Eukaryotic transposable elements as “cargo carriers”: the forging of metal resistance in the fungus *Paecilomyces variotii*

Andrew S. Urquhart¹, Nicholas F. Chong¹, Yongqing Yang^{1,2}, Alexander Idnurm¹

¹School of BioSciences, the University of Melbourne, VIC, 3010, Australia

²Inner Mongolia Academy of Agricultural & Animal Husbandry Sciences, Hohhot 010031, China

Correspondence: A. Idnurm; email alexander.idnurm@unimelb.edu.au | phone +61 03 8344 2221

ABSTRACT

The horizontal transfer of large gene clusters by mobile elements is a key driver of prokaryotic adaptation in response to environmental stresses. Eukaryotic microbes face similar environmental stresses yet a parallel role for mobile elements has not yet been established. A stress faced by all microorganisms is the prevalence of toxic metals in their environment. In fungi, identified mechanisms for protection against metals generally rely on genes that are dispersed within an organism's genome. Here we have discovered a large (~85 kb) region that confers resistance to several metals in the genomes of some, but not all, strains of a fungus, *Paecilomyces variotii*. We name this region *HEPHAESTUS* (*H ϕ*) and present evidence that this region is mobile within the *P. variotii* genome with features highly characteristic of a transposable element. While large gene clusters including those for the synthesis of secondary metabolites have been widely reported in fungi, these are not mobile within fungal genomes. *HEPHAESTUS* contains the greatest complement of host-beneficial genes carried by a transposable element in eukaryotes. This suggests that eukaryotic transposable elements might play a role analogous to their bacterial counterparts in the horizontal transfer of large regions of host-beneficial DNA. Genes within *HEPHAESTUS* responsible for individual metal resistances include those encoding a P-type ATPase transporter, PcaA, required for cadmium and lead resistance, a transporter, ZrcA, providing resistance to zinc, and a multicopper oxidase, McoA, conferring resistance to copper. Additionally, a subregion of *H ϕ* conferring resistance to arsenate was identified. The presence of a strikingly similar cluster in the genome of another fungus, *Penicillium fuscoglaucum*, suggests that *HEPHAESTUS* arrived in *P. variotii* via horizontal gene transfer.

KEYWORDS

Eurotiales, gene cluster, metal homeostasis, horizontal gene transfer, transposon

INTRODUCTION

Metals are used within diverse fields such as agriculture, construction, electronics and pharmaceutical science. A number of these, such as cadmium, lead and arsenic, are toxic even at low levels of exposure and thus environmental contamination resulting from their use is an increasing concern ¹. The prevalence of toxic metals in the environment is of particular relevance to fungal biology. Some fungi are able to accumulate metals to high concentrations, making such fungi dangerous for human consumption ^{2,3}, allowing bioremediation of contaminated sites ⁴ or concentrating valuable metals for extraction ⁵. Zinc and copper play essential roles in many cellular functions and are often important in fungal virulence ⁶. However, at high concentrations these metals can also be toxic; for example, copper-based molecules are frequently used as fungicides in agriculture.

Compared to prokaryotes, our understanding of the evolutionary mechanisms through which eukaryotes have adapted to environmental stresses – such as to metal toxicity – is relatively limited. One fungal example is the recent expansion of the *CUP1* genes encoding metallothioneins in wine and sake-producing strains of *Saccharomyces cerevisiae*. This gene expansion is hypothesized to be a response to the use of copper-based materials during fermentation and/or copper sulfate fungicides in vineyards ^{7,8}.

In prokaryotes, horizontal transfer of integrative and conjugative elements (ICEs) has played a key role in the evolution of heavy metal tolerance. ICEs are mobile pieces of DNA that are inserted into the host genome and encode the genes necessary for conjugation thus allowing their spread between bacterial hosts. These elements frequently carry additional host-beneficial cargo genes facilitating the transfer of key traits such as resistance to antibiotics. There are several examples of ICE elements carrying the genes required for tolerance to heavy metals; for example, tolerance to copper in *Pseudomonas syringae* ⁹, copper and arsenic in *Salmonella enterica* ¹⁰, and copper and zinc in *Histophilus somni* ¹¹. The acquisition of ICEs is thus a means by which bacterial strains can evolve resistance to heavy metals. The potential role of horizontal gene transfer events in the evolution of metal tolerance has not been explored in eukaryotes. However, horizontal gene transfer, while less extensive than in prokaryotes, is now recognized as an important evolutionary force in other areas of eukaryote biology including fungal virulence and nutrient utilization ^{12,13}.

The horizontal transfer of multigene traits is facilitated by the fact that filamentous fungi feature distinct regions in which genes associated with the same process are physically clustered, especially for the synthesis of secondary metabolites¹⁴⁻¹⁶. Other gene clusters play roles in primary metabolism such as nitrate assimilation¹⁷ and biotin synthesis¹⁸. Small clusters of genes involved in metal homeostasis such as iron uptake and arsenic detoxification have been identified in fungi¹⁹. In *S. cerevisiae* three linked genes mediate resistance to arsenicals²⁰. How such clusters are transferred between species is not understood. Although it has recently been hypothesized that gene clusters are horizontally transferred between eukaryotic species on large transposable elements^{21,22}, transposons carrying large regions of host-beneficial DNA have not been reported.

Paecilomyces variotii is a cosmopolitan member of the Eurotiales that is capable of growing under a wide range of environmental conditions. For example, it is noteworthy for its high level of resistance to heat which causes it to be a frequent cause of food spoilage in heat treated foods^{23,24}. It has also been isolated from diverse environments including as a plant endophyte²⁵, marine sources²⁶, human infections²⁷, and indoors. The ecology of this species is currently poorly understood.

Recently, a genetic transformation system has been developed and the genomes of three *P. variotii* strains have been sequenced^{28,29}. Here, we describe the discovery and functional analysis of part of the genome, absent from some strains, that features genes encoding proteins involved in metal detoxification or protection. This gene cluster is distinct from previously described clusters associated with heavy metal resistance because it encodes > 30 genes and provides resistance to at least five different metals: arsenic, cadmium, copper, lead and zinc. Furthermore, we present evidence that this region is mobile within the *P. variotii* genome and may have arisen through horizontal gene transfer, suggesting a role for transposable elements in the transfer of large genomic regions conferring beneficial traits between fungi.

RESULTS

The HEPHAESTUS region is only present in certain *Paecilomyces variotii* strains.

Comparison of two *P. variotii* strains revealed a region of the genome present within strain CBS 144490 but absent from CBS 101075 (Figure 1a). Examination of the predicted genes within

this region revealed that many have similarities to those involved in protection against various toxic metals (Supplementary Table 1). We name this genetic element *HEPHAESTUS* (*H ϕ*), after the Greek god of metallurgy. Genes within the cluster were annotated based on sequence similarity to homologs in other fungi. Sequencing four additional *P. variotii* strains revealed the presence of the cluster in strains FRR 1658, FRR 3823 and FRR 5287, and its absence from strain FRR 2889 (Figure 1b). A recently sequenced strain of *P. variotii*, MTDF-01, also lacks this region²⁸. Neither of the other published *Paecilomyces* genomes, *P. formosus* strain number 5³⁰ or *Paecilomyces niveus* strain CO7³¹, have the cluster. “*Byssochlamys*” (anamorph *Paecilomyces*) strain BYSS01³² has part of the cluster, but this genome appears to represent a *Monascus* species rather than *Paecilomyces* based on BLAST searches of phylogenetic markers (data not shown).

***HEPHAESTUS* is a mobile element within the *Paecilomyces* genome.**

The edges of the cluster feature an inverted repeat AATAT[A/T]GGGACCTGTGA, where the [A/T] is different between the two copies. Furthermore, one gene encodes a Tc-1 like transposase, another contains an RNase H domain homologous to that found in reverse transcriptases, and there are also remnants of other transposable elements. To explore the possibility that *H ϕ* is a mobile element we compared the genomic location of *H ϕ* within the four sequenced *P. variotii* strains that have the cluster (*H ϕ* + strains). Each contained *H ϕ* at an independent location on a different assembly contig (Figure 1c). Examination of the *H ϕ* integration sites revealed clear patterns: *H ϕ* appears to have been inserted into either a TTAC site (in FRR 1658, FRR 3823, CBS 144490) or a TAAC site (in FRR 5287) (Figure 1d). In each case the target site showed microhomology to the edges of *H ϕ* suggesting that microhomology is important for *H ϕ* transposition. A further distinctive feature was the presence of the “ACAAC TAG” motif next to the right terminal repeat. We hypothesize that this sequence must be carried as part of the *H ϕ* sequence or synthesized during integration.

Fig. 1 | HEPHAESTUS ($H\phi$), a large gene cluster in *P. variotii*. **a**, Map of the HEPHAESTUS region and flanking DNA in *P. variotii* CBS 144490. Genes with putative roles in metal resistance based on homology are highlighted in orange. Genes with similarity to those involved in transposition of transposons are highlighted in red. Flanking genes conserved in the $H\phi$ -strain CBS 101075 in black. TIR; terminal inverted repeats. Tick marks represent 1 kb of nucleotide sequence. **b**, Read mapping to the $H\phi$ region shows that of the four newly sequenced strains FRR1658, FRR 3823 and FRR 5287 contain $H\phi$ but FRR 2889 does not. **c**, Genomic location of *Hephaestus* in four strains annotated onto the 8 largest contigs of the *P. variotii* strain MTDf-01 assembly. Three other sequenced strains, CBS 101075, FRR 2889 and MTDf-01, do not contain $H\phi$. **d**, Sequences of the insertion sites in the four $H\phi$ + *P. variotii* strains. Features include the terminal inverted repeats (TIR; red), the “ACAACTAG” motif (purple), the TTAC/TAAC sequence (bold) and regions of microhomology (green).

HEPHAESTUS shows evidence of horizontal gene transfer.

Examination of other Eurotiales genomes in GenBank via BLAST identified a remarkably similar gene cluster in *Penicillium fuscoglaucum* strain FM041³³. The shared genes between this cluster and HEPHAESTUS show a high level of conservation (Figure 2C) suggesting the occurrence of a horizontal gene transfer event in the evolutionary history of this region. While the left border of the *P. fuscoglaucum* cluster lies beyond the edge of the sequencing contig, the right hand terminal inverted repeat is preserved and the “ACAACTAG” motif is replaced by the similar sequence ACAACATG. Adjacent to this ACAACATG sequence are two thymine nucleotides presumably derived from an TTAC target site. This provides evidence that the *P. fuscoglaucum* cluster may be or have been similarly mobile.

Fig. 2 | a, Comparison of the *Penicillium fuscoglaucum* gene cluster and *P. variotii* $H\phi$. Homologous genes share the same color. **b**, The right terminal repeat is conserved in *P. fuscoglaucum* along with sequence similar to the “ACAACTAG” motif and the TT sequence presumably resulting from integration into a TTAC site. **c**, Graph of the top results of a BLAST search for all predicted *P. variotii* genes against the *P. fuscoglaucum* assembly scaffolds. Common genes between the *P. fuscoglaucum* putative metal resistance cluster and $H\phi$ (colored-coded as in panel **a**) show greater than expected conservation.

The presence of *HEPHAESTUS* increases resistance of *P. variotii* to cadmium, lead, zinc, arsenic and copper.

To test the hypothesis that the presence of *HEPHAESTUS* impacts the fitness in the presence of toxic metals we tested six *P. variotii* strains CBS 144490, FRR 1658, FRR 3823, and FRR 5287 ($H\phi$ +), and FRR 2889 and CBS 101075 ($H\phi$ -) for their ability to grow on different metal ions (Supplemental Figure 1). The $H\phi$ + strains have increased resistance to zinc, cadmium, lead and copper ions. To confirm that these resistances were conferred by *HEPHAESTUS* we examined the segregation of the resistance phenotype in the progeny of a sexual cross between two strains, CBS 144490 ($H\phi$ +) and CBS 101075 ($H\phi$ -) (Figure 3). Resistance to all four metal ions segregated with the presence of *HEPHAESTUS*. Furthermore, while characterizing resistance to arsenate we noticed a difference in early spore germination and hyphal growth between strains CBS 101075 ($H\phi$ -) and CBS 144490 ($H\phi$ +) in the presence of arsenate after 24 hours. This difference was less pronounced by four days (Figure 3).

Fig. S1 | Growth for four days at 30 °C of six *P. variotii* strains on a range of concentrations of various metals. Strains with the *Hφ* cluster are more resistant to cadmium, lead, copper and zinc than those strains lacking it.

Fig. 3 | Resistance to five metal ions, Cu^{2+} , Cd^{2+} , Zn^{2+} , Pb^{2+} and arsenate, segregates with the presence of *Hφ* in progeny of a cross between CBS 144490 (*Hφ*+) and CBS 101075 (*Hφ*-). **a**, PCR analysis of the presence or absence of *Hφ* in progeny, with the mating type locus used as a control for independent segregation. **b**, Growth at 30 °C of the two parental and 22 ascospore-derived progeny strains in cleared Campbells V8 juice plates (CV8) in presence of CuSO_4 , CdSO_4 or ZnSO_4 for 2 days; PbNO_3 for 3 days. Strains were grown overnight with Na_2HAsO_4 in liquid media and then stained with Congo Red to increase contrast.

A P-type ATPase encoded by *pcaA* within *HEPHAESTUS* confers resistance to cadmium and lead.

Having established that *Hφ* provides resistance to toxic metals we sought to identify the specific genes conferring these resistances. The *pcaA* gene in *Hφ* was investigated as a candidate for cadmium resistance because it shows striking sequence similarity to a previously characterized cadmium transporters in *Penicillium janthinellum*³⁴ and *Aspergillus fumigatus*³⁵. Consistent with this hypothesis, transformation of the *pcaA* gene from *Hφ* into the *Hφ*- strain CBS 101075 resulted in a strain with increased resistance to Cd²⁺ and Pb²⁺ (Figure 4a). Replacement of the *pcaA* coding sequence with Green Fluorescent Protein (GFP) in CBS 144490 resulted in sensitivity to cadmium. These experiments confirmed that *pcaA* in the *Hφ* cluster is responsible for cadmium resistance.

An unexpected finding was that in addition to conferring cadmium resistance, PcaA also confers resistance to Pb²⁺ (Figure 4a). The growth pattern differs from that of sensitivity to other metals, with rapid radial growth in strains without *Hφ* at some concentrations of lead ions, with its role in resistance most clearly illustrated at 5 mM or higher. Previous analysis of close homologs^{34,35} did not investigate the transport of lead; however, P-type ATPases which transport both cadmium and lead have been identified, for example AtHMA3 in the plant *Arabidopsis thaliana*³⁶. We cannot yet conclude whether the selective advantage of PcaA is primarily in response to cadmium or lead (or both).

To determine if the expression of the *pcaA* promoter in *P. variotii* CBS 144490 is regulated by these metals, a strain in which the reading frame of *pcaA* had been replaced by GFP was examined. This strain displayed green fluorescence when exposed to cadmium or lead ions, whilst low levels of fluorescence, similar to the background fluorescence of an untransformed strain, were observed in cultures without cadmium (Figure 4b).

Fig. 4 | a, Deletion of *pcaA* in CBS14490 decreases resistance to cadmium and lead ions, and conversely transformation of *pcaA* into CBS101075 increases resistance to cadmium and lead ions demonstrating the role of *pcaA* in zinc resistance. **b**, GFP expression in the *pcaA*-GFP replacement strain increases in response to both cadmium and lead demonstrating that expression of the *pcaA* gene increases in response to high concentrations of cadmium or lead.

Zinc resistance is conferred by the *zrcA* gene.

In contrast to cadmium, which is a highly-toxic non-essential metal, zinc is essential for many cellular functions³⁷. However, high levels of zinc are toxic both to humans and microorganisms³⁸. BLAST searches identified a gene, *zrcA*, with homology (36% amino acid identity) to the transporter Zrc1 that confers resistance to zinc in *Saccharomyces cerevisiae*³⁹. We investigated whether this gene was primarily responsible for the zinc resistance conferred by *Hφ*. Transformation of *P. variotii* strain CBS 101075 with *zrcA* resulted in a strain with increased levels of resistance to zinc ions and replacement of the *zrcA* reading frame with GFP in CBS 144490 resulted in sensitivity to zinc (Figure 5a). Furthermore, the *zrcA* cDNA was able to complement the sensitivity of a *zrc1Δ cot1Δ S. cerevisiae* mutant to 5 mM ZnSO₄ (Figure 5b). Fluorescence in the GFP replacement strain increased when zinc was added to the media, demonstrating that *zrcA* is upregulated in response to high concentrations of zinc (Figure 5c).

A strain of *P. variotii* in which GFP was fused behind ZrcA and expressed under the constitutive *Leptosphaeria maculans* actin promoter was developed to investigate the subcellular localization of ZrcA. This strain showed GFP fluorescence consistent with plasma membrane localization, as well as within the vacuole (Figure 5d). This contrasts to the known localization of Zrc1 homologs. In *S. cerevisiae* and *Cryptococcus neoformans* Zrc1 is localized to the vacuolar membrane and leads to the accumulation of zinc in the vacuole^{40,41}. However, in *Candida albicans* and *Hebeloma cylindrosporum* the Zrc1 homologs appear to be localized to the endoplasmic reticulum and leads to zinc storage in so called “Zincosomes”^{42,43}. The presence of *P. variotii* ZrcA in this membrane suggests a possible role in exporting zinc ions out of the cell rather than in involvement in intracellular zinc storage as in the case of Zrc1 in *S. cerevisiae*, *C. albicans* and *H. cylindrosporum*.

Fig. 5 | a, Deletion of *zrcA* in CBS 144490 decreases resistance to zinc ions and addition of *zrcA* to CBS 101075 increases resistance to zinc ions. **b**, A *zrcA*-containing construct complements the zinc sensitivity of a *zrc1Δ cot1Δ* *S. cerevisiae* mutant, unlike the empty vector (*pTH19*). **c**, GFP expression in the *zrcA*-GFP replacement strain increases in response to zinc demonstrating that expression of the *zrcA* gene is increased in response to high concentrations of zinc ions. **d**, A *zrcA*-GFP fusion protein localizes largely to the cell membrane indicating a possible role in export of zinc from the cell.

A multigene region of *Hφ* confers arsenic resistance.

A sub-region of *Hφ* contains a cluster of genes putatively involved in arsenic resistance, including homologs of *arsH*⁴⁴, *arsC*⁴⁵, and *arsB*⁴⁶ as well as genes encoding a cyclin similar to Pho80⁴⁷ and an arsenic methyltransferase⁴⁸ (Figure 6a). Deletion of a 9,478 bp region containing most of these genes resulted in sensitivity to arsenic and complementation of this mutant with these genes restored arsenic resistance (Figure 6b). The key gene(s) responsible for the increase in arsenic tolerance associated with *Hφ* within this region have yet to be identified and it is possible that some of these genes act redundantly or synergistically. In addition to the genes in the cluster, other copies of a subset of these genes are found in the *P.*

variotii genome. Further studies of this region in *P. variotii* or other species will provide greater insight into the mechanisms of arsenic resistance in filamentous fungi.

One explanation for the presence of the genes for arsenic detoxification within *H ϕ* is insertion of the region into an ancestor sequence similar to the *Penicillium fuscoglaucum* cluster, which lacks these genes. This hypothesis is supported by the presence of two putative pseudogenes that have been created at the right junction of this hypothesized insertion. The first is *hhpH*, which is a truncated putative iron transporter. The *P. fuscoglaucum* cluster contains a full-length homolog spanning across this junction, suggesting that the arsenic region was mostly likely inserted into the ancestral cluster rather than deleted from it. The second pseudogene is *hhpI*, which is a truncated homolog of a conserved group of proteins containing a domain of unknown function (DUF3435). A model describing how such a junction may have formed is presented in Figure 6c.

Fig. 6 | a, The putative arsenic-associated region contains 10 genes (in addition to *hhpI*, which is possibly a pseudogene) in *P. variotii*. A sub-region containing six of these genes was deleted by homologous recombination. **b**, This deletion resulted in strains (Ars KO) with significantly reduced growth (mycelial weight) in 7 mM arsenate after 48 h. Complementation of this mutant (Ars Comp) restored arsenic resistance. **c**, Model explaining the origin of the arsenic region in *H ϕ* as an insertion of a preformed arsenic cluster into a *Penicillium fuscoglaucum* cluster-like ancestor.

The multicopper oxidase McoA confers resistance to copper.

Despite the important use of copper-based fungicides for over 100 years, knowledge of copper resistance in filamentous fungi is limited. We hypothesized that the *mcoA* gene encoding a putative multicopper oxidase within the *HEPHAESTUS* region might confer resistance to copper ions, since other multicopper oxidases confer resistance to copper ions in both bacteria (e.g. CueO in *Escherichia coli*⁴⁹) and fungi (e.g. Fet3 in *S. cerevisiae*⁵⁰). Consistent with this hypothesis, transformation of CBS 101075 with *mcoA* conferred resistance to copper ions (Fig. 7a). Furthermore, qPCR analysis showed that expression of *mcoA* in 1.5 mM copper was 4-5 times higher than when strain CBS 144490 was grown without copper (Fig. 7b). This is the first multicopper oxidase conferring resistance to copper reported from a filamentous fungus.

Fig. 7 | The multicopper oxidase McoA encoded within *Hφ* confers resistance to copper ions. **a**, Transformation of *mcoA* into CBS 101075 (*Hφ*⁻) provides increased resistance to copper, suggesting that McoA contributes to the copper resistance conferred by *Hφ*. **b**, qPCR analysis shows increased expression of *mcoA* in the presence of copper in strain CBS 144490.

*Differences between expression at different copper concentrations are statistically significant ($P < 0.01$) as assessed using a two-tailed t-tests assuming unequal variance. Independent biological replicates were measured, 0 mM and 1.5 mM $n = 4$; 0.8 mM $n = 3$. Error bars represent plus or minus one standard deviation from the mean.

DISCUSSION

Most cases of horizontal DNA transfer in eukaryotes involve transposons⁵¹. Transposons have been widely characterized in eukaryotes since their discovery in maize 70 years ago⁵². However, the potential benefits of genes carried within them to the host are not usually apparent. One exception is a putative transposon, ToxhAT, involved in the movement of the

ToxA host specificity toxin between plant pathogens⁵³. ToxhAT contains 14 kb of DNA flanked by TIRs and appears to have horizontally transferred as a transposon between *Parastagonospora/Pyrenophora* and *Bipolaris*. However recent mobility in the *Bipolaris sorokiniana* genome appears to be as part of a broader genomic island rather than through its own transposition. An important difference between ToxhAT and *Hφ* is that a single gene within ToxhAT, namely ToxA, confers the phenotype as opposed to a multitude of genes. Recently, increasingly large transposable elements have been found in eukaryotes. These include helitrons in maize capable of capturing up to nine gene fragments and totaling up to 39 kb⁵⁴, and the tetratron elements of fish that are well over 150 kb²¹. The discovery of these large transposable elements in eukaryotes has prompted the hypothesis that such transposons might be capable of carrying greater amounts of useful “cargo” including entire gene clusters²².

Here we describe *Hφ*, a large (~85 kb) region in a fungal genome that contains a number of genes involved in resistance to toxic metal ions. These genes include a P-type ATPase conferring resistance to cadmium and lead, a *zrc1* homolog conferring resistance to zinc, a multicopper oxidase conferring resistance to copper and subregion of the cluster conferring resistance to arsenic compounds. Further, we present evidence that *Hφ* is a transposable element. The four *Hφ*⁺ strains sequenced each contain *Hφ* at a unique genomic location on separate assembly contigs demonstrating mobility within the *P. variotii* genome (Figure 1c). Moreover, examination of the flanking sequences of *Hφ* suggests an active mechanism of transposition given the perfect preservation of the *Hφ* border sequences (including terminal inverted repeats) and preference for particular target sites (Figure 1d). Often, transposition of DNA transposons involves target site duplication⁵⁵. *Hφ* does not appear to employ such duplications but rather makes use of microhomology to the target site. This is reminiscent of Crypton transposons, which also make use of microhomology to the target site⁵⁶. However, Cryptons differ from *Hφ* in key respects including that they lack TIRs. There is also precedent for an absence of target site duplication, namely the *Spy* transposons⁵⁷. Another distinctive feature is that *Hφ* carries an “ACAAC TAG” motif outside one of the terminal inverted repeats; we are not aware of a similar feature in other transposons.

Hφ carries genes with homology to those involved in transposition, including a putative Tc-1 like transposase, and a reverse transcriptase homolog, which may play a role in the movement of *Hφ*, as might the pair of highly conserved internal repeats which initially frustrated DNA

sequencing assembly. While the mechanisms of *Hφ* transposition will require further studies to illuminate them, *Hφ* is likely capable of transposition by a specific and possibly novel mechanism given that the integration sites do show a consistent pattern different from known transposons. At 85 kb *Hφ* represents an exceptionally large fungal transposon (although smaller than the Teratorn element in fish ²¹) and carries substantially more host beneficial genes than any previously described transposable element. Other examples of beneficial gene clusters in fungi (including those involved in secondary metabolite synthesis) have also been hypothesized to move via horizontal gene transfer ¹⁶ but do not show evidence of transposition within the genome.

A comparable region encoding for metal resistance has not previously been described in fungi, with *Hφ* being set apart both in the number of genes encoded (> 30), the number of metal ion resistances conferred (≥ 5), and evidence of mobility. However, there are parallels with bacterial operons for metal ion detoxification. For instance, ICEs can contain large sets of genes that are involved in the detoxification of multiple metal ions, and such elements are able to spread through replication as plasmids and via conjugation with other bacteria ^{9,10,58}.

Evidence for a past horizontal DNA transfer event in the evolutionary history of the *Hφ* cluster is provided by the strikingly high nucleotide similarity between elements of this cluster and a similar, albeit incomplete, cluster in *P. fuscoglaucum*. Curiously, *P. fuscoglaucum* is from a group of fungi noted for a high level of HGT ^{16,59,60}. The high level of nucleotide identity between the two clusters (~95%) is inconsistent with vertical gene transmission from an ancient common ancestor. We thus suggest that *HEPHAESTUS* most likely represents a recent horizontal DNA transfer event conferring resistance to various metals in certain *P. variotii* strains. This implies that as in the case of prokaryote ICEs ⁹⁻¹¹, HGT is likely to play a role in evolutionary adaptation to metal stress in fungi.

The *P. variotii* strains in this study were not accompanied by detailed collection data so we were unable to draw correlations between the presence of *HEPHAESTUS* and the ecological source of the strains. However, given the seemingly diverse and poorly understood ecological niches which *P. variotii* appears to fill, an important subject for future studies will be the distribution of the cluster from a population biology perspective in a greater number of strains.

An implication of carrying host beneficial genes on a mobile element in *P. variotii* is that we expect that such regions will be subjected to Repeat Induced Point (RIP) mutation should they become duplicated in the *P. variotii* genome²⁹. RIP is a fungal defense against transposable elements that inactivates duplicated DNA during sexual reproduction through C to T mutations⁶¹. Thus, we predict that the “domestication” of these genes in a stable genomic location may be favored. This selective pressure may limit the life cycle of positively selected transposons. Supporting this is the presence of a P-type ATPase in *Aspergillus clavatus* strain NRRL 1 with unusually high homology (93% nucleotide identity) to the *pcaA* gene of *Hφ* (GenBank XM_001276247.1⁶²) yet not located in a similar genomic context. We speculate that this gene may represent a domesticated relic of an element similar to *Hφ*. If these elements are not evolutionarily stable, then the window of time in which they can be observed may be narrow and thus their importance as drivers of evolutionary change obscured.

In summary, we have identified a new mobile element in fungi involved in tolerance to toxic metal ions. This region contains over 30 genes many of which appear to have potential roles in metal detoxification, some of which we have characterized in this study. *Hφ* is the first eukaryotic transposable element found to carry such a significant complement of host-beneficial genes and highlights the possible role of large transposable elements in the horizontal transfer of large genomic regions between eukaryotes. Future studies should explore its origins in related Eurotiales species and the mechanism by which such a large region is able to transpose.

METHODS

DNA sequence refinement and annotation of the *Hφ* region. Examination of the next generation sequencing reads of *P. variotii* CBS 144490 assembled to contig 123, where the *Hφ* gene cluster is located, revealed a number of regions of ambiguity. Furthermore, comparison of the genes by BLAST against fungal genome sequences in GenBank identified a related cluster found in the cheese-associated species *Penicillium fuscoglaucum*³³, whose arrangement of genes also included homologs in *P. variotii* found on contig 14, which is another region found in strain CBS 144490 and not CBS 101075. Hence, specific regions were amplified by PCR and sequenced to provide a high confidence level of assembly, which joins contig 14 into the middle of the cluster on contig 123. The reason that this was not assembled by the program Velvet is likely due to ~300 bp of perfectly repeated DNA flanking this region. The complete *Hφ* sequence

from strain CBS 144490 has been deposited to GenBank (accession MT022027). Wild type strains can be obtained from culture collections and genetically manipulated strains can be provided by the authors on request.

Strains and culturing. The two *P. variotii* strains with their genomes sequenced are CBS 101075 and CBS 144490; these strains were routinely cultured on cleared Campbell's V8 juice (CV8) agar. Four additional strains FRR 1658 (*H ϕ* +, ITS GenBank MN880081), FRR 2889 (*H ϕ* -, ITS GenBank MN880082), FRR 3823 (*H ϕ* +, ITS GenBank MN880083) and FRR 5287 (*H ϕ* +, ITS GenBank MN880084) were obtained from the FRR culture collection maintained by CSIRO (North Ryde, New South Wales, Australia). Correct identification of the four FRR strains was confirmed using an ITS phylogenetic tree (Supplementary Figure 2) generated in MrBayes⁶³ based on a previous taxonomic study of the *Paecilomyces* genus⁶⁴.

Escherichia coli strain NEB® 5-alpha was used for propagation of plasmid DNA. *Agrobacterium tumefaciens* strain EHA105 was used for transformation of *P. variotii*. The bacterial strains were maintained on Luria broth (LB) agar, supplemented with either kanamycin (50 mg/L) or ampicillin (50 mg/L) in plasmid-carrying strains. *Saccharomyces cerevisiae* strain BY4742 and derivatives was used for *zrcA* complementation experiments. *S. cerevisiae* was cultured on yeast extract peptone dextrose (YPD), YPD +G418 (for selection of the KanMX construct), yeast nitrogen base (YNB) -leucine (for selection of the *leu2* construct) or YNB -uracil (for selection of the pTH19 plasmid).

Fig. S2 | Phylogeny of *Paecilomyces* strains based on the ITS region, illustrating that four of the FRR culture collection strains are *P. variotii*. Strains in which the presence or absence of *Hφ* is known are in bold font. Posterior probabilities are given for each clade.

Genetic segregation. Crosses between strains CBS 101075 and CBS 144490 were conducted as previously described, and 22 ascospore-derived colonies obtained^{24,29}. The mating type of each progeny was determined by PCR using primers MAI0448, MAI0449, MAI0450 and MAI0451. The presence or absence of *Hφ* was assessed using primer pairs AP89 and AP90, specific for *Hφ*, and AP92 and AP93, which amplify across the *Hφ* insertion site. Primer sequences are provided in Table S2.

Cloning and DNA manipulations. The *pcaA* and *zrcA* open reading frames were amplified with primer pairs AP104 and AP105 and AP297 and AP298, and cloned into plasmid PLA2 linearized with BglII, which enables constitutive expression from the actin promoter of *Leptosphaeria maculans*^{29,65}. A construct expressing a ZrcA-GFP fusion protein was generated by amplifying the *zrcA* reading frame with primers AP351 and AP352, and the *GFP* reading frame with primers AP353 and AP354 and cloning into the BglII site of plasmid PLA2.

Constructs for replacement of the *pcaA* and *zrcA* genes with Green Fluorescent Protein (GFP) via homologous recombination were made by PCR amplification of the required gene regions followed by cloning of these PCR products into plasmid pPZP-201BK⁶⁶ linearized with EcoRI and HindIII. For *pcaA* the regions of homology were amplified using primer pairs AP219 and AP220 and AP223 and AP224. The *GFP-hph* region was amplified from a GFP-tagging construct generated previously⁶⁷ using primers AP221 and AP222. For *zrcA* the regions of homology were amplified using primer pairs AP299 and AP302 and AP305 and AP307. The *GFP-hph* was amplified using primers AP303 and AP304.

A construct for the deletion of a 9,478 bp region containing genes putatively involved in arsenic resistance was similarly constructed using the NEBuilder assembly kit (New England Biolabs, Ipswich, MA, USA). Genomic DNA on the 5' flank of the region was amplified using primers AP34 and AP35, and on the 3' flank with primers AP36 and AP37. The *hph* expression cassette conferring hygromycin resistance was amplified with primers AP63 and AP64. The three fragments were assembled into plasmid pPZP-201BK linearized with EcoRI and HindIII. A construct to complement this deletion was generated by amplifying the required region in 4 pieces by PCR using primer pairs AP308 and AP309, AP310 and AP314, AP315 and AP311 and AP312 and AP313. These fragments were cloned into plasmid pMAI2⁶⁵ linearized with XbaI to recreate a 12,176 bp fragment of the genome.

Transformation of *P. variotii* and screening for correct integration events. Donor DNA was cloned into T-DNAs and transformed via *Agrobacterium tumefaciens* into the genome of *P. variotii*, as described previously²⁹.

In cases where targeted integration of the construct was required, DNA was prepared using a CTAB buffer protocol⁶⁸ and correct integration of the construct was assessed by PCR. For *pcaA*, primer pairs AP235 and AP251 and AP236 and AP252 were used. For *zrcA*, primer pairs AP323 and AP251 and AP324 and AP252 were used. Deletion of the arsenic region was assessed using primers pairs AP116 and MAI0023 and AP117 and MAI0022.

Yeast complementation. Homologous recombination using a PCR product was used to disrupt the *ZRC1* and its paralog *COT1* in *S. cerevisiae*. In the case of *ZRC1*, the *LEU2* allele was amplified using AP316 and AP317 off the plasmid pGAD⁶⁹ with appropriate homologous sequence added the primers for targeting to the *ZRC1* locus. In the case of *COT1*, primers

AP327 and AP328 were used to amplify the KanMX expression cassette from plasmid pFA6a-GFP(S65T)-kanMX6⁷⁰ with appropriate homologous sequence added the primers for targeting to the *COT1* locus. These PCR products were sequentially transformed as described previously⁷¹ into *S. cerevisiae* strain BY4742, selecting on either media lacking leucine in the case of *LEU2* or with 200 µg/ml G418 in the case of KanMX. DNA was prepared from the transformants⁷² and correct integration of the deletion construct was assessed by PCR using primers AP325 and AP326 for *ZRC1*, and AP341 and AP342 for *COT1*.

The *zrcA* cDNA was cloned into plasmid pTH19 to express ZrcA under the control of the constitutive alcohol dehydrogenase 1 promoter in *S. cerevisiae*⁷³, as follows. RNA was extracted using Trizol and treated with DnaseI (New England Biolabs). The RNA was reverse transcribed using AMV reverse transcriptase (New England Biolabs) and an oligo-dT primer. The *Zrc1* cDNA was amplified using Q5 Polymerase Master Mix (New England Biolabs) with primers AP337 and AP340. The PCR product was then cloned into pTH19 linearized with BamHI and HindIII using the NEBuilder assembly kit (New England Biolabs). The plasmid was transformed into yeast as described previously⁷¹ selecting for transformants on media lacking uracil.

Quantitative PCR. RNA was extracted using Trizol from 48 hours old cultures grown in 10% Campbells V8 juice supplemented with different concentrations of copper sulfate. The RNA was reverse transcribed using AMV reverse transcriptase (New England Biolabs) using oligo-dT. The *mcoA* gene was amplified using primers AP369 and AP370 and the constitutive beta-tubulin gene was amplified as described previously⁶⁷.

Microscopy. Strains were examined under a Leica DM6000 B microscope equipped with a Leica DFC 450 C camera. GFP fluorescence was visualized using the “I3” filter cube.

Genome sequencing and assembly. Genomic DNA was extracted from stationary liquid cultures of *P. variotii* strains FRR 1658, FRR 2889, FRR 3823 and FRR 5287, as described previously⁶⁸. Genomes were sequenced using an Illumina NovaSeq 6000 System at the Murdoch Children’s Research Institute (Melbourne, Australia). Reads were trimmed using BBduk and mapped to the *Hø* region⁷⁴ using Geneious version 11.1.5 to identify junctions. Raw reads have been deposited in the GenBank SRA database (BioProject ID: PRJNA604095).

ACKNOWLEDGEMENTS

We thank the University of Melbourne and Australasian Mycological Society for support for this research, Mark Wilson (CSIRO) for providing fungal strains, and Barbara Howlett for comments on the manuscript.

REFERENCES

- 1 Tchounwou, P. B., Yedjou, C. G., Patlolla, A. K. & Sutton, D. J. in *Molecular, Clinical and Environmental Toxicology Experientia Supplementum* (ed A. Luch) (Springer, 2012).
- 2 Campos, J. A., Tejera, N. A. & Sánchez, C. J. Substrate role in the accumulation of heavy metals in sporocarps of wild fungi. *BioMetals* **22**, 835-841 (2009).
- 3 Cocchi, L., Vescovi, L., Petrini, L. E. & Petrini, O. Heavy metals in edible mushrooms in Italy. *Food Chem.* **98**, 277-284 (2006).
- 4 Deshmukh, R., Khardenavis, A. A. & Purohit, H. J. Diverse metabolic capacities of fungi for bioremediation. *Indian J. Microbiol.* **56**, 247-264 (2016).
- 5 Schippers, A. *et al.* Biomining: metal recovery from ores with microorganisms. *Geobiotechnology I: Metal-Related Issues* **141**, 1-47 (2014).
- 6 Gerwien, F., Skrahina, V., Kasper, L., Hube, B. & Brunke, S. Metals in fungal virulence. *FEMS Microbiol. Rev.* **42**, 1-21 (2018).
- 7 Fay, J. C., McCullough, H. L., Sniegowski, P. D. & Eisen, M. B. Population genetic variation in gene expression is associated with phenotypic variation in *Saccharomyces cerevisiae*. *Genome Biol.* **5**, R26 (2004).
- 8 Warringer, J. *et al.* Trait variation in yeast is defined by population history. *PLoS Genet.* **7**, e1002111 (2011).
- 9 Colombi, E. *et al.* Evolution of copper resistance in the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* through acquisition of integrative conjugative elements and plasmids. *Environ. Microbiol.* **19**, 819-832 (2017).
- 10 Bhatt, K., Timsit, E., Rawlyk, N., Potter, A. & Liljebjelke, K. Integrative conjugative element ICEHs1 encodes for antimicrobial resistance and metal tolerance in *Histophilus somni*. *Front. Vet. Sci.* **5**, 153 (2018).
- 11 Arai, N. *et al.* *Salmonella* genomic island 3 is an integrative and conjugative element and contributes to copper and arsenic tolerance of *Salmonella enterica*. *Antimicrob. Agents Chemother.* **63**, e00429-00419 (2019).
- 12 Keeling, P. J. & Palmer, J. D. Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* **9**, 605-618 (2008).
- 13 Andersson, J. O. Lateral gene transfer in eukaryotes. *Cell. Mol. Life Sci.* **62**, 1182-1197 (2005).
- 14 Rokas, A., Wisecaver, J. H. & Lind, A. L. The birth, evolution and death of metabolic gene clusters in fungi. *Nat. Rev. Microbiol.* **16**, 731-744 (2018).
- 15 Walton, J. D. Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fungal Genet. Biol.* **30**, 167-171 (2000).
- 16 Wisecaver, J. H. & Rokas, A. Fungal metabolic gene clusters-caravans traveling across genomes and environments. *Front. Microbiol.* **6**, 161 (2015).
- 17 Johnstone, I. L. *et al.* Isolation and characterization of the *crnA-niiA-niaD* gene cluster for nitrate assimilation in *Aspergillus nidulans*. *Gene* **90**, 181-192 (1990).

- 18 Hall, C. & Dietrich, F. S. The reacquisition of biotin prototrophy in *Saccharomyces cerevisiae* involved horizontal gene transfer, gene duplication and gene clustering. *Genetics* **177**, 2293-2307 (2007).
- 19 Le Govic, Y. *et al.* Genomic organization and expression of iron metabolism genes in the emerging pathogenic mold *Scedosporium apiospermum*. *Front. Microbiol.* **9**, 827 (2018).
- 20 Bobrowicz, P., Wysocki, R., Owsianik, G., Goffeau, A. & Ułaszewski, S. Isolation of three contiguous genes, *ACR1*, *ACR2* and *ACR3*, involved in resistance to arsenic compounds in the yeast *Saccharomyces cerevisiae*. *Yeast* **13**, 819-828 (1997).
- 21 Inoue, Y. *et al.* Complete fusion of a transposon and herpesvirus created the Teratorn mobile element in medaka fish. *Nat. Commun.* **8**, 551 (2017).
- 22 Arkhipova, I. R. & Yushenova, I. A. Giant transposons in eukaryotes: is bigger better? *Genome Biol Evol.* **11**, 906-918 (2019).
- 23 van den Brule, T. *et al.* The most heat-resistant conidia observed to date are formed by distinct strains of *Paecilomyces variotii*. *Environ. Microbiol.*, in press (2020).
- 24 Houbaken, J., Varga, J., Rico-Munoz, E., Johnson, S. & Samson, R. A. Sexual reproduction as the cause of heat resistance in the food spoilage fungus *Byssoschlamys spectabilis* (anamorph *Paecilomyces variotii*). *Appl. Environ. Microbiol.* **74**, 1613-1619 (2008).
- 25 Rodrigo, S., Santamaria, O., Halecker, S., Lledo, S. & Stadler, M. Antagonism between *Byssoschlamys spectabilis* (anamorph *Paecilomyces variotii*) and plant pathogens: Involvement of the bioactive compounds produced by the endophyte. *Ann. Appl. Biol.* **171**, 464-476 (2017).
- 26 Marante, F. J. T. *et al.* Structural characterization and metabolite profiling of the facultative marine fungus *Paecilomyces variotii*. *Annals of Microbiology* **62**, 1601-1607 (2012).
- 27 Houbaken, J., Verweij, P. E., Rijs, A., Borman, A. M. & Samson, R. A. Identification of *Paecilomyces variotii* in clinical samples and settings. *J. Clin. Microbiol.* **48**, 2754-2761 (2010).
- 28 Wang, H. *et al.* Whole-genome sequencing and analysis of *Paecilomyces variotii* MTDF-01 from *Moutai*-flavor liquor brewing process. *Food Science* **40**, 185-192 (2019).
- 29 Urquhart, A. *et al.* Genomic and genetic insights into a cosmopolitan fungus, *Paecilomyces variotii* (Eurotiales). *Front. Microbiol.* **9**, 3058 (2018).
- 30 Oka, T., Ekino, K., Fukuda, K. & Nomura, Y. Draft genome sequence of the formaldehyde-resistant fungus *Byssoschlamys spectabilis* No. 5 (anamorph *Paecilomyces variotii* No. 5) (NBRC109023). *Genome Announc* **2**, e01162-01113 (2014).
- 31 Biango-Daniels, M. N., Wang, T. W. & Hodge, K. T. Draft genome sequence of the patulin-producing fungus *Paecilomyces niveus* strain CO7. *Genome Announc.* **6**, e00556-00518 (2018).
- 32 Radwan, O., Gunasekera, T. S. & Ruiz, O. N. Draft genome sequence of *Byssoschlamys sp.* isolate BYSS01, a filamentous fungus adapted to the fuel environment. *Genome Announc* **6**, e00164-00118 (2018).
- 33 Ropars, J. *et al.* Adaptive horizontal gene transfers between multiple cheese-associated fungi. *Curr. Biol.* **25**, 2562-2569 (2015).
- 34 Lai, H. *et al.* Functional characterization of the copper-transporting P-type ATPase gene of *Penicillium janthinellum* strain GXCR. *Journal of Microbiology* **47**, 736-745 (2009).
- 35 Bakti, F., Sasse, C., Heinekamp, T., Pócsi, I. & Braus, G. H. Heavy metal-induced expression of PcaA provides cadmium tolerance to *Aspergillus fumigatus* and supports its virulence in the *Galleria mellonella* model. *Front. Microbiol.* **9**, 744 (2018).
- 36 Gravot, A. *et al.* AtHMA3, a plant P1B-ATPase, functions as a Cd/Pb transporter in yeast. *FEBS Lett.* **561**, 22-28 (2004).

- 37 Frassinetti, S., Bronzetti, G., Caltavuturo, L., Cini, M. & Della Croce, C. The role of zinc in life: A review. *J. Environ. Pathol. Toxicol. Oncol.* **25**, 597-610 (2006).
- 38 Babich, H. & Stotzky, G. Toxicity of zinc to fungi, bacteria, and coliphages: influence of chloride ions. *Appl. Environ. Microbiol.* **36**, 906-914 (1978).
- 39 Kamazono, A., Nishizawa, M., Teranishi, Y., Murata, K. & Kimura, A. Identification of a gene conferring resistance to zinc and cadmium ions in the yeast *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **219**, 161-167 (1989).
- 40 MacDiarmid, C. W., Gaither, L. A. & Eide, D. Zinc transporters that regulate vacuolar zinc storage in *Saccharomyces cerevisiae*. *EMBO J.* **19**, 2845-2855 (2000).
- 41 Cho, M. *et al.* Vacuolar zinc transporter Zrc1 is required for detoxification of excess intracellular zinc in the human fungal pathogen *Cryptococcus neoformans*. *Journal of Microbiology* **56**, 65-71 (2018).
- 42 Crawford, A. C. *et al.* Biphasic zinc compartmentalisation in a human fungal pathogen. *PLoS Path.* **14**, e1007013 (2018).
- 43 Blaudez, D. & Chalot, M. Characterization of the ER-located zinc transporter Znt1 and identification of a vesicular zinc storage compartment in *Hebeloma cylindrosporum*. *Fungal Genet Biol.* **48**, 496-503 (2011).
- 44 López-Maury, L., Florencio, F. J. & Reyes, J. C. Arsenic sensing and resistance system in the cyanobacterium *Synechocystis sp* strain PCC 6803. *J. Bacteriol.* **185**, 5363-5371 (2003).
- 45 Rosenstein, R., Peschel, A., Wieland, B. & Götz, F. Expression and regulation of the antimonite, arsenite, and arsenate resistance operon of *Staphylococcus xylosus* plasmid pSX267. *J. Bacteriol.* **174**, 3676-3683 (1992).
- 46 Sato, T. & Kobayashi, Y. The *ars* operon in the *skin* element of *Bacillus subtilis* confers resistance to arsenate and arsenite. *J. Bacteriol.* **180**, 1655-1661 (1998).
- 47 Zhou, X. *et al.* A genome-wide screen in *Saccharomyces cerevisiae* reveals pathways affected by arsenic toxicity. *Genomics* **94**, 294-307 (2009).
- 48 Qin, J. *et al.* Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2075-2080 (2006).
- 49 Grass, G. & Rensing, C. CueO is a multi-copper oxidase that confers copper tolerance in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **286**, 902-908 (2001).
- 50 Shi, X. L., Stoj, C., Romeo, A., Kosman, D. J. & Zhu, Z. W. Fre1p Cu²⁺ reduction and Fet3p Cu¹⁺ oxidation modulate copper toxicity in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 50309-50315 (2003).
- 51 Schaack, S., Gilbert, C. & Feschotte, C. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* **25**, 537-546 (2010).
- 52 McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U. S. A.* **36**, 344-355 (1950).
- 53 McDonald, M. C. *et al.* Transposon-mediated horizontal transfer of the host-specific virulence protein ToxA between three fungal wheat pathogens. *mBio* **10**, e01515-01519 (2019).
- 54 Castanera, R. *et al.* Highly expressed captured genes and cross-kingdom domains present in Helitrons create novel diversity in *Pleurotus ostreatus* and other fungi. *BMC Genomics* **15**, 1071 (2014).
- 55 Weinert, T. A., Schaus, N. A. & Grindley, N. D. F. Insertion sequence duplication in transpositional recombination. *Science* **222**, 755-765 (1983).
- 56 Goodwin, T. J. D., Butler, M. I. & Poulter, R. T. M. Cryptons: a group of tyrosine-recombinase-encoding DNA transposons from pathogenic fungi. *Microbiology* **149**, 3099-3109 (2003).

- 57 Han, M. J., Xu, H. E., Zhang, H. H., Feschotte, C. & Zhang, Z. Spy: A New Group of Eukaryotic DNA Transposons without Target Site Duplications. *Genome Biol. Evol.* **6**, 1748-1757 (2014).
- 58 Johnson, C. M. & Grossman, A. D. Integrative and conjugative elements (ICEs): what they do and how they work. *Annu. Rev. Genet.* **49**, 577-601 (2015).
- 59 Wisecaver, J. H., Slot, J. C. & Rokas, A. The evolution of fungal metabolic pathways. *PLoS Genet.* **10**, e1004816 (2014).
- 60 Cheeseman, K. *et al.* Multiple recent horizontal transfers of a large genomic region in cheese making fungi. *Nat. Commun.* **5**, 2876 (2014).
- 61 Hane, J. K., Williams, A. H., Taranto, A. P., Solomon, P. S. & Oliver, R. P. in *Genetic Transformation Systems in Fungi, Vol 2* (eds M. A. Van Den Berg & K. Maruthachalam) 55-68 (Springer, 2015).
- 62 Arnaud, M. B. *et al.* The Aspergillus Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Res.* **40**, D653-D659 (2012).
- 63 Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755 (2001).
- 64 Samson, R. A., Houbraken, J., Varga, J. & Frisvad, J. C. Polyphasic taxonomy of the heat resistant ascomycete genus *Byssoschlamys* and its *Paecilomyces* anamorphs. *Persoonia* **22**, 14-27 (2009).
- 65 Idnurm, A. *et al.* Spontaneous and CRISPR/Cas9-induced mutation of the osmosensor histidine kinase of the canola pathogen *Leptosphaeria maculans*. *Fungal Biol Biotechnol.* **4**, 12 (2017).
- 66 Covert, S. F., Kapoor, P., Lee, M. H., Briley, A. & Nairn, C. J. *Agrobacterium tumefaciens*-mediated transformation of *Fusarium circinatum*. *Mycol. Res.* **105**, 259-264 (2001).
- 67 Urquhart, A. S., Hu, J., Chooi, Y. H. & Idnurm, A. The fungal gene cluster for biosynthesis of the antibacterial agent viriditoxin. *Fungal Biol Biotechnol.* **6**, 10 (2019).
- 68 Pitkin, J. W., Panaccione, D. G. & Walton, J. D. A putative cyclic peptide efflux pump encoded by the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonum*. *Microbiology* **142**, 1557-1565 (1996).
- 69 James, P., Halladay, J. & Craig, E. A. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436 (1996).
- 70 Bähler, J. *et al.* Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**, 943-951 (1998).
- 71 Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols* **2**, 31-34 (2007).
- 72 Lööke, M., Kristjuhan, K. & Kristjuhan, A. Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques* **50**, 325-328 (2011).
- 73 Harashima, T. & Heitman, J. G alpha subunit Gpa2 recruits kelch repeat subunits that inhibit receptor-G protein coupling during cAMP-induced dimorphic transitions in *Saccharomyces cerevisiae*. *Mol. Biol. Cell.* **16**, 4557-4571 (2005).
- 74 Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821-829 (2008).

748 Table S1 – Genes present within the *Hφ* region.
749

Protein ID in JGI	Gene name	Putative function
195182	<i>hhpA</i>	DUF3435
167128	<i>hhpB</i>	Transcription factor, MADS-box. Possible pseudogene.
290670	<i>hhpC</i>	Reverse transcriptase
276811	<i>hhpE</i>	Mitochondrial Fe ²⁺ transporter MMT1
NA		Tn pseudogene
232099	<i>hhpF</i>	
179937	<i>zrcA</i>	Zinc transporter
232097	<i>hhpG</i>	
276812	<i>mcoA</i>	Multicopper oxidase
290673	<i>hhpH</i>	Low affinity Fe ²⁺ transporter
290674	<i>hhpI</i>	Truncated DUF3435 domain protein. Possible pseudogene.
290675	<i>hhpJ</i>	
276813	<i>arsH</i>	NADPH-dependent FMN reductase-like arsH
239850	<i>arsX</i>	Unknown gene conserved in ars clusters
239851	<i>arsC1</i>	Arsenate reductase, arsC
239852	<i>phoD2</i>	Cyclin, PHO80
290679	<i>arsB</i>	Arsenite efflux transporter
239854	<i>arsC2</i>	Arsenate reductase
290681	<i>arsY</i>	Intermediate filament Protein. Next to Cyt19 in other fungal genomes.
290682	<i>asmA</i>	Arsenic methyltransferase; like Cyt19
290683	<i>hhpK</i>	K ⁺ -dependent Ca ²⁺ /Na ⁺ exchanger
239858	<i>hhpL</i>	Metal transporting P-type ATPase
195275	<i>pcaA</i>	Metal transporting P-type ATPase
276819	<i>hhpM</i>	Transcription factor, MADS-box
276820	<i>hhpN</i>	Tc1 transposon
290686	<i>hhpO</i>	

290687	<i>hhpP</i>	
290688	<i>hhpQ</i>	
195756	<i>hhpR</i>	metalloprotease
290691	<i>hhpS</i>	AAA ATPase; mitochondrial chaperone BCS1
290692	<i>hhpT</i>	
290693	<i>hhpU</i>	
290694	<i>hhpV</i>	
239869	<i>hhpW</i>	
239870	<i>hhpX</i>	

750

751

752 Table S2 – primers used in this study.

Name	Sequence (5'-3')
AP34	AACAGCTATGACATGATTACGTTCAAAATCGGCTTCCATCG
AP35	GTCGTCGGGTACCGTGGTTTCGAGTGCCTTGG
AP36	AAACCACGGTACCCGACGACAGTCGCGCGCTCC
AP37	GTAAACGACGGCCAGTGCCAGGTACGGAACTTGCAGAGACC
AP63	GCGCACTCGAAACCACGGTACGCTGGGATTGCCCCCTCGATGC
AP64	CGCGCGACTGTCGTCGGGTACGCTACTGAACGTTATGAC
AP89	GCGGTAGATTGTAATTGTCG
AP90	TGTTGTATGCGCGGAATTCTG
AP92	TTCGAAGAAGATCATATTCC
AP93	GGTCGAAAACGCGCACATCG
AP104	TCGAAACCTAATCAATCAACATGGAATGTGATGATTGCGG
AP105	TGCTCATAGTCACATCCCTCAGATCTTCGACCAACGCAG
AP116	GGTTCGAGCGGCGTACAGC
AP117	CAGGATAGTGGACCGTGCACG
AP219	ACAGCTATGACATGATTACGGTCACCAACCAAGGAATTGC
AP220	TATTGAAGTGAGCCTTTGC
AP221	GCAAAGGCTCACTTCAATAATGGTGAGCAAGGGCGAGGAGC
AP222	GCAATAGTAGTGCCCTGTGACCTACTGAACGTTATGACG
AP223	TCACAGGGCACTACTATTGC
AP224	TAAAACGACGGCCAGTGCCACCCAAGCTTCTACAGAGTCC
AP235	CGATGGCTGAGTAGCCATCC
AP236	GAATCCGACAGAGACGAAGC
AP251	CCGGTGAACAGCTCCTCGCC
AP252	CGTCATAACGTTCACTAGG
AP297	TCGAAACCTAATCAATCAACATGGCGCTTCACTTGACTCAG
AP298	TGCTCATAGTCACATCCCTCAGCCACAGCAGGTCAAGATC
AP299	ACAGCTATGACATGATTACGGTAATGATTATGCAAACTGC
AP302	GGTTGAGAACAGGATGTGG
AP303	CCACATCCTGTTCTCAACCATGGTGAGCAAGGGCGAGGAGC
AP304	GGCGTACTCAAACCTCAGCCCTACTGAACGTTATGACG
AP305	GCTGAGGTTTGAGTACGCC
AP307	TAAAACGACGGCCAGTGCCATCTGAGAGTGTATGCGAAGG
AP308	TAGGCCTCTGCAGGTCGACTGCTGTCCACACACATACC
AP309	GGTAGGGTGACTCTCATTCTGG
AP310	CCGAATGAGAGTCACCCTACC
AP311	GGTTTGTTTACAAGTTGGCTGCC
AP312	GGCAGCCAACCTGTAACCAAACC
AP313	AGCTCGGTACCCGGGGATCCTGGTTGGATCCTTCTGAAGC
AP314	GCTCCTTACCATTGACGTTCC
AP315	GGAACGTCAATGGTGAAGGAGC
AP316	TGATCACCGGTAAGAATTGAGAATCATCTCTTTTGACCTTAGACACGGTTTTTTTCAACTGTGGGAATACTCAGG
AP317	TTACAGGCAATTGGAAGTATTGCAGTTTACAGCGTCATCTACAATACAATTAGATGATGCTTAAGCAAGGATTTTCTTAAC
AP323	AACAAAGCCGCTAGGTAGG
AP324	CTCCTCCAAATATAGATAGG
AP325	CTTGGCGCAGTAATTTTCCC
AP326	TGTTCTCTGTAGAACCATGG
AP327	AATAGTTCTGCATAGCTATAGAAAGAAAGTTAACACAAAGTACGGAAGATTGAGTAAATGACATGGAGGCCAGGAATACC
AP328	CCGTATAACGATTTTTAAAGTATTTAATTCTTCACGCTTTTTCGTTATAAAGTCCTCAGTATAGCGACCAGCATTCTC
AP337	CAAGCATACAATCAACTCCAATGGCGCTTCACTTGACTCAG
AP340	AGCTTGGCTGCAGGTCGACGCTAGCCACAGCAGGTCAAGATC

AP341	AGCTCGAAAAGCATTCTGG
AP342	TGATCATCATTATTATCACC
AP351	TCGAAACCTAATCAATCAACATGGCGCTTCACTTGACTCAG
AP352	TTGCTCACCCCTCCGCCACCGCCACAGCAGGTCAAGATC
AP353	GGTGGCGGAGGGGTGAGCAAGGGCGAGGAGC
AP354	TGCTCATAGTCACATCCCTCACTTGTACAGCTCGTCCATG
AP369	ACCCCAGGAAGTTCAGGTCC
AP370	CGTAGTCGTCATATTTGTCCG
MAI0448	GTCCACTTATTACAATGGC
MAI0449	CCGTTCAATATGGCGATC
MAI0450	GAACTGTCATTTGTCGG
MAI0451	TCCCATGTTCCCACTGAAG

753