Neuroactive compound discovery by high-content screens in zebrafish

by
Douglas Myers-Turnbull

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biological and Medical Informatics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Michael Keiser
E6422EC87FC3457...

Michael Keiser
Chair

DocuSigned by:
Jason Gestwicki
[illegible]E5...

Jason Gestwicki

DocuSigned by:
Jason Sello
2F30A400EE504B5...

Jason Sello

Committee Members

ACKNOWLEDGMENTS

I would first like to thank my advisor, Mike Keiser, as well as Dave Kokel for offering their mentorship and granting me a home for science at UCSF. I am also grateful to my committee members, Jason Gestwicki and Jason Sello, who have been exceptionally supportive.

And thanks to former Kokel Lab members – Cole Helsell, Jack Taylor, Tia Tummino, Matt McCarroll, Darya Cheng, Amanda Carbajal, Reid Kinser, Chris Ki, and Ethan Fertsch. Thanks to Cole for the early golden years of pair-programming and debugging badly documented firmware, Jack for unrelenting experimental sorcery and general cleverness, Tia for brilliance and her being Tia, and Matt for guidance beyond the call of duty.

I can't undersell the Keiser Lab's spirit, especially during the last two years, nor can I think of how to capture it in text. You would have had to be there. Thanks, Mike, and everyone else, for cultivating it. It's a big lab, but I want to be sure to mention Elena Caceres by name. I also want to thank to Leo Gendelev, Garrett Gaskins, and Wren Saylor.

Thanks also to a small army for animal husbandry and protocol development, including Louie Ramos, Julian Castaneda, Veronica Manzo, Nolan Wong, Maddie Burns, Vy Nguyen, and Alexis Marquez. Thanks as well to Paul Schnier, Shigenari Hayashi, John Kornak, and others at UCSF for their expertise and collaborations.

I next want to highlight the contributions of outstanding interns I've worked with: Ashley Oh, Rebekah Alexander, and Shreya Ravikumar. Ashley contributed a large and valuable data set; Rebekah collected high-quality data referenced in this dissertation; and Shreya integrated a computer vision model discussed in the last chapter.

I last want to thank my family and friends for their encouragement, and for their friendship.

CONTRIBUTIONS

Some material in chapters 1 and 2 has been adapted from a manuscript in submission titled 'Simultaneous analysis of behavior-modulating compounds in zebrafish with SauronX highlights avenues in neuroactive drug discovery' (doi:10.1101/2020.01.01.891432v1).

ABSTRACT

Neuroactive compound discovery by high-content screens in zebrafish

Douglas Myers-Turnbull

Neuroactive compounds are crucial tools in drug discovery and neuroscience. However, discovering mechanistically novel drugs has proven challenging. Behavioral screens in larval zebrafish have helped researchers discover compounds with novel mechanisms.

In Chapter 2, we introduce and evaluate an open platform for behavioral screening, SauronX. This instrument records movement behaviors in multiwell plates, capturing high-resolution video data at high framerate (100 Hz) under complex photic and acoustic stimuli. To test, we trained machine learning models to resolve phenotypes caused by compounds with diverse mechanisms in fully randomized screens. First, we benchmarked the system with 14 quality–control (QC) compounds and found that all 14 could be distinguished from each other and from vehicle controls. We then extended to a set of reference phenotypic readouts from 648 neuroactive compounds.

The hardware and software system has been used in studies by several research groups, so far limited to direct collaborations. In this work, we have sought to document the platform fully, providing 3D diagrams, component information, and source code. We have also deposited 7 years of phenotypic data for 3.2 million animals and 34,000 compounds. The data are curated, structured, tied to extensive metadata, and available under a permissive Creative Commons (CC-BY) license.

CONTENTS

FIGURES

TABLES

ABBREVIATIONS

AAALAC Association for Assessment and Accreditation of Laboratory Animal Care

CAD computer-aided design

CMO confusion matrix ordering

CNS central nervous system

CQP Constant Quantization Parameter

DAT dopamine transporter

DMSO dimethyl sulfoxide

FDA U.S. Food and Drug Administration

$GABA_AR$ GABA ionotropic receptor

HEVC High-Efficiency Video Encoding

IACUC Institutional Animal Care Use Committee

isoDMT N,N-Dimethylaminoisotryptamine

k-NN k-nearest neighbor

KDE kernel density estimate

LC liquid chromatography

LC–MS liquid chromatography-–mass spectrometry

LED light-emitting diode

MCX Mixed-mode Cation eXchange

MOA mechanism of action

MS mass spectrometry

nano-ESI nano-electrospray ionization

NMP N-methyl-2-pyrrolidone

NVMe Non-Volatile Memory Host Controller Interface Specification

p-ERK phospho-ERK

QC quality–control

RF Random Forests

ROI region of interest

SDK software development kit

SPL sound pressure level

SSRI selective serotonin reuptake inhibitor

SVM support vector machine

t-SNE t-distributed stochastic neighbor embedding

TRPA1 transient receptor potential channel A1

UCSF University of California, San Francisco

USAMRICD U.S. Army Medical Research Institute of Chemical Defense

XIC extracted ion chromatogram

YOLOv5 You Only Look Once v5

CHAPTER 1

Introduction

1.1 Phenotypic screens for CNS drug discovery

Disorders of the central nervous system (CNS) affect 100 million Americans at an economic burden of $920 billion per year [1]. Despite this, CNS drug discovery rates have declined [2]. Most projects screen for high-affinity interaction with one target [3]. Although extremely high-throughput, these screens require prior knowledge of the disease-linked targets, which is especially limited for CNS disorders [4, 5]. Although most projects are target-first, most first-in-class drugs approved by the U.S. Food and Drug Administration (FDA) from 1999–2008 were discovered phenotype-first [6], suggesting that many CNS drug discovery projects would benefit from phenotype-first screens.

In contrast to target-based screens, phenotypic screens require less understanding of pathogenesis and can identify compounds with previously unknown and multitarget pharmacological actions. In many historical cases, a drug was discovered first, and its mechanism only later [7, 8]. For example, the antidepressant activities of tricyclics and monoamine oxidase inhibitors were discovered in psychiatric hospitals by observing patients. These discoveries implicated serotonin in depression and lent to the development of selective serotonin reuptake inhibitors (SSRIs) [9]. Such phenomenological discoveries are responsible for most prototypical neuroactive drugs. Scaling phenomenological discovery to higher-throughput using animal models is yielding a powerful approach to CNS drug discovery.

1.2 Phenotypic profiling in zebrafish

Zebrafish larvae and embryos have long been used to assay environmental toxicants [10, 11] and have made waves in neuroscience as models for vision [12, 13, 14, 15, 16], threat response [17], memory [18], algesia [19, 20, 21], and sleep [22, 23, 24]. These successes in the laboratory have extended to the clinic: In a rare example of bench-to-bedside, the FDA approved lorcaserin as an antiepileptic, based significantly on evidence in zebrafish [25]. More recently, a zebrafish model was used in the life-saving treatment of a 12-year-old patient [26]. Genetic and compound-induced disease models in zebrafish larvae have shown promising consistency with rodent models [27, 28].

Zebrafish are well-suited for *phenotypic profiling*, a quantitative, high-throughput approach to phenotype-first compound discovery [29, 23]. Profiles are quantitative readouts of aggregate animal movements in multiwell plates. These experiments often employ acoustic, photic (light-based), and other stimuli to perturb the animals' behavior in an effort to reveal more compound-induced behavioral changes. Previous screens identified new neuroactive compounds and predicted their targets, later supported by in vitro assays [30, 31, 23, 32, 33]. Diverse compounds have been identified, including photoactivatable transient receptor potential channel A1 (TRPA1) ligands [34], antiepileptics [35], antipsychotics [36], appetite modulators [37], and anesthetic-like compounds [38, 39].

One way to predict the pharmacology of a mechanistically novel compound is by association to a compound of known pharmacology. This *guilt-by-association* approach links novel compounds to known ligands, but it requires both *reference profiles* for compounds with known pharmacology and a way to measure similarity between profiles.

The following chapter describes a phenotypic profiling hardware/software system, benchmarks it for three criteria using quality–control data, and applies it on a set of 648 recognized CNS-target ligands to build a preliminary set of reference profiles for guilt-by-association

discovery campaigns.

References

[1] Chris Delvecchio, Jens Tiefenbach, and Henry M Krause. "The zebrafish: a powerful platform for in vivo, HTS drug discovery." In: *Assay Drug Dev. Technol.* 9.4 (Aug. 2011), pp. 354–361. ISSN: 1540-658X. DOI: 10.1089/adt.2010.0346.

[2] Aaron S Kesselheim, Thomas J Hwang, and Jessica M Franklin. "Two decades of new drug development for central nervous system disorders." In: *Nat. Rev. Drug Discov.* 14.12 (2015), pp. 815–816. ISSN: 1474-1776. DOI: 10.1038/nrd4793.

[3] Mark A Lindsay. "Target discovery." In: *Nat. Rev. Drug Discov.* 2.10 (Oct. 2003), pp. 831–838. ISSN: 1474-1776. DOI: 10.1038/nrd1202.

[4] Menelas N Pangalos, Lee E Schechter, and Orest Hurko. "Drug development for CNS disorders: strategies for balancing risk and reducing attrition." In: *Nat. Rev. Drug Discov.* 6.7 (July 2007), pp. 521–532. ISSN: 1474-1776. DOI: 10.1038/nrd2094.

[5] Yves Agid et al. "How can drug discovery for psychiatric disorders be improved?" In: *Nat. Rev. Drug Discov.* 6.3 (Mar. 2007), pp. 189–201. ISSN: 1474-1776. DOI: 10.1038/nrd2217.

[6] David C Swinney. "Phenotypic vs. Target-Based Drug Discovery for First-in-Class Medicines." In: *Clinical Pharmacology & Therapeutics* 93.4 (2013), pp. 299–301. ISSN: 1532-6535. DOI: 10.1038/clpt.2012.236.

[7] Alan Wayne Jones. "Early drug discovery and the rise of pharmaceutical chemistry." In: *Drug Test. Anal.* 3.6 (June 2011), pp. 337–344. ISSN: 1942-7603, 1942-7611. DOI: 10.1002/dta.301.

[8] Jürgen Drews. "Drug discovery: a historical perspective." In: *Science* 287.5460 (Mar. 2000), pp. 1960–1964. ISSN: 0036-8075. DOI: 10.1126/science.287.5460.1960.

[9] Chaitra T Ramachandraih et al. “Antidepressants: From MAOIs to SSRIs and more.” In: *Indian J. Psychiatry* 53.2 (Apr. 2011), pp. 180–182. ISSN: 0019-5545, 1998-3794. DOI: 10.4103/0019-5545.82567.

[10] Sudhakar Deeti, Sean O’Farrell, and Breandán N Kennedy. “Early safety assessment of human oculotoxic drugs using the zebrafish visualmotor response.” In: *J. Pharmacol. Toxicol. Methods* 69.1 (Jan. 2014), pp. 1–8. ISSN: 1056-8719. DOI: 10.1016/j.vascn.2013.09.002.

[11] Anjali K Nath et al. “Chemical and metabolomic screens identify novel biomarkers and antidotes for cyanide exposure.” In: *FASEB J.* 27.5 (May 2013), pp. 1928–1938. ISSN: 0892-6638. DOI: 10.1096/fj.12-225037.

[12] Florian A Dehmelt et al. “Spherical arena reveals optokinetic response tuning to stimulus location, size and frequency across entire visual field of larval zebrafish.” In: *bioRxiv* 1 (Sept. 2019). DOI: 10.1101/754408.

[13] Owen Randlett et al. “Distributed Plasticity Drives Visual Habituation Learning in Larval Zebrafish.” In: *Curr. Biol.* 29.8 (Apr. 2019), 1337–1345.e4. ISSN: 0960-9822, 1879-0445. DOI: 10.1016/j.cub.2019.02.039.

[14] Andreas M Kist and Ruben Portugues. “Optomotor Swimming in Larval Zebrafish Is Driven by Global Whole-Field Visual Motion and Local Light-Dark Transitions.” In: *Cell Rep.* 29.3 (Oct. 2019), 659–670.e3. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2019.09.024.

[15] Rebecca Ward et al. “Pharmacological restoration of visual function in a zebrafish model of von-Hippel Lindau disease.” In: *Dev. Biol.* 457 (Feb. 2019). ISSN: 0012-1606, 1095-564X. DOI: 10.1016/j.ydbio.2019.02.008.

[16] Conor Daly et al. “A Brain-Derived Neurotrophic Factor Mimetic Is Sufficient to Restore Cone Photoreceptor Visual Function in an Inherited Blindness Model.” In: *Sci. Rep.* 7.1 (Sept. 2017), p. 11320. ISSN: 2045-2322. DOI: 10.1038/s41598-017-11513-5.

[17] Andrew J Rennekamp et al. " 1 receptor ligands control a switch between passive and active threat responses." In: *Nat. Chem. Biol.* 12.7 (July 2016), pp. 552–558. ISSN: 1552-4450. DOI: 10.1038/nchembio.2089.

[18] Marc A Wolman et al. "Chemical modulation of memory formation in larval zebrafish." In: *PNAS* 108.37 (Sept. 2011), pp. 15468–15473. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1107156108.

[19] Peter J Steenbergen and Nabila Bardine. "Antinociceptive effects of buprenorphine in zebrafish larvae: An alternative for rodent models to study pain and nociception?" In: *Appl. Anim. Behav. Sci.* 152.Supplement C (Mar. 2014), pp. 92–99. ISSN: 0168-1591. DOI: 10.1016/j.applanim.2013.12.001.

[20] Andrew Curtright et al. "Modeling nociception in zebrafish: a way forward for unbiased analgesic discovery." In: *PLOS One* 10.1 (Jan. 2015), e0116766. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0116766.

[21] Martin Haesemeyer et al. "A Brain-wide Circuit Model of Heat-Evoked Swimming Behavior in Larval Zebrafish." In: *Neuron* 98.4 (May 2018), 817–831.e6. ISSN: 0896-6273, 1097-4199. DOI: 10.1016/j.neuron.2018.04.013.

[22] David A Prober et al. "Hypocretin/orexin overexpression induces an insomnia-like phenotype in zebrafish." In: *J. Neurosci.* 26.51 (Dec. 2006), pp. 13400–13410. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.4332-06.2006.

[23] Jason Rihel et al. "Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation." In: *Science* 327.5963 (Jan. 2010), pp. 348–351. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1183090.

[24] Chanpreet Singh, Grigorios Oikonomou, and David A Prober. "Norepinephrine is required to promote wakefulness and for hypocretin-induced arousal in zebrafish." In: *Elife* 4.September (2015), pp. 1–22. ISSN: 2050-084X. DOI: 10.7554/eLife.07000.001.

[25] Aliesha Griffin et al. “Clemizole and modulators of serotonin signalling suppress seizures in Dravet syndrome.” In: *Brain* 140.3 (Mar. 2017), pp. 669–683. ISSN: 0006-8950. DOI: 10.1093/brain/aww342.

[26] Dong Li et al. “ARAF recurrent mutation causes central conducting lymphatic anomaly treatable with a MEK inhibitor.” In: *Nat. Med.* 25 (July 2019). ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-019-0479-2.

[27] Matthew P Shaw et al. “Stable transgenic C9orf72 zebrafish model key aspects of the ALS/FTD phenotype and reveal novel pathological features.” In: *Acta Neuropathol Commun* 6.1 (Nov. 2018), p. 125. ISSN: 2051-5960. DOI: 10.1186/s40478-018-0629-7.

[28] Jessica R Morrice, Cheryl Y Gregory-Evans, and Christopher A Shaw. “Modeling Environmentally-Induced Motor Neuron Degeneration in Zebrafish.” In: *Sci. Rep.* 8.1 (Mar. 2018), p. 4890. ISSN: 2045-2322. DOI: 10.1038/s41598-018-23018-w.

[29] Randall T Peterson and Mark C Fishman. “Discovery and use of small molecules for probing biological processes in zebrafish.” In: *Methods Cell Biol.* 76 (2004), pp. 569–591. ISSN: 0091-679X. DOI: 10.1016/s0091-679x(04)76026-4.

[30] David Kokel et al. “Rapid behavior-based identification of neuroactive small molecules in the zebrafish.” In: *Nat. Chem. Biol.* 6.3 (2010), pp. 231–237. ISSN: 1552-4450, 1552-4469. DOI: 10.1038/nchembio.307.

[31] Christian Laggner et al. “Chemical informatics and target identification in a zebrafish phenotypic screen.” In: *Nat. Chem. Biol.* 8.2 (Dec. 2011), pp. 144–146. ISSN: 1552-4450, 1552-4469. DOI: 10.1038/nchembio.732.

[32] Nathalie Jeanray et al. “Phenotype classification of zebrafish embryos by supervised learning.” In: *PLOS One* 10.1 (Jan. 2015), e0116989. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0116989.

[33] Olivier Mirat et al. “ZebraZoom: an automated program for high-throughput behavioral analysis and categorization.” In: *Front. Neural Circuits* 7 (2013), p. 107. DOI: 10.3389/fncir.2013.00107.

[34] Darya Cheng et al. “Identification of compounds producing non-visual photosensation via TRPA1 in zebrafish.” In: *BioRxiv* 1 (June 2020). DOI: 10.1101/2020.06.10.111203. URL: https://www.biorxiv.org/content/10.1101/2020.06.10.111203v1.

[35] Scott C Baraban, Matthew T Dinday, and Gabriela A Hortopan. “Drug screening in Scn1a zebrafish mutant identifies clemizole as a potential Dravet syndrome treatment.” In: *Nat. Commun.* 4.2410 (2013), p. 2410. DOI: 10.1038/ncomms3410.

[36] Giancarlo Bruni et al. “Zebrafish behavioral profiling identifies multitarget antipsychotic-like compounds.” In: *Nat. Chem. Biol.* 12 (May 2016), pp. 1–25. ISSN: 1552-4450. DOI: 10.1038/nchembio.2097.

[37] Josua Jordi et al. “High-throughput screening for selective appetite modulators: A multibehavioral and translational drug discovery strategy.” In: *Sci Adv* 4.10 (Oct. 2018). ISSN: 2375-2548. DOI: 10.1126/sciadv.aav1966.

[38] Matthew N McCarroll et al. “Zebrafish behavioural profiling identifies GABA and serotonin receptor ligands related to sedation and paradoxical excitation.” In: *Nat. Commun.* 10.1 (Sept. 2019), p. 4078. ISSN: 2041-1723. DOI: 10.1038/s41467-019-11936-w.

[39] Xiaoxuan Yang et al. “High-throughput Screening in Larval Zebrafish Identifies Novel Potent Sedative-hypnotics.” In: *Anesthesiology* 129.3 (Sept. 2018), pp. 459–476. ISSN: 0003-3022, 1528-1175. DOI: 10.1097/ALN.0000000000002281.

CHAPTER 2

Phenotypic screening with SauronX

2.1 Preface

This chapter has been adapted from the following manuscript. References to supplemental items have been omitted from this dissertation but are available under the same DOI. All material is open access.

'Simultaneous analysis of behavior-modulating compounds in zebrafish with SauronX highlights avenues in neuroactive drug discovery." (doi:10.1101/2020.01.01.891432v1) The authors are Douglas Myers-Turnbull, Jack C Taylor, Cole Helsell, Matthew N McCarroll, Chris S Ki,Tia A Tummino,Shreya Ravikumar,Reid Kinser,Leo Gendelev, Rebekah Alexander, Michael J Keiser, and David Kokel.

2.2 Results

Here, we describe our platform in detail. We then benchmark our system in a machine learning approach on 14 quality–control (QC) compounds, and then establish preliminary reference phenotypic profiles from a chemical library of 648 known-neuroactive compounds.

2.2.1 An open phenotypic profiling platform

In our work to develop an open platform for behavioral profiling, we identified three requirements: the ability to screen without interruption, reproducibility of analyses, and extensibility to add or remove hardware. We modified an existing system [1]. The new hardware/software platform has been used to assay N,N-Dimethylaminoisotryptamine (isoDMT) analogs [2], a

non-hallucinogenic ibogaine analog with therapeutic potential [3], and toxicants at the U.S. Army Medical Research Institute of Chemical Defense (USAMRICD).

The setup is shown in Figure 2.1A on the following page. Plates are positioned on a flat translucent stage, fixed in a groove so that sound propagated through the stage contacts the plate uniformly. The plates are illuminated from the bottom with infrared light through an acrylic diffuser and recorded with an overhead camera while light and sound stimuli are applied. The digital camera is mounted to a telecentric lens with an infrared pass filter so that the photic stimuli do not affect the video. The lens eliminates parallax, resulting in all wells having the same apparent dimensions, simplifying feature calculations and eliminating parallax corrections as potential confounding variables. The camera captures 1 Mpx to 6 Mpx (16-bit depth) images at a preset frame rate of 100 Hz to 150 Hz. Nanosecond-resolved timestamps corresponding to the image sensor acquisition are used to precisely synchronize captured frames with stimuli. Computer-aided design (CAD) files and related information are available in the Data Repository [4].

To expand the repertoire of observable behavioral responses, stimuli are applied during capture. Photic stimuli are delivered via 6 overhead LED arrays. Acoustic stimuli from audio files are delivered through surface transducers mounted on the stage. A microphone, photosensor, and secondary camera verify the delivery and timing of stimuli. These stimuli evoke compound-dependent behaviors that would not otherwise be observed. For example, we observed a compound-modulated 'step' response to 355 nm ultraviolet light, which is visible to zebrafish [5]. This response differs markedly from 400 nm light.

We use a 4-step workflow (Figure 2.1B on the next page). Animals are anesthetized in cold water and dispensed into the wells of a multiwell plate, dosed, and incubated for 1 hr. The plates are placed in the instrument, and the animals are acclimated in darkness for 5 minutes. A battery of stimuli is then applied while video is recorded. Videos can be analyzed in many ways, including tracking of animals. For the experiments in this manuscript, we used

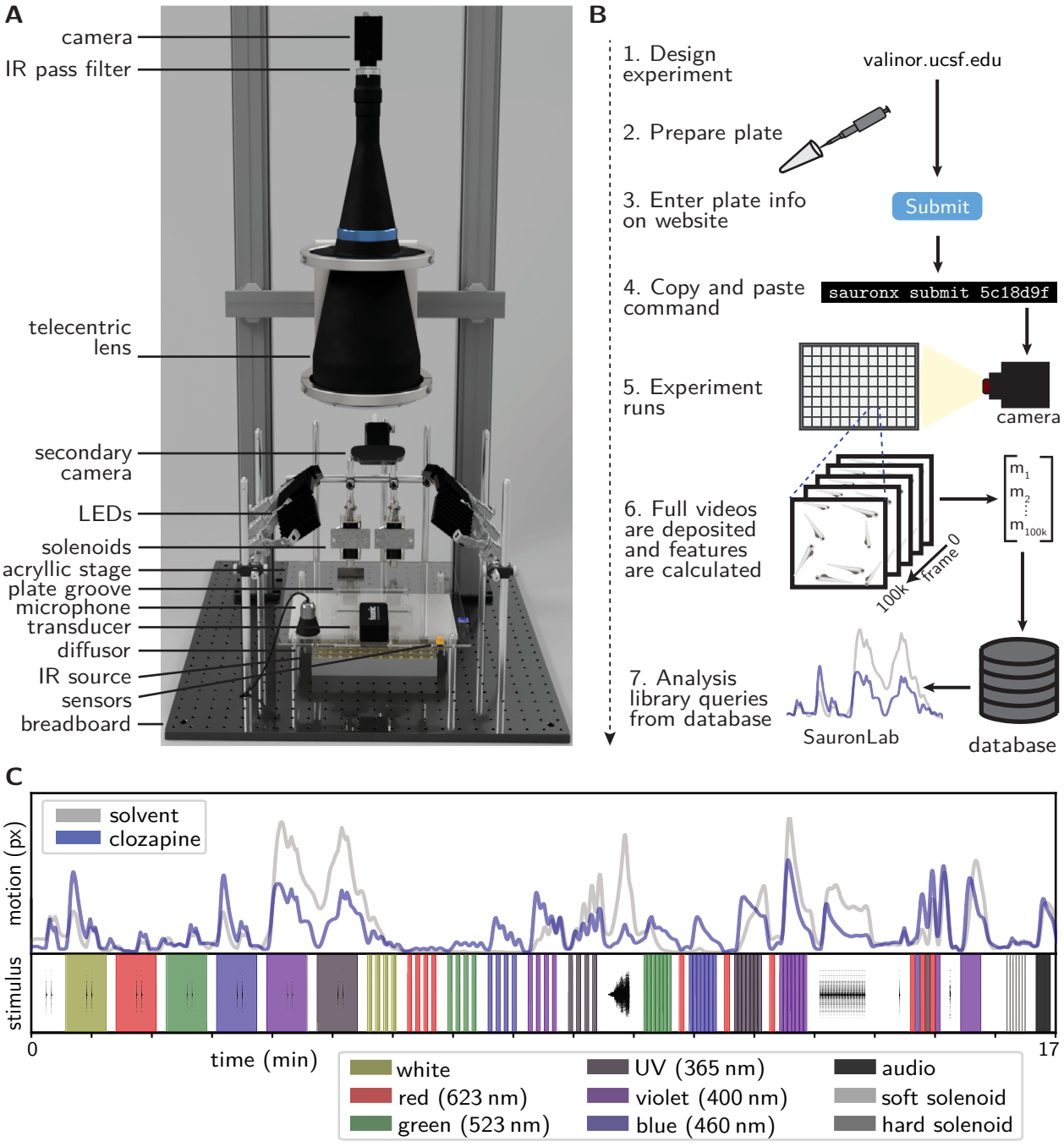

Figure 2.1: Overview of hardware/software platform. (**A**) Front view of the instrument. (**B**) Stages of the experimental/computational pipeline. (**C**) Example motion-trace for wells treated with vehicle (dimethyl sulfoxide (DMSO)) or clozapine at $50\,\mu\mathrm{M}$. Top: motion within the well as a function of time. Bottom: stimuli applied over time. The shaded colors represent application of high-intensity light, the black lines depict the waveforms of audio assays, and the gray vertical lines (at the end) denote the application of acoustic stimuli by solenoids. N = 12 wells / condition.

multiple (8) animals per well and calculated a simple feature (*motion-trace*) of aggregate locomotor activity over time (Equation 2.1 on page 26). Although using multiple animals per well complicates per-animal tracking, it resulted in much higher algorithm performance (discussed later). Figure 2.1C on the previous page shows example traces under a standard battery for vehicle (DMSO solvent) or the antipsychotic clozapine.

A website is used to design plate layouts, stimulus batteries, and experiments, as well as to organize and search for genetic constructs and compound stocks. A custom language called Gale can be used to design assays from simple expressions (but is not required). The hardware is driven by custom software. Post-processing of data is not coupled to capture, allowing many plates to be run in sequence without interruption. After a run completes, the videos are compressed and archived permanently, and data is inserted into a relational database on a remote server.

The database incorporates coarse-grained and fine-grained data. The coarse-grained data, such as hierarchical grouping of experiments, simplifies search. The fine-grained data is included for reproducibility and post-hoc diagnostics. For example, compound treatments are indicated by 'batch', with supplier information and lot numbers. In developing the system, we identified information required to conduct reproducible, audit-able analyses. In accordance, we propose a minimum information standard [6] at https://osf.io/nyhpc/.

These data are used in an open source analysis platform (sauronlab), which provides tools for search and analysis. Analyses include quantifying the strength of phenotypes, classifying and clustering phenotypes, analyzing mechanism of actions (MOAs), and searching for similar phenotypes. All analyses are tied to a timestamp that restricts the data queried from the database, ensuring that results do not change when new data is added. Notebooks illustrating these analyses with code and output are available at sauronlab-publication.

In contrast to commercial phenotyping systems, the hardware, data storage, and analysis are uncoupled. Videos are efficiently compressed and can be stored indefinitely and analyzed

with additional methods at any point. Although the hardware is larger than most commercial systems at 61 cm×61 cm×114 cm, this simplified construction and enabled rapid iteration between analyzing data and adapting hardware. As part of an effort to develop an open alternative to commercial systems, we benchmarked the platform's ability to distinguish compound-induced phenotypes.

2.2.2 Discrimination of phenotypes for quality–control compounds

We wanted to evaluate the platform in a way that is not constrained to a single phenotype. Specifically, we sought to test the ability to detect compound-induced phenotypes (*detection criterion*), identify phenotypes caused by the same compounds while distinguishing those caused by different compounds (*distinction criterion*), and group compounds with similar mechanisms or effects (*grouping criterion*).

First, we curated a set of 14 compounds with diverse structures and MOAs. The lethal control used a high dose of the anesthetic eugenol, which is routinely used as a humane method to euthanize fish [7, 8]. These 14 compounds and 2 controls formed the QC set. Experiments were run using 7-day-old wild-type zebrafish 1 h post-treatment under a standard battery (Figure 2.1C on page 10).

For each compound, we selected a 5-point logarithmic concentration gradient to capture the range between phenotypic inactivity and lethality. 8 replicate plates were screened, with all compounds and concentrations on each plate in random positions. For each compound and concentration, a binary treatment–vehicle Random Forests (RF) classifier was trained to assign the motion vectors as either *treatment* or *vehicle*. The same procedure was used for treatment–lethal models. We plotted the resulting out-of-bag accuracy values in concentration–response curves. Due to the high dimensionality, such curves are not expected to be sigmoidal or even monotone increasing. For most compounds, treatment–vehicle accuracy increased with concentration, while treatment–lethal accuracy dropped sharply at high concentrations

Table 2.1: Quality–control compounds with optimal concentrations.

Compound	Concentration (µM)	Primary mechanism of action
almorexant	90	OX_1, OX_2 antagonist
bromocriptine	16	D_2, D_3 agonist
clozapine	50	D_2, 5-HT_{2A} antagonist
donepezil	16	AChE inhibitor
endosulfan	0.32	$GABA_AR$ antagonist
etomidate	6.25	$GABA_AR$ agonist
haloperidol	25	D_2 antagonist
indoxacarb	6.25	Na_V inhibitor
(S)+ketamine	100	NMDAR antagonist
lidocaine	1200	Na_V inhibitor
optovin	6.25	TRPA1 opener
(+)-sertraline	25	SERT inhibitor
tiagabine	100	GAT inhibitor
tracazolate	25	$GABA_AR$ modulator

orexin receptor; $D_1/D_2/D_3$ (dopamine receptors 1/2/3); 5-HT_1/5-HT_2 (serotonin receptors); acetylcholinesterase; $GABA_AR$ (GABA ionotropic receptor); voltage-gated sodium channel; N-Methyl-D-aspartate receptor; TRPA1 (transient receptor potential channel A1); serotonin transporter; GABA transporter.

(Figure 2.2A on the next page). Notably, treatment–lethal accuracy was high even for sedating doses of the anesthetic etomidate [9], indicating that sedation and lethality were distinguished.

Using these data, we set an 'optimal' concentration per compound by balancing phenotypic strength with non-lethality (Table 2.1). This *optimal-concentration* set was screened in 15 replicate plates, with 6 replicates of each compound per plate. In compound–vehicle models, the mean accuracy was 93%. In contrast, randomly false-labeling controls yielded 49% for vehicle–vehicle comparisons. This established that compounds could be separated from controls, meeting the detection criterion.

The phenotypes were then visualized together using t-distributed stochastic neighbor embedding (t-SNE) [10]. Each compound generated a cloud of replicate profiles generally separate from the controls and other compounds (Figure 2.3 on the next page), indicating an ability to identify phenotypes caused by the same compounds and distinguish them from others. A RF multiclass classifier was trained to quantify this.

The out-of-bag predictions were visualized in a confusion matrix (Figure 2.4A on page 15).

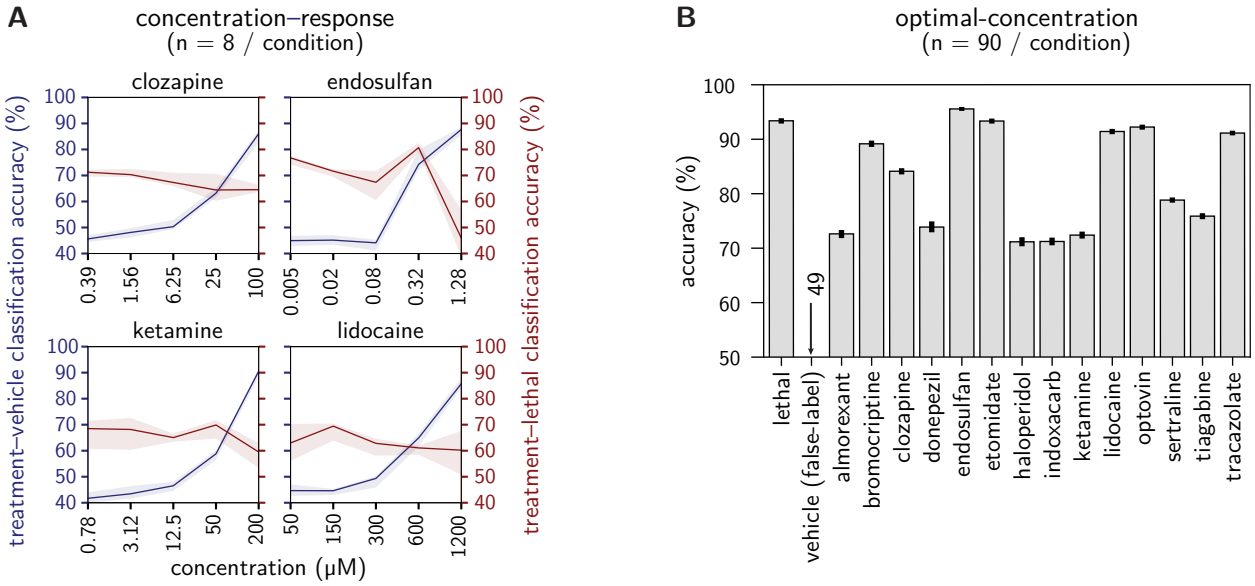

Figure 2.2: Results for the QC experiments. (**A**) Concentration–response curves for treatment–vehicle (left axis; blue) and treatment–lethal (right axis; red) accuracy. Opaque lines denote the median accuracy, and shaded regions denote a 95th percentile confidence interval by bootstrap. N = 8 wells/condition. (**B**) Treatment–solvent classification accuracy by compound on the optimal-concentration QC set. Error bars signify a 90th-percentile confidence interval computed by repeat training on subsampled wells (n = 27+27=54 wells/model). N = 90 wells/condition.

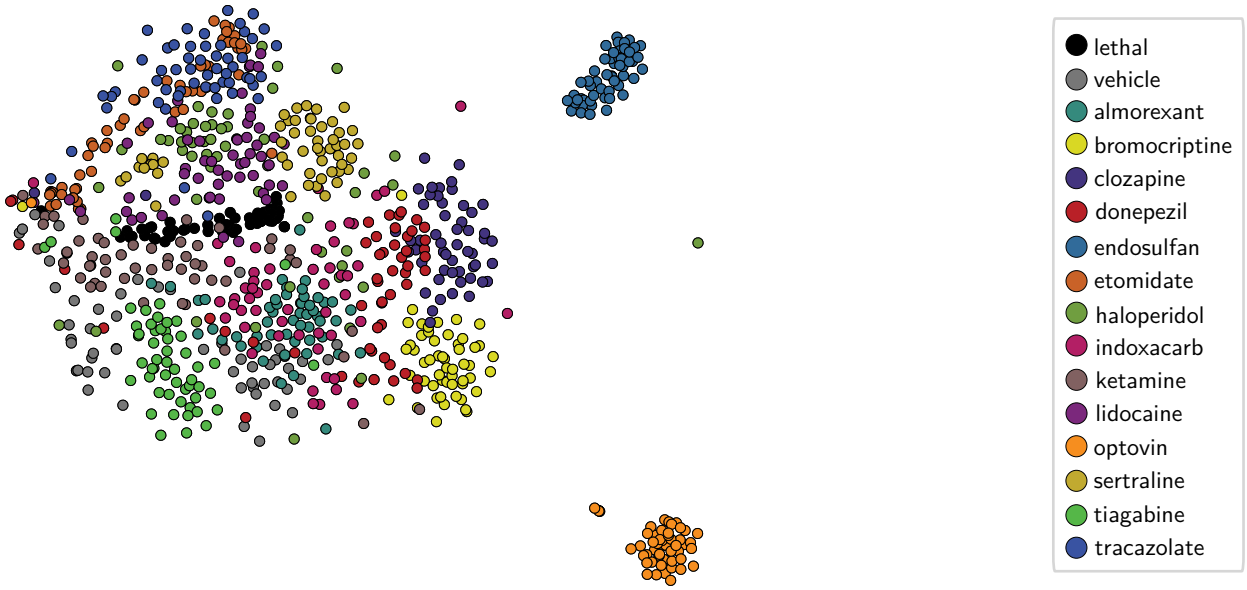

Figure 2.3: T-SNE projection of motion vectors in the optimal-concentration QC set. Each point denotes one well. N = 90 wells/condition.

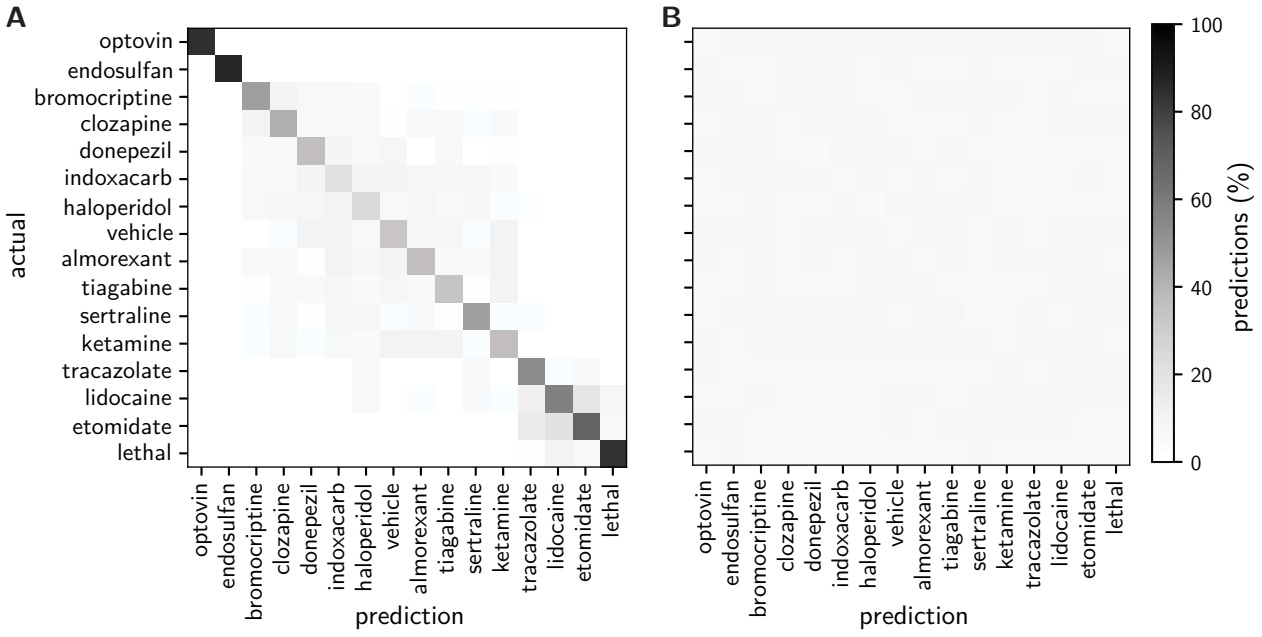

Figure 2.4: Confusion matrix of QC compound phenotypes. (**A**) Confusion matrix from a multiclass classification model (Random Forests) on the optimal-concentration QC set. N = 90 wells/condition. (**B**) Confusion matrix from a corresponding model trained on false-labeled vehicle-treated wells. N = 18 wells/condition.

The labels were sorted by an algorithm that maximized block–diagonal structures, grouping like phenotypes. The diagonal was high (mean=94%), reflecting accurate self-classification and phenotypic uniqueness. The classifier distinguished several compounds, such as almorexant and tiagabine, that were poorly separated in Figure 2.3 on the previous page. As an adversarial experiment, we collected a dataset of only vehicle-treated wells and false-labeled them to mimic the real dataset. Classifiers were unable to distinguish the false-labeled treatments (Figure 2.4B), supporting the distinction criterion.

Grouping of compounds (grouping criterion) was harder to assess with few compounds. Although generally distinguishable, lidocaine, etomidate, and tracazolate were sorted nearby. These compounds reduced movement, analogous to their effects in humans, but they evoked noticeably distinct responses to stimuli. This offered anecdotal but encouraging support for grouping.

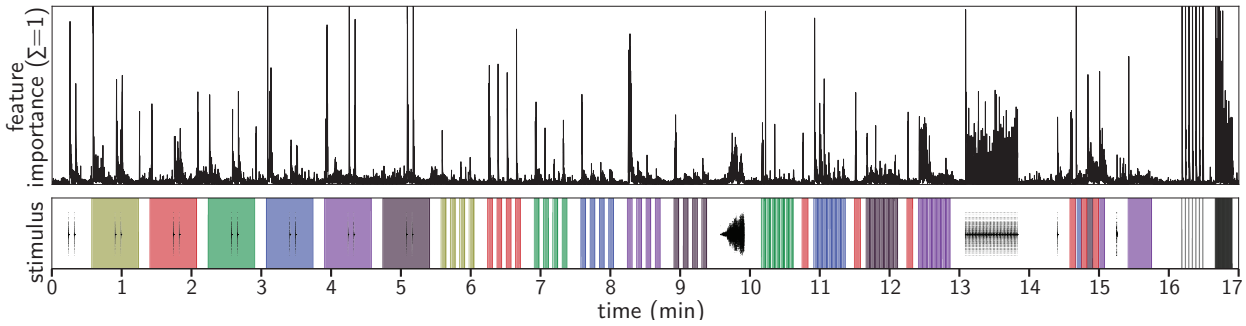

Figure 2.5: Top: Frame-by-frame feature weights the optimal-concentration multiclass model (values sum to 1). Some extreme values were cropped for visual clarity. Bottom: Stimulus battery as in Figure 2.1C on page 10. N = 90 wells/condition.

2.2.3 Data-driven protocol optimization

We hypothesized that this approach of classification on a QC set served as a general evaluation method to guide experimental design. We applied it to design a stimulus battery, optimize experimental and computational methods, and quantify the impact of potentially confounding variables.

In a data-driven approach to design a battery, we compared 53 30 s to 60 s behavioral assay. Assays that provided high classification accuracy were included in the final battery. Background (stimulus-free) assays had notably low performance and most of the acoustic assays with pure tones yielded little information. While pure tones are commonly used [11], acoustic assays generated from complex environmental sounds resulted in higher accuracy. Assays with high-frequency light stimuli and those with simultaneous photic and acoustic stimuli also performed well. Likewise, the optimal-concentration classifier (Figure 2.4A on the previous page), heavily weighted frames occurred near stimuli (Figure 2.5), directly highlighting their importance. Although these experiments were based on few compounds, this data-driven approach eliminated assays that provided low phenotypic information and suggested that complex assays may be more useful for resolving compound-induced phenotypes.

Next, we evaluated how performance changed under different experimental conditions. Using a separate data set, we evaluated using 3, 4, 6, 8, and 10 animals per well. Accuracy

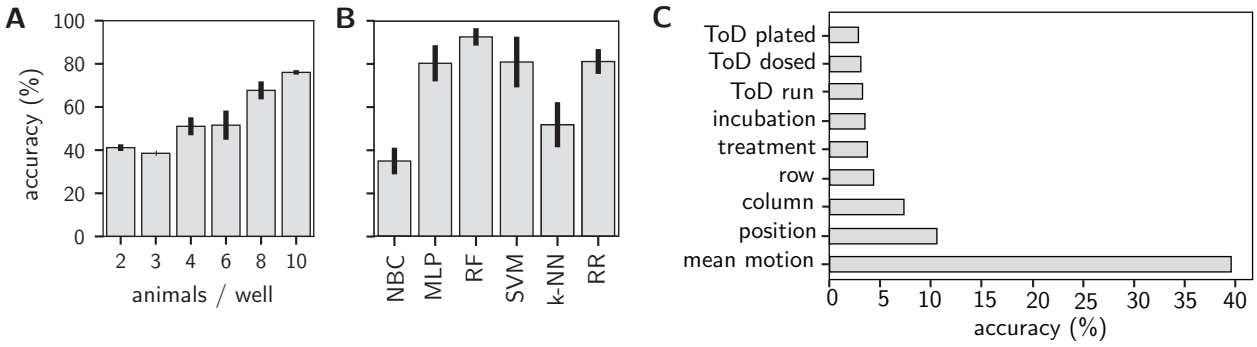

Figure 2.6: Protocol optimization and confounding analysis. (**A**) Treatment–treatment accuracy (Random Forests) by number of animals per well in a dedicated experiment. Error bars denote an 80% confidence interval by bootstrap over wells. N = 12 wells/condition. (**B**) Treatment–treatment accuracy as evaluated by different classification algorithms in the optimal-concentration QC set. The extents of the error bars mark the values for the two individual plates. N = 90 wells/condition. (**C**) Multiclass classification of quality–control compounds from metadata variables and average motion.

increased with the number of animals (Figure 2.6A). These two results showed a trade-off between higher performance and the logistics and ethics of using more animals.

Similar to this optimization of experimental protocols, computational methods could be benchmarked. We benchmarked several classification models, testing across hyperparameter sets (Figure 2.6B). Neural networks, random forests, and support vector machines (SVMs) outperformed simple models like linear classifiers and k-nearest neighbors (k-NNs), suggesting that powerful models were needed.

Finally, we considered the impact of potentially confounding variables, such as time of day (hour:minute) and exact duration of compound treatment (deviation from 1 h). None of the variables we tested were predictive (Figure 2.6C. In contrast, the arrangement of compounds within a plate significantly affected phenotypic readouts, demonstrating a critical need to control for positional confounding. These experiments pointed to a general procedure to compare and optimize protocols.

2.2.4 A reference set of CNS-target ligands

To predict mechanisms for novel compounds, we sought a set of reference profiles from compounds with diverse pharmacological actions. We used the SCREEN-WELL Neurotransmitter

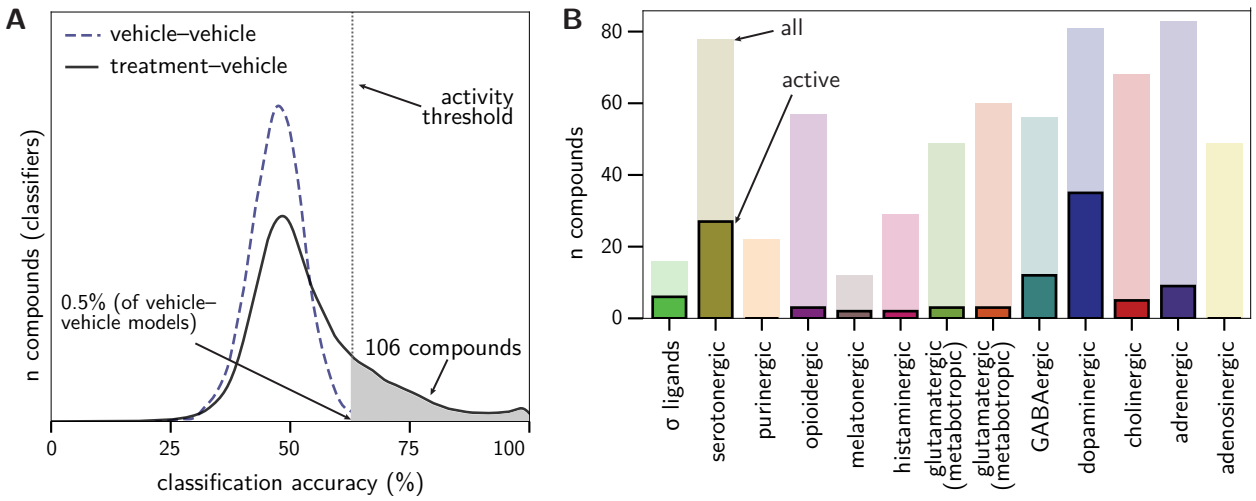

Figure 2.7: NT-650 hit-calling. (**A**) Distribution of accuracy for treatment–vehicle (black, solid) and vehicle–vehicle (blue, dashed). The hit threshold is shown at x=63%. N = 648 non-control wells for treatment–vehicle. (N = 200 classifiers for vehicle–vehicle.) (**B**) Distribution of hits (opaque) and total compounds (translucent) per major neurotransmitter system. N = 104 compounds.

library (*NT-650*, Enzo Life Sciences), which contained 648 CNS ligands. A fully randomized screen was performed, generating approximately 7 replicates per compound at 33 µM in a 648-compound reference set.

To identify hit compounds, per-compound treatment–vehicle and treatment–lethal classifiers were trained as per the QC set. Visualizing the accuracy values, vehicle–vehicle values were centered near 50% (Figure 2.7A) The treatment–vehicle values were long-tailed, indicating that compounds falling within this tail were likely active. To select phenotypically active compounds (hits), we applied an accuracy threshold that excluded 99.5% of vehicle–vehicle, yielding 106 nonlethal hit compounds. Only 1 compound, tetrahydrodeoxycorticosterone, was lethal at the concentration tested, though other compounds may have been toxic but nonlethal.

To know which types of compounds were more phenotypically active, we grouped compounds by the neurotransmitter systems that they primarily target (Figure 2.7B). Dopaminergic and serotonergic systems were enriched for phenotypic activity, while adenosinergic, purinergic, and glutamatergic were depleted, though all 13 had at least one hit. Aside from

the GABA ionotropic receptor ($GABA_AR$) , enriched targets were mostly monoaminergic, including monoamine transporters and dopamine, serotonin, histamine, and muscarinic receptors.

Multiclass models were then trained to distinguish between the 104 treatments, along with controls. The results were visualized in a sorted confusion matrix (Figure 2.4A on page 15). A strong diagonal indicated that the compounds were phenotypically coherent. Several clusters were observed, including $GABA_AR$, dopamine transporter (DAT), and dopamine, glutamate, and melatonin receptor ligands. These data indicate that the behavioral screening paradigm is capable of distinguishing neuroactive molecules that interact with discrete neurotransmitter systems. Importantly, multiple chemical scaffolds were present per cluster, indicating scaffold hopping [12] and illustrating the potential for this approach to be used to discover structural starting points for new drugs.

2.3 Discussion

Here we presented an open platform for behavioral phenotyping in zebrafish and posted complete specifications. To facilitate data mining, we deposited phenotypic data for 34,000 compounds and 3.2 million animals.

Prior studies validate behavioral profiling as a way to discover and characterize neuroactive compounds. Different hardware, zebrafish strains, and computational methods have been applied, and this diversity calls for quantitative evaluations. We found that classification in a quality–control set provided an intuitive and powerful metric to summarize performance. This approach has immediate applications, such as optimizing protocols and assessing the impact of confounding variables. In particular, positional confounding can significantly affect results, supporting a need for treatment randomization. The SCREEN-WELL Neurotransmitter library provides compounds physically arranged by their major pathways, illustrating how this confounding could solely explain a promising result. We provide a lower bound on performance

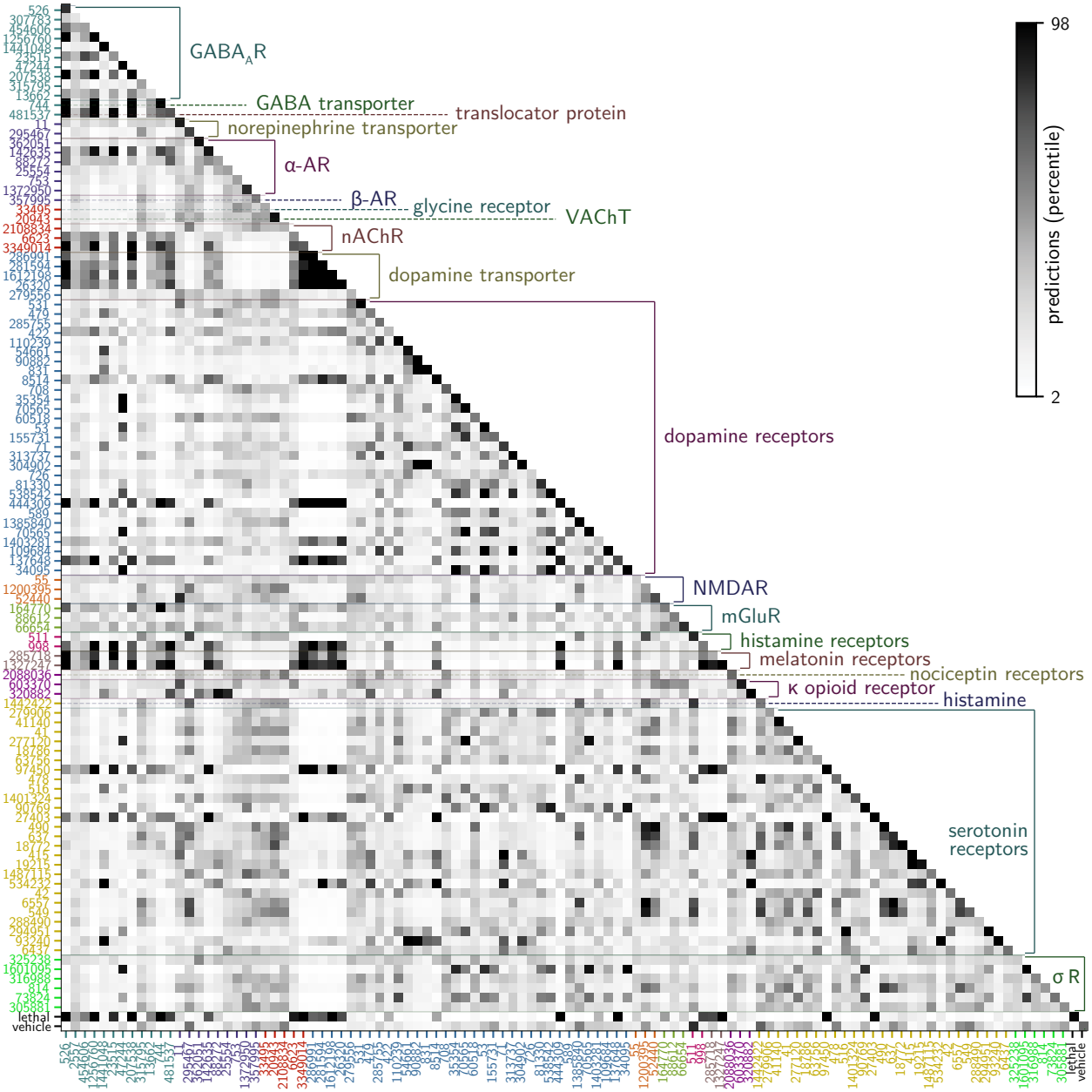

Figure 2.8: Confusion matrix of the 106 NT-650 hit compounds, plus vehicle and lethal controls. Sorting by MOA targets as provided by Enzo. Range from the 2nd percentile to 98th percentile. Labels are ChEMBL IDs; axis colors indicate the class as per Figure 2.7B on page 18. Colors on the diagonal are arbitrary and used to distinguish adjacent labels. N = 240 wells. GABA$_A$R (GABA ionotropic receptor); α adrenergic receptor; β adrenergic receptor; Vesicular acetylcholine transporter; nicotinic acetylcholine receptor; N-Methyl-D-aspartate receptor; metabotropic glutamate receptor; σR (σreceptor).

and hope this will invite comparisons using the same approach or the development of superior or complimentary benchmarks.

Certain modifications could expand the observable subset of compound-induced movement behaviors. First, we used a concentration of 33 µM, but the concentration–response experiments indicated that some compounds were phenotypically inactive below 100 µM. Second, affecting complex states such as aggression, addiction, or learning may improve resolution. We used a trivial readout for high-dimensional movement behaviors, but tracking [13, 14, 15], optical flow [16], probabilistic models [17], and deep learning [18] have been successful in analyzing similar data.

Finally, technologies like RNA-seq and mass spectrometry could be applied in concert with behavioral experiments as powerful, high-throughput, and high-dimensional approaches to delineate the mechanisms underlying behavioral modifications. Future studies will likely leverage advances in many of these areas to improve the resolution of behavioral profiling.

Although data from SauronX has been published in recent studies, these have been limited to use at the University of California, San Francisco (UCSF) and through direct collaborations. We suggest that other groups may benefit from this platform, and we hope to spur comparison and further development of open systems for behavioral phenotyping.

2.4 Methods

2.4.1 Extra: Object detection for discrete phenotypes

Seeking a way to identify lethal compound treatments that could be interpreted more directly, we trained You Only Look Once v5 (YOLOv5) [19, 20] deep-learning object-detection models. Single frames from 46 wells were annotated by drawing rectangles around individual animals and labeling them *alive* or *deceased*, based on morphology. For potential future applications, we also included other phenotypes, labeling lateral orientation (as a sign of loss of righting

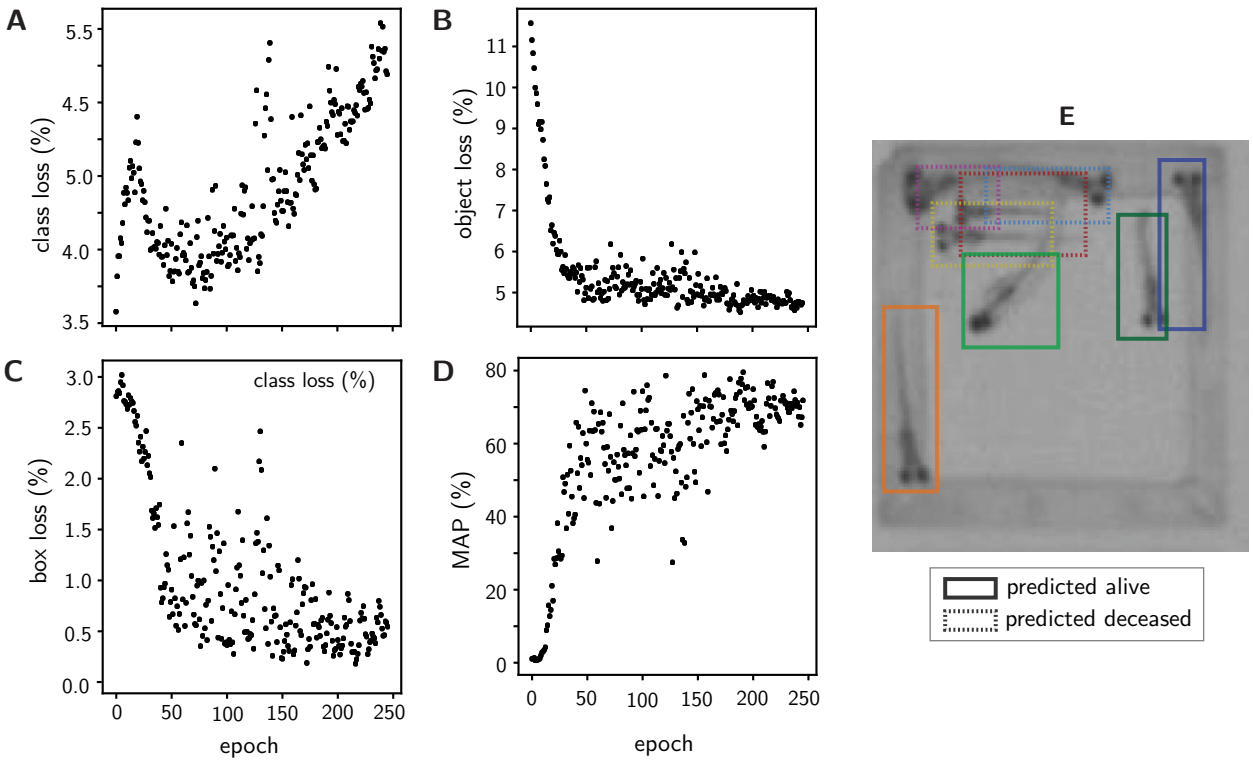

Figure 2.9: Results for the YOLOv5 models. (**A–C:**) test set loss curves for object (*A*), bounding box (*B*), classification of alive versus deceased (*C*). (**D**) Mean average precision in validation set for intersection over union 0.5. (**E**) Example well with detected objects and labels.

reflex) and curvature (as a sign of active motion). Considering only *live*/*deceased*, *live* were detected with 93% precision (Figure 2.9), and *deceased* with 57%. The low precision for *deceased* likely resulted from a relative paucity of training examples (lethal concentrations are preferentially avoided), but the number of deceased animals were counted as $8 -$ alive.

YOLOv5 (Git tag v5.0; commit `f5b8f7d54c9f`) models were trained on images at time 16:55 from 46 wells and 11 plates. Boxes were drawn around animals and labeled using labelImg. Augmented images were generated under D_4 symmetry operations. Cross-validation was performed with a 3:1 train:test split.

2.4.2 Animal husbandry

Zebrafish husbandry was performed as described [21]. Embryos were from group matings of wild-type zebrafish (Singapore, ZFIN:ZDB-GENO-980210-24) raised on a 14/10-hour

Table 2.2: LED stimuli.

color	chromaticity	intensity	manufacturer	part number
red	625 nm	4.6×10^2 lm	Osram Sylvania	LZ4-40R108-0000
green	525 nm	6.6×10^2 lm	Osram Sylvania	LZ4-40G108-0000
blue	650 nm	3.9×10^3 mW	Osram Sylvania	LZ4-40B208-0000
violet	400 nm	3.0×10^3 mW	LED Engin	LZ4-40UB00-00U7
UV	365 nm	8.8×10^2 mW	New Energy	LST1-01G01-UV01-00
white	4000 K	1.6×10^3 lm	New Energy	[1]

'Intensity' values are (total) radiant or luminous flux reported in documentation, not measured. [1] XHP70A-00-0000-0D0BN240E-SB01

light/dark cycle in 28 °C egg water (Instant Ocean (003746) with $NaHCO_3$ to pH 7.0–7.4) [21] until 7 dpf. Animals were maintained in a facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). Experiments were performed in accordance with protocols approved by UCSF's Institutional Animal Care Use Committee (IACUC) and in accordance with the Guide for the Care and Use of Laboratory Animals [22].

2.4.3 Software and data availability

Hardware information, protocols, links to software repositories, extended supplemental data, and the full database are available at https://osf.io/nyhpc. Software is released under an Apache 2.0 license.

2.4.4 Instrument

Note that most of this information is derived from the supplemental methods file for doi:10.1101/2020.01.01.891432v1. A PointGrey Grasshopper GS3-U3-41C6M-C camera (FLIR Integrated Imaging Solutions) and infrared pass filter were used (LE8744 polyester #87, LEE Filters). Six light-emitting diode (LED) arrays were positioned overhead, with 4 LEDs per array.

Two surface transducers were fastened on the stage (5 W transducer, Generic) and used

with a 150 W amplifier (APA150, Dayton Audio). Two 36 V push–pull solenoids (SparkFun Electronics) were positioned near the top of the plate, one contacting the stage directly, and the other contacting a 1 mm-deep strip of synthetic felt. Audio files are provided along with sound pressure level (SPL) measurements in the supplement.

An Arduino Mega 2560 rev 3 (Arduino.cc) drove the LEDs, solenoids, and small sensors while a computer directly controlled the microphone, transducers, and camera.

2.4.5 Video acquisition and processing

Videos were captured for a fixed framerate of 100 Hz and fixed region of interest (ROI), 1600 × 1068 in 8-bit grayscale. Video data was streamed from the camera to a high-performance Non-Volatile Memory Host Controller Interface Specification (NVMe) drive (Samsung EVO 970 PRO M.2 1 TB, Samsung) as raw image sensor ('RAW') data. Acquisition was handled via custom driver code based on the Spinnaker C++ software development kit (SDK) version 1 (FLIR Systems). Acquisition for preset framerate up to 150 Hz was generally stable (neither capture nor transfer were throttled), but we noted that both a fast NVMe drive and transfer via USB 3.1 Gen 1 (i.e. not USB 3.1) was required for this throughput. We chose 100 Hz as our default framerate given the observation that the NVMe drive would need to store raw data for *two* experiments at the same time (one being compressed while the other is being acquired).

After acquisition, timestamps from the firmware clock corresponding to image sensor acquisition were mapped to a `std::chrono::high_resolution_clock` system clock in the drivers. These were then used to trim the video frames to the exact start and end of the battery using timestamps from the image sensor. After these excess frames were deleted, data were compressed with High-Efficiency Video Encoding (HEVC) via ffmpeg 4 using Quick Sync Video (Intel) hardware encoding on an Intel i7-9700K Coffee Lake 8-Core 3.6 GHz processor. This was found to provide better performance with no perceptible loss in quality over both

software encoding and hardware encoding on a GeForce GTX 1060 (NVIDIA) graphics card. We used Constant Quantization Parameter (CQP) (15) to maintain a constant loss suitable for feature calculation. Quick Sync was called with keyframe interval 100 and preset 'veryfast'. These parameters were selected after extensive testing; significant compression artifacts were not observed for these values. A typical video was circa 200 GB in image sensor output (RAW files) and circa 5 GB of compressed data for a 17-minute video. Compression ratios were noticeably higher for videos with less motion. Data were finally partitioned into ROIs for wells for feature calculation.

2.4.6 Data collection and filtration

Healthy larvae were sorted and then immobilized with cold egg water with 25 mL of 4 °C added to 12 mL room-temperature egg water in a 100 cm petri dish containing about 1,000 fish. 8 larvae in 300 µL were then distributed by pipette into the wells of 96-well plates, using trimmed tips to avoid injuring the animals. Plates were incubated at room temperature for 1 hr, at which animals were mobile.

For QC experiments, compound plates and aliquots were stored at −20 °C. 2.0 µL of solvent-dissolved compound was then added to each well. Solvents were dimethyl sulfoxide (DMSO) except for donepezil (water). Some donepezil wells had less than 2.0 µL remaining due to evaporation (annotated in the database). Each concentration–response curve included 5 concentrations on a logarithmic scale with an additional hypothesized ideal concentration. The optimal-concentration QC set was replicated across 15 plates, applying 6 replicates of the 14 compounds and 2 controls ($16 \times 6 = 96$). The vehicle-only adversarial control experiment was collected with 3 plates using earlier hardware and a different battery. However, optimal-concentration QC accuracy was high when subsampled to 3 plates. 5/14 optimal-concentration plates and 1/9 concentration–response plates were excluded because hardware sensors flagged them for potential problems.

The SCREEN-WELL Neurotransmitter library (Enzo Life Sciences) was purchased in solution at 10 mM (peptides 100 µM) in 2015 and stored at −80 °C. A Biomek FX^{P} (Beckman Coulter) was used for randomization.

For NT-650, 1 µL was added per well to yield 33 µM, except for peptides at 0.33 µM. Treatments were randomized across plates and wells. Each plate contained 14 DMSO, 8 water, and 6 lethal eugenol controls, except for 1 of every 7 plates due to an uneven split. 7 replicates were screened per compound, with deviation from 7 due to a subsequent filtration. 13/80 plates were excluded based on sensor readout. We also filtered 23/7680 wells that had insufficient volume of compound in the daughter plate. DIVERSet was screened as provided at 33 µM/well, 1 µL.

2.4.7 Phenotype analysis

Pre-interpolation motion vectors were

$$m'(I^t) = \sum_{ij} \mathbb{1}\left|I_{ij}^t - I_{ij}^{t-1} \geq 10\right|, \tag{2.1}$$

where I^t is the image matrix at 1-indexed frame t. The threshold 10 was chosen by comparing a histogram of pixel intensity changes in wells with and without fish. The final motion m was then quantified by linear interpolation of m' values and image sensor acquisition timestamps to align frames and stimuli. The Scipy 1.3.0 [23] function `interp1d` was called with `kind=previous` and `fill_value=extrapolate`.

RFs were trained with scikit-learn 0.24.1 (RRID:SCR_002577) [24] with default hyperparameters except for the number of estimators, which was 20,000 for treatment–vehicle and 40,000 for multiclass classification. Reported accuracies were out-of-bag.

For NT-650 hit-calling, 4 replicate treatment–vehicle classifiers were trained per compound. Per classifier, all replicate treatment wells were compared with the same number of randomly sampled vehicle wells, restricted to the plates containing the compound treatment and with

the same solvent (DMSO or water). Amoxapine (CHEMBL1113) was dissolved in N-methyl-2-pyrrolidone (NMP); it was compared to DMSO. Lethality was detected by an analogous procedure. For the NT-650 multiclass problem, the mean was taken over 5 confusion matrices, each trained on a stratified subset with 4 wells per compound.

2.4.8 Visualization

Motion-trace visualizations were smoothed from 100 Hz to 10 Hz with a sliding window. T-SNE parameters were scikit-learn defaults. Concentration–response curves were computed with 1,000 bootstrap samples. kernel density estimate (KDE) were Gaussian, calculated with statsmodels 0.10 (RRID:016074) [25] by calling `kdensityfft` with `kernel=gau` and `bw=normal_reference`. Matrix sorting used confusion matrix ordering (CMO) [26] via clana version 4.0; `simulated_annealing` was called with default arguments.

2.4.9 Control experiments and battery design

Assays subject to constraints were generated exhaustively, using LED assays, pure tones and environmental sounds, and combinations. Assays were ranked by the 80th percentile of their accuracy over the 16 unique treatments. The number-of-fish experiment used 2 randomized plates, 2 plates / condition.

2.5 Author contributions

D.M-T. designed the experiments, performed the analyses, and wrote the paper. D.M-T., C.H., C.S.K., and D.K. developed the hardware and drivers. D.M-T., C.H., and C.S.K. wrote the software. J.C.T. collected the data and assisted design. R.A., T.A.T., and D.M-T. collected preceding data. R.K. compiled the CAD diagrams. S.R. and D.M.-T. trained convolutional neural networks. D.M-T., J.C.T., and L.G. performed the randomization. D.K. and M.J.K. provided financial support and supervision.

2.6 Acknowledgments

The authors thank Louie Ramos, Julian Castaneda, Veronica Manzo, Nolan Wong, Madison Burns, Vy Nguyen, and Alexis Marquez for animal husbandry; Giancarlo Bruni for hardware; Capria Rinaldi and Ashley Oh for referenced data; Steven Chen, Michelle Arkin, Adam Renslo for Biomek access and handling; the Dow Chemical Company and Robert Tombari for providing compounds; and Eric Lam for fabrication assistance. Funding was provided by the National Institute on Alcohol Abuse and Alcoholism, the Paul G. Allen Family Foundation, the Genentech Fellowship Program, and National Institutes of Health grant 4T32GM6754714.

References

[1] Giancarlo Bruni et al. "Zebrafish behavioral profiling identifies multitarget antipsychotic-like compounds." In: *Nat. Chem. Biol.* 12 (May 2016), pp. 1–25. ISSN: 1552-4450. DOI: 10.1038/nchembio.2097.

[2] Lee E Dunlap et al. "Identification of Psychoplastogenic N,N-Dimethylaminoisotryptamine (isoDMT) Analogues through Structure–Activity Relationship Studies." In: *J. Med. Chem.* 63.3 (Feb. 2020), pp. 1142–1155. ISSN: 0022-2623, 1520-4804. DOI: 10.1021/acs.jmedchem.9b01404.

[3] Lindsay P. Cameron et al. "A Non-Hallucinogenic Psychedelic Analog with Therapeutic Potential." In: *Nature* 589.7842 (Jan. 2021), pp. 474–479. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-020-3008-z.

[4] Douglas Myers-Turnbull et al. *Zebrafish behavioral profiling using Sauron*. 2020. DOI: 10.17605/OSF.IO/NYHPC.

[5] Duco Endeman, Lauw J Klaassen, and Maarten Kamermans. "Action Spectra of Zebrafish Cone Photoreceptors." In: *PLOS One* 8.7 (2013). ISSN: 1932-6203. DOI: 10.1371/journal.pone.0068540.

[6] Chris F Taylor et al. “Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project.” In: *Nat. Biotechnol.* 26.8 (Aug. 2008), pp. 889–896. ISSN: 1087-0156, 1546-1696. DOI: 10.1038/nbt.1411.

[7] Members of the Panel on Euthanasia and AVMA Staff Consultants. *Guidelines for the Euthanasia of Animals*. Tech. rep. American Veterinary Medical Association, 2020. URL: https://www.avma.org/sites/default/files/2020-02/Guidelines-on-Euthanasia-2020.pdf.

[8] Donald L Neiffer and M Andrew Stamper. “Fish sedation, analgesia, anesthesia, and euthanasia: considerations, methods, and types of drugs.” In: *ILAR J.* 50.4 (2009), pp. 343–360. ISSN: 1084-2020, 1930-6180. DOI: 10.1093/ilar.50.4.343.

[9] Matthew N McCarroll et al. “Zebrafish behavioural profiling identifies GABA and serotonin receptor ligands related to sedation and paradoxical excitation.” In: *Nat. Commun.* 10.1 (Sept. 2019), p. 4078. ISSN: 2041-1723. DOI: 10.1038/s41467-019-11936-w.

[10] Laurens van der Maaten and Geoffrey Hinton. “Visualizing Data using t-SNE.” In: *JMLR* 9.Nov (2008), pp. 2579–2605. ISSN: 1532-4435, 1533-7928. URL: https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf.

[11] Ashwin A Bhandiwad et al. “Auditory sensitivity of larval zebrafish (Danio rerio) measured using a behavioral prepulse inhibition assay.” In: *J. Exp. Biol.* 216.Pt 18 (Sept. 2013), pp. 3504–3513. ISSN: 0022-0949, 1477-9145. DOI: 10.1242/jeb.087635.

[12] Hans-Joachim Böhm, Alexander Flohr, and Martin Stahl. “Scaffold hopping.” In: *Drug Dis. Today Tech.* 1.3 (Dec. 2004), pp. 217–224. ISSN: 1740-6749. DOI: 10.1016/j.ddtec.2004.10.009.

[13] Francisco Romero-Ferrero et al. “idtracker.ai: tracking all individuals in small or large collectives of unmarked animals.” In: *Nat. Met.* 16.2 (Feb. 2019), pp. 179–182. ISSN: 1548-7091, 1548-7105. DOI: 10.1038/s41592-018-0295-5.

[14] Xiaoying Wang et al. “Automatic multiple zebrafish larvae tracking in unconstrained microscopic video conditions.” In: *Sci. Rep.* 7.1 (Dec. 2017), p. 17596. ISSN: 2045-2322. DOI: 10.1038/s41598-017-17894-x.

[15] Vilim Štih et al. “Stytra: An open-source, integrated system for stimulation, tracking and closed-loop behavioral experiments.” In: *PLOS Comput. Biol.* 15.4 (Apr. 2019), e1006699. ISSN: 1553-734X, 1553-7358. DOI: 10.1371/journal.pcbi.1006699.

[16] Tabitha S Rudin-Bitterli et al. “Combining Motion Analysis and Microfluidics – A Novel Approach for Detecting Whole-Animal Responses to Test Substances.” In: *PLOS One* 9.12 (Dec. 2014), e113235. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0113235.

[17] Robert Evan Johnson et al. “Probabilistic Models of Larval Zebrafish Behavior Reveal Structure on Many Scales.” In: *Curr. Biol.* 30 (Dec. 2019). ISSN: 0960-9822. DOI: 10.1016/j.cub.2019.11.026.

[18] Omer Ishaq, Sajith Kecheril Sadanandan, and Carolina Wählby. “Deep Fish.” In: *SLAS Discov* 22.1 (Jan. 2017), pp. 102–107. ISSN: 2472-5560, 2472-5552. DOI: 10.1177/1087057116667894.

[19] Glenn Jocher et al. *ultralytics/yolov5: v5.0 - YOLOv5-P6 1280 models, AWS, Supervise.ly and YouTube integrations*. Apr. 2021. DOI: 10.5281/zenodo.4679653. URL: https://zenodo.org/record/4679653.

[20] Joseph Redmon and Ali Farhadi. “YOLO9000: Better, Faster, Stronger.” In: *arXiv* 1 (Dec. 2016). arXiv: 1612.08242 [cs.CV].

[21] Monte Westerfield. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio Rerio)*. University of Oregon Press, 2000. URL: https://market.android.com/details?id=book-Iy8PngEACAAJ.

[22] National Research Council (US) Committee for the Update of the Guide for the Care and Use of Laboratory Animals. *Guide for the Care and Use of Laboratory Animals*. 8th ed.

Washington (DC): National Academies Press (US), May 2011. ISBN: 9780309154000. DOI: 10.17226/12910.

[23] Eric Jones, Travis Oliphant, Pearu Peterson, et al. *SciPy: Open source scientific tools for Python*. 2001. URL: https://www.scipy.org/.

[24] Fabian Pedregosa. "Scikit-learn: Machine Learning in Python." In: *JMLR* 12 (2011), pp. 2825–2830. ISSN: 1532-4435. URL: https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf.

[25] Skipper Seabold and Josef Perktold. "Statsmodels: Econometric and statistical modeling with python." In: *9th Python in Science Conference*. 2010.

[26] Martin Thoma. "Analysis and Optimization of Convolutional Neural Network Architectures." In: *ArXiv* 1 (July 2017), pp. 34–36. arXiv: 1707.09725 [cs.CV]. URL: https://arxiv.org/abs/1707.09725.

CHAPTER 3

Parallel neurochemical and neurophysiological data

3.1 Preface

Intrigued by the image of a *space* of phenotypic profiles in a reference set such as NT-650, we sought data on the post-treatment zebrafish brain. The desire was to paint a neuropharmacological landscape through multiple data types captured in the same model animal. Temporally between a compound's putative MOAs (i.e. from literature) and its phenotypic profile are data types that can be accessed with high- and moderate-throughput in vivo techniques. Here, we discuss two: phospho-ERK imaging, and neurotransmitter profiling by mass spectrometry. Matt McCarroll, Paul Schnier, and Shigenari Hayashi contributed to this work.

3.2 Phospho-ERK imaging

Phospho-ERK (p-ERK) can be used as a proxy to locate active neurons and active neural networks in zebrafish. We took this approach with confocal microscopy, generating whole-brain activity maps in larval zebrafish following treatment with QC compounds (Figures 3.1A and 3.1B on the following page). P-ERK can be detected in active neurons within minutes following an action potential. To ensure accumulation of this biomarker, we treated the animals with compound for one hour, the same treatment duration used in our QC SauronX data. The resultant maps outlined neuroanatomical regions and specific neurons associated with QC treatments, summarized in Figure 3.2 on page 34.

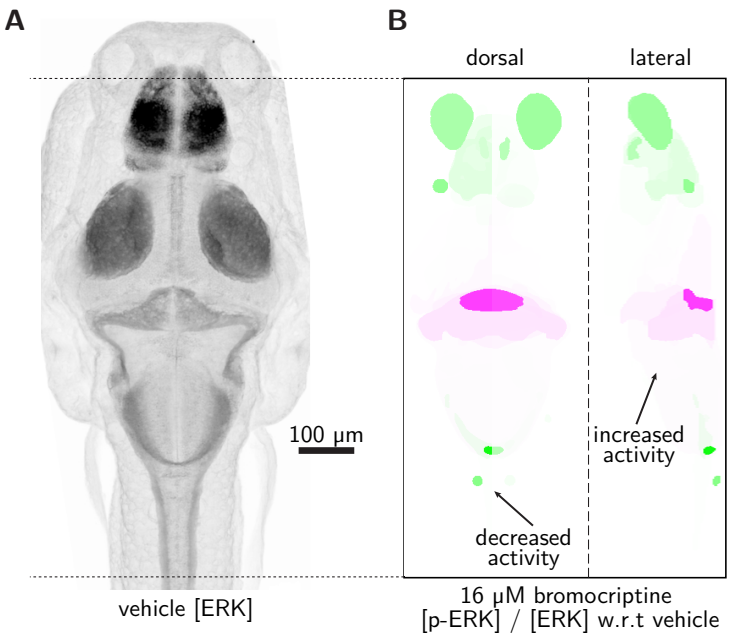

Figure 3.1: Phospho-ERK imaging of zebrafish under compound treatment. (**A**) Total ERK levels in a vehicle-treated adolescent zebrafish. n = 1. Confocal projections showing the average fluorescent intensity of image-registered larval brains stained with anti-phospho-ERK. (**B**) The ratio of phopho-ERK to total ERK levels for compound-treated zebrafish, divide by the same ratio for vehicle-treated animals. n = 10 animals / condition. Pixels are white where $p \geq 10^{-}4$.

3.3 Neurotransmitter profiling by mass spectrometry

In a higher throughput setup, we quantified global (whole-animal) levels of neurotransmitters by mass spectrometry (Figure 3.3 on page 35).

3.4 Methods

3.4.1 Phospho-ERK imaging

P-ERK imaging was performed as described [1]. Animals were fixed, washed, permeabilized, incubated with antibodies, and imaged on a Leica SP8 confocal microscope (RRID:SCR_018169). Images were registered, averaged, and adjusted for brightness and contrast. Heatmaps were computed via Z-Brain MAPmap, Git commit 423ba96b with arguments `nPermutes = 500; FDRThresh = 0.0005; UsingERK = 1`. Brightness and contrast were

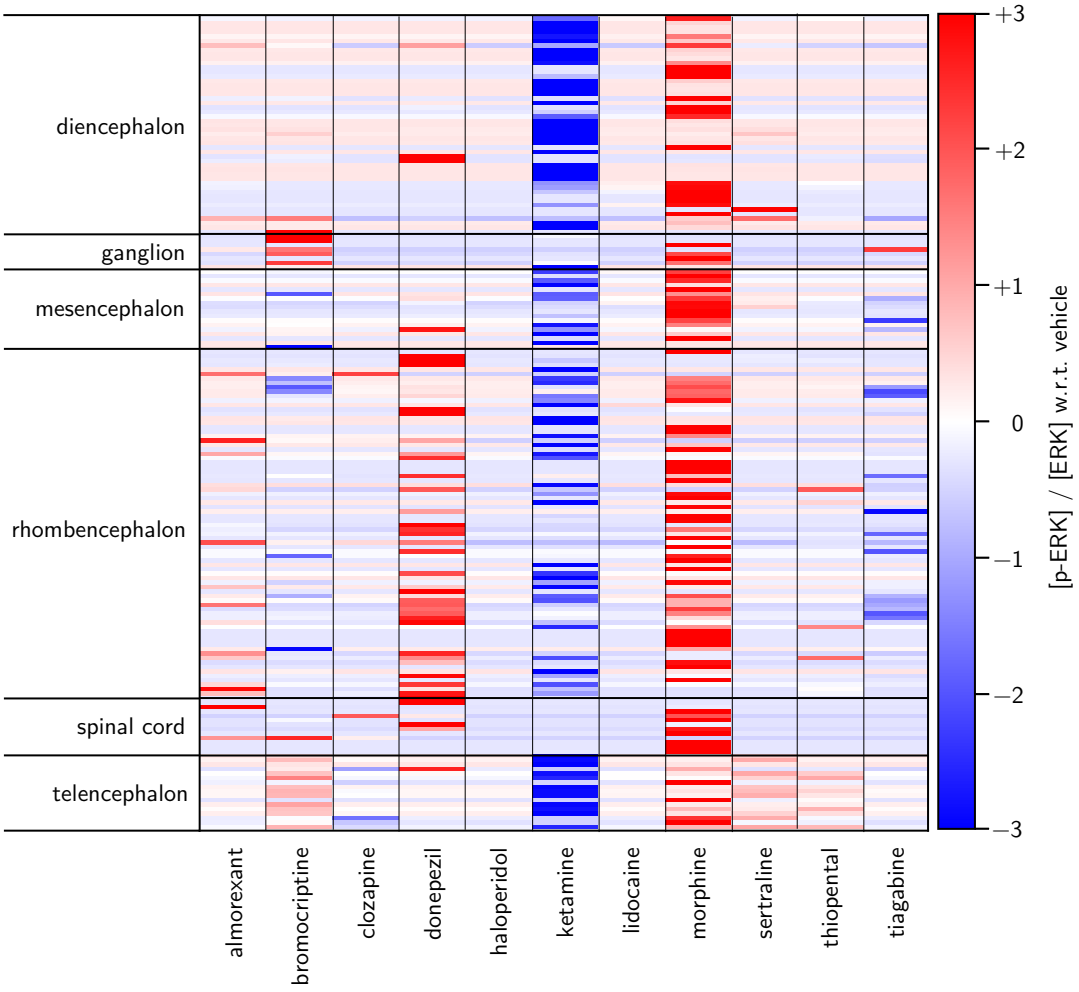

Figure 3.2: Phospho-ERK to total ERK with respect to vehicle controls shown per brain region. Sub-regions are sorted alphabetically within each category. Concentrations were according to Table 2.1 on page 13; morphine was at $100\,\mu\mathrm{M}$. n = 10 animals / condition.

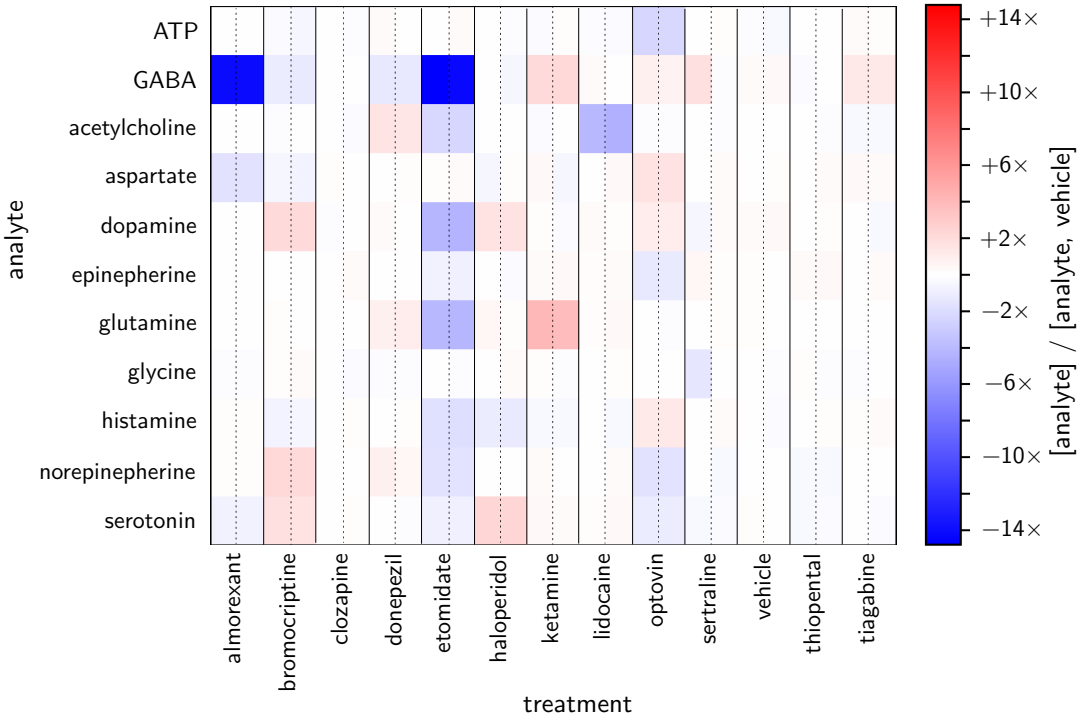

Figure 3.3: Optovin ($6.25\,\mu M$, clozapine ($50\,\mu M$), ketamine ($100\,\mu M$), and tracazolate ($25\,\mu M$) were tested at the same concentrations as in Table 2.1 on page 13. Thiopental was tested at $50\,\mu M$. Haloperidol ($6.25\,\mu M$), donepezil ($50\,\mu M$), lidocaine ($200\,\mu M$), sertraline ($12.5\,\mu M$), almorexant ($50\,\mu M$), and bromocriptine ($50\,\mu M$). Both tracazolate-treated replicates failed quality–control checks. n = 100 animals per sample.

adjusted using Fiji (RRID:SCR_002285). The following antibodies were used: anti-tERK (1:750, Cell Signaling #4696) and anti-pERK (1:750, Cell Signaling #4370).

3.4.2 Neurotransmitter profiling

Lysate samples were purified by an Oasis Mixed-mode Cation eXchange (MCX) exchange extraction column (Waters) and dried in a SpeedVac vacuum concentrator (Thermo Fisher Scientific). Samples were reconstituted in 10 µL 70% acetonitrile. Liquid chromatography–mass spectrometry (LC–MS) experiments were performed on an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (RRID:SCR_020562)) interfaced to a nano-liquid chromatography (LC) system (Thermo Fisher Scientific).

Mobile phases A and B consisted of water and acetonitrile, respectively. The loading chromatography conditions were 70% B, which was increased linearly to 97% over 25 min using a ZIC Hilic column (SeQuant) column with 1 µL injection volume. The nano-electrospray ionization (nano-ESI) source was operated in both positive and negative ionization modes with a capillary voltage of $\pm$ 1900 V. Mass spectra were acquired in the m/z range of 70 Da to 700 Da. Xcalibur (Thermo Fisher Scientific) was used for mass spectrometry (MS) control, data collection, and analysis. Relative quantification of analytes was obtained by generating an extracted ion chromatogram (XIC) for each ion and comparing the average area for 3 runs per sample.

Appendix A

Supplemental items

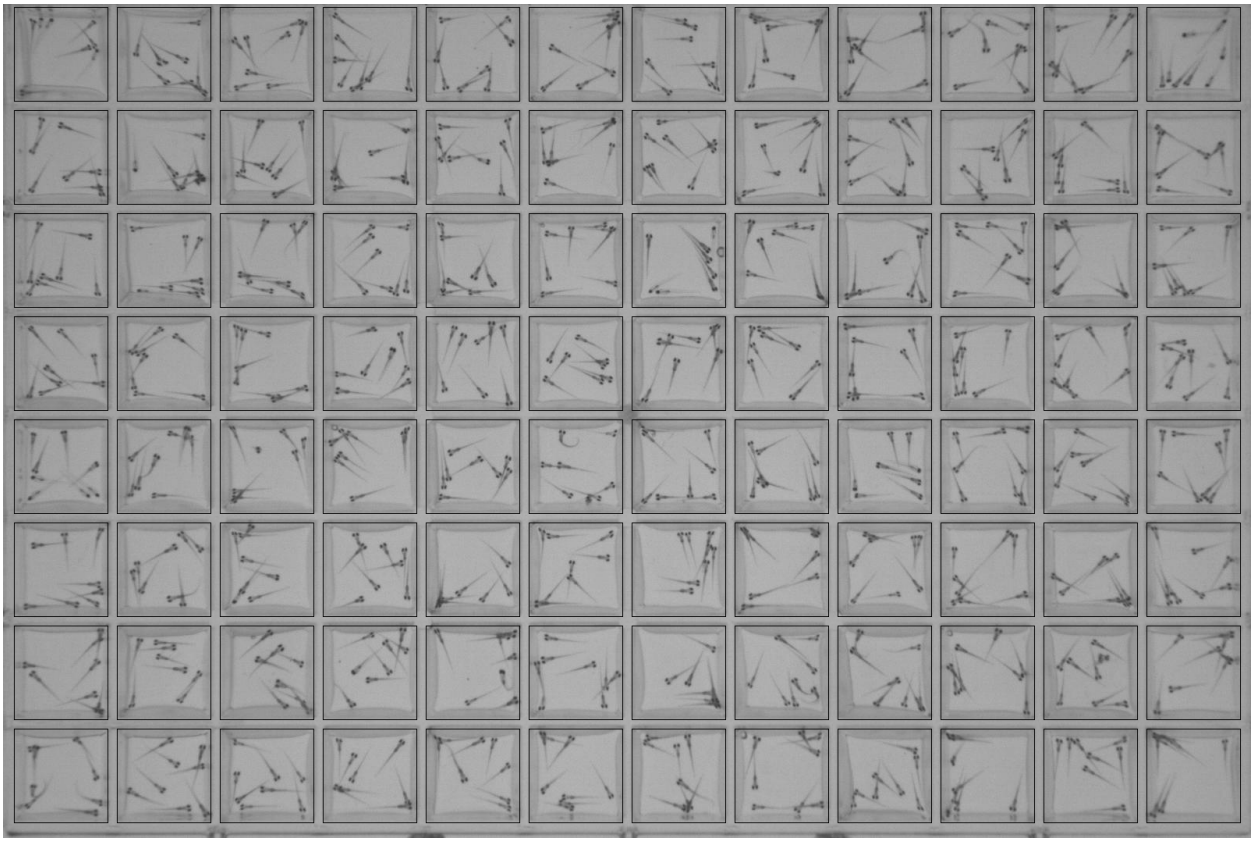

Figure A.1: Example captured frame with ROI overlay.

Table A.1: Database identifiers for referenced SauronX data.

Dataset	Runs
C–R QC	6883,6885,6887,6888,6894,7029,7030,7031
O-D QC	7327,7329,7330,7331,7349,7473,7521,7522,7605
NT-650	7667,7638,7706,7649,7971,7944,7697,7705,7672,7693,7975,7701, 7666,7972,7710,7681,7974,7694,7686,7709,7679,7682,7687,7731, 7743,7983,7751,7987,7970,7771,7759,7980,7783,7762,7940,7957, 7772,7801,7806,7824,7823,7977,7831,7984,7828,8232,8233,8231, 8238,8234,8226,8225,7943,8235,8237,8239,8241,7988,8240,8229, 8230,8228,8227,8236,7833,7832,7829,7830,7836,7834,7849,7842, 7841,7847,7840,7850,7882,7894,7900,7895
Prestwick	8433,8434,8435,8436,8443,8444,8445,8448,8459,8460,8461,8462, 8463,8464,8465,8466
DIVERSet	6579,6580,6634,6636,6647,6664,6666,6669,6671,6673,6679,6682, 6684,6685,6688,6692,6695,6698,6701,6703,6705,6706,6708,6712, 6714,6716,6720,6722,6725,6727,6728,6730,6756,6759,6761,6766, 6768,6773,6775,6777,6778,6783,6801,6803,6804,6808,6811,6813, 6815,6825,6845,6848,6851,6855,6865,7035,7036,7038,7083,7088, 7097,7098,7140,7143,7147,7174,7180,7218,7219,7222,7257,7258, 7259,7264,7280,7282,7283,7287,7290,7293,7294,7296,7311,7314, 7315,7320,7325,7376,7380,7382,7385,7387,7390,7399,7401,7403, 7404,7406,7410,7412,7413,7414,7418,7424,7435,7436,7441,7446, 7448,7450,7453,7457,7459,7461,7464,7468,7469,7470,7477,7479, 7491,7493,7495,7507,7509,7512,7927,7928

Identifiers correspond to the `tag` field of the database table `runs`.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Douglas Myers-Turnbull
534F1557CCA34AD...
Author Signature

12/17/2021
Date