



OPEN

High-throughput ligand profile characterization in novel cell lines expressing seven heterologous insect olfactory receptors for the detection of volatile plant biomarkers

Katalin Zboray^{1,2}, Adam V. Toth^{1,3}, Tímea D. Miskolczi^{1,3}, Krisztina Pesti^{2,3,6}, Emilio Casanova⁴, Emanuel Kreidl^{4,7}, Arpad Mike^{1,3}, Áron Szenes⁵, László Sági^{6,8} & Peter Lukacs^{1,6,8}✉

Agriculturally important crop plants emit a multitude of volatile organic compounds (VOCs), which are excellent indicators of their health status and their interactions with pathogens and pests. In this study, we have developed a novel cellular olfactory panel for detecting fungal pathogen-related VOCs we had identified in the field, as well as during controlled inoculations of several crop plants. The olfactory panel consists of seven stable HEK293 cell lines each expressing a functional *Drosophila* olfactory receptor as a biosensing element along with GCaMP6, a fluorescent calcium indicator protein. An automated 384-well microplate reader was used to characterize the olfactory receptor cell lines for their sensitivity to reference VOCs. Subsequently, we profiled a set of 66 VOCs on all cell lines, covering a concentration range from 1 to 100 μM . Results showed that 49 VOCs (74.2%) elicited a response in at least one olfactory receptor cell line. Some VOCs activated the cell lines even at nanomolar (ppb) concentrations. The interaction profiles obtained here will support the development of biosensors for agricultural applications. Additionally, the olfactory receptor proteins can be purified from these cell lines with sufficient yields for further processing, such as structure determination or integration with sensor devices.

The composition of volatile organic compounds (VOCs) in airborne emissions provides valuable information for various applications, including the detection of explosives or toxic materials¹ as well as the differentiation between healthy and diseased biological samples based on specific VOC blends².

In the animal kingdom, VOCs are bound and recognized by odorant or olfactory receptors (ORs), although several other peptide molecules such as Odorant Binding Proteins (OBP) are also capable of specific VOC recognition.

It is due to these receptors that animals, such as drug-sniffing dogs, can detect and trace marker volatiles with unprecedented sensitivity. Mammals have several hundred to a thousand ORs, which are G-protein-coupled receptors (GPCRs)³. Numerous research programs have focused on the development of biosensors based on such mammalian receptors. However, due to their action mechanism, the receptors themselves are not sufficient for odorant perception, and additional signal transduction molecules are required for their proper function⁴.

¹Plant Protection Institute, Centre for Agricultural Research, Martonvásár, Hungary. ²TetraLab Ltd., Budapest, Hungary. ³Department of Biochemistry, Eötvös Loránd University, Budapest, Hungary. ⁴Department of Pharmacology, Center of Physiology and Pharmacology, Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria. ⁵Department of Pathology, University of Veterinary Medicine, Budapest, Hungary. ⁶Agricultural Institute, Centre for Agricultural Research, Martonvásár, Hungary. ⁷Present address: Novartis AG, 6336 Langkampfen, Austria. ⁸These authors contributed equally: László Sági and Peter Lukacs. ✉email: lukacs.peter@atk.hu

In contrast, insect ORs have a substantially different structure and mode of operation. For insects, too, odor perception is critical for feeding, oviposition, mate recognition, and predator avoidance⁵; therefore, insects are also able to sense relevant odorants at minute concentrations. Moreover, they possess a much narrower receptor repertoire compared to mammals, and—unlike GPCRs—do not need additional components for odorant recognition⁶.

Insect ORs are typically assembled from two types of proteins: the tuning OR, which is responsible for VOC binding and ligand-specificity, and the Olfactory Receptor Co-receptor (ORCO⁷) which is essential for proper OR folding and function. While the amino acid sequence of OR proteins is very diverse, that of ORCO proteins is highly conserved even in diverse species. In many cases OR and ORCO protein combinations from different species can form functional channels. Cryo-EM structure determination on the ORCO protein from a parasitic fig wasp (*Apocrypta bakeri*) revealed a homotetrameric structure in the resulting autonomous cation channel⁸.

Regarding their action mechanism insect ORs predominantly work as ligand-gated ion channels^{9,10}, although G-protein-mediated signaling was also indicated in several studies^{9,11–14}. VOC ligand binding induces channel opening and an influx of cations. Channel permeability is much higher for monovalent (e.g., Na⁺, K⁺, and Cs⁺) than for divalent (Ca²⁺ and Mg²⁺) cations, and permeability ratios also depend on the identity of the OR⁸. Under physiological conditions, ORs are localized in the dendrites of olfactory sensory neurons, where they depolarize the membrane upon ligand binding. This depolarization activates voltage-gated ion channels and changes the firing rate of the neuron¹⁵.

Using the antenna of the fruit fly it was possible to distinguish between the smell of healthy and cancer cell lines based on their different OR activation patterns¹⁶. The receptor repertoire was mapped earlier for *Drosophila melanogaster*¹⁷. In vivo measurements, however, are not compatible with a sensor for the analysis of multiple samples. An olfactory panel consisting of several ORs selected for the specific sensing of the target molecules could work effectively as a biosensor.

Several research groups conducted experiments with heterologously expressed or purified ORs linked to sensor devices, but a complex insect receptor panel has not yet been developed. As a first step to this direction the ligand profiles of the sensory proteins must be characterized.

Here we present an improved method for the generation of OR-expressing cell lines in which odorant binding is easy to follow by GFP fluorescence change without the need of loading the cells with a dye. As an example of the methodology, we generated seven stable cell lines in HEK293 cells all expressing a different *Drosophila* OR, ORCO, and GCaMP6 as a genetically encoded calcium ion sensor protein. Using these cell lines as an olfactory panel, the concentration–response profile of a large set of odorants was mapped here for the first time. The set of 66 tested VOC molecules was compiled in relation to our research on plant- and plant pathogen-derived VOCs¹⁸. VOC-specific cell responses were determined with a high-throughput microplate reader assay over a concentration range of three orders of magnitude. The new OR cell lines and hitherto untested plant disease-related volatiles represent original contributions to the present panel and ligand list of the ORs examined.

Results and discussion

Generation of the stable OR cell lines

In vivo measured ligand profile data derived from the DoOR database (<http://neuro.uni-konstanz.de/DoOR/default.html>)^{19,20} were used to select 12 ORs, which can be appropriate to detect the pathogen-related VOCs relevant for our study. To assemble a novel cell-based olfactory panel, 11 stable cell lines were successfully generated which expressed different *Drosophila* ORs and the fluorescent calcium ion indicator protein GCaMP6²¹ in HEK293 cells. Out of these cell lines, seven were responsive to the respective OR-specific VOCs. Using these functional lines we constructed an olfactory assay in which binding of the examined odorants can be quantified by a fluorescence signal.

Four different proteins were expressed in each of the cell lines: the respective ORs, the ORCO universal co-receptor, the mCherry fluorescent protein as a marker for OR expression, and the GCaMP6 fluorescent calcium ion indicator protein. To ensure stable and high-level expression, bacterial artificial chromosome (BAC) vectors containing an approximately 70-kb long open chromatin region (BAC^{Rosa26}) were used to deliver the transgenes into the cells.

In comparison to conventional expression vectors, the use of a full-length, 210 kb BAC^{Rosa26} expression vector could increase up to nine-fold the yield of heterologous secreted proteins²². It also helped to express high amount of recombinant sodium channels in CHO cells²³. However, transfection of this large DNA was much less efficient compared to shorter plasmids. Therefore we generated a truncated 70 kb version of BAC^{Rosa26} by homologous recombination and FLP recombinase mediated mutagenesis. (Supplementary Fig. 4).

Transfection efficiency and long-term protein production capacity were tested for this truncated 70 kb BAC^{Rosa26} by GFP expression. Chinese hamster ovary (CHO) cells were nucleoporated with a plasmid, original 210 kb BAC^{Rosa26} and truncated 70 kb BAC^{Rosa26} vector all containing CAG promoter, GFP and Neomycin resistance gene as a selection marker. The ratio of GFP positive cells were approximately 3-times higher 24 h post transfection for truncated 70 kb BAC^{Rosa26} transfected cells (12% GFP positive of living cells) than for original 210 kb BAC^{Rosa26} transfected ones (4.1%) while it was less compared to plasmid transfected cells (17.1%) (Supplementary Fig. 4). The same cell pools were cultured for an additional 2 months in the presence of antibiotic selection pressure. Both BAC^{Rosa26} transfected cell pools showed high GFP intensity levels (Supplementary Fig. 4) with 88.3% GFP positive cells for the original BAC^{Rosa26} and 99.3% of the truncated BAC^{Rosa26} while in the plasmid transfected cell pool only 42.7% of the cells showed GFP fluorescence despite of constant selection pressure.

As the truncated 70 kb BAC^{Rosa26} could ensure high protein expression level but combined with more efficient transfection compared to the original version we used this as an expression vector in this study.

The expression of the OR and mCherry proteins was driven by the strong and constitutive CAG promoter²⁴. The remaining two proteins, ORCO and GCaMP6, were expressed as a single fusion protein by linking GCaMP6 N-terminally to ORCO to allow membrane localization. Based on the experience of Corcoran et al.²⁵ constitutive ORCO expression can lead to cell death after several weeks. In order to ensure long-term viability of our OR cell lines, the doxycycline-inducible Ptet-T6 promoter²⁶ was chosen to drive the expression of the GCaMP6-ORCO fusion protein.

To test if the arrangement of genetic elements has an impact on gene expression or not, we cloned and transfected them in two different formats. In the OR13a, OR47a, OR85b, and OR98a cell lines the two expression units were physically integrated tandem (Fig. 1A) into BAC^{Rosa26}. In this arrangement the Ptet-T6 promoter exhibited no baseline expression. Based on binding experiments with the ORCO-activating artificial ligand VUAA1²⁷ there was no GCaMP6-ORCO expression without doxycycline treatment whereas a strong induction was recorded 48 h after doxycycline treatment (Supplementary Fig. 1A). The OR10a, OR49b, and OR71a cell lines, on the other hand, were generated by cotransfection of the same two expression units (Fig. 1A).

In these cell lines, the Ptet-T6 promoter had a substantial baseline activity which was further increased upon doxycycline induction (Supplementary Fig. S1B). In the light of these results, we concluded that tandem integration of the expression units is the method to choose for future cell lines as it prevented uninduced Ptet-T6 promoter activity. Nevertheless, this leaky promoter activity did not hinder the maintenance of these cell lines as their responsiveness was stable during the 60 days monitored (Supplementary Fig. S2). Thus, the previous observation that constitutive ORCO expression was detrimental to long-term cell culture²⁵ could not be confirmed for *Drosophila* OR10a, OR49b, and OR71a + ORCO combinations in our experiments. The OR cell lines may have different sensitivity to constitutive ORCO expression because the ion permeability of the channel depends on the OR identity⁸. Alternatively, components in the culture medium may activate the ORs in some cell lines, but not in others, due to their different ligand profiles. In four stable cell lines (OR7a, OR19a, OR69a, and OR47b) none of the tested VOCs could activate the ORs despite the fact that OR7a, OR19a and OR69a showed strong responses to multiple VOCs according to the DoOR database. Similarly to the seven responding cell lines, the mCherry fluorescence signal was strongly visible indicating the presence of at least the OR transcripts as mCherry was transcribed from the same mRNA as the ORs. All cell lines responded to VUAA1 indicating the presence of functional GCaMP6-ORCO fusion proteins. Lack of response of some ORs in heterologous expression systems was reported in several previous studies^{17,28–32}. It was shown that translation can be a bottleneck in these cases. Functional OR protein expression can be improved by codon optimization of OR genes instead of using wild type sequences³³. In our study, we used a codon-optimized version of ORCO, however, we did not optimize any of the OR genes. OR release from the ER and trafficking to the plasma membrane could have been hindered in non-responding cell lines, too. The use of signal peptide tags might help in these cell lines as it was previously shown that their use could significantly enhance the intensity of the Ca²⁺ response of the *Drosophila* OR47a/ORCO transfected cells in transient expression studies³⁴.

Out of the seven responsive OR cell lines reported here only one, OR13a was stably expressed so far in an insect cell line (Sf21)³⁵. This and other ORs were expressed transiently either by mRNA injection into *Xenopus* oocytes or by transfection into immortalized cell lines: OR13a³⁶, OR10a and OR71a^{37,38}, OR47a^{9,34,39}, OR49b^{39,40}, and OR85b^{41,42}. In contrast, all our stable cell lines can be used for months for direct measurements. Moreover, the cell cultures can be scaled up, then cryopreserved and thawed again at any time for subsequent measurements.

The responsiveness of the OR cell lines to their reference ligands

The functionality of the OR cell lines was verified by the artificial ORCO agonist VUAA1 and at least one OR-specific odorant (Fig. 2). To identify OR-specific reference ligands, two or three strong ligand candidates were chosen based on the DoOR database (<http://neuro.uni-konstanz.de/DoOR/default.html>)²⁰ and tested at 100 μM final concentration. The compound that triggered the highest response was chosen as the reference ligand for the given OR. For four cell lines (OR49b, OR71a, OR85b, and OR98a)—from our plant disease-related ligand list—the strongest agonists were not the ones which were suggested by the DoOR database. We used styrene instead of 2-methylphenol for OR49b, 6-methyl-5-hepten-2-one instead of 4-ethylguaiaicol for OR71a and instead of ethyl benzoate for OR98a, and 2-heptanone instead of butyl acetate for OR85b as reference ligands (Fig. 2B, Supplementary Table 2).

We used the GFP-linked calcium ion indicator GCaMP6, in order to measure the response of the cells as a fluorescence signal (Fig. 1B, C).

The maximum ΔF values measured on the OR cell lines were exceptionally high, ranging between 124 and 462 ΔF% in response to 50 μM VUAA1 and 29–339 ΔF% to their reference ligands during long term stability testing (Supplementary Fig. S2), in contrast to 20–60% previously measured in comparable experimental systems^{25,35,43}. OR expression level of the cells can show some variation between measurement days. To minimize any VOC response differences due to this variation, the measured ΔF values were normalized to average cell line-specific ΔF values in response to 50 μM VUAA1 measured on the same day and were counted as 1. These VUAA1 normalized ΔF values were used to compare the different receptors and ligands.

We tested for the origin of the Ca²⁺ ions on two cell lines, OR47a and (GCaMP6-GSG)ORCO, with assay buffer lacking Ca²⁺ and assay buffer lacking Ca²⁺ and containing EGTA. Neither of the cell lines responded to the ORCO agonist ligand (Supplementary Fig. S3), indicating that the signal is caused by extracellular Ca²⁺.

The time requirement for the measurement of one VOC sample in the seven OR cell lines was 13 min when measured in triplicates (three wells/sample/cell line). It was possible to measure 16 samples within approximately 3.5 h in a whole 384-well microplate.

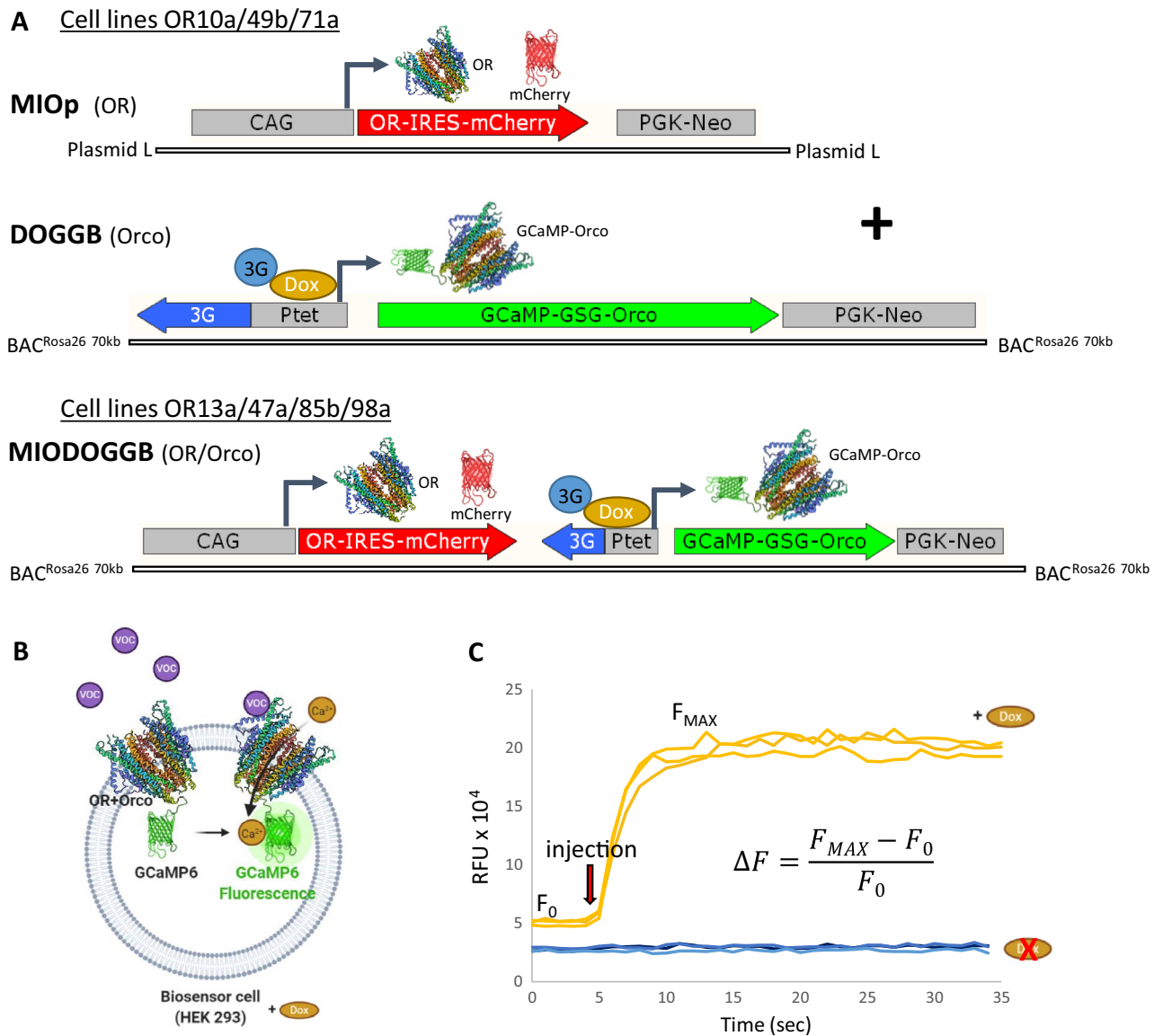


Figure 1. Generation and characterization of OR-expressing cell lines. **(A)** DNA vector constructs for OR cell line production: MIOp plasmid containing the corresponding OR and mCherry, and DOGGB BAC expression vector coding for the GCaMP6-Orco fusion protein were cotransfected to generate OR10a, OR49b and OR71a cell lines (upper panel); MIODOGGB BAC expression vectors containing the corresponding OR and all other components were used to generate OR13a, OR47a, OR85b and OR98a cell lines (lower panel). The inducible Ptet bidirectional promoter only activates GCaMP6-Orco protein expression upon doxycycline (Dox) treatment. **(B)** Detection mechanism of the cell lines: the OR+Orco receptor complex opens upon VOC ligand binding, the resulting Ca^{2+} influx is detectable as GCaMP6-mediated fluorescence. VOC volatile organic compound. **(C)** Fluorescence signal intensity change of the OR47a cell line 48 h after doxycycline induction (orange curves) and without doxycycline (blue curves) in response to 50 μM VUAA1. Each curve represents a technical replicate measured in separate wells of a 384-well microplate. RFU relative fluorescence unit, F_{MAX} maximal RFU value during the measurement, F_0 average baseline fluorescence, ΔF was calculated according to the equation shown.

Comparison of ligand sensitivity with previously published OR-expressing cell lines

The concentration-dependent effect of the reference ligands on their cognate cell lines was examined in more detail. Twelve different ligand concentrations were measured on each cell line in an overall 0.01 μM to 100 μM concentration range with 5 repetitions with three technical replicates each time. The half-maximal effective concentration (EC50), and the maximal response compared to 50 μM VUAA1 were calculated (Fig. 2B).

Based on these parameters, a comparative evaluation of the OR cell lines was performed. There are a number of studies in which HEK293 cells have been used for functional characterization of insect ORs^{25,33,34,44–46}. In these studies, typical EC50 values and ligand sensitivity of the ORs were very similar to our results, but our system shows a much higher fluorescence intensity change (ΔF) in response to VOC ligands.

In an insect cell line stably expressing OR13a, 1-octen-3-ol evoked concentration-dependent Ca^{2+} signals with an EC50 of 4.33 μM ³⁵. The effect of the same ligand has been investigated recently on heterologously expressed

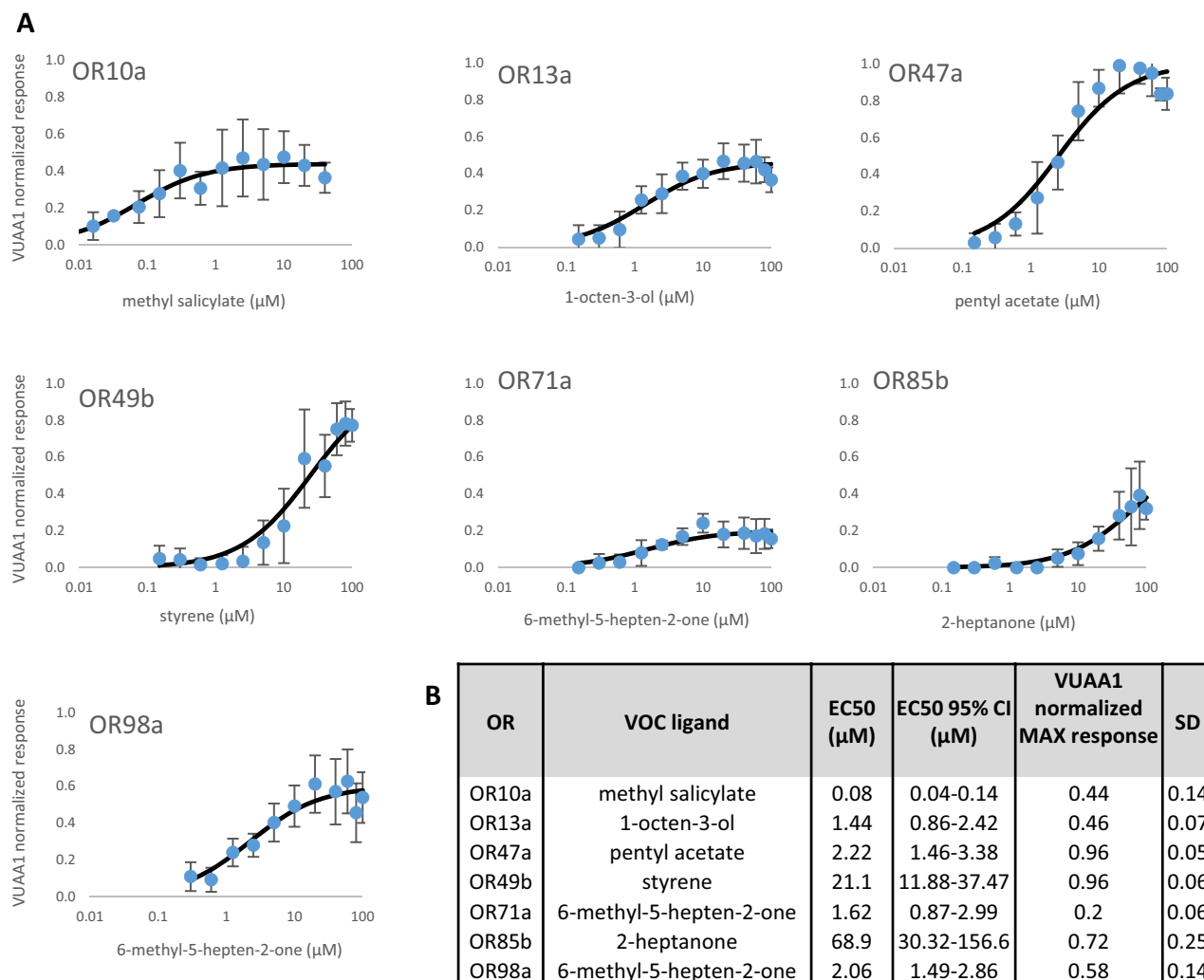


Figure 2. Concentration-dependent response of the OR cell lines to the respective reference VOC ligands. (A) VUAA1-normalized response was calculated by dividing VOC response with the average of VUAA1 response on the day of the measurement for the corresponding OR cell line. Concentration–response curves are shown as the average of 5 measurements, error bars represent the standard error (SD). (B) EC₅₀ values and maximal VUAA1 normalized responses were calculated from the fitted Hill equation, where the Hill coefficient was constrained to 1 and the maximal response was maximum 1. The calculations and the original traces are available in the supplement.

OR13a receptors of a different fruit fly species (*Bactrocera dorsalis*) in HEK293 cells. An EC₅₀ of 11.28 μM in fluorescent measurements, and 1.268 μM in electrophysiology was detected³⁶. Our OR13a expressing cells had similar sensitivity to 1-octen-3-ol with an average EC₅₀ of 1.44 μM (Fig. 2B).

The sensitivity of the OR47a to pentyl acetate varied between different studies. In *Xenopus* oocytes, a range of 50–300 μM could trigger receptor responses⁹, while in a similar system the calculated EC₅₀ was 10.7 ± 2.0 μM³⁹ and 30.8 ± 1.28 μM³⁴. In our OR47a cell line, the EC₅₀ was 2.22 μM for pentyl acetate (Fig. 2B), which indicates about five times higher sensitivity.

For the OR49b cell line, we used styrene as a reference ligand (see "Plant disease-related VOC measurements on the OR cell lines". below) but measured the EC₅₀ for 2-methylphenol (*o*-cresol), too. With 13.95 μM (Supplementary file Data summary.xlsx) of EC₅₀ our OR49b cell line was 17-times more sensitive to this ligand than previously measured in *Xenopus* oocytes in which the calculated EC₅₀ was 239 ± 76 μM³⁹. In a recent study, the same OR was expressed transiently in HEK293T cells, which reacted to 2-methylphenol above the limit of detection in a 2 μM to 800 μM concentration range. Though the EC₅₀ was not calculated, 80 μM concentration triggered a 38.5% response⁴⁰.

The OR85b receptor response to 2-heptanone was measured previously in *Xenopus* oocytes by Nichols and Luetje⁴² with an EC₅₀ of 70 ± 20 μM, and by Misawa et al.⁴¹ with an average EC₅₀ of 45.6 μM compared to the average 68.9 μM value we obtained (Fig. 2B).

Plant disease-related VOC measurements on the OR cell lines

Once responsive OR cell lines were selected, we tested them for sensing fungal pathogens (powdery mildew, *Botrytis cinerea*, *Fusarium* spp., *Pyrenophora* spp., and gray mold) of important crop plants (wheat, barley, grape, lettuce, rape, spinach, and strawberry). To this end, we had compiled a pathogen-related VOC catalog primarily based on our GC–MS analyses of infected samples (published results¹⁸ and unpublished data of our laboratory) and also systematically retrieved from published literature data^{47–50}.

Together with strong ligands for the seven ORs according to the DoOR database, a set of 66 compounds from our compilation was measured on all OR cell lines in 1 μ M, 10 μ M, and 100 μ M concentrations (Fig. 3, Supplementary Table S2).

Among the 1386 VOC concentration-OR combinations measured, 156 (11%) gave positive odorant responses. As the reference ligand (6-methyl-5-hepten-2-one) with the lowest overall VUAA1-normalized response elicited a 0.2 VUAA1-normalized response (OR71a, Fig. 2), this value was taken as the threshold to separate strong/medium reactions from low intensity ones. The majority (60.8%) of the responses were of low intensity, i.e., less than this 20% threshold of the VUAA1-normalized response. Reactions with strong (at least 40% of the VUAA1 response) or medium (20–40% of the VUAA1 response) intensity were recorded in 17% and 22.2% of all positive cases, respectively.

In 11 (2.4%) of the 462 VOC-OR combinations tested, the OR cell lines responded over the whole concentration range measured. High sensitivity to a ligand at the lowest concentration did not always guarantee the typical logistic relationship between ligand concentration and response intensity. Several odorants only activated ORs at the medium concentration, with moderate increase in response at higher concentrations. As an example, isopentyl acetate first activated the OR47a cell line at 10 μ M and increasing the ligand concentration to 100 μ M increased the response intensity from 12 to 18% (Supplementary Table S2). Some cell lines (and their ORs) showed a well-defined affinity for structurally similar compounds. For instance, OR49b was highly specific for aromatic compounds, as only these elicited strong or medium responses, with the exception of butyl acetate. Figure 3 summarizes the VOC ligand profiles of the OR cell lines and the chemical structures of strong and medium ligands.

We have a special focus on VOCs induced by wheat powdery mildew infections. Quantitative experimental data from consecutive years have therefore been collected in open field plots as well as from controlled inoculations of several cultivars. Six powdery mildew-specific volatile biomarkers have been identified and quantified in the headspace of healthy and diseased wheat plants¹⁸. Four of these biomarker VOCs were commercially available and three of them, 1-octen-3-ol, (5Z)-octa-1,5-dien-3-ol, and 3-octanone were in the detectable range of two OR cell lines in samples from diseased wheat plants (Table 1), while the concentration of 1-heptanol was below the detection limit of the OR13a cell line.

Comparison of the OR cell line ligand profiles with the DoOR database

The DoOR database-derived response profiles, which are collections of in vivo measured datasets, served as a good starting point for the identification of candidate ORs and the prediction of their response profiles to our VOC ligands. The DoOR database sets the scale by comparing the measured data to each other, while we took the maximum response value for each OR induced by VUAA1 as 1, and compared the effect of all other ligands to this. This essentially creates a database with a similar response profile ranging from 0 to 1, but this way of calculating response magnitudes is clearly different from the data in the DoOR database, therefore response magnitude values cannot be directly compared. Nevertheless, it is an important issue to investigate how much functional responses in an expression system correspond with in vivo data.

Twenty-two out of the 66 VOCs tested in the present study have not been tested in the DoOR database, so receptor responses to these VOCs represent new information. OR10a was the best-studied receptor for our ligand set; the DoOR database contained response values for 41 VOCs, while OR71a was the least studied with only 21 previously measured VOCs. The OR10a, OR47a, and OR49b cell lines responded at 100 μ M to all VOCs that were predicted as a strong ligand (at least 0.4 response) for these ORs according to the DoOR database (5/5, 8/8 and 2/2, resp. in Supplementary Table S2). Similarly, a high matching rate (11/15, 73%) was also found for the OR98a cell line. However, for three cell lines, OR13a, OR85b and OR71a the respective success rates were only 57% (4/7), 30% (3/10), and 25% (1/4). These observations indicate the importance of testing the ligand profile and sensitivity in different expression systems, and emphasize the potential role of cellular environment in determining sensitivity profiles of individual ORs.

Concluding remarks

We generated responsive OR cell lines which stably express seven functional *Drosophila* ORs together with GCaMP6, a fluorescent calcium ion indicator protein. These cell lines were produced by transfection of HEK293 cells with a novel BAC expression vector carrying an open chromatin region to ensure higher and more stable expression levels in a long term. The OR cell lines detected odorant molecules from the ppb concentration in the liquid phase. The response was quick with the cells reaching maximum fluorescence intensity change within 10 s. Peaks of fluorescence intensity change were as high as 462% for the ORCO agonist VUAA1 and 339% for reference ligands, which is outstanding compared to the published data of ca. 50%. The concentration–response profiles of the OR cell lines were mapped for 66 plant- and plant pathogen-derived VOCs over a concentration range of three orders of magnitude. We recorded ligand-specific fluorescence responses in 11% of all 1386 measured combinations of VOC concentrations and OR cell lines, similarly to the 17% described by Hallem et al.¹⁷. In terms of plant pathogen-related VOC sensing, three major volatile biomarkers specific for powdery mildew infection were in detectable concentration in diseased wheat samples¹⁸ for the OR13a and OR47a cell lines. The OR proteins expressed here can be purified from the cells and coupled to other signal-transducing elements. For

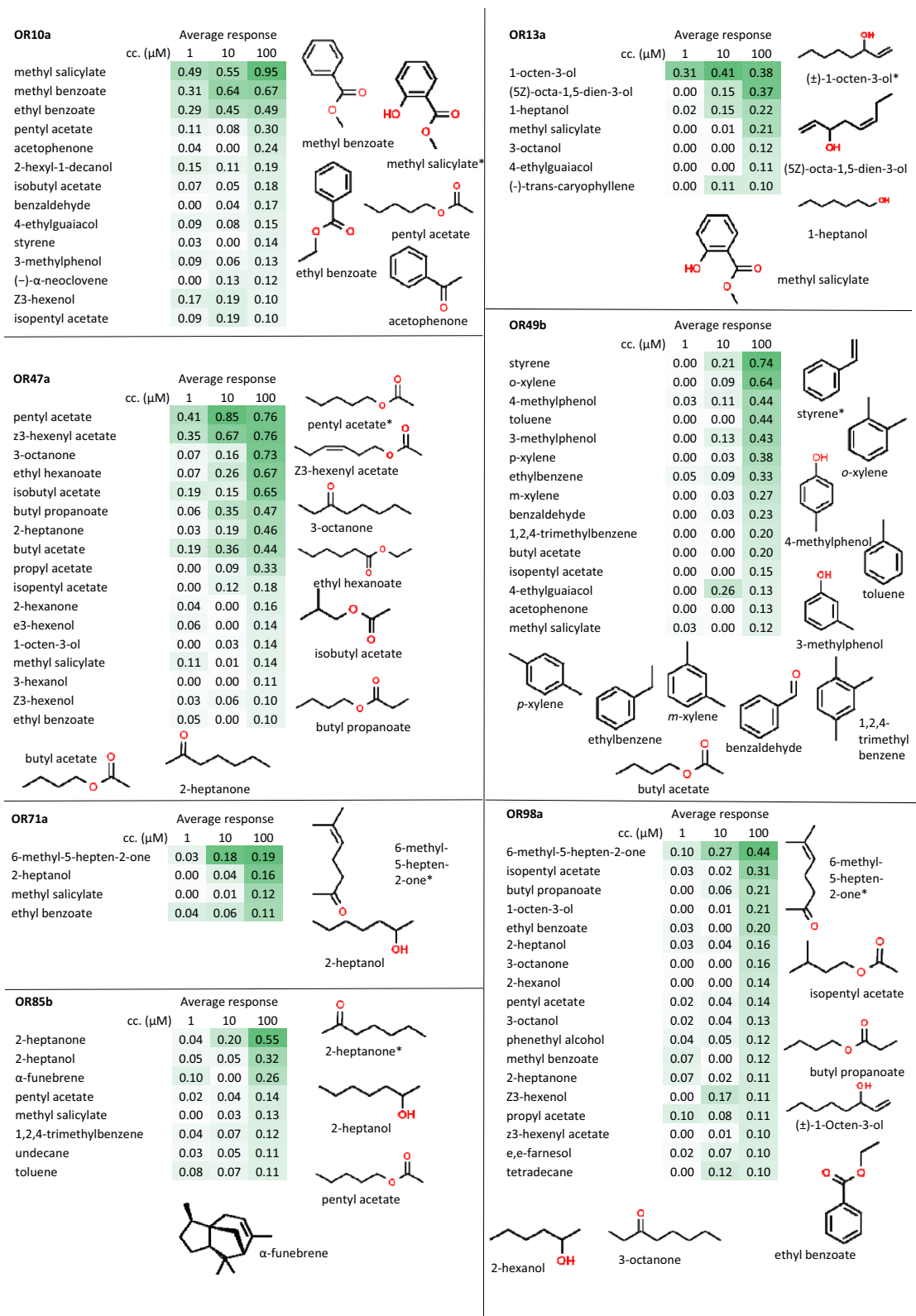


Figure 3. VOC ligand profile of the OR cell lines. Each VOC was measured at 1, 10, and 100 μM final concentration on each receptor. VUAA1-normalized response was calculated as described in methods. Values are averages of at least three biological replicates for responsive combinations and two for the non-responsive combinations. The chemical structures of the strong and medium ligands are shown for each cell line. Heat map: white-green colors correspond to the lowest-highest values measured for each cell line.

Biomarker VOC	VOC concentration of DMSO samples (μM)		EC50	
	Min	Max	OR13a	OR47a
1-Octen-3-ol	0.30	79.24	1.44	
(5Z)-Octa-1,5-dien-3-ol	0.16	93.64	3.05	
3-Octanone	0.21	12.48		2.23

Table 1. Concentration range and detectability of major wheat powdery mildew biomarker VOCs by two OR cell lines.

instance, systems like the ones developed earlier^{37,51} may be utilized for our assay too, and such coupling steps could further improve sensitivity for the ligands explored here.

Materials and methods

Molecular cloning of OR genes into plasmid and BAC expression vectors

The OR genes OR7a (FlyBase ID FBtr0071186), OR47a (FBtr0088111), and a codon-optimized version of ORCO (FBtr0113193) were synthesized (Thermo Fisher Scientific) as double-stranded linear gene fragments. The cDNAs of the OR10a, OR13a, OR19a, OR47b, OR49b, OR67b, OR69a, OR71a, OR85b, OR98a genes were cloned from the wild-type Canton S strain of *D. melanogaster*. Total RNA was extracted from adult flies by TRIzol Reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. Total RNA (1 μg) was treated with DNase I, then reverse transcribed with oligo(dT)_{12–18} primers and SuperScript IV enzyme (Invitrogen). The OR genes were amplified with gene-specific primers containing an *AscI* recognition sequence on both primers (except *SwaI* for OR19a) and a Kozak consensus sequence on the forward primers (Supplementary Table S1).

The above 12 OR genes linked to a wild-type internal ribosome entry site from the encephalomyocarditis virus (EMCV IRES) and the *mCherry* fluorescent protein gene⁵² were inserted into a modified pBlueScript KS (Agilent/Stratagene) called plasmid “L”²². The OR genes were inserted into this vector at the *AscI* (*SwaI*) restriction site, while IRES-*mCherry* at the *PmeI* restriction site and the resulting plasmids were named as MIO(receptor name)p. The N-terminus of ORCO was fused to GCaMP6²¹ via a Gly₂SerGly₃SerGly linker. Subsequently, this fusion gene was cloned into a modified plasmid “L” at the *AscI* restriction site, in which the CAG promoter²⁴ was replaced by a Doxycycline-inducible Ptet-T6 bidirectional promoter²⁶ and its cognate transactivator protein (3G) gene (derived from AAVS1_Puro_Tet3G_3xFLAG_Twin_Strep, a gift from Yannick Doyon: Addgene plasmid #92099⁵³). The resulting plasmid was named as DOGGp in this work. MIO13a/47a/85b/98aDOGGp plasmids were generated by ligating *Sall*- and *SrfI*-digested MIOp with *Bsu36I*-digested DOGGp (Fig. 1A). All enzymes were purchased from Thermo Fisher Scientific (except for *SrfI*, which was from New England Biolabs). Cloning was performed in the *E. coli* DH10B strain. All plasmid sequences are available upon request.

For the generation of the BAC expression vectors, a modified version of the murine *Rosa26* BAC DNA (BAC PAC Resources Children's Hospital Oakland Research Institute clone number: RP24–85I15) was used as the vector backbone. Briefly, the original 210-kb BAC was shortened by removing approximately 70–70 kb DNA sequences from both ends of the genomic insert. First, two targeting constructs carrying 50 bp long homologous sequence regions (HR) on both ends (5' targeting construct: HR–*frt*–Kanamycin–*frt*–HR, 3' targeting constructs: HR–*frt3*–Ampicillin–*frt3*–HR) were generated by PCR (primers in Supplementary Table S1). The 5' targeting construct was recombined into the targeted region replacing the original sequence by ET-cloning/recombineering^{54,55}, then the cassette was excised by FLPe recombinase⁵⁶. Subsequently, recombination and cassette excision were repeated with the 3' targeting construct.

For ET-cloning DH10B *E. coli* harboring the original 210-kb BAC^{Rosa26} were electroporated with the temperature sensitive pSC101-BAD-gbaA (pRed/ET plasmid on Supplementary Fig. 4) plasmid which carries the recombinase proteins necessary for homologous recombination. Cells harboring the pSC101-BAD-gbaA plasmid were selected in tetracycline (5 $\mu\text{g}/\text{ml}$) at 30 °C overnight. Bacterial cells derived from one single positive colony were cultured overnight at 30 °C and transferred to 50 ml of fresh medium next day. At an optical density (OD600) of 0.2, the expression of the recombinogenic proteins was induced by the addition of L-arabinose (to a 0.3–0.4% final concentration) and by shifting the temperature to 37 °C. After one additional hour, cells were harvested and electrocompetent cells were prepared by a double wash with ice cold distilled water. Cells were resuspended in 10% ice cold glycerol solution and aliquoted to Eppendorf tubes (100 μl) prior snap freezing in liquid N₂ or directly used for electroporation. 100 ng of linear DNA was electroporated to these cells with a Gene Pulser Xcell (BioRad) electroporator. Cells were incubated for 70 min at 37 °C without antibiotics in LB medium and plated on LB-agar containing the respective antibiotic (15 $\mu\text{g}/\text{ml}$ kanamycin after 5' targeting construct recombination and 100 $\mu\text{g}/\text{ml}$ ampicillin after 3' targeting construct recombination) at 37 °C overnight.

In the next step, bacterial cells derived from one single positive colony were cultured overnight at 37 °C in LB/antibiotic (ampicillin/kanamycin). The next day electrocompetent cells were prepared the same way as for ET-cloning and were electroporated with the expression plasmid 706 FLP which expresses the FLPe recombinase⁵⁶ to excise the antibiotic cassette flanked by *frt*/*frt3* sites in the 5'/3' targeting constructs. Cells were resuspended in LB medium without antibiotics after electroporation and incubated for 70 min at 30 °C. 100 μl of cell suspension was plated into LB-agar containing tetracycline (3 $\mu\text{g}/\text{ml}$) and the appropriate antibiotics for the BAC (chloramphenicol 12.5 $\mu\text{g}/\text{ml}$) and incubated overnight at 30 °C. The plate was incubated at 37 °C for 4–6 h. 10 single colonies were picked and streaked out to another LB agar/chloramphenicol plate from the left to the right, each in one row and cultured 37 °C overnight. Single colonies were picked from this plate and we plate them on

replica plates either containing only chloramphenicol or chloramphenicol and the antibiotic derived from 5'/3' targeting constructs (ampicillin/kanamycin). Some of the colonies will not grow on ampicillin/kanamycin plates anymore, here the recombination was successful.

After recombination and cassette excision was complete with the 5' targeting construct it was repeated the same way with the 3' targeting construct (see experimental outline in Supplementary Fig. 4).

DOGp and MIO13a/47a/85b/98aDOGp plasmids were linearized by *Sfa*AI and *Pac*I restriction enzymes to generate the linear fragments required for BAC recombination into the 2nd exon of the *Rosa26* gene by ET-cloning/Recombineering as described in detail²². The obtained BAC vectors were designated as DOGGB and MIO13a/47a/85b/98aDOGGB, respectively (Fig. 1A).

Cell culture and stable cell line generation

HEK293 cells (ATCC CRL-1573™) were cultured in growth medium (Dulbecco's Modified Eagle Medium with high glucose, L-glutamine, and sodium pyruvate supplemented with 10% tetracycline-free Fetal Bovine Serum from Biosera and penicillin–streptomycin mixture). Cells were kept at 5% CO₂ and 37 °C in a humidified incubator and were passaged two–three times a week. Transfection-grade BAC DNA was purified with the NucleoBond Xtra BAC DNA purification kit (Macherey–Nagel) and was stored at 4 °C. Plasmid DNA was purified with Genopure Plasmid Midi Kit (Roche). In all cases, a total amount of 9.375 µg DNA was transfected into HEK293 cells reaching 60–70% confluency, in T-25 flasks with FuGene HD transfection reagent (Promega) at a 1:3 DNA:FuGene HD ratio. G418 antibiotic (Thermo Fisher Scientific) was added to the cells 48 h after transfection at 100 µg/mL concentration. For stable cell pool generation, cells were cultured for two weeks, then the antibiotic concentration was increased to 150 µg/mL and cells were cultured for one more week or until they divided vigorously without any visible sign of cell death.

Actively growing transfected cell pools were trypsinized and resuspended in 1 mL growth medium for FACS sorting. The OR13a/47a/85b/98a cell lines were sorted for mCherry fluorescence without doxycycline addition, while the OR10a/49b/71a cell lines were sorted 48 h after doxycycline addition for GFP and mCherry fluorescence with a FACSAria III cell sorter (BD Biosciences) into 96-well cell culture microplates. Clonal cell lines were cultured in growth medium supplemented with 150 µg/mL G418 and observed under an Eclipse Ti2 inverted fluorescence microscope (Nikon) when starting to grow in the 96-well microplates. Four to six actively growing mCherry positive cell clones were selected and expanded for testing their VOC response. Odorant-responsive clonal cell lines were selected and cryopreserved for subsequent experiments.

Calcium ion fluorescence assay

Twelve thousand to 18,000 cells were plated into each well of a black wall, transparent bottom 384-well microplate (Greiner) 48 h or 72 h before the measurements and were simultaneously treated with 1 µg/mL doxycycline. Typically, three columns were filled with one OR expressing cell line, each row was treated with the same compound (three technical replicates for each OR), and each responsive VOC-OR combination was measured at least three times on different days. The plate has 24 columns, thus we could record responses in up to eight different cell lines in each plate. On the day of the experiment, the growth medium was replaced with 20 µL assay buffer/well (140 mM NaCl, 5 mM KCl, 2 mM CaCl₂, 1 mM MgCl₂, 5 mM HEPES-Na, 20 mM glucose, pH 7.3, 340 mOsm). The fluorescence was excited at 485 nm and the emission was recorded at 535 nm with a SpectraMax iD3 multi-mode microplate reader (Molecular Devices) mounted with a two-channel injector. Baseline fluorescence (F₀) was recorded for 5 s, after which 7 µL VOC buffer (140 mM NaCl, 5 mM KCl, 2 mM CaCl₂, 1 mM MgCl₂, 5 mM HEPES-Na, 31 mM glucose, 0.1% DMSO, pH 7.3, 340 mOsm) was injected into each well in one row and the fluorescence response was recorded for an additional 30 s, at a rate of one read/sec. VOC solutions (VOC buffer with 4× concentration of the final measured VOC concentration, 9 µL/well was injected) were measured on the same wells in a second round with the same protocol. Each individual concentration of a specific VOC sample was recorded in three wells for the same cell line on the same microplate. The raw data were exported in xls file format, F₀ and F_{MAX} values, and the fluorescence intensity changes (ΔF) were calculated in Excel (Microsoft) (Supplementary file Data summary.xlsx) using the following equation: $\Delta F = \frac{F_{MAX} - F_0}{F_0}$ (see also Fig. 1C). On each plate, we recorded the response of each cell line to 50 µM VUAA1. The average of these ΔF values of each cell line were used to calculate the VUAA1-normalized values for the rest of the plate, by dividing VOC response ΔF values by the average VUAA1 response ΔF value of the corresponding OR cell line. Each VOC-OR combination was measured three times as technical replicates. The outliers were removed based on the suggestion of Grubbs outlier test. The average of these replicates were used as one ΔF and VUAA1-normalized ΔF in the following analysis.

In the next rows we tested the reference VOCs (or other VOCs known to activate the corresponding cell lines), in order to test if all cell lines are responsive. The remaining rows were used to test VOCs with unknown effect.

The recorded F_{MAX}, F₀, ΔF, VUAA1-normalized ΔF values, as well as the OR name, the name of the tested sample, and the final concentration for each well were copied into supplementary file Data summary.xlsx, that was used to summarize all the plates recorded. Here, the averages of the measurements were calculated for each receptor-VOC-concentration combination. Responses where the VUAA1-normalized ΔF values did not exceed 0.05 were considered as 0.

Chemicals and chemical dilutions

All VOCs were purchased from Sigma-Aldrich (except for (5Z)-octa-1,5-dien-3-ol from Toronto Research Chemicals) in the highest available purity. VOCs were diluted either directly in the VOC dilution buffer or first in DMSO for water-insoluble compounds.

Data availability

The datasets generated in the current study are available in our local computers. Access to the datasets can be granted upon reasonable request. Researchers interested in accessing the data may contact the corresponding author for inquiries and requests.

Received: 12 June 2023; Accepted: 14 November 2023

Published online: 08 December 2023

References

- Glatz, R. & Bailey-Hill, K. Mimicking nature's noses: From receptor deorphaning to olfactory biosensing. *Prog. Neurobiol.* **93**, 270–296 (2011).
- Wilson, A. D. Recent progress in the design and clinical development of electronic-nose technologies. *Nanobiosens. Dis. Diagn.* **5**, 15–27 (2016).
- Kato, A. & Touhara, K. Mammalian olfactory receptors: Pharmacology, G protein coupling and desensitization. *Cell. Mol. Life Sci.* **66**, 3743–3753 (2009).
- Peterlin, Z., Firestein, S. & Rogers, M. E. The state of the art of odorant receptor deorphanization: A report from the orphanage. *J. Gen. Physiol.* **143**, 527–542 (2014).
- Carey, A. F. & Carlson, J. R. Insect olfaction from model systems to disease control. *Proc. Natl. Acad. Sci.* **108**, 12987–12995 (2011).
- Liu, Y., Chen, Q., Man, Y. H. & Wu, W. J. Insect olfactory receptors as essential detectors for volatile chemicals in biomimetic odorant sensors. *Appl. Mech. Mater.* **461**, 822–828 (2014).
- Vosshall, L. B. & Hansson, B. S. A unified nomenclature system for the insect olfactory coreceptor. *Chem. Senses* **36**, 497–498 (2011).
- Butterwick, J. A. *et al.* Cryo-EM structure of the insect olfactory receptor Orco. *Nature* **560**, 447–452 (2018).
- Sato, K. *et al.* Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002–1006 (2008).
- Wicher, D. *et al.* Drosophila odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**, 1007–1011 (2008).
- Getahun, M. N., Olsson, S. B., Lavista-Llanos, S., Hansson, B. S. & Wicher, D. Insect odorant response sensitivity is tuned by metabotropically autoregulated olfactory receptors. *PLoS One* **8**, e58889 (2013).
- Nakagawa, T. & Vosshall, L. B. Controversy and consensus: Noncanonical signaling mechanisms in the insect olfactory system. *Curr. Opin. Neurobiol.* **19**, 284–292 (2009).
- Deng, Y. *et al.* The stimulatory Gas protein is involved in olfactory signal transduction in Drosophila. *PLoS One* **6**, e18605 (2011).
- Silbering, A. F. & Benton, R. Ionotropic and metabotropic mechanisms in chemoreception: 'Chance or design?'. *EMBO Rep.* **11**, 173–179 (2010).
- Schmidt, H. R. & Benton, R. Molecular mechanisms of olfactory detection in insects: Beyond receptors. *Open Biol.* **10**, 200252 (2020).
- Strauch, M. *et al.* More than apples and oranges—Detecting cancer with a fruit fly's antenna. *Sci. Rep.* **4**, 1–9 (2014).
- Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125**, 143–160 (2006).
- Hamow, K. A. *et al.* Emission of novel volatile biomarkers for wheat powdery mildew. *Sci. Total Environ.* **781**, 146767 (2021).
- Galizia, C. G. & Rössler, W. Parallel olfactory systems in insects: Anatomy and function. *Annu. Rev. Entomol.* **55**, 399–420 (2010).
- Münch, D. & Galizia, C. G. DoOR 2.0—Comprehensive mapping of *Drosophila melanogaster* odorant responses. *Sci. Rep.* **6**, 21841 (2016).
- Ohkura, M. *et al.* Genetically encoded green fluorescent Ca²⁺ indicators with improved detectability for neuronal Ca²⁺ signals. *PLoS One* **7**, e51286 (2012).
- Zboray, K. *et al.* Heterologous protein production using euchromatin-containing expression vectors in mammalian cells. *Nucleic Acids Res.* **43**, e102 (2015).
- Lukacs, P. *et al.* Non-blocking modulation contributes to sodium channel inhibition by a covalently attached photoreactive riluzole analog. *Sci. Rep.* **8**, 1–11 (2018).
- Miyazaki, J. *et al.* Expression vector system based on the chicken β -actin promoter directs efficient production of interleukin-5. *Gene* **79**, 269–277 (1989).
- Corcoran, J. A., Jordan, M. D., Carraher, C. & Newcomb, R. D. A novel method to study insect olfactory receptor function using HEK293 cells. *Insect Biochem. Mol. Biol.* **54**, 22–32 (2014).
- Loew, R., Heinz, N., Hampf, M., Bujard, H. & Gossen, M. Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnol.* **10**, 81 (2010).
- Jones, P. L., Pask, G. M., Rinker, D. C. & Zwiebel, L. J. Functional agonism of insect odorant receptor ion channels. *Proc. Natl. Acad. Sci. USA* **108**, 8821–8825 (2011).
- Wang, G., Carey, A. F., Carlson, J. R. & Zwiebel, L. J. Molecular basis of odor coding in the malaria vector mosquito *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA* **107**, 4418–4423 (2010).
- Andersson, M. N. *et al.* A sex pheromone receptor in the hessian fly *Mayetiola destructor* (Diptera, Cecidomyiidae). *Front. Cell. Neurosci.* **10**, 1–12 (2016).
- de Fouchier, A. *et al.* Functional evolution of Lepidoptera olfactory receptors revealed by deorphanization of a moth repertoire. *Nat. Commun.* **8**, 15709 (2017).
- Carey, A. F., Wang, G., Su, C. Y., Zwiebel, L. J. & Carlson, J. R. Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* **464**, 66–71 (2010).
- Yuvaraj, J. K. *et al.* Characterization of odorant receptors from a non-ditrysian moth, *Eriocrania semipurpurella* sheds light on the origin of sex pheromone receptors in Lepidoptera. *Mol. Biol. Evol.* **34**, 2733–2746 (2017).
- Roberts, R. E., Yuvaraj, J. K. & Andersson, M. N. Codon optimization of insect odorant receptor genes may increase their stable expression for functional characterization in HEK293 cells. *Front. Cell. Neurosci.* **15**, 1–10 (2021).
- Miazi, F. *et al.* Optimization of insect odorant receptor trafficking and functional expression via transient transfection in HEK293 cells. *Chem. Senses* **44**, 673–682 (2019).
- Termtanasombat, M. *et al.* Cell-based odorant sensor array for odor discrimination based on insect odorant receptors. *J. Chem. Ecol.* **42**, 716–724 (2016).
- Xu, L. *et al.* Two odorant receptors regulate 1-octen-3-ol induced oviposition behavior in the oriental fruit fly. *Commun. Biol.* **6**, 1–8 (2023).
- Murugathas, T. *et al.* Biosensing with insect odorant receptor nanodiscs and carbon nanotube field-effect transistors. *ACS Appl. Mater. Interfaces* **11**, 9530–9538 (2019).
- Khadka, R. *et al.* Data on preparation and characterization of an insect odorant receptor based biosensor. *Data Br.* **21**, 2142–2148 (2018).
- Röllecke, K. *et al.* Amiloride derivatives are effective blockers of insect odorant receptors. *Chem. Senses* **38**, 231–236 (2013).

40. Kolesov, D. V. *et al.* Impacts of OrX and cAMP-insensitive Orco to the insect olfactory heteromer activity. *Mol. Biol. Rep.* **48**, 4549–4561 (2021).
41. Misawa, N., Mitsuno, H., Kanzaki, R. & Takeuchi, S. Highly sensitive and selective odorant sensor using living cells expressing insect olfactory receptors. *Proc. Natl. Acad. Sci.* **107**, 15340–15344 (2010).
42. Nichols, A. S. & Luetje, C. W. Transmembrane segment 3 of *Drosophila melanogaster* odorant receptor subunit 85b contributes to ligand–receptor interactions. *J. Biol. Chem.* **285**, 11854–11862 (2010).
43. Mitsuno, H., Sakurai, T., Namiki, S., Mitsuhashi, H. & Kanzaki, R. Novel cell-based odorant sensor elements based on insect odorant receptors. *Biosens. Bioelectron.* **65**, 287–294 (2015).
44. Liu, H. *et al.* Functional analysis of Orco and odorant receptors in odor recognition in *Aedes albopictus*. *Parasit. Vectors* **9**, 363 (2016).
45. Hou, X. *et al.* Functional characterization of odorant receptors from the moth *Eriocrania semipurpurella*: A comparison of results in the xenopus oocyte and HEK cell systems. *Insect Biochem. Mol. Biol.* **117**, 103289 (2020).
46. del Marmol, J., Yedlin, M. A. & Ruta, V. The structural basis of odorant recognition in insect olfactory receptors. *Nature* **597**, 126–131 (2021).
47. Korpi, A., Järnberg, J. & Pasanen, A. L. Microbial volatile organic compounds. *Crit. Rev. Toxicol.* **39**, 139–193 (2009).
48. Khaleel, C., Tabanca, N. & Buchbauer, G. α -Terpineol, a natural monoterpene: A review of its biological properties. *Open Chem.* **16**, 349–361 (2018).
49. Jansen, R. M. *et al.* Detection of diseased plants by analysis of volatile organic compound emission. *Annu. Rev. Phytopathol.* **49**, 157–174 (2011).
50. Lemfack, M. C., Nickel, J., Dunkel, M., Preissner, R. & Piechulla, B. mVOC: A database of microbial volatiles. *Nucleic Acids Res.* **42**, D744–D748 (2014).
51. Yamada, T. *et al.* Highly sensitive VOC detectors using insect olfactory receptors reconstituted into lipid bilayers. *Sci. Adv.* **7**, eabd2013 (2021).
52. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567–1572 (2004).
53. Dalvai, M. *et al.* A scalable genome-editing-based approach for mapping multiprotein complexes in human cells. *Cell Rep.* **13**, 621–633 (2015).
54. Muyrers, J. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res.* **27**, 1555–1557 (1999).
55. Zhang, Y., Buchholz, F., Muyrers, J. P. P. & Stewart, A. F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* **20**, 123–128 (1998).
56. Buchholz, F., Angrand, P.-O. & Stewart, A. F. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat. Biotechnol.* **16**, 657–662 (1998).

Acknowledgements

This work was supported by the Economic Development and Innovation Operational Program (Grant No. GINOP-2.3.2-15-2016-00051) (K.Z., A.V.T., T.M., Á.S., K.P., AM, L.S., P.L.), the Hungarian Scientific Research fund (PD 143055), the Austrian Science Fund (FWF) Doc Fund DOC59-B33 (E.C.), the FWF standalone grants P 33430 and P 32900 (E.C.), and by the City of Vienna Fund for innovative, interdisciplinary Cancer Research (E.C.). Part of this work was done in collaboration with Tetralab Ltd., Budapest. Cell sorting was conducted by György Várady at the Laboratory of Flow Cytometry of the ELKH Research Centre for Natural Sciences (Budapest). We thank Bettina Menich and Zainab Quddoos for technical assistance in cell maintenance and microplate reader measurements as well as Orsolya Németh and Anikó Keszöcze for help in OR cloning.

Author contributions

K.Z., E.C., A.M. and P.L. conceived the experiments, K.Z., A.V.T., T.D.M., E.K., K.P. and P.L. conducted the experiments, K.Z., Á.S. and P.L. analyzed the data, K.Z., A.M., L.S. and P.L. wrote the manuscript. All authors have read and approved the manuscript.

Competing interests

E.C. holds a patent on the BAC expression system.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-47455-4>.

Correspondence and requests for materials should be addressed to P.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023