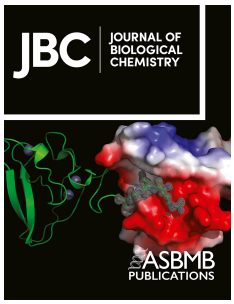
Journal Pre-proof



Natural volatiles preventing mosquito biting: an integrated screening platform for accelerated discovery of ORco antagonists

Georgia Kythreoti, Trias Thireou, Christos Karoussiotis, Zafiroula Georgoussi, Panagiota GV. Liggri, Dimitrios P. Papachristos, Antonios Michaelakis, Vasileios Karras, Spyros E. Zographos, Stefan Schulz, Kostas Iatrou

PII: S0021-9258(24)02441-4

DOI: https://doi.org/10.1016/j.jbc.2024.107939

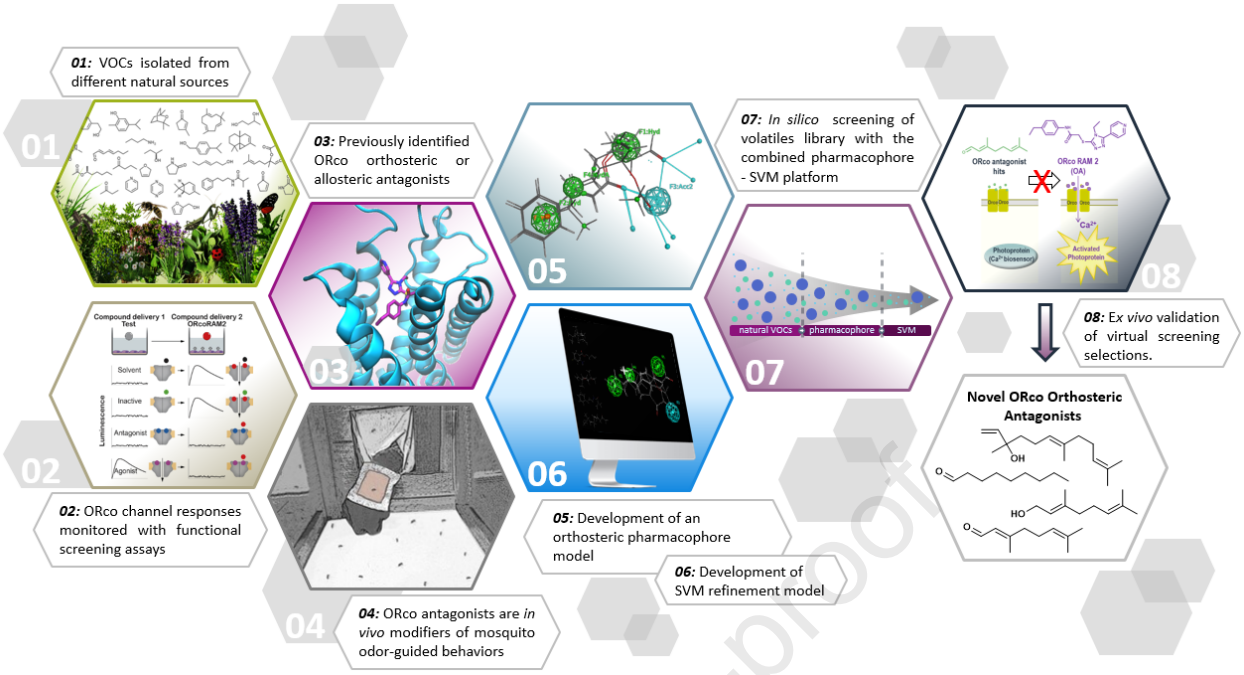
Reference: JBC 107939

To appear in: *Journal of Biological Chemistry*

Received Date: 9 July 2024

Revised Date: 16 October 2024

Please cite this article as: Kythreoti G, Thireou T, Karoussiotis C, Georgoussi Z, Liggri PG, Papachristos DP, Michaelakis A, Karras V, Zographos SE, Schulz S, Iatrou K, Natural volatiles preventing mosquito biting: an integrated screening platform for accelerated discovery of ORco antagonists, *Journal of Biological Chemistry* (2024), doi: https://doi.org/10.1016/j.jbc.2024.107939.



Natural volatiles preventing mosquito biting: an integrated screening platform for accelerated discovery of ORco antagonists

Georgia Kythreoti*[1b], Trias Thireou*[2], Christos Karoussiotis[1], Zafiroula Georgoussi[1], Panagiota GV Liggri[3,4], Dimitrios P Papachristos[4], Antonios Michaelakis[4], Vasileios Karras[4], Spyros E Zographos[3#], Stefan Schulz[5], Kostas Iatrou[1#]

[1]Institute of Biosciences and Applications, National Centre for Scientific Research "Demokritos", Athens, Greece; [2]Department of Biotechnology, Agricultural University of Athens, Athens, Greece; [3]Institute of Chemical Biology, National Hellenic Research Foundation, Athens, Greece; [4]Scientific Directorate of Entomology and Agricultural Zoology, Benaki Phytopathological Institute, 145 61 Kifissia, Greece; [5]Institute of Organic Chemistry, Technische Universität Braunschweig, Braunschweig, Germany;

[b]Current address: Department of Science and Mathematics, Deree – The American College of Greece, Athens, Greece.

* Authors with equal contribution

#Communicating authors: email iatrou@bio.demokritos.gr & sez@eie.gr

ABSTRACT

Insect olfactory receptors are heteromeric ligand-gated cation channels composed of an obligatory receptor subunit, ORco, and one of many variable subunits, ORx, in as yet undefined molar ratios. When expressed alone *ex vivo*, ORco forms homotetrameric channels gated by ORco-specific ligands acting as channel agonists. Using an insect cell-based system as a functional platform for expressing mosquito odorant receptors *ex vivo,* we identified small molecules of natural origin acting as specific ORco channel antagonists, orthosteric or allosteric relative to a postulated ORco agonist binding site, which cause severe inhibition of olfactory function in mosquitoes. In the present communication, we have compiled common structural features of such orthosteric antagonists and developed a ligand-based pharmacophore whose properties are deemed necessary for binding to the agonist binding site and causing inhibition of ORco's biological function. *In silico* screening of an available collection of natural volatile compounds with the pharmacophore resulted in identification of several ORco antagonist hits. Cell-based functional screening of the same compound collection resulted in the identification of several compounds acting as orthosteric and allosteric antagonists of ORco channel function *ex vivo* and inducing anosmic behaviors to *Aedes albopictus* mosquitoes *in vivo.* Comparison of the *in silico* screening results with those of the functional assays revealed that the pharmacophore predicted correctly 7 out of the 8 confirmed orthosteric antagonists and none of the allosteric ones. Because the pharmacophore screen produced additional hits that did not cause inhibition of the ORco channel function, we also generated a Support Vector Machine (SVM) model based on two descriptors of all pharmacophore hits. Training of the SVM on the *ex vivo* validated compound collection resulted in the selection of the confirmed orthosteric antagonists with a very low

cross-validation out-of-sample misclassification rate. Employment of the combined pharmacophore-SVM platform for *in silico* screening of a larger collection of olfaction-relevant volatiles produced several new hits. Functional validation of randomly selected hits and rejected compounds from this screen confirmed the power of this virtual screening platform as a convenient tool for accelerating the pace of discovery of novel vector control agents. To the best of our knowledge, this study is the first one that combines a pharmacophore with a SVM model for identification of AgamORco antagonists and specifically orthosteric ones.

INTRODUCTION

Many insect species have the potential to transmit a wide range of pathogens to humans and animals, causing a variety of vector-borne diseases (VBDs). According to the World Health Organization (WHO), VBDs account for more than 17% of all infectious diseases, causing more than 700,000 deaths annually (1). Therefore, they pose a significant threat to global public and animal health and have substantial socioeconomic impacts. Although effective control of insect disease vectors (IDVs) is crucial, it is also quite challenging. One powerful, effective and safe control method involves the use of long lasting and environmentally friendly repellents and anosmia-inducing agents. These agents interfere with the olfactory capacity of blood-feeding insects and reduce the frequency of their biting host organisms and transmitting pathogens to them.

Insects rely on their olfactory system to sense volatile chemicals that regulate various behaviors, including social interactions, mate and oviposition site selection, food source location and enemy recognition (2,3). Insect odor receptors, expressed in olfactory sensory neurons, are heteromeric ligand-gated cation channels. They are composed of one of many variable subunits, ORx, which confer specificity in the recognition of the odor molecules, and an obligatory receptor subunit, ORco, which is necessary for channel formation and signal transduction (4,5), in as yet undetermined molar ratios. In contrast to the variable ORx subunits, ORco is highly conserved amongst different insect orders, spanning many hundreds of millions of years of evolution (2,6,7). Moreover, we and others have shown that in the absence of a co-expressed ORx subunit, ORco can form *in vitro* homotetrameric cation channels (8,9) whose function may be activated or suppressed by synthetic ORco agonists and antagonists (10-15). Additionally, ORco antagonists have broad inhibitory activities on the majority of ORs of a variety of insects. Consequently, their binding site(s) on ORco may serve as "universal" modulatory site(s) for volatile compounds. Given such considerations, we set out to identify new ORco antagonists interrupting insect–host recognition and thus reducing and preventing the spread of VBDs.

Discovery of bioactive molecules through *in vivo* screening of large compound collections is an expensive and time-consuming process. The complexity of this process may be greatly reduced by the availability of appropriate *in vitro* or *ex vivo* functional assays and, even more so, by the undertaking of initial virtual screening (VS) steps that use the physico-chemical and

structural properties of compounds and/or target proteins to generate predictive models for identification of candidate bioactive molecules. Hence, VS methods narrow the search space and upon combination with experimentally verified bio-interaction information reduce the time and cost required for a screening project.

Several techniques are currently used for VS. Among them, the pharmacophore method and machine learning (ML) hold prominent roles. The pharmacophore is an ensemble of steric and electronic features that ensure optimal supramolecular interactions with a specific biological target structure that may lead to activation or blocking of its biological response (16). The simplicity and abstract nature of the pharmacophore concept enables the complexity of interactions between ligands and receptors to be reduced to a small set of features (17). Thus, pharmacophore-based techniques have become an integral part of computer-aided drug design and have been successfully applied for virtual screening, *de novo* design, and lead optimization (18). Pharmacophore models can be derived from experimentally determined protein-ligand complexes (receptor-based pharmacophores) or known active compounds (ligand-based pharmacophores). On the other hand, ML has established itself as a VS methodology in its own right and is constantly growing in popularity. Both conventional machine learning methods, such as Support Vector Machines (SVMs) and deep learning methods are used (19-21). A support vector machine (SVM) is a supervised learning algorithm with a growing number of applications in precision medicine and drug discovery (22,23). In a SVM binary classification problem, a high dimension decision surface is constructed (24,25). Several different kernels are introduced to map the data to the featured space, making SVMs able to handle various nonlinear problems with improved generalization characteristics.

In the present study, we are reporting on the development of a two-step VS protocol that achieves the goal of accelerating the discovery of new bioactive molecules that prevent mosquitoes from obtaining blood meals from their hosts by virtue of acting as antagonists of the ORco channel. In the first step, a pharmacophore model was constructed based on a set of small ligands that we have previously determined to function as specific ORco channel antagonists, orthosteric or allosteric relative to the ORcoRAM2 agonist binding site (14,26) and cause severe inhibition of olfactory function in mosquitoes (26,27). Sequentially, a SVM model was applied to refine the results and to better prioritize the compounds for experimental validation. The usefulness of the specific VS protocol is assessed by *ex vivo* assays using a previously developed cell-based functional platform (10-15).

RESULTS

Development of a ligand-based pharmacophore for accelerated discovery of ORco orthosteric antagonists

Our previous studies on a limited collection of 54 volatile organic compounds (VOCs) of natural origin have led to the identification of several ORco ligands, which acted as

antagonists of the homomerized ORco subunit (14,26,27). Some of the identified antagonists were also shown to possess powerful repellent activities for different mosquito species (26,27). Moreover, based on competition assays against a previously characterized ORco agonist, OrcoRAM2, which was predicted to bind to each ORco subunit of a homotetrameric ORco channel at a hypothesized site (8,9,28,29), the identified antagonists, shown in in **Table 1,** were classified as orthosteric or allosteric relative to the OrcoRAM2 binding site.

In order to identify compounds with putative ORco orthosteric antagonist activities in available VOC collections of natural or synthetic origin, we sought to develop a ligand-based pharmacophore that could describe orthosteric antagonist features necessary for blocking ORco's biological response. If successful, the specific pharmacophore could be employed as probe for an initial virtual screening of available compound collections prior to carrying out relevant functional screens.

i. Development of the pharmacophore model

Using as a training set the previously characterized collection of 54 VOCs, which included 4 positive examples (the orthosteric antagonists, shown in **Table 1)** and 50 negative ones (the 3 allosteric antagonists shown in **Table 1** and 47 inactive compounds shown in **Table S1**), a ligand-based pharmacophore has been developed that described the 3D arrangement of orthosteric antagonist features necessary for blocking ORco's biological response. The specific pharmacophore model has been required to match all orthosteric input molecules, while keeping the number of false positives (allosteric antagonists and inactive compounds) at a minimum. Four features (**Figure 1**) were found to meet these requirements best. These included one atom-centered hydrophobic feature “HydA”, two centroid hydrophobic features “Hyd” and one projected location of potential H-bond donors “Acc2”. Hydrogen bond Acc2 projected annotations are added to those heavy atoms that qualify as H-bond acceptors and are given Acc annotations (MOE 2016, Pharmacophore Annotation Schemes; see Experimental Procedures). The statistical significance of our model was estimated at -4.4626 (MOE 2016, The Pharmacophore Elucidator; see Experimental Procedures).

ii. Initial testing of the pharmacophore model - hit validation

The results of the training process of the specific pharmacophore model on the collection of 54 VOCs of **Table S1**, are shown in **Table 2**. For validation purposes, the previously reported *ex vivo* functional activities of all obtained hit compounds of **Table 2** were considered [26].

As may be seen in **Table 2**, screening of the training set with the selected pharmacophore resulted in the expected recognition of the four previously identified ORco orthosteric antagonists, compounds II, IV, 4 and 39 (ORco *ex vivo* inhibition of >40%; (26). In addition, however, the screening identified four more hits (compounds 33, 40, 42, 43), which either did not display any antagonist activities in our *ex vivo* activity screens (compound 42) or caused only minor inhibition of ORco activity, in the order of 15-20% (compounds 33, 40 and 43) (26). The remaining 46 compounds, including the three previously identified ORco allosteric antagonists shown in **Table 1**, were not selected by the pharmacophore.

iii. Virtual and functional screening of a new VOC collection

The specific pharmacophore model was also employed for an *in silico* screen of a new, previously "unseen" collection of 49 natural VOCs (**Table S2**). In this collection, the pharmacophore model uncovered the presence of the 24 hits shown in **Table 3.**

To evaluate the performance of the pharmacophore model in the *in silico* screen, the same VOC collection was functionally screened in parallel using the previously described cell-based platform for determining the % inhibition in ORco agonist activity *ex vivo*. The functional screen uncovered the presence of 12 active compounds in this collection of natural VOCs, which caused a substantial, equal or greater than 40%, degree of inhibition in the activity of the homomeric ORco channels. The results of the cell-based activity screen are shown in **Figure 2**. The structure of the 7 *ex vivo* active pharmacophore hits is shown superimposed onto the pharmacophre model in Supplementary **Figure S1**.

Finally, the bioactive VOCs were also subjected to competition tests against the ORco agonist OrcoRAM2 to deduce IC_{50} values and distinguish orthosteric from allosteric antagonists. The competition assays were carried out using as competitors three different concentrations the ORco agonist OrcoRAM2 (50, 100 and 150 µM). These assays, representatives of which are shown in **Figure 3,** provided the measure of inhibitory activities, in terms of IC_{50} values, for the confirmed antagonists, vis-à-vis the *ex vivo* ORco activity normally induced by the presence of 100 µM of OA. Secondly, they allowed the distinction between ORco allosteric and orthosteric antagonists relative to the ORco agonist binding site.

All *ex vivo* validated orthosteric and allosteric antagonists, 8 and 4, respectively, present in the new, virtually screened VOC collection, 8 and 4, respectively, together with their IC_{50} values, are listed in **Tables 3** (compounds in bold) and **4**, respectively.

Overall performance of the pharmacophore model

The 24 pharmacophore hits shown in **Table 3** included all but one (#74) of the eight orthosteric antagonists identified through the cell-based activity screening and competition assays presented in **Figures 2** and **3** (compounds #54, 60, 77, 83, 88, 98 and 99). None of the allosteric antagonists shown in **Table 4** were identified as antagonist hits by the pharmacophore. Thus, the sensitivity of the pharmacophore model for *in silico* prediction of actual orthosteric antagonists present in the specific collection of natural VOCs (**Table S3**) has been an impressive 0.88. However, the remaining 17 hits shown in **Table 3** were found to be either not active against ORco in the *ex vivo* assays or to inhibit Orco activity by substantially less than the previously defined useful inhibition cutoff point of 40% (**Figure 2**). Accordingly, the specificity of the pharmacophore screen has been 0.59 (see Experimental Procedures), a value that may be unsustainable in terms of experimental effort, especially for virtual screening of large libraries. Overall, the performance of the pharmacophore model described above was evaluated using the Power metric (PM) value (30), since this value might estimate better the performance of a virtual screening when few experiments can be carried out. The PM value for the pharmacophore VS was equal to 0.68, leaving room for improvement. For this reason, a second filtering step was added to the *in silico* screening pipeline.

Generation of a SVM model based on 2D descriptors for pharmacophore prediction filtering

To improve on the reliability of predictions for the identification of orthosteric antagonist hits, we trained several SVM models using the 32 pharmacophore hits shown in **Tables 2** (8 hits)

and **3** (24 hits). The set of 2D descriptors has been calculated in MOE (see Experimental Procedures). The descriptor pairs that resulted in the best SVM model included the KierA2 and SlogP_VSA1 structural features. KierA2 or second alpha modified shape index is a topological descriptor that encodes the branching of a molecule. In general, for straight chain molecules, KierA2=A-1 (where A is the atom count). SlogP_VSA1, on the other hand, describes the sum of the accessible van der Waals surface area for each atom whose logarithm of the octanol/water partition coefficient is in the range (-0.4 to -0.2] or, in other words, the extent of hydrophobic or hydrophilic effects on the surface area of the molecule. The SVM model with the lowest out-of-sample misclassification rate was selected and optimized, yielding a cross-validation loss equal to 0.032 (see below). The results of applying the selected SVM filters on the 32 pharmacophore hits are detailed in **Table 5** and shown diagrammatically in **Figure 4**.

The data points of the training classes together with the decision boundaries that separate them in the feature space are visualized in the classification map shown in **Figure 4**. The radial basis function (RBF) kernel handled the non-linearly separable data creating curved decision boundaries.

VOCs antagonizing ORco function act as spatial, mosquito anosmia-like inducing agents

The functionalities of the new *ex vivo*-validated ORco antagonists, orthosteric and allosteric ones, except that for #99 (**Table 3**) due to unavailability of sufficient quantity, were subsequently assessed *in vivo* against *Aedes albopictus* as previously described (26,31), at different concentrations ranging from a high of 200 to a low of 50 nmole/cm^2. At such concentrations, all *ex vivo* validated antagonists were found to cause *in vivo* inhibition in the numbers of mosquitoes that landed on the exposed hand areas to various extents (data not shown). Seeking potent repellents, compounds showing significant repellency (RI >50%) at the dose of 50 nmole/cm^2 were subsequently tested at an even lower dose of 10 nmole/cm^2. Thus, while compounds #39, #54, #77 and the allosteric antagonist #62 that exhibited mild repellent activity (RI 30%-50%; data not shown) were excluded from further testing, 7 new antagonists displaying high activities in the preliminary *in vivo* tests, 4 orthosteric (#60, 83, 88 and 98) and 3 allosteric ones (#64, 66 and 75) were assessed at the low compound dose of 10 nmole/cm^2 (**Figure 5**).

As can be seen from the results presented in **Figure 5 and Table S4**, even at the very low dose of 10 nmole of compound per cm^2 of naked hand area, mosquitoes exposed to all but one (#88) tested ORco orthosteric antagonists, identified through the combined employment of *in silico* and *ex vivo* screening, were found to display noticeably reduced attraction responses to the human smell emissions. Of particular note has been the orthosteric antagonist #60 [(2E,4E)-Decadienal; RI=0.71± 0.05] and allosteric antagonist #66 (2-Methylquinoline; RI=0.93±0.01), which caused aversion to the hand emissions comparable to that of DEET (RI=0.84±0.01).

***In silico* screening by the combined pharmacophore/SVM model for discovery of additional ORco orthosteric antagonists**

To examine the combined power of the optimized 2-step *in silico* screening protocol, we virtually screened a new collection of 241 compounds, most of them olfaction-relevant volatiles [**Supplemental spreadsheet**; (32,33)] for the presence of additional orthosteric

antagonists of ORco. Initial application of the specific pharmacophore model on this VOC collection resulted in the identification of 100 hits (**Figure 6**), while subsequent application of the SVM filter excluded another 56 compounds. Thus, the two-step protocol predicted the presence of 44 putative orthosteric antagonists in this compound collection (**Figure 6**).

Subsequently, a set of 15 compounds comprised of 5 randomly selected *in silico* hits (putative orthosteric antagonists) and 10 randomly selected workflow-rejected compounds was selected for *ex vivo* functional testing. The mapping of the selected 15 compounds, relative to the established SVM and SlogP_VSA1 classification map boundaries, is shown diagrammatically in **Figure 7**.

Validation of the combined orhtosteric pharmacophore/SVM model

The results of the *ex vivo* functional testing for the 15 selected representatives, whose SVM mapping coordinates have been presented in **Figure 7**, are shown in **Figure 8**.

As is shown in **Figure 8**, 3 out of the 5 retained hits showed >40% inhibitory activities, whereas the remaining 2, nonanal and citral, displayed reduced activities bordering the arbitrary cutoff inhibition limit. On the other hand, 9 out of 10 workflow-excluded compounds exhibited no or low (≤40%) inhibitory activity. Of the 10 excluded compounds that were selected for validation, only hexanoic acid was found to have a marginal inhibitory activity in the *ex vivo* assay.

A summary of the overall structural properties and *ex vivo* functionality of the validated compounds selected from the combined pharmacophore/SVM *in silico* screening is shown in **Table 6**.

DISCUSSION

Demand for novel agents to control harmful insects

Many insect species, including mosquitoes, have the potential to transmit a wide variety of pathogens to humans and animals, leading to VBDs with substantial socioeconomic impacts. Vector control, mainly through the use of insecticides, has been the principal method of preventing vector-borne infectious diseases for over 100 years and remains highly effective when comprehensively applied and sustained. However, given the current climatic changes worldwide, which have brought substantial temperature increases in geographic regions with temperate climates, and the concurrent increases in movements of people due to easier travel conditions, such diseases are spreading at an alarming rate in countries where they were previously absent. Therefore, there is a growing demand for novel, long lasting and environmentally friendly means of control that include repellents and anosmia-inducing agents. Yet, the classical research methods for discovery of new protective agents against insect bites, particularly in a spatial context that does not involve direct application on human or animal skin surfaces, is a time consuming and expensive task that prevents the expedient development of novel control measures.

The insect ORco receptor as target for discovery of host-seeking disruptors

Progress in the rate of discovery of protective agents for humans and domestic animals against various insect disease vectors, particularly mosquitoes, has been achieved relatively

recently through the employment of *ex vivo* expression systems developed from cultured amphibian oocytes (35-37), and mammalian (36,38-42) or insect cell cultures (15,43-45) upon coupling to relevant bioactivity reporter assays. Further enhancement in the rate of discovery of relevant bioactive compounds has been achieved recently through the exploitation of the seminal discovery that upon *ex vivo* expression, the evolutionarily conserved, obligatory odor receptor co-receptor ORco forms homomeric cation channels (11,13,46) whose function may be activated by specific agonists such as VUAA1, OrcoRAM2 and derivatives (11,13,47) and inhibited by structurally-related antagonists (10,12,13,48). Equally important has been the demonstrated inhibition of general odorant receptor function *in vivo* as a consequence of specific mutations in the ORco subunit (39,49-55). These findings led to the notion that inhibition of the olfactory functions producing anosmia-like phenotypes on targeted insect species may also be achieved by the binding of volatile ORco antagonists, preferably of natural origin, to the ORco subunit in nearly all ORx/ORco heteromeric receptor complexes in live insects. This notion has been amply proven by the demonstration that the great majority of natural volatiles causing inhibition of ORco function *ex vivo*, are capable of inhibiting the olfactory functions in laboratory and field mosquito populations in a spatial context (26,27).

Accelerating the discovery of new ORco antagonists

The employment of some of the *ex vivo* expression-activity detection systems mentioned above in throughput formats has allowed a significant acceleration in the rate of discovery of relevant bioactive compounds by activity screening of small size compound collections. For the specific insect cell-based expression-activity testing system employed in the current study (26,33,43), we note that it may be used for the detection of compounds with both agonist and antagonist activities against ORco. Indeed, the presence of a few agonists amongst the compounds examined for activities against ORco has been detected (33) but to date, all detected agonists had activities lower than those of the available synthetic ones (VUAA1 or OrcoRAM2). Moreover, because of our current focus on the study ORco antagonists, the identified agonist hits have yet to be studied in detail. Given the small size and nature of the libraries screened in this study, *i.e.*, ~300 natural volatiles of relatively low molecular weights, averaging ~150 Da, it is not surprising that no potent ORco agonists were identified. In contrast, VUAA1, the non-volatile synthetic compound and only ORco agonist discovered by *ex vivo* screening, was identified amongst the members of a library of 118,000 small molecules (Vanderbilt Small Molecule Library), typically used in drug development (11,56). In contrast, the structure-related compounds in the OrcoRAM agonist (and antagonist) series (10,46) have been generated by rational chemical synthesis. Therefore, the likelihood of identifying a complex aromatic compound resembling the structural features and binding mode of VUAA1 from the libraries used in this study, was inherently low. Nevertheless, due to its versatility, our *ex vivo* assay is suitable for screening more diverse libraries, as opposed to only focused ones. In the future, such an approach may reveal a number of novel agonists including potent ones.

Despite the flexibility of the currently available protein expression-activity detection systems, when the requirements for screens of large compound collections are considered in terms of time and material costs, such systems are still not adequate by themselves for fast discovery of adequate numbers of new, highly active compounds.

A combined orthosteric pharmacophore/SVM model for optimized predictions

To expedite further the search for new bioactive molecules in large compound collections, computational screening methods could be applied as virtual pre-screening tools that might reduce the number of molecules to be functionally screened *ex vivo* to a reasonable level. Toward this goal, we have employed in this study a two-step, ligand-based *in silico* pipeline consisting of a first pharmacophore screening (Step-1) and a subsequent SVM filtering step (Step-2). This pipeline showed highly satisfactory performance in predicting active orthosteric antagonists for ORco, as confirmed by follow-up functional validation.

Pharmacophores are frequently used in virtual screening projects, due to their simplicity and their ability to speed up the *in silico* process (18,57,58). Moreover, since they do not depend on specific functional or structural groups, they can identify chemically divergent molecules. Creating successful pharmacophores depends on the generation of sufficient ligand conformers, so that the bioactive conformation is approximated (57), and on using an adequate number of inactive molecules, to balance the sensitivity and specificity of the models. Pharmacophores are commonly combined with other computational techniques such as Support Vector Machines (SVMs), to improve the accuracy of the results (59). Support Vector Machines are well established in bioinformatics and chemoinformatics, since they can handle high-dimensional data and small datasets, and they can model non-linear decision boundaries. They are also adaptable and versatile. Feature selection and hyper-parameters optimization are critical for SVM high performance (60,61). However, SVMs can also be computationally expensive for large datasets (62,63). For this reason, we have employed the SVM filtering after the pharmacophore screening step in our workflow.

Our virtual screening pipeline achieved 0.75 and 0.82 sensitivity and specificity, respectively, resulting in an overall performance of 0.8 for predicting orthosteric antagonists that caused more than 40% inhibition to ORco (**Figure 6**). Such a performance is notable because elimination of more than 80% of the number of compounds to be tested translates in commensurate time-, material- and labor-cost savings for *ex vivo* and *in vivo* tests. Thus, our pipeline can both save resources and accelerate the discovery of novel agents. Moreover, although similar virtual screening protocols for discovery of novel drugs with defined specificities have been reported recently (see below for discussion), to the best of our knowledge, our study is the first one that combines a pharmacophore with a SVM model for identification of AgamORco antagonists and specifically orthosteric ones that are advantageous for future site-specific, ORco structure-based screening as compared to blind-docking trials.

Our pharmacophore model (Step-1) resembles the model previously proposed by Bhattacharjee and colleagues (64). That model consisted of a hydrogen-bond acceptor site, two aliphatic and one aromatic hydrophobic site. It was successfully used for virtual screening of an in-house compound database that resulted in four new potential insect repellent candidates. Other studies on insect olfactory ligands (65), employed a Laplacian-corrected Naïve Bayesian machine learning, ligand-based, approach to predict novel volatile *Anopheles gambiae* ORco antagonists. Selected hit compounds were further evaluated for their ability to inhibit electrophysiological responses in adult *Drosophila melanogaster* flies and in behavioral attraction assays against D*. melanogaster* larvae. In contrast to our study, the model was not trained to discriminate between orthosteric and allosteric antagonists. Electroantennography (EAG) recordings of two selected hits, 2-tert-Butyl-6-methylphenol

(BMP) and Linalyl formate (LF) suggested an allosteric and non-competitive ORco-dependent mechanism, which was further confirmed by concentration-inhibition analysis of BMP in *Xenopus laevis* oocytes expressing AgamORco. Machine learning techniques such as RandomForest and kNN classifier have also been successfully employed to predict new receptor agonists other than ORco, *i.e.*, SlitOR24 and SlitOR25 from *Spodoptera littoralis* (66). A Supported Vector Machine (SVM) model, (such as Step-2 in our pipeline) has been used for identification of agonists for SlitOR25 (67).

Structural features-activity relationships

While our approach is characterized by high performance, as with any other prediction method, it could not be 100 percent accurate. For example, hexanoic acid that has been rejected by our workflow at the pharmacophore selection step, showed antagonist activity (**Figure 7**) and has thus been considered as a false negative compound. A meta-analysis of the structure-activity relationship of the hits listed in **Table 6**, has revealed that hexanoic has the smaller length (6 carbon atoms) among the linear hits. In its most extended conformation, the distance between the two centroid hydrophobic features Hyd (carbon atoms) is 6.4 Å, which does not conform with the pharmacophore model shown in **Figure 1**, where the optimum Hyd1-Hyd2 distance has been determined to be 7.2 Å. Given that the initial set of orthosteric antagonists (**Table1**) as well as the set used for pharmacophore training (**Tables 2 and S2**) are dominated by longer chain linear compounds (8 to 10 carbon atoms), that can obtain conformations satisfying pharmacophore distances as well as bulky cyclic and aromatic compounds, it is possible that the pharmacophore model is negatively biased toward molecules of smaller length. Such inconsistencies of the model could be eliminated by incorporating more experimental data on short-length agonists. Furthermore, *ex vivo* concentration-inhibition analysis remains to be performed to exclude that hexanoic acid cannot act as an allosteric antagonist, *i.e.*, that it is a true negative result (as per terminology of our platform for the allosteric *ex vivo* active compounds). On the other hand, two compounds, nonanal and citral that have been retained by our workflow (**Figure 7**), showed borderline activities in the *ex vivo* experiments (**Figure 8**) and were classified as false positives. These two aldehydes can participate in only one hydrogen bond through their carbonyl group (hydrogen bond acceptor), in contrast with the other three active compounds in the series, which can participate in two hydrogen bonding interactions. In particular, nerolidol and geraniol bear a hydroxyl group that can act as a hydrogen bond acceptor/donor whereas linalyl acetate bears an acetate ester with two oxygen atoms in proximity that can act as hydrogen bond acceptors (**Table 6**).

Concerning the pipeline-rejected hits linalool and α-terpineol, both tertiary alcohols of molecular weights 154.25, with very similar cLogP values of 2.468 and 2.369, respectively, and identical polar surface area (PSA) 20.23 (SlogP_VSA1=0), despite the high variability of the *ex vivo* obtained response values, they are considered no- or low-activity inhibitors (**Figure 8**). Moreover, the rejected carvone and fenchone, can participate in only one hydrogen bond, while limonene lacks a functional group for participation in hydrogen bonds (**Table 6**). All three compounds are relatively compact cyclic molecules. The observed activities could therefore, upon further investigation, be a result of allosteric binding. Finally, the rejected geranyl acetone, similar to the retained nonanal and citral, can only participate in one hydrogen bond, while myrcene contains no polar functionalities. Geranyl acetate and Octanoic acid have similar KierA2 and SlogP_VSA1 parameters located well outside the

decision boundaries for active orthosteric antagonists (**Figure 7**). Hexanoic acid, which does not conform to the pharmacophore model (*vide supra*) but lies inside the decision boundaries, has identical SlogP_VSA1 to geranyl acetate and octanoic acid but different KierA2, due to the different spatial density of atoms in this shorter molecule. Similarly, despite their similar KierA2, the inactive geranyl acetate, with SlogP_VSA1=7.74, has both larger SPA (26.3) and more hydrophobic character (cLogP 3.264) than the active geraniol with SlogP_VSA1=0 (SPA=20.23 and cLogP=2.524).

We are noting that compound #88 [ethyl (E/Z)-2-(cyclohex-2-en-1-ylidene) acetate], the single orthosteric antagonist that caused minimal behavioral effects at the dose of 10 nmole/cm^2, was found to have the highest IC_{50} (195.7 µM) amongst its active counterparts in the *ex vivo* tests (**Tables 3** and **5**). Therefore, its low *in vivo* activity may be due to its low inhibitory potency against ORco or/and its relatively high calculated volatility (VP= 0.207 mmHg), that might affect its performance under the 5-min experimental timescale of the behavioral assays. Concerning its weak *ex vivo* binding to ORco and its low *in vivo* activity, it should be kept in mind that this compound has been tested as a mixture of E/Z isomers. It is very likely that ORco selectively binds one of the two isomers, as has been shown be the case for compounds binding to other olfactory receptors (68,69). In support of this notion, **Figure S2** showcases the explicit orientations of either the sp^2 or the sp^3 hybridization carbons towards the spheroid F4 (**Figure 1**). Among the two isomers, the Z is better fitting the specific pharmacophore model because the saturated carbons bearing two hydrogens orient to the larger spheroid F2, whilst the unsaturated (sp^2) carbon with its one hydrogen is oriented to spheroid F4 providing better occupancy. Moreover, the cyclohexene ring carbons holding a sp^3 hybridization are also bended, thus contributing to the model complementarity in this isomer. Hence, it is possible that the E isomer is a weak or a non-ORco binder, resulting in the apparent weak inhibitory and behavioral activity of the mixture. While the test of the individual isomers is beyond the scope of this study, the ORco specialization against multiple geometric, diastereomeric or enantiomeric isomers of an olfactory ligand is worth investigating in future studies. Such information can reveal the role of ORco on the remarkable selectivity of insect olfaction and be further exploited in ORco-based *in silico* and *ex vivo* screening approaches.

We also note that compound #74 (Bisabolene; also a mix of isomers) that was found to be active in the *ex vivo* screens (**Figure 2**), escaped detection by the pharmacophore (**Table 3**). Nevertheless, subsequent analyses showed it to have values placing it within the SVM boundaries (KierA2 = 5.4685, SlogP_VSA1 = 0) and also be marginally active in the *in vivo assays* at a dose of 50 nmole/cm^2 (data not shown). Accordingly, based on the results of the initial pharmacophore screen, we consider it to be a false negative result of our screening pipeline. Moreover, compounds #39 (2,4-Octadienal), #54 [(*Z*)-3-Nonen-1-ol], #74 (Bisabolene) and #77 (α-Bisabolol) and the allosteric antagonist #62 (α-Pinene oxide), which exhibited mild repellent activities (RI 30%-50%) at the same dose (data not shown), were not tested at the lower dose of 10nmole/cm^2. Future studies should aim to include a more comprehensive evaluation of all *ex vivo*-tested compounds to determine their minimum effective doses and thus provide a more complete understanding of their structure-activity relationships.

Combined pharmacophore-SVM approaches for drug discovery - advantages, limitations and future prospects

Our computational pipeline was successful in predicting the presence of at least two strong AgamORco orthosteric antagonists in the collection of 241 odorant compounds, Nerolidol and Geraniol and also confirming the presence of a third one, Linalyl acetate, that had been identified previously as such (26). These findings assert its validity as a screening tool for accelerating discovery of AgamORco orthosteric antagonists.

An advantage of the two-step ligand-based approach presented here is that it can select a diverse pool of predicted binders in a short time without the need of previous target-specific knowledge. Combining the pharmacophore with the SVM filter could facilitate the exploration of big data, improve the screening performance of virtual screening and help to study the mechanism of ligands biological activity. Pharmacophore models and SMVs have been previously combined in virtual screening pipelines to address other drug discovery problems. For example, Chen et al (70), used a Pharmacophore Ensemble/SVM approach that predicted the activation of human pregnane X receptor (hPXR) by 160 molecules of known activities (EC_{50} values). That study showed that the combined Pharmacophore/SVM model executed extremely well for the 32 molecules in the training set, 120 compounds in the test set and 8 additional ones in the outlier set, which were structurally distinct from those in the training set. Importantly, the combined model performed better than any of the individual pharmacophore models in the ensemble and was thus established as a powerful predictive tool to facilitate drug discovery. In a newer study, Cieślak et al (71) screened the ZINC database in search of monoamine oxidase (MAO) inhibitors that might function as potential antidepressants and agents for slowing down the progression of Parkinson's or Alzheimer's diseases. During the training step, ligands with known MAO enzyme inhibitory activity values were docked against the MAO-A and MAO-B subtype crystal structures and the best results were used to generate the pharmacophore hypotheses. Additionally, several ML models were trained in order to predict the docking scores. The five ML models with the best performance, including SVMs, were selected. Following training, ZINC database compounds fitting the pharmacophore hypothesis were ranked according to the consensus scoring of the ML models. Twenty-four of the top diverse ligands were then experimentally tested and low molecular weight week inhibitors were found. Thus, compared to docking-only VS, the combined protocol accelerated the discovery of potential MAO inhibitors.

The number of active and inactive compounds with known activity values is critical for the success of the VS process, as they influence the performance of both the pharmacophore hypotheses and the SVM algorithms. Moreover, activity values should preferentially originate from the same experimental procedure. The diversity of ligand structures is equally important for training unbiased models and employing scaffold hopping. Future, optimization of our computational pipeline by incorporation of more experimental data could significantly improve its performance. Moreover, as inferred above, should more experimental data on agonists become available, the same pipeline can be adapted to the discovery of such classes of compounds.

To conclude, any ligand-based approach is bound to exhibit some limitations. The shape and electrostatic potential of the ORco binding site and the conformation, hydrophobicity, polarity, and hydrogen bonding potential of the interacting amino acid residues are the

determining factors for discrimination of even subtle differences in physicochemical properties and active conformations between inactive compounds and physiologically relevant ligands. Given the recent availability of 3D-structures of ORco of the parasitic fig wasp *Apocrypta bakeri*, pea aphid *Acyrthosiphon pisum*, and the structural homologue MhOR5 from the jumping bristletail *Machilis hrabei* (8,9,72,73), reliable AgamORco homology models of apo- and liganded form can be created (28) and combined with our *in silico* ligand-based pipeline and *ex vivo* evaluation platform. To this end, our pipeline can constitute the first step for screening large chemical libraries and proposing candidates for subsequent site-specific molecular docking and MD simulations against AgamORco homology models. Such an approach is currently underway for seeking both novel active compounds and gaining structural insights on ligand recognition mechanism by AgamORco.

EXPERIMENTAL PROCEDURES

Pharmacophore model development. Based on the previously published orthosteric antagonists and inactive or allosteric compounds, several pharmacophore models were developed using Molecular Operating Environment software (MOE v. 2016.0801; Chemical Computing Group Inc., 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2016). The Unified annotation scheme was employed including H-bond Donors and Acceptors, as well as Hydrophobic Atoms and Hydrophobic Centroids. The radius of all features was set to 1Å, except for the radius of Hydrophobic Atoms, which was set to 0.7 Å. Query Spacing and Active Coverage were set to 0.9 and 1 correspondingly. Therefore, the generated pharmacophore models were required to match all orthosteric input molecules, while keeping the number of false positives to a minimum. The selected pharmacophore model was used to screen a collection of small molecules of natural origin to identify orthosteric ORco antagonists.

Sensitivity, Specificity and Power Metric (PM) (30) were used for the evaluation of virtual screening performance. They are defined as,

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$PM = \frac{TPR}{TPR + FPR}$$

where TPR= TP/(TP+FN) and FPR = FP/(FP+TN) are the true positive rate and the false positive rate correspondingly, TP the number of true positives, FN the number of false negatives, FP the number of false positives and TN the number of true negatives. PM ranges from 0 to +1. Values around 0.5 indicate poor to random models, whereas values between 0.9 and 1.0 are calculated for high quality models. PM is statistically robust with respect to the ratio of actives

to the total number of compounds and can be safely applied in early-recognition virtual screening problems.

To improve the performance of the pharmacophore model and to further understand the key features of orthosteric antagonists, we implemented the following procedure. Using MOE, we calculated all the 2D QuaSAR-Descriptors of the molecules identified by the selected pharmacophore model. For each combination of two calculated descriptors, we generated in MATLAB a support vector machine (SVM) with a Gaussian or radial basis function kernel, to classify the orthosteric antagonists from the rest of the molecules. The SVM with the lowest out-of-sample misclassification rate was subsequently optimized and the ten-fold cross-validation loss was reported.

Chemicals. Compounds analyzed in this study, VOCs and known repellents are presented in **Tables S1 and S2.** Carvacrol (CRV, **I**), linalyl acetate (LA, **4**), (2E,4E)-2,4-octadienal (OCT, **39**) and ethyl cinnamate (EC, **IV**) were purchased from Sigma Aldrich; isopropyl cinnamate (IPC, **II**) from Alfa Aesar; cumin alcohol (CA, **III**) from Acros Organics; N-(4-ethylphenyl)-2-[[4-ethyl-5-(3-pyridinyl)-4H-1,2,4-triazol-3-yl]thio}acetamide (ORco Receptor Agonist ORcoRAM2; OA) from Asinex Corporation and Vitas M Chemical Ltd; *N,N*-diethyl-3-methylbenzamide (DEET; **V**) from Sigma-Aldrich; and coelenterazine from Biosynth. All other VOCs were provided by the Institute of Organic Chemistry, Technische Universität Braunschweig, Germany. For the insect cell-based screening assay, initial stock solutions were prepared as needed and stored at −20 ^{o}C. The ORco agonist ORcoRAM2, stocks were prepared in dimethyl sulfoxide (DMSO) whereas the VOCs and coelenterazine stocks were prepared in ethanol. The assay was performed in modified Ringer's buffer (25 mM NaCl, 190 mM KCl, 3 mM CaCl2, 3 mM MgCl2, 20 mM Hepes, 22.5 mM glucose, pH 6.5; 35), so that the final concentration of DMSO used not to exceed the range of 0.2% to 0.35%.

Transformation of Bm5 cells for AgamORco and Photina expression and Ca^{2+} influx assays. An insect cell-based assay was employed as a screening platform for the identification and analysis of novel ORco ligands capable of modifying olfaction-mediated mosquito behaviors. Lepidopteran cultured cells (Bombyx mori Bm5; (74), constitutively expressing the AgamORco ligand-gated ion channel were employed, along with a reporter photoprotein Photina (75). Briefly, Bm5 cells were transformed to stably express cDNAs for AgamORco and Photina from high-expression-level pEIA plasmid vectors as previously described (44,76-78). Upon ligand binding activation of the ORco channel, Ca^{2+} ions entering the cells in turn activate the photoprotein, resulting in an increase in luminescence. Cell lines were grown in IPL-41 insect cell culture medium (Genaxxon Bioscience GmbH) supplemented with 10% fetal bovine serum (Biosera) in the presence of 10 µg/ml puromycin and maintained at 28 ^{o}C. The ligand binding to the ORco channel and subsequent functional effects were monitored via luminescence emission of the Ca^{2+} influx Photina biosensor, as previously reported (14,34). Specifically, insect cells resuspended in modified Ringer's buffer were seeded in white 96-well plates (200,000–300,000 cells/well), and incubated at room temperature in the dark with 5 µM coelenterazine. Luminescence emissions were then recorded in an Infinite M200 microplate reader (Tecan) at 4s intervals for up to 20s, using buffer and 1% Triton-X100 as baseline and maximum intensities, respectively. Tested compounds were initially added at a 100 µM final concentration and the ORco channel response was monitored for 10s at 4s intervals. Cells were allowed to return to baseline, allowing for the monitoring of the

secondary effect of ligand binding (4s intervals for 80s), resulting from the addition of 100 µM OA activating the ion channel. Luminescence data were acquired using i-Control 1.3 software by Tecan and normalized by considering ORco agonist luminescent response as the maximal (100%) receptor response for each experimental set. Independent experiments were run in triplicate and repeated at least three times.

Binding assays. ORco response inhibitions of identified antagonists were further analyzed to determine orthosteric or allosteric binding, relative to the OA (ORcoRAM2) binding site. Solvent or identified antagonists were added to insect cells, constitutively expressing AgamORco and Photina, at concentrations ranging from 1 µM to 1 mM. A 96-well format assay was also employed as described above, and the induced luminescence, if any, was measured. Subsequent addition of OA at different concentrations, 50, 100 or 150 µM were carried out as antagonist dose-dependent inhibition assays, illuminating the type of ligand binding on ORco. OriginPro 8 software, by OriginLab Corporation, was used for curve fitting and IC_{50} value calculations. Dose–response curves were plotted by fitting the normalized data into the equation, where A_1 and A_2 are the bottom and top asymptotes, respectively, p is the Hillslope, y is the percent response at a given concentration, and x is logarithm of ligand concentration. Independent experiments were run in triplicate and repeated at least three times.

Laboratory rearing of Aedes albopictus. Adult *Ae. albopictus* mosquitoes were obtained from the laboratory colony of the Benaki Phytopathological Institute (Kifissia, Greece). The colony is maintained under specific laboratory conditions (25 ± 2 °C, 80% relative humidity, and a 16/8-hour light/dark photoperiod). Larvae were reared in cylindrical enamel pans filled with tap water, with approximately 400 larvae per pan. They were fed ad libitum with powdered fish food (JBL Novo Tom 10% Artemia) until they emerged as adults (26).

Repellence Bioassays. For the in vivo determination of the repellent activity of tested compounds, the assessment was based on human hand landing counts using cages (33×33×33 cm) equipped with a 32×32 mesh on one side. Each cage contained 100 adult mosquitoes (5 to 10 days old, sex ratio 1:1) starved for 12 hours at 25 ± 2 °C and 70–80% relative humidity (31). Tested compounds were applied on chromatography paper (Whatman), covering a total area of 24 cm^2, at dose equivalent to 50 nmole/cm^2, diluted with dichloromethane (DCM). Data concerning the repellency indices were analyzed using the Kruskal–Wallis test (79). When significant differences were detected, Mann–Whitney U tests were carried out for pair-wise comparison with a Bonferroni correction for adjustment of p-values (80). Mosquito landings for each treatment were counted over 5-minute periods. Each treatment was repeated eight times and four human volunteers were used. Landing numbers were converted to repellency indices (RI ± SE) using the following equation: RI = [1 - T/C] x 100, where C is the number of landings in the control and T is the number of landings in the treatment (26).

DATA AVAILABILITY

All data are contained within the manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

ACKNOWLEDGMENTS

The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "1st Call for H.F.R.I. Research Projects to support Faculty Members & Researchers and the Procurement of High-and the procurement of high-cost research equipment grant" (Project Number: HFRI-FM17-637).

REFERENCES

1. (WHO, 2020) World Health Organization. Vector-borne diseases fact sheet Vector-borne diseases. WHO website [online; Accessed 1 January 2024] https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases.
2. Wicher, D., and Miazzi, F. (2021) Functional properties of insect olfactory receptors: ionotropic receptors and odorant receptors. *Cell Tissue Res* **383**, 7-19
3. Fan, X. B., Mo, B. T., Li, G. C., Huang, L. Q., Guo, H., Gong, X. L., and Wang, C. Z. (2022) Mutagenesis of the odorant receptor co-receptor (Orco) reveals severe olfactory defects in the crop pest moth Helicoverpa armigera. *BMC Biol* **20**, 214
4. Mier, P., Fontaine, J. F., Stoldt, M., Libbrecht, R., Martelli, C., Foitzik, S., and Andrade-Navarro, M. A. (2022) Annotation and Analysis of 3902 Odorant Receptor Protein Sequences from 21 Insect Species Provide Insights into the Evolution of Odorant Receptor Gene Families in Solitary and Social Insects. *Genes (Basel)* **13**
5. Yan, H., Jafari, S., Pask, G., Zhou, X., Reinberg, D., and Desplan, C. (2020) Evolution, developmental expression and function of odorant receptors in insects. *J Exp Biol* **223**
6. Jones, W. D., Nguyen, T. A., Kloss, B., Lee, K. J., and Vosshall, L. B. (2005) Functional conservation of an insect odorant receptor gene across 250 million years of evolution. *Curr Biol* **15**, R119-121
7. Hansson, B. S., and Stensmyr, M. C. (2011) Evolution of insect olfaction. *Neuron* **72**, 698-711
8. Butterwick, J. A., Del Marmol, J., Kim, K. H., Kahlson, M. A., Rogow, J. A., Walz, T., and Ruta, V. (2018) Cryo-EM structure of the insect olfactory receptor Orco. *Nature* **560**, 447-452
9. Del Marmol, J., Yedlin, M. A., and Ruta, V. (2021) The structural basis of odorant recognition in insect olfactory receptors. *Nature* **597**, 126-131
10. Jones, P. L., Pask, G. M., Romaine, I. M., Taylor, R. W., Reid, P. R., Waterson, A. G., Sulikowski, G. A., and Zwiebel, L. J. (2012) Allosteric antagonism of insect odorant receptor ion channels. *PLoS One* **7**, e30304
11. Jones, P. L., Pask, G. M., Rinker, D. C., and Zwiebel, L. J. (2011) Functional agonism of insect odorant receptor ion channels. *Proc Natl Acad Sci U S A* **108**, 8821-8825
12. Chen, S., and Luetje, C. W. (2013) Phenylthiophenecarboxamide antagonists of the olfactory receptor co-receptor subunit from a mosquito. *PLoS One* **8**, e84575
13. Chen, S., and Luetje, C. W. (2012) Identification of new agonists and antagonists of the insect odorant receptor co-receptor subunit. *PLoS One* **7**, e36784
14. Tsitoura, P., Koussis, K., and Iatrou, K. (2015) Inhibition of Anopheles gambiae odorant receptor function by mosquito repellents. *J Biol Chem* **290**, 7961-7972

15. Tsitoura, P., Andronopoulou, E., Tsikou, D., Agalou, A., Papakonstantinou, M. P., Kotzia, G. A., Labropoulou, V., Swevers, L., Georgoussi, Z., and Iatrou, K. (2010) Expression and membrane topology of Anopheles gambiae odorant receptors in lepidopteran insect cells. *PLoS One* **5**, e15428
16. Wermuth, G., Ganellin, C. R., Lindberg, P., and Mitscher, L. A. (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure and Applied Chemistry* **70**, 1129-1143
17. Sanders, M. P. A., McGuire, R., Roumen, L., de Esch, I. J. P., de Vlieg, J., Klomp, J. P. G., and de Graaf, C. (2012) From the protein's perspective: the benefits and challenges of protein structure-based pharmacophore modeling. *Medchemcomm* **3**, 28-38
18. Seidel, T., Wieder, O., Garon, A., and Langer, T. (2020) Applications of the Pharmacophore Concept in Natural Product inspired Drug Design. *Mol Inform* **39**, e2000059
19. Tripathi, M. K., Nath, A., Singh, T. P., Ethayathulla, A. S., and Kaur, P. (2021) Evolving scenario of big data and Artificial Intelligence (AI) in drug discovery. *Mol Divers* **25**, 1439-1460
20. Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., and Kumar, P. (2021) Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers* **25**, 1315-1360
21. Gentile, F., Yaacoub, J. C., Gleave, J., Fernandez, M., Ton, A. T., Ban, F., Stern, A., and Cherkasov, A. (2022) Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. *Nat Protoc* **17**, 672-697
22. Heikamp, K., and Bajorath, J. (2013) The future of virtual compound screening. *Chem Biol Drug Des* **81**, 33-40
23. Nayarisseri, A., Khandelwal, R., Tanwar, P., Madhavi, M., Sharma, D., Thakur, G., Speck-Planche, A., and Singh, S. K. (2021) Artificial Intelligence, Big Data and Machine Learning Approaches in Precision Medicine & Drug Discovery. *Curr Drug Targets* **22**, 631-655
24. Cortes, C., and Vapnik, V. (1995) Support-vector networks. *Machine Learning* **20**, 273-297
25. Bernhard E. Boser, Guyon, M., and Vapnik, V. N. (1992) A training algorithm for optimal margin classifiers. *COLT '92: Proceedings of the fifth annual workshop on Computational learning theory*, 144–152
26. Kythreoti, G., Sdralia, N., Tsitoura, P., Papachristos, D. P., Michaelakis, A., Karras, V., Ruel, D. M., Yakir, E., Bohbot, J. D., Schulz, S., and Iatrou, K. (2021) Volatile allosteric antagonists of mosquito odorant receptors inhibit human-host attraction. *J Biol Chem* **296**, 100172
27. Kröber, T., Koussis, K., Bourquina, M., Tsitoura, P., Konstantopoulou, M., Awolola, T. S., Danie, F. R., Qiao, H., Pelosi, P., Iatrou, K., and Guerin, P. M. (2018) Odorant-binding protein-based identification of natural spatial repellents for the Africa malaria mosquito Anopheles gambiae. *Insect Biochemistry and Molecular Biology* **96**, 36-50
28. Chidambara Thanu, V., Jabeen, A., and Ranganathan, S. (2024) iBio-GATS—A Semi-Automated Workflow for Structural Modelling of Insect Odorant Receptors. *Int J Mol Sci* **25**, 3055
29. Pacalon, J., Audic, G., Magnat, J., Philip, M., Golebiowski, J., Moreau, C. J., and Topin, J. (2023) Elucidation of the structural basis for ligand binding and translocation in conserved insect odorant receptor co-receptors. *Nat Commun* **14**, 8182
30. Lopes, J. C. D., Dos Santos, F. M., Martins-Jose, A., Augustyns, K., and De Winter, H. (2017) The power metric: a new statistically robust enrichment-type metric for virtual screening applications with early recovery capability. *J Cheminform* **9**, 7

31. Giatropoulos, A., Emmanouel, N., Koliopoulos, G., and Michaelakis, A. (2012) A study on distribution and seasonal abundance of Aedes albopictus (Diptera: Culicidae) population in Athens, Greece. *J Med Entomol* **49**, 262-269
32. Xu, P., Choo, Y. M., Chen, Z., Zeng, F., Tan, K., Chen, T. Y., Cornel, A. J., Liu, N., and Leal, W. S. (2019) Odorant Inhibition in Mosquito Olfaction. *iScience* **19**, 25-38
33. Xu, P., Choo, Y. M., An, S., Leal, G. M., and Leal, W. S. (2022) Mosquito odorant receptor sensitive to natural spatial repellents and inhibitory compounds. *Insect Biochem Mol Biol* **144**, 103763
34. Tsitoura, P., and Iatrou, K. (2016) Positive Allosteric Modulation of Insect Olfactory Receptor Function by ORco Agonists. *Front Cell Neurosci* **10**, 275
35. Bohbot, J. D., and Dickens, J. C. (2009) Characterization of an enantioselective odorant receptor in the yellow fever mosquito Aedes aegypti. *PLoS One* **4**, e7032
36. Bohbot, J. D., Jones, P. L., Wang, G., Pitts, R. J., Pask, G. M., and Zwiebel, L. J. (2011) Conservation of indole responsive odorant receptors in mosquitoes reveals an ancient olfactory trait. *Chem Senses* **36**, 149-160
37. Wang, G. R., Carey, A. F., Carlson, J. R., and Zwiebel, L. J. (2010) Molecular basis of odor coding in the malaria vector mosquito Anopheles gambiae. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4418-4423
38. Bobkov, Y. V., Walker Iii, W. B., and Cattaneo, A. M. (2021) Altered functional properties of the codling moth Orco mutagenized in the intracellular loop-3. *Sci Rep* **11**, 3893
39. Liu, Q., Liu, W., Zeng, B., Wang, G., Hao, D., and Huang, Y. (2017) Deletion of the Bombyx mori odorant receptor co-receptor (BmOrco) impairs olfactory sensitivity in silkworms. *Insect Biochem Mol Biol* **86**, 58-67
40. Liu, H., Liu, T., Xie, L., Wang, X., Deng, Y., Chen, C. H., James, A. A., and Chen, X. G. (2016) Functional analysis of Orco and odorant receptors in odor recognition in Aedes albopictus. *Parasit Vectors* **9**, 363
41. Turner, R. M., Derryberry, S. L., Kumar, B. N., Brittain, T., Zwiebel, L. J., Newcomb, R. D., and Christie, D. L. (2014) Mutational analysis of cysteine residues of the insect odorant co-receptor (Orco) from Drosophila melanogaster reveals differential effects on agonist- and odorant-tuning receptor-dependent activation. *J Biol Chem* **289**, 31837-31845
42. Halty-deLeon, L., Miazzi, F., Kaltofen, S., Hansson, B. S., and Wicher, D. (2016) The mouse receptor transporting protein RTP1S and the fly SNMP1 support the functional expression of the Drosophila odorant coreceptor Orco in mammalian culture cells. *J Neurosci Methods* **271**, 149-153
43. Anderson, A. R., Wanner, K. W., Trowell, S. C., Warr, C. G., Jaquin-Joly, E., Zagatti, P., Robertson, H., and Newcomb, R. D. (2009) Molecular basis of female-specific odorant responses in Bombyx mori. *Insect Biochem Mol Biol* **39**, 189-197
44. Douris, V., Swevers, L., Labropoulou, V., Andronopoulou, E., Georgoussi, Z., and Iatrou, K. (2006) Stably transformed insect cell lines: tools for expression of secreted and membrane-anchored proteins and high-throughput screening platforms for drug and insecticide discovery. *Adv Virus Res* **68**, 113-156
45. Kiely, A., Authier, A., Kralicek, A. V., Warr, C. G., and Newcomb, R. D. (2007) Functional analysis of a Drosophila melanogaster olfactory receptor expressed in Sf9 cells. *J Neurosci Methods* **159**, 189-194
46. Romaine, I. M., Taylor, R. W., Saidu, S. P., Kim, K., Sulikowski, G. A., Zwiebel, L. J., and Waterson, A. G. (2014) Narrow SAR in odorant sensing Orco receptor agonists. *Bioorg Med Chem Lett* **24**, 2613-2616

47. Taylor, R. W., Romaine, I. M., Liu, C., Murthi, P., Jones, P. L., Waterson, A. G., Sulikowski, G. A., and Zwiebel, L. J. (2012) Structure-activity relationship of a broad-spectrum insect odorant receptor agonist. *ACS Chem Biol* **7**, 1647-1652
48. Pask, G. M., Bobkov, Y. V., Corey, E. A., Ache, B. W., and Zwiebel, L. J. (2013) Blockade of insect odorant receptor currents by amiloride derivatives. *Chem Senses* **38**, 221-229
49. DeGennaro, M., McBride, C. S., Seeholzer, L., Nakagawa, T., Dennis, E. J., Goldman, C., Jasinskiene, N., James, A. A., and Vosshall, L. B. (2013) orco mutant mosquitoes lose strong preference for humans and are not repelled by volatile DEET. *Nature* **498**, 487-491
50. Franco, T. A., Oliveira, D. S., Moreira, M. F., Leal, W. S., and Melo, A. C. (2016) Silencing the odorant receptor co-receptor RproOrco affects the physiology and behavior of the Chagas disease vector Rhodnius prolixus. *Insect Biochem Mol Biol* **69**, 82-90
51. Koutroumpa, F. A., Monsempes, C., Francois, M. C., de Cian, A., Royer, C., Concordet, J. P., and Jacquin-Joly, E. (2016) Heritable genome editing with CRISPR/Cas9 induces anosmia in a crop pest moth. *Sci Rep* **6**, 29620
52. Soffan, A., Antony, B., Abdelazim, M., Shukla, P., Witjaksono, W., Aldosari, S. A., and Aldawood, A. S. (2016) Silencing the Olfactory Co-Receptor RferOrco Reduces the Response to Pheromones in the Red Palm Weevil, Rhynchophorus ferrugineus. *PLoS One* **11**, e0162203
53. Trible, W., Olivos-Cisneros, L., McKenzie, S. K., Saragosti, J., Chang, N. C., Matthews, B. J., Oxley, P. R., and Kronauer, D. J. C. (2017) orco Mutagenesis Causes Loss of Antennal Lobe Glomeruli and Impaired Social Behavior in Ants. *Cell* **170**, 727-735 e710
54. Yan, H., Opachaloemphan, C., Mancini, G., Yang, H., Gallitto, M., Mlejnek, J., Leibholz, A., Haight, K., Ghaninia, M., Huo, L., Perry, M., Slone, J., Zhou, X., Traficante, M., Penick, C. A., Dolezal, K., Gokhale, K., Stevens, K., Fetter-Pruneda, I., Bonasio, R., Zwiebel, L. J., Berger, S. L., Liebig, J., Reinberg, D., and Desplan, C. (2017) An Engineered orco Mutation Produces Aberrant Social Behavior and Defective Neural Development in Ants. *Cell* **170**, 736-747 e739
55. Zhang, R., Gao, G., and Chen, H. (2016) Silencing of the olfactory co-receptor gene in Dendroctonus armandi leads to EAG response declining to major host volatiles. *Sci Rep* **6**, 23136
56. Rinker, D. C., Jones, P. L., Pitts, R. J., Rutzler, M., Camp, G., Sun, L. J., Xu, P. X., Dorset, D. C., Weaver, D., and Zwiebel, L. J. (2012) Novel high-throughput screens of Anopheles gambiae odorant receptors reveal candidate behaviour-modifying chemicals for mosquitoes. *Physiological Entomology* **37**, 33-41
57. Giordano, D., Biancaniello, C., Argenio, M. A., and Facchiano, A. (2022) Drug Design by Pharmacophore and Virtual Screening Approach. *Pharmaceuticals (Basel)* **15**
58. Lu, S. H., Wu, J. W., Liu, H. L., Zhao, J. H., Liu, K. T., Chuang, C. K., Lin, H. Y., Tsai, W. B., and Ho, Y. (2011) The discovery of potential acetylcholinesterase inhibitors: a combination of pharmacophore modeling, virtual screening, and molecular docking studies. *J Biomed Sci* **18**, 8
59. Singh, M., Rajawat, J., Kuldeep, J., Shukla, N., Mishra, D. P., and Siddiqi, M. I. (2022) Integrated support vector machine and pharmacophore based virtual screening driven identification of thiophene carboxamide scaffold containing compound as potential PARP1 inhibitor. *J Biomol Struct Dyn* **40**, 8494-8507
60. Wu, H., Liu, J., Zhang, R., Lu, Y., Cui, G., Cui, Z., and Ding, Y. (2024) A review of deep learning methods for ligand based drug virtual screening. *Fundam Res* **4**, 715-737

61. Sun, H., Pan, P., Tian, S., Xu, L., Kong, X., Li, Y., Dan, L., and Hou, T. (2016) Constructing and Validating High-Performance MIEC-SVM Models in Virtual Screening for Kinases: A Better Way for Actives Discovery. *Sci Rep* **6**, 24817
62. Rodriguez-Perez, R., and Bajorath, J. (2022) Evolution of Support Vector Machine and Regression Modeling in Chemoinformatics and Drug Discovery. *J Comput Aided Mol Des* **36**, 355-362
63. Niazi, S. K., and Mariam, Z. (2023) Recent Advances in Machine-Learning-Based Chemoinformatics: A Comprehensive Review. *Int J Mol Sci* **24**
64. Bhattacharjee, A. K., Dheranetra, W., Nichols, D. A., and Gupta, R. K. (2005) 3D pharmacophore model for insect repellent activity and discovery of new repellent candidates. *Qsar & Combinatorial Science* **24**, 593-602
65. Kepchia, D., Xu, P., Terryn, R., Castro, A., Schurer, S. C., Leal, W. S., and Luetje, C. W. (2019) Use of machine learning to identify novel, behaviorally active antagonists of the insect odorant receptor co-receptor (Orco) subunit. *Sci Rep* **9**, 4055
66. Caballero-Vidal, G., Bouysset, C., Gevar, J., Mbouzid, H., Nara, C., Delaroche, J., Golebiowski, J., Montagne, N., Fiorucci, S., and Jacquin-Joly, E. (2021) Reverse chemical ecology in a moth: machine learning on odorant receptors identifies new behaviorally active agonists. *Cell Mol Life Sci* **78**, 6593-6603
67. Caballero-Vidal, G., Bouysset, C., Grunig, H., Fiorucci, S., Montagne, N., Golebiowski, J., and Jacquin-Joly, E. (2020) Machine learning decodes chemical features to identify novel agonists of a moth odorant receptor. *Sci Rep* **10**, 1655
68. Sims, C., Birkett, M. A., and Withall, D. M. (2022) Enantiomeric Discrimination in Insects: The Role of OBPs and ORs. *Insects* **13**
69. Bohbot, J. D., and Pitts, R. J. (2015) The narrowing olfactory landscape of insect odorant receptors. **3**
70. Chen, C. N., Shih, Y. H., Ding, Y. L., and Leong, M. K. (2011) Predicting activation of the promiscuous human pregnane X receptor by pharmacophore ensemble/support vector machine approach. *Chem Res Toxicol* **24**, 1765-1778
71. Cieslak, M., Danel, T., Krzysztynska-Kuleta, O., and Kalinowska-Tluscik, J. (2024) Machine learning accelerates pharmacophore-based virtual screening of MAO inhibitors. *Sci Rep* **14**, 8228
72. Zhao, J., Chen, A. Q., Ryu, J., and Del Marmol, J. (2024) Structural basis of odor sensing by insect heteromeric odorant receptors. *Science* **384**, 1460-1467
73. Wang, Y., Qiu, L., Wang, B., Guan, Z., Dong, Z., Zhang, J., Cao, S., Yang, L., Wang, B., Gong, Z., Zhang, L., Ma, W., Liu, Z., Zhang, D., Wang, G., and Yin, P. (2024) Structural basis for odorant recognition of the insect odorant receptor OR-Orco heterocomplex. *Science* **384**, 1453-1460
74. Grace, T. D. (1967) Establishment of a line of cells from the silkworm Bombyx mori. *Nature* **216**, 613
75. Bovolenta, S., Foti, M., Lohmer, S., and Corazza, S. (2007) Development of a Ca(2+)-activated photoprotein, Photina, and its application to high-throughput screening. *J Biomol Screen* **12**, 694-704
76. Farrell, P. J., Lu, M., Prevost, J., Brown, C., Behie, L., and Iatrou, K. (1998) High-level expression of secreted glycoproteins in transformed lepidopteran insect cells using a novel expression vector. *Biotechnol Bioeng* **60**, 656-663
77. Lu, M., Johnson, R. R., and Iatrou, K. (1996) Trans-activation of a cell housekeeping gene promoter by the IE1 gene product of baculoviruses. *Virology* **218**, 103-113

78. Lu, M., Farrell, P. J., Johnson, R., and Iatrou, K. (1997) A Baculovirus (Bombyx mori Nuclear Polyhedrosis Virus) Repeat Element Functions as a Powerful Constitutive Enhancer in Transfected Insect Cells*. *Journal of Biological Chemistry* **272**, 30724-30728
79. Kruskal, W. H., and Wallis, W. A. (1952) Use of Ranks in One-Criterion Variance Analysis. *J Am Stat Assoc* **47**, 583-621
80. Hazra, A., and Gogtay, N. (2016) Biostatistics Series Module 3: Comparing Groups: Numerical Variables. *Indian J Dermatol* **61**, 251-260

FIGURE LEGENDS

Figure 1. The pharmacophore model. *Left*: The ligand-based pharmacophore and the 4 orthosteric antagonists of **Table 1** (Isopropyl cinnamate, Ethyl cinnamate, Linalyl acetate, 2,4-octadienal) used to generate it. The features include one atom-centered hydrophobic feature HydA (green), two centroid hydrophobic features Hyd (green) and one projected location of potential H-bond donors Acc2 (blue). *Right:* Spacing distances between the specific pharmacophore features.

Figure 2. *Ex vivo* screening results. All compounds were tested at a final concentration of 100µM. The primary compound additions (white bars) did not induce significant ORco channel function, while secondary additions of the OA (ORcoRAM2) to wells containing primary additions of functionally inactive compounds produced responses (grey bars) equal to at least 80% of the full response obtained in the control wells (OA only added, set as 100%; red bar at right of each panel). ORco antagonist hits (green bars) produced significantly lower secondary responses, set arbitrarily at ⩽60% of the normal channel response, upon OA addition. Arabic numbers correspond to those of the compounds listed in **Table S2,** while roman numbers are those of the previously characterized ORco antagonists (blue bars) shown in **Table 1**. Error bars indicate mean±SE. Mean values report independent experiments run in triplicate, technical repetition, and replicated at least three times, biological repetition.

Figure 3. Competition plots for 8 active compounds [orthosteric (upper panels) and allosteric antagonists (lower panels)] showing the % response as a function of ligands concentration in the presence of 50, 100 and 150µM of ORcoRAM2. Error bars indicate mean±SE. Mean values report independent experiments run in triplicate, technical repetition, and replicated at least three times, biological repetition. (For additional data on pIC_{50} and R^2, please see **Table S3**)

Figure 4. The selected SVM model. The SVM model was trained on all pharmacophore hits (**Tables 2 and 3**) using the 2D descriptors KierA2 and SlogP_VSA1 (MOE software). *Ex vivo* active orthosteric antagonists are represented by green diamonds whereas non-active compounds are indicated by red circles. Compounds with coordinates (KierA2, SlogP_VSA1) that lie within the areas delineated by the decision boundaries (solid lines) are predicted to be orthosteric antagonists of ORco.

Figure 5. Box plots depicting repellency indices (RIs) against *Aedes albopictus* mosquitoes in "hand in a cage" repellence assays. **(A)** Selected orthosteric and **(B)** allosteric antagonists (green) and the widely used insect repellent DEET (red) were examined using 10nmole of each tested compound per cm^2 of exposed hand area (240nmole/24cm^2 total exposed area). Each treatment was replicated eight times, utilizing four human volunteers. The previously characterized antagonists isopropyl cinnamate (IPC; blue) and carvacrol (CRV; blue) (26) served as controls for the tested orthosteric and allosteric antagonists, respectively. The box plots represent the mean values with upper and lower quartiles, and the range of outliers within 1.5IQR are indicated by error bars. Compound identities are listed in **Table S1** and **S2**. Different letters (a, b, etc.) indicate statistically significant differences between tested compounds ($p<0.05$), Mann–Whitney U test with Bonferroni correction (adjusted p values a=0.005 and a=0.003 for the orthosteric and allosteric group, respectively).

Figure 6. *in silico* screening of a new VOC library (33) for orthosteric Orco antagonists. Starting from 241 VOCs, the pharmacophore identified 100 hits, 44 of which were retained by the SVM filter. Sensitivity, Specificity and Virtual Screening PM performance were calculated as described under Experimental Procedures.

Figure 7. The SVM classification maps. The locations of the five workflow-retained hits (green diamonds) and the ten workflow-rejected compounds (red circles) (see also **Table 6**) are shown in the diagram in the context of their inclusion within or exclusion from the defined SVM boundaries. Hexanoic acid, which was rejected by the pharmacophore model but was found to be marginally active as an antagonist in the *ex vivo* assays, is indicated by an orange circle inside the upper SVM boundary of bioactive hits. The *ex vivo* activities of the fifteen compounds are shown in **Figure 8**.

Figure 8. Validation of virtual screening results. *Ex vivo* functional assays were carried out for 15 randomly selected compounds that included 5 workflow-retained (green) and 10 workflow-rejected (orange) compounds**.** The Orco/Photina activity platform (14,26,34) was employed using isopropyl cinnamate (IPC; blue) as antagonist activity standard (52% response or 48% inhibition of the normal activity in the presence of 100µM OA). The primary compound additions, each at a 100µM concentration, did not induce significant ORco channel function (<15% for all of them; data not shown). The cutoff response point for antagonistic activity against 100µM OA was arbitrarily set at <60% (>40% inhibition of the 100% activity obtained by addition of 100µM Orco plus solvent shown in red). Box plots depict mean values, with upper and lower quartiles, and the range of outliers within 1.5IQR are indicated by error bars. The response values for the tested compounds are listed in **Table 6.**

Table 1. Previously identified ORco orthosteric and allosteric antagonists. Structural features of previously identified AgamORco orthosteric and allosteric antagonists. Compound numbers are the same as those presented in (26).

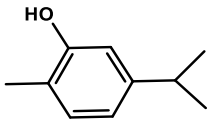
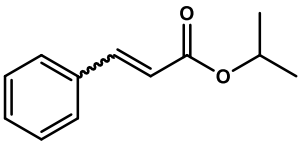
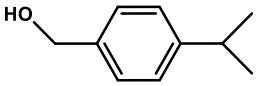
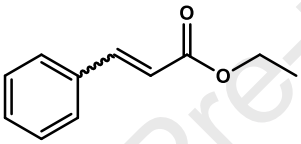
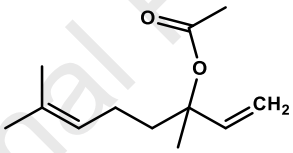
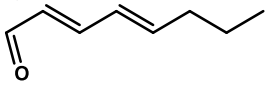
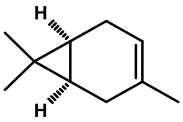
No	Compound	Structure	Chemical class	Antagonist type
I	Carvacrol (CRV)		Monoterpene alcohol	allosteric
II	Isopropyl cinnamate (IPC)		cinnamate ester	orthosteric
III	Cumin alcohol (CA)		Monoterpene alcohol	allosteric
IV	Ethyl cinnamate (EC)		cinnamate ester	orthosteric
4	Linalyl acetate (LA)		monoterpene ester	orthosteric
39	2,4-octadienal (OCT)		fatty aldehyde	orthosteric
45	(1S)-3-Carene (CAR)		monoterpene	allosteric

Table 2. ORco orthosteric antagonist hits and *ex vivo* validation. The training set for the selected pharmacophore model consisted of the four confirmed orthosteric antagonists shown in **Table 1** and fifty negative examples (3 allosteric and 47 inactive) shown in **Table 1** and **Table S1**. Compound numbering is as per (26) and **Table S1**; functionally validated (bioactive) hits are shown in bold, while inactive ones are shown in italics; √: >40% inhibition; NA: not active (<40% inhibition).

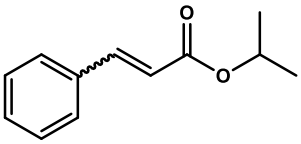
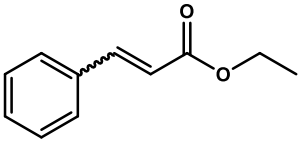
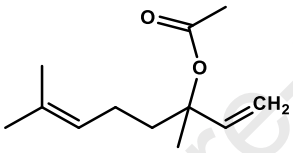
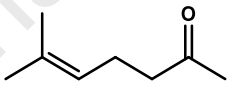
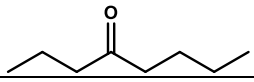
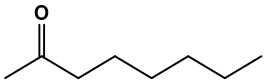
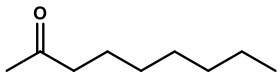
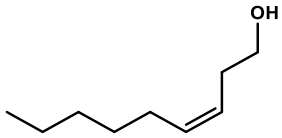
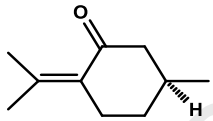
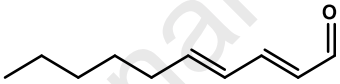
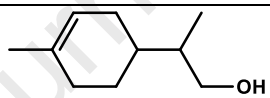
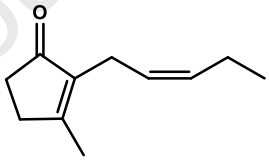
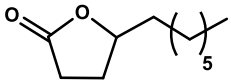
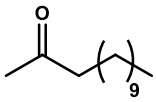
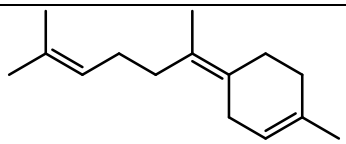
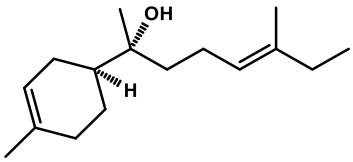
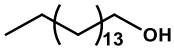
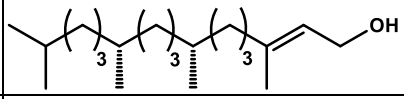
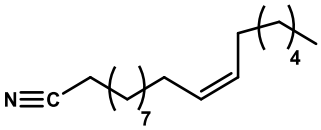
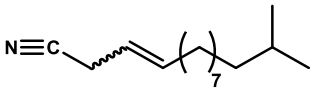
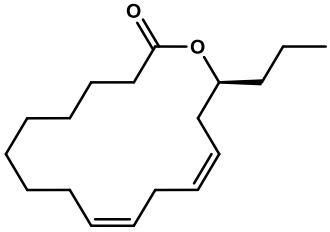
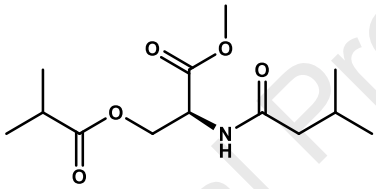
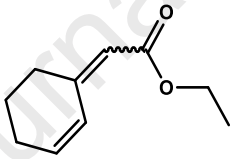
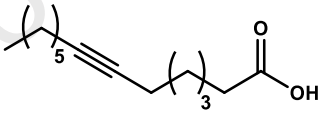
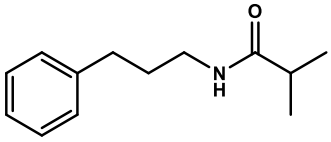
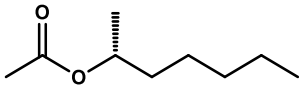
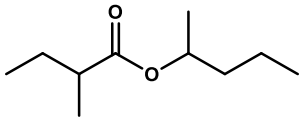
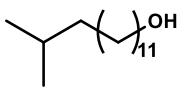
No	Compound	Structure	Chemical class	*Ex vivo* validation
II	**Isopropyl cinnamate**		cinnamate ester	√
IV	**Ethyl cinnamate**		cinnamate ester	√
4	**Linalyl acetate**		monoterpene ester	√
39	**2,4-octadienal**		fatty aldehyde	√
33	*2-Heptanone*		ketone	NA
40	*6-Methyl-5- hepten-2-one*		ketone	NA
42	*4-Octanone*		ketone	NA
43	*2-Octanone*		ketone	NA

Table 3. *Virtual screening* of a new compound collection with the specific pharmacophore model. Validated antagonist hits have been defined as VOCs causing at least 40% inhibition of ORco activity in the *ex vivo* assays shown in **Figure 2**. ***X***: not detected by the *in silico* screen; √: pharmacophore antagonist hits; NA: Not active or less than 40% maximum inhibition in the competition assay. IC_{50}: Concentration of 50% inhibition in the presence of 100µM ORco agonist. Bold font indicates *ex vivo* active pharmacophore hits whereas italics indicates the *ex vivo* active compound that escaped detection by the pharmacophore.

No	Compound	Structure	Pharmacophore hits	Chemical class	*Ex vivo* validation (IC_{50})
53	2-Nonanone		√	ketone	NA
54	**(Z)-3-Nonen-1-ol**		√	**aliphatic alcohol**	**√ (48.9µM)**
57	Pulegone		√	monoterpene ketone	NA
59	Limonene oxide (cis/trans mix)		√	monoterpene epoxide	NA
60	**(2E,4E)-Decadienal**		√	**fatty aldehyde**	**√ (66.7µM)**
65	p-Menth-1-en-9-ol		√	monoterpene alcohol	NA
68	cis-Jasmone		√	terpenoid	NA
71	γ-Undecalactone		√	lactone	NA
72	2-Tridecanone		√	ketone	NA
74	*Bisabolene (mix of isomers)*		***X***	*sesquiterpene*	*√ (47.7µM)*
77	**α-Bisabolol**		√	**sesquiterpene alcohol**	**√ (47µM)**
78	1-Hexadecanol		√	fatty alcohol	NA

No	Compound	Structure	Pharmacophore hits	Chemical class	*Ex vivo* validation (IC_{50})
79	Phytol		√	terpenoid	NA
81	(Z)-Octadec-11-ene nitrile		√	fatty nitrile	NA
83	**13-Methyl tetradec-3-ene nitrile**		√	**fatty nitrile**	√ **(25µM)**
84	(9Z,12Z,15S)-Octadeca-9,12-dien-15-olide		√	Macrocyclic unsaturated lactone	NA
85	N-(3-Methyl butyryl)-O-(2-methyl propionyl)-L-serine methyl ester		√	diester amide	NA
88	**ethyl (E/Z)-2-(cyclohex-2-en-1-ylidene) acetate (*mix of isomers*)**		√	**ester**	√ **(195.7µM)**
89	7-Tetradecynoic acid		√	unsaturated fatty acid	NA
93	N-Phenylethyl-2-methyl propionic acid amide		√	peptide	NA
94	(R)-2-Heptyl acetate		√	ester	NA
95	2-Pentyl 2-methylbutanoate		√	ester	NA
96	13-Methyl tetradecane-1-ol		√	fatty alcohol	NA

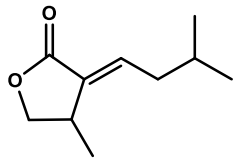
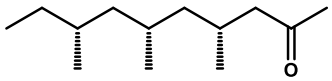
No	Compound	Structure	Pharmacophore hits	Chemical class	*Ex vivo* validation (IC_{50})
98	(E)-3-Methyl-2-(3-methylbutyliden)-4-butanolide		√	lactone	√ (43.2µM)
99	(4R,6R,8R)-trimethyldecan-2-one		√	ketone	√ (57µM)

Table 4. *Ex vivo* validated ORco allosteric antagonists. Structural features and chemical classes of identified AgamORco allosteric antagonists. IC_{50} values shown here are those obtained in the presence of 100 µM ORco agonist.

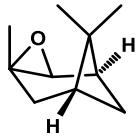
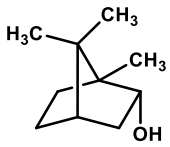
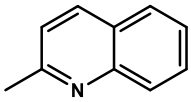
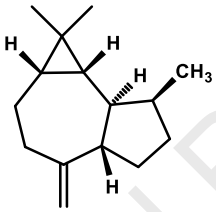
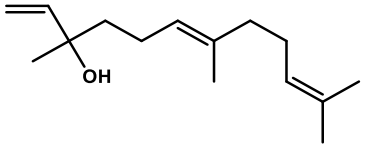
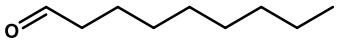
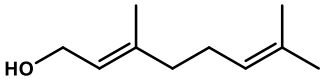
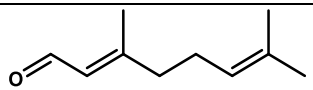
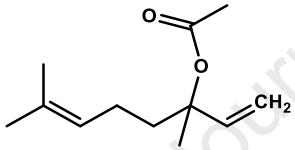
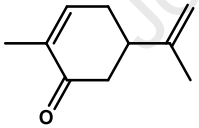
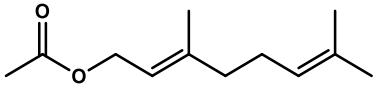
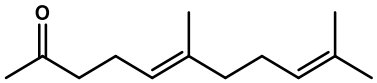
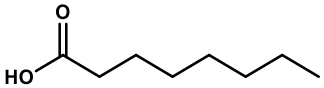
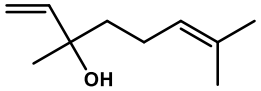
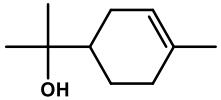
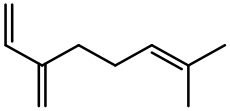
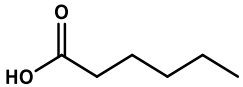
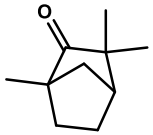
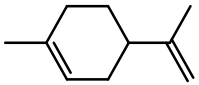
No	Compound	Structure	Pharmacophore hits	Chemical class	IC_{50}
62	α-Pinene oxide		X	Monoterpene epoxide	52µM
64	Borneol		X	Monoterpene alcohol	86.5µM
66	2-Methylquinoline		X	hydroquinoline	115µM
75	Aromadendrene		X	sesquiterpene	26.2µM

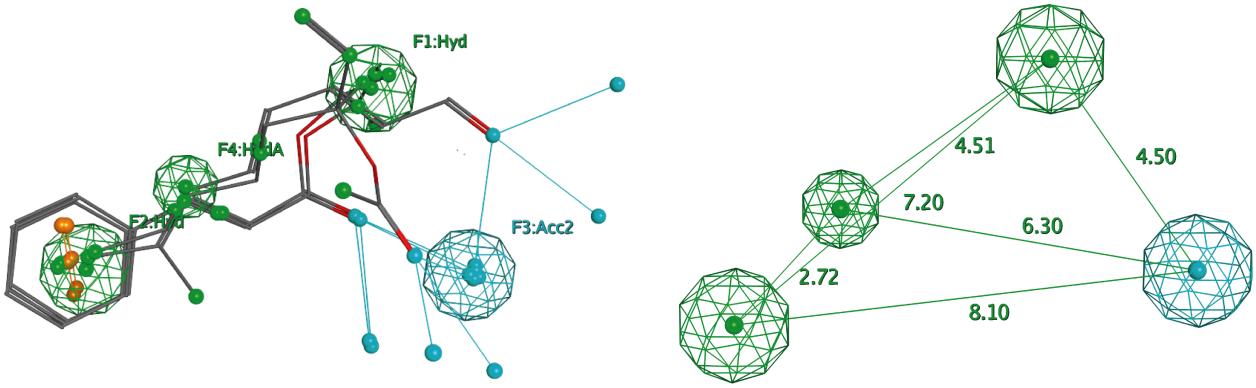
Table 5. Results of the selected SVM classification of the 32 pharmacophore hits. The table lists the orthosteric antagonist hits identified by the pharmacophore model after the application of the selected SVM model on the training set of compounds. The *ex vivo* active compound #99 was laid outside the decision boundaries of the SVM model (**Figure 4**) and was defined as a false-negative result yielding a cross-validation loss of 0.032. NA: not or marginally active (≤40% inhibition) in the *ex vivo* assays.

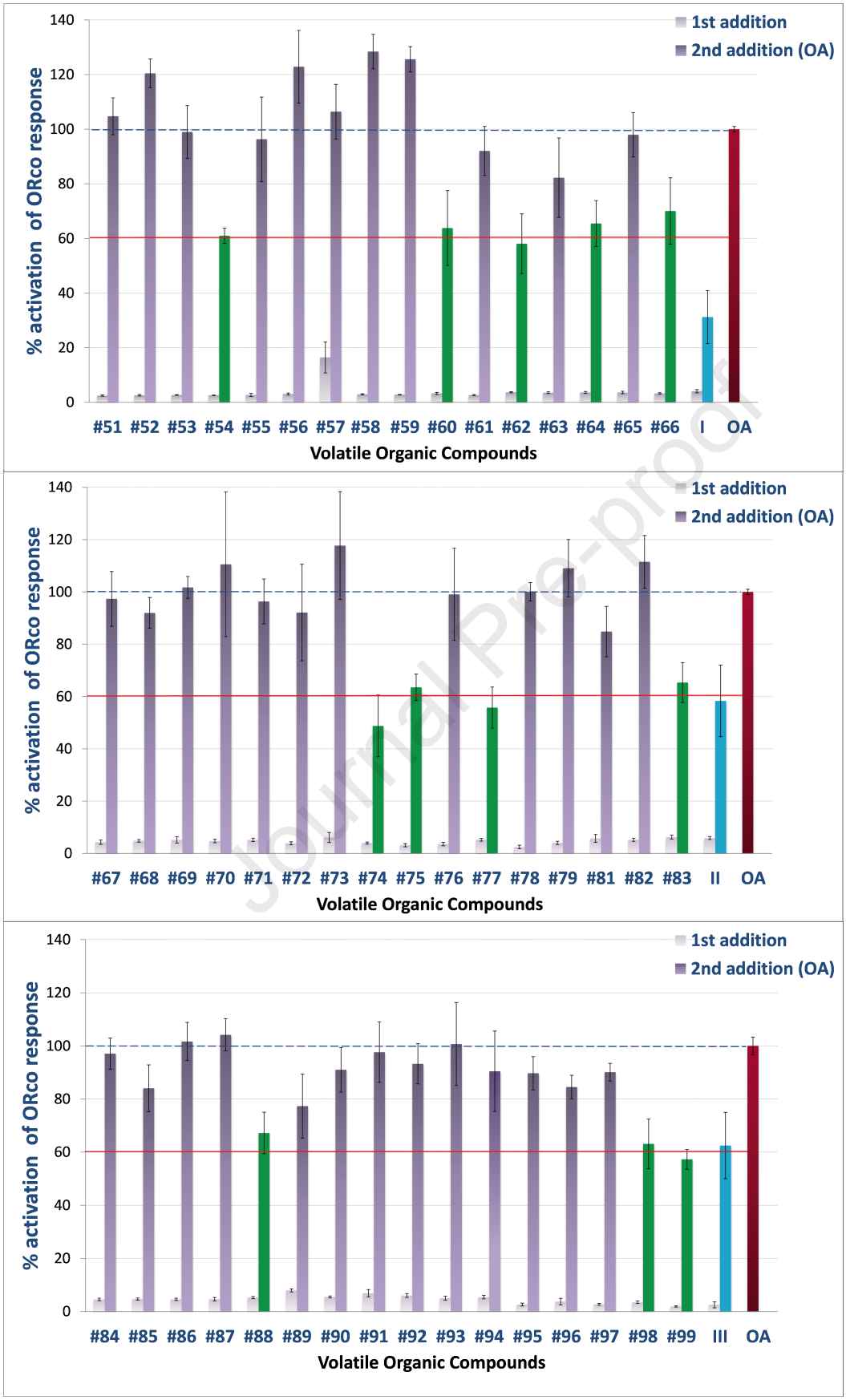
Compound	Cell-based activity (IC_{50} in µM)	KierA2	SlogP_VSA1	within SVM decision boundaries
II	41.7	4.5847445	7.7454643	**Yes**
IV	64.5	4.4210858	7.7454643	**Yes**
4	67.7	5.2678456	7.7454643	**Yes**
33	NA	4.4425101	5.6876111	No
39	59.8	5.5963559	0	**Yes**
40	NA	3.809427	5.6876111	No
42	NA	5.4008284	5.6876111	No
43	NA	5.4008284	5.6876111	No
53	NA	6.368185	5.6876111	No
54	48.9	8.1811314	0	**Yes**
57	NA	2.9886453	5.6876111	No
59	NA	2.0609839	0	No
60	66.7	7.5834055	0	**Yes**
65	NA	3.795996	0	No
68	NA	3.8128579	5.6876111	No
71	NA	5.9534798	7.7454643	No
72	NA	10.287	5.6876111	No
77	47	5.4362946	0	**Yes**
78	NA	16	0	No
79	NA	13.104808	0	No
81	NA	16.544603	0	No
83	25	11.143562	0	**Yes**
84	NA	12.104386	7.7454643	No
85	NA	8.1214361	20.749712	No
88	195.7	5.1365461	7.7454643	**Yes**
89	NA	11.143562	7.7454643	No
93	NA	5.4330149	5.2587838	No
94	NA	6.0677109	7.7454643	No
95	NA	5.9571776	7.7454643	No
96	NA	13.066667	0	No
98	43.2	3.932668	7.7454643	**Yes**
99	57	6.7910275	5.6876111	No

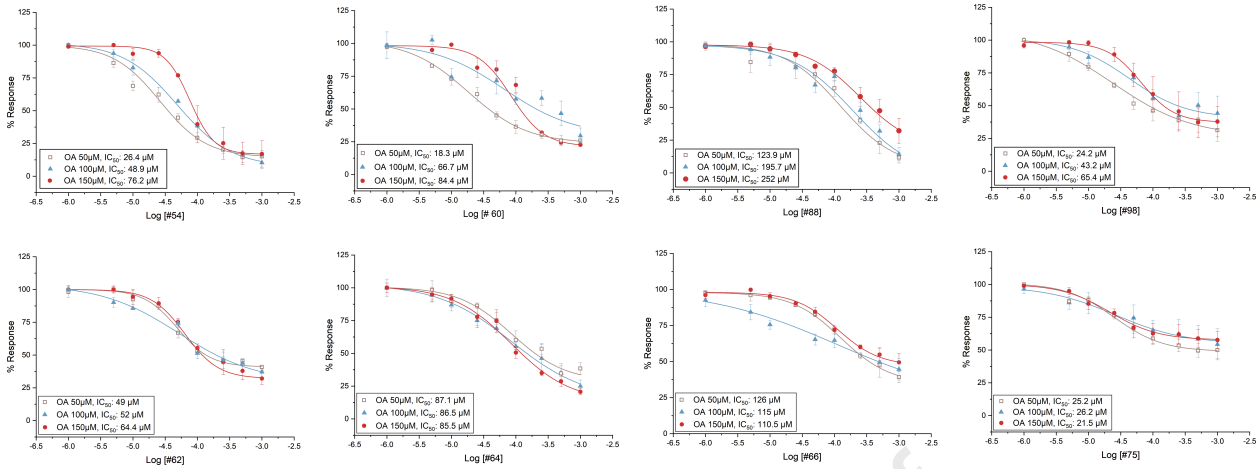
Table 6. Selected Pharmacophore-SVM pipeline hits and initial functional testing. Compounds causing ≤60% response are indicated in bold.

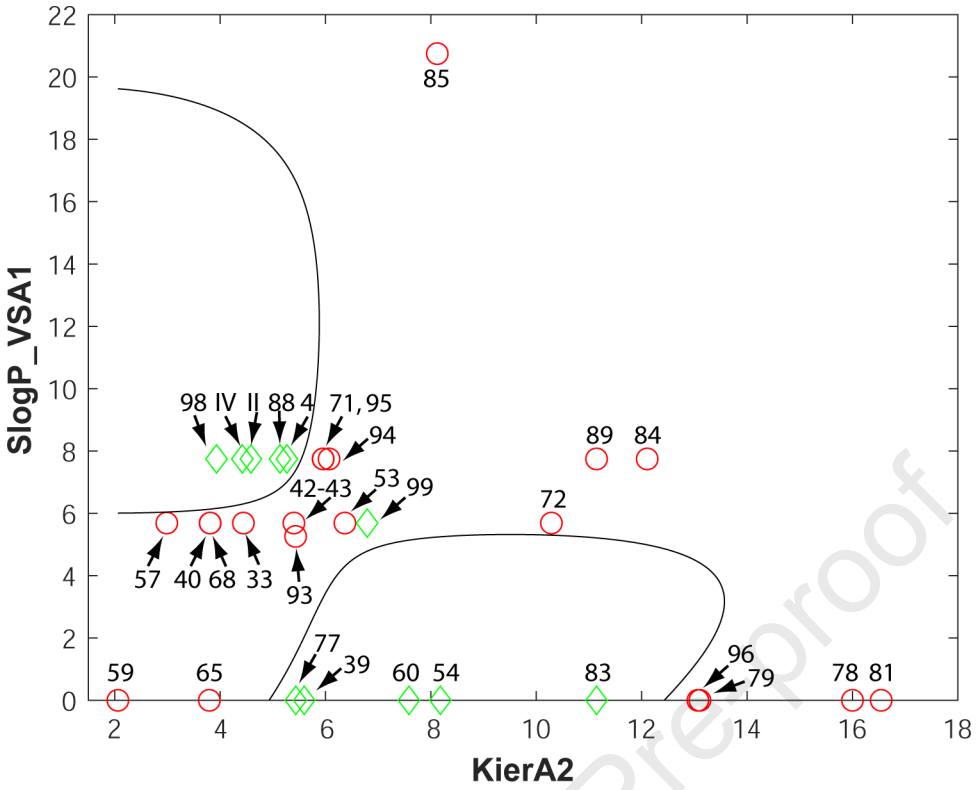
PubChem ID	Compound	Structure	Pipeline Retained	Chemical class	% Response at 100µM OA
5284507	**Nerolidol**		Y	sesquiterpene alcohol	51
31289	Nonanal		Y	fatty aldehyde	65
637566	**Geraniol**		Y	Acyclic Monoterpene alcohol	43
638011	Citral		Y	Monoterpene aldehyde	69
8294	**Linalyl acetate**		Y	Monoterpene ester	59
7439	Carvone		N	monoterpene ketone	71
1549026	Geranyl acetate		N	Monoterpene ester	67
1549778	Geranyl acetone		N	monoterpene ketone	92
379	Octanoic Acid		N	fatty acid	81

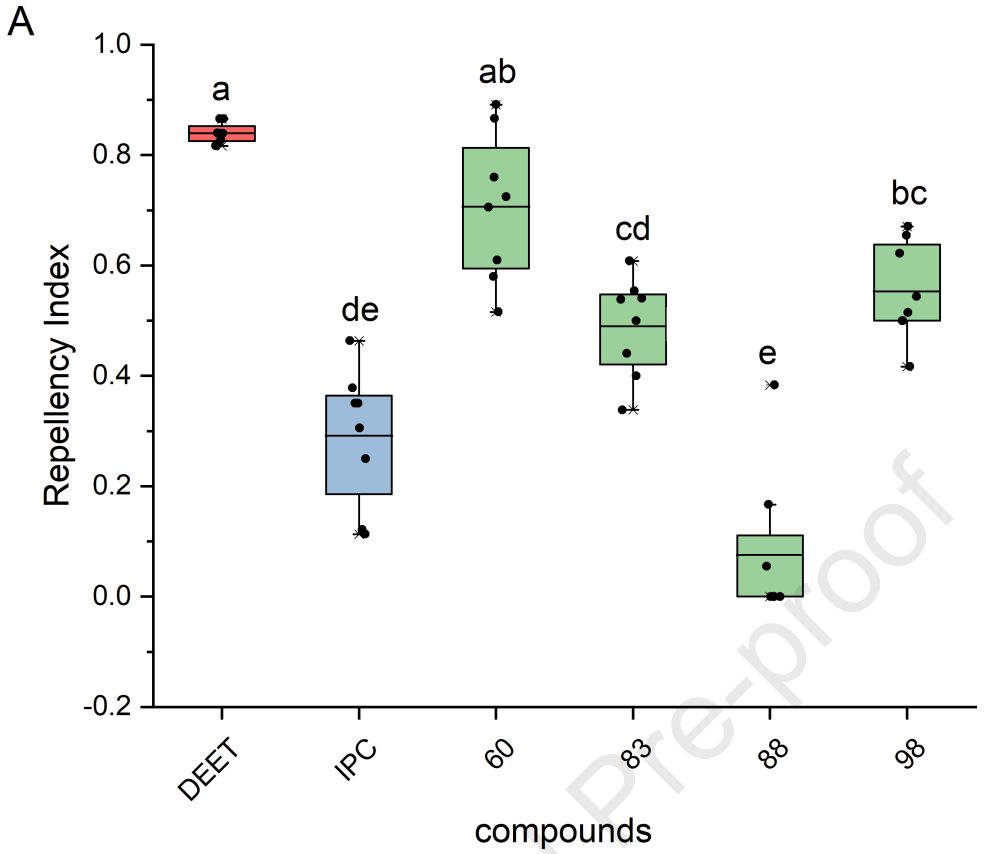
6549	Linalool		N	Monoterpene alcohol	98
17100	α-Terpineol		N	monoterpene alcohol	100
31253	Myrcene		N	monoterpene	71
8892	**Hexanoic Acid**		N	fatty acid	59
14525	Fenchone		N	monoterpene ketone	67
22311	Limonene		N	monoterpene	82

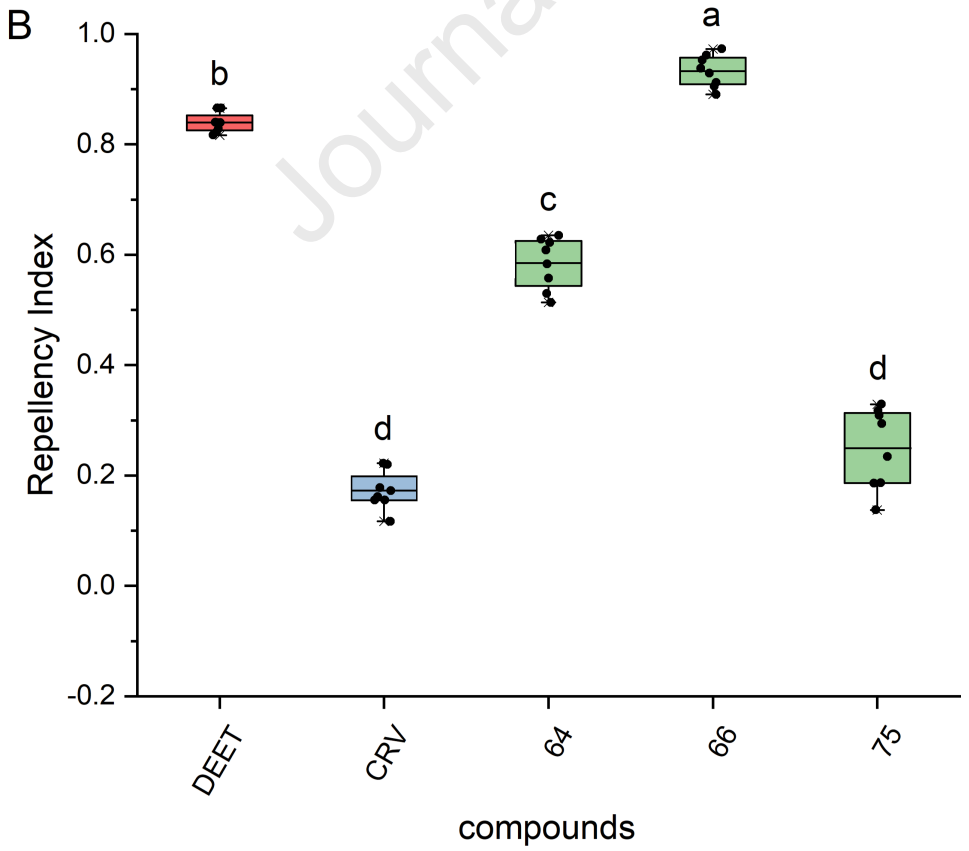


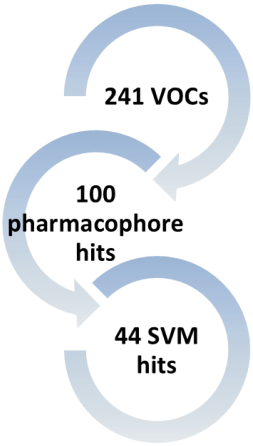




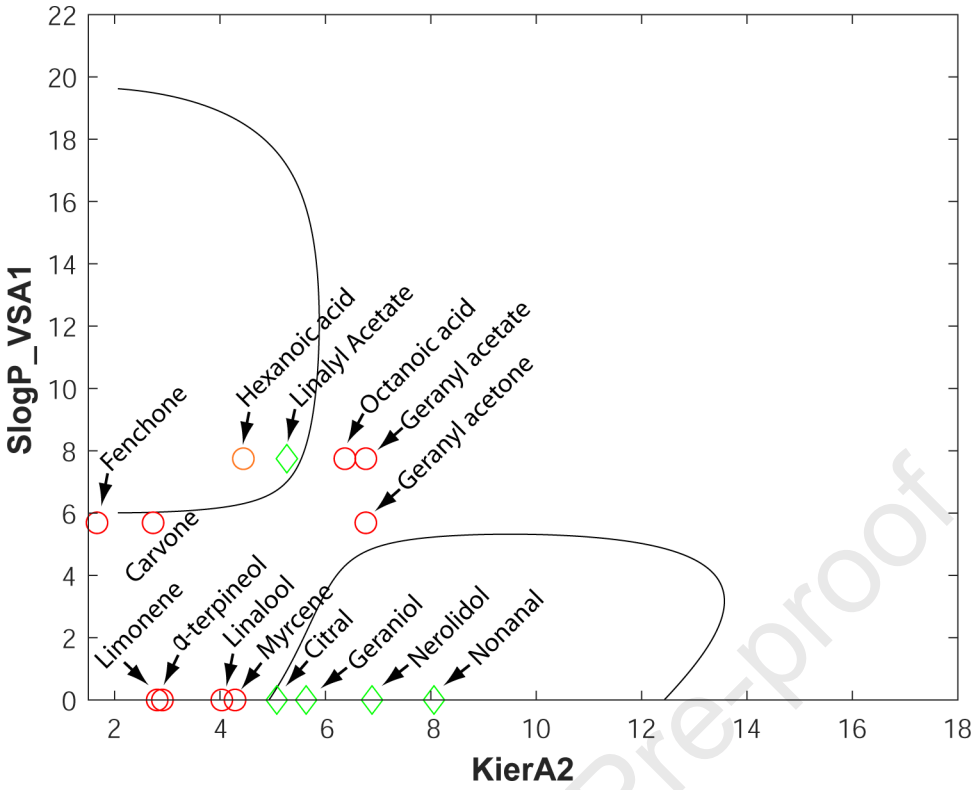


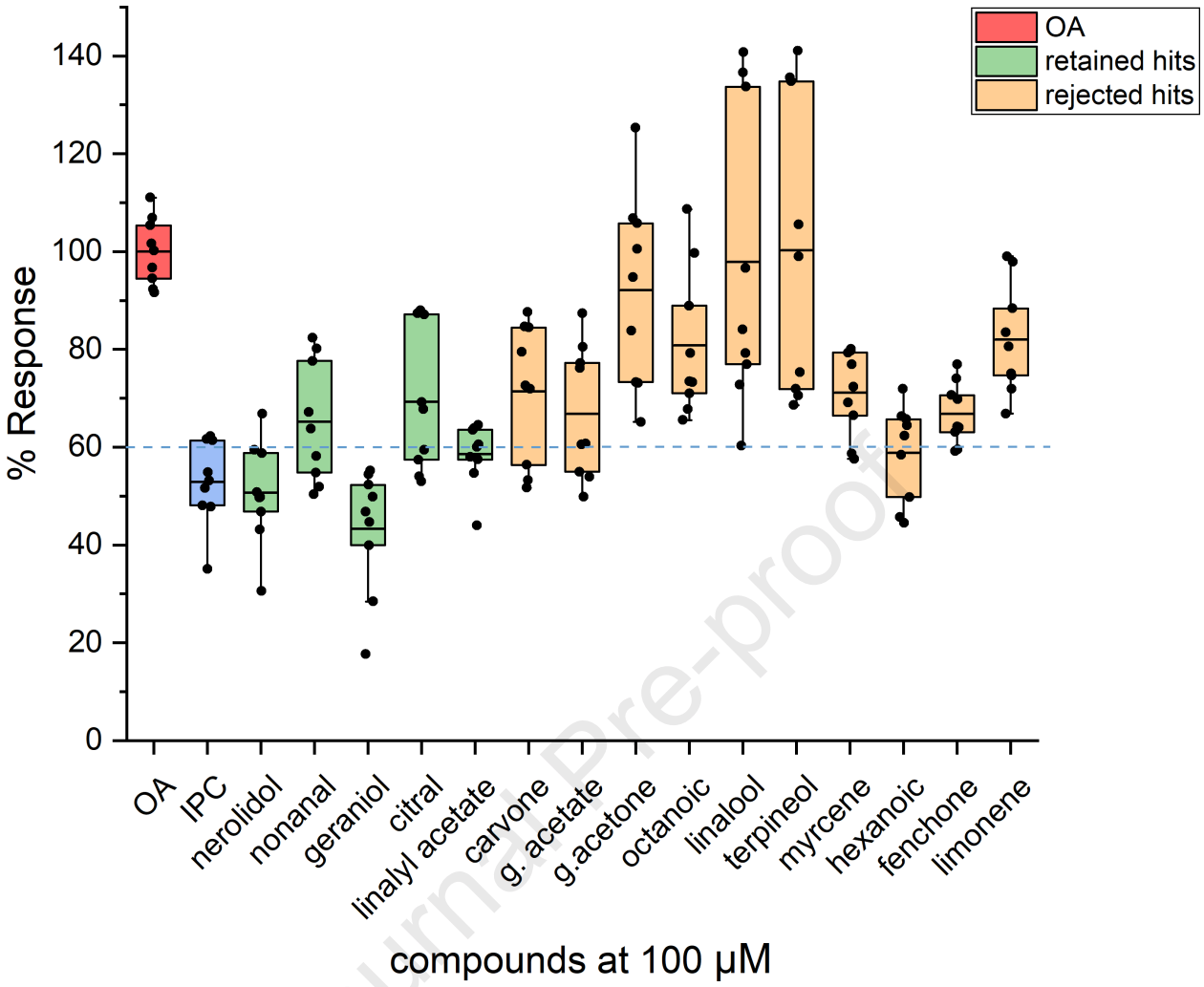






In silico screened compounds (pharmacophore)	241
Pharmacophore hits	100 of 241 (41.5%)
SVM hits	44 of 100 (44%)
Total workflow *in silico* hits	**44 of 241 (18.3%)**
Selected for functional testing	– 15 compounds – 5 predicted active – 10 predicted inactive
ex vivo active / *ex vivo* inactive	3 of 5 / 9 of 10
Sensitivity	0.75
Specificity	0.82
Virtual Screening Performance PM	0.81





Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: