



Article

Differential Expression Profile of lncRNAs from Primary Human Hepatocytes Following DEET and Fipronil Exposure

Robert D. Mitchell III ¹ , Andrew D. Wallace ², Ernest Hodgson ^{2,3} and R. Michael Roe ^{1,*}

¹ Department of Entomology and Plant Pathology, North Carolina Agromedicine Institute, Campus Box 7647, 3230 Ligon Street, North Carolina State University, Raleigh, NC 27695, USA; rdmitche@ncsu.edu

² Toxicology Program, Department of Biology, North Carolina State University, Raleigh, NC 27695, USA; adwallacnc@gmail.com (A.D.W.); ernest_hodgson@ncsu.edu (E.H.)

³ Department of Applied Ecology, Toxicology Program, Department of Biology, North Carolina Agromedicine Institute, North Carolina State University, Raleigh, NC 27695, USA

* Correspondence: michael_roe@ncsu.edu; Tel.: +1-919-515-4325

Received: 15 August 2017; Accepted: 4 October 2017; Published: 7 October 2017

Abstract: While the synthesis and use of new chemical compounds is at an all-time high, the study of their potential impact on human health is quickly falling behind, and new methods are needed to assess their impact. We chose to examine the effects of two common environmental chemicals, the insect repellent *N,N*-diethyl-*m*-toluamide (DEET) and the insecticide fluocyanobenpyrazole (fipronil), on transcript levels of long non-protein coding RNAs (lncRNAs) in primary human hepatocytes using a global RNA-Seq approach. While lncRNAs are believed to play a critical role in numerous important biological processes, many still remain uncharacterized, and their functions and modes of action remain largely unclear, especially in relation to environmental chemicals. RNA-Seq showed that 100 μ M DEET significantly increased transcript levels for 2 lncRNAs and lowered transcript levels for 18 lncRNAs, while fipronil at 10 μ M increased transcript levels for 76 lncRNAs and decreased levels for 193 lncRNAs. A mixture of 100 μ M DEET and 10 μ M fipronil increased transcript levels for 75 lncRNAs and lowered transcript levels for 258 lncRNAs. This indicates a more-than-additive effect on lncRNA transcript expression when the two chemicals were presented in combination versus each chemical alone. Differentially expressed lncRNA genes were mapped to chromosomes, analyzed by proximity to neighboring protein-coding genes, and functionally characterized via gene ontology and molecular mapping algorithms. While further testing is required to assess the organismal impact of changes in transcript levels, this initial analysis links several of the dysregulated lncRNAs to processes and pathways critical to proper cellular function, such as the innate and adaptive immune response and the p53 signaling pathway.

Keywords: DEET; fipronil; long non-coding RNA; lncRNA; primary liver cells; epigenetics; RNA-Seq; transcriptomics; Zika virus; human hepatocytes

1. Introduction

The study of the impact of environmental chemicals on human health has fallen significantly behind the rate at which we synthesize new chemical compounds. Recent studies suggest that while we generate nearly 10 million new chemical compounds annually, factors such as a lack of funding and reduced public interest and awareness have diminished the amount of research into the potential consequences of human exposure to environmental chemicals [1]. To fully understand the impact of chemicals on human health and provide more rapid and comprehensive methods to evaluate risk, we must take advantage of recent advances in high-throughput DNA sequencing, annotation of the

human genome, and bioinformatics, while at the same time using global approaches relating effects of chemical exposure on molecular pathways to whole organism function.

The insect repellent *N,N*-diethyl-*m*-toluamide (DEET) and the insecticide fluocyanobenzopyrazole (fipronil) are pesticides that have a high potential for human exposure. DEET is applied at a concentration of 5% to 100% to skin to repel insects and other arthropods by approximately 30% of the U.S. population annually [2]. Recently, fetal Zika virus infections in the Americas linked to severe birth defects including microcephaly [3] have further increased our utilization and dependency on DEET. Expectant mothers are encouraged to use this repellent on their skin at a minimum concentration of 30% active ingredient every time they are at risk of being bitten by mosquitoes before and during pregnancy [4]. This level of repetitive use of DEET on human health has never been considered before. Fipronil is an insecticide used to treat companion animals for fleas and ticks as well as around the home for other arthropod control, e.g., termites, roaches, and ants [5]. It is also used as a pesticide in numerous countries around the world to protect crops such as corn and cotton [6–8]. However, several countries in the European Union have restricted the use of fipronil citing its potential lethality to honeybees [6]. High doses of fipronil in humans have in some cases resulted in severe vomiting, agitation, and seizures [9].

While both DEET and fipronil have been available commercially for many years (DEET since 1957 and fipronil since 1993), the study of these compounds directly in human systems can improve our understanding of their toxicology [10] and also provide a model for developing a new approach to risk assessment for environmental chemicals in general. Previous studies have measured DEET and fipronil metabolite levels in blood and urine, but no work has been conducted at the DNA/RNA level in relation to human health [11,12]. DEET is metabolized in humans by cytochrome P450 enzymes into the primary metabolites *N,N*-diethyl-*m*-hydroxymethylbenzamide (BALC) and *N*-ethyl-*m*-toluamide (ET). While several P450s have been demonstrated to be active in DEET metabolism, CYP2B6 is the principal P450 responsible for the conversion of DEET to BALC, and CYP2C19 for the conversion of DEET to ET [13]. DEET metabolites are primarily excreted from the human body in urine, but can also be expelled in feces [14]. The predominant metabolite of fipronil is fipronil sulfone (5-amino-1-(2,6-dichloro-4-trifluoromethylphenyl)-3-cyano-4-rifluoromethylsulfonyl-pyrazole), which is primarily metabolized by the cytochrome P450 enzyme CYP3A4 [15,16]. Unlike DEET, fipronil is primarily eliminated in the feces [17]. Our group recently found that exposure of primary human hepatocytes to 100 μ M DEET, 10 μ M fipronil, and a mixture of 100 μ M DEET and 10 μ M fipronil significantly altered transcript levels for numerous protein-coding and non-protein coding genes [18]. Here, the goal was to determine what epigenetic elements, i.e., long non-protein coding RNA (lncRNA) transcripts, were significantly differentially expressed after exposure to DEET and fipronil at the same concentrations, and what role they may play in human function.

Long non-protein coding RNAs (lncRNAs) are RNA transcripts greater than 200 nucleotides long which rarely code for protein. They are the largest class of noncoding genes and are processed much like messenger RNA (mRNA), i.e., they are transcribed from active chromatin and can have a 5' cap and a poly-A tail. The three-dimensional (3-D) structure of lncRNAs predominantly determine what RNA, DNA, or proteins they influence [19]. We now know that protein-coding genes account for less than 2% of the human genome and the majority of transcripts do not code for protein but perform other essential functions [20]. There are several distinct sub-categories of lncRNAs based on their configuration in the genome (i.e., location and proximity to protein-coding genes [21]) that play a critical role at every level of gene regulation [19,22,23]; they serve as transcription signals, transcription factor decoys, guides for chromatin-modifying enzymes, and molecular scaffolds facilitating ribonucleoprotein complex formation [20]. Little is known of their mode(s) of action in general, but even less is known about their function in response to environmental chemicals [24].

Considering an almost complete lack of knowledge of the role of lncRNAs in animal and human responses to environmental chemicals, the objective of this study was to use DEET and fipronil, alone and in a mixture, as models to examine their impact on lncRNA transcript levels in primary

human hepatocytes. The research also included an analysis of the potential interaction of lncRNA transcription with that for coding genes to provide leads for future assessments of risks to chemical exposure. Studying DEET and fipronil both alone and in combination in primary liver cells can provide insight on whether processes and pathways are shared or unique among the exposure conditions examined. The use of primary human cells, in general, provides the closest possible estimate of the chemical–human global molecular interaction versus utilization of immortalized cell lines or animal models.

2. Results and Discussion

2.1. Effects of DEET and Fipronil on lncRNA Versus Protein-Coding Gene Transcription

Primary human hepatocytes were treated with either 100 μM DEET, 10 μM fipronil, or a mixture of 100 μM DEET and 10 μM fipronil. At a significance level of $p \leq 0.01$, we observed transcripts for 2 lncRNA genes that were upregulated and 18 that were downregulated by 100 μM DEET. This accounted for 0.04% of the total number of coding and noncoding genes identified in the latest Ensembl release (Ensembl release 87) of the annotated human genome [25]. Specifically, of all of the annotated categories recognized by the Ensembl project that include coding genes, noncoding genes (defined as small noncoding genes, long noncoding genes, and miscellaneous noncoding genes), and pseudogenes, 20 of the 56,384 genes were lncRNA genes whose transcripts were differentially expressed after exposure to 100 μM DEET. When primary human hepatocytes were treated with 10 μM fipronil, there were 76 lncRNA genes whose transcripts were upregulated and 193 downregulated, accounting for 0.48% of the total number of coding and noncoding genes identified in the latest human genome annotation (i.e., 269 of 56,384 genes). When primary human hepatocytes were treated with a mixture of 100 μM DEET and 10 μM fipronil, we observed 75 lncRNA genes whose transcripts were upregulated and 258 downregulated. This accounted for 0.59% of the total number of coding and noncoding genes identified in the latest human genome annotation (i.e., 333 of 56,384 genes). We observed a more-than-additive effect as the sum of the lncRNAs dysregulated by the 100 μM DEET treatment and the 10 μM fipronil treatment was 289, but the two chemicals together elicited up- or downregulation of 333 lncRNAs. We define “more-than-additive” as the number of dysregulated transcripts when cells were treated with DEET and fipronil together that was greater than the number of dysregulated transcripts when cells were treated separately with each compound with the concentration of each compound remaining the same in all treatments. This definition does not include a detailed dose–response evaluation, but is simply meant as a statement describing a mathematical calculation.

In this study, we included transcribed pseudogenes as lncRNAs and specifically defined an lncRNA as any non-protein coding gene whose transcripts were ≥ 200 nucleotides long. A pseudogene is a highly similar copy of a protein-coding gene. A protein-coding gene that is similar to a specific pseudogene is termed a parental gene to that pseudogene and no longer produces a functional protein product in most cases [26]. Pseudogenes typically regulate parental genes as lncRNA transcripts, and previous studies have established that pseudogenes, when transcribed, function as drivers of gene regulation like that for other lncRNAs. Not all pseudogenes are actively transcribed (although the ones identified in this study were transcribed); some estimate that only 2–20% of pseudogenes in the human genome are actively transcribed [26–28]. Pseudogenes function as regulators of target genes when their transcripts interact with target gene promoters and are processed into short noncoding RNAs that hybridize to the protein-coding sense strand transcripts [28]. Experimental evidence supports the role of transcribed pseudogenes as regulating messenger RNAs (mRNAs) via small interfering RNAs [29], regulating other lncRNA transcripts [30], and functioning as microRNA (miRNA) decoys [31].

One of the two lncRNA genes whose transcripts were upregulated by 100 μM DEET was a pseudogene, and 5 of the 18 (28%) downregulated transcripts were pseudogenes. Thirty-four of the 76 (45%) lncRNA genes whose transcripts were upregulated by 10 μM fipronil were pseudogenes,

and 72 of the 193 (37%) downregulated transcripts were pseudogenes. Thirty-two of the 75 (43%) lncRNA genes whose transcripts were upregulated by the 100 μ M DEET plus 10 μ M fipronil mixture were pseudogenes, and 97 of the 258 (38%) downregulated transcripts were pseudogenes. Therefore, approximately 40% of the dysregulated lncRNAs in each treatment were pseudogenes. These findings indicate that many genes that once coded for active proteins, but were thought to currently be inactive (pseudogenes, also known as lncRNAs), may still play a prominent role in gene regulation. This may have arisen from the accumulation of multiple mutations over time, rendering a protein-coding gene inactive. Then, these lncRNAs were “repurposed” to influence the activity of other noncoding and coding elements without coding for proteins themselves. However, our understanding of how lncRNA transcripts interact amongst themselves, other epigenetic elements, protein-coding genes, and gene products is still in its infancy.

Figure 1A shows that 20 of the lncRNAs whose transcripts were significantly differentially expressed ($p \leq 0.01$) were shared among all three treatment conditions (100 μ M DEET, 10 μ M fipronil, and 100 μ M DEET combined with 10 μ M fipronil). Interestingly, there were no lncRNAs whose differentially expressed transcripts were unique only to the 100 μ M DEET treatment. There were 204 lncRNA genes whose transcripts were differentially expressed that were shared between the 10 μ M fipronil treatment and the 100 μ M DEET plus 10 μ M fipronil treatment, 45 that were unique to the fipronil-only treatment, and 109 that were unique to the DEET plus fipronil mixture. Figure 1B shows the relationship between the statistically significant up- and downregulated protein-coding genes at $p \leq 0.01$, which we defined as any genes that were processed into messenger RNA (mRNA) and deemed protein-coding via the National Center for Biotechnology Information (NCBI) database [32]. This definition excludes any noncoding RNAs or the few pseudogenes that are now thought to code for protein and may be reclassified in the near future [33].

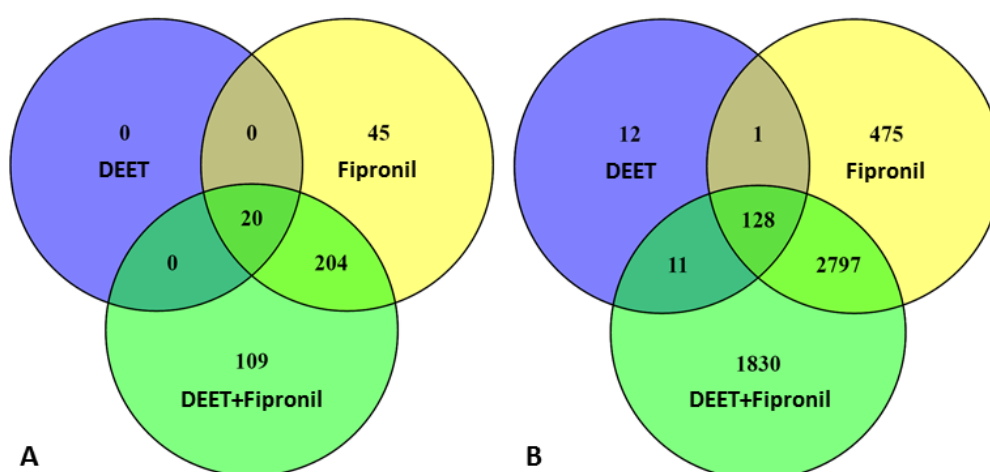


Figure 1. Relationships between the number of long non-protein coding RNAs (lncRNAs) whose transcripts were significantly ($p \leq 0.01$) differentially up- and downregulated (A) and protein-coding genes whose transcripts were significantly ($p \leq 0.01$) differentially expressed (B) when primary human hepatocytes were treated with DEET (100 μ M), fipronil (10 μ M), or a mixture of the two (100 μ M DEET and 10 μ M fipronil) for 72 h.

Table 1 shows the 20 lncRNAs whose transcripts were differentially expressed across all three treatment conditions. This group of affected lncRNAs was interesting because the chemistry and mode of action of DEET and fipronil are completely different and yet the same lncRNA transcripts were affected. Of the 20, only 5 lncRNAs (25%) had assigned Gene Ontology (GO) terms at the biological process, cellular component, or molecular function level. These will be discussed in more detail below, but it is worth noting that the log₂ fold change (log₂FC), a common metric for differential expression values [34], was very similar and in the same direction, either up- or downregulated,

for each lncRNA among the 100 μ M DEET, 10 μ M fipronil, and 100 μ M DEET plus 10 μ M fipronil conditions. While all 20 examples in Table 1 were up- or downregulated in the same manner, this terminology differs from the more-than-additive effect we observed. The “more-than-additive” effect refers to a generalization and did not pertain to every single transcript, but rather the overall number of genes whose transcripts were up- or downregulated under a specific treatment condition. Table S3 provides additional information about these 20 lncRNAs that were dysregulated across all three treatment conditions, including chromosomal coordinates and gene accession numbers.

Table 1. Long non-protein coding RNAs (lncRNAs) whose transcripts were significantly ($p \leq 0.01$) differentially expressed (Diff. Exp.; up (+) or down (–) regulated) in primary human hepatocytes after exposure to 100 μ M DEET, 10 μ M fipronil, or a mixture of 100 μ M DEET and 10 μ M fipronil for 72 h.

Differential Expression	Gene Symbol ^a	Gene Name ^b	Log2FC (DT)	Log2FC (Fip)	Log2FC (DT+Fip)
Up	<i>CYP2B7P</i>	Cytochrome P450 family 2 subfamily B member 7, pseudogene	+2.96	+2.42	+2.79
	<i>HCP5</i>	HLA complex P5	+0.78	+0.93	+1.38
Down	<i>MALAT1</i>	Metastasis associated lung adenocarcinoma transcript 1	–1.34	–1.87	–2.28
	<i>NEAT1</i>	Nuclear paraspeckle assembly transcript 1	–1.18	–1.22	–1.52
	<i>LINC01554</i>	Long intergenic non-protein coding RNA 1554	–1.1	–1.37	–1.33
	<i>LINC01004</i>	Long intergenic non-protein coding RNA 1004	–1.58	–2.01	–2.26
	<i>PSORS1C3</i>	Psoriasis susceptibility 1 candidate 3	–1.58	–0.85	–1.39
	<i>AQP7P1</i>	Aquaporin 7 pseudogene 1	–0.8	–1.05	–0.55
	<i>SCART1</i>	Scavenger receptor protein family member	–1.49	–1.24	–1.89
	<i>PDXDC2P</i>	Pyridoxal dependent decarboxylase domain containing 2	–0.78	–1.14	–1.41
	<i>LINC00893</i>	Long intergenic non-protein coding RNA 893	–1.21	–1.18	–1.51
	<i>WASH5P</i>	WAS protein family homolog 5 pseudogene	–0.79	–0.76	–1.06
	<i>PFN1P2</i>	Profilin 1 pseudogene 2	–1.24	–1.33	–1.78
	<i>LINC00482</i>	Long intergenic non-protein coding RNA 482	–1.01	–0.91	–1.06
	<i>ERVK13-1</i>	Endogenous retrovirus group K13 member 1	–0.84	–0.78	–1.25
	<i>LOC100289230</i>	Uncharacterized LOC100289230	–1.66	–1.8	–1.44
	<i>LOC728040</i>	HCG1813624	–1.37	–2.5	–3.47
	<i>LOC100190986</i>	Uncharacterized LOC100190986	–0.91	–0.96	–1.29
	<i>LINC01000</i>	Long intergenic non-protein coding RNA 1000	–0.71	–0.53	–0.88
<i>LOC100272217</i>	Uncharacterized LOC100272217	–1.46	–1.29	–1.54	

^a HUGO gene nomenclature committee (HGNC) gene symbol [35]; ^b National Center for Biotechnology Information (NCBI) gene description; ^c FC = log2 fold change; DT = 100 μ M DEET; Fip = 10 μ M fipronil; DT + Fip = 100 μ M DEET plus 10 μ M fipronil mixture.

2.2. Chromosomal Mapping of Dysregulated LncRNAs

RNA-Seq (transcriptomic) data were analyzed to reveal the identity and genomic location of lncRNAs relative to protein-coding genes whose transcripts were significantly differentially expressed ($p \leq 0.01$) after exposure to 100 μ M DEET, 10 μ M fipronil, and a mixture of 100 μ M DEET with 10 μ M fipronil. Figure 2 shows the distribution and magnitude (in $\pm \log_2$ FC values) of all differentially expressed lncRNA transcripts across all human chromosomes. Every human chromosome was affected in some way by one or more of the test treatments, and the magnitude of up- or downregulation of dysregulated lncRNA gene transcripts was typically similar across treatments as well.

Genomic location data were analyzed with the Idiographica algorithm [36] to visualize the location and orientation of dysregulated lncRNA genes in relation to dysregulated protein-coding genes across all human chromosomes. There were five instances where a dysregulated lncRNA (from the 100 μ M DEET treatment) occupied genomic space within 1000 kilobases (kb) of a dysregulated protein-coding gene transcription start site (TSS). This 1000 kb range from the center of an lncRNA gene to a neighboring protein-coding gene TSS constituted a regulatory region as defined by the Genomic Regions Enrichment of Annotations Tool (GREAT) [37]; this will be discussed in more detail later. A TSS is the first nucleotide base in a section of DNA where RNA polymerase II begins to synthesize a complementary RNA transcript at the 5' end of a gene [38]. In the 100 μ M DEET treatment, there were even two examples of a dysregulated lncRNA gene residing within 300 kb of another dysregulated

lncRNA gene, suggesting that some of these long noncoding elements may actually influence the activity of one another and/or the protein-coding genes with which each interacts. Figure 3A shows the chromosomal distribution of the dysregulated lncRNAs (20 genes; $p \leq 0.01$) and dysregulated protein-coding genes (152 genes, $p \leq 0.01$) together after 100 μM DEET treatment. Chromosomes 1, 4, 5, 6, 7, 9, 11, 16, 17, 19, and X had lncRNAs that were dysregulated by the treatment, and we could clearly visualize (via the ideogram) the protein-coding genes that were closely oriented with those 20 lncRNA genes. Figure 3B–C shows the lncRNA genes and protein-coding genes whose transcripts were differentially expressed when primary human hepatocytes were treated with 10 μM fipronil or a mixture of 100 μM DEET plus 10 μM fipronil, respectively. There were too many genes listed to label gene symbols, so the dysregulated lncRNA genes are simply labeled with blue bars and the dysregulated protein-coding genes are labeled with orange bars. All gene markers on the chromosomal maps represent their relative size (in base pairs) and position in relation to size and orientation to all the known coding and noncoding genes on each human chromosome.

A chromosome-by-chromosome analysis was conducted in an attempt to determine if any correlation existed between the distribution of dysregulated lncRNAs and dysregulated protein-coding genes. The findings are reported in Supplementary Tables S1 and S2. Supplementary Sections S1 and S2 list the chromosomes with the highest and lowest number of lncRNAs and protein-coding genes that were dysregulated by all three of our treatments. We found no clear correlation for each chromosome between the highest percentages of dysregulated lncRNAs versus chromosomes with the highest percentage of dysregulated coding genes. However, there are several instances of dysregulated lncRNAs occurring within a regulatory region of a dysregulated protein-coding gene, some of which will be displayed below.

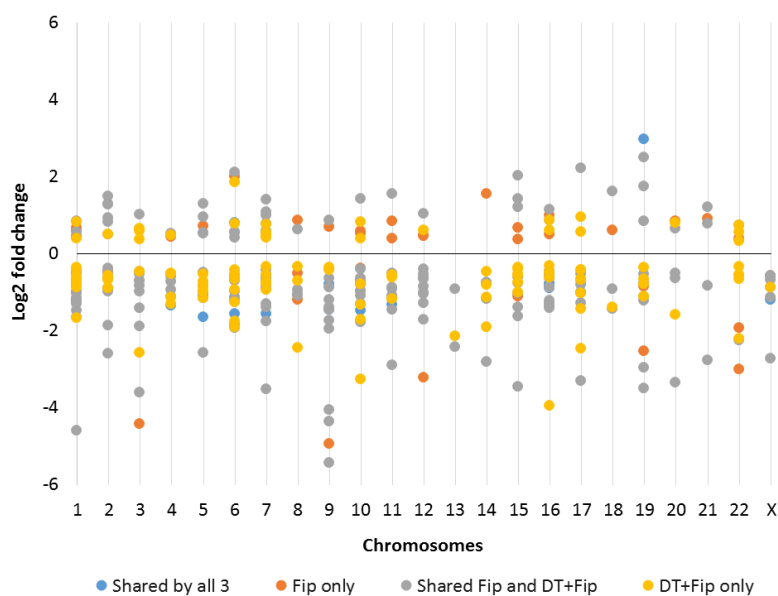
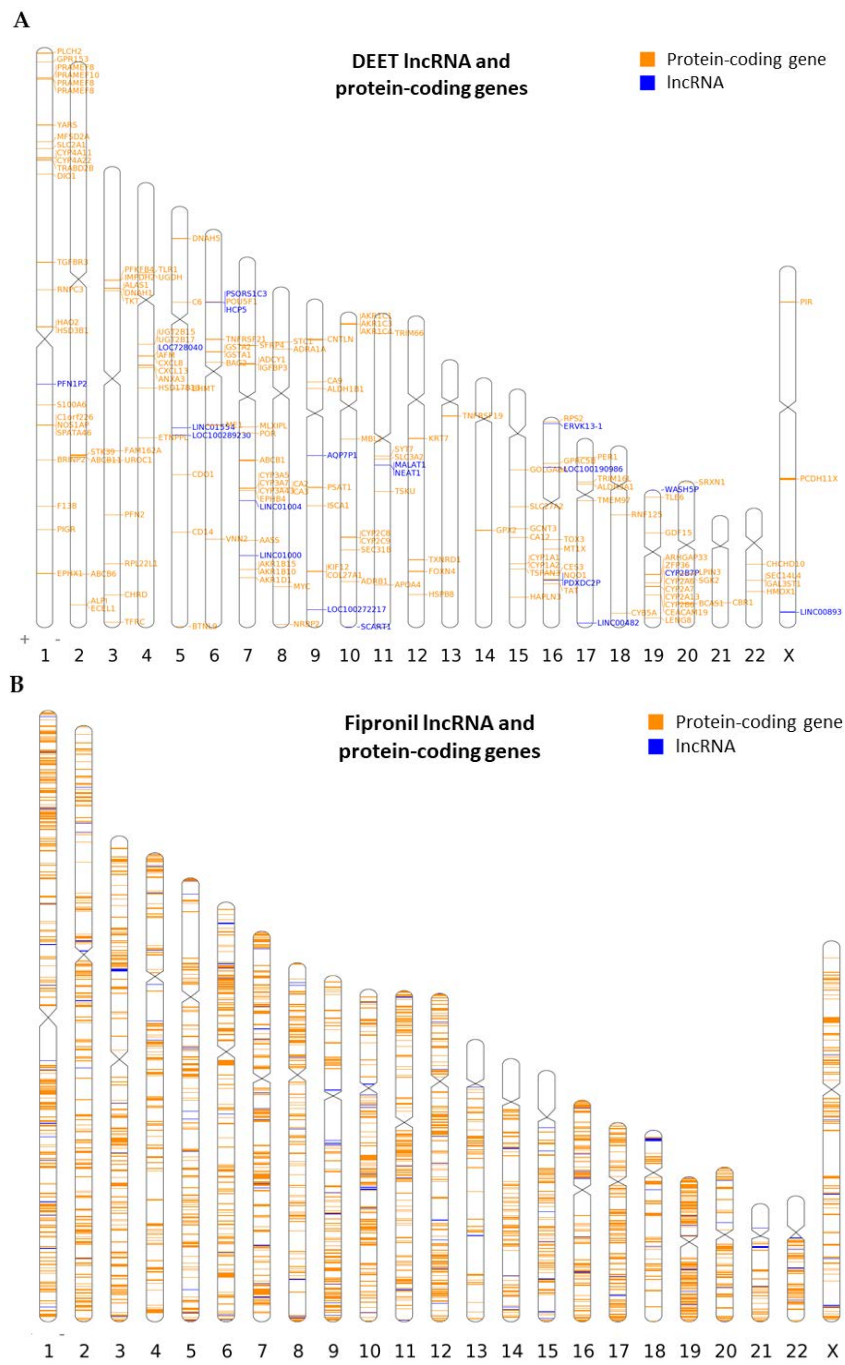


Figure 2. Log₂ fold change of transcripts significantly differentially expressed from long non-protein coding RNA genes by chromosome ($p \leq 0.01$). Shared by all 3 means those transcripts were differentially expressed in all three treatments; Fip only means those transcripts were only differentially expressed when hepatocytes treated with 10 μM fipronil; Shared Fip and DT+Fip means those transcripts were only differentially expressed when hepatocytes treated with fipronil or a combination of DEET and fipronil, but not 100 μM DEET alone; and DT+Fip only means those transcripts were only differentially expressed with the combination of DEET and fipronil, but not each treatment alone. A single representative log₂ fold change value was used for transcripts that were differentially expressed under more than one treatment condition. * A representative log₂ fold change refers to the average log₂ fold change for any genes whose transcript expression was affected by more than one treatment, like shared by all 3, where the same genes were dysregulated by all three treatment conditions.



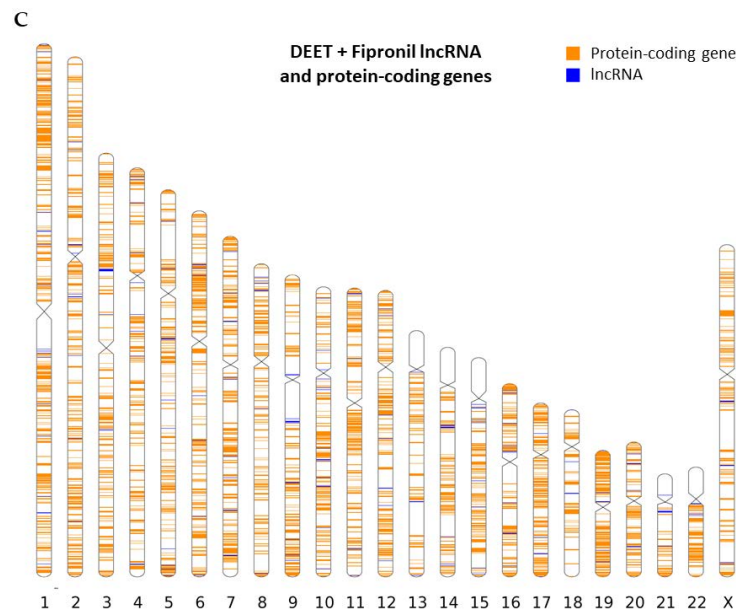


Figure 3. Chromosome maps showing location of lncRNAs with differentially expressed transcripts in relation to dysregulated protein-coding genes after exposure of primary human hepatocytes to 100 μM DEET, 10 μM fipronil, or a mixture of 100 μM DEET and 10 μM fipronil. (A) Chromosomal location of lncRNAs and protein-coding genes significantly dysregulated ($p \leq 0.01$) when hepatocytes exposed to 100 μM DEET; (B) chromosomal location of lncRNAs and protein-coding genes significantly dysregulated ($p \leq 0.01$) when hepatocytes exposed to 10 μM fipronil; (C) chromosomal location of lncRNAs and protein-coding genes significantly dysregulated ($p \leq 0.01$) when hepatocytes exposed to a mixture of 100 μM DEET and 10 μM fipronil. For (A), blue bars and gene symbols denote identity and location of lncRNAs with differentially expressed transcripts (both up- and downregulated). Orange bars and gene symbols denote identity and location of lncRNAs with differentially expressed protein-coding gene transcripts (both up- and downregulated). For (B,C), blue bars denote location of lncRNAs with differentially expressed transcripts (both up- and downregulated). Orange bars denote location of protein-coding genes with differentially expressed transcripts (both up- and downregulated).

2.3. Association of lncRNA Genes with Protein-Coding Genes

To identify which protein-coding genes were associated (i.e., within a pre-determined distance and with a potentially functional relationship) with lncRNAs whose transcripts were differentially expressed after the 100 μM DEET treatment, the 10 μM fipronil treatment, and the 100 μM DEET plus 10 μM fipronil treatment, we used the Genomic Regions Enrichment of Annotations Tool (GREAT). Neighboring protein-coding genes were defined by the GREAT algorithm as those within 1000 kb of an input genomic region (lncRNAs we found to be dysregulated by our treatments in this case). The GREAT algorithm assumed that lncRNAs within 1000 kb of a neighboring gene transcription start site (TSS) could affect the transcription of that neighboring gene, thereby assigning putative functionality to the input lncRNA. The program only considered neighboring protein-coding genes and did not associate lncRNA genes with other noncoding gene transcription regions [37]. Supplementary Figure S1 shows the readings obtained when the dysregulated lncRNAs for each of our three treatments were interrogated with the GREAT algorithm for their proximity (within 1000 kb) to protein-coding genes of known function. A short explanation of the readings is discussed in Supplementary Section S3.

Table 2 shows the 20 lncRNAs whose transcripts were differentially expressed after 100 μM DEET exposure and their neighboring protein-coding genes and lncRNA genes. Included in Table 2 are the protein-coding genes that were dysregulated (shown by an asterisk) and have been described before [18]. Supplementary Table S3 elaborates on information about the specific lncRNAs presented

in Table 2. Supplementary Table S3 provides the GenBank accession number, chromosome coordinates, and whether the lncRNA was significantly up- or downregulated after treatment for each of the 20 dysregulated lncRNAs. Supplementary Table S4 shows the same information as Supplementary Table S3 for the 10 μ M fipronil treatment, and Supplementary Table S5 shows the same from the 100 μ M DEET plus 10 μ M fipronil treatment. After GREAT analysis was utilized to establish the identity of protein-coding genes that were associated (within 1000 kb) with the dysregulated lncRNAs from each treatment condition; we used these genes as input into a functional analysis algorithm to determine within what biological processes and pathways these genes may be active. The findings of the functional characterization are presented later in Section 2.4.

Table 2. Protein-coding and non-protein coding genes neighboring (within 1000 kb) the 20 lncRNAs whose transcripts were up- or downregulated by 100 μ M DEET using GREAT algorithm parameters. The GREAT algorithm defines neighboring genes as those whose transcription start site (TSS) is within 1000 kb of the input lncRNAs. lncRNA = long non-protein coding RNA; kb = kilobases. All gene names are HUGO gene nomenclature committee (HGNC) gene symbols.

lncRNA	Gene(s) within 1000 kb of lncRNA
CYP2B7P	CYP2A7 (−54710), CYP2B6 (−53837), CYP2A6 *, CYP2A13 *
HCP5	MICB (−33621), MICA (+60915)
AQP7P1	ANKRD20A1 (−646908)
MALAT1	SCYL1 (−22962), FRMD8 (+115516), NEAT1 *
SCART1	CYP2E1 (−59218), MTG1 (+67017)
PFN1P2	PPIAL4B (−247525), NBPF9 (−199977)
PDXDC2P	PDPR (−92503), CLEC18A (+69943), NQO1 *
LINC01000	CALU (−88173), METTL2B (+174390)
LOC100190986	METTL9 (−166237), NPIP3 (−13482)
PSORS1C3	POU5F1 (−5124), HLA-C (+96269), HCP5 *
LINC01554	GLRX (−33468), ELL2 (+105889)
LOC100272217	FUBP3 (−1184)
LINC00893	IDS (−28345), CXorf40A (−6965)
NEAT1	SCYL1 (−100412), FRMD8 (+38066), MALAT1 *
WASH5P	OR4F17 (−41513)
LOC100289230	CHD1 (−3535)
LINC00482	SLC38A10 (−10731), TMEM105 (+24638)
LINC01004	KMT2E (−27723), LHFPL3 (+657799)
ERVK13-1	KCTD5 (−16561), PDPK1 (+127950), RPS2 *
LOC728040	AFM (+36985), RASSF6 (+101963), CXCL8 *

* Neighboring differentially expressed protein-coding gene or lncRNA found previously to be within 1000 kb of dysregulated lncRNA (after 100 μ M DEET treatment) before the GREAT algorithm parameters were implemented.

Figure 4A displays magnified regions on specific chromosomes (via the Idiographica algorithm) where lncRNAs with differentially expressed transcripts were closely associated (based on distance in base pairs) with neighboring dysregulated protein-coding genes for the 100 μ M DEET treatment. The three specific regions selected are examples from a total of five cases of dysregulated lncRNAs that were within 1000 kb of protein-coding genes that were dysregulated by exposure to 100 μ M DEET. Figure 4B shows the dysregulated lncRNAs and protein-coding genes from the 10 μ M fipronil treatment in the same regions that were highlighted with the DEET-only data in Figure 4A. The same dysregulated protein-coding and lncRNA genes that were neighbors in the 100 μ M DEET treatment were visible for the fipronil treatment in addition to many other protein-coding genes that were not dysregulated by the 100 μ M DEET treatment; recall the fipronil treatment activated 13.5X as many differentially expressed lncRNA gene transcripts and 21.3X as many dysregulated protein-coding gene transcripts as the DEET-only treatment. The analysis presented in the following sections includes the coding and noncoding genes that were within 1000 kb of neighboring lncRNAs (termed *cis* activation) because the GREAT algorithm that we used did not focus on long-distance relationships (*trans* activation). It is known that certain lncRNAs can affect distant coding and noncoding genomic

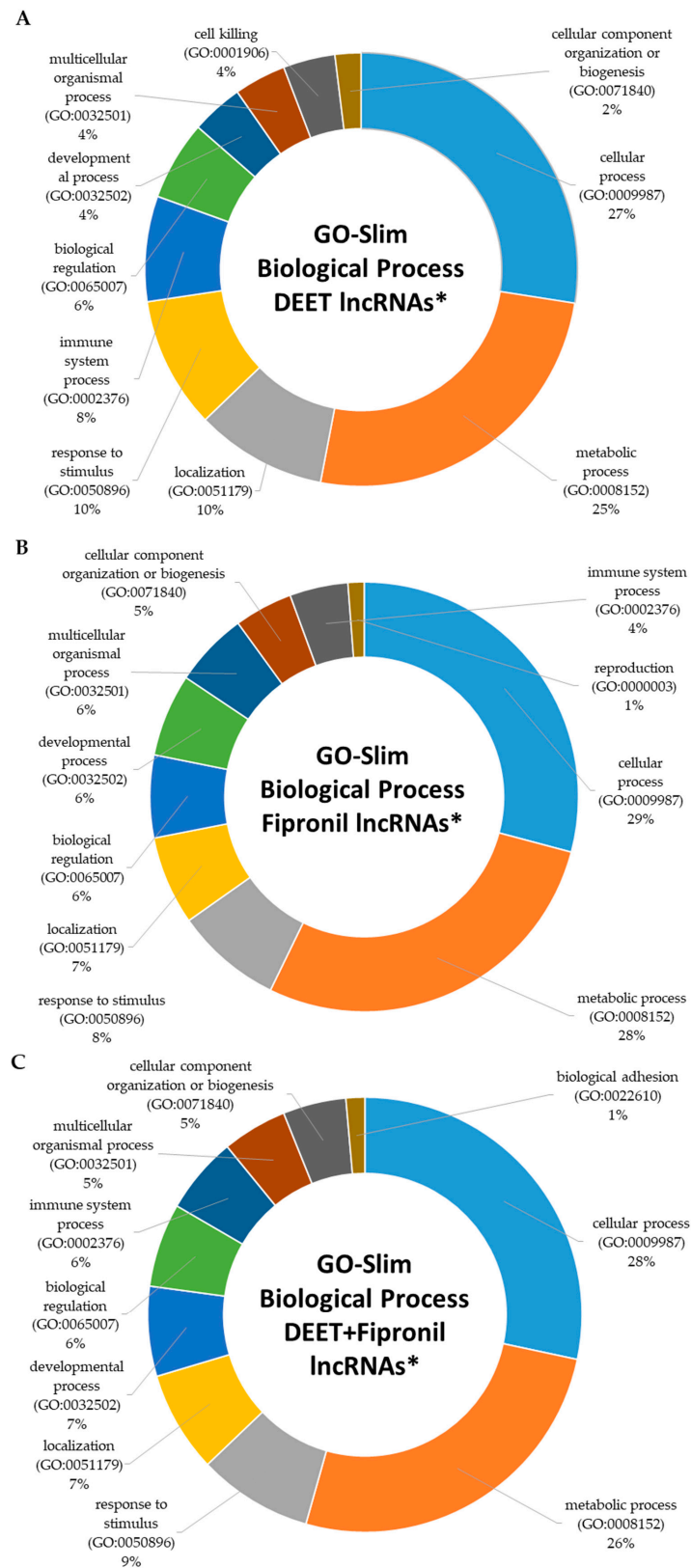


Figure 5. Top 10 biological processes affected by exposure of primary human hepatocytes to (A) 100 μ M DEET; (B) 10 μ M fipronil; and (C) 100 μ M DEET plus 10 μ M fipronil. * Assignments were made using GO-Slim analysis of dysregulated lncRNAs and “associated” protein-coding genes (as determined using the GREAT algorithm as described in methods section). GO = gene ontology assignment number.

2.5. Metabolic Signaling Pathways Associated with Dysregulated LncRNAs

In our previous work, we established that several protein-coding genes whose transcription was up- and downregulated by DEET and fipronil exposure were involved in critical metabolic signaling pathways [18]. Here, we establish that DEET and fipronil influenced the transcription of many lncRNAs that were either directly or indirectly involved in the activity of signaling pathways key to normal cellular function. Figure 6A displays the top 10 signaling pathways (based on the number of genes included in each pathway) affected by exposure to 100 μ M DEET. The 11th signaling pathway affected by 100 μ M DEET that is not displayed in Figure 6A is the Ras pathway. Figure 6B–C displays the top 10 signaling pathways affected by 10 μ M fipronil and a mixture of 100 μ M DEET and 10 μ M fipronil, respectively. In total, the 10 μ M fipronil treatment affected 45 signaling pathways (Supplementary Table S9) and the mixture of 100 μ M DEET and 10 μ M fipronil affected 68 signaling pathways (Supplementary Table S10).

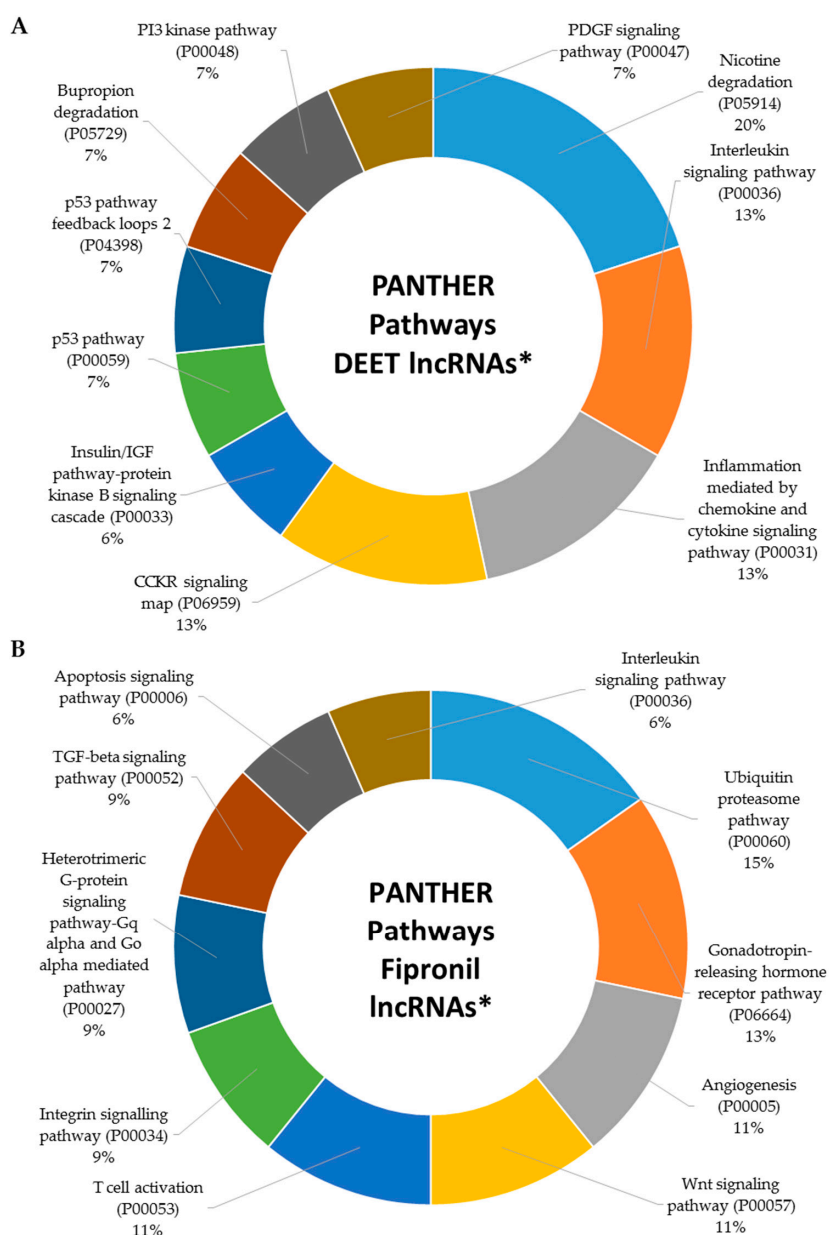


Figure 6. Cont.

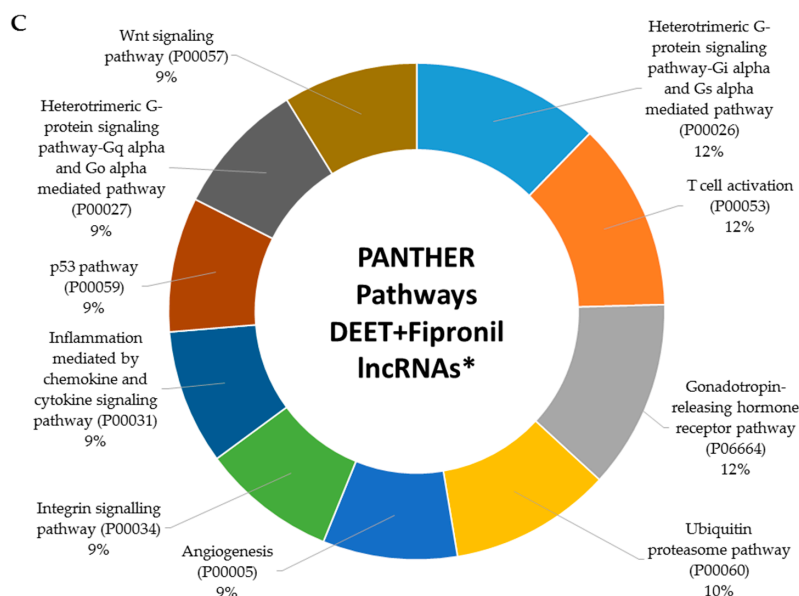


Figure 6. Top 10 signaling pathways affected by exposure of primary human hepatocytes to (A) 100 μ M DEET; (B) 10 μ M fipronil; and (C) 100 μ M DEET plus 10 μ M fipronil. * Assignments were made using PANTHER analysis of dysregulated lncRNAs and “associated” protein-coding genes (as determined using the GREAT algorithm as described in the Methods section). Letters and numbers in parenthesis are PANTHER identifiers.

When we searched the signaling pathway database with the 10 μ M fipronil data, we obtained 45 matches that included all 11 pathways affected by the 100 μ M DEET treatment. This was expected, since all of the dysregulated lncRNAs found in the 100 μ M DEET treatment were also affected by the 10 μ M fipronil treatment. However, the fipronil-only treatment activated 34 additional signaling pathways. The mixture of 100 μ M DEET and 10 μ M fipronil activated many more signaling pathways than either DEET or fipronil alone. We found 68 total signaling pathways associated with the mixture, of which 11 were shared with the 100 μ M DEET and 10 μ M fipronil treatments and the 34 additional matches from the fipronil-only treatment were shared, leaving 23 pathways unique to the response of primary human hepatocytes to a mixture of 100 μ M DEET and 10 μ M fipronil. Some of the pathways most affected by all three treatments included the immune system, p53, Ras, and Wnt signaling pathways. Each is critical to normal cellular function, and interference in these processes is directly linked to disease presentation and progression.

There is much cross talk between metabolic signaling and the immune system or the p53 signaling pathway. Initiation of a targeted immune response, lymphocyte activation for example, is bioenergetically expensive. Therefore, precise control of the interplay between metabolic signaling with the immune system (or other signaling processes we identified as dysregulated in this work) is essential. Interaction between these complex systems is now a major focus in the study of metabolic disorders and cancer [41,42]. The p53 tumor suppressor gene regulates critical metabolic changes in cells. It carefully balances glycolysis and oxidative phosphorylation as well as the autophagy pathway; dysregulation of the p53 pathway can have a profound influence on the behavior of normal and diseased cells [43,44].

2.5.1. Dysregulated LncRNAs Involved in Innate and Adaptive Immunity

The immune system is a defense mechanism that protects the body from foreign invaders, such as bacteria and viruses. Deficiencies in this network can result in autoimmune diseases, inflammatory diseases, and cancer. There is great concern nearly tantamount to a crisis situation regarding the current emergence of drug-resistant bacteria, of which our bodies can no longer successfully combat via both

innate and adaptive immunity [45]. Environmental chemicals that further weaken the immune system exacerbate an already dire dilemma. Transcripts of lncRNA NEAT1 were significantly downregulated in all three treatments (see Table 1), and are known to influence immune gene expression and immune cell functions. NEAT1 does this by binding to splicing factor proline and glutamine rich (SFPQ), which in turn activates the transcription of the gene that codes for chemokine interleukin 8 or IL8 [19]. We did not see differential expression in *SFPQ*, but we did observe the upregulation of *IL8* in the 100 μ M DEET and 100 μ M DEET plus 10 μ M fipronil mixture but not in the 10 μ M fipronil condition. The lncRNA growth arrest-specific transcript 5 (GAS5), whose transcripts were upregulated in the 10 μ M fipronil and 100 μ M DEET plus 10 μ M fipronil mixture treatments, is critical in regulation of the cell cycle and apoptotic control of T cells, and its normal expression is linked to tumor suppression [46]. However, GAS5 overexpression has been shown to inhibit the growth of T cells and promote spontaneous apoptosis [47]. Noncoding repressor of NFAT (NRON) is known to inhibit nuclear factor of activated T cells (NFAT). While we did not observe dysregulation of *NRON*, we did see significant downregulation of *NEAT5* transcripts in both the fipronil-only and DEET plus fipronil mixture treatments [19].

Using GO-Slim annotations in PANTHER, we identified four additional genes neighboring (within 1000 kb) lncRNAs that were involved in either antigen processing and presentation or the immune response in the 100 μ M DEET treatment. MHC class I polypeptide-related sequence A and B (*MICA* and *MICB*) were two of these genes neighboring the lncRNA HLA Complex P5 gene (*HCP5*). Products of the C-X-C Motif Chemokine Ligand 8 gene (*CXCL8*), also associated with *HCP5*, and the Major Histocompatibility Complex, Class I, C protein gene (*HLA-C*) were identified as immune response components. *HLA-C* is associated with the dysregulated lncRNAs psoriasis susceptibility 1 candidate 3 (*PSORS1C3*) and *HCP5*. *HCP5* is an endogenous retrovirus that has become part of the human genome, and is specifically associated with HIV-1 viral load where the expression of one variant of the final protein product of *HCP5* is shown to interact with HIV-1 and reduce its viral presence [48]. When we expanded our signaling pathway search to the dysregulated lncRNAs and neighboring protein-coding genes from the 10 μ M fipronil treatment we obtained matches for 29 immune system components, and if we expand to the 100 μ M DEET plus 10 μ M fipronil lncRNAs and associated protein-coding genes we obtain 45 matches to immune functions. There were several dysregulated genes from our previous study that were included, but which the current study was not limited to, since they also function within the immune response pathway (*FOS*, *JUNB*, *LCP2*, *TLR1*, *TLR2*, *TLR3*, and *TLR4*). Serine/threonine-protein kinase B-Raf (BRAF), whose gene neighbors the differentially expressed lncRNA *NDUFB-AS1* (whose transcripts were downregulated by both the 10 μ M fipronil and 100 μ M DEET plus 10 μ M fipronil treatments), is a protein that transmits signals from the outside to the inside of a cell and is ultimately involved in cell growth and proliferation. Mutations in this gene have been shown to lead to cancer, since the accumulation of mutations in *BRAF*, like many other genes critical to normal human processes, contributes to the development of cancer [49]. The Hallmarks of Cancer, published in 2000 (updated in 2011), was a seminal paper that established six cellular alterations necessary to dictate malignant growth that we still follow today [50,51]. These alterations can arise from perturbations of components in any of the key metabolic processes discussed here.

2.5.2. Dysregulated lncRNAs Involved in the Transformation-Related Protein 53 (p53) Signaling Pathway

The p53 pathway helps the body to respond to stress and prevents genome mutations by activating cell cycle arrest, cellular senescence, DNA repair, or apoptosis. Transformation-related protein 53 (p53) is any isoform of the protein coded from the *tp53* gene and has been referred to as the “guardian of the genome” largely due to its function as a tumor suppressor gene [52]. While p53 regulates a large set of genes, the p53 pathway is itself under the control of multiple self-regulatory pathways, including seven negative and three positive feedback loops [53]. Several lncRNAs have already been implicated in regulation of the p53 pathway at various levels, and we found that some previously

described lncRNAs and neighboring protein-coding genes were affected by all three of our treatment conditions in some capacity. The lncRNA genes *MALAT1* and *MEG3*, whose transcripts are termed p53 regulators, were dysregulated by both DEET and fipronil as described previously. The lncRNA *H19*, whose transcripts are considered p53 effectors, was also dysregulated [20]. We identified one protein-coding gene neighboring the dysregulated lncRNA gene *ERVK13-1* that was connected with p53 activity in the 100 μ M DEET treatment called *PDPK1*. In the 10 μ M fipronil and 100 μ M DEET plus 10 μ M fipronil treatments, there were five (*MDM2*, *RCHY1*, *PDPK1*, *CDKN2B*, and *PRKAB2*) and seven (*PIK3C3*, *HDAC2*, *MDM2*, *RCHY1*, *PDPK1*, *CDKN2B*, and *PRKAB2*) lncRNA genes or lncRNA-associated protein-coding genes linked to p53 activity, one of which was mouse double minute 2 homolog (*MDM2*). The MDM2 protein, associated with the dysregulated lncRNA gene *LOC100130075*, is a well-known regulator of the p53 pathway as it controls six of the ten known feedback loops mentioned above. The presence of MDM2 limits the growth-suppressive functions of p53 by degrading the protein, and MDM2 levels decrease when p53 must respond to stress [54]. In this study, *LOC100130075* transcripts were downregulated by the 10 μ M fipronil and 100 μ M DEET plus 10 μ M fipronil treatments, which could have affected MDM2 activity in relation to p53. The protein 3-phosphoinositide-dependent protein kinase 1 (PDPK1), whose TSS neighbors the lncRNA gene *ERVK13-1* whose transcripts were downregulated in all three treatment conditions, is also connected with p53 activity. PDPK1 is called the “master kinase” because of its importance in signaling pathways tied to growth factors, hormones, and insulin [55]. PDPK1 is a negative regulator of p53, and its levels are elevated in several different types of cancer, including prostate, liver, and breast cancer. Its inhibition has been shown to hinder tumor growth, and it is a promising candidate for cancer intervention [56]. We also noted that several of the protein-coding transcripts that we identified as significantly up- or downregulated previously were directly or indirectly related to p53 functionality, including *ATM*, *SIRT1*, *CDKN2A*, *CDKN2B*, *CREBBP*, *PAK2*, *PDRG1*, *TP53INP2*, *CDIP1*, *PERP*, *RRM2B*, *TRIAP1*, *CSNK2A1*, *CSNK2A2*, and *HIPK2*, which implies that the lncRNAs linked to p53 activity whose transcripts were up- or downregulated by our treatments could have had an influence on these genes or their protein products as well, since they all play some role in the same molecular pathway.

2.6. Specific Well-Studied LncRNAs Dysregulated by DEET and Fipronil

We identified several genes (coding and noncoding) that were significantly affected ($p \leq 0.01$) by 100 μ M DEET, 10 μ M fipronil, and a mixture of 100 μ M DEET plus 10 μ M fipronil. Many of those genes were linked to critical biological processes and signaling pathways as described previously. In Supplementary Section S4, we discuss a subset of lncRNAs whose transcript expression was affected in our experiments that have been well-studied in the field; these are focused on in more detail than other known lncRNAs, and a summary of their expression across our three treatment categories is displayed in Table 3.

Table 3. Subset of well-studied lncRNAs whose transcripts were significantly ($p \leq 0.01$) differentially expressed (Diff. Exp.; up (+) or down (–) regulated) in primary human hepatocytes after exposure to 100 μ M DEET, 10 μ M fipronil, or a mixture of 100 μ M DEET and 10 μ M fipronil for 72 h.

Diff. Exp.	Gene Symbol ^a	Gene Name ^b	FC ^c (DT)	FC ^c (Fip)	FC ^c (DT+Fip)
Up	<i>H19</i>	Long Intergenic Non-Protein Coding RNA 8	-	+1.55	+0.59
Down	<i>HULC</i>	Highly Up-Regulated In Liver Cancer	-	-0.57	-
	<i>MALAT1</i>	Metastasis Associated Lung Adenocarcinoma Transcript 1	-1.34	-1.87	-2.28
	<i>NEAT1</i>	Nuclear Enriched Abundant Transcript 1	-1.10	-1.22	-1.52
	<i>XIST</i>	X (Inactive)-Specific Transcript	-	-0.60	-0.93
	<i>TSIX</i>	TSIX Transcript, XIST Antisense RNA	-	-0.59	-0.97
	<i>MEG3</i>	Maternally Expressed 3	-	-1.19	-1.70

^a HUGO gene nomenclature committee (HGNC) gene symbol [35]; ^b National Center for Biotechnology Information (NCBI) gene description; ^c FC = log₂ fold change; DT = 100 μ M DEET; Fip = 10 μ M fipronil; DT + Fip = 100 μ M DEET plus 10 μ M fipronil mixture.

3. Materials and Methods

3.1. Primary Human Hepatocytes

Plated primary human hepatocytes were obtained from Life Technologies Corporation (Carlsbad, CA, USA). Upon arrival, the medium the cells were shipped in was removed and replaced with fresh, sterile William's E Medium supplemented with additives necessary to properly maintain primary human hepatocytes in culture as previously reported [15,18,57]. The plated cells were then placed in a humidified incubator (relative humidity of 95%) at 5% CO₂/95% air at a temperature of 37 °C for a total of 48 h (media replaced once after 24 h to assess viability and quality).

3.2. DEET and Fipronil Treatments

Treatments with DEET and fipronil began 48 h after the cells arrived. On the same plate, three wells of primary human hepatocytes were each inoculated with DEET (purity >98%; Cat. No. F2284, Chem Service, Inc., West Chester, PA, USA; final concentration 100 µM in each well), fipronil (purity >98%; Cat. No. PS2136, Chem Service, Inc., West Chester, PA, USA; 10 µM), DEET and fipronil (mixed together before adding; 100 µM DEET and 10 µM fipronil), and DMSO. The insecticides were added to the culture media dissolved (*wt/vol*) in DMSO (dimethyl sulfoxide; ≥99.7% pure; Cat. No. BP231-100, Fisher Scientific International, Inc., Hampton, NH, USA). The amount of DMSO (0.1% final concentration) was the same for all treatments, and was previously shown to produce minimal cytotoxicity or changes in gene expression for hepatocytes in culture (LeCluyse et al. [58]).

The concentration levels chosen for the insecticides were determined from the dose–response data for DEET and fipronil, respectively, from Das et al. [15,57] using the same experimental assay conditions for primary human hepatocytes. The DEET concentration chosen (100 µM) is in the intoxication range of what would be expected in human blood within 8 hours after a dermal treatment, 6.7-fold greater than what would be expected in human blood when DEET is appropriately applied at a maximum dose, and approximately one-fifth that for a person subjected to an acute intentional oral overdose of DEET [59]. No data are available on fipronil levels in human blood; the fipronil treatment level in this study was one-tenth that of DEET. Once all cells received the treatment or carrier only, they were incubated undisturbed for 72 h in a humidified incubator under the environmental conditions previously described.

3.3. RNA Isolation and Quality Assessment

After 72 h of treatment, hepatocytes were washed three times with 1X phosphate-buffered saline (PBS, Cat. No. 10010-023, Life Technologies, Carlsbad, CA, USA) and harvested from each individual well (three DEET wells, three fipronil wells, three DEET plus fipronil wells, and three control wells) in a lysis buffer suspension. The suspension was stored at –80 °C or processed immediately via the RNeasy Mini Kit (Cat. No. 74104, Qiagen Inc., Valencia, CA, USA) RNA isolation protocol. Each isolated total RNA sample was separately analyzed for purity on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) by the North Carolina State Genome Sciences Laboratory (Raleigh, NC, USA). No samples with an RNA Integrity Number (RIN) of less than 9.0 were used for sequencing; the lowest RIN obtained was 9.4.

3.4. Illumina Sequencing

Sequencing of all treatments and the controls was performed on the Illumina HiSeq 2000 platform (Illumina, Inc., San Diego, CA, USA) at the Beijing Genomics Institute collaborative genome center at the Children's Hospital of Philadelphia (BGI@CHOP, Philadelphia, PA, USA) and has been described before [18].

3.5. RNA-Seq Analysis

Elements of the Tuxedo suite pipeline [60] were used to analyze the RNA-Seq data. Each of the fastq files was aligned to the hg19 build of the human genome with Cufflinks [61,62], and quality-control and result plots were generated from the Cummebund package as described before [18]. Quality control steps investigating variability between replicates were conducted, and there were no indications of outliers (i.e., negligible variability between replicates observed).

3.6. Dysregulated LncRNA Characterization

The protein and non-protein coding genes whose transcripts were indicated to have been differentially expressed, at a significance level of $p \leq 0.01$, were arranged in Venn diagrams to explore shared and unique genes among our treatment conditions. The lncRNAs with differentially expressed transcripts, at a significance level of $p \leq 0.01$, were extracted from the total transcriptome data set and further classified using the Genomic Regions Enrichment of Annotations Tool (GREAT) to map lncRNAs whose transcripts were up- and downregulated to potential target genes based on proximity to a transcription start site (TSS) and gene annotations of the neighboring proteins [37]. The GREAT algorithm assumes that lncRNA transcription sites within 1000 kilobases (kb) of a “neighboring” gene TSS can affect the transcription of that neighboring gene. If the nearest TSS is over 1000 kb away, then no neighboring protein-coding genes are assigned. Neighboring or associated protein-coding genes are defined by the GREAT algorithm as those within 1000 kb of an input lncRNA transcription site. GREAT calculates the distance between input sequences (lncRNA genes in this case) and target TSSs by measuring the distance in nucleotide base pairs from the middle of each input sequence to the closest TSS of a protein-coding gene. GREAT was run with the binomial and hypergeometric functions disabled, but all other filters intact, since much of the input data (i.e., lncRNAs with up- and downregulated transcripts) were largely unannotated. Dysregulated lncRNAs and their associated protein-coding genes, as refined in the GREAT algorithm, were analyzed for potential functions using the Protein ANalysis THrough Evolutionary Relationships (PANTHER version 11) classification system. PANTHER uses Gene Ontology (GO)-slim terms to classify genes based on annotations established by the Gene Ontology Consortium. GO-slim is useful with large data sets, and is a viable choice when a more broad classification of gene products is desired. Gene symbols (official HUGO gene nomenclature committee (HGNC) gene symbols [35]) from our lncRNAs and neighboring protein-coding genes were also fed into the PANTHER “gene list analysis” tool to visualize and further annotate GO associations and conserved signaling pathways [40].

The web application, Idiographica version 2.3 (<http://www.ncrna.org/idiographica>), was used to develop chromosome maps of genes dysregulated by selected treatments and Venny version 2.0 (<http://bioinfogp.cnb.csic.es/tools/venny>) was used to generate Venn diagrams [36,63]. We also used the Biological Database Network’s “Database to Database Conversions” tool (<https://biodbnet-abcc.ncifcrf.gov/db/db2db.php>) to convert between various identifiers necessary to run specific algorithms [64] and as inputs in Microsoft Excel, Word, and Powerpoint (2013 version).

4. Conclusions

As seen in Supplementary Tables S6–S10, there were many other biological processes and pathways that were not discussed that are also very important in normal cellular function. Transcriptomic analyses revealed to us possible lncRNA-coding gene partnerships and their putative functions, but there is more work remaining to establish a cause and effect relationship. However, the research reported provides potential risks. Our analysis is the first of its kind to link lncRNAs to neighboring protein-coding genes (and other lncRNAs) and functions related to DEET and fipronil exposure, either alone or in combination, in primary human hepatocytes. It can help us to begin to understand the complex molecular interactions that are responsible for the human liver’s response to two common environmental chemicals at the epigenetic level. While this study was performed using

primary human liver cells, we acknowledge that our findings could be different for different cell types or in different organs and tissues. Also, there are questions regarding actual tissue concentrations of insect repellents in humans, as Roy et al. [65] and others have suggested, that have yet to be resolved.

Although many of the lncRNAs we found were uncharacterized, we inferred function from factors such as chromosomal position in relation to protein-coding and noncoding genes that had been previously assigned function and associated with key molecular pathways. We also determined the relative impact of DEET and fipronil on human hepatocytes based on lncRNA expression profiles associated with each. All of this may be useful for future studies that aim to use lncRNAs in measurement of exposure to environmental chemicals as well as prognostic and diagnostic indicators of overexposure and disease. In addition, specific lncRNAs could be utilized for prevention of disease or treatments related to these and other chemicals.

This type of information is becoming more essential, as millions of new chemical compounds are synthesized each year and released into the environment at a wide range of concentrations (for varying durations) as repellents and pesticides, among other uses. We do not know what the response of another chemical or chemical combination would look like, but we felt that DEET and fipronil were a good place to start since they have different modes of action and are commonly used. The rise of the Zika epidemic and associated microcephaly in children, along with other debilitating birth defects, has prompted governmental agencies such as the Centers for Disease Control (CDC) to recommend repellents that are effective at preventing disease transmission. Expecting mothers are encouraged to use DEET on their skin at a minimum concentration of 30% active ingredient every time they are at risk of being bitten by mosquitoes before and during pregnancy [4]. However, we do not really know the potential long term effects of repeated and prolonged use of these chemicals. In addition, fear and misinformation may encourage people to use DEET at higher concentrations over a longer duration to avoid contracting the disease. This could potentially be extremely harmful to unborn or infant children, the elderly, or immuno-compromised individuals along with normal healthy individuals. Finally, even though we tested specific concentrations of two common environmental chemicals, we have no data on the effects of repetitive or long-term exposure to DEET (and fipronil) on human health.

Supplementary Materials: Supplementary materials can be found at www.mdpi.com/1422-0067/18/10/2104/s1.

Acknowledgments: The authors gratefully acknowledge Jeffrey Roach, Ph.D. from the University of North Carolina-Chapel Hill Information Technology Services and Elizabeth Scholl, Ph.D. from the North Carolina State University Bioinformatics Consulting Core for their assistance with bioinformatics analyses. This research was supported in part by the U.S. Central Appalachian Regional Education and Research Center (CARERC) Pilot Study and the North Carolina Agromedicine Institute (Raleigh, NC). Robert D. Mitchell III was supported by an Entomology Department Teaching Assistantship and a Graduate School Doctoral Dissertation Completion Grant at North Carolina State University.

Author Contributions: Ernest Hodgson, R. Michael Roe, Robert D. Mitchell III, and Andrew D. Wallace conceived and designed the experiments; Robert D. Mitchell III and Andrew D. Wallace performed the experiments; Robert D. Mitchell III analyzed the data; Andrew D. Wallace, R. Michael Roe, and Ernest Hodgson contributed reagents/materials/analysis tools; Robert D. Mitchell III and R. Michael Roe wrote the paper.

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

Abbreviations

DEET	<i>N,N</i> -diethyl- <i>m</i> -toluamide
NCBI	National Center for Biotechnology Information
TSS	Transcription start site
HUGO	Human Genome Organisation
HGNC	HUGO Gene Nomenclature Committee
GREAT	Genomic regions enrichment of annotations tool
PANTHER	Protein analysis through evolutionary relationships

References

1. Bernhardt, E.S.; Rosi, E.J.; Gessner, M.O. Synthetic chemicals as agents of global change. *Front. Ecol. Environ.* **2017**, *15*, 84–90. [[CrossRef](#)]
2. Veltri, J.C.; Osimitz, T.G.; Bradford, D.C.; Page, B.C. Retrospective analysis of calls to poison control centers resulting from exposure to the insect repellent *N,N*-diethyl-*m*-toluamide (deet) from 1985–1989. *J. Toxicol. Clin. Toxicol.* **1994**, *32*, 1–16. [[CrossRef](#)] [[PubMed](#)]
3. Mlakar, J.; Korva, M.; Tul, N.; Popović, M.; Poljšak-Prijatelj, M.; Mraz, J.; Kolenc, M.; Resman Rus, K.; Vesnaver Vipotnik, T.; Fabjan Vodusek, V. Zika virus associated with microcephaly. *N. Engl. J. Med.* **2016**, *2016*, 951–958. [[CrossRef](#)] [[PubMed](#)]
4. Morris, H. Zika: The Latest Advice for Travellers. Available online: <http://www.telegraph.co.uk/travel/news/Zika-virus-medical-advice-for-travellers/> (accessed on 22 February 2017).
5. Vargo, E.L.; Parman, V. Effect of fipronil on subterranean termite colonies (isoptera: Rhinotermitidae) in the field. *J. Econ. Entomol.* **2012**, *105*, 523–532. [[CrossRef](#)] [[PubMed](#)]
6. Carrington, D. EU to Ban Fipronil to Protect Honeybees. Available online: <https://www.theguardian.com/environment/2013/jul/16/eu-fipronil-ban-bees> (accessed on 22 February 2017).
7. Hamon, N.; Gamboa, H.; Ernesto, J.; Garcia, M. Fipronil: A major advance for the control of boll weevil in colombia. In Proceedings of the Beltwide Cotton Conferences, Nashville, TN, USA, 9–12 January 1996.
8. Tingle, C.C.; Rother, J.A.; Dewhurst, C.F.; Lauer, S.; King, W.J. Fipronil: Environmental fate, ecotoxicology, and human health concerns. In *Reviews of Environmental Contamination and Toxicology*; Ware, G.W., Ed.; Springer: New York, NY, USA, 2003; pp. 1–66, ISBN 978-1-4899-7283-5.
9. Mohamed, F.; Senarathna, L.; Percy, A.; Abeyewardene, M.; Eaglesham, G.; Cheng, R.; Azher, S.; Hittarage, A.; Dissanayake, W.; Sheriff, M.R. Acute human self-poisoning with the *N*-phenylpyrazole insecticide fipronil—A gabaa-gated chloride channel blocker. *J. Toxicol. Clin. Toxicol.* **2004**, *42*, 955–963. [[CrossRef](#)] [[PubMed](#)]
10. Schoenig, G.P.; Osimitz, T.G.; Gabriel, K.L.; Hartnagel, R.; Gill, M.W.; Goldenthal, E.I. Evaluation of the chronic toxicity and oncogenicity of *N,N*-diethyl-*m*-toluamide (deet). *Toxicol. Sci.* **1999**, *47*, 99–109. [[CrossRef](#)] [[PubMed](#)]
11. Heffernan, A.; English, K.; Toms, L.; Calafat, A.; Valentin-Blasini, L.; Hobson, P.; Broomhall, S.; Ware, R.; Jagals, P.; Sly, P. Cross-sectional biomonitoring study of pesticide exposures in queensland, australia, using pooled urine samples. *Environ. Sci. Pollut. Res.* **2016**, *23*, 23436–23448. [[CrossRef](#)] [[PubMed](#)]
12. Herin, F.; Boutet-Robinet, E.; Levant, A.; Dulaurent, S.; Manika, M.; Galatry-Bouju, F.; Caron, P.; Soulat, J.-M. Thyroid function tests in persons with occupational exposure to fipronil. *Thyroid* **2011**, *21*, 701–706. [[CrossRef](#)] [[PubMed](#)]
13. Usmani, K.A.; Rose, R.L.; Goldstein, J.A.; Taylor, W.G.; Brimfield, A.A.; Hodgson, E. In vitro human metabolism and interactions of repellent *N,N*-diethyl-*m*-toluamide. *Drug Metab. Dispos.* **2002**, *30*, 289–294. [[CrossRef](#)] [[PubMed](#)]
14. Selim, S.; Hartnagel, R.E.; Osimitz, T.G.; Gabriel, K.L.; Schoenig, G.P. Absorption, metabolism, and excretion of *N,N*-diethyl-*m*-toluamide following dermal application to human volunteers. *Fundam. Appl. Toxicol.* **1995**, *25*, 95–100. [[CrossRef](#)] [[PubMed](#)]
15. Das, P.C.; Cao, Y.; Cherrington, N.; Hodgson, E.; Rose, R.L. Fipronil induces cyp isoforms and cytotoxicity in human hepatocytes. *Chem. Biol. Interact.* **2006**, *164*, 200–214. [[CrossRef](#)] [[PubMed](#)]
16. Tang, J.; Usmani, K.A.; Hodgson, E.; Rose, R.L. In vitro metabolism of fipronil by human and rat cytochrome p450 and its interactions with testosterone and diazepam. *Chem. Biol. Interact.* **2004**, *147*, 319–329. [[CrossRef](#)] [[PubMed](#)]
17. Toxicology Data Network. Fipronil. Available online: <https://toxnet.nlm.nih.gov/cgi-bin/sis/search/a?dbs+hsdb:@term+@DOCNO+7051> (accessed on 10 April 2017).
18. Mitchell, R.D.; Dhammi, A.; Wallace, A.; Hodgson, E.; Roe, R.M. Impact of environmental chemicals on the transcriptome of primary human hepatocytes: Potential for health effects. *J. Biochem. Mol. Toxicol.* **2016**, *30*, 375–395. [[CrossRef](#)] [[PubMed](#)]
19. Zhang, Y.; Cao, X. Long noncoding RNAs in innate immunity. *Cell. Mol. Immunol.* **2016**, *13*, 138–147. [[CrossRef](#)] [[PubMed](#)]

20. Zhang, A.; Xu, M.; Mo, Y.-Y. Role of the lncRNA–p53 regulatory network in cancer. *J. Mol. Cell Biol.* **2014**, *6*, 181–191. [[CrossRef](#)] [[PubMed](#)]
21. Villegas, V.E.; Zaphiropoulos, P.G. Neighboring gene regulation by antisense long non-coding RNAs. *Int. J. Mol. Sci.* **2015**, *16*, 3251–3266. [[CrossRef](#)] [[PubMed](#)]
22. Rinn, J.L.; Chang, H.Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **2012**, *81*, 145–166. [[CrossRef](#)] [[PubMed](#)]
23. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914. [[CrossRef](#)] [[PubMed](#)]
24. Kung, J.T.; Colognori, D.; Lee, J.T. Long noncoding RNAs: Past, present, and future. *Genetics* **2013**, *193*, 651–669. [[CrossRef](#)] [[PubMed](#)]
25. Aken, B.L.; Ayling, S.; Barrell, D.; Clarke, L.; Curwen, V.; Fairley, S.; Banet, J.F.; Billis, K.; Girón, C.G.; Hourlier, T. The ensembl gene annotation system. *Database* **2016**, *2016*, baw093. [[CrossRef](#)] [[PubMed](#)]
26. Milligan, M.J.; Lipovich, L. Pseudogene-derived lncRNAs: Emerging regulators of gene expression. *Front. Genet.* **2015**, *5*, 476. [[CrossRef](#)] [[PubMed](#)]
27. Harrison, P.M.; Zheng, D.; Zhang, Z.; Carriero, N.; Gerstein, M. Transcribed processed pseudogenes in the human genome: An intermediate form of expressed retrosequence lacking protein-coding ability. *Nucleic Acids Res.* **2005**, *33*, 2374–2383. [[CrossRef](#)] [[PubMed](#)]
28. Milligan, M.J.; Harvey, E.; Yu, A.; Morgan, A.L.; Smith, D.L.; Zhang, E.; Berengut, J.; Sivananthan, J.; Subramaniam, R.; Skoric, A. Global intersection of long non-coding RNAs with processed and unprocessed pseudogenes in the human genome. *Front. Genet.* **2016**, *7*, 26. [[CrossRef](#)] [[PubMed](#)]
29. Tam, O.H.; Aravin, A.A.; Stein, P.; Girard, A.; Murchison, E.P.; Cheloufi, S.; Hodges, E.; Anger, M.; Sachidanandam, R.; Schultz, R.M. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **2008**, *453*, 534–538. [[CrossRef](#)] [[PubMed](#)]
30. Watanabe, T.; Cheng, E.-C.; Zhong, M.; Lin, H. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Res.* **2015**, *25*, 368–380. [[CrossRef](#)] [[PubMed](#)]
31. Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W.J.; Pandolfi, P.P. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **2010**, *465*, 1033–1038. [[CrossRef](#)] [[PubMed](#)]
32. NCBI, R.C. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **2017**, *45*, D12.
33. Branca, R.M.; Orre, L.M.; Johansson, H.J.; Granholm, V.; Huss, M.; Pérez-Bercoff, Å.; Forshed, J.; Käll, L.; Lehtiö, J. Hirief lc-ms enables deep proteome coverage and unbiased proteogenomics. *Nat. Methods* **2014**, *11*, 59–62. [[CrossRef](#)] [[PubMed](#)]
34. Dündar, F.; Skrabanek, L.; Zumbo, P. Introduction to differential gene expression analysis using RNA-seq. *Appl. Bioinform.* **2015**, 1–67.
35. Gray, K.A.; Yates, B.; Seal, R.L.; Wright, M.W.; Bruford, E.A. Genenames. org: The HGNC resources in 2015. *Nucleic Acids Res.* **2014**, *43*, 1079–1085. [[CrossRef](#)] [[PubMed](#)]
36. Kin, T.; Ono, Y. Idiographica: A general-purpose web application to build idiograms on-demand for human, mouse and rat. *Bioinformatics* **2007**, *23*, 2945–2946. [[CrossRef](#)] [[PubMed](#)]
37. McLean, C.Y.; Bristor, D.; Hiller, M.; Clarke, S.L.; Schaar, B.T.; Lowe, C.B.; Wenger, A.M.; Bejerano, G. Great improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **2010**, *28*, 495–501. [[CrossRef](#)] [[PubMed](#)]
38. Stamatoyannopoulos, J.A. Illuminating eukaryotic transcription start sites. *Nat. Methods* **2010**, *7*, 501–503. [[CrossRef](#)] [[PubMed](#)]
39. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874. [[CrossRef](#)] [[PubMed](#)]
40. Mi, H.; Huang, X.; Muruganujan, A.; Tang, H.; Mills, C.; Kang, D.; Thomas, P.D. PANTHER version 11: Expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **2016**, *45*, 183–189. [[CrossRef](#)] [[PubMed](#)]
41. Ganeshan, K.; Chawla, A. Metabolic regulation of immune responses. *Annu. Rev. Immunol.* **2014**, *32*, 609–634. [[CrossRef](#)] [[PubMed](#)]
42. Tomar, N.; De, R.K. Cross talk between the metabolic and immune systems. *Immunoinformatics* **2014**, *1184*, 13–21.
43. Puzio-Kuter, A.M. The role of p53 in metabolic regulation. *Genes Cancer* **2011**, *2*, 385–391. [[CrossRef](#)] [[PubMed](#)]

44. Zhang, X.-D.; Qin, Z.-H.; Wang, J. The role of p53 in cell metabolism. *Acta Pharmacol. Sin.* **2010**, *31*, 1208. [CrossRef] [PubMed]
45. WHO. U.N. Issues List of 12 Most Worrying Drug-Resistant Bacteria. Available online: <http://www.cbsnews.com/news/un-list-12-most-worrying-bacteria-antibiotic-resistant-superbugs/> (accessed on 27 February 2017).
46. Liu, Y.; Zhao, J.; Zhang, W.; Gan, J.; Hu, C.; Huang, G.; Zhang, Y. LncRNA gas5 enhances g1 cell cycle arrest via binding to ybx1 to regulate p21 expression in stomach cancer. *Sci. Rep.* **2015**, *5*, 10159. [CrossRef] [PubMed]
47. Mourtada-Maarabouni, M.; Hedge, V.L.; Kirkham, L.; Farzaneh, F.; Williams, G.T. Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (gas5). *J. Cell Sci.* **2008**, *121*, 939–946. [CrossRef] [PubMed]
48. Van Manen, D.; Kootstra, N.A.; Boeser-Nunnink, B.; Handulle, M.A.; van't Wout, A.B.; Schuitemaker, H. Association of HLA-c and HCP5 gene regions with the clinical course of HIV-1 infection. *Aids* **2009**, *23*, 19–28. [CrossRef] [PubMed]
49. Davies, H.; Bignell, G.R.; Cox, C.; Stephens, P.; Edkins, S.; Clegg, S.; Teague, J.; Woffendin, H.; Garnett, M.J.; Bottomley, W. Mutations of the braf gene in human cancer. *Nature* **2002**, *417*, 949–954. [CrossRef] [PubMed]
50. Hanahan, D.; Weinberg, R.A. The hallmarks of cancer. *Cell* **2000**, *100*, 57–70. [CrossRef]
51. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [CrossRef] [PubMed]
52. Read, A.; Strachan, T. Chapter 18: Cancer genetics. *Hum. Mol. Genet.* **1999**, *2*.
53. Harris, S.L.; Levine, A.J. The p53 pathway: Positive and negative feedback loops. *Oncogene* **2005**, *24*, 2899–2908. [CrossRef] [PubMed]
54. Moll, U.M.; Petrenko, O. The mdm2-p53 interaction. *Mol. Cancer Res.* **2003**, *1*, 1001–1008. [PubMed]
55. Mora, A.; Komander, D.; van Aalten, D.M.; Alessi, D.R. Pdk1, the master regulator of agc kinase signal transduction. *Semin. Cell Dev. Biol.* **2004**, 161–170. [CrossRef]
56. Mäemets-Allas, K.; Viil, J.; Jaks, V. A novel inhibitor of akt1–pdk1 interaction efficiently suppresses the activity of akt pathway and restricts tumor growth in vivo. *Mol. Cancer Ther.* **2015**, *14*, 2486–2496. [CrossRef] [PubMed]
57. Das, P.C.; Cao, Y.; Rose, R.L.; Cherrington, N.; Hodgson, E. Enzyme induction and cytotoxicity in human hepatocytes by chlorpyrifos and *N,N*-diethyl-*m*-toluamide (deet). *Drug Metabol. Drug Interact.* **2008**, *23*, 237–260. [CrossRef] [PubMed]
58. LeCluyse, E.L.; Madan, A.; Hamilton, G.; Carroll, K.; DeHaan, R.; Parkinson, A. Expression and regulation of cytochrome p450 enzymes in primary cultures of human hepatocytes. *J. Biochem. Mol. Toxicol.* **2000**, *14*, 177–188. [CrossRef]
59. Baselt, R.C. *Disposition of Toxic Drugs and Chemicals in Man*, 4th ed.; Preston Publications Inc.: Niles, IL, USA, 1994; ISBN 0962652318.
60. Trapnell, C.; Roberts, A.; Goff, L.; Pertea, G.; Kim, D.; Kelley, D.R.; Pimentel, H.; Salzberg, S.L.; Rinn, J.L.; Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with tophat and cufflinks. *Nat. Protoc.* **2012**, *7*, 562–578. [CrossRef] [PubMed]
61. Trapnell, C.; Hendrickson, D.G.; Sauvageau, M.; Goff, L.; Rinn, J.L.; Pachter, L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* **2013**, *31*, 46–53. [CrossRef] [PubMed]
62. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; Van Baren, M.J.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **2010**, *28*, 511–515. [CrossRef] [PubMed]
63. Oliveros, J. Venny. An Interactive Tool for Comparing Lists With Venn Diagrams. 2007. Available online: <http://www.citeulike.org/user/hroest/article/6994833> (accessed on 15 August 2017).
64. Mudunuri, U.; Che, A.; Yi, M.; Stephens, R.M. Biodbnet: The biological database network. *Bioinformatics* **2009**, *25*, 555–556. [CrossRef] [PubMed]
65. Roy, D.N.; Goswami, R.; Pal, A. The insect repellents: A silent environmental chemical toxicant to the health. *Environ. Toxicol. Pharmacol.* **2017**, *50*, 91–102. [CrossRef] [PubMed]

