acids; t test, $P > 0.05$; *SI Appendix*, Fig. S34*B*). In contrast, although biomasses were again similar ($P > 0.05$, one-way ANOVA; *SI Appendix*, Fig. S35), *T-MAA_01413* released significantly ($P < 0.05$) more peptides and amino acids from *G. mellonella* cuticle than did the WT (Fig. 5*B*). *T-MAA_01413* also produced significantly (t test, $P < 0.05$) higher levels of total extracellular protease, Pr1 subtilisin protease, and chitinase activity on *G. mellonella* cuticle (Fig. 5*C*) while producing WT levels of native enzymes when cultured on locust cuticle (*SI Appendix*, Fig. S34*C*).

Discussion

Seve5
comparedleve188lx0y881-110Td()Tj/T1_01Tf1.1o3lud(enzy-110isiHGT_010

for emergence of fungal pathogens in natural ecosystems remain largely unexplored.

In the present study, we used a combination of phylogenetic tools to identify 18 genes acquired by ancestors of *M. robertsii* through multiple independent HGT events. Most of the HGT genes are adapted to the recipient because their GC and GC3 content resemble *M. robertsii* rather than their putative donors (*SI Appendix*, Fig. S36). This contrasts with very recent HGT events characterizing the emergence of new fungal pathogens in agricultural ecosystems (19). Most of the bacterial-origin genes were also present in a small minority of sequenced ascomycete genomes. Previous interpretations of a patchy phylogenetic distribution of genes in related fungi have evoked independent gene transfer events (21, 22), but phylogenetic artifacts and/or unappreciated gene loss events may have influenced some of these interpretations. Two genes (glyoxalase MAA_06871 and hypothetical MAA_06534) showed conserved synteny in *Metarhizium* and some other hypocrealean entomopathogens, so their HGT events likely happened in a common ancestor. For the remaining 16 HGT genes, synteny conservation in *Metarhizium* species and lack of synteny between *Metarhizium* and non-*Metarhizium* fungi provide support for multiple independent transfer events. It is also possible that synteny conservation between multiple *Metarhizium* species is the result of secondary HGT between *Metarhizium* species after an ancient HGT. However, three of the genes (acetyltransferase MAA_00129 and lipid carriers MAA_05244 and MAA_03817) clustered with several non-*Metarhizium* fungal sequences to produce gene trees generally consistent with known ascomycete phylogenies, e.g., *Beauveria* and *Cordyceps* clustered together and distinct from *Metarhizium* sequences. Retention of sexual reproduction in *Cordyceps* and its loss in *Beauveria* have eliminated syntenic relationships (23). It is thus likely that these three genes were acquired by ancestral ascomycetes, and orthologs have subsequently been lost from multiple lineages, presumably reflecting lack of selection for their retention. The specialist *Metarhizium* species *M. album* and *M. acridum* lack homologs of the protease MAA_01413 that are present in generalist *Metarhizium* spp and *Fusarium*, suggesting that gene loss as well as gene gain has contributed to the diversification of *Metarhizium* spp. Three genes (polysaccharide lyase MAA_06640, β-xylosidase MAA_07693, and pyroglutamyl peptidase MAA_09637) were found in only *Metarhizium* spp; the pyroglutamyl peptidase was absent in *M. acridum* and *M. album*, and the other two were absent in *M. album*. It is possible that these three genes entered *Metarhizium* during the host-range expansion, but it is also possible that they were lost in the specialists because gene loss and gain are dynamic in *Metarhizium* (7).

It is noteworthy that only ∼1% of the genomes screened for this study were from entomopathogens (24), but this 1% of entomopathogens from diverse genera was the source of 37% of the non-*Metarhizium* fungal homologs of HGT genes. The non-entomopathogenic fungi harboring these genes included many with plant associations, including *Fusarium oxysporum* that infects the roots of more than 100 plant species in diverse ecological niches. Broad insect host-range *Metarhizium* species, including *M. robertsii*, also colonize roots, and the shared ecological and evolutionary lineage of these fungi could have been enabling factors in the acquisition and retention of genes in these genera. Shared ecological niches could also explain why half of the HGT genes clustered with actinobacterial sequences. However, soil bacteria are very diverse, suggesting that there may be unknown mechanism(s) driving preferential acquisition of actinobacterial genes.

The majority (∼70%) of the HGT genes were up-regulated on insect cuticle, and 50% of these were involved only in cuticle penetration, suggesting that the barrier imposed by insect cuticle serves as a major selective pressure for retention of HGT genes in entomopathogens. That the majority of the putative HGT genes have virulence phenotypes suggests that searching for HGT genes provides a mechanism for identifying unsuspected sets of proteins that sustain niche adaptation to insects. Consistent with the composition of insect cuticle, HGT genes important for cuticle penetration are fatty acid carriers and proteases. MAA_05244 and MAA_05652 acquired from arthropods act synergistically to utilize lipids on the outer layer of the insect cuticle. Insect pathogenic fungi use diverse P450 enzymes to hydroxylate epicuticular hydrocarbons into fatty acids (25, 26), and MAA_05244 and MAA_05652 will allow the resulting fatty acids to be metabolized. The third arthropod-derived HGT gene (MAA_03817) allows *M. robertsii* to take up hemocoel sterols (14). The triple gene-deletion mutant of the three lipid-carrier genes is almost avirulent, indicating that exploitation of host lipids via the three arthropod-derived lipid carriers is necessary for infection.

The inner layer (procuticle) of the insect cuticle is mainly composed of proteins and chitins; *M. robertsii* secretes numerous proteases and chitinases to degrade this layer and thereby gain entry into the hemocoel. It has long been thought that cuticle-degrading proteases function redundantly, so that disruption of one gene will only marginally impact fungal virulence (17). However, we found that the three HGT proteases are individually important in virulence. The pyroglutamyl peptidase MAA_09637 is responsible for stripping pyroglutamyl from cuticular proteins. Proteins that are N-terminally blocked with pyroglutamyl are protected from proteolysis (27). The endoproteases (MAA_00986 and MAA_01413) degrade insect cuticle, although to a smaller extent than proteinase K. Proteinase K shows 61% sequence similarity to Pr1, the most highly expressed *M. robertsii* subtilisin, but Pr1 is 33-fold more effective than proteinase K at degrading structural cuticle proteins (28). It is therefore likely that the virulence phenotype of MAA_00986 and MAA_01413 depends more on their ability to increase expression of chitinases and other cuticle-degrading proteases than it does their own ability to solubilize cuticle. *Metarhizium* proteases, including Pr1, are induced by protein-degradation products (29), which could result from the unique specificities of MAA_00986 and MAA_01413. Similarly, indirect induction of chitinases by the HGT proteases could follow proteolysis of the protein matrix covering cuticular chitin; the exposed chitin induces the expression of chitinases (30). Therefore, the three HGT proteases in *M. robertsii* are integrated into the complex cuticle-degrading enzyme system and influence the expression of other enzymes in the system. The mutant lacking the three protease genes and a lipid carrier (MAA_05652) was almost nonpathogenic, suggesting the collective actions of the HGT genes during penetration of cuticle are necessary for infection of insects.

Orthologs of one of the HGT proteases (MAA_00986) and the three HGT lipid carriers are present in all seven sequenced *Metarhizium* species. They are essential for infectivity to insects, suggesting that their acquisition and retention contributed to the emergence of entomopathogenicity in *Metarhizium*. Entomopathogenicity is advantageous, as it enables *Metarhizium* to escape competition from other microbes and build up population levels greater than the carrying capacity of the rhizosphere (31). Heterologous expression of the HGT protease (MAA_01413) up-regulated expression of *M. acridum*'s native cuticle-degrading enzymes on nonhost cuticle. This enabled the specialist to penetrate a nonnatural host cuticle and expand host range, suggesting that the triggers that regulate production of cuticle-degrading enzymes contribute to the host range of entomopathogenic fungi. In addition to *MAA_01413*, introduction of four other HGT genes also expanded the host range of *M. acridum*. Homologs of the five HGT genes identified in *M. robertsii* were present in all three generalists, but were lacking in the two specialists (*M. acridum* and *M. album*). *M. majus* and *M. guizhouense*, with intermediate host range, lacked three and two of the genes, respectively. In

Downloaded at Shanghai Jiao Tong University on January 5, 2021

conclusion, HGT was a key mechanism in the emergence of entomopathogenicity from a plant-associated ancestor of *Metarhizium*, and in subsequent broadening of host range in some *Metarhizium* lineages.

Materials and Methods

Further details of the study materials and methods are shown in *SI A S M M*.

Identification of Putative HGT Genes, Gene-Expression Analysis, and Construction of Gene-Deletion Mutants. Genomewide identification of HGT genes using a BLASTP-based method and Darkhorse were performed as described previously (21, 32). Phylogenetic analyses with maximum likelihood, Bayesian inference, or distance-based neighbor-joining were conducted as described previously (33–35). Tree topology comparisons using CONSEL were conducted as described previously (36, 37). GC content was calculated by using CodonW (38). K_a/K_s ratio was calculated by using KaKs_Calculator v2.0 (39). qRT-PCR analysis of gene expression was conducted as described previously (15). Gene deletion and gene knockdown using antisense RNA method were conducted as described previously (40, 41).

Bioassays. Bioassays were conducted using *G. m* larvae (Ruiqing Bait) (14).

Assaying Activities of Cuticle Degrading Enzymes. Quantification of amino acids and peptides released during enzymolysis of insect cuticle and assays of chitinase activity were conducted with commercially available kits. Activities of Pr1 proteases and total extracellular proteases were assayed as described previously (42).

The activity of pyroglutamyl peptidase MAA_09367 was tested by using L-pyroglutamyl-naphthylamine (27). The substrate specificities of the proteases MAA_01413 and MAA_00986 were assayed as described previously (42).

ACKNOWLEDGMENTS. This work was supported by National Natural Science Foundation of China Grants 31471818 and 31672078, and National Key R&D Program of China Grant 2017YFD0200400.

1. Frick WF, et al. (2010) An emerging disease causes regional population collapse of a common North American bat species. *S* 329:679–682.
2. Berger L, et al. (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *P N A S USA* 95:9031–9036.
3. Kim K, Harvell CD (2004) The rise and fall of a six-year coral-fungal epizootic. *A N* 164:S52–S63.
4. Cameron SA, et al. (2011) Patterns of widespread decline in North American bumble bees. *P N A S USA* 108:662–667.
5. Archie EA, Luikart G, Ezenwa VO (2009) Infecting epidemiology with genetics: A new frontier in disease ecology. *T E E* 24:21–30.
6. Raffaele S, et al. (2010) Genome evolution following host jumps in the Irish potato famine pathogen lineage. *S* 330:1540–1543.
7. Hu X, et al. (2014) Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. *P N A S USA* 111:16796–16801.
8. Zhang X, St Leger RJ, Fang W (2017) Pyruvate accumulation is the first line of cell defense against heat stress in a fungus. *MB* 8:e01284-17.
9. Gao Q, et al. (2011) Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *M* and *M.* *PL S G* 7:e1001264.
10. Barelli L, Moonjely S, Behie SW, Bidochka MJ (2016) Fungi with multifunctional lifestyles: Endophytic insect pathogenic fungi. *P M B* 90:657–664.
11. Lu HL, St Leger RJ (2016) Insect immunity to entomopathogenic fungi. *A G* 94:251–285.
12. Gluck-Thaler E, Slot JC (2015) Dimensions of horizontal gene transfer in eukaryotic microbial pathogens. *PL S P* 11:e1005156.
13. Soanes D, Richards TA (2014) Horizontal gene transfer in eukaryotic plant pathogens. *A R P* 52:583–614.
14. Zhao H, et al. (2014) Host-to-pathogen gene transfer facilitated infection of insects by a pathogenic fungus. *PL S P* 10:e1004009.
15. Guo N, et al. (2017) Alternative transcription start site selection in Mr-OPY2 controls lifestyle transitions in the fungus *M*. *N C* 8:1565.
16. Fang W, Pava-ripoll M, Wang S, St Leger RJ (2009) Protein kinase A regulates production of virulence determinants by the entomopathogenic fungus, *M*. *G B* 46:277–285.
17. St Leger RJ, Charnley AK, Cooper RM (1986) Cuticle-degrading enzymes of entomopathogenic fungi: Mechanisms of interaction between pathogen enzymes and insect cuticle. *J I P* 47:295–302.
18. Spear RN, Cullen D, Andrews JH (1999) Fluorescent labels, confocal microscopy, and quantitative image analysis in study of fungal biology. *M E* 307:607–623.
19. Friesen TL, et al. (2006) Emergence of a new disease as a result of interspecific virulence gene transfer. *N G* 38:953–956.
20. Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *A R P* 46:75–100.
21. Sun BF, et al. (2013) Multiple interkingdom horizontal gene transfers in *P* and closely related species and their contributions to phytopathogenic lifestyles. *PL S O* 8:e60029.
22. de Jonge R, et al. (2012) Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. *P N A S USA* 109:5110–5115.
23. Xiao G, et al. (2012) Genomic perspectives on the evolution of fungal entomopathogenicity in *B*. *S R* 2:483.
24. Wang C, Wang S (2017) Insect pathogenic fungi: Genomics, molecular interactions, and genetic improvements. *A R E* 62:73–90.
25. Lin L, et al. (2011) The MrCYP52 cytochrome P450 monoxygenase gene of *M* is important for utilizing insect epicuticular hydrocarbons. *PL S O* 6:e28984.
26. Zhang S, et al. (2012) CYP52X1, representing new cytochrome P450 subfamily, displays fatty acid hydroxylase activity and contributes to virulence and growth on insect cuticular substrates in entomopathogenic fungus *B*. *J B C* 287:13477–13486.
27. Dando PM, Fortunato M, Strand GB, Smith TS, Barrett AJ (2003) Pyroglutamyl-peptidase I: Cloning, sequencing, and characterisation of the recombinant human enzyme. *P E* 28:111–119.
28. St Leger RJ, Frank DC, Roberts DW, Staples RC (1992) Molecular cloning and regulatory analysis of the cuticle-degrading-protease structural gene from the entomopathogenic fungus *M*. *J B* 204:991–1001.
29. Freimoser FM, Screen S, Bagga S, Hu G, St Leger RJ (2003) Expressed sequence tag (EST) analysis of two subspecies of *M* reveals a plethora of secreted proteins with potential activity in insect hosts. *M* 149:239–247.
30. Screen SE, Hu G, St Leger RJ (2001) Transformants of *M* sf. overexpressing chitinase from *M* sf. show 2 4 7 . (- j 9