ARTICLE

doi:10.1038/nature12594

Odour receptors and neurons for DEET and new insect repellents

Pinky Kain[1]*, Sean Michael Boyle[2]*, Sana Khalid Tharadra[1], Tom Guda[1], Christine Pham[3], Anupama Dahanukar[1,2,3,4] & Anandasankar Ray[1,2,3,4]

There are major impediments to finding improved DEET alternatives because the receptors causing olfactory repellency are unknown, and new chemicals require exorbitant costs to determine safety for human use. Here we identify DEET-sensitive neurons in a pit-like structure in the *Drosophila melanogaster* antenna called the sacculus. They express a highly conserved receptor, Ir40a, and flies in which these neurons are silenced or *Ir40a* is knocked down lose avoidance to DEET. We used a computational structure–activity screen of >400,000 compounds that identified >100 natural compounds as candidate repellents. We tested several and found that most activate *Ir40a*$^+$ neurons and are repellents for *Drosophila*. These compounds are also strong repellents for mosquitoes. The candidates contain chemicals that do not dissolve plastic, are affordable and smell mildly like grapes, with three considered safe in human foods. Our findings pave the way to discover new generations of repellents that will help fight deadly insect-borne diseases worldwide.

Blood-feeding insects transmit deadly diseases such as malaria, dengue, lymphatic filariasis and West Nile fever to hundreds of millions of people, causing immense suffering and more than a million deaths every year. Insect repellents can be very effective in reducing disease transmission by blocking contact between blood-seeking insects and humans.

N,N-diethyl-*meta*-toluamide (DEET) has remained the primary insect repellent used for more than 60 years. However, DEET has little effect on disease control in endemic regions due to high costs and inconvenience of continuous application on skin at high concentrations. DEET also dissolves some plastics, synthetic fabrics and painted surfaces[1]. Additionally, DEET inhibits mammalian acetylcholinesterase[2]. Instances of DEET resistance have also been reported in flies[3] and mosquitoes[4,5]. However, the main barriers in developing improved repellents are the estimated cost for identification[6] and the subsequent cost of safety analyses for new chemistries.

A significant challenge in finding improved DEET substitutes is that the target receptors through which it repels insects are unknown. Recent studies have given rise to many different models of DEET action. Pure DEET causes inhibition[7,8] or mild electrophysiological modification of neural responses to weakly-activating odours in *Drosophila* antennal olfactory neurons[9], but whether these effects contribute to repellency is unknown. Mosquitoes can also directly detect DEET[10] and mutations in the *orco* co-receptor gene in *Aedes aegypti* cause reduction in repellency[11]. Some DEET-sensitive olfactory neurons have been identified in *Culex quinquefasciatus*[10] and *A. aegypti*[5], but it is not yet known whether they are responsible for repellency or which odour receptors they express. A broadly tuned larval odour receptor responds to DEET[12,13]; however, its role in avoidance in larval or adult mosquitoes has not been demonstrated. Not only can more than one pathway contribute to olfactory repellency, analyses are further confounded by the observation that DEET also activates bitter taste neurons that mediate contact-avoidance in *Drosophila*[14,15].

DEET is detected by neurons of the sacculus

To identify the elusive DEET-sensing neurons of the olfactory system in an unbiased manner, we used the nuclear factor of activated T cells (NFAT)-based system to report DEET-evoked neural activity through expression of green fluorescent protein (GFP) in *Drosophila melanogaster*[16] (Fig. 1a). Exposure to 10% DEET resulted in an increase in expression of GFP in neurons that innervate sensilla within the sacculus, a pit-like structure in the antenna (Fig. 1b, c, Supplementary Fig. 1a and Supplementary Video 1). The dendrites of GFP$^+$ neurons primarily innervated the most distal chamber (I) of the sacculus (Fig. 1c and Supplementary Fig. 1b). Previous studies of DEET overlooked the sacculus because it is intractable to traditional electrophysiology methods.

Contrary to expectations from a previous report[17], we were unable to find DEET-activated reporter expression in odorant receptor neurons (ORNs) of the maxillary palps (Fig. 1b). We therefore performed single-sensillum electrophysiology analyses and found that the previously reported *Or42a*$^+$ pb1A neurons responded poorly to DEET, but strongly to hexane that was used as solvent in the previous study (Supplementary Fig. 2a, b).

ORNs innervating the sacculus do not express *Or* genes, but instead members of a conserved ionotropic receptor (*IR*) gene family[18–21]. In the antennal lobes robust DEET-dependent GFP was detected in the characteristic 'column' glomerulus (Fig. 1d and Supplementary Fig. 3a), which is innervated by axons of *Ir40a*-expressing neurons of the sacculus[18]. Faint GFP was also observed in the *Or67d*$^+$ DA1 glomerulus, which is probably caused by exposure to male pheromone cis-vaccenyl acetate (cVA) in the assay, because the cVA-responsive *Or67d*$^+$ at1 neuron did not respond to DEET (Supplementary Fig. 2c). The DC4 glomerulus, which is innervated by other sacculus ORNs that express *Ir64a*[19], showed a very faint signal as well (Supplementary Fig. 3a). The simplest interpretation of these results is that *Ir40a*$^+$ sacculus ORNs innervating chamber I and projecting to the column glomerulus may represent a chief olfactory detection pathway for DEET.

Consistent with previous electrophysiological analyses[14,15], we found DEET-dependent GFP expression in gustatory neurons of the labellum

[1]Department of Entomology, University of California, Riverside, California 92521, USA. [2]Genetics, Genomics and Bioinformatics Program, University of California, Riverside, California 92521, USA. [3]Neuroscience Program, University of California, Riverside, California 92521, USA. [4]Institute of Integrative Genome Biology, University of California, Riverside, California 92521, USA.
*These authors contributed equally to this work.

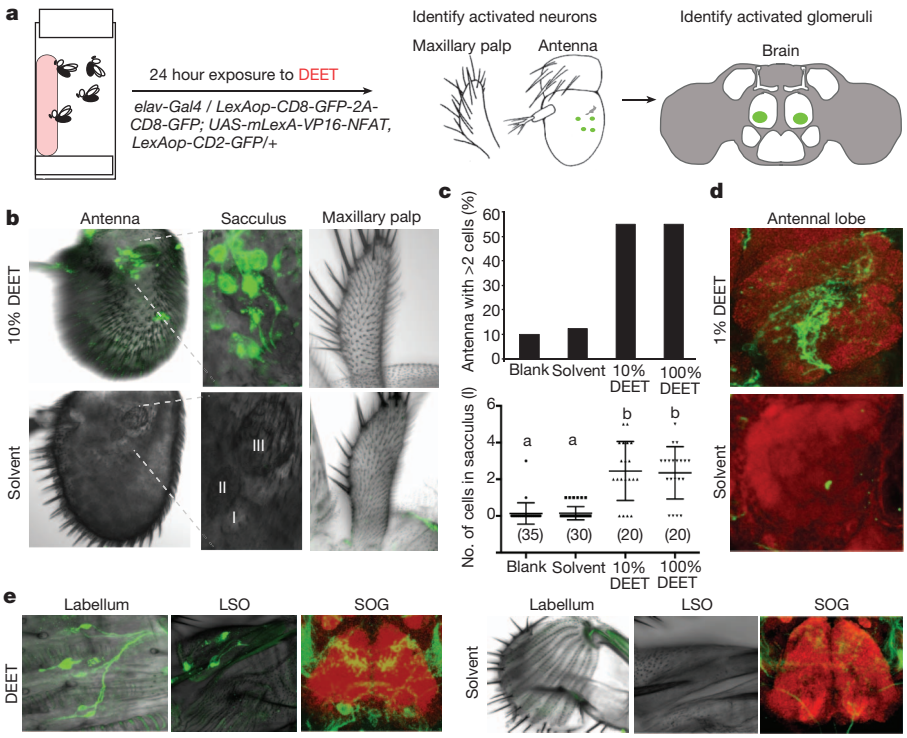


Figure 1 | DEET is detected by *Ir40a*$^+$ sacculus neurons. **a**, Schematic of the NFAT (CaLexA)-based method to label neurons activated by DEET. **b**, Confocal micrographs of olfactory organs from flies stimulated with 10% DEET or solvent (acetone). **c**, Quantification of GFP$^+$ antennae (upper graph) and mean numbers of GFP$^+$ cells in chamber I (lower graph). $n = 35$ (blank), $n = 30$ (solvent), $n = 20$ (10% DEET), $n = 20$ (100% DEET). $P < 0.0001$, one-way ANOVA with Tukey's post hoc test. **d**, GFP$^+$ axonal termini in antennal lobes of flies treated as indicated. **e**, **f**, Expression of GFP in the labellum, labral sense organ (LSO) and suboesophageal ganglion (SOG). Anti-GFP (green) and anti-nc82 (red). For SOG, dorsal is top. Error bars indicate s.e.m.

(Fig. 1e). In addition, we observed DEET-dependent GFP in neurons innervating the labral sense organ (LSO) of the pharynx (Fig. 1e). The DEET activity mapped to neurons marked by *Gr33a* and *Gr89a*, which are bitter-sensing deterrent neurons (Supplementary Fig. 3b). Axonal projections of DEET-sensitive gustatory neurons in the suboesophageal ganglion (SOG) revealed arborization patterns similar to those of taste neurons originating in the labellum and the pharynx (Fig. 1e, Supplementary Fig. 3b).

In order to directly test physiological responses of the sacculus *Ir40a*$^+$ ORNs to DEET we performed *in vivo* calcium imaging in flies expressing GCaMP3 using *Ir40a-Gal4*[18,22,23]. *Ir40a* neurons show robust activation in response to a puff of DEET delivered from an atomizer but not to control solvent dimethylsulphoxide (DMSO) (Fig. 2a, b). Moreover, the DEET response is dependent on *Ir40a* (Fig. 2c).

In order to test whether the *Ir40a*$^+$ ORNs are required for DEET repellency we blocked synaptic transmission in these neurons using *Ir40a-Gal4* to express the active form of tetanus toxin (TNTG)[24]. We used a trap lured by 10% apple cider vinegar (ACV) in which a DEET-treated filter paper was placed inside the trap. Avoidance was significantly decreased in *Ir40a-TNTG* flies as compared to various controls, including a non-functional version of the tetanus toxin (IMPTV), suggesting that *Ir40a*$^+$ neurons are required for DEET repellency (Fig. 2d). All genotypes exhibited attraction to 10% ACV in two-choice trap assays (Supplementary Fig. 4a).

Ir40a is necessary for DEET avoidance

To test directly whether *Ir40a* is required for olfactory avoidance to DEET, we examined the behaviour of flies in which *Ir40a* was knocked down pan-neuronally using an *elav-Gal4* driver to express a *UAS-Ir40a* RNA interference (RNAi) construct. In two-choice trap assays (Fig. 3a), we found a significant loss of DEET avoidance in the *Ir40a* RNAi flies compared to control flies (Fig. 3b). Similar results were obtained when *Ir40a* RNAi was executed selectively in *Ir40a*$^+$ ORNs using two independent *UAS-Ir40a* RNAi transgenes (Fig. 3c). Not only was avoidance completely abolished, *Ir40a* knockdown flies actually showed a mild attraction to the DEET trap. Attraction to ACV was unaffected (Supplementary Fig. 4b, c).

We next wanted to rule out the possibility of a developmental role for *Ir40a*. We therefore suppressed expression of *Ir40a-RNAi* during development using a temperature-sensitive *Gal80*^{ts} transgene (Fig. 3d). Flies were raised at the permissive temperature (18 °C) until just before adult eclosion, at which point they were left at 18 °C (RNAi off) or shifted to the Gal80^{ts} restrictive temperature 29 °C (RNAi on). Behavioural assays performed four days after the temperature shift showed that *Ir40a* RNAi in the adult was sufficient to abolish DEET avoidance when RNAi was induced in *Ir40a*$^+$ ORNs (Fig. 3e, knockdown).

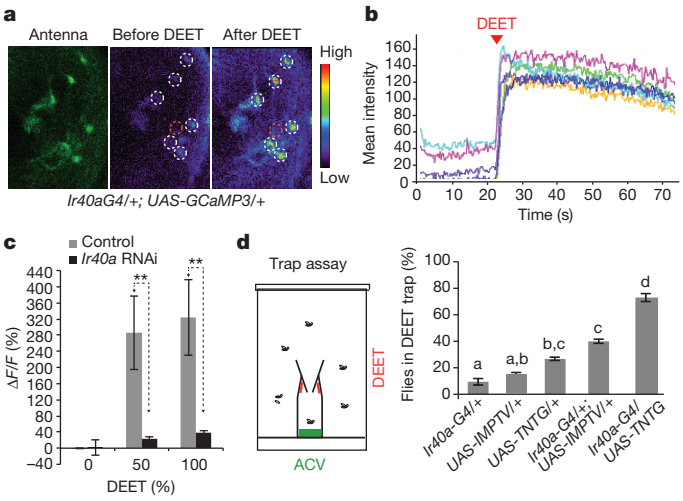


Figure 2 | *Ir40a* neurons detect DEET and are required for repellency. **a**, Images of calcium activity in *Ir40a-Gal4/+;UAS-GCaMP3/+* neurons colour-coded as indicated (right). Measurements taken from areas in dashed circles: cells (white), background (red). **b**, Mean fluorescence intensities for six different cells. Red arrowhead indicates onset of ~2-s puff of DEET. **c**, Mean percentage change in fluorescence intensity after application of ~2-s indicated stimulus; genotypes were *Ir40a-Gal4/+;UAS-GCaMP3/+* (control) and *Ir40a-Gal4/Ir40a-Gal4;UAS-GCaMP3/UAS-Ir40a* RNAi (line number 2) indicated as (*Ir40a*-RNAi). $n = 10$–13. $^{**}P < 0.01$, Student's *t*-test. ACV, apple cider vinegar. **d**, Schematic (left) and results (right) for DEET-treated trap assays for indicated genotypes. $n = 6$ trials, 20 flies per trial for each genotype. Letters indicate statistical significance, $P \leq 0.008$, one-way ANOVA with Tukey's post hoc analysis. Error bars represent s.e.m.

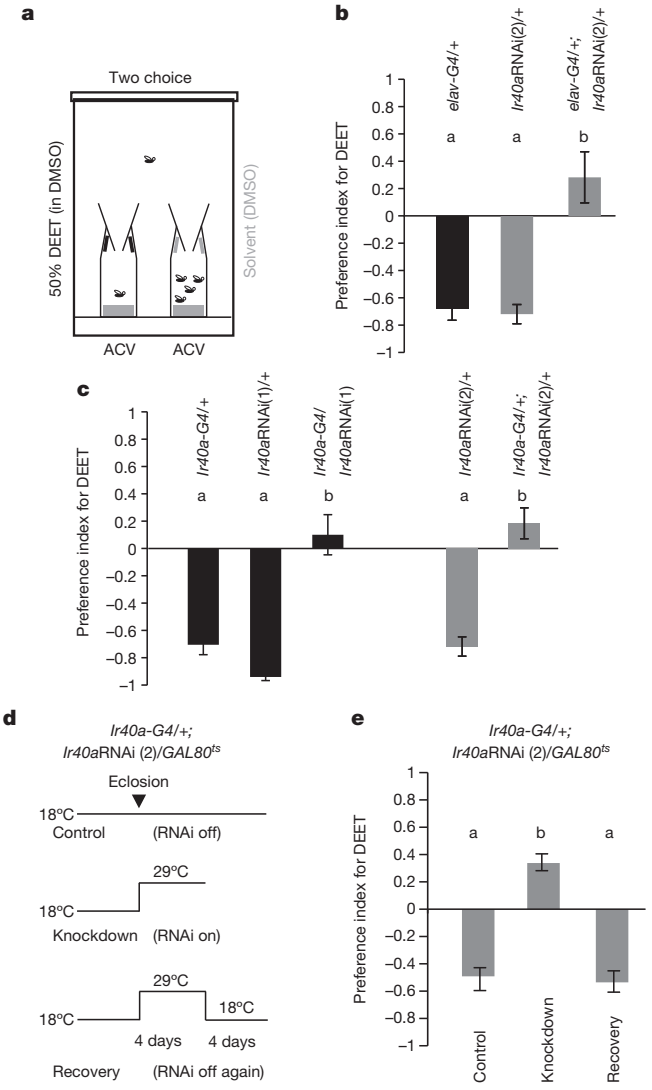


Figure 3 | *Ir40a* is required for DEET avoidance. **a**, Set-up for behavioural two-choice assay. **b**, **c**, Mean preference index of indicated genotypes for DEET in two-choice assays using *elav-Gal4* (***b***) and *Ir40a-Gal4* (***c***). $n = 6$ trails (20 flies per trial) except *elav-Gal4/+;Ir40aRNAi(2)* $n = 10$ trials and RNAi experiments with *Ir40a-Gal4* $n = 12$ trials each. **d**, Genotype and schematic for post-developmental knockdown and recovery of *Ir40a*. **e**, Mean DEET preference index of flies derived from indicated treatments in two-choice assays. $n = 6$ trials for all conditions, with 20 flies per trial. For **b**–**e**, $P < 0.001$, one-way ANOVA with Tukey's post hoc analysis. Error bars represent s.e.m.

Moreover, DEET avoidance was completely restored when flies were returned to the Gal80^{ts} permissive temperature (Fig. 3e, recovery). Attraction to ACV was unaffected (Supplementary Fig. 4d). Taken together, these experiments demonstrate that *Ir40a* is required in adult *Ir40a*$^{+}$ sacculus ORNs for olfactory avoidance of DEET.

In silico prediction of new repellents

Identification of DEET receptors and neurons offers a powerful system to screen for improved repellents. However, volatile chemical space that can be exploited to find DEET substitutes is vast and therefore poses unfeasible requirements in terms of cost and time to screen. The receptor structure is unavailable for screening and the most effective repellents may require detection by both olfactory and gustatory pathways. To circumvent these limitations we developed a high-throughput chemical informatics screen. Previous studies using such structure–activity approaches have given encouraging results[25].

We identified structural features shared by DEET and other known repellents and used them to screen a vast library of compounds *in silico* for the presence of these features. We assembled a training set of known repellents that included: the two commercially approved repellents DEET and picaridin; 34 *N*-acyl piperidines[25] that were identified by structural relatedness to picaridin; natural repellents eucalyptol, linalool, alpha-thujone and beta-thujone[10,26,27]; and a structurally diverse panel of other odours as negatives[28,29]. We focused on a descriptor-based computational approach and using a sequential-forward-selection method[30] we incrementally identified a unique subset of 18 descriptors that were highly correlated with repellency (correlation of 0.912) (Fig. 4a and Supplementary Table 1). The repellents clustered together if the optimized descriptor subset was used to calculate Euclidean distances amongst odorants of the training set (Fig. 4b).

The optimized descriptor set was used to train a support vector machine (SVM), which is a well-known supervised learning approach[31], to predict compounds that shared optimized structural features with known repellents (Fig. 4a). A fivefold cross-validation on the training set of repellents was performed and a mean receiver-operating-characteristic (ROC) analysis curve generated. The area under curve (AUC) was determined to be high (0.994), indicating that the *in-silico* approach was extremely effective at predicting repellents from compounds that were excluded from the training set (Fig. 4c).

We next used the 18-optimized-descriptor and SVM method to screen *in silico* a large virtual chemical library consisting of >440,000 volatile-like chemicals. Inspection of the top 1,000 predicted repellents (0.23% of hits) revealed a diverse group of chemicals that retain some structural features of the known repellents (Fig. 4d, e). We computed partition coefficient (log*P*) values of the 1,000 compounds to exclude those predicted to be lipophilic (log*P*>4.5) and therefore more likely to pass through the skin barrier in topical applications[32] (Fig. 4e). We also computed predicted vapour pressures of these chemicals, because volatility may be a useful predictor of spatial volume of repellency (Fig. 4e).

Although the *in silico* screen was feasible, a more significant challenge lies in identifying safe and effective DEET substitutes that can be rapidly approved for human use. To identify such compounds, we applied our *in silico* screen to an assembled natural odour library consisting of >3,000 chemicals identified as originating from plants, insects or vertebrate species, and compounds already approved for human use as fragrances, cosmetics or flavours (Supplementary Information). Although many of the top 200 hits share structural features with known repellents from the training set, they also represent structurally diverse chemicals, allowing targeted exploration of previously untested chemical space (Fig. 4f). For example, several anthranilates and pyrazines were identified, even though such compounds were largely missing from the training set.

Ir40a^{+} cells are activated by new repellents

We selected four compounds from the list: methyl *N*,*N*-dimethyl anthranilate (MDA), ethyl anthranilate (EA), butyl anthranilate (BA) and 2,3-dimethyl-5-isobutyl pyrazine (DIP), of which the first three have a mild grape-like aroma, excellent safety profiles and have been thoroughly tested and approved for human consumption or oral inhalation by the Food and Drug Administration (FDA), World Health Organization and European Food Safety Authority, and have been listed in the 'generally recognized as safe' (GRAS) list by the Flavour and Extract Manufacturer's Association (Fig. 4g and Supplementary Table 2). The fourth, a pyrazine, is an ant trail pheromone[33]. The anthranilate and pyrazine classes also contain a large diversity of chemicals found in nature and therefore present attractive repositories of structural substitutes.

For all four chemicals we found robust activation of sacculus ORNs (Fig. 5a, Supplementary Video 2) that innervate the *Ir40a*$^{+}$ 'column' glomerulus (Fig. 5b, as shown for BA). They also activated gustatory neurons that project to similar areas of the SOG as DEET (Fig. 5b, as shown for BA). GCaMP3 imaging in *Ir40a*$^{+}$ neurons showed robust responses to these chemicals, whereas several other classes of common

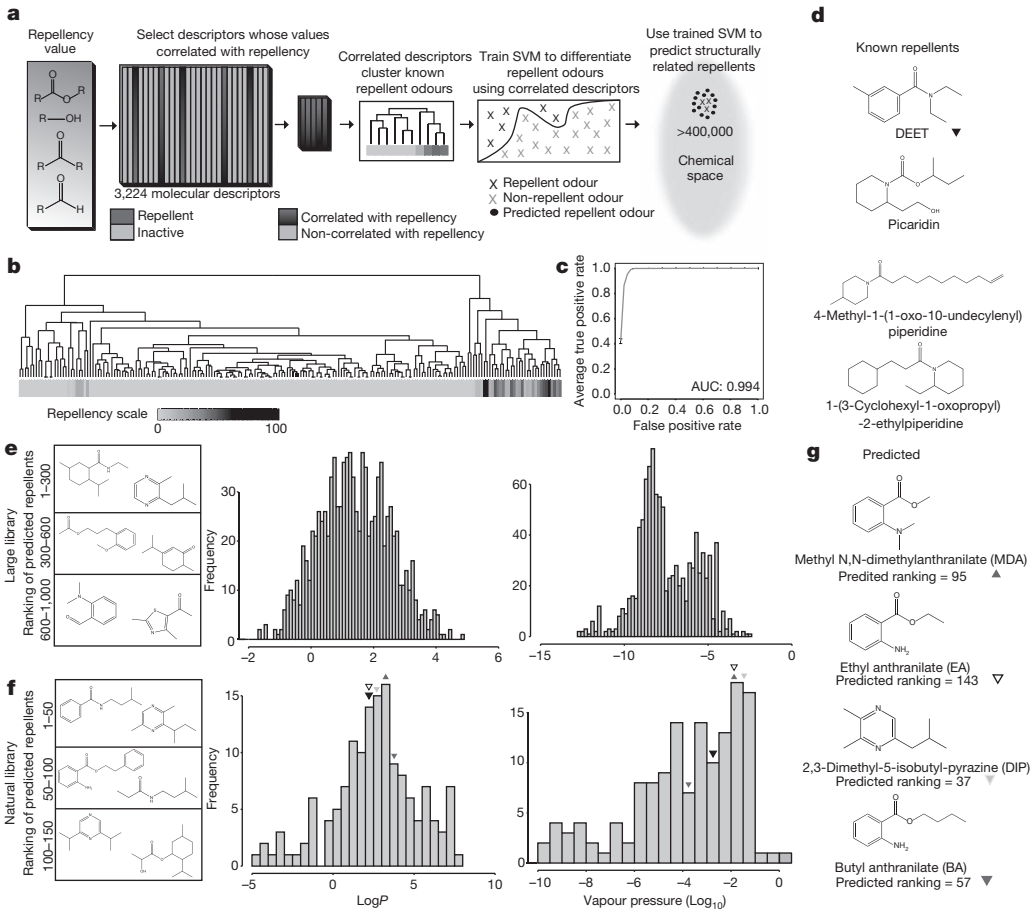


Figure 4 | Chemical informatics prediction of new repellents. **a**, Cheminformatics discovery pipeline to identify novel DEET-like repellents. **b**, Hierarchical cluster analysis of 201 training set odorants using optimized descriptors to calculate distances in chemical space. **c**, Receiver operating characteristic (ROC) curve representing computational validation of repellent predictive ability from 20 independent fivefold cross validations. AUC, area under the curve. **d**, DEET, picaridin and two unapproved repellents[25]. **e**, Representative predicted repellents from >400,000 odorant library (left) and computationally determined values for 1,000 top-ranked predicted repellents (right). **f**, Representative predicted repellents from >3,000 natural odour library (left) and computationally determined values for 150 top-ranked predicted repellents (right). Arrowheads indicate values for DEET and selected odours shown in **g**.

odorants did not (Fig. 5c and Supplementary Fig. 5). These results demonstrate that the computationally predicted chemicals activate the same chemosensory pathways as DEET and are therefore ideal candidates for new repellents.

In order to test the effect of these compounds on behaviour we used a two-choice trap assay in which flies can sense a DEET-treated filter paper positioned at the entrance of a trap through both olfactory and gustatory systems[3,17] (Fig. 5d). All four compounds had strong dose-dependent repellent effects on *D. melanogaster* (Fig. 5d). Measurements were taken at 24 h and 48 h after the start of the assay, and were found to be consistent. Six additional predicted repellents were tested in a similar manner, at least four of which elicited strong repellency similar to DEET (Supplementary Fig. 6).

To confirm the role of *Ir40a*$^+$ neurons in mediating avoidance to these new repellents, we examined behavioural avoidance of flies in which synaptic activity of *Ir40a*+ neurons was silenced using TNTG as before. We found that avoidance of chemical treated traps was substantially decreased in *Ir40a-TNTG* flies as compared to control flies (Fig. 5d), showing that *Ir40a*$^+$ neurons are required for repellency to the four chemicals.

Mosquitoes avoid predicted repellents

To test the effects of the identified chemicals on mosquito behaviour, we adapted an arm-in-cage assay that allows quantitative analysis of chemical repellency on mosquitoes attracted to a human arm (described in Methods) (Fig. 6a, Supplementary Fig. 7). Female *A. aegypti* mosquitoes showed strong avoidance behaviour to DEET, irrespective of whether or not they could directly contact DEET (Fig. 6b). However, for sporadic landings the average time spent on the net before escape although not significant ($P = 0.203$ for 10% DEET and $P = 0.06$ for 1% DEET, Student's *t*-test) was reduced when direct contact with DEET was permitted, particularly at the lower concentration (Fig. 6c). Although it is difficult to assess from these experiments the direct contribution of the gustatory system alone, it demonstrates that mosquitoes can avoid DEET strongly at close range, even without making direct contact with it.

In order to test whether the four newly identified *Drosophila* repellents were also olfactory repellents to mosquitoes, we performed behaviour trials using the non-contact version of the assay. Notably, we found that all four compounds applied at 10% concentration demonstrated substantial repellency (Fig. 6d). The fraction of mosquitoes present on the net throughout the duration of the assay (Fig. 6d), as well as the cumulative number of mosquitoes present on the net were substantially decreased in the presence of the test compounds (Fig. 6e). For the mosquitoes that did land on the repellent treatment, the escape index, as measured by the frequency of take-off, was substantially higher as compared to those landing on controls (Supplementary Figs 8 and 9).

One of the major disadvantages of DEET is its property of solubilizing plastics and synthetic materials[1], which affects its usefulness. We tested the ability of the four repellents to dissolve a 3 × 3 mm square of vinyl. While the vinyl completely disappeared in DEET within 6 h, there was no significant difference in the weight of the vinyl squares immersed in the four DEET substitutes after 6 h or 30 h (Fig. 6f).

Discussion

The unbiased strategy to use a genetic-reporter of neural activity was instrumental in identifying DEET-sensitive *Ir40a*$^+$ neurons. These reside in the pit-like sacculus that could protect neurons from harsh chemicals. Both olfactory and gustatory systems are activated by DEET, with additional modes of detection in the antenna being mediated by *orco*[11] and a yet to be identified tuning *Or* gene (Fig. 6h). Additionally, DEET has been reported to have a mild enhancing or suppressing effect on the activity of various *Or*-expressing

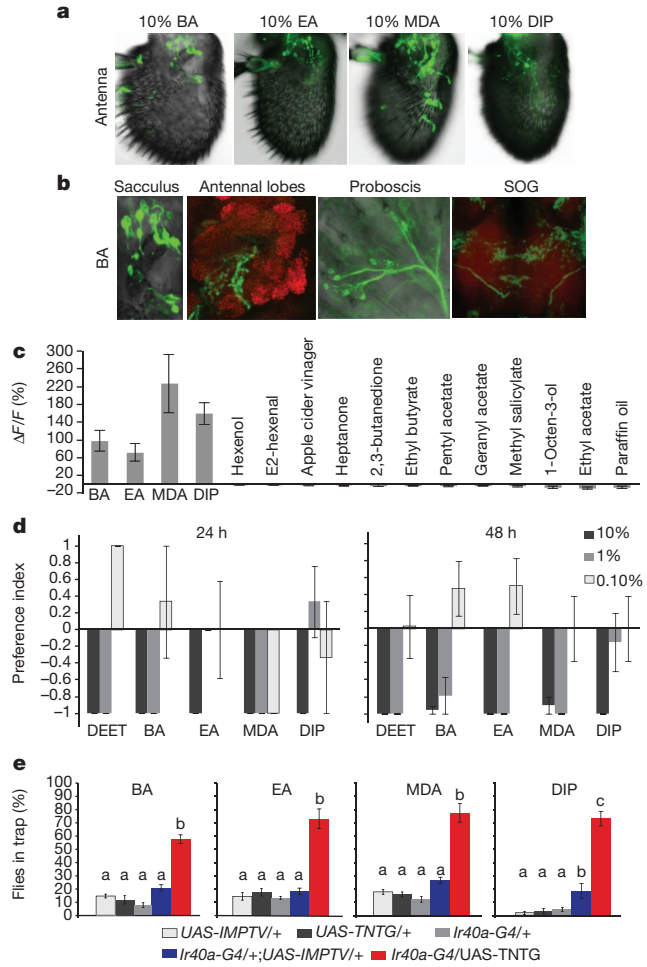
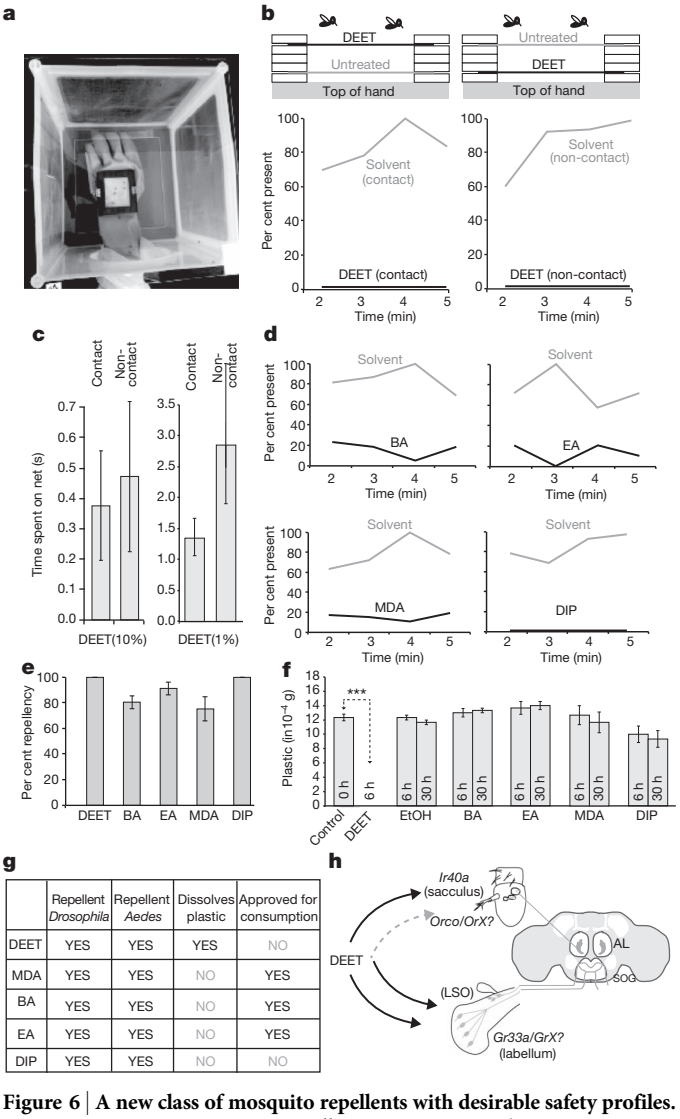


Figure 5 | Predicted repellents activate Ir40a neurons and are strong repellents for *Drosophila*. **a**, Images of antenna of *elav-Gal4/ LexAop-CD8-GFP-2A-CD8-GFP; UAS-mLexA-VP16- NFAT, LexAop-CD2-GFP/+* flies exposed to indicated stimuli for 24 h. **b**, BA-activated GFP^+ neurons in indicated tissues. **c**, Mean changes in fluorescence intensity in *Ir40a-Gal4/+;UAS-GCaMP3/+* cells after ~2-s application of indicated odorants. $n = 9–17$. **d**, Mean responses of flies to predicted repellents in two-choice olfactory and gustatory trap assays measured at 24 h and 48 h. $n = 3–10$ trials (24 h) and 7–10 (48 h); 10 flies per trial, trials with <40% participation were excluded. **e**, Quantification of flies of indicated genotypes entering repellent-treated traps. $n = 6$ trials for each genotype, ~20 flies for each trial. $P < 0.001$, one-way ANOVA with Tukey's post hoc test. For **c**–**e**, error bars represent s.e.m.



	Repellent *Drosophila*	Repellent *Aedes*	Dissolves plastic	Approved for consumption
DEET	YES	YES	YES	NO
MDA	YES	YES	NO	YES
BA	YES	YES	NO	YES
EA	YES	YES	NO	YES
DIP	YES	YES	NO	NO

Figure 6 | A new class of mosquito repellents with desirable safety profiles. **a**, Arm-in-cage assay to measure repellency in mosquitoes. **b**, Mean percentage of female *A. aegypti* present for >5 s on top net at indicated times to 10% DEET (black line) or solvent controls performed separately (grey line) in a contact (left) or non-contact (right) assay. **c**, Average time on net for each landing event in **b**. **d**, Mean percentage of female *A. aegypti* present for >5 s on top net in non-contact assay at indicated times. **e**, Cumulative repellency summed across minutes 2–5 of indicated non-contact treatment (10%) in comparison to appropriate solvent control. Forty mosquitoes were used per trial, $n = 5$ trials per treatment for **b**–**e**. **f**, Mean weight of vinyl pieces following submersion in indicated compounds or ethanol (control) for indicated amount of time. $n = 3$, $^{***}P < 10^{-5}$, Student's *t*-test. Error bars represent s.e.m. **g**, Properties of new repellents. **h**, Model for DEET detection and processing in *Drosophila*.

neurons of antennal basiconics in *Drosophila*, although a causal relationship between this effect and repellency has not been established[9]. DEET also has a solvent effect that slows down volatile odour release, potentially also from skin[10]. Thus, several pathways and mechanisms are likely to participate in overall repellency.

Ir40a can account for the widespread effect of DEET olfactory repellency because it is highly conserved in species that show strong avoidance to it including *Drosophila*, mosquitoes, head lice[34] and tribolium[35], but not in the honey bee[36]. Ir40a orthologues are conserved across many insects, with several regions of amino acid similarity across the length of the protein (Supplementary Fig. 10). This degree of conservation may better explain the repellent effects of DEET across several insect species compared to *Or* pathways that are not as well conserved. The *Ir40a* pathway therefore has important implications in the development of safe and affordable strategies to control several types of insects and arthropods that are disease vectors of animals and plants or are plant pests.

The chemical informatics enabled us to identify a number of affordable and safe potential repellents that are good candidates for regulatory approval for human use (Fig. 6g). This screen identified ~1,000 compounds and >100 additional natural compounds, many approved for use in human food and cosmetics, which may lead to other effective repellents. The repellency strategy may also have promise for use in combination with other behaviour control strategies, such as masking of CO_2-mediated attraction behaviour or population control by trapping as a part of an integrated pull–mask–push strategy[37,38]. Moreover, these DEET substitutes may be of value in controlling DEET-resistant strains as well. Because several of the new repellents are affordable, activate both the olfactory and bitter gustatory neurons, are approved for human consumption and are strong repellents for fruit flies, they may also have important implications for control of agricultural pest insects that cause enormous crop loss. Novel repellents that are safe and affordable can be used

to limit insect-human contact in disease-endemic areas of the world and to provide an important line of defence against deadly vector-borne diseases.

METHODS SUMMARY

Physiological experiments. NFAT-based neural tracing[16] and GCaMP3-based calcium imaging[22,23] were performed as previously described with some modifications (see Methods). Single-unit recordings from olfactory sensilla were performed as described previously[37].

Behavioural experiments. For olfactory trap assays, 20 *Drosophila* were released in cylindrical arenas containing Eppendorf tube traps (Figs 2d and 3a) with 10% apple cider vinegar as a lure. Repellents were presented on filter papers placed near the trap openings in a manner that did not allow physical contact with the fly before its entering the trap. Trap assays to measure repellency when both olfactory and gustatory inputs were possible were performed as described previously[3]. Mosquito arm-in-cage avoidance assays were performed with 40 mated *A. aegypti* females held in a cage and presented a human arm that was inserted in a glove containing a window covered with a double-layer of netting. Test compounds were applied to the nettings. Attraction towards the arm was measured using video recordings and analysts were blind to treatments.

Chemical informatics. Optimized molecular descriptors were selected from 3,224 Dragon descriptors based on their ability to increase the correlation between descriptor values and repellency. The repellency-optimized descriptor set was used to first train a support vector machine to predict repellents and then applied to predict new repellents from large compound libraries.

Insects. Fly lines were obtained from the Bloomington *Drosophila* Stock Center for TNT and GCaMP3 experiments, the Vienna *Drosophila* RNAi Center for *UAS-Ir40a* RNAi, J. Wang (UC San Diego) for NFAT tracing, and R. Benton (University of Lausanne) for *Ir40a*-Gal4. Flies were grown on standard cornmeal-dextrose media, at 25 °C unless otherwise noted and mosquitoes at 27 °C and 70% RH.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 January; accepted 15 August 2013.

Published online 2 October 2013.

1. Krajick, K. Medical entomology–Keeping the bugs at bay. *Science* **313,** 36–38 (2006).
2. Corbel, V. *et al.* Evidence for inhibition of cholinesterases in insect and mammalian nervous systems by the insect repellent deet. *BMC Biol.* **7,** 47 (2009).
3. Reeder, N. L., Ganz, P. J., Carlson, J. R. & Saunders, C. W. Isolation of a DEET-insensitive mutant of *Drosophila melanogaster* (Diptera: Drosophilidae). *J. Econ. Entomol.* **94,** 1584–1588 (2001).
4. Klun, J. A. *et al.* Comparative resistance of *Anopheles albimanus* and *Aedes aegypti* to *N,N*-diethyl-3-methylbenzamide (Deet) and 2-methylpiperidinyl-3-cyclohexen-1-carboxamide (AI3–37220) in laboratory human-volunteer repellent assays. *J. Med. Entomol.* **41,** 418–422 (2004).
5. Stanczyk, N. M., Brookfield, J. F., Ignell, R., Logan, J. G. & Field, L. M. Behavioral insensitivity to DEET in *Aedes aegypti* is a genetically determined trait residing in changes in sensillum function. *Proc. Natl Acad. Sci. USA* **107,** 8575–8580 (2010).
6. Gupta, R. K. & Bhattacharjee A. K. in *Insect Repellents: Principles, Methods, and Uses* (eds M. Debboun, Frances, S.P. & Strickman, D.) 195–228 (Taylor & Francis Group, 2007).
7. Ditzen, M., Pellegrino, M. & Vosshall, L. B. Insect odorant receptors are molecular targets of the insect repellent DEET. *Science* **319,** 1838–1842 (2008).
8. Bohbot, J. D. & Dickens, J. C. Odorant receptor modulation: ternary paradigm for mode of action of insect repellents. *Neuropharmacology* **62,** 2086–2095 (2012).
9. Pellegrino, M., Steinbach, N., Stensmyr, M. C., Hansson, B. S. & Vosshall, L. B. A natural polymorphism alters odour and DEET sensitivity in an insect odorant receptor. *Nature* **478,** 511–514 (2011).
10. Syed, Z. & Leal, W. S. Mosquitoes smell and avoid the insect repellent DEET. *Proc. Natl Acad. Sci. USA* **105,** 13598–13603 (2008).
11. DeGennaro, M. *et al. orco* mutant mosquitoes lose strong preference for humans and are not repelled by volatile DEET. *Nature* **498,** 487–491 (2013).
12. Xia, Y. *et al.* The molecular and cellular basis of olfactory-driven behavior in *Anopheles gambiae* larvae. *Proc. Natl Acad. Sci. USA* **105,** 6433–6438 (2008).
13. Liu, C. *et al.* Distinct olfactory signaling mechanisms in the malaria vector mosquito *Anopheles gambiae*. *PLoS Biol.* **8** (2010).
14. Lee, Y., Kim, S. H. & Montell, C. Avoiding DEET through insect gustatory receptors. *Neuron* **67,** 555–561 (2010).
15. Weiss, L. A., Dahanukar, A., Kwon, J. Y., Banerjee, D. & Carlson, J. R. The molecular and cellular basis of bitter taste in *Drosophila*. *Neuron* **69,** 258–272 (2011).
16. Masuyama, K., Zhang, Y., Rao, Y. & Wang, J. W. Mapping neural circuits with activity-dependent nuclear import of a transcription factor. *J. Neurogenet.* **26,** 89–102 (2012).
17. Syed, Z., Pelletier, J., Flounders, E., Chitolina, R. F. & Leal, W. S. Generic insect repellent detector from the fruit fly *Drosophila melanogaster*. *PLoS ONE* **6,** e17705 (2011).
18. Silbering, A. F. *et al.* Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems. *J. Neurosci.* **31,** 13357–13375 (2011).
19. Ai, M. *et al.* Acid sensing by the *Drosophila* olfactory system. *Nature* **468,** 691–695 (2010).
20. Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136,** 149–162 (2009).
21. Abuin, L. *et al.* Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* **69,** 44–60 (2011).
22. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6,** 875–881 (2009).
23. Pelz, D., Roeske, T., Syed, Z., de Bruyne, M. & Galizia, C. G. The molecular receptive range of an olfactory receptor *in vivo* (*Drosophila melanogaster* Or22a). *J. Neurobiol.* **66,** 1544–1563 (2006).
24. Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. & O'Kane, C. J. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14,** 341–351 (1995).
25. Katritzky, A. R. *et al.* Synthesis and bioassay of improved mosquito repellents predicted from chemical structure. *Proc. Natl Acad. Sci. USA* **105,** 7359–7364 (2008).
26. Klocke, J. A., Darlington, M. V. & Balandrin, M. F. 1,8-Cineole (Eucalyptol), a mosquito feeding and ovipositional repellent from volatileoil of *Hemizonia fitchii* (Asteraceae). *J. Chem. Ecol.* **13,** 2131–2141 (1987).
27. Kline, D. L., Bernier, U. R., Posey, K. H. & Barnard, D. R. Olfactometric evaluation of spatial repellents for *Aedes aegypti*. *J. Med. Entomol.* **40,** 463–467 (2003).
28. Carey, A. F., Wang, G., Su, C. Y., Zwiebel, L. J. & Carlson, J. R. Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* **464,** 66–71 (2010).
29. Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125,** 143–160 (2006).
30. Haddad, R. *et al.* A metric for odorant comparison. *Nature Methods* **5,** 425–429 (2008).
31. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20,** 273–297 (1995).
32. Walker, J. D., Rodford, R. & Patlewicz, G. Quantitative structure–activity relationships for predicting percutaneous absorption rates. *Environ. Toxicol. Chem.* **22,** 1870–1884 (2003).
33. Tentschert, J., Bestmann, H. J., Holldobler, B. & Heinze, J. 2,3-dimethyl-5-(2-methylpropyl)pyrazine, a trail pheromone component of *Eutetramorium mocquerysi* Emery (1899) (Hymenoptera: Formicidae). *Naturwissenschaften* **87,** 377–380 (2000).
34. Mumcuoglu, K. Y., Galun, R., Bach, U., Miller, J. & Magdassi, S. Repellency of essential oils and their components to the human body louse, *Pediculus humanus humanus*. *Entomol. Exp. Appl.* **78,** 309–314 (1996).
35. Hou, X. W., Fields, P. & Taylor, W. The effect of repellents on penetration into packaging by stored-product insects. *J. Stored Prod. Res.* **40,** 47–54 (2004).
36. Abramson, C. I. *et al.* Proboscis conditioning experiments with honeybees, *Apis mellifera caucasica*, with butyric acid and DEET mixture as conditioned and unconditioned stimuli. *J. Insect Sci.* **10,** 1–17 (2010).
37. Turner, S. L. & Ray, A. Modification of CO_2 avoidance behaviour in *Drosophila* by inhibitory odorants. *Nature* **461,** 277–281 (2009).
38. Turner, S. L. *et al.* Ultra-prolonged activation of CO_2-sensing neurons disorients mosquitoes. *Nature* **474,** 87–91 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. Ganguly and D. Carter for help with calcium imaging; Z. Wisotsky for help with gustatory experiments; D. MacWilliam for help with olfactory experiments; J. Wang for sharing the NFAT transgenic fly line; and R. Benton for sharing the *Ir40a-Gal4* fly line. This work was partly funded by a Whitehall Foundation grant to A.D., an R21NS074332 (NINDS) to A.D. and A.R., and an R56AI099778 (NIAID) and R01AI087785 (NIAID) to A.R. The granting agencies had no role in experimental design or analysis.

Author Contributions S.M.B. planned and performed the chemical informatics and solubility experiment, and helped design the behaviour experiments. P.K. planned and performed the NFAT imaging, Ca^{2+} imaging and *Drosophila* behaviour experiments. S.K.T. performed Ca^{2+} imaging, electrophysiology and some behaviour experiments. T.G. performed the arm-in-cage experiments. C.P. performed behaviour analysis. S.M.B., P.K. and S.K.T. helped prepare drafts of the manuscript and figures. A.D. planned experiments and helped write the manuscript. A.R. planned experiments, managed the project, and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.R. (anand.ray@ucr.edu).

METHODS

Fly stocks. Wild type flies were w^{1118} backcrossed to Canton-S for 5 generations. *UAS-GCaMP3* (BL#32236), *UAS-TNTG* (BL#28838), *UAS-IMPTV* (BL#28840) and *Tub-PGal80^{ts}* (BL#7017) were obtained from the Bloomington *Drosophila* Stock Center. The following stocks were generously provided: *LexAop-CD8-GFP-2A-CD8-GFP*; *UAS-mLexA-VP16-NFAT, LexAop-CD2-GFP* by J. Wang (UC San Diego, CA), *Ir40a-Gal4* by R. Benton (University of Lausanne, Switzerland), and *elav-Gal4* by L. Luo (Stanford, CA). *UAS-Ir40a* RNAi (line 1) (v101725) and *UAS-Ir40a* RNAi (line 2) (v3960) lines were obtained from the Vienna *Drosophila* RNAi Center. *Ir40a* RNAi is predicted to have no off-targets. Fly stocks were grown on standard cornmeal-dextrose media at 25 °C unless otherwise noted. Flies of appropriate genotypes for behaviour experiments were randomly sorted from populations before performing behavioural or electrophysiological experiments.

NFAT-based neural tracing. Late dark *Drosophila* pupae ready to emerge of genotype *elav-Gal4/ LexAop-CD8-GFP-2A-CD8- GFP; UAS-mLexA-VP16-NFAT, LexAop-CD2-GFP/+* [16] were collected on moist filter paper strips in culture vials which contained 2 Kimwipes soaked in 5 ml of water in a relatively odour-free environment. A 100 µl sample of odour at the indicated concentration was dissolved in acetone, spread on a filter strip (~1 cm × 3 cm), dried for 1 min and placed in a vial with 10–15 pupae. The exposure was given for 24 h and the filter paper strip with odour was replaced at ~12–14 h with fresh odour.

Calcium imaging using GCaMP3. DEET, DMSO, hexane and candidate compounds were purchased from Sigma-Aldrich or the eMolecules database (http://www.emolecules.com) from Enamine, Vitas M Labs or Chembridge and were of the highest purity available. Approximately 10–12-day-old flies raised at 29 °C (to improve Gal4 activity) were anaesthetized and secured by their wings on double-sided sticky tape (ventral side up) on a Petri dish (BD Falcon, 50 × 9 mm). The fly proboscis, head and body were immobilized by sticky tape as shown (Supplementary Fig. 11). One antenna was stably held down using a glass electrode on thin layer of 70% glycerol that enhanced imaging of fluorescence. The antenna was orientated with the arista and sacculus pointing upwards accessible to odours. Odorants were delivered using 5 ml plastic syringes containing 2 Whatman filter paper strips (2 × 3 cm). A fine mist of DEET at indicated concentrations in DMSO was sprayed into the syringe using an atomizer. Fresh atomized odour syringes were prepared immediately before odour delivery. For DEET substitutes (BA, EA, MDA and DIP), a 100 µl of 50% dilution in DMSO was applied to the filter paper directly and for other odorants 100 µl of 10^{-2} solution in paraffin or water for apple cider vinegar (ACV) was applied directly on the filter paper. The odour puff (~2 s) was delivered using the syringe over the antenna manually. For imaging odour-evoked activity from the antenna using GCaMP3, a Leica SP5 inverted confocal microscope was used. A filter block with 488 nm excitation filter and 500–535 nm emission filter was used and images were acquired at 3.3 frames per second with a resolution of 330 × 330 pixels using a 10× objective. The settings were optimized to capture odour-induced responses of GCaMP3 with high spatial and temporal resolution while limiting reporter bleaching.

Data analysis for calcium imaging was performed using the Leica SP5 LAS AF software (in quantify mode) to obtain the heat map images and fluorescence intensity changes. The $\Delta F/F$ percentage was calculated separately for each selected cell body by taking the mean intensity value of all frames for 5 s before the odour puff (F_{pre}) and taking the mean intensity value of all frames for 5 s around the peak responses (F_{post}) after the end of the ~2 s of stimulus delivery period. Similarly, the mean intensity values were taken for a background area in the vicinity of the cells.

The $\Delta F/F$ percentage was calculated according to the formula:

$$\Delta F/F(\%) = \frac{(F_{post} - F_{background(post)}) - (F_{pre} - F_{background(pre)})}{F_{pre} - F_{background(pre)}} \times 100$$

Immunohistochemistry. After 24 h exposure to either odour or solvent (control), flies were anaesthetized on ice and the tissue dissected in chilled 1× PBS and fixed for 30 min in 4% PFA (0.3% Triton X-100) at room temperature. After washes with PBST (PBS with 0.3% Triton X-100) brains were blocked using PBST with 5% bovine serum albumin (BSA). Rabbit anti-GFP (1:1,000, Invitrogen) and anti-nc82 (1:10 Developmental Studies Hybridoma Bank) were used as primary antibodies and samples were incubated for 3 nights at 4 °C. Alexa Fluor 488 anti-rabbit immunoglobulin G (IgG) (Invitrogen; 1:200) and Alexa Fluor 546 anti-mouse IgG (Invitrogen; 1:200) were used as secondary antibodies, respectively, followed by overnight incubation at 4 °C. Images were acquired with a Zeiss or Leica SP5 confocal microscope and images processing was done using ImageJ and Photoshop software. Data analysis was performed offline, and the investigator was blind to the treatment while counting GFP$^+$ antennal neurons in the confocal micrographs.

Temperature sensitive Gal80^{ts} experiment. For the two-choice behaviour assay in Fig. 3 and Supplementary Fig. 4, flies (10 males and 10 females) with genotypes *Ir40a-Gal4/+; UAS-Ir40a* RNAi*(2)/Gal80^{ts}* were grown throughout at 18 °C (permissive temperature) where Gal80 is active and RNAi is off. Such flies were treated as control. In parallel, flies of the same genotype were shifted to 29 °C (non-permissive temperature) from 18 °C as late black pupae for 4 days to activate Gal4 and switch on RNAi. These flies were used as knockdown flies. A subset of flies that were shifted to 29 °C was shifted back to 18 °C for 4 additional days to turn off the RNAi and these were used as recovery flies.

Electrophysiology. Flies used were 4–7-days-old and raised on cornmeal food at 25 °C. Extracellular recordings were made by inserting a glass electrode into the base of a palp sensillum as done previously[37,39]. Odorants were diluted in hexane or DMSO, at indicated concentrations (made fresh for every stimulus). For DEET stimulation, 10 µl of diluted odorant was applied to a filter paper strip, the hexane solvent was evaporated for 30 s (as in a previous study[17]) or for 5 min, and placed into a glass pasture pipette cartridge, and each cartridge was only used once. The evaporation of hexane from the filter paper strip was much slower upon mixing with DEET and lingering dampness of the filter paper could be observed visually as well.

Behavioural testing of *Drosophila* olfactory avoidance assay for DEET. For each trial, flies that were 3–6 days old (10 males and 10 females) were starved for 18 h.

For the trap assay, flies were transferred to a cylindrical 38.1 mm (diameter) × 84.1 mm (height) chamber containing a trap fashioned from an upturned 1.5 ml microcentrifuge tube with 2 mm removed from the tapered end. A pipette tip (1,000 µl) was cut 2.5 cm from the narrow end and 0.5 cm from top and inserted into the bottom of the inverted microcentrifuge tube. A 15 mm × 16 mm #1 Whatmann filter paper was inserted in between the pipette tip and tip of microcentrifuge tube so that entering flies could not make physical contact with it. A 25 µl sample of test compound was applied to filter paper and 125 µl of 10% ACV was applied to the upturned lid of the microcentrifuge tube as attractant. Trials were run for 24 h and the numbers of flies entering the trap counted (Fig. 2d).

In the two-choice test, two 10% ACV (125 µl) lured traps as described above were placed in the cylinder, one with 50 µl solvent (DMSO) and another with 50 µl the test odorant at 50% applied to the filter paper (Fig. 3). The more volatile DIP was tested at a lower concentration of 25%. For positive control tests in Supplementary Fig. 4, 125 µl of 10% ACV in test traps and 125 µl of water in control traps was added in the upturned microcentrifuge tube lid. Both traps contained filter papers as before with 50 µl solvent (DMSO). All trials were run for 24 h, positions randomized, and counted. Only trials with >35% participation were considered.

$$\text{Preference index} = \frac{\text{Number of flies in treated trap} - \text{number in control trap}}{\text{Number of flies in treated} + \text{control traps}} \quad (1)$$

***Drosophila* olfactory and gustatory avoidance assay for DEET.** Repellency was tested in Fig. 5d and Supplementary Fig. 6 using a *Drosophila melanogaster* two-choice trap assay as described previously[3,17] with minor modifications. Briefly, traps were made with two 1.5 ml microcentrifuge tubes (USA Scientific) and 20 ml pipette tips (USA Scientific), each cap contained standard cornmeal medium. A T-shaped piece of filter paper (Whatman #1) was impregnated with 5 µl of acetone (control) or 5 µl of 10%, 1% or 0.10% test odour, diluted in acetone. Traps were placed within a Petri dish (100 × 15 mm, Fisher) containing 10 ml of 1% agarose to provide moisture. Ten wild-type Canton-S flies 4–7-days-old were used per trial, which lasted 48 h, by which time point nearly all flies in the assays had made a choice. For the 24 h time point data were considered only if >35% of flies had made a choice; at 48 h the majority of flies had made choices. The preference index was calculated as in equation (1) above.

Mosquito arm-in-cage avoidance assay for DEET. Repellency was tested in mated and starved *A. aegypti* females using an arm-in-cage assay. *A. aegypti* mosquitoes (eggs obtained from Benzon Research) were maintained at ~27 °C and 70% relative humidity on 14 h:10 h light:dark cycle. Behavioural tests were done with 40 mated, non-blood fed, ~24 h starved, 4–10-day-old females in 30 cm × 30 cm × 30 cm cages with a glass top to allow for video recording (Fig. 6a, Supplementary Fig. 7). The experimental protocol was reviewed and approved by the Institutional Review Board (IRB) Compliance Analyst at UCR and determined not to require additional Human Research Review Board approval. Each test compound solution (500 µl) of 10% concentration in acetone solvent was applied evenly to a white rectangular 7 cm × 6 cm polyester netting (mesh size 26 × 22 holes per square inch) in a glass Petri dish and suspended in the air for 30 min to allow solvent evaporation. The more volatile 2,3-dimethyl-5-isobutyl

pyrazine was dissolved in paraffin oil. Acetone or paraffin oil (500 µl) served as control. A nitrile glove (Sol-vex) was modified as described in Supplementary Fig. 7 such that a 5.8 cm × 5 cm window was present for skin odour exposure. A set of magnetic window frames were designed to secure the treated net ~1.5 mm above skin, and a second untreated netting ~4.5 mm above the treated net in a manner so that mosquitoes were attracted to skin emanations in the open window but unable to contact treated nets with tarsi, or contact and pierce skin. Additionally the test compound had minimal contact with skin. A clean set of glove and magnets were used for every trial. Care was taken that the experimenter did not use cosmetics such as soap on arms. For each trial the arm was first inserted for 5 min and the number of mosquitoes landing or escaping test window recorded on video for a 5-min period. Solvent controls were always tested before a treatment. Mosquitoes showed robust attraction to a solvent treated arm when offered a second time after a gap of 5 mins providing a rigorous test for the treatments to be tested second. No cage was tested more than once within 1 h of a testing session and not more than twice on any single day. Videos were analysed blind and the numbers of mosquitoes present for a 5-s continuous duration were counted every minute. Mosquitoes reliably started accumulating in controls at the 2 min point and data from this time point were considered for analysis.

Percentage present was calculated as the average number of mosquitoes on the window for 5 s at a given time point across trials. All values were normalized to percentage of the highest value for the comparison, which was assigned a 100 per cent present.

Percentage repellency = (1−(mean cumulative number of mosquitoes on the window of treatment for 5 s at time points 2, 3, 4, 5 min / mean cumulative number of mosquitoes that remained on window of solvent treatment for 5 s at time points 2, 3, 4, 5 min)) × 100.

Escape index = (average number of mosquitoes in treatment that landed yet left the mesh during a five second window over the following time points: 2 min, 3 min, 4 min, 5 min) / (average number of mosquitoes that landed yet left the mesh during a 5 s window over the same time points in (treatment + control))

Each time point had $n = 5$ trials, 40 mosquitoes per trial, except for EA, in which $n = 4$.

Chemical Informatics. A single energy-minimized three-dimensional structure was predicted for each compound using the Omega2 software package[40]. The commercially available software package Dragon (3,224 individual descriptors) from Talete was used to calculate molecular descriptors[41]. Descriptor values were normalized across compounds to standard scores by subtracting the mean value for each descriptor type and dividing by the standard deviation. Molecular descriptors that did not show variation across compounds were removed.

For our analysis, compounds from different studies were approximated into a single metric of 'protection duration' as a rough indicator of repellency. The non-repellent diversifying training set of odours were assigned protection times of zero, whereas the approved repellents DEET and picaridin were assigned the highest value since we made the assumption that these would have structural properties important for regulatory approval. Compounds were clustered using Euclidean distance and hierarchical clustering based on differences in repellency values, and a set of 5 compounds with the highest activity that clustered together was classified as 'training repellents'.

A compound-by-compound repellency distance matrix was calculated from repellency data. A separate compound-by-compound descriptor distance matrix was calculated using the 3,224 descriptor values calculated by the Dragon software package. Using a sequential forward selection (SFS) approach, all descriptors are individually compared and selected for their ability to increase the correlation between descriptor values and repellency. The descriptor that correlates best is retained and each further iteration adds an additional descriptor to improve the correlation values. This process is continued until additional descriptors fail to improve the correlation value from the previous step. This process results in a unique descriptor set that is optimized for repellency.

This repellency-optimized descriptor set was used to train a support vector machine (SVM) using regression and a radial basis function kernel available in the R package e1071, which integrates libsvm[42,43]. Optimal gamma and cost values were determined using the tune.SVM function. The resulting trained SVM was then applied to predict activity for compounds from two libraries *in silico*, a natural compound library of ~3,200 volatiles and a > 440,000 compounds library.

For the natural compound library we assembled a subset of 3,197 volatile compounds from defined origins including plants, humans, insects[44], food flavours and a fragrance collection[45] including fruit and floral volatiles[46–53]. For the larger library we assembled a subset of >440,000 small molecules from the eMolecules database that have properties of volatile odourants. (Molecular weight <325 grams per mole and atoms: C, O, N, H, S.)

We performed a fivefold cross-validation by dividing the data set randomly into five equal sized partitions. Four of the partitions were applied to train the SVM and the remaining partition, which was not used for training, was used to test predictive ability. This process was repeated five times, each trial excluding a different subset of compounds as the training set and assigning the remainder as the test set. The whole process was repeated 20 times to improve consistency. A receiver operating characteristics (ROC) analysis was then used to analyse the performance of our computational repellency prediction. The overall predictive ability was calculated as a single receiver operating characteristic (ROC) curve for all 20 independent validations.

Calculation of Log*P* and vapour pressure values. SMILES structures of the predicted repellent odours were used with EPI Suite (http://www.epa.gov/oppt/exposure/pubs/episuite.htm) to calculate predicted Log*P* and vapour pressure values.

Vinyl solubility test. One 3 × 3 mm square of 4 gauge vinyl was submerged in 1 ml of each test compound in a glass container, stirred at a constant rate on a shaker and checked every 30 min until the vinyl square in DEET was completely dissolved (6 h). The vinyl pieces in each of the other compounds were removed, rinsed in ethanol and weighed. The process was repeated at 30 h (24 h after the vinyl square completely disappeared in DEET).

Statistical analyses. For behaviour experiments with preference index, arcsine-transformed data were analysed. Tests used are indicated in the figure legends and they are Student's *t*-test, one-way ANOVA and Tukey's post hoc analysis. Statistical tests for each experimental category and sample trails sizes were selected on the basis of previously published studies using similar assays, which are cited throughout the manuscript. For all graphs, error bars indicate s.e.m.

39. de Bruyne, M., Clyne, P. J. & Carlson, J. R. Odor coding in a model olfactory organ: the *Drosophila* maxillary palp. *J. Neurosci.* **19,** 4520–4532 (1999).
40. Boström, J., Greenwood, J. R. & Gottfries, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graph Model.* **21,** 449–462 (2003).
41. DRAGON software for Windows for molecular descriptor calculations v.5.5 (Talete, 2007).
42. Chang, C. & Lin, C. Libsvm: a library for support vector machines. (http://www.csie.ntu.edu.tw/~cjlin/libsvm) (2001).
43. Karatzoglou, A., Meyer, D. & Hornik, K. Support vector machines in R. *J. Stat. Softw.* http://www.jstatsoft.org/v15/i09 (2006).
44. El-Sayed, A. M. The Pherobase database of insect pheromones and semiochemicals http://www.pherobase.com/ (2009).
45. Flavors and fragrances 2007–2008 catalogue (Sigma-Aldrich, 2007).
46. Cork, A. & Park, K. C. Identification of electrophysiologically-active compounds for the malaria mosquito, *Anopheles gambiae*, in human sweat extracts. *Med. Vet. Entomol.* **10,** 269–276 (1996).
47. Curran, A. M., Rabin, S. I., Prada, P. A. & Furton, K. G. Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. *J. Chem. Ecol.* **31,** 1607–1619 (2005).
48. Gallagher, M. *et al.* Analyses of volatile organic compounds from human skin. *Br. J. Dermatol.* **159,** 780–791 (2008).
49. Knudsen, J. T., Eriksson, R., Gershenzon, J. & Ståhl, B. Diversity and distribution of floral scent. *Bot. Rev.* **72,** 1–120 (2006).
50. Logan, J. G. *et al.* Identification of human-derived volatile chemicals that interfere with attraction of *Aedes aegypti* mosquitoes. *J. Chem. Ecol.* **34,** 308–322 (2008).
51. Meijerink, J. *et al.* Identification of olfactory stimulants for *Anopheles gambiae* from human sweat samples. *J. Chem. Ecol.* **26,** 1367–1382 (2000).
52. Zeng, X. N. *et al.* Analysis of characteristic odors from human male axillae. *J. Chem. Ecol.* **17,** 1469–1492 (1991).
53. Zeng, X. N., Leyden, J. J., Spielman, A. I. & Preti, G. Analysis of characteristic human female axillary odors: qualitative comparison to males. *J. Chem. Ecol.* **22,** 237–257 (1996).