Received: 25 January 2018 | Revised: 8 June 2018 | Accepted: 20 June 2018

DOI: 10.1002/jez.b.22824

RESEARCH ARTICLE

WILEY JEZ-B MOLECULAR AND DEVELOPMENTAL EVOLUTION

Expansions of key protein families in the German cockroach highlight the molecular basis of its remarkable success as a global indoor pest

Mark C. Harrison[1] | Nicolas Arning[1] | Lukas P. M. Kremer[1] | Guillem Ylla[2] | Xavier Belles[2] | Erich Bornberg-Bauer[1] | Ann Kathrin Huylmans[3] | Evelien Jongepier[1] | Maria-Dolors Piulachs[2] | Stephen Richards[4] | Coby Schal[5]

[1]Institute for Evolution and Biodiversity, University of Münster, Münster, Germany

[2]Institut de Biologia Evolutiva, CSIC-University Pompeu Fabra, Barcelona, Spain

[3]Institute of Science and Technology Austria, Klosterneuburg, Austria

[4]Human Genome Sequencing Center, Department of Human and Molecular Genetics, Baylor College of Medicine, Houston, Texas

[5]Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, North Carolina

Correspondence
Mark C Harrison, Institute for Evolution and Biodiversity, University of Münster, Münster, Hüfferstrae 1, D-48149, Germany.
EMail: m.harrison@wwu.de

Funding information
M.C.H was supported by DFG grant no. BO2544/11-1 to E.B.B. C.S. by United States National Science Foundation grant IOS1557864. X.B. and M.-D.P. were supported by the Spanish Ministry of Economy and Competitiveness (grants CGL2015-64727-P to X.B. and CGL2016-76011-R to M.-D.P, both with FEDER funds) and Catalan Government (grant 2017 SGR 1030).

*Mark C. Harrison, Nicolas Arning, and Lukas P.M. Kremer contributed equally to this work.

Abstract

The German cockroach, *Blattella germanica*, is a worldwide pest that infests buildings, including homes, restaurants, and hospitals, often living in unsanitary conditions. As a disease vector and producer of allergens, this species has major health and economic impacts on humans. Factors contributing to the success of the German cockroach include its resistance to a broad range of insecticides, immunity to many pathogens, and its ability, as an extreme generalist omnivore, to survive on most food sources. The recently published genome shows that *B. germanica* has an exceptionally high number of protein coding genes. In this study, we investigate the functions of the 93 significantly expanded gene families with the aim to better understand the success of *B. germanica* as a major pest despite such inhospitable conditions. We find major expansions in gene families with functions related to the detoxification of insecticides and allelochemicals, defense against pathogens, digestion, sensory perception, and gene regulation. These expansions might have allowed *B. germanica* to develop multiple resistance mechanisms to insecticides and pathogens, and enabled a broad, flexible diet, thus explaining its success in unsanitary conditions and under recurrent chemical control. The findings and resources presented here provide insights for better understanding molecular mechanisms that will facilitate more effective cockroach control.

1 | INTRODUCTION

The German cockroach, *Blattella germanica*, is an obligatory commensal with humans and a quintessential generalist omnivore that also engages in detritivory, saprophagy, coprophagy, and cannibalism (Schal, C., Gautier, J.-Y., & Bell, 1984; Schal, 2011). *B. germanica* is a perennial pest in residential settings and other human-built structures, including restaurants, hospitals, schools, transportation networks, and even structures housing confined animals (e.g., poultry, pigs). The German cockroach is also a major public health pest, mainly because it mechanically vectors disease agents associated with unsanitary areas (Schal, 2011; Brenner, 1995; Ahmad, Ghosh, Schal, & Zurek, 2011) and as producer of potent allergens that cause asthma morbidity (Gore & Schal, 2007). To mitigate these effects, cockroach infestations are frequently targeted with a broad array of insecticides, which in turn have selected for multiple resistance mechanisms to all organic insecticides within as little as 3 years of their deployment (Schal & Hamilton, 1990). Resistance to DDT was documented in 1951, to organophosphates in 1964, carbamates in 1968, and to pyrethroids in 1989 (Cochran, 1995). Since then, resistance has been documented to hydramethylnon, fipronil, sulfluramid, and various neonicotinoids (Cochran, 1995). All these resistance mechanisms involve physiological mechanisms, including decreased penetration of the cuticle, increased sequestration and excretion, upregulation of insecticide metabolism, and mutations at insecticide target sites (Cochran, 1995).

Multiple resistance mechanisms to insecticides have been documented for a broad range of pest species, such as several disease carrying mosquito species (reviewed in Liu, 2015), or major crop pests like the peach potato aphid, *Myzus persicae*, (Bass et al., 2014) and the tobacco cutworm, *Spodoptera litura* (Cheng et al., 2017). An important mechanism that can lead to the development of resistance to multiple insecticides is the amplification of genes involved in metabolic detoxification (Bass et al., 2014; Cheng et al., 2017; Liu, 2015). For example, it was found that *S. litura* has adapted not only to diverse host plant chemistry and ecological conditions, but also to recurrent onslaughts with a diverse array of insecticides by dramatically expanding gene families encoding detoxification enzymes (e.g., cytochrome P450s, carboxylesterases, and glutathione *S*-transferases) and gustatory receptors for bitter or toxic tastants (Cheng et al., 2017).

The recently released genome of *B. germanica* revealed an enormous proteome comprising almost 30,000 protein coding genes, mainly as a result of 93 significant gene family expansions without any detected significant contractions (Harrison et al., 2018). It is conceivable, in a similar mode as reported for other pest species, that this large proteome allowed *B. germanica* to evolve such a broad range of resistance to toxins and pathogens. In fact, a de novo transcriptome of *B. germanica* revealed a large repertoire of transcripts important for the metabolism of toxins and defense against pathogens (Zhou et al., 2014). These include genes in both the Toll and Imd pathways, including serpins (47), cathepsins (32), lipopolysaccharide-binding proteins (21), transferrins (20), Gram negative bacteria binding proteins (4), Toll-like receptors (16), peptidoglycan recognition proteins (15), lysozymes (8), Imd proteins (3), and others including 115 lectin-like proteins and 18 β-glucanases (Zhou et al., 2014). Also, in the recently described genome of the American cockroach, *Periplaneta americana*, another major human pest species, gene family expansions could be related to insecticide resistance, most notably cytochrome P450 monooxygenases (Li et al., 2018).

In this study, we investigate the putative functions of the expanded gene families in the genome of *B. germanica*. We also analyze the expression of these gene families across 11 developmental stages (Ylla, Piulachs, & Belles, 2017), in order to establish a possible diversification of function among gene copies within expanded families. This study provides findings and resources for a better understanding of the molecular mechanisms involved in the global success of this major commensal indoor pest despite concerted efforts to eradicate it.

2 | RESULTS & DISCUSSION

2.1 | Expanded gene families

Ninety-three gene families are significantly expanded in the *B. germanica* genome with up to 21-fold increase in size compared to an estimated ancestral state (Supporting Information Table S1). The ancestral size of gene families was estimated based on a phylogeny containing 19 insect species (see methods of Harrison et al. 2018 and their Fig. S2 for further details). Interestingly, no gene families were contracted, leading to one of the largest so far reported insect proteomes (29,216 proteins; Harrison et al. 2018). The accuracy of this large proteome was confirmed by strong evidential support and many manually curated gene families.

The enrichment of gene ontology (GO) terms associated with the genes in the expanded families was carried out (Figure 1). The majority of the enriched GO terms within the Cellular Component category were related to cell structure and transport, such as "microtubule associated complex", "dynein complex", and "cytoskeleton" (Figure 1). The strong enrichment of "flavin adenine dinucleotide binding" (59 annotations, expected 14.3) can be attributed to the large expansion of glucose dehydrogenases (64 genes, ancestrally 29), since this GO term is mapped to the pfam domain "GMC_oxred_N" (glucose-methanol-choline oxidoreductase, one of the two functional domains of glucose dehydrogenase). Several enriched terms can be linked to a heightened ability of the German cockroach to detect chemical signals, such as "olfactory receptor activity", "odorant binding", "sensory perception of smell", and "G-protein coupled receptor signalling pathway", and are associated with the large expansion of ionotropic receptors, as previously reported (Harrison et al., 2018; Robertson et al., 2018). Many other functions relate to protein processing, redox reactions, and developmental processes (Figure 1).

The majority of the 93 gene families can be grouped into eight functional categories, with more than half contained in the five functional categories "detoxification", "defense against pathogens", "digestion", "sensory perception", and "gene regulation" (Table 1). In the following sections, we discuss how the gene family expansions related to these functions might have allowed the German cockroach to become such a successful pest of the human-built environment, despite extremely challenging conditions.

2.2 | Metabolism of insecticides

As a household pest with strong effects on human health, the German cockroach is systematically attacked with a wide range of insecticides. This species has, however, developed resistance to many types of insecticides (Cochran, 1995). One important pathway in insecticide resistance is metabolic degradation. Three important gene families in this respect are cytochrome P450 monooxygenases (P450s), carboxyl esterases, and glutathione *S*-transferases (GSTs) (Liu, 2015; Ranson et al., 2002). All three gene families are significantly expanded in *B. germanica*, possibly explaining its resistance to many insecticides.

P450s, which catalyze the oxidation of a broad range of insecticides (Silva et al., 2012), are highly abundant in the genome of *B. germanica* with 158 genes. Within our 21 species set of 20 insects and the centipede *Strigamia maritima*, only the genome of the yellow fever mosquito, *Aedes aegypti*, contained more P450s genes with 178. Both *B. germanica* and *A. aegypti* are pests of humans, experience strong selection from insecticides, and therefore have evolved high resistance to insecticides (Vontas et al., 2012). Interestingly, the next largest repertoires of P450s in our species set also belonged to pest species. These were the Florida carpenter ant, *Camponotus floridanus* (127); the red flour beetle, *Tribolium castaneum* (124); the migratory locust, *Locusta migratoria* (122); and the kissing bug, *Rhodnius prolixus* (119), an important vector of the Chagas parasite. In the American

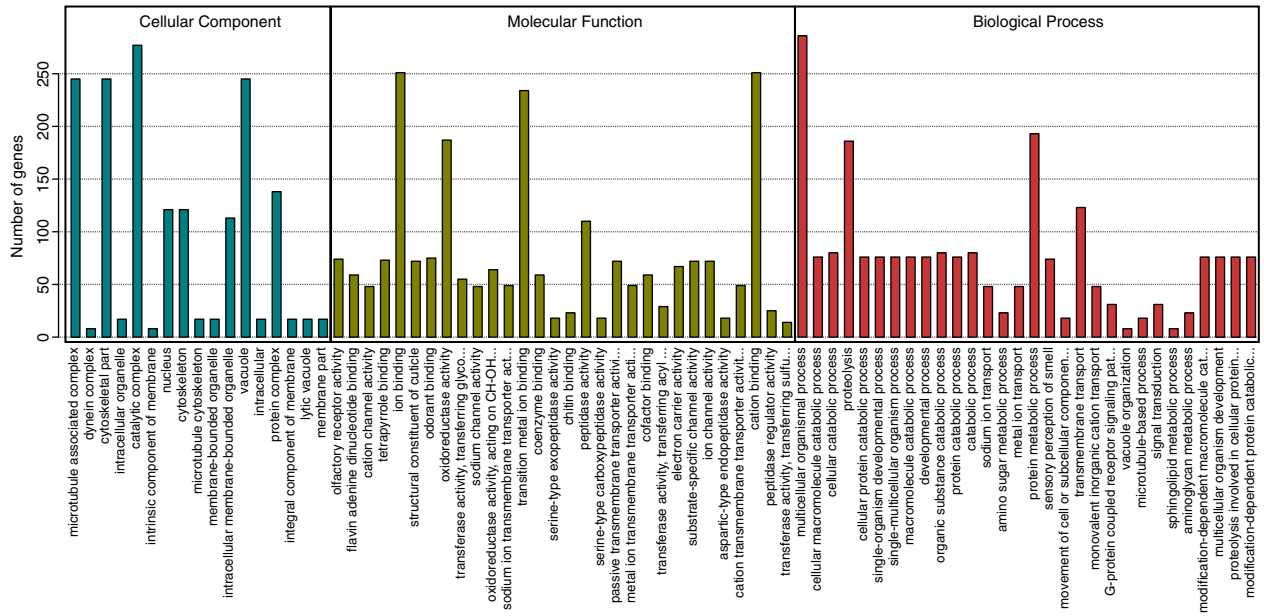

FIGURE 1 Enriched Gene Ontology (GO) terms within the expanded gene families. Only significantly enriched terms (*P*-value < 0.05) are represented, while for molecular function and biological process only top 30 terms, based on *P*-value, are shown [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Expanded gene families summarized by functional category

Function	Number of clusters	Number of genes	
		B. germanica	(ancestrally)
Detoxification of insecticides	7	252	(139)
Microbial defense and immune response	10	195	(51)
Digestion	9	274	(109)
Perception	13	398	(128)
Gene regulation	13	306	(126)
Cell structure	7	204	(95)
Protein processing	7	208	(69)
Development	5	120	(66)
Others	20	503	(201)

cockroach, P450s are also expanded at 178 gene copies (Li et al., 2018). An expansion of P450s can be expected to help these species metabolize a broader range of insecticides. It is important to note however, that the carpenter ant, a social insect, has not been demonstrated to express resistance to any insecticides, underscoring that P450s participate not only in detoxification of allelochemicals and xenobiotics, but also in many unrelated physiological processes.

In *B. germanica*, especially P450s from the subgroup CYP4 are expanded (59, ancestrally 30, Supporting Information Table S1), where we found evidence for transposable element (TE)-assisted tandem duplications (Janoušek, Karn, & Laukaitis, 2013). This is based on a significantly higher TE content within 10 kb flanking regions of these genes (mean 51.4%) compared to the whole genome level (46.3%; *P*-value = 0.001; Wilcoxon rank sum test; Figure 2). The majority of the TEs were long interspersed nuclear elements (LINEs) (mean 34.0% compared to 29.3% in whole genome; *P*-value < 0.001; Wilcoxon rank sum test). Further evidence for tandem duplications is offered by 20 out of the 59 *B. germanica* CYP4 genes sitting in close synteny (one 700 kb region) on the same scaffold (Figure 2, scaffold 637) and containing an even higher TE content in flanking regions (mean 54.7% TE content; 36.4% LINES). Most other CYP4 genes also contained higher than expected TE content in flanking regions, although their synteny could not be established due to their distribution across different scaffolds. A subgroup of CYP6 genes was also expanded (eight, ancestrally three). Both CYP4 and CYP6 are known to have important roles in detoxification and insecticide resistance, for example, in the peach-potato aphid, *Myzus persicae* (Silva et al., 2012), and in the bed bug, *Cimex lectularius* (Zhu et al., 2013). A TE assisted expansion of P450s may therefore have allowed *B. germanica* to successfully develop resistance against many different insecticides. We note as well that CYP4 and CYP6 participate in other physiological functions, including, among others, biosynthesis of cuticular hydrocarbons (Qiu et al., 2012) and clearance of odorants in the antennae (Keeling et al., 2013).

Carboxylesterases hydrolyze carboxylic esters and are therefore important enzymes for the metabolism of organophosphorus (OP) and pyrethroid insecticides (Hemingway & Karunaratne, 1998). Nevertheless, carboxylesterases are also important for general digestion that requires degrading ester bonds, such as cellulose metabolism. A higher activity of esterases could be observed in the presence of insecticides for resistant compared to susceptible strains of *B. germanica* (Prabhakaran & Kamble, 1993). We have found a significant expansion of E4 esterases (62, ancestral state: 41) in the genome of *B. germanica*, possibly allowing multiple resistance to OPs and pyrethroids. An amplification of E4 and closely related FE4 enzymes in the peach potato aphid, *M. persicae*, where up to 80 gene copies exist, has led to resistance against OPs via an overexpression of these enzymes (Bass et al., 2014; Silva et al., 2012).

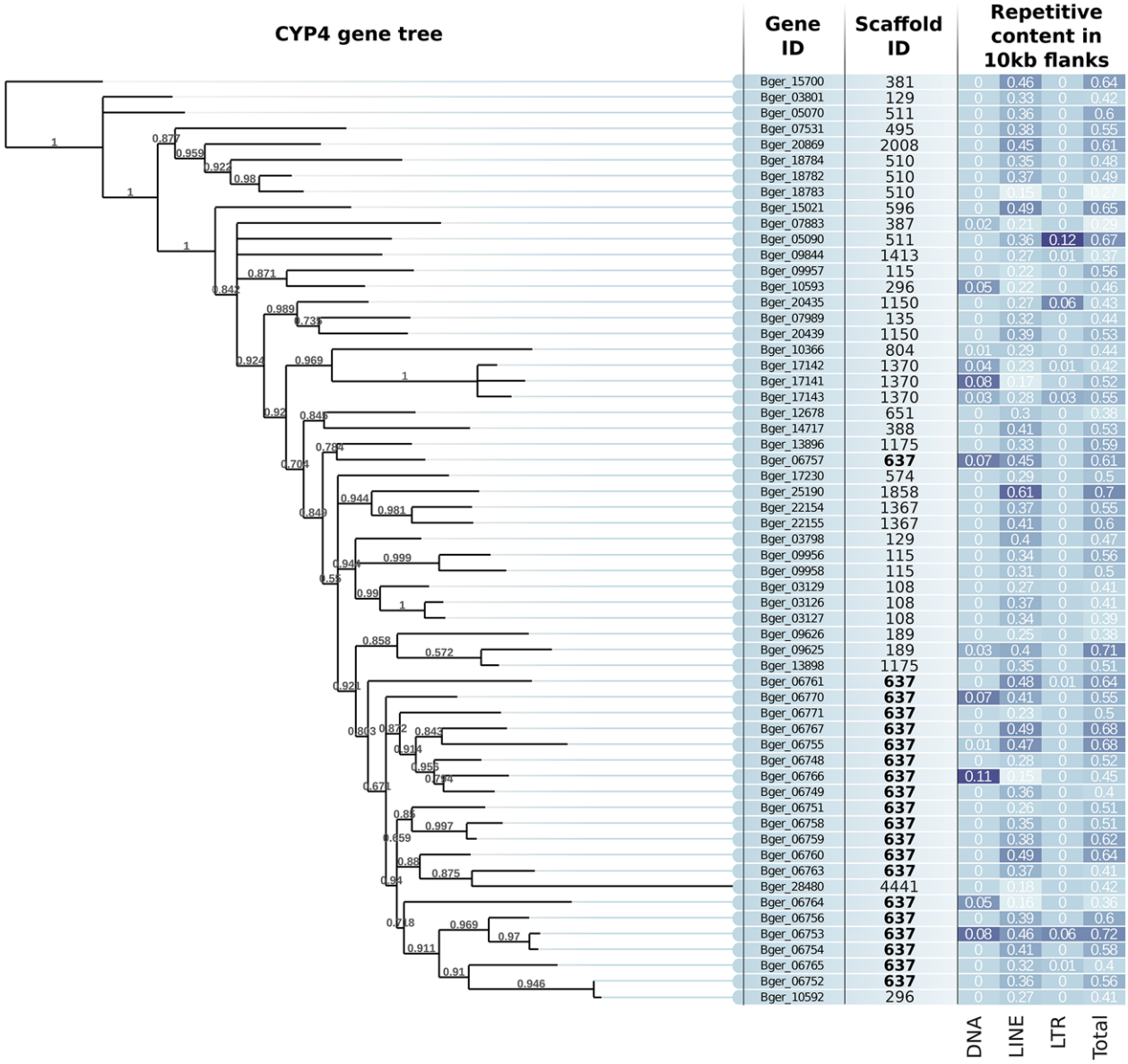

FIGURE 2 Transposable element content in CYP4 genes. On the left is a gene tree of the 59 CYP4 genes. The central column contains the scaffold IDs on which the genes are situated. The heatmap shows three categories of repetitive content (DNA transposons, LINEs and LTR retrotransposons) in flanking regions (10 kb either side of genes). The colour shading is based on *Z*-scores calculated per category across all genes, with darker colours indicating an enrichment [Color figure can be viewed at wileyonlinelibrary.com]

GSTs are implicated in resistance to several insecticides including DDT and OPs by converting toxic metabolites into water soluble conjugates that can be more readily excreted (Enayati, Ranson, & Hemingway, 2005; Silva et al., 2012). A subgroup of the GSTs was significantly expanded in *B. germanica* (23, ancestrally 10), which is also expanded in the malaria mosquito, *Anopheles gambiae* (Ranson et al., 2002). GSTs also offer additional protection against the effects of insecticides by reducing oxidative stress (Silva et al., 2012; Vontas, Graham, & Hemingway, 2001). It is also important to note that GSTs from the German cockroach (Bla g 5) and the American cockroach as well as from other arthropods and nematodes, are clinically relevant human allergens (Arruda et al., 1997; Sookrung et al., 2018).

UDP-glucuronosyltransferases (UGT) are also involved in the conjugation of insecticides and show elevated expression in insecticide resistant mosquitoes when exposed to permethrin (Vontas et al., 2012). We annotated 61 UGTs in the *B. germanica* genome, which is more than twice the number estimated for the ancestral state (30). The expansion of UGTs appears to be a hallmark of not only an evolutionary response to xenobiotics, including insecticides, but also to polyphagy and omnivory, as demonstrated in a locust (Wang et al., 2014).

Gene family expansions and a corresponding increase in protein abundance is considered one of the main mechanisms that allow the development of insecticide resistance (Liu, 2015). The expansions of these four gene families, and others such as the ABC transporters (see below), are therefore most likely responsible for this pest's ability to develop resistance to many different insecticides, including organochlorines, OPs, carbamates, pyrethroids, and newer classes like neonicotinoids and phenylpyrazoles. Interestingly, alternative

mechanisms that allow the development of insecticide resistance exist, such as mutations within the functional regions or changes in expression of insecticide resistance genes. This is the case for the bed bug, *Cimex lectularius*, which is resistant to many insecticides but for which these gene families are not expanded (Benoit et al., 2016; Rosenfeld et al., 2016).

2.3 | Defense against pathogens

The German cockroach thrives in unsanitary conditions, leading to its status as a major public health pest and vector of disease agents (Brenner, 1995; Ahmad et al., 2011). We found a number of expansions in gene families with functions related to microbial defense and immune response, which may have allowed *B. germanica* to adapt to septic conditions.

We discovered an expansion within a subgroup of ATP-binding cassette (ABC) transporters (39; ancestrally 25; total number of ABC transporters: 81) in the *B. germanica* genome. ABC-transporters are membrane proteins, responsible for the efflux of molecules such as xenobiotics and lipids from eukaryotic cells (and import in prokaryotes) (Rees, Johnson, & Lewinson, 2009). Their expansion may allow *B. germanica* to efficiently remove microbial toxins. These genes show varied expression profiles throughout development, where different gene copies may offer protection at different developmental stages (Figure 3a). Interestingly, at least three gene copies appear to be maternally provided within the unfertilized egg. This offers support for a diversification of specificity and possibly function among the ABC transporter gene copies. ABC-transporters have also been implicated in resistance to insecticides (Gahan, Pauchet, Vogel, & Heckel, 2010; Silva et al., 2012; Dermauw & Van Leeuwen, 2014) and, accordingly, are also expanded in the American cockroach (Li et al., 2018), but they may have much more diverse functions (Broehan, Kroeger, Lorenzen, & Merzendorfer, 2013).

A massively expanded family of defense proteins are the hemolymph lipopolysaccharide (LPS)-binding proteins (86; ancestrally 14), which have been found to be important for bacteria ingestion by hemocytes in the American cockroach, *P. americana* (Jomori & Natori, 1992). Many of these proteins are highly expressed within nymphal stages and much fewer during embryonic development (Figure 3b). A high number of these LPS-binding proteins are expressed at an above average level in adults (z-value $>$ 0; Fig. 3b). These results support the notion that LPS-binding proteins can be important for protecting against bacterial infection, the risk of which increases in nymphal stages and especially adulthood when individuals are less protected.

Two expanded families can be linked specifically to defense against fungal infection. There are 10 drosomycin copies in the *B. germanica* genome compared to an estimated single ancestral copy. Drosomycin is an antifungal peptide, which in *Drosophila melanogaster* is secreted into the hemolymph as a reaction to infection (Ferrandon et al., 1998; Lemaitre, Reichhart, & Hoffmann, 1997). We found eight glucosylceramidases (ancestrally two), which may be important for metabolizing glucosylceramides found in the membrane of pathogenic fungi, such as *Cryptococcus neoformans* and *Cryptococcus gattii* (Watanabe et al., 2015).

Further expansions can be related to more general immune responses, such as DSCAM (9) and serpins (24), among others (Supporting Information Table S1). The expanded family of tenascin-X genes (16 genes) may allow improved wound healing (Egging et al., 2007), while further families such as MOXD1-like proteins (8) (Dowling et al., 2012; de Boer et al., 2013) and catalases (16) (Finkel & Holbrook, 2000) are important for repairing or preventing cell damage caused by oxidative stress.

2.4 | Digestive enzymes

As with the defense proteins, the expansions of nine digestive protein families may have supported the success of this human-commensal pest. The German cockroach is an extreme omnivore with a diverse and adaptive diet (Jensen, Schal, & Silverman, 2015), for which a broad repertoire of digestive enzymes would be required. Accordingly, we find a large set of α-glucosidase genes (18), which allow the breakdown of starch and disaccharides to glucose (Sørensen, Norén, Sjöström, & Danielsen, 1982), while 20 β-glucosidases allow these cockroaches to digest large amounts of plant material (Sticklen, 2008). The members of both families of glucosidases show varied and partially complementary expression especially during nymphal and adult stages (Figures 3c and d). The higher expression of α- compared to β- glucosidase in these results is likely related to the lack of plant material in the diet of the lab-raised animals. However, the variation in expression profiles between gene copies suggests a diversification in substrate specificity of these enzymes. One gene in this α-glucosidase gene family (gene id: Bger_16909) was ubiquitously highly expressed throughout all developmental stages. This gene is likely a heteromeric amino acid transporter (HAT), which is known to have homology to α-glucosidase, since beside an α-amylase domain it also contains an SLC3A2_N domain (4F2 cell-surface antigen heavy chain) (Palacín & Kanai, 2004).

Glucose dehydrogenases, which metabolize glucose to provide an organism with energy (Neijssel, Hommes, Postma, & Tempest, 1989), are also expanded in *B. germanica* (64). Glucose-dependent insulinotropic receptors, of which 51 copies exist in the genome of *B. germanica* compared to nine in the ancestral state, also play an important role in the metabolism of glucose by controlling the release of insulin.

The two further typical energy sources, fats and proteins, can be metabolized by large numbers of lipases and trypsins, respectively (Supporting Information Table S1). Both expanded groups of enzymes are predominantly over-expressed during nymphal and adult stages when individuals were feeding on a fat- and protein-rich diet (Figures 3e and f). However, expression profiles differ considerably between gene copies indicating a variation in substrate specificity. Interestingly, transcripts of two lipase gene copies (Bger_00361 and Bger_17567) were highly abundant only within the nonfertilized egg (Figure 3 e), suggesting they were maternally deposited and may play an important role in releasing energy from lipids for the first cell divisions in early embryonic development (Ziegler & Van Antwerpen, 2006). Another lipase gene copy (Bger_25734) was overexpressed only at day 6 of embryonic development suggesting a specific function in lipid metabolism for later embryo development. This occurs prior to

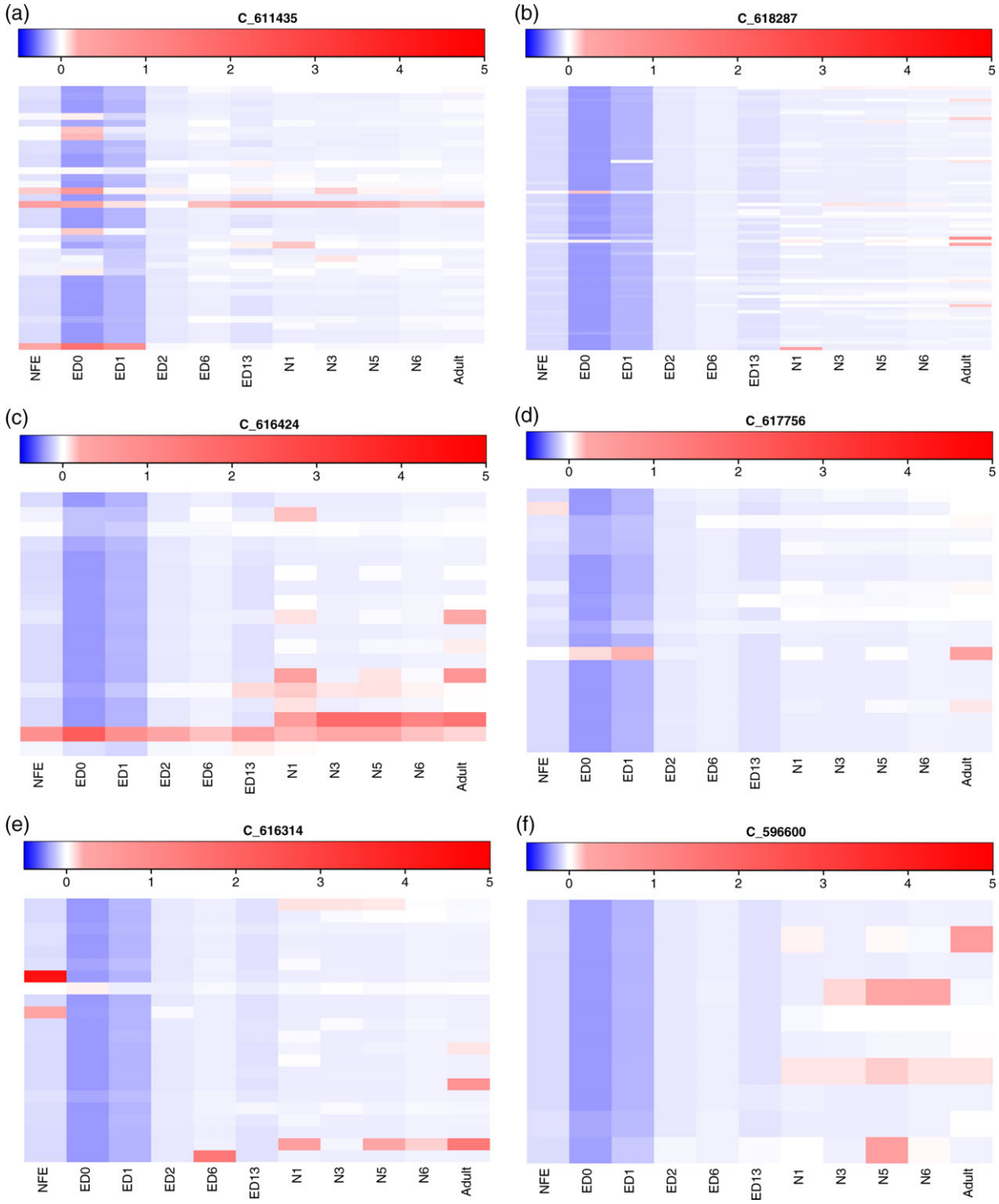

FIGURE 3 Expression of expanded families of defence and digestive proteins. Each row in the heatmap represents an individual gene within the gene family. The colour shows the *Z*-score of expression at each developmental stage. The *Z*-scores are calculated within each developmental stage among all genes, and the values represent the number of standard deviations from the standardized mean of 0. NFE, nonfertilised egg; ED0,1,2,6, embryos in days after fertilisation; N1,3,5,6, nymphal stages. (a) ABC transporters; (b) hemolymph lipopolysaccharide-binding proteins; (c) α-glucosidases; (d) β-glucosidases; (e) lipase 3; and (f) trypsins [Color figure can be viewed at wileyonlinelibrary.com]

a rupturing of the amnion and serosa, after which the embryo has direct contact and access to the yolk, which contains lipid reserves (Tanaka, 1976). The dramatic morphogenetic processes that follow require extensive lipid metabolism, which could explain the burst of Bger_25734.

With these gene family expansions, the German cockroach is able to digest and metabolize a wide range of carbohydrate, fat, and protein sources throughout its development. This explains the ability of this species to adapt its diet to ever changing nutritional conditions, both in quantity and quality.

2.5 | Sensory perception

Thirteen gene families with functions related to sensory perception were significantly expanded in *B. germanica*. These include gene families related to chemosensation of odorants and tastants such as ionotropic, gustatory, and odorant receptors (IRs, GRs, & ORs, respectively), of which the IRs are known to be particularly expanded (Harrison et al., 2018), as well as odorant binding proteins (Robertson et al., 2018) (Supporting Information Table S1).

Further expanded gene families related to perception include transient receptor potential cation channel A1 (TRPA1, 22, ancestrally 11), which has reported functions in mechanosensation (Nilius, Owsianik, Voets, & Peters, 2007), olfaction, vision, hearing, and thermosensation (in *Helicoverpa armigera* a TRPA1 channel senses thermal stimulus and irritating chemicals) (Wei et al., 2015). Cyclic nucleotide gated cation channels (22, ancestrally 12) are important for photo- or olfactory sensation (Kaupp & Seifert, 2002) and amiloride-sensitive sodium channel proteins (48, ancestrally 18) may be important for detecting sodium and potassium salts (Liu et al., 2003).

This high abundance of genes related to perception may allow this generalist to detect not only many types of food sources but also to detect and avoid dangerous toxins and pathogens, which abound in their habitats. Avoidance may therefore be an additional survival strategy to the defence mechanisms described above.

2.6 | Gene regulation

Thirteen gene families with putative functions in gene regulation were significantly expanded from a total of 126 ancestrally to 306 genes in the genome of *B. germanica* (Table 1). Ten of these expanded families comprise genes containing zinc finger (ZF) domains (expanded from 118 to 258 in total; Supporting Information Table S1). The majority of these genes (193) contain between one and 19 ZF domains of the type C2H2. C2H2 ZF domains generally have a DNA binding function and are often present in transcription factors (Tadepally, Burger, & Aubry, 2008; Schmitz, Zimmer, & Bornberg-Bauer, 2016). The amplification of these ZF genes is not particularly remarkable, since such expansions are common among metazoa, often facilitated by the repetitive nature of the domains (Schmitz et al., 2016). However, this increase in ZF transcription factors most likely supported the regulation of the many expanded gene families described above, thus allowing a diversification of function, expression level, or specificity of the expanded genes, for example, to a broader range of substrates, such as toxins or food sources. Lineage-specific expansions in ZF gene families have also been reported for other pest species, such as the Colorado potato beetle (Schoville et al., 2018), the large milkweed bug, and the pea aphid (Panfilio et al., 2017).

The role of myb/SANT-like DNA-binding domain-containing protein 3-like, which was strongly expanded from 6 to 39 in the *B. germanica* genome, is not conclusively resolved. However, it also appears to regulate gene expression and protein synthesis in mammals (Barasch et al., 2017).

3 | SUMMARY

Many of the 93 expanded gene families within the genome of the German cockroach, *B. germanica*, may help explain how this indoor pest manages to thrive in unsanitary, inhospitable conditions. One mechanism contributing to the expansion of gene families may be TE assisted tandem duplication, for which evidence is presented in this study for the important detoxification genes CYP6. Seven expanded gene families comprising a total of 252 genes (ancestrally 139) allow this cockroach to metabolize a broad range of toxins, thus explaining its resistance to many insecticides, in which these toxins are employed. A further 10 expanded gene families (195 genes, ancestrally 51) function in microbial defense and immune response, which illustrates how this species thrives in septic conditions, often carrying many pathogens, which threaten human health. Another major factor leading to the success of this species, are its very generalist, omnivorous feeding capabilities, that can in part be explained by the large expansion of nine gene families related to the digestion of carbohydrates, proteins, and fats (274 genes, ancestrally 109). A massive expansion in gene families related to sensory perception (13 gene families, 398 genes, ancestrally 128), especially gustatory and odorant perception (in particular ionotropic receptors; Harrison et al. 2018; Robertson et al. 2018), allow the detection and differentiation of a very broad range of chemical signals, thus possibly enabling the German cockroach to accurately distinguish between toxins and food sources. These gene family expansions were accompanied by a large expansion of C2H2 zinc finger transcription factors within the genome of *B. germanica*, allowing an accurate, flexible regulation of these important gene families. We report that members of several of these expanded gene families show differential and complementary expression throughout development. It is conceivable that a greater range of expression differences may become apparent when individuals are challenged with different toxin or dietary conditions. The resources presented here may help researchers to better understand the mechanisms involved in the resistance of *B. germanica* to insecticides and pathogens, thus allowing the development of more specific and efficient strategies for their control. Moreover, future comparisons with its sister species, the Asian cockroach *Blattella asahinai*, may provide insight into the evolution of synanthropy and specialized adaptations that make the German cockroach such a successful commensal in human structures.

4 | MATERIALS AND METHODS

4.1 | Expanded gene families

The proteomes of 19 insect species (*Nasonia vitripennis*, *Polistes canadensis*, *Apis mellifera*, *Acromyrmex echinatior*, *Atta cephalotes*, *Solenopsis invicta*, *Pogonomyrmex barbatus*, *Camponotus floridanus*, *Linepithema humile*, *Harpegnathos saltator*, *Tribolium castaneum*, *Aedes aegypti*, *Drosophila melanogaster*, *Rhodnius prolixus*, *Macrotermes natalensis*, *Cryptotermes secundus*, *Zootermopsis nevadensis*, *Blattella germanica*, and *Locusta migratoria*; sources: Table 2) were clustered using the hierarchical clustering algorithm MC-UPGMA (Loewenstein, Portugaly, Fromer, & Linial, 2008). Significant expansions and

TABLE 2 Source websites and version numbers of 20 arthropod proteomes that were used for estimating gene family expansions

Species	Source website	Data files
P. barbatus	http://hymenopteragenome.org	genome 1.0, gff 1.2 (fixed)
D. melanogaster	ftp://ftp.flybase.net/releases/FB2016_04/	CDS 6.12
H. saltator	http://hymenopteragenome.org	CDS 3.3
R. prolixus	ftp://ftp.ensemblgenomes.org/pub/metazoa/release-32/	genome RproC1, gff 1.32
T. castaneum	https://www.ncbi.nlm.nih.gov/genome/216	genome 5.2, gff 5
A. mellifera	ftp://ftp.ensemblgenomes.org/pub/metazoa/release-32/	genome 4.5, gff 4.5
L. migratoria	http://159.226.67.243/download.htm	genome 2.4.1, gff 2.4.1
N. vitripennis	http://arthropods.eugenes.org/EvidentialGene/nasonia/genes/	CDS 2.1
P. canadensis	https://www.ncbi.nlm.nih.gov/genome/?term=txid91411	GCF_001313835.1 ASM131383v1
A. cephalotes	http://hymenopteragenome.org	genome 1.0, gff 1.2
C. floridanus	http://hymenopteragenome.org	genome 3.3, gff 3.3
L. humile	http://hymenopteragenome.org	genome 1.0, gff 1.2
S. invicta	http://hymenopteragenome.org	genome 1.0, gff 2.2.3
A. echinatior	http://hymenopteragenome.org	CDS 3.8
A. aegypti	ftp://ftp.ensemblgenomes.org/pub/metazoa/release-32/	CDS 3
M. natalensis	http://gigadb.org/dataset/100057	genome 1, gff 1.2
Z. nevadensis	http://termitegenome.org	genome 1.0, gff 2.2
C. secundus	Harrison et al. (2018)	genome 1.0, gff 1.0
B. germanica	Harrison et al. (2018)	genome 1.1, gff 1.1

contractions ($P < 0.05$) were identified with CAFE v3.0 (Han, Thomas, Lugo-Martinez, & Hahn, 2013). For further details see Harrison et al. (2018).

4.2 | Identification of function

The putative function of genes within the expanded gene families was estimated based on functional domains and sequence similarity to known proteins in public databases. The proteome of *B. germanica* was annotated with functional domains using PfamScan v31 (Finn et al., 2016). All protein sequences were blasted against the nonredundant NCBI database (last accessed: 29-07-2014) using blastp. GO terms were obtained with Pfam2GO and mapped to the genes of expanded families. The enrichment of GO-terms associated with genes of expanded families compared to all genes in the *B. germanica* proteome was analyzed using the R package TopGO (Alexa & Rahnenfuhrer, 2010), using the parent–child Fisher test.

4.3 | Analysis of expression profiles

Expression patterns were based on the 11 developmental stages (nonfertilized egg, five embryonic stages, four nymph stages, and 5-day-old adult females) described by Ylla et al. (2017). For each developmental stage, there were two biological replicates, with each replicate containing a pool of individuals. All insects were fed ad libitum with dog food and water. The expression count data were generated for the corresponding transcriptomes of these 11 developmental stages, which were prepared by Ylla, Piulachs, & Belles (2018). These transcriptomes are publicly available at Gene Expression Omnibus under the accession code GSE99785.

The raw counts were normalized using DESeq2 (Love, Huber, & Anders, 2014). *Z*-scores were calculated within each developmental stage across all genes in expanded families using the scale function and plotted in R (version 3.4.2) (R Core Team, 2017).

4.4 | Repetitive content

Repeat content had previously been annotated for the *B. germanica* genome (Harrison et al., 2018). In this study, we assessed the repeat content of the 10 kb flanking regions of genes as follows: (i) The flanking regions of each gene were extracted; (ii) any CDS from neighbouring genes was removed such that only intergenic sequences were considered as flanking sequences; (iii) the proportion of flanking sequences covered by repeats was calculated, differentiating between low complexity repeats, simple repeats, and the three major classes of interspersed repeats: LINEs, LTRs, and DNA-transposons; and (iv) the repeat content was *Z*-score transformed, such that heat maps reflect deviation in repeat content from species averages. Steps (i)–(iii) were carried out with the bedtools (Quinlan & Hall, 2010) commands *flank*, *subtract*, and *coverage*, respectively. The alignments for these analyses were carried out with pasta (Mirarab et al., 2015) and phylogenetic trees constructed with fasttree v2.1.7 (Price, Dehal, & Arkin, 2010).

ACKNOWLEDGMENTS

We thank the Deutsche Forschungsgemeinschaft for providing funding for this research to M.C.H. via grant BO2544/11-1 to E.B.B., and acknowledge United States National Science Foundation grant IOS1557864 to C.S. X.B. & M.-D.P. were supported by the Spanish Ministry of Economy and Competitiveness (grants CGL2015-64727-P to

X.B. and CGL2016-76011-R to M.-D.P., both with FEDER funds) and Catalan Government (grant 2017 SGR 1030).

AUTHOR CONTRIBUTIONS

M.C.H. carried out analyses and wrote the manuscript. N.A. annotated TEs and performed TE related analyses. L.P.M.K. identified and annotated expanded gene families. G.Y., X.B., & M.-D.P. developed and provided transcriptome data. E.B.B. managed and coordinated the genomic analysis project. A.-K.H. annotated P450s. E.J. annotated the *B. germanica* genome and assisted with TE annotations. S.R. conceived and managed the *B. germanica* sequencing project. C.S. provided biological material for the sequencing project and assisted in interpretation of results and writing the manuscript. All authors have read, corrected, and commented on the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Mark C. Harrison http://orcid.org/0000-0003-3095-019X
Xavier Belles http://orcid.org/0000-0002-1566-303X
Erich Bornberg-Bauer http://orcid.org/0000-0002-1826-3576
Evelien Jongepier http://orcid.org/0000-0003-0206-5396
Maria-Dolors Piulachs http://orcid.org/0000-0002-4500-2718

REFERENCES

Ahmad, A., Ghosh, A., Schal, C., & Zurek, L. (2011). Insects in confined swine operations carry a large antibiotic resistant and potentially virulent enterococcal community. *BMC Microbiology*, *11*(1), 23.

Alexa, A., & Rahnenfuhrer, J. (2010). Topgo: Enrichment analysis for gene ontology. *R package version*, *2*(0).

Arruda, L. K., Vailes, L. D., Platts-Mills, T. A., Hayden, M. L., & Chapman, M. D. (1997). Induction of ige antibody responses by glutathiones-transferase from the german cockroach (*Blattella germanica*). *Journal of Biological Chemistry*, *272*(33), 20907–20912.

Barasch, N., Gong, X., Kwei, K. A., Varma, S., Biscocho, J., Qu, K., … Pollack, J. R. (2017). Recurrent rearrangements of the myb/sant-like dna-binding domain containing 3 gene (msantd3) in salivary gland acinic cell carcinoma. *PloS One*, *12*(2), e0171265.

Bass, C., Puinean, A. M., Zimmer, C. T., Denholm, I., Field, L. M., Foster, S. P., … Williamson, M. S. (2014). The evolution of insecticide resistance in the peach potato aphid, *Myzus persicae*. *Insect Biochemistry and Molecular Biology*, *51*, 41–51.

Benoit, J. B., Adelman, Z. N., Reinhardt, K., Dolan, A., Poelchau, M., Jennings, E. C., … Richards, S. (2016). Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nature communications*, *7*, ncomms10165.

Brenner, R. (1995). *Economics and medical importance of german cockroaches. Understanding and controlling the German cockroach* (pp. 77–92). New York, NY: Oxford University Press.

Broehan, G., Kroeger, T., Lorenzen, M., & Merzendorfer, H. (2013). Functional analysis of the ATP-binding cassette (abc) transporter gene family of *Tribolium castaneum*. *BMC Genomics*, *14*(1), 6.

Cheng, T., Wu, J., Wu, Y., Chilukuri, R. V., Huang, L., Yamamoto, K., … Mita, K. (2017). Genomic adaptation to polyphagy and insecticides in a major east asian noctuid pest. *Nature Ecology & Evolution*, *1*(11), 1747.

Cochran, D. G. (1995). Insecticide resistance. In M. K. Rust, J. M. Owens, D. R., editor, *Understanding and controlling the German cockroach*, pp. 179–192. Oxford, England: Oxford University Press.

de Boer, M. E., Ellers, J., van Gestel, C. A., den Dunnen, J. T., van Straalen, N. M., & Roelofs, D. (2013). Transcriptional responses indicate attenuated oxidative stress in the springtail *Folsomia candida* exposed to mixtures of cadmium and phenanthrene. *Ecotoxicology*, *22*(4), 619–631.

Dermauw, W. & van Leeuwen, T. (2014). The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance. *Insect Biochemistry and Molecular Biology*, *45*, 89–110.

Dowling, J. J., Arbogast, S., Hur, J., Nelson, D. D., McEvoy, A., Waugh, T., … Ferreiro, A. (2012). Oxidative stress and successful antioxidant treatment in models of ryr1-related myopathy. *Brain*, *135*(4), 1115–1127.

Egging, D., van Vlijmen-Willems, I., van Tongeren, T., Schalkwijk, J., & Peeters, A. (2007). Wound healing in tenascin-x deficient mice suggests that tenascin-x is involved in matrix maturation rather than matrix deposition. *Connective Tissue Research*, *48*(2), 93–98.

Enayati, A. A., Ranson, H., & Hemingway, J. (2005). Insect glutathione transferases and insecticide resistance. *Insect Molecular Biology*, *14*(1), 3–8.

Ferrandon, D., Jung, A., Criqui, M.-C., Lemaitre, B., Uttenweiler-Joseph, S., Michaut, L., … Hoffmann, J. (1998). A drosomycin–gfp reporter transgene reveals a local immune response in *Drosophila* that is not dependent on the toll pathway. *The EMBO Journal*, *17*(5), 1217–1227.

Finkel, T. & Holbrook, N. J. (2000). Oxidants, oxidative stress and the biology of ageing. *Nature*, *408*(6809), 239–247.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., … Bateman, A. (2016). The pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*, *44*(D1), D279–D285.

Gahan, L. J., Pauchet, Y., Vogel, H., & Heckel, D. G. (2010). An abc transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* cry1ac toxin. *PLoS Genetics*, *6*(12), e1001248.

Gore, J. C. & Schal, C. (2007). Cockroach allergen biology and mitigation in the indoor environment. *Annual Review of Entomology*, *52*, 439–463.

Han, M. V., Thomas, G. W., Lugo-Martinez, J., & Hahn, M. W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using cafe 3. *Molecular Biology and Evolution*, *30*(8), 1987–1997.

Harrison, M. C., Jongepier, E., Robertson, H. M., Arning, N., Bitard-Feildel, T., Chao, H., … Bornberg-Bauer, E. (2018). Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nature Ecology & Evolution*, *2*(3), 557.

Hemingway, J. & Karunaratne, S. (1998). Mosquito carboxylesterases: A review of the molecular biology and biochemistry of a major insecticide resistance mechanism. *Medical and Veterinary Entomology*, *12*(1), 1–12.

Janoušek, V., Karn, R. C., & Laukaitis, C. M. (2013). The role of retrotransposons in gene family expansions: insights from the mouse abp gene family. *BMC Evolutionary Biology*, *13*(1), 107.

Jensen, K., Schal, C., & Silverman, J. (2015). Adaptive contraction of diet breadth affects sexual maturation and specific nutrient consumption in an extreme generalist omnivore. *Journal of Evolutionary Biology*, *28*(4), 906–916.

Jomori, T. & Natori, S. (1992). Function of the lipopolysaccharide-binding protein of *Periplaneta americana* as an opsonin. *FEBS Letters*, *296*(3), 283–286.

Kaupp, U. B. & Seifert, R. (2002). Cyclic nucleotide-gated ion channels. *Physiological Reviews*, *82*(3), 769–824.

Keeling, C. I., Henderson, H., Li, M., Dullat, H. K., Ohnishi, T., & Bohlmann, J. (2013). Cyp345e2, an antenna-specific cytochrome p450 from the mountain pine beetle, *Dendroctonus ponderosae* hopkins, catalyses the

oxidation of pine host monoterpene volatiles. *Insect Biochemistry and Molecular Biology*, *43*(12), 1142–1151.

Lemaitre, B., Reichhart, J.-M., & Hoffmann, J. A. (1997). *Drosophila* host defense: Differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(26), 14614–14619.

Li, S., Zhu, S., Jia, Q., Yuan, D., Ren, C., Li, K., … Zhan, S. (2018). The genomic and functional landscapes of developmental plasticity in the American cockroach. *Nature Communications*, *9*(1), 1008.

Liu, L., Leonard, A. S., Motto, D. G., Feller, M. A., Price, M. P., Johnson, W. A., & Welsh, M. J. (2003). Contribution of *Drosophila* deg/enac genes to salt taste. *Neuron*, *39*(1), 133–146.

Liu, N. (2015). Insecticide resistance in mosquitoes: impact, mechanisms, and research directions. *Annual Review of Entomology*, *60*, 537–559.

Loewenstein, Y., Portugaly, E., Fromer, M., and Linial, M. (2008). Efficient algorithms for accurate hierarchical clustering of huge datasets: Tackling the entire protein space. *Bioinformatics*, *24*(13), i41–i49.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biology*, *15*(12), 550.

Mirarab, S., Nguyen, N., Guo, S., Wang, L.-S., Kim, J., & Warnow, T. (2015). Pasta: ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. *Journal of Computational Biology*, *22*(5), 377–386.

Neijssel, O. M., Hommes, R. W., Postma, P. W., & Tempest, D. W. (1989). Physiological significance and bioenergetic aspects of glucose dehydrogenase. *Antonie van Leeuwenhoek*, *56*(1), 51–61.

Nilius, B., Owsianik, G., Voets, T., & Peters, J. A. (2007). Transient receptor potential cation channels in disease. *Physiological Reviews*, *87*(1), 165–217.

Palacín, M. & Kanai, Y. (2004). The ancillary proteins of HATs: SLC3 family of amino acid transporters. *Pflügers Archiv*, *447*(5), 490–494.

Panfilio, K. A., Jentzsch, I. M. V., Benoit, J. B., Erezyilmaz, D., Suzuki, Y., Colella, S., … Richards, S. (2017). Molecular evolutionary trends and feeding ecology diversification in the hemiptera, anchored by the milkweed bug genome. *bioRxiv*, page 201731.

Prabhakaran, S. K. & Kamble, S. T. (1993). Activity and electrophoretic characterization of esterases in insecticide-resistant and susceptible strains of german cockroach (dictyoptera: Blattellidae). *Journal of Economic Entomology*, *86*(4), 1009–1013.

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). Fasttree 2–approximately maximum-likelihood trees for large alignments. *PloS One*, *5*(3), e9490.

Qiu, Y., Tittiger, C., Wicker-Thomas, C., Le Goff, G., Young, S., Wajnberg, E., … Feyereisen, R. (2012). An insect-specific p450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(37), 14858–14863.

Quinlan, A. R. & Hall, I. M. (2010). Bedtools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M. V., … Feyereisen, R. (2002). Evolution of supergene families associated with insecticide resistance. *Science*, *298*(5591), 179–181.

Rees, D. C., Johnson, E., & Lewinson, O. (2009). Abc transporters: The power to change. *Nature Reviews Molecular Cell Biology*, *10*(3), 218–227.

Robertson, H. M., Baits, R. L., Walden, K. K., Wada-Katsumata, A., & Schal, C. (2018). Enormous expansion of the chemosensory gene repertoire in the omnivorous german cockroach *Blattella germanica*. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*. https://doi.org/10.1002/jez.b.22797.

Rosenfeld, J. A., Reeves, D., Brugler, M. R., Narechania, A., Simon, S., Durrett, R., … Mason, C. E. (2016). Genome assembly and geospatial phylogenomics of the bed bug cimex lectularius. *Nature communications*, *7*, ncomms10164.

Schal, C. (2011). Cockroaches. In Hedges, S. & Moreland, D., editors, *Handbook of Pest Control*, chapter 2, pages 150–291. Pennsylvania State University: GIE Media.

Schal, C., Gautier, J.-Y., & Bell, W. J. (1984). Behavioural ecology of cockroaches. *Biological Reviews*, *59*(2), 209–254.

Schal, C. & Hamilton, R. (1990). Integrated suppression of synanthropic cockroaches. *Annual Review of Entomology*, *35*(1), 521–551.

Schmitz, J. F., Zimmer, F., & Bornberg-Bauer, E. (2016). Mechanisms of transcription factor evolution in metazoa. *Nucleic Acids Research*, *44*(13), 6287–6297.

Schoville, S. D., Chen, Y. H., Andersson, M. N., Benoit, J. B., Bhandari, A., Bowsher, J. H., … Richards, S. (2018). A model species for agricultural pest genomics: The genome of the colorado potato beetle, *Leptinotarsa decemlineata* (coleoptera: Chrysomelidae). *Scientific reports*, *8*(1), 1931.

Silva, A. X., Jander, G., Samaniego, H., Ramsey, J. S., & Figueroa, C. C. (2012). Insecticide resistance mechanisms in the green peach aphid *Myzus persicae* (hemiptera: Aphididae) i: A transcriptomic survey. *PloS One*, *7*(6), e36366.

Sookrung, N., Reamtong, O., Poolphol, R., Indrawattana, N., Seesuay, W., Saelim, N., … Tungtrongchitr, A. (2018). Glutathione s-transferase (gst) of American cockroach, *Periplaneta americana*: Classes, isoforms, and allergenicity. *Scientific Reports*, *8*(1), 484.

Sørensen, S. H., Norén, O., Sjöström, H., & Danielsen, E. M. (1982). Amphiphilic pig intestinal microvillus maltase/glucoamylase. *The FEBS Journal*, *126*(3), 559–568.

Sticklen, M. B. (2008). Plant genetic engineering for biofuel production: Towards affordable cellulosic ethanol. *Nature Reviews Genetics*, *9*(6), 433–443.

Tadepally, H. D., Burger, G., & Aubry, M. (2008). Evolution of c2h2-zinc finger genes and subfamilies in mammals: Species-specific duplication and loss of clusters, genes and effector domains. *BMC Evolutionary Biology*, *8*(1), 176.

Tanaka, A. (1976). Stages in the embryonic development of the German cockroach, *Blattella germanica* Linné (Blattaria, Blattellidae). *Konty*, *44*(4), 512–525.

Vontas, J., Kioulos, E., Pavlidi, N., Morou, E., Della Torre, A., & Ranson, H. (2012). Insecticide resistance in the major dengue vectors *Aedes albopictus* and *Aedes aegypti*. *Pesticide Biochemistry and Physiology*, *104*(2), 126–131.

Vontas, J. G., Graham, J., & Hemingway, J. (2001). Glutathione *S*-transferases as antioxidant defence agents confer pyrethroid resistance in *Nilaparvata lugens*. *Biochemical Journal*, *357*(1), 65–72.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., … Kang, L. (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nature Communications*, *5*, 2957.

Watanabe, T., Ito, T., Goda, H. M., Ishibashi, Y., Miyamoto, T., Ikeda, K., Taguchi, R., Okino, N., & Ito, M. (2015). Sterylglucoside catabolism in *Cryptococcus neoformans* with endoglycoceramidase-related protein 2 (egcrp2), the first steryl-β-glucosidase identified in fungi. *Journal of Biological Chemistry*, *290*(2), 1005–1019.

Wei, J. J., Fu, T., Yang, T., Liu, Y., & Wang, G. R. (2015). A trpa1 channel that senses thermal stimulus and irritating chemicals in *Helicoverpa armigera*. *Insect Molecular Biology*, *24*(4), 412–421.

Ylla, G., Piulachs, M.-D., & Belles, X. (2017). Comparative analysis of miRNA expression during the development of insects of different

metamorphosis modes and germ-band types. *BMC Genomics*, *18*(1), 774.

Ylla, G., Piulachs, M.-D., & Belles, X. (2018). Comparative transcriptomics in two extreme neopterans reveals general trends in the evolution of modern insects. *iScience*, *4*, 164–179.

Zhou, X., Qian, K., Tong, Y., Zhu, J. J., Qiu, X., & Zeng, X. (2014). De novo transcriptome of the hemimetabolous German cockroach (*Blattella germanica*). *PLoS One*, *9*(9), e106932.

Zhu, F., Gujar, H., Gordon, J. R., Haynes, K. F., Potter, M. F., & Palli, S. R. (2013). Bed bugs evolved unique adaptive strategy to resist pyrethroid insecticides. *Scientific Reports*, *3*, 1456.

Ziegler, R. & van Antwerpen, R. (2006). Lipid uptake by insect oocytes. *Insect Biochemistry and Molecular Biology*, *36*(4), 264–272.

SUPPORTING INFORMATION

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

How to cite this article: Harrison MC, Arning N, Kremer LM, et al. Expansions of key protein families in the German cockroach highlight the molecular basis of its remarkable success as a global indoor pest. *J Exp Zool (Mol Dev Evol)*. 2018;1–11. https://doi.org//jezb.22824