Received: 11 August 2017 | Revised: 14 December 2017 | Accepted: 19 December 2017

DOI: 10.1111/mec.14480

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Role of cytochrome P450 2B sequence variation and gene copy number in facilitating dietary specialization in mammalian herbivores

Smiljka Kitanovic[1] | Teri J. Orr[1] | Daniel Spalink[1] | Granger B. Cocke[1] | Katharina Schramm[2] | P. Ross Wilderman[3] | James R. Halpert[3] | M. Denise Dearing[1]

[1]Department of Biology, University of Utah, Salt Lake City, UT, USA

[2]Department of Botany, Weber State University, Ogden, UT, USA

[3]School of Pharmacy, University of Connecticut, Storrs, CT, USA

Correspondence
Smiljka Kitanovic, Department of Biology, University of Utah, Salt Lake City, UT, USA.
Email: kika.kitanovic@utah.edu

Funding information
National Science Foundation, Grant/Award Number: DEB 1457366, IOS 1256383, IOS 1256840

Abstract

Theory postulates that dietary specialization in mammalian herbivores is enabled by a specialized set of liver enzymes that process the high concentrations of similar plant secondary metabolites (PSMs) in the diets of specialists. To investigate whether qualitative and quantitative differences in detoxification mechanisms distinguish dietary specialists from generalists, we compared the sequence diversity and gene copy number of detoxification enzymes in two woodrat species: a generalist, the white-throated woodrat (*Neotoma albigula*) and a juniper specialist, Stephens' woodrat (*N. stephensi*). We focused on enzymes in the cytochrome P450 subfamily 2B (CYP2B), because previous research suggests this subfamily plays a key role in the processing of PSMs. For both woodrat species, we obtained and sequenced *CYP2B* cDNA, generated *CYP2B* phylogenies, estimated *CYP2B* gene copy number and created a homology model of the active site. We found that the specialist possessed on average ~5 more *CYP2B* gene copies than the generalist, but the specialist's *CYP2B* sequences were less diverse. Phylogenetic analysis of putative CYP2B homologs resolved woodrat species as reciprocally monophyletic and suggested evolutionary convergence of distinct homologs on similar key amino acid residues in both species. Homology modelling of the CYP2B enzyme suggests that interspecific differences in substrate preference and function likely result from amino acid differences in the enzyme active site. The characteristics of *CYP2B* in the specialist, that is greater gene copy number coupled with less sequence variation, are consistent with specialization to a narrow range of dietary toxins.

KEYWORDS
CYP2B, cytochrome P450, detoxification, diet, herbivores, woodrats

1 | INTRODUCTION

Although a common feeding strategy among mammals, herbivory can have serious consequences (Dearing, Foley, & McLean, 2005; Ley et al., 2008). Most plants produce plant secondary metabolites (PSMs) to defend themselves from consumption; thus, herbivores must routinely cope with dietary toxins. More than 40 years ago, Freeland and Janzen proposed that detoxification enzymes mediate the interactions between mammalian herbivores and the plants on which they feed and that there are fundamental physiological constraints in the detoxification systems of generalists and specialists (Freeland & Janzen, 1974). The generalist feeding strategy of most

mammalian herbivores was surmised to result from limitations of a hepatic detoxification system that had evolved to process a wide range of substrates with low catalytic activity towards any particular substrate. Generalists were proposed to harbour a broad diversity of biotransformation ("detoxification") enzymes to metabolize the variety of toxins in their diet. In contrast, the dozen or so mammalian species capable of specializing on toxic plants were predicted to have evolved detoxification enzymes with enhanced function to facilitate efficient catalysis of a narrow set of PSMs in the diet.

The idea that enzymatic biochemical trade-offs govern the diet breadth of herbivores has become entrenched in the field of plant–animal interactions. Several studies performed on herbivorous insects are largely consistent with the predictions of the biochemical trade-off hypothesis (but also see Sezutsu, Le Goff, & Feyereisen, 2013; for a different perspective). Seminal work on insects revealed that a generalist insect has more detoxification enzymes with lower catalytic activity towards a wider variety of substrates than a specialist (Li, Baudry, Berenbaum, & Schuler, 2004; Li, Berenbaum, & Schuler, 2003; Li, Zangerl, Schuler, & Berenbaum, 2004). Moreover, the diversification of detoxification enzymes in relation to host plants appears adaptive and nonrandom (Calla et al., 2017). The regulation of cytochrome P450 genes is also consistent with host-plant specialization in herbivorous *Drosophila* (Bono, Matzkin, Castrezana, & Markow, 2008). However, to date this hypothesis has not been as rigorously tested in mammalian systems (Marsh, Wallis, McLean, Sorensen, & Foley, 2006; Ngo, Kong, Kirlich, McKinnon, & Stupans, 2000).

In addition to having more efficient detoxification enzymes, dietary specialists are predicted to produce higher quantities of detoxification enzymes than generalists to adequately respond to high toxin loads in their diet (Freeland & Janzen, 1974). Gene duplication is one mechanism that enables greater enzyme production (Perry et al., 2007). Multiple copies of functional detoxification genes have been documented across various species, such as *CYP2B6* in human populations (Bass et al., 2014; Malenke, Magnanou, Thomas, & Dearing, 2012; Nelson et al., 2004; Zanger et al., 2007). Variation in the copy number of detoxification genes in the cytochrome P450 superfamily is known to affect the ability of humans to metabolize pharmaceuticals (Ingelman-Sundberg, 2005). Moreover, adaptation to pesticides by herbivorous insects can be enabled via increased copy number of cytochrome P450 genes (Bass et al., 2014).

To further our understanding of the mechanisms underlying dietary specialization in herbivorous mammals, we have been conducting studies of cytochrome P450 enzymes, particularly those within the subfamily 2B (CYP2B), in specialist and generalist woodrats (genus *Neotoma*). Our previous work indicates that cytochrome P450 detoxification enzymes are critical in the ability of woodrats to ingest juniper (Skopec, Malenke, Halpert, & Dearing, 2013). Junipers are high in terpenes, a class of PSMs known for their toxicity to the central nervous system (Waidyanatha et al., 2013). Juniper feeding induces expression of hepatic CYP2B enzymes in several species of woodrats, with specialist woodrats exhibiting greater CYP2B activity relative to generalist woodrats (Haley, Lamb, Franklin, Constance, & Dearing, 2007; Magnanou, Malenke, & Dearing, 2009; Skopec, Haley, & Dearing, 2007). Purified woodrat CYP2B enzymes display high affinities towards specific terpenes found in juniper, suggesting the importance of these enzymes in the metabolism of juniper PSMs (Wilderman, Shah, Jang, Stout, & Halpert, 2013; Wilderman et al., 2014). Finally, woodrats have multiple gene copies of *CYP2B* that encode enzymes with varying substrate specificities (Malenke et al., 2012; Shah et al., 2016).

To advance our understanding of the role of CYP2B enzymes with respect to dietary specialization, we compared *CYP2B* diversity in a juniper specialist, Stephens' woodrat (*Neotoma stephensi*), to that of a generalist, the white-throated woodrat (*N. albigula*). The specialist consumes almost exclusively one-seeded juniper (*Juniperus monosperma*) foliage (80%–95% of diet; Dial, 1988; Vaughan, 1982), which is abundant in terpenes, particularly α-pinene, whereas the generalist consumes far less juniper (18%–34% of diet; Dial, 1988). The balance of the generalist's diet does not contain significant quantities of other plants with high levels of terpenes (Dial, 1988). For example, the generalist's diet may also include around 29% yucca, and 3%–7% of other plants, such as rabbitbrush, sumac, apache plume, sage, saltbush and ephedra (Dial, 1988;). In some portions of the generalist's range, cacti may constitute up to 80% of its diet (Orr, Newsome, & Wolf, 2015).

Implicit in the biochemical trade-off hypothesis is that specialists will have less variation in their detoxification machinery, compared to generalists. We tested this prediction by isolating, sequencing and analysing cDNA of *CYP2B* in both species. To test the enzyme quantity hypothesis and the associated prediction that the specialist will possess more *CYP2B* gene copies than the generalist, we estimated *CYP2B* gene copy numbers. Finally, we generated phylogenies to place the *CYP2B* sequence variation within the context of the evolutionary history of these genes.

2 | MATERIALS AND METHODS

2.1 | Woodrat capture and husbandry

The specialist (*Neotoma stephensi*) was collected on Woodhouse Mesa, Coconino County, Arizona, near Flagstaff, Arizona, USA (35°30′N, 111°27′W), whereas the generalist (*Neotoma albigula*) was collected in Castle Valley, Grand County, Utah, USA (38°30′N, 109°18′W). These habitats are floristically similar, as both are categorized as Great Basin desert, which is abundant in juniper. Adult woodrats were caught in Sherman live traps (7.62 × 89 × 22.86 cm) and transported to the University of Utah Department of Biology Animal Facility. All individuals were housed singly in solid-bottom shoebox cages (48 × 27 × 20 cm) containing shavings and a plastic tube. The animal room was maintained at approximately 28°C, average humidity 15%–20% and a constant light/dark cycle at 12L/12D (Kohl, Weiss, Cox, Dale, & Dearing, 2014; Skopec, Kohl, Schramm, Halpert, & Dearing, 2015). Woodrats consumed water ad libitum and were fed high-fibre rabbit chow (Harlan Teklad formula 2031). All procedures had prior approval

from the University of Utah Animal Care and Use Committee (IACUC 10-01013 and 12-12010).

2.2 | Woodrat diets

Enzymes in the CYP2B subfamily are inducible (Nannelli, Chirulli, Longo, & Gervasi, 2008; Pustylnyak, Pivovarova, Slynko, Gulyaeva, & Lyakhovich, 2009). Therefore, to ensure that there was sufficient *CYP2B* transcript for cloning, expression was induced by feeding juniper to the animals. On the day of diet presentation, frozen juniper foliage collected from the Arizona site was ground with dry ice and added to the chow. All generalist animals were fed a diet of 30% juniper and 70% rabbit chow for 3 days (Harlan Teklad formula 2031 in ground form; Table 1). This percentage of juniper was selected because it represents the natural diet, whereas higher doses of juniper result in depressed food intake (Dial, 1988; Sorensen, McLister, & Dearing, 2005). In contrast, the specialist is capable of maintaining body mass in the laboratory on much higher levels of juniper (Torregrossa, Azzara, & Dearing, 2011). Therefore, in addition to being fed 30% juniper ($N = 1$), the specialists were fed higher levels of juniper to better represent their range of juniper ingestion in the wild (Dial, 1988). Three animals were fed 60% juniper (Table 1), and the animals were dispatched after 3 days on the juniper diets. Immediately after capture, one specialist individual was maintained on an exclusively juniper diet (100%) for 1 day before dispatch. To determine *CYP2B* gene copy number, liver samples were taken from additional animals (Table 1). A set of liver samples from each individual was placed in RNAlater (Thermo Fisher Scientific) and stored in −80°C, whereas another set of liver samples was placed directly into microfuge tubes and stored in −80°C until analysis.

2.3 | CYP2B cloning & sequencing

Total RNA was isolated from 16 to 27 mg of generalist and specialist ($N = 5$ individuals per species) liver tissue (RNAqueous®-4PCR Kit, Ambion). Liver tissue was preserved in RNAlater (Thermo Fisher Scientific) and stored at −80°C until extraction of total RNA. The isolated RNA served as template for *CYP2B* cDNA synthesis using a specific reverse primer (Table 2, High-Capacity cDNA Reverse Transcription Kit, Thermo Fisher Scientific). *CYP2B* cDNA was amplified in PCR with Herculase II Fusion DNA Polymerase (Agilent Technologies). Primers L6ab and H7 (Table 2) had been previously designed to study *CYP2B* variants in the desert woodrat (*N. lepida*), using *CYP2B* cDNA sequence alignments from the Norway rat (*Rattus norvegicus*) and house mouse (*Mus musculus*) in GENBANK (Malenke et al., 2012). A range of PCR conditions were applied for the amplification of generalist and specialist *CYP2B* cDNA: 95°C: 1–5′, followed by 95°C: 20–30″, 62°C: 30–40″, 68°C: 1′30″–2′ for 35–40 cycles. Final extension was performed at 68°C for 0–4′. Because Herculase II Fusion DNA Polymerase generates blunt-ended PCR products, 3′ A overhangs were added with Taq Polymerase (Thermo Fisher Scientific) to PCR products purified with the GeneJET PCR Purification Kit (Thermo Fisher Scientific). Purified PCR products were then introduced into the Topoisomerase I-activated pCR™2.1-TOPO® vector (Thermo Fisher Scientific). TOP10 cells (One Shot® Chemically Competent™ *E. coli* cells, Thermo Fisher Scientific) were transformed with pCR™2.1-TOPO® vector containing ligated *CYP2B* cDNA, and blue-white colony screening was used to identify TOP10 colonies with *CYP2B* cDNA. Plasmid DNA (*CYP2B*-pCR2.1-TOPO) was isolated from positive (white) TOP10 colonies with the QIAprep® Spin Miniprep Kit (Qiagen) or ZR Plasmid Miniprep–Classic Kit (Zymo

TABLE 1 Summary of analysed *CYP2B* clones in individual woodrats (genus *Neotoma*)

Species	ID	Sex	Diet (% juniper)	# analysed *CYP2B* clones	# unique *CYP2B* clones/individual	# SRS variants
N. albigula	658	F	30	12	10	9
N. albigula	663	M	30	32	18	11
N. albigula	666	F	30	14	12	8
N. albigula	670	M	30	16	14	11
N. albigula	780	F	30	13	13	10
N. albigula	784	M	30	N/A	N/A	N/A
N. albigula	925	F	30	N/A	N/A	N/A
N. stephensi	703	M	85–100	28	19	11
N. stephensi	708	M	85–100	N/A	N/A	N/A
N. stephensi	711	F	85–100	N/A	N/A	N/A
N. stephensi	716	M	60	12	6	5
N. stephensi	728	F	60	16	10	7
N. stephensi	737	M	30	13	4	3
N. stephensi	739	F	60	11	7	4

To determine *CYP2B* sequence variation and quantify gene copy number, five and seven individuals from each woodrat species were used, respectively. SRS variants are CYP2B variants classified according to the identity of key "substrate recognition site" residues in the enzyme active site. N/A —not applicable; refers to individuals used only for analysis of *CYP2B* gene copy number.

TABLE 2 List of woodrat primer names, sequences and associated uses

Primer name	Primer sequence (5′→3′)	Use
NL_CYP2B_H7 (reverse)	GACACCTGGCCACCTCAG	Synthesis of *CYP2B* cDNA
NL_CYP2B_L6ab (forward)	GRYCASACCAGGACCATGGRG	Amplification of *CYP2B* cDNA
NA_NS_CYP2B_L6d (forward)	TGRYYASACCAGGACCATG	Amplification of *CYP2B* cDNA
NA_NS_CYP2B_L6d2 (forward)	TGRYTACACCAGGACCATG	Amplification of *CYP2B* cDNA
NA_NS_CYP2B_H7b (reverse)	ACCTGGCCACCTCAGCCCAG	Amplification of *CYP2B* cDNA
NA_CYP2B_exon1 (forward)	TGGGCTTTGTGCTGCTCC	qPCR primers for *N. albigula CYP2B*
NA_CYP2B_exon1 (reverse)	GCCTCCTCTGTCCATCTGC	qPCR primers for *N. albigula CYP2B*
NS_CYP2B_exon8 (forward)	TCCTGAGTTCAGCTCTCCATG	qPCR primers for *N. stephensi CYP2B*
NS_CYP2B_exon8 (reverse)	GCATCCAGGAAGTGGTCAGG	qPCR primers for *N. stephensi CYP2B*
NA_SOD1_exon4_3b (forward)	TGTGGGAGACCTGGGC	qPCR primers for *N. albigula SOD1*
NA_SOD1_exon4 _H4e1 (reverse)	CAATGGTCTCCTGAGAGTGAG	qPCR primers for *N. albigula SOD1*
NS_SOD1_exon4 _2b (forward)	GCATGTGGGAGACCTGG	qPCR primers for *N. stephensi SOD1*
NS_SOD1_exon4 _3b (reverse)	GAGATCACATGATCTTCAATGG	qPCR primers for *N. stephensi SOD1*
NA_NS_SOD1_deg1 (forward)	GAYATCAKTGCTTATCCACC	Amplification of internal portion of *SOD1*
NA_NS_SOD1_deg1 (reverse)	GTCATCTTSTTTCTCRTGGAC	Amplification of internal portion of *SOD1*

Research), and sequenced at the University of Utah DNA Sequencing Core Facility (Salt Lake City, Utah), Eton Bioscience Inc. (San Diego, California), or Molecular Cloning Laboratories (South San Francisco, California). *CYP2B* nucleotide sequences were assembled with the ApE application (http://biologylabs.utah.edu/jorgensen/wayned/ape/) and translated into amino acid sequences using the ExPASy Bioinformatics Resource Portal (http://www.expasy.org/). To estimate *CYP2B* diversity within an individual, we sequenced clones from one individual of each species more extensively, whereas four other individuals of each species had fewer clones sequenced (Table 1). A *Z* test for proportions was applied to determine whether there were significant differences in the proportion of unique *CYP2B* sequences between the two species (Sokal & Rohlf, 1995; Zar, 1999).

2.4 | Identification of CYP2B substrate recognition site variants

CYP2B enzyme substrate specificity is governed by substrate recognition sites (SRSs; Gotoh, 1992). Mutation of ten key residues in these SRSs alters substrate recognition or turnover in CYP2B enzymes (Domanski & Halpert, 2001; He, Harlow, Szklarz, & Halpert, 1998; Huo et al., 2017; Kumar, Chen, Waxman, & Halpert, 2005; Strobel & Halpert, 1997; Wilderman et al., 2014). We categorized CYP2B clones into SRS variants according to amino acid identity at nine of these residue positions (Table 3). CYP2B SRS variants were then organized based on the volume of amino acid side chain at each key SRS residue position (Zamyatnin, 1972), beginning with position 101—from larger isoleucine (Ile) to smaller valine (Val)—and then proceeding with the residues (from larger to smaller) at the remaining eight key SRS residue positions (Table 3).

We used a Student's *t* test (SPSS version 20) to determine whether the specialist and generalist differed significantly in their cumulative side chain volumes at SRS residue positions 101, 104, 108, 114 and 209. In addition, we calculated the active site volume (ASV) of homology models for each CYP2B SRS variant. Using the X-ray crystal structures of woodrat CYP2B35 (PDB ID: 5E58, www.rcsb.org) and CYP2B37 (PDB ID: 5E0E), homology models were made for each CYP2B SRS variant using MODELLER 9.18 (Webb & Sali, 2016), and the ASV was calculated using VOIDOO (Kleywegt & Jones, 1994; Uppsala Software Factory, Uppsala, Sweden).

2.5 | CYP2B sequence analysis

All cloned *CYP2B* nucleotide sequences were imported into GENEIOUS version 9.1.7 (Biomatters, available from http://www.geneious.com). As these sequences potentially represented multiple gene copies, we used an iterative approach to first bin all *CYP2B* nucleotide sequences into putative homologous sets to avoid analysing alignments of genes representing paralogous genes with independent and discordant evolutionary histories. We first aligned all sequences using the MUSCLE (Edgar, 2004) plugin in GENEIOUS, calculated a distance matrix based on aligned sequence similarity and then parsed the sequences into segregate copies based on these distances. *CYP2B* clone sequences were binned into the following groups: 97.0% identical within individuals, 93% identical among individuals of the same species and 88% identical among individuals of different species. This categorization is intended to represent putative alleles of the same gene copy, or recently duplicated homologs. We recognize that this estimate is imperfect because some *CYP2B* genes share products that are 98% or more identical; for example, the products of *CYP2B1* and *CYP2B2* in the rat can be more than 98% identical (Nelson, 2009). Nonetheless, this method reduces the sequence complexity to illustrate higher-level patterns related to evolutionary history. Alignments of putative individual *CYP2B* gene copies were

TABLE 3 Summary of woodrat specialist and generalist CYP2B SRS variants [Colour table can be viewed at wileyonlinelibrary.com]

	Key CYP2B substrate recognition site residues											
Variants	101	104	108	114	209	290	363	367	477	CSCV (Å^3)	ASV (Å^3)	# individuals
1	I	I	F	I	I	I	I	V	F	856.7	191.4	3 NA
2	I	I	F	I	I	I	A	A	F	856.7	253.1	1NA
3	I	I	F	I	I	N	A	A	I	856.7	285.0	1 NA
4	I	I	I	I	I	I	I	V	I	833.5	200.3	1 NS
5	I	I	I	I	I	I	A	A	F	833.5	240.8	1 NS
6	I	I	I	I	I	I	A	A	I	833.5	285.7	1 NS
7	I	I	I	I	I	N	A	A	F	833.5	242.8	1 NS
8	I	I	I	I	I	N	A	A	I	833.5	320.0	1 NS
9	I	I	I	I	M	I	I	V	F	829.7	218.5	1 NS
10	I	I	I	I	M	I	A	A	I	829.7	297.0	2 NS
11	I	I	I	I	M	N	I	V	F	829.7	184.2	1 NS
12	I	I	I	I	M	N	A	A	F	829.7	271.0	1 NS
13	I	I	I	I	M	N	A	A	I	829.7	284.0	5 NS
14	I	I	I	I	S	I	I	V	F	755.8	203.1	1 NS
15	I	I	I	I	S	N	A	A	I	755.8	279.9	1 NS
16	I	V	F	I	I	I	I	V	F	830	206.9	1 NA + 3 NS
17	I	V	F	I	I	I	V	V	F	830	193.5	1 NS
18	I	V	F	I	I	I	A	A	I	830	294.3	1 NS
19	I	V	F	I	M	I	I	V	F	826.2	186.6	1 NS
20	I	V	F	I	M	I	A	A	F	826.2	273.7	1 NS
21	I	V	F	I	T	I	I	V	F	779.4	228.8	1NA
22	I	V	F	I	T	N	I	V	F	779.4	233.6	1 NA
23	I	V	F	V	I	I	I	V	F	803.3	260.7	5 NA
24	I	V	F	V	I	I	I	V	I	803.3	254.5	1 NA
25	I	V	F	V	I	I	A	A	I	803.3	326.9	1 NA
26	I	V	F	V	I	N	A	A	I	803.3	379.7	1 NA
27	I	V	F	V	M	N	A	A	F	799.5	308.7	1 NA
28	I	V	F	V	V	I	I	I	F	776.6	217.5	1 NA
29	I	V	F	V	T	I	I	V	F	752.7	250.7	1 NA
30	V	I	I	I	I	I	I	V	F	806.8	204.1	2 NA
31	V	I	I	I	M	I	I	V	F	803	191.7	1 NA
32	V	I	I	I	M	I	A	A	I	803	300.1	2 NA
33	V	I	I	I	M	N	I	V	F	803	192.8	1 NA
34	V	I	I	I	M	N	A	A	F	803	287.1	3 NA
35	V	I	I	I	M	N	A	A	I	803	303.2	4 NA
36	V	V	F	I	I	I	I	V	F	803.3	208.2	3NA + 1 NS
37	V	V	F	I	I	I	I	V	I	803.3	191.7	1 NA
38	V	V	F	I	I	N	I	V	I	803.3	260.3	1 NS
39	V	V	F	I	M	I	I	V	F	799.5	166.0	1 NA
40	V	V	F	I	M	N	I	V	F	799.5	210.3	1 NA
41	V	V	F	I	M	N	A	A	F	799.5	268.9	1 NA
42	V	V	F	I	M	N	A	A	I	799.5	306.6	1 NA
43	V	V	F	I	T	I	I	V	F	752.7	209.2	5 NA
44	V	V	F	I	T	I	I	V	I	752.7	253.1	1 NA
45	V	V	F	I	T	I	A	A	I	752.7	327.9	2 NA

(Continues)

TABLE 3 (Continued)

	Key CYP2B substrate recognition site residues											
Variants	101	104	108	114	209	290	363	367	477	CSCV (Å^3)	ASV (Å^3)	# individuals
46	V	V	F	I	T	N	A	A	F	752.7	262.1	1 NA
47	V	V	F	I	S	I	I	V	F	725.6	237.7	4 NS

Woodrat CYP2B SRS (substrate recognition site) variants are organized according to the volumes of their side chains at key SRS residue positions. CSCV represents "cumulative side chain volume" for CYP2B residues 101, 104, 108, 114 and 209. ASV is active site volume of CYP2B SRS variants. Variants that share the same side chains in first four or five SRS residues are shaded in a particular colour, which relates the variant to CYP2B homologs in Figure 3 and Table S1. NA = *N. albigula,* NS = *N. stephensi.*

then subjected to phylogenetic analysis when they included four or more sequences representing at least two individual samples. Phylogenetic analysis was conducted under a Maximum-likelihood framework using RAXML version 8 (Stamatakis, 2014). All analyses included *Homo sapiens* and *M. musculus* as outgroups, as well as representative sequences of *N. lepida*. We then constructed a representative species tree based on the coalescence of the topologies of individual putative homologs using *BEAST version 2.4.0 (Bouckaert et al., 2014). We ran the MCMC for 100 million generations and eliminated 20% of the posterior as burn-in. In the species tree, homologs were associated with the CYP2B SRS variants they encode, which were grouped and colour-coded according to the identity of their first four or five key SRS residues (Tables 3 and S1; Figure 3).

2.6 | Estimation of CYP2B gene copy number

We used quantitative PCR (qPCR) to estimate *CYP2B* gene copy number in generalist and specialist genomic DNA ($N = 7$ individuals per species). Genomic DNA was isolated from 10 to 15 mg of woodrat liver tissue (Quick-DNA Universal Kit, Genesee Scientific, and Quick-gDNA MiniPrep Kit, Zymo Research), which had been stored at −80°C. Using our alignments of 60 generalist and 41 specialist woodrat unique *CYP2B* cDNA sequences (MUSCLE, http://www.ebi.ac.uk/Tools/msa/muscle/), we designed primers to amplify an 113 base pair (bp) fragment of exon 1 in generalist *CYP2B* and a 76-bp fragment of exon 8 in specialist *CYP2B* (Tables 2 and S2). *CYP2B* copy number was estimated relative to a single copy reference gene–superoxide dismutase 1 (*SOD1*). *SOD1*, also known as Cu/Zn superoxide dismutase, has been identified as a single copy gene in various species, including human (Levanon et al., 1985), rat (Puga & Oates, 1987), mouse (Gu, Morales, & Hecht, 1995), helminths *Taenia solium* and *T. crassiceps* (Parra-Unda, Vaca-Paniagua, Jiménez, & Landa, 2012) and parasitic fungus *Cordiceps militaris* (Park et al., 2005). To design primers to amplify fragments of *SOD1* in qPCR, exon 4 of woodrat *SOD1* was first amplified using degenerate primers and then sequenced (Table 2). The forward and reverse *SOD1* degenerate primers, with sequences based on alignments of *N. lepida*, *R. norvegicus* and *M. musculus SOD1*, annealed to intron 3 and exon 5 of woodrat *SOD1*, respectively. Primers were then designed for amplification of *SOD1* exon 4 in generalist and specialist to generate 93- and 78-bp *SOD1* amplicons in qPCR, respectively (Tables 2 and S2). Reactions for qPCR, in triplicate for each individual, were prepared in 96-well plates using 2× Apex qPCR GREEN Master Mix (Apex BioResearch Products), which was followed by cycling in the CFX96 Real-Time System C1000 Touch Thermal Cycler (Bio-Rad) at 95°C: 10′; then 95°C: 30″ and 60°C: 30″ for 40 cycles. Efficiencies of qPCR for *CYP2B* and *SOD1* primer pairs were between 92 and 109% (Livak & Schmittgen, 2001). Using the $2^{-\Delta C'T}$ method (Livak & Schmittgen, 2001), we evaluated *CYP2B* copy number in seven individuals of each woodrat species; five of those individuals had previously been used to obtain *CYP2B* sequences. We observed a single melt peak for each qPCR sample, as well as single bands on a 1.2% agarose gel for selected qPCR samples from each species, indicating amplification of a single product. Sequences of *CYP2B* and *SOD1* amplicons were confirmed for a representative individual of each species. Differences in mean *CYP2B* gene copy number between the two woodrat species were evaluated with an independent samples *t* test in SPSS (version 20). Specialist *CYP2B* and *SOD1* amplicons generated in qPCRs had comparable lengths and GC content, which is a key requirement for valid quantification of relative gene copy number when using SYBR Green as a fluorescent dye. Although the generalist *CYP2B* amplicons had 22% greater length and 12%–22% greater GC content than the generalist *SOD1* amplicon, such small differences in target and reference amplicon characteristics have either little (Colborn, Byrd, Koita, & Krogstad, 2008) or no (Rutledge & Stewart, 2010; Spandidos et al., 2008) impact on the determination of *CYP2B* gene copy number relative to *SOD1* reference gene.

3 | RESULTS

3.1 | Woodrat CYP2B characterization

The cDNA obtained for *CYP2B* was 1,476 bp from start to stop codon and translated to a protein sequence of 491 amino acid residues, as was previously found for desert woodrat, *N. lepida* (Malenke et al., 2012). The best *CYP2B* match in the nucleotide database at NCBI for both species was *N. lepida*. Coverage ranges for generalist and specialist *CYP2B* sequences were 95%–98% and 95%–97%, respectively, with *e*-values of 0.0.

3.2 | Evaluation of CYP2B sequence variation

We obtained high-quality sequences for 87 *CYP2B* cDNA clones from the generalist and 80 clones from the specialist. The number

TABLE 4 Summary of CYP2B sequence diversity found in each woodrat species

Species	# analysed *CYP2B* clones	# unique *CYP2B* nucleotide sequences	# unique CYP2B amino acid sequences	# SRS variants	# SRS variants in >1 individual
N. albigula	87	60	57	27	8
N. stephensi	80	41	39	18	3
Shared	N/A	0	0	2	2
Total	167	101	96	47	13

of *CYP2B* clones sequenced per animal varied and is provided in Table 1. The specialist and generalist differed significantly with respect to diversity of CYP2B sequences at both the nucleotide and amino acid levels. The generalist displayed greater diversity in *CYP2B* nucleotide sequences with 69% (60) of the clones representing unique sequences vs. 51% (41 clones) in the specialist (p = .009, Z = 2.3393, 1-tailed; Table 4). Similarly, there was more diversity in the CYP2B amino acid sequences of the generalist with 66% (57 clones) of unique CYP2B clones vs. 49% (39 clones) of those in specialist (p = .014, Z = 2.1896, 1-tailed; Table 4). No CYP2B nucleotide or amino acid sequences were shared between the two species.

3.3 | Evaluation of CYP2B SRS variants

Based on the identity of amino acids in nine SRS residue positions, we found a total of 47 CYP2B SRS variants (Table 3); some variants include several nucleotide and amino acid sequences that differ from each other outside the SRSs. Only 13 of these SRS variants were detected in more than one individual, whereas the remaining variants were found in single individuals. Two CYP2B SRS variants (16 and 36) were shared between the species (Table 3). The specialist had 18 CYP2B SRS variants, whereas the generalist had 27, but this was not significantly different (p = .107, Z = 1.24118, 1-tailed). Most specialist CYP2B SRS variants possessed cumulatively larger side chains at residues 101, 104, 108, 114 and 209, compared to those of the generalist. The sum of side chain volumes at these five sites (cumulative side chain volume, CSCV) was significantly different between the two woodrat species (p = .03, 1-tailed), with the CSCV of specialist (mean ± *SE*: 815.1 ± 6.9 Å^3) being larger than that of generalist (mean ± *SE*: 797.9 ± 5.2 Å^3) (Table 3).

The active site volume (ASV) of CYP2B SRS variants did not significantly differ between the woodrat species (generalist [mean ± *SE*: 251.1 ± 9.5 Å^3] vs. specialist [mean ± *SE*: 244.4 ± 9.2 Å^3], p = .69, 2-tailed; Table 3). However, ASV significantly differed in CYP2B SRS variants depending on the presence or absence of Ala at residues 363 and 367 (Ala mean ± *SE*: 290.8 ± 6.5 Å^3 vs. non-Ala mean ± *SE*: 214.2 ± 5.2 Å^3, $p = 3.0 \times 10^{-11}$, 2-tailed; Figure 1). ASV also depends on the presence or absence of Ala at these residues in CYP2B SRS variants of the same species (generalist: Ala mean ± *SE*: 300.8 ± 9.9 Å^3, vs. non-Ala mean ± *SE*: 216.0 ± 6.6 Å^3, $p = 6.3 \times 10^{-7}$; specialist: Ala mean ± *SE*: 278.9 ± 7.6 Å^3, vs. non-Ala mean ± *SE*: 209.9 ± 7.5 Å^3, $p = 4.2 \times 10^{-6}$).

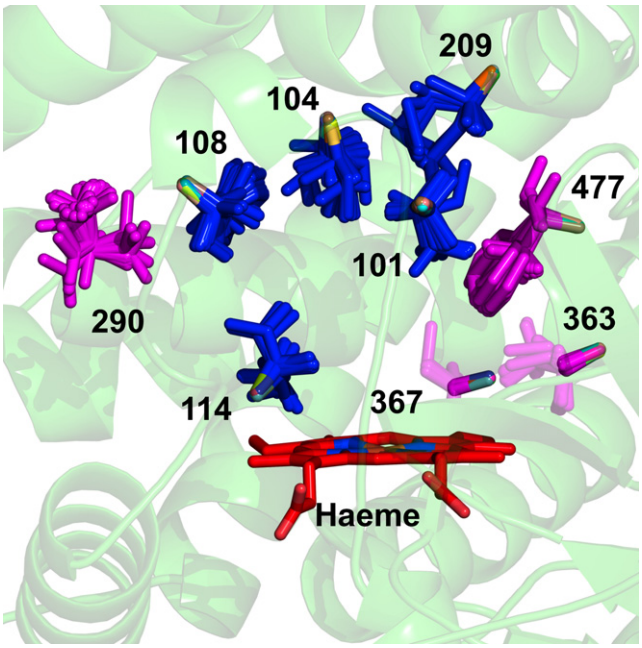

FIGURE 1 Composite homology model of the active site with key SRS residues in woodrat specialist and generalist CYP2Bs. Composite model for active site of woodrat CYP2B variants is based on X-ray crystallography data for *N. lepida* CYP2B35 and CYP2B37 enzymes. Each CYP2B SRS (substrate recognition site) residue position above displays superimposed residues found in corresponding SRS residue positions of all 47 CYP2B generalist and specialist woodrat variants. At positions 363 and 367, solid colours emphasize Ala residues, whereas muted colours represent Ile and Val [Colour figure can be viewed at wileyonlinelibrary.com]

In addition, we identified CYP2B variants with polar residues at residue positions 209 and 290 in both species. To date, only hydrophobic residues have been found at residue position 209 in CYP2B sequences of other species, and dog CYP2B11 Asp290 is the only known hydrophilic residue found at residue position 290 (Malenke et al., 2012; Nelson, 2009; Okamatsu et al., 2017). Three specialist CYP2B SRS variants (CYP2B SRS variants 14, 15, 47; Table 3) contained a serine (Ser) at residue position 209, whereas seven generalist CYP2B SRS variants (CYP2B SRS variants 21, 22, 29, 43-46, Table 3) contained a threonine (Thr) at the same position. We found seven specialist and eleven generalist CYP2B variants with an asparagine (Asn) at residue position 290 (Table 3).

A sequence similar to CYP2B37, with a Val at residue position 114 (Malenke et al., 2012), was found only in the generalist and

included a variant shared among all five tested generalist individuals (CYP2B SRS variant 23, Table 3). Val 114 was also found in six other generalist CYP2B SRS variants (SRS variants 24-29, Table 3), but not in any specialist CYP2B SRS variants, which contained Ile at this site.

3.4 | CYP2B sequence analysis

We estimated a minimum of 17 putative homologous sets of CYP2B genes across both *Neotoma* species based on the distance metrics we used to bin homologous sequences. We produced phylogenies for five of these purported homologs (see homolog 1 in Figure 2; homologs 2, 3, 6 and 7 in Figures S1–S4). *CYP2B* gene copies within a species were generally recovered as monophyletic (or paraphyletic in two instances), although putative alleles sequenced within the same individual were not necessarily each other's closest relatives. The coalescent species tree recovered both species as monophyletic with strong support (Figure 3). *Neotoma lepida* was recovered as sister to a clade containing generalist *N. albigula* and specialist *N. stephensi*. The number of putative homologous loci varied among individuals, with 5–10 loci expressed in each generalist individual and 1–8 loci expressed in each specialist individual. A single locus (homolog 6) was expressed in all individuals.

Comparisons of CYP2B SRS variant and homolog associations revealed the following patterns: first, there was clear differentiation in the SRS patterns within a homolog between species. For example, in the case of homolog 3, all individuals within a species shared the similar SRS variants; however, these variants differed between the specialist and the generalist (Figure 3; Table S1). This pattern in replicated for homolog 6 (Figure 3; Table S1). Second, similar CYP2B SRS variants, which share the first five key SRSs, were encoded by different homologs. In the generalist, CYP2B SRS variant 43 was encoded by three different homologs—3, 5 and 9 (individual NA780)—and the similar CYP2B SRS variant 45 was encoded by homolog 11 (individual NA780; dark green boxes, Figure 3; Table S1). Likewise, in the specialist, highly similar CYP2B SRS variants 9 and 10 were encoded by homologs 9 and 6, respectively (individual NS728; yellow boxes, Figure 3; Table S1).

3.5 | Estimation of CYP2B gene copy number

We found that individuals of both species possessed multiple copies of the *CYP2B* gene. Moreover, there was a significant difference in gene copy between species ($t = -3.745$, $df = 12$, $p = .03$). On average, specialists had ~5 more *CYP2B* gene copies per haploid genome compared to generalists (mean gene copy number $\pm SE$: specialist = 12.9 ± 1.0, generalist = 7.9 ± 0.9; Figure 4).

4 | DISCUSSION

The physiological factors that facilitate dietary specialization in mammals are largely unknown. To address this gap in knowledge, we explored the molecular basis for dietary specialization with a detailed study of CYP2B sequence variation in a juniper specialist,

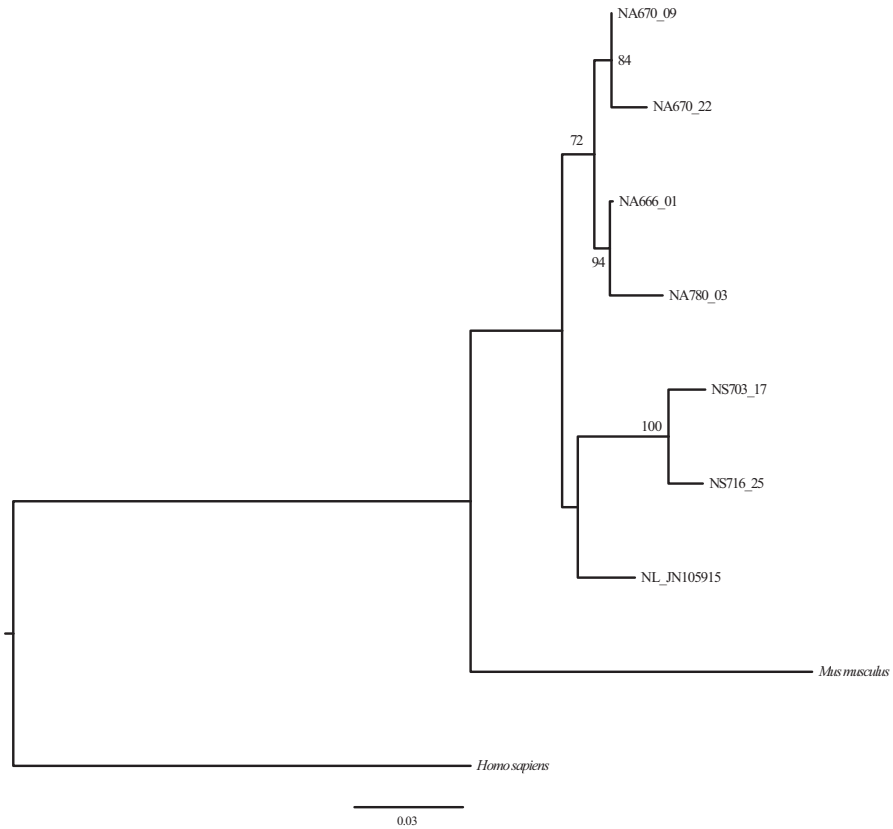

FIGURE 2 Phylogeny of putative *CYP2B* homolog 1 in *N. albigula* and *N. stephensi*. Each *CYP2B* clone is labelled by species name (NA = *N. albigula*, NS = *N. stephensi*, NL = *N. lepida*), followed by the individual's identification number. Listed after each underscore symbol is the *CYP2B* clone identification number for a particular individual

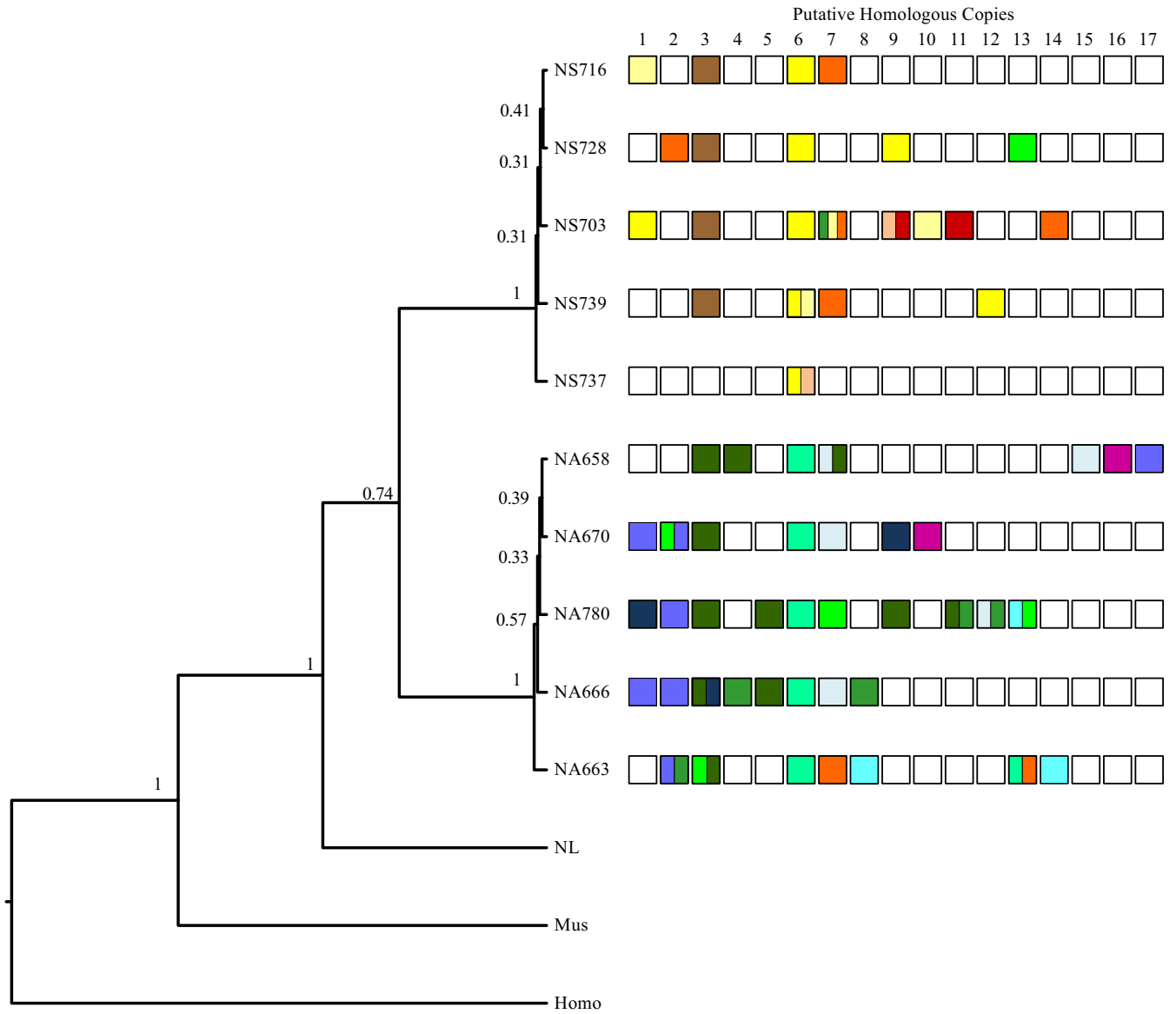

FIGURE 3 Phylogeny of putative *CYP2B* homologs in *N. albigula* and *N. stephensi*. Colour coding in boxes relates *CYP2B* homologs to CYP2B SRS (substrate recognition site) variants in Table 3 and Table S1. Blank boxes indicate that particular homologs are not found in an individual. NA = *N. albigula*, NS = *N. stephensi*, NL = *N. lepida* [Colour figure can be viewed at wileyonlinelibrary.com]

N. stephensi, and generalist, *N. albigula*, by testing the biochemical trade-offs (Freeland & Janzen, 1974) and enzyme quantity hypotheses. In general, our results are consistent with both hypotheses. There appeared to be less variation in the *CYP2B* nucleotide and amino acid sequences of the specialist compared to the generalist. We also found that the specialist had more gene copies of *CYP2B* than the generalist. We discuss the implications of these findings with respect to dietary specialization in the subsequent paragraphs.

4.1 | Structural variability in the CYP2B active site

CYP2B enzymes are known for their plasticity, which allows a single CYP2B active site to accommodate substrates widely ranging in size (Mr ~80–800; Wilderman & Halpert, 2012). However, species-specific CYP2B amino acid sequences, especially in key SRS residues, may affect the CYP2B active site shape and surface area available for substrate binding and processing, possibly underlying these species' abilities to process diets with varying levels and suites of toxins. We found that CYP2B SRS variants organized by SRS residue volume generally clustered according to species. At SRS residue positions 101, 104, 108, 114 and 209, most specialist CYP2B SRS variants possess cumulative side chain volumes greater than those of the generalist (Table 3). This finding suggests that substrates are likely positioned slightly differently in CYP2B enzymes of the generalist vs. the specialist, which likely contributes to different substrate preferences, and perhaps different diet preferences. In contrast, differences in the active site volume of CYP2B SRS variants were not species specific. However, variation in active site volume did correlate to the amino acid identity of SRS residue positions 363 and 367 (Figure 1; Table 3). As the presence or absence of Ala at these two

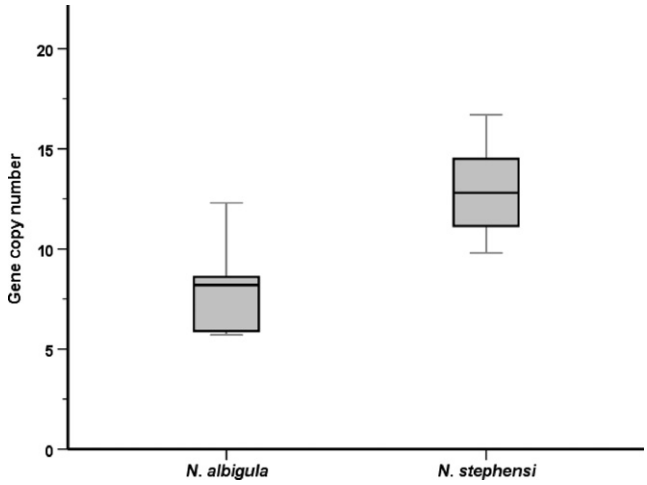

FIGURE 4 Box plot of haploid *CYP2B* gene copy number in woodrats. Haploid *CYP2B* gene copy number was estimated relative to single-copy *SOD1* gene in seven individuals of each woodrat species (*N. albigula* and *N. stephensi*). Mean values for *CYP2B* gene copy number are indicated by horizontal lines within the boxes, whereas the bottom and top 25% of data distribution are indicated with lines extending vertically from the boxes (whiskers)

residue positions led to significant differences in active site volume for all CYP2B SRS variants, as well as within species, we conclude that at least two copies of the *CYP2B* gene were present prior to the split of *N. albigula* and *N. stephensi*. Differences in amino acid identity at SRS residue positions 101, 104, 108, 114 and 209 likely arose after the speciation event. This inference is supported by homologs that are shared between species and possess the same or similar amino acids at SRS residue positions 363 and 367. Indeed, we found that CYP2B variants encoded by homolog 3 (variants 43, 44 and 45 in generalist, and variant 47 in specialist) did not contain Ala residues at these two SRS residue positions, whereas most CYP2B variants encoded by homolog 6 (variants 32, 34, 35, but not 33 in generalist, and variants 8, 10, 12, 13, 15, but not 11 in specialist) contained Ala residues at residue positions 363 and 367 (Table S1).

We previously identified and expressed a *N. lepida* CYP2B enzyme, CYP2B37, which displayed high affinity towards α-pinene, relative to other CYP2B enzymes from *N. lepida* (Malenke et al., 2012; Wilderman et al., 2014). CYP2B37 is characterized by a Val in residue position 114, which plays a key role in high affinity binding to α-pinene (Wilderman et al., 2014). This residue was found in seven generalist CYP2B SRS variants, but not in specialist CYP2B SRS variants, which all contained Ile at this key SRS residue position (Table 3). The specialist is likely to encounter greater concentrations of α-pinene in its diet, relative to the generalist, and would thus require CYP2B enzymes with high affinities for this toxic substrate. We propose that other residues in specialist CYP2B enzymes play a role in high affinity binding of α-pinene, which may be tested in future functional assays.

Additionally, the unique presence of polar residues at SRS residue positions 209 and 290 in several woodrat CYP2B variants may elevate CYP2B enzyme specificity for the terpenoids in their juniper diets. A distinction between some woodrat CYP2B SRS variants occurs at position 209, at which seven generalist CYP2B SRS variants contain Thr, whereas three specialist CYP2B SRS variants contain a Ser (Table 3). Although both Thr and Ser side chains possess the –OH group, disparity in their sizes may differentially affect the binding affinity and processing of substrates.

Functional studies support that single amino acid changes are sufficient to alter CYP2B substrate specificities (Huo et al., 2017; Wilderman et al., 2014). For example, substitution of Ala at SRS residue positions 363 or 367 in *N. lepida* CYP2B35 enzyme with the corresponding residues from CYP2B37 (CYP2B35 A363I and A367V) altered the CYP2B35 7-alkoxycoumarin activity profile towards that of CYP2B37 (Huo et al., 2017). Furthermore, the CYP2B35 double mutant A363I/A367V mimics the 7-alkoxycoumarin substrate specificity profile of CYP2B37. Thus, *CYP2B* sequence variation revealed in this study may underlie critical species-specific differences in CYP2B enzyme function. These results warrant further investigation into the functional differences of these CYP2B SRS variants.

4.2 | Biochemical trade-off hypothesis

Since its publication over 40 years ago, the biochemical trade-off hypothesis of Freeland and Janzen (1974) has inspired substantial research efforts into the molecular underpinnings of dietary specialization. For example, studies of cytochrome P450-mediated detoxification in mammalian specialists have revealed that many PSMs may forego phase II reactions (Iason, 2005; McLean et al., 1993). Additionally, coding sequences and regulation of cytochrome P450 genes have been found to support host-plant specialization (Bono et al., 2008). This study represents the first to explore the Freeland and Janzen hypothesis in herbivorous mammals by comparing cytochrome P450 gene sequence diversity and copy numbers between a mammalian specialist and generalist.

We found that CYP2B nucleotide and amino acid sequences in the specialist woodrat exhibited 26% less variation than those of the generalist (Table 4). This is manifested, for example, in the number of putative homologous loci that were expressed in each species. Of the 17 estimated homologs, 16 were present in generalist, but only 11 in specialist (Figure 2). The relatively low diversity in CYP2B sequences of specialists is consistent with the biochemical trade-off hypothesis (Freeland & Janzen, 1974). Thus, the smaller set of unique CYP2B sequences in the specialist may be highly specialized for the narrow range of toxins found in juniper, such as α-pinene and other monoterpenes, whereas the larger set of unique CYP2B sequences in the generalist may support detoxification of miscellaneous compounds found in their diet.

4.3 | Enzyme quantity hypothesis

Enzyme quantity, which also impacts an organism's ability to process dietary substrates, is often modulated by gene copy number. Dietary specialists may leverage this mechanism, that is high gene copy

number, to increase efficiency of PSM detoxification. The greater copy number of *CYP2B* genes found in the specialist may allow it to express higher quantities of CYP2B enzymes than the generalist, thereby promoting more efficient processing of high PSM levels in the juniper diet. Indeed, expression of *CYP2B* increases in the specialist when fed increasing quantities of juniper (Skopec et al., 2007), which is consistent with the notion that the specialist can increase the quantity of CYP2B enzymes to match the level of PSMs in the diet. However, it is worth noting that regulatory responses, and not gene copy number, can also modulate enzyme quantity as demonstrated in specialist and generalist insects (Li et al., 2003; Li, Zangerl, et al., 2004). The quantity hypothesis warrants further testing with additional approaches, given the limitations of qPCR in estimating copy number of functional genes or gene expression.

The *CYP2B* haploid gene copy numbers estimated in the specialist and generalist are consistent with those reported in another woodrat species, *N. lepida* (mean ± *SE* = 8.6 ± 0.9; Malenke et al., 2012), as well as with *CYP2B* gene copy numbers reported in rat and mouse (Nelson et al., 2004). The rat (*R. norvegicus*) possesses 11 *CYP2B* loci, with 6–7 loci that are functional (Hu et al., 2008). The *CYP2B* gene copy number for the generalist (up to nine copies per haploid genome) corresponds to the number of unique *CYP2B* sequences in the individual most extensively sequenced, NA663 (Table 1), in which 18 unique *CYP2B* nucleotide sequences were found. However, some of these "unique" *CYP2B* sequences, which differ from other *CYP2B* sequences by no more than 1–2 nucleotides, may reflect processing errors. In addition, some unique *CYP2B* sequences may result from alternative splicing of internal duplicated-intact exons (Kondrashov & Koonin, 2001), a process possibly occurring in some *CYPs*, such as *CYP2C39*. However, alternative splicing has not yet been described for any *CYP2Bs* (Nelson et al., 2004). Tandem chimerism also increases protein diversity, but it is unknown whether this process contributes to *CYP2B* variation (Parra et al., 2006).

4.4 | CYP2B evolution in dietary specialists and generalists

We propose that *CYP2B* genes in these two species of woodrats have been under differential selection, with the acknowledgement that this is a limited comparison between two species. Members of the family Cricetidae, including woodrats, diverged from the family Muridae (rats and mice) about 22–25 mya, and likely inherited 1–4 *CYP2B* copies from their most recent common ancestor (Steppan, Adkins, & Anderson, 2004). The greater abundance of *CYP2B* unique nucleotide sequences in the generalist is consistent with the hypothesis that these genes likely evolved via neofunctionalization, a process in which duplicated genes mutate and assume new function, such as altered substrate specificity (Hahn, 2009). In contrast, the *CYP2B* genes in the specialist may have evolved predominantly by conservation of duplicated gene copies. Supporting this hypothesis, the phylogenetic distribution of the presence of putative homologous loci indicates only a single instance where a putative homolog is present in the specialist, but not the generalist; in contrast, six homologs are restricted to the generalist clade (Figure 3).

In both woodrat species, we found that several related CYP2B SRS variants were encoded by different putative homologs, suggesting that some duplicated genes preserved their sequences at the key SRS residue positions, but acquired mutations in other regions of *CYP2B* that were not under selection. Preservation of SRS sequences may enable some duplicated CYP2B enzymes to retain their substrate specificities.

The CYP2B enzymes of the specialist and generalist appear to have diverged from each other possibly as a result of selection for different functions related to their diets. This is supported by the different CYP2B SRS variants encoded by putative homologs; for example, homologs 3 and 6 each encode the same or similar CYP2B SRS variants in almost all individuals within each species, but exhibit different SRS variants between species. Sequence differences in the SRS variants between the species likely result in CYP2B substrate specificities and activities that are unique for each species. Genomic information coupled with functional analyses is necessary to test this hypothesis.

In summary, this study represents one of the first experimental investigations of the biochemical trade-off hypothesis in herbivorous mammals. The results are congruent with the biochemical trade-off hypothesis in that the specialist herbivore appears to have less functional diversity in its CYP2B subfamily of P450 enzymes compared to the generalist. In addition, the specialist has more gene copies of *CYP2B*, which could facilitate higher expression of these enzymes to permit more efficient biotransformation of the elevated levels of PSMs in its diet of juniper. The phylogenetic analyses of putative homologs of *CYP2B* are suggestive of differential selection pressures on these two species, perhaps because of differences in their diet. We recognize that this comparison is limited in its power given that it is restricted to a single specialist and generalist. Moreover, it is important to recognize that the detoxification system consists of many more enzymes than were considered in this study. Further genomic and functional analyses are needed to advance our understanding of dietary specialization in mammalian herbivores and to thoroughly test the biochemical trade-off and enzyme quantity hypotheses.

ACKNOWLEDGEMENTS

We thank Quincy A. Parkes for technical assistance in *CYP2B* cloning and sequence analysis, Dr. James S. Ruff for his insightful suggestions on statistical analysis of the data and Dr. Leslie E. Sieburth for assistance with the qPCR assay. We also thank two anonymous reviewers for suggestions that improved the quality of the manuscript. Support for this study was provided by NSF IOS 1256383 to MDD, NSF IOS 1256840 to JRH, and NSF DEB 1457366 to Dr. Lynn Bohs.

DATA ACCESSIBILITY

N. albigula and *N. stephensi CYP2B* cDNA sequence data are publicly accessible in GenBank and can be found under Accession nos MG459305–MG459417.

AUTHOR CONTRIBUTION

S.K. and M.D.D. designed research; S.K., G.B.C. and T.J.O. performed research; K.S., P.R.W., D.S. and J.R.H. contributed reagents or analytical tools; S.K., M.D.D., G.B.C., T.J.O., D.S. and P.R.W. analysed data; S.K., M.D.D., D.S., T.J.O. and P.R.W. wrote manuscript.

ORCID

Smiljka Kitanovic http://orcid.org/0000-0001-9151-6421
Daniel Spalink http://orcid.org/0000-0003-3250-568X

REFERENCES

Bass, C., Puinean, A. M., Zimmer, C. T., Denholm, I., Field, L. M., Foster, S. P., … Williamson, M. S. (2014). The evolution of insecticide resistance in the peach potato aphid, *Myzus persicae*. *Insect Biochemistry and Molecular Biology*, *51*, 41–51. https://doi.org/10.1016/j.ibmb.2014.05.003

Bono, J. M., Matzkin, L. M., Castrezana, S., & Markow, T. A. (2008). Molecular evolution and population genetics of two *Drosophila mettleri* cytochrome P450 genes involved in host plant utilization. *Molecular Ecology*, *17*(13), 3211–3221. https://doi.org/10.1111/j.1365-294X.2008.03823.x

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, *10*, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Calla, B., Noble, K., Johnson, R. M., Walden, K. K., Schuler, M. A., Robertson, H. M., & Berenbaum, M. R. (2017). Cytochrome P450 diversification and hostplant utilization patterns in specialist and generalist moths: Birth, death and adaptation. *Molecular Ecology*, *26*(21), 6021–6035. https://doi.org/10.1111/mec.14348

Colborn, J. M., Byrd, B. D., Koita, O. A., & Krogstad, D. J. (2008). Estimation of copy number using SYBR Green: Confounding by AT-rich DNA and by variation in amplicon length. *American Journal of Tropical Medicine and Hygiene*, *79*(6), 887–892.

Dearing, M. D., Foley, W. J., & McLean, S. (2005). The influence of plant secondary metabolites on the nutritional ecology of herbivorous terrestrial vertebrates. *Annual Review of Ecology, Evolution and Systematics*, *36*, 169–189. https://doi.org/10.1146/annurev.36.102003.152617

Dial, K. P. (1988). Three sympatric species of *Neotoma*: Dietary specialization and co-existence. *Oecologia*, *76*(4), 531–537. https://doi.org/10.1007/BF00397865

Domanski, T. L., & Halpert, J. R. (2001). Analysis of mammalian cytochrome P450 structure and function by site-directed mutagenesis. *Current Drug Metabolism*, *2*(2), 117–137. https://doi.org/10.2174/1389200013338612

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Freeland, W. J., & Janzen, D. H. (1974). Strategies in herbivory by mammals: The role of plant secondary compounds. *The American Naturalist*, *108*(961), 269–289. https://doi.org/10.1086/282907

Gotoh, O. (1992). Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid coding nucleotide sequences. *The Journal of Biological Chemistry*, *267*(1), 83–90.

Gu, W., Morales, C., & Hecht, N. B. (1995). In male mouse germ cells, copper-zinc superoxide dismutase utilizes alternative promoters that produce multiple transcripts with different translation potential. *The Journal of Biological Chemistry*, *270*(1), 236–243. https://doi.org/10.1074/jbc.270.1.236

Hahn, M. W. (2009). Distinguishing among evolutionary models for the maintenance of gene duplicates. *Journal of Heredity*, *100*(5), 605–617. https://doi.org/10.1093/jhered/esp047

Haley, S. L., Lamb, J. G., Franklin, M. R., Constance, J. E., & Dearing, M. D. (2007). Xenobiotic metabolism of plant secondary compounds in juniper (*Juniperus monosperma*) by specialist and generalist woodrat herbivores, genus *Neotoma*. *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology*, *146*(4), 552–560. https://doi.org/10.1016/j.cbpc.2007.06.007

He, Y.-Q., Harlow, G. R., Szklarz, G. D., & Halpert, J. R. (1998). Structural determinants of progesterone hydroxylation by cytochrome P450 2B5: The role of nonsubstrate recognition site residues. *Archives of Biochemistry and Biophysics*, *350*(2), 333–339. https://doi.org/10.1006/abbi.1997.0516

Hu, S., Wang, H., Knisely, A. A., Reddy, S., Kovacevic, D., Liu, Z., & Hoffman, S. M. G. (2008). Evolution of the CYP2ABFGST gene cluster in rat, and a fine-scale comparison among rodent and primate species. *Genetica*, *133*(2), 215–226. https://doi.org/10.1007/s10709-007-9206-x

Huo, L., Liu, J., Szklarz, G. D., Dearing, M. D., Halpert, J. R., & Wilderman, P. R. (2017). Rational re-engineering of the O-dealkylation of 7-alkoxycoumarin derivatives by cytochromes P450 from the desert woodrat *Neotoma lepida*. *Biochemistry*, *56*(16), 2238–2246. https://doi.org/10.1021/acs.biochem.7b00097

Iason, G. (2005). The role of plant secondary metabolites in mammalian herbivory: Ecological perspectives. *Proceedings of the Nutrition Society*, *64*(1), 123–131. https://doi.org/10.1079/PNS2004415

Ingelman-Sundberg, M. (2005). Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): Clinical consequences, evolutionary aspects and functional diversity. *The Pharmacogenomics Journal*, *5*(1), 6–13. https://doi.org/10.1038/sj.tpj.6500285

Kleywegt, G. J., & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallographica Section D-Biological Crystallography*, *50*(2), 178–185. https://doi.org/10.1107/S0907444993011333

Kohl, K. D., Weiss, R. B., Cox, J., Dale, C., & Dearing, M. D. (2014). Gut microbes of mammalian herbivores facilitate intake of plant toxins. *Ecology Letters*, *17*(10), 1238–1246. https://doi.org/10.1111/ele.12329

Kondrashov, F. A., & Koonin, E. V. (2001). Origin of alternative splicing by tandem exon duplication. *Human Molecular Genetics*, *10*(23), 2661–2669. https://doi.org/10.1093/hmg/10.23.2661

Kumar, S., Chen, C. S., Waxman, D. J., & Halpert, J. R. (2005). Directed evolution of mammalian cytochrome P450 2B1. *The Journal of Biological Chemistry*, *280*(20), 19569–19575. https://doi.org/10.1074/jbc.M500158200

Levanon, D., Lieman-Hurwitz, J., Dafni, N., Wigderson, M., Sherman, L., Bernstein, Y., … Groner, Y. (1985). Architecture and anatomy of the chromosomal locus in human chromosome 21 encoding the Cu/Zn superoxide dismutase. *The EMBO Journal*, *4*(1), 77–84.

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., & Bircher, J. S. (2008). Evolution of mammals and their gut microbes. *Science*, *320*(5883), 1647–1651. https://doi.org/10.1126/science.1155725

Li, X., Baudry, J., Berenbaum, M., & Schuler, M. (2004). Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. *Proceedings of the National Academy of Science of the USA*, *101*(9), 2939–2944. https://doi.org/10.1073/pnas.0308691101

Li, W., Berenbaum, M. R., & Schuler, M. A. (2003). Diversification of furanocoumarin-metabolizing cytochrome P450s in two papilionids: Specificity and substrate encounter rate. *Proceedings of the National Academy of Science of the USA*, *100*(2), 14593–14598. https://doi.org/10.1073/pnas.1934643100

Li, W., Zangerl, A. R., Schuler, M. A., & Berenbaum, M. R. (2004). Characterization and evolution of furanocoumarin inducible cytochrome P450s in the parsnip webworm, *Depressaria pastinacella*. *Insect Molecular Biology*, *13*(6), 603–613. https://doi.org/10.1111/j.0962-1075.2004.00518.x

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Magnanou, E., Malenke, J., & Dearing, M. D. (2009). Expression of biotransformation genes in woodrat (*Neotoma*) herbivores on novel and ancestral diets: Identification of candidate genes responsible for dietary shifts. *Molecular Ecology*, *18*(11), 2401–2414. https://doi.org/10.1111/j.1365-294X.2009.04171.x

Malenke, J. R., Magnanou, E., Thomas, K., & Dearing, M. D. (2012). Cytochrome P450 diversity and dietary novelty in the herbivorous, desert woodrat (*Neotoma lepida*). *PLoS ONE*, *7*(8), e41510. https://doi.org/10.1371/journal.pone.0041510

Marsh, K. J., Wallis, I. R., McLean, S., Sorensen, J. S., & Foley, W. J. (2006). Conflicting demands on detoxification pathways influence how common brushtail possums choose their diets. *Ecology*, *87*(8), 2103–2112. https://doi.org/10.1890/0012-9658(2006) 87[2103:CDODPI]2.0.CO;2

McLean, S., Foley, W. J., Davies, N. W., Brandon, S., Duo, L., & Blackman, A. J. (1993). Metabolic fate of dietary terpenes from *Eucalyptus radiata* in common ringtail possum (*Pseudocheirus peregrinus*). *Journal of Chemical Ecology*, *19*(8), 1625–1643. https://doi.org/10.1007/BF00982297

Nannelli, A., Chirulli, V., Longo, V., & Gervasi, P. G. (2008). Expression and induction by rifampicin of CAR- and PXR-regulated CYP2B and CYP3A in liver, kidney and airways of pig. *Toxicology*, *252*(1–3), 105–112. https://doi.org/10.1016/j.tox.2008.08.004

Nelson, D. R. (2009). The cytochrome P450 homepage. *Human Genomics*, *4*(1), 59–65. https://doi.org/10.1186/1479-7364-4-1-59

Nelson, D. R., Zeldin, D. C., Hoffman, S. M. G., Maltais, L. J., Wain, H. M., & Nebert, D. W. (2004). Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative splice variants. *Pharmacogenetics*, *14*(1), 1–18. https://doi.org/10.1097/00008571-200401000-00001

Ngo, S., Kong, S., Kirlich, A., McKinnon, R. A., & Stupans, I. (2000). Cytochrome P450 4A, peroxisomal enzymes and nicotinamide cofactors in koala liver. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology*, *127*(3), 327–334. https://doi.org/10.1016/S0742-8413(00)00160-2

Okamatsu, G., Komatsu, T., Ono, Y., Inoue, H., Uchide, T., Onaga, T., … Teraoka, H. (2017). Characterization of feline cytochrome P450 2B6. *Xenobiotica*, *47*(2), 93–102. https://doi.org/10.3109/00498254.2016.1145754

Orr, T. J., Newsome, S. D., & Wolf, B. O. (2015). Cacti supply limited nutrients to a desert rodent community. *Oecologia*, *178*(4), 1045–1062. https://doi.org/10.1007/s00442-015-3304-8

Park, N. S., Lee, K. S., Sohn, H. D., Kim, D. H., Lee, S. M., Park, E., … Jin, B. R. (2005). Molecular cloning, expression, and characterization of the Cu, Zn superoxide dismutase (SOD1) gene from the entomopathogenic fungus Cordyceps militaris. *Mycologia*, *97*(1), 130–138. https://doi.org/10.1080/15572536.2006.11832846

Parra, G., Reymond, A., Dabbouseh, N., Dermitzakis, E. T., Castelo, R., Thomson, T. M., … Guigó, R. (2006). Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Research*, *16*(1), 37–44. https://doi.org/10.1101/gr.4145906

Parra-Unda, R., Vaca-Paniagua, F., Jiménez, L., & Landa, A. (2012). Cu, Zn superoxide dismutase: Cloning and analysis of the Taenia solium gene and Taenia crassiceps cDNA. *Experimental Parasitology*, *130*(1), 32–38. https://doi.org/10.1016/j.exppara.2011.10.002

Perry, G. H., Dominy, N. J., Claw, K. G., Lee, A. S., Fiegler, H., Redon, R., … Stone, A. C. (2007). Diet and evolution of human amylase gene copy number variation. *Nature Genetics*, *39*(10), 1256–1260. https://doi.org/10.1038/ng2123

Puga, A., & Oates, E. L. (1987). Isolation and nucleotide sequence of rat Cu/Zn superoxide dismutase cDNA clones. *Free Radical Research Communications*, *3*(6), 337–346. https://doi.org/10.3109/10715768709088074

Pustylnyak, V., Pivovarova, E., Slynko, N., Gulyaeva, L., & Lyakhovich, V. (2009). Species-specific induction of CYP2B by 2,4,6-tryphenyldioxane-1,3 (TPD). *Life Sciences*, *85*(23–26), 815–821. https://doi.org/10.1016/j.lfs.2009.100.015

Rutledge, R. G., & Stewart, D. (2010). Assessing the performance capabilities of LRE-based assays for absolute quantitative real-time PCR. *PLoS ONE*, *5*(3), e9731. https://doi.org/10.1371/journal.pone.0009731

Sezutsu, H., Le Goff, G., & Feyereisen, R. (2013). Origins of P450 diversity. *Philosophical Transactions of the Royal Society B*, *368*(1612), 20120428. https://doi.org/10.1098/rstb.2012.0428

Shah, M. B., Liu, J., Huo, L., Zhang, Q., Dearing, M. D., Wilderman, P. R., … Halpert, J. R. (2016). Structure-function analysis of mammalian CYP2B enzymes using 7-substituted coumarin derivatives and probes: Utility of crystal structures and molecular modelling in understanding xenobiotic metabolism. *Molecular Pharmacology*, *89*(4), 435–445. https://doi.org/10.1124/mol.115.102111

Skopec, M. M., Haley, S., & Dearing, M. D. (2007). Differential hepatic gene expression of a dietary specialist (*Neotoma stephensi*) and generalist (*Neotoma albigula*) in response to juniper (*Juniperus monosperma*) ingestion. *Comparative Biochemistry and Physiology, Part D*, *2*(1), 34–43. https://doi.org/10.1016/j.cbd.2006.11.001

Skopec, M. M., Kohl, K. D., Schramm, K., Halpert, J. R., & Dearing, M. D. (2015). Using the specialization framework to determine degree of dietary specialization in a herbivorous woodrat. *Journal of Chemical Ecology*, *41*(12), 1059–1068. https://doi.org/10.1007/s10886-015-0654-y

Skopec, M. M., Malenke, J., Halpert, J. R., & Dearing, M. D. (2013). An in vivo assay for elucidating the importance of cytochromes P450 for the ability of a wild mammalian herbivore (*Neotoma lepida*) to consume toxic plants. *Physiological and Biochemical Zoology*, *5*, 593–601. https://doi.org/10.1086/672212

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The principles and practice of statistics in biological research*. New York, NY: W.H. Freeman and Company.

Sorensen, J. S., McLister, J. D., & Dearing, M. D. (2005). Plant secondary metabolites compromise the energy budgets of specialist and generalist mammalian herbivores. *Ecology*, *86*(1), 125–139. https://doi.org/10.1890/03-0627

Spandidos, A., Wang, X., Wang, H., Dragnev, S., Thurber, T., & Seed, B. (2008). A comprehensive collection of experimentally validated primers for Polymerase Chain Reaction quantitation of murine transcript abundance. *BMC Genomics*, *9*(663), https://doi.org/10.1186/1471-2164-9-633

Stamatakis, A. (2014). RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Steppan, S., Adkins, R., & Anderson, J. (2004). Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Systematic Biology*, *53*(4), 533–553. https://doi.org/10.1080/10635150490468701

Strobel, S. M., & Halpert, J. R. (1997). Reassessment of cytochrome P450 2B2: Catalytic specificity and identification of four active-site residues. *Biochemistry*, *36*, 11697–11706. https://doi.org/10.1021/bi9710176

Torregrossa, A.-M., Azzara, A. V., & Dearing, M. D. (2011). Differential regulation of plant secondary compounds by herbivorous rodents. *Functional Ecology*, *25*(6), 1232–1240. https://doi.org/10.1111/j.1365-2435.2011.01896.x

Vaughan, T. A. (1982). Stephens' woodrat, a dietary specialist. *Journal of Mammalogy*, *63*(1), 53–62. https://doi.org/10.2307/1380671

Waidyanatha, S., Johnson, J. D., Hong, S. P., Robinson, V. G., Gibbs, S., Graves, S. W., … Smith, C. S. (2013). Toxicokinetics of α-thujone following intravenous and gavage administration of α-thujone or α- and ß-thujone mixture in male and female F344/N rats and B6C3F1 mice. *Toxicology and Applied Pharmacology*, *271*(2), 216–228. https://doi.org/10.1016/j.taap.2013.05.001

Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current Protocols in Protein Science*, *86*, 2.9.1–2.9.37. https://doi.org/10.1002/cpps.20

Wilderman, P. R., & Halpert, J. R. (2012). Plasticity of CYP2B enzymes: Structural and solution biophysical methods. *Current Drug Metabolism*, *13*(2), 167–176. https://doi.org/10.2174/138920012798918417

Wilderman, P. R., Jang, H.-H., Malenke, J. R., Salib, M., Angermeier, E., Lamime, S., … Halpert, J. R. (2014). Functional characterization of cytochromes P450 2B from the desert woodrat *Neotoma lepida*. *Toxicology and Applied Pharmacology*, *274*, 393–401. https://doi.org/10.1016/j.taap.2013.12.005

Wilderman, P. R., Shah, M. B., Jang, H.-H., Stout, C. D., & Halpert, J. R. (2013). Structural and thermodynamic basis of (+)-α-pinene binding to human cytochrome P450 2B6. *Journal of the American Chemical Society*, *135*(28), 10433–10440. https://doi.org/10.1021/ja403042k

Zamyatnin, A. A. (1972). Protein volume in solution. *Progress in Biophysics and Molecular Biology*, *24*, 107–123. https://doi.org/10.1016/0079-6107(72)90005-3

Zanger, U. M., Klein, K., Saussele, T., Blievernicht, J., Hofmann, M. H., & Schwab, M. (2007). Polymorphic CYP2B6: Molecular mechanisms and emerging clinical significance. *Pharmacogenomics*, *8*, 743–759. https://doi.org/10.2217/14622416.8.7.743

Zar, J. H. (1999). *Biostatistical analysis*, 4th ed. Upper Saddle River, NJ: Prentice Hall.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Kitanovic S, Orr TJ, Spalink D, et al. Role of cytochrome P450 2B sequence variation and gene copy number in facilitating dietary specialization in mammalian herbivores. *Mol Ecol*. 2018;27:723–736. https://doi.org/10.1111/mec.14480