

Origin and Spread of de Novo Genes in *Drosophila melanogaster* Populations

Li Zhao,^{1*} Perot Saelao,¹ Corbin D. Jones,² David J. Begun^{1*}

Comparative genomic analyses have revealed that genes may arise from ancestrally nongenic sequence. However, the origin and spread of these de novo genes within populations remain obscure. We identified 142 segregating and 106 fixed testis-expressed de novo genes in a population sample of *Drosophila melanogaster*. These genes appear to derive primarily from ancestral intergenic, unexpressed open reading frames, with natural selection playing a significant role in their spread. These results reveal a heretofore unappreciated dynamism of gene content.

Although the vast majority of genes present in any species descend from a gene present in an ancestor, recent analyses suggest that some genes originate from ancestrally nongenic sequences (1–3). Evidence for these “de novo” genes has generally derived from a combination of phylogenetic and genomic/transcriptomic analyses that reveal evidence of lineage- or species-specific transcripts associated with nongenic orthologous sequences in sister species. De novo genes, which were first identified in *Drosophila* (1–3), have also been identified in humans, rodents, rice, and yeast (4–9). In *Drosophila*, de novo genes tend to be specifically expressed in tissues associated with male reproduction (2, 10), which suggests that sexual or gametic selection may be important (1–3, 9), although other functional roles may evolve (10, 11). Because previous studies of de novo gene evolution used comparative rather than population genetic approaches, the earliest steps in de novo gene origination remain mysterious. Here, we used population genomic and transcriptomic data from *Drosophila melanogaster* and its close relatives to investigate the origin and spread of de novo genes within populations.

Illumina paired-end RNA sequencing (RNA-seq) and de novo and reference-guided assembly and alignment were used to characterize the testis transcriptome of six previously sequenced inbred Raleigh (RAL) *D. melanogaster* strains (12); an average of 65 million paired-end reads were produced for each strain (table S1). We inferred (13) the presence of 142 polymorphic de novo candidate genes that are expressed in at least one RAL strain but are not known on the basis of publicly available data from *D. melanogaster*. The median number of segregating de novo genes carried per strain was 49. Reverse transcription polymerase chain reaction (RT-PCR) and 5' and 3' rapid amplification of cDNA ends (RACE) in a subset of genes supported inferences from RNA-seq analysis (table S2). These candidate polymorphic genes correspond to unique, intergenic sequence in the *D. melanogaster* reference sequence (table S3), are alignable to unique orthologous regions

in the *D. simulans* and *D. yakuba* reference sequences, and show no significant BLASTP hits to the NCBI nr (nonredundant) protein database. The candidate genes exhibited expression neither in testis RNA-seq data from three *D. simulans* and two *D. yakuba* strains (table S1 and fig. S1) nor in whole male and female RNA-seq data from 59 *D. simulans* strains (13). None of the candidates showed significant expression in whole females from the same *D. melanogaster* strains used for testis RNA-seq (table S4). These data support the hypothesis that the 142 candidates are new, male-specific, de novo genes still segregating in *D. melanogaster*. Expression levels of the candidate genes greatly exceed levels of background transcription in intergenic sequence (fig. S2) (13); several additional attributes of these genes, as described below, support the hypothesis that the observed transcripts are biologically meaningful.

Segregating de novo genes were moderately expressed (Fig. 1A and Table 1), but their ex-

pression was significantly lower than that of annotated male-biased genes (13) (Table 1) or annotated genes (table S6). We observed no enrichment of polymorphic de novo genes near annotated male-biased genes and no significant correlation between the strand (+/–) of polymorphic de novo genes and that of their immediate annotated neighbors [χ^2 test, $P > 0.1$ (table S5 and fig. S3), supported by simulations (13)]. There was a marginally significant underrepresentation of X chromosome segregating de novo genes relative to annotated male-biased genes (10 genes are X-linked; t exact test, $P = 0.01$; Fig. 1B). This result stands in contrast to speculation based on a small sample of older, fixed de novo genes (2, 3) that de novo male-biased genes are overrepresented on the X chromosome.

As expected, de novo genes were significantly shorter and simpler than annotated genes and annotated male-biased genes (Table 1 and table S6). This pattern is likely due mostly to the larger proportion of polymorphic de novo genes that are single-exon (57.0%) compared to the proportion of annotated single-exon (table S6) or single-exon male-biased genes (Table 1) (13). Among the 61 multi-exon de novo genes, the majority of splice events (98%) were associated with canonical sites; rare noncanonical splice sites were found in four genes as minor isoform splice events, which were similar to those previously observed in *D. melanogaster* (14). Alternative splicing was observed in 20 of the 61 multi-exon segregating de novo genes (table S7), with conserved reading frames across alternative isoforms. Genes associated with alternative splicing generally exhibited

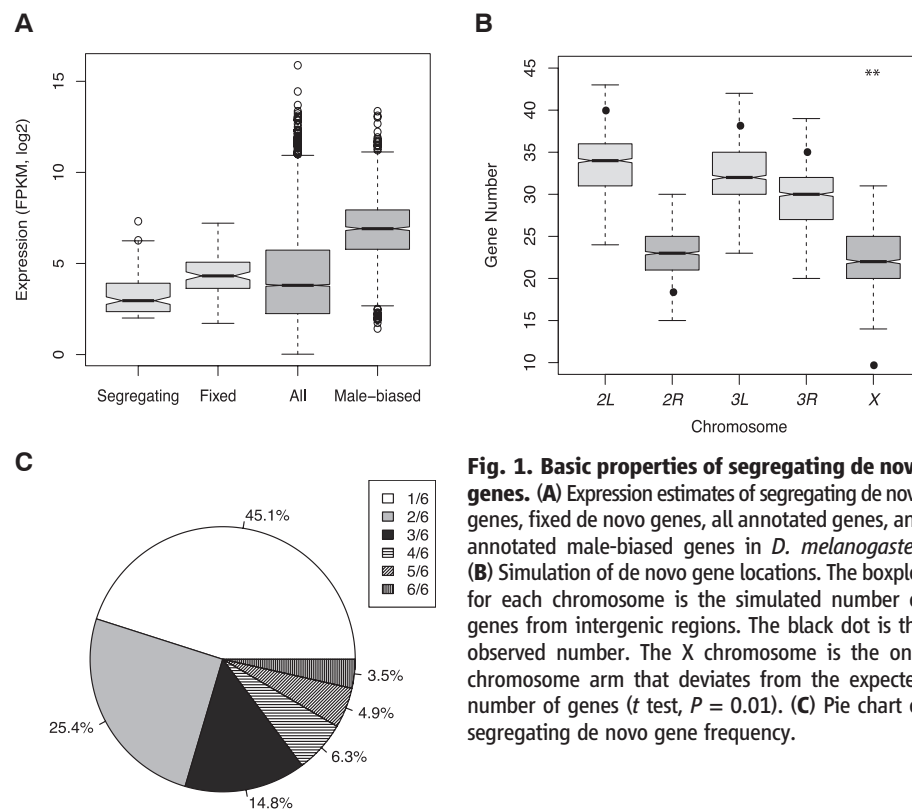


Fig. 1. Basic properties of segregating de novo genes. (A) Expression estimates of segregating de novo genes, fixed de novo genes, all annotated genes, and annotated male-biased genes in *D. melanogaster*. (B) Simulation of de novo gene locations. The boxplot for each chromosome is the simulated number of genes from intergenic regions. The black dot is the observed number. The X chromosome is the only chromosome arm that deviates from the expected number of genes (t test, $P = 0.01$). (C) Pie chart of segregating de novo gene frequency.

¹Department of Evolution and Ecology, University of California, Davis, CA 95616, USA. ²Department of Biology and Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC 27599, USA.

*Corresponding author. E-mail: lizhao@ucdavis.edu (L.Z.); djbegun@ucdavis.edu (D.J.B.)

multiple isoforms across strains that expressed the corresponding gene, with no evidence of genetic variation for alternative splice use.

Of 142 polymorphic genes, 134 (94%) had a minimum open reading frame (ORF) of at least 150 base pairs (bp) and were classified as potentially coding. To determine the likelihood that the high proportion of genes harboring long ORFs occurred by chance, we investigated the coding potential of intergenic regions in the reference sequence, focusing on single-exon ORFs. We observed that 59.9% of random 800-bp intergenic sequences were associated with a ≥ 150 -bp single-exon ORF, whereas 97.5% of the observed single-exon de novo genes were associated with such an ORF ($P < 0.01$). Moreover, the mean length of single-exon de novo gene ORFs was substantially greater than that expected in random intergenic sequence ($P < 0.05$). These observations further support the idea that the observed transcripts are unlikely to be explained simply as random noise. The eight polymorphic de novo genes that did not satisfy our arbitrary minimum ORF criterion were autosomal and slightly smaller (mean transcript length = 743 bp) than ORF-containing polymorphic genes. Because orthologous sequences from expressing and non-expressing *D. melanogaster* lines have similar coding potential, most segregating de novo genes are likely to have resulted from the recruitment of small, preexisting, unexpressed ORFs (1). For *D. simulans* and *D. yakuba* orthologous sequences, 70% and 45%, respectively, contained ORFs similar to those observed for segregating genes in *D. melanogaster*. Of the 134 predicted de novo proteins, 41.8% may be intrinsically unfolded (fig. S4, A to D) and 50% of these have predicted binding regions (fig. S4E); both observations are consistent with potential biological function (15). For putative protein-coding genes, the average 5' and 3' untranslated region (5'UTR and 3'UTR) lengths—248 bp and 364 bp, respectively—were slightly shorter than the average lengths for annotated *D. melanogaster* genes but were slightly longer than the averages for annotated male-biased genes (Table 1). The incidence of the two major polyadenylation signals (AAUAAA and AUUAAA) in or near the putative 3'UTRs of segregating de novo genes was similar to, but slightly lower than, the incidence in the whole genome (table S8). Overall, polymorphic de novo genes have structural organization consistent with small protein-coding genes in the species.

Segregating de novo genes either were expressed at a relatively high level in expressing strains or showed almost no evidence of expression in other strains. Hartigan's dip test on transcript abundance estimates rejected unimodality for 134 of 142 genes and was consistent with bimodal expression across lines for most genes. We used a cutoff of two fragments per kilobase of exon per million fragments mapped (FPKM > 2) for inferring expression of a transcript in a line (16) to determine the proportion of strains, from 0.17 (1/6) to 1.0 (6/6) expressing each transcript.

Because no candidates show expression in the reference sequence strain, the genes expressed in all six RAL strains are considered to be polymorphic in the species. More than half the genes (55%) were not rare in the Raleigh sample, as they were expressed in at least two of the six RAL strains (Fig. 1C); 29.5% were definitely common, being expressed in three or more strains, which is inconsistent with mutation-selection balance. We observed 106 unannotated male-specific transcripts expressed in all six strains and in the reference strain (table S9) but not in the outgroup strains. The corresponding "fixed" de novo genes were not included in downstream analyses relating to segregating genes.

We extracted the 100 bp upstream and 50 bp downstream of the inferred transcription start site (TSS) from the genome sequences of the expressing strains for each of the 61 multi-exon genes. MDscan identified and clustered motifs in these flanking sequences; sequence logos were then generated. We observed four common consensus sequence motifs (8 or 10 bp; Fig. 2A), each of which was found associated with roughly half the segregating de novo genes (13) (table S10). In total, 371 annotated male-biased genes (23.3%) were also associated with at least one of these motifs, which suggests that the de novo genes share regulatory features with known male-biased genes. We identified 67 annotated male-biased genes (table S11) that have two or more motifs in the 5' regions. However, GO (Gene Ontology) enrichment analysis (fig. S5) provided no insight into the possible functions of de novo genes.

These data support the hypothesis that de novo gene expression is influenced by cis-acting variants in the regions corresponding to the 5' flanking regions of expressing chromosomes. In the simplest case that de novo gene expression is due to a single noncoding nucleotide change, one would predict an excess of fixed differences between expressing and non-expressing chromosomes in flanking regions compared to random samples of intergenic sequences. We focused on the 32 genes expressed in more than two strains and for which our genetic analysis (13) supported cis-

acting variation driving de novo gene expression. Of these genes, 31.2% exhibited a fixed, derived single-nucleotide polymorphism (SNP) within 500 bp upstream of the TSS, whereas only 8.43% of simulated "genes" (intergenic regions defined by harboring derived SNPs with the same frequency distribution as the 32 observed genes) exhibited a fixed SNP in the comparable 5' region ($P < 0.01$). More generally, divergence between expressing and non-expressing chromosomes for these 500-bp regions was significantly greater than divergence in simulated data ($P = 0.048$); this finding supports the hypothesis that cis-regulatory changes play a role in de novo gene origination.

Under this hypothesis, segregating genes should be associated with allele-specific expression. We thus measured allelic imbalance (17, 18) in the testis in a set of three unique F₁ genotypes created by crossing the six RAL strains (table S1) (13). For the 59 autosomal genes for which one parent expressed the gene and the other did not, expression patterns in the heterozygote for 51 genes were explained completely by cis-acting variation (i.e., allelic imbalance was complete); 7 genes showed evidence of regulation by both cis-acting and trans-acting factors. Only 1 of the 59 genes showed no evidence of allelic imbalance, consistent with expression driven solely by trans-acting variation (table S12). More generally, for genes expressed in both parents, the expression of alleles in F₁ was consistent with expression levels in each parental line (table S13), further supporting the importance of cis-acting expression variants. The roughly bimodal expression patterns and the dominant role of cis effects support the idea that the proportion of lines expressing a gene provides an estimate of its population frequency.

One population genetic explanation for polymorphic de novo genes is that singleton genes (45% of genes) are primarily deleterious and that higher-frequency genes are primarily neutral. If the deleterious nature of de novo genes were due to the cost of transcription or translation, or from toxic interactions of the resulting RNAs or proteins with other molecules, then lower-frequency genes should be more abundantly expressed and

Table 1. Properties of segregating and fixed de novo genes and comparison with annotated male-biased genes in *D. melanogaster*. Wilcoxon test, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$; ns, not significant. For segregating de novo genes, P values are comparisons of segregating versus fixed genes and segregating versus male-biased genes. For fixed de novo genes, P values are comparisons of fixed de novo genes versus male-biased genes. Male-biased genes are as defined in (13). All estimates are medians, except for exon number (mean).

	Segregating de novo genes	Fixed de novo genes	Male-biased genes
Number	142	106	1595
Transcript length (bp)	801 ns/***	1013**	1184
Exon length (bp)	518*/***	512***	355
Exon number	1.47*/***	1.79*	2.37
Intron length (bp)	91*/***	70.5***	77
5'UTR length (bp)	248*/***	267.5***	170
3'UTR length (bp)	364 ns/***	337***	267
Single-exon gene (%)	57*/***	48.1***	35.8
Expression (FPKM)	7.78***/***	19.96***	66.54

longer than higher-frequency genes. However, contrary to this expectation, lower-frequency genes were expressed at a lower level, were shorter, and were less complex than higher-frequency genes (table S6) (13). The different properties of rare versus common de novo genes (Table 2) (13) supports the idea that de novo genes having certain properties (e.g., greater expression, longer transcripts, more exons) are more likely to spread under selection.

We investigated the role of directional selection on polymorphic de novo genes by determining whether they are associated with reduced

nucleotide diversity (19, 20). For each de novo gene expressed in at least two strains, we compared the nucleotide diversity (π) for expressed sequence (strains) versus non-expressed orthologous sequence (non-expressing strains) and compared the observed differences to a frequency-corrected expected value from resampling of intergenic sequence from the six RAL strains (13). For 46 of 65 genes, π was lower in the expressed lines (mean = 0.0060) than in the non-expressed lines (mean = 0.0092) and exhibited a roughly 38% reduction relative to non-expressed orthologous sequence

over the 65 genes (Wilcoxon test, $P=0.003$). For 30 genes, π was significantly lower in the expressed lines (Wilcoxon test, $P < 0.05$). The region of reduced heterozygosity near expressed sequences is on the scale of 5 kb or less (Fig. 2B and fig. S6), which is contrary to the expectation of strong selection on new mutations (19) but consistent with weaker selection (20) or soft selective sweeps (21) (Fig. 2, C and D). Polymorphic de novo genes were significantly (Wilcoxon test, $P < 0.001$) (13) more likely to be differentially expressed between populations (29 of 142, or 17%) relative to annotated genes (4.5%) and male-biased genes (6.3%), which also supports the idea that selection may play a role in their spread.

We used the Hudson-Kreitman-Aguade-like (HKA1) test statistic (22, 23) to compare the heterozygosity/divergence ratio for genomic regions associated with fixed de novo genes to that observed for appropriately sampled intergenic regions (13, 20). The HKA1 for fixed regions (mean = -0.48) was significantly smaller than that expected for comparable random intergenic regions (mean = 0.12; Wilcoxon test, $P < 0.001$). Moreover, regions corresponding to fixed genes associated with higher expression (FPKM > 10) exhibited a smaller HKA1 statistic relative to regions associated with fixed genes having lower (FPKM \leq 10) expression (HKA1 = -0.33 versus -0.86; Wilcoxon test, $P < 0.001$). These observations also support the hypothesis that de novo genes have been influenced by directional selection.

Our analyses suggest that there are many polymorphic de novo male-specific genes in *D. melanogaster* populations, likely recruited by selection primarily from ancestral, unexpressed ORFs (fig. S7). Given the small number of genotypes investigated for a single tissue and our strict filtering criteria, we have likely substantially underestimated the number of polymorphic de novo genes. Our results also suggest the existence of many more fixed de novo *D. melanogaster* genes than previously inferred (2, 4, 10), which supports the idea that a substantial genetic component of male reproductive biology in this species remains completely unexplored.

More generally, our results suggest that important attributes of an organism's biology cannot be accurately represented or investigated without knowledge of de novo gene variation within species. In the absence of gene loss, de novo gene gain would lead to a long-term increase in gene number. Although our analyses are consistent with substantial numbers of polymorphic gene losses, we observed no population genetic evidence that losses result from directional selection (13). Thus, de novo genes may often spread under selection, while gene loss may occur primarily as a result of drift associated with loss of ancestral gene function. However, important details of such processes remain obscure, and much additional work is required to clarify the dynamics, biochemical and genetic properties, and phenotypic effects of young de novo genes and the processes underlying gene loss in natural populations.

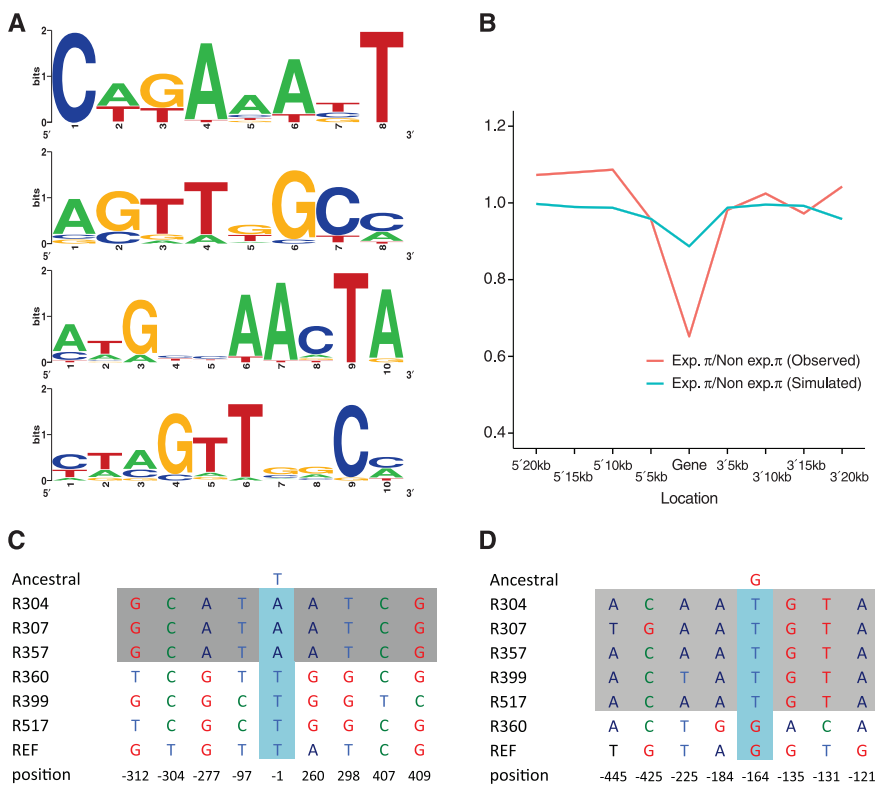


Fig. 2. Regulation and population genetics of segregating de novo genes. (A) Potential cis-regulatory elements. The most common shared 8- and 10-bp consensus motifs in 5' flanking regions are listed. From top to bottom, 34, 29, 25, and 30 multiple-exon genes show these motifs. (B) Nucleotide diversity (π) for de novo genes and flanking regions. The red line is the ratio of π -expressing lines to π -non-expressing lines; the green line shows expected values from resampling of intergenic DNA conditional on the same derived allele frequency distribution as the observed de novo genes. π estimates for 5' and 3' flanking regions of genes were incremented in 5-kb windows. (C) A gene (Gene_X_141) that may have experienced a hard selective sweep. Gray box denotes expressing lines. The TSS region contains a derived allele fixed in expressing strains and absent in non-expressing strains; flanking regions are homozygous in expressing strains. (D) A gene (Gene_3L_079) showing no evidence of hard sweep. Gray box denotes expressing lines. The TSS region includes a derived allele fixed in expressing lines, but the flanking regions of expressing chromosomes retain nucleotide variation.

Table 2. Properties of segregating genes differ across frequency classes. Wilcoxon test, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$; P values are comparison of singleton versus nonsingleton genes and singleton versus high-frequency ($\geq 3/6$) genes. FPKM and transcript length estimates are medians; exon numbers are means.

	Singleton	Nonsingleton	High-frequency
FPKM	5.76***/**	9.91	12.31
Transcript length (bp)	723**/**	869	1312
Exon number	1.38**/**	1.53	1.81

References and Notes

- D. J. Begun, H. A. Lindfors, M. E. Thompson, A. K. Holloway, *Genetics* **172**, 1675–1681 (2006).
- M. T. Levine, C. D. Jones, A. D. Kern, H. A. Lindfors, D. J. Begun, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9935–9939 (2006).
- D. J. Begun, H. A. Lindfors, A. D. Kern, C. D. Jones, *Genetics* **176**, 1131–1137 (2007).
- Q. Zhou *et al.*, *Genome Res.* **18**, 1446–1455 (2008).
- D. G. Knowles, A. McLysaght, *Genome Res.* **19**, 1752–1759 (2009).
- W. Xiao *et al.*, *PLOS ONE* **4**, e4603 (2009).
- J. Cai, R. Zhao, H. Jiang, W. Wang, *Genetics* **179**, 487–496 (2008).
- A. R. Carvunis *et al.*, *Nature* **487**, 370–374 (2012).
- T. J. A. J. Heinen, F. Staubach, D. Häming, D. Tautz, *Curr. Biol.* **19**, 1527–1531 (2009).
- S. Chen, Y. E. Zhang, M. Long, *Science* **330**, 1682–1685 (2010).
- J. A. Reinhardt *et al.*, *PLOS Genet.* **9**, e1003860 (2013).
- T. F. C. Mackay *et al.*, *Nature* **482**, 173–178 (2012).
- See supplementary materials on *Science Online*.
- N. Sheth *et al.*, *Nucleic Acids Res.* **34**, 3955–3967 (2006).
- V. N. Uversky, *Protein Sci.* **11**, 739–756 (2002).
- D. Hebenstreit *et al.*, *Mol. Syst. Biol.* **7**, 497 (2011).
- D. Schwartz, *Genetics* **67**, 411–425 (1971).
- R. M. Graze *et al.*, *Mol. Biol. Evol.* **29**, 1521–1532 (2012).
- J. M. Smith, J. Haigh, *Genet. Res.* **23**, 23–35 (1974).
- N. L. Kaplan, R. R. Hudson, C. H. Langley, *Genetics* **123**, 887–899 (1989).
- J. Hermisson, P. S. Pennings, *Genetics* **169**, 2335–2352 (2005).
- D. J. Begun *et al.*, *PLOS Biol.* **5**, e310 (2007).
- C. H. Langley *et al.*, *Genetics* **192**, 533–598 (2012).

Acknowledgments: We thank the Begun lab for valuable comments, especially N. Svetec, J. Cridland, A. Sedghifar, T. Seher, and Y. Brandvain. We also thank three anonymous reviewers and M. W. Hahn, C. H. Langley, J. Anderson, and A. Kopp for comments and suggestions. We acknowledge Phyllis Wescott (1950–2011) for more than 30 years of service

in the *Drosophila* media kitchen at UC Davis. Funded by NIH grant GM084056 (D.J.B.) and NSF grant 0920090 (C.D.J. and D.J.B.). Author contributions: L.Z. and D.J.B. conceived and designed the experiments; L.Z., P.S., and C.D.J. generated data; L.Z. performed data analysis; P.S. and L.Z. did molecular experiments; L.Z. and D.J.B. wrote the paper; and L.Z., D.J.B., and C.D.J. revised the paper. Illumina reads produced in this study are deposited at NCBI BioProject under accession number PRJNA210329.

Supplementary Materials

www.sciencemag.org/content/343/6172/769/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S17
References (24–58)

11 November 2013; accepted 7 January 2014
Published online 23 January 2014;
10.1126/science.1248286

Crude Oil Impairs Cardiac Excitation-Contraction Coupling in Fish

Fabien Brette,¹ Ben Machado,¹ Caroline Cros,¹ John P. Incardona,² Nathaniel L. Scholz,² Barbara A. Block^{1*}

Crude oil is known to disrupt cardiac function in fish embryos. Large oil spills, such as the Deepwater Horizon (DWH) disaster that occurred in 2010 in the Gulf of Mexico, could severely affect fish at impacted spawning sites. The physiological mechanisms underlying such potential cardiotoxic effects remain unclear. Here, we show that crude oil samples collected from the DWH spill prolonged the action potential of isolated cardiomyocytes from juvenile bluefin and yellowfin tunas, through the blocking of the delayed rectifier potassium current (I_{Kr}). Crude oil exposure also decreased calcium current (I_{Ca}) and calcium cycling, which disrupted excitation-contraction coupling in cardiomyocytes. Our findings demonstrate a cardiotoxic mechanism by which crude oil affects the regulation of cellular excitability, with implications for life-threatening arrhythmias in vertebrates.

Crude oil is a complex chemical mixture containing hydrocarbons (aliphatic and aromatic) and other dissolved-phase organic compounds. Toxicity research on crude oil constituents has focused mainly on polycyclic aromatic hydrocarbons (PAHs) (1, 2), pervasive environmental contaminants that are also found in coal tar, creosote, air pollution, and land-based runoff. In the aftermath of oil spills, PAHs can persist for many years in marine habitats and thereby create pathways for lingering biological exposure and associated adverse effects.

PAH toxicity is structure-dependent, and the carcinogenic, mutagenic, and teratogenic properties of many individual PAHs are known (3, 4). Developing fish are particularly vulnerable to dissolved PAHs in the range of ~100 parts per billion (ppb or $\mu\text{g/liter}$) down to $\leq 10 \mu\text{g/liter}$. Consequently, PAH toxicity to fish early life stages is an important contributor to both acute and long-term impacts of environmental disasters (2, 5).

Numerous studies on crude oils and PAHs, particularly in the aftermath of the Exxon Valdez spill, have described embryonic heart failure, bradycardia, arrhythmias, reduction of contractility, and a syndrome of cardiogenic fluid accumulation (edema) in exposed fish embryos (6, 7). These severe effects are lethal to embryos and larval fishes (8–10) and could be due to atrioventricular conduction block (11).

Despite recent progress using zebrafish and other experimental models to study PAH cardiotoxicity (12), the mechanisms that underpin the physiological effects on cardiac function and changes in cardiac morphology during development are not known. The Deepwater Horizon (DWH) oil spill released >4 million barrels of crude oil during the peak spawning months for Atlantic bluefin tuna (*Thunnus thynnus*) in 2010. This large and long-lived species reaches a mass of 650 kg over a life span of 35 years or more (13), and the Gulf of Mexico population of bluefin tuna is severely depleted (14). Electronic-tagging data confirm that bluefin tuna spawn in the vicinity of the DWH spill, which indicates that bluefin tuna embryos, larvae, juveniles, and adults were likely exposed to crude oil-derived PAHs (14). Many other Gulf of Mexico pelagics

may have spawned in oiled habitats, including yellowfin tuna, dolphin fish, blue marlin, and swordfish (15).

To more precisely define the mechanisms of crude oil cardiotoxicity and to evaluate the potential vulnerability of eggs, larvae, and juveniles in the vicinity of the DWH spill, we assessed the impact of field-collected DWH oil samples on *in vitro* cardiomyocyte preparations dissociated from the hearts of bluefin tuna (*T. orientalis*) and yellowfin tuna (*T. albacares*). Juvenile tunas were caught at sea and held in captivity at the Tuna Research Conservation Center and the Monterey Bay Aquarium (16).

The cardiotoxic effects of four distinct environmental samples of MC252 crude oil were assessed as water-accommodated fractions (WAFs) prepared in Ringer solution for marine fish (16). Oil samples were collected under chain of custody during the DWH spill response effort. The samples included riser “source” oil (sample 072610-03), riser oil that was “artificially weathered” by heating at 90° to 105°C (sample 072610-W-A), and two skimmed oil samples: “slick A” (sample CTC02404-02), collected 29 July 2010, and “slick B” (sample GU2888-A0719-OE701), collected 19 July 2010 by the U.S. Coast Guard cutter *Juniper*. High-energy WAFs were prepared in a commercial blender that dispersed oil droplets to mimic release conditions at the MC252 well head (16). As expected from previous studies (11, 12), the total sum (Σ) of PAHs declined in WAFs from source oil to the surface-weathered samples, owing to loss of naphthalenes, whereas the total concentrations of three-ringed PAHs (e.g., phenanthrenes) increased proportionately (fig. S1 and table S1). PAH concentrations were in a range expected to cause cardiotoxicity in intact embryos and consistent with the Σ PAHs measured in some surface water samples during the DWH oil spill (up to 84 $\mu\text{g/liter}$) (16, 17). WAFs in Ringer solution were perfused over freshly dissociated, isolated tuna cardiomyocytes, and we assessed the effects of these oil-containing solutions on excitation-contraction (EC) coupling using electrophysiological and Ca^{2+} -imaging techniques.

¹Department of Biology, Stanford University, Hopkins Marine Station, Pacific Grove, CA 93950, USA. ²Northwest Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, WA 98112, USA.

*Corresponding author. E-mail: bblock@stanford.edu



Origin and Spread of de Novo Genes in *Drosophila melanogaster* Populations

Li Zhao *et al.*

Science **343**, 769 (2014);

DOI: 10.1126/science.1248286

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of March 20, 2015):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/343/6172/769.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2014/01/22/science.1248286.DC1.html>

This article **cites 55 articles**, 28 of which can be accessed free:

<http://www.sciencemag.org/content/343/6172/769.full.html#ref-list-1>

This article has been **cited by** 3 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/343/6172/769.full.html#related-urls>

This article appears in the following **subject collections**:

Evolution

<http://www.sciencemag.org/cgi/collection/evolution>