INVESTIGATING THE MOLECULAR BASIS OF RESISTANCE TO 1080 POISON IN THE AUSTRALIAN BUSH RAT (*RATTUS FUSCIPES*).

Kate E. Penwarden

School of Animal Biology, The University of Western Australia

This paper is presented in partial fulfilment of the degree of Zoology Honours. 2010.

Abstract.

Sodium fluoroacetate or "1080" is widely used in Australia to control vertebrate pests. It is especially effective in Western Australia where native fauna is naturally tolerant to 1080. However, recently, resistance in target pest species has been detected in wild rabbit populations as well as in manipulative laboratory breeding experiments, threatening to make pest control by 1080 redundant. This research sought to define the molecular basis for 1080 resistance in a Western Australian vertebrate, the Australian bush rat, *Rattus fuscipes*. To this end, I made use of natural variation in 1080 resistance in different bush rat populations and used a targeted gene approach and manipulative experiments to test specific hypotheses concerning regulatory or nucleotide sequence changes that may confer resistance. Plasma citrate analysis suggested that the one population of rats may not have been as tolerant as predicted, influencing the interpretation of some results. Sequencing of five genes suggested that a mutation at one site in the Cyp3a18 gene was involved in tolerance. Fold change data suggested that regulation of Gstm7, Slc11a1 and possibly Gstm3 genes were involved in differences between the populations that may or may not be involved with tolerance. The expression of Gstm7 and Slc11a1 was also found to be correlated with plasma citrate levels, an indicator of tolerance. This study has concluded that the most likely mechanisms of tolerance in *R. fuscipes* are particular sequence mutations and regulation changes, especially in several 'classic resistance genes' involved in detoxification that have known relationships to resistance of other toxins. We can now look for these same molecular changes in pest species to determine if they share the same mechanism/s underlying tolerance. Once the genomic basis of 1080 tolerance in both native and pest species is understood, strategies to minimise evolution of resistance in pests can be established.

Introduction.

The highly toxic 1080 poison (sodium fluoroacetate) is widely used to control vertebrate pests (Twigg & King 1991). It is especially effective in WA where native fauna has developed varying degrees of tolerance to the toxin as it is naturally found in two native leguminous plant genera (*Gastrolobium* and *Oxylobium*) with which fauna has co-evolved (Mead *et al.* 1985; Twigg & King 1991; Twigg *et al.* 2003; Goncharov *et al.* 2006). Variation in tolerance of native species can be found between classes, species, populations and even within populations, and is believed to reflect a combination of factors relating to exposure to the toxin (Oliver *et al.* 1979; McIlroy 1982; McIlroy 1984; King 1990; Twigg & King 1991; Twigg *et al.* 2002). In WA, the tolerance of native fauna contrasts with the sensitivity of target pest species, such as rabbits, foxes, wild dogs/dingoes and cats (Twigg & King 1991). This resulting specificity is integral to the effectiveness of 1080, in that poison baits can be liberally dispersed in the environment with virtually no risk of poisoning non-target, native species (Oliver *et al.* 1979; King 1990; Twigg *et al.* 2003; Sherley 2004). There is, however, evidence of pests becoming tolerant to 1080 thus threatening to make redundant Australia's most effective method of vertebrate pest control (rabbits -Twigg *et al.* 2002; laboratory studies - Kandel & Chenoweth 1952; Howard *et al.* 1973). The discovery of the molecular basis underlying the evolution of 1080 resistance may hold the key to delaying or preventing widespread resistance in pests.

Three major mechanisms of evolved tolerance are recognised: reducing the availability of the toxin, increasing target site insensitivity or improving the capability of metabolic detoxification systems (Li *et al.* 2007; Puinean *et al.* 2010; see Penwarden 2010 for examples). The corresponding biochemical avenues underlying tolerance to 1080 could be one of the following:

1) Penetration, involving how toxins are taken into the body –the main mechanism in vertebrates;
2) Protein amino acid mutations, which may alter toxin binding affinity and sensitivity of the target site; or
3) Regulatory changes in expression of gene products at the RNA or protein level which may increase efficiency of detoxification mechanisms (Reviewed in Wilson 2001; Li *et al.* 2007; Puinean *et al.* 2010).

In the case of 1080 resistance, there is indisputable evidence that there is a genetic basis. Firstly, some isolated populations have considerably higher tolerance levels than any non-isolated populations (Mead *et al.* 1985; King 1990) and secondly, genetic crosses of tolerant (western) and sensitive (eastern) *R. fuscipes* have resulted in offspring with tolerance mid-way between the two parents (Mead *et al.* 1985; Twigg *et al.* 2003). Therefore, the most likely mechanisms of resistance are genetic: –protein amino acid mutations and regulatory changes in gene products.

The present study aimed to determine the molecular mechanism/s underlying the evolution of 1080 tolerance in Western Australian *Rattus fuscipes.* The Australian bush rat, *R. fuscipes*, was selected as a model species as it is a common and widespread species across southern Australia and populations vary geographically in their susceptibility to 1080, especially between south-eastern and south-western populations (Oliver *et al.* 1979; McIlroy 1984; Mead *et al.* 1985; Twigg & King 1991; Twigg *et al.* 2003). Once the mechanism of tolerance in native species is determined, we will have something for which to test in pest species. Discovering the mechanism of tolerance in a native species is the logical first step in discovering the mechanism in pests as, unlike in pest populations, tolerance levels in native species and variation among populations are relatively well characterised.

Due to *R. fuscipes*' close evolutionary relationship to the widely used laboratory rat, *R. norvegicus*, the extensive genomic resources available for laboratory rats can be readily transferred to *R. fuscipes* (Hinten *et al.* 2007; Berry, O., Rodger, J. 2008 unpublished). I used a targeted gene approach, making use of *R. fuscipes* populations differing in tolerance, sequencing genes in these populations and conducting manipulative experiments to examine regulatory changes in RNA. I made use of *R. fuscipes* individuals from three populations that were predicted to differ in their degree of tolerance to 1080 –a sensitive population, a tolerant population and a "super-tolerant" island population predicted to have much higher tolerance than the tolerant population (See methods). Using super-tolerant rats, I expected to see only regulatory changes that were not involved with a generic stress response that might have been observed in the less-tolerant WA rats.

I hypothesised that by comparing populations of different tolerances, differences in amino acid sequences and gene regulation between populations would be highlighted as possible causes of tolerance. Any changes in gene regulation or nucleotide sequences in tolerant and super-tolerant rats would not be expected to be present in sensitive rats if that change were conferring tolerance.

Methods.

Previous analysis and candidate gene selection

Prior to this study, microarray analysis of gene expression profiles was conducted on RNA extracted from liver tissue of 1080-tolerant WA rats (Berry, O., Rodger, J. 2008 unpublished). That analysis used Affymetrix rat gene chips (GeneChip Rat Genome 230 2.0 Array) to screen for differential gene expression in around 29,000 genes in the liver, where most enzymatic activity likely to confer 1080 tolerance occurs. The rats from which RNA was extracted were subject to one of two experimental treatments: 1) a sub-lethal dose of 1080 poison, or 2) a sterile saline control (n=4 each) and the difference in gene expression between the two treatments was measured.

Microarray analysis allows quick and efficient evaluation of expression of a great number of genes in a single experiment, however it is not as reliable as techniques focussed on a smaller number of genes. Therefore, it is standard practice for microarray results to be verified with real-time quantitative Reverse Transcription PCR (qRT-PCR) looking at a select number of genes (Kammenga *et al.* 2007). Expression arrays identified approximately 200 genes (refined from around 29,000 genes) that exhibited a significant difference in gene expression between the two experimental treatments. Of the approximate 200 genes up- or down- regulated in the presence of 1080, 13 were selected from the top 50 regulated genes for further investigation as the most likely candidates to be involved in 1080 tolerance. These 13 were chosen based on three criteria:

1) The difference in gene regulation between the experimental and control rats identified by gene expression profiling (microarrays);
2) Whether the genes affect systems that are also affected by the action of 1080 poisoning; and

3) Whether the genes had been found in the literature to play a role in tolerance of other toxins.

Full gene names and more detailed explanations of why the 13 genes were selected are presented in Appendix I.

Populations

This study used three populations of *R. fuscipes* with differing tolerances:

1) "WA" (tolerant). Boranup, south-west WA, 34 ^{o}09'S, 115 ^{o}02'E (collected in 2008 for the microarray study), near Pine Creek where tolerance in *R. fuscipes* was documented by Mead *et al.* (1985). (n=8);
2) "WI" (predicted to be more tolerant than WA rats –"super-tolerant"). Woody Island, Recherche Archipelago off Esperance on the southern coast of WA, 33°56'S, 122 ^{o}02'E, in the same island group as Mondrain Island where bush rats are known to have amongst the highest levels of 1080 tolerance in vertebrates (Mead *et al.* 1985; Twigg & King 1991). (n=8);
3) "SA" (sensitive). South Australian rats from which liver tissue was obtained from the Evolutionary Biology Unit of the South Australian Museum, Adelaide. (n=4).

Fluoroacetate is highly toxic to most animals, the approximate lethal dose for most un-adapted mammals being under 2 mg/kg body weight (Atzert 1971 cited by Twigg & King 1991). South-western Australian *R. fuscipes* have a variable LD_{50} (median lethal dose) ranging from 1-80 mg/kg (King 1990) and are much more tolerant than eastern Australian individuals of the same species (LD_{50} around 1 mg/kg, similar to that of introduced *Rattus* species, including *R. rattus* and *R. norvegicus*) (Oliver *et al.* 1979; McIlroy 1982; McIlroy 1984; King 1990).

The mitochondrial COI (cytochrome c oxidase subunit I) gene in the Woody Island rats and four of the WA rats was sequenced to confirm that all the Woody Island rats were in fact *R. fuscipes* (primers from Folmer *et al.* 1994 used, see Appendix II). A neighbour-joining phylogenetic tree was constructed, comparing WI sequences with existing sequences from two introduced rats and one WA rat species to verify that all WI rats were *R. fuscipes* (Fig 1).

Experimental dosing with 1080 poison

In August 2010, eight *R. fuscipes* (four male, four female) were trapped using Elliot traps on Woody Island and brought back alive to UWA. A Regulation 17 Licence to take fauna for scientific purposes (No. SF007600) and a Regulation 4 Authority to enter CALM land and/or waters (No. CE002941) were obtained for this field work. Rats were kept for 13 days before commencing the experiment. Animals were collected and maintained under strict conditions in accordance with animal ethics (RA 3/100/956).

Four rats received an intraperitoneal injection with a sub-lethal dose of 1080 (10 mg/kg body weight). Three animals received an equivalent volume of sterile physiological saline (0.9% NaCl) as a procedural control. After dosing, rats were housed in subdued lighting and provided with shelter. Rats were euthanased two hours post-injection by isofluorane anaesthesia followed by cervical dislocation. Blood was taken via cardiac puncture for plasma citrate analysis and liver tissue was taken and stored in RNAlater at -80°C. This procedure, including the plasma citrate analysis, replicated the protocol followed for WA rats used for microarray analysis (Berry, O., Rodger, J. 2008 unpublished). Blood plasma was separated and analysed for plasma citrate by the DAFWA Animal Health Laboratories.

RNA isolation and cDNA synthesis

In order to look at expressed genes, total RNA was extracted from *R. fuscipes* liver tissue from SA (n=4) and WI (n=7) rats using TRIZOL^{TM} Reagent according to the manufacturer's protocol (Invitrogen, Carlsbad, CA). RNA left over from the microarray analysis (also extracted using TRIZOL^{TM} Reagent) in 2008 was used for the WA rats (n=8) as well as two *R. norvegicus* wistar rats used previously as a control. To exclude genomic contamination, RNA from all rats was DNase I treated using RQ1 RNase-free DNase (Promega, Madison, WI) and then steps 2 onwards of Zymo Research DNA-free RNA kit ™ were performed in accordance with manufacturer's protocol (RNA eluted twice in 10µl) (Zymo Research, Orange, CA). 2µg diluted DNase I treated mRNA from each of the 21 rats was transcribed into first-strand cDNA using avian mycloblastosis virus (AMV) reverse transcriptase, an oligo (dT) primer and dNTPs according to manufacturer's protocol (Promega, Madison, WI). The resulting cDNA was used as a template for sequencing PCR and qRT-PCR amplifications. Nucleic acid concentration of the cDNA was quantified by a NanoDrop® ND-1000 Spectrophotometer (Biolab Nanodrop Technologies, Wilmington, DE).

Primer design for sequencing and qRT-PCR

Sequencing and qRT-PCR primers were designed based on the *R. norvegicus* mRNA reference sequences (see appendix II and III) for each of the 13 genes plus a housekeeping gene for qRT-PCR [see 'regulatory changes' section of methods], plus aconitase for sequencing, using the NCBI website (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). Sequencing primers were designed to amplify the longest possible cDNA fragment, with a minimum, optimum and maximum primer melting temperature of 53°C, 55°C and 63°C respectively. For qRT-PCR primers, amplicon length was limited to between 100 and 300 base pairs and primer melting temperatures were set to a minimum, optimum and maximum of 53°C, 55°C and 57°C respectively. All primers spanned an exon-exon junction to avoid amplifying genomic DNA (gDNA) except for the Dusp6 sequencing primer pair, for which there were no possible primers that spanned this type of junction. GC content was set to between 50% to 65% for all primers. To sequence aconitase, for which I aimed to sequence the whole gene, six primers were required to amplify its length. These primers were designed using Primer3 (Rozen & Skaletsky 2000) to regions conserved between *R. norvegicus* and *Mus domesticus.* A complete list of qRT-PCR and sequencing primer sequences can be found in Appendix II and III. Although *R. fuscipes* is closely related to *R. norvegicus*, on which primers were designed (Hinten *et al.* 2007), sequence differences are present and rendered some sequencing and qRT-PCR primers inactive.

Nucleotide sequence changes: Sequencing

Mitochondrial aconitase (Aco2) was sequenced in all three populations as the toxin interferes with the function of this enzyme. 1080 shuts down metabolism by entering the Krebs cycle and acting as a "suicide substrate" for aconitase (see Penwarden 2010), effectively 'stealing' the enzyme required for the completion of the cycle, and this results in depletion of energy, halting all active cellular processes, accompanied by an excess of citrate (Clarke 1991; Goncharov *et al.* 2006). Attempts were made to obtain sequences in the WA and SA rats for all 13 target genes plus aconitase.

The adequate amplification of correct products was achieved for five genes (Aco2, Cyp3a18, Gstm7, Gstm2, Car3). PCR amplification reaction mixtures used Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, CA) and gene-specific oligonucleotide primers in a 25µl reaction. PCR was performed using a Mastercycler® pro Thermal Cycler (Eppendorf, Westbury, NY). To identify correct DNA products, 4µL aliquots of

the PCR reaction mixture were electrophoresed on agarose gel (Fig 2). Amplification of correct product was also verified by alignment with each gene in *R. norvegicus*.

Sequencing reactions were carried out in forward and reverse directions with Big Dye versions 3.1 dye terminators and sequences were obtained with an ABI3730XL (Applied Biosystems, Foster City, CA) sequencer. Sequences were checked by eye with SEQUENCHER® software (Gene Codes Corporation, Ann Arbor, MI) (see appendix IV) before being exported to *MEGA* version 4 (Tumara *et al.* 2007) for alignment via the ClustalW algorithm (Thompson *et al.* 1994). For each gene, the sequences of all the rats, plus a GenBank reference sequence for the gene in *R. norvegicus*, were aligned, studied for sequence variation between rats, then translated into proteins to identify which nucleotide mutations resulted in amino acid substitution (i.e. non-synonymous mutations).

Regulatory changes (RNA): Gene-specific confirmation by qRT-PCR

qRT-PCR measures the abundance of specific gene products generated during each cycle of the PCR process (Heid *et al.* 1996). qRT-PCR amplified and examined the expression profile of 13 selected genes and one housekeeping gene from 17 samples (WA rats -dosed [n=4], control [n=4]; *R. norvegicus* wistar [n=2]; and WI rats –dosed [n=4], control [3]).

In order to correct for variation between samples in qRT-PCR, normalization is conventionally performed by including a housekeeping/reference gene that will not differ in expression across different experimental conditions to act as a control (Pohjanvirta *et al.* 2006). Phosphoglycerate kinase 1 (Pgk1) was trialled and selected as the housekeeping gene as, out of the three genes trialled, it had the smallest change in expression between control and 1080 treated rats after amplification by qRT-PCR (threshold cycle mean =0.97). Pgk1 and two other genes were coamplified on one plate by qRT-PCR each time. Following standard procedures, primers were optimised and samples and controls were run in duplicate in two to four separate runs to reduce variability.

Each qRT-PCR 20µl reaction contained the following components: 4 µl of 5 x HOT FIREPol® EvaGreen® qPCR Mix Plus (ROX), 4µl cDNA from one rat and 2.5µl each of forward and reverse primer as recommended by the manufacturer (Solis BioDyne,

Tartu, Estonia). Two no template controls for each gene on the plate were also included. These reactions were performed using an iQ™5 Multicolor Real-Time PCR Detection System machine (Bio-Rad, Richmond, CA). Cycling conditions were as follows: 15 min denaturing at 95°C, 40 cycles of 15 sec at 95°C, 20 sec at 60°C, 20 sec at 72°C, followed by 1 min at 95°C, 1 min at 55°C then 81 cycles of 10 sec at 55°C for the melt curve (0.5°C temperature step-up). A final melt-curve step was performed by the qRT-PCR detection machine to confirm the absence of amplification of non-specific products. At least two runs in duplicate were obtained for all genes except Ppp1r3b and Gck which were not amplified successfully.

qRT-PCR data analysis

Gene regulation was examined in two ways: 1) baseline gene expression and 2) fold change. Baseline gene expression is the baseline expression normalised as a ratio of the housekeeping gene. Baseline expression was analysed because the different expression levels of genes are clearly reflected at this stage before calculating fold change. Fold change is the proportional increase from the level in the control to the level in the treated animal and therefore cancels out any difference in baseline expression between populations so that the impact of treatment (i.e. 1080 response) can be more easily seen.

1) Baseline gene expression: For each rat, the mean C_T (threshold cycle, the cycle at which the fluorescence of the sample crosses the threshold) of the replicates for each gene was used to calculate the expression level of that gene. These means were then $1/(2^{\wedge} C_T)$ treated and the ratios of the 1080 treated data to the housekeeping gene data were calculated (as in Dong *et al.* 2009). For each gene an average of treated and control animals was taken. A two-way ANOVA assuming equal variances looked for significant differences between treatment and population (WA and WI) in each gene. A sequential Bonferroni correction set the significance threshold at P =0.05/(n-1) for each gene when genes were ranked from most significant to least significant and "n-1" is the number of genes ranked below and including that gene. Scheffe post-hoc tests were used to identify the specific treatment groups exhibiting differences of gene expression. This test was used as it is the most stringent post-hoc test, doesn't need an equal number of animals in each treatment and is robust to non-homogeneity of variances and non-normality. Power analyses were conducted online (Retrospective power calculations, http://statpages.org/postpowr.html) to determine how big a difference could be detected and considered significant with our sample size of four animals per treatment.

2) Fold change: For each gene, the delta-delta C_T formula, $2^{-\Delta\Delta C_T}$ (Pfaffl 2001) was used on each treatment animal to calculate fold change. Then, an average was taken over all four treated animals to get an average measure of fold change in each gene. Fold changes in WA and WI rats were plotted alongside fold change data from the microarrays (WA rats). An unpaired t-test grouped by population looked for significant differences in fold change between populations, as these could be associated with the difference in tolerance between populations. Fold change in the WA and WI rats was examined for correlation with the degree of 1080 tolerance among individual rats exposed to 1080, as indicated by plasma citrate concentration.

Results.

Plasma citrate analysis

The citrate levels between the treatments and populations were compared by a two-way ANOVA assuming equal variances. There was a significant difference in the citrate levels between treatments (df =1, 11, F =19.059, P =0.0011) but no difference between populations and no significant interaction of treatment and population (Fig 3).

Nucleotide sequences

After COI sequencing confirmed that the WI rats were all *R. fuscipes* (Fig 1), sequencing of Aco2 (3,287 base pairs), Cyp3a18 (2,005-bp), Gstm7 (1,208-bp), Gstm2 (657-bp) and Car3 (802-bp) revealed nucleotide and amino acid mutations between the populations (Table 1). Differences only between *R. norvegicus* and *R. fuscipes* were ignored but are shown in appendix V with a full list of differences in each gene. With the exception of COI in Fig 1, full sequences of genes are not shown due to space restrictions and small number of mutations. A total of 5, 1, 0, 1 and 1 nucleotide mutations were identified within *R. fuscipes* in Aco2, Cyp3a18, Gstm7, Gstm2 and Car3 respectively (Table 1). Of these, the majority were synonymous (silent) mutations. Two non-synonymous (missense) mutations were identified in Cyp3a18 (site 1256) and Aco2 (site 2324) (Table 1). Non-synonymous mutations were distributed according to population (i.e. tolerance level) at the site in Cyp3a18 but not at the site in Aco2. Some of the synonymous mutations also exhibited evidence of population structure e.g. Aco2

site 906 and 1266, and Gstm2 site 363 (Table 1). At the sites of non-synonymous mutations, sequence variation *within* WA or WI rats could not be accounted for by tolerance of each individual measured by plasma citrate level.

Regulation

Baseline gene expression analysis:

The baseline gene expression was compared in the WA and WI control and treated rats (Fig 4). Overall, gene expression was generally higher in WI rats than WA rats. A two-way analysis of variance assuming equal variance revealed significant differences only between populations in the expression of Lpin1, Gstm2, Dusp6 and Gstm3 (Table 2). The expression of Lpin1, Gstm2 and Gstm3 were significantly higher in the WI population while the expression of Dusp6 was significantly higher in the WA population (Fig 4). Using the WA control rats as a base for the expression of each gene, the significant differences in baseline gene expression between the populations were large, with some genes being up-regulated compared to the WA control rats in one population and down-regulated compared to WA control rats in the other (Fig 4). The same situation (that is, different expression between populations) was seen for fold change, where expression was measured as the increase above the control group (Fig 5).

For the genes that had significantly different expression indicated by the two-way ANOVA, Scheffe post-hoc tests identified the specific treatment groups exhibiting differences of gene expression (Table 3). All differences were between populations rather than between treatments (Table 3). The differences in raw gene expression between populations are on such a comparatively large scale that the differences within populations (i.e. treatment) are concealed.

To give an indication of how well the sample size of four animals per treatment was representing real differences between groups, power analyses were conducted. Of the insignificant differences, the power analyses showed that with a sample size of four animals per treatment, the minimum detectable difference was large, at least 86.7% of the mean, and ranged between genes between that and 702,555.6% of the mean.

Fold change:

Fold change (the change in expression from the control to the treatment) obtained from analysis of qRT-PCR was compared between the WA and WI rats alongside the fold

change values obtained from microarrays of the same WA rats (Fig 5). qRT-PCR confirmed the microarray results for all but two genes (Dusp6 and Tbx3). Unpaired t-tests grouped by population indicated that fold change was significantly different between populations for Gstm7 (df =6, t = -2.730, P =0.0342) and Slc11a1 (df =6, t =2.552, P =0.0433), and there was a strong trend in Gstm3 (df =6, t = -2.427, P =0.0514).

The fold changes in the genes of the two populations were plotted against plasma citrate levels to look for a correlation between the amount of citrate in the plasma and the regulation of genes that may be associated with citrate levels and tolerance (Fig 6). In all but two genes, no obvious pattern was apparent and the two factors had very weak correlation. However, Gstm7 and Slc11a1 had strong relationships between fold change and citrate level (Fig 6). Correlations were significant in WI rats in Gstm7 (R =0.989) and Slc11a1 (R =0.988). When baseline gene expression for treated rats (rather than fold change) was plotted against citrate levels, results were very similar to those shown in Fig 6.

Discussion.

Plasma citrate analysis

Plasma citrate levels reflect sensitivity to fluoroacetate and are an accepted measure of comparing tolerance within and between phylogenetically similar groups (Oliver *et al.* 1979; Twigg & King 1991). It is known that SA rats have significantly higher plasma citrate levels than WA rats when dosed with an equivalent volume of 1080 (Mead *et al.* 1985). We are less sure about citrate differences between the Western Australian populations. In the WA and WI rats, the difference in citrates found between treatments was to be expected as an excess of citrate is one of the symptoms of 1080 poisoning. There was also variation found between individual rats of the same treatment in the same population. Surprisingly, the plasma citrate analysis of WA and WI rats revealed no significant difference between populations. This suggested that either the Woody Island rats were not as tolerant as predicted based on its proximity with the super-tolerant Mondrain Island or the mechanism of tolerance does not reduce the excess of citrate, instead simply allows the animal to deal better with the excess. Studying citrate

levels in *R. fuscipes* from Mondrain Island (known to be super-tolerant; Mead *et al.* 1985) could shed light on this issue. If super-tolerant Mondrain Island rats had similar citrates to WA populations, it could be concluded that, in this case, plasma citrate levels did not reflect tolerance. If, on the other hand, they had citrate levels much lower than the WA population, that would support the argument that citrate does predict tolerance and that the tolerance of the WI population was not significantly different from the WA population. A greater sample number from each population (as suggested by the power analysis) would also show significant differences in expression more confidently.

As we are now unsure of the tolerance of the Woody Island rats, this affects how ensuing analyses are approached as well as what conclusions can be drawn from the results. The most likely contender for amino acid mutation conferring tolerance (Cyp3a18) had sequences from SA and WA only, not WI, and therefore the conclusions drawn from this will not be affected by uncertain tolerance of the WI population. However, the possibility that WI rats may have the same tolerance as WA rats will have to be considered when looking at regulation differences between the WA and WI populations.

While differences in tolerance between the two populations are unclear, the two populations regulated differently genes involved in detoxification suggesting that they may have different detoxification responses which could otherwise imply different tolerance levels between the two rat populations. However, in this case, this is not a confident interpretation as it is likely these differences in regulation are unrelated to tolerance. The baseline gene expression of some genes was significantly different between populations and strong relationships between citrate levels and expression of two genes (Gstm7 and Slc11a1) were opposed (positive/negative) in the two populations. For these genes, the correlation between citrate and expression was stronger for the WI rats. We can therefore conclude that, although we are unsure of tolerance differences between the populations, there are notable differences in their response to 1080.

Nucleotide sequences

Firstly, it is important to note that genes chosen for sequencing were primarily selected based on how they were regulated (microarray expression data), rather than on the

likelihood of these genes having sequence mutations, as this is not measured by microarrays. Of the five genes sequenced, non-synonymous mutations at site 1256 in Cyp3a18 (2,005 base pairs) and site 2324 in Aco2 (3,287 base pairs) were the most likely to be involved in tolerance. At the 1256 site in Cyp3a18, the tolerant WA population is associated with a G base which is absent in sensitive rats, thus the amino acid substitution that results could be conferring tolerance in this population. It should be noted that these changes were seen in a small sample size as sequences from only two tolerant rats covered this site. The 3-dimensional protein structure of Cyp3a18 is unknown so it is unclear whether the site where a mutation occurs is an important part of the protein (e.g. binding site, swivel etc.) (NCBI 2010). Sequence changes in cytochrome P450s (e.g. Cyp3a18) have been found to be the cause of multiple cases of toxin resistance (see 'Comparison with results of other resistance mechanism studies' below for examples) and could play a similar role here. As well as the non-synonymous mutation in Cyp3a18, some of the synonymous mutations also exhibited evidence of population structure, e.g. Aco2 site 906 and 1266 and Gstm2 site 363. The evidence of population structure in synonymous mutations makes the population structure found in non-synonymous mutations more plausible. By contrast, at the 2324 site in the Aco2 gene, the non-synonymous mutation was found in both tolerant and sensitive populations and therefore is not likely to be causing tolerance. The 3-dimensional protein structure in Aco2 is better understood, however the 2324 site is not located at a significant site in the protein (NCBI 2010). Again, the inclusion of more samples will strengthen these conclusions.

The presence of PCR incorporation errors is likely for the gene sequences obtained in this study. For Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, CA), the highest error rate applicable to this study was 7.2×10^{-5} errors/bp (Ling et. al. 1991). Therefore, based on the longest amplicon length, there could have been up to 0.11 errors per PCR reaction. The issue of errors could be solved by performing multiple separate PCRs on a number of identical reaction replicates, as the chance of all replicates having the same error in the sequence is low.

Regulation

Fold change data is more representative of the differences in regulation of genes than baseline gene expression, which had large inter-population effects swamping any possible effect of treatment. Baseline gene expression identified regulation differences

between populations but proved less relevant in discovering tolerance mechanisms and thus is not discussed further. Fold change measured by qRT-PCR in the WA rats confirmed the results of the microarrays except for two genes (Dusp6 and Tbx3 which were regulated in the opposite direction). Fold change significantly differed between populations for Gstm7 (WA down, WI up) and Slc11a1 (WA up, WI down) and there was a strong trend in Gstm3 (WA down, WI up). If the citrates do in fact accurately reflect similar tolerance of the two populations, these regulatory differences would be a result of differing responses to 1080 that do not affect tolerance. Considering that we expect all native species to share the same resistance mechanism/s, it is highly unlikely that the two populations would have different mechanisms of resistance.

The magnitude of fold change measured by qRT-PCR was consistently lower than fold change in the same genes measured by microarray analysis. Surprisingly, this result does not reflect the reported underestimation of fold change by microarrays compared to qRT-PCR (Yuen *et al.* 2002). In addition, the magnitude of fold change by qRT-PCR found in this study is lower than that seen in most other examples where regulation, measured by qRT-PCR is the mechanism providing resistance (e.g. Ding *et al.* 2003 – GSTs in DDT-resistant *Anopheles gambiae* (mosquito); Nikou *et al.* 2003 –cytochrome P450 in pyrethroid-resistant *Anopheles gambiae*; McGrath *et al.* 2005 –plant disease resistance conferred by transcription factors; Vontas *et al.* 2005 –multiple genes suggested to be involved in insecticide resistance in *Anopheles gambiae*).

General conclusions

Plasma citrate was variable within treatments and between treatments, however there was no difference between populations. It is likely that the WI rats are not significantly more tolerant than WA rats, however the two populations had different regulatory responses to 1080. The nucleotide sequence data point to a mutation at the 1256 site in Cyp3a18 as being the most likely mutation to be conferring tolerance. More samples from tolerant and sensitive populations are required to strengthen this conclusion. Fold change data suggested that regulation of Gstm7, Slc11a1 and possibly Gstm3were different between populations. As we cannot be sure that the two populations compared had different tolerance levels, it cannot be concluded with any certainty that the regulation of these genes are likely to be involved in tolerance.

As with all genes suggested by various methods in this study as being possibly involved in tolerance, it has not been shown that the regulation of these genes is the *cause* of tolerance, only likely to be involved in a tolerance response. However, the changes could also simply be genetic divergence over a time of isolation and due to other selection pressures or gene flow with surrounding populations. Also, as citrates indicated that the two populations do not differ in tolerance, the possibility of regulation changes being stress responses unrelated to tolerance cannot be ruled out. Unfortunately, differences due to stress, differences involved with a tolerance response and differences simply due to genetic drift could not be distinguished with the present information. Using populations that are known to be "super-tolerant" (as was attempted here) may be able to rule out stress responses as more tolerant rats should be less stressed by the effects of poisoning. Nevertheless, most of the genes implicated by sequencing and qRT-PCR are 'classic resistance genes' involved in detoxification suggesting they are likely to play a role in 1080 resistance (Cyp3a18, Gstm7, Gstm3; see appendix I). Future studies of protein expression and biochemical assays, as well as similar studies with a greater number of samples and populations with known tolerance, will resolve whether the mechanisms suggested here are involved in tolerance of 1080 in *R. fuscipes*.

***Comparison with results of other resistance mechanism studies*.**

The mechanisms implicated in this study (amino acid mutation in one gene and regulation of three genes, most of which are common resistance genes involved in detoxification) are similar to other mechanisms found in the literature on toxin resistance. The regulation of Slc11a1, implicated in this study, has not been found to have a role in resistance to other toxins, although an amino acid mutation in the gene causes resistance to *Salmonella typhimurium* in mice (White *et al.* 2005). By contrast, the involvement of cytochrome P450s (e.g. Cyp3a18) and glutathione S-transferases (GSTs e.g. Gstm7 and Gstm3) in conferring resistance to toxins, both by single amino acid mutations in the proteins as well as regulation of the gene products as was suggested by this study, is common. For example, amino acid mutations in these classic resistance genes conferring resistance to: DDT and pyrethroids (Pittendrigh *et al.* 1997; Brengues *et al.* 2003); Bt toxins (Tabashnik *et al.* 1997; Griffitts *et al.* 2001); organophosphates and carbamates (Wilson 2001); cyclodiene insecticides (ffrench-constant *et al.* 1990; ffrench-constant & Roush 1991); warfarin (Hermodson *et al.* 1969; Kohn *et al.* 2000); and changes in regulation of classic resistance genes conferring

tolerance to: DDT and pyrethroids (GSTs -Ottea & Plapp 1984; Grant & Hammock 1992; Prapanthadara *et al.* 1993; P450s - Cuany *et al.* 1990; Brun *et al.* 1996); organophosphates (P450s -Houpt *et al.* 1988; P450s and GSTs -Li *et al.* 2007); carbamates (P450s -Woo Cha *et al.* 2000); neonicotinoid insecticides (P450 –Puinean *et al.* 2010); warfarin and bromodiolone (P450s - Markussen *et al.* 2008).

Resistance management

The evolution of resistance in target pest species in Australia could make current pest control by 1080 redundant, with dire consequences for biodiversity conservation and agricultural production (Sherley 2004). Understanding the molecular basis of resistance can assist strategising to avoid evolved resistance. Strategies may include revising baiting patterns and doses or altering the toxin itself. One of the best examples of management of toxin resistance in pests is resistance in insects to Bt toxins in transgenic crops. Theories and strategies extrapolated from this example (such as 'gene stacking' and the high dose/refuge (HDR) strategy which apply fundamental evolutionary principles –see Liu & Tabashnik 1997; Tabashnik *et al.* 1997; Jurat-Fuentes *et al.* 2003) could give insights into, and be altered to manage, evolved resistance to 1080. However, in the case of 1080, resistance in pests must be managed whilst not affecting resistance in native species, to conserve the specificity of the toxin. Therefore, whether pests and native species share the same resistance mechanisms has implications on what strategies can be used to manage resistance in pests.

Benefits of this research and its place in future resistance management.

There is a limited understanding of the genomic basis of 1080 tolerance in both native and pest species. Therefore, this study aimed to broaden understanding of these mechanisms of tolerance in a native animal with the view to extrapolate this information to pest species when attempting to identify tolerance mechanisms in pests. The preliminary expression profiles (microarrays; Berry & Rodger 2008, unpublished) along with this study are the first investigations into molecular mechanisms that may confer tolerance to 1080. An amino acid mutation and possibly regulatory changes to several genes mainly involved in detoxification were suggested by this study to be the most likely to play a substantial role in conferring tolerance to 1080. As changes in detoxification have been implicated here as the tolerance mechanism in *R. fuscipes*, these detoxification genes should be the first option investigated in pest species when trying to determine the mechanism underlying evolution of 1080 tolerance in pests.

Once the genomic basis of 1080 tolerance in both native and pest species is understood, it potentially offers both a means of monitoring the process as well as adapting management strategies so that the opportunities for resistance to evolve are minimised.

Acknowledgements

Firstly, I would like to express my immense gratitude to my supervisors Dr. Jenny Rodger (UWA), Dr. Oliver Berry and Dr. Tom Walsh (CSIRO) whose assistance with laboratory procedures and writing is greatly appreciated, and without whose involvement, this project would not have been possible. Special thanks go to Dr. Oliver Berry who also came all the way to Woody Island to help me catch rats. I also thank the members of my marking panel, comprising Dr. Jenny Rodger, Dr. Dale Roberts and Dr. David Hunt, who gave much constructive feedback on my literature review and final thesis, as well as honours coordinator Dr. Leigh Simmons. I am very appreciative of CSIRO for the use of their facilities where I carried out most of the work for this study and for generous funding of the project. I am very grateful also for the funding provided by the School of Animal Biology, UWA.

I especially want to extend my thanks to Steve Hoddy and Heidi Norwood from MacKenzies Cruises, Esperance for kindly facilitating our access to Woody Island and for the wonderful wildlife tour on the way. Special thanks go to Yvette Hitchen who carried out the COI sequencing and was always happy to help when I had so many questions. Many thanks to Gary Martin and Simone Ross who helped dose the rats with 1080 and euthanase them. Also to Malcolm McGrath who performed the plasma citrate analysis at DAFWA. Thank you to South Australian Museum for sending South Australian rat liver tissue for use in this study. I gratefully acknowledge many individuals at the School of Animal Biology including Roberta Bencini, Brian Chambers, Rick Roberts and Jason Kennington for their assistance in things such as the loan of traps, setting up a CT room and assistance with statistical analysis. I also wish to thank the other honours students for their moral support and the rats involved in this study who gave their lives for science.

References

Brengues, C., Hawkes, N. J., Chandre, F., Mc Carroll, L., Duchon, S., Guillet, P., Manguin, S., Morgan, J. C., Hemingway, J. 2003. Pyrethroid and DDT cross-resistance in *Aedes aegypti* is correlated with novel mutations in the voltage-gated sodium channel gene. *Med. Vet. Entomol.*, 17, 87-94.

Brun, A., Cuany, A., Le Mouel, T., Berge, J.B., Amichot, M. 1996. Inducibility of the *Drosophila melanogaster* cytochrome P450 gene, *CYP6A2*, by phenobarbital in insecticide susceptible or resistant strains. *Insect Biochem. Mol. Biol.* 26, 697–703.

Clarke, D.D. 1991. Fluoroacetate and fluorocitrate: mechanism of action. *Neurochem. Res.*, 16, 1055-1058.

Cuany, A., Pralavorio, M., Pauron, M., Berge, J.B., Fournier, D., Blais, C., Lafont, R., Saluan, J.P., Weissbart, D., Larroque, C., Lange, R. 1990. Characterization of microsomal oxidative activities in a wild-type and in a DDT resistant strain of *Drosophila melanogaster*. *Pestic.Biochem. Physiol.* 37, 293–302.

Ding, Y., Ortelli, F., Rossiter, L.C., Hemingway, J., Ranson, H. 2003. The *Anopheles Gambiae* glutathione transferase supergene family: annotation, phylogeny and expression profiles. *BMC Genomics*, 13, 35-51.

Dong, S., Mulders, W.H.A.M., Rodger, J., Robertson, D. 2009. Changes in neuronal acitivity and gene expression in guinea-pig auditory brainstem after unilateral partial hearing loss. *Neuroscience,* 159, 1164-1174.

ffrench-constant, R.H., Roush, R.T., Mortlock,D., Dively, G.P. 1990. Isolation of dieldrin resistance from field populations of *D. melanogaster* (Diptera: Drosophilidae). *J. Econ. Entomol.* 83, 1733–37.

ffrench-constant, R.H. & Roush, R.T. 1991. Gene mapping and cross-resistance in cyclodiene insecticide-resistant *Drosophila melanogaster* (Mg.). *Genet. Res.*, 57, 17-21.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.,* 3, 294-299.

Goncharov, N.V., Jenkins, R.O., Radilov, A.S. 2006. Toxicology of fluoroacetate: a review, with possible directions for therapy research. *J. appl. Toxicol.*, 26, 148-161.

Grant, D.F., Hammock, B.D. 1992. Genetic and molecular evidence for a trans-acting regulatory locus controlling glutathione stransferase-2 expression in *Aedes aegypti*. *Mol. Gen. Genet.* 234, 169–176.

Griffitts, J.S., Whitacre, J.L., Stevens, D.E., Aroian, R.V. 2001. *Bt* toxin resistance from loss of a putative carbohydrate-modifying enzyme. *Science,* 293, 860–864.

Heid, C.A., Stevens, J., Livak, K.J., Williams, P.M. 1996. Real time quantitative PCR. *Genome Res*., 6, 986-994.

Hermodson, M.A., Suttie, J.W., Link, K.P. 1969. Warfarin metabolism and vitamin K requirement in the warfarin-resistant rat. *Am. J. Physiol*., 217, 1316-1319.

Hinten, G.N., Rossetto, M., Baverstock, P.R. 2007. New microsatellite markers for the bush rat, *Rattus fuscipes greyii*: characterisation and cross-species amplification. *Mol. Ecol*., 7, 1254-1257.

Houpt, D.R., Pursey, J.C., Morton, R.A. 1988. Genes controlling malathion resistance in a laboratory-selected population of *Drosophila melanogaster*. *Genome* 30, 844-853.

Howard, W.E., Marsh, R.E., Palmateer, S.D. 1973. Selective breeding of rats for resistance to sodium monofluoroacetate. *J. Appl.Ecol.,* 10, 731-737.

Jurat-Fuentes, J.L., Gould, F.L., Adang, M.J. 2003. Dual resistance to *Bacillus thuringiensis* Cry1Ac and Cry2Aa toxins in *Heliothis virescens* suggests multiple mechanisms of resistance. *Appl. Environ. Microbiol.*, 69, 5898-5906.

Kandel, A. & Chenoweth, M.B. 1952. Tolerance to fluoroacetate and fluorobility rate in rats. *J. Pharmacol. Expt. Ther.*, 104, 248-252.

Kammenga, J.E., Herman, M.A., Joop Ouborg, N., Johnson, L., Breitling, R. 2007. Microarray challenges in ecology. *TREE*, 22, 273-279.

King, D.R. 1990. *1080 and Australian Fauna.* Agriculture Protection Board Technical Series No. 8.

Kohn, M.H., Pelz, H-J., Wayne, R.K. 2000. Natural selection mapping of the warfarin-resistance gene. *Proc. Natl. Acad. Sci. USA*, 97, 7911-7915.

Li, X., Schuler, M.A., Berenbaum, M.R. 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu. Rev. Entomol*., 52, 231-253.

Ling, L.L., Keohavong, P., Dias, C., Thilly, W.G. 1991. Optimization of the polymerase chain reaction with regard to fidelity: modified T7, Taq, and Vent DNA polymerases. *PCR Methods Appl*., 1, 63-69.

Liu, Y-B. & Tabashnik, B.E. 1997. Experimental evidence that refuges delay insect adaptation to *Bacillus thuringiensis*. *Proc. R. Soc. Lond. Ser. B,* 264,605-610.

Markussen, M.D., Heiberg, A., Fredholm, M., Kristensen, M. 2008. Identification of cytochrome P450 differentiated expression related to developmental stages in bromodiolone ressitance in rats (*Rattus norvegicus*). *Pest. Biochem. Physiol.*, 91, 147-152.

Mead, R.J., Oliver, A.J., King, D.R., Hubach, P.H. 1985. The co-evolutionary role of fluoroacetate in plant-animal interactions in Australia. *Oikos*, 44, 55-60.

McGrath, K.C., Dombrecht, B., Manners, J.M., Schenk, P.M., Edgar, C.I., Maclean, D.J., Scheible, W-R., Udvardi, M.K., Kezan, K. 2005. Repressor- and activator-type ethylene response factors functioning in Jasmonate signaling and disease resistance identified via a genome-wide screen of *Arabidopsis* transcription factor gene expression. *Plant Physiol.*, 139, 949-959.

McIlroy, J.C. 1982. The sensitivity of Australian animals to 1080 poison. IV. Native and introduced rodents. *Aust. Wildl. Res.*, 9, 505-517.

McIlroy, J.C. 1984. The sensitivity of Australian animals to 1080 poison. VII. Native and introduced birds. *Aust. Wildl. Res.,* 11, 373-385.

NCBI (National Center for Biotechnology Information). Available from: < http://www.ncbi.nlm.nih.gov/>. [October 2010].

Nikou, D., Ranson, H., Hemingway, J. 2003. An adult-specific CYP6 P450 gene is overexpressed in a pyrethroid-resistant strain of the malaria vector, *Anopheles gambiae*. *Gene*, 318, 91-102.

Oliver, A.J., King, D.R., Mead, R.J. 1979. Fluoroacetate tolerance, a genetic marker in some Australian mammals. *Aust. J. Zool.,* 27, 363-372.

Ottea, J.A., Plapp, F.W. 1984. Glutathione S-transferase in the house fly: biochemical and genetic changes associated with induction and insecticide resistance. *Pest. Biochem. Physiol.* 22, 203–208.

Penwarden, K.E. 2010. The evolution of tolerance to 1080 poison in Western Australian animals: genomic basis and implications for pest control. Literature review submitted in partial fulfilment of the degree of Zoology Honours, School of Animal Biology, The University of Western Australia.

Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, 29, 2002-2007.

Pittendrigh, B., Reenan, R., ffrench-Constant, R.H., Ganetzky, B. 1997. Point mutations in the *Drosophila* sodium channel gene *para* associated with resistance to DDT and pyrethroid insecticides. *Mol. Gen. Genet.* 256, 602–610.

Pohjanvirta, R., Niittynen, M., Lindén, J., Boutros, P.C., Moffat, I.D., Okey, A.B. 2006. Evaluation of various housekeeping genes for their applicability for normalization of mRNA expression in dioxin-treated rats. *Chem. Biol. Interact.,* 160, 134-149.

Prapanthadara, L.A., Hemingway, J., Ketterman, A.J. 1993. Partial purification and characterization of glutathione S-transferases involved in DDT resistance from the mosquito *Anopheles gambiae*. *Pesticide Biochem. Physiol.*, 47, 119-133.

Puinean, A.M., Foster, S.P., Oliphant, L., Denholm, I., Field, L.M., Millar, N.S., Williamson, M.S., Bass, C. 2010. Amplification of a cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in the aphid *Myzus persicae. PLOS*, 6, e1000999.

Rozen, S. & Skaletsky, H. J. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology (Ed. by S. Krawetz & S. Misener), pp 365-386. New Jersey, Humana Press.

Sherley, M. 2004. The traditional categories of fluoroacetate poisoning signs and symptoms belie substantial underlying similarities. *Toxicol. Lett.*, 151, 399-406.

Tabashnik, B.E., Liu, Y-B., Finson, N., Masson, L., Heckel, D.G. 1997. One gene in diamondback moth confers resistance to four *Bacillus thuringiensis* toxins. *Proc. Natl. Acad. Sci. USA*, 94, 1640-1644.

Tamura, K., Dudley, J., Nei, M., Kumar, S. 2007. *MEGA4*: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.,* 24, 1596-1599.

Thompson, J.D., Higgins, D.G., Gibson, T.J. 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22, 4673-4680.

Twigg, L.E. & King, D.R. 1991. The impact of fluoroacetate-bearing vegetation on native Australian fauna: a review. *Oikos*, 61, 412-430.

Twigg, L.E., Martin, G.R., Lowe, T. 2002. Evidence of pesticide resistance in medium-sized mammalian pests: a case study with 1080 poison and Australian rabbits. *J. Appl. Ecol.*, 39, 549-560.

Twigg, L. E., Martin, G. R., Eastman, A. F., King, D. R., Kirkpatrick, W. E. 2003. Sensitivity of some Australian animals to sodium fluoroacetate (1080): additional species and populations, and some ecological considerations. *Aust. J. Zool.,* 51, 515-531.

Vontas, J., Blass, C., Koutsos, A.C., David, J-P., Kafatos, F.C., Louis, C., Hemingway, J., Christophides, G.K., Ranson, H. 2005. Gene expression in insecticide resistant and susceptible *Anopheles gambiae* strains constitutively or after insecticide exposure. *Insect Mol. Biol.*, 14, 509-521.

White, J.K., Mastroeni, P., Popoff, J-F., Evans, C.A.W., Blackwell, J.M. 2005. Slc11a1-mediated resistance to *Salmonella enteric* serovar T*yphimurium* and *Leishmania donovani* infections does not require functional inducible nitric oxide synthase or phagocyte oxidase activity. *J. Leucoc. Biol.*, 77, 311-320.

Wilson, T.G. 2001. Resistance of *Drosophila* to toxins. *Annu. Rev. Entomol.*, 46, 545-571.

Woo Cha, S., Kyoung Gu, H., Poong Lee, K., Han Lee, M., Seop Han, S., Cheon Jeong, T. 2000. Immunotoxicity of ethyl carbamate in female BALB/c mice: role of esterase and cytochrome P450. *Toxicol. Lett.*, 115, 173-181.

Yuen, T., Wurmbach, E., Pfeffer, R.L., Ebersole, B.J., Sealfon, S.C. 2002. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res.,*30, e48.

Appendices

Appendix I. The 13 genes likely to be involved in 1080 tolerance that were selected for further study. Column 4 briefly outlines gene function and how the gene could be involved in tolerance and also includes examples (if any) of the gene being implicated in cases of resistance to other toxins.

Primary function of the genes	Gene name	Gene abbreviation	Main reasons they are likely to be involved in tolerance of 1080 [1] microarray list ordered from largest difference in regulation between experimental and control rats to the smallest difference.
Classic resistance genes (detoxification)	Cytochrome P450 3a18	Cyp3a18	Cytochrome P450s shown to be involved in a number of cases of insecticide, herbicide, pollutant and drug **tolerance.** (e.g. Brun *et al.* 1996; Cuany *et al.* 1990 Houpt *et al.* 1988; Ohkawa *et al.* 1999; Markussen *et al.* 2008; Werck-Reichhart *et al.* 2000; Feyereisen 2005). Down-regulated in the presence of 1080 14th on microarray list [1]
	Glutathione S-transferase mu 2	Gstm2	GSTs involved in **tolerance and detoxification** of other toxins (Fotouhi-Ardakani *et al.* 2000; Wilson 2001; ffrench-constant *et al.* 2004). **Defluorination** occurs mainly in the liver via a major detoxification enzyme (fluoroacetate-specific defluorinase –FSD) which is glutathione-dependent and catalysed by a glutathione S-transferase (Twigg & King 1991; Tu *et al.* 2006).
	Glutathione S-transferase mu 7	Gstm7	
	Glutathione S-transferase mu 3	Gstm3	

			Defluorination is followed by reduced glutathione in the liver (Twigg & King 1991). Down-regulated in the presence of 1080 Gstm2: 28th on microarray list [1] Gstm3 and 7: 17th on the microarray list[1]
Metabolism/Krebs cycle	Carbonic anhydrase III	Car3	Role in cellular transport and metabolism (Shang *et al.* 2009) Down-regulated in the presence of 1080 **1st on microarray list** [1]
	Glucokinase	Gck	Linked to **Krebs cycle** and glycolysis (where 1080 acts; Yang *et al.* 2007) Down-regulated in the presence of 1080 23rd on microarray list [1]
	Protein phosphatatase I, regulatory (inhibitor) subunit 3B	Ppp1r3b	Predicted to be linked to **Krebs cycle** (where 1080 acts), increases glycogen synthesis by suppressing inactivation of glycogen phosphorylase and increasing rate of activation of glycogen synthase (NCBI 2010). Down-regulated in the presence of 1080 27th on microarray list [1]
	Lipin 1	Lpin1	Predicted to be structurally **related to** enzymes that metabolise haloacids which **degrade nematocides** (NCBI 2010). Up-regulated in the presence of 1080

			31st on microarray list [1]
Cell death	Complement component 6	C6	Predicted to be involved in cell death/membrane attack complex (NCBI 2010). Possible scenario: C6 enters cells, flags the cells affected by 1080 → cell death, thereby getting rid of 1080. Up-regulated in the presence of 1080 **3rd on microarray list** [1]
More specific actions	Dual specificity phosphotase 6	Dusp6	Intracellular (unlike the other genes chosen). Differentially **regulate MAPK isoforms**, which are signals from outside the cell to the nucleus that let the cell know what it has ‘bumped into’ (Owens & Keyes 2007). MAPK is involved in resistance to myxoma virus (Vilcek 2004). Down-regulated in the presence of 1080 20th on microarray list [1]
	Solute carrier family 11 (protein-coupled divalent metal ion transporters), member 1	Slc11a1	**Transcription factor**, regulates transcription, controls other genes that may be involved in tolerance mechanisms (Chen *et al.* 2007). Involved in resistance to *Salmonella typhimurium* in mice (White *et al.* 2005). Up-regulated in the presence of 1080 15th on microarray list [1]
	T-box 3	Tbx3	**Transcription factor**, controls other genes that may be involved in tolerance mechanisms (Zhang & King 1996).

			Down-regulated in the presence of 1080 30th on microarray list [1]
Unknown	Xm_001081053 (gene identifier) (similar to schlafen 3 & 4)	Slfn4	In the mitochondria, unknown what it does, but known to be in rats and some structure known. Schlafen 3 and 4 –growth regulatory gene (Schwartz *et al.* 1998). Up-regulated in the presence of 1080 **7th on microarray list** [1]

Note: some of the genes on the microarray list were unknown genes and therefore could not be studied further. However, one unknown gene chosen for further study was noted to be similar to Schlafen 3 and 4 so it was investigated as primers for Schlafen 3 and 4 could be used. Some genes that were high on the microarray list were not chosen for further study as their function was not related to resistance, for example some genes were involved in a generic stress response which was less likely to be the mechanism of tolerance than the chosen alternatives.

Additional references from appendix I:

Chen, Y., Lin, C., Ou, T., Wu, C., Tsai, W., Liu, H., Yen, J. 2007. Solute carrier family 11 member A1 gene polymorphisms in reactive arthritis. *J. Clin. Immunol.*, 27, 46-52.

Feyereisen, R. 2005. Insect cytochrome P450. In: Comprehensive Molecular Insect Science Vol. 4 (Ed. by L.I. Gilbert, K. Latrou, S.S. Gill), pp. 1–77.Oxford: Elsevier.

Fotouhi-Ardakani, N., Schecter, R.L., Batist, G. 2000. Evidence for the genomic duplication of the glutathione transferase A3 gene in genus *Rattus*. *Mol. Biol. Evol.*, 17, 331-335.

ffrench-constant, R.H., Daborn, P.J., Le Goff, G. 2004. The genetics and genomics of insecticide resistance. *TIG*, 20, 163-170.

Ohkawa, H., Tsujii, H., Ohkawa, Y. 1999. The use of cytochrome P450 genes to introduce herbicide tolerance in crops: a review. *Pesticide Science*, 55, 867-874.

Owens, D.M & Keyse, S.M. 2007. Differential regulation of MAP kinase signaling by dual-specificity protein phosphatases. *Oncogene*, 26, 3203-3213.

Schwartz, D.A., Katayama, C.D., Hedrick, S.M. 1998. *Schlafen*, a new family of growth regulatory genes that affect thymocyte development. *Immunity*, 9, 657-668.

Shang X., Chen, S., Ren, H., Huang, H. 2009. Carbonic anhydrase III: The new hope for the elimination of exercise-induced muscle fatigue. *Med. Hypotheses,* 72, 427-429.

Tu, L.Q., Wright, P.F.A., Rix, C.J., Ahokas, J.T. 2006. Is fluoroacetate-specific defluorinase a glutathione S-transferase? *Comp. Biochem. Physiol.,* 143, 59-66.

Vilcek, J. 2004. Why are rabbits uniquely sensitive to myxoma virus? Cherchez l'interferon!. *Nature Immunol.*, 5, 1205-1206.

Werck-Reichhart, D., Hehn, A., Didierjean, L. 2000. Cytochromes P450 for engineering herbicide tolerance. *Trends Plant Sci.*, 5, 116-123.

Yang, X-J., Mastaitis, J., Mizuno, T., Mobbs, C. 2007. Glucokinase regulates reproductive function, glucocotricoid secretion, food intake, and hypothalamic gene expression. *Endocrinology*, 148, 1928-1932.

Zhang, J. & King, M.L. 1996. *Xenopus VegT* RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. *Development,* 122, 4119-41.

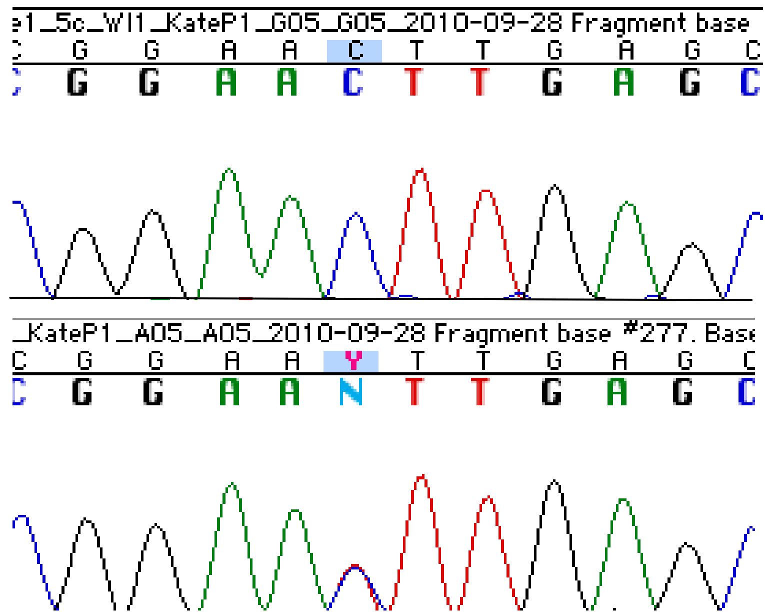
Appendix II. Sequencing primers for the 5 target genes that were successfully sequenced designed on *R. norvegicus* (GenBank numbers written below gene name), plus COI primers from (Folmer *et al.* 1994). Annealing temperature used shown in column 3.

Gene	*Forward and reverse primer sequences*	*Annealing temp.*	*Product length (bp)*
Aco2			
Aco2-143-F	5’-ACC TGC TAG AGA GAA GAA CAT TAA CAT -3’	60°C	Depends on combination of F and R.
Aco2-1229-F	5’-AAG TGC AAG TCT CAG TTC ACC -3’	53°C	
Aco2-735-F	5’-CCA AAG ATG TGA TCC TGA AAG -3’	60°C	
Aco2-1543-R	5’-GAA GTC AGT TTC TGG GTT GAA -3’	60°C	
Aco2-2140-R	5’-GAT GAA TCT TGT TGT ACT CAG AGG -3’	53°C	
Aco2-2636-R	5’-AGT GCT GTC ATC AAA AAT AAA TAC A -3’	53°C	
NM_024398.2			
Cyp3a18	F 5’- ACG GTG ATG GCA TGT GGA AA - 3’	65°C	1526
NM_145782	R 5’- TGA TAC ACC GCA GAG CCA CT - 3’		
Gstm7	F 5’- CCT GGA CTT CCC CAA TCT GC - 3’	61°C	981
NM_031154.1	R 5’- AGC AGC AGG AAG AAA GAG CG - 3’		
Gstm2	F 5’- TGG GTT ACT GGG ACA TCC GT - 3’	61°C	564
NM_177426	R 5’-TCT TCA GGC CCT CAA ACC GA - 3’		
Car3	F 5’- CCA GCC ACA ATG GTC CTG AG - 3’	65°C	709
NM_019292	R 5’- CAA TTC CCC ACC AGA GGC AC - 3’		
COI			
LCO1490:	5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3'	40°C	710
HC02198:	5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3'		

Appendix III. qRT-PCR primers for the 13 target genes and housekeeping gene designed on *R. norvegicus* (GenBank numbers written below gene name).

Gene	*Forward and reverse primer sequences*		*Product length (bp)*
Car3	F	5'-GAC GGG AGA AAG GCG AGT TC -3'	225
NM_019292	R	5'-GCT TGG CCA TCT GGT CTG AG -3'	
Ppp1r3b	F	5'-AGC CAG AGA GCA GGT GAG AG -3'	163
NM_138912.2	R	5'-GGC AGG CTA GAA GTC CGG TA -3'	
Gck	F	5'-CAT GAT TGT GGG CAC TGG CT -3'	275
NM_012565.1	R	5'-GAA GGT TCT CGT CCA CCA GC -3'	
Lpin1	F	5'-CGC CTT GCA CAG AGA AGT GA -3'	224
NM_001012111.1	R	5'-TGT TGG TCT TGG CAT GCT CC -3'	
C6	F	5'-ATG GGC GCT GGG TTT CAT TT -3'	133
NM_176074.2	R	5'-CCA GAT TGG CTG GAA CTC GG -3'	
Gstm7	F	5'-ACA ACC TGT GTG GGG AGA CA -3'	158
NM_031154.1	R	5'-CCG CAT CAT TCC AGG CAG TT -3'	
Gstm2	F	5'-TAT GGA CAC CCG CCT ACA GT -3'	265
NM_177426	R	5'-CTT CAG GCC CTC AAA CCG AG -3'	
Dusp6	F	5'-ACC CCC AAT CTG CCC AAT CT -3'	137
NM_053883.2	R	5'-TTG CCT CGG GCT TCA TCT AT -3'	
Slc11a1	F	5'-AAT TGC ACG CGT CCT TCT CA -3'	139
NM_001031658.1	R	5'-CAC AGC AAA GGG CAG CAG TA -3'	
Tbx3	F	5'-AGA CCG GCA TCC CTT TCT CA -3'	184
NM_181638.1	R	5'-GGG AAC ATT CGC CTT CCT GA -3'	
Cyp3a18	F	5'-CAA ACC GTC GGT GTT TTG GG -3'	203
NM_145782	R	5'-GAT GGG CTC CCC TTT TGC TT -3'	
Slfn4	F	5'-CAG TTG CTC TGG GCA CAG TT -3'	155
XM_577117.3	R	5'-ACA AAG CCA TGC AGG GTC AG -3'	
Gstm3	F	5'-GTT TGC AGG GGA CAA GGT CA -3'	228
NM_020540.1	R	5'-ATG GCA GGG GCC TAA TCA GT -3'	
Pgk1	F	5'-GAA-GGG-AAG-GGA-AAA-GAT-GC -3'	180
NM_053291.3	R	5'-AAA-TCC-ACC-AGC-CTT-CTG-TG -3'	

Appendix IV: Example of chromatograms with one heterozygous rat and one homozygous rat showing the method of determining what base/s were at each site. The highlighted base in the upper chromatogram would be called a C (blue) while the same base in the lower chromatogram would be called Y –a combination of C (blue) and T (red), making it a heterozygote.



Appendix V. Variable sites among aligned nucleotide sequences for the rats sequenced, including differences between *R. norvegicus* and *R. fuscipes*. There are a number of non-synonymous changes, however only two of these are within *R. fuscipes*. 'Reference' is *R. norvegicus* reference sequence for each gene (see appendix II and II for Genbank reference numbers), W1, W2 are *R. norvegicus* sequenced in this study, 1-4 are SA rats, F#, M# are WA rats, WI# are WI rats. Spaces/gaps in the place of nucleotides are where sequence wasn't obtained at that site. a.a =amino acid. In 'change in amino acid' row, blank space =no change and '?' =gap in one of the sequences that can't didn't translate to an amino acid. 3D structure of protein ('part of gene' row) known only for Aco2. A: adenine, T: thymine, C: cytosine, G: guanine. Heterozygotes (multiple nucleotides at the one site): R =G + A, W =A + T, Y =C + T.

Aco2 Reading frame 3. S.B.S =substrate binding site, Sw =swivel.

Rat\Site	426	492	498	513	699	906	1095	1212	1266	1548	1560	1821	1938	2025
Reference	A	T	C	A	C	C	T	A	T	A	T	G	T	C
W1										G	T			T
W2										G	T	G	C	T
2									C	G	C	G	C	T
3						C	C	G	C	G	C	G	C	T
4	G	C	T	G	T	C	C	G	C	G	C	G	C	T
F1										G	T	R	C	T
F5										G	T	R	C	T
F20										G	T	R	C	T
M21												R		
WI1									T	G	T	G	C	T
WI2						T	C	G	T	G	T	G	C	T
WI5										G	T	G	C	T
WI6										G	T	G	C	T
WI7						T	C	G	T	G	T	G	C	T
WI8	G	C	T	G	T	T	C	G	T	G	T	G	C	T
a.a	142	164	166	171	233	302	365	404	422	516	520	607	646	675
Change in a.a.														
Part of gene												Sw S.B.S	Sw	Sw

Aco2 continued...

Rat\Site	2046	2082	2133	2199	2250	2324	2344	2365	2366,7	2388	2393	2406	2428
Reference	T	A	T	C	C	A	C	C	- -	C	G	C	C
W1	T												
W2	C												
2	C	C	C	C	G	W							
3	C												
4	C	C	C	C	G	A							
F1	C	C	C	C	G	W							
F5	C												
F20	C	C	C										
M21						W							
WI1	C												
WI2	C	C	C	T	G	W	Y	G	T C	T	T	T	A
WI5	C	C	C	T	G								
WI6	C	C	C	T	G								
WI7	C	C	C	T	G	A							
WI8	C	C	C	T	G	W	Y	G	T C	T	T	T	A
a.a	682	694	711	733	750	775	782	789	789	796	797	802	809
Change in a.a						YES	YES	YES	?	YES			
Part of gene	Sw	Sw	Sw	Sw									

Cyp3a18 Reading frame 2.

Rat\Site	303	454	555	618	669	936	1053	1056	1068	1197	1256	1335	1458	1482	1520	1572
Reference	T	A	A	C	C	T	C	G	T	G	A	C	G	T	G	G
1	G	C	G	T	T	C	T	A	C	A	A	T	T	C	A	A
3	G	C	G	T	T	C	T	A	C	A	A	T	T	C	A	A
4	G	C	G	T	T	C	T	A	C	A	A	T	T	C	A	A
M2										A	G	T	T	C	A	A
F1											R					
a.a	101	152	185	206	223	312	351	352	356	399	419	444	486	494	507	524
Change in a.a	?										YES		YES		YES	

Gstm7 Reading frame 2. (Ignoring beginning of sequences which didn't sequence well)

Rat\Site	619	651	732	774	775	792	838	956	964	993	995	1058	1059
Reference	T	A	A	T	G	C	-	C	A	C	C	C	C
W1	T	A	A	T	G	C	A	C	A	C	C		
W2	T	A	A	T	G	C	A	C	A	C	C		
3	A	G	G	C	A	T	A	T	G	T	-		
4	A	G	G	C	A	T	A	T	G	T	-	-	-
F1	A	G	G										
M2	A	G	G	C	A	T	A	T	G	T	-		
M3													
M21	A	G	G	C	A	T	A	T	G	T	-		
a.a	207	217	244	258	259	264	280	319	322	331	332	353	353
Change in a.a	YES				YES		?	YES	YES			?	?

Gstm2 Reading frame 1.

Rat\Site	78	117	120	192	208	222	243	246	256	258	295	363	432	437	438	462	498	553
Reference	A	C	C	T	A	T	C	G	C	T	G	G	G	C	A	G	A	G
W2	A	C	C	T	A	T	C	G	C	T	G	G	G	C	A	G	A	G
1	C	T	T	C	G	C	T	C	T	G	A	G	G	T	G	C	G	A
2	C	T	T	C	G	C	T	C	T	G	A	G	G	T	G	C	G	A
3	C	T	T	C	G	C	T	C	T	G	A	G	G	T	G	C	G	A
4	C	T	T	C	G	C	T	C	T	G	A	G	G	T	G	C	G	A
F1	C	T	T	C	G	C	T	C	T	G	A	R	G	T	G	C	G	A
F4	C	T	T	C	G	C	T	C	T	G	A	R	G	T	G	C	G	A
F5	C	T	T	C	G	C	T	C	T	G	A	R	G	T	G	C	G	A
F20	C	T	T	C	G	C	T	C	T	G	A	R	G	T	G	C	G	A
M2	C	T	T	C	G	C	T	C	T	G	A	G	G	T	G	C	G	A
M3	C	T	T	C	G	C	T	C	T	G	A	A	G	T	G	C	G	A
M21	C	T	T	C	G	C	T	C	T	G	A	G	A	T	G	C	G	A
a.a	26	39	40	64	70	74	81	82	86	86	99	121	144	146	146	154	166	185
Change in a.a					YES						YES			YES	YES			YES

Car3 Reading frame 3.

Rat\Site	135	312	423	474	555	612
Reference	T	G	C	C	G	A
W1	T	G	C	C	G	A
W2	T	G	C	C	G	A
1	C	C	T	C	A	G
2	C	C	T	Y	A	G
3	C	C	T	Y	A	G
4	C	C	T	C	A	G
F1	C	C	T	C	A	G
F4	C	C	T	C	A	G
F5	C	C	T	C	A	G
F20	C	C	T	C	A	G
M2	C	C	T	C	A	G
M3	C	C	T	C	A	G
M6	C	C	T	C	A	G
M21	C	C	T	C	A	G
a.a	45	104	141	158	185	204
Change in a.a						

Tables

Table 1. Nucleotides at sites in four sequenced genes displaying sequence mutations between the four types of rats (excluding mutations only between *R. norvegicus* and *R. fuscipes*). Spaces/gaps are rats for which sequence wasn't obtained at that site, i.e. nucleotides are unknown. The non-synonymous changes are highlighted in bold. (Reference sequence, W1, W2 =*R. norvegicus*; 1-4 =sensitive SA rats; F# and M# =WA rats; WI# =Woody Island rats). The lengths of each gene are Aco2: 3,287-bp (base pairs), Cyp3a18: 2,005-bp, Gstm2: 657-bp, Car3: 802-bp. Excluding the reference sequence, up to 68.15, 66.28, 82.95 and 84.54% of the gene for Aco2, Cyp3a18, Gstm2 and Car3 respectively was sequenced. A: adenine, T: thymine, C: cytosine, G: guanine. Heterozygotes (multiple nucleotides at the one site): R =G + A, W =A + T, Y =C + T.

		Aco2					Cyp3a18	Gstm2	Car3
Pop.	**Rat**	906	1266	1560	1821	**2324**	**1256**	363	474
***R. norvegicus* (wistar rat)**	Reference sequence	C	T	T	G	**A**	**A**	G	C
	W1			T	G				C
	W2			T	G			G	C
SA rats	1				G		**A**	G	C
	2		C	C	G	**W (=A+T)**		G	Y (=C+T)
	3	C	C	C	G		**A**	G	Y
	4	C	C	C	G	**A**	**A**	G	C
WA rats	F1			T	R (=G+A)	**W**	**R (=G+A)**	R (=G+A)	C
	F4							R	C
	F5			T	R			R	C
	F20			T	R			R	C
	M2						**G**	G	C
	M3							A	C
	M6							G	C
	M21				R	**W**			C
WI rats	WI1		T	T	G				
	WI2	T	T	T	G	**W**			
	WI5			T	G				
	WI6			T	G				
	WI7	T	T	T	G	**A**			
	WI8	T	T	T	G	**W**			
	Amino acid position	302	422	520	607	**775**	**419**	121	158
	Amino acid substitution?	No	No	No	No	**Yes**	**Yes**	No	No
	Interesting part of the 3D protein structure?				Mt swivel, Substrate binding site				

GenBank reference sequences (each gene in *R. norvegicus*) – Aco2: NM_024398.2, Cyp3a18: NM_145782, Gstm7 (no sequence mutations within *R. fuscipes*): NM_031154.1, Gstm2: NM_177426, Car3: NM_019292. (Aco2: mitochondrial aconitase, Cyp3a18: cytochrome P450 3a18, Gstm2: glutathione S-transferase mu 2, Car3: carbonic anhydrase III).

Table 2. Results of a two-way ANOVA assuming equal variance looking at the main effects of population and treatment in the baseline gene expression of WA and WI rats displaying P-values for each gene. A sequential Bonferroni correction set the P-value required for significance at 0.05/(n-1), calculated in column 4.

Gene	DF	F value	P-value	Significant
Car3	1	2.442	0.1464	
Lpin1	1	29.730	0.0002	Yes
C6	1	3.982	0.0713	
Gstm7	1	8.826	0.0127	
Gstm2	1	33.752	0.0001	Yes
Dusp6	1	13.225	0.0039	Yes
Slc11a1	1	8.819	0.0128	
Cyp3a18	1	6.496	0.0271	
Slfn4	1	0.286	0.6033	
Gstm3	1	27.981	0.0003	Yes
Tbx3	1	9.063	0.0119	

(Car3: carbonic anhydrase III, Lpin1: Lipin 1, C6: complement component 6, Gstm7: glutathione S-transferase mu 7, Gstm2: glutathione S-transferase mu 2, Dusp6: Dual specificity phosphotase 6, Slc11a1: solute carrier family 11 member 1, Cyp3a18: cytochrome P450 3a18, Slfn4: schlafen 4, Gstm3: glutathione S-transferase mu 3, Tbx3: T-box 3.)

Table 3. P-values from Scheffe post-hoc tests performed on significant differences found by the previous ANOVA in baseline gene expression between population and treatment in WA and WI rats, indicating the specific treatment groups exhibiting differences in gene expression.

Gene	Groups	P-value
Lpin1	control WA, control WI	0.0213
	control WA, treated WI	0.0167
	control WI, treated WA	0.0246
	treated WA, treated WI	0.0194
Gstm2	control WA, control WI	0.0403
	control WA, treated WI	0.0045
	control WI, treated WA	0.0396
	treated WA, treated WI	0.0044
Dusp6	control WI, treated WA	0.0480
Gstm3	control WA, treated WI	0.0057
	treated WA, treated WI	0.0054

(Lpin1: Lipin 1, Gstm2: glutathione S-transferase mu 2, Dusp6: Dual specificity phosphotase 6, Gstm3: glutathione S-transferase mu 3.)

Figures

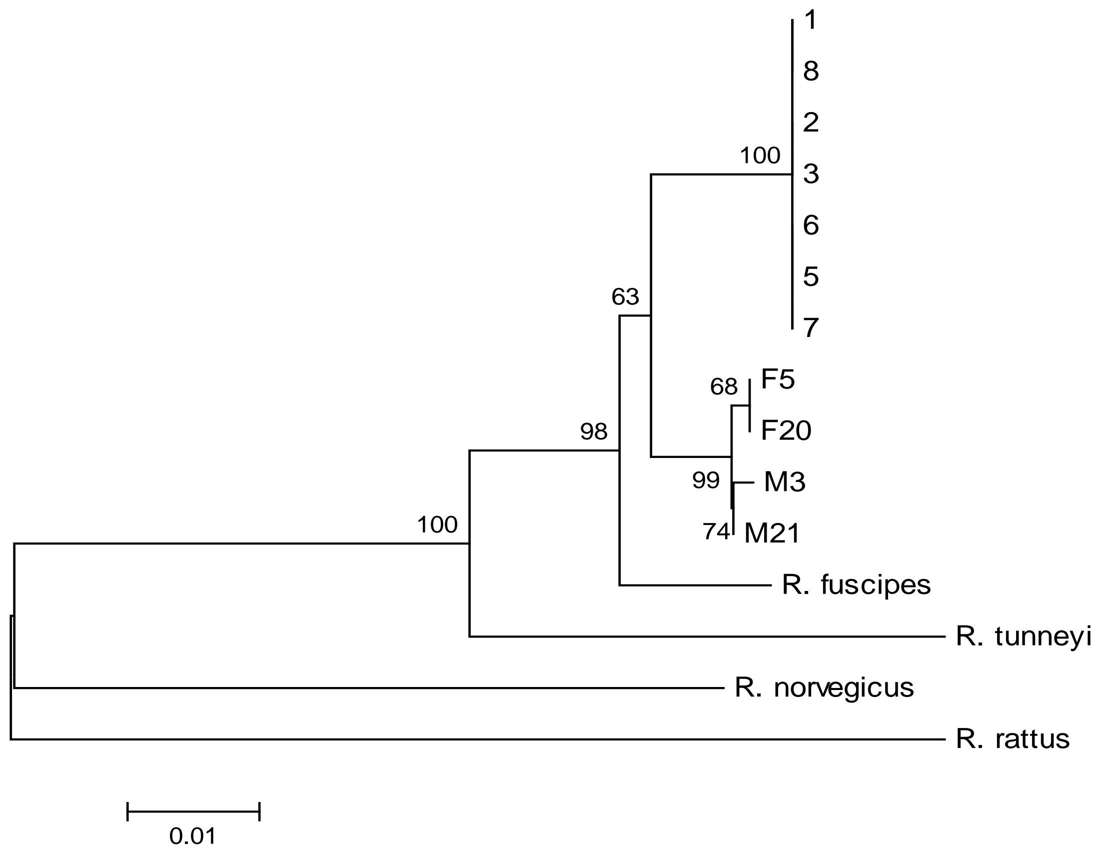


Fig 1. Phylogenetic neighbour-joining tree of two introduced rat species in WA (*R. norvegicus* and *R. rattus*), two native WA species (*R. fuscipes* and *R. tunneyi*) and *R. fuscipes* individuals used in this study. Bootstrap (999 replicates), nucleotide Kimura 2-parameter model used (most common model assuming transitions and transversions occur at different rates). Constructed on nucleotide sequences (including *R. norvegicus* COI sequence below) to confirm by DNA barcoding that all WI rats were *R. fuscipes*. (1-8 are WI; F# and M# are WA). GenBank numbers of reference sequences used were the COI regions in *R. norvegicus*: NC_001665.2, *R. rattus*: FJ355927.1 , *R. tunneyi*: EF186633.1, *R. fuscipes*: EF186554.1). <u>Below:</u> COI nucleotide and amino acid sequence in *R. norvegicus*. Nucleotides that are different in *R. fuscipes* are highlighted in grey. Didn't obtain sequence for my *R. fuscipes* after 716-bp (highlighted in red). No non-synonymous mutations (amino acid substitutions) were evident between the two *Rattus* species.

```
 1   M  L  V  N  R  -  L  F  S  T  N  H  K  D  I  G  T  L  Y  L
 1  ATGCTCGTAAACCGTTGACTCTTTTCAACTAACCACAAAGATATCGGAACCCTCTACCTA
21   L  F  G  A  -  A  G  I  V  G  T  A  L  S  I  L  I  R  A  E
61  TTATTTGGAGCCTGAGCAGGAATAGTAGGGACAGCTTTAAGTATTCTAATTCGAGCTGAA
41   L  G  Q  P  G  A  L  L  G  D  D  Q  I  Y  N  V  I  V  T  A
```

```
121   CTAGGACAGCCAGGCGCACTCCTAGGAGATGACCAAATCTATAATGTCATCGTCACAGCC
 61    H  A  F  V  I  I  F  F  I  V  I  P  I  I  I  G  G  F  G  N
181   CATGCATTCGTAATAATTTTCTTTATAGTAATACCTATAATAATTGGAGGCTTCGGGAAC
 81    -  L  V  P  L  I  I  G  A  P  D  I  A  F  P  R  I  N  N  I
241   TGACTTGTACCACTAATAATTGGAGCCCCTGATATAGCATTCCCACGAATAAATAACATA
101    S  F  -  L  L  P  P  S  F  L  L  L  L  A  S  S  I  V  E  A
301   AGCTTTTGACTGCTTCCTCCATCATTTCTACTCCTTTTAGCATCCTCCATAGTAGAAGCT
121    G  A  G  T  G  -  T  V  Y  P  P  L  A  G  N  L  A  H  A  G
361   GGAGCTGGAACAGGATGAACAGTATATCCCCCCTTAGCCGGAAACCTAGCCCATGCTGGA
141    A  S  V  D  L  T  I  F  S  L  H  L  A  G  V  S  S  I  L  G
421   GCATCCGTAGATTTAACTATTTTTTCCCTCCACCTAGCCGGGGTGTCTTCTATCTTAGGA
161    A  I  N  F  I  T  T  I  I  N  I  K  P  P  A  I  T  Q  Y  Q
481   GCTATCAACTTTATCACCACTATCATTAATATAAAACCCCCTGCTATAACCCAATATCAG
181    T  P  L  F  V  -  S  V  L  I  T  A  V  L  L  L  L  S  L  P
541   ACACCTCTCTTTGTATGATCCGTACTAATTACAGCCGTCCTACTACTTCTCTCACTGCCA
201    V  L  A  A  G  I  T  I  L  L  T  D  R  N  L  N  T  T  F  F
601   GTATTAGCAGCAGGTATCACTATACTCCTTACAGACCGAAATCTAAATACTACTTTCTTC
221    D  P  A  G  G  G  D  P  I  L  Y  Q  H  L  F  -  F  F  G  H
661   GACCCCGCTGGAGGTGGAGACCCAATCCTTTATCAACACCTATTCTGATTCTTCGGCCAC
241    P  E  V  Y  I  L  I  L  P  G  F  G  I  I  S  H  V  V  T  Y
721   CCAGAAGTGTACATCTTAATTCTTCCAGGGTTTGGAATTATTTCACATGTAGTTACCTAT
261    Y  S  G  K  K  E  P  F  G  Y  I  G  M  V  -  A  I  I  S  I
781   TACTCTGGAAAAAAAGAACCCTTCGGATATATAGGTATGGTATGAGCCATAATATCTATT
281    G  F  L  G  F  I  V  -  A  H  H  I  F  T  V  G  L  D  V  D
841   GGCTTCCTAGGATTTATTGTATGAGCACATCACATATTCACAGTAGGCCTAGATGTAGAC
301    T  R  A  Y  F  T  S  A  T  I  I  I  A  I  P  T  G  V  K  V
901   ACCCGAGCCTACTTTACATCTGCCACTATAATTATCGCAATTCCTACAGGCGTAAAAGTA
321    F  S  -  L  A  T  L  H  G  G  N  I  K  -  S  P  A  I  L  -
961   TTCAGCTGACTCGCTACACTACATGGAGGAAATATCAAATGATCCCCCGCCATATTATGA
341    A  L  G  F  I  F  L  F  T  V  G  G  L  T  G  I  V  L  S  N
1021  GCCTTAGGGTTTATCTTCTTATTCACAGTAGGGGGCCTAACAGGGATCGTACTATCTAAC
361    S  S  L  D  I  V  L  H  D  T  Y  Y  V  V  A  H  F  H  Y  V
1081  TCATCCCTTGACATTGTACTTCATGATACATACTATGTAGTAGCTCACTTCCACTATGTC
381    L  S  I  G  A  V  F  A  I  I  A  G  F  V  H  -  F  P  L  F
1141  TTATCTATAGGAGCAGTATTCGCCATCATAGCTGGCTTCGTCCACTGATTCCCACTATTC
401    S  G  Y  T  L  N  D  T  -  A  K  A  H  F  A  I  I  F  V  G
1201  TCAGGCTATACCCTAAATGACACATGAGCAAAAGCCCACTTTGCCATTATATTTGTAGGT
421    V  N  I  T  F  F  P  Q  H  F  L  G  L  A  G  I  P  R  R  Y
1261  GTAAACATAACATTCTTTCCTCAACACTTCCTAGGATTAGCAGGGATACCTCGTCGTTAC
441    S  D  Y  P  D  A  Y  T  T  -  N  T  V  S  S  I  G  S  F  I
1321  TCTGATTATCCAGATGCTTACACCACATGAAATACAGTCTCCTCTATAGGCTCATTCATC
461    S  L  T  A  V  L  V  M  I  F  M  I  -  E  A  F  A  S  K  R
1381  TCACTTACGGCCGTCCTTGTAATGATCTTCATGATTTGAGAAGCCTTCGCATCAAAACGA
481    E  V  L  S  I  S  Y  S  S  T  N  L  E  -  L  H  G  C  P  P
1441  GAAGTACTCTCAATTTCCTACTCCTCAACTAACCTAGAATGACTGCATGGATGCCCCCCA
501    P  Y  H  T  F  E  -  I
1501  CCCTACCACACATTCGAATAAATAA
```

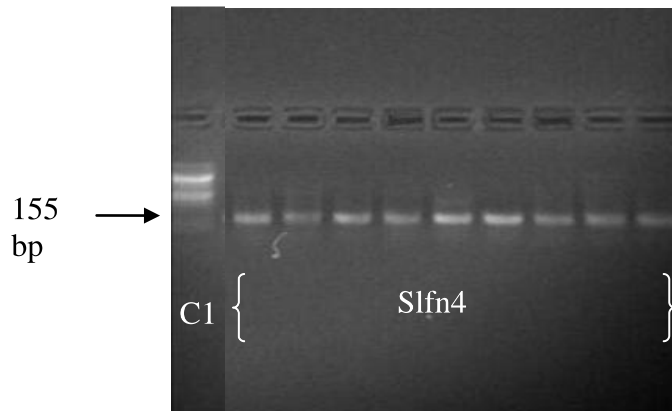


Fig 2. Gel electrophoresis of the PCR-amplified products in order to verify specificity of the primers. TBE 2% agarose gel with Sybr® Safe DNA gel stain (Invitrogen, Carlsbad, CA). Column 1 (C1): molecular weight standard (100-bp ladder DNA marker) (Axygen biosciences, Union City, CA). Slfn4 gene amplified in 9 rats (one rat per column). All columns have only a single band indicating that primers are amplifying only one gene product. The gene products are 155 base pairs long –the correct length for the fragments for Slfn4, indicating that the primers have amplified the correct product.

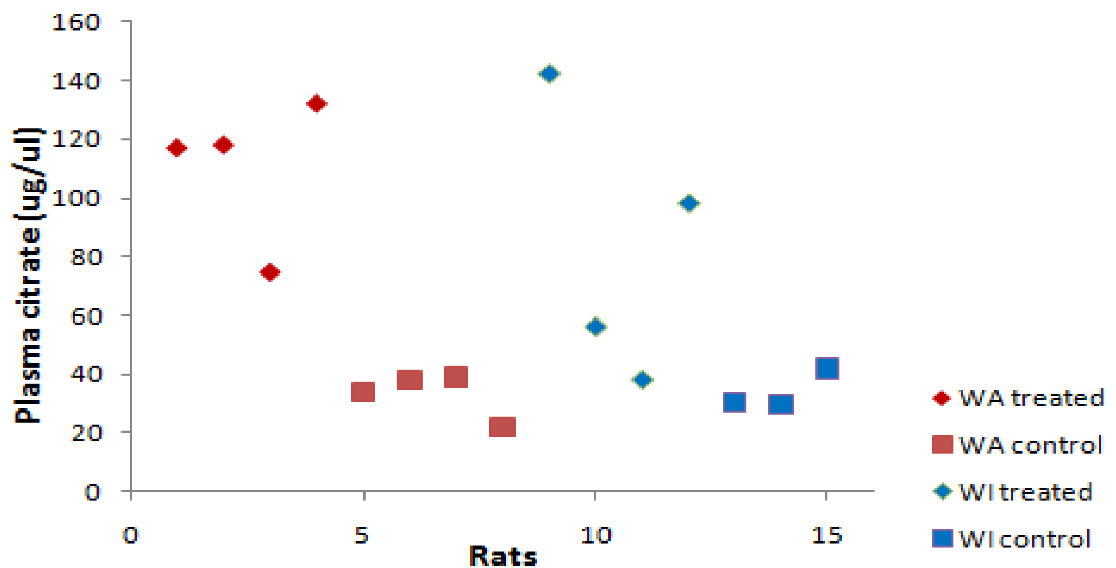


Fig 3. Plasma citrate levels measured in µg/µl for each of the WA (red) and WI (blue) control (square) and treated (diamond) rats.

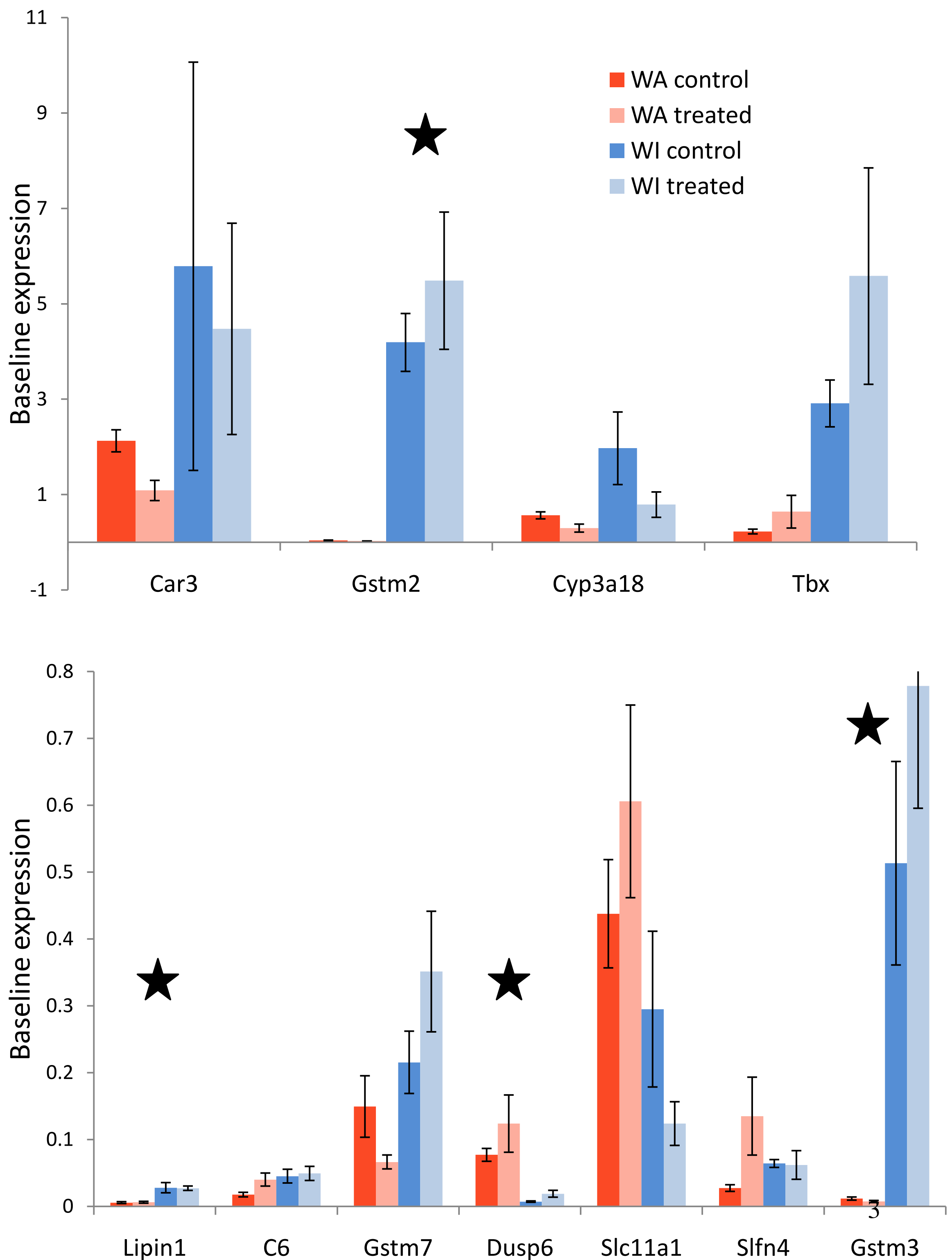


Fig 4. Averages of baseline gene expression data for each population and treatment combination for each gene. Above: Genes with higher expression on a larger y axis scale. Below: Genes with lower expression on a magnified y axis scale. Black star represents a significant difference between populations in a gene. Sequential Bonferroni correction set the significance level at P =0.05/(n-1). Error bars are standard errors.

(Car3: carbonic anhydrase III, Lpin1: Lipin 1, C6: complement component 6, Gstm7: glutathione S-transferase mu 7, Gstm2: glutathione S-transferase mu 2, Dusp6: Dual specificity phosphotase 6, Slc11a1: solute carrier family 11 member 1, Cyp3a18: cytochrome P450 3a18, Slfn4: schlafen 4, Gstm3: glutathione S-transferase mu 3, Tbx3: T-box 3.)

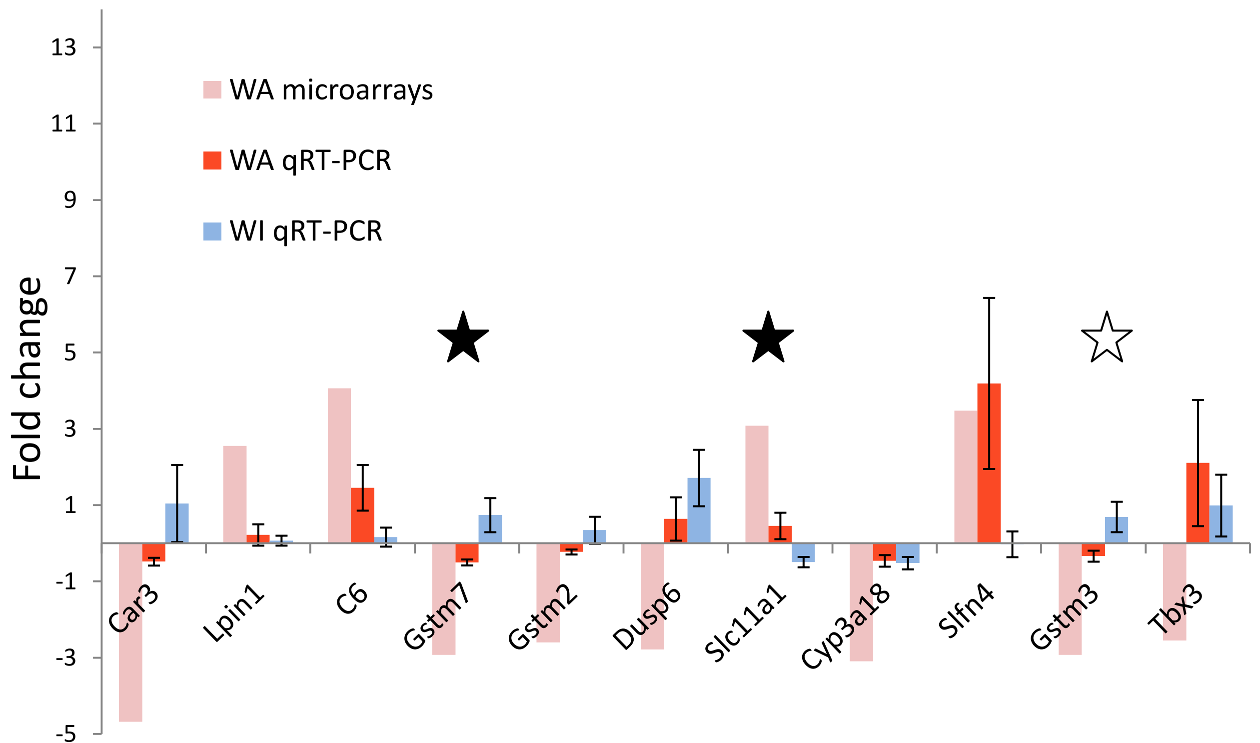


Fig 5. Fold change centred around zero =no change, measured in WA and WI rats by qRT-PCR and WA by microarrays. WA qRT-PCR confirms the WA microarray results for most genes. A filled black star indicates a significant difference between populations for that gene, the star above Gstm3 indicates a strong trend. Error bars are standard errors. (Car3: carbonic anhydrase III, Lpin1: Lipin 1, C6: complement component 6, Gstm7: glutathione S-transferase mu 7, Gstm2: glutathione S-transferase mu 2, Dusp6: Dual specificity phosphotase 6, Slc11a1: solute carrier family 11 member 1, Cyp3a18: cytochrome P450 3a18, Slfn4: schlafen 4, Gstm3: glutathione S-transferase mu 3, Tbx3: T-box 3.)

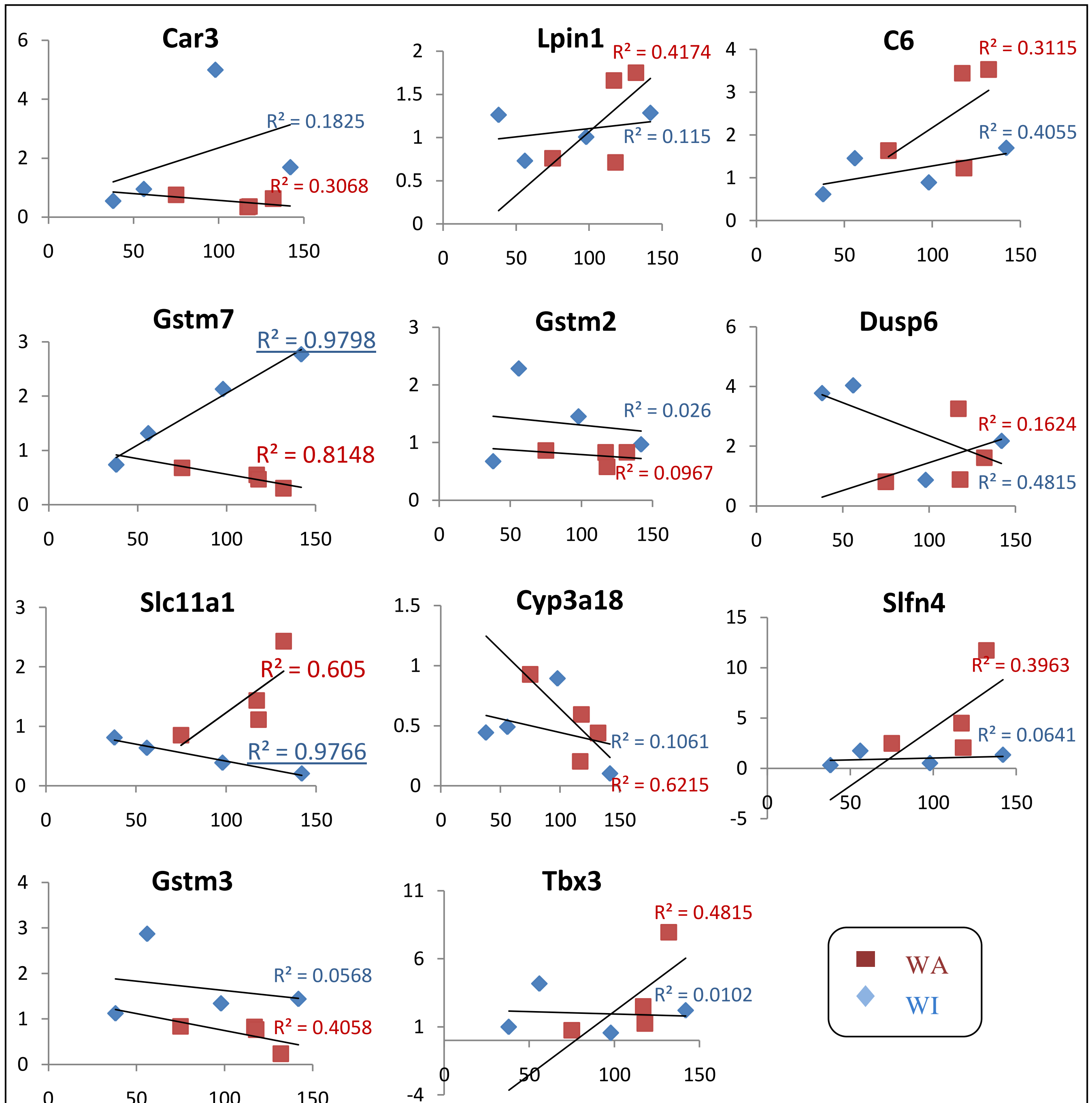


Fig 6. Fold change (y axis) in each of the four treatment rats in each population plotted against plasma citrate levels (x axis) for WI and WA treated rats. Plots show the lack of relationship between the two variables in all except Gstm7 and Slc11a1 for which there are strong relationships. Correlations (R) for the two genes were Gstm7: 0.903 (WA) and 0.989 (WI), Slc11a1: 0.778 (WA) and 0.988 (WI). Correlations were significant in WI rats for Gstm7 and Slc11a1 (R >0.95). (Car3: carbonic anhydrase III, Lpin1: Lipin 1, C6: complement component 6, Gstm7: glutathione S-transferase mu 7, Gstm2: glutathione S-transferase mu 2, Dusp6: Dual specificity phosphotase 6, Slc11a1: solute carrier family 11 member 1, Cyp3a18: cytochrome P450 3a18, Slfn4: schlafen 4, Gstm3: glutathione S-transferase mu 3, Tbx3: T-box 3).