Article

Differential Selection on Caste-Associated Genes in a Subterranean Termite

Julianne M. Radford, David Chen, Anna M. Chernyshova, Cambrie Taylor, Alex W. Guoth, Tian Wu, Kathleen A. Hill and Graham J. Thompson *

Department of Biology, Western University, London, ON N6A 5B7, Canada; jradfor8@uwo.ca (J.M.R.); dchen362@uwo.ca (D.C.); achernys@uwo.ca (A.M.C.); ctayl65@uwo.ca (C.T.); aguoth@uwo.ca (A.W.G.); twu075@gmail.com (T.W.); khill22@uwo.ca (K.A.H.)
* Correspondence: graham.thompson@uwo.ca

Simple Summary: Social insects can sometimes boldly invade new habitats, including areas of human habitation where they can become unwanted domestic or agricultural pests. In this study, we use molecular sequence analysis to study genetic patterns associated with the invasion and division of labour in the Eastern subterranean termite *Reticulitermes flavipes*. By studying how genes vary by caste and by population, we show that even termites invasive to a metropolitan city can still harbour plenty of genetic variation, as much or more as native termite populations. We suggest therefore that invasive termites do not necessarily suffer long term loss-of-variation upon invasion. Second, we show that genes associated in their expression with the soldier caste evolve approximately twice as fast as genes expressed by other castes of this species, regardless of from what population the castes were sampled. Why termite soldier genes evolve quickly is not known, but it seems unrelated to invasion or the invaded habitat. Given that soldiers are sterile and thus have no direct fitness, the evidence for gene-level selection in the soldier caste is an intriguing example of kin selection.

Abstract: Analyzing the information-rich content of RNA can help uncover genetic events associated with social insect castes or other social polymorphisms. Here, we exploit a series of cDNA libraries previously derived from whole-body tissue of different castes as well as from three behaviourally distinct populations of the Eastern subterranean termite *Reticulitermes flavipes*. We found that the number (~0.5 M) of single nucleotide variants (SNVs) was roughly equal between nymph, worker and soldier caste libraries, but d_N/d_S (ratio of nonsynonymous to synonymous substitutions) analysis suggested that some of these variants confer a caste-specific advantage. Specifically, the d_N/d_S ratio was high (4.3) for genes expressed in the defensively specialized soldier caste, relative to genes expressed by other castes (~1.7–1.8) and regardless of the North American population (Toronto, Raleigh, Boston) from which the castes were sampled. The populations, meanwhile, did show a large difference in SNV count but not in the manner expected from known demographic and behavioural differences; the highly invasive unicolonial population from Toronto was not the least diverse and did not show any other unique substitution patterns, suggesting any past bottleneck associated with invasion or with current unicoloniality has become obscured at the RNA level. Our study raises two important hypotheses relevant to termite sociobiology. First, the positive selection ($d_N/d_S > 1$) inferred for soldier-biased genes is presumably indirect and of the type mediated through kin selection, and second, the behavioural changes that accompany some social insect urban invasions (i.e., 'unicoloniality') may be detached from the loss-of-diversity expected from invasion bottlenecks.

Keywords: social insect invasions; caste differentiation; kin selection; urban entomology

1. Introduction

The study of social insect biology often uses behavioural genetics theory to better understand the evolution of castes and associated social behavior. Inclusive fitness theory

provides a resident body of work that explains why societies function in a coordinated manner or collapse under disjointed conflict [1,2]. Central to the theory is an argument for gene-level selection, and efforts to test its many predictions are befitting to population genetic data [3,4]. The mass adoption of genomic technologies into the field of social insect biology represents a timely transition towards new types of data that can build upon classic population genetics models to incorporate the broadest patterns of genetic and epigenetic variation [5–7]. This emerging field of sociogenomics has not yet been unequivocally aligned with inclusive fitness thinking [8,9], but there are opportunities to leverage the far-reaching theory against genomic and transcriptomic data sets.

In this study, we take a step in this direction by utilizing a termite transcriptome to examine how the nucleotide composition of expressed gene sets varies as a function of caste and population-level behaviour of the Eastern subterranean termite *Reticulitermes flavipes*. Termites are an all-eusocial clade (Termitoidea; ~3000 spp.) of cockroaches (Blattodea) that are distant relatives to any social Hymenoptera [10,11], and thus any member species presents a phylogenetically understudied opportunity to relate genomic profiles to social differences between individuals or populations [12,13]. The Eastern subterranean termite is a widely-studied species for which genomic resources are becoming more readily available, including a non-annotated draft genome [14]. This temperate North American species [15] is native to the eastern United States and south to contiguous parts of Mexico, but its affinity for human industry and habitation has contributed to its vicariant distribution elsewhere as a highly invasive pest [16,17].

The basic social structure of the Eastern subterranean termite involves instars of both sexes that differentiate along one of two principal pathways, which are: the reproductive pathway from eggs and larva into wing-budded nymphs that can further differentiate into imagine queens and kings, and the non-reproductive pathway from eggs and larva into workers, which can further differentiate into soldiers (Figure 1). Because workers and soldiers have low or no direct fitness of their own and ultimately labour on behalf of reproducing relatives, they can rightly be considered reproductively altruistic [18,19]. As for most social insects, the process of caste differentiation is likely mediated by environmental cues that trigger physiological restrictions on development [20,21]. However, genetic differences too could influence an individual termite's differentiation into one caste or another [22], and do so along either pathway. For *R. flavipes*, colonies are typically headed by a pair of outbred primary reproductives [23], but in invasive habitats the species tends toward a unicolonial structure with spatially diffuse nests, low aggression and rampant neotenic reproduction with potentially hundreds of egg-laying nymphs [24]. These changes to the social biology of invasive *Reticulitermes* may represent an exaggerated subsample of their native-range behaviour [25] but, regardless, unicoloniality makes invasive termites difficult to eradicate, with economic impact in North America scaling up to billions of dollars annually [26].
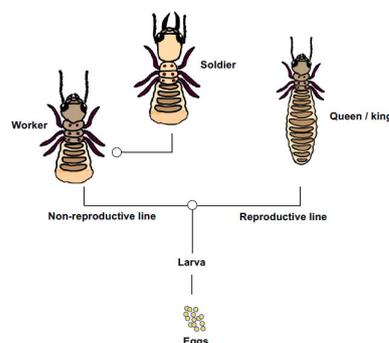

Figure 1. A simplified representation of caste developmental pathway in *Reticulitermes flavipes* [18,27]. Larvae differentiate into non-reproductive (worker and soldier) or reproductive (nymphs, primary queens and kings) castes. Workers can further differentiate into soldiers. These two points of caste differentiation or 'switches' sensu [28] are shown by open circles.

Similar to the process of caste differentiation, this social transition from colonial to unicolonial lifestyle may be mediated by environmental cues, but genetic bottlenecks and other sweeping changes to genetic diversity upon invasion may also play a role, especially if they dampen nestmate recognition cues that in turn lowers aggression [29]. Here, we exploit a series of cDNA libraries previously derived from whole-body tissue of three different castes as well as from behaviourally distinct colonial and unicolonial populations of the Eastern subterranean termite *Reticulitermes flavipes*. From caste- and population-specific libraries initially assembled by Wu et al. [30], we test for differences in expressed gene diversity between one reproductively selfish (wing-padded nymphs) and two reproductively altruistic (worker, soldier) castes. If caste differences are partly explained by genotype, then we expect caste-specific libraries to differ in the number, type or pattern of single nucleotide variants (SNVs). Likewise, if invasive transitions to unicoloniality in *R. flavipes* are associated with genetic bottlenecks or other systemic changes to nucleotide diversity, then we expect termites collected from invasive populations to harbour fewer SNVs or to otherwise show unique mutational profiles compared to their native counterparts. We test these two qualitative predictions by comparing the relative number and type of nucleotide substitutions against a consensus reference assembly.

2. Methods

2.1. Acquiring the RNA-Seq Data Set

The analyzed dataset consisted of nine RNA-Seq libraries, which were generated from three castes (nymph, soldier, worker). The samples include representatives of both sexes and were collected from each of three separate colonies across three geographic populations: one invasive (Toronto, ON, Canada) and two that are within the native range (near Raleigh, NC; near Boston, MA, USA). The details of sample collection, processing and sequencing are described in Wu et al. [30] and in Behl et al. [31]. For the purposes of this study, we accessed the National Center for Biotechnology Information Sequence Read Archive (NCBI-SRA) to download each of the nine termite transcriptome libraries (Table 1). Next, we transferred them onto the bioinformatics platform GALAXY [32] for analysis. Each NCBI-SRA file consisted of paired sets of forward and reverse Illumina HiSeq 2000 sequence reads that correspond to a whole transcriptome of the cDNA template. Alongside each of these ~30 Gb FASTQ files, we used FILEZILLA to transfer the corresponding reference assembly of 29,641 unique transcripts as a multi-FASTA file into the GALAXY environment. These and all other major steps in data analysis are shown in Figure 2.

2.2. Galaxy Workflow

Within GALAXY, we configured a workflow to pre-process and map ~4.3 M sequence reads for automated variant calling. Our workflow adopted industry standard quality criteria [33]; it incorporated checking for Phred quality scores [34] and filtering out input files with incorrect or unreliable nucleotide base sequence calls that are sometimes generated due to the technical limitations of sequencing platforms. After this standardization step, we used the TRINITY method for the *de novo* assembly to map all suitable transcripts to a 14 K-gene reference transcriptome to visualize all aligned regions of reads, as in [35]. Following this, we produced variant call files (VCFs) that we screened for SNVs associated with population or caste. To reduce sequence-length bias, we employed NORMALIZE-FASTA tool to trim, to the same maximum length, all the reads aligned to the reference transcriptome. We further curated our dataset according to recommended practices [33] by implementing the CUTADAPT tool to locate and remove duplicate sequences from the FASTQ files.

The output files with reference transcriptome were then fed into the BWA-MEM mapping tool, which makes use of the Burrow–Wheeler transform [36], and a block-sorting compression algorithm to generate alignments. Next, we removed duplicate sequences using MARKDUPLICATES, then sorted the remaining sequences with the SORTBAM tool and compressed all aligned sequences into sets of binary BAM files. By uploading BAM files

into the Integrative Genomics Viewer, we were able to visualize the reads mapped to the reference. To generate VCFs, we input the normalized reference and aligned-sequence BAM files into FREEBAYES [37], which implements Bayesian criteria to assess the likelihood of a variant call at each position in the transcriptome.

Table 1. Summary statistics for *Reticulitermes flavipes* SNV count by population and caste.

|	SRA File Accession	SNV Count Mean ± SD
Caste		
Nymph	SAMN06579168 SAMN06579169 SAMN06579170	145,824 ± 12,414
Soldier	SAMN06579171 SAMN06579172 SAMN06579173	144,752 ± 14,243
Worker	SAMN06579174 SAMN06579175 SAMN06579176	141,981 ± 20,306
Population		
Raleigh	SAMN06579175 SAMN06579172 SAMN06579169	155,606 ± 870
Boston	SAMN06579174 SAMN06579171 SAMN06579168	126,342 ± 6726
Toronto	SAMN06579176 SAMN06579173 SAMN06579170	150,608 ± 86.4
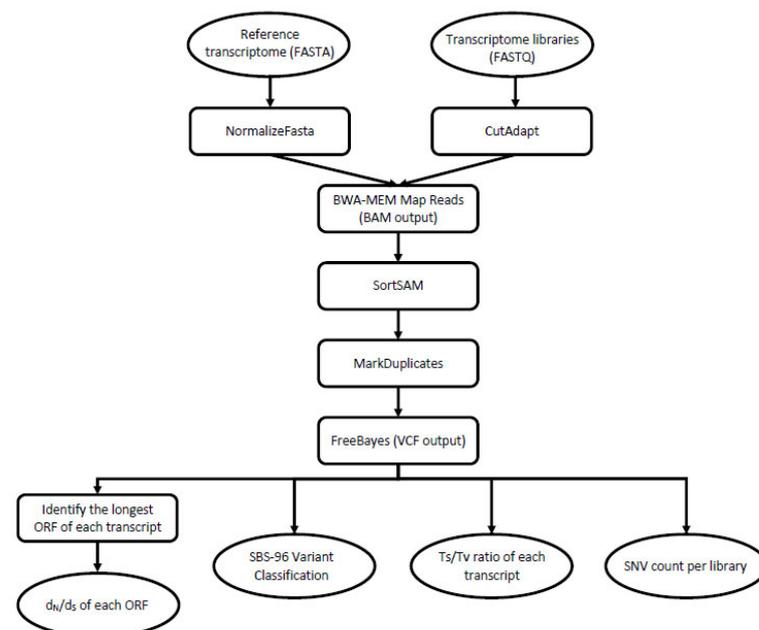
Figure 2. A graphical workflow depicting the major steps that we undertook in data processing. Oval shapes represent input/output, and the rectangular shapes represent the tools/processing steps. To begin, we accessed the nine NCBI-SRA library files (FASTQ format) and the corresponding reference assembly (FASTA format). From there, we used these data files in combination to generate the four principal analyses shown at the bottom of the flow chart.

2.3. Analysis of Single Nucleotide Variation

To analyze generated VCFs, we first imported them into a local UNIX environment and performed analyses stepwise. We first used the SPLITVCFS tool available in the PICARD package of GATK [38] to isolate SNVs and exclude other variant types. Next, we implemented the COUNTVARIANTS tool built-in to GATK [38] to estimate the total number of SNVs for each population or caste-specific library. We then imported the compressed BGZIP and indexed TABIX VCFs into R. Next, within the RSTUDIO environment, we used the VCFR [39], RESHAPE2 [40] and GGPLOT2 [41] packages to filter out putative variants if the read depth was above a minimum threshold value of '10', corresponding to our targeted minimum FDR value ≤ 0.1. Using R, we then tested if counts of SNVs differed between the three populations and castes.

To compare the broader mutational profile between caste- and population-specific libraries, we generated SBS-6 mutational catalogues that profiled the total count of all possible single base substitution mutation types referenced by the pyrimidine base (C > A, C > G, C > T, T > A, T > C and T > G) using a custom PYTHON script. We also developed SBS-96 mutational catalogues that profiled the count of all possible single base substitution mutation types and flanking 5′ and 3′ bases, for a total of 96 trinucleotide mutation types [42].

To quantify selection pressures using the d_N/d_S ratio of the entire transcriptome, we identified the longest open reading frame in each transcript bound by a start (AUG) and stop (UAA, UAG, UGA) codon. From this representative template, we estimated the proportion of synonymous (p_S) and nonsynonymous (p_N) substitutions by dividing the count of p_S and p_N for each transcript by the count of synonymous and nonsynonymous sites, respectively. Specifically, we estimated the number of synonymous (d_S) and nonsynonymous (d_N) substitutions per site using Equations (1) and (2) of Nei and Gojobori [43].

$$d_S = -\frac{3}{4}ln\left(1 - \frac{4p_S}{3}\right) \tag{1}$$

$$d_N = -\frac{3}{4}ln\left(1 - \frac{4p_N}{3}\right) \tag{2}$$

The d_N/d_S ratio was calculated by dividing d_N by d_S for each of the nine samples. The d_N/d_S ratio was also calculated from a subset of $n = 570$ genes that are known to strongly co-express with caste (minimum two-fold change and FDR < 0.05; [30]).

3. Results

We extracted SNVs from the whole-transcriptome libraries and compared the corresponding VCFs for each caste and population. The variant read depth coverage was comparable between all nine libraries (Figure 3) and revealed extensive SNV counts (~144 K SNVs per library; Table 1), especially in the 10–100 read depth range, indicating that the variants were inferred with low inter-sample bias and high technical accuracy [44].

The distribution of SNVs did not differ by caste (Kruskal–Wallis Test, $\chi^2 = 0.089$, df = 2, $p = 0.96$; Figure 4). By comparison to a common reference transcriptome, worker, soldier and nymph libraries showed remarkably similar numbers of SNVs (~15 K), with the worker library having a slightly higher variance across the three sampled populations. The mean transition/transversion ratio was also nearly identical between castes (worker = 1.71, soldier = 1.70, nymph = 1.70; one-way ANOVA $F = 0.68$, $p = 0.54$), as were the overall SBS-96 mutational profiles (Figure 5).
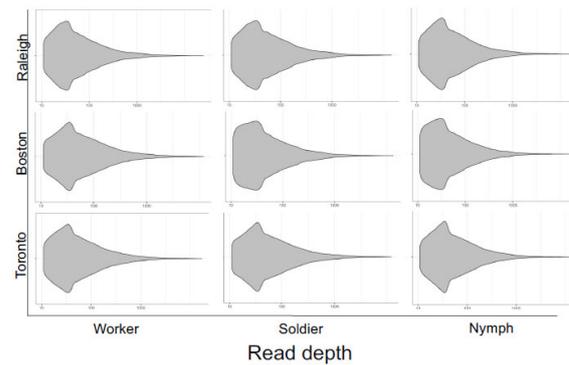
Figure 3. Read depth of high-quality (depth > 10) SNVs extracted from variant call files for three termite castes (worker, soldier, nymph) sampled from three populations (Raleigh, Boston, Toronto). The overall profiles of read depth coverage are similar between libraries, suggesting consistently high qualities.
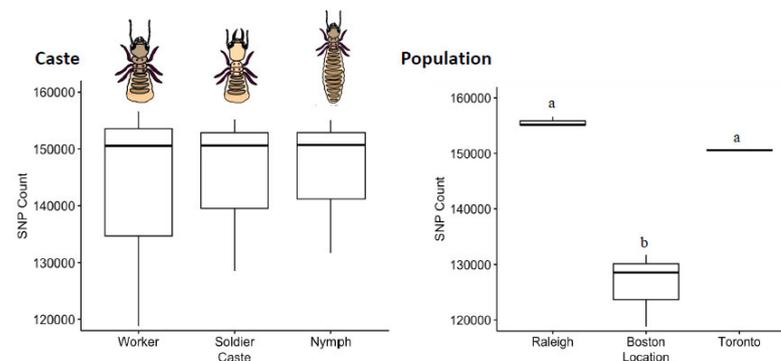

Figure 4. Counts of SNVs by caste and population. Termite castes are relatively uniform in SNV diversity. A pair-wise post-hoc analysis indicated that workers, soldiers and nymphs all harbour nearly identical levels of diversity in their transcribed sequences (Pairwise post-hoc Dunn Test, $p > 0.4$ in all cases). Among populations, however, termites collected from Raleigh were more variable than were those collected from Boston (Dunn tests; $p = 0.0109$), with the Toronto population being intermediate (Dunn test; $p > 0.08$ in both cases). Error bars are 95% CI. Different letters above boxes indicate a significant difference. Note that for the Toronto termite population, the three caste libraries used had a very similar SNV counts, and the estimate of variance is miniscule.

The distribution of SNVs did differ by population (Kruskal–Wallis, $\chi_2 = 7.2$, df = 2, $p = 0.027$). By comparison to a common reference transcriptome, Raleigh and Toronto termites harbored significantly more expressed-gene diversity than termites sampled from Boston. Notably, the invasive unicolonial Toronto population was not the least diverse, and thus by comparative analysis does not show the obvious loss-of-diversity expected from a recent invasion bottleneck (Figure 4). The mean transition/transversion ratio was nearly identical among populations (Toronto = 1.69, Boston = 1.71, Raleigh = 1.70; one-way ANOVA $F = 3.37$, $p = 0.104$), though mutational profiles did reveal an excess of transitions (C > T and T > C) across all the nine libraries (Figure 5).

The mean d_N/d_S ratio did not vary by population (one-way ANOVA $F = 3.99$, $p = 0.079$) or caste (one-way ANOVA $F = 0.56$, $p = 0.601$; Figure 6). However, when we performed this same analysis on a 230-gene subset of caste-associated genes identified by Wu et al. (2018), the d_N/d_S ratio did differ by caste (one-way ANOVA $F = 35.004$, $p = 0.0004$) but not by population (one-way ANOVA $F = 0.047$, $p = 0.95$). Specifically, genes associated in their expression with the defensively specialized and sterile soldier caste appear to evolve at a higher rate ($d_N/d_S = 4.3$) than genes associated with the nymph ($d_N/d_S = 1.8$) or worker ($d_N/d_S = 1.7$) castes (Figure 6).
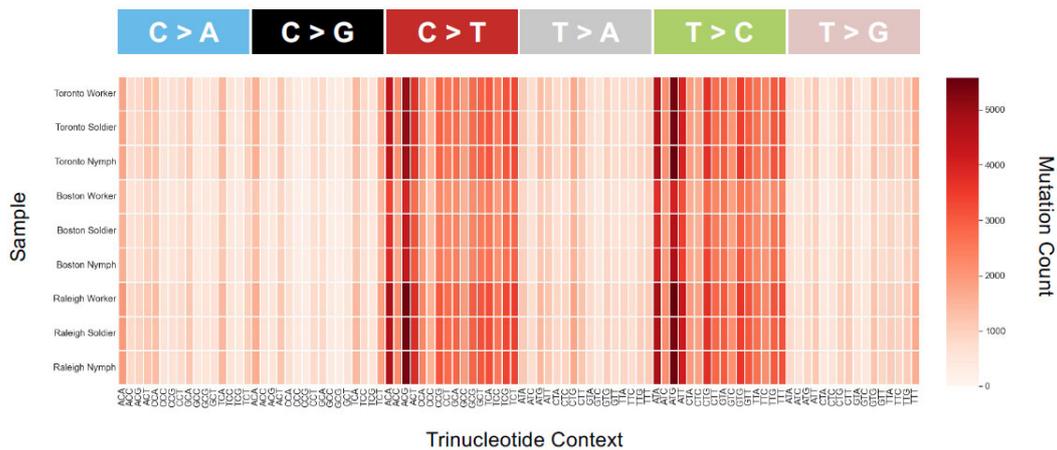
Figure 5. Heatmap of SBS-96 mutation counts for each of the nine samples. All 96 possible mutations are plotted (format available at https://cancer.sanger.ac.uk/cosmic/signatures/SBS, (accessed on 1 November 2021). Most plots show similar biases in mutation types and counts—for example, an excess of transitions across all samples. An SBS-96 plot with purine as a lead base is very similar qualitatively (not shown), consistent with similar mutational mechanisms across populations and castes. There are nonetheless slight differences in the actual numbers of SBSs between caste and population. Additionally, the scale for the Boston population has a smaller range than for Toronto and Raleigh, which is consistent with fewer variants found there (Figure 4B).
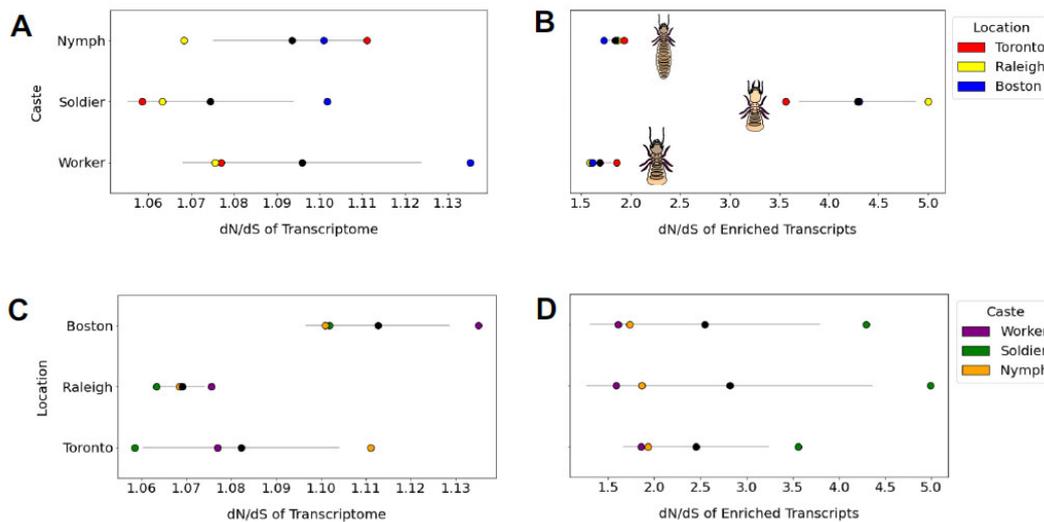

Figure 6. The d_N/d_S ratios of whole transcriptome and a subset of $n = 230$ caste-specific transcripts are grouped by caste or population. For each graph, mean and standard deviation are displayed as a black circle and grey line, respectively. We observed no transcriptome-wide differences in mean d_N/d_S between castes (shown in (**A**)). However, when caste-specific transcripts were tested separately (**B**), we found d_N/d_S of soldier-expressed genes to be significantly higher than d_N/d_S for genes uniquely expressed in nymphs (mean difference = 2.45, Tukey's post hoc test $p = 0.001$) or workers (mean difference = 2.599, $p = 0.01$). We found no difference between populations in mean d_N/d_S at the whole-transcriptome (**C**) or caste-specific transcriptome (**D**) level.

4. Discussion

In this study, we used RNA-Seq analysis to reveal genetic events associated with caste and invasiveness from comparably sampled North American populations of the Eastern subterranean termite. Our whole transcriptome-wide screen, spanning 29,641 transcripts that represent 13,755 genes, showed a remarkably consistent SNV count for each caste library, but the nature of these variants suggest that some confer a caste-specific advantage.

A codon-based analysis revealed a relatively high d_N/d_S ratio (~4.3) in genes associated with the reproductively altruistic soldier caste, regardless of the population from which the individuals were sampled. Given that soldiers are sterile and thus have no direct fitness, any evidence for selection on genes functionally associated with this caste is de facto evidence for indirect selection, of the type mediated through kin selection. Additionally, we found differences in SNV diversity across three termite populations but not in the manner expected if the invasive population was recently bottlenecked. The Toronto population was not the least diverse; it harboured a similar or higher number of SNVs as one of the native range populations without revealing an obvious mark of a bottleneck or any other substitution patterns associated with the invaded habitat. We suggest that a 90-year timeframe, since the Eastern subterranean termite invasion of the city of Toronto, may be sufficient to obscure any obvious signature of a past bottleneck. By extension, we suggest that the unicolonial behaviour of invasive termites observed in the city is no longer strictly associated with the loss of genetic diversity.

4.1. Test for Caste Differences

Despite a consensus that genetic effects on social insect caste ought to be rare, we nonetheless find occasional and sometimes complex exceptions to this rule [45,46], including from termites [47,48] and specifically from *Reticulitermes* [25,49,50]. There are a growing number of examples of weak or strong genetic effects on caste and subcaste differentiation. This is especially evident in social insect species that mate more than once (i.e., polyandry), have loci that constrain development into one caste or another, hybridize between reproductively isolated lineages, or recurrently use some form of thelytokous pathenogenesis to clone the queen [28]. Genotype-caste associations can therefore evolve in strikingly different and complex ways.

For our focal species, environmentally mediated gene expression is an important and well-established factor in caste differentiation [30,51,52], yet no test has prior examined how variants of expressed genes might correlate with caste. We found that our selected measure of nucleotide diversity (i.e., the number of SNVs) does not covary with caste, at least not in the three castes sampled from three populations. Instead, variation levels within and between castes are similar. Therefore, no transcriptome-wide association between caste and SNVs at either of the developmental switches was detected in *R. flavipes* caste development. That is, there is no widespread difference in RNA diversity between castes representative of the reproductive (nymph) and non-reproductive lines (workers, soldiers), or between the workers and soldiers themselves. The d_N/d_S analysis did, however, reveal an intriguing difference between caste-associated genes. Specifically, a subset of genes known to associate in their expression with the soldier caste differed in d_N/d_S ratio from genes uniquely expressed in the nymph and worker caste. This caste-biased pattern in nucleotide substitution suggests that selection acts differently on genes that are functionally associated with different castes. Our analysis does not include primary reproductive—'kings' and 'queens'. First, these castes can be difficult to locate and extract from diffuse underground nests but, also, they are effectively absent in northern and invasive populations such as Toronto. In our study, we use late-instar nymphs to represent the (neotenic) reproductive caste.

Prior studies demonstrate that soldier-biased genes are taxonomically restricted [31] and enriched for biologically novel processes [30]. These patterns further indicate that genes associated with the soldier caste in *Reticulitermes* evolve under grossly different selection pressures than genes associated with other castes. The Nei and Gojobori estimate of selection was developed well before population genomic data. Our use of this estimate here may therefore differ from its initially intended use and violate some of its assumptions regarding the nature of mutations [53], which in our case are SNVs. Its simple formulae nonetheless remain in widespread use and provides an intuitive measure of selection. The precise signature of caste-mediated selection is not yet well characterised for termites [54] or any social insect [55], but our RNA-Seq analysis clearly reveals evidence for differential selection of soldier-associated genes. Why soldier-associated genes

have a two-fold higher estimate of d_N/d_S than worker- or nymph-associated genes in our dataset remain unknown but appears to be unrelated to habitat and therefore unrelated to invasion. Nevertheless, soldiers of *Reticulitermes* have no reproductive potential of their own—they are sterile—and thus gain fitness only indirectly through assisted production of non-descendent kin (i.e., siblings, half-siblings, etc.). The novel evidence for selection on soldier-expressed genes in this study is presumably indirect in nature and commonly referred to as kin selection.

4.2. Test for Population Differences

Reticulitermes are typically colonial and thus live in simple or extended families [25], yet for some populations nestmate recognition appears weak or absent, notably in populations invasive to France [29,56] or other parts of the world [17]. In the city of Toronto, and other Canadian municipalities where *R. flavipes* has invaded [57], the unicolonial populations are characterized, as in France [56], by low aggression and apparent intermixing of colonies that lack clear kin boundaries [58,59]. Moreover, unicolonial termite populations retain a high capacity for secondary neotenic reproduction via brachypterous (wing-budded) nymphs that can form independently reproductive sub-colonies [23]. These invasive qualities from a social insect [60,61] likely contribute to the success of *R. flavipes* in urban environments, rendering them more difficult to eradicate.

Early research suggests that the introduction of *R. flavipes* to the city of Toronto was a single event along the city's industrial foreshore between the years 1935 and 1938 [62]. How many reproductives were introduced to 'the shores of Toronto Bay' is not known, but whatever genetic diversity is currently present in this large metropolitan population of termites may have arisen from a bottleneck that occurred over the last ~90 years. Previous genetic analysis of microsatellite DNA showed that the termites of Toronto are genetically distinct from other subterranean termite populations in the province, e.g., [63] and have presumably spread across the city's metropolitan area (~630 sq. km) via nest budding and human-mediated dispersal [64]. Despite the invasiveness of the urban-dwelling termites of Toronto, our analysis reveals as many or more polymorphic transcripts than southernly forest-dwelling populations in the species' native range. Time, large effective populations sizes and the possibility of repeated invasions from the United States [64] or secondary invasions via Europe [17] may therefore have restored any initial loss of diversity upon invasion.

In our study, we have used a de novo consensus assembly as a generic reference for all comparisons. At present, no annotated model genome that might have served as an out-of-study neutral comparator is available for *R. flavipes*. Because our reference sequence does not correspond to any single individual, but rather is a consensus sequence generated from multiple library reads, we do not attribute importance to any specific functional difference. We do, however, identify broad patterns of diversity that qualitatively distinguish libraries and serve to test our predictions for one population or caste to have fewer (or more) silent or non-silent SNVs than another. Collectively, our findings suggest that unicolonialism in the one invasive population studied here is actually not linked to any single transcriptome-wide pattern of diversity, as may be the case elsewhere for *Reticulitermes* [29,65] or for other social insects [66–68]. Instead, unicoloniality in Toronto termites may depend on other aspects of their biology that we have not measured here. Potentially, these may include flexible social organization [60], cold tolerance [69], diversity in cues involved in nestmate recognition [70] and ecological conditions that may have promoted unicoloniality [71]. For example, short summers at the northern latitudes of Ontario may select for local reliance on relatively rapid neotenic reproduction within natal nests, instead of the production of dispersing alates and nest founding [72]. Further, the absence of primary sexuals in unicolonies may in turn dampen the normal pheromone-induced incentive for nestmate recognition and inter-colony defence [73], resulting in relatively 'open' societies [74] that consist of interconnected nests with mixed kin groups. Our work thus suggests an important role

for both environmental (climatic constraints) and social (pheromonal) factors that are implicated in maintaining unicolony formation in northern invasive subterranean termites.

Author Contributions: J.M.R., D.C., A.M.C., K.A.H. and G.J.T. conceived, developed and managed the project; J.M.R. and D.C. performed the analyses and generated the figures, with data processing assistance from C.T., A.W.G., T.W. and A.M.C.; J.M.R. and G.J.T. drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in this study are publicly available through the National Center for Biotechnology Information Sequence Read Archive (NCBI-SRA) under the accession numbers presented in Table 1.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Strassmann, J.E.; Queller, D.C. Insect societies as divided organisms: The complexities of purpose and cross-purpose. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8619–8626. [CrossRef] [PubMed]
2. Sundström, L.; Boomsma, J. Conflicts and alliances in insect families. *Heredity* **2001**, *86*, 515–521. [CrossRef] [PubMed]
3. Pamilo, P.; Gertsch, P.; Thorén, P.; Seppä, P. Molecular population genetics of social insects. *Annu. Rev. Ecol. Syst.* **1997**, *28*, 1–25. [CrossRef]
4. Thompson, G.J.; Hurd, P.L.; Crespi, B.J. Genes underlying altruism. *Biol. Lett.* **2013**, *9*, 20130395. [CrossRef] [PubMed]
5. Toth, A.L.; Rehan, S.M. Molecular evolution of insect sociality: An Eco-Evo-Devo perspective. *Annu. Rev. Entomol.* **2017**, *62*, 419–442. [CrossRef]
6. Weitekamp, C.A.; Libbrecht, R.; Keller, L. Genetics and evolution of social behavior in insects. *Annu. Rev. Genet.* **2017**, *51*, 219–239. [CrossRef]
7. Libbrecht, R.; Oxley, P.R.; Kronauer, D.J.; Keller, L. Ant genomics sheds light on the molecular regulation of social organization. *Genome Biol.* **2013**, *14*, 212. [CrossRef]
8. Linksvayer, T.A.; Wade, M.J. Theoretical predictions for sociogenomic data: The effects of kin selection and sex-limited expression on the evolution of social insect genomes. *Front. Ecol. Evol.* **2016**, *4*, 65. [CrossRef]
9. Mullen, E.K.; Thompson, G.J. Understanding honey bee worker self-sacrifice: A conceptual-empirical framework. *Adv. Insect Physiol.* **2015**, *48*, 325–354.
10. Krishna, K.; Grimaldi, D.A.; Krishna, V.; Engel, M.S. Treatise on the Isoptera of the World: Basal Families. *Bull. Am. Mus. Nat.* **2013**, *2013*, 200–623. [CrossRef]
11. Bourguignon, T.; Lo, N.; Cameron, S.L.; Šobotník, J.; Hayashi, Y.; Shigenobu, S.; Watanabe, D.; Roisin, Y.; Miura, T.; Evans, T.A. The evolutionary history of termites as inferred from 66 mitochondrial genomes. *Mol. Biol. Evol.* **2014**, *32*, 406–421. [CrossRef] [PubMed]
12. Scharf, M.E. Omic research in termites: An overview and a roadmap. *Front. Genet.* **2015**, *6*, 76. [CrossRef] [PubMed]
13. Harrison, M.C.; Jongepier, E.; Robertson, H.M.; Arning, N.; Bitard-Feildel, T.; Chao, H.; Childers, C.P.; Dinh, H.; Doddapaneni, H.; Dugan, S. Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nat. Ecol. Evol.* **2018**, *2*, 557–566. [CrossRef]
14. Romiguier, J.; Gayral, P.; Ballenghien, M.; Bernard, A.; Cahais, V.; Chenuil, A.; Chiari, Y.; Dernat, R.; Duret, L.; Faivre, N.; et al. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* **2014**, *515*, 261–263. [CrossRef] [PubMed]
15. Weesner, F.M. Termites of the Nearctic Region. In *Biology of Termites*; Krishna, K., Weesner, F.M., Eds.; Academic Press: New York, NY, USA, 1970; Volume 2, pp. 477–525.
16. Evans, T.A.; Forschler, B.T.; Grace, J.K. Biology of invasive termites: A worldwide review. *Annu. Rev. Entomol.* **2013**, *58*, 455–474. [CrossRef]
17. Eyer, P.A.; Blumenfeld, A.J.; Johnson, L.N.; Perdereau, E.; Shults, P.; Wang, S.; Dedeine, F.; Dupont, S.; Bagnères, A.G.; Vargo, E.L. Extensive human-mediated jump dispersal within and across the native and introduced ranges of the invasive termite *Reticulitermes flavipes*. *Mol. Ecol.* **2021**, *30*, 3948–3964. [CrossRef] [PubMed]
18. Roisin, Y. Diversity and Evolution of Caste Patterns. In *Termites: Evolution, Sociality, Symbiosis, Ecology*; Abe, T., Bignell, D.E., Higashi, M., Eds.; Kluwer Academic: Dordrecht, The Netherlands, 2000; pp. 95–120.

19. Higashi, M.; Yamamura, N.; Abe, T. Theories on the Sociality of Termites. In *Termites: Evolution, Sociality, Symbiosis, Ecology*; Abe, T., Bignell, D.E., Higashi, M., Eds.; Kluwer Academic: Dordrecht, The Netherlands, 2000; pp. 169–188.

20. Nijhout, H.F. Development and evolution of adaptive polyphenisms. *Evol. Dev.* **2003**, *5*, 9–18. [CrossRef]

21. Wheeler, D.E. Developmental and physiological determinants of caste in social Hymenoptera: Evolutionary implications. *Am. Nat.* **1986**, *128*, 13–34. [CrossRef]

22. Crozier, R.H.; Schluns, H. Genetic caste determination in termites: Out of the shade but not from Mars. *Bioessays* **2008**, *30*, 299–302. [CrossRef]

23. Thorne, B.L.; Traniello, J.F.A.; Adams, E.S.; Bulmer, M. Reproductive dynamics and colony structure of subterranean termites of the genus *Reticulitermes* (Isoptera Rhinotermitidae): A review of the evidence from behavioral, ecological, and genetic studies. *Ethol. Ecol. Evol.* **1999**, *11*, 149–169. [CrossRef]

24. Perdereau, E.; Bagnères, A.G.; Vargo, E.L.; Baudouin, G.; Xu, Y.; Labadie, P.; Dupont, S.; Dedeine, F. Relationship between invasion success and colony breeding structure in a subterranean termite. *Mol. Ecol.* **2015**, *24*, 2125–2142. [CrossRef]

25. Vargo, E.L. Diversity of termite breeding systems. *Insects* **2019**, *10*, 52. [CrossRef] [PubMed]

26. Rust, M.K.; Su, N.Y. Managing social insects of urban importance. *Annu. Rev. Entomol.* **2012**, *57*, 355–375. [CrossRef] [PubMed]

27. Lainé, L.V.; Wright, D.J. The life cycle of *Reticulitermes* spp. (Isoptera: Rhinotermitidae): What do we know? *Bull. Entomol. Res.* **2003**, *93*, 267–278. [CrossRef]

28. Thompson, G.J.; Chernyshova, A.M. Caste Differentiation: Genetic and Epigenetic Factors. In *Encyclopedia of Social Insects*; Starr, C., Ed.; Springer: Cham, Switzerland, 2021; pp. 165–176.

29. Dronnet, S.; Chapuisat, M.; Vargo, E.L.; Lohou, C.; Bagnères, A.G. Genetic analysis of the breeding system of an invasive subterranean termite, *Reticulitermes santonensis*, in urban and natural habitats. *Mol. Ecol.* **2005**, *14*, 1311–1320. [CrossRef]

30. Wu, T.; Dhami, G.K.; Thompson, G.J. Soldier-biased gene expression in a subterranean termite implies functional specialization of the defensive caste. *Evol. Dev.* **2018**, *20*, 3–16. [CrossRef]

31. Behl, S.; Wu, T.; Chernyshova, A.M.; Thompson, G.J. Caste-biased genes in a subterranean termite are taxonomically restricted: Implications for novel gene recruitment during termite caste evolution. *Insect. Soc.* **2018**, *65*, 593–599. [CrossRef]

32. Jalili, V.; Afgan, E.; Gu, Q.; Clements, D.; Blankenberg, D.; Goecks, J.; Taylor, J.; Nekrutenko, A. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Res.* **2020**, *48*, W395–W402. [CrossRef] [PubMed]

33. Conesa, A.; Madrigal, P.; Tarazona, S.; Gomez-Cabrero, D.; Cervera, A.; McPherson, A.; Szcześniak, M.W.; Gaffney, D.J.; Elo, L.L.; Zhang, X. A survey of best practices for RNA-seq data analysis. *Genome Biol.* **2016**, *17*, 13. [CrossRef]

34. Ewing, B.; Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **1998**, *8*, 186–194. [CrossRef]

35. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.D.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [CrossRef] [PubMed]

36. Mantaci, S.; Restivo, A.; Rosone, G.; Sciortino, M. An extension of the Burrows–Wheeler transform. *Theor. Comput. Sci.* **2007**, *387*, 298–312. [CrossRef]

37. Garrison, E.; Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv* **2012**, arXiv:1207.3907.

38. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; Del Angel, G.; Rivas, M.A.; Hanna, M. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491–498. [CrossRef]

39. Knaus, B.J.; Grünwald, N.J. VCFR: A package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* **2017**, *17*, 44–53. [CrossRef] [PubMed]

40. Wickham, H. Reshaping data with the reshape package. *J. Stat. Softw.* **2007**, *21*, 1–20. [CrossRef]

41. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*, 2nd ed.; Springer: New York, NY, USA, 2016.

42. Alexandrov, L.B.; Kim, J.; Haradhvala, N.J.; Huang, M.N.; Tian Ng, A.W.; Wu, Y.; Boot, A.; Covington, K.R.; Gordenin, D.A.; Bergstrom, E.N.; et al. The repertoire of mutational signatures in human cancer. *Nature* **2020**, *578*, 94–101. [CrossRef]

43. Nei, M.; Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **1986**, *3*, 418–426.

44. Sims, D.; Sudbery, I.; Ilott, N.E.; Heger, A.; Ponting, C.P. Sequencing depth and coverage: Key considerations in genomic analyses. *Nat. Rev. Genet.* **2014**, *15*, 121–132. [CrossRef]

45. Anderson, K.E.; Linksvayer, T.A.; Smith, C.R. The causes and consequences of genetic caste determination in ants (Hymenoptera: Formicidae). *Myrmecol. News* **2008**, *11*, 119–132.

46. Schwander, T.; Lo, N.; Beekman, M.; Oldroyd, B.P.; Keller, L. Nature versus nurture in social insect caste differentiation. *Trends Ecol. Evol.* **2010**, *25*, 275–282. [CrossRef] [PubMed]

47. Goodisman, M.A.D.; Crozier, R.H. Association between caste and genotype in the termite *Mastotermes darwiniensis* Froggatt (Isoptera: Mastotermitidae). *Aust. J. Entomol.* **2003**, *42*, 1–5. [CrossRef]

48. Fougeyrollas, R.; Dolejšová, K.; Sillam-Dussès, D.; Roy, V.; Poteaux, C.; Hanus, R.; Roisin, Y. Asexual queen succession in the higher termite *Embiratermes neotenicus*. *Proc. Royal Soc. B* **2015**, *282*, 20150260. [CrossRef]

49. Hayashi, Y.; Lo, N.; Miyata, H.; Kitade, O. Sex-linked genetic influence on caste determination in a termite. *Science* **2007**, *318*, 985–987. [CrossRef] [PubMed]

50. Yamamoto, Y.; Matsuura, K. Genetic influence on caste determination underlying the asexual queen succession system in a termite. *Behav. Ecol. Sociobiol.* **2012**, *66*, 39–46. [CrossRef]

51. Dedeine, F.; Weinert, L.A.; Bigot, D.; Josse, T.; Ballenghien, M.; Cahais, V.; Galtier, N.; Gayral, P. Comparative analysis of transcriptomes from secondary reproductives of three *Reticulitermes* termite species. *PLoS ONE* **2015**, *10*, e0145596.

52. Miura, T.; Scharf, M.E. Molecular basis underlying caste differentiation in termites. In *Biology of Termites: A Modern Synthesis*; Springer Science+Business Media B.V.: Dordrecht, The Netherlands, 2011; pp. 211–253.

53. Kryazhimskiy, S.; Plotkin, J.B. The population genetics of dN/dS. *PLoS Genet.* **2008**, *4*, e1000304. [CrossRef]

54. Harrison, M.C.; Chernyshova, A.M.; Thompson, G.J. No obvious transcriptome-wide signature of indirect selection in termites. *J. Evol. Biol.* **2020**, *34*, 403–415. [CrossRef]

55. Helanterä, H.; Uller, T. Neutral and adaptive explanations for an association between caste-biased gene expression and rate of sequence evolution. *Front. Genet.* **2014**, *5*, 297. [CrossRef]

56. Leniaud, L.; Pichon, A.; Uva, P.; Bagneres, A.G. Unicoloniality in *Reticulitermes urbis*: A novel feature in a potentially invasive termite species. *Bull. Entomol. Res.* **2009**, *99*, 1–10. [CrossRef]

57. Kirby, C.S. The distribution of termites in Ontario after 25 years. *Can. Entomol.* **1965**, *97*, 310–314. [CrossRef]

58. Grace, J.K. Absence of overt agonistic behavior in a northern population of *Reticulitermes flavipes* (Isoptera, Rhinotermitidae). *Sociobiology* **1996**, *28*, 103–110.

59. Simkovic, V.; Thompson, G.J.; McNeil, J.N. Testing for aggression and nestmate recognition in the Eastern subterranean termite (*Reticulitermes flavipes*). *Insect. Soc.* **2018**, *65*, 281–288. [CrossRef]

60. Moller, H. Lessons for invasion theory from social insects. *Biol. Conserv.* **1996**, *78*, 125–142. [CrossRef]

61. Eyer, P.-A.; Vargo, E.L. Breeding structure and invasiveness in social insects. *Curr. Opin. Insect. Sci.* **2021**, *46*, 24–30. [CrossRef] [PubMed]

62. Urquhart, F.A. The introduction of the termite into Ontario. *Canad. Entomol.* **1953**, *85*, 292–293. [CrossRef]

63. Raffoul, M.; Hecnar, S.J.; Prezioso, S.; Hecnar, D.R.; Thompson, G.J. Trap response and genetic structure of Eastern subterranean termites (Isoptera, Rhinotermitidae) in Point Pelee National Park, Ontario, Canada. *Can. Entomol.* **2011**, *143*, 263–271. [CrossRef]

64. Scaduto, D.A.; Garner, S.R.; Leach, E.L.; Thompson, G.J. Genetic evidence for multiple invasions of the Eastern subterranean termite into Canada. *Environ. Entomol.* **2012**, *41*, 1680–1686. [CrossRef]

65. Perdereau, E.; Bagneres, A.G.; Bankhead-Dronnet, S.; Dupont, S.; Zimmermann, M.; Vargo, E.L.; Dedeine, F. Global genetic analysis reveals the putative native source of the invasive termite, *Reticulitermes flavipes*, in France. *Mol. Ecol.* **2013**, *22*, 1105–1119. [CrossRef]

66. Helanterä, H.; Strassmann, J.E.; Carrillo, J.; Queller, D.C. Unicolonial ants: Where do they come from, what are they and where are they going? *Trends Ecol. Evol.* **2009**, *24*, 341–349. [CrossRef]

67. Holway, D.A.; Lach, L.; Suarez, A.V.; Tsutsui, N.D.; Case, T.J. The causes and consequences of ant invasions. *Annu. Rev. Ecol. Syst.* **2002**, *33*, 181–233. [CrossRef]

68. Eyer, P.-A.; McDowell, B.; Johnson, L.N.; Calcaterra, L.A.; Fernandez, M.B.; Shoemaker, D.; Puckett, R.T.; Vargo, E.L. Supercolonial structure of invasive populations of the tawny crazy ant *Nylanderia fulva* in the US. *BMC Evol. Biol.* **2018**, *18*, 209. [CrossRef] [PubMed]

69. Clarke, M.W.; Thompson, G.J.; Sinclair, B.J. Cold tolerance of the Eastern subterranean termite, *Reticulitermes flavipes* (Isoptera: Rhinotermitidae), in Ontario. *Environ. Entomol.* **2013**, *42*, 805–810. [CrossRef] [PubMed]

70. Perdereau, E.; Dedeine, F.; Christides, J.P.; Bagnères, A.G. Variations in worker cuticular hydrocarbons and soldier isoprenoid defensive secretions within and among introduced and native populations of the subterranean termite, *Reticulitermes flavipes*. *J. Chem. Ecol.* **2010**, *36*, 1189–1198. [CrossRef]

71. Chapman, R.E.; Bourke, A.F.G. The influence of sociality on the conservation biology of social insects. *Ecol. Lett.* **2001**, *4*, 650–662. [CrossRef]

72. Esenther, G.R. Termites in Wisconsin. *Ann. Entomol. Soc. Am.* **1969**, *62*, 1274–1284. [CrossRef]

73. Thompson, G.J. Termites reigned by royals close ranks. *Insect. Soc.* **2021**, *68*, 1–2. [CrossRef]

74. Clément, J.-L. Open and closed societies in *Reticulitermes* termites (Isoptera, Rhinotermitidae): Geographic and seasonal variations. *Sociobiology* **1986**, *11*, 311–323.