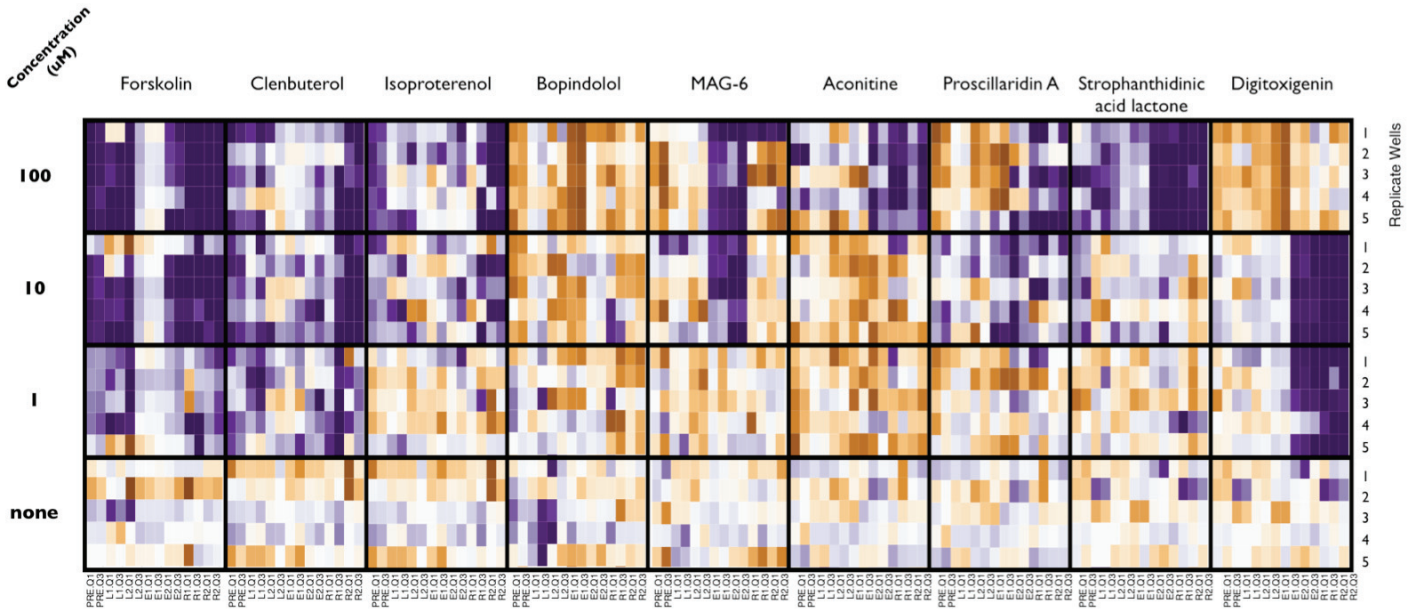
Rapid behavior-based identification of neuroactive small molecules in the zebrafish

David Kokel[1,2], Jennifer Bryan[3,3b], Christian Laggner[4], Rick White[3], Chung Yan J. Cheung[1,2], Rita Mateus[1,2], David Healey[1,2], Sonia Kim[1,2], Andreas A. Werdich[1], Stephen J. Haggarty[2,5,6], Calum A. MacRae[1], Brian Shoichet[4], and Randall T. Peterson[1,2]

SUPPLEMENTARY RESULTS

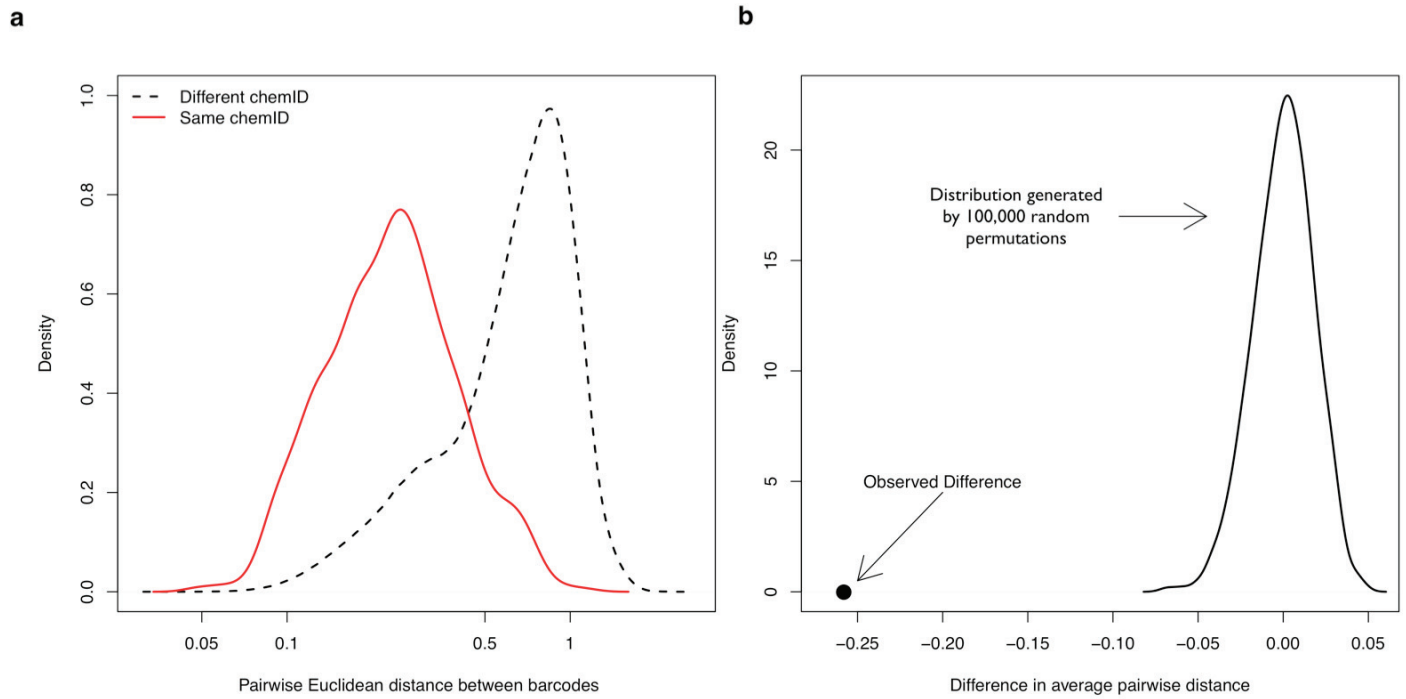
Supplementary Figure 1.



Supplementary Figure 1. Dose response relationships for forskolin, clenbuterol, isoproterenol, bopindolol, MAG-6, aconitine, proscillaridin A, strophanthidinic acid lactone, and digitoxigenin. The behavioral barcodes of five replicate wells are shown for each concentration indicated.

Nature Chemical Biology: doi:10.1038/nchembio.307

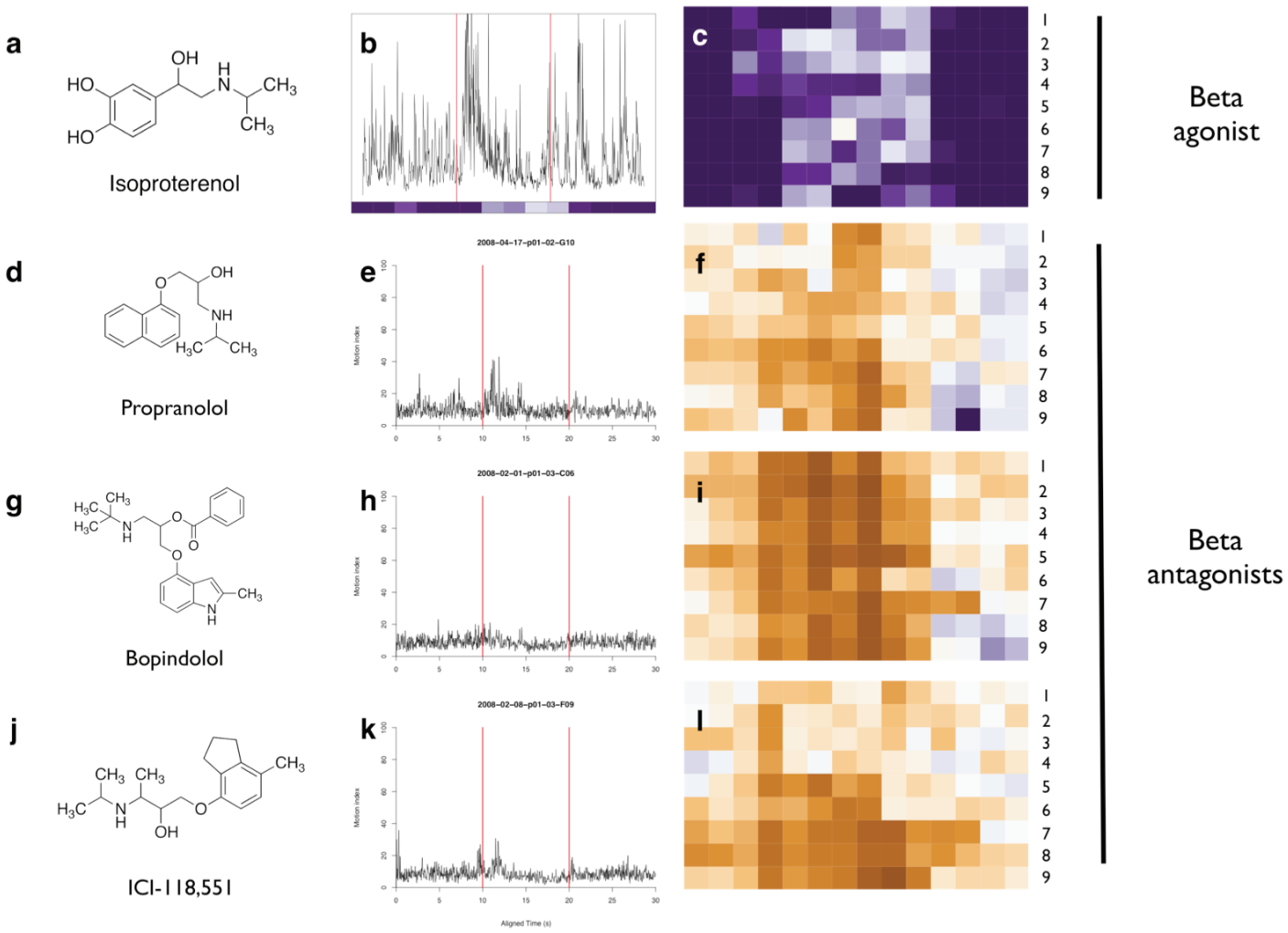
Supplementary Figure 2.



Supplementary Figure 2. (**a**) Smoothed histograms depicting the distribution of pairwise distances between barcodes that have the same (red solid line) and different (black dotted line) chemID, which identifies an individual well of an underlying source library. Prior to computing Euclidean distance, the barcodes were transformed with the cumulative normal probability distribution (mean = 0, sd = 7), exactly as was done prior to clustering. Note the distances are analyzed on the log scale, to increase symmetry. (**b**) Assessment of the null hypothesis of no difference in the average log-transformed distance for barcode pairs with the same vs. different chemID. Null distribution approximated by a distribution obtained from 100,000 random permutations of 1,000 randomly selected chemIDs. An upper bound for the p-value of the observed difference (-0.258) can be obtained non-parametrically ($p<10^{-5} = 10^{-54}$) or using a normal approximation ($p<9.98e-71$).

Nature Chemical Biology: doi:10.1038/nchembio.307

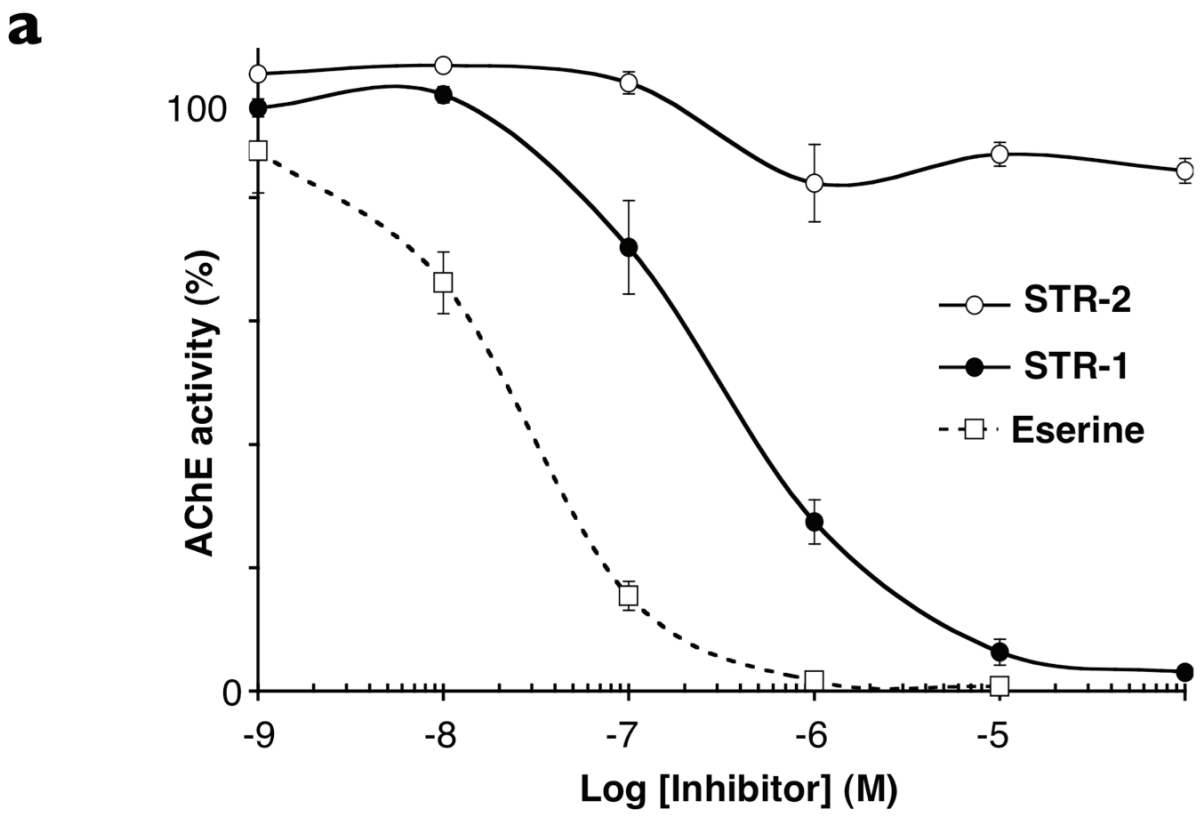
Supplementary Figure 3.



Supplementary Figure 3. Chemicals with opposing mechanisms of action cause opposite behavioral phenotypes. (**a-c**) Isoproterenol, an ADBR agonist, increases zebrafish motor activity during all phases of the PMR. (**d-l**) Three ADBR antagonists, propranolol, bopindolol and compound ICI-118,557, reduce motor activity during all phases of the PMR.

Nature Chemical Biology: doi:10.1038/nchembio.307

Supplementary Figure 4.

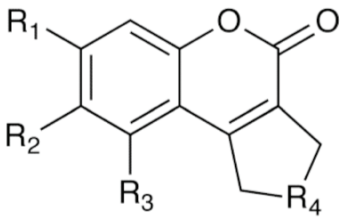


Supplementary Figure 4. Activities of AChEIs *in vitro*. (**a**) Quantification of purified AChE activity with eserine, STR-1, or STR-2 at the indicated concentrations. Data points and error bars represent mean values ± standard deviation.

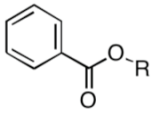
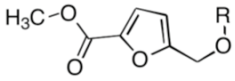
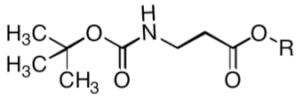
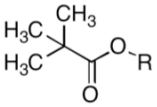
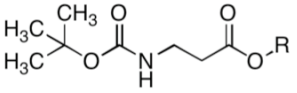
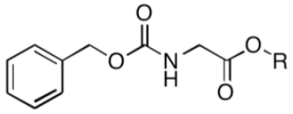
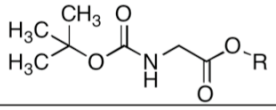
Nature Chemical Biology: doi:10.1038/nchembio.307

Supplementary Figure 5.

a



b

MAG	IC50 (M)	R1	R2	R3	R4
MAG-1	1.E-09		H	H	C2H4
MAG-2	nd	CH3	H		C2H4
MAG-3	1.E-06		Cl	H	CH2
MAG-4	nd		H	H	C2H4
MAG-5	1.E-05		Cl	H	C2H4
MAG-6	0% (10 uM)		Cl	H	CH2
MAG-7	5.E-07		Cl	H	CH2

Supplementary Figure 5. MAG compounds are MAO inhibitors. Chemical structures of (**a**) the common chemical scaffold shared by MAG compounds (**b**) Substitution patterns of the various MAG compounds and their half maximal inhibitory concentrations (IC_{50}) for MAO.

Nature Chemical Biology: doi:10.1038/nchembio.307

Table S1. Statistical analysis of the phenotype-based clusters

Group	Main ref.	No. of stat.sig. fingerprint elements (alpha= 0.05)*	Minimum p-value*
Isoproterenol	Fig 2 d-f	13	6.53E-08
Diazepam	Fig 2 g-i	8	2.33E-07
Apomorphine	Fig 2 j-l	12	1.88E-06
Digitoxigenin	Fig 2 m-o	6	6.03E-07
6-nitroquipazine	Fig 2 p-r	3	6.73E-07
Sodium cluster	Fig 3 b, c	11	4.71E-07
Adrenergic cluster	Fig 3 d, e	11	9.94E-14
Dopaminergic cluster	Fig 3 f-g	10	< 2.2e-16
Adenosine cluster	Fig 3 h, i	13	< 2.2e-16
STR cluster	Fig 4 a	4	0.00034

* based on Bonferroni-adjusted p-values

Nature Chemical Biology: doi:10.1038/nchembio.307

Table S2. Small molecule screening data

Category	Parameter	Description
Assay	Type of assay	In vivo zebrafish behavioral assay
	Primary measurement	Automated detection of motor responses to light stimuli by embryonic zebrafish
	Key reagents	Wild-type zebrafish embryos, Ekkwill or TuAB strains
	Assay protocol	See Methods, Automated PMR assay and measurement
	Additional comments	
Library	Library size	Natural products, drugs, rule-of-5 compliant molecules, and other diverse molecules.
	Library composition	14,000 small molecules
	Source	Biomol International, Prestwick Chemical, Microsource Discovery Systems, and Chembridge Corporation
	Additional comments	Compounds arrayed in 96-well plates as single compounds in DMSO
HTS process	Format	96-well plates, flat, clear bottoms, opaque walls (Corning)
	Concentration(s) tested	10-100 micromolar, 0.3% DMSO
	Plate controls	None. Untreated and DMSO-treated controls were tested on separate plates.
	Reagent/ compound dispensing system	Multi-channel pipette
	Detection instrument and software	Nikon TE200 microscope, Prior motorized stage, Hamamatsu ORCA-ER camera, and customized Metamorph Software. See text for details.
	Correction factors	Time scale correction. See text.
	Normalization	Log scale normalization and standardization of data were performed to reduce the impact of plate effects and other experimental artifacts. See text.
	Additional comments	
Post-HTS analysis	Hit criteria	Test compounds were compared to 1598 control wells and identified as active if at least one of the 14 behavioral barcode features exhibited a pseudo z-score absolute value greater than 5. See text.
	Hit rate	7%
	Additional assay(s)	Some assay plates performed multiple times for statistical comparison. Cell free AChE and MAO-B assays.
	Confirmation of hit purity and structure	Repurchase.
	Additional comments	

Nature Chemical Biology: doi:10.1038/nchembio.307

SUPPLEMENTARY MOVIES

Movie S1. Movie of the PMR in zebrafish embryos in a petri dish at low magnification. A single pulse of light is seen as a series of white frames at ~10 seconds after the start of the movie.

Movie S2. Movie of the PMR at higher magnification.

Movie S3. Movie of the PMR beahvior at 30 hpf, showing that animals do not normally respond to a second pulse of light.

Movie S4. Movie of the robotic screening hardware delivering light pulses to the individual wells of a 96 well plate.

Movie S5. Movie of an untreated control well in the ETR assay.

Movie S6. Movie of the slow to relax (STR) phenotype in a well treated with STR-1 during the ETR assay.

Nature Chemical Biology: doi:10.1038/nchembio.307

SUPPLEMENTARY METHODS

Chemical treatment

Chemical libraries were screened at a 1:360 dilution in E3 buffer for a final concentration between 10 and 100 μM depending on stock concentrations in the library. Library stocks were dissolved in either water or DMSO. DMSO was the solvent for greater than 90% of compounds in the libraries. Stock solutions were added directly to zebrafish in the wells of a 96 well plate, mixed, and allowed to incubate for 2-10 hours in the dark prior to behavioral evaluation in the PMR and ETR assays. The final concentration of DMSO in screening wells was less than 1%.

The light stimulus

Light stimuli lasting 1 s each were generated with a 300Watt xenon bulb housed in a Sutter Lambda LS illuminator and delivered to the well 10s and 20s after the start of each video. A cold mirror (reflectance between 300 nm and 700 nm) on the Sutter illuminator was used to block wavelengths outside of this range. Light intensity was measured using a Flashmate L-3083S light meter (Sekonic) and found to be 58,000 lux (lm/m^2).

Normalization, behavioral barcoding, and clustering

The time scale for all wells was aligned globally. Slight discrepancies in timing were reconciled by enforcing the following nominal landmarks: observation begins at 0s, the first light pulse occurs at 10s, the second light pulse occurs at 20s, and observation ends at 30s. The 24,996 behavioral profiles were derived from 437 individual plate scans. To reduce the impact of potential experimental artifacts, such as plate effects, the profiles were normalized to each other. Normalization and standardization (described below) were carried out on the log scale, due to greater symmetry, and normalized data were back-transformed, i.e. exponentiated, for

Nature Chemical Biology: doi:10.1038/nchembio.307

all figures depicting the motion index. A plate profile was computed using a loess smooth with a span of 5% with time as the explanatory variable and the combined motion index data from all wells on the plate as the response variable. A global profile was computed by averaging the plate profiles. At each time point, the difference between the plate profile and the global profile is the putative time-specific plate effect. The motion index for each well on the plate was then normalized by subtracting these plate effects.

The motion index profile for each well was partitioned into seven periods, according to the following global segmentation: PRE = 0 to 10 secs, L1 = 10 to 11 secs, E1 = 11 to 12 secs, E2 = 12 to 16 secs, E3 = 16 to 20 secs, R1 = 20 to 22 secs, R2 = 22 to 30 secs. For each time period, the motion index values for each well were summarized in two features, Q1 and Q3, which are the first and third quartiles, respectively. Among the ~25K wells, 1,598 wells were reserved as untreated controls. Features from treated wells were standardized by subtracting the median control feature and dividing by the control median absolute deviation, yielding a pseudo z-score, though we made no assumption of normality. This fourteen number summary formed the behavioral barcode.

For the purposes of phenoclustering, we targeted only those wells that were distinctly different from control, for at least one combination of time period and feature. After filtering for wells with one or more scores greater than 5 in absolute value, 1,627 behavioral profiles were retained as PMR candidate hits. Prior to clustering, we transformed pseudo z-scores with the normal cumulative probability function (mean = 0, sd = 7) to map the scores to the interval [0, 1] and to de-emphasize differences in the extreme tails. The wells were clustered using Euclidean distance and average linkage clustering. In all heatmaps, the cell representing an individual pseudo z-score is shaded according the associated cumulative normal probability, with mean = 0 and sd = 3.

All computations and figures were carried out within the R statistical programming environment [1] using the add-on packages lattice [2], RColorBrewer [3], and amap [4].

Nature Chemical Biology: doi:10.1038/nchembio.307

Statistical analysis

Analyses were performed at the level of (1) individual barcodes, (2) groups of barcodes, and (3) the entire collection of barcodes.

Individual barcodes— Among the approximately 25,000 wells, 1,598 wells were reserved as untreated controls. The statistical meaning of each individual barcode is implicit; by definition, each element of the barcode reports the difference between an observed behavioral feature under treatment and that seen in untreated controls, in standard deviation units (described fully below in methods). We defined significant barcodes as those that were distinctly different from control, for at least one combination of time period and feature, with a pseudo z-score greater than 5 in absolute value. Using this measure of significance, 1,627 behavioral profiles were retained as PMR candidate hits.

Groups of barcodes—Various groups of wells were featured in sub-analyses and/or figures as representing a coherent set of behavioral barcodes, distinguishable from those typical of untreated controls. Table S1 presents statistical evidence for these implicit claims. There were 1,598 control wells that were not treated with any small molecule. For each group of featured wells, we assessed whether the elements of the behavioral barcode were statistically distinct from these controls. This resulted in 14 p-values per comparison, due to the multivariate nature of the underlying barcode. To address the problem of multiple testing, we achieved strong control of the family-wise Type I error rate with the Bonferroni correction. For each element of the barcode, we tested for equality of distribution with the two-sided, two-sample Kolmogorov-Smirnov test, a nonparametric method with good power to detect differences in both the location and shape of distributions. Each featured group of wells is statistically distinguishable from control at a level of $\leq$ 0.00034, since each group has at least one barcode element with a p-value as small or smaller. Furthermore, the number of barcode

Nature Chemical Biology: doi:10.1038/nchembio.307

elements at which statistical distinction from control is unequivocal, ranging from a minimum of 3 (**Fig. 2 p-r**) to a maximum of 13 (**Fig. 2b-c**) out of a possible 14 (**Table S1**).

Entire dataset—To determine if the barcodes generated by the high-throughput PMR assay are reproducible, we looked at the similarity of independent wells treated with the same small molecule. We posed the null hypothesis that well-to-well differences were completely random, i.e. that the dissimilarity of a pair of wells treated with the same chemical is statistically the same as a pair of wells treated with different chemicals. **Fig S2a** shows the distribution of pairwise distances between barcodes for pairs of wells treated with the same and with different small molecules. Distances between pairs of barcodes arising from treatment with the same compound are clearly smaller than those observed among pairs treated with distinct compounds and we summarized this with the difference in average pairwise distance. **Fig S2** illustrates the statistical significance of the observed result, i.e. p-value < 9.98e-71. This was computed from the empirical distribution of the difference in average pairwise distance (same chemical – different chemical) for 100,000 random permutations of the chemical identifiers.

AChE and MAO enzymatic assays

AChE activity was assayed in vitro and in larval lystates as described [5,6]. MAO-B activity was assayed as described [7].

Supplementary references

1. Team, R. D. C. R: A Language and Environment for Statistical Computing. (2008).
2. Sarkar, D. *lattice: Lattice Graphics*. *R package version 0.17-4*. 2008).
3. Neuwirth, E. RColorBrewer: ColorBrewer palettes. R package version 1.0-2. (2007).
4. Lucas, A. *amap: Another Multidimensional Analysis Package*. *R package version 0.8-2*. 2007).
5. Behra, M. et al. Acetylcholinesterase is required for neuronal and muscular development in the zebrafish embryo. *Nat Neurosci* **5**, 111-118 (2002).
6. Behra, M., Etard, C., Cousin, X. & Strahle, U. The use of zebrafish mutants to identify secondary target effects of acetylcholine esterase inhibitors. *Toxicol Sci* **77**, 325-333 (2004).
7. Holt, A. & Palcic, M. M. A peroxidase-coupled continuous absorbance plate-reader assay for flavin monoamine oxidases, copper-containing amine oxidases and related enzymes. *Nature protocols* **1**, 2498 (2006).

Nature Chemical Biology: doi:10.1038/nchembio.307