



(51) International Patent Classification:

G06V 10/771 (2022.01) G06V 20/69 (2022.01)
C12Q 1/02 (2006.01)

(21) International Application Number:

PCT/GB2024/051496

(22) International Filing Date:

12 June 2024 (12.06.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2310406.0 06 July 2023 (06.07.2023) GB

(71) Applicant: **IMPERIAL COLLEGE INNOVATIONS LIMITED** [GB/GB]; Level 1 Faculty Building, c/o Imperial College, Exhibition Road, London, Greater London SW7 2AZ (GB).

(72) Inventor: **BROWN, Andre**; c/o Imperial College Innovations Limited, Level 1 Faculty Building, c/o Imperial College, Exhibition Road, London, Greater London SW7 2AZ (GB).

(74) Agent: **PIOTROWICZ, Pawel**; Venner Shipley LLP, 406 Cambridge Science Park, Milton Road, Cambridge CB4 0WW (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: DETERMINING THE EFFECTS OF MODIFIED PEPTIDES

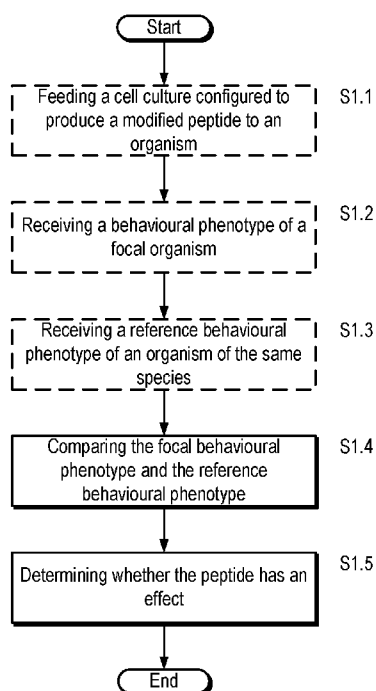


Fig. 26

(57) Abstract: A method of determining an effect of a modified peptide is disclosed. The method comprises comparing a behavioural phenotype of an organism, the organism having ingested a cell culture, wherein cells in the cell culture are configured to produce a modified peptide, with a reference behavioural phenotype of organisms of the same species, and determining whether the peptide has an effect based on the difference between the behavioural phenotype and the reference behavioural phenotype.

Published:

— *with international search report (Art. 21(3))*

DETERMINING THE EFFECTS OF MODIFIED PEPTIDES

Field

The present invention relates to methods and hardware for identifying effects of modified peptides by comparing behavioural phenotypes of organisms.

Background

Modified peptides such as ribosomally synthesized and post-translationally modified peptides (RiPPs), possess a wide range of biological functions and are increasingly considered to be important for the discovery of new drugs, medicants and other therapeutics. RiPPs are modified peptides of ribosomal origin and are produced by a variety of organisms, including prokaryotes, eukaryotes, and archaea.

RiPPs comprise or consist of a ribosomally-produced peptide which undergoes enzymatic post-translational modification. The combination of peptide translation by the ribosome and subsequent modification is referred to as "post-ribosomal peptide synthesis" (PRPS).

Modified peptides can be classified into several different sub-groups with, for example, different structures, or different properties. These different structures or properties may have different effects on organisms.

Next-generation sequencing methods has made the discovery of new RiPPs in genomes common. Genome modification has in turn made the engineering of RiPPs in new host organisms possible.

The uses and biological activities of RiPPs are diverse. However, identifying the effect of these modified peptides can be difficult, in part due to the large quantities of high concentrations required to perform analyses.

Therefore, there is a need for a method of identifying the effects or potential effects of these modified peptides, for example, on a range of organisms.

Summary

According to a first aspect of the invention, there is provided a method of determining an effect of a modified peptide. The method comprises comparing a behavioural phenotype of an organism, the organism having ingested a cell culture,
5 wherein cells in the cell culture is configured to produce a modified peptide, with a reference behavioural phenotype of organisms of the same species, and determining whether the peptide has an effect based on the difference between the behavioural phenotype and the reference behavioural phenotype.

10 There may be a plurality of reference behavioural phenotypes. The reference phenotype may be an average, a weighted average, a mean, a median or a mode of a plurality of reference phenotypes.

Produce may mean express.

15

The organism may be an animal, for example, a nematode (e.g., *Caenorhabditis elegans*), an insect larva, (e.g., *Drosophila melanogaster*), zebrafish (*Danio rerio*), *Xenopus* sp. etc.).

20 The cell culture may comprise or consist of a bacterial cell culture, for example *Escherichia coli*, a fungal cell culture (e.g., *Saccharomyces cerevisiae*, *Pichia pastoris*, or *Schizosaccharomyces pombe*), a plant cell culture, or a protozoa cell culture.

25 The effect of the peptide may be a known biological effect, or a known chemical effect. The cell culture may be or comprise a genetically modified cell culture. The cell culture may have been treated with a compound, drug, pharmaceutical, or therapeutic. The cell culture may have a coating of a compound, for example, an active compound.

30

The modified peptide may be a ribosomally synthesised and post-translationally modified peptide.

The modified peptide may be a peptide modified by an enzyme.

35

The peptide may be a designed peptide e.g., a structure-guided or bioinformatic-guided peptide.

The reference behavioural phenotype may comprise a behavioural phenotype from an organism which has undergone a treatment.

- 5 The treatment may include a control treatment, for example, exposure to a solvent or a buffer solution and the like.

The treatment may comprise organisms having ingested a second cell culture, wherein cells in the second cell culture are configured to produce either non-
10 modified peptides, or modified peptides having a specified sequence.

The reference organism(s) may produce peptides producing a random sequence of amino acids, and may only produce peptides producing a random sequence of amino acids.

15 Identifying the peptide as having an effect may comprise identifying that the peptide has a similar effect as a known peptide, for example, by having two behavioural phenotypes which are very similar, where the reference behavioural phenotypes are from organisms which have been treated with or have ingested
20 compounds with known effects.

The second cell culture may be or comprise a genetically modified cell culture. The second cell culture may have been treated with a compound, drug, pharmaceutical, or therapeutic. The second cell culture may have a coating of a compound, for
25 example, an active compound.

The treatment may comprise an organism having been exposed to a compound.

The compound may be a chemical compound, a drug, for example, a
30 pharmaceutical drug, a nutraceutical, or a medicament, and the like.

For the reference behavioural phenotype, the organism having ingested a second cell culture, wherein cells in the second cell culture produce a second peptide, the second peptide may have a known effect.

35 The known effect may be a known effect on the behavioural phenotype of the organism, the known effect may be a known effect on a different organism, a cell, a

group of cells, a tissue, or an organ. For example, the known effect may be a physiological effect on a, or part of a, human, a non-human animal, plant, a fungus, a bacterium or archaeon. For example, the effect may be as a stimulant, as a depressant, or another physiological effect.

5

The organism and/or the reference organism may be a genetically modified organism.

The behavioural phenotype may comprise one or more features.

10

The behavioural phenotype may comprise any suitable number of features, for example, 4 features, 16 features, 64 features, 256 features, 1,000 features, 2,000 features, 3,000 features, 4,000 features, or more than 4,000 features.

15

The behavioural phenotype may comprise one or more traits, each trait may comprise one or more of:

body part location;

morphology;

body posture;

20

speed;

velocity;

turn data; and

trajectory data.

25

Morphology may comprise length, area, and width of the organism. Body posture may comprise curvature, curvature may be in relation to major and minor axes.

Trajectory data may comprise path curvature, path grid data, path area data.

Velocity data may comprise angular velocity, radial velocity. Turn data may comprise count data, frequency data, and/or curvature data.

30

The body part location may comprise anterior, posterior, or mid-body location data.

The body part location may comprise head location, neck location, hip location, and tail location.

35

A feature may comprise one or more traits.

Each feature may comprise trait data. For example, curvature of a particular part of the body, the speed relative to body length etc.

The behavioural phenotype may comprise frequency data.

5

Comparing the behavioural phenotypes may comprise selecting one or more features from the behavioural phenotypes.

Feature selection may comprise recursive feature elimination.

10

The feature selection may comprise using a regression model.

The regression model may be a logistic regression.

15 The difference between the behavioural phenotype and the reference behavioural phenotype may be less than a pre-determined threshold.

The behavioural phenotype may be obtained using an image of the organism.

20 The behavioural phenotype may be obtained using video of the organism.

The video may be for any suitable length of time.

25 The behavioural phenotype may be obtained by using image analysis to extract one or more traits from an image or a series of images.

30 According to a second aspect of the invention, there is provided a system comprising a camera configured to capture one or more images of an organism, a processor configured to compare a behavioural phenotype of an organism, the organism having ingested a cell culture, wherein cells in the cell culture are configured to produce a modified peptide, with a reference behavioural phenotype of organisms of the same species, and wherein the processor is configured to determine whether the peptide has an effect based on the difference between the behavioural phenotype and the reference behavioural phenotype.

35

Brief description of the drawings

Certain embodiments of the present invention will now be described, by way of example, with reference to the accompanying drawings, in which:

- 5 Figure 1 is a schematic of a system for recording images of an organism;
Figure 2 is a schematic block diagram of a system for recording images of an organism;
Figure 3 A-D are schematics of (A) insect larvae, (B) zebrafish, (C) roundworms, and (D) rat in respective arenas for image acquisition;
- 10 Figure 4 is a schematic of the core features of the body parts and morphology of *Caenorhabditis elegans*;
Figure 5 is a schematic of the core features of the path (trajectory) and postures of *Caenorhabditis elegans*;
Figure 6 a schematic of the core features of the speed and velocities of
- 15 *Caenorhabditis elegans*;
Figure 7 is a block diagram of operations of each of the core features;
Figure 8 are the results of recursive feature elimination;
Figure 9 are eight boxplots of N2 and three mutants using the Tierpsy_8 subset of features. The small set of features facilitates a visual comparison between
- 20 different strains;
Figure 10 is an example of a two-dimensional histogram of the effect of short light pulses for different strains and features for speed;
Figure 11 is an example of a two-dimensional histogram of the effect of short light pulses for different strains and features for radial tail up velocity relative to hips;
- 25 Figure 12a and 12b are comparisons between the changes of behaviour among the different strains under blue light stimulation for selected features;
Figure 13 is a cluster map of the median value of the Jensen–Shannon divergence between different strains among the ATR plates for short light pulses;
Figure 14 is a cluster map of the median value of the Jensen–Shannon divergence
- 30 between different strains among the ATR plates for long light pulses;
Figure 15a is a photograph of a megapixel camera array suitable for collecting images and video of organism behaviour;
Figure 15b is a 3D schematic drawing of a single imaging unit;
Figure 15c is a schematic drawing (front and side view) of an imaging unit
- 35 annotated with dimensions in millimetres;
Figure 15d to 15g are (15d) images of 480 wells recorded simultaneously from multiple cameras, (15e) an image from a single camera, (15f) a single well, and

(15g) a worm level showing sufficient resolution to precisely track the nematodes.

Each square well

measures 8×8 mm;

Figure 16a to 16c are examples of features describing morphology, movement, and posture. Each box shows median, 25th and 75th percentile (central mark, left and right edge, respectively), while whiskers show the rest of the distribution except for

outliers (outside 1.5 times the IQR above the 75th and below the 25th percentile);

Figure 17 is a boxplot of quantitative behavioural phenotypes;

Figure 18 is a plot predicted v true labels of worm strains in a held-out test;

10 Figure 19a is a plot of PCA analysis of escape responses to photostimulation for N2 and CB4856 in the 10 sec immediately before (left) and immediately after (right) a 10 s

stimulus showing detectable behavioural responses;

Figure 19b are plots showing behavioural responses of N2 and CB4856 strains to

15 photostimulation;

Figure 19c is a plot of repeated photostimulation responses in N2 worms;

Figure 19d is a plot of the fraction of worms triggered to move forwards by each stimulus;

Figures 20a and 20b are plots of principal component analysis of 256 extracted

20 behavioural features from standard (20a) or blue light (20b) imaging conditions;

Figure 20c is a plot of overall fraction of forward locomotion under blue-light imaging conditions;

Figure 20d and 20e are plots showing changes in the overall fraction of forward (20d) or paused (20e) locomotion upon blue light stimulation;

25 Figure 21 is a heatmap (a) of behavioural fingerprints of worms in response to treatment with drugs, heatmaps (b–d) of some compounds from the same class cluster together according to their behavioural response including inhibitors of muscle contraction (b), antibiotics (c), and antiparkinsonian (anticholinergic) drugs (d);

30 Figure 22a is a figure of a photograph of an entire 96-well plate with a megapixel camera array with enough resolution and high enough frame rate to track, segment, and estimate the posture of *C. elegans*;

Figure 22b is a plot of all compound doses in the speed/tail curvature space;

Figure 22c is a plot of sample worm skeletons over time showing the effect of the

35 compounds highlighted in (Figure 22b) on motion;

Figure 22d is a series of plots representing the number of features significantly different from the DMSO control at a false discovery rate of 1% for each compound, grouped by mode of action;

5 Figure 23a is a hierarchical clustering of behavioural fingerprints highlighting structure in the responses to different compounds. Each row of the heat map represents the mean dose fingerprint of a specific compound described by 256 pre-selected features from each blue light condition;

Figure 23b is a plot of cluster purity as a function of the hierarchical cluster distance;

10 Figure 23c is a plot of compounds in the same class which can have different dose-response curves;

Figure 24a is a 3D plot of toy data illustrating the potential benefit of normalization in correcting for potency differences within mode-of-action classes;

15 Figure 24b is a confusion matrix obtained through cross-validation for the best performing feature set (1,024 features) and logistic regression classifier following feature selection and hyperparameter tuning on the training data;

Figure 24c is a confusion matrix for the classifier trained in (Figure 24b) applied to previously unseen test compounds without any further tuning;

20 Figure 24d is the novelty score assigned to novel test compounds with a mode of action not seen during training compared to the novelty score of compounds from the test set in Figure 24c;

Figure 25a is a confusion matrix showing cross-validation performance of a classifier trained to distinguish serotonin receptor antagonists from each other;

Figure 25b is a confusion matrix for the mitochondrial inhibitors;

25 Figure 26 is a process flow diagram of method of determining whether a modified peptide has an effect;

Figure 27a is a boxplot of motion behaviour of worms fed with lanthipeptide-expressing E. coli and a control sample;

30 Figure 27b is a boxplot of the width of worms fed with lanthipeptide-expressing E. coli and a control sample; and

Figure 27c is a boxplot of the curvature of worms fed with lanthipeptide-expressing E. coli and a control sample.

Detailed description of certain embodiments

35 Behaviour is a sensitive and integrative readout of nervous system function and therefore an attractive measure for assessing the effects of mutation or drug treatment on animals. The behaviour of an organism is therefore an excellent

candidate for identifying or assessing the effects of modified peptides, for example, RIPPs on organisms. Behaviour has been used as a tool to assess the effects of mutations and drugs in a variety of organisms, for example, nematodes (e.g., *Caenorhabditis elegans*), insects or insect larvae, (e.g., fruit flies, *Drosophila melanogaster*), fish (e.g., zebrafish, *Danio rerio*), amphibians (e.g., African clawed frog, *Xenopus laevis* etc.), birds (e.g., chick, *Gallus gallus domesticus*), and mammals, (e.g., mouse, *Mus musculus*, rat, *Rattus norvegicus*).

Collecting behavioural phenotype data

10 Video data provide a rich but high-dimensional representation of behaviour, and so the first step of analysis is often some form of tracking and feature extraction to reduce dimensionality while maintaining relevant information. Modern machine-learning methods are powerful but notoriously difficult to interpret, while handcrafted features are interpretable but do not always perform as well. As an
15 example, a set of handcrafted features is presented to compactly quantify the roundworm *Caenorhabditis elegans* behaviour. However, hand crafted features are available for other organisms, or could be identified or compiled using a similar method as will be described hereinafter. These handcrafted features however may be used directly for similar organisms, for example, other roundworms, insect
20 larvae, amphibians or amphibian larvae, and fish, including fish larvae or fry.

The features are designed to be interpretable but to capture as much of the phenotypic differences between worms as possible. The full feature set is more powerful than a previously defined feature set in classifying mutant strains. A
25 combination of automated and manual feature selection is used to define a core set of interpretable features that still provides sufficient power to detect behavioural differences between mutant strains and the wild-type. The new features are used to detect time-resolved behavioural differences in a series of optogenetic experiments targeting different neural subsets.

30 Referring to Figure 1, a system 1 for capturing video of organisms 2 in their environment comprises a camera 3 attached to a lens 4, which is in turn connected to a computer 5 via a connector 6 configured to allow communication between the camera 3 and the computer 5. The camera 3 may be a digital camera. The
35 organisms 2 are placed in an arena 7 to allow them to exhibit one or more behaviours. The arena may comprise suitable substrate 8 for the organisms 2 to move on. For example, for roundworms (e.g., *C. elegans*), the substrate 8 may be

a buffered agar-based medium (e.g., Nematode Growth Medium (NGM) see
www.wormbook.org), but different organisms will have different substrate
requirements. The substrate 8 may have food 9 for the organism arranged either
on or in the substrate 8. The food 9 may comprise a cell culture. As explained
5 herein after, the cell culture may be configured (e.g., through selection or genetic
modification) to produce or express one or more modified peptides, e.g.,
ribosomally synthesized and post-translationally modified peptides (RiPPs).

The organism 2 may be placed in the arena 7 and allowed to exhibit one or more
10 behaviours. The camera 3 captures video or sequential images of the organism 2
and sends them to the computer 5 via the connector 6.

Referring also to Figure 3, the computer comprises a graphics processing unit 11 for
processing the image data received from the camera 4, a central processing unit 12
15 for other processing tasks, an input 13 and output 14, an optional display 15, and a
bus 16 connecting these components to each other within the computer 5.

Referring to Figure 3, four different organism 2 arenas 7, A to D show examples of
organisms 2 which can have their behaviours recorded using the system 1
20 described hereinbefore. A is an example of insect larvae moving on a petri dish or
plate as the arena 7, B is an example of a top-down view of zebrafish in an
aquarium as the arena 7, C is an example of a roundworm on an NGM plate, in this
case with the food 9 L broth also shown, and D is an example of a rat in their arena
7. While these are some typical examples, any suitable organism 2 which is capable
25 of exhibiting behaviour can be recorded using the system 1 described earlier. Body
position and postures can be extracted using a variety of standard well-known
image analysis tools such as background subtraction, thresholding, edge detection
and the like to isolate the body of the organism 2. These positions and postures can
be combined with time data to determine various trajectory information such as
30 path, speed, and velocities of the organism 2.

Behavioural features for quantitative phenotyping – Introduction to measuring phenotypes

Measuring phenotypes is essential in most areas of biology, but there are no rules
35 that determine which aspects of a phenotype to focus on. This has led to calls for
more exhaustive characterizations of phenotype under the umbrella term
phenomics. Imaging is well suited to phenomics because images can capture

complex morphological differences and videos provide a natural extension to measure dynamics. However, the raw pixel intensities in images do not map directly to most quantities of interest and so an element of choice in representation remains. The success of deep learning approaches demonstrates the usefulness of automatically learned features on image analysis problems. However, the tasks solved in deep learning have well-defined objectives such as minimizing cross-entropy loss. When the objective is scientific understanding or hypothesis generation, high performance depends not just on accuracy but on interpretability and the nonlinear combination of many features through neural networks does not typically lead to highly interpretable features. The inventor's objective is to find a middle ground using a range of interpretable features optimized to quantify *Caenorhabditis elegans* morphology and behaviour.

In *C. elegans*, phenotyping morphology and behaviour, both of which are readily captured using imaging, have a long history dating back to Brenner's original paper describing the isolation of the first visible mutants (Brenner S. 1974 The genetics of *Caenorhabditis elegans*. Genetics 77, 71–94). Subsequent work has pioneered the quantitative analysis of nematode behaviour in *C. elegans* and other species. A desire to increase throughput, sensitivity and the ability to quantify multiple phenotypes from the same recording have led to the development of many worm trackers over the intervening decades. A previous tracker could track the centroid position of 25 worms in real time at 1 Hz. The tracker was used to study oxygen and carbon dioxide responses and responses to a variety of chemicals.

The next generation of trackers were used to quantify speed or the behavioural components of chemotaxis, still at relatively low resolution. High-resolution single-worm trackers were developed first simply to record single animals for long periods for subsequent manual annotation of egg laying, but were quickly adapted for use in high-dimensional quantitative phenotyping. Throughput was increased by using multiple single-worm trackers in parallel. At the same time, multi-worm trackers that tracked many animals at lower resolution were developed to increase throughput using a single camera. Improvements in camera technology eventually led to the development of a high-resolution multi-worm tracker that operates in real time and records worm outline and skeleton at 30 Hz. New trackers continue to be developed for specific applications or with new features. The inventor has also recently developed a high-resolution multi-worm tracker to store not just worm outline and skeleton but also worm pixels to achieve compression without losing

information about worms and their surroundings. Keeping a portion of the image data enables reanalysis using improved computer vision algorithms or manual annotation. In summary, there is no shortage of methods for collecting worm behaviour data and several options for quantifying behavioural phenotypes. Given
5 the large set of possible approaches and features, a principled way of selecting useful features would be helpful. Herein, the inventor introduce a set of handcrafted features that can be measured from single or multi-worm tracking data, provided there is sufficient resolution to quantify worm posture. The features are inspired by phenomics to be as exhaustive as possible, but with an explicit bias towards
10 interpretability to support exploratory analyses, generate hypotheses and guide mechanistic studies. The inventor uses a large database of videos of mutant worms and wild isolates to select feature subsets that balance explanatory power and interpretability. The inventor also analyses a new set of optogenetics experiments and show that the same feature set can be used to find differences in behavioural
15 dynamics that reveal a range of behavioural responses to optogenetic stimulation of different neural circuits in worms.

Behavioural features for quantitative phenotyping – Material and methods

The data from the mutants and wild isolates are from two previously published
20 studies and are available online from the OpenWorm Movement Database community page on Zenodo <https://zenodo.org/communities/open-worm-movement-database/>. As described previously, the worms in these videos were young adults recorded for 15 min crawling on agar on a patch of Escherichia coli OP50 which serves as a food source. Worms were allowed to habituate to the
25 tracking plates for 30 min before recording.

The optogenetics experiments were performed on young adults on OP50 that were allowed to habituate for 30 min prior to recording. Worms were recorded for 7 min without perturbation and then stimulated with blue LEDs (peak intensity at 467 nm)
30 with five 5 s pulses and one 90 s pulse, each separated by 60 s. Tracking plates were prepared with 300 ml OP50 liquid culture mixed with all-trans-retinal (ATR) dissolved in ethanol to a final plate concentration of 83 mM ATR (0.25 ml ATR each 300 ml of OP50). Control plates were prepared identically but using 100% ethanol without ATR. Plates were left for 48 h to dry with the lids on and stored for up to 5
35 days at 48 °C. See table 2 for a list of strains used in the optogenetics experiments. All worms were segmented, tracked and skeletonized using Tierpsy Tracker. [34]

Binaries, source code and documentation are available at
<http://ver228.github.io/tierpsy-tracker/>.

Behavioural features for quantitative phenotyping – Results

5 (a) Feature definition

For the initial parameterization of the worm, the inventor focused on defining features that cover as much of the range of observable phenotypes as possible while remaining interpretable. The inventor identified a range of feature classes that the inventor then characterized by defining multiple features for each class to
10 ensure phenotypic breadth. Following previous efforts at handcrafting features for quantifying *C. elegans*, the classes cover morphology, path, posture, and relative and absolute velocities. Interpretability is subjective and depends on the assessor's background. As a rule of thumb, the inventor considered more derived features—that is, features requiring more computational steps or longer algorithms to
15 calculate—to be less interpretable. However, for the initial parametrization, the inventor erred on the side of including more features because if a feature is subsequently found to be important for detecting strain differences, it could be worth sacrificing some interpretability.

20 The starting features are shown schematically in Figures 4 to 6 and are described in more detail in Javer et al., 2018, "Powerful and interpretable behavioural features for quantitative phenotyping of *Caenorhabditis elegans*", Phil. Trans. Roy. Soc. B 373: 20170375 and at
<https://royalsocietypublishing.org/doi/suppl/10.1098/rstb.2017.0375>.

25

(b) Feature expansion

To further increase the breadth of phenotyping, the inventor performs a series of operations that expands the total number of features (Figure 7). First, any feature that can be localized to a part of the body, for example, curvature, is calculated
30 separately for five segments along the worm (colloquially: head, neck, midbody, hips and tail). Velocities are calculated additionally at the head tip and tail tip because motion at the extremities, especially the head, is often informative. Second, the inventor calculates the derivatives of any time series features (i.e., features that are calculated in each frame). For example, the inventor calculates
35 the rate of change of curvature for each segment. Third, the inventor subdivides features according to motion state (forward, backward and pause) leading to features such as midbody curvature during reversals. Finally, the distributions of

these features are quantified by calculating the 10th percentile, median and 90th percentile values.

5 The final phenotypic fingerprint thus derived has 4083 values. Although they are derived in several steps, they remain describable in words. For example, the feature curvature_tail_w_forward_abs_90th is the 90th percentile of the absolute value of the tail curvature measured while the worm is moving forwards. The large number of features belies an underlying simplicity: there are only 16 basic features, each subjected to similar operations during expansion.

10

(c) Feature selection

Feature selection is useful to (i) remove noisy or irrelevant features and (ii) choose one from a set of highly correlated and therefore redundant features. Removing irrelevant features can improve performance while removing highly correlated
15 features reduces the complexity of the representation without hurting performance. The features defined above involve relatively simple computations and were chosen based on the inventor's prior notions of what would be relevant for worm phenotyping so the inventor does not expect many features to fall into the first category. On the other hand, the expansion procedure is likely to produce sets of
20 correlated features that capture redundant information about the phenotype. In any case, relevance and redundancy are defined with respect to a particular dataset. The inventor therefore chose to quantify the usefulness of the full set of features on a classification task on diverse previously published datasets consisting of a total of 11 406 individual worms from 358 strains drawn from mutants affecting
25 neurodevelopment, synaptic and extra-synaptic signalling, muscle function and morphology as well as wild isolates representing some of the natural diversity of *C. elegans* strains around the world.

As a preprocessing step, the inventor cleaned the data by removing any feature
30 assigned as not a number (NaN) in more than 2.5% of the worms in the full set. On this dataset, only paused motion state features were eliminated because in 14.3% of the videos, the worms never paused and therefore, the subdivision is not defined. Any remaining NaN values are imputed using the population mean value of the corresponding feature. The inventor then z-normalized the data by subtracting
35 the feature mean and dividing by its standard deviation to make features with different units comparable on the same scale.

The inventor divided the data into training and validation sets by randomly splitting the data per strain into 80–20% for training and testing, respectively. The inventor then used recursive feature elimination to identify useful feature sets using the following procedure: (i) the inventor fit a logistic regression model using stochastic
5 gradient descent on the categorical cross-entropy loss. (ii) Each feature is ranked in importance by removing it from the fitted model and calculating the change in the loss. More important features will increase the loss when removed, while less important features will have little effect or even decrease the loss. (iii) The least
10 important features are dropped until the next power of two is reached, e.g. if there are 3000 features, 952 features will be dropped leaving 2048, or 211. The inventor repeated this procedure 10 times for different random subsets of worms and plotted the classification accuracy as a function of feature number for the inventor newly defined features (Tierpsy features), the previously defined features from Yemini et al. (Yemini E, Jucikas T, Grundy LJ, Brown AEX, Schafer WR. 2013 "A
15 database of *Caenorhabditis elegans* behavioural phenotypes". Nat. Methods 10, 877–879) and the combination of both feature sets (see Figure 8a).

The Tierpsy features perform better than the Yemini et al. features (peak accuracy of 56.21% at 1024 features compared to 47.82% at 256 features). The combined
20 feature set shows the best performance, but the improvement over the Tierpsy features is small. This suggests that the Tierpsy features capture almost all of the phenotypic information in the Yemini et al. features as well as some new information that was missed. There is no drop in performance (and even a slight increase) as the first several thousand features are eliminated. The shape of the
25 accuracy curve is highly reproducible when different subsets of worms are used for classification while the identity of the most useful features is highly variable, suggesting that many of the features in the total set are redundant. That is, within a set of correlated features, it makes little difference to classification accuracy which feature is kept, and which are dropped.

30 The features that are selected also depend on nature of the strains used for feature selection. For example, if only wild isolate strains (rather than the full set of mutants which contain several strains with severe locomotion defects) are used for feature selection, the performance on classification on the mutant data is reduced.
35 By contrast, if all strains or only mutants are used in feature selection, the performance on classifying wild isolates is unaffected. This may be because the nature of the variation between wild isolates is represented by the differences

between some mutants, whereas there are differences between mutants that are not observed in wild isolates. This supports the choice of using a set of strains with a broad range of phenotypes in feature selection if the goal is to find a feature set that has the best chance of generalizing to unseen worm strains.

5

While classification accuracy does not allow us to prioritize features within correlated groups, interpretability can provide a guide. To bias the results towards simple interpretable features, the inventor started by eliminating classes of features and steps in the feature expansion. The inventor found that removing derivatives and the subdivision by motion state both significantly reduced classification accuracy confirming that these are useful operations (Figure 8b). However, the inventor found it was possible to eliminate eigenworm features and the normalization by worm length, and to use only the absolute value of features that had previously been signed as positive or negative based on dorsoventral orientation (e.g. curvature was originally defined as positive or negative depending on whether the body bend was dorsal or ventral). Together, removing these features reduces the total number by almost half with no detectable effect on accuracy. The inventor labels this reduced set of features the Tierpsy_2 k.

10

15

20

25

30

35

For accurate classification or for clustering applications where the full spectrum of differences and similarities would be useful, the inventor recommends a reduced set of 256 features, which the inventor labels the Tierpsy_256 (electronic supplementary material, table S1), that balances completeness and compactness of the representation. Many phenotyping tasks occur on a smaller scale, with just a handful of strains compared to a reference (such as several mutants compared to a wildtype strain). If there are specific hypotheses for relevant phenotypic differences, having a large number of features to choose from makes it more likely that the hypotheses will be testable without having to code new features. However, for exploratory work, 256 features can lead to a larger number of differences than are needed to guide experiments and the large number of features increases the burden of multiple testing corrections. The inventor has therefore used a combination of classification power, subjective interpretability and coverage of feature classes to define the Tierpsy_8 and Tierpsy_16 (table 1), which give classification accuracies of 20.37+ 0.41% and 28.67+0.45%, respectively (mean+standard deviation).

Pre-selecting these smaller sets of features before performing a new analysis reduces the multiple testing burden and results in phenotypic fingerprints that can be visualized and understood at a glance (Figure 9).

5 (d) Direct analysis of time series for optogenetic experiments

The feature expansion and summarization above captures some dynamic aspects of phenotype, but is intended for comparisons where the relevant differences are not localized in time and could occur at any point during a recording. For optogenetic experiments where the stimulation has a clear start and stop, it makes more sense
10 to align time series and look for differences directly rather than summarizing the entire experiment in a feature vector.

As above, the inventor wanted to analyse data with a range of phenotypic differences. The inventor therefore collected data from 11 strains expressing
15 channel rhodopsin in different neural subsets (table 2). The inventor separate the data for long segment) and calculate histograms for each feature in each set. The behaviour may vary during or after a pulse and therefore, it is useful to use multiple time bins to capture transient effects. The inventor pooled the histograms for each strain with ATR and the controls without ATR. Because worms do not make
20 ATR, any behavioural effects observed in the no-ATR condition are more likely to be generic blue light responses rather than ChR2-specific effects. The blue light response is most clearly observable during the 90 s pulses.

A selection of responses to 5 s optogenetic activation are shown in Figures 10 and
25 11. There are clear differences between treatment and controls in several features for the three strains expressing ChR2 in different neural circuits. To systematically find features that respond differently to blue light between treatment and no-ATR controls, the inventor calculated the Jensen–Shannon divergence between the treatment and control distributions for each feature and used a permutation test to
30 determine a p-value for the comparison. Finally, the inventor corrected for multiple comparisons within a given strain using the Benjamini–Hochberg procedure to control the false discovery rate at 0.05. The results are summarized in figure 6. Only five strains show p-values smaller than 0.05 after correction (AQ2235, AQ2052, AQ2232, HBR180, HBR520) in the short pulses (figure 6a). On the other
35 hand, three additional strains show differences (HBR222, HBR520, AQ2050, AQ2028) when analysing the long pulse data (figure 6b), suggesting a slower response to optogenetic stimulation in these strains. Finally, it is worth noting that

the three strains that did not show a significant difference compared to the control do not have a *lite-1* deletion mutation, so it is possible that an underlying optogenetic effect is masked by their aversive blue light response. The permutation tests show that behaviour is significantly affected by optogenetic stimulation in several strains, but do not show how the features change to allow a comparison between strains. In order to quantitatively compare the behavioural responses between strains, the inventor calculated a distance matrix between each strain in each condition (control and ATR) using the Jensen–Shannon divergence among the corresponding feature histograms. The results for the samples with ATR are shown in Figure 13 and 14. Two-thirds of the strains are close to N2 in the long pulses but fewer than half are close to N2 in the short pulses. The observed clustering pattern suggests that there is a range of distinct behavioural phenotypes induced by the optogenetic stimulation of different neural subsets. The equivalent plots for the control plates are shown in electronic supplementary material, figure S4 of Javer et al., 2018, “Powerful and interpretable behavioural features for quantitative phenotyping of *Caenorhabditis elegans*”.

Behavioural features for quantitative phenotyping – Discussion

The features the inventor has defined here are intended to provide a range of options to balance power and interpretability in behaviour representation from a small set of easily interpretable features to a large set of features that provide improved classification accuracy compared a previously defined set of handcrafted features. Using a diverse set of worm behaviour data from mutants and wild isolates, the inventor found, as expected, that the inventor’s feature definitions led to groups of correlated features that contained redundant information. The inventor took advantage of this redundancy to favour interpretable features over equally useful but less interpretable features.

The inventor found that two categories of features could be eliminated entirely with little effect on classification accuracy: those derived from the distribution of eigenworm amplitudes and those based on dorsoventral asymmetries. There is no contradiction between the elimination of eigenworm features and their usefulness in other applications. The implication of this result is simply that the information present in the distribution of the eigenworm amplitudes taken across an entire video is captured using other more interpretable features. The eigenworm representation remains useful for many other applications, especially where an understanding of postural time series is important, rather than phenotypic

summary based on posture distributions each video into short pulses (five 5 s long segments) and long pulses (one 90 s calculated for an entire video. Similarly, the known asymmetry between dorsal and ventral turns will clearly be important in some studies. It is just that on average, for distinguishing worm strains, it is not a critical distinction and most of the information is present in the symmetrized data. This is a positive result for multi-worm tracking data where the dorsoventral orientation of worms is difficult to determine. The inventor's results suggest that the Tierpsy features will be as useful for high-resolution multi-worm tracking as they are for single-worm tracking.

The data the inventor used to perform feature selection cover a range of mutant phenotypes and the natural variation of wild isolates. The inventor's goal was that the feature subsets the inventor selected will be useful for capturing as complete a range of phenotypic variation as possible. For applications where interpretability is paramount, they provide a useful starting point. However, for any new application with a different kind of phenotypic variation such as new mutants or worms in different experimental conditions, a different subset of features could be more appropriate. Therefore, for applications where there are sufficient data to use a training set on feature selection and a hold-out set for testing, the inventor would recommend repeating the feature selection procedure starting from the full set of features or the Tierpsy_2 k. Alternatively, if prior knowledge or specific hypotheses suggest a certain class of features is important, manually selected features can be simply added to one of the smaller feature subsets to capture the relevant effect without unduly increasing the burden of multiple testing.

Behavioural features for quantitative phenotyping – Conclusion

A critical step in any phenotyping project is choosing the right representation for the problem at hand. Phenomics is based on the assumption that the right representation can be difficult to determine a priori, and that it is therefore useful to measure phenotypes as exhaustively as possible. The inventor has adopted this approach to define a large number of behavioural features that are then selected based on how well they explain data from a diverse set of strains and based on a subjective assessment of their interpretability. The direct parameterization of behaviour the inventor describes here is just one approach to the larger problem of the quantitative analysis of behaviour characteristic of computational ethology. The inventor has found that this approach has reasonably good power to detect subtle

behavioural differences and is particularly useful in cases where interpretability is paramount.

One of the reasons the phenomic approach to behaviour analysis is useful is that animal behaviour is complex and the effects of perturbations are difficult to predict. The same is increasingly true for artificial agents controlled by artificial neural networks, which has led to calls for an ethology of artificial agents to understand their behaviour. When applied to understanding the output of increasingly detailed simulations of *C. elegans*, the inventor believes that a high-dimensional representation of behaviour will be essential to provide enough constraints to fit model parameters that are difficult to estimate directly from experiments. The inventor proposes that the features defined here are a useful starting point for performing quantitative model validation for cellular-level simulations based on the *C. elegans* connectome.

15

Megapixel camera arrays enable high-resolution animal tracking in multiwell plates

With reference to Barlow et al. "*Megapixel camera arrays enable high-resolution animal tracking in multiwell plates*", Communications Biology, 2022, 5:253.

20

Tracking small laboratory animals such as flies, fish, and worms is used for phenotyping in neuroscience, genetics, disease modelling, and drug discovery. An imaging system with sufficient throughput and spatiotemporal resolution would be capable of imaging a large number of animals, estimating their pose, and quantifying detailed behavioural differences at a scale where hundreds of treatments could be tested simultaneously. The inventor presents an array of six 12-megapixel cameras used to record all the wells of a 96-well plate with sufficient resolution to estimate the pose of *C. elegans* worms and to extract high-dimensional phenotypic fingerprints similar to those described hereinbefore. This system, which may be similar to the system 1 described hereinbefore can be used to study behavioural variability across wild isolates, the sensitisation of worms to repeated blue light stimulation, the phenotypes of worm disease models, and worms' behavioural responses to drug treatment. Because the system is compatible with standard multiwell plates, it makes computational ethological approaches accessible in existing high-throughput pipelines.

35

Recording and quantifying animal behaviour is a core method in neuroscience, behavioural genetics, disease modelling, and psychiatric drug discovery. Both the scale of behaviour experiments and the information that can be extracted from them have increased dramatically. However, further increases in throughput are possible and would enable entirely new kinds of experiments. Therefore, there is a need for a system to image freely behaving animals that would maximise both phenotypic content and experimental throughput. In terms of phenotypic content, a key parameter is the resolution of the recording. If there is sufficient spatial and temporal resolution, then body parts can be identified and tracked, the animal's pose can be estimated, and the full suite of computational ethology methods can be applied to analyse any behaviour of interest. Because of its simple morphology, detailed pose estimation is well-established for the roundworm *C. elegans* and previous work has shown the usefulness of detailed behavioural phenotyping in several domains including, for example, classifying mutants, studying chemotaxis and thermotaxis, quantifying escape responses, and addressing basic questions in computational ethology and the physics of behaviour. Maintaining sufficient resolution for pose estimation was therefore the first design constraint the inventor required. In early worm trackers, maintaining high resolution required a motorised stage to keep a single worm in the field of view of a low-resolution camera, but the availability of inexpensive megapixel cameras enabled multiworm tracking with sufficient resolution to estimate each worms' pose and determine its head position. High spatial resolution and throughput has been demonstrated using flatbed scanners to quantify worm lifespans and behavioural decline with age. However, the low temporal resolution (1 frame per 15 min) precludes the detection of behavioural phenotypes which happen at shorter time scales.

To maximise experimental throughput, off-the-shelf multiwell plates were used so that any behaviour screening pipeline would still be compatible with existing pipeline elements, such as liquid and plate-handling robots as well as small animal sorting machines. Because behaviour occurs over time, a standard plate-scanning approach in which each well of a multiwell plate is imaged in turn using a motorised stage limits throughput regardless of scan speed since each well must be recorded long enough to observe the behaviour of interest. Moreover, mechanical arrangements with moving parts introduce higher maintenance costs and have a higher risk of failure compared to a static camera system. Therefore, the second design constraint was that the system should be able to image all of the wells of a multiwell plate simultaneously without move parts.

The inventor's solution to simultaneously image a large area with high resolution was to use an array of machine vision cameras that are small enough to be arranged in close proximity to one another with partially overlapping fields of view at a resolution sufficient to track small animal pose. Here the inventor presents the design of an array of six 12-megapixel cameras that uses a near- infrared light panel for illumination, a set of high-intensity blue LEDs for photostimulation, and the associated open-source software for automatically identifying wells and keeping track of metadata. The software is fully integrated into the inventor's existing Tierpsy Tracker software (Javer, A. et al. An open-source platform for analysing and sharing worm behaviour data. Nat. Methods 15, 645–646 (2018)), including a graphical user interface for reviewing tracking data, joining trajectories, and annotating problematic wells to discard from analysis, as well as a neural network for distinguishing worms from non-worm objects. The inventor demonstrates the potential of the new tracking system in neuroscience, disease modelling, genetics, and phenotypic drug screening.

Results

Megapixel camera array design. Based on the inventor's previous work with single-worm tracking, the inventor set a target resolution of at least 75 pixels/mm and a recording rate of 25 frames per second in order to accurately estimate worm pose and identify the head from the tail which requires the measurement of the small head swinging behaviour that is often referred to as 'foraging' in the worm community. These constraints require a total of 8100×5400 pixels, or about 44 megapixels with a 3:2 aspect ratio, to cover a standard 96-well plate. Single cameras with this resolution that can record at 25 frames per second are not available commercially. The inventor therefore considered arrays of cameras and found that six Basler aca4024 cameras (Basler AG, Ahrensburg, Germany) in a 3×2 array equipped with Fujinon HF3520-12M lenses (Fujifilm Holding Corporation, Tokyo, Japan) was an optimal solution: This combination of lenses and cameras allowed mounting the cameras in close proximity, whereas cameras with higher resolution or larger sensors would have required significantly larger lenses. Imaging a multiwell plate with multiple cameras significantly reduces the blind spots caused by vertical separators between wells, compared with using a single camera with a conventional lens. A similar effect could be obtained by using a single camera with a telecentric lens, but the multi-camera approach remains a more compact and cost-effective way of achieving the required resolution. To provide uniform

illumination whilst mitigating light-avoidance response, the inventor used a dedicated light system using 850 nm LEDs (Loopbio GmbH, Vienna, Austria, see Material and Methods for more details). Blue light stimulation is provided by a custom LED array (Marine Breeding Systems, St. Gallen, Switzerland) using four

5 Luminus CBT-90 TE light-emitting diodes (bin J101, 456 nm peak wavelength, 10.3 W peak radiometric flux each) with user-tuneable intensity. The camera lenses are equipped with long pass filters (Schneider- Kreuznach IF 092, Schneider- Kreuznach, Germany, and Midopt LP610, Midwest Optical Systems Inc, Palatine, IL, USA) to block the photostimulation light while allowing the brightfield IR signal to

10 reach the camera sensors. A schematic of the imaging system is shown in Figures 15a–c. To further increase throughput, the inventor built five units of the camera arrays which can be operated in parallel (Kastl-High-Res, Loopbio GmbH, Vienna, Austria).

15 Choice of suitable multiwell plate design. Because the fields of view of the six cameras partially overlap, the imaging system provides flexibility in selecting a multiwell plate with any number of wells. For the present purposes, 96-well plates with square wells provided a good balance between imaging area and the number of wells (Fig. 15d, e). Plates with smaller numbers of wells would reduce imaging

20 throughput, while 96-well plates with circular wells would reduce the area available for worms to behave and increase shadowing around the well edges (Supplementary Figures 1a, b of Barlow et al. 2022, “Megapixel camera arrays enable high-resolution animal tracking in multiwell plates”, Communications Biology volume 5, Article number: 253). Using square well plates (Whatman 96 well plate

25 with flat bottom, GE Healthcare, Chicago, IL, USA) significantly increases the efficiency of the system: in the inventor’s tests, in plates with circular wells only 21% of the imaging area is available for capturing useful data, while the rest is outside of any well or lost in shadows. For square wells, 43% is available for behaviour. This has important implication for tracking and experimental design, as

30 not all worms placed in a well will be always visible. For example, when imaging worms of the N2 control strain, all worms in a well are simultaneously tracked 40% of the time, and this figure can depend on the strain, with worms of the wild isolate strain CB4856 being all visible simultaneously only 9% of the time (Supplementary Fig. 5 of Barlow et al. 2022). The fraction of the imaging area available for tracking

35 can be further increased by using custom plates with thinner wall dividers and shallower wells to reduce the shadowing (Supplementary Fig. 1c), but this comes at an increased cost of manufacture.

The output of the combined system is 30 videos tiling across the five imaged multiwell plates corresponding to 480 simultaneous behaviour assays (Fig. 15d). Expanding the image to the level of a single well (Fig. 15f) and a single animal (Fig. 15g) shows that the resolution is sufficient to estimate the pose and identify the head of single worms, which can reveal detailed trajectory differences between individuals that are the basis for quantitative behavioural phenotyping.

High throughput imaging. Due to the high amount of raw image data produced by USB3 cameras at full bandwidth (6 cameras recording at 25 fps produce approximately 6.5TB/hour of raw footage), live compression on a dedicated system was required. To achieve this, the inventor used a total of 10 Motif Recording Units (Motif— Video Recording System, Loopbio GmbH, Vienna, Austria) equipped with Nvidia Quadro P2000 GPUs (Nvidia Corporation, Santa Clara, CA, USA), each recording from 3 cameras. The two recording units with cameras from the same system were set up in a parent-child configuration.

The Motif software acquires and compresses images on the fly and stores them in the open imgstore format (<https://github.com/loopbio/imgstore>) along with timestamps and frame numbers for each individual frame, as well as continuous and synchronised recordings of environmental data (for each unit this was: outside temperature and inside temperature, humidity, and light level). Recording the time and frame number for each image allows precise timekeeping over a long recording duration as it removes temporal drift due to skipped or dropped frames and due to differences in camera clocks. Additional metadata for each recording is saved with the imgstore, including the camera serial numbers, camera and system settings, and any user-defined data. A single workstation manages all imaging experiments on all units across the whole system, from experimental parameter tuning (including the intensity of the photostimuli) to video collection to data transfer, by accessing the Motif user interfaces using the web browser of the parent machines. Given the large number of high-resolution cameras, the control workstation was connected to a large monitor (the inventor uses a 43-inch 4k monitor) to facilitate camera focussing and sample positioning.

In addition to providing a web-accessible user interface, Motif allows complete control of the camera arrays and arbitrarily complex scheduling of data acquisition and photostimulation programmatically, via an API

(<https://github.com/loopbio/python-motifapi>). This allows us to run imaging experiments on all camera arrays by executing a single Python script on the monitoring workstation. Encoding the parameters of experiments in a script improves reproducibility by making settings consistent over time by default.

5

Updates to Tierpsy Tracker, and companion software, for multiwell imaging format. In the present camera array system, each camera records multiple wells which complicates metadata handling since there is no longer a one-to-one correspondence between a video file and a particular experimental condition. The

10 inventor has updated Tierpsy Tracker to handle videos with multiple wells: it can automatically identify wells from the video (Fig. 15f), and return results on a well-by-well basis. In the Viewer, the user can see the names and boundaries of the wells and have the option of marking any well as "bad" if necessary. This flag is propagated to the final tracking results so that the contribution of "bad" wells can
15 be filtered out for downstream analysis.

To keep track of the experimental conditions of each well, the inventor has developed an open-source module in Python to automatically handle experimental metadata (github.com/Tierpsy/tierpsy-tools-python). For each experiment, a series
20 of csv files specifies the worms and compounds (if applicable) that were added to each well. This can include information on how a COPAS worm sorter (Union Biometrica) was used to dispense different strains in the wells of the imaging plates, which compound source plate was replicated onto each imaging plate, or any column shuffling performed by a liquid handling robot. These tables are then
25 combined to create a mapping between each well in an imaging plate (identified by a unique ID) and an experimental condition. For each imaging run, the user needs to log the camera array used for each imaging plate at the time of the experiment. This information is then mapped to video file names to create a final metadata table suitable for subsequent analysis (see Material and Methods and Supplementary Fig.
30 3 of Barlow et al. 2022 for more details).

Another key software improvement the inventor incorporated is a convolutional neural network (CNN) to exclude non-worm objects from subsequent analysis. While the inventor previously used contrast-based segmentation and size-based
35 filtering for worm detection in the inventor's analysis, introducing the CNN into Tierpsy Tracker improves the quality of the tracking data and the subsequent

analysis results as well as the speed of the analysis because fewer objects are analysed in subsequent steps (see Methods for more details).

5 Tierpsy Tracker does not maintain the identity of the worms across gaps in tracking which can occur when worms leave the field of view or cross each other. The number of unique tracks is thus typically greater than the number of worms. While Tierpsy initially calculates features on a track-by-track basis, the inventor uses a single averaged feature vector per well because of the uncertainty in how the tracks map to individual worms. The natural unit of measurement of sample size with the
10 setup described in this work is therefore a well, rather than an individual worm.

Rapid assessment of natural variation in behaviour. The inventor tracked the behaviour of N2 and wild isolates of the divergent set in the *C. elegans* Natural Diversity Resource (CeNDR) strain collection with the present system to detect
15 natural variation in behaviour. To further increase the dimensionality of the behavioral phenotypes, the inventor included a blue light stimulation protocol using a set of four bright blue LEDs. Each tracking experiment is divided into three parts: 1) a 5-minute pre-stimulus recording, 2) a 6-minute stimulus recording with three 10-second blue light pulses starting at 60, 160, and 260 sec, and 3) a 5-minute
20 post-stimulus recording. Blue light can elicit an escape response in worms, thus expanding the range of observable behaviours. Programmable blue light stimulation is reproducible, compatible with high throughput assays, and is also useful for optogenetic stimulation.

25 The inventor tracked on average 20 wells per strain. Given the high throughput achieved with the present new system, this experiment can be performed within a few hours. The recordings of the camera array maintain enough resolution to extract the full set of Tierpsy features, which describe in detail the morphology, movement, and posture of the worms, including subdivision by motion mode
30 (forward, backward, and paused) and body part. The inventor extracts a set of 3076 summary features per well for each recording period (pre- stimulus, blue light, and post-stimulus), resulting in a total of 9228 features for each well. This allows us to detect fine differences in the morphology, posture and movement of the worms which varies in a nonuniform way among wild isolates (Fig. 16a-c). The
35 neck curvature of wild isolates tends to show more significant differences to N2 worms than the curvature of other parts of the body, which might be related to differences in foraging behaviour between N2 and wild isolates (Fig. 16b). However,

not all strains show the same curvature pattern across the body indicating natural variation in posture. All the wild isolates move on average faster than N2 worms but their response to blue light varies (with some being more and others less sensitive to blue light), showing that the blue light stimulus increases the
5 dimensionality of behavioural differences (Fig. 16c).

To assess how well a worm strain can be predicted based on its behavioural fingerprint, the inventor estimated the classification accuracy using a random forest classifier. The inventor first splits the data into a training/tuning set and a held-out
10 test set. The inventor used the training set to select features using recursive feature elimination (RFE) and tune the hyperparameters of the model. Figure 17 shows the highest cross-validation accuracy achieved for different sizes of selected feature sets. The accuracy improves when the inventor selects features increasing the samples-to-features ratio, as this helps control overfitting and, in parallel,
15 reduces the correlation between features. Combining features from different blue light conditions (blue curve) increases the dimensionality of the data and the classification accuracy. Using the best performing features and hyperparameters, the inventor trained a classifier with the entire training/tuning set and used it to make predictions in the test set. The test accuracy the inventor achieved is 66%
20 which is significantly higher than random (9%).

Temporal response and sensitisation to aversive blue light stimulation. Having established that blue light stimulation can be leveraged to improve classification accuracy, the inventor moved to further investigated the response elicited by blue
25 light in N2 and CB4856 at a higher temporal resolution.

The inventor imaged with blue light stimulation on both the N2 and CB4856 strains, and observed different behavioural responses between these two strains. The inventor extracted the same set of 3076 previously defined features with a time
30 resolution of 10 sec and used these to construct the behaviour phenotype space. N2 and CB4856 have well-known behavioural differences and are expected to occupy different regions of the phenotype space. Principal component analysis (PCA) shows that application of blue light stimulation moves the strains from their already distinct positions in the plane defined by the first two principal components (PCs) to new positions, indicating detectable responses to the stimulus in both strains (Fig.
35 19a). Blue light- induced displacement through phenotype space led to better separation between the two strains (Fig. 19a, right), confirming that the addition of

the stimulus can reveal further behavioural differences between two strains already known to be distinct.

The *C. elegans* escape response is characterised by a combination of increased forward locomotion and decreased spontaneous reversals. The differences in blue light-induced escape response between the N2 and CB4856 strains can thus also be seen by simply examining the fraction of the worm population moving forwards, moving backwards, or remaining stationary. For both strains, a single photostimulus triggers a sharp and steady increase in the fraction of worms moving forwards, followed by a relaxation towards the pre-stimulus level once the stimulus ceases. The fraction of worms moving backwards has a slight increase at the beginning of the stimulus, and then decreases without increasing again until the stimulus is over. Finally, the fraction of stationary worms declines rapidly during the stimulus and is restored after the stimulus ends (Fig. 3b). However, while in CB4856 the rate at which the population fractions return to the pre-stimulus levels is steady, N2 shows a sharp initial decline in forward-moving worms over several seconds (and a corresponding sharp increase in stationary worms and backward-moving worms) before relaxing steadily. Repeated photostimulation (twenty pulses of 10 s on, 90 s off) of N2 worms causes sensitisation, as light pulses trigger a progressively increasing fraction of worms to move forwards (Fig. 19a to d). Meanwhile, between light pulses, worms recover to a progressively decreasing baseline level of forward locomotion (and conversely, progressively increasing stationary fraction), possibly due to fatigue from increased activities during the pulse. This reduced forward locomotion fraction persists in the absence of photostimulation, with no obvious return towards pre-stimulus levels over a 6.5 min period after the final pulse (Fig. 19c, e). The combined effect of sensitisation and fatigue leads to a roughly linearly increasing response over multiple light pulses, as illustrated by taking the difference between the fraction of worms moving forward before and after stimulation (Fig. 19d).

Photostimulation can thus better distinguish between worm strains using existing predefined feature sets, as well as create new features for quantifying the details of the escape behaviour. Similar experiments on habituation to repeated mechanical stimulation have been used extensively to study learning in *C. elegans*. Aversive blue light stimulation acts through different sensory neurons and converges on the same motor circuits and so may provide useful comparative data to investigate the genetics and neuroscience of learning mechanisms.

The addition of these new and interpretable features increases the dimensionality of the worm behavioural phenotypic space, which may be useful for phenotyping applications.

5

Behavioural phenotypes of ALS disease models in response to blue light. A previous study generated several Amyotrophic Lateral Sclerosis (ALS) disease model strains that carry patient amino acid changes in the *C. elegans* *sod-1* gene. This study found that the disease model strains have no obvious behavioural defects unless they are exposed to oxidative stress by overnight treatment with paraquat.

10

The inventor phenotyped these ALS disease model strains on the present system and saw similar results. PCA of a pre-defined set of 256 Tierpsy features under standard imaging conditions (5 min of spontaneous behaviour) does not show clear differences between the strains (Fig. 20a). Adding blue light pulses (three 10-second blue light pulses over six minutes) leads to better separation between the strains in PC space (Fig. 20b). Although the SOD-1(+) wild-type control strain (blue) and the SOD-1(A4V) mutant disease strain (orange) clearly separate into their own clusters, SOD-1(H71Y), SOD-1(G85R) and SOD-1(0) null strains cluster together, suggesting that their overall responses to blue light are similar to each other. The clustering of SOD-1(H71Y), SOD-1(G85R), and SOD-1(0) strains upon blue light stimulation is consistent with the previous finding that all three strains have loss of *sod-1* function in glutamatergic neurons. By contrast, the SOD-1(A4V) strain has overexpression of *sod-1* in cholinergic neurons without affecting glutamatergic neurons⁴⁵, and this disease strain forms its own cluster in the blue light PC space (Fig. 4b).

15

20

25

In the previous study and in the inventor's results, no difference is observed in the baseline behaviour of the strains. However, exposure to aversive conditions, possibly acting through very different mechanisms, highlights the difference between strains by exposing an otherwise cryptic phenotype. Upon blue light stimulation, SOD-1(H71Y), SOD-1(G85R), and SOD-1(0) strains show significantly bigger increases in forward locomotion compared to the SOD-1(+) control strain and the SOD-1(A4V) disease strain (Fig. 20c,d). This increase in forward movement appears to be primarily at the expense of stationary (Fig. 20e) rather than backwards locomotion (Supplementary Fig. 2 of Barlow et al. 2022). Nevertheless, a closer look at reversal frequencies at a finer temporal resolution reveals

30

35

decreased reversals in the three *sod-1* loss-of-function strains but not the other two strains (Supplementary Fig. 2b of Barlow et al. 2022).

Phenotypic screen of human-approved drugs. The inventor used a library of 245
5 drugs to quantify worms' responses to human-approved drugs across multiple
behavioural features. 240 drugs were from the Prestwick *C. elegans* library, a
collection of small-molecule out-of-patent drugs selected by the supplier (Prestwick
Chemical, Illkirch-Graffenstaden, France) for their chemical structural diversity and
good tolerability in worms. To these, the inventor added a set of 4 antipsychotics
10 and an insecticide that have a strong phenotype that the inventor could use as
positive controls and for checking the automated metadata handling code. Three
worms were added to each well of 96-well plates and were left on the drug for four
hours before imaging. The inventor processed the videos using Tierpsy Tracker,
extracting 3016 behavioural features from each imaging condition (pre-stimulus,
15 blue light stimulus, post-stimulus) and concatenated the feature vectors so that
each well was represented by a 9048-dimensional feature vector. The inventor used
a linear mixed model to identify compounds that had a significant effect on
behaviour in at least one feature. The linear mixed model used the imaging day as
a random effect to account for day-to-day variation in the data. The 153
20 compounds that had a detectable effect were kept for further analysis. The inventor
restricted the feature space to a subset of 256 features (Tierpsy256) that the
inventor selected in the previous work for each imaging condition, so that each well
was now represented by a 768-dimensional feature vector. The features were then
z-normalised and both features and samples were hierarchically clustered using
25 complete linkage and correlation as the similarity measure (Fig. 21a).

The compounds in the library are mostly well-characterised with known modes of
action. By examining clusters in detail, the inventor found several clusters that
included multiple compounds from the same mode of action (Fig. 21b-d). One of
30 the identified clusters contains several compounds that inhibit muscle contraction
and/ or are related to vasodilation (Fig. 21b). The most clearly defined cluster (Fig.
21) contains antibiotics, and most of them share the same mode of action
(ribosomal protein synthesis inhibition). Because the worms were imaged on a lawn
of bacterial food, the most likely cause of these behavioural differences is a change
35 in the bacterial food lawn that worms sense and respond to, but a direct effect on
the worms is not impossible since *C. elegans* do respond to some antibiotics. A third
cluster is formed by compounds used to treat the symptoms of Parkinson's

Disease(anticholinergics). Previous studies have shown that anticholinergic drugs can affect locomotion in *C. elegans* and also induce motor activity in other model animals.

- 5 Most of the compounds had a detectable effect on behaviour, but many of the effects were less obvious than a library of invertebrate-targeting compounds that the inventor screened recently using the same method. A part of the explanation is likely to be a lack of conservation of some drug targets between humans and worms, although it should be noted that many are sufficiently conserved that
- 10 human-targeted drugs have effects through the expected receptor class. Another reason some compounds do not have a detectable effect is drug uptake which is known to be an issue for drug screens in worms, highlighting the continuing need for improved drug delivery to maximise the usefulness of worms in drug screening.

15 Discussion

- The inventor has developed a megapixel camera array system to enable high throughput, high content imaging of worms in standard multiwell plates. By partially overlapping the fields of view of six cameras, the inventor can image an entire multiwell plate at spatial and temporal resolutions that are sufficient for
- 20 tracking *C. elegans* and extracting high-dimensional phenotypic fingerprints. While the experiments presented in this work were carried out in 96-well plates, the imaging system can easily support 24- and 48-well plates as well. The inventor has added features to Tierpsy Tracker to make it compatible with the multiwell imaging format, so that each well is detected and analysed separately. The inventor
- 25 incorporated strong blue LED lights into the camera array system to provide precise and repeatable photostimulation and found that this leads to better separation between wild isolates and ALS disease model strains, in the latter case revealing phenotypes that could not be detected in standard unstimulated assays. Repeated blue light stimulation also revealed a novel sensitisation phenotype in N2 worms, in
- 30 marked contrast to the habituation in the reversal response reported in previous experiments on repeated mechanical stimuli which are used to study learning in worms.

- Our imaging hardware and analysis software are designed to support high
- 35 throughput phenotypic screening, as the multiwell format allows for a large number of experiments to be conducted simultaneously. Furthermore, the inventor's experimental pipeline uses liquid handling robots for dispensing agar, food, drugs,

and worms, in order to streamline the workflow for large-scale phenotypic screening. On a typical eight-hour imaging day, a single experimenter can operate five runs on all five-camera array units, thus collecting imaging data from 2400 independent wells in a 96- well plate format. Typical post-acquisition processing
5 time for this volume of data (assuming the standard 16 min video length at 25 fps, three worms per well) is 50-85 h using a MacPro (Processor: 2.7 GHz 12-Core Intel Xeon E5; Memory: 64GB 1866 MHz DDR3) or 5-11 h on a local cluster to go from raw video data to fully extracted behavioural features. Processing time increases significantly with object number and depends on the quality of the video (good
10 contrast, lack of debris, etc.).

Multiple cameras have previously been used to record large areas for worm tracking, but with an emphasis on long time recording at lower resolution compared to the applications presented here. Multicamera imaging systems have also been
15 used to record the behaviour of other species. When imaging animals in the field, for example, experimenters have employed multiple cameras to simultaneously image locations of interest. The most common scenario sees the use of multiple cameras pointed at the same animals for tracking position and pose in three dimensions. Multiple camera systems have also been used to increase coverage and
20 throughput in *Drosophila* imaging experiments.

There are several options available to record data from multiple cameras, from the video acquisition tools and SDK provided by camera manufacturers to third-party tools like The Recorder (MultiCamera.Systems LLC, Houston, TX, USA) or the open-
25 source Micro-Manager. The introduction of the GenICam Standard, developed by the European Machine Vision Association, has provided developers with a common programming interface that is not strictly coupled to the interface technology of the different cameras, thus making it easier to develop user interfaces.

30 Setting up a user-friendly multicamera system from scratch still requires a considerable investment in terms of time and know-how. A Motif camera system provides a compact imaging setup and a computer system with matched specifications and includes source code in Python, making it more transparent than some other proprietary systems.

35 A main strength of the inventor's camera array system is its scalability. Screening throughput can be readily expanded with additional imaging units, as the system is modular, and each camera array has a relatively small physical footprint. On-the-fly

compression of raw videos provided by the software keeps the data volume manageable. Post-acquisition analysis is easily parallelised since videos can be analysed independently and processing time can be decreased linearly by allocating more computational cores to the task (e.g. by using a high-performance cluster).

5

The megapixel camera arrays the inventor describes here represent a natural progression in worm tracking hardware where advances in the past have come from multiplexing to increase throughput and increasing resolution to get more information from multi- worm trackers. The inventor's system may allow higher throughput screening with a resolution that enables the full suite of computational ethology tools to be brought to bear on phenotyping. The inventor anticipates this will open new directions in large- scale behaviour quantification with applications in genetics, dis- ease modelling, and drug screening.

10

15 *Megapixel camera arrays screening - Methods*

Worm strains. *C. elegans* strains used in this work are listed in Supplementary Table 1. Worms are cultured on Nematode Growth Medium (NGM) agar at 20 °C and fed with *E. coli* OP50 following standard procedures.

20 Standard phenotyping assay. The standard phenotyping assay was used for most experiments in this work unless otherwise noted (detailed protocol: <https://doi.org/10.17504/protocols.io.bsicncaw>). See Supplementary Table 2 for the detailed protocols used to collect the data shown in each figure panel.

25 Briefly, Day 1 adult worms were obtained by bleach-synchronisation (detailed protocol: <https://doi.org/10.17504/protocols.io.2bzgap6>) and used for all imaging experiments. Imaging plates were prepared by filling 96 well plates with 200 µL of low peptone (0.013% Difco Bacto) NGM agar per well using an Integra VIAFILL reagent dispenser (INTEGRA Biosciences Ltd, UK) (detailed protocol: <https://doi.org/10.17504/protocols.io.bmxbk7in>), and stored at 4 °C until use. On the day before imaging, plates were placed in a LEEC BC2 drying cabinet (LEEC Ltd, Nottingham, UK) to lose 3–5% weight (starting weight 59 g without lid, target weight 56 g, this takes between 2 and 3 h). Each plate was then seeded with 5 µL per well of 1:10 diluted OP50 using VIAFILL, and stored at room temperature overnight.

35

On imaging day, synchronised Day 1 adults were washed in M9 (detailed protocol: <https://doi.org/10.17504/protocols.io.bfqbjmsn>) and dispensed into imaging plate wells using COPAS 500 Flow Pilot worm sorter (detailed protocol: <https://doi.org/10.17504/protocols.io.bfc9jiz6>). Three worms were placed into each well unless noted otherwise. Plates were returned to a 20 °C incubator for 1 h to dry following liquid handling, and then placed onto the multi-camera tracker for 0.5 h to acclimatise prior to image acquisition.

Drug experiments. Drug experiments followed the standard phenotyping assay workflow, but with a few modifications. A detailed protocol can be found at <https://doi.org/10.17504/protocols.io.bs6znhf6>.

Briefly, imaging plates were prepared with drugs the day before imaging and stored in the dark overnight at 4 °C. Using a COPAS 500 Flow Pilot, three worms were dispensed into each well of 96-well plates. Following liquid handling, plates were kept in a 20 °C incubator for an extra 3 hours to allow drug exposure (total drug exposure time was thus four hours).

Bright field illumination. To prevent light-avoidance response while illuminating the sample in bright field, the inventor used a dedicated 850 nm light system (Loopbio GmbH, Vienna, Austria), consisting of a 200 × 200 mm edge-lit LED panel. Briefly, the edge-lit configuration has the LEDs attached to the frame of the panel and shining into a horizontal light guide plate, which homogenises and diffuses the light. This setup was characterised to provide a minimum radiance of 15 W m⁻² sr⁻¹, a minimum uniformity of 95 ± 10%, and a half-angle (the angle at which the measured intensity falls to 50% its maximum value) is 30°.

The panel is significantly larger (200 × 200 mm) than the sample area. This achieves the double goal of minimising edge effects and thermally insulating the sample from the LED panel, as this can be placed at a relatively large distance (65 mm). To further improve light collimation, two computer monitor privacy filters (3 M, Saint Paul, MN, USA) placed at right angles are inserted between the LED panel and the sample area.

Image acquisition. All videos were acquired at 25 fps on the trackers in a temperature-controlled room at 20 °C, with a shutter time of 25 ms, and 12.4 µm

px–1 resolution. For all experiments unless otherwise noted, three sequential videos

were taken, run in series by a script: a 5-minute pre-stimulus video, a 6-minute blue light recording with 10-second blue light pulses set to 100% intensity at the 5 60, 160, and 260 s mark, and a 5-minute post-stimulus recording. The timing of recordings and photostimulation was controlled using Loopbio's API for Motif software [<https://github.com/loopbio/python-motifapi>] in a script.

For the serial blue light stimulation experiments, the plates were continuously imaged for 43 min and 20 s in the following stimulation pattern: 5 min off, 20 x 10 (10 s on at 100% intensity, 90 s off), 5 min off.

Image processing and quality control. Segmentation, tracking, and pose estimation over time was performed using Tierpsy Tracker. Each video was checked using Tierpsy Tracker's Viewer, and wells with visible contamination, agar damage, 15 or excess liquid (from worm sorter, so that worms swim rather than crawl) were marked as bad and excluded from the analysis.

Convolutional neural network to exclude nonworm objects. The inventor improved Tierpsy tracking by incorporating a CNN classifier after segmentation to exclude 20 nonworm objects from being analysed and skewing the results.

In the video compression step at the beginning of the Tierpsy analysis pipeline, a segmentation algorithm detects putative worm objects according to a set of user-defined parameters. The pixels in the frame that are further away than a threshold 25 from any of the putative worms are set to 0, creating a "Masked Video". The objects selected by the masking algorithm are tracked throughout the video, but now if only they pass the filtering step powered by a CNN classifier.

The classifier was trained on a dataset of 43,561 grey-scale "masked" images 30 measuring 80×80 pixels each, collected across several imaging systems in the inventor's lab. All images were manually annotated and objects were marked as either "worm" or "nonworm" by two independent researchers, so a consensus could be sought. The annotated dataset was split into training, validation, and test sets containing 80, 10, and 10% of the images, respectively, while keeping the classes 35 balanced in each set. All images were pre-processed in two steps. First, the background pixels set to 0 by the masking algorithm were shifted to the top 95 percentile of the grey values in the unmasked area. This prevents the artificial edge

between the masked and nonmasked area from disproportionately influencing the classifier. Second, all pixel values were scaled to the range of 0 to 1 by min-max normalisation, to reduce the influence of variable illumination and contrast in different imaging setups.

5

The architecture of the CNN is a shallower adaptation of VGG16, featuring eight convolution layers with 3×3 filter size and stride 1, each followed by a rectified linear activation unit, four max-pool layers (filter size 2×2 , stride 2) applied every two convolution layers, and a fully connected layer. Batch normalisation is applied to the third and seventh convolution layer to accelerate training by reducing internal covariate shift⁶⁸, and a Dropout layer is added before the fully connected layer to prevent overfitting⁶⁹. In total, the CNN has about 1.78 million trainable parameters.

10

15

The CNN classifier was implemented in PyTorch 1.6, and was trained with the cross-entropy loss function and the Adam optimisation algorithm at a learning rate of 10^{-4} . It achieved an accuracy of 97.68% and F1 score of 97.98% as measured on the independent test set.

20

To improve performance at the inference step, the inventor applies the CNN to a subset (one image per second) of all the images featuring the same putative worm object. This yields, per snapshot, the probability of the object to be a valid worm. If the median of this probability over time is higher than 0.5, the object is classified as a valid worm.

25

Video processing with multiple wells. Using multiwell plates for imaging significantly increased the experimental throughput, but also introduced challenges for data analysis as each video output contains 16 separate wells. Further software engineering was thus warranted to process multiwell videos, so that wells are detected and analysed separately.

30

To achieve this, the inventor implemented an algorithm in Tierpsy Tracker that automatically detects multiple wells in a field of view and stores the coordinates of well boundaries. Briefly, the inventor created a template that approximates the appearances of a well in the video, and replicated it on a lattice to simulate the grid of wells. The overall dimensions of the lattice are defined in Tierpsy's configuration file, but the lattice spacing parameters were chosen, via SciPy's differential

35

evolution routine, to minimise the differences between the video's first frame (or its static background, if Tierpsy was instructed to calculate it) and the simulated grid of wells.

- 5 Automatic extraction of behavioural features was then performed on a per-worm basis, before worms were sorted into their respective wells based on their (x, y) coordinates in order to obtain well-averaged behavioural features.

Data provenance. Tracking multiwell plates complicates the handling of metadata,
10 since there isn't a unique mapping between videos and experimental conditions. When well shuffling is performed using the liquid handling robot, the well contents in the imaging plate also needs to be tracked. To handle experimental metadata for imaging with the camera arrays, the records that need to be compiled manually during the experiments was standardised and an open-source module in Python
15 (<https://github.com/Tierpsy/tierpsy-tools-python/hydra>) was developed to combine the experimental records to create a full metadata table with the experimental conditions for each well (Supplementary Fig. 3).

The experimental records are typically compiled in the form of csv files. In each
20 tracking day, the experimenter needs to record: i) information about the media type and the bacterial food present on the imaging plates, and the worm strains that were dispensed into the wells of the plates (this is recorded in a summarized way in the wormsorter.csv file), ii) information about the experimental runs, including the unique IDs of the imaging plates, the instrument name where each
25 plate was imaged, and the environmental conditions (manual_metadata.csv), iii) if applicable, information about the contents of the compound source plates (sourceplate.csv) and the mapping between imaging plates and source plates (if the liquid handling robot was used for column shuffling, this mapping will be recorded automatically in the robotlog.csv; if there was no shuffling, this will be recorded in
30 imaging2source.csv).

Using the functions in the hydra module, firstly a plate metadata table is created to contain all the well-specific experimental conditions for every well of each unique imaging plate, including the compound contents if applicable. Then, the information
35 about the experimental runs is merged with the plate metadata to create a final metadata table with the complete experimental conditions for every recording of every well. At this stage, the video filenames are also matched to the sample based

on the camera array instrument ID. For example, scripts showing metadata handling, see https://github.com/Tierpsy/tierpsy-tools-python/tree/master/examples/hydra_metadata.

- 5 Statistics and reproducibility. Each well contains multiple worms (either 2 or 3, indicated in relevant captions) but worm identity is not necessarily maintained across the duration of tracking. The inventor therefore uses the number of wells as the sample size for statistical analysis rather than the number of worms. All experiments were repeated across at least three independent tracking days.
- 10 Feature data and analysis scripts are available on Zenodo.

Analysis of time-resolved response to photostimulation. Tierpsy Tracker was used to calculate a set of 3076 summary features for each well for each non-overlapping 10 s interval of the 6-minute stimulus recording (with three 10-second blue light

15 pulses starting at 60, 160, and 260 sec). Samples where more than 40% of the features failed to be calculated were excluded from the analysis, and so was any feature that failed to be calculated for more than 20% of the samples in any of the 10 s intervals. Missing values were then imputed by averaging the valid values within each time interval. The feature matrix (all wells, in all time intervals) was

20 then scaled by applying z-normalisation. Principal Components were then calculated using the whole feature matrix. Figure 3a shows a density plot of the measurements collected in the 10 s immediately before (left) and immediately after (right) a 10-second stimulus, projected onto the plane defined by the first two principal components.

25 To investigate the response to photostimulation with higher temporal resolution, Tierpsy Tracker¹⁷ was used to detect the motion mode (forwards, backwards, stationary) of each worm over time. To calculate the fraction of worms in each motion mode over time (Fig. 3b), the number of worms in each motion mode at

30 each time point in each well was divided by the total number of tracked worms at each time point in each well. This gave the fraction of worms in each motion mode, at each time point, for each well, so that an average could be taken across all wells. The 95% confidence interval for the average was obtained by nonparametric bootstrap ($n = 1000$ resamplings, with replacement).

35 For the longer experiments in Fig. 3c-e, the motion mode detected by Tierpsy Tracker for each worm overtime was first down-sampled to 0.5 Hz by dividing the

video into nonoverlapping two seconds intervals and taking the prevalent motion mode in each interval. The fraction of the worm population in each motion mode over time was calculated by counting the number of worms in each motion mode and then dividing by the total number of worms detected at each time point. The
 5 95% confidence interval was calculated via nonparametric bootstrap by the seaborn Python library.

Classification of wild isolates. For the classification of the divergent set the inventor used a random forest classifier as implemented in scikit-learn. For feature selection,
 10 the inventor used recursive feature elimination with a random forest estimator (RFE), as implemented in scikit-learn. The inventor started by splitting the data randomly in a training/tuning set and a test set, with 20% of the data from each strain assigned to the test set. The inventor used the training/tuning set for feature selection. The inventor tried specific candidate feature set sizes $\{2i, \text{for } i = 7:11\}$.
 15 For each size, the inventor performed cross-validation and: i) used each training fold to select N features and train a classifier with the selected features; ii) used each test fold to estimate the classification accuracy. The inventor repeated the process 20 times to get statistical estimates of the mean cross-validation accuracy for each size and selected the best performing size Nbest. The inventor then
 20 selected Nbest features using the entire training/tuning set and used this set for downstream analysis. At a second stage, the inventor tuned the hyperparameters of the random forest classifier using grid search with cross-validation as implemented in scikit-learn with the grid shown in Table 1. The best-performing parameters are reported in Table 1. Finally, the inventor trained a classifier on the
 25 entire training set using the selected features and hyperparameters and used it to make predictions on the test set.

Table 1 - Tested parameter grid for random forest classifier.

Parameter	Values	Selected
n_estimators	200:200:2000	1000
max_features	'auto', 'sqrt'	'sqrt'
max_depth	10:10:110, None	50
min_samples_split	2, 5, 10	2
min_samples_leaf	1, 2, 4	2

30 *Behavioral fingerprints predict insecticide and anthelmintic mode of action.*

With reference to McDermott-Rouse, A. et al. "Behavioral fingerprints predict insecticide and anthelmintic mode of action" Molecular Systems Biology 17: e10267 | 2021.

- 5 Novel invertebrate-killing compounds are required in agriculture and medicine to overcome resistance to existing treatments. Because insecticides and anthelmintics are discovered in phenotypic screens, a crucial step in the discovery process is determining the mode of action of hits. Visible whole-organism symptoms are combined with molecular and physiological data to determine mode of action.
- 10 However, manual symptomology is laborious and requires symptoms that are strong enough to see by eye. Here, high-throughput imaging and quantitative phenotyping is used to measure *Caenorhabditis elegans* behavioral responses to compounds and train a classifier that predicts mode of action with an accuracy of 88% for a set of ten common modes of action. Compounds are also classified within
- 15 each mode of action to discover substructure that is not captured in broad mode-of-action labels. High-throughput imaging and automated phenotyping could therefore accelerate mode-of-action discovery in invertebrate targeting compound development and help to refine modification categories.
- 20 Behavioral fingerprints – Introduction
- Invertebrate pests including insects, mites, and nematodes damage crops, decrease livestock productivity, and cause disease in humans. Nematodes alone infect over 1 billion people and lead to the loss of 5 million disability-adjusted life-years annually. In livestock, they infect sheep, goats, cattle, and horses causing gastroenteritis that
- 25 leads to diarrhea, reduced growth, and weight loss. Nematodes that parasitize crops have been estimated to cause well over \$100 billion in annual crop losses. Crop loss due to insects is measured in tens of metric megatons and is predicted to increase due to climate change. Compounds that kill or impair invertebrates are one of the primary means of defense in human and veterinary medicine and in crop
- 30 protection. However, resistance is widespread in nematodes and insects and drives continuing efforts to discover new invertebrate-targeting compounds.

To date, most currently approved treatments for infections in humans and livestock and for crop protection in the field have been discovered through phenotypic

35 screens. That is, compounds are first screened for the ability to kill or impair a target species without any hypothesized molecular target. A critical problem is then determining hit compounds' mode of action, which is important for understanding

resistance mechanisms, avoiding pathways where resistance is already common, and subsequent lead optimization. Despite advances in biochemical and genetic methods for determining mode of action, direct observation of the symptoms induced by compounds remains a key step in mode-of-action discovery. Because
5 most insecticides and anthelmintics target the neuromuscular system, behavioural symptoms are a particularly important class of phenotypes to consider, but manual observation of behaviour is time-consuming, insensitive to subtle phenotypes, and prone to inter-operator variability and bias. The inventor therefore sought to develop more automated and quantitative methods to do mode-of-action prediction
10 from phenotypic screens of freely behaving invertebrates.

Behavioural fingerprints can be used to discover neuroactive compounds and that behavioural fingerprints correlate with compound mode of action in zebrafish (*Danio rerio*). However, this approach has not yet been applied to invertebrate animals—
15 the targets of insecticides and anthelmintics—at a large scale. Furthermore, although previous zebrafish screens were high throughput, their spatial resolution was low and phenotypes were limited to activity levels in response to stimuli. Recent work in computational ethology has shown the power of moving beyond point representations of animal behaviour to include information on posture. From
20 previous symptomology work, it is clear that detailed postural information can be useful for resolving mode of action. The inventor chose *C. elegans* as our model system because it is small and compatible with multiwell plates and automated liquid handling. It is sensitive to anthelmintics and insecticides and has played an important role in mode-of-action discovery in the past (Brenner, 1974).

25 To combine the benefits of high throughput and high resolution, the inventor used megapixel camera arrays to record the behavioural responses of worms to a library of 110 compounds covering 22 distinct modes of action. The inventor simultaneously recorded all of the wells of 96-well plates with sufficient resolution
30 to extract the pose of each animal and a high-dimensional behavioural fingerprint that captures aspects of posture, motion, and path. The inventor show that worms have diverse dose-dependent behavioural responses to insecticides and anthelmintics and develop a machine learning approach that shares information across replicates and doses to accurately predict the mode of action of previously
35 unseen test compounds. Furthermore, the inventor shows that a novelty detection algorithm can provide an indication that a compound belongs to a mode of action not seen in the training set. This novelty score can be used as a measure of

confidence in the class prediction, suggesting a way to prioritize compounds with potentially novel modes of action early in the development process. These results demonstrate that high-throughput phenotyping in *C. elegans* is a promising approach for assisting target deconvolution in anthelmintic and insecticide
5 discovery. Finally, the inventor shows that our prediction accuracy might be limited by uncertainties in the class definition rather than noise or phenotypic dimensionality. Specifically, the inventor shows that the inventor can classify compounds even within a mode-of-action class, suggesting that there are limitations in our knowledge of the relevant pharmacology rather than limitations in
10 our ability to reproducibly detect compound-induced phenotypes.

Behavioural fingerprints – Results

Insecticides affect phenotypes in multiple behavioural dimensions. The inventor assembled a library of 110 insecticides and anthelmintics with diverse targets to
15 sample a range of modes of action used medically and commercially (see Dataset EV1 from McDermott-Rouse 2021 for full list). The modes of action represented in the library cover 70% of the market of insecticides used in the field and several important classes of anthelmintics used in veterinary and human medicine (Nixon et al, 2020). To quantify the effects of the compounds on behaviour, the inventor
20 recorded worms using megapixel camera arrays that simultaneously image all of the wells of 96-well plates (Fig 22A). The inventor recorded at least 10 replicates at three doses for each compound with enough resolution to extract high-dimensional behavioural fingerprints following segmentation, pose estimation, and tracking (Fig 22A). The behavioural fingerprints are vectors of posture and motion features that
25 are subdivided by body segment and motion state including “midbody curvature during forward crawling” or “angular velocity of the head with respect to the tail while the worm is paused”. The inventor has previously shown that similar features can detect even subtle behavioural differences that can be difficult to detect by eye and that the combined feature set has sufficient dimensionality to accurately
30 classify worms with diverse behavioural differences caused by genetic variation and optogenetic perturbation.

Referring to Figure 22, insecticides affect phenotypes in multiple behavioural dimensions.

35

A The inventor images an entire 96-well plate with a megapixel camera array with enough resolution and high enough frame rate to track, segment, and estimate the posture of *C. elegans* over time.

5 B All compound doses in the speed/tail curvature space, with points and lines showing the mean and standard deviation of biological dose replicates. On average, 12 biological replicates were collected per compound dose, together with 601 DMSO replicates, across at least 3 different tracking days for each condition. Several compounds, including the serotonin receptor antagonist mianserin (blue),
10 glutamate-gated chloride channel activator emamectin benzoate (purple), and vesicular acetylcholine transporter inhibitor SY1713 (red), have a strong effect on the worms' behavioural phenotype. They can be distinguished from the DMSO control (black) and from each other based on speed and tail curvature alone. Not all compounds are well separated in these two dimensions (gray points). Inset images
15 are samples that show postural differences.

C Sample worm skeletons over time show the effect of the compounds highlighted in (B) on motion.

20 D Number of features significantly different from the DMSO control at a false discovery rate of 1% for each compound, grouped by mode of action. The pre-stimulus, blue light stimulus, and post-stimulus data are shown separately (a total of 3,020 features are tested for each assay period). The percentage of significantly different features is highest for the blue light stimulus recording.

25 As expected, several compound classes have strong visible effects on *C. elegans* behaviour including the glutamate-gated chloride channel activator emamectin benzoate, the spiroindoline vesicular acetylcholine transporter inhibitor SY1713, and the serotonin receptor antagonist mianserin. All three compounds at specific doses
30 can be distinguished from DMSO controls and from each other in a simple two-dimensional space defined by speed and body curvature (Fig 22B). The large differences in curvature and in motion caused by some compounds are observable by eye, as shown in the inset images and in Fig 22C. However, not all compounds are well separated in these two dimensions; the gray points in Fig 22B show the
35 dose means and standard deviation of all the compound doses in the speed/tail curvature space, which largely overlap. Some of the screened compounds might not have detectable effects on *C. elegans* and therefore cannot be used for phenotypic

mode-of-action prediction. To find the compounds with no effect, the inventor compared the behavioural fingerprints of treated worms with DMSO controls using univariate statistical tests for each feature and correcting for multiple comparisons with the Benjamini–Yekutieli procedure. To account for random day-to-day variation
5 in the experiments, the inventor used a linear mixed model for these statistical tests, where the fixed effect is the drug dose and the day of the experiment is added as a random effect. The number of features that are significantly different at a false discovery rate of 1% between the behaviour of worms treated with each compound and the DMSO controls is summarized in a heat map (Fig 22D).

10

To further increase the dimensionality of the behavioural phenotypes, the inventor included a blue light stimulation protocol. Each tracking experiment is divided into three parts: (i) a 5 min pre-stimulus recording, (ii) a 6-min stimulus recording with three 10-s blue light pulses starting at 60, 160, and 260 s, and (iii) a 5 min post-
15 stimulus recording. Blue light is aversive to *C. elegans* (Edwards et al, 2008) and so it can help to distinguish between animals that are simply pausing and those that are not able to move. Behavioural differences are observed in each assay period, but the stimulus period shows the most differences (Fig 22D). Even within mode-of-action classes, compound potency can be highly variable. The largest potency
20 difference is observed for the octopamine agonists where amidine affects 0.08% of features and oxazoline affects 75% of features. Overall, 86% of compounds have a detectable effect on behaviour in at least one feature. The 17 compounds that showed no detectable effect in any stimulus period were not included in subsequent analysis.

25

Compounds with the same mode of action have similar effects on behaviour. Having established that *C. elegans* shows diverse behavioural responses to insecticides and anthelmintics, the inventor next sought to determine to what extent the responses are mode-of-action-specific. For the initial clustering, the
30 inventor used 256 features from each blue light condition. These features were selected for their usefulness in classifying mutant worms in a previous paper (Javer et al, 2018). For all clustering and classification tasks, the inventor first z-normalize each feature to put them on a common scale and to prevent arbitrary choices of units from impacting the analysis. The inventor used hierarchical clustering to
35 visualize the relationships between the behavioural responses to different compounds at different doses (Fig 23A). Each row of the heat map is the average of all of the replicates of a given compound at a specific dose. The inventor also

included the averages of six subsets of the DMSO replicates randomly partitioned across tracking days as control points. Several of the compound classes show clear clustering, including the AChE inhibitors, vAChT inhibitors, GluCl agonists, and mAChR agonists. The DMSO averages also cluster closely together. The degree of mode-of-action clustering is greater than expected by chance, which can be seen in a plot of the cluster purity observed in the data compared with random clustering (Fig 23B). However, the distance between compounds that share the same mode of action can be large, even for classes that cluster well overall, in part because behavioural fingerprints change with dose. It is also not always possible to align feature vectors using doses because compounds can have very different potencies: A low dose for one compound could be a high dose for another. Furthermore, the compound concentration inside the worm is likely to be much lower than the concentration in the media because of *C. elegans*' considerable xenobiotic defences and the degree of uptake will also vary across compounds even within a mode-of-action class. For this reason, the centre-top part of the heat map in Fig 23A is populated with low doses and compounds that either have a low potency or low uptake in the worm, which do not form distinct clusters based on their mode of action, but are rather clustered around the DMSO averages.

These effects can be seen in dose-response plots for individual features. The three mitochondrial inhibitors in Fig 23C all decrease angular velocity, but they do it at different doses. At 3 μ M, only SY1048 has a strong effect, while at 30 μ M, rotenone has a similarly strong effect. Clustering based on angular velocity would lead to qualitatively different conclusions about nearest neighbours at these different doses. For the spiroindolines, similar differences in dose-response are observed for body curvature with the added difference that the effect of SY1786 is non-monotonic and returns to baseline at high doses. These non-monotonic effects can be due to compounds precipitating from solution at high doses or due to intrinsically complex compound effects such as a compound that causes an increase in speed at low doses but is lethal at high doses. Regardless of the cause, complex dose-response curves present challenges for mode-of-action prediction since supervised machine learning algorithms rely on differences in feature distributions to learn decision boundaries and dose-response effects spread out the distributions and increase the overlap between classes.

Combining classifiers by voting enables mode-of-action prediction

The behavioural fingerprints of compounds with the same mode of action have the same direction in the phenotypic space and can be used for classification in mode-of-action classes. For the classification task, a minimum number of compounds per class can be used to get an accurate representation of the class distribution. Out of
5 the compounds with detectable effects in *C. elegans*, the inventor chooses only the classes with at least five compounds (10 classes with 76 compounds). The inventor takes advantage of the fact that several replicates are recorded per condition and resample with replacement from the multiple replicates for each dose to create a set of average behavioural fingerprints. This effectively smooths the data reducing
10 the effect of outliers. At the same time, it provides a simple method for balancing classes before classifier training. For classes with fewer compounds, the inventor resamples more times so that each class contains the same number of points (see Fig EV1). To partially mitigate the effect of compound potency, the inventor then normalizes each behavioural fingerprint to unit magnitude. This normalization is
15 done row-wise on each sample in contrast to the z-normalization described above which is done column-wise on each feature. Rescaling in this way brings compounds with similar effect profiles but different potencies closer together in feature space (Figs 24A and EV2), but because of nonlinearities in the dose-response profiles, the overlap is not perfect even after rescaling.

20 Predictions must be combined across doses and replicates to make a single prediction for the mode of action of a given compound. Inspired by an analogy with the multi-sensor fusion problem (Singh et al, 2019), the inventor uses a voting procedure to make a final prediction. However, in contrast to multi-sensor fusion,
25 the inventor cannot train different classifiers for each dose because 1 μ M for one compound is not equivalent to 1 μ M for another compound. Instead, the inventor trains a single classifier for all doses and make predictions for each data point. Each data point contributes a vote for a compound's class and the class with the most votes wins.

30 The inventor splits our data into a training/tuning set consisting of 60 compounds and a hold-out/reporting dataset consisting of 16 compounds containing at least one compound for each mode-of-action class. For mode-of-action prediction, the inventor started with the full set of features output by Tierpsy (3,020 per blue light
35 condition for 9,060 in total) and used the training set to determine an appropriate classifier, select features, and tune hyperparameters using cross-validation. The inventor achieved the highest cross-validation accuracy with 1,024 features

selected using recursive feature elimination with a logistic regression estimator. The hyperparameters of the classifier were also tuned using cross-validation, and the best version of the estimator was multinomial logistic regression with L2 regularization and penalty parameter $C = 10$. Using a regularized linear classifier helps to control overfitting in this high-dimensional feature space, which boosts the cross-validation accuracy. The confusion matrix from cross-validation using the best performing feature set and classifier is shown in Fig 24. To determine whether the classifier could generalize to unseen compounds, the inventor applied it to the test data without further tuning. The classifier predicted the correct mode of action for the unseen compounds 88% of the time (Fig 24C). The inventor generated a null model by partitioning the DMSO data randomly across tracking days to 10 classes. Following the same steps as for the compound-treated data, the inventor obtained a maximum cross-validation accuracy of 10% using the training set, while the prediction accuracy in the test set was 12.5%.

In addition to the 10 modes of action that were represented by at least five compounds in our dataset, the inventor had 17 compounds with a detectable effect on *C. elegans* that belonged to 11 sparsely populated mode-of-action classes. The inventor used these additional compounds to simulate another use case for our approach: detecting screening hits that represent potentially novel modes of action that do not fall into known classes. The inventor uses the term "novel test set" to describe these compounds, since their modes of actions are unknown (novel) to the trained classifier. Using a novelty detection algorithm with some modifications, the inventor assigned a novelty score to each of the test and novel test compounds based on their affinity to each of the existing classes. To obtain the novelty score, the inventor uses an ensemble of support vector machine (SVM) classifiers that flag novel compounds based on the confidence values of the main multinomial logistic regression classifier used for the predictions of known classes. The ensemble of SVM classifiers is trained using partitions of the training set into presumed-known and presumed-unknown classes. The novelty score is defined as the weighted average of the output of this ensemble. Most of the novel compounds were assigned novelty scores above 0.8 (Fig 24D). Several of the non-novel compounds—those that come from a class that is present in the training data—have high novelty scores, but this includes the two test compounds that were incorrectly classified. In this case, the high novelty score correctly indicates low confidence in the prediction of the classifier. To explore the origin of the high novelty score for

the incorrectly classified compounds, the inventor looked for differences between the effects of compounds within a class.

Mode-of-action deconvolution within classes

5 Although compounds are categorized into broad mode-of-action classes, most compounds will have some degree of off-target engagement. If the off-target effects are different for compounds within a mode-of-action class or if the compounds have differences in pharmacokinetics, they may lead to different phenotypes. In this case, it may be possible to use behavioural fingerprinting to
10 further deconvolve mode-of-action classes revealing hidden compound heterogeneity. To test for phenotypic differences within mode-of-action classes, the inventor trained a classifier to distinguish the replicates from each compound within a class from the replicates of the other compounds in the class. The inventor then used cross-validation accuracy to quantify the distinguishability of the compounds
15 within a class. The inventor hypothesized that for classes without mechanistic sub-classes, the classifier would perform similarly to random guessing. In contrast, if compounds with the same mode of action had different off-target profiles or different pharmacokinetics, the classifier would be able to reliably distinguish individual compounds or subsets of compounds with the same broad mode of
20 action.

In all classes, the compound-level classifiers performed better than random guessing, in some cases by a large margin. One of the incorrectly classified compounds in the test set was ritanserin, which was also assigned a high novelty
25 score. The within-class classifier shows that it is indeed clearly distinguishable from the other 5-HT receptor antagonists (Fig 25A). Although ritanserin is known to be a 5-HT receptor antagonist, it is also known to affect multiple other targets. In addition to detecting outlying compounds, the deviations from random guessing revealed substructures within the classes. For example, one group contains the two
30 antidepressants, which are nearest neighbours in terms of structural similarity (atom pair Tanimoto coefficient of 0.44 between the antidepressants compared with a Tanimoto coefficient of 0.20 ± 0.09 (mean $\pm$ SD) for the other pairwise comparisons within the class). As with ritanserin, the other four compounds in this class have known polypharmacology, which could be driving their clustering.
35 Another class with interesting substructure is that consisting of mitochondrial inhibitors, which also separates into distinguishable groups (Fig 25B). In this case, the phenotypically distinct groups separate the complex I inhibitors from the

complex II and complex III inhibitors, which appear phenotypically more similar. The mectins the inventor tested are structurally similar and are known to share the same binding site, suggesting they would be difficult to separate into subgroups. Consistent with this expectation, the compound-level mectin classifier performs
5 only slightly better than chance (Fig EV3).

Lanthipeptides expressed in *E. coli* affect *C. elegans* behaviour

The inventor has prepared a library of *E. coli* expressing diverse lanthipeptides as previously described. This library was then diluted (1:1000000) and plated on to
10 150 mm round plates containing LB agar with carbenicillin (100 µg/ml), and grown overnight at 37 °C. The bacterial colonies from the library were picked into 384-well plates containing liquid LB with carbenicillin (100 µg/ml) using a colony picker (PIXL, Singer Instruments) to have one colony per well. The liquid culture was grown overnight at 37 °C and the next morning the OD600 was recorded. Glycerol
15 stocks of the library were prepared by adding 100 µl of 30% glycerol per well using a VIAFILL reagent dispenser (INTEGRA Biosciences Ltd, UK), incubation at 37 °C for 1h before storage at -80 °C.

For screening the library, the frozen stocks were replicated into 384-well plates
20 containing liquid LB with carbenicillin (100 µg/ml), using plastic pin replicators. These cultures were grown overnight at 37 °C and the next morning the OD600 was recorded. Imaging plates were prepared by filling 96-well plates with 200 µL nematode growth medium agar (<https://wormbook.org/>), carbenicillin (100 µg/ml) and IPTG (1 mM) per well using a VIAFILL reagent dispenser (INTEGRA Biosciences
25 Ltd, UK), and stored at 4 °C until use. Before imaging, the 96-well plates were placed in a LEEC BC2 drying cabinet (LEEC Ltd, Nottingham, UK) to lose 3–5% weight. Two days before tracking 10 µl of the 384-well plate (lanthipeptide library) liquid culture was seeded on 4 x 96-well imaging plates using a VIAFLO96 (INTEGRA Biosciences Ltd, UK) to have one colony per well. Plates were dried in a
30 laminar flow hood for 1.5 hr and then stored at 20 °C overnight. The following day synchronised *C. elegans* at the fourth larval stage were dispensed on to the plates. An average of ten worms was placed into each well using an INTEGRA VIAFILL. Plates were dried in a laminar flow hood for 1.5 hr and then placed in an incubator at 20 °C overnight. After 24 hours from dispensing the worms, the imaging plates
35 were placed in the room of the megapixel camera array tracker for 0.5 h to acclimatise prior to image acquisition.

For all experiments three sequential videos were taken, run in series by a script: a 5-minute pre-stimulus video, a 6-minute blue light recording with 10-second blue light pulses at the 60, 160, and 260 s mark, and a 5-minute post-stimulus recording. Segmentation, tracking, and pose estimation over time was performed using Tierpsy Tracker. Each video was checked using Tierpsy Tracker's Viewer, and wells with visible contamination, agar damage, or excess liquid were marked as bad and excluded from the analysis. The video processing with multiple wells and the automatic extraction of behavioural features were performed as described hereinbefore.

10

Referring to Figures 27a, 27b, and 27c, the behavioural fingerprints of worms fed lanthipeptide-expressing *E. coli* was compared to control *E. coli* that were not expressing a lanthipeptide. For each of the Tierpsy_16 features the inventor found multiple lanthipeptide clones that affected worm physiology. The effects for different clones are reflected in changes in motion, morphology, and posture.

15

Behavioral fingerprints – Discussion

Worms have diverse responses to insecticides and anthelmintics and the corresponding behavioural fingerprints can be used to cluster compounds with similar modes of action. With appropriate normalization and by combining information across doses and replicates through voting, the inventor can also accurately predict the mode of action of previously unseen compounds despite differences in compound potency and uptake into the worm.

20

Expanding the range of organisms included in the training data is one way to increase the coverage of compounds investigated. There are methods for deriving multidimensional behavioral fingerprints that incorporate postural information from flies and zebrafish larvae, and both organisms are compatible with high-throughput screening.

25

Next-generation sequencing methods have made the discovery of new RiPPs in genomes common. Genome modification advances, for example, CRISPR and related techniques, have made the production of RiPPs in new host organisms possible. For example, bacterial cells (e.g., *Escherichia coli*), fungal cells (e.g., *Saccharomyces cerevisiae*, *Pichia pastoris*, or *Schizosaccharomyces pombe*), plant cells culture, or a protozoa cell) may be modified so that they produce modified peptides such as RiPPs. These modified peptides may have been known from a

30

35

different organism, or may be a peptide that has previously been unknown. In turn feeding these cells, for example in the form of a cell culture, or a cell lawn, to different (e.g., "higher") organisms, for example, protozoa (e.g., predatory protozoa), roundworms, insects, insect larvae, amphibians or amphibian larvae, and fish, including fish larvae or fry, has the effect of concentrating the RiPPs to a sufficient level that measurable behavioural differences can be identified using the techniques described earlier.

Thus, the inventor has surprisingly found that by measuring the behavioural phenotype of an organism that has ingested a cell culture configured (for example by genetic modification) to produce a modified peptide and comparing this "focal" behavioural phenotype with one or more reference behavioural phenotypes of organisms of the same species, it is possible to identify whether the modified peptide (RiPP) has an effect on the organism. Further, this technique may allow the identification of the effect of the modified peptide on the focal organism. The one or more reference behavioural phenotypes may include wildtype behavioural phenotypes, allowing the assessment of the diet of modified peptide producing organisms to be compared against an organism that has had no treatment or modification. Thus, determining whether certain behavioural features are more or less common, or more or less pronounced in the focal behavioural phenotype, may indicate that the modified peptide is having an effect.

Using the megapixel camera array and system described above, or an equivalent video-capture system suitable for other organisms, a high-throughput system can be used to record the behaviour of one or more focal organisms, extract features from the behaviours of the organisms and use these features compared against reference behavioural phenotypes to identify modified peptides that have effects on the organisms. Using the system 1 described here, it is possible to perform this method more quickly and using fewer physical and computing resources than has previously been possible.

Identifying the peptide as having an effect may comprise identifying that the peptide has a similar effect as a known peptide, for example, by identifying that two behavioural phenotypes are similar, where the reference behavioural phenotypes are from organisms which have been treated with or have ingested compounds with known effects. The treatment may comprise organisms having ingested a second cell culture, wherein cells in the second cell culture are

configured to produce either non-modified peptides, or modified peptides having a specified sequence. Alternatively, or additionally, the treatment may comprise the reference organism(s) having ingested a second cell culture, wherein cells in the second cell culture produce a second peptide which has a known effect. The known effect may be a known effect on the behavioural phenotype of the organism, the known effect may be a known effect on a different organism, a cell, a group of cells, a tissue, or an organ. For example, the known effect may be a physiological effect on a, or part of a, human, a non-human animal, plant, a fungus, a bacterium or archaeon. For example, the effect may be as a stimulant, as a depressant, or another physiological effect.

New therapeutics, or possible or potential new therapeutics, may be identified using this system and technique. For example, the one or more reference behavioural phenotypes may include those from organisms which have had a known effect, treatment, or modification. Reference behavioural phenotypes may include those from organisms which have been subject to a treatment from a compound (e.g., exposure to a solvent, a buffer solution, a therapeutic, a pharmaceutical drug, a chemical compound, a nutraceutical, or a medicament, and the like), those which have had a genetic modification (e.g., post transcriptional or genome modification), or organisms which have had a particular life history, diet, development, physical treatment (e.g., different or varying light, temperature, sound, or vibration exposure), or organisms which have ingested a cell culture which is known to express certain compounds or molecules, for example, known peptides. Behavioural phenotypes of organisms (for example a database of behavioural phenotypes) which have had these treatments or modifications may allow for the identification of modified peptides which have similar effects to these treatments or modifications, if the features of the focal and reference behavioural phenotypes are similar within an envelope or past a certain threshold or thresholds.

Thus, the features extracted from a behavioural phenotype of a focal organism may be compared with features extracted from a behavioural phenotype of one or more reference organisms. The distance between the extracted features may allow for the identification of the similarity of an effect of the modified peptide in the focal organism with that of the known effect of a treatment or modification in the reference organism. For example, if the focal organism having been fed a cell culture expressing or producing a particular modified peptide shows a similar behavioural phenotype to that of a reference organism that has had a treatment

with a known pharmaceutical drug, it may be that the modified peptide has a similar effect as the known pharmaceutical drug.

5 The cell culture used to feed the organisms for the focal and/or the reference behavioural phenotypes may consist or comprise a genetically modified cell culture. The cell culture may have been treated with a compound, drug, pharmaceutical, or therapeutic. The cell culture may have a coating of a compound, for example, an active compound.

10 Referring to Figure 26, a cell culture is prepared comprising or consisting of cells which express or produce a modified peptide. These cells may comprise or consist of any cell which is suitable for ingestion by a target organism, for example, the cells can be used as a feed or a feed supplement for the target organism. These cell cultures may then be fed to a target organism (step S1.1). Example target
15 organisms include nematodes (e.g., *Caenorhabditis elegans*), insects or insect larvae, (e.g., fruit flies, *Drosophila melanogaster*), fish (e.g., zebrafish, *Danio rerio*), amphibians (e.g., African clawed frog, *Xenopus laevis* etc.), birds (e.g., chick, *Gallus gallus domesticus*), and mammals, (e.g., mouse, *Mus musculus*, or brown rat, *Rattus norvegicus*). In a specific example, a fungal cell (e.g.,
20 *Saccharomyces cerevisiae*, *Pichia pastoris*, or *Schizosaccharomyces pombe*) or a bacteria/prokaryote cell culture, (e.g., *Escherichia coli*) may be modified to express or produce one or more RiPPs and then fed to the roundworm *Caenorhabditis elegans* either in suspension or as a bacterial or fungal lawn on an agar-based plate (e.g., see www.wormbook.org). The organisms may be fed any suitable amount of
25 cell culture. How much cell culture is required to produce an effect on the behaviour of the organism will depend on the modified peptide, and the focal organism.

A behavioural phenotype can then be captured by recording the movements of the organism after having ingested the cell culture. A specific example of this for *C.*
30 *elegans* is described hereinbefore, but this technique is readily applied to organisms with similarly elongate bodies and undulating movements, for example, fish, fish larvae, amphibian larvae, protozoa, and some insect larvae. Behavioural phenotypes can be captured for other organisms, and features identified and extracted in a similar way to that described above. This behavioural phenotype is
35 then received by a user for further analysis (step S1.2).

One or more reference behavioural phenotype(s) of the same species of organism as the focal organism may be acquired in a similar way or may be received from a database of suitable reference behavioural phenotypes (step S1.3). As described above, the reference behavioural phenotype may be of an unmodified wild-type organism, a genetically modified organism, or an organism that has undergone a treatment.

The features of the focal behavioural phenotype and the reference behavioural phenotype(s) may then be compared with each other (step S1.4). Based on whether differences or similarities between the behavioural phenotypes, for example whether certain features of the behavioural phenotype are more or less common, or more or less pronounced in the focal behavioural phenotype, it can be determined whether the modified peptide has an effect on the organism (step S1.5). The effects or potential effects of these modified peptides, may be, for example, their effects on a range of organisms.

The reference phenotype may be a plurality of reference behavioural phenotypes. The reference phenotype may be an average, a weighted average, a mean, a median or a mode of a plurality of reference phenotypes. Likewise, the focal behavioural phenotype may consist or comprise a behavioural phenotype from one or more of focal organisms, and again, these phenotypes may be averaged (e.g., a weighted average, a mean, a median or a mode) or modelled in a suitable way.

The effect of the modified peptide may be a known biological effect, or a known chemical effect.

The modified peptide may be a peptide which has been modified by an enzyme. The modified peptide may be a designed peptide for example, a structure-guided or bioinformatic-guided peptide.

The behavioural phenotype may comprise one or more features, for example, 4 features, 16 features, 64 features, 256 features, 1,000 features, 2,000 features, 3,000 features, 4,000 features, or more than 4,000 features. The features may be extracted from analysis of the video of the organisms. Each feature may comprise one or more traits. The behavioural phenotype may comprise one or more traits, each trait may comprise one or more of: body part location, morphology, body posture, speed, velocity, turn data, and/or trajectory data. Each feature may

comprise trait data. For example, curvature of a particular part of the body, the speed relative to body length etc.

5 Morphology may comprise length, area, and width of the organism. Body posture may comprise curvature, curvature may be in relation to major and minor axes. Trajectory data may comprise path curvature, path grid data, path area data. Velocity data may comprise angular velocity, radial velocity. Turn data may comprise count data, frequency data, and/or curvature data.

10 The body part location may comprise anterior, posterior, or mid-body location data. The body part location may comprise head, location, neck location, hip location, and tail location.

15 The behavioural phenotype data may comprise frequency and/or amplitude data.

Comparing the behavioural phenotypes may comprise selecting one or more features from the behavioural phenotypes, and then comparing these features. This feature selection may include recursive feature elimination, for example, using a regression model (e.g., a logistic regression).

20

Modifications

It will be appreciated that various modifications may be made to the embodiments hereinbefore described. Such modifications may involve equivalent and other features which are already known in the design of methods to identify effects of modified peptides using behavioural phenotypes of organisms which may be used instead of or in addition to features already described herein. Features of one embodiment may be replaced or supplemented by features of another embodiment.

30 Although claims have been formulated in this application to particular combinations of features, it should be understood that the scope of the disclosure of the present invention also includes any novel features or any novel combination of features disclosed herein either explicitly or implicitly or any generalization thereof, whether or not it relates to the same invention as presently claimed in any claim and whether or not it mitigates any or all of the same technical problems as does the present invention. The applicants hereby give notice that new claims may be formulated to such features and/or combinations of such features during the

prosecution of the present application or of any further application derived therefrom.

Claims

1. A method of determining an effect of a modified peptide, the method comprising:
 - 5 comparing a behavioural phenotype of an organism, the organism having ingested a cell culture, wherein cells in the cell culture is configured to produce a modified peptide, with a reference behavioural phenotype of organisms of the same species; and
 - determining whether the peptide has an effect based on the difference
 - 10 between the behavioural phenotype and the reference behavioural phenotype.
2. The method of claim 1, wherein the modified peptide is a ribosomally synthesised and post-translationally modified peptide.
- 15 3. The method of claim 1, wherein the modified peptide is a peptide modified by an enzyme.
4. The method of any of claims 1 to 3, wherein the reference behavioural phenotype comprises a behavioural phenotype from an organism which has
- 20 undergone a treatment.
5. The method of any of claim 4, wherein the treatment comprises organisms having ingested a second cell culture, wherein cells in the second cell culture are configured to produce either non-modified peptides, or modified peptides having a
- 25 specified sequence.
6. The method of claim 4 or 5, wherein the treatment comprises an organism having been exposed to a compound.
- 30 7. The method of any of claims 1 to 6, wherein for the reference behavioural phenotype, the organism has ingested a second cell culture, wherein cells in the second cell culture produce a second peptide, the second peptide having a known effect.
- 35 8. The method of any of claims 1 to 7, wherein the organism and/or the reference organism is a genetically modified organism.

9. The method of any of claims 1 to 8, wherein the behavioural phenotype comprises one or more features.

10. The method of claim 9 wherein the behavioural phenotype comprises one or
5 more traits, each trait may comprise one or more of:

body part location;

morphology;

body posture;

speed;

10 velocity;

turn data; and

trajectory data.

11. The method of claim 10, wherein a feature comprises one or more traits.
15

12. The method of any of claims 1 to 11, wherein the behavioural phenotype comprises frequency data.

13. The method of any of claims 1 to 12, wherein comparing the behavioural
20 phenotypes comprises selecting one or more features from the behavioural phenotypes.

14. The method of claim 13, wherein feature selection comprises recursive
feature elimination.
25

15. The method of claim 13 or 14, wherein the feature selection comprises using
a regression model.

16. The method of any of claims 1 to 15, wherein the difference between the
30 behavioural phenotype and the reference behavioural phenotype is less than a pre-determined threshold.

17. The method of any of claims 1 to 16, wherein the behavioural phenotype is
obtained using an image of the organism.
35

18. The method of claim 19, wherein the behavioural phenotype is obtained
using video of the organism.

19. The method of any of claims 1 to 17 wherein the behavioural phenotype is obtained by using image analysis to extract one or more traits from an image or a series of images.

5

20. A system comprising:
a camera *configured* to capture one or more images of an organism; and
a processor configured to compare a behavioural phenotype of an organism,
the organism having ingested a cell culture, wherein cells in the cell culture are
10 configured to produce a modified peptide, with a reference behavioural phenotype
of organisms of the same species, and wherein the processor is configured to
determine whether the peptide has an effect based on the difference between the
behavioural phenotype and the reference behavioural phenotype.

15

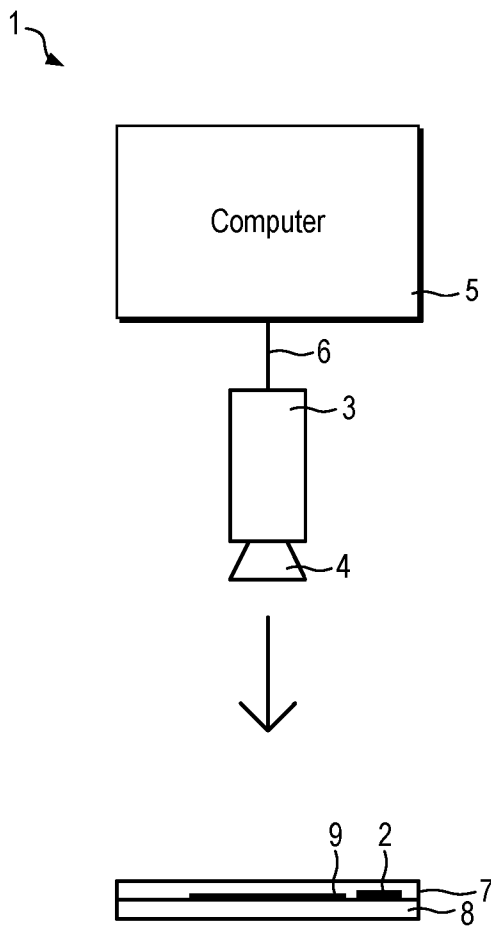


Fig. 1

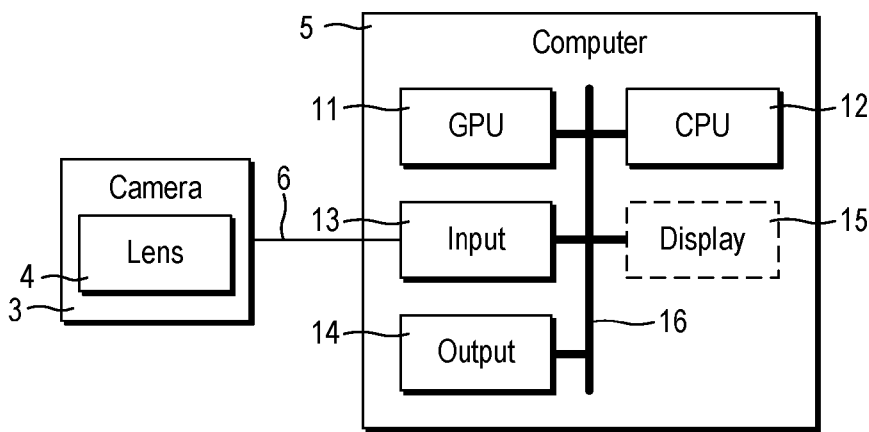


Fig. 2

2/30

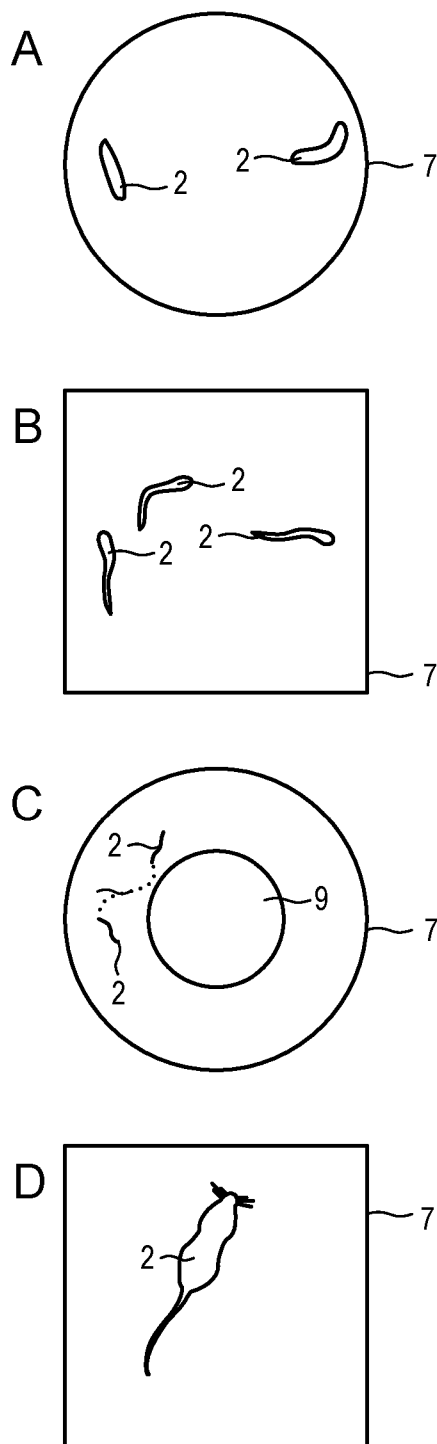


Fig. 3

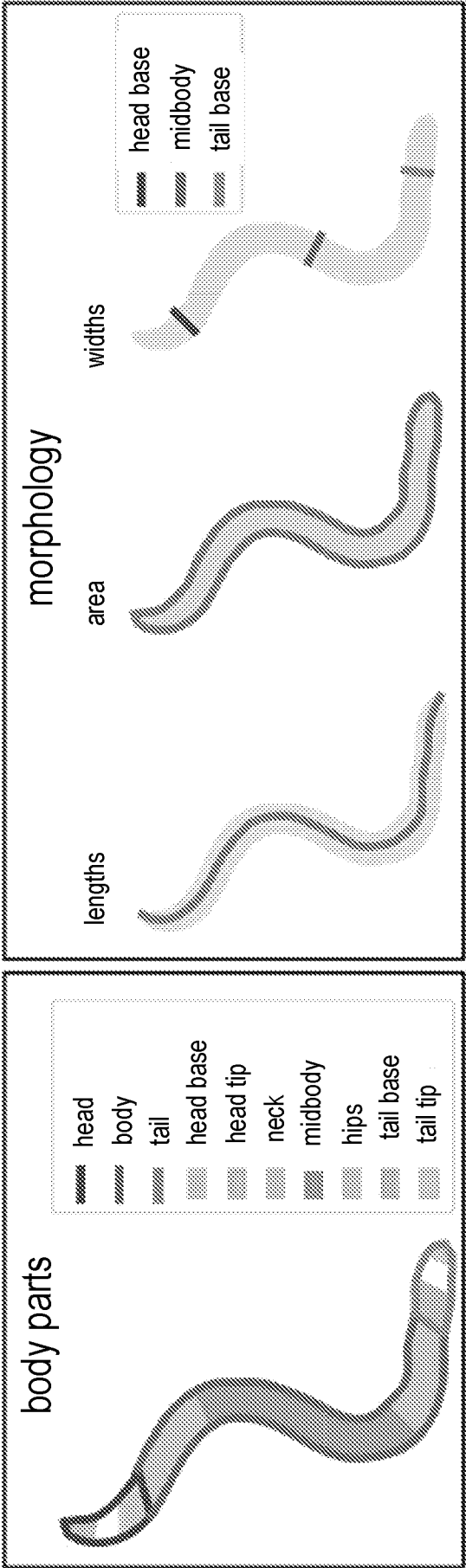


Fig. 4

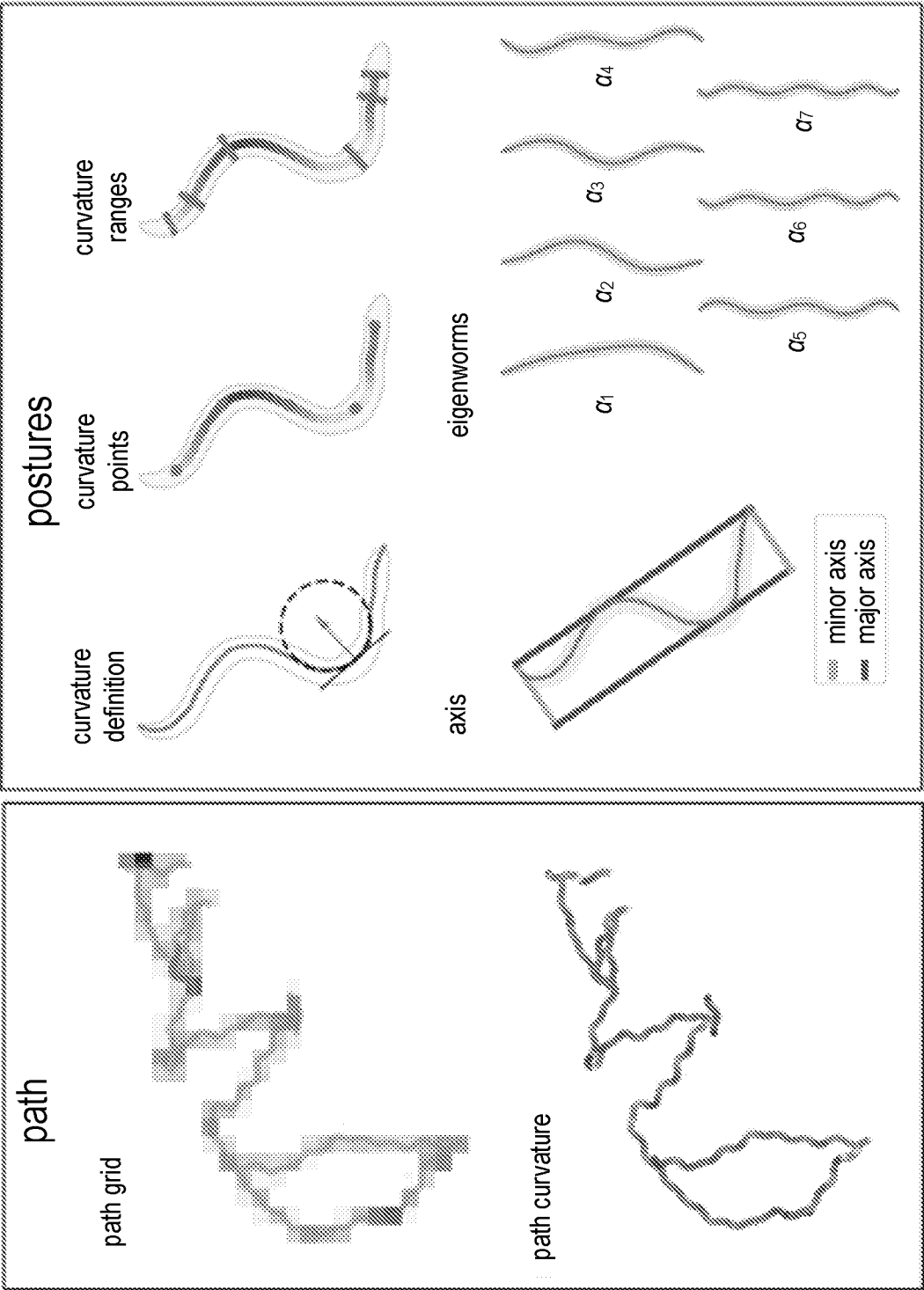


Fig. 5

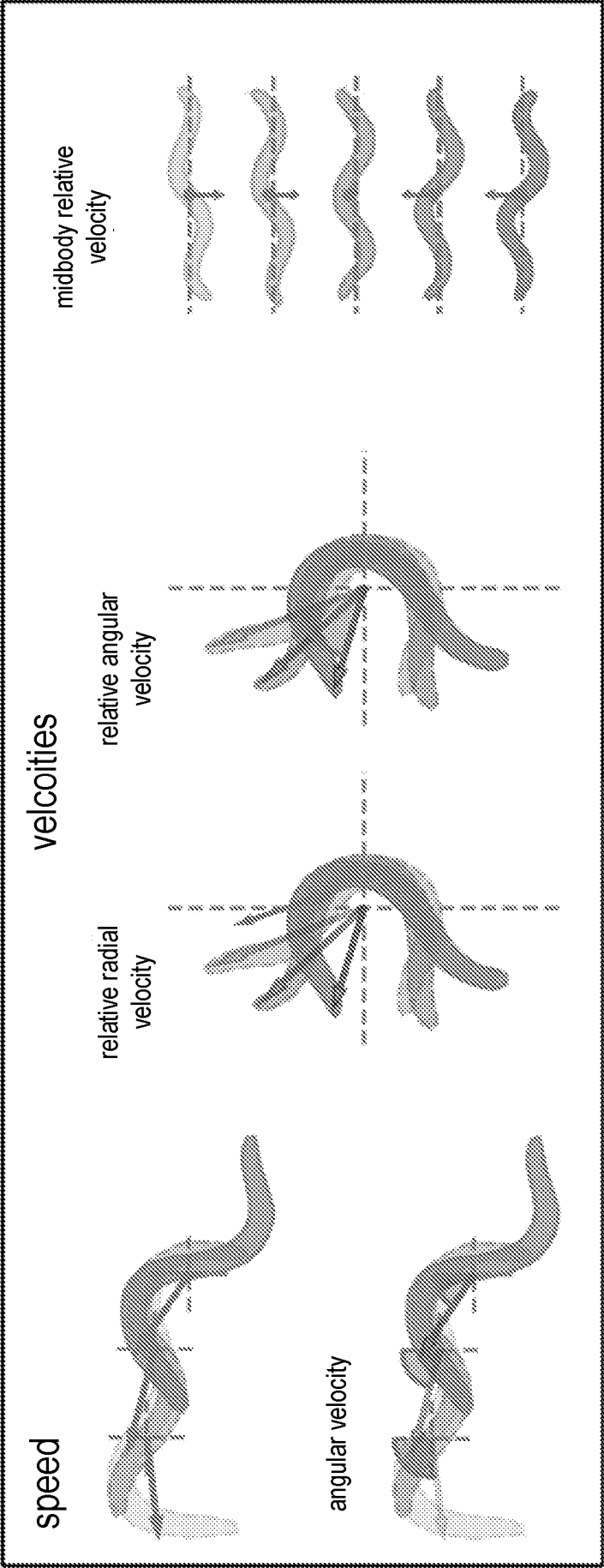


Fig. 6

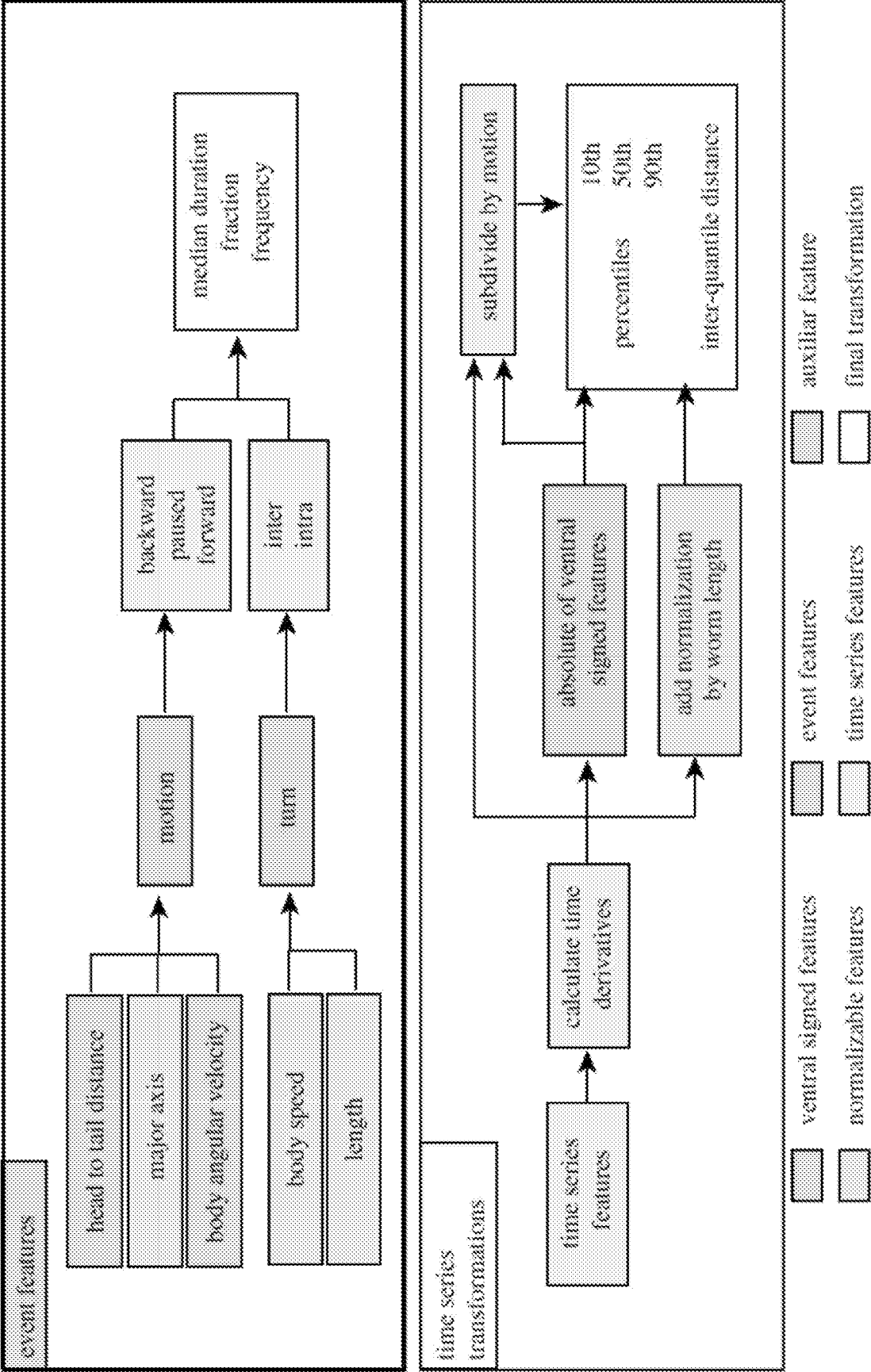


Fig. 7

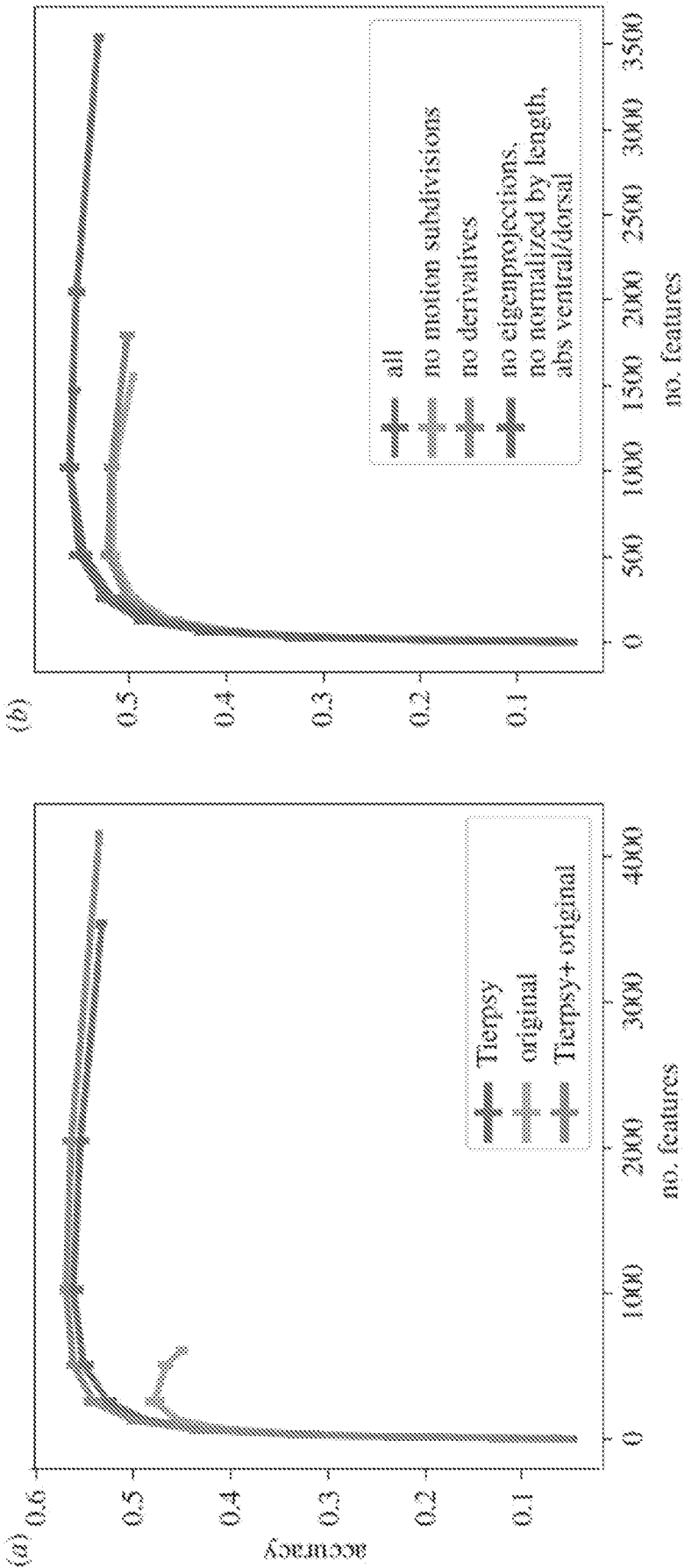


Fig. 8

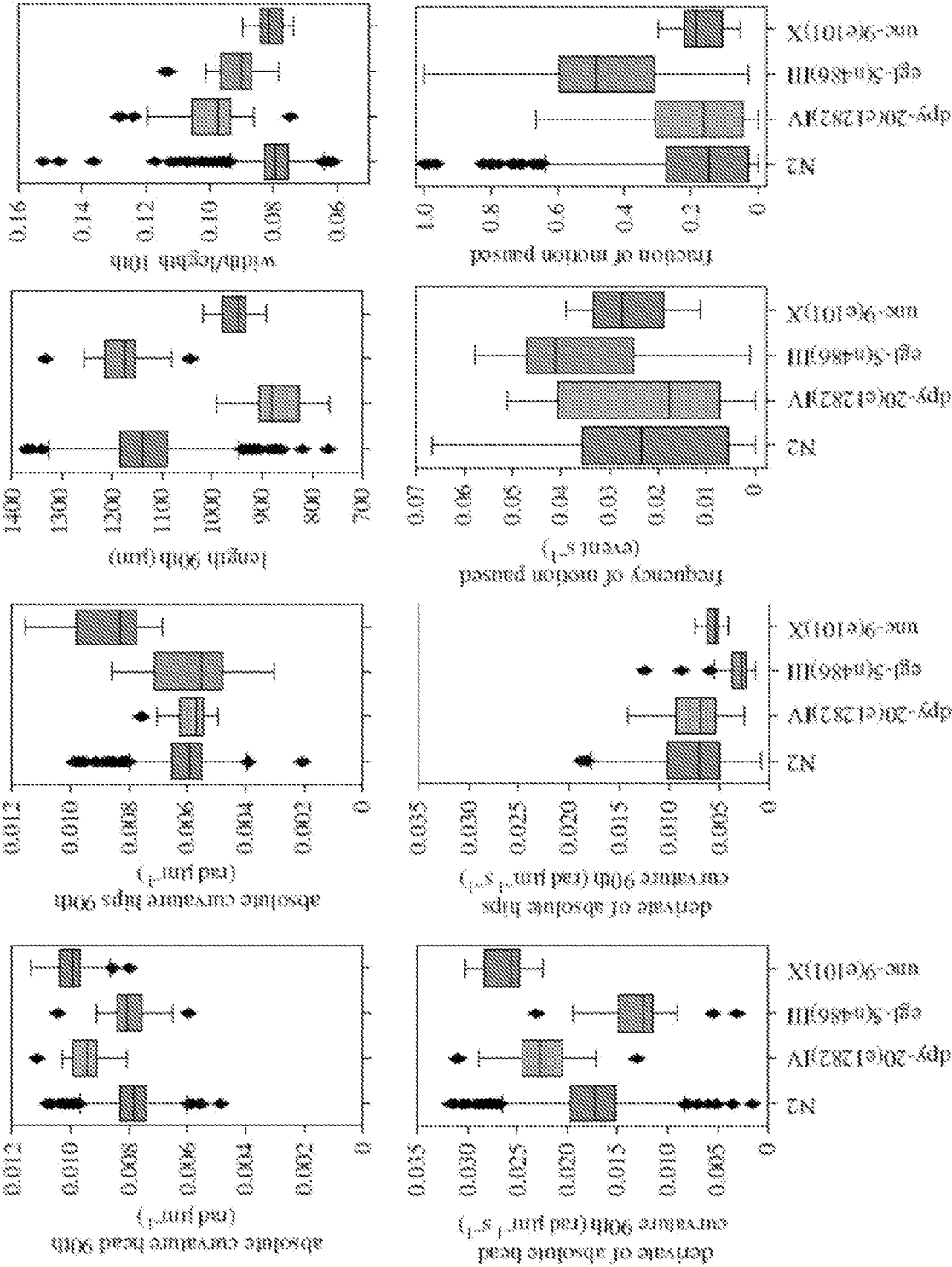


Fig. 9

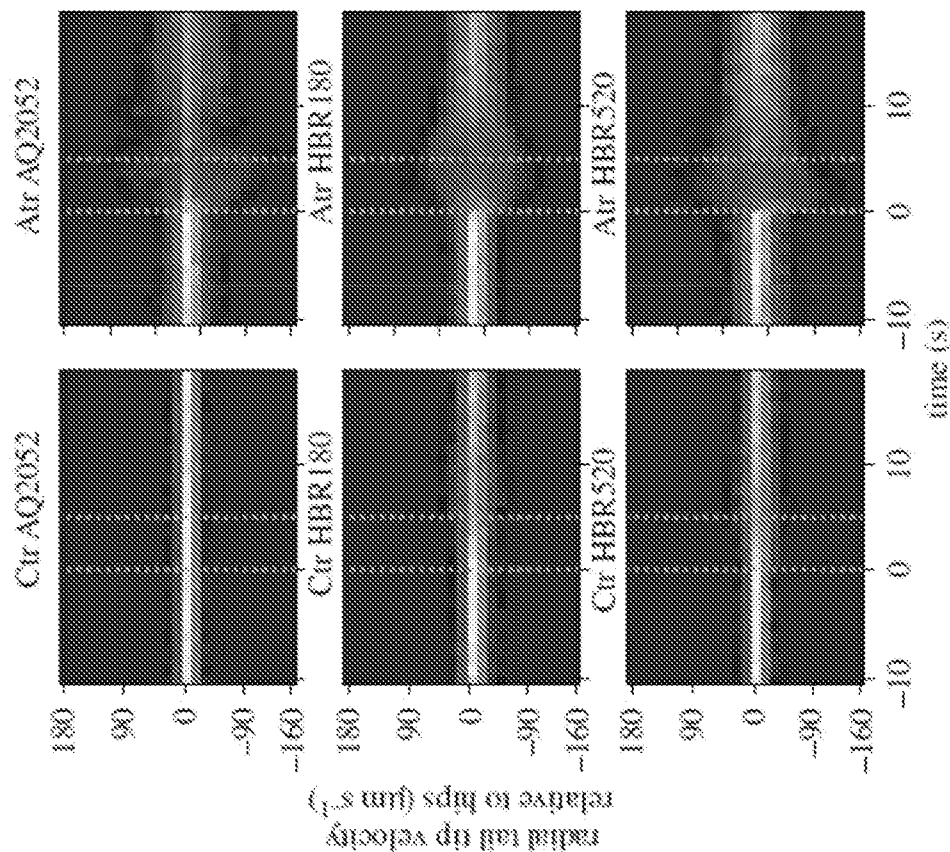


Fig. 11

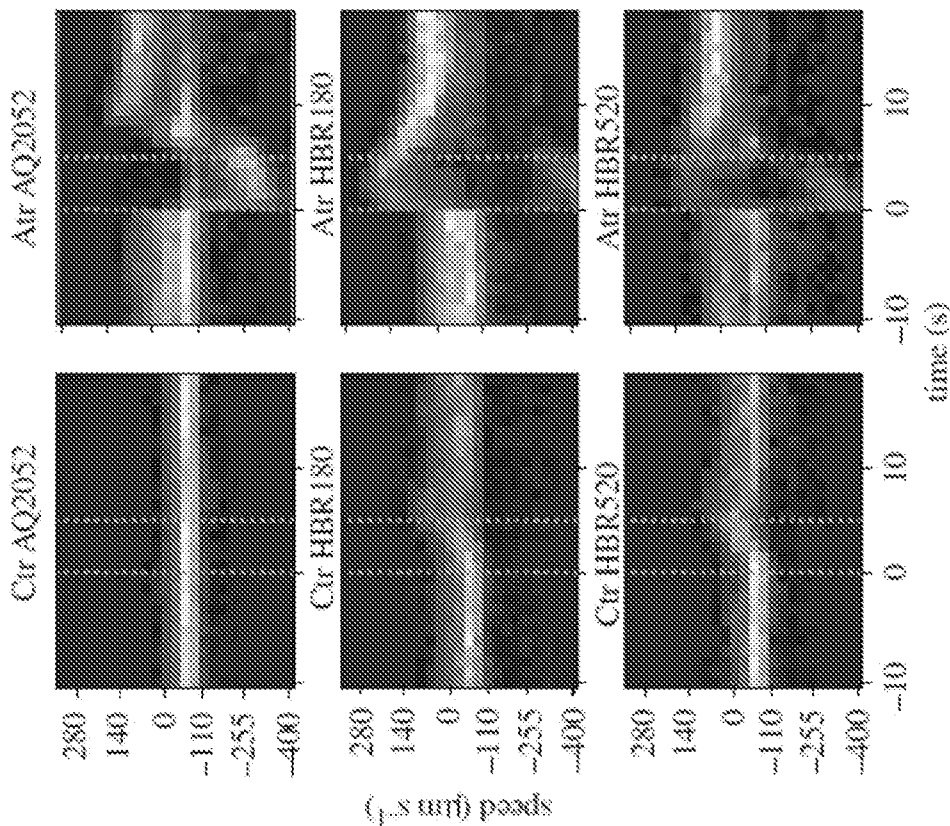


Fig. 10

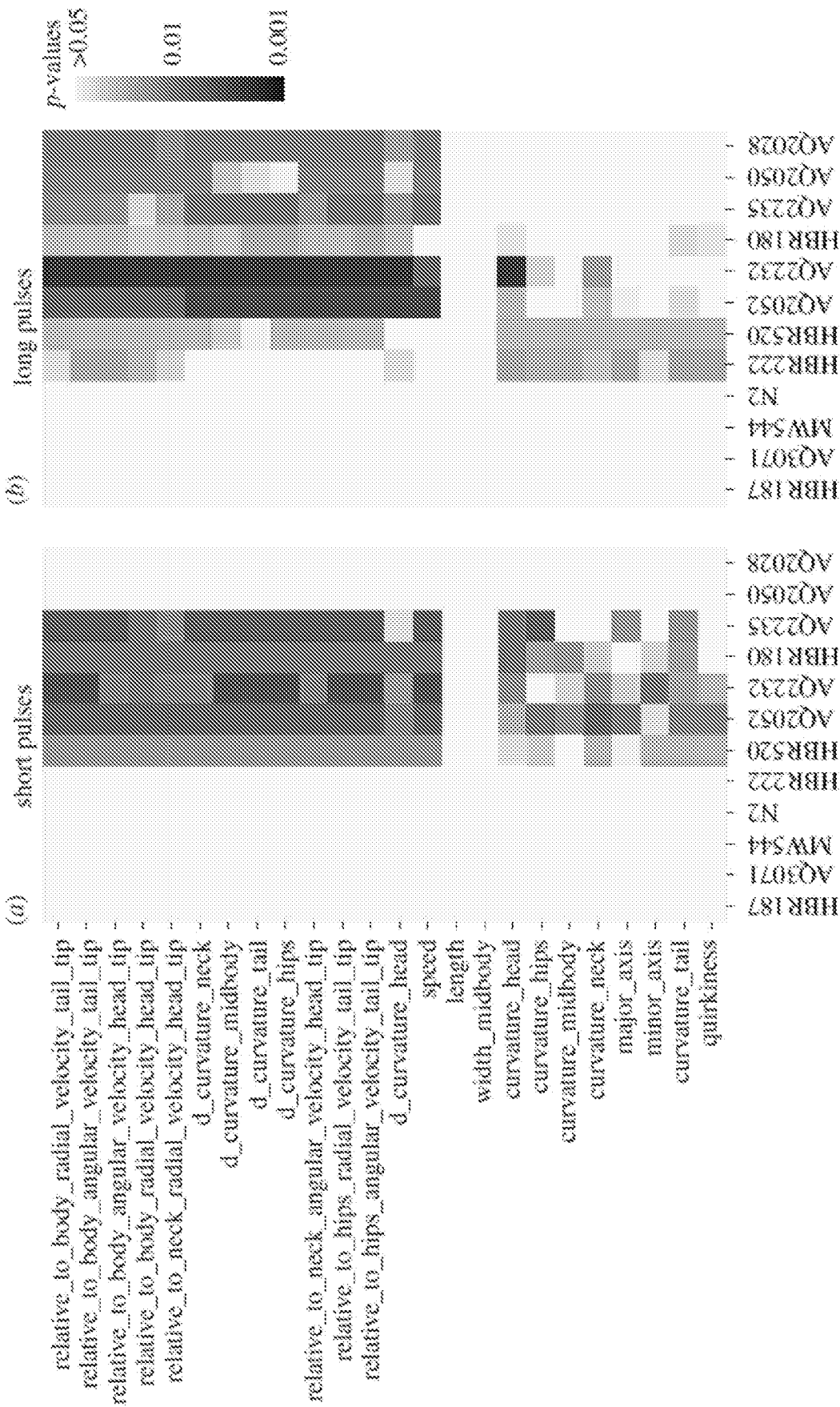


Fig. 12a

Fig. 12b

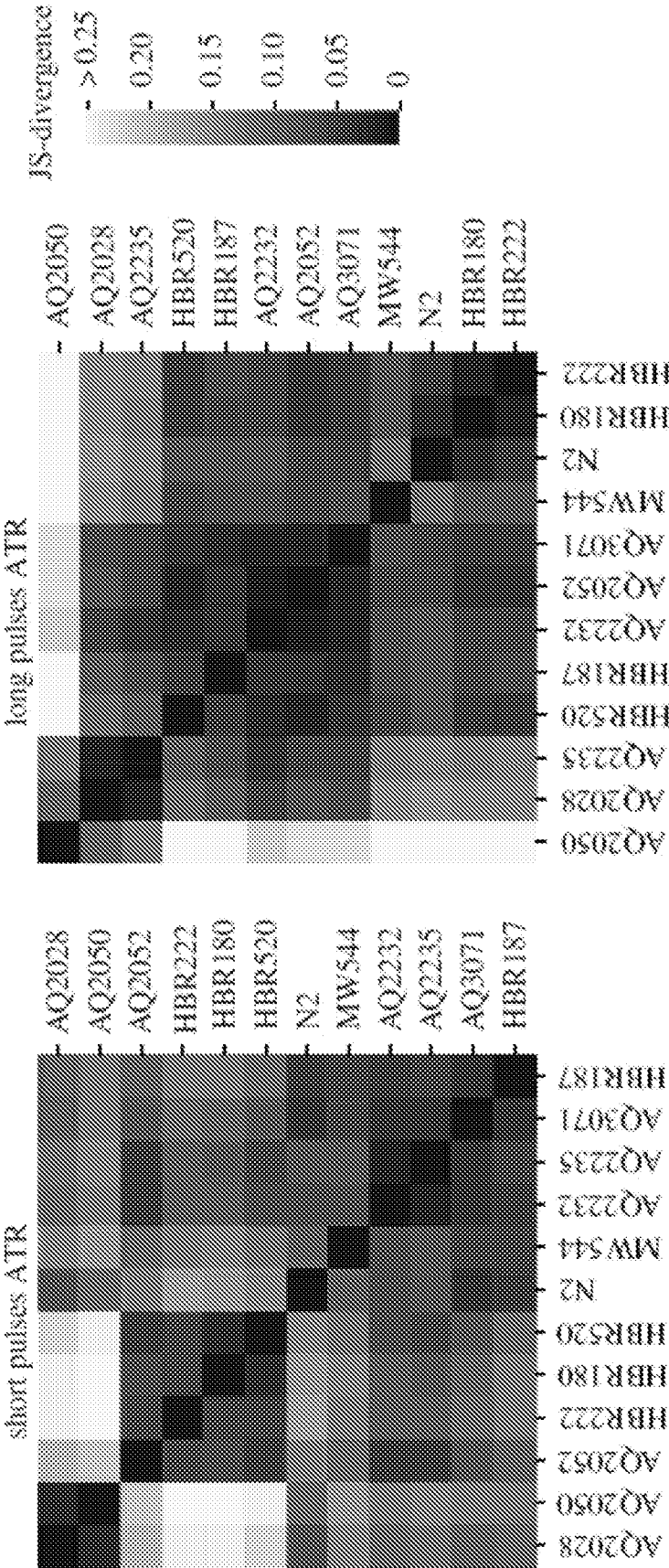


Fig. 14

Fig. 13

12/30

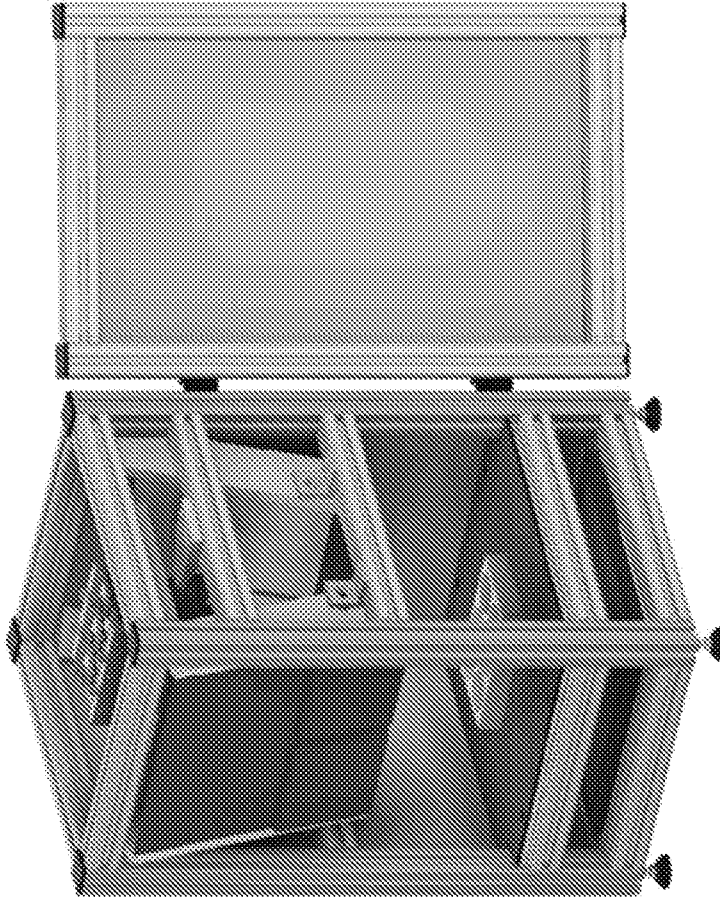


Fig. 15b

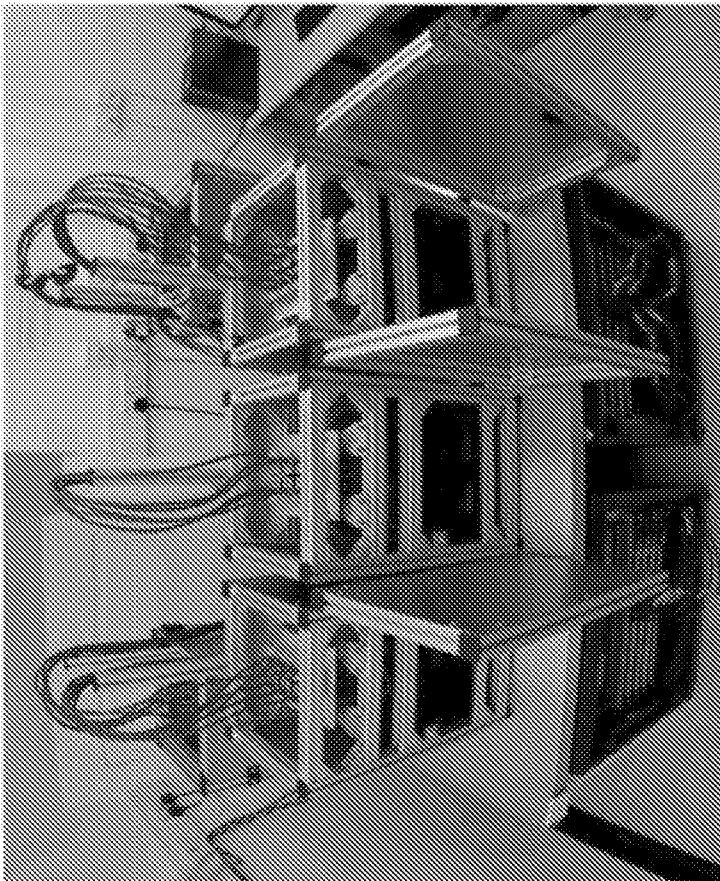


Fig. 15a

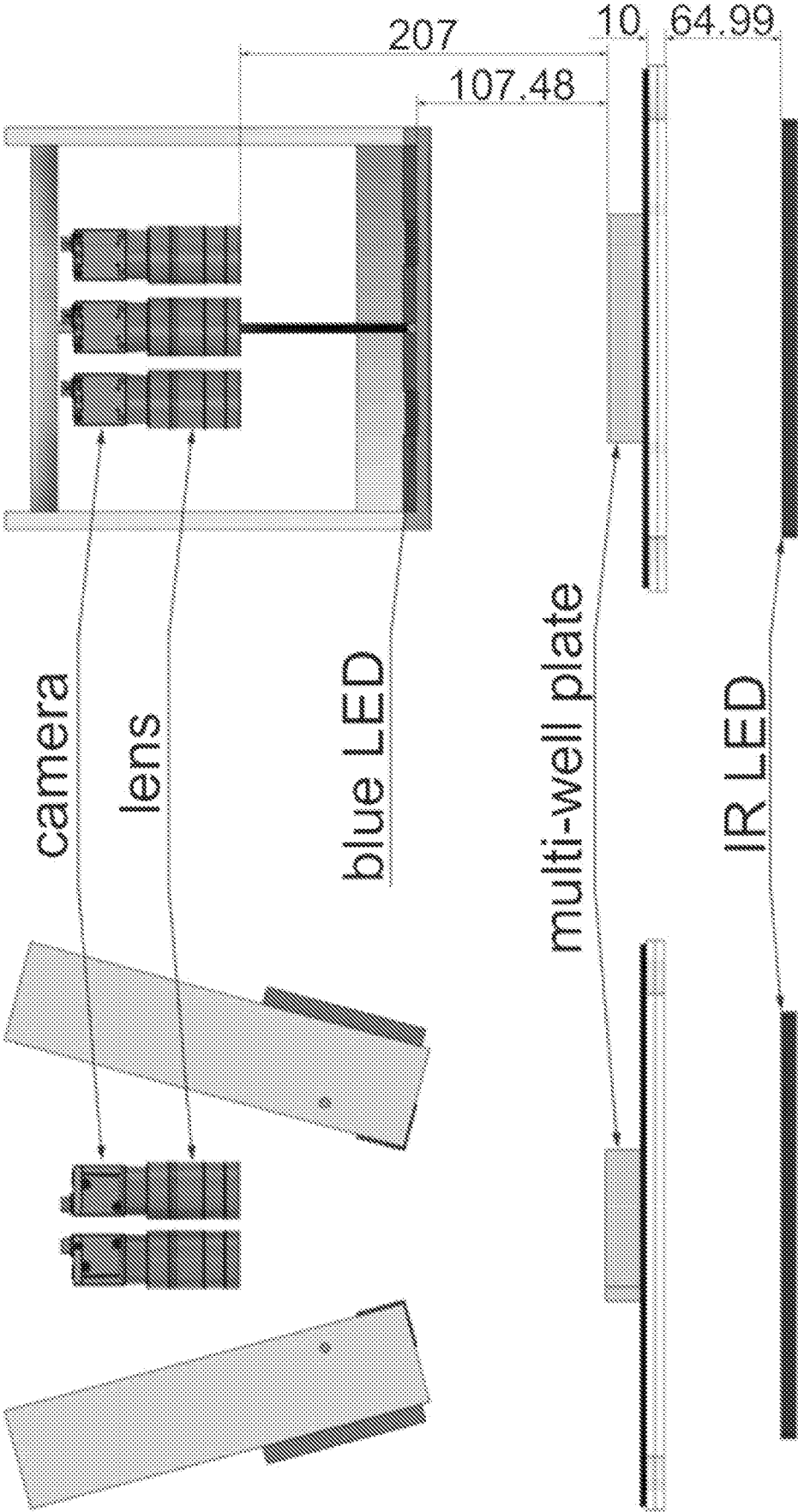


Fig. 15c

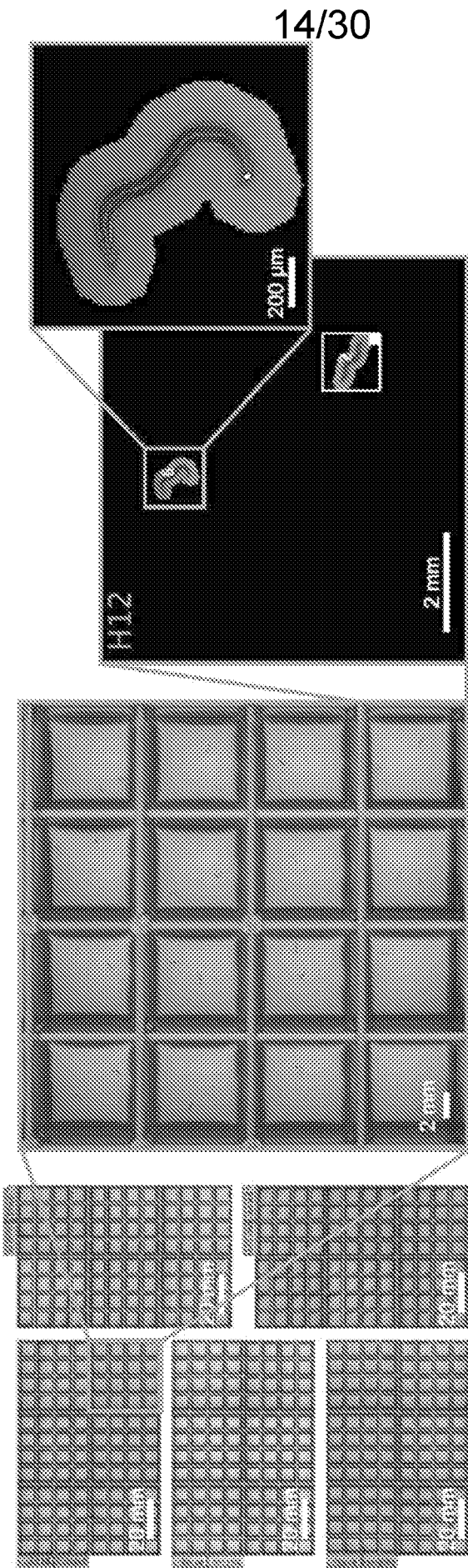


Fig. 15d

Fig. 15e

Fig. 15f

Fig. 15g

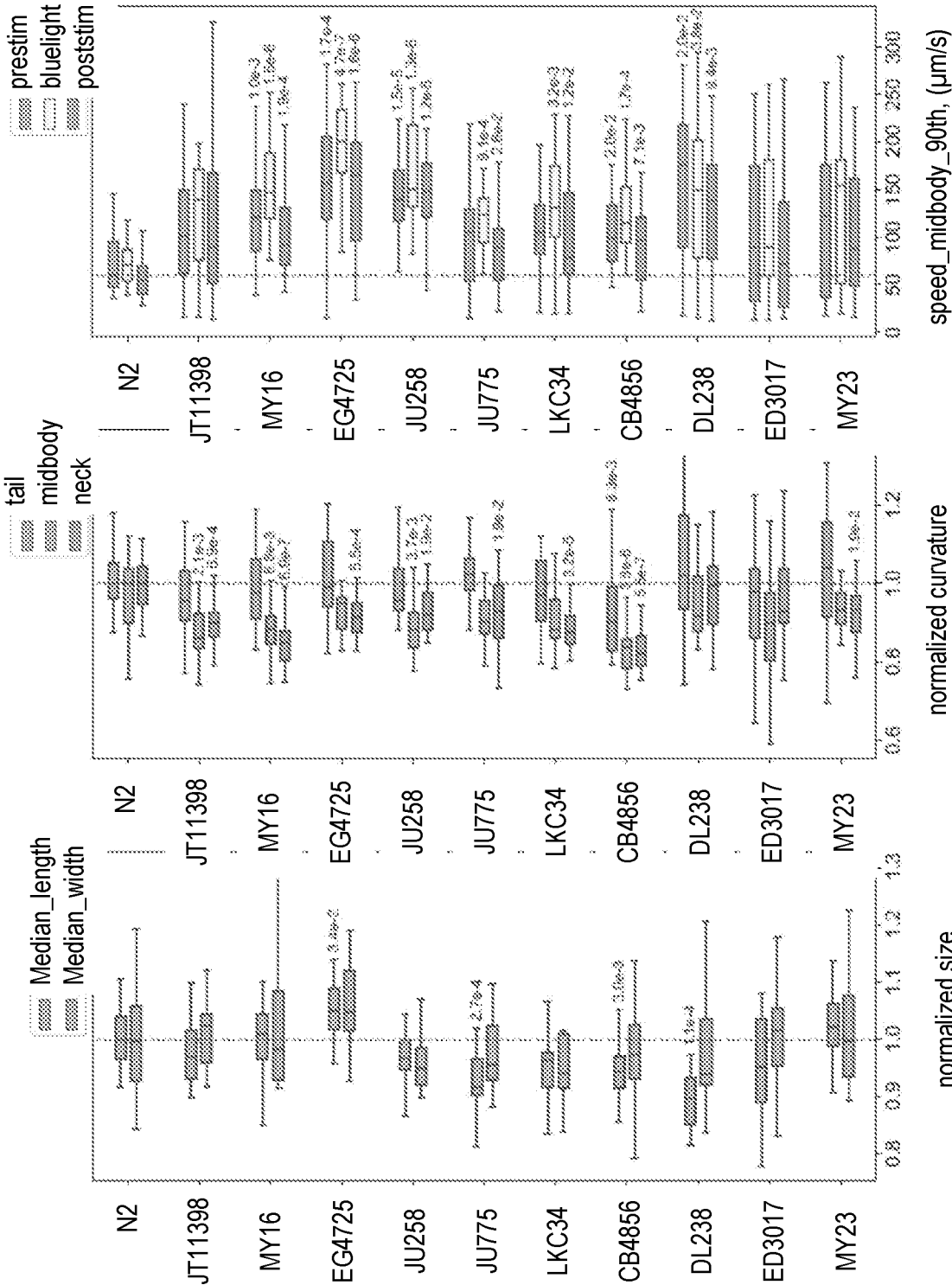


Fig. 16c

Fig. 16b

Fig. 16a

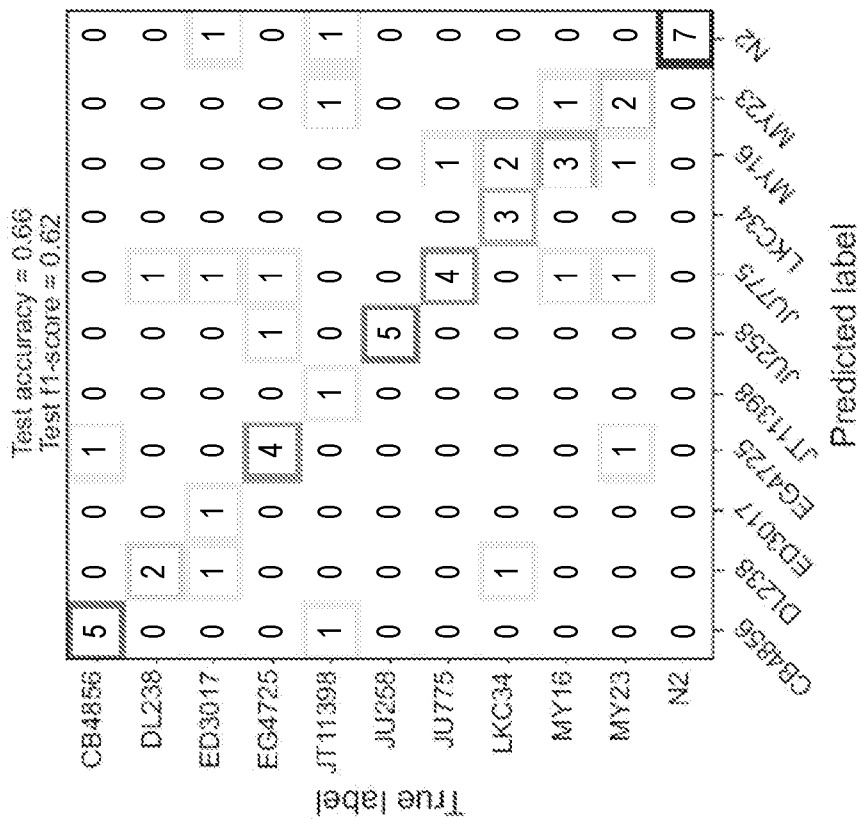


Fig. 18

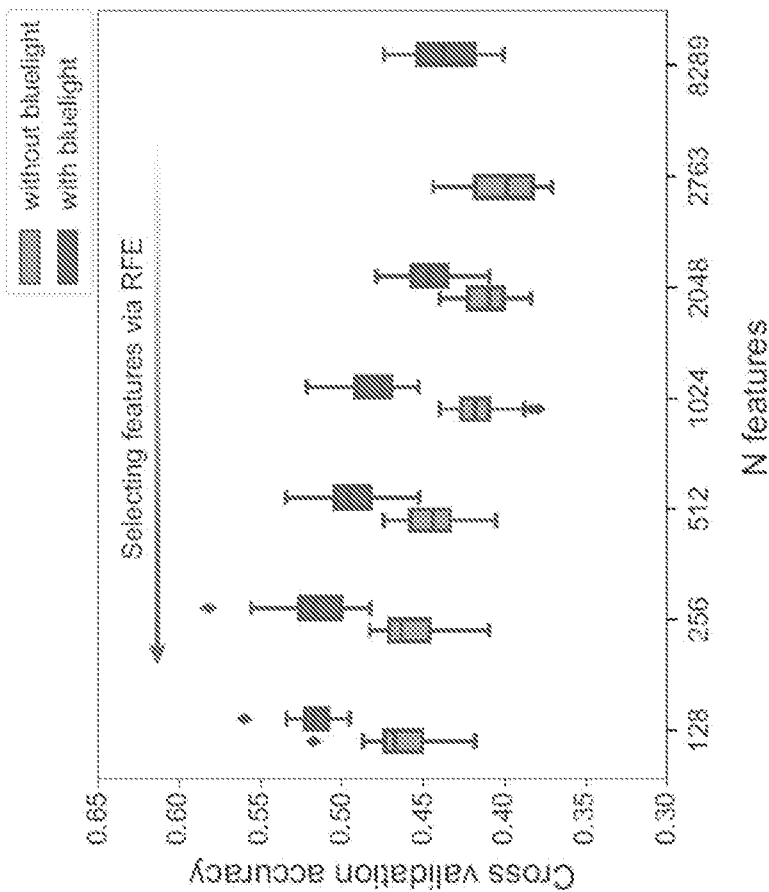


Fig. 17

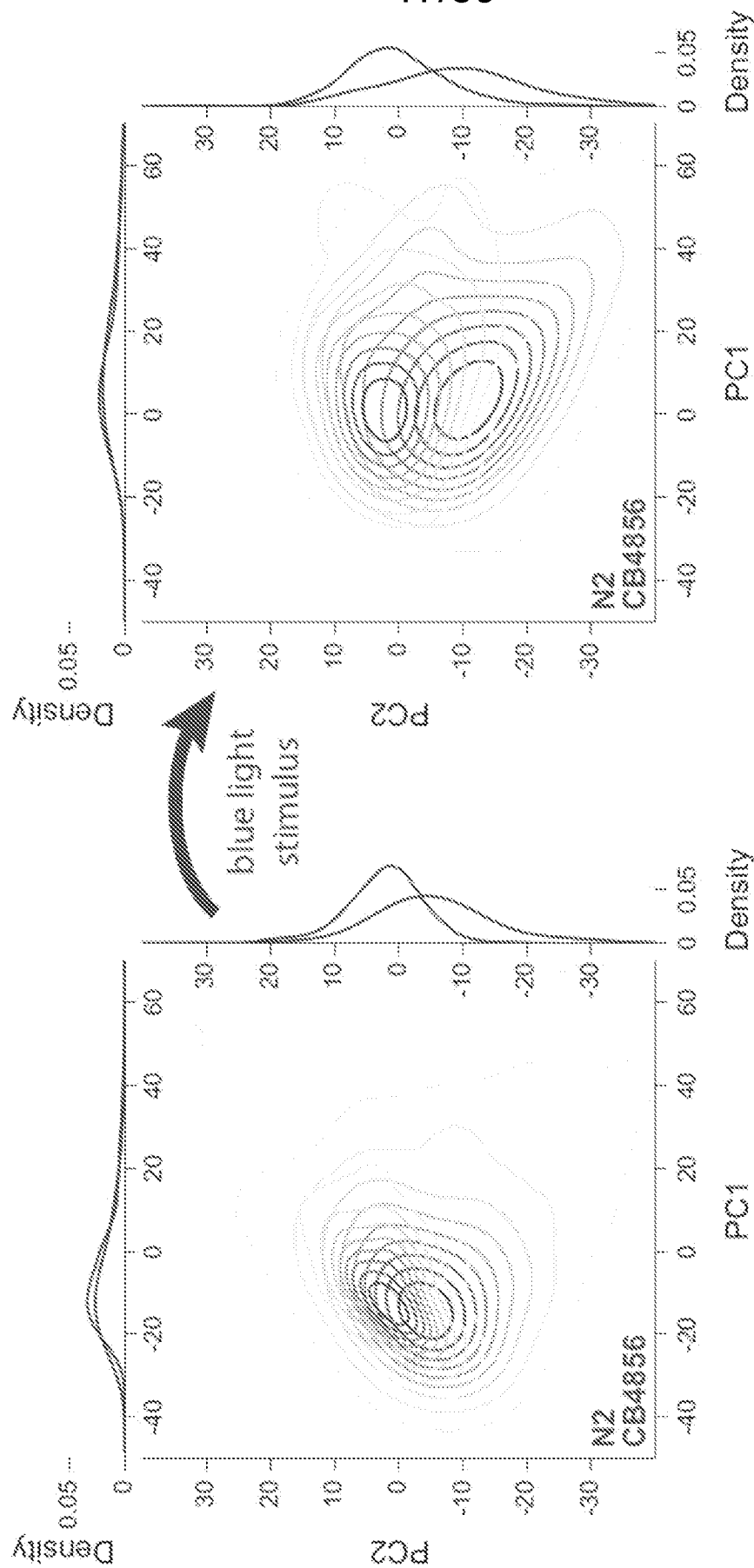


Fig. 19a

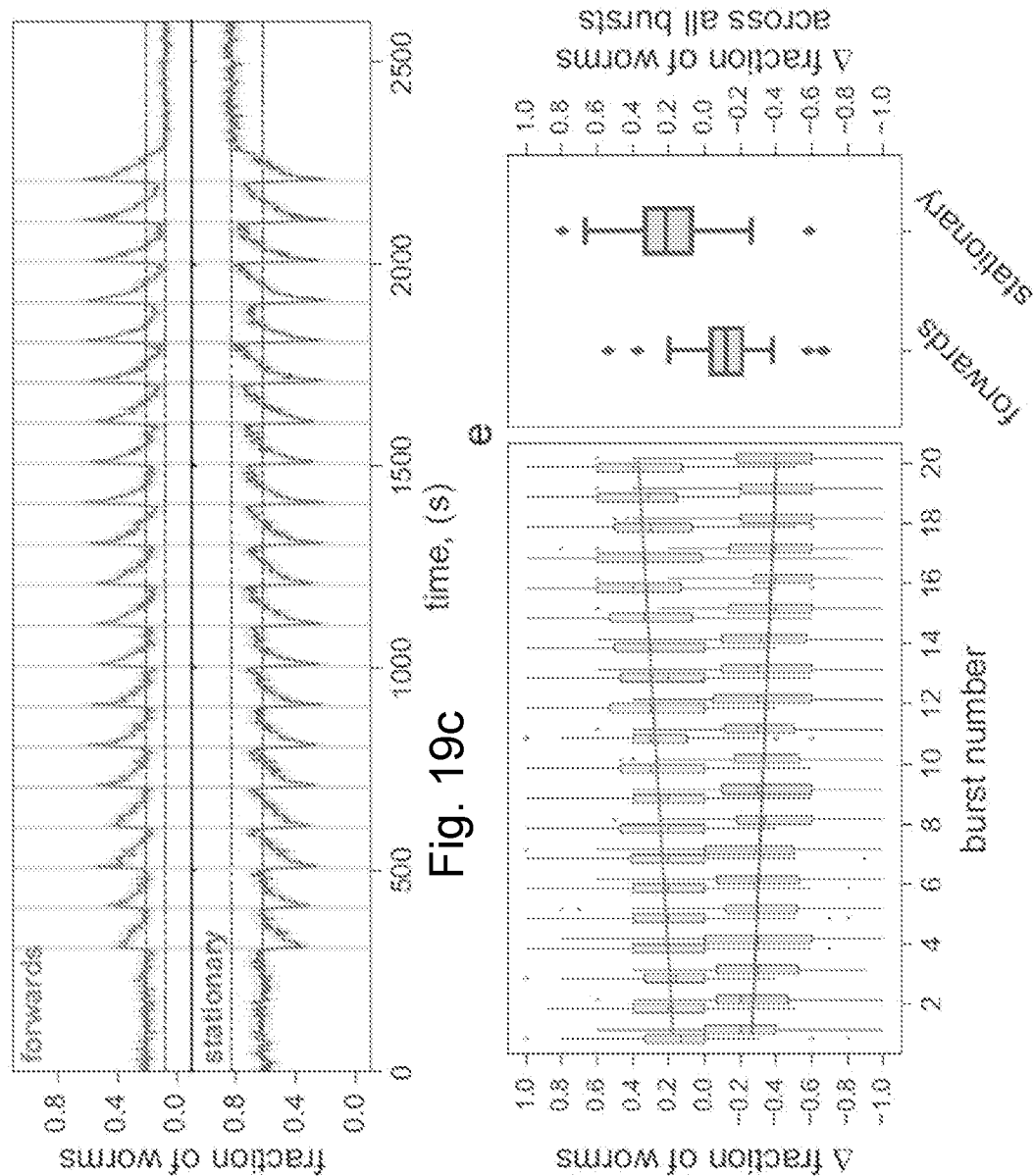


Fig. 19d

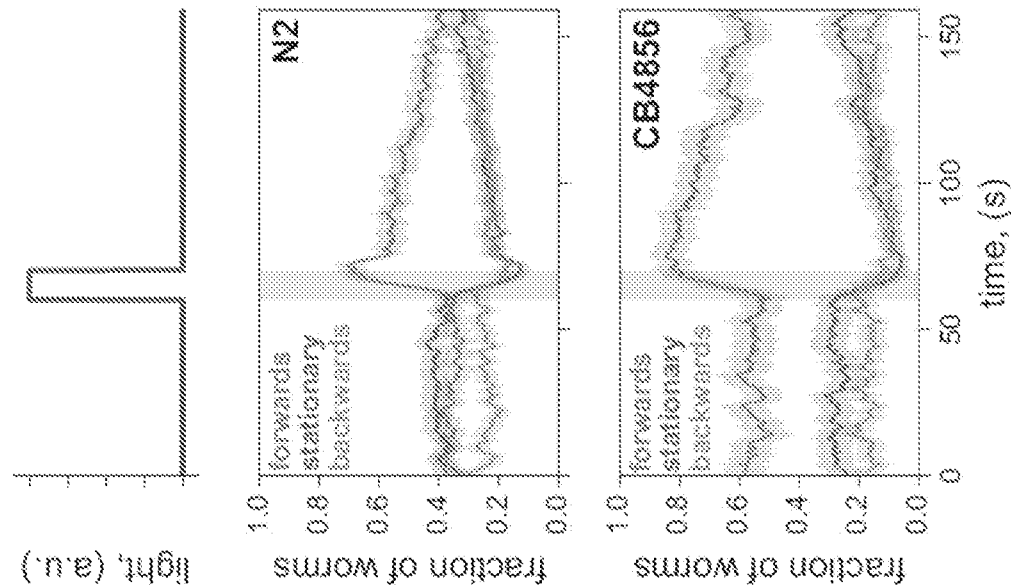


Fig. 19b

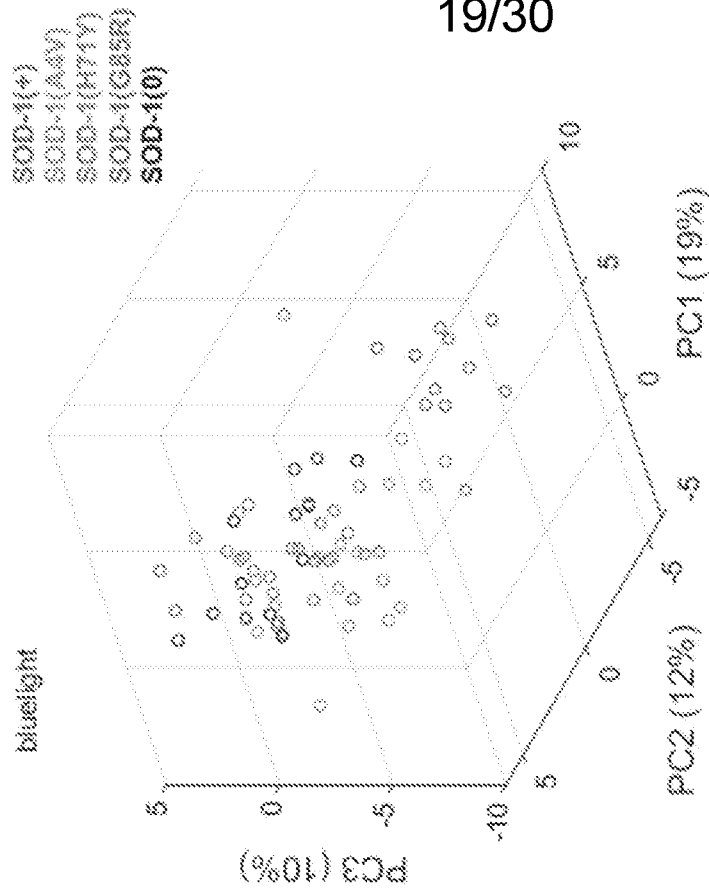


Fig. 20b

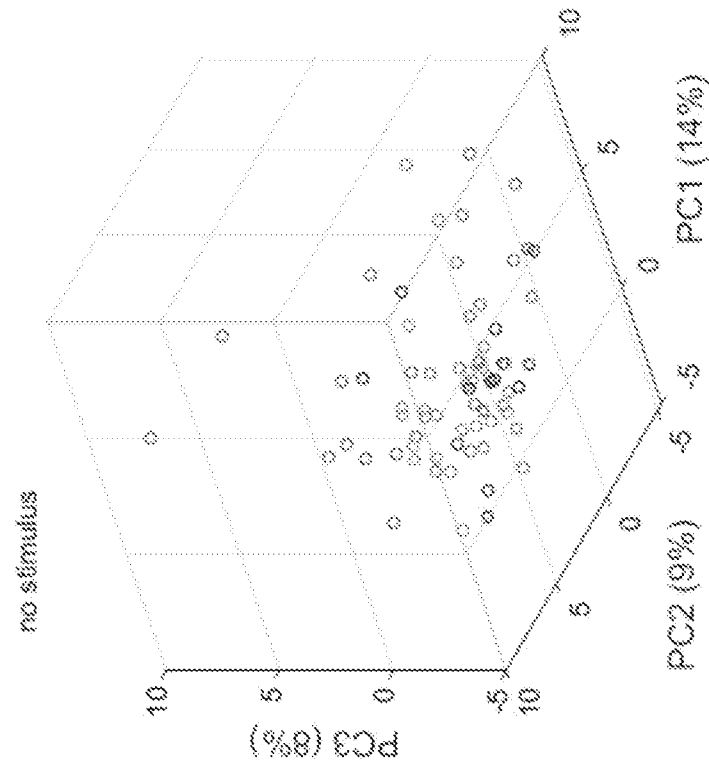


Fig. 20a

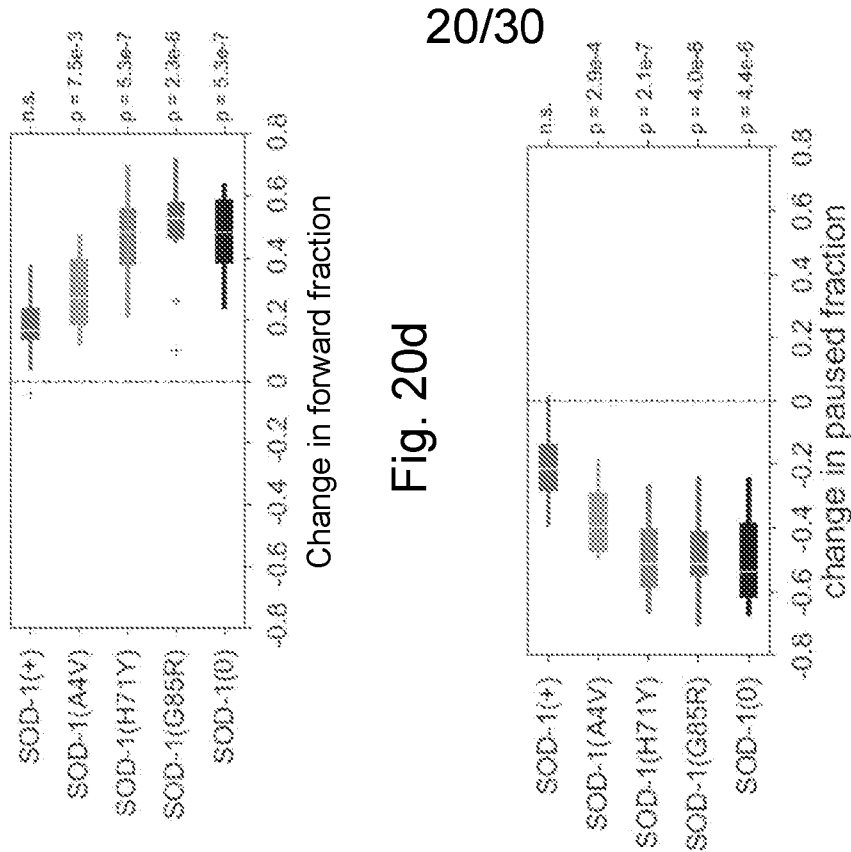


Fig. 20d

Fig. 20e

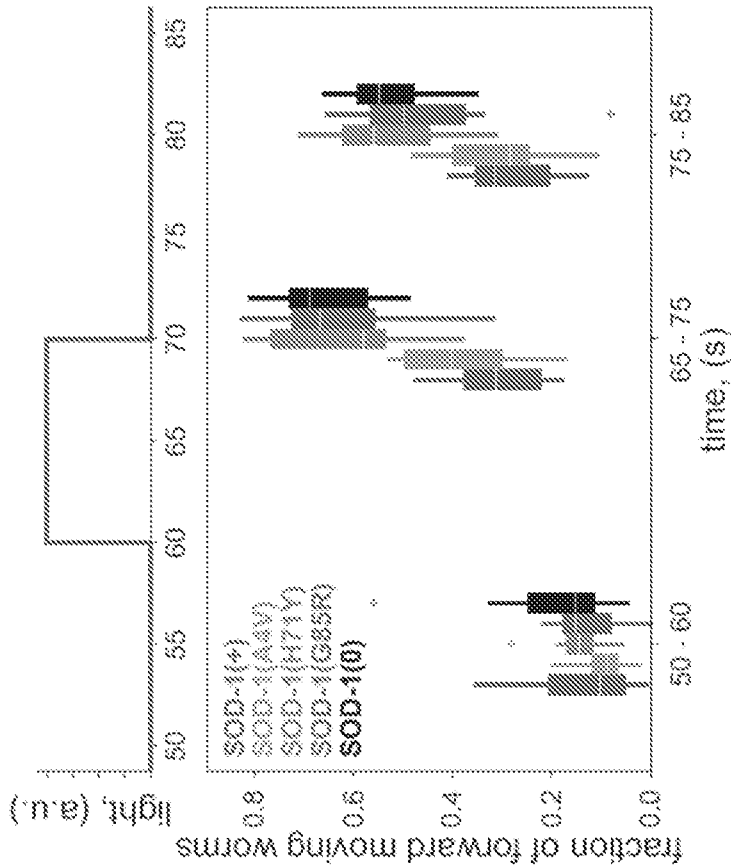


Fig. 20c

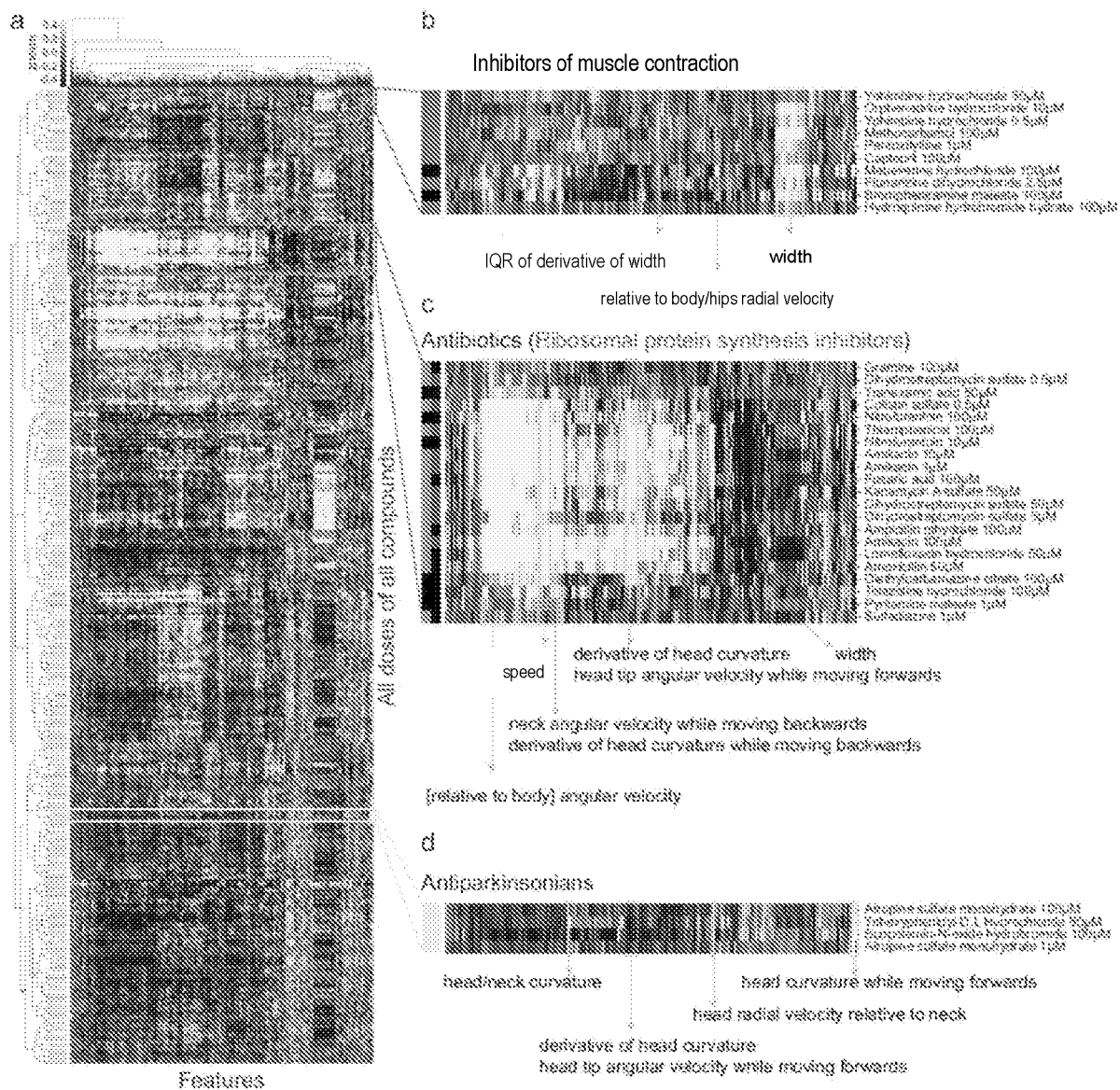


Fig. 21

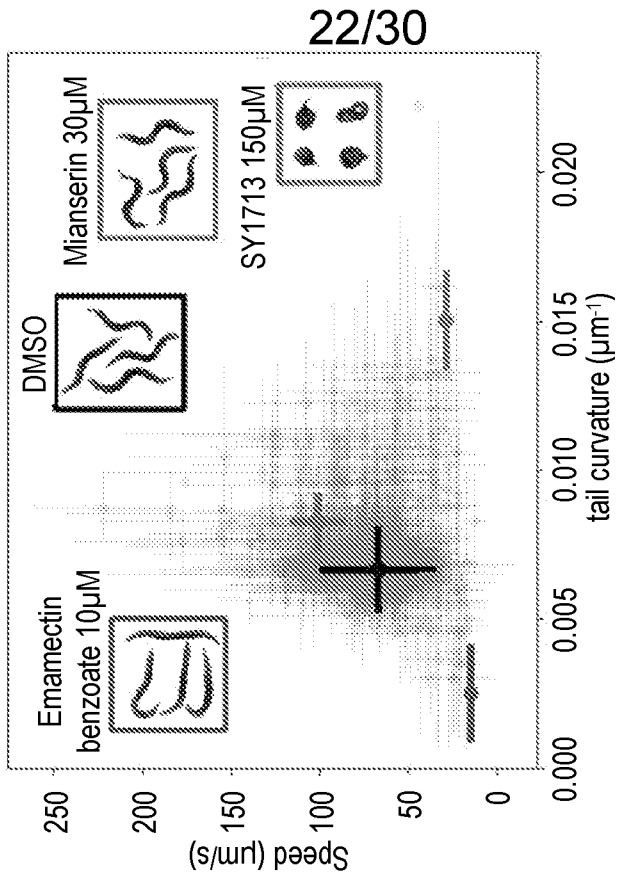
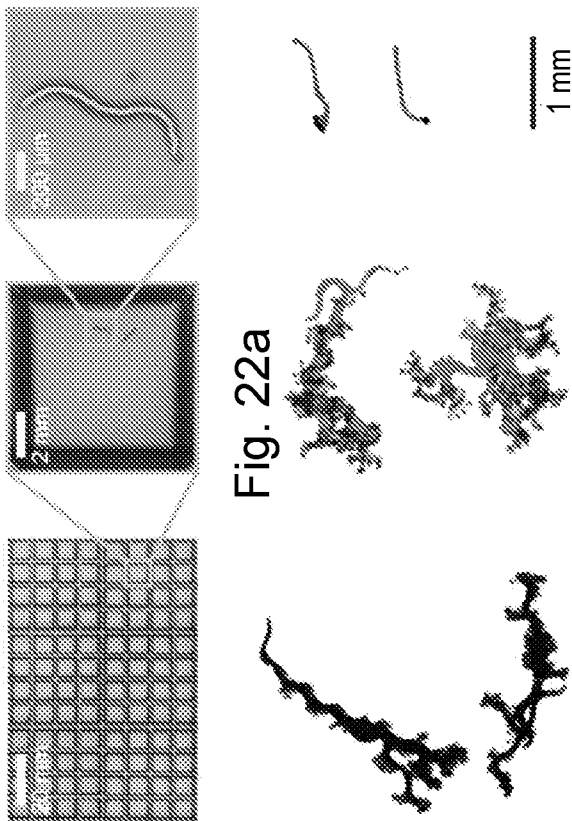


Fig. 22b



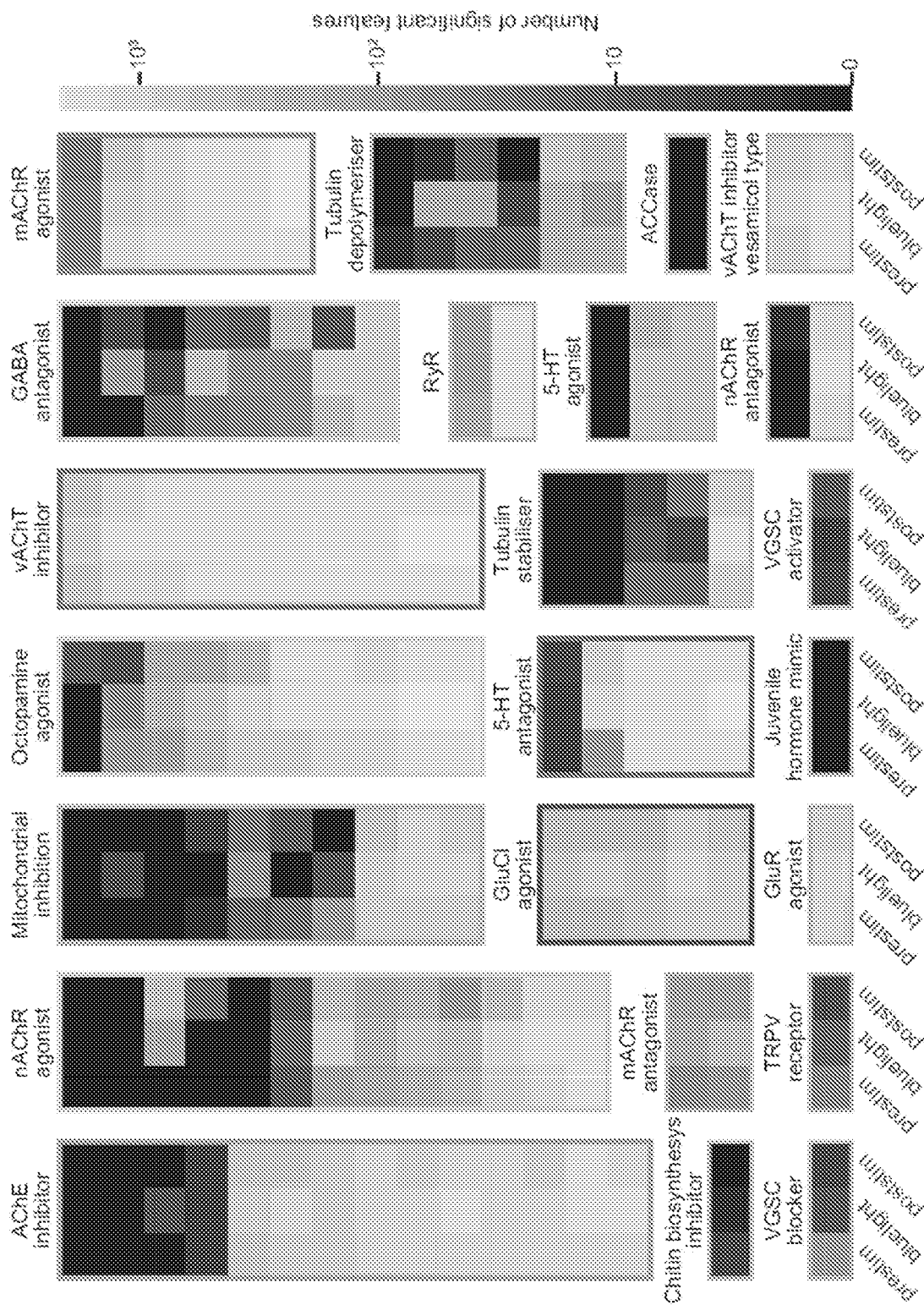


Fig. 22d

24/30

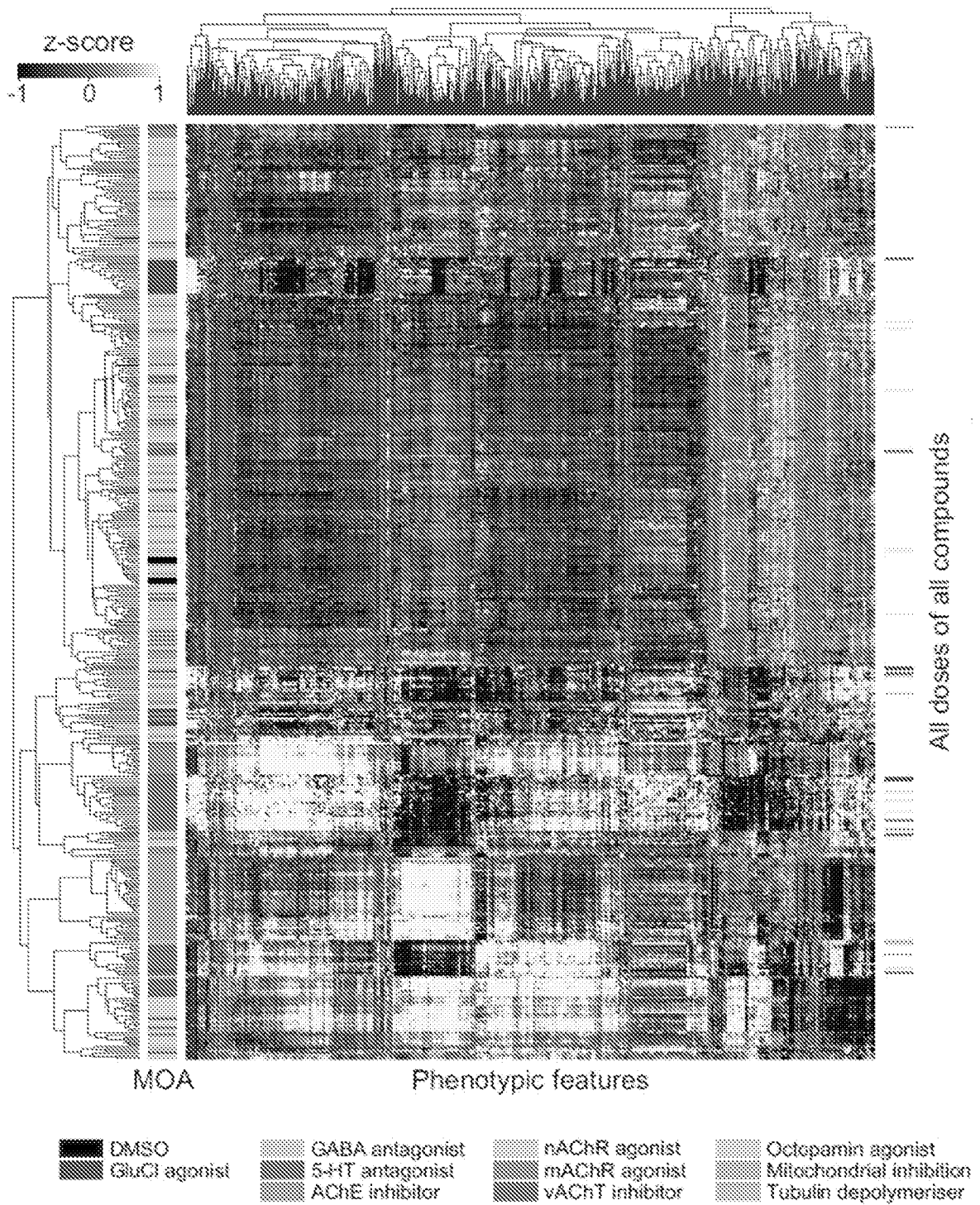


Fig. 23a

25/30

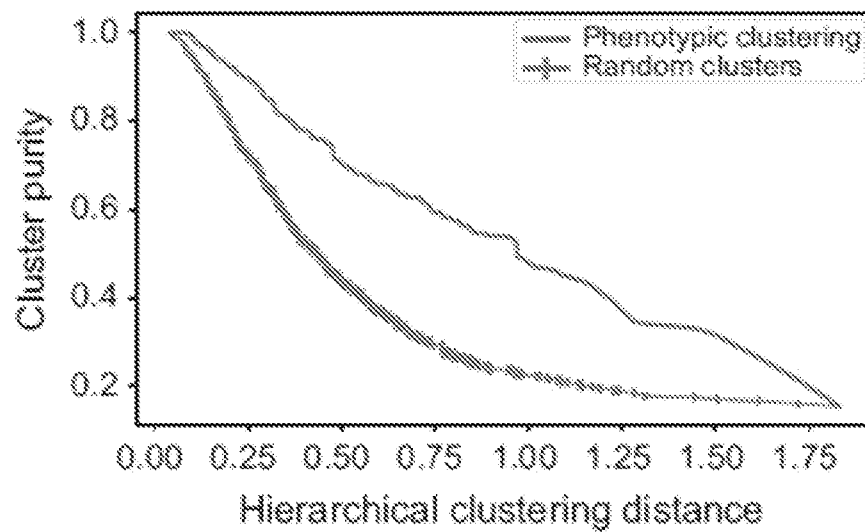


Fig. 23b

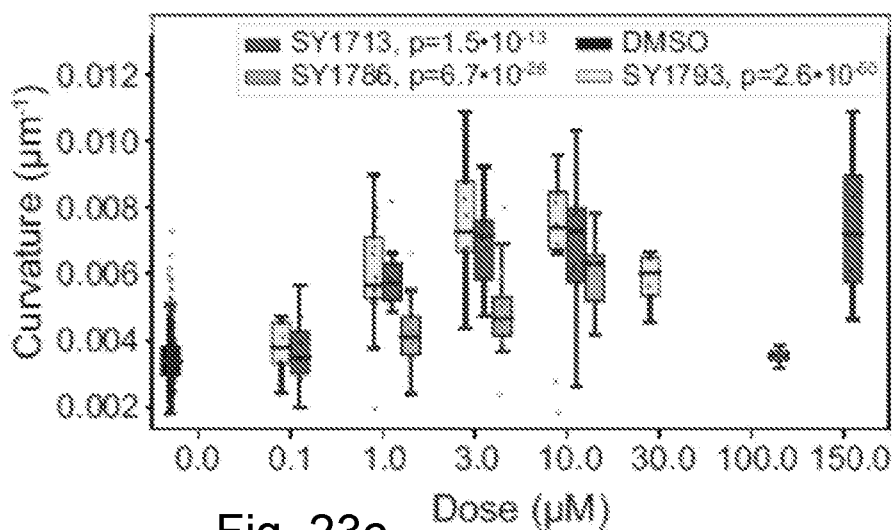
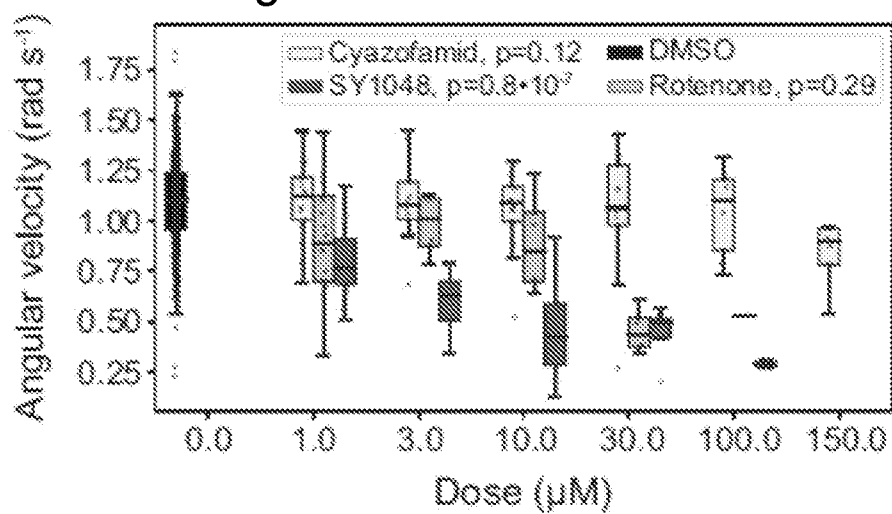


Fig. 23c

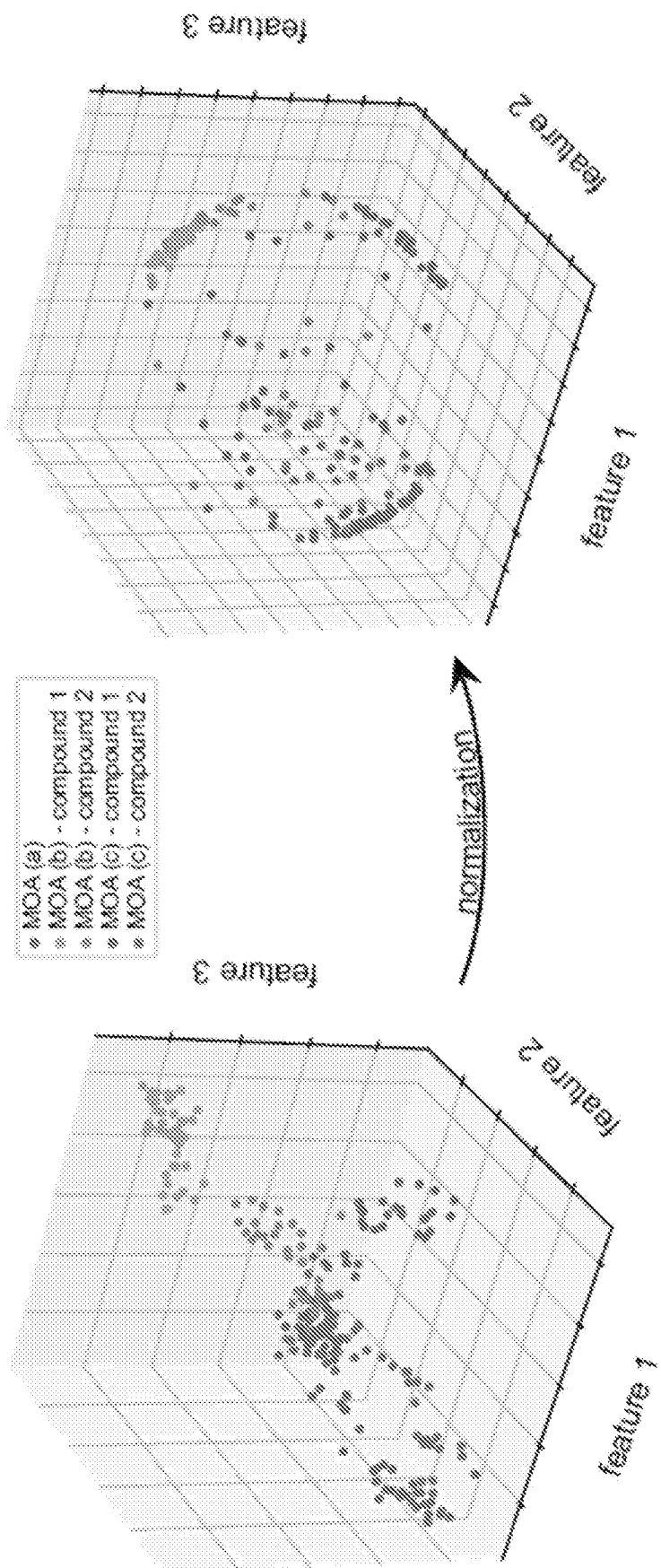


Fig. 24a

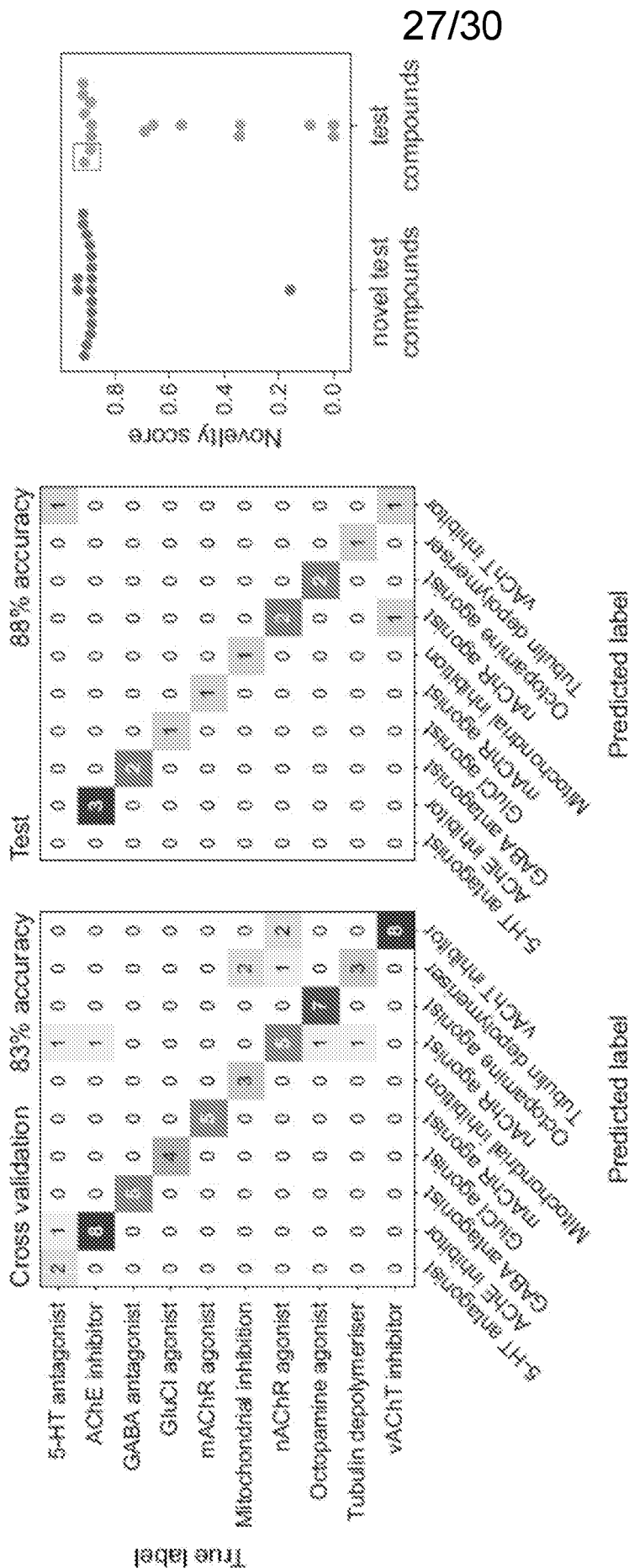


Fig. 24d

Fig. 24c

Fig. 24b

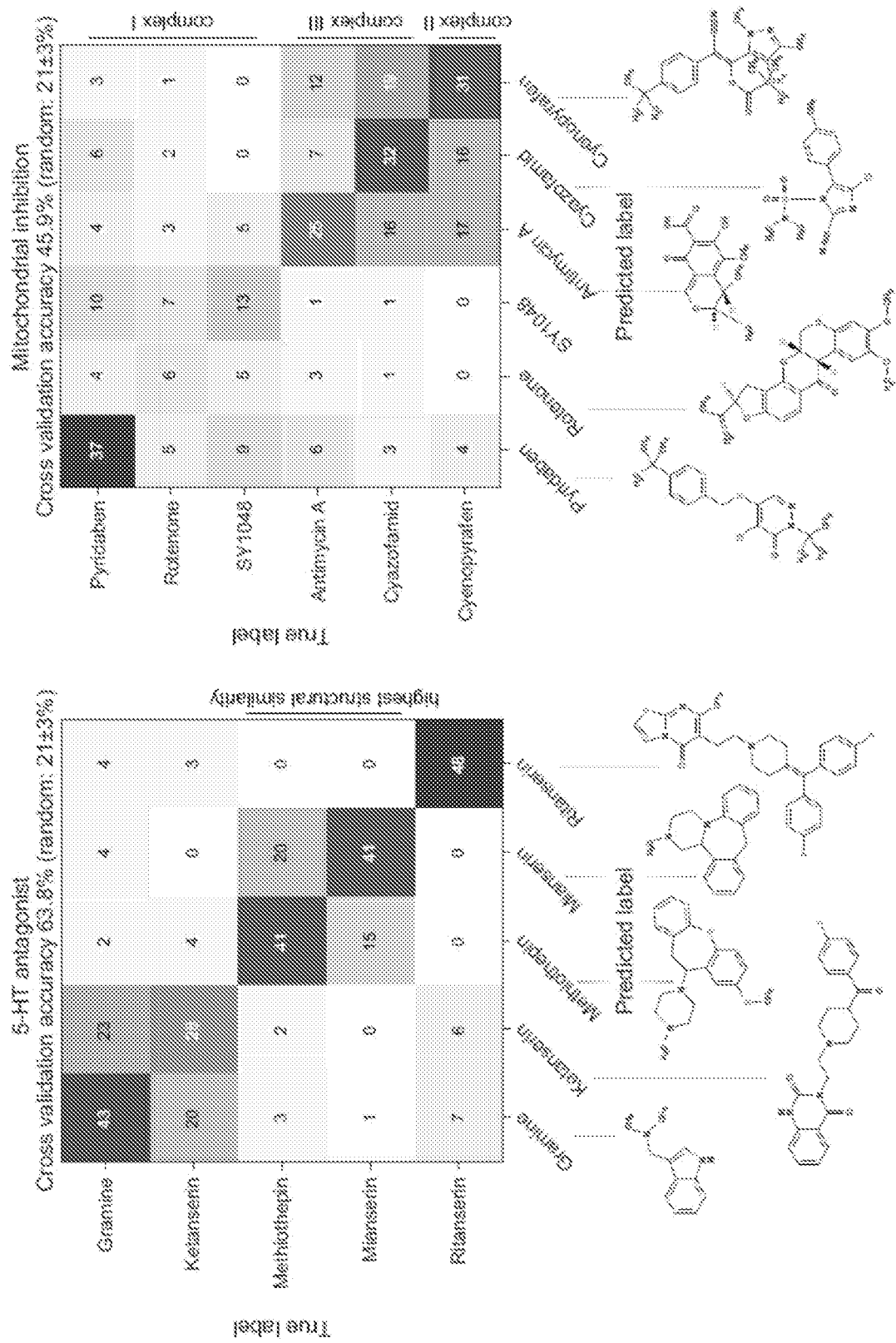


Fig. 25a

Fig. 25b

29/30

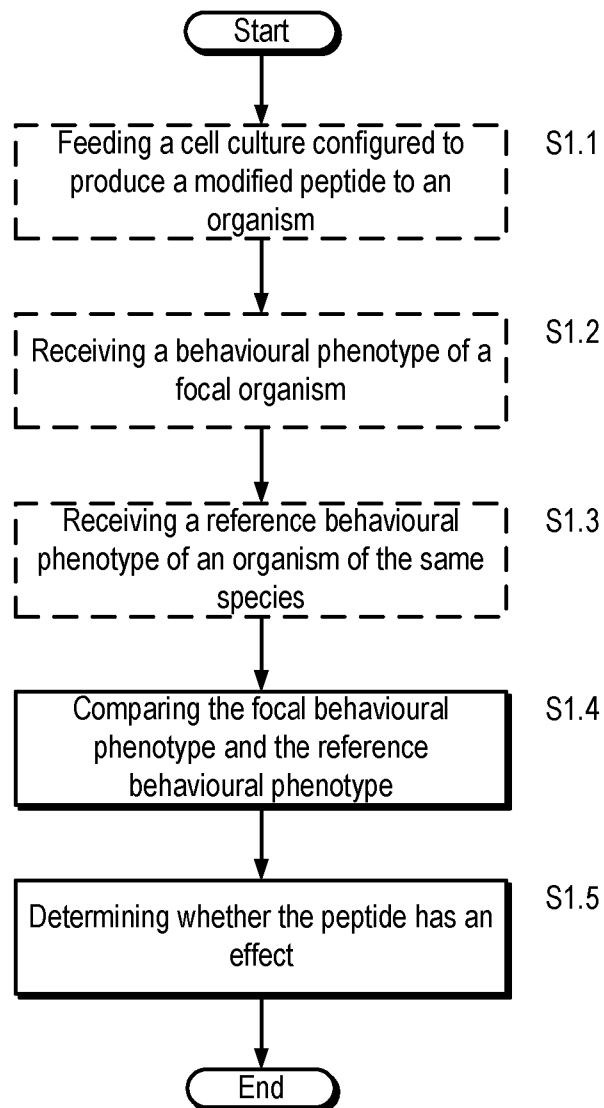


Fig. 26

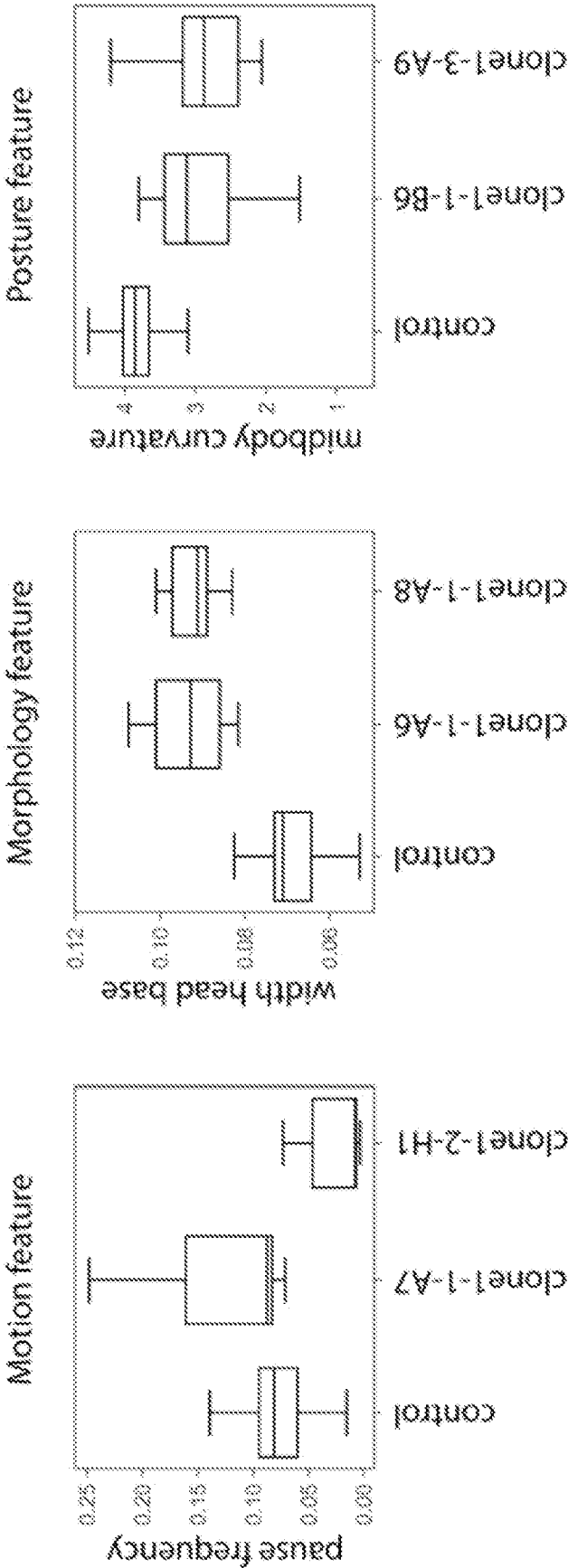


Fig. 27c

Fig. 27b

Fig. 27a

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2024/051496

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06V10/771 C12Q1/02 G06V20/69
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06V G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 929 691 B1 (CADUS PHARMACEUTICAL CORP [US]) 15 December 2004 (2004-12-15) paragraph [0087] paragraph [0228] paragraph [0257] paragraph [0264] ----- - / - -	1 - 20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

18 September 2024

Date of mailing of the international search report

26/09/2024

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

De Coi, Juri

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2024/051496

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TIMMONS LISA ET AL: "Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in <i>Caenorhabditis elegans</i>", GENE, ELSEVIER AMSTERDAM, NL, vol. 263, no. 1, 19 May 2017 (2017-05-19), pages 103-112, XP085030665, ISSN: 0378-1119, DOI: 10.1016/S0378-1119(00)00579-5 abstract section 3.2 section 3.4	1-20
Y	----- BARLOW IDA L. ET AL: "Megapixel camera arrays enable high-resolution animal tracking in multiwell plates", COMMUNICATIONS BIOLOGY, [Online] vol. 5, 23 March 2022 (2022-03-23), pages 1-13, XP093184514, ISSN: 2399-3642, DOI: 10.1038/s42003-022-03206-1 Retrieved from the Internet: URL:https://www.nature.com/articles/s42003-022-03206-1> [retrieved on 2024-09-10] page 8, column 1, line 13 - line 27 -----	17-19

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2024/051496

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0929691	B1	15-12-2004	AT E284972 T1 15-01-2005
			AU 733007 B2 03-05-2001
			CA 2266543 A1 02-04-1998
			DE 69731960 T2 15-12-2005
			EP 0929691 A2 21-07-1999
			JP 2001501097 A 30-01-2001
			US 6159705 A 12-12-2000
			WO 9813513 A2 02-04-1998
