SCIENTIFIC REPORTS

OPEN

Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*

Guohua Xiao[1]*, Sheng-Hua Ying[2]*, Peng Zheng[1]*, Zheng-Liang Wang[2]*, Siwei Zhang[1], Xue-Qin Xie[2], Yanfang Shang[1], Raymond J. St. Leger[3], Guo-Ping Zhao[4], Chengshu Wang[1,4] & Ming-Guang Feng[2]

[1]Key Laboratory of Insect Developmental and Evolutionary Biology, Institute of Plant physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China, [2]College of Life Sciences, Zhejiang University, Hangzhou 310058, China, [3]Department of Entomology, University of Maryland, College Park, Maryland 20742, USA, [4]Key Laboratory of Synthetic Biology, Institute of Plant physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China.

The ascomycete fungus *Beauveria bassiana* is a pathogen of hundreds of insect species and is commercially produced as an environmentally friendly mycoinsecticide. We sequenced the genome of *B. bassiana* and a phylogenomic analysis confirmed that ascomycete entomopathogenicity is polyphyletic, but also revealed convergent evolution to insect pathogenicity. We also found many species-specific virulence genes and gene family expansions and contractions that correlate with host ranges and pathogenic strategies. These include *B. bassiana* having many more bacterial-like toxins (suggesting an unsuspected potential for oral toxicity) and effector-type proteins. The genome also revealed that *B. bassiana* resembles the closely related *Cordyceps militaris* in being heterothallic, although its sexual stage is rarely observed. A high throughput RNA-seq transcriptomic analysis revealed that *B. bassiana* could sense and adapt to different environmental niches by activating well-defined gene sets. The information from this study will facilitate further development of *B. bassiana* as a cost-effective mycoinsecticide.

Fungi are the commonest cause of insect disease in nature and approximately 1,000 fungal species are reported to kill insects, spiders or mites[1]. Fungi are particularly well suited for development as biopesticides because unlike bacteria and viruses they infect insects by direct penetration of the cuticle and so function as contact insecticides. *Beauveria* is one of the best-known genera of entomopathogenic fungi and worldwide numerous registered mycoinsecticide formulations based on *Beauveria bassiana* (Bb) and *B. brongniartii* are used for control of insect pests[2]. Bb has a particularly wide host range (over 700 species) allowing it to be used against vectors of human disease and a wide range of insect pests[3]. For example, in China, approximately one million hectares a year are treated with Bb to control forest insects such as the pine caterpillar *Dendrolimus punctatus*[4,5].

B. bassiana was discovered by Agostinio Bassi in 1835 as the cause of the devastating muscardine disease of silkworms[6]. The ability of insects to defend against *Beauveria* has illuminated many aspects of innate immunity with direct relevance to human immunology[7]. Bb is also a well-known biocatalyst in chemical and industrial applications[8]. The important role of Bb as a plant endophyte and antagonist of plant pathogenic fungi has only become apparent in the last 20 years[9]. Furthermore, as shown by their pathogenicity to soil amoeba[10], at least some Bb isolates have additional unpredicted flexibility in their trophic capabilities. However, the mechanisms under-lying the physiological plasticity of Bb are still poorly understood. In addition, although the sexual stage of Bb has been identified as *Cordyceps bassiana*[11], it is very rarely observed and the role of sexuality in Bb is unknown.

Beauveria is well known for producing a large array of biologically active secondary metabolites including non-peptide pigments and polyketides (e.g., oosporein, bassianin and tenellin), nonribosomally synthesized peptides (e.g., beauvericin, bassianolides and beauveriolides), and secreted metabolites involved in pathogenesis and virulence (e.g., oxalic acid) that have potential or realized industrial, pharmaceutical and agricultural uses[12]. Silkworm larvae infected by Bb (batryticated silkworms), have for centuries been a traditional Chinese medicine. The medicinal potential of batryticated silkworms has been validated by modern technologies e.g., water extract of batryticated silkworms protect against β-amyloid induced neurotoxicity[13].

Expressed sequence tag analyses, insertion mutagenesis and gene functional studies of Bb have already identified some of the genes involved in fungal development, virulence, detoxification, insect immune avoidance and stress responses[14–16]. To facilitate further comprehensive understanding of Bb pathogenesis and interactions between insects and plants, we sequenced the genome of Bb strain ARSEF 2860, and performed a comparative study with the sequenced genomes of ascomycete insect pathogens, *Metarhizium robertsii* (Mr), *M. acridum* (Ma)[17] and *Cordyceps militaris* (Cm)[18]. The comparison revealed a common set of gene family expansions that distinguish them from plant pathogens and saprophytes, as well as species-specific gain or loss of genes that correlate with different pathogenic strategies. Transcriptional responses of Bb to insect cuticles, insect hemocoel and plant root exudates were studied using an RNA_seq technique and demonstrated modulation of genes involved in signal transduction, secreted proteins and metabolism.

Results

General features. The genome of Bb strain ARSEF 2860 was shotgun sequenced to 76.6 × coverage using a Roche 454 system and Illumina paired-end sequencing. The assembly resulted in a total genome size of 33.7 mega bytes (Mb), which is similar to that of Cm (32.2 Mb), but smaller than Mr (39.0 Mb) and Ma (38.1 Mb) (Table 1). By mapping 13,412 EST sequences to the scaffolds, the Bb genome was estimated to be 96.1% complete. The genome was predicted to encode 10,366 protein genes, which is more than Cm (9,684) and Ma (9,849) but fewer than Mr (10,582). In terms of gene density (genes per Mb), Bb has a more compact genome structure relative to the other insect pathogens (Table 1), but it is less compact than plant pathogenic *Fusarium* spp.[19].

An Interproscan analysis identified 3,002 protein families in Bb (containing 7,283 total predicted proteins), which is approximately 200 more than *Metarhizium* spp. and Cm[17,18]. Bb also has a higher proportion (18.2%) of genes encoding putatively secreted proteins than Cm (16.2%), Mr (17.6%) and Ma (15.1%) (Table 1). All the insect pathogens have a 2 to 3-fold higher proportion of their genome devoted to secreted products than other Ascomycetes including plant pathogens and the mycoparasitic *Trichoderma* spp.[20]. Similar to other insect pathogens (average 17%), 17.6% of Bb proteins are in the pathogen-host interaction (PHI) database (Supplementary Table S1). Functional analysis indicated that relative to Cm and Mr, Bb has more proteins involved in cell metabolism, energy, cell cycle, transcription, transportation, signal transduction and cell differentiation, but it has fewer proteins than *Metarhizium* spp. involved in localization and formation of organelles, virulence and detoxification (Fig. 1A). Whole genome reciprocal analysis indicated that more than 80% of Bb genes show orthologous relationships with those of Cm and Mr, and its genome harbors the fewest number of species-specific genes (Fig. 1B).

A Blast score ratio analysis showed that Bb is more closely related to Cm than Mr (Fig. 2A). Sequence identity between Bb and Cm orthologs was 76%, as compared to 58% with *Metarhizium* spp. Although Bb and Cm are different from each other in conidiogenesis, infection structure formation, life cycle and host range (Fig. 2C), their relatedness is similar to that between the mycoparasitic fungi *Trichoderma virens* and *T. atroviride* (74%)[20], *Aspergillus fumigatus* and *A. niger* (69%)[21] and fish and humans (75%)[22]. A phylogenomic analysis based on 1,915 orthologous protein sequences showed that Bb and Cm diverged after a split with mycoparasitic *Trichoderma* spp. (Fig. 2B), and reinforced our previous analysis[18], suggesting that the split between the *Cordyceps* spp. (including Bb) and *Metarhizium* lineages occurred before *Metarhizium* diverged from the plant endophytic *Epichloë* lineage.

In contrast to two *Metarhizium* species[17], there are no obvious syntenic relationships between the genome structures of Bb and Cm (Fig. 3A). Analysis of Bb and Cm paired paralogs showing >70% nucleotide identity found a stronger C:G to T:A mutation bias in Cm than in Bb (Supplementary Fig. S1). Thus, similar to Mr[17], but unlike Cm[18] and *Neurospora crassa*[23], Bb may not use repeat-induced point mutations (RIP) for genome defense against repetitive sequences. This is consistent with expanded gene families and more transposons in the Bb and Mr genomes relative to Cm and Ma (Supplementary Table S2). RIP only occurs during meiosis[23], so its apparent absence in Bb suggests the sexual cycle is rare in this fungus. Unlike Cm which is specific to lepidopteran pupae[18], Bb has a wide host range. The wide host range Mr also has more gene families and larger gene families than the locust specialist Ma, and likewise 61 families were expanded in Bb relative to Cm (Supplementary Table S3). These included subtilisins and trypsins involved in degrading insect cuticles. Relative to other fungi, there are expansion/contraction of different protein families in Bb (Supplementary Table S4).

Bacterial-like toxins. In contrast to entomopathogenic bacteria and viruses, entomopathogenic fungi infect insects via cuticular penetration, and are usually assumed to lack per os infectivity[24]. However, the Bb genome contains many more bacterial-like toxins than other fungi (Supplementary Table S5). For example, Bb has 13 heat-labile enterotoxins compared to six in Mr and one or none in

| Table 1 | Comparison of genome features between *B. bassiana* and other insect pathogens. |||||
|---|---|---|---|---|
| Features | B. bassiana | C. militaris | M. robertsii | M. acridum |
| Size (Mb) | 33.7 | 32.2 | 39.0 | 38.1 |
| Coverage (fold) | 76.6× | 147× | 100× | 107× |
| Scaffold No. (> 1 kb) | 242 | 13 | 176 | 241 |
| Scaffold N50 (Mb) | 0.73 | 4.55 | 1.96 | 0.33 |
| % G+C content | 51.5 | 51.4 | 51.5 | 50.0 |
| % G+C in coding gene | 56.6 | 58.6 | 54.4 | 54.1 |
| % Repeat rate | 2.03 | 3.04 | 0.98 | 1.52 |
| Protein-coding genes | 10,366 | 9,684 | 10,582 | 9,849 |
| Protein families[a] | 3,002 | 2,736 | 2,797 | 2,746 |
| (protein no.) | (7,283) | (6,725) | (7,556) | (6,948) |
| Gene density (gene per Mb) | 308 | 301 | 271 | 259 |
| Exons per gene | 2.7 | 3.0 | 2.8 | 2.7 |
| % Secreted proteins | 18.2 | 16.2 | 17.6 | 15.1 |
| tRNA | 113 | 136 | 141 | 122 |
| Pseudogenes | 304 | 102 | 363 | 440 |
| NCBI accession | ADAH00000000 | AEVU00000000 | ADNJ00000000 | ADNI00000000 |

[a]InterproScan analysis data.

2
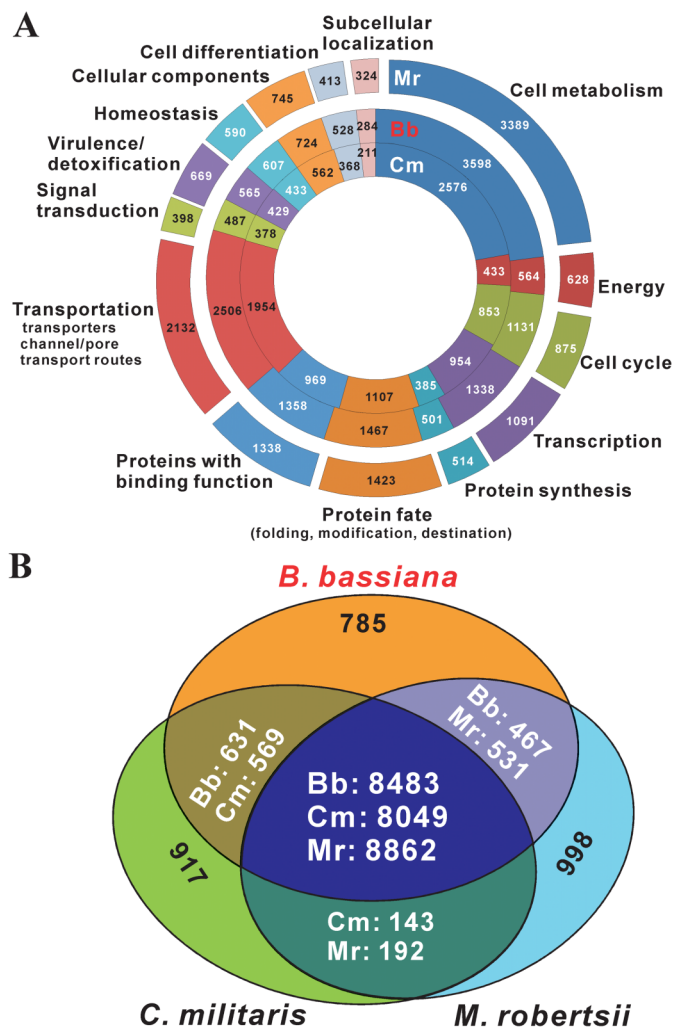
Figure 1 | Comparative genomics analysis of three insect pathogens. (A) Functional classification and comparison of *B. bassiana* (Bb), *C. militaris* (Cm) and *M. robertsii* (Mr) proteins. Each circle represents the relative fraction of genes represented in each of the categories for each genome. (B) Reciprocal blast analysis of the predicted proteins among the three insect pathogens. The cut-off *E* value is at ≤ 1e-5.

other fungi. Bb strains expressing a vegetative insecticidal toxin gene (*Vip3A*) from *Bacillus thuringiensis* (Bt) exhibited enhanced oral toxicity[24]. Unexpectedly, we also found that the Bb genome encodes eight genes showing similarities (<1e-20) to Bt Cry-like delta endotoxins while other fungi have at most one of these genes. This suggests that Bb may possess greater oral toxicity than other fungi. Most fungi lack genes for bacterial zeta toxin-like proteins but Bb has three, and the other insect pathogens one, suggesting that insect pathogens may use the bacterial toxin-antitoxin system to control cell stasis or death[25]. Interestingly, the TC-like insecticidal toxins present in *Trichoderma* species are absent in other fungi, including the entomopathogens (Supplementary Table S5).

Proteases. The genomes of Cm and *Metarhizium* spp. code for many more proteolytic enzymes than do non-insect pathogens. This may reflect an increased range of functions required when infecting insects[17,18]. Likewise, the Bb genome also codes for significantly (*P*<0.005) more trypsins (23 vs. an average of 2 in plant pathogens), subtilisins (43 vs. 17), aspartic proteases (21 vs. 18) and carboxypeptidases (52 vs. 32). However, Bb and the plant pathogens have similar numbers of cysteine peptidases (47 vs. 43), threonine peptidases (20 vs. 19) and metallopeptidases (98 vs. 89)

(Supplementary Table S6). The broad host range Mr has more proteases than the narrow host range Ma and Cm[17,18]. The dramatic expansion of proteases in Bb resembles Mr, suggesting that this is an adaptation to a broad host range. Thus, Bb has more subtilisins (43 vs. 35), and trypsins (23 vs. 12) than Cm, and has four subfamilies that Cm lacks (Supplementary Table S7). Bb has two trypsins (BBA_07369 and BBA_03951) that only show high identities (>60%) to bacterial trypsins, and a trypsin (BBA_02900) with a Cm homolog that clusters with insect trypsins (Supplementary Fig. S2). This is consistent with horizontal gene transfer in fungi[26]. However, a patchy distribution of genes can also be due to lineage specific gene loss.

Carbohydrate active enzymes. The number of glycoside hydrolases (GH) possessed by Bb (145) resembles other insect pathogens (average 141), rather than the endophyte *E. festucae* (98), but is significantly (*P*=0.0069) less than plant pathogens (average 199) (Supplementary Table S8). All the insect pathogens lack several families of cellulases (GH6, GH7, GH12, GH45 and GH61) and other enzymes involved in degrading plant cell walls (GH11, GH30, GH51, GH53, GH62, GH67 and GH115). Compared to *Metarhizium* spp., Bb and Cm have fewer xyloglucosyl transferases (GH16) responsible for degrading xylan oligomers and polymeric xylan[27]. Unlike Cm[18], Bb has a phosphoketolase required for xylose metabolism and full virulence in *Metarhizium* (BBA_09253 vs. MAA_04563, 41% identity)[28]. Thus, in contrast to Cm, Bb could germinate and grow on xylose medium, albeit very weakly when compared to Mr (Fig. 4A and 4B). Because they have fewer cellulases and hemicellulases, the number of carbohydrate-binding module 1 (CBM1) domains is significantly less in insect pathogens (average 10) than plant pathogens (25) (Supplementary Table S9). Insect pathogens also have significantly fewer putative oxidative lignin enzymes (average 29 in insect pathogens vs. 40 in plant pathogens, *P*=0.0016) (Supplementary Table S10), carbohydrate esterases (9 vs. 33, *P*=0.0025) (Supplementary Table S11), cutinases (4 vs. 12, *P*=0.0034) and pectin lyases (8 vs. 20, *P*=0.0122) (Supplementary Table S4). Cutinases and pectin lyases in particular are known to be virulence factors for plant pathogens[29].

Chitin is the second most abundant polymer in insect cuticle. The necessity to degrade it is reflected in an abundance of GH18 family chitinases in Bb (20) and the three other insect pathogens (average 19) compared to plant pathogens (average 11) (Supplementary Table S8). Fungal chitinases are subdivided into three subgroups[30]. Our analysis indicated that eight of the 20 Bb chitinases belong to sub-group A (without a chitin-binding domain, CBM), four belong to subgroup B (one CBM at the C-terminal), and eight are subgroup C chitinases (possessing CBM18 and CBM50 LysM chitin-binding modules) (Supplementary Table S12). Insect and plant pathogens have similar (*P*=0.2258) numbers of subgroup A chitinases but the entomopathogens have many more chitinases with CBMs (average 11 vs. 2, *P*=0). Phylogenetic analyses of subgroup B and C chitinases revealed that most of the gene duplication events have occurred since Bb/Cm, *Metarhizium* spp., and *Trichoderma* spp. diverged from a common ancestor, suggesting their abundance in each clade is due to convergent evolution (Supplementary Fig. S3).

Cytochrome P450s (CYPs). CYPs are involved in many essential cellular processes and play diverse roles in detoxification, degradation of xenobiotics, and the biosynthesis of pathogenesis related secondary metabolites[31]. Plant pathogens (average 118) and *Metarhizium* spp. (average 111) have more CYPs than Bb (83) and Cm (57), and 24 CYP families present in *Metarhizium* spp. are absent in Bb and Cm (Supplementary Table S13). These include nitric oxide reductases (CYP55, NOR) used for anaerobic denitrification[32]. Thus, unlike *Metarhizium* spp., Bb and Cm may not be able to respond to hypoxic conditions by nitrate or ammonia fermentation. Bb and Cm also lack CYP619 for biosynthesis of the tetraketide mycotoxin
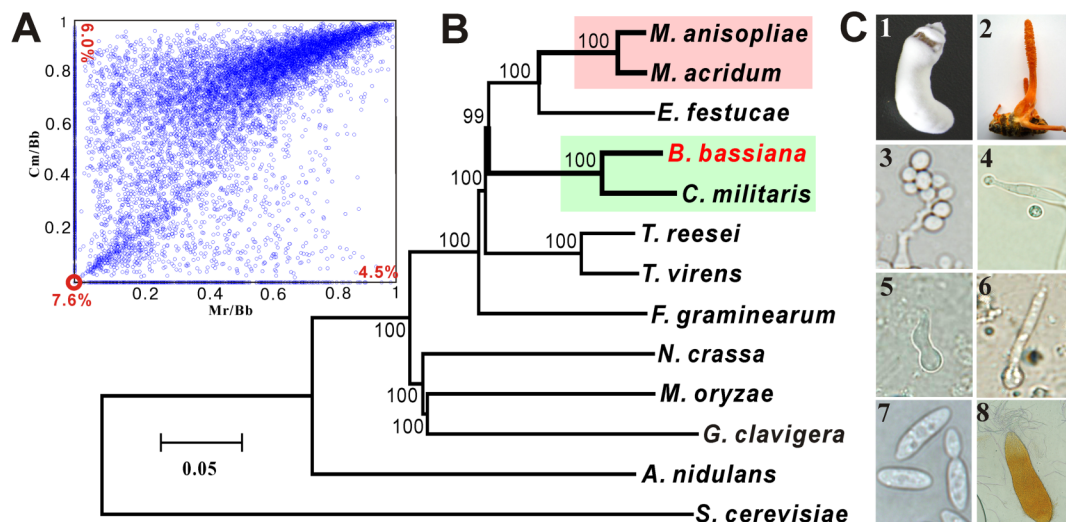
3

Figure 2 | Comparative genomics and phylogenomic analyses of *B. bassiana* with other fungi. (A) Scatter plots of Blast score ratio analysis of *B. bassiana* (Bb), *C. militaris* (Cm) and *M. robertsii* (Mr) showing that Bb is more closely related to Cm. The numbers in red in the lower left corners indicate the percentages of Bb species-specific sequences and the numbers in the upper left or lower right show the percentages of lineage-specific genes between pairs of genomes as indicated. (B) A maximum likelihood phylogenomic tree constructed using the Dayhoff amino acid substitution model showing the evolutionary relationship of Bb with other fungal species. (C) Phenotypic and morphological comparison of Bb and Cm. Panel 1, a wax moth (*Galleria mellonella*) larva killed and mycosed by Bb; Panel 2, Cm fruits on a silkworm (*Bombyx mori*) pupa; Panel 3, Bb conidial formation on a zig-zag shared conidiophore; Panel 4, Cm conidial formation on a phialide-type conidiophore; Panel 5, a Bb germ tube on a locust hind wing showing tip swelling; Panel 6, a Cm germ tube on a locust hind wing without tip swelling; Panel 7, Bb hyphal bodies harvested from cotton bollworm 48 hrs post infection; Panel 8, a perithecium produced on the Cm fruiting body.

patulin[33]. Relative to Cm (57), the Bb genome encodes ~1.5-fold more CYP genes (83), including 16 additional subfamilies. Conversely, there are six families present in Cm that are absent in Bb. In particular, Bb lacks CYP567 whereas Cm has three CYP567s, among which CCM_03048 and CCM_03052 are located in a putative sesquiterpene cyclase gene (CCM_03050) cluster. The fungal CYP51 system is involved in sterol biosynthesis and detoxification of azole antifungals[34]. The four entomopathogens each have two CYP51 genes. There are 6 ω-hydroxylase CYP52 enzymes in Bb as compared to 4 in Cm, 1 in Mr and 4 in Ma (Supplementary Table S13). CYP52s catalyze the first step in the ω-oxidation pathway of alkanes, so their presence is consistent with efficiently metabolizing insect epicuticle alkanes[35,36].

Of the 83 CYPs in the Bb genome, 18 are within gene clusters involved in the biosynthesis of secondary metabolites. For example,

a CYP617 protein (BBA_02632) is proximal to a non-ribosomal peptide synthetase (NRPS, BBA_02630) for bassianolide biosynthesis[37]. CYP5293 (BBA_09720) is within the beauvericin (NRPS, BBA_09727) biosynthetic cluster[38]. CYP655 (BBA_07335) and CYP623 (BBA_07336) are close to a polyketide synthase (PKS)-NRPS hybrid (BBA_07338) for tenellin biosynthesis[39]. *Metarhizium* CYP65 DtxS2 is a multifunctional enzyme involved in hydroxylation, desaturation, oxidation and epoxidation of intermediate insecticidal toxin destruxins[40]. Bb has a cluster, absent in Cm, containing two CYP65s (BBA_08705 and BBA_08706) and one CYP58 (BBA_08698) for biosynthesis of as yet unknown metabolite(s). The CYP6001 subfamily produce oxylipins for signaling and secondary metabolism in fungi[41]. Cm lacks CYP6001 but it is present as a single copy gene in Bb and *Metarhizium* spp.

Small secreted cysteine-rich proteins (SSCPs). Most of the secreted effector-type proteins of plant pathogens are small (<300 amino acids) and contain four or more cysteine residues[42]. We surveyed and compared fungal SSCPs. In total, 396 clusters were obtained by a Blastclust analysis of Bb and 11 other fungal species. Of these, 12 clusters contain SSCPs shared by insect pathogens and the plant endophyte *E. festucae,* and 26 are found in insect and plant pathogens. Of the 91 clusters specific to insect pathogens, 52 contain genes from Bb. Relative to other insect pathogens (average 307), the Bb genome encodes more SSCPs (373), and many of them are species specific (154 vs. an average of 95) (Supplementary Table S14).

As with other fungi[20,42], most of the entomopathogen SSCPs are of unknown function. Of 373 Bb SSCPs, only 130 contained conserved domains recognized by an Interproscan analysis. Some of these had homologs in the PHI database of verified virulence determinants, e.g. five putative cutinases and five trypsins that may be used by Bb to target plant and insect cuticle components, respectively. Lectins are used by mammalian pathogenic fungi to evade detection by host receptors[43]. Six Bb SSCPs were identified as concanavalin A-like lectins, and potentially could function in interactions with both

Table 2 | The core genes involved in the biosynthesis of secondary metabolites in insect pathogens

Core gene[a]	B. bassiana	C. militaris	M. robertsii	M. acridum
DMAT	0	1	5	3
TC	3	3	3	3
TS	7	2	8	6
FAS	2	1	2	2
GGPS	0	3	4	4
NRPS	13	5	15	13
NRPS-like	7	8	9	8
PKS	12	9	24	13
PKS-like	1	2	3	4
HYBRID	3	3	5	1
Sum	**45**	**37**	**78**	**57**

[a]Abbreviations: DMAT, Dimethylallyl tryptophan synthase; TC, terpene cyclase; TS, terpene synthase; FAS, fatty-acid synthase; GGPS, geranylgeranyl diphosphate synthase; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; HYBRID, hybrid PKS–NRPS enzyme.
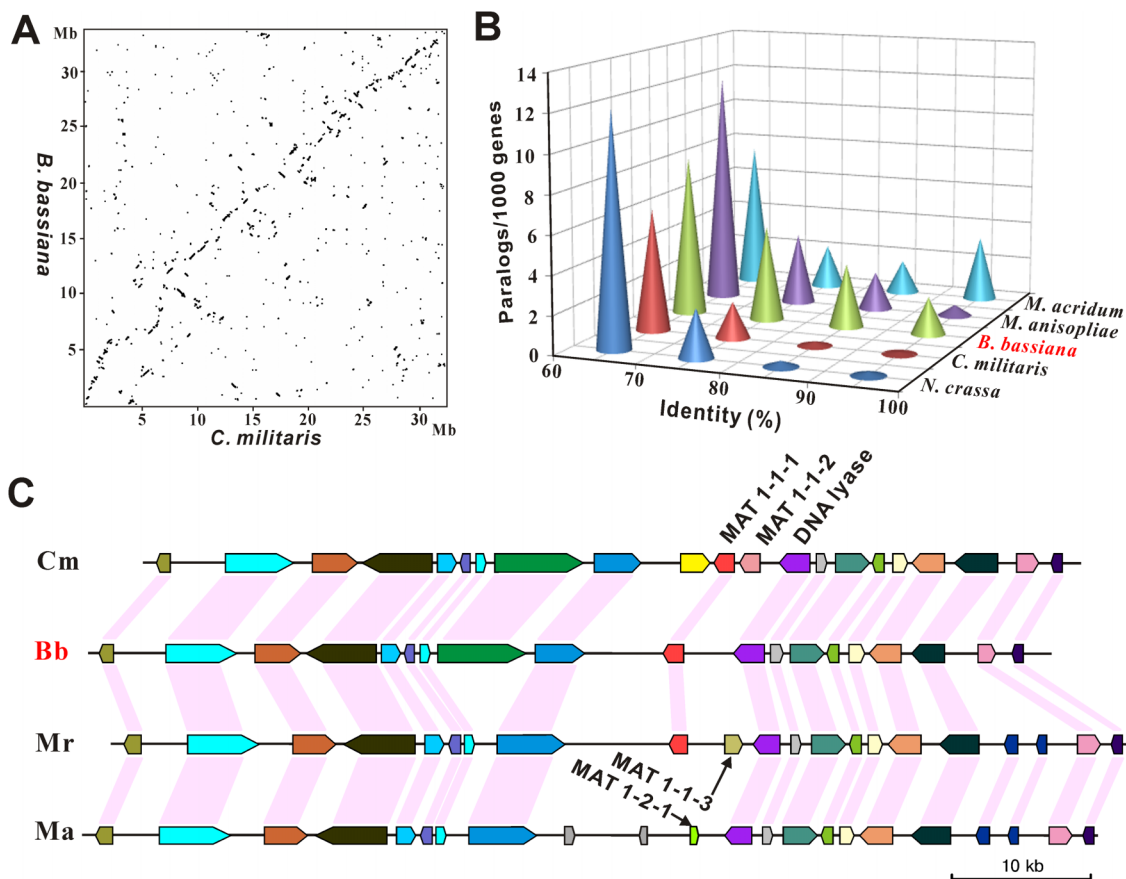
Figure 3 | Comparative genomic analysis. (A) Dot blot analysis of *B. bassiana* and *C. militaris* genome structures using ordered scaffold data. (B) Distribution of paralogous genes with different levels of nucleotide similarity in *B. bassiana* (highlighted in red) and other fungi. (C) Syntenic relationships of the MAT loci and their flanking regions between *B. bassiana* (Bb), *C. militaris* (Cm), *M. robertsii* (Mr) and *M. acridum* (Ma). Loci in the same color show orthologous relationships.

insects and plants. Bb has four genes encoding proteins with eight cysteine-containing extracellular membrane (CFEM) domains resembling pathogenicity determinants in plant pathogens[44].

Signal transduction. The PHI dataset contains large numbers of G-protein coupled receptors (GPCRs), protein kinases and transcription factors and that have similar sequences in the entomopathogen genomes (Supplementary Table S1). Fungal GPCRs sense extracellular cues and transmit the signals to distinct trimeric G-protein subunits[45]. Besides the conserved pheromone receptors and cAMP receptors, most of the insect pathogen GPCRs resemble the rice-blast fungus *M. oryzae* Pth11-like proteins (Supplementary Table S15). Relative to *Metarhizium* spp. (average 47), Bb and Cm have fewer Pth11-like receptors (average 21) and lack a GPR1-like GPCR which in yeast is activated during nitrogen starvation[46]. Thus, Bb and *Metarhizium* have evolved different mechanisms for nutrient sensing.

Functional kinome analysis of the plant pathogen *F. graminearum* indicated that many protein kinases (PKs) are involved in fungal growth, conidiation, pathogenesis, stress responses, toxin production and/or sexual reproduction[47]. The insect pathogens have more PKs (average 170, $P<0.05$) than the plant pathogens (average 145) (Supplementary Table S4). The specialist Ma has more PKs (193) than the generalist Mr (161), which may allow a more stringent discrimination between potential insect hosts and subsequent control of cell differentiation[17]. Consistent with this, Bb has fewer PKs (159) than Cm (167). Like other fungi, insect pathogens have large numbers of transcription factors (TFs) (Supplementary Table S16). However, Bb has more (10) GATA-type TFs than Cm (5) and *Metarhizium* species (4–5). Fungal GATA-type TFs are involved in

multiple functions, including nitrogen metabolism, light induction, siderophore biosynthesis, mating-type switching and chromatin rearrangement[48]. Overall, these results imply that signal controls vary as much between the insect pathogens as they do between insect and plant pathogens.

Cryptic sexuality. The teleomorph of Bb was identified as *C. bassiana*[11], but its sexual reproduction is seldom observed in nature nor to date is it inducible in the laboratory. A previous PCR based analysis showed that individual isolates of Bb carried either MAT1-1 or MAT1-2 mating-type genes[49]. Analysis of the mating-type locus indicated that the sequenced Bb strain is the MAT1-1 type (Fig. 3C). Thus, like Cm, Bb is a heterothallic and outcrossing fungus. Syntenic analysis of Bb, Cm and the two *Metarhizium* species showed that except for the idiomorphic regions, the genes flanking the mating-type locus are highly conserved, especially between Bb and Cm. However, unlike Bb and *Metarhizium* species, Cm commonly performs sexual reproduction[18]. To probe the cryptic sexuality of Bb, we surveyed Bb homologs of sex-related genes that have been functionally verified in *A. nidulans* and *N. crassa*[50] (Supplementary Table S17). Many genes functioning in mating processes, karyogamy, meiosis and fruiting-body development in *A. nidulans* and *N. crassa* are also present in Bb and Cm. However, a meiosis-specific topoisomerase Spo11 present in Cm (CCM_09527), *A. nidulans* and *N. crassa* is absent in Bb and *Metarhizium* spp. Spo11 is crucial for initiating meiotic recombination by generating DNA double-strand breaks[51]. Thus, lack of Spo11-like protein in Bb and *Metarhizium* spp. may contribute, at least in part, to an infrequent sexual cycle.

Secondary metabolism. A plethora of insecticidal and other bioactive secondary metabolites has been identified from Bb, e.g. the
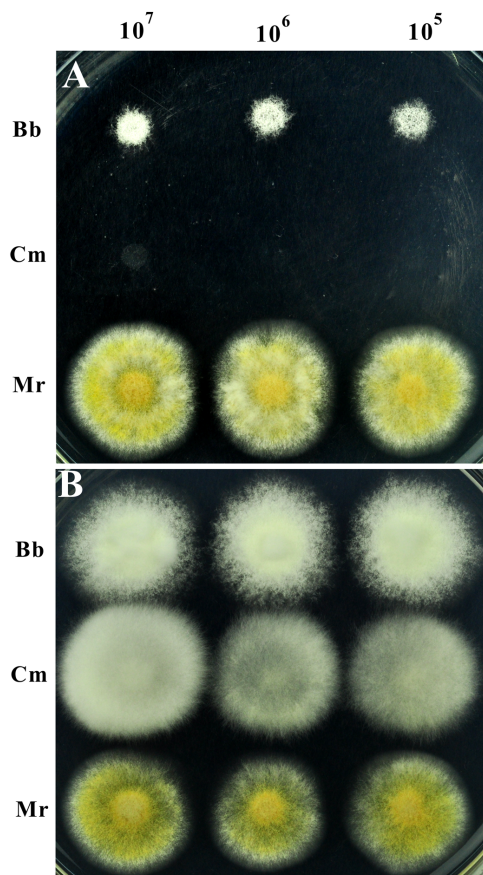
Figure 4 | Growth tests. Conidial suspensions (2.5 μl) of *B. bassiana* (Bb), *C. militaris* (Cm) and *M. robertsii* (Mr) were inoculated onto minimal medium supplemented with 1% xylose (A) or 1% glucose (B) and incubated for 10 days at 25°C. In contrast to Mr, Cm failed to grow on xylose medium while Bb germinated and grew poorly.

cyclopeptides beauvericin, bassianolide and beauverolide, the yellow pigment pyridines tenellin and bassiatin and the diben-zoquinone oosporein[52]. Only genes involved in the biosynthesis of beauvericin[37], bassianolide[38] and tenellin[39] have been functionally verified. Our genome survey found that there are 45 non-ribosomal peptide synthetase (NRPS), polyketide synthase (PKS) and terpenoid synthase/cyclase core genes in the Bb genome, which is more than Cm but fewer than *Metarhizium* spp. (Table 2). Three of the putative biosynthesis clusters are highly conserved in the four insect pathogens but are absent in other fungi, i.e. NPRS (BBA_05020), PKS (BBA_09745) (Supplementary Fig. S4A) and terpene synthase (BBA_06542). This finding implies that the evolution of fungal entomopathogenicity may be associated with the production of some similar secondary metabolites. Otherwise, there are four clusters conserved between insect pathogens and *T. reesei*, including two NRPS-like proteins (BBA_06997 and BBA_04028), one NRPS (BBA_05020) and one terpene cyclase (BBA_06542) (Supplementary Fig. S4B). Two clusters are limited to Bb, Cm and *T. reesei*, i.e. a type I PKS cluster and a PKS-NRPS hybrid for tenellin biosynthesis[39] (Supplementary Fig. S4C). Ten clusters are specifically conserved between Bb and Cm. However, the NPRS gene clusters involved in the biosynthesis of insecticidal toxins beauvericin by an iterative NRPS BbBEAS[37] and bassianolide by an iterative NRPS BbBSLS[38] are unique to the Bb genome.

Biotransformation and detoxification. Bb is a well-known whole-cell catalyst in industrial applications[8]. We compared Bb with other fungi for enzymes putatively involved in bioconversions to

determine if their presence or absence was associated with particular fungal lifestyles (Supplementary Table S18). Relative to plant pathogens, insect pathogens in general have more acyl-CoA dehydrogenases (average 14 vs. 9, $P=0.0139$), fatty acid hydroxylases (average 14 vs. 5, $P=0.0002$), amidohydrolases (average 14 vs. 9, $P=0.0210$), glyoxalases (average 4 vs. 1, $P=0.0023$) and mo-nooxygenases (average 24 vs. 13, $P=0.0445$). The acyl-CoA dehydrogenases are responsible for β-oxidation of fatty acids while amidohydrolases are involved in biotransformation of N-heterocyclic compounds[53]. The glyoxalase system detoxifies methylglyoxal[54]. Consistent with its unique sulfoxidation ability[55], Bb has three aryl sulfotransferases that are absent in other fungi. Among insect pathogens, Bb has the highest numbers of epoxide hydrolases (7 vs. 4–5), nitrilases (11 vs. 7–10 in others) and monooxygenases (36 vs. 16–30 in others). Epoxide hydrolases are highly versatile biocatalysts for the hydrolysis of epoxides[55]. Nitrilases are valuable alternatives to chemical catalysts for biotransformation of various organic nitriles[56].

Transcriptional responses of Bb to different environmental niches. Bb must overcome a number of environmental challenges in plants and insects. We used the high throughput Illumina RNA-seq method to investigate the range of Bb transcriptional responses. After sequencing, we obtained a total of 2.3 to 3.7 million tags per treatment, and these were mapped to predicted genes. We determined that, of 10,366 annotated Bb genes, 68%, 52% and 78% were transcriptionally active during growth on locust hind wings (LW), in cotton bollworm blood (CB) and in corn root exudates (RE), respectively (Supplementary Table S19). As reported in previous studies[17,18], individual gene expression values were normalized to the index of transcripts per million tags (TPM). Pairwise analysis of gene expression data was used to determine that large numbers of genes were significantly ($P<0.05$) up- or down-regulated by Bb during different treatments (Fig. 5A). For example, Bb up-regulated 1,150 genes and down regulated 2,868 genes in CB as compared to LW. When referenced to gene expression in RE, growth in CW and LW induced up-regulation of 3,302 and 4,530 genes, respectively. Of the top 100 genes expressed in CB, RE and LW, 74, 79 and 81, respectively, were not highly expressed in the other media (Fig. 5B), and more secreted proteins were in the top 100 expressed by Bb on LW (25) than in CB (15) and RE (14) (Supplementary Table S20). Of the 373 identified SSCPs, 92 were expressed on LW, 128 in CB and 156 in RE. The bacterial-like toxins were expressed at very low levels in all three environments.

The analysis showed that Bb differentially expressed transcription factors (TFs) for gene regulation (Fig. 5C), GPCR genes for niche recognition (Supplementary Fig. S5A) and kinases for signal trans-duction (Supplementary Fig. S5B). For example, a Pth11-like GPCR (BBA_03214) and a C2H2-type zinc finger TF (BBA_00971) were highly expressed by Bb in RE (TPM=3,163 for GPCR and TPM=1,648 for TF), whereas only low or trace expression occurred in CB (TPM=137 for GPCR and TPM=18 for TF) and on LW (TPM=4 for GPCR and TPM=5 for TF). A functional category analysis of the highly expressed genes indicated that around one third are involved in catabolism or anabolism as defined by FunCat (Supplementary Fig. S6A). Genes involved in amino acid metabolism were more highly transcribed by Bb on LW (30%) than in CB (19%) and RE (11%). Most of these were involved in catabolism, consistent with utilization of amino acids on the cuticle surface, and/or mobil-ization of internal nitrogenous nutrients. Fourteen percent of the most highly expressed genes in CB were involved in protein syn-thesis, as compared to 6% on LW and 2% in RE. This could be associated with the quick propagation of Bb cells in insect hemocoel by yeast-type budding[57]. Root exudates are carbohydrate rich, and 49% of the highly expressed genes in RE were for carbohydrate metabolism versus 14% on LW and 30% in CB (Supplmentary Fig.
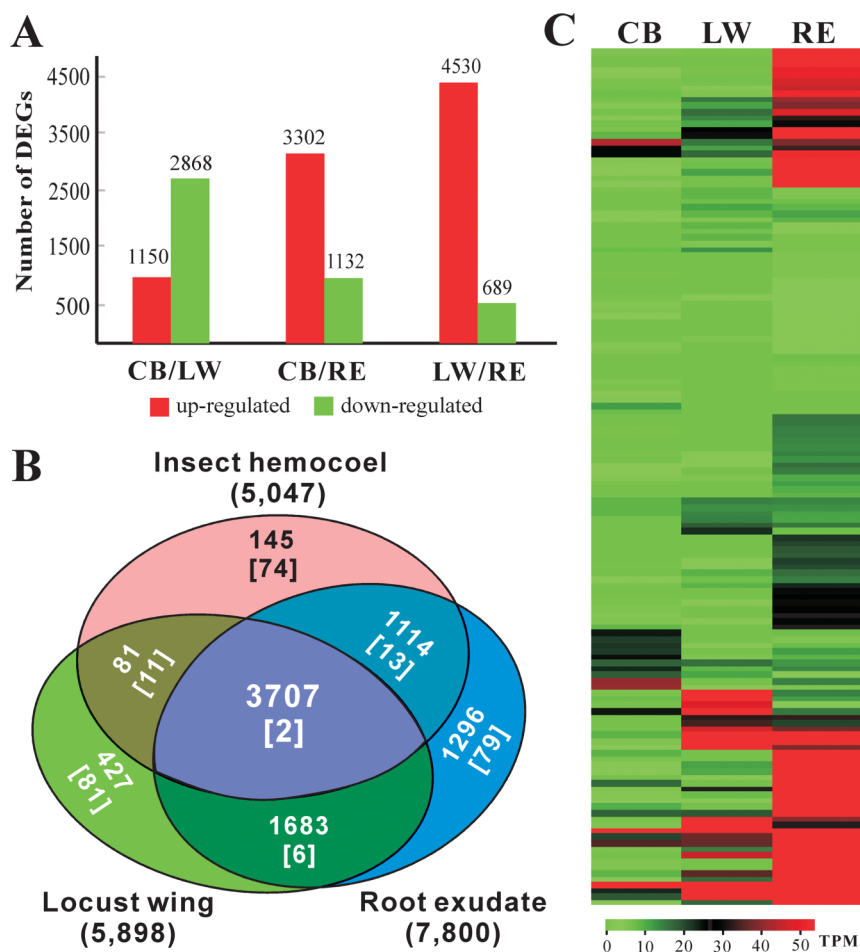
Figure 5 | RNA-seq analysis of differentially expressed genes by *B. bassiana* adapting to different environmental niches. (A) Summary of significantly up- and down-regulated genes between different libraries. (B) Venn diagram analysis of co-expressed genes in different libraries. The numbers in parenthesis show total expressed genes with more than two transcripts. The numbers in square brackets show those of the 100 most highly expressed genes. (C) Heat map of the transcription factors differentially expressed by Bb in different conditions. The fungus was harvested from cotton bollworm hemocoel (CB), locust hind wings (LW) and corn root exudates (RE) for RNA extraction and RNA-seq analysis (TPM, transcripts per million tags).

S6B). Six heat-shock proteins were among the top 100 genes expressed by Bb in RE, compared to four in CB and one on LW, suggesting that plant root exudates provide the most stressful growth conditions.

Discussion

We report here a genomic analysis of *B. bassiana*, one of the best-studied and most widely used insect biocontrol agents. A comparative analysis with the genome sequences of three other insect pathogens demonstrated that Bb and Cm are closely related and evolved into insect pathogens independently of the *Metarhizium* lineage. We assume therefore that similar expansion of certain gene families, such as proteases and chitinases, is associated with functions necessary for insect pathogenesis and reflects convergent evolution. Likewise, plant pathogens have expanded families of glycoside hydrolases, carbohydrate esterases, cutinases and pectin lyases in order to degrade plant materials[58]. Mammalian pathogens are enriched for aspartyl proteases and phospholipases[59]. Mycoparasitic fungi have expanded numbers of chitinases[20].

Besides proteases and chitinases[1], virulence-related genes already characterized in Bb include a MAP kinase BBSLT2 (BBA_03334) mediating cell growth[60], a neuronal calcium sensor BBCSA1 (BBA_05195) regulating extracellular acidification[61], a cytochrome P450 enzyme CYP52X1 degrading cuticular fatty acids (BBA_02428)[36] and a GH73 family of β-1,3-glucanosyltransferase BBGAS1 (BBA_04640) maintaining cell well integrity[62]. All of these genes were mapped in the sequenced

Bb genome. Several other experimentally verified virulence genes in *Metarhizium* spp. also have orthologs in Bb, e.g., a perilipin-like protein (BBA_08759, 63% identity) that controls cellular lipid storage and appressorium penetration[63], an osmosensor (BBA_08887, 59% identity) to mediate adaptation to the insect hemocoel[64], adhesins MAD1 (BBA_02419, 34% identity) and MAD2 (BBA_02379, 47% identity) to mediate spore adhesion to insect and plant surfaces[65]. The mechanism of adhesion used by Bb also involves hydrophobins that mediate cell surface hydrophobicity and virulence[66]. Like other insect pathogens, Bb has two class I hydrophobins but it has additional class II hydrophobins (Supplementary Table S4). The presence of these genes in both Bb and *Metarhizium* spp. suggests some shared strategies for interacting with plants and insects. Ninety-one of the 397 SSCP clusters are shared exclusively by the insect pathogens suggesting many shared strategies are currently unknown. However, *Beauveria* lacks a homolog to the collagen-like protein used by *Metarhizium* to evade the insect immune system[67]. Bb also lacks the *Metarhizium dtxS1* (MAA_10043) gene cluster involved in biosynthesis of the insecticidal destruxins[40]. It is likely that some of the highly expressed genes unique to Bb will play important and novel roles as Bb overcomes challenges in the dynamic microenvironments it will encounter in insects. As an endophyte, Bb presumably possesses mechanisms to avoid stimulating plant defenses. Fungal endoxylanases (GH11) are known to trigger plant immune responses and induce necrosis of infected plant tissue[68]. Therefore, lack of GH11 in Bb and *E. festucae* could be an adaptation limiting induction of necrosis by these endophytes, and facilitating immune evasion. Like

the basidiomycete plant symbiont *L. bicolor*[42], Bb has a large battery of SSCPs and about half of them (154/373) are species-specific, implying that many specific functions are required for specialization to insect pathogenesis and endophytism.

If the finding that asexual *Aspergillus* species usually arise from sexual lineages[69] is broadly applicable to fungi then *Beauveria* spp. are probably an asexual derivation from a *Cordyceps* lineage. Host switching is particularly common in *Cordyceps* spp. accounting for their wide variety of associations with animals, plants and fungi[70]. Some *Trichoderma* spp., such as *T. strigosum*, have a *Cordyceps* teleomorph. Our phylogenomic data suggests that the insect pathogenic Cm and Bb diverged from mycoparasitic *Trichoderma* 74–97 million years ago (Fig. 2B), so potentially mycoparasitism could have evolved from insect pathogenicity or vice versa. The high degree of genome structure divergence between Bb and Cm is unexpected given their close phylogenetic relationship. Transposable elements (TEs) are a major force driving genetic variation and genome evolution[71]. Bb has many more TEs than Cm, apparently because Bb lacks the genome defense mechanism of repeat-induced point mutations. However, the genomes of *Metarhizium* species are highly syntenic in spite of a similar difference in the number of TEs[17]. Most field populations of *Beauveria* and *Metarhizium* species reproduce clonally[72]. In contrast, Cm readily reproduces sexually[18], thereby facilitating genome structure reorganization due to frequent genetic and/or chromosomal recombination. Thus, differences in life cycle might have led to the genome structure disparities between Bb and Cm.

High throughput transcriptomics demonstrated that Bb finely tunes gene transcription to adapt to different environmental niches. Bb up-regulated proteases on LW and carbohydrate hydrolases in RE. For example, a subtilisin-like protease (BBA_00443, TPM=230) was highly expressed by Bb on LW but not in CB or in RE (Supplementary Table S20). The insect epicuticle or waxy layer comprises a heterogeneous mixture of long-chain alkanes, wax esters and fatty acids, and represents the first barrier against fungal attack. Bb highly transcribed CYP52s (BBA_02428 and BBA_09022) and lipase genes (BBA_01783 and BBA_08812) which should target epicuticular hydrocarbons and lipids[36]. Chitin constitutes up to 40% of the procuticle but is absent from the epicuticular layer[73]. As with *Metarhizium*[17], Bb does not significantly up-regulate chitinase genes on the epicuticle. Two SSCPs (GH75 chitosanase BBA_06270 and CFEM protein BBA_09339) were highly and specifically expressed in CB (Supplementary Table S20). Potentially, chitosanase BBA_06270 could be involved in remodeling cell wall structure to evade host immune recognition[74]. Homologs of CFEM BBA_09339 are involved in plant pathogenesis[44]. Future functional verifications of these genes will benefit our understanding of fungal pathogenesis.

In conclusion, we have sequenced the genome of the well-known insect pathogenic fungus *B. bassiana* and used RNA-seq to generate expression profiles of Bb growing in different host or environmental niches. The resulting information will benefit future molecular studies of insect-fungus interactions, and facilitate the development of *Beauveria* as cost-effective mycoinsecticides and microbial biocatalysts.

Methods

Fungal strain and maintenance. *B. bassiana* strain ARSEF 2860 was selected for genome sequencing as it shows commercial potential for biological controls of aphids, planthoppers, and spider mites[75]. To compare fungal abilities to utilize pentose and hexose, *M. robertsii* strain ARSEF 23 and *C. militaris* Cm01 were included with Bb for growth tests on minimum medium (MM, NaNO$_3$ 6 g/l, KCl 0.52 g/l, MgSO$_4$.7H$_2$O 0.52 g/l, KH$_2$PO$_4$ 0.25 g/l) amended with 1% xylose and glucose, respectively. Otherwise, fungal cultures were maintained on potato dextrose agar (Difco).

Genome sequencing and assembly. The genome of Bb ARSEF 2860 strain was shotgun sequenced using a Roche 454 GS FLX system for massively parallel pyrosequencing at the Chinese National Human Genome Center (Shanghai, China). This resulted in 930 Mb of sequence data (with an average read length of 385 bp). Assembly was performed using the Newbler software (Ver. 2.3) within the Roche 454 suite package[76], which produced 1,764 contigs with a total size of 33.7 Mb. For

sequence scaffolding, a DNA library of 2–5 kb inserts was generated and sequenced with an Illumina system. This resulted in 1.3 Gb of paired-end reads and by mapping these reads to contigs, 1,764 contigs were assembled into 242 scaffolds. The whole project has been deposited at DDBJ/EMBL/GenBank under the accession no. ADAH00000000.

Genome annotation. To maximize accuracy, the gene structures of *B. bassiana* were predicted with a combination of different algorithms[17,18]. The inconsistent ORFs were individually subject to Blast searches against the NCBI curated refseq_protein database and manually inspected. Previously acquired ESTs[14] were used to verify and complete the predicted gene models. All predicted gene models were annotated by InterproScan analysis (http://www.ebi.ac.uk/Tools/pfa/iprscan/). The potential secreted proteins were predicted by SignaIP 3.0 (http://www.cbs.dtu.dk/services/SignalP/) and TargetP (http://www.cbs.dtu.dk/services/TargetP/) analysis. Genome repetitive elements were analyzed by Blast against the RepeatMasker library (Open 3.2.9) (http://www.repeatmasker.org/) and with the Tandem Repeat Finder (http://tandem.bu.edu/trf/trf.html). The transposases/retrotransposases were classified by Blastp analysis against the Repbase (http://www.girinst.org/repbase/) plus manual inspections. Putative *Beauveria* virulence factors were identified by searching against the pathogen-host interaction database (http://www.phi-base.org/about.php) with a cut-off *E* value of 1e-5, plus additional searches of known virulence genes reported in entomopathogenic fungi. One tail *t*-tests were conducted to compare the difference of protein family sizes between insect pathogens and other fungi.

Comparative genomic analysis. For genome structure comparison of Bb and Cm, the scaffolds of both genomes were oriented by MEGABLAST for dot plotting and a pair-wise comparison with an Argo Genome Browser[77]. A Blast Score Ratio (BSR) test[19] was conducted to compare the differences between Bb, Cm and Mr genomes. The BSR index for each reference protein is calculated by dividing the query score by the reference score and normalized from 0 to 1. A score of 1 indicates a perfect match while a score of 0 indicates no Blast match of a query protein in the reference proteome. The normalized pairs of BSR indices were then plotted using the Matlab (ver. 7.0) program.

Paralogy, orthology and phylogenomic analysis. The best candidate paralogs in examined fungal genomes were identified by reciprocal Blastn analysis of the coding DNA sequences with a cut-off *E* value of 1e-20 and more than 60% coverage of Blast alignment length. For Bb and Cm paralogs with more than 70% identity, the paired sequences were aligned with ClustalW and the nucleotide mutation ratios were estimated and compared. Ortholog conservation in fungi was characterized with Inparanoid 7.0 (http://inparanoid.sbc.su.se/cgi-bin/index.cgi). Corresponding orthologous protein sequences were aligned with Clustal X 2.0 and the concatenated amino acid sequences were used to generate a maximum likelihood phylogenomic tree with the program TREE-PUZZLE[78] using a Dayhoff model. Based on the constructed phylogenomic tree, protein family size variation (expansion or contraction) between Bb and Cm was analyzed using the program CAFE[79] by referencing against the most closely related species, *T. virens*.

Protein family classifications. To identify the gene clusters and their proteins responsible for the biosynthesis of secondary metabolites, the whole genome data set was subject to analysis with the programs SMURF (http://jcvi.org/smurf/index.php) and antiSMASH (http://antismash.secondarymetabolites.org/) with default settings. The families of proteases were identified by Blastp searching against the MEROPS peptidase database Release 9.4 (http://merops.sanger.ac.uk/) with a cutoff *E* value of 1e-100 plus manual inspections with the InterproScan analysis results. Fungal trypsins were selected for phylogenetic analysis with the program MEGA 5.0[80] using a Dayoff model, 1,000 replicates for bootstrap analysis and a pairwise deletion for gaps or missing data. The cytochrome P450s were named according to the classifications collected at the P450 database (http://blast.uthsc.edu/). Kinases were classified by Blastp against the KinBase (http://kinase.com/) with a cutoff *E* value of 1e-30. Carbohydrate-active enzymes were classified by local Blastp searching against a library of catalytic and carbohydrate-binding module enzymes (http://www.cazy.org/). G-proteins were identified from PFAM domain scanning. G-protein coupled receptors were selected from the best hits to GPCRDB sequences (http://www.gpcr.org/7tm/) and by confirmation that they contained seven transmembrane helices with the N-terminus outside and the C-terminus inside the plasma membrane. To identify the small secreted cysteine-rich proteins, the proteins less than or equal to 300 amino acid with secreted signals obtained above and those containing four or more cysteine residues were included for Blastclust analysis (http://toolkit.tuebingen.mpg.de/blastclust) at the cutoffs of coverage 80% and identity 20%[20].

Transcriptome analysis. Conidia of *B. bassiana* ARSEF 2860 strain were harvested from 14-day old potato dextrose agar and used for different assays. To examine gene induction on insect cuticle, locust (*Locusta migratoria*) hind wings were collected, air-dried and surface sterilized in 10% H$_2$O$_2$ (10 min). The wings were washed in sterile water (twice) and immersed in a Bb conidial suspension (2 × 10^7 spores per ml) for 20 seconds[17]. The inoculated wings were placed on 1% water agar and incubated at 25°C for 24 hrs for total RNA extraction. For analysis of transcriptional adaptation to insect hemocoel, the 5th instar cotton bollworm (*Helicoverpa armigera*) larvae were each injected with 10 μl of a spore suspension (10^8 spores/ml). Hemolymph from infected insects 48 hours post inoculation was collected on ice and immediately

applied on top of a step gradient of 25 and 50% Centricoll (Sigma). The fungal cells were purified for RNA extraction by centrifugation at 10,000 g for 10 min at 4°C. For analysis of transcriptional adaptation to plant root exudates, mycelia harvested from 36 hour Sabouraud dextrose broth were incubated in corn root exudates for another 24 hours before being used for RNA extraction. Root exudates were prepared as described before[81]. RNA was extracted with a Qiagen RNeasy kit plus on-column treatment with RNase-free DNase I. Messenger RNA was purified, and after reverse transcription into cDNA libraries were constructed for tag preparation according to the massively parallel signature sequencing protocol[82]. The tags were sequenced with an Illumina technique. We omitted tags from further analysis if only one copy was detected or it could be mapped to a different transcript. Other tags were mapped to the genome or annotated genes if they possessed no more than one nucleotide mismatch[17,18]. The level of gene transcription was converted to transcripts per million tags (TPM) for each mapped gene for expressional comparison between samples. The RNA_seq expression dataset is available at the Gene Expression Omnibus under the accession GSE32699.

1. St. Leger, R. J. & Wang, C. S. Genetic engineering of fungal biocontrol agents to achieve greater efficacy against insect pests. *Appl Microbiol Biotechnol* **85**, 901–907 (2010).
2. Thomas, M. B. & Read, A. F. Can fungal biopesticides control malaria? *Nat Rev Microbiol* **5**, 377–383 (2007).
3. de Faria, M. R. & Wraight, S. P. Mycoinsecticides and Mycoacaricides: A comprehensive list with worldwide coverage and international classification of formulation types. *Biol Control* **43**, 237–256 (2007).
4. Wang, C. S., Fan, M. Z., Li, Z. Z. & Butt, T. M. Molecular monitoring and evaluation of the application of the insect-pathogenic fungus *Beauveria bassiana* in southeast China. *J Appl Microbiol* **96**, 861–870 (2004).
5. Li, Z. Z. *et al.* Biological control of insects in Brazil and China: history, current programs and reasons for their successes using entomopathogenic fungi. *Biocontrol Sci Technol* **20**, 117–136 (2010).
6. Steinhaus, E. A. Microbial control-The emergence of an idea: A brief history of insect pathology through the nineteenth century. *Hilgardia* **26**, 107–160 (1956).
7. Gottar, M. *et al.* Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. *Cell* **127**, 1425–1437 (2006).
8. Griffiths, D. A., Brown, D. E. & Jezequel, S. G. Metabolism of xenobiotics by *Beauveria bassiana*. *Xenobiotica* **23**, 1085–100 (1993).
9. Ownley, B. H. *et al.* Beauveria bassiana: endophytic colonization and plant disease control. *J Invertebr Pathol* **98**, 267–270 (2008).
10. Bidochka, M. J., Clark, D. C., Lewis, M. W. & Keyhani, N. O. Could insect phagocytic avoidance by entomogenous fungi have evolved via selection against soil amoeboid predators? *Microbiology* **156**, 2164–2171 (2010).
11. Li, Z. Z. *et al.* Discovery and demonstration of the teleomorph of *Beauveria bassiana* (Bals.) Vuill., an important entomogenous fungus. *Chinese Sci Bull* **46**, 751–753 (2001).
12. Xu, Y. *et al.* Biosynthesis of the cyclooligomer depsipeptide bassianolide, an insecticidal virulence factor of *Beauveria bassiana*. *Fungal Genet Biol* **46**, 353–364 (2009).
13. Koo, B. S. *et al.* Bombycis corpus extract (BCE) protects hippocampal neurons against excitatory amino acid-induced neurotoxicity. *Immunopharmacol Immunotoxicol* **25**, 191–201 (2003).
14. Cho, E. M., Liu, L., Farmerie, W. & Keyhani, N. O. EST analysis of cDNA libraries from the entomopathogenic fungus *Beauveria* (*Cordyceps*) *bassiana*. I. Evidence for stage-specific gene expression in aerial conidia, *in vitro* blastospores and submerged conidia. *Microbiology* **152**, 2843–2854 (2006).
15. Wang, Z. L., Lu, J. D. & Feng, M. G. Primary roles of two dehydrogenases in the mannitol metabolism and multi-stress tolerance of entomopathogenic fungus *Beauveria bassiana*. *Environ Microbiol* doi: 10.1111/j.1462-2920.2011.02654.x (2011).
16. Fan, Y., Zhang, S., Kruer, N. & Keyhani, N. O. High-throughput insertion mutagenesis and functional screening in the entomopathogenic fungus Beauveria bassiana. *J Invertebr Pathol* **106**, 274–279 (2011).
17. Gao, Q. *et al.* Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. *PLoS Genet* **7**, e1001264 (2011).
18. Zheng, P. *et al.* Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional Chinese medicine. *Genome Biol* **12**, R116 (2011).
19. Ma, L. J. *et al.* Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* **464**, 367–373 (2010).
20. Kubicek, C. P. *et al.* Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. *Genome Biol* **12**, R40. (2011).
21. Galagan, J. E. *et al.* Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* **438**, 11020055–1115.
22. Fedorova, N. D. *et al.* Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. *PLoS Genet* **4**, e1000046 (2008).
23. Galagan, J. E. *et al.* The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**, 859–868 (2003).
24. Qin, Y. *et al.* Integration of insecticidal protein Vip3Aa1 into *Beauveria bassiana* enhances fungal virulence to *Spodoptera litura* larvae by cuticle and per Os infection. *Appl Environ Microbiol* **76**, 4611–4618 (2010).
25. Mutschler, H. *et al.* A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biol* **9**, e1001033 (2011).
26. Fitzpatrick, D. A. Horizontal gene transfer in fungi. *FEMS Microbiol Lett* **329**, 1–8 (2012).
27. Noda, J., Brito, N. & González, C. The *Botrytis cinerea* xylanase Xyn11A contributes to virulence with its necrotizing activity, not with its catalytic activity. *BMC Plant Biol* **10**, 38 (2010).
28. Duan, Z. B., Shang, Y. F., Gao, Q., Zheng, P. & Wang, C. S. A phosphoketolase Mpk1 of bacterial origin is adaptively required for full virulence in the insect-pathogenic fungus *Metarhizium anisopliae*. *Environ Microbiol* **11**, 2351–2360 (2009).
29. Hugouvieux-Cotte-Pattat, N., Condemine, G., Nasser, W. & Reverchon, S. Regulation of pectinolysis in *Erwinia chrysanthemi*. *Ann Rev Microbiol* **50**, 213–257 (1996).
30. Seidl, V. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. *Fungal Biol Rev* **22**, 36–42 (2008).
31. Crešnar, B. & Petrič, S. Cytochrome P450 enzymes in the fungal kingdom. *Biochim Biophys Acta* **1814**, 29–35 (2011).
32. Takaya, N., Uchimura, H., Lai, Y. & Shoun, H. Transcriptional control of nitric oxide reductase gene (CYP55) in the fungal denitrifier *Fusarium oxysporum*. *Biosci Biotechnol Biochem* **66**, 1039–1045 (2002).
33. Artigot, M. P. *et al.* Molecular cloning and functional characterization of two CYP619 cytochrome P450s involved in biosynthesis of patulin in *Aspergillus clavatus*. *Microbiology* **155**, 1738–1747 (2009).
34. Ferreira, M. E. *et al.* The ergosterol biosynthesis pathway, transporter genes, and azole resistance in *Aspergillus fumigatus*. *Med Mycol* **43**, S313–319 (2005).
35. Lin, L. C. *et al.* The *MrCYP52* cytochrome P450 monoxygenase gene of *Metarhizium robertsii* is important for utilizing insect epicuticular hydrocarbons. *PLoS ONE* **6**, e28984 (2011).
36. Zhang, S. *et al.* CYP52X1, representing new cytochrome P450 subfamily, displays fatty acid hydroxylase activity and contributes to virulence and growth on insect cuticular substrates in entomopathogenic fungus *Beauveria bassiana*. *J Biol Chem* **287**, 13477–13486 (2012).
37. Xu, Y. *et al.* Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus *Beauveria bassiana*. *Chem Biol* **15**, 898–907 (2008).
38. Xu, Y. *et al.* Biosynthesis of the cyclooligomer depsipeptide bassianolide, an insecticidal virulence factor of *Beauveria bassiana*. *Fungal Genet Biol* **46**, 353–364 (2009).
39. Eley, K. L. *et al.* Biosynthesis of the 2-pyridone tenellin in the insect pathogenic fungus *Beauveria bassiana*. *Chembiochem* **8**, 289–297 (2007).
40. Wang, B., Kang, Q., Lu, Y. Z., Bai, L. Q. & Wang, C. S. Unveiling the biosynthetic puzzle of destruxins in *Metarhizium* species. *Proc Natl Acad Sci USA* **109**, 1287–1292 (2012).
41. van Bogaert, I. N., Groeneboer, S., Saerens, K. & Soetaert, W. The role of cytochrome P450 monooxygenases in microbial fatty acid metabolism. *FEBS J* **278**, 206–221 (2011).
42. Martin, F. *et al.* The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* **452**, 88–92 (2008).
43. Rappleye, C. A. & Goldman, W. E. Fungal stealth technology. *Trends Immunol* **29**, 18–24 (2008).
44. Kulkarni, R. D., Kelkar, H. S. & Dean, R. A. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends Biochem Sci* **28**, 118–121 (2003).
45. Xue, C., Hsueh, Y. P. & Heitman, J. Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. *FEMS Microbiol Rev* **32**, 1010–1032 (2008).
46. Xue, Y., Batlle, M. & Hirsch, J. P. GPR1 encodes a putative G protein-coupled receptor that associates with the Gpa2p Galpha subunit and functions in a Ras-independent pathway. *EMBO J* **17**, 1996–2007(1998).
47. Wang, C. *et al.* Functional analysis of the kinome of the wheat scab fungus *Fusarium graminearum*. *PLoS Pathog* **7**, e1002460 (2011).
48. Scazzocchio, C. The fungal GATA factors. *Curr Opin Microbiol* **3**, 126–131 (2000).
49. Yokoyama, E., Arakawa, M., Yamagishi, K. & Hara, A. Phylogenetic and structural analyses of the mating-type loci in Clavicipitaceae. *FEMS Microbiol Lett* **264**, 182–191 (2006).
50. Dyer, P. S. & O'Gorman, C. M. Sexual development and cryptic sexuality in fungi: insights from *Aspergillus* species. *FEMS Microbiol Rev* **36**, 165–192 (2012).
51. Panizza, S. *et al.* Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. *Cell* **146**, 372–383 (2011).
52. Molnár, I., Gibson, D. M. & Krasnoff, S. B. Secondary metabolites from entomopathogenic Hypocrealean fungi. *Nat Prod Rep* **27**, 1241–1275 (2010).
53. Poirier, Y., Antonenkov, V. D., Glumoff, T. & Hiltunen, J. K. Peroxisomal beta-oxidation-a metabolic pathway with multiple functions. *Biochim Biophys Acta* **1763**, 1413–1426 (2006).
54. Inoue, Y., Maeta, K. & Nomura, W. Glyoxalase system in yeasts: structure, function, and physiology. *Semin Cell Dev Biol* **22**, 278–284 (2011).
55. Orru, R. V., Archelas, A., Furstoss, R. & Faber, K. Epoxide hydrolases and their synthetic applications. *Adv Biochem Eng Biotechnol* **63**, 145–167 (1999).
56. Martínková, L. & Kren, V. Biotransformations with nitrilases. *Curr Opin Chem Biol* **14**, 130–137 (2010).

57. Boucherie, H. Protein synthesis during transition and stationary phases under glucose limitation in *Saccharomyces cerevisiae*. *J Bacteriol* **161**, 385–392(1985).

58. Xu, J. R., Peng, Y. L., Dickman, M. B. & Sharon, A. The dawn of fungal pathogen genomics. *Annu Rev Phytopathol* **44**, 337–366 (2006).

59. van Asbeck, E. C., Clemons, K. V. & Stevens, D. A. *Candida* parapsilosis: a review of its epidemiology, pathogenesis, clinical aspects, typing and antimicrobial susceptibility. *Crit Rev Microbiol* **35**, 283–309 (2009).

60. Luo, X. *et al.* The MAP kinase Bbslt2 controls growth, conidiation, cell wall integrity, and virulence in the insect pathogenic fungus *Beauveria bassiana*. *Fungal Genet Biol* doi: 10.1016/j.fgb.2012.05.002 (2012).

61. Fan, Y., Ortiz-Urquiza, A., Kudia, R. A. & Keyhani, N. O. Bbcsa1, a fungal homolog of neuronal calcium sensor-1 regulates extracellular acidification and contributes to virulence in the entomopathogenic fungus *Beauveria bassiana*. *Microbiology* doi:10.1099/mic.0.058867-0 (2012).

62. Zhang, S., Xia, Y. X. & Keyhani, N. O. Contribution of the gas1 gene of the entomopathogenic fungus *Beauveria bassiana*, encoding a putative glycosylphosphatidylinositol-anchored beta-1,3-glucanosyltransferase, to conidial thermotolerance and virulence. *Appl Environ Microbiol* **77**, 2676-2684 (2011).

63. Wang, C. S. & St. Leger, R. J. The *Metarhizium anisopliae* perilipin homolog MPL1 regulates lipid metabolism, appressorial turgor pressure, and virulence. *J Biol Chem* **282**, 21110–21115 (2007).

64. Wang, C. S., Duan, Z. B. & St. Leger, R. J. The MOS1 osmosensor of *Metarhizium anisopliae* is required for adaptation to insect host hemolymph. *Eukaryot Cell* **7**, 302–309 (2008).

65. Wang, C. S. & St. Leger, R. J. The MAD1 adhesin of *Metarhizium anisopliae* links adhesion with blastospore production and virulence to insects, and the MAD2 adhesin enables attachment to plants. *Eukaryot Cell* **6**, 808–816 (2007).

66. Zhang, S. Z., Xia, Y. X., Kim, B. & Keyhani, N. O. Two hydrophobins are involved in fungal spore coat rodlet layer assembly and each play distinct roles in surface interactions, development and pathogenesis in the entomopathogenic fungus, *Beauveria bassiana*. *Mol Microbiol* **80**, 811–826 (2011).

67. Wang, C. S. & St. Leger, R. J. A collagenous protective coat enables *Metarhizium anisopliae* to evade insect immune responses. *Proc Natl Acad Sci USA* **103**, 6647–6652 (2006).

68. Noda, J., Brito, N. & González, C. The *Botrytis cinerea* xylanase Xyn11A contributes to virulence with its necrotizing activity, not with its catalytic activity. *BMC Plant Biol* **10**, 38 (2010).

69. Geiser, D. M., Timberlake, W. E. & Arnold, M. L. Loss of meiosis in *Aspergillus*. *Mol Biol Evol* **13**, 809–817 (1996).

70. Spatafora, J. W. *et al*. Phylogenetic evidence for an animal pathogen origin of ergot and the grass endophytes. *Mol Ecol* **16**, 1701–1711 (2007).

71. Daboussi, M. J. & Capy, P. Transposable elements in filamentous fungi. *Annu Rev Microbiol* **57**, 275–299 (2003).

72. Meyling, N. V. *et al*. Community composition, host range and genetic structure of the fungal entomopathogen *Beauveria* in adjoining agricultural and seminatural habitats. *Mol Ecol* **18**, 1282–1293 (2009).

73. Merzendorfer, H. & Zimoch, L. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J Exp Biol* **206**, 4393–4412 (2003).

74. Nahar, P., Ghormade, V. & Deshpande, M. V. The extracellular constitutive production of chitin deacetylase in *Metarhizium anisopliae*: a possible edge to entomopathogenic fungi in the biological control of insect fungi. *J Invertebr Pathol* **85**, 80–88 (2004).

75. Feng, M. G., Johnson, J. B. & Kish, L. P. Survey of entomopathogenic fungi naturally infecting cereal aphids in irrigated cereal crops in southwestern Idaho. *Environ Entomol* **19**, 1534–1542 (1990).

76. Schatz, M. C., Delcher, A. L. & Salzberg, S. L. Assembly of large genomes using second-generation sequencing. *Genome Res* **20**, 1165–1173(2010).

77. Engels, R. *et al*. Combo: a whole genome comparative browser. *Bioinformatics* **22**, 1782–1783 (2006).

78. Schmidt, H. A., Petzold, E., Vingron, M. & von Haeseler A. Molecular phylogenetics: parallelized parameter estimation and quartet puzzling. *J Parallel Distrib Comput* **63**, 719–727 (2003).

79. de Bie, T., Cristianini, N., Demuth, J. P. & Hahn, M. W. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269–1271 (2006).

80. Tamura, K. *et al*. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739 (2011).

81. Wang, C. S., Hu, G. & St. Leger, R. J. Differential gene expression by *Metarhizium anisopliae* growing in root exudate and host (*Manduca sexta*) cuticle or hemolymph reveals mechanisms of physiological adaptation. *Fungal Genet Biol* **42**, 704–718 (2005).

82. Brenner, S. *et al*. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol* **18**, 630–634 (2000).

Acknowledgments

Author contributions

CW and MGF designed the research. GX, SHY, RJS, GPZ and CW analyzed the data. PZ, ZLW, SZ, XQX and YS performed the experiments. RJS, CW and MGF wrote the paper. All authors reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xiao, G. *et al*. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Sci. Rep.* **2**, 483; DOI:10.1038/srep00483 (2012).