



OPEN

More than apples and oranges - Detecting cancer with a fruit fly's antenna

SUBJECT AREAS:
OLFACTORY RECEPTORS
BREAST CANCER
OPTICAL IMAGING

Martin Strauch^{1*}, Alja Lüdke^{1*}, Daniel Münch^{1*}, Thomas Laudes¹, C. Giovanni Galizia^{1,2}, Eugenio Martinelli², Luca Lavra³, Roberto Paolesse⁴, Alessandra Ulivieri², Alexandro Catini², Rosamaria Capuano² & Corrado Di Natale²

Received
1 October 2013

Accepted
6 December 2013

Published
6 January 2014

Correspondence and
requests for materials
should be addressed to
C.G.G. (Giovanni.
Galizia@uni-konstanz.
de)

* These authors
contributed equally to
this work.

¹Neurobiology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany, ²Department of Electronic Engineering, University of Rome Tor Vergata, Via del Politecnico 1, 00133 Roma, Italy, ³Department of Clinical and Molecular Medicine, University of Rome La Sapienza, Via di Grottarossa 1035, 00189 Roma, Italy, ⁴Department of Chemistry, University of Rome Tor Vergata, Via della Ricerca Scientifica, 00133 Roma, Italy.

Cancer cells and non-cancer cells differ in their metabolism and they emit distinct volatile compound profiles, allowing to recognise cancer cells by their scent. Insect odorant receptors are excellent chemosensors with high sensitivity and a broad receptive range unmatched by current gas sensors. We thus investigated the potential of utilising the fruit fly's olfactory system to detect cancer cells. Using *in vivo* calcium imaging, we recorded an array of olfactory receptor neurons on the fruit fly's antenna. We performed multidimensional analysis of antenna responses, finding that cell volatiles from different cell types lead to characteristic response vectors. The distances between these response vectors are conserved across flies and can be used to discriminate healthy mammary epithelial cells from different types of breast cancer cells. This may expand the repertoire of clinical diagnostics, and it is the first step towards electronic noses equipped with biological sensors, integrating artificial and biological olfaction.

CANCER cells exhibit a metabolism that is fundamentally altered as compared to that of normal cells^{1,2}, leading to changes in the tumor's microenvironment, in lipid peroxidation activity^{3,4} and to a variety of potential intra- and extracellular cancer-specific markers. Thus, quantitative and qualitative variations of metabolites provide important information on cell condition⁵. As a consequence, metabolomics is emerging as a novel diagnostic methodology. The detection of cancer-related marker molecules has been focussed on examining the composition of cell serum or culture media^{6,7}. Since many metabolic products are small molecules, the analysis of volatile organic compounds (VOCs), e.g. in breath, is a promising non-invasive alternative for identifying cancer tissues⁸⁻¹⁰.

Several studies investigated VOC profiles in the headspace of cancer cell cultures¹¹⁻¹³. Traditionally, analyses of cancer cell VOCs have been performed using gas chromatography - mass spectroscopy (GC-MS), the gold standard for *sequential* and *analytic* VOC analysis. In recent years, VOC analysis has been complemented by gas sensor arrays^{11,14,15} which offer a *simultaneous* and *synthetic* representation of the VOC profile in real time. Detection of cancer with gas sensor arrays has been demonstrated both *in vitro* using the headspace of cell cultures^{11,16} and *in vivo* by analysing breath samples^{14,15}.

Gas sensor arrays are also referred to as electronic noses. They resemble natural olfactory systems in that they consist of an array of differentially selective units¹⁷. However, electronic noses are still limited with respect to sensitivity and receptive range¹⁸, a criterion for which they can perform worse than biological systems¹⁹.

The observation that animals, e.g. dogs, can recognise various types of cancer, set off a new wave of experiments^{20,21} (reviewed in^{22,23}). Dogs were able to detect melanoma tissue in skin samples²⁴, bladder cancer²⁵ and prostate cancer²⁶ in urine samples, lung and breast cancer in breath samples^{27,28}, ovarian carcinoma in tissue and blood samples^{29,30}, and colorectal cancer in breath and watery stool samples³¹. These and other applications, where dogs are employed e.g. to detect illicit substances, show how well-suited natural noses are for chemosensing, even if the target odours do not occur in the animal's natural environment. Similarly, fruit flies can distinguish odours that do not belong to their normal habitat³².

Using dogs for chemosensing has the limitation that access to the signals is only indirect, via the animal's behaviour in combination with a human trainer and interpreter. Avoiding a behavioural readout, we here show that the fruit fly's antenna can be accessed directly by calcium imaging, giving rise to a natural chemosensor array that can detect cancer-cell specific volatile profiles.



Calcium imaging of the fruit fly's antenna. The structure of most olfactory systems is similar: Olfactory receptor neurons (ORNs) express a given olfactory receptor (OR) type. In the fruit fly *Drosophila melanogaster* there are about 50 different types of OR expressing ORNs. The expressed type of OR defines the characteristic spectrum of ligands the ORN will respond to³³. A particular odour elicits an ensemble response which consists of the differentially activated, inhibited or non-activated ORNs^{34,35}. This ensemble response can be observed by calcium imaging on the antenna surface of the fly (Fig. 1a) and it is conveyed to higher brain areas for further processing (Fig. 1b).

Calcium imaging relies on the fact that the intracellular Ca^{2+} concentration is correlated with action potential rate and can be used as a proxy for neuronal activity^{36,37}. The cytosolic Ca^{2+} concentration rises mainly through extracellular Ca^{2+} influx via voltage-gated Ca^{2+} - or cation-channels^{36,38,39}, or via influx from intracellular stores³⁹. For *Drosophila* ORNs, influx of extracellular Ca^{2+} via ligand-gated channels is discussed^{40,41} (Fig. 1c). The change in Ca^{2+} concentration can be measured with reporter proteins like GCaMP^{42,37}. GCaMP consists of a Ca^{2+} binding calmodulin domain and a green fluorescent protein. Upon Ca^{2+} binding, GCaMP changes its conformation and fluorescence properties, now emitting more light at 530 nm (Fig. 1c).

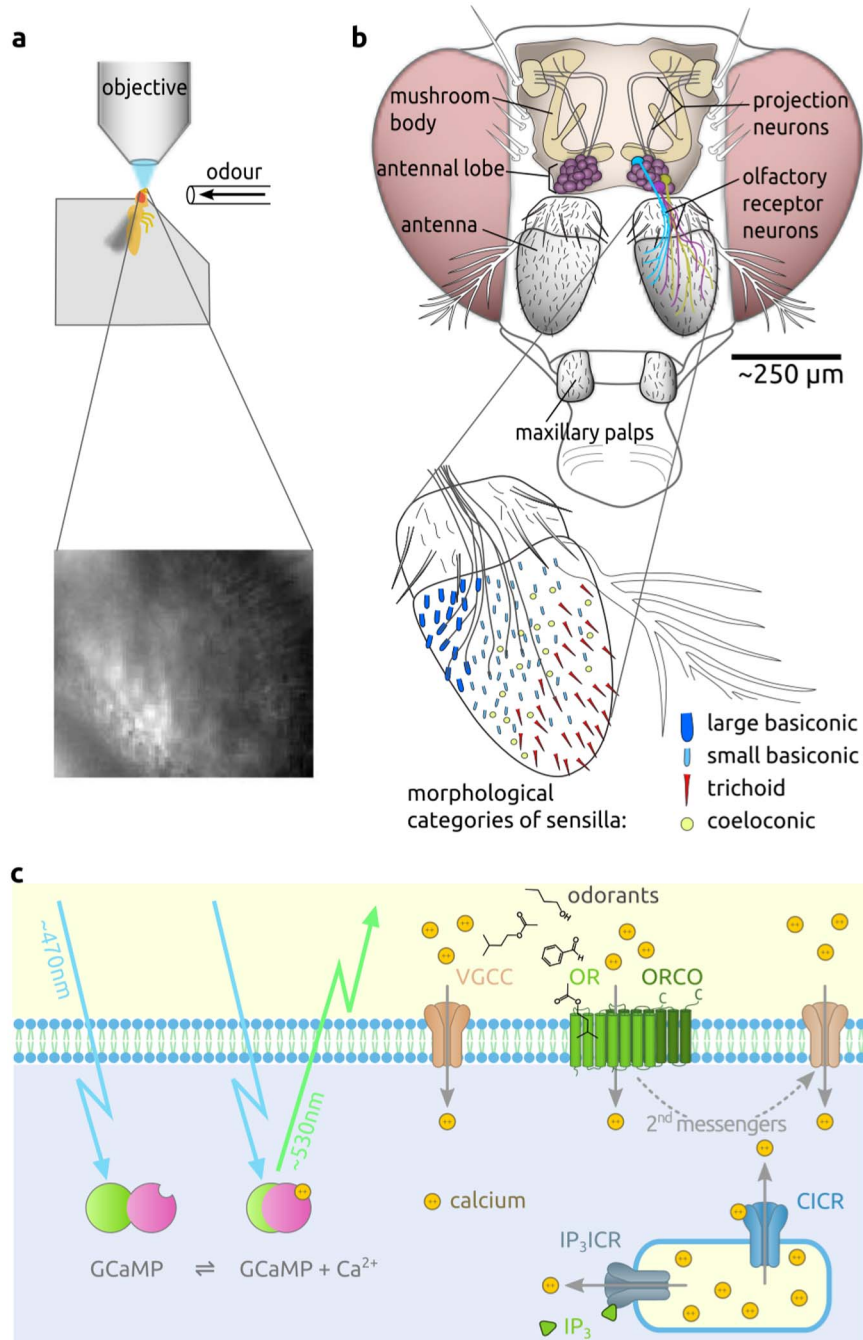


Figure 1 | Calcium imaging of the fly's antenna. (a) *Top:* The antenna of a fixed fruit fly is exposed to fluorescent light. *Bottom:* Fluorescent image of the left antenna of an Orco-GCaMP3 fly. (b) *Top:* The peripheral olfactory organs comprise the antennae and the maxillary palps. Antennae were used in this study. *Bottom:* Morphological categories of sensilla (which house the ORN's dendrites) and their distribution on the third antennal segment. The fluorescent area on the antennal surface in (a) overlaps with the area of large and small basiconic sensilla. (c) Schematic of GCaMP function (left) and Ca^{2+} sources (right). VGCC voltage gated cation channel, OR odorant receptor, ORCO OR co-receptor, CICR Ca^{2+} induced Ca^{2+} release, IP_3/ICR IP_3 (second messenger molecule) induced Ca^{2+} release.



Here, we expressed GCaMP3 in approx. 30 (of the 50) antennal ORN types in *Drosophila melanogaster* by utilising the promoter of the naturally expressed co-receptor Orco in these ORNs⁴³. We could thus record odour evoked activation patterns on the antenna by calcium imaging (see Fig. 1a)⁴⁴.

A sensitive, natural receptor array. Insect odorant receptors are highly sensitive chemosensors^{45–47}. Some studies have accessed the insect olfactory periphery through measuring an antennal compound signal. For instance, the antenna of the Colorado potato beetle *Leptinotarsa decemlineata* was integrated into a field effect transistor (FET)^{48,49} which could resolve different concentrations (in the 10^{-9} order) of Z-3-hexen-1-ol, a green-leaf odour⁴⁸.

FET signals as in^{48,49} and electroantennograms (EAGs) measure summed antennal activity, providing only a single feature. Reading out an array of olfactory receptors would lead to multiple features, enabling odour discrimination as by artificial sensor arrays. In a comparative study¹⁹, a natural *Drosophila* receptor array outperformed a metal oxide semiconductor sensor array, a popular gas sensor type. The fruit fly's natural sensors proved to be less correlated with each other and had a broader receptive range than their artificial counterparts, effectively covering larger parts of odour space. While Berna et al.¹⁹ relied on a collection of single sensillum recordings from the literature to evaluate the *Drosophila* receptor array *in silicio*, Park et al.⁵⁰ developed an *in vivo* sensor array: Simultaneous EAGs from four different insect species resulted in a sensor array where each of the four sensors had a distinct odour response profile.

Here, we show that a large receptor array on the *Drosophila* antenna can be read out *in vivo* by calcium imaging, with a quality that is sufficient to identify cancer cell lines.

Results

Cancer odours elicit responses on the antenna. By selectively expressing GCaMP in all Orco bearing cells we measured odour

responses on the *Drosophila* antenna (Fig. 2). The seven test odours (Methods) were taken from the headspace of the culture media of five different cancer cell line samples (*canc1* to *canc5*), of a healthy control cell sample (*healthy control*) and of a cell free medium sample (*medium control*). The test odours were given as 1 s double pulses and in random order in each fly. Additionally, butanol and N₂ (odourless) were measured as references at the beginning and at the end of each experiment. Odour delivery resulted in strong intracellular Ca²⁺ increases, whereas no response was visible during N₂ applications (Fig. 2). Spatial response patterns differed between odours, but were reproducible for repeated stimulation with the same odour (Supplementary Fig. S1).

Example spatial response patterns for *healthy control*, *canc2* and *canc1* are shown in Fig. 2 (complete odour set: see Supplementary Fig. S1). All three response patterns differed clearly from the N₂ control. The cancer odours exhibited a spatially broader and higher amplitude response than *healthy control* (Fig. 2).

Formalising the visual impression, we computed distances between odour response patterns based on a subset of the pixels: As noisy or non-responding pixels can obscure odour distances, we employed unsupervised feature selection (Methods) to select a set of 300 informative response spots on the antenna of each fly. A response pattern then comprises the activity values of all response spots in the same fly at a particular point in time. Fig. 3a visualises the positions of such response spots on one antenna (same fly as in Fig. 2). Thus, we reduced the CCD-camera readout of 80×60 pixels down to 300 representative spots. Each spot represents a place on the antenna, and thus the response of an unknown number of ORNs. Because ORNs are not uniformly distributed on the antenna, responses differed for different locations.

Next, we clustered the response spots based on response time-series similarity (Methods). Fig. 3a reveals several, spatially contiguous clusters on the surface of the antenna. Several clusters exhibited responses to all odours, with slight differences in response amplitude.

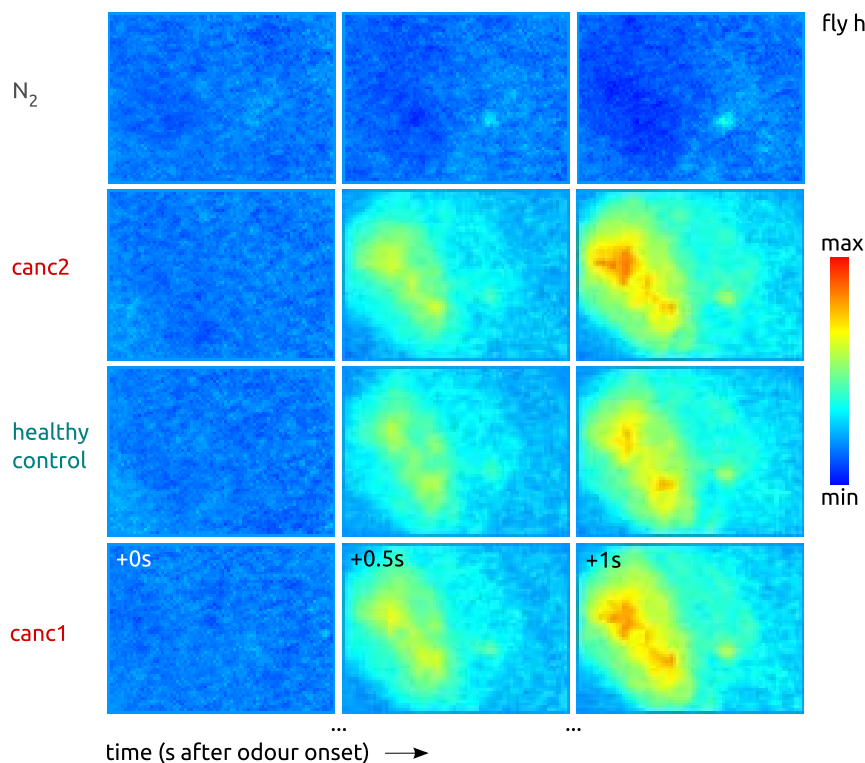


Figure 2 | Spatial odour response patterns on the antenna differ for different cancer cell lines: Fluorescence changes $((F - F_0)/F_0)$ in response to odour stimulation. Images are spatially filtered with a Gaussian kernel (width = 3). The colour scale is min-max (with respect to the entire movie: see Supplementary Fig. S1).

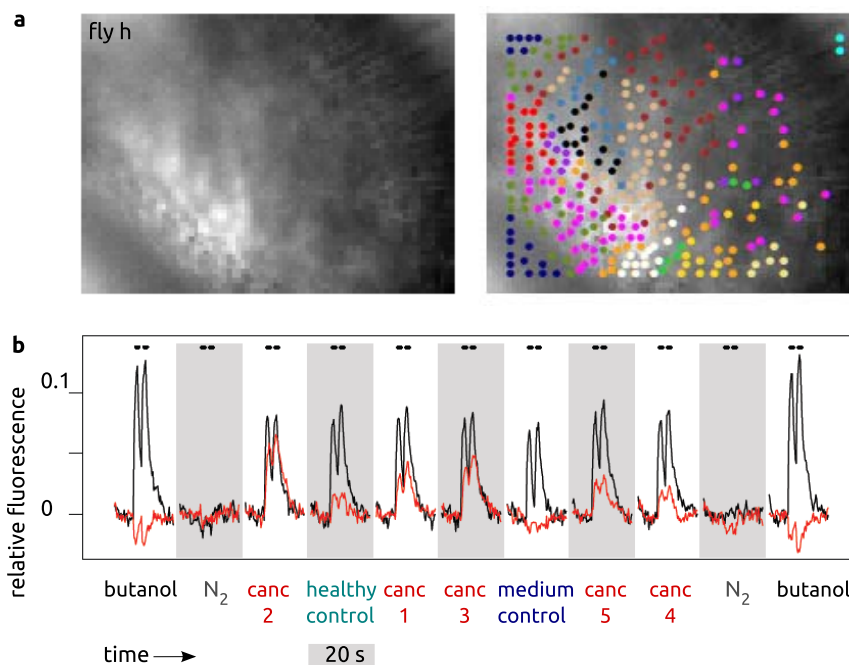


Figure 3 | Response spots form spatially contiguous clusters with common response properties. (a) Left: Raw fluorescence image showing the *Drosophila* antenna. Right: Clustering overlaid (see Methods). (b) Mean time series $((F - F_0)/F_0)$ of the black and red clusters from (a). Note that the black cluster responds to butanol and all test odours, while the red cluster responds to cancer cells, but not to clean medium or butanol. The sequence of odour application (double pulses, 1 s each) is indicated on the x-axis. In the experiments, odours were separated by a pause of one minute. The black bars at the top mark the time points (5 for each response peak) selected for further analysis.

Others showed distinguished, odour-specific responses. For example, in Fig. 3b the response spots in the black cluster exhibited strong responses to all stimuli except for the N_2 control, while the spots in the red cluster showed higher responses to the cancer cell samples than to *healthy control* and *medium control*.

Response spot positions are roughly conserved between flies. We next compared responses across flies. To this end, for each measurement we used the time windows where responses peaked (frames 30–34 and 40–44; marked as black bars in Fig. 3b). These values, across odour responses, were z-score normalised to correct for animal-specific differences in signal amplitude and variability (Methods), resulting in what we now call “response profiles”.

We clustered these response profiles across all flies (Fig. 4). Within each fly, the resulting clusters were spatially contiguous, confirming the findings above (Fig. 3a). Every fly contributed to most clusters, and the approximate spatial location of these clusters was conserved across flies (Fig. 4a, s.a. Supplementary Fig. S2).

Clusters differed in their odour response profiles. For example, cluster 11 exhibited stronger responses to the reference odour butanol than to the test odours, while cluster 12 responded to the cancer odours but less to *healthy control* (Fig. 4b, c). Similarly, the response profiles of cluster 5 and 12 WERE complementary, while cluster 13 RESPONDED to all test odours, but more strongly to the test odours than to the butanol reference (Fig. 4d). While distances between odour response patterns are based on many response spots, already a single, spatially confined region on the antenna (e.g. cluster 12) might contribute significantly to a cancer detection task.

Note that the clustering was performed on response profiles, i.e. using odour response information only. The spatial aspects, contiguity within a fly and conserved position across different flies, emerged solely as a consequence of response profile similarity. The odour set employed in the experiments has induced a functional segmentation of the antenna surface: Control and cancer odours lead to distinguished and reproducible odour responses on the antenna with clusters of response spots that are functionally similar across animals.

Odour distances are strongly conserved between flies. We investigated whether distances between odour response profiles allowed for discrimination of *healthy control* and cancer odours. For a single antenna, Fig. 5a visualises the distances between odour response profiles by a two-dimensional PCA projection (Methods): The high-amplitude responses to butanol and the (odourless) N_2 ended up at opposite sides of the 2D space, with the seven test odours in between, approximately aligned along an axis.

For Fig. 5b, PCA was computed on a matrix containing only responses to the seven test odours: Uninfluenced by the strong amplitude difference between the butanol and N_2 responses, the two-dimensional projection better resolved the smaller distances between the test odours, leading to a clear separation of the two control odours (*medium control* and *healthy control*) from the cancer odours (*canc1* to *canc5*). We obtained comparable results on pooled data, considering all response spots from all antennae (Supplementary Fig. S3).

We further assessed dimensionality of the odour responses by scree plots (Fig. 5c,d), showing that 10 or 6 dimensions, respectively, contributed to the variance in the odour responses. How significant are these differences across animals? If the observed distances are based on characteristic odour responses, correlation should be high at odour response and low otherwise, while variability across animals should be low at odour response and high otherwise. This is indeed what we found: Based on the signals of the $c = 300$ response spots in each animal and for each time point of the measurements, we calculated the 7×7 Euclidean distance matrices between the seven test odours. These were then correlated (Methods) to quantify odour distance conservation across animals (Fig. 5e). We found two pronounced peaks in mean correlation that coincided with the two odour pulses. Maximum mean Pearson correlations were greater than 0.9 (Fig. 5e).

Next, we quantified the distances (using full-dimensional response profiles) of *healthy control* and of the cancer odours to *medium control*. *Healthy control* consistently had the smallest distance to

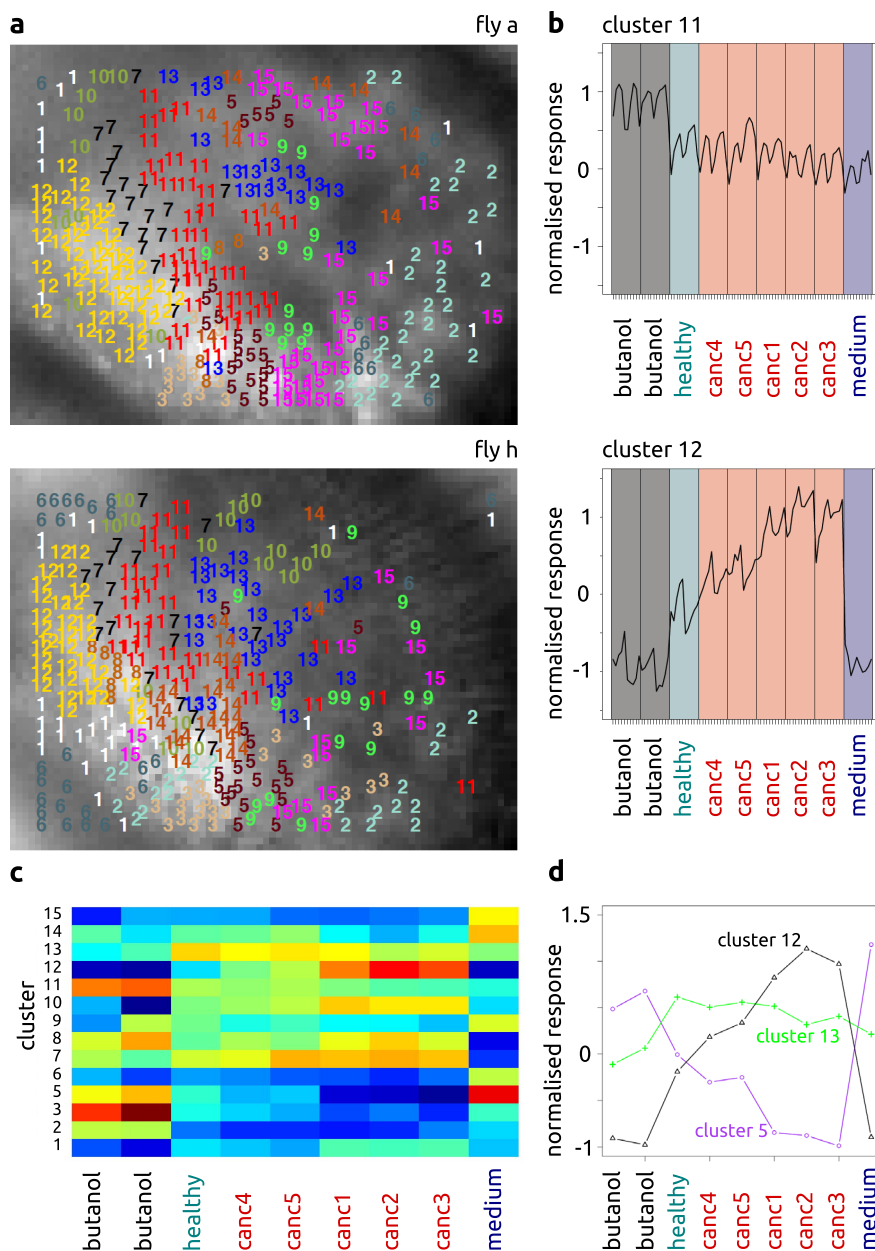


Figure 4 | Response spot clusters are roughly conserved across flies. (a) Clustering of response spots across all flies. The 14 (out of 15) clusters with members in each fly are superimposed onto raw fluorescence images from the respective antenna movies. For an extended version of this figure, see Supplementary Fig. S2. (b) Cluster centers for two of the clusters from (a). The cluster centers are normalised odour response profiles containing 10 time points for each odour presentation. (c) False-colour coded response magnitudes (mean response over all 10 time points during the odour responses) of each cluster for all test odours. (d) Response magnitudes across odours for clusters 5, 12 and 13 (same data as in c). Note that cluster 12 has no response to butanol, but responds to the different test odours, while cluster 5 is strongly responsive to butanol and control medium, but less to cancer cell odours. Cluster 13 has a less specific response profile.

medium control, while the distance to the cancer cell samples was much larger and significantly different (Kruskal-Wallis, $p = 0.00003$; post-hoc: pairwise Wilcoxon with Holm correction). Fig. 5f summarises ranked distances to *medium control*. We also found high inter-rater concordance for the distance ranking as measured by Kendall's coefficient of concordance ($W = 0.81$, $p = 0.0002$). Thus, based on distances between odours as measured by calcium imaging of the antenna, it is possible to discriminate *healthy control* cells from cancer cells.

In order to assess the reliability of this approach, we performed linear discriminant analysis (LDA, MASS package for R) with leave-one-out cross validation: One odour/animal combination was omitted, then LDA was performed, and the odour class of the omitted

odour/animal combination was predicted. We found that classification success was high between the groups {*canc1*, *canc2*, *canc3*} and {*canc4*, *canc5*}, but misclassification was frequent within these groups (see Supplementary Fig. S3c). This indicates that it is not possible to reliably discriminate *canc4* from *canc5*, or *canc1* from either *canc2* or *canc3*, but that discrimination between these two groups is robust. Importantly, in all cases healthy cells could be discriminated from cancer cells (LDA and leave-one-out cross validation, Supplementary Fig. S3d).

Both in the PCA projection (Fig. 5a, b) and in the distance analysis (Fig. 5f) two odour clusters became visible, grouping {*canc1*, *canc2*, *canc3*} and {*canc4*, *canc5*} together. We note that this clustering corresponds to differences in cell proliferation: The same grouping

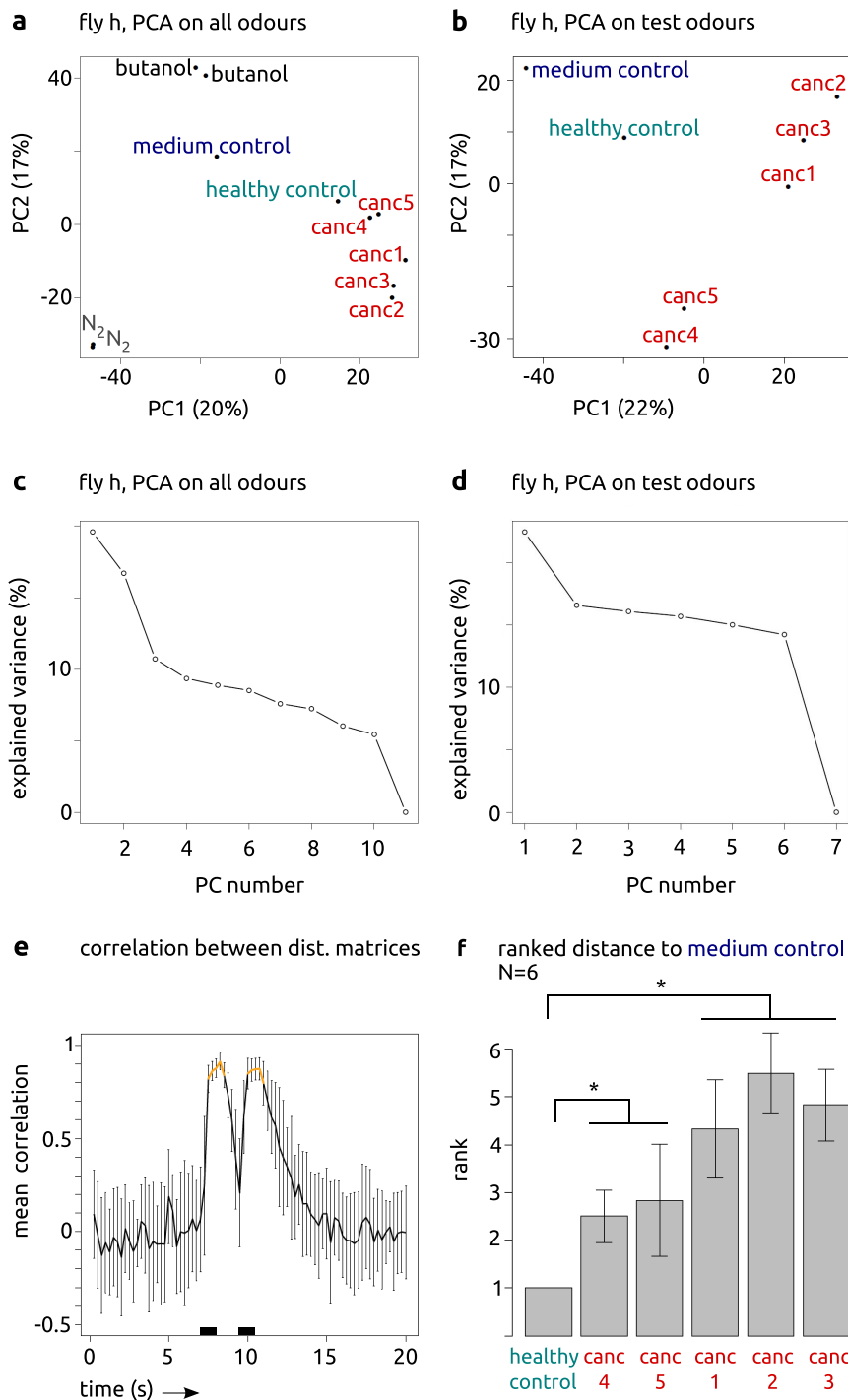


Figure 5 | Cancer odours can be detected using *Drosophila* antennae. (a) Single fly (same fly as in Fig. 2): Both reference and test odours in a 2D space spanned by PC1 and PC2. (b) Only test odours in a 2D space spanned by PC1 and PC2. (c) Scree plot of the PCA analysis shown in (a). Variability accounted for remains high for the first 10 dimensions. (d) Scree plot of the PCA analysis shown in (b). Variability accounted for remains high for the first 6 dimensions. (e) Mean (and standard deviation) Pearson correlation between the 7×7 test odour distance matrices over time. Each odour was applied two times (black bars). Time points marked in orange were used for the response profiles. (f) Ranked distances (mean and std. dev.) of all stimuli that contain the culture medium to *medium control*. Stimuli are significantly different (Kruskal-Wallis rank sum test, $p = 0.00003$). Post-hoc testing: Significant differences (corrected $p < 0.05$; pairwise Wilcoxon rank sum tests with Holm correction) are marked with stars.

appears in the proliferation curves (Supplementary Fig. S4a), as *canc4* and *canc5* derive from faster growing cell lines (lower doubling time, Supplementary Fig. S4b) than *canc1*, *canc2* and *canc3*.

While our protocol controlled for cell density by inoculating an appropriate number of cells (Supplementary Fig. S4a, c and Methods), this observation could also suggest that faster growing cells produce quantitatively more metabolites, rather than

qualitatively different metabolites. As a result, our analysis would detect quantitative differences. Therefore, we quantified the total amount of substances in the probe overhead spaces using GC-MS. We found that *canc3* had slightly less total VOC abundance than the other probes, while *canc4* had slightly more total VOC abundance. Thus, VOC abundance does not correlate with cell proliferation (Supplementary Fig. S4b, d), indicating that quantitative differences



in metabolite production cannot explain the results reported here. Possibly, cell proliferation speed in cancer and *healthy control* cells (highest doubling time, Supplementary Fig. S4b) translates into metabolic differences that affect the qualitative chemical composition of the culture medium headspace.

Discussion

We have shown that an array of odorant receptors on the fruit fly's antenna can be read out by calcium imaging, allowing to access a highly sensitive, natural chemosensor array. These *Drosophila* sensors are capable of detecting and discriminating medically relevant odours at low concentrations. Odours from culture media imprinted by different cancerous and healthy cell samples were represented by distinct odour response patterns on the antenna, and distances between response patterns were conserved across flies. The finding that cancer cells produce different VOC profiles opens up the possibility of employing *Drosophila* sensors for analysing volatiles emitted from human breath, skin or other bodily products.

Our results confirm earlier findings³² indicating that odorant receptors with broad receptive ranges can provide distinguished responses to a variety of chemicals even if these chemicals do not occur in the animal's natural environment. This generality of olfactory sensing is also the basis for the reported cases of cancer classification by other animals, such as dogs²³. Direct access to an array of odorant receptors, as established in this work, is an objective and standardisable way of circumventing the behavioural readout from animal systems that is prone to errors and that requires training of animals and human operators.

In contrast to electrophysiological recordings of single sensilla³², imaging a receptor array allows to record many receptors simultaneously. Initially, imaging the antenna surface results in data from a number of sensors (pixels) that is much larger than the number of accessible receptor types (approx. 30). Pixel responses may contain single receptor signals or unknown mixtures of multiple receptor signals. Pixels can be redundant or they may contain only noise. To obtain a diverse set of informative features for distance analysis, without using information about odour identity, we thus employed an unsupervised feature selection approach (Methods).

Relying on a natural receptor array differs from test assays targeted at cancer-specific metabolites or markers that have to be known beforehand. Our array strategy works also in the absence of such knowledge, allowing to discriminate cancer cells from healthy cells based on differences in their chemical markup that need not be known *a priori*. An array response comprises multiple, to some degree independent, features that can be combined and that increase the sensor's robustness. Differences between cancer cells and healthy cells that consist not of the presence or absence of key compounds, but rather of modified relative concentrations, are easily detected in a combinatorial array response pattern, while they can be confounded with sample concentration in a single-detector assay.

Distances between odours were conserved across flies, and differences in cell growth rate may account in part for the observed distances and clusters. However, there was no complete correspondence between cell growth rate (Supplementary Fig. S4) and odour distance (Fig. 5): *canc4* and *canc5*, which had the highest growth rate, had a smaller distance to *healthy control* (lowest growth rate) than *canc1* - *canc3* (intermediate growth rate). Thus, growth rate and cell density may contribute to the odour distances, but they cannot explain them entirely.

Based on the odour response distances, we can state that 1) all cell types leave traces in headspace VOC profiles that distinguish them from *medium control*, that 2) larger differences occur for the cancer cells than for *healthy control*, and that 3) at least one chemical occurs in different concentration in cancer cells with respect to *healthy control* cells. The latter difference may be decreased or increased concentration. The increased responses to the cancer odours in

cluster 12 suggests that the concentration of at least one substance was increased.

The findings of this study may also lead to a more targeted test assay for the breast cancer types. This would involve identification of several key odorant receptors that contribute most to the odour response distances. We found that, across flies, a spatially confined region on the antenna shows graded responses to the test odours (cluster 12 in Fig. 4b). Judging from the position on the antenna, it appears likely that the responses in this region stem from large basiconic sensilla⁵¹ (s.a. Figure 1b), a hypothesis that needs to be tested with dedicated experiments on single receptor cell lines.

Detection of cancer odours with artificial chemosensing systems has been demonstrated before, e.g. by GC-MS, by metalloporphyrin-coated quartz microbalances or by gold nanoparticle sensor arrays^{12,14,11}. While sensor arrays can provide real-time responses, gas chromatography is only sequential and often requires pre-concentrating the samples in order to achieve sufficient sensitivity. Electronic noses are not yet universal detectors, and they are usually limited in their receptive range and their sensitivity¹⁸. Even though the *Drosophila* sensor array is bound to be limited in its receptive range as well, harnessing the capabilities of biological sensor systems can ultimately lead to sensitive chemosensors with a broader receptive range than available today, and it might help to complement existing electronic noses, filling the gaps that cannot be reached by artificial systems.

Here, we have presented a proof of concept, showing that the *Drosophila* sensor array is suitable for medical applications. Future work includes integrating odorant receptors into artificial systems with the potential for real-time readouts from a sensor array with high sensitivity and a broad receptive range.

Methods

Cell cultures. Six different human breast cell lines were used in the experiments: MDA-231, MCF-7, SKBR3 (kindly provided by Prof. Giannini G., Department of Molecular Medicine, "Sapienza" University of Rome, Rome, Italy), BT474, ZR75-1 and MCF-10A (kindly supplied by Dr. Falcioni R., Department of Experimental Oncology, Regina Elena National Cancer Institute, Rome, Italy).

The immortalised, non-transformed human mammary epithelial cell line MCF-10A, referred to as *healthy control*, was grown in DMEM/F12 medium (Sigma-Aldrich) supplemented with 5% fetal bovine serum, 20 ng/ml epidermal growth factor (EGF), 10 µg/ml insulin, 0.5 µg/ml hydrocortisone (Sigma-Aldrich), 100 units/ml penicillin and 100 µg/ml streptomycin (Sigma-Aldrich), as previously described⁵².

The five human breast cancer cell lines (*canc1* to *canc5*) were derived from different breast cancer histotypes: MDA-231 (*canc4*), MCF-7 (*canc5*) and SKBR3 (*canc1*) cell lines from metastatic breast adenocarcinoma (MetAC), BT474 cells (*canc2*) from invasive ductal carcinoma (IDC) and ZR75-1 cells (*canc3*) from metastatic invasive ductal carcinoma (MetIDC) (see ATCC.org website). These cancer cells were grown in DMEM culture medium (DMEM high-glucose medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum (Sigma-Aldrich), 100 units/ml penicillin and 100 µg/ml streptomycin (Sigma-Aldrich)). All cells were cultured under standard conditions at 37°C in humidified atmosphere containing 5% CO₂.

Analysis of cell growth rate. In order to inoculate the media for the later VOC experiments with a number of cells that will result in a similar cell number for all cell lines after 24 h, the proliferation rate of *healthy control* and cancer cell lines was analysed beforehand by cell count over 96 h (Supplementary Fig. S4). Each cell line was seeded in triplicate in its specific culture medium at an initial concentration of 2.5×10^5 cells/flask (25 cm²), and cell growth was monitored after 24, 48, 72 and 96 h. Data from individual growth curves was used to calculate growth rate and doubling time (t_d) with <http://www.doubling-time.com/compute.php>. The proliferation of *healthy control* cells was also analysed in DMEM culture medium (as used in the later VOC analysis) to check for a possible alteration of growth rate: No significant changes (compared to their growth in the specific culture medium) were observed up to 96 h of culture (Supplementary Fig. S4c).

Sample preparation. For VOCs analysis, *healthy control* and cancer lines (*canc1* to *canc5*) were seeded in triplicate in culture flasks (25 cm²) in 5 mL of their specific culture medium and were grown for 24 h. The number of plated cells was chosen based on the specific doubling time of every cell line, in order to obtain a comparable cell number at the end of the incubation of 24 h. After 24 h, the specific culture medium was removed and replaced with 5 mL of the DMEM culture medium. Cells were grown in these conditions for the next 96 h, up to a confluence of 50%-60% (around 1.5×10^6 cells/flask). After this incubation period, the DMEM culture



medium was harvested, centrifuged at 1200 rpm for 5 min and collected in sterilised glass vials. Note that with this procedure, all samples derive from flasks with comparable cell density.

The *medium control* was obtained by incubating DMEM culture medium in the same conditions as the cell samples, but without seeded cells. Thus, *medium control* contains the same background odour as the cell samples, but without the influence of cells. Cell viability was evaluated by Trypan Blue exclusion test in order to assess the effect of any cell stress during the incubation time.

Animals. *Drosophila melanogaster* were kept at 25°C on a 12/12 light/dark cycle. Flies were reared on standard medium (100 ml contain: 0.7 g of agar, 2.4 g yeast, 2.1 g of sugar beet syrup, 7.1 g of cornmeal, 6.7 g of fructose, 1.4 ml of Nipagin (10%), 0.6 ml of propionic acid).

Flies were of genotype w; P[Orco:Gal4]; P[UAS:GCaMP3]attP40, expressing the Ca²⁺ reporter GCaMP3^{42,37} in all Orco bearing cells (UAS-GCaMP3 flies were provided by Loren L. Looger, Howard Hughes Medical Institute, Janelia Farm Research Campus, Ashburn, Virginia, USA). 1–5 day old female flies were used for experiments.

Odorant preparation. The reference odorant 1-butanol was purchased from Sigma (Sigma-Aldrich, Steinheim, Germany; CAS: 71-36-3) in ≥99.5% purity and diluted in 5 ml mineral oil (Sigma-Aldrich, Steinheim, Germany; CAS: 8042-47-5) to a concentration of 10⁻⁴ vol/vol. All odours were prepared in 20 ml headspace vials, covered with nitrogen and sealed with a Teflon septum (Axel Semrau, Germany). Cancer samples were used in 1 ml aliquots. Nitrogen was directly taken from an injector connected to a gas bottle.

Calcium imaging. Calcium imaging was performed as described elsewhere^{44,53}. In brief, we used a fluorescence microscope (BX50WI, Olympus, Tokyo, Japan) equipped with a 50× air lens (Olympus LM Plan FI 50×/0.5). A CCD camera (SensiCam, PCO, Kelheim, Germany) was mounted on the microscope recording with 8 × 8 pixel on-chip binning, which resulted in 80 × 60 pixel sized images. For each stimulus, recordings of 20 s at a rate of 4 Hz were performed using TILLVISION (TILL Photonics, Gräfelfing, Germany). A monochromator (Polychrome V, TILL Photonics, Gräfelfing, Germany) produced excitation light of 470 nm wavelength which was directed onto the antenna via a 500 nm low-pass filter and a 495 nm dichroic mirror. Emission light was filtered through a 505 nm high-pass emission filter.

Stimulus application. Odours were applied automatically using a computer-controlled autosampler (PAL, CTC Switzerland). 2 ml of headspace was injected in two 1 ml portions at time points 6 s and 9 s with an injection speed of 1 ml/s into a continuous flow of purified air flowing at 60 ml/min. The 1 s stimulus was directed onto the antenna of the fly via a Teflon-tube (inner diameter 1 mm, length 38 cm).

Experiments were performed double-blind. The seven test odours (*healthy control*, *medium control*, *cancl1*, *cancl2*, *cancl3*, *cancl4*, *cancl5*) were measured in random order. Reference odours (1-butanol and N₂) were measured before and after a full block of test odours in order to ensure reproducibility and viability of the preparation (1-butanol), and to exclude contamination of the system (N₂). The autosampler syringe was flushed with purified air for 1 min after each injection and washed with pentane (Merck, Darmstadt, Germany) automatically after each application of 1-butanol.

GC-MS measurements. Vials containing samples were placed in a water bath equilibrated at 40°C. Headspace VOCs were preconcentrated onto a SPME fibre (50/30 μm divinylbenzene/carboxen/PDMS, SUPELCO, Bellefonte, PA, USA) manually exposed to sample vapours for 1 h. The fibre with extracted VOCs was transferred to the GCMS (GCMS-QP 2010 Shimadzu) and desorbed at 250°C for 3 minutes in the injection port of the GC. The instrument is equipped with an EQUITY-5 (poly(5% diphenyl/95% dimethyl siloxane) phase, SUPELCO, Bellefonte, PA, USA) capillary column, 30 m length × 0.25 mm I.D. × 0.25 μm thickness. The analysis was conducted in splitless mode using ultra-high purity helium as carrier gas. Carrier gas constant liner velocity was kept constant at 30.2 cm/min. The oven temperature was kept at 40°C for 5 min, then increased by 7°C/min up to 220°C and then at 15°C/min up to 300°C. The final temperature was held for 2 min (total run time: 39 min). The mass spectrometer was used in the full scan mode over a mass range of 40–450 m/z. The detector voltage was 0.7 kV. The temperature of interface and ion source was kept constant at 250°C. The GC area of the samples was calculated using the section GCMS post-run analysis of the GCMS solutions software (version 2.4, Shimadzu Corporation).

Unsupervised feature selection. Antenna imaging movies were processed in a KNIME (www.knime.org) workflow using the ImageBee plugin⁵⁴ for insect neuroimage data (http://tech.knime.org/community/image-processing). For feature selection, all 11 movies recorded in one fly (reference and test odours) were concatenated, resulting in a movie matrix *A* with dimensions (*m* = 80 × 11 time points) × (*n* = 80 × 60 pixels). All images were aligned by cross correlation to correct for animal movement.

In principle, all *n* pixels could be used for distance computations, however at the cost of including many unresponsive or noisy pixels that can obscure odour distances. We thus selected *c* = 300 pixels (features) based on their contribution to the norm of *A*.

Before feature selection, data was preprocessed to ensure that the norm of *A* was not dominated by e.g. unspecific background fluorescence: 1) Background fluorescence

was removed by subtracting the mean image separately for each of the 11 movies. 2) Photon shot noise was reduced by smoothing individual images with a Gaussian kernel (width 9). Then, we selected *c* column vectors (pixels) into the *m* × *c* matrix *C* such that the Frobenius norm error $\|A - CC^+ A\|_F$ was minimised (where *C*⁺ is the pseudoinverse). The objective criterion was optimised with the convex cone algorithm⁵⁵. Minimising the Frobenius norm error in this way corresponds to the norm error minimisation objective of PCA, and it helps to identify a diverse set of pixels that contribute a lot to the variance in *A*. I.e., instead of selecting pixels by an equally-spaced grid, pixel selection was biased towards areas that actually responded to the odour stimuli.

Response spot time series. Unsupervised feature selection from the movie matrix *A* (with *m* time points and *n* pixels) resulted in a set of *c* = 300 pixel indices. Before extracting the pixels from (the movement-corrected, but otherwise unprocessed) *A*, dye bleaching and other global intensity changes were reduced by histogram normalisation: For each of the 11 movies in *A*, we took the first image as a reference, and matched all histograms of the remaining images to the histogram of the first image, obtaining the processed movie \hat{A} . Then, the *c* pixels were extracted from \hat{A} . In order to reduce noise, a postprocessing step described in⁵⁵ replaced a pixel *p_j* by the average of those pixels that have a time series more similar to *p_j* than to any of the other *c* - 1 pixels.

In summary, from each antenna recording we obtained a set of *c* representative *response spots*, i.e. pixel positions and the corresponding averaged and processed time series, in an unsupervised fashion. Response spots were normalised to the prestimulus interval, separately for each of the 11 stimuli, by computing $(F_i - F_0)/F_0$ for each time point *i*. Here, *F₀* is the mean fluorescence during 20 time points before stimulus application and *F_i* is the fluorescence value at time point *i*.

For each time series, we selected five time points from each of the two response peaks (marked in Fig. 3b), resulting in *t* = 10 time points for each of the *s* = 11 stimuli, i.e. 110 time points. Together, each fly contributed a *c* × (*s* × *t*) *response profile* matrix *M* which was then z-score normalised: From all *i* rows (response spots), the mean μ_i was subtracted and rows were divided by the standard deviation σ_i . Likewise, from all *j* columns (time points), the mean μ_j was subtracted and columns were divided by the standard deviation σ_j .

Clustering, distance matrices, PCA. For Fig. 3, clustering was performed on complete (unnormalised) response spot time series from a single fly. For Fig. 4, clustering of (normalised) response profiles was performed on the row-concatenated

matrix $M_{\text{pooled}} = \begin{Bmatrix} M_1 \\ \vdots \\ M_N \end{Bmatrix}$ (all *a* = 1, ..., *N* flies pooled). In both cases, we used the *k*-

means clustering algorithm (stats package for R, default settings, 1000 restarts). The number of clusters (15) was estimated based on a scree plot of the overall within-cluster sum of squares error.

For analysis of odour distances, the *M_a* were reshaped as *s* × (*t* × *c*). Odour × odour (*s* × *s*) Euclidean distance matrices (Fig. 5e, f) were computed on the *M_a*. For correlation analysis, we regarded only the 7 × 7 distance matrices for the 7 test odours, enabling us to state explicitly that distances between the relevant test odours are correlated. We correlated, for each time point, the individual distance matrices: Fig. 5e shows the mean of all (*N* * (*N* - 1))/2 pairwise correlations between the *N* flies over time. By correlation we refer to the Pearson product moment correlation coefficient. Only the lower diagonal submatrices (without the diagonal) of the distance matrices were correlated.

PCA was computed on the full *s* × (*t* × *c*) matrix *M* from one fly (Fig. 5a), and on a modified matrix *M'*, from which all data points stemming from the reference odours butanol and N₂ were removed (Fig. 5b). PCA for pooled data (Supplementary Fig. S3) was computed on the column-concatenated matrices $M_{\text{pooled}} = \{M_1, \dots, M_N\}$ and $M'_{\text{pooled}} = \{M'_1, \dots, M'_N\}$, respectively.

1. Tennant, D. A., Durán, R. V. & Gottlieb, E. Targeting metabolic transformation for cancer therapy. *Nat. Rev. Cancer* **10**, 267–277 (2010).
2. Ward, P. S. & Thompson, C. B. Metabolic reprogramming: A cancer hallmark even Warburg did not anticipate. *Cancer Cell* **21**, 297–308 (2012).
3. Kneepkens, C. M., Lepage, G. & Roy, C. C. The potential of the hydrocarbon breath test as a measure of lipid peroxidation. *Free Radical Biol. Med.* **17**, 127–160 (1994).
4. Singer, S. J. & Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. *Science* **175**, 720–731 (1972).
5. Jerby, L. *et al.* Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer. *Cancer Res.* **72**, 5712–5720 (2012).
6. Hoffman, S. A., Joo, W.-A., Echan, L. A. & Speicher, D. W. Higher dimensional (hi-d) separation strategies dramatically improve the potential for cancer biomarker detection in serum and plasma. *J. Chromatogr. B* **849**, 43–52 (2007).
7. Solassol, J., Du-Thanh, A., Maudelonde, T. & Guillot, B. Serum proteomic profiling reveals potential biomarkers for cutaneous malignant melanoma. *Int. J. Biol. Markers* **26**, 82–87 (2011).
8. Phillips, M. *et al.* Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. *Lancet* **353**, 1930–1933 (1999).
9. Amann, A., Spanel, P. & Smith, D. Breath analysis: The approach towards clinical applications. *Mini-Rev. Med. Chem.* **7**, 115–129 (2007).



10. Hakim, M. *et al.* Volatile organic compounds of lung cancer and possible biochemical pathways. *Chem. Rev.* **112**, 5949–5966 (2012).
11. Barash, O., Peled, N., Hirsch, F. R. & Haick, H. Sniffing the unique “odor print” of non-small-cell lung cancer with gold nanoparticles. *Small* **5**, 2618–2624 (2009).
12. Filipiak, W. *et al.* Release of volatile organic compounds (VOCs) from the lung cancer cell line CALU-1 in vitro. *Cancer Cell Int.* **8**, 17 (2008).
13. Bartolazzi, A. *et al.* A sensor array and GC study about VOCs and cancer cells. *Sens. Actuators, B* **146**, 483–488 (2010).
14. Di Natale, C. *et al.* Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors. *Biosens. Bioelectron.* **18**, 1209–1218 (2003).
15. Peng, G. *et al.* Diagnosing lung cancer in exhaled breath using gold nanoparticles. *Nat. Nanotechnol.* **4**, 669–673 (2009).
16. Kwak, J. *et al.* Volatile biomarkers from human melanoma cells. *J. Chromatogr. B* **931**, 90–96 (2013).
17. Persaud, K. & Dodd, G. H. Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. *Nature* **299**, 352–355 (1982).
18. Röck, F., Barsan, N. & Weimar, U. Electronic nose: Current status and future trends. *Chem. Rev.* **108**, 705–725 (2008).
19. Berna, A. Z., Anderson, A. R. & Trowell, S. C. Bio-benchmarking of electronic nose sensors. *PLoS ONE* **4**, e6406 (2009).
20. Williams, H. & Pembroke, A. Sniffer dogs in the melanom clinic? *Lancet* **333**, 734 (1989).
21. Church, J. & Williams, H. Another sniffer dog for the clinic? *Lancet* **358**, 930 (2001).
22. Lippi, G. & Cervellin, G. Canine olfactory detection of cancer versus laboratory testing: myth or opportunity? *Clin. Chem. Lab. Med.* **50**, 435–439 (2012).
23. Boedecker, E., Friedel, G. & Walles, T. Sniffer dogs as part of a bimodal bionic research approach to develop a lung cancer screening. *Interact. Cardiovasc. Thorac. Surg.* **14**, 511–515 (2012).
24. Pickel, D., Manucy, G. P., Walker, D. B., Hall, S. B. & Walker, J. C. Evidence for canine olfactory detection of melanoma. *Appl. Anim. Behav. Sci.* **89**, 107–116 (2004).
25. Willis, C. M. Olfactory detection of human bladder cancer by dogs: proof of principle study. *BMJ* **329**, 712–0 (2004).
26. Cornu, J.-N., Cancel-Tassin, G., Ondet, V., Girardet, C. & Cussenot, O. Olfactory detection of prostate cancer by dogs sniffing urine: A step forward in early diagnosis. *Eur. Urol.* **59**, 197–201 (2011).
27. McCulloch, M. *et al.* Diagnostic accuracy of canine scent detection in early- and late-stage lung and breast cancers. *Integr. Cancer. Ther.* **5**, 30–39 (2006).
28. Ehmman, R. *et al.* Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *Eur. Respir. J.* **39**, 669–676 (2012).
29. Horvath, G., Järverud, G. a. K., Järverud, S. & Horvath, I. Human ovarian carcinomas detected by specific odor. *Integr. Cancer. Ther.* **7**, 76–80 (2008).
30. Horvath, G., Andersson, H. & Paulsson, G. Characteristic odour in the blood reveals ovarian carcinoma. *BMC Cancer* **10**, 643 (2010).
31. Sonoda, H. *et al.* Colorectal cancer screening with odour material by canine scent detection. *Gut* **60**, 814–819 (2011).
32. Marshall, B., Warr, C. G. & de Bruyne, M. Detection of volatile indicators of illicit substances by the olfactory receptors of *Drosophila melanogaster*. *Chem Senses* **35**, 613–625 (2010).
33. Hallem, E. A., Ho, M. G. & Carlson, J. R. The molecular basis of odor coding in the *Drosophila* antenna. *Cell* **117**, 965–79 (2004).
34. Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125**, 143–60 (2006).
35. Galizia, C. G., Münch, D., Strauch, M., Nissler, A. & Ma, S. Integrating Heterogeneous Odor Response Data into a Common Response Model: A DoOR to the Complete Olfactome. *Chem. Senses* **35**, 551–563 (2010).
36. Charpak, S., Mertz, J., Beaufreire, E., Moreaux, L. & Delaney, K. Odor-evoked calcium signals in dendrites of rat mitral cells. *PNAS* **98**, 1230–1234 (2001).
37. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* **6**, 875–81 (2009).
38. Lucas, P. & Shimahara, T. Voltage- and calcium-activated currents in cultured olfactory receptor neurons of male mamestra brassicae (Lepidoptera). *Chem. Senses* **27**, 599–610 (2002).
39. Grienberger, C. & Konnerth, A. Imaging calcium in neurons. *Neuron* **73**, 862–885 (2012).
40. Wicher, D. *et al.* *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**, 1007–11 (2008).
41. Sato, K., Pellegrino, M., Nakagawa, T. T., Vosshall, L. B. & Touhara, K. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002–6 (2008).
42. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca²⁺ probe composed of a single green fluorescent protein. *Nat. Biotechnol.* **19**, 137–41 (2001).
43. Couto, A., Alenius, M. & Dickson, B. J. Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. *Curr. Biol.* **15**, 1535–1547 (2005).
44. Pelz, D., Roeske, T., Syed, Z., Bruyne, M. d. & Galizia, C. G. The molecular receptive range of an olfactory receptor in vivo (*Drosophila melanogaster* Or22a). *J. Neurobiol.* **66**, 1544–1563 (2006).
45. Angioy, A. M., Desogus, A., Barbarossa, I. T., Anderson, P. & Hansson, B. S. Extreme sensitivity in an olfactory system. *Chem. Senses* **28**, 279–284 (2003).
46. Weissbecker, B., Holighaus, G. & Schütz, S. Gas chromatography with mass spectrometric and electroantennographic detection: analysis of wood odors by direct coupling of insect olfaction and mass spectrometry. *J. Chromatogr. A* **1056**, 209–216 (2004).
47. Park, K. C., Zhu, J., Harris, J., Ochieng, S. A. & Baker, T. C. Electroantennogram responses of a parasitic wasp, *Microplitis croceipes*, to host-related volatile and anthropogenic compounds. *Physiol. Ent.* **26**, 69–77 (2001).
48. Schöning, M. J., Schroth, P. & Schütz, S. The use of insect chemoreceptors for the assembly of biosensors based on semiconductor field-effect transistors. *Electroanalysis* **12**, 645–652 (2000).
49. Schütz, S. *et al.* An insect-based BioFET as a bioelectronic nose. *Sens. Actuators, B* **65**, 291–295 (2000).
50. Park, K. C., Ochieng, S. A., Zhu, J. & Baker, T. C. Odor discrimination using insect electroantennogram responses from an insect antennal array. *Chem. Senses* **27**, 343–352 (2002).
51. de Bruyne, M., Foster, K. & Carlson, J. R. Odor coding in the *Drosophila* antenna. *Neuron* **30**, 537–552 (2001).
52. Giannini, G. *et al.* EGF- and cell-cycle-regulated STAG1/PMEPA1/ERG1.2 belongs to a conserved gene family and is overexpressed and amplified in breast and ovarian cancer. *Mol. Carcinog.* **38**, 188–200 (2003).
53. Münch, D., Schmeichel, B., Silbering, A. F. & Galizia, C. G. Weaker ligands can dominate an odor blend due to syntopic interactions. *Chem. Senses* **38**, 293–304 (2013).
54. Strauch, M., Rein, J., Lutz, C. & Galizia, C. G. Signal extraction from movies of honeybee brain activity by convex analysis: The ImageBee plugin for KNIME. *BMC Bioinformatics* **14**(Suppl 18):S4 (2013).
55. Strauch, M., Rein, J. & Galizia, C. G. Signal extraction from movies of honeybee brain activity by convex analysis. In Mandoiu, I., Pop, M., Rajasekaran, S. & Spouge, J. (eds.) *Proceedings of ICCABS, Feb 23–25 2012, Las Vegas, USA, IEEE* (2012).

Acknowledgments

We acknowledge support by Fondazione Veronesi (2011–2013), German Research Foundation DFG and Zukunftskolleg, Universität Konstanz.

Author contributions

M.S. analysed data. A.L., D.M. and T.L. performed imaging/fruit fly experiments. L.L. and A.U. performed cell culture experiments. R.P., A.C. and R.C. performed GC-MS analysis. M.S., A.L., D.M., C.G.G., E.M. and C.D. designed research and wrote the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Strauch, M. *et al.* More than apples and oranges - Detecting cancer with a fruit fly's antenna. *Sci. Rep.* **4**, 3576; DOI:10.1038/srep03576 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0>