Wilding, CS

Regulating resistance: CncC:Maf, antioxidant response elements and the overexpression of detoxification genes in insecticide resistance

http://researchonline.ljmu.ac.uk/8702/

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)


LJMU has developed LJMU Research Online for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

http://researchonline.ljmu.ac.uk/
Regulating resistance: CncC:Maf, antioxidant response elements and the overexpression of detoxification genes in insecticide resistance

Craig S. Wilding
School of Natural Sciences and Psychology,
Liverpool John Moores University,
Liverpool,
L3 3AF,
UK
Tel: +44(0)151 231 2500
While genetic and genomic tools have greatly furthered our understanding of resistance-associated mutations in molecular target sites of insecticides, the genomic basis of transcriptional regulation of detoxification loci in insect pests and vectors remains relatively unexplored. Recent work using RNAi, reporter assays and comparative genomics are beginning to reveal the molecular architecture of this response, identifying critical transcription factors and their binding sites. Central to this is the insect ortholog of the mammalian transcription factor Nrf2, Cap ‘n’ Collar isoform-C (CncC) which as a heterodimer with Maf-S regulates the transcription of phase I, II and III detoxification loci in a range of insects with CncC knockdown or upregulation directly affecting phenotypic resistance. CncC:Maf binds to specific antioxidant response element sequences upstream of detoxification genes to initiate transcription. Recent work is now identifying these binding sites for resistance-associated loci and, coupled with genome sequence data and reporter assays, enabling identification of polymorphisms in the CncC:Maf binding site which regulate the insecticide resistance phenotype.
Exposure to insecticide instigates a complex response through which insects sequester, detoxify or excrete toxins before they reach their target or have other adverse consequences. The battery of detoxification genes and those elements which control their coordinated response has been labelled ‘the defensome’ [1]. The insect defensome must cope with a variety of assaults from foodstuffs e.g. haem breakdown products or plant allelochemicals, but has been latterly co-opted to deal with xenobiotic insecticidal challenge. Whilst the mammalian xenobiotic response has received much attention, a detailed understanding of the mechanistic basis of detoxifying enzyme upregulation in the insecticide resistance response of insects has been lacking. Recent work on both model, and non-model insects is beginning to redress this imbalance.

**CncC:Maf regulates insecticide resistance and resistance-associated genes**

Gene expression is regulated by a complex of transcriptional activators that bind to regions upstream of transcription start sites recruiting chromatin-modifying factors and the RNA polymerase II containing transcription initiation apparatus. Core RNA polymerase is capable of DNA dependant RNA synthesis *in vitro* but incapable of specific promoter recognition in the absence of additional factors.

In eukaryotes a key transcriptional activator in the response to a wide variety of stressors is encoded by nuclear factor, erythroid 2-like *Nfe2l2 (Nrf2)* [2-5], a mammalian bZIP family transcription factor that binds to specific promoter motifs – termed antioxidant response elements (AREs) - stimulating transcription. In mammals, Nrf2 is a key regulator of both developmental pathways and the rather nebulously titled ‘stress response’ [2-4]. Under normal conditions, Nrf2 is retained cytoplasmically, bound to the cytoskeletal ubiquitin ligase Keap1. Upon stress exposure, Nrf2 releases, translocates to the nucleus, and forms a heterodimer with a small Muscle Aponeurosis Fibromatosis (Maf-S) protein [6] binding to AREs upstream of a battery of antioxidant genes (Figure 1) including GSTs [7-9], carboxylesterases [10], cytochromes p450 [11] and ABC transporters [12] and is involved in regulation of the proteasome, serving to degrade damaged proteins and enzymes following stress-induced damage [13]. In *Drosophila* the insect Nrf2 ortholog *Cap ‘n’ collar isoform C*, (CncC), is known to have
a central role in both development and the ‘stress response’ [5,14]. Xenobiotic exposure, including insecticidal challenge falls under this banner. If CncC:Maf regulates the expression of insecticide-resistance associated genes then perturbations to CncC levels, or ARE polymorphisms should alter both phenotypic insecticide resistance and detoxification gene expression. Thus, a regulatory role of CncC:Maf in the response to insecticides may occur through a variety of mechanisms: upregulation of CncC/Maf (leading to increased target transcription), down-regulation of Keap1 (increasing nuclear translocation of CncC:Maf), mutations in key domains of these proteins, or mutations in AREs upstream of target genes affecting promoter activity. Metabolic insecticide resistance can occur due to either changes in enzyme activity resulting from coding polymorphisms or due to constitutive upregulation of detoxification genes. In either case, those transcription factors initiating expression of detoxification genes must themselves be constitutively expressed. CncC is itself constitutively activated in DDT-resistant Drosophila strains [15] as is Nrf2 is in mammals [16] although these transcription factors do have a relatively short half-life (<20 min) [17]. Constitutive CncC overexpression is also seen in a number of arthropods e.g. resistant Tribolium [18], Anopheles stephensi [1] and spider mites [19] suggesting that this may underlie the resistant phenotype in some instances. Although mutations to CncC, Maf-S or Keap1 may have phenotypic effects, e.g. deletion of the NHB1 domain can result in induced expression of CncC targets [20] there is, as yet, no evidence that naturally occurring mutations to these highly conserved TFs underpin resistance. Initial studies in Drosophila [21] demonstrated that either overexpressing CncC, or introducing a loss of function Keap1 mutation not only upregulated the detoxifying enzyme gstD1, a gene with an upstream ARE, but also significantly increased survival to the toxic herbicide paraquat. By contrast, RNAi knockdown (KD) of CncC decreased both gstD1 expression and survival demonstrating the importance of CncC:Maf for insect survival in the face of xenobiotic exposure. The first work to study the role of CncC:Maf in a true resistant phenotype used tissue specific Keap1 KD (releasing CncC for cytoplasmic transposition) demonstrating a significant increase in resistance to the organophosphate malathion in Drosophila melanogaster [22]. The same study showed that >70% of genes upregulated
following phenobarbital (a prototypical inducer of the xenobiotic response) exposure are also upregulated by ectopic CncC exposure [22] demonstrating the breadth of effect of this TF. Recent work now shows the universality of the role of CncC:Maf in insecticide resistance with studies on *Drosophila*, flour beetles [18], Colorado potato beetles [23], *Aphis gosypii* [24] and spider mites (Arachnidae) [19] all showing that perturbing the CncC:Maf balance affects resistance to a variety of insecticides and alters the expression of key genes previously demonstrated to be involved in this resistance (Table 1). These studies have used a variety of approaches including CncC/Maf knockdown through RNAi, targeted GAL4/UAS overexpression of CncC/Maf and loss-of-function mutations in Keap1.

The decreasing cost of sequencing now enables understanding the whole transcriptomic response of perturbing CncC:Maf. In *Tribolium*, RNASeq analysis after CncC KD showed 662 genes had increased expression and 91 downregulation including a range of phase I, II and III genes [25]. It is unlikely that all have AREs and are under direct influence of CncC but that disturbing the CncC:Maf balance instigates a cascade response. Ingham *et al.* also knocked-down MAF in a multi-insecticide resistant strain of *Anopheles gambiae* [26]. KD increased mortality to DDT and pyrethroids (it did not redress full susceptibility but this strain is nearly fixed for target-site resistance mechanisms) and, through microarray analysis, the transcriptomic response to MAF KD was determined. Here, genes expressed differentially were correlated with a mined dataset of differentially expressed genes from multiple IR studies to identify transcripts upregulated in microarrays and correlated with CncC:Maf-S expression including the key Anopheline detoxification candidates cyp6m2 and Gstd1.

**Antioxidant Response Elements and Insecticide Resistance**

Mammalian studies have identified a consensus ARE motif to which CncC:Maf binds: 5′-TMAnnRTGAYnnnGCRwww-3′ [27]. The experimentally determined *Drosophila* motif is similar but whilst demonstrating a consensus exhibits substantial variability (Figure 2). This motif conservation enables its genome-wide identification computationally through positional matrix screening (see Fig
2) e.g. using Motifdb [28]. However, insects are a diverse and ancient Class (the time from the
Drosophila-Anopheles MRCA is 265MY and Drosophila-Myzus 358MY c.f. 90MY between human and
mouse) [29]. Since in mammalian systems a “universally applicable consensus sequence cannot be
derived” [30], the presumption that the Drosophila positional matrix is appropriate for other insects
remains untested. However, differences in Tribolium AREs [18] versus Drosophila (Figure 2) suggest
AREs in other insects require experimental identification. The ideal method of identifying binding sites
for CncC:Maf involves ChIP-Seq as undertaken in Drosophila [31-33]. A constraining factor on the
ability to undertake ChIP-Seq for other insects is the lack of validated CncC or Maf antibodies (although
ModEncode [34] circumvented such difficulties through use of ChIP-seq on transgenic flies expressing
CncC-eGFP fusion proteins with immunoprecipitation performed using an anti-GFP antibody).

Both in vivo and in vitro reporter assays have been used to detect the functionality of AREs. Whilst
such reporter assays clearly show AREs drive expression, in the absence of CncC:Maf overexpression,
it is polymorphisms differentiating resistant from susceptible animals which will be causal of
resistance and of use for resistance management [35]. Sometimes these may be gross polymorphisms.
Inserted transposable elements (TEs) can carry TFBSs e.g. the Bari-Jheh TE brings new AREs upstream
of two juvenile hormone epoxy hydrolase genes mediating survival to malathion and paraquat [36]
and AREs are found in other Drosophila TEs [36]. SNPs are also a likely source. In humans, ARE
sequence polymorphisms underlie inter-individual gene expression variation [27,37,38] with even
single base changes affecting ARE functionality. Insects have much higher levels of sequence diversity
than humans e.g. in Anopheles π=1.53% for a typical autosome within 1kbp upstream of genes where
AREs would reside and across the genome there is 1 variant base every 2bp [39]. Thus it seems likely
that ARE SNPs may affect expression and that there is a reservoir of SNPs in AREs which may be
selected following insecticide challenge. Experimentally introduced ARE SNPs can be shown to affect
detoxification gene expression e.g. mutagenesis of the ARE upstream of a gstD1-GFP reporter
demonstrated only the WT ARE was inducible by stress (e.g. paraquat or H202) indicating the effect of
polymorphisms on promoter activity [21]. Kalsi and Palli [18] also examined reporter activity of various
CYP6B gene promoters from *Tribolium* demonstrating that SNPs can significantly affect expression. For *D. melanogaster* strains differing in DDT resistance levels a 15bp deletion in a CncC:Maf binding site exhibiting between-strain polymorphism correlated with DDT susceptibility [40] although when association studies of DDT resistance levels were conducted on the *Drosophila* Genetics Reference Panel, this variant was not associated with DDT resistance [41]. Whilst these studies demonstrate promoter activity of AREs, what is clearly needed is an understanding of the effect of ARE SNPs on resistance and expression e.g. using Crispr [42] driven disruption of AREs in defined genetic backgrounds.

**Role of other TFs in resistance**

The transcription initiation machinery is complex and a CncC and ARE focus may be short-sighted. Kalsi and Palli [23] conducted RNAi knockdown studies in *Tribolium* on members of three superfamilies bHLH/PAS, bZIP and Nuclear Receptors (Table 1). KD of CncC, Maf or Methoprene tolerant all caused significant increases in mortality to the pyrethroid deltamethrin but crucially, only CncC and Maf KD also significantly altered the expression of key detoxification genes of the Cyp6BQ family. Whilst this appears to indicate the CncC:Maf pathway is more important in this phenotype, other transcription factors may be involved in other resistance phenotypes e.g. RNAi KD of the *Aphis gossypii* aryl hydrocarbon receptor affected the gossypol resistance associated Cyp6AD2 [43], and reduced deltamethrin resistance in *T. castaneum* [18], the FOXA TF is implicated in Bti resistance in the Lepidopterans *Helicoverpa* and *Spodoptera* [44], and putative TF binding sites such as members of HNF family (also KD screened in [23]) have been identified in sequencing studies of resistant *Aedes* [45] and TFBSs identified in TEs inserted upstream of detoxification genes in *Drosophila* [46]. However, for these studies there has been no follow-up to identify and characterise their binding sites. This may be complicated since binding sites for other TFs may not be proximal (as are AREs) since upstream of genes lies both the proximal promoter and various cis-regulatory modules. The methods for identification and characterisation of TFBSs in CREs have been reviewed [47,48] and application of
these methods will address this knowledge gap. In *Drosophila* a large body of work is accumulating to develop a comprehensive map of transcription factors and transcription factor binding sites (TFBSs) [48-50] empowering computational approaches for TFBS identification e.g. [51]. Such work needs to extend also into other insects given the economic and societal impacts of insecticide resistance. The first step in this is knowledge of the TF repertoire and which genes are *cis*-regulated. Genome sequencing efforts have enabled annotation of, for example, bHLH transcription factors in lice [52], Psyllidae [53], *Nasonia* [54], *Nilaparvata* [55] and vector mosquitoes [56] and further work to identification their roles and binding sites is necessary. As genome-wide allelic imbalance studies are now demonstrably feasible and affordable for insects [57] identification of *cis*-regulated genes in resistant insects will aid the honing of the search.

**Conclusions and future directions**

It is clear that CncC:Maf has an important role in insecticide resistance and that CncC upregulation and/or polymorphisms in its response elements directly affect regulation of detoxification genes. The high levels of phenotypic resistance seen in many insects to a range of insecticides cautions that other transcription factors and enhancers are likely involved. The relative ease of study of CncC and its proximal ARE should not draw attention away from searching for other TFs and characterising these in the way that has started to occur for CncC:Maf. Concerted efforts employing comparative genomics, true GWAS, CHiP-Seq and Crispr to further our understanding of this complex phenotype is needed.
Figure 1. Under normal conditions CncC is held in the cytoplasm by the ubiquitin ligase Keap1 and degraded through the proteasome pathway. Under oxidative stress such as insecticidal exposure, CncC dissociates from Keap1, translocates to the nucleus and forms a heterodimer with Maf-S. The CncC/Maf heterodimer binds to antioxidant response elements (AREs) upstream of target genes and initiates transcription, in the example here of a cytochrome P450.
Figure 2. Variability in the antioxidant response element sequence. 2A. Sequence logo for CncC:Maf-S ARE binding site in *Drosophila melanogaster* identified through ChIP-seq experiments. Logo generated at jaspar.genereg.net (Matrix ID: MA0530.1) [58]. Figure 2B. Alignment of AREs identified upstream of key cytochrome P450 genes of insecticide resistant *Tribolium castaneum* [18]. Note that whereas the sequence logo for *Drosophila* indicates a high likelihood for a C at position 11, at the equivalent position in the *Tribolium* AREs is a T (boxed). Note that Position 1 in Figure 2A = base five of the mammalian ARE (5′-TMAntRTGAYnnnGCRwwww-3′)
Table 1. RNAi knockdown of transcription factors involved in insecticide resistance in insect and arachnid species.  

<table>
<thead>
<tr>
<th>Species</th>
<th>Phenotype</th>
<th>KD target</th>
<th>Effect on phenotypic resistance</th>
<th>Effect on gene expression</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemiptera</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aphis gossypii</td>
<td>gossypol tolerance</td>
<td>CncC</td>
<td>Increased gossypol tolerance</td>
<td>Cyp6AD2 downregulated (qPCR)</td>
<td>[24]</td>
</tr>
<tr>
<td>Aphis gossypii</td>
<td>gossypol tolerance</td>
<td>Ahr, Arnt</td>
<td>Increased gossypol tolerance</td>
<td>Cyp6AD2 downregulated (qPCR)</td>
<td>[43]</td>
</tr>
<tr>
<td>Coleoptera</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tribolium castaneum</td>
<td>Deltamethrin resistance</td>
<td>CncC, Maf-S, Met, HNF4, HR96, Spineless, USP</td>
<td>Increased mortality</td>
<td>CncC KD: Cyp6BQ2, Cyp6BQ4, Cyp6BQ6, Cyp6BQ7, Cyp6BQ9, Cyp6BQ11, Cyp6BQ12 (qPCR)</td>
<td>[18]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Increased mortality</td>
<td>MAF: Cyp6BQ2, Cyp6BQ3, Cyp6BQ4, Cyp6BQ5, Cyp6BQ6, Cyp6BQ7, Cyp6BQ9, Cyp6BQ10, Cyp6BQ12 (qPCR)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Increased mortality</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No significant effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No significant effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No significant effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No significant effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tribolium castaneum</td>
<td>Deltamethrin resistance</td>
<td>CncC</td>
<td>Not tested, but see above</td>
<td>662 genes upregulated, 91 downregulated (RNASeq). CnCC, Cyp6BQ2, Cyp6BQ6, Cyp6BQ7, Cyp6BQ9 (qPCR)</td>
<td>[25]</td>
</tr>
<tr>
<td>Leptinotarsa decemlineata</td>
<td>Imidacloprid resistance</td>
<td>CncC</td>
<td>Survival decreased from 54% to 5% following KD</td>
<td>Cyp9Z25, Cyp9Z29, Cyp6BJ1v1, Cyp6BJa/b</td>
<td>[23]</td>
</tr>
<tr>
<td>Lepidoptera</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helicoverpa armigera</td>
<td>Bti resistance (Cry1AC toxin)</td>
<td>Fox-A</td>
<td>Lower Bti mortality and higher pupation following KD</td>
<td>ABCC2, ABCC3 (qPCR)</td>
<td>[44]</td>
</tr>
</tbody>
</table>
### Diptera

<table>
<thead>
<tr>
<th>Species</th>
<th>Resistance</th>
<th>Gene/Expression</th>
<th>Effect</th>
<th>Reduced expression of Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Anopheles gambiae</em></td>
<td>Permethrin, deltamethrin, DDT resistance</td>
<td><em>Maf-S</em></td>
<td>Increased mortality to DDT, permethrin, deltamethrin. No effect on bendiocarb mortality. Decreased mortality to malathion.</td>
<td>Cyp6M2, GstD1, GstD3, Jheh1, Jheh2, Gnmt. Increased expression of Cyp4H17 [26]</td>
</tr>
<tr>
<td><em>Culex quinquefasciatus</em></td>
<td>Permethrin resistance</td>
<td><em>GSαS Adenylyl cyclase</em></td>
<td>Increased susceptibility of permethrin</td>
<td>GSαS KD: Cyp9M10, Cyp6AA7, Cyp9J34 (qPCR) AC KD: Cyp9M10, Cyp9J34, Cyp9J40, Cyp6AA7 (qPCR) [59]</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>Paraquat survival</td>
<td><em>CncC Keap1</em></td>
<td>Decreased paraquat survival</td>
<td>gstD1 expression reduced. gstD1 expression increased [21]</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td></td>
<td><em>CncC Keap1</em></td>
<td>Increased malathion resistance</td>
<td>Reduced expression of Cyp6a2, Cyp6a8, gstD2, gstD7, Jheh1 (qPCR) [22]</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>DDT resistance</td>
<td><em>CncC</em></td>
<td></td>
<td>Reduced expression of Cyp6a2, Cyp6a8 (qPCR) [15]</td>
</tr>
</tbody>
</table>

### Acari

<table>
<thead>
<tr>
<th>Species</th>
<th>Resistance</th>
<th>Gene/Expression</th>
<th>Effect</th>
<th>Reduced expression of Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Tetranychus cinnabarinus</em></td>
<td>Fenpropathrin resistance</td>
<td><em>CncC Maf-S</em></td>
<td>LC30 increased from 12.75% to 19.5%</td>
<td>CncC KD: decreased expression of Cyp389B1, Cyp391A1, Cyp392A28. MAF KD: Cyp389B1, Cyp392A28. [19]</td>
</tr>
</tbody>
</table>


31. Deng H: Multiple roles of Nrf2-Keap1 signaling. *Fly* 2014, **8**:7-12.

32. Deng H, Kerppola TK: Regulation of *Drosophila* metamorphosis by xenobiotic response regulators. *PLOS Genetics* 2013, **9**:e1003263.


---

**15. Insecticide resistance mediated through elevated expression of detoxification genes is a constitutive rather than an induced phenomenon. Misra *et al.* show that CncC is constitutively expressed in resistant strains of *Drosophila* and that this constitutively expressed gene causes upregulation of key detoxification genes.**

**18. Kalsi and Palli knocked-down a variety of transcription factors and demonstrated that it is CncC/MAF that controls upregulation of the CYP6BQ genes, previously implicated in pyrethroid resistance in flour beetles but also that ARE elements in the CYP6BQ promoter promote expression in reporter assays co-transfected with CncC and Maf.**

**21. An older but comprehensive study of the role of CncC in *Drosophila*. A molecular biology *tour de force* employing a variety of methods to show how CncC is involved in detoxification and aging.**

**25. Following injection of dsRNA (CncC or GFP) RNASeq was used by Kalsi and Palli to understand the role of CncC in the transcriptomic response in insecticide resistant *Tribolium*. This is the only study to use RNASeq to study the role of CncC/Maf.**

**26. Ingham *et al.* use RNAi knockdown of Maf-S in the Tiassalé strain of *Anopheles gambiae* followed by whole-genome microarrays to identify genes regulated by CnCC/Maf. They then compare the differentially regulated genes to those genes identified as differentially expressed across a number of transcriptomic studies of the insecticide resistance phenotype in mosquitoes.**

**37. Although not a study of the insects or insecticide resistance, Kuosmanen *et al.* utilised a variety of approaches (molecular modelling, analysis of CHIP datasets and protein binding microarrays) to show how sequence variation in AREs can affect NRF2 binding and be associated with disease**
resistance. Such work is now needed for the insecticide resistance phenotype in insect genomic databases.

**47. This excellent and comprehensive review covers experimental and computational approaches for identifying regulatory motifs in genomes. It focuses on more distal cis-regulatory elements which are likely to be more problematical to identify than proximal AREs. Application of these methods to insect species beyond Drosophila may identify other TFs (other than CncC) and their binding sites involved in the insecticide resistance phenotype.