ELSEVIER

The P450 genes of the cat flea, *Ctenocephalides felis:* a CYPome in flux

René Feyereisen [a,b,*]

[a] Department of Plant and Environmental Sciences, University of Copenhagen, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark
[b] Laboratory of Agrozoology, Department of Plants and Crops, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, 9000, Ghent, Belgium

ARTICLE INFO

ABSTRACT

The genome of the cat flea, an ectoparasite of major veterinary importance and the first representative of the Siphonaptera, is highly unusual among arthropod genomes in showing a variable size and a very large number of gene duplications (Driscoll et al., 2020). The cat flea is the target of several classes of insecticides, justifying the description of its CYPome, the complement of P450s that are an important family of detoxification enzymes. 103 P450 genes were annotated on the nine chromosomes, with an additional 12 genes on small, extrachromosomal scaffolds. Only 34 genes were found as single sequences, with 47 duplicated two to four-fold. This included duplication of genes that are mostly single copy P450 genes in other arthropods. Large clusters of mitochondrial clan P450s were observed, resulting in a CYP12 bloom within this clan to 34 genes, a number of mitochondrial P450s not seen in other animals so far. The variable geometry of the cat flea CYPome poses a challenge to the study of P450 function in this species, and raises the question of the underlying causes of single copy control versus multicopy licence of P450 genes.

1. Introduction

The cat flea *Ctenocephalides felis* (Siphonaptera, Pulicidae) is a worldwide obligate ectoparasite of cats and dogs as well as wildlife (Rust and Dryden, 1997; Rust, 2016). It is a carrier of a number of pathogens, from *Rickettsia* to *Bartonella*, and intermediate host of parasites such as tapeworms. As blood feeder, the cat flea is also responsible for allergic dermatitis. A well known, related member of the Order Siphonaptera is the Oriental rat flea, vector of the bubonic plague bacterium *Yersinia pestis*. The medical and veterinary importance of fleas, and the cat flea in particular, has made the protection of household pets a major industry. With few viable alternatives (Rust, 2020), reliance on chemical control is predominant, both oral administration and topical ("spot-on") treatment of the flea host with a number of active ingredients from several classes. Challenges in the toxicology of flea control chemicals are primarily the prevention of toxicity to the hosts (cat or dog) and to their owners (particularly children). Selective toxicity is mainly achieved by maximizing the differential sensitivity of the insecticidal target in the flea with that in the vertebrates, when the target is in the nervous system (Gassel et al., 2014). Even for highly selective agents whose insecticidal target is absent from vertebrates (such as the insect growth regulators methoprene or lufenuron), potential toxic side effects need to be avoided (Gaens et al., 2019). Efficacy at recommended doses therefore depends on a delicate balance between bioavailability of the insecticide and its metabolic inactivation. Efficacy can be compromised by resistance, a

major problem facing chemical control of arthropods (Feyereisen et al., 2015; Nauen et al., 2022). Relatively few (28) cases of resistance have been reported for the cat flea (Mota-Sanchez and Wise, 2021). Target site resistance (*kdr* and *Rdl*) predominate, but metabolic resistance has also been inferred in some cases. For instance, synergism by PBO reduced malathion resistance (Bossard et al., 1998), an indirect sign of cytochrome P450 involvement. No metabolic resistance to isoxazolines such as fluralaner has been reported to date. The metabolism of flea control agents is therefore an important aspect to understand their efficacy and to manage resistance. Yet xenobiotic metabolism in fleas is relatively understudied, but certainly relies on classical processes such as phase I, the introduction of functional groups and phase II, conjugation.

The cat flea genome (Driscoll et al., 2020) now makes it possible to analyze in detail the cytochrome P450 genes in this species that encode the principal phase I enzymes of oxidative metabolism. Here, the full complement of P450 genes (the CYPome) of the cat flea is described, to serve as basis for future research on toxicology of this important pest. A second goal of this study relates to the evolution of the P450 gene family. It is one of the very large families of genes in eukaryotic genomes (Nelson, 2018). The enzymes they encode are mostly characterized by their monooxygenase function, activating molecular oxygen to insert one atom in a substrate while reducing the other to water. This general scheme translates into a great variety of chemical reactions much beyond simple hydroxylation (Guengerich and Munro, 2013). However, the nature of the substrate, from hormone to pesticide or

---

environmental toxicant, brings P450 enzymes in multiple pathways subject to widely different evolutionary constraints and selection regimes. CYPome composition is generally characterized by many genes in few CYP families, and few genes in many CYP families (Sezutsu et al., 2013; Dermauw et al., 2020). A close look at a genome characterized by a high proportion of gene duplicates ("genome in flux") (Driscoll et al., 2020) might reveal if this pattern of duplication is random or restricted to some P450 subfamilies or genes. Furthermore, the CYPome of the cat flea may inform on the relationship between host use and detoxification enzyme complements (Rane et al., 2019). Here, the P450 genes of the cat flea are presented. Reflecting the whole genome, the CYPome comprises many gene duplicates. Furthermore, the cat flea CYPome shows a great expansion of the mitochondrial clan P450s that has not been observed previously in animals.

2. Materials and Methods

2.1. Annotation of P450 genes

The cat flea *Ctenocephalides felis* genome (QVAO01, GCA_003426905.1) was searched by blast with representative P450 sequences from Diptera and Coleoptera (Dermauw et al., 2020), and matched when appropriate with RefSeq models (Annotation Release 100). This resulted in 171 annotations of genes, pseudogenes, loose exons and gene fragments (Supplementary Table 1). Although 40 gene models from RefSeq appeared to be correct, there were 10 fusions of two genes into one gene model, and 2 fusions of three. There were also 2 genes split into two halves, and 17 RefSeq gene models that corresponded to partial pseudogenes (see Supplementary Table 1). The high frequency of frameshifts caused by single nucleotide indels was an additional annotation problem, and evidence from transcripts was not abundant enough to reliably correct all these frameshifts. It was therefore difficult to decide whether these frameshifts were errors from PacBio sequencing (only 25x coverage), or a biological reality of this peculiar genome, with many duplications resulting in some genes being in the process of pseudogenization. The CYPome presented here includes 23 (near) full length sequences with at least one frameshift, of which six were designated with the P (pseudogene) suffix because of lacking transcript evidence, three as a result of short genomic gaps, and one as a result of a premature stop codon (Supplementary Table 1).

2.2. Naming P450 sequences and P450 phylogeny

Full length sequences were given names following the official P450 nomenclature. While this was straightforward for 34 unique sequences represented by a single gene, the remaining 81 sequences consisted of two, three or four duplicates of an additional 34 unique sequences. These sequences were on average 98.3 ± 1.5% identical and all were >95% identical. Driscoll et al. (2020) considered genes with >90% identity as duplicates, so the criteria used in this study were more stringent. Such duplicate sequences were given a,b,c,d suffixes after the gene number.

Sequences were aligned by MAFFT (Katoh et al., 2019) and maximum likelihood phylogeny constructed by iq-Tree (Minh et al., 2020). The trees were drawn by FigTree (http://tree.bio.ed.ac.uk/).

3. Results

3.1. P450 sequences

The cat flea genome is characterized as "plastic" with rampant gene duplication (Driscoll et al., 2020). This was abundantly apparent in this annotation of the CYPome, the first one from the order Siphonaptera. 115 P450 genes were annotated, from a set of 68 unique CYP sequences, of which only half were present as single copy genes. For the remainder, 25 were duplicated once, 5 present in 3 copies and 4 present in 4
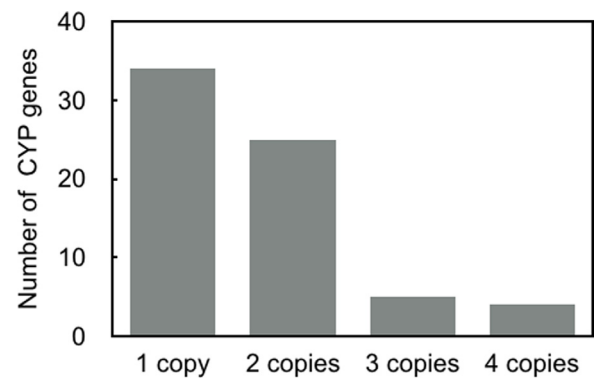

Figure 1. Duplicated P450 genes in the cat flea. Sequences that are >95% identical were considered to be duplicates.

copies (Table 1 and Figure 1). 103 of the 115 genes were distributed on the "big 9 scaffolds" which represent the nine core chromosomes. Mostly tandem duplications, with evidence for segmental duplications covering several genes were observed (Supplementary Table 1). Twelve additional P450 genes were distributed on seven smaller scaffolds. The precise nature of these smaller scaffolds which map to homologous regions on the assembled chromosomes is unclear. All 12 P450 genes (and many gene fragments) on these small scaffolds were duplicates of genes found on the assembled chromosomes. As the genome sequence was obtained from pooled females, and as no two cat fleas have the same genome size (Driscoll et al., 2020), these smaller scaffolds may represent copy number variation (CNV), or optional genome complements.

3.2. Distribution into CYP families

The P450 sequences were distributed in the four major insect CYP clans (Dermauw et al., 2020) and 22 families (including 2 new families in the CYP3 clan). Notably, the cat flea CYPome had a particularly large number of mitochondrial clan P450 with 34 sequences and just 13 CYP4 clan P450s. The pattern of P450 gene duplications in the cat flea is consistent with that observed at the whole genome level. The proportion of duplicated single copy genes at the genome level (38%) (Driscoll et al., 2020) is the same as that observed for the CYPome (47 / 115). At the genome level, these gene duplications represent CNV resulting from unequal crossing over rather than artefacts from misassembled allelic variants. This is also shown for the CYPome where 55 truncated genes and loose exons were annotated. However, the distribution of these gene fragments among CYP clans was uneven, with a disproportionately high number of gene fragments for the two smaller CYP clans (CYP2 and CYP4 clans), and fewer for the larger clans with higher number of full length duplicates (CYP3 and MITO clans).

Arthropod CYPomes generally consist of many families with few genes, and few families with many genes. The many families with few genes often fall into conserved orthologous groups. In Holometabola (Dermauw et al., 2020), these consist of just a dozen P450 genes, including those involved in ecdysteroid metabolism. These genes are all conserved in the cat flea, but several are duplicated: in the mitochondrial clan, CYP314A1 and CYP315A1 are both tandemly duplicated (with also a CYP315A1 pseudogene), while CYP302A1 has one chromosomal copy and one extrachromosomal copy (Figure 2). In the CYP4 clan, both types of CYP4G genes (Feyereisen, 2020) are duplicated, so that the only single copy, conserved orthologs are in CYP2 clan, CYP18A1, CYP306A1, CYP303A1, CYP15A1, CYP307A1 and B1, and a CYP305, as well as the mitochondrial clan CYP49A1 and 301A1.

On the other hand, the few (sub)families with many genes are also represented, such as the CYP9FH, CYP6WN and CYP6WN which show notable blooms. The mitochondrial clan CYP12 family has 25 members in the cat flea, many more than e.g. *Anopheles gambiae* with just four
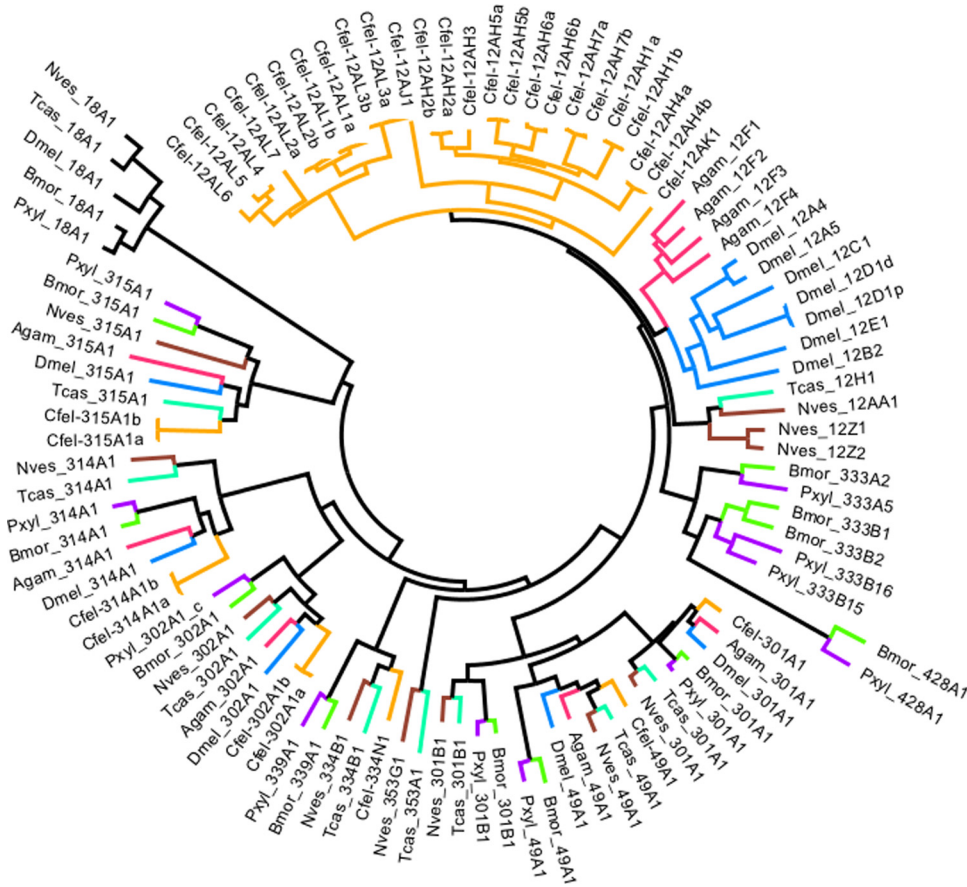
Table 1

CYP numbers by clan in the cat flea. Total CYP genes is the number of genes on chromosomes and on the small scaffolds. This total consists of unique sequences represented either by single genes or by duplicated genes, where duplicates are > 95% identical in sequence to each other.

	CYP2 clan	CYP3 clan	CYP4 clan	mitochondrial clan	total
CYP genes on "big 9" scaffolds (chromosomes)	11	47	13	32	103
CYP genes on small scaffolds	**0**	10	0	2	12
total CYP genes	**11**	57	13	34	115
unique CYP sequences	**10**	27	9	22	68
as single genes	9	10	5	10	34
as duplicated genes (2-4x)	1	30	4	12	47
gene fragments	**13**	15	17	10	55

Figure 2. Maximum likelihood phylogeny of cat flea P450s. The four CYP clans are marked in blue (CYP2 clan), green (CYP3 clan), orange (CYP4 clan) and red (mitochondrial clan). P450s marked with an asterisk * are found on small scaffolds rather than on the chromosomes (big 9 scaffolds).

(Ranson et al., 2002). Figure 3 shows the phylogenetic relationships of mitochondrial clan sequences from two Coleoptera, two Lepidoptera and two Diptera, as well as the cat flea and illustrates this CYP12 bloom as well as the sets of conserved orthologs in the other families.

3.3. CYP gene clusters

Gene duplications are well known to be at the origin of CYP gene clusters in arthropods (Feyereisen, 2012; Sezutsu et al, 2013). The cat flea genome is therefore not surprisingly harbouring many CYP gene clusters. Figure 4 illustrates two such clusters. One, on big scaffold 8 shows a small cluster of CYP315A1 genes, usually a single copy gene in arthropods (Dermauw et al., 2020). Instead, in the cat flea, there is a tandem duplication of CYP315A1 along with another gene (loc 3035/3038) within 14 kb, with one of the copies of CYP315A1 pseudogenized (stop

codon and frameshift in exon 1). There is a second duplication with inversion of the CYP315A1 gene. The two CYP315A1 genes differ by just seven amino acids. The other cluster, on big scaffold 9, shows seven CYP9FH genes within 410 kb. It seems to have resulted from a segmental duplication and inversion of three genes, with subsequent duplication of one gene. This gene, CYP9FH3 is 93% identical to the adjoining CYP9FH4b gene. Supplementary Table 1 shows several more clusters, often containing loose exons or partial genes.

4. Discussion

4.1. CYPome size

CYPome sizes as estimated from automated or manual annotations are definite numbers, attempting to represent the reality of a "type"

Figure 3. Maximum likelihood tree of mitochondrial clan P450s. The cat flea mitochondrial clan P450s are shown along with those from two Diptera (*Drosophila melanogaster* and *Anopheles gambiae*), two Lepidoptera (*Bombyx mori* and *Plutella xylostella*) as well as two Coleoptera (*Tribolium castaneum* and *Nicrophorus vespilloides*). CYP18 sequences (CYP2 clan) are shown as outgroup. The cat flea sequences are from this work, the others are taken from Dermauw et al. (2020).
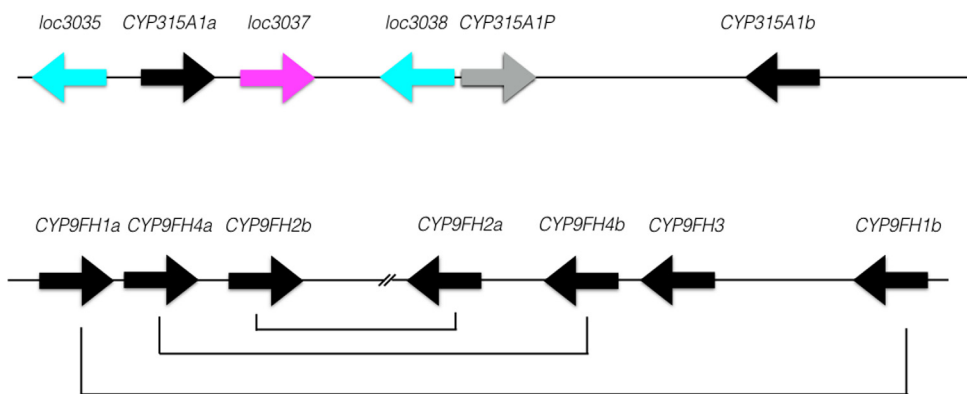

Figure 4. Clusters of CYP genes. Top: CYP315A1 cluster on scaffold 8; bottom: CYP9FH cluster on scaffold 9. Supplementary Table 1 gives the coordinates of more clusters and tandem duplications of CYP genes.

genome (Rane et al., 2019; Dermauw et al., 2020). The inclusion of pseudogenes, gene fragments, or just the use of P450 domain counts (e.g. from Interpro) can bias CYPome size from one study to another and criteria should be well defined before comparisons can be made (Dermauw et al., 2020, Vertacnik et al., 2021). CYPome size is subject to change when the genome assembly improves, or when several genomes of the same species are compared. In that case, changes are a result of gene copy number variation (CNV) in that species. CNV is often an intermediate step (Innan and Kondrashov, 2010) in the gene duplication process observed when comparing sister species. Recently duplicated genes are often found in tandem and can be identical or differing by one or a few residues, e.g. CYP6B4 and CYP6B5 in *Papilio glaucus* (Hung et al., 1996) or CYP6AF1 and CYP6AF2 in *Anopheles gambiae* (Ranson et al., 2002). CNV in insect CYPomes is not well documented yet, variable copy number of CYP6G1 (Schmidt et al., 2010) and CYP6CY3 (Bass et al., 2013) being good examples of CNV associated with insecticide resis-

tance. Indeed, the study of resistance is now providing more examples of P450 gene CNV and gene amplification (Nauen et al., 2022).

The CYPome size of the cat flea is an extreme example of this ambiguity of assigning a precise number to the CYPome size. In part, this may be due to sub-optimal sequencing quality, shown by a relatively high number of single frameshifts, some that can be corrected with the help of transcript sequences. This sub-optimal quality of sequencing probably translates in a sub-optimal assembly as well, and future studies may resolve such deficiencies, as they have, indeed, with sequential improvements of genomes such as Drosophila over the years. In greater part, however, the number of CYP genes in the cat flea is ambiguous because the genomes of Siphonaptera appear to be plastic, with a core of nine chromosomes covering 654 Mb in the cat flea, and another 120 Mb being dispersed in unassembled small scaffolds (Driscoll et al., 2020).

Even on the nine assembled chromosomes (big 9 scaffolds), there is a great number of almost identical copies of P450 genes, probably reflect-

Table 2

CYPome size and life history traits of arthropod ectoparasites. Generalist indicates more than one host species documented, specialist indicates a single known host species. CYPome sizes may to some extent depend on the criteria used by the authors of the studies.

species	CYPome size	reference	host species specialist /generalist	larval feeding	adult feeding (female/male)
Ctenocephalides felis	103	this work	generalist	blood	blood
Glossina sp.	62-82	Rane et al., 2019	generalist	maternal "milk"	blood
Stomoxys calcitrans	202	Rane et al., 2019	generalist	organic matter	blood
Lutzomyia longipalpis	100	Rane et al., 2019	generalist	organic matter	blood & plant sugars / plant sugars
Phlebotomus papatasi	94	Rane et al., 2019	generalist	organic matter	blood & plant sugars / plant sugars
Philornis downsi	102	Romine et al., 2021	generalist	organic matter & blood ?	blood
Cimex lectularius	57	Dermauw et al., 2020	generalist	blood	blood
Rhodnius prolixus	117	Rane et al., 2019	generalist	blood	blood
Pediculus humanus	36	Dermauw et al., 2020	specialist	blood	blood
Lepeophtheirus salmonis	21	Humble et al., 2019	generalist	skin & blood	skin & blood
Varroa destructor	26	Vlogiannitis et al., 2021	specialist	hemolymph	hemolymph
Dermanyssus gallinae	74	this work	generalist	blood	blood
Ixodes scapularis	199	Dermauw et al., 2020	generalist	blood	blood

ing active, present CNV, caught as a snapshot as it were of the genomes of the pooled females that were sequenced. Yet, the genome assemblies of other insects that were obtained from pooled individuals do not show such high numbers of nearly identical sequences (i.e. recent CNV) nor so many gene fragments and pseudogenes (i.e. ongoing selection). The cat flea CYPome is therefore exceptional, reflecting the dynamic nature of the genome as a whole.

Such genomic plasticity has also been observed in the rotifer *Brachionus asplanchnoidis*, with an extreme (2-fold) variation in genome size (Stelzer et al., 2021). However, in that case, the variation is due to satellite DNA rather than active coding genes. Perhaps a closer analog would be in fungi. The plant pathogenic *Nectria haematococca* carries three "optional" chromosomes (conditionally-dispensable supernumerary chromosomes) enriched with duplicated genes (Coleman et al., 2009), accounting for variation between isolates, and perhaps adaptation to diverse ecological niches.

4.2. CYPome size and life history traits

With 103 (chromosomal) CYP genes, the cat flea as obligate ectoparasite does not have a reduced CYPome size when compared to free living insects in sister clades (Diptera and Lepidoptera) for which CYPomes have been annotated (Rane et al., 2019; Dermauw et al. 2020). Other arthropods feeding exclusively on blood have CYPomes sizes ranging from 36 (*Pediculus humanus*), to 199 (*Ixodes scapularis*) (Table 2). Hasty conclusions about a possible relation between parasitic lifestyle and CYPome size are therefore premature.

4.3. CYP gene CNV and duplications

The great number of very closely related sequences seen in the CYPome from the cat flea genome assembly from pooled individuals makes the distinction between interindividual CNV and species-wide gene duplication semantically ambiguous. It is commonly observed that single genes that show no loss or duplications between closely related species, i.e. "stable" genes, are mostly those encoding P450 enzymes with essential physiological function. This is the case for vertebrates (Thomas, 2007) and for 12 Drosophila species (Good et al., 2014). The way this stability is measured, by reference to a common ancestor, is however relative to the placement of that ancestor, as shown over the scale of the arthropod lineage (Dermauw et al., 2020). The association of gene essentiality with gene conservation or phylogenetic stability is intuitively attractive. Yet the two concepts are not fully congruent as shown by the genome-wide study of Waterhouse et al. (2011) in fungi, vertebrates and arthropods. They suggested that the "quantitative distinction between genes with known essential functions and those without is substantially less prominent than the distinction between single-copy constrained genes and those with a multicopy license."

Relatively little is known about which P450 genes have essential functions in insects (and which have not). P450s involved in hormone metabolism or cuticular hydrocarbon biosynthesis are thought to be essential. Yet gene essentiality is probably not a "static and binary property" (Rancati et al., 2018), and is measured by a variety of criteria. On the other hand, single copy control and multicopy licence can be measured by more objective criteria once an evolutionary timescale is defined (see above). This was done elegantly for the timescale of the 12 Drosophila radiation (Good et al., 2014).

The very high proportion of recently duplicated / copy number variable CYP genes in the cat flea therefore offers the opportunity to identify genes under single copy control and multicopy license. Can this then be correlated with gene essentiality ? The P450s known in *Bombyx mori*, *Anopheles gambiae* and Drosophila (and thus presumably in the cat flea as well) to catalyze various steps in ecdysteroid synthesis (CYP307, 306, 302, 315, 314) and inactivation (CYP18) do not fall clearly in the single copy control (or "stable") category: three of them are duplicated (CYP302, 315 and 314, and with a full length pseudogene for CYP315), while they are "stable" in the Drosophila lineage (Good et al., 2014). Conversely CYP307, unstable in Drosophila, has only single copies of the 307A and 307B genes found in the holometabolous ancestor (Sezutsu et al., 2013). It may be relevant that the CYP306A1 and CYP18A1 genes are under single copy control, as they are the product of an ancient gene duplication (basal to Mandibulata), maintained in close synteny (only 5 kb apart in the cat flea), and play opposing roles in determining the ecdysteroid titer (Dermauw et al., 2020). These genes may therefore remain singletons because dosage imbalance would be deleterious and lead to developmental defects. The multicopy license of the CYP302, 315 and 314 genes is not easy to explain.

All higher insects (Neoptera) studied so far have at least one CYP4G gene essential for cuticular hydrocarbon biosynthesis and lethal when knocked out. There are two copies each of the CYP4G1- and CYP4G15-type genes as well as ten gene fragments in the cat flea. The CYP4G subfamily is notoriously "unstable" in insects, with a revolving door pattern of gene birth and death (Feyereisen, 2020) although they are stable in the Drosophila lineage (Good et al., 2014).

Among the genes generally considered non-essential, or with "accessory functions associated with unstable environmental interactions" (Thomas, 2007), while the majority of genes are under multicopy license, some in the CYP3 and CYP4 clans are not. With unequal crossing over as major mechanism of gene duplication, the process leads to a variable size of the duplicated (and lost) chromosomal segment, leading to the duplication of more than one gene (e.g. Figure 4), or the generation of gene fragments. The CYP gene fragments in the cat flea are not evenly distributed among the four CYP clans, with the CYP2 and CYP4 clans having proportionally more gene fragments than the CYP3 and mitochondrial clans, even though they are the two clans with the lowest number of genes.

These observations about which genes are duplicated and which are not leaves an overall picture that does not support a strong link between essentiality and stability. While selection would primarily determine essentiality, the proximal cause(s) of copy number control probably include factors other than selection. Moreover, the fact that the duplications in the CYPome closely mirror what is observed at the whole proteome level (Driscoll et al., 2020) points to an idiosyncrasy of the cat flea genome, and possibly that of other Pulicidae. Of course, a better knowledge of the actual function of each P450 would allow a finer analysis.

4.4. Toxicological relevance

P450s play an important role in insecticide metabolism and resistance (Nauen et al., 2022). While currently there is nothing known about the precise function of any cat flea P450 beyond presumptions that are based on conserved orthology, a few points merit attention. In the sister clades to Siphonaptera, Diptera and Lepidoptera, many studies point to members of the CYP3 clan as key players in insecticide metabolism and resistance (Vontas et al., 2020; Katsavou et al., 2022). However, members of the CYP2 and mito clans have also been shown to metabolize xenobiotics (Shi et al., 2022). The high number of closely related sequences in the cat flea, and the probable variability in numbers between populations, perhaps even individuals, will make it difficult to assign insecticide metabolism to a definite gene, and, for resistance, make the design of molecular probes for overexpression or amplification very difficult. The proliferation of mitochondrial clan P450s of the CYP12 family is particularly noteworthy. In the house fly, CYP12A1 is constitutively overexpressed in a diazinon-resistant strain, and it metabolizes diazinon, cyclodienes and a variety of xenobiotics (Guzov et al. 1998). In Drosophila, CYP12A4 overexpression confers lufenuron resistance (Bogwitz et al., 2005). Conversely, CYP12A5 is responsible for the bioactivation of nitenpyram (Harrop et al., 2018). CYP12D1, a recently duplicated gene of Drosophila, is commonly observed in transcriptomic surveys to be associated with xenobiotic resistance, and its transgenic overexpression confers resistance to DDT and dicyclanil (Daborn et al., 2007). The mitochondrial clan CYP353D1 of *Laodelphax striatellus* metabolizes buprofezin and imidacloprid (Elzaki et al., 2017a; Elzaki et al., 2017b). Because of the focus on CYP3 clan P450s in studies of resistance or adaptation to host plants in Diptera and Lepidoptera, the toxicological relevance of insect mitochondrial P450s is often overlooked. Recent work on *Helicoverpa armigera* (Shi et al., 2022) also points to the capacity of five mitochondrial clan P450s to metabolize xenobiotics.

5. Conclusion

Resistance, and in particular metabolic resistance, is currently not a major problem in the cat flea (Rust, 2016), with control failures mainly attributable to application problems. But resistance is always "around the corner", and in the case of P450-dependent detoxification (or activation), the variable geometry of the cat flea CYPome may prove to hinder identification of the responsible gene. In turn, this will delay the development of molecular diagnostics. These are more easily deployed in the case of target site resistance based on point mutations in known targets.

The very large number of closely related sequences that point to high CNV and rate of gene duplication make the cat flea CYPome a standout case for arthropods, and possibly animals in general. The mitochondrial clan P450s as the second largest clan after the CYP3 clan is unprecedented and calls for greater attention to the functions of these P450 enzymes.

Data availability

The full CYPome annotation of genes, pseudogenes and gene fragments is provided in Supplementary Table 1, and a FASTA format file of all full length sequences is provided in Supplementary File 1.

References

Bass, C., Zimmer, C.T., Riveron, J.M., Wilding, C.S., Wondji, C.S., Kaussmann, M., Field, L.M., Williamson, M.S., Nauen, R., 2013. Gene amplification and microsatellite polymorphism underlie a recent insect host shift. Proceedings of the National Academy of Sciences 110, 19460–19465 doi.org/10.1073/pnas.1314122110.

Bogwitz, M.R., Chung, H., Magoc, L., Rigby, S., Wong, W., O'Keefe, M., McKenzie, J.A., Batterham, P., Daborn, P.J., 2005. Cyp12a4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. Proceedings of the National Academy of Sciences 102, 12807–12812 doi.org/10.1073/pnas.0503709102.

Bossard, R.L., Hinkle, N.C., Rust, M.K., 1998. Review of Insecticide Resistance in Cat Fleas (Siphonaptera: Pulicidae). Journal of Medical Entomology 35, 415–422 doi.org/10.1093/jmedent/35.4.415.

Coleman, J.J., Rounsley, S.D., Rodriguez-Carres, M., Kuo, A., Wasmann, C.C., Grimwood, J., Schmutz, J., Taga, M., White, G.J., Zhou, S., Schwartz, D.C., Freitag, M., Ma, L., Danchin, E.G.J., Henrissat, B., Coutinho, P.M., Nelson, D.R., Straney, D., Napoli, C.A., Barker, B.M., Gribskov, M., Rep, M., Kroken, S., Molnár, I., Rensing, C., Kennell, J.C., Zamora, J., Farman, M.L., Selker, E.U., Salamov, A., Shapiro, H., Pangilinan, J., Lindquist, E., Lamers, C., Grigoriev, I.V., Geiser, D.M., Covert, S.F., Temporini, E., VanEtten, H.D., 2009. The Genome of *Nectria haematococca*: Contribution of Supernumerary Chromosomes to Gene Expansion. PLoS Genet 5, e1000618 doi.org/10.1371/journal.pgen.1000618.

Daborn, P.J., Lumb, C., Harrop, T.W.R., Boey, A., Wong, W., ffrench-Constant, R.H., Batterham, P., 2007. Evaluating the insecticide resistance potential of eight *Drosophila melanogaster* cytochrome P450 genes by transgenic over-expression. Insect Biochemistry and Molecular Biology 37, 512–519 doi.org/10.1016/j.ibmb.2007.09.003.

Dermauw, W., Van Leeuwen, T., Feyereisen, R., 2020. Diversity and evolution of the P450 family in arthropods. Insect Biochemistry and Molecular Biology 127, 103490 doi.org/10.1016/j.ibmb.2020.103490.

Driscoll, T.P., Verhoeve, V.I., Gillespie, J.J., Johnston, J.S., Guillotte, M.L., Rennoll-Bankert, K.E., Rahman, M.S., Hagen, D., Elsik, C.G., Macaluso, K.R., Azad, A.F., 2020. A chromosome-level assembly of the cat flea genome uncovers rampant gene duplication and genome size plasticity. BMC Biol 18, 70 doi.org/10.1186/s12915-020-00802-7.

Elzaki, M.E.A., Miah, M.A., Han, Z., 2017a. Buprofezin Is metabolized by CYP353D1v2, a cytochrome P450 associated with imidacloprid resistance in *Laodelphax striatellus*. Int. J. Mol. Sci. 18, 2564.

Elzaki, M.E.A., Miah, M.A., Wu, M., Zhang, H., Pu, J., Jiang, L., Han, Z., 2017b. Imidacloprid is degraded by CYP353D1v2, a cytochrome P450 overexpressed in a resistant strain of *Laodelphax striatellus*. Pest Manag Sci 73, 1358–1363. doi:10.1002/ps.4570.

Feyereisen, R., 2012. Insect CYP genes and P450 enzymes. In: Gilbert, L.I. (Ed.), Insect Molecular Biology and Biochemistry. Elsevier, pp. 236–316.

Feyereisen, R., 2020. Origin and evolution of the CYP4G subfamily in insects, cytochrome P450 enzymes involved in cuticular hydrocarbon synthesis. Molecular Phylogenetics and Evolution 143, 106695 doi.org/10.1016/j.ympev.2019.106695.

Feyereisen, R., Dermauw, W., Van Leeuwen, T., 2015. Genotype to phenotype, the molecular and physiological dimensions of resistance in arthropods. Pesticide Biochemistry and Physiology 121, 61–77 doi.org/10.1016/j.pestbp.2015.01.004.

Gaens, D., Rummel, C., Schmidt, M., Hamann, M., Geyer, J., 2019. Suspected neurological toxicity after oral application of fluralaner (Bravecto®) in a Kooikerhondje dog. BMC Vet Res 15, 283 doi.org/10.1186/s12917-019-2016-4.

Gassel, M., Wolf, C., Noack, S., Williams, H., Ilg, T., 2014. The novel isoxazoline ectoparasiticide fluralaner: Selective inhibition of arthropod γ-aminobutyric acid- and l-glutamate-gated chloride channels and insecticidal/acaricidal activity. Insect Biochemistry and Molecular Biology 45, 111–124 doi.org/10.1016/j.ibmb.2013.11.009.

Good, R.T., Gramzow, L., Battlay, P., Sztal, T., Batterham, P., Robin, C., 2014. The molecular evolution of cytochrome P450 genes within and between drosophila species. Genome Biol Evol. 20, 1118–1134. doi:10.1093/gbe/evu083.

Guengerich, F.P., Munro, A.W., 2013. Unusual Cytochrome P450 Enzymes and Reactions. Journal of Biological Chemistry 288, 17065–17073 doi.org/10.1074/jbc.R113.462275.

Guzov, V.M., Unnithan, G.C., Chernogolov, A.A., Feyereisen, R., 1998. CYP12A1, a Mitochondrial Cytochrome P450 from the House Fly. Archives of Biochemistry and Biophysics 359, 231–240 doi.org/10.1006/abbi.1998.0901.

Harrop, T.W., Denecke, S., Yang, Y.T., Chan, J., Daborn, P.J., Perry, T., Batterham, P., 2018. Evidence for activation of nitenpyram by a mitochondrial cytochrome P450 in *Drosophila melanogaster*. Pest. Manag. Sci 74, 1616–1622 doi.org/10.1002/ps.4852.

Humble, J.L., Carmona-Antoñanzas, G., McNair, C.M., Nelson, D.R., Bassett, D.I., Egholm, I., Bron, J.E., Bekaert, M., Sturm, A., 2019. Genome-wide survey of cytochrome P450 genes in the salmon louse *Lepeophtheirus salmonis* (Krøyer, 1837). Parasites Vectors 12, 563 doi.org/10.1186/s13071-019-3808-x.

Hung, C.F., Holzmacher, R., Connolly, E., Berenbaum, M.R., Schuler, M.A., 1996. Conserved promoter elements in the CYP6B gene family suggest common ancestry for cytochrome P450 monooxygenases mediating furanocoumarin detoxification. Proc Natl Acad Sci U S A. 93, 12200–12205. doi:10.1073/pnas.93.22.12200.

Innan, H., Kondrashov, F., 2010. The evolution of gene duplications: classifying and distinguishing between models. Nature Reviews Genetics 11, 97–108 doi.org/10.1038/nrg2689.

Katoh, K., Rozewicki, J., Yamada, K.D, 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics 20, 1160–1166.

Katsavou, E., Riga, M., Ioannidis, P., King, R., Zimmer, C.T., Vontas, J., 2022. Functionally characterized arthropod pest and pollinator cytochrome P450s associated with xenobiotic metabolism. Pestic Biochem Physiol 181, 105005. doi:10.1016/j.pestbp.2021.105005.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol 37, 1530–1534.

Mota-Sanchez, D. and J.C. Wise. 2021. The Arthropod Pesticide Resistance Database. Michigan State University. On-line at: http://www.pesticideresistance.org

Nauen, R, Bass, C., Feyereisen, R., Vontas, J., 2022. The role of cytochrome P450s in insect toxicology and resistance. Annu. Rev. Entomol. 67, 105–124 doi.org/10.1146/annurev-ento-070621-061328.

Nelson, D.R., 2018. Cytochrome P450 diversity in the tree of life. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1866, 141–154 doi.org/10.1016/j.bbapap.2017.05.003.

Rancati, G., Moffat, J., Typas, A., Pavelka, N., 2018. Emerging and evolving concepts in gene essentiality. Nat Rev Genet 19, 34–49 doi.org/10.1038/nrg.2017.74.

Rane, R.V., Ghodke, A.B., Hoffmann, A.A., Edwards, O.R., Walsh, T.K., Oakeshott, J.G., 2019. Detoxifying enzyme complements and host use phenotypes in 160 insect species. Current Opinion in Insect Science 31, 131–138 doi.org/10.1016/j.cois.2018.12.008.

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M.V., Unger, M.F., Collins, F.H., Feyereisen, R., 2002. Evolution of Supergene Families Associated with Insecticide Resistance. Science 298, 179–181 doi.org/10.1126/science.1076781.

Romine, M., Knutie, S.A., Crow, C.M., Vaziri G.J., Chaves, J., Koop, J.A.H., Lamichhaney, S. 2021. The genome sequence of the avian vampire fly (*Philornis downsi*), an invasive nest parasite of Darwin's finches in Galápagos. doi.org/10.1101/2021.06.09.447800.

Rust, M., 2016. Insecticide Resistance in Fleas. Insects 7, 10 doi.org/10.3390/insects7010010.

Rust, M.K., 2020. Recent Advancements in the Control of Cat Fleas. Insects 11, 668 doi.org/10.3390/insects11100668.

Rust, M.K., Dryden, M.W., 1997. The biology, ecology, and management of the cat flea. Annu. Rev. Entomol. 42, 451–473 doi.org/10.1146/annurev.ento.42.1.451.

Schmidt, J.M., Good, R.T., Appleton, B., Sherrard, J., Raymant, G.C., Bogwitz, M.R., Martin, J., Daborn, P.J., Goddard, M.E., Batterham, P., Robin, C., 2010. Copy Number Variation and Transposable Elements Feature in Recent, Ongoing Adaptation at the Cyp6g1 Locus. PLoS Genet 6, e1000998 doi.org/10.1371/journal.pgen.1000998.

Sezutsu, H., Le Goff, G., Feyereisen, R., 2013. Origins of P450 diversity. Phil. Trans. R. Soc. B 368, 20120428 doi.org/10.1098/rstb.2012.0428.

Shi, Y., Qu, Q., Wang, C., He, Y., Yang, Y., Wu, Y., 2022. Involvement of CYP2 and mitochondrial clan P450s of *Helicoverpa armigera* in xenobiotic metabolism. Insect Biochemistry and Molecular Biology 140, 103696 doi.org/10.1016/j.ibmb.2021.103696.

Stelzer, C.P., Blommaert, J., Waldvogel, A.M., Pichler, M., Hecox-Lea, B., Mark Welch, D.B., 2021. Genome structure of *Brachionus asplanchnoidis* , a Eukaryote with intrapopulation variation in genome size (preprint). Evolutionary Biology doi.org/10.1101/2021.03.09.434534.

Thomas, J.H., 2007. Rapid Birth–Death Evolution Specific to Xenobiotic Cytochrome P450 Genes in Vertebrates. PLos Genetics 3, e67.

Vertacnik, K., Herrig, D., Godfrey, R.K., Hill, T., Geib, S., Unckless, R., Nelson, D., Linnen, C.R., 2021. Ecological correlates of gene family size: the draft genome of the redheaded pine sawfly *Neodiprion lecontei* (preprint). Evolutionary Biology doi.org/10.1101/2021.03.14.435331.

Vlogiannitis, S., Mavridis, K., Dermauw, W., Snoeck, S., Katsavou, E., Morou, E., Harizanis, P., Swevers, L., Hemingway, J., Feyereisen, R., Van Leeuwen, T., Vontas, J., 2021. Reduced proinsecticide activation by cytochrome P450 confers coumaphos resistance in the major bee parasite *Varroa destructor*. Proc Natl Acad Sci USA 118 e2020380118doi.org/10.1073/pnas.2020380118.

Vontas, J., Katsavou, E., Mavridis, K., 2020. Cytochrome P450-based metabolic insecticide resistance in *Anopheles* and *Aedes* mosquito vectors: Muddying the waters. Pesticide Biochemistry and Physiology 170, 104666 doi.org/10.1016/j.pestbp.2020.104666.

Waterhouse, R.M., Zdobnov, E.M., Kriventseva, E.V., 2011. Correlating Traits of Gene Retention, Sequence Divergence, Duplicability and Essentiality in Vertebrates, Arthropods, and Fungi. Genome Biology and Evolution 3, 75–86 doi.org/10.1093/gbe/evq083.