Title
Deep phenotypic profiling of neuroactive drugs in larval zebrafish

Author
Gendelev, Leo

Publication Date
2019

Peer reviewed|Thesis/dissertation

Deep phenotypic profiling of neuro-active drugs in larval Zebrafish

by
Lev Gendelev

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

in

Biophysics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Michael Keiser
4DF1BD06D670465...
Michael Keiser
Chair

DocuSigned by:
DAVID KOKEL
DocuSigned by:
Jim Wells
5F2F4D1A06164C2...
DAVID KOKEL
Jim Wells

Committee Members

In Loving Memory of My Mother,
Lucy Gendelev

Acknowledgements

First of all, I thank my advisor, Michael Keiser, for his unwavering support through the most difficult of times. I will above all else reminisce over our brainstorming sessions, where my half-wrought white-board scribbles were met with enthusiasm and sometimes even developed into exciting side-projects. My PhD experience felt less like work and more like a fantastical scientific journey. For this bottomless support and experience that will surely will never repeat again, I am forever grateful.

I thank David Kokel, who was very much the enabler of this journey with rare scientific generosity and collaborative enthusiasm; for encouraging me to push the high-risk, high-reward Deepfish project forward, not just with words but with the impressive Zebrafish assays and talented personnel of his lab; for allowing me to fully indulge in my love affair for small molecules and neural networks. Thank you for this unique opportunity to work on a truly cutting-edge, exciting, and highly-impactful project.

I thank Jim Wells for helping us navigate this deep-fishing ship safely to shore and for letting an Astrophysicist play with a pipette and even purify proteins as a rotation student.

I thank Bill DeGrado for infusing my mortal brain with a glimpse of what it's like to design proteins in one's head and for getting me to unconditionally surrender a pairwise view of molecular interactions; Matt Jacobson for sharing his encyclopedic insights into computational chemistry both as a field and industry; Michael Grabe for his support and being a key mentor for me over the years.

I thank members of the Kokel lab for all the exciting science and support over the years; Matt McCarroll for testing enough dual-target predictions to map out a sizable pairwise view of the Zebrafish pharmacological landscape, and for making sure the soft-tap project stayed on track, even when I couldn't physically be there to do it myself; Douglas Turnbull for his passion and key

insights and contributions to the Deepfish project and for countless troubleshooting sessions; Cole Helsell for his enthusiasm in discussing details of every aspect of Zebrafish science; Chris Ki for his round-the-clock tech support; Jack Taylor, for not missing a fish and enthusiastically running more plates than most would consider reasonable.

I thank Steve Chen from the Small Molecule Discovery Center for his masterful and unwavering help with randomizing plate layouts.

I thank all members of the Keiser Lab for contributing to a truly unique and special scientific environment; Ben Wong for all his cluster-support; Nick Mew and Elena Caceres for key scientific discussions and ideas; Kangway Chuang for his encyclopedic knowledge, insights, and discussions into machine learning approaches; Garrett Gaskins for being there with me from the very beginning. Here's to all the new and talented additions to the Keiser Lab who will continue to push the frontiers of machine learning in the biosciences.

I dedicate this thesis to my mother, who had unwavering faith in me not only as a scientist, but as a person, and in whose memory I will strive to continue my research in the biological sciences. To my father, thank you for everything. I'm sorry I didn't turn out to be an Astrophysicist, but I hope this will do as a compromise. To my sister; thank you for your love and demonstrating to me that a PhD is attainable - with enough sweat and tears. I thank my friend, Ilan Chemmama, for making sure I made it in mostly one piece. To UCSF, for believing in and accepting an Astrophysicist to the Biophysics PhD program, and for giving us 11 more years with mom than we would have had otherwise.

The text of Chapter 1 is a reprint of the material as it appears in:

> McCarroll, M., Gendelev, L., et al, Leveraging Large-scale Behavioral Profiling in Zebrafish to Explore Neuroactive Polypharmacology, *ACS Chem. Biol.*, (2016)

It appears here with permission from the authors.

The text of Chapter 2 is a reprint of the material as will appear in:

> McCarroll, M.*, Gendelev, L.*, et al, Large-scale behavior-based chemical screening identifies ligands and targets related to paradoxical excitation in zebrafish, *Nature Communications* (2019); accepted. *co-first authors

It appears here with permission from the authors. The supplementary material from this paper has been included as Appendix A.1.

Abstract

Deep phenotypic profiling of neuro-active drugs in larval Zebrafish

Lev (Leo) Gendelev

In-vivo phenotypic screening in larval zebrafish has shown much promise for neuroactive drug discovery, but unleashing its full potential remains a challenge. How do we robustly quantify how drugs modulate zebrafish behavior, and in parallel, how do we unravel which targets or pathways they act by? We develop a phenotypic screening and computational pipeline to begin meeting these challenges. Starting with a motion index (MI) as a readout for phenotype, and correlation distance (CD) as a measure of phenotypic similarity, we extend the similarity ensemble approach to computationally predict targets for *sets* of phenotypic screening hits. Using this approach, we predict an "antipsychotic" target profile for previously uncharacterized hit compounds with MI's matching those of known antipsychotic compounds. For a novel phenotype associated with sedation and paradoxical excitation caused by anesthetics such as etomidate and propofol, we predict not only the canonical GABAergic pathway, but a novel target entirely; the serotonin-6 receptor, which we validate with both in-vitro and in-vivo experiments.

However, our initial attempts at extending this approach to other known drug classes such as stimulants and convulsants are met with unexpected challenges; we hypothesize that the MI signatures and the CD used to compare might not be robust enough for these more subtle phenotypes. And so the Deepfish project is born. We train Siamese neural networks (SNNs) on a highly replicated screen of 650 known neuroactive drugs to learn a custom distance metric for comparing MI. This new distance metric scores higher than CD at the task of separating same-drug replicate pairs versus different-drug pairs, all while *generalizing* to a quality control screen done

months prior. In that arena, the new distance metric gets higher classification accuracy on average, but also strikingly outperforms CD for 3 drugs with more subtle phenotypes.

Armed with a way of training robust distance metrics, we make progress with using unsupervised deep-learning approaches to find more robust representations of behavior. We discover that for computing similarities between these high-dimensional embedded fingerprints, training custom distance metrics is even more imperative. However, we see signs that overfitting is possible with the Siamese networks on our highly-replicated dataset - both with the raw MI and high-dimensional embedded representations - so we design and perform a version of the screen with fully randomized drug layouts, which we will use to benchmark our methods in the near future.

Table of Contents

List of Figures

List of Tables

Introduction

Historically, the very first drugs were discovered serendipitously; in many cases, they were derivatives of plants with some desirable effects in humans[1]. Later on, advances in chemistry and chemical synthesis produced interesting compounds that were subsequently tested on animals or biological tissues for activity[2]. Eventually, on the heels of the revolution in basic biology, genomics, and biotechnology over the last few decades, target based drug discovery began to take over[3]. However, attrition rates and R&D productivity in the era of target-based discovery have been exceedingly high[3]. At the same time, the number of first-in-class drugs - or new-molecular-entities - discovered with phenotypic screening between 1999 and 2008 was almost double the number discovered with target-based screening [3]. For central nervous system (CNS) disorders in particular, target-based approaches have been especially slow, likely due to our limited understanding of CNS physiology and neurocircuitry[4]. This highlights the ever-important role of phenotypic screening in drug-discovery, perhaps due to its ability to discover complex MOAs (mechanisms of action) for a variety of disorders, especially those of the CNS.

However, not all phenotypic screens are equally powerful. Recently, a set of 3 criteria were proposed that any worthwhile phenotypic screen should strive to meet[5]. These include a) the model system's relevance to actual disease; will activity in a biological tissue translate to an in-vivo system? b) Stimulus relevance; is the use of broad cytotoxants to cause cellular injury in a phenotypic model relevant? and c) Proximity of assay to clinical end-point. It is an arduous task for any phenotypic screening setup to meet all of these requirements fully. For instance, many anticancer screens rely on in-vitro cell-lines where the readout is cell-death; so they tend to enrich for cytotoxants. With Alzheimer's disease, our understanding of the underlying disease biology is limited, therefore developing a robust phenotypic assay is challenging[6,7].

For neurological and behavioral disorders, most in-vitro models don't really come close to meeting any of those 3 criteria, which explains why neuroactive drug discovery is probably the slowest of all. I will add another important criteria to the list; d) scale. Mouse cancer models might get relatively high marks on the first 3 criteria, but scaling them up to high-throughput early-stage drug discovery is infeasible.

With these criteria in mind, phenotypic screening in larval Zebrafish is uniquely suited for CNS drug discovery. Larval Zebrafish meet criterion (a) possessing 82% of the orthologues of human genes associated with disease[8], criterion (b) they have specific reactions to human drugs and human-disease relevant stimuli[9], criterion (c) they are an in-vivo system more closely positioned to the clinical end-point than any in-vitro model, and even criterion (d) they can be scaled to high throughput screening (HTS) of large chemical libraries.

I. A platform for phenotypic screening in larval zebrafish

Over the past decade, the Kokel lab has developed and optimized technology that allows for the high-throughput interrogation of neuro-active behavior as a response to small-molecule perturbation. Larval Zebrafish are screened on 96-well plates (usually 8 fish per well) under the influence of small-molecule treatment and various stimuli, and video is recorded of their swimming behavior over the course of 15-30 minutes (**Figure i.1a-c**). Videos are then converted to a "motion index time series" (from here-on referred to in this manuscript as MI) by taking pixel-wise difference maps between consecutive frames, capturing the average motion of fish over time (**Figure i.1d**). This motion index then serves as a phenotypic "fingerprint" of small molecules, a proxy for how the compounds affect fish behavior.

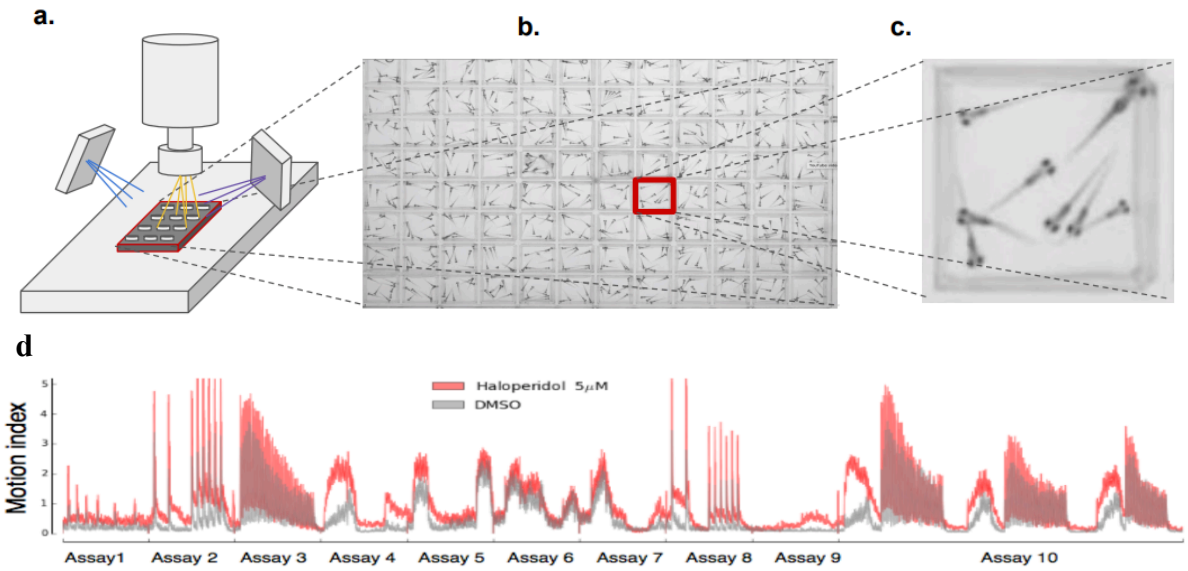


Figure i.1 Zebrafish behavioral screening

A cartoon depicting the phenotypic screening pipeline developed over the last decade in the Kokel Lab. (a) Larval zebrafish are loaded onto 96-well plates and placed under a video-recording apparatus. (b) Videos are recorded of the fish swimming as a response to various stimuli such as a soft solenoid tap or ultra-violet light. Different wells may contain different drug treatments or different doses. (c) Zoomed in view showing an individual well. In most experiments, each well contains 8 fish. (d) Motion index time-series representation. Subsequent frames from the video are subtracted from each other and the sum over pixel differences is computed. Haloperidol, a prototypical antipsychotic shown in red, elicits a unique motion signature in the fish that is distinct from DMSO (negative control) treated fish. Figure reprinted with permission from Bruni et al, 2016[10].

Armed with motion index fingerprints of entire libraries of compounds, we can start asking crucial questions; what *are* the most prolific pharmacologically modulated behaviors exhibited by Zebrafish? Are these phenotypes clinically relevant for humans? Are there novel compounds that cause these phenotypes, too, and if so, what are their MOAs? Are they the canonical or novel entities associated with these behaviors? This MI-based analysis has led to a whole body of work and multiple publications, including early successes on an antipsychotic phenotype[10] and paradoxical excitation in the context of sedation and anesthesia (chapter 2). Investigations of other interesting phenotypes by members of the Kokel Lab are ongoing affairs.

II. A pipeline: from phenotype to target prediction

In the process of working on these specific phenotypes, we created a pipeline for phenotypic drug discovery in larval zebrafish, which goes something like this: starting with the motion index signature for a known neuroactive drug of interest (the query drug), the correlation distance (CD) between the query and each compound from a high-throughput screen of unknown chemical matter is calculated. The list is ranked by CD, hits are triaged based on some reasonable CD cutoff, and remaining compounds form sets of hits defined at increasing stringency cutoffs. Targets are computationally predicted for the sets using the Similarity Ensemble Approach (SEA) [11], followed by Enrichment Factor (EF) calculations[12] (see the methods section V) to compute the set-wise enrichment over random background.

III. Beyond motion index and correlation distance

But as successful as motion index has been for these "low-hanging fruits", subtler phenotypes have been elusive. Stimulants, opioids, analgesics, anxiolytics, muscle relaxants, convulsants and anti-consultants and other drug classes didn't seem to have readily actionable phenotypes in larval fish. In principle, there could be 3 possible explanations for this; (a) these classes of drugs simply have no effect on the behaviors of larval Zebrafish, (b) MI was not a robust enough fingerprint of fish behavior, or (c) CD was not the right tool for the job. Even if option (a) were true, we as scientists must try to disprove it, making the task of exploring (b) and (c) evermore worthwhile. In reality, the answer might be hidden in some combination of (b) and (c); that is, MI is not the most robust representation of fish behavior, CD is an imperfect tool for the job, and the imperfection from both stacks together to provide an overwhelming obstacle for drug discovery.

IV. Guide to the chapters

This thesis is divided into three chapters. They follow the natural progression of scientific inquiry and development that took place during the course of our collaboration with the Kokel Lab over the years.

Chapter 1 is a review that we wrote for ACS Chemical Biology in 2016. It is based on the phenotypic screening pipeline we were in the early stages of developing at the time. This review highlights the major underlying idea behind much of our work; combining larval zebrafish phenotypic screens with computational target identification.

Chapter 2 is a co-first authored story of an *intriguing* phenotype in larval zebrafish that we originally called the "soft-tap phenotype" thanks to the name of the assay that triggered the behavior, but later realized was actually the discovery of a novel model for *paradoxical excitation* in larval zebrafish. In this paper - recently accepted for publication in Nature Communication - we demonstrated a successful application of our pipeline outlined in Chapter 1.

Chapter 3 was chronologically inspired by our early work in chapter 4, where we realized that CD was not an appropriate metric for comparing high-dimensional embeddings of fish motion. However, it turns out that using Siamese neural networks (SNNs) to learn a custom distance function on motion index produces promising results, so we describe this work in chapter 3; unpublished work that will serve as the basis for a paper to be submitted for review soon.

Chapter 4 is a work in progress; it will take the reader through some of the deep-learning approaches we've been exploring to learn more meaningful representations of fish behavior. It will end with a short segue into our initial attempts to employ our methods from chapter 3 to learn a custom Siamese distance function on these high-dimensional representations.

V. Birds-eye view

Our ultimate goal is to be able to combine all the work presented in this thesis as individual chapters into one final opus; training robust distance metrics both on MI and more advanced representations of fish behavior, and applying them to our phenoblasting (phenotype-to-target) pipeline, enabling the discovery of novel targets and compounds for all kinds of interesting neuroactive phenotypes and pharmacological classes.

V. Methods

I. Similarity Ensemble Approach (SEA)

The Similarity Ensemble Approach (SEA)[11] was a technique borne out of Michael Keiser's work in the Shoichet lab at UCSF, over a decade ago, that has found tremendous success in applications both in academia and industry alike. It has been used for predicting off-target effects, to discovering new targets for old drugs[13], and even to driving new biological hypothesis and experiments. I will refer the reader to the original text for details, but the key insight behind SEA is that targets could be related by their shared ligand chemistry. A simple sum over all the pairwise tanimoto similarities across pools of target ligands proved insufficient to relate them; the key idea that made SEA so successful over previous attempts was that the raw tanimoto sum had to be fit to an extreme value distribution. This can be abstracted to relating individual ligands – which can be thought of as targets with a ligand pool of size 1 - to all other targets. Statistics are computed on ligands and targets from the ChEMBL database[14], a massive dataset of ligand-target associations (binding affinities, functional assays, etc) published over the last several decades both by academic and private entities.

II. Enrichment factor calculations

Originally designed for use as "guilt-by-association" EFs in the context of adverse-drug-reactions (ADRs)[12], we extrapolate the fundamental idea of calculating enrichment. In the original case, enrichment was calculated on ADRs; in our case, these are targets. The formula for calculating EFs is given as $E_{xy} = \frac{n}{\frac{A \times T}{N}}$. For a given set to target association (set x, target y), we compute the enrichment as the number of times target y is associated with any compounds from set x (n), divided by the number of times target y is associated with random background sets (T), and the number of times set x is associated with other targets (A). The EF is scaled by N, a matrix normalization factor. In simple terms, the EF *enriches* for specific set-to-target associations while penalizing for non-specific associations. These non-specific associations can come out of two possible realms. In one, they are the associations between a set of compounds and other targets from ChEMBL. In the other, they are associations between a target and random sets of compounds from the screen. Both of these cases are considered, and the EF is penalized by how many times these non-specific associations are made.

IV. References

1. Takenaka, T. Classical vs reverse pharmacology in drug discovery. *BJU Int.* **88**, 7–10 (2008).

2. Lazo, J. S. Rear-view mirrors and crystal balls: a brief reflection on drug discovery. *Mol. Interv.* **8**, 60–63 (2008).

3. Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nat. Rev. Drug Discov.* **10**, 507–519 (2011).

4. Enna, S. J. & Williams, M. Challenges in the search for drugs to treat central nervous system disorders. *J. Pharmacol. Exp. Ther.* **329**, 404–411 (2009).

5. Vincent, F. *et al.* Developing predictive assays: the phenotypic screening 'rule of 3'. *Sci. Transl. Med.* **7**, 293ps15 (2015).

6. Lowe, D. Phenotypic Screening: The State of the Art. *In the Pipeline* (2017). Available at: https://blogs.sciencemag.org/pipeline/archives/2017/07/25/phenotypic-screening-the-state-of-the-art. (Accessed: 1st September 2019)

7. Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug Discov.* **16**, 531–543 (2017).

8. MacRae, C. A. & Peterson, R. T. Zebrafish as tools for drug discovery. *Nat. Rev. Drug Discov.* **14**, 721–731 (2015).

9. Richendrfer, H., Pelkowski, S. D., Colwill, R. M. & Creton, R. On the edge: pharmacological evidence for anxiety-related behavior in zebrafish larvae. *Behav. Brain Res.* **228**, 99–106 (2012).

10. Bruni, G. *et al.* Zebrafish behavioral profiling identifies multitarget antipsychotic-like compounds. *Nat. Chem. Biol.* **12**, 559–566 (2016).

11. Keiser, M. J. *et al.* Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* **25**, 197–206 (2007).

12. Lounkine, E. *et al.* Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **486**, 361–367 (2012).

13. Keiser, M. J. *et al.* Predicting new molecular targets for known drugs. *Nature* **462**, 175–181 (2009).

14. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–7 (2012).

Preamble to Chapter 1

In the early days of the Keiser – Kokel Lab collaboration, everything seemed possible and the world of polypharmacological phenotypic drug discovery in larval zebrafish seemed ripe for the picking. We *could* predict polypharmacological profiles for novel antipsychotic compounds. We *could* even predict novel single-target mechanisms of action for screening hits. We *thought* we could predict novel multi-target mechanisms of action for the paradoxical excitation phenotype (more on that in chapter 2). As it often turns out in the biological sciences, reality is not as simple.

In chapter 1 we suggest that computational target predictions using SEA and EFs are like a bridge allowing us to cross over from compound to target. For the antipsychotics, the SEA and EF calculations seemed to work like a magic torch, illuminating a complex polypharmacological landscape. But as I will explain in the preamble to Chapter 2, although this bridge has led us safely to new continents for single-targets, there are some important pylons still missing from the bridge to the multi-target archipelago.

One of these challenges is likely a result of some of the virtues of SEA itself. SEA is incredibly proficient at predicting "off-target" mechanisms; but the "off-target" prediction might actually be the "on-target" mechanism for a different indication. Methadone, an opioid used for chronic pain management, was predicted by SEA and experimentally validated to antagonize the muscarinic M3, adrenergic α2, and neurokinin NK2 receptors[1], but it's also one of the most common treatments for opioid addiction[2]. Could some of its off-target receptors - or some polypharmacological combination of them – be responsible for its effectiveness in that capacity? In the context of predicting targets for sets of phenotypic hit compounds, how do we know which of the predicted targets is actually responsible for the phenotype?

This is very much an active area of research. I had the chance to dive into this highly complex and important space during the course of our research on the paradoxical excitation phenotype; I will discuss some of my efforts and thoughts about where we currently stand on this in the preamble to chapter 2. For now, I invite the reader to come along for the ride and explore chapter 1, which makes our case for performing phenotypic screening in larval zebrafish and using computational target identification to unravel MOAs.

I. References

1. Keiser, M. J., Irwin, J. J. & Shoichet, B. K. The chemical basis of pharmacology. *Biochemistry* **49**, 10267–10276 (2010).

2. Bodnar, R. J. Opioid addiction. *Peptides* **116**, 68–70 (2019).

Chapter 1:

Leveraging Large-scale Behavioral Profiling in Zebrafish to Explore Neuroactive Polypharmacology

Matthew N. McCarroll, Leo Gendelev, Michael J Keiser*, David Kokel*

University of California San Francisco, Institute of Neurodegenerative Diseases, 675 Nelson Rising Lane, San Francisco, California 94143, United States

* Corresponding authors (MJK and DK)

1.1 Abstract

Many psychiatric drugs modulate the nervous system through multitarget mechanisms. However, systematic identification of multitarget compounds has been difficult using traditional *in vitro* screening assays. New approaches to phenotypic profiling in zebrafish can help researchers identify novel compounds with complex polypharmacology. For example, large-scale behavior-based chemical screens can rapidly identify large numbers of structurally diverse and phenotype-related compounds. Once these compounds have been identified, a systems-level analysis of their structures may help to identify statistically enriched target pathways. Together, systematic behavioral profiling and multitarget predictions may help researchers identify new behavior-modifying pathways and CNS therapeutics.

1.2 Introduction

Polypharmacology is both a challenge and an opportunity in central nervous system (CNS) drug discovery[1-3]. Although drugs are frequently associated with "magic bullets" against single targets, most CNS drugs act on multiple targets simultaneously.[2] Polypharmacology complicates drug discovery efforts because compounds with complex mechanisms are difficult to identify, understand, and optimize. Despite these challenges, polypharmacology also represents an opportunity to discover new compounds with mechanisms of action that have not already been exhaustively exploited *in vitro*. Here, we review two approaches to large-scale behavior-based chemical biology—phenotypic profiling and predictive multitarget enrichments—that can help researchers identify novel compounds with complex polypharmacology.

Psychiatric disorders such as psychosis, depression, and mood disorders are thought to have multigenic and multifactorial etiologies[4-6]. Typically, drug discovery paradigms focus on identifying compounds with single target mechanisms. This single-drug–single-target approach is very effective

for diseases caused by a single mutated gene or deregulated protein. However, most psychiatric drugs including antipsychotics, antidepressants, and anxiolytics are thought to exert their therapeutic effects *via* multiple targets[2,7,8]. As shown by the Psychoactive Drug Screening Program (PDSP), these compounds tend to have complicated target interaction profiles and complex mechanisms of action[9]. For example, antipsychotic drugs are thought to exert their efficacy through a constellation of multiple targets. Attempts to improve these compounds by maximizing single target selectivity have been largely unsuccessful[2]. For CNS drug discovery, more promiscuous compounds are often more effective[2,7,8].

1.3 High Content Compounds

How can multitarget compounds be identified? Unlike *in vitro* assays that identify compounds acting on single predefined targets, phenotypic assays encompass a broader target space. For example, high-content cell-based screens are powerful tools for studying intracellular signaling pathways[10]. Similarly, behavioral screens in whole organisms are a powerful approach to understanding neuronal phenotypes that require integration of a multitude of cell types, sensory systems, and neuronal circuits across an entire organism[11,12]. Behavioral screens are an effective way to identify neuroactive compounds, but understanding their mechanisms is a major challenge. Once identified, how can these compounds be understood? One approach may be to leverage the hit compounds themselves.

Different types of screening assays identify hit compounds with varying extents of biological information. (The term "hit compounds" can have different meanings in different contexts. Here, we use the term "hit" to refer to compounds that have been identified in a screen. In some cases, we may also use the terms "primary hit" and "confirmed hit" to refer to these compounds pre- and post-validation, respectively.) Hit compounds from target-based assays are expected to primarily

contain information about the target they were screened against (Figure 1a). *In vitro* assays can be used to screen millions of compounds and identify hits with maximum potency and selectivity[13,14]. However, since the screening assays are performed in a simplified system (and typically are aimed against a single target), the hit compounds from *in vitro* assays do not contain meaningful information about anything except the original screening target.

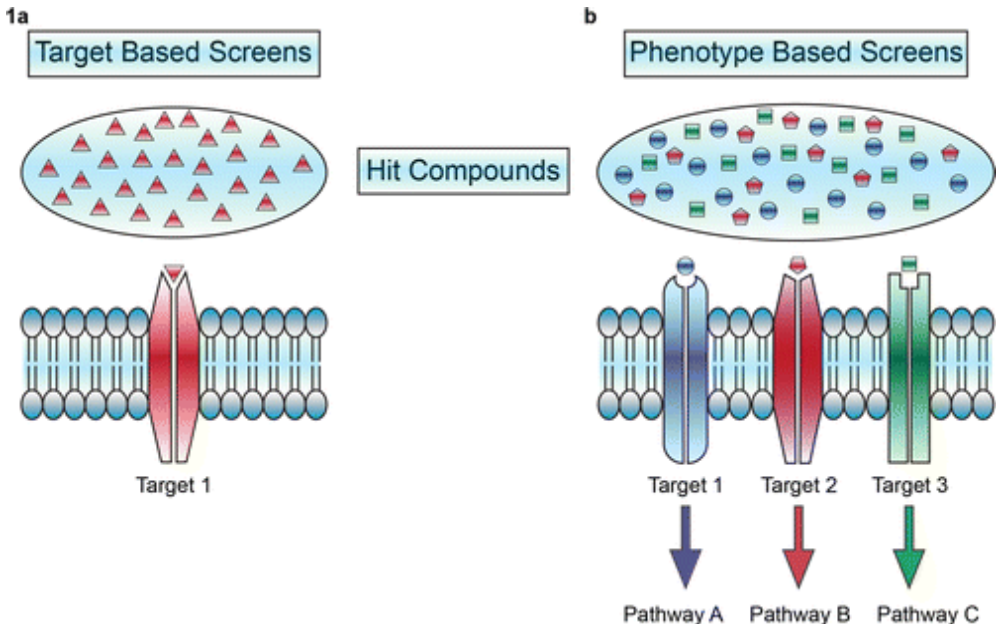


Figure 1.1 Hit compounds from phenotype-based assays showing more target content

Hit compounds from phenotype-based assays showing more target content. (a) Hit compounds from target-based assays are expected to contain biological information about only the target they are screened against. (b) Hit compounds from phenotype-based assays may contain information about multiple targets and pathways that contribute to the phenotype. As a result, sets of structurally diverse and phenotypically related compounds are a valuable tool for understanding complex phenotypes.

By contrast, compounds identified in vivo benefit from more biological context and information than those identified in vitro. These compounds are likely to act on multiple targets and pathways to cause a given phenotype (**Figure 1b**). Although some phenotypes may depend on a

single target15, many phenotypes depend on multiple targets. This is one reason why target identification is such a major challenge in phenotypic screening and why phenotype-based assays are so effective at identifying compounds with complex binding profiles. Any single confirmed hit compound can provide some clues about its target pathways. However, large numbers of phenotypically related compounds contain more information than the sum of their parts. Using large numbers of structurally diverse primary hit compounds, it should be possible to enrich for clues about the signaling pathway networks that underlie complex multitarget signaling pathways in vivo. How can large numbers of phenotypically related hits be identified? This requires a model organism that is well suited for large-scale chemical biology.

1.4 Behavior-based Drug Screening in Zebrafish

Zebrafish are uniquely suited to whole-organism phenotype-based chemicalscreening[16,17], and therefore to the discovery of new drugs with polypharmacological effects. Zebrafish exhibit a wealth of complex behaviors including anxiety/fear[18-21], mating[22,23], feeding[24,25], pain[26-28], sensory[29,30], and sleep behaviors[31-33] (over 190 catalogued[34]). The majority of these responses are robust and conserved and in some instances resemble those of mammals[35,36]. For example, the acoustic startle response is a defensive reaction to potentially threatening stimuli in multiple species of vertebrates, including zebrafish[37,38]. Compounds that modify a particular phenotype, such as zebrafish acoustic startle response, could act on common targets. Alternatively, they could act on different, or multiple, targets in the same, or parallel, signaling pathways. Although this relatively simple zebrafish behavior does not directly simulate human CNS pathologies, the response is controlled by evolutionarily conserved neurotransmitter signaling pathways[38]. As a result, this kind of simple reflex can be useful for identifying compounds with the potential to modify neuronal signaling and behavioral circuits in humans.

Studies investigating how psychoactive compounds affect zebrafish behavior have been successfully conducted in both adult and larval animals[31,39-46]. Many classes of psychoactive compounds can modulate zebrafish behavior including hallucinogens, stimulants, sedatives, antipsychotics, alcohol, as well as other drugs of abuse[36,42,44,47-50]. Similar to their effects on humans, low doses of alcohol and amphetamine can increase zebrafish locomotor behavior, while higher doses of alcohol reduce locomotion[47,51]. Benzodiazepines and barbiturates are sedatives in humans and also reduce motor activity in larval zebrafish[12,40,48,52,53]. In addition, hallucinogens such as ibogaine change adult behavior in light/dark preference assays, promote novel tank exploration, and mirror exploration[42]. These findings suggest that zebrafish are an effective means to identify and characterize psychoactive compounds.

The first high-throughput behavior-based screens in zebrafish assayed thousands of compounds to determine their effects on the photo-motor response (a stereotyped motor behavior in early zebrafish embryos initiated by high intensity light[54]) and sleep/wake cycles[31,55]. Behavior-based screens do not require a priori knowledge of precise neurological mechanisms but rather utilize change in animal behavior as a read out. This is advantageous because much of neuronal network signaling in both healthy and pathological conditions is still unknown. Systematic behavioral profiling enabled identification of unique phenotypes. Importantly, different neurological circuits control similar behavioral phenotypes[56,57], indicating various mechanisms of action through which screening compounds could act to alter behavior. Subsequent clustering of known and novel compounds by phenotypic similarity was then used to predict mechanisms of action[31,55].

Although any single assay may access a limited number of signaling pathways, a large battery of assays may provide a higher-resolution readout of many neurological pathways. Animal behavior is the output of a complex network of neurological signaling pathways[18,29,58]. As a result, many behaviors can be used to identify neuroactive compounds and study their mechanisms[11]. A high-

resolution behavioral battery would increase the scope of neurological systems being assayed as well as further increase the multidimensional behavioral profile generated for specific small molecules. Because zebrafish screens are scalable, they can be used to generate large databases of behavioral information. The more behavioral and chemical space covered in a screen, the greater the predictive power of the system to classify new neuroactive compounds and identify their mechanisms of action.

1.5 Identifying Phenotypically Related Compounds

Phenotypes, and their relationships, can be challenging to measure. As more compounds are profiled against an expanding set of assays, the predictive power of the resulting databases will likely increase. However, as the number of phenotypic dimensions increase, the relationships between phenotypes may become more challenging to identify. To systematically identify the target pathways that give rise to a given phenotype, it is first necessary to identify and categorize phenotypically related compounds from a screen. Here, we describe three approaches to phenotype classification: manual annotation, similarity ranking, and cluster analysis.

One way to classify compounds with robust phenotypes is by manual annotation. For example, assume a researcher is looking for drugs that modify animals' acoustic startle response. Whereas control animals do not respond to a particular acoustic stimulus (**Figure 2a**, upper panel), a subset of screened compounds causes the animals to respond in a robust and reproducible way (**Figure 2a**, lower panel). Expert analysis of recorded movies is one way to identify compounds that cause this behavior. Manual phenotyping can be a beneficial first step because the human brain is a powerful pattern recognition tool, and the experience gives the researcher a first-hand appreciation for the entirety of the screening data set. However, despite the power of expert classification, there

are many limitations to this method, such as the time constraints of scoring increasingly large data sets, user fatigue, and lack of scalability.

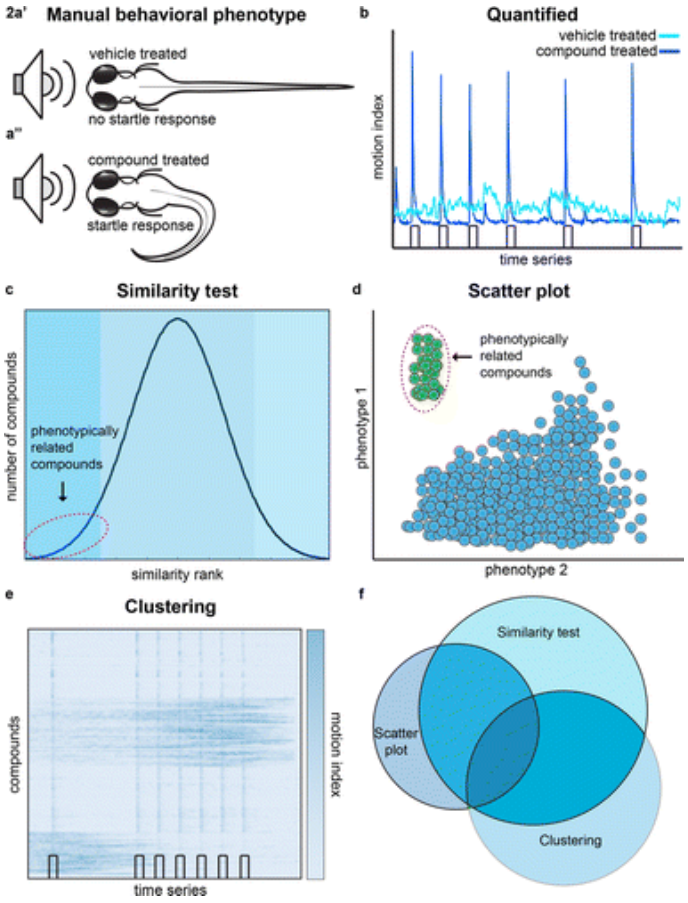


Figure 1.2 Manual and objective quantification methods to identify phenotypically related compounds

---

(a) Manual observation of behavioral phenotypes, for example from an acoustic startle response assay, can be used to identify neuroactive compounds in zebrafish. (a') A sub-threshold acoustic stimulus does not initiate the acoustic startle response in vehicle treated zebrafish but (a'') does cause a startle response in a subset of compound treated animals. (b) An example of behavioral quantification. The motion index (y-axis) is plotted against time (x-axis). Black rectangles represent the timing and duration of the stimulus. Vehicle treated controls (light blue) display no significant change in motion index in response to the stimulus. By contrast, animals treated with some hit compounds (dark blue) show large changes in motion index due to a sensitized startle response. (c) Normal distribution of phenotypic distances relative to a query profile. A small number of related compounds will match the query phenotype (outlined in red). (d) Scatter plot of two separate phenotypes; related compounds group together as having high magnitude in phenotype 1 but low magnitude in phenotype 2. (e) Hierarchical clustering algorithms can be used to group together screening compounds with similar behavioral phenotypes, potentially revealing patterns not recognized by other methods. (f) Venn diagram illustrating potential overlap of phenotypically related compounds identified from different methods.

Objective quantification of behavioral features can help researchers to identify phenotypically related compounds. Many different similarity metrics can be used to represent phenotypic similarity (**Figure 2b**). Approaches to identifying phenotypically related compounds include feature extraction, hierarchical clustering, and calculating distances between time series[31,55]. For example, a plot of distances between a query profile and the entire data set would typically show a normal (or bimodal) distribution, where a few compounds in the data set may be defined as hit compounds depending on how closely they match the query profile (**Figure 2c**). Specific behavioral features can also be used to identify subsets of related hit compounds (**Figure 2d**). And, clustering approaches can be used to identify major clusters of phenotypically related compounds (**Figure 2e**). All of these approaches can be used to organize hit compounds into phenotypically related sets. No single method is best in all situations. One can expect that each of these approaches will result in a large but not complete set of overlap (**Figure 2f**). A combination of approaches may give the most comprehensive and useful results.

1.6 Mining Phenotypically Related Compounds for Multitarget Mechanisms

Phenotype-based screens can identify hundreds of primary hit compounds, and it may be impractical to follow up on all of them. Once phenotypically related compounds have been identified, an important question is how to make sense of them. Common questions include: Are the primary hits reproducible? How should compounds be prioritized? What are their mechanisms of action? Here, we review two approaches for understanding hit compounds and their targets—structure-based clustering and computational target predictions—and introduce the concept of multitarget enrichment factors.

Structure-based clustering is one way that biologists can quickly organize their primary hit compounds into meaningful groups. Structurally related compounds are often prioritized for follow-up studies because chemical substructures that are identified multiple times are likely to indicate truly reproducible hit compounds. However, while structure-based clustering can help researchers to focus on specific compounds and compound series, this alone does not always bring the investigator any closer to understanding the biological mechanisms about how these compounds are working. Computational approaches from the field of systems pharmacology have been successfully used to predict single targets of single compounds. For example, the Similarity Ensemble Approach (SEA) has been used to both identify off-target interactions[59,60] and predict targets of novel compounds identified from *in vivo* phenotypic screens in *C. elegans*[61] and zebrafish[62]. Given the power of SEA to predict targets of single compounds, it may also be possible to identify predicted targets that are statistically enriched among the most phenotypically related primary hit compounds. This approach could enumerate testable hypotheses about how poorly understood and novel compounds affect behavioral phenotypes (**Figure 3c**). For example, in a recent study, we predicted targets for compounds found to modify *C. elegans* feeding behavior[61]. In the first stage, a screen yielded 84 phenotypically related but structurally diverse compounds, which we compared against more than 2000 human targets. Of the 84 compounds, SEA thereby predicted 79 to have one or more human targets in ChEMBL, with 572 compound–target pairs in total. Sixteen of these pairs were tested *in vitro* to validate their putative mammalian targets, of which nine had strong activity. These mechanistic hypotheses were then analyzed *in vivo* against both biologically and pharmacologically homologous proteins through combined genetic and pharmacological perturbations. Together, these data illustrate the use of simple model organisms with target predictions to gain mechanistic information from phenotypic screening compounds. Notably, the pharmacological target predictions

arising from SEA were articulated by ligands known to human targets, but it was possible to map the results back to *C. elegans* biology.

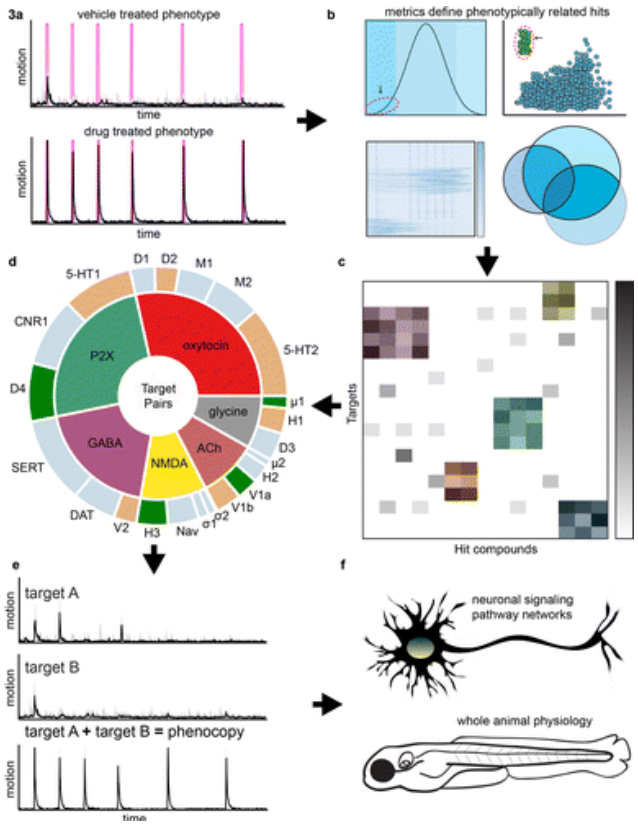


Figure 1.3 Proposed workflow for identifying phenotypically related neuroactive drugs and their targets

(a) Example of a behavioral phenotype used to identify hit compounds from a phenotypic screen; magenta bars represent acoustic stimulus, motion index in black. (b) Similarity ranking, phenotypic features, or clustering methods can then be used to identify related hit compounds. (c) SEA algorithms used to predict hit compound–target space, where subsets of compounds could be predicted to interact with certain targets. (d) Joint enrichment factors (EFs) are calculated from full sets of phenotypically related screening compounds to predict multiple targets potentially required for the phenotype. Plot reads outward: Inner wedges represent the first target of the enriched target-pair, and outer wedges the second target. (e) Validation of the single or multitarget predictions by i*n vivo* treatment and phenocopy of animals with drugs of known pharmacology. Experimentally validated target pairs are colored green in d, while pairs that do not validate are orange. Untested pairs are light blue. (f) Information gained from complete sets of phenotypically related compounds, multitarget predictions, and subsequent phenocopy validation will aid in understanding networks of neuronal circuits, cellular signaling pathways, and whole animal neuropharmacology. Abbreviations: 5-HT1-2, serotonin receptor 1–2; ACh, acetylcholine; CNR1, cannabinoid receptor 1; D1–4, dopamine receptor 1–4; DAT, dopamine active transporter; GABA, gamma-aminobutyric acid; H1–3, histamine 1–3 receptor; μ1–2, mu-opioid receptor 1–2; M1–2, muscarinic acetylcholine receptor 1–2; σ1–2, sigma receptor 1–2; Nav, sodium ion channel; NMDA, *N*-methyl-d-aspartate receptor; P2X, P2X purinoreceptor; SERT, serotonin transporter; V(1a,1b,2), vasopressin (1a,1b,2).

In the future, large databases of structurally diverse compounds and their behavioral profiles may enable researchers to identify multitarget drugs with complex pharmacological profiles[63]. Beyond identifying novel molecules, one can leverage the entire set of phenotypically related compounds in order to gain insight into single targets by which hit compounds modify the system. For drug side-effect prediction, enrichment factors (EFs[64]) have been used to link target singletons to drug adverse events. Perhaps even more exciting is the possibility of being able to use clusters of targets to determine multitarget mechanisms[62,65-67]. For example, by looking at how hit compounds cluster with their predicted targets (**Figure 3c**), one may be able to predict potential combinations of targets that form mechanistic hypotheses. Joint enrichment factors could be developed to predict combinations of targets that underlie a given phenotype (**Figure 3d**). These kinds of chemoinformatic target predictions can help prioritize compounds and identify targets or target combinations that would be difficult to identify by any other means.

There are many reasons why neuroactive drug phenotypes may not translate from zebrafish to humans including issues related to target engagement, the blood–brain barrier, and the evolutionary conservation of specific receptor and neural pathways. Neuroactive compounds are expected to interact with different targets at different concentrations, and saturating levels may engage targets above the clinical dose, thus complicating phenotypic predictions. In zebrafish, many compounds are frequently tested in the μM range (high compound doses >100 μM are often toxic while low doses <100 nM frequently have no effect[31,55]). Although these concentrations may differ from the clinical dose, full dose response curves can be generated to identify behavioral profiles at specific concentrations. For example, if one wanted to identify novel antipsychotic-like compounds it might be more useful to focus on phenotypes caused by relatively low doses (because high-dose phenotypes may be caused by toxic side effects.) Ultimately, compounds identified in zebrafish

should be tested in rodent and other mammalian models to fully understand their pharmacokinetics and brain penetration.

Another issue that pertains to dose is the existence of a blood–brain barrier in the larval animal. Markers for the presence of a protective layer exist in zebrafish as early as 3 days post-fertilization including epithelial cells and the presence of tight junctions[68-71]. The presence of a functional blood–brain barrier in the screening model may increase the chance that any hit compounds will also cross the blood–brain barrier in humans. However, some zebrafish receptors may differ from their human and rodent orthologs in functionally important ways. For example, delta opioid receptors in frogs and zebrafish are insensitive to high affinity ligands at the human delta opioid receptor due to a single amino acid variation[72]. While studies like this one do raise concern for the translatability of compounds identified from a zebrafish screen, other compounds do work on human and zebrafish receptors. For example, a reversible TrpA1 ligand, optovin[15], was discovered in a zebrafish behavioral screen and found to work on both mouse and human orthologs *in vivo* and *in vitro*, respectively. Thus, confirmed hit compounds identified in a zebrafish screen can translate to mammalian neuropharmacology.

A hypothetical multitarget drug discovery workflow might work as follows: First, an interesting behavioral phenotype is defined, and a method of quantification is established (**Figure 3a**). Next, *in vivo* phenotypic screens are performed; hit compounds are identified and clustered into phenotypically related sets (**Figure 3b**). These related hit compounds are then analyzed *via* target-prediction algorithms to determine the potential target space, and joint EFs are calculated to generate hypotheses of target combinations and larger network pathways acting in concert to produce the phenotype (**Figure 3c,d**). To validate these hypotheses, one could then use combinations of small molecules with established pharmacology to test these hypotheses *in vivo* (**Figure 3e**). In this way, researchers could use structurally diverse yet phenotypically related

compounds to gain insight into cellular pathways, neurological networks, and how neuroactive compounds with complex multitarget mechanisms affect the brain and behavior in intact living organisms (**Figure 3f**).

Phenotypically related compounds contain a wealth of biological information. By combining *in vivo* high throughput screening with multidimensional phenotyping, it should be possible to identify large sets of diverse compounds with enough statistical power to drive target prediction and identification. Examination of large and diverse chemical libraries against an extensive repertoire of zebrafish behavioral assays could identify novel neuroactive compounds with complex target interaction profiles, while systems pharmacology methods can predict mechanisms of action for these novel compounds to generate hypotheses for further *in vivo* validation. Together, this approach may describe target-signaling pathways within the cell and circuitry within neuronal networks and even reveal mechanisms of whole animal behavioral neuropharmacology.

Funding Information:

D.K. and M.M. received funding from National Institutes of Health grants AA022583 and K01MH091449. M.K. received funding from National Institutes of Health R44GM093456. L.G. received funding from National Institutes of Health training grant T32EB009383. D.K. and M.K. received funding from the Glenn Foundation for Medical Research.

The authors declare the following competing financial interest(s): D.K. is a consultant for Teleos Therapeutics, which uses zebrafish for neuroactive drug discovery.

1.7 Acknowledgements

The authors would like to thank E. Fertsch, J. Taylor, and members of their research groups for helpful comments on the manuscript.

1.8 Glossary

Phenotypic screening	A screening approach used in biological research to identify molecules or genes that change a cell or animal's phenotype in a desired or interesting way
Neuropharmacology	The study of small molecules or peptides that affect the nervous system, and how these interactions alter behavior
Polypharmacology	The identification and implementation of compounds that interact with multiple biological targets and or disease pathways
Primary hit compound	A compound that has been identified in a screen but has not been validated
Confirmed hit compound	A primary hit compound that has been validated in subsequent assays
In vivo:	Studies performed on whole organisms, Latin for "within the living"
Photomotor response	A stereotyped sensorimotor behavior in zebrafish provoked by visual light but that is not transduced by the eyes or pineal gland
Acoustic startle response	An evolutionarily conserved defensive behavioral response to adverse acoustic stimulus
Enrichment factor	A "guilt-by-association" metric that relates a set of compounds to a protein target, after correction against a distribution of random sets; in this Review, the sets are specifically groups of compounds selected from a phenotypic screen by their ability to trigger a certain behavior
Similarity ensemble approach	Statistical method to predict biological targets for compounds; it uses chemical "fingerprints" to rapidly compare the 2D structure of a query

	compound against structures of ligands experimentally known to bind to approximately 2500 protein targets (using ChEMBL) and outputs an E value for each association.
Systems pharmacology	A network view of drug action, rather than the canonical "one drug, one target" view; seeks to answer questions such as which groups of targets or pathways a drug must modulate in order to achieve a therapeutic effect, as well as to understand how drug or protein target combinations can work by triggering different nodes in a multiscale biological network.

1.9 References

1. Anighoro, A., Bajorath, J., and Rastelli, G. (**2014**) Polypharmacology: Challenges and Opportunities in Drug Discovery J. Med. Chem. 57, 7874– 7887

2. Roth, B. L., Sheffler, D. J., and Kroeze, W. K. (**2004**) Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia Nat. Rev. Drug Discovery 3, 353– 359

3 Besnard, J. *et al.* (**2012**) Automated design of ligands to polypharmacological profiles Nature 492, 215– 220

4 Schizophrenia Working Group of the Psychiatric Genomics Consortium (**2014**) Biological insights from 108 schizophrenia-associated genetic loci Nature 511, 421– 427

5 Kendler, K. S., Aggen, S. H., and Neale, M. C. (**2013**) EVidence for multiple genetic factors underlying dsm-iv criteria for major depression JAMA Psychiatry 70, 599– 607

6 Craddock, N. and Sklar, P. (**2013**) Genetics of bipolar disorder Lancet 381, 1654– 1662

7 Roth, B. L., Sheffler, D., and Potkin, S. G. (**2003**) Atypical antipsychotic drug actions: unitary or multiple mechanisms for "atypicality"? Clin. Neurosci. Res. 3, 108– 117

8 Griebel, G. and Holmes, A. (**2013**) 50 years of hurdles and hope in anxiolytic drug discovery Nat. Rev. Drug Discovery 12, 667– 687

9 Roth, B. L., Lopez, E., Patel, S., and Kroeze, W. K. (**2000**) The Multiplicity of Serotonin Receptors: Uselessly Diverse Molecules or an Embarrassment of Riches? Neuroscientist 6, 252– 262

10 Singh, D. K. *et al.* (**2010**) Patterns of basal signaling heterogeneity can distinguish cellular populations with different drug sensitivities Mol. Syst. Biol. 6, 369

11 Kokel, D. and Peterson, R. T. (**2011**) Chapter 22 - Using the Zebrafish Photomotor Response for Psychotropic Drug Screening, in Methods in Cell Biology (Detrich, H. W., Westerfield, M., and Zon, L. I., Ed.), pp 517– 524

12 Rihel, J. and Schier, A. F. (**2012**) Behavioral screening for neuroactive drugs in zebrafish Dev. Neurobiol. 72, 373– 385

13 Gleeson, M. P., Hersey, A., Montanari, D., and Overington, J. (**2011**) Probing the links between in vitro potency, ADMET and physicochemical parameters Nat. Rev. Drug Discovery 10, 197– 208

14 Szymański, P., Markowicz, M., and Mikiciuk-Olasik, E. (**2012**) Adaptation of High-Throughput Screening in Drug Discovery—Toxicological Screening Tests Int. J. Mol. Sci. 13, 427– 452

15 Kokel, D. *et al.* (**2013**) Photochemical activation of TRPA1 channels in neurons and animals Nat. Chem. Biol. 9, 257– 263

16 Peterson, R. T., Link, B. A., Dowling, J. E., and Schreiber, S. L. (**2000**) Small molecule developmental screens reveal the logic and timing of vertebrate development Proc. Natl. Acad. Sci. U. S. A. 97, 12965– 12969

17 Rennekamp, A. J. and Peterson, R. T. (**2015**) 15 years of zebrafish chemical screening Curr. Opin. Chem. Biol. 24, 58– 70

18 Agetsuma, M. *et al.* (**2010**) The habenula is crucial for experience-dependent modification of fear responses in zebrafish Nat. Neurosci. 13, 1354– 1356

19 Nowicki, M. *et al.* (**2014**) Serotonin antagonists induce anxiolytic and anxiogenic-like behavior in zebrafish in a receptor-subtype dependent manner Pharmacol., Biochem. Behav. 126, 170– 180

20 Maximino, C. *et al.* (**2014**) Fingerprinting of Psychoactive Drugs in Zebrafish Anxiety-Like Behaviors PLoS One 9, e103943

21 Maximino, C. *et al.* (**2010**) Scototaxis as anxiety-like behavior in fish Nat. Protoc. 5, 209– 216

22 Pradhan, A. and Olsson, P.-E. (**2015**) Zebrafish sexual behavior: role of sex steroid hormones and prostaglandins Behav. Brain Funct. 11, 23

23 Gumm, J. M., Snekser, J. L., and Iovine, M. K. (**2009**) Fin-mutant female zebrafish (Danio rerio) exhibit differences in association preferences for male fin length Behav. Processes 80, 35– 38

24 De Marco, R. J. *et al.* (**2014**) The behavior of larval zebrafish reveals stressor-mediated anorexia during early vertebrate development Front. Behav. Neurosci. 8, 367

25 Jordi, J. *et al.* (**2015**) A high-throughput assay for quantifying appetite and digestive dynamics Am. J. Physiol. - Regul. Integr. Comp. Physiol. 309, R345– R357

26 Curtright, A. *et al.* (**2015**) Modeling nociception in zebrafish: a way forward for unbiased analgesic discovery PLoS One 10, e0116766

27 Malafoglia, V. *et al.* (**2013**) The zebrafish as a model for nociception studies J. Cell. Physiol. 228, 1956– 1966

28 Malafoglia, V. *et al.* (**2014**) Extreme thermal noxious stimuli induce pain responses in zebrafish larvae J. Cell. Physiol. 229, 300– 308

29 Del Bene, F. *et al.* (**2010**) Filtering of Visual Information in the Tectum by an Identified Neural Circuit Science 330, 669– 673

30 Lacoste, A. *et al.* (**2015**) A Convergent and Essential Interneuron Pathway for Mauthner-Cell-Mediated Escapes Curr. Biol. 25, 1526– 1534

31 Rihel, J. *et al.* (**2010**) Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation Science 327, 348– 351

32 Gandhi, A. V., Mosser, E. A., Oikonomou, G., and Prober, D. A. (**2015**) Melatonin Is Required for the Circadian Regulation of Sleep Neuron 85, 1193– 1199

33 Singh, C., Oikonomou, G., and Prober, D. A. (**2015**) Norepinephrine is required to promote wakefulness and for hypocretin-induced arousal in zebrafish eLife 4, e07000

34 Kalueff, A. V. *et al.* (**2013**) Towards a Comprehensive Catalog of Zebrafish Behavior 1.0 and Beyond Zebrafish 10, 70– 86

35 Panula, P. *et al.* (**2010**) The comparative neuroanatomy and neurochemistry of zebrafish CNS systems of relevance to human neuropsychiatric diseases Neurobiol. Dis. 40, 46– 57

36 Mathur, P. and Guo, S. (**2010**) Use of zebrafish as a model to understand mechanisms of addiction and complex neurobehavioral phenotypes Neurobiol. Dis. 40, 66– 72

37 Burgess, H. A. and Granato, M. (**2007**) Sensorimotor Gating in Larval Zebrafish J. Neurosci. 27, 4984– 4994

38 Wolman, M. A. *et al.* (**2015**) A Genome-wide Screen Identifies PAPP-AA-Mediated IGFR Signaling as a Novel Regulator of Habituation Learning Neuron 85, 1200– 1211

39 Kokel, D. and Peterson, R. T. (**2008**) Chemobehavioural phenomics and behaviour-based psychiatric drug discovery in the zebrafish Briefings Funct. Genomics Proteomics 7, 483– 490

40 Baraban, S. C., Dinday, M. T., and Hortopan, G. A. (**2013**) Drug screening in Scn1a zebrafish mutant identifies clemizole as a potential Dravet syndrome treatment Nat. Commun. 42410

41 Neelkantan, N. *et al.* (**2013**) Perspectives on Zebrafish Models of Hallucinogenic Drugs and Related Psychotropic Compounds ACS Chem. Neurosci. 4, 1137– 1150

42 Cachat, J. *et al.* (**2013**) Unique and potent effects of acute ibogaine on zebrafish: The developing utility of novel aquatic models for hallucinogenic drug research Behav. Brain Res. 236, 258– 269

43 Cachat, J. *et al.* (**2011**) Three-dimensional neurophenotyping of adult zebrafish behavior PLoS One 6, e17597

44 Grossman, L. *et al.* (**2010**) Characterization of behavioral and endocrine effects of LSD on zebrafish Behav. Brain Res. 214, 277– 284

45 Tran, S., Muraleetharan, A., Fulcher, N., Chatterjee, D., and Gerlai, R. (**2015**) MK-801 increases locomotor activity in a context-dependent manner in zebrafish Behav. Brain Res. 296, 26– 29

46 Tran, S., Nowicki, M., Muraleetharan, A., Chatterjee, D., and Gerlai, R. (**2015**) Differential effects of acute administration of SCH-23390, a D1 receptor antagonist, and of ethanol on swimming activity, anxiety-related responses, and neurochemistry of zebrafish Psychopharmacology (Berl.) 232, 3709– 3718

47 Guo, N. *et al.* (**2015**) Influences of acute ethanol exposure on locomotor activities of

zebrafish larvae under different illumination Alcohol 49, 727

48 Gebauer, D. L. *et al* (**2011**) Effects of anxiolytics in zebrafish: Similarities and differences between benzodiazepines, buspirone and ethanol Pharmacol., Biochem. Behav. 99, 480– 486

49 Kyzar, E. J. *et al.* (**2012**) Effects of hallucinogenic agents mescaline and phencyclidine on zebrafish behavior and physiology Prog. Neuro-Psychopharmacol. Biol. Psychiatry 37, 194– 202

50 Stewart, A. *et al.* (**2011**) Zebrafish models to study drug abuse-related phenotypes Rev. Neurosci. 22, 95– 105

51 Kyzar, E. *et al.* (**2013**) Behavioral effects of bidirectional modulators of brain monoamines reserpine and d-amphetamine in zebrafish Brain Res. 1527, 108– 116

52 Renier, C. *et al.* (**2007**) Genomic and functional conservation of sedative-hypnotic targets in the zebrafish Pharmacogenet. Genomics 17, 237– 253

53 Chen, F. *et al.* (**2015**) Effects of lorazepam and WAY-200070 in larval zebrafish light/dark choice test Neuropharmacology 95, 226– 233

54 Kokel, D. *et al.* (**2013**) Identification of Nonvisual Photomotor Response Cells in the Vertebrate Hindbrain J. Neurosci. 33, 3834– 3843

55 Kokel, D. *et al.* (**2010**) Rapid behavior—based identification of neuroactive small molecules in the zebrafish Nat. Chem. Biol. 6, 231– 237

56 Ahrens, M. B. *et al.* (**2012**) Brain-wide neuronal dynamics during motor adaptation in zebrafish Nature 485, 471– 477

57 Yokogawa, T., Hannan, M. C., and Burgess, H. A. (**2012**) The Dorsal Raphe Modulates Sensory Responsiveness during Arousal in Zebrafish J. Neurosci. 32, 15205– 15215

58 Wyart, C. *et al.* (**2009**) Optogenetic dissection of a behavioural module in the vertebrate spinal cord Nature 461, 407– 410

59 Keiser, M. J. *et al.* (**2007**) Relating protein pharmacology by ligand chemistry Nat. Biotechnol. 25, 197– 206

60 Keiser, M. J. *et al.* (**2009**) Predicting new molecular targets for known drugs Nature 462, 175– 181

61 Lemieux, G. A. *et al.* (**2013**) In Silico Molecular Comparisons of C. elegans and Mammalian Pharmacology Identify Distinct Targets That Regulate Feeding PLoS Biol. 11, e1001712

62 Laggner, C. *et al.* (**2012**) Chemical informatics and target identification in a zebrafish phenotypic screen Nat. Chem. Biol. 8, 144– 146

63 Kokel, D. *et al.* (**2012**) Behavioral barcoding in the cloud: embracing data-intensive digital phenotyping in neuropharmacology Trends Biotechnol. 30, 421– 425

64 Lounkine, E., *et al.* (**2012**) Large-scale prediction and testing of drug activity on side-effect targets Nature 486, 361– 367

65 Ciceri, P. *et al.* (**2014**) Dual kinase-bromodomain inhibitors for rationally designed polypharmacology Nat. Chem. Biol. 10, 305– 312

66 Carrieri, A., Pérez-Nueno, V. I., Lentini, G., and Ritchie, D. W. (**2013**) Recent trends and future prospects in computational GPCR drug discovery: from virtual screening to polypharmacology Curr. Top. Med. Chem. 13, 1069– 1097

67 Zhao, S. and Iyengar, R. (**2012**) Systems Pharmacology: Network Analysis to Identify Multiscale Mechanisms of Drug Action Annu. Rev. Pharmacol. Toxicol. 52, 505– 521

68 Eliceiri, B. P., Gonzalez, A. M., and Baird, A. (**2011**) Zebrafish Model of the Blood-Brain Barrier: Morphological and Permeability Studies Methods Mol. Biol. 686, 371– 378

69 Jeong, J.-Y. *et al.* (**2008**) Functional and developmental analysis of the blood–brain barrier in zebrafish Brain Res. Bull. 75, 619– 628

70 Xie, J., Farage, E., Sugimoto, M., and Anand-Apte, B. (**2010**) A novel transgenic zebrafish model for blood-brain and blood-retinal barrier development BMC Dev. Biol. 10, 76

71 Zhang, J. *et al.* (**2010**) Establishment of a neuroepithelial barrier by Claudin5a is essential for zebrafish brain ventricular lumen expansion Proc. Natl. Acad. Sci. U. S. A. 107, 1425– 1430

72 Vardy, E. *et al.* (**2015**) Single Amino Acid Variation Underlies Species-Specific Sensitivity to Amphibian Skin-Derived Opioid-like Peptides Chem. Biol. 22, 764– 775

Preamble to Chapter 2

In chapter 1, we made our case for phenotypic screening in larval zebrafish and proposed a pipeline for leveraging its unique strengths. Chapter 2 will take the reader into our investigation of a very specific and distinct phenotype characterized by sedation and paradoxical excitation, induced by anesthetics such as etomidate and propofol. Originally called the "soft-tap" phenotype (see **Figure 2.1a**) – thanks to the soft-solenoid whose tapping triggered the behavior – we eventually realized it was a model for *paradoxical excitation* in larval zebrafish. We will show that using phenotypic screening in larval zebrafish, we discover 125 structurally diverse "hits" that cause this unique phenotype, and with SEA and EF predictions are able to predict not only the canonical GABAergic mechanism, but a novel biological target entirely - the serotonin-6 (HT6) receptor - and to some extent, the metabotropic glutamate 4 receptor (mGluR4).

Chapter 2 might read like a self-contained success story for our platform. But in this preamble, I'd like to focus instead on some of the interesting discoveries and challenges that didn't make it into that story, but merit further investigation, and might contain one or two nuggets of value for the scientific community. Perhaps in the future our group, or even other groups, might unravel the mysteries contained therein.

I. Dual-target EF predictions lead to exciting experimental results

After we had confirmed the novel single target HT6, we were quite motivated and set our sights on the prized multi-target real estate. In our initial dual-target EF calculations, one of the most prominent predictions was the combination of Epidermal Growth Factor Receptor (ERBB2) with Voltage-Gated Sodium Channel (SCN9A).

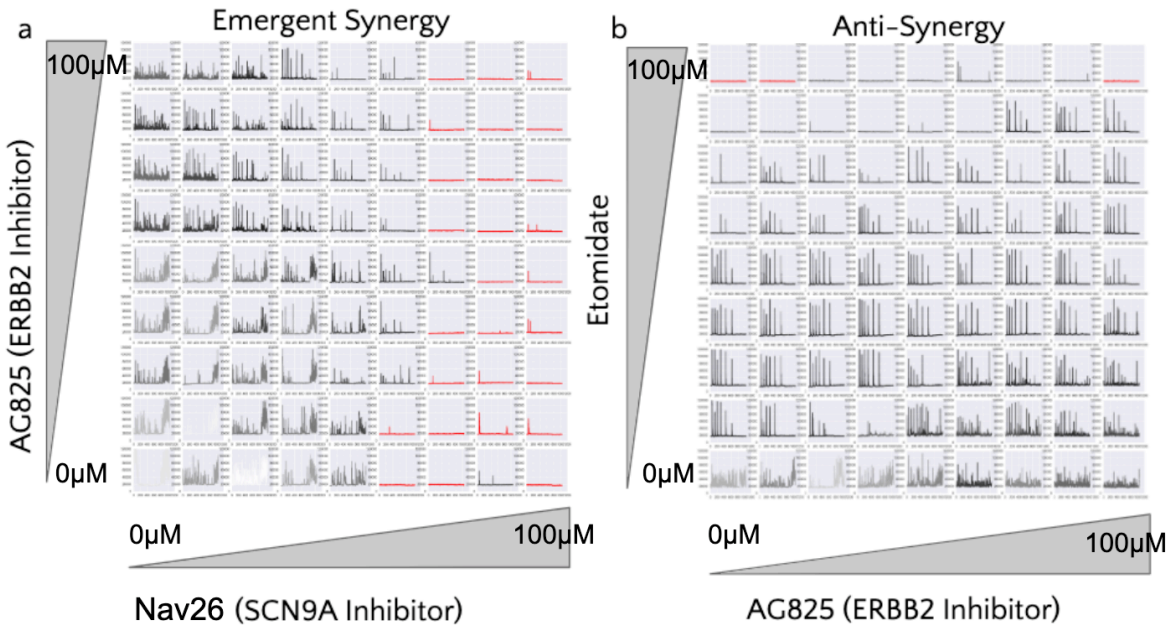


Fig 2.i.1 Dual-target motion-index phenocopy experiments.

Two examples of wildly different effects with the same ERBB2 inhibitor, AG 825. Each cell in both matrices have a sample of MI from two key assays for the soft-tap phenotype; the soft-tap and the blue-light assays a) x-axis; increasing titration of Nav26, an SCN9A inhibitor, y-axis; increasing titration of AG 825. Neither compound causes the soft-tap phenotype by itself (left column, or bottom row). Emergent synergy is seen towards the middle of the panel. b) Increasing the dose of AG 825 weakens the soft-tap effect from Etomidate (anti-synergy). Red traces indicate toxic dose.

The sodium channel prediction made sense in the context of sedation; these ion channels could be placed in pathways that modulate neuronal excitation and inhibition. However, the ERBB2 prediction was more unconventional and harder to place in the context of the CNS. But we were intrigued. We did a simple experimental test to see if this prediction had any merit; we attempted to *phenocopy* the behavior with known inhibitors of these two predicted targets. When we purchased and profiled a known inhibitor of SCN9A, Nav 26, and AG 825, a known inhibitor of ERBB2, in the fish, neither compounds caused the soft-tap response individually, but did cause a response in combination (**Figure 2.i.1a)**.

We were excited. Combination therapy[1,2] is considered one of the most promising directions in drug-discovery because of the ability to reduce doses (synergistic potency[3]). Ours was a case of

emergent synergy, arguably the most sought after prize in drug discovery; where neither drug causes an effect on its own but in combination the drugs cause a strong response. To our best knowledge, ours was the first case of the specific prediction of *emergent synergy* in the context of a phenotypic screen.

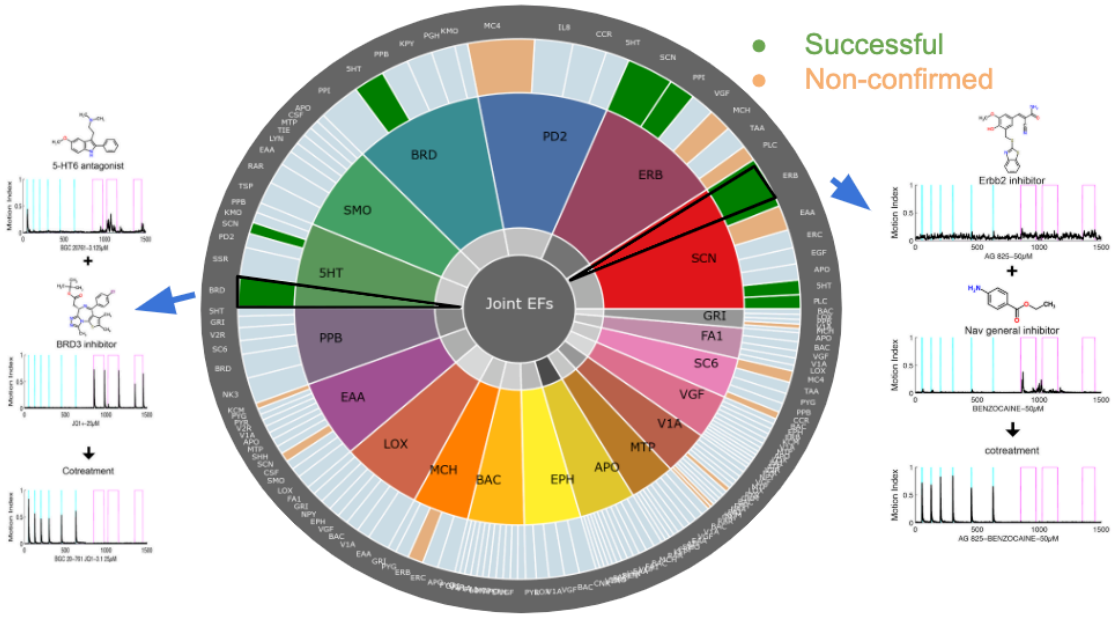


Fig 2.i.2 Dual-target EF predictions

Sunburst diagram for dual target EF predictions for the top 173 hit compounds. The inner rainbow-colored wheel contains the first target and the outer light-blue colored wheel contains the second target of the dual target pair predictions. Outer ring slices filled in green are successfully retested predictions; orange filled outer ring slices are for unsuccessful predictions.

And so the fishing expedition began. Matthew McCarroll from the Kokel Lab tested dozens of dual-target predictions (**Figure 2.i.2**). There were several more successful cases, such as some synergism between an Nav 26, the SCN9A inhibitor, and JQ1, a bromo-domain 3 (BRD3) inhibitor, but hit rates were relatively low. There was even the intriguing case of *negative-synergy* between etomidate and AG 825, the same ERBB2 inhibitor that together with Nav 26 caused the emergent-synergy phenomenon (**Figure 2.i.1b)**. Here, increasing concentrations of AG 825 caused a weakening of the soft-tap response triggered by etomidate. It seemed we had the Strange Case of Dr Jekyll and Mr

Hyde on our hands; AG 825 played a synergistic role in the morning, but exhibited the exact opposite behavior at night.

II. Exciting results are followed by confusion, frustration, and exploration of new methods

Much time was spent on literature review; searching for signs that our predictions and experimental results could be placed in some biological framework. Some theories were developed, but we weren't too confident in them. More importantly, the compounds that were underlying the original SCN9A+ERBB2 prediction, failed to validate *in-vitro*. Were the predictions completely wrong or did we test the compounds against the wrong subtypes of the receptor? Even more intriguing was the possibility that the predictions were actually the *off-target* effects of the screening hits; that is, perhaps there were targets that we didn't test that were the actual *drivers* of the phenotype, and the targets we did test just came along for the ride.

I tested the latter possibility by constructing a new SEA[4], where targets were first broken up into target *fragments*. These fragments were based on a structural clustering of ChEMBL[5] target ligand pools using the Butina clustering algorithm[6]. SEA predictions were then performed on a large subset of compounds from ChEMBL, and fragments were clustered based on the compound to fragment graph using the MCL network clustering algorithm[7].

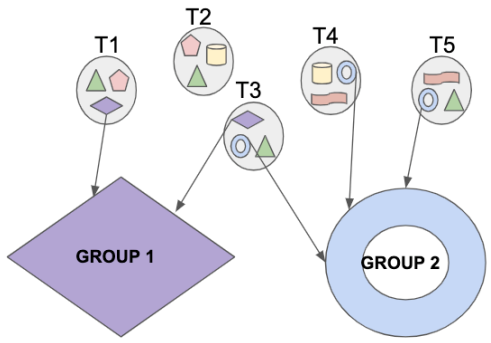


Fig 2.i.3 Target-fragment clustering.

Illustration of the ChEMBL target-fragment clustering scheme. We break up ChEMBL targets into fragments by clustering their associated ligands. The fragments are grouped by clustering them based on a SEA network of ligands and their target-fragment associations. A new SEA library is then constructed based on these "fragment groups" instead of the usual ChEMBL targets.

The resulting clusters represented "groups of fragments" – let's just call them "groups" - that shared ligand chemistry (**Figure 2.i.3**). The new SEA predictions were for these "groups" instead of the original ChEMBL targets. This way, both the on and off-target predictions might be present in an enriched "group". The idea was that these enriched "groups" could generate testable hypothesis. For example, in the paradoxical-excitation case, when we performed SEA-group calculations on the hit compounds, the top enrichment was for group 12, which, as it turned out, contained various $GABA_AR$ subtypes, Peripheral-type benzodiazepine receptor (TSPO), as well as the arachidonate lipoxegenase-5 (LOX5) (**Table 2.i.1**).

Table 2.i.1 Target fragments in group 12

Group#	Target fragment
12	GABA-A receptor; alpha-1/beta-3/gamma-2_cluster_1
	GABA receptor alpha-2 subunit_cluster_0
	Arachidonate 5-lipoxygenase_cluster_15
	…
	Peripheral-type benzodiazepine receptor_cluster_12
	GABA-A receptor; alpha-6/beta-3/gamma-2_cluster_1
	Metabotropic glutamate receptor 5_cluster_1

Perhaps this could explain the AG 825 – etomidate anti-synergy observation; the prediction that we originally thought was ERBB2 was actually driven by LOX5, thanks to some shared ligand chemistry between the receptors. A literature review was able to place the 5,12 lipoxygenases in the context of GABA neurotransmission[8,9]. Indeed, there was another "group" with both ERBB2 and lipoxegenase-12 (LOX12), suggesting that the same shared ligand chemistry that had LOX5 taking the role of ERBB2 in its antagonism of $GABA_AR$ allowed it to take the role of ERBB2 in its synergism with SCN9A. We tested this hypothesis by ordering several LOX5 inhibitors and attempted to phenocopy the emergent synergy and anti-synergy phenotypes by performing co-treatment experiments with the SNC9A inhibitor and Etomidate. These initial results were negative.

III. Discussion

We couldn't disprove the prediction that ERBB2 was driving the phenotype. However, it is not impossible that testing additional known LOX5 inhibitors could have changed the story. In fact, in the case of the serotonin-6 prediction in chapter 2, multiple known inhibitors of HT6 failed to phenocopy; suggesting that differences between the Zebrafish and Human orthologs make it difficult to predict which specific drugs will work in fish, and warrant further investigation of the LOX5 and LOX12 hypothesis.

Another discovery we made along the way was that calculating EFs on multiple sets of compounds - assembled by increasingly stringent correlation distance threshold - drastically improved our signal to noise ratio for single-target predictions, but in the process caused the disappearance of our ERBB2 signal. This was additional evidence that ERBB2 was not the target responsible for the observed synergy with SCN9A and the observed anti-synergy with $GABA_AR$.

Although we weren't able to explain our observations with the LOX5/12 hypothesis, they point to the fact that the paradoxical excitation phenotype is a polypharmacological one by nature,

offering a fertile breeding ground for testing bleeding-edge chemoinformatic approaches for those unphased by the challenges of unravelling its secrets. From the scientific and biological perspective, it would be of great value to the community to demystify these observations; potentially new and completely novel pathways related to sedation and paradoxical excitation *should* exist based on our data; and might be discovered with the right set of experiments and techniques. But for now, I hope the reader will forgive us some of our shortcomings in unraveling the full beauty of the polypharmacological multi-target story, and enjoy Chapter 2; the story of how we discovered a novel phenotype in larval zebrafish - paradoxical excitation - and unraveled some of the novel single-target-biology and pharmacology that caused it using our combined phenotypic screening and computational target identification approach outlined in Chapter 1.

I. References

1. Pemovska, T., Bigenzahn, J. W. & Superti-Furga, G. Recent advances in combinatorial drug screening and synergy scoring. *Curr. Opin. Pharmacol.* **42**, 102–110 (2018).
2. Palmer, A. C. & Sorger, P. K. Combination Cancer Therapy Can Confer Benefit via Patient-to-Patient Variability without Drug Additivity or Synergy. *Cell* **171**, 1678–1691.e13 (2017).
3. Meyer, C. T. *et al.* Quantifying Drug Combination Synergy along Potency and Efficacy Axes. *Cell Syst* **8**, 97–108.e16 (2019).
4. Keiser, M. J. *et al.* Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* **25**, 197–206 (2007).
5. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–7 (2012).
6. Butina, D. Unsupervised Data Base Clustering Based on Daylight's Fingerprint and Tanimoto Similarity: A Fast and Automated Way To Cluster Small and Large Data Sets. *J. Chem. Inf. Comput. Sci.* **39**, 747–750 (1999).
7. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002).
8. Zhang, Z. & Pan, Z. Z. Signaling cascades for δ-opioid receptor-mediated inhibition of GABA synaptic transmission and behavioral antinociception. *Mol. Pharmacol.* **81**, 375–383 (2012).
9. Saxena, N. C. Inhibition of GABA(A) receptor (GABAR) currents by arachidonic acid in HEK 293 cells stably transfected with alpha1beta2gamma2 GABAR subunits. *Pflugers Arch.* **440**, 380–392 (2000).

Chapter 2:

Zebrafish behavioral profiling identifies ligands, targets, and neurons related to sedation and paradoxical excitation

Matthew N. Mc Carroll[1†], Leo Gendelev[1†], Reid Kinser[1], Jack Taylor[1], Giancarlo Bruni[2], Douglas Myers-Turnbull[1], Cole Helsell[1], Amanda Carbajal[3], Capria Rinaldi[1], Hye Jin Kang[4], Jung Ho Gong[5] Jason K. Sello[5], Susumu Tomita[6], Randall T. Peterson[2,7], Michael J. Keiser[1,8*,] David Kokel[1,9*]

[1] Institute for Neurodegenerative Diseases, University of California, San Francisco, California 94143, USA

[2] Cardiovascular Research Center and Division of Cardiology, Department of Medicine, Massachusetts General Hospital and Department of Systems Biology, Harvard Medical School, 149 13th Street, Charlestown, Massachusetts, 02129, USA

[3] Department of Biology, San Francisco State University, San Francisco, CA, USA

[4] Department of Pharmacology and NIMH Psychoactive Drug Screening Program, University of North Carolina Chapel Hill Medical School, Chapel Hill, NC 27759, USA

[5] Department of Chemistry, Brown University, Providence, RI, 02912

[6] Department of Cellular and Molecular Physiology, Department of Neuroscience, Yale University School of Medicine, New Haven, CT 06510, USA

[7] Current address: Department of Pharmacology and Toxicology, College of Pharmacy, University of Utah, Salt Lake City, UT, 84112, USA. randall.peterson@pharm.utah.edu

[8] Departments of Pharmaceutical Chemistry and of Bioengineering & Therapeutic Sciences, University of California, San Francisco, California 94158, USA

[9] Department of Physiology, University of California, San Francisco, California 94158, USA

† Co-First Author

* Corresponding author. E-mail: david.kokel@ucsf.edu, keiser@keiserlab.org

2.1 Abstract

Anesthetics are generally associated with sedation, but some anesthetics can also increase brain and motor activity — a phenomenon known as paradoxical excitation. Previous studies have identified GABAA receptors as the primary targets of most anesthetic drugs, but how these compounds produce paradoxical excitation is poorly understood. To identify and understand such compounds, we applied a behavior-based drug profiling approach. Here, we show that a subset of central nervous system depressants cause paradoxical excitation in zebrafish. Using this behavior as a readout, we screened thousands of compounds and identified dozens of hits that caused paradoxical excitation. Many hit compounds modulated human GABAA receptors, while others appeared to modulate different neuronal targets, including the human serotonin-6 receptor. Ligands at these receptors generally decreased neuronal activity, but paradoxically increased activity in the caudal hindbrain. Together, these studies identify ligands, targets, and neurons affecting sedation and paradoxical excitation in vivo in zebrafish.

2.2 Introduction

Anesthetics and other central nervous system (CNS) depressants primarily suppress neural activity, but sometimes they also cause paradoxical excitation [1]. During paradoxical excitation, brain activity increases [2,3] and produces clinical features such as confusion, anxiety, aggression, suicidal behavior, seizures, and aggravated rage [4,5]. These symptoms primarily affect small but vulnerable patient populations including psychiatric, pediatric, and elderly patients [6,7]. Understanding paradoxical excitation is important for discovering, understanding, and developing CNS depressants and for understanding how small molecules affect the vertebrate nervous system. However, relatively few

compounds have been identified that cause paradoxical excitation, and few model systems have been identified that reproducibly model paradoxical excitation *in vivo*.

Many ligands that cause paradoxical excitation are agonists or positive allosteric modulators (PAMs) of $GABA_A$ receptors ($GABA_ARs$), the major type of inhibitory receptors in the CNS [8]. However, it is likely that other mechanisms also affect paradoxical excitation. One such mechanism may involve serotonin imbalance, which affects behavioral disinhibition [9,10], and has paradoxical effects on neuronal circuit output [11]. For example, the serotonin-6 receptor (HTR6) is an excitatory G protein-coupled receptor (GPCR) reported to modulate cholinergic and glutamatergic systems by disinhibiting GABAergic neurons [12]. In serotonin syndrome, excessive serotonergic signaling causes agitation, convulsions, and muscle rigidity. Despite these excitatory effects of excessive serotonergic signaling, several serotonin receptor agonists are used as anxiolytics, hypnotics, and anticonvulsants [13]. Examples include clemizole and fenfluramine, which promote 5-HT signaling and have anticonvulsant properties in humans and zebrafish [13,14]. By contrast, serotonin antagonists and inverse agonists improve sleep and are used for treating insomnia [15]. Furthermore, serotonin receptors are secondary and tertiary targets of some anesthetics, suggesting that 5-HT receptors may contribute to sedation [16]. However, the potential impact of serotonin receptors on anesthesia and paradoxical excitation is poorly understood.

In principle, large-scale behavior-based chemical screens would be a powerful way to identify compounds that cause sedation and paradoxical excitation. The reason is that phenotype-based screens are not restricted to predefined target-based assays. Rather, phenotype-based screens can be used to identify targets and pathways that produce poorly understood phenotypes. Indeed, virtually all CNS and anesthetic drug prototypes were originally discovered based on their behavioral effects before their targets were known [17]. Furthermore, these compounds were valuable tools for understanding the mechanisms of anesthesia and sedation. Although behavior-based chemical

screens in vertebrates would be most relevant for human biology, behavioral assays in mice, primates, and other mammals are difficult to scale.

Zebrafish are uniquely well-suited for studying the chemical biology of sedation and paradoxical excitation. Zebrafish are vertebrate animals with complex brains and behaviors, they are small enough to fit in 96-well plates, and they readily absorb compounds dissolved in the fish water. Compared with humans, zebrafish share many conserved genes and neurotransmitter signaling pathways [18]. For example, the zebrafish genome contains orthologs for all but two human $GABA_AR$ subunit isoforms [19]. The α-isoform family is the largest and most diverse family of $GABA_AR$ subunits in both humans and zebrafish [20]. The zebrafish genome also encodes orthologs of serotonin receptors, including orthologs of HTR6 [21,22]. Additionally, there are several important differences between the species. One such difference is that the zebrafish genome encodes a $GABA_AR$ β_4-subunit, which does not have a clear ortholog in mammals [19]. Another difference is that whereas mammals have six $GABA_AR$ α-subunit isoforms, zebrafish have eight [20]. Previously, drug profiling studies in zebrafish have identified neuroactive compounds related to antipsychotics, fear, sleep, and learning [23–26]. However, specific behavioral profiles for compounds that cause paradoxical excitation have not been previously described in zebrafish.

The purpose of these studies is to identify and understand compounds that cause paradoxical excitation. First, we develop a scalable model of paradoxical excitation in zebrafish. Then, we use this model in large-scale chemical screens to identify dozens of compounds that cause paradoxical excitation. Third, we use these compounds as research tools to identify receptors affecting sedation and paradoxical excitation. Finally, we map whole-brain activity patterns during these behavioral states. Together, these studies improve our understanding of how small molecules cause sedation and paradoxical excitation and may help to accelerate the pace of CNS drug discovery.

2.3 Results

I. $GABA_AR$ ligands produce paradoxical excitation in zebrafish

To determine how sedatives affect zebrafish behavior, we assembled a set of 27 CNS depressants in ten functional classes (Fig. 1a, Supplementary Table 1) and tested these compounds in a battery of automated behavioral assays. The behavioral assays were originally devised to discriminate between a broad range of neuroactive compounds[23]. Here, the assays were used to profile anesthetics and other CNS-depressants. In one assay, we used excitatory violet light stimuli to identify compounds that reduce motor activity (Fig. 1b, Supplementary Movie 1). In another assay, we used low-volume acoustic stimuli to identify compounds that enhance startle sensitivity (Fig. 1b, Supplementary Movie 2). Most CNS depressants caused a dose-dependent decrease in animals' average motion index (MI) (Supplementary Figure 1), however we noticed a striking exception.

Two anesthetic $GABA_AR$ ligands, etomidate and propofol, caused enhanced acoustic startle responses (eASRs). These eASRs occurred in response to a specific low-volume acoustic stimulus, but not to other stimuli (Fig. 1a-b, Supplementary Figure 2, 3). Unlike vehicle-treated controls, all the animals in a well treated with etomidate showed robust eASRs (Supplementary Movie 3, 4). High speed video revealed that the eASRs resembled short latency C-bends (Fig. 1c). Multiple eASRs could be elicited with multiple stimuli (Fig. 1d). Etomidate's half maximal effective concentration (EC_{50}) for causing eASRs was 1 μM, consistent with its EC_{50} against $GABA_ARs$ *in vitro* (Fig. 1e) [27]. Neither etomidate or propofol were toxic at the concentrations that

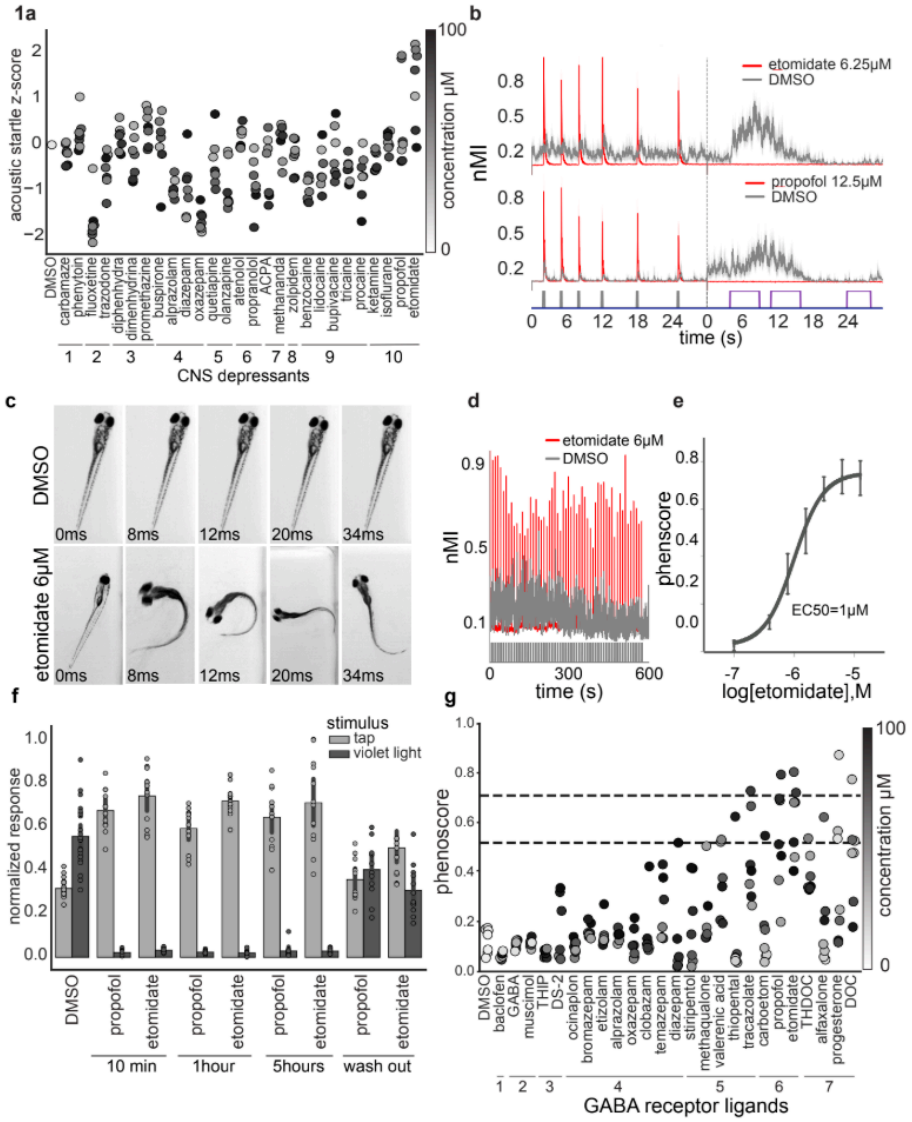


Fig 2.1 GABA$_A$R PAMs enhance acoustic startle in zebrafish.

Zebrafish were treated with the indicated compounds and analyzed for changes in behavioral responses. (a) The scatter plot quantifies acoustic startle response as a z-score (y-axis) in zebrafish treated with the indicated CNS- depressants (x-axis) at the indicated concentrations (colorbar). Each point represents the average of n = 12 wells and 6 experimental replicates (also listed in Supplementary Table 1). (b) These plots show how the indicated compounds impact zebrafish motor activity (y-axis) over time (x-axis) (n = 12 wells, shaded boundary = 95% confidence interval). Colored bars above the x-axis represent the timing and duration of low-volume acoustic stimuli (grey bars) and violet light stimuli (purple bars). The vertical dotted line indicates where the first assay ends and the second begins. (c)
Representative images of animals treated with the indicated compounds. Time stamps indicate the
time elapsed from the initial presentation of a low-volume acoustic stimulus. (d) These plots compare the motor activity (y-axis) over time (x-axis) of animals treated with DMSO (grey) or etomidate (red) (n= 50 larvae). Consecutive stimuli (n = 60) are indicated by vertical grey bars. (e) Dose-response curve showing phenoscores at the indicated concentrations (each point represents n =12 wells/dose, error bars: ± SD). (f) Bar plot showing normalized response to the indicated stimulus (tap or violet light) of animals treated with

DMSO, 6μM propofol, or 6μM etomidate (n = 12 wells, error bars: ± SD) for the indicated durations. (g) Average phenoscores (y-axis) of zebrafish treated with the indicated compounds. Dashed lines intersecting the y-axis at 0.51 and 0.71 correspond respectively to 1% and 5% significance cutoffs, as determined from statistical simulations. Compounds are grouped by ligand class: 1) $GABA_AR$ agonist, 2) $GABA_AR$ orthosteric agonist, 3) PAM of δ-subunit containing $GABA_ARs$, 4) $GABA_AR$ BZ-site PAM, 5) $GABA_AR$ R non-BZ-site PAM, 6) $GABA_AR$ neurosteroid PAM, 7) $GABA_AR$ anesthetic PAM.

caused eASRs (Supplementary Table 2).The eASRs persisted for several hours and rapidly reversed after drug washout (Fig. 1b, f). Curiously, not all anesthetics caused eASR behaviors in zebrafish, including isoflurane (a volatile inhalational anesthetic that is relatively toxic in zebrafish), dexmedetomidine (a veterinary anesthetic and alpha-adrenergic agonist), and tricaine (a local anesthetic and sodium channel blocker) (Fig. 1a). Together, these data suggest that a subset of $GABA_AR$ ligands produce sedation and paradoxical excitation in zebrafish.

To determine if other $GABA_AR$ ligands caused similar phenotypes, we used the phenoscore metric to quantify similarities between the archetypal profile caused by etomidate (6.5 uM) and a diverse range of GABAergic compounds (Supplementary Table 3). Average phenoscores of DMSO-treated negative controls were significantly less than etomidate-treated positive controls (0.2 versus 0.71, $P < 10^{-10}$, Kolmogorov-Smirnov test) (Supplementary Figure 20). Average phenoscores for the test compounds fell on a continuum between the positive and negative controls (Fig. 1g). Based on statistical simulations, these phenoscores were subdivided into three categories: weak, intermediate, and strong (Supplementary Note 1). Compounds with the strongest phenoscores ($0.71 < x < 1$) included several anesthetic and neurosteroid PAMs including etomidate, propofol, progesterone, and DOC (Fig. 1g). The highest scoring treatments for these compounds produced behavioral profiles that were both strongly sedating and produced high-magnitude eASRs (Supplementary Figure 2). These profiles were not statistically different from the positive controls ($P > 0.05$, Kolmogorov-Smirnov test, Supplementary Figure 20). Together, these data suggest that a subset of $GABA_AR$ PAMs cause sedation and paradoxical excitation in zebrafish. However, due to the

overlapping pharmacology of numerous $GABA_AR$ subtypes, these data do not clearly point to any specific subset of receptor subtypes as being necessary or sufficient for these behaviors.

In humans, the M-current is a low-threshold, non-inactivating, voltage-dependent current that limits repetitive action potentials and has been implicated in propofol-induced paradoxical excitation [28,29]. To determine if M-currents affect eASRs in zebrafish, we tested several M-current activators and inhibitors. In animals treated with M-current activators (flupirtine [30,31] and ICA-069673 [32]), eASR magnitudes significantly decreased (Supplementary Figure 6, $P < 0.01$, two-tailed *t*-test, $n = 6$ wells; 8 fish/well). By contrast, in animals treated with M-current inhibitors (linopirdine, XE-991, and oxotremorine [33]) eASR magnitudes significantly increased (Supplementary Figure 6, $P < 0.000001$, two-tailed *t*-test, $n = 6$ wells; 8 fish/well). These data suggest that zebrafish eASRs are a form of paradoxical excitation affected by potassium channel M-currents.

II. Large-scale behavioral screening identifies hit compounds

To prepare for large-scale screening, we calculated phenoscores for hundreds of positive and negative control wells (treated with etomidate or DMSO, respectively). The average phenoscores of the positive controls were significantly greater than the negative controls (0.7, ± 0.11 SD versus 0.1 ± 0.05 SD), suggesting that a large-scale screen would have an expected false positive and negative rate of 2% and 0.4%, respectively (at a threshold of 3 SD) (Fig. 2a, Z-factor = 0.7, $n = 944$ wells). Then, we screened a library of 9,512 structurally-diverse compounds plus 2,336 DMSO-treated negative controls, and analyzed their effects on sedation and paradoxical excitation. Visualized as a contour plot, the highest density of phenoscores occurred in three major regions (Fig. 2b). The first region contained 11,679 compounds and DMSO-treated control wells that did not phenocopy etomidate. The second region contained 44 potentially toxic compounds that immobilized zebrafish

but did not cause paradoxical excitation. The third region contained 125 compounds that both produced immobilization and phenocopied etomidate and were considered primary screening hits (Supplementary Table 4, Supplementary Figure 7).

To organize these hit compounds by structural similarity, we clustered them based on Tanimoto similarities and visualized the results as a dendrogram that contained 14 clusters (Fig. 2c). Several clusters included compounds previously associated with $GABA_A$Rs (Fig. 2c, d). For example, Cluster 10 included several dihydro/quinazolinones that are structurally-related to methaqualone, a sedative hypnotic drug (Fig. 2c, d). A second cluster, Cluster 14, included several isoflavonoids, which are structurally-related to flavonoid sedatives [34]. Overall, we selected a broad range of 57 primary hit compounds across all the clusters to re-purchase and re-test (Supplementary Table 4). Each compound was re-tested in a dose-response format from 0.1 to 100 μM and scored based on its ability to immobilize zebrafish and increase eASRs. Together, 81% of these primary hit compounds (46/57) caused reproducible eASR phenotypes at one or more concentrations (Fig. 2e, Supplementary Figure 8, and Supplementary Table 4), indicating a high rate of reproducibility from the primary screen.

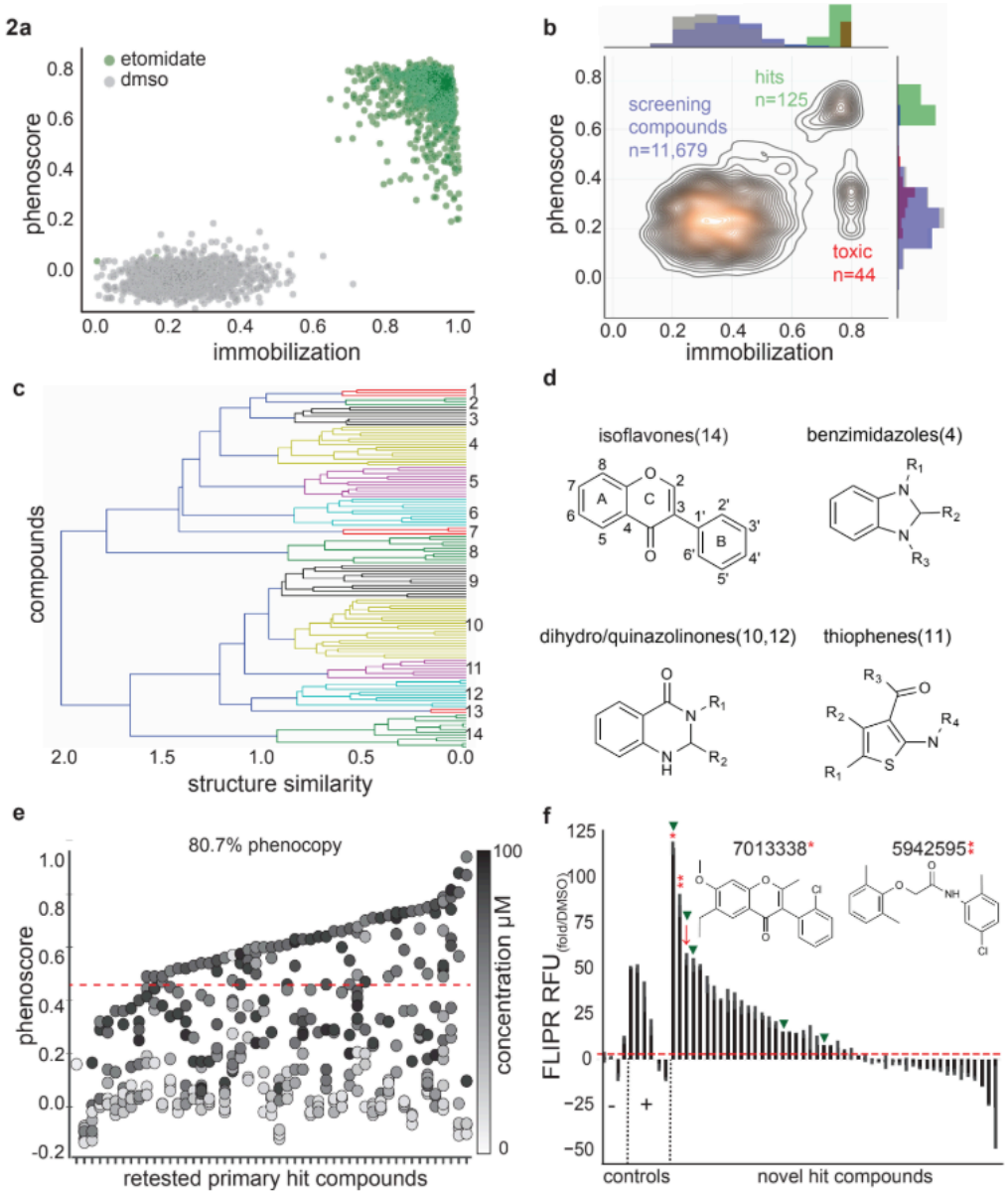


Fig 2.2 A high-throughput behavioral screen identifies GABAergic compounds..

Zebrafish were treated with various compounds and analyzed for anesthetic-related behaviors. (a) This scatter plot compares phenoscores of individual wells treated with DMSO or etomidate (6.25 µM) (Z-factor= 0.7, n = 944 wells). (b) This contour plot scores each well from the large-scale behavior-based chemical screen (11,679 compounds, 2,336 DMSO controls) by its phenoscore (y-axis) and immobilization index (x-axis). Labels indicate regions with 125 hit compounds (green), 44 toxic compounds (red), and the remaining screening compounds and DMSO controls (blue and gray respectively). (c) Structural clustering of the top 125 hit compounds (y-axis) forms 14 clusters using a Tanimoto similarity metric (x-axis). (d) Example structures of selected compounds in the indicated clusters. (e) This scatter plot shows an 80.7% reproducibility rate for 57 primary hit compounds. Each point represents the average phenoscore of n = 12 wells at the indicated concentrations (colorbar).
The first column represents DMSO controls; the order of other compounds are listed in
Supplementary Table 4. (f) Human $GABA_AR$ activation (y-axis) was measured by FLIPR analysis. Of 47 hit compounds, 23 potentiated $GABA_ARs$. Compounds

7013338 and 5942595 potentiated $GABA_AR$s significantly greater than positive controls (red asterisk = 7013338, two red asterisks = 5942595, P < 0.0001, two-tailed t-test, n = 2-4 replicates as indicated). The hit threshold was defined as 2x the average DMSO control group. Picrotoxin, BGC 20-761, progesterone, and DMSO were used as negative controls (x-axis) while etomidate, tracazolate, propofol, diazepam, and thiopental 24 were used as positive controls. Arrows indicate compounds that were predicted by SEA to bind $GABA_AR$s (red arrows) and compounds that bound to TSPO in vitro (green arrowheads).

To determine if these compounds targeted human $GABA_AR$s, we tested them in a fluorescent imaging plate reader (FLIPR) assay on HEK293 cells transfected with $\alpha_1\beta_2$ and $\alpha_1\beta_2\gamma_2$ human $GABA_AR$ subtypes. In this cell line, etomidate, tracazolate, and propofol increased fluorescence in the presence of GABA, as expected for $GABA_AR$ PAMs. In addition, half of the tested hit compounds (23/46) also showed PAM activity (Fig. 2f, Supplementary Figure 9). By contrast, PAM activity was not observed with negative control compounds including BGC 20-761 (an HTR6 antagonist) and PTX (a $GABA_AR$ channel blocker) which likely reduced $GABA_AR$ activity due to inhibition of constitutively active $GABA_AR$s in the system. Interestingly, the PAM activity of two hit compounds, 7013338 and 5942595, was significantly greater than the positive controls (Fig. 2f, $P < 0.0001$, two-tailed *t*-test, $n = 4$). While some of the compounds appeared to function in this assay as negative allosteric modulators (NAMs), reductions in fluorescence were likely due to toxicity-induced cell loss (Fig. 2f, Supplementary Table 5). In addition to the PAM assay, four hit compounds directly activated $GABA_AR$s in the absence of GABA, including 5860357, 6091285, 5835629, and 7284610 (Supplementary Figure 9). The strongest direct activator, 5835629, did not further enhance $GABA_AR$ activation in the PAM assay, presumably because the cells were already maximally activated by the compound before GABA was added. These data suggest that behavioral screens in zebrafish can enrich for compounds with activity at specific human receptors. In addition, these data suggest that many of the hit compounds identified in the screen cause sedation and paradoxical excitation via $GABA_AR$s.

III. Hit compounds act on targets including GABAAR and HTR6

To determine if any of the hit compounds acted on non-$GABA_AR$ targets, we used the Similarity Ensemble Approach (SEA)[35] to predict targets based on 'guilt-by-association' enrichment factor scores (EFs). Among the top-ranked 1,000 screening compounds, 150 targets were enriched (Supplementary Table 6). As we analyzed subsets of hit compounds with increasing phenotypic stringency, the number of enriched targets decreased (Fig. 3a, b). The most stringent set of 30 top-ranked hit compounds contained 25 enriched targets including $GABA_ARs$, 5α-reductase, mGluRs, and 5-HTRs (Supplementary Table 7). By contrast, this stringent set of hit compounds were not enriched for other sporadically predicted targets such as histone deacetylase, matrix metalloproteinase, and carbonic anhydrase (Fig. 3b). As additional negative controls, we tested 48 reference compounds targeting receptors with relatively low EF scores and confirmed they did not cause eASR phenotypes at any concentration (Supplementary Table 8, Supplementary Figure 10). Together, these data suggest that the hit compounds may act on $GABA_ARs$, 5α-reductase, mGluRs, or 5-HTRs (Supplementary Note 2).

A second approach to target identification was to test the binding affinity of hit compounds against a panel of 43 human and rodent neurotransmitter-related targets. Of 46 hit compounds, 33 of them bound to at least one of 19 receptors at a $K_i < 10\mu M$ (Fig. 3c, d). Several hit

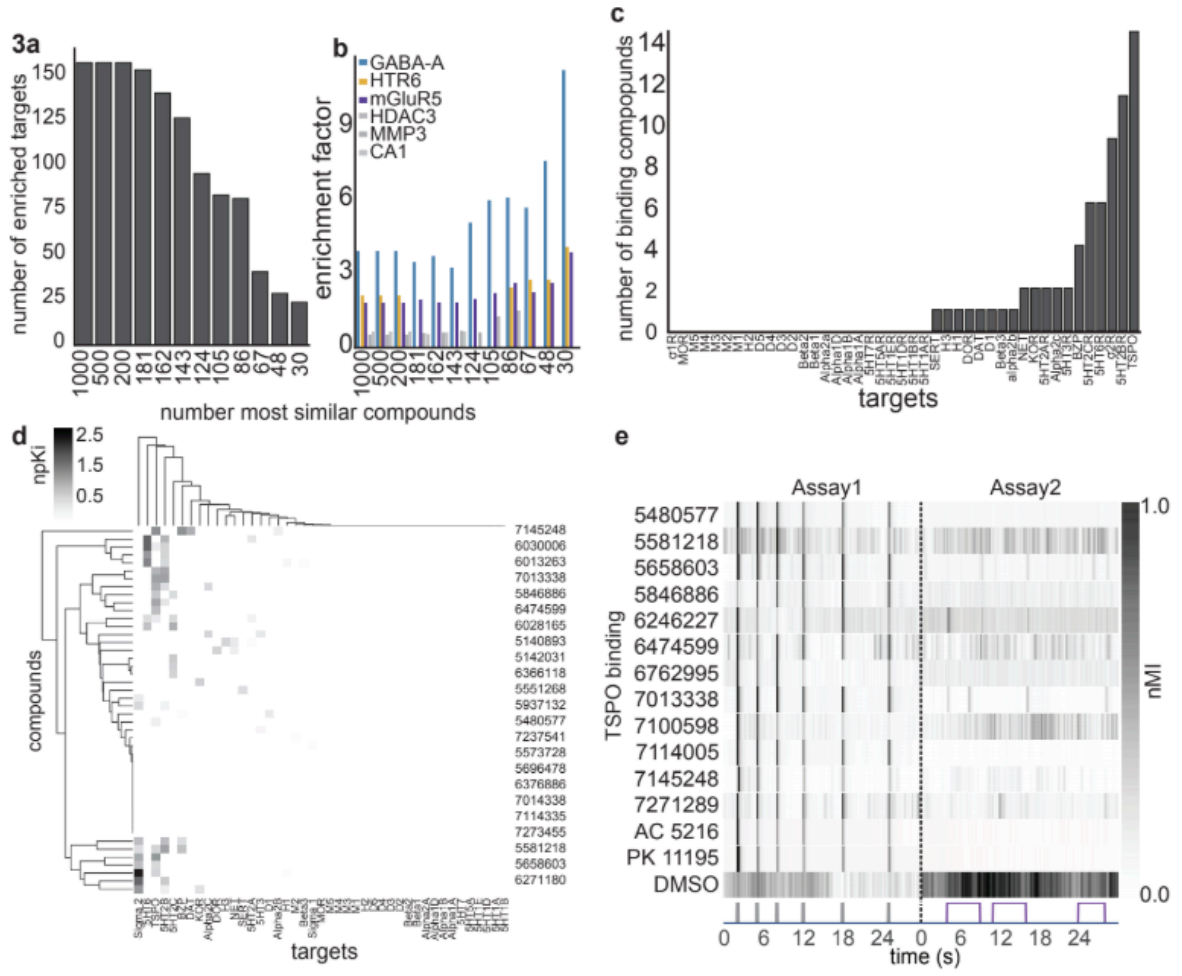


Fig 2.3 Potential targets include GABA$_A$R, mGluR, TSPO, and HTR6.

(a) SEA analysis was performed on decreasing numbers of hit compounds (1,000-30). The bar plot shows the number of SEA enriched targets decreasing as the analysis focuses on the top 30 hit compounds. (b) The bar plot shows increasing enrichment of GABA$_A$R, HTR6, and mGluR5 as the top targets predicted for the top 30 hit compounds. (c) This bar plot shows the number of 46 primary hit compounds (y-axis) that bound to the indicated CNS receptors (x-axis). (d) The clustergram shows binding affinity profiles at the indicated CNS receptors. The colorbar indicates normalized Ki (npKI) (e) Heatmap of average motor activity profiles for TSPO binding compounds (y-axis) over time (x-axis) (n = 12 wells). Assay 1 is comprised of 6 low amplitude acoustic stimuli (grey); Assay 2 is a series of 3 violet light pulses (violet). These two assays are separated by a dotted line. AC 5216 and PK 11195 are TSPO binding compounds. Abbreviations: nMI, normalized motion index; MMP3, Matrix Metallopeptidase 3; CA1, Carbonic anhydrase 1; HDAC3, Histone Deacetylase 3.

compounds bound to multiple targets, including compound 7145248, which bound to TSPO, the benzodiazepine receptor (BZP), dopamine transporter, and the alpha 2b receptor (Fig. 3d). The

most common targets (binding > 5 compounds) included BZP, sigma 2, HTR2A/B/C, HTR6, and TSPO (Fig 3c, d). TSPO, previously known as the peripheral benzodiazepine receptor (PBR) (because it was originally identified as a binding site for benzodiazepine anxiolytic drugs), is a mitochondrial protein that supplies cholesterol to steroid-producing enzymes in the brain [36]. TSPO ligands are thought to enhance GABAergic signaling by increasing neurosteroid production. However, some TSPO ligands, including benzodiazepines and zolpidem, also bind to $GABA_A$Rs directly [37]. We found that 14 hit compounds bound to TSPO *in vitro* (Fig. 3c, d), and that two TSPO reference ligands, PK 11195 and AC 5216, both caused eASRs *in vivo* (Fig. 3e). Of the 14 compounds that bound to TSPO *in vitro*, 5 compounds potentiated $GABA_A$R in FLIPR assays (Fig. 3c; 2f, green arrowheads). These data suggest that TSPO ligands promote anesthetic-related phenotypes via direct interactions with $GABA_A$Rs, indirect effects on neurosteroidogenesis, or both.

Both target identification approaches, SEA and the *in vitro* receptor binding assays, implicated HTR6. For example, SEA predicted that seven hit compounds, six benzenesulfonamides and one piperazine, targeted HTR6 (Fig. 4a, Supplementary Table 9). These compounds reproducibly caused eASRs *in vivo* (Fig. 4b). *In vitro*, six of these hit compounds bound to HTR6 at nanomolar concentrations (Ki = 54-807 nM) (Fig. 4c). To determine their functional effects, we tested them for agonist and antagonist activity in G-protein and β-arrestin pathways at eight serotonin receptor subtypes (1A, 2A, 2B, 2C, 4, 5A, 6, and 7A). Six of the compounds antagonized HTR6 *in vitro*. Most of them antagonized both G-protein and β-arrestin pathways,

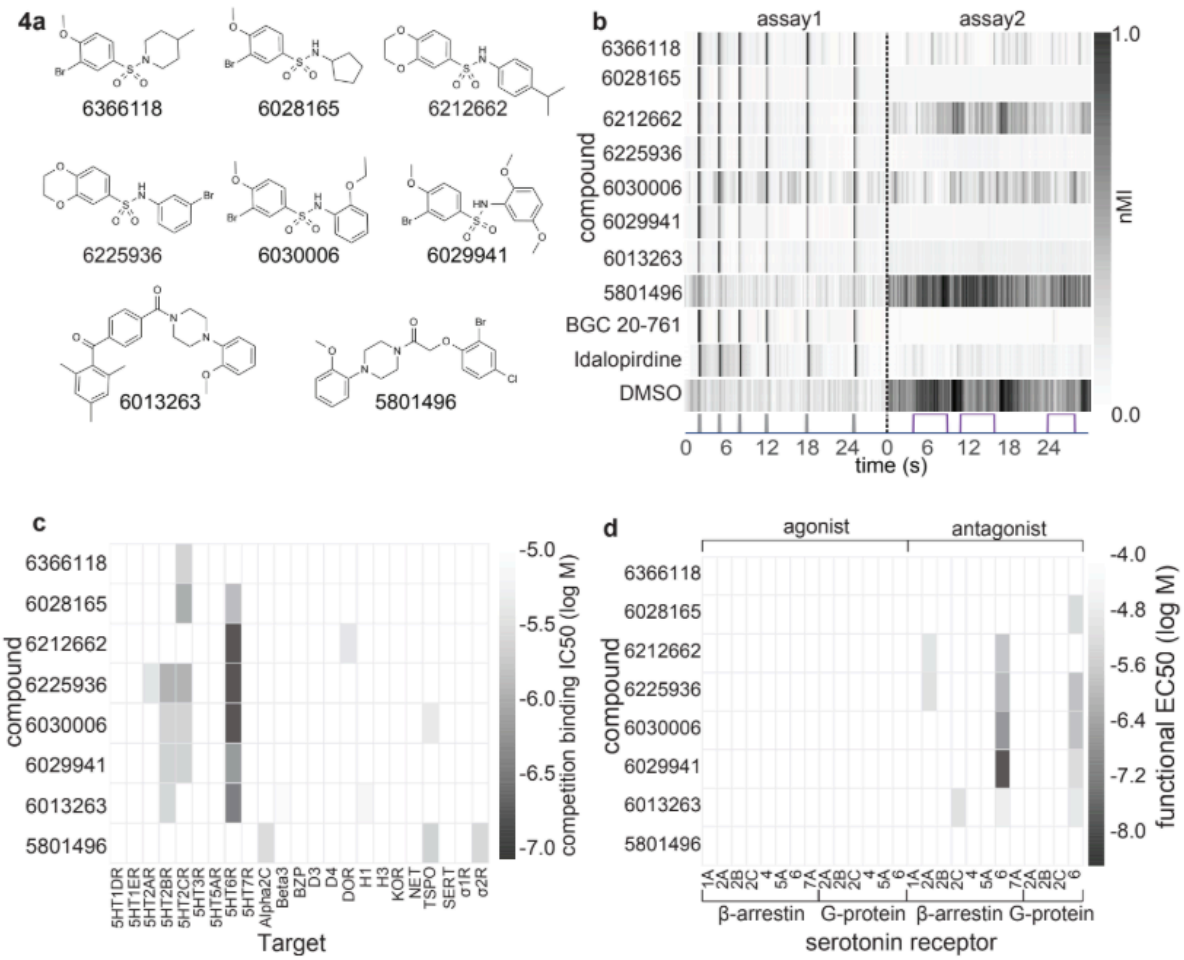


Fig 2.4 A subset of hit compounds are HTR6 antagonists.

(a) Structures of eight primary hit compounds predicted to bind HTR6. (b) Heatmap showing average (n = 12) motor activity profiles over time (x-axis) for compounds predicted to bind HTR6 (y-axis). Assay 1 is comprised of 6 low- amplitude acoustic stimuli; Assay 2 is a series of 3 violet light pulses (as indicated), these two assays are separated by a dotted line. BGC 20-761 and Idalopirdine are previously annotated HTR6 antagonists. (c) Heatmap showing binding affinities of primary hit compounds at 23 CNS receptors (x-axis). (d) Heatmap showing functional activity of primary hit compounds at the indicated GPCRs. (nMI, normalized motion index).

suggesting that the compounds were unbiased HTR6 antagonists (activity range 3.30nM-18.2μM) (Fig. 4d). By contrast, a structurally related piperazine, 5801496, did not cause eASRs *in vivo*. To determine if any annotated HTR6 antagonists also caused eASRs, we analyzed six structurally-diverse HTR6 reference antagonists. Two of these reference antagonists, BGC 20-761 and idalopirdine, reproducibly caused eASRs *in vivo* (Fig. 4b). It is unclear why only 2/6 reference HTR6

antagonists caused eASRs in zebrafish, but it may be related to issues with absorption, metabolic stability, and/or structural differences between human and zebrafish receptors. A panel of 36 additional 5-HT modulating ligands at various serotonergic targets did not cause eASRs at any concentration tested (Supplementary Table 9). Together, these data suggest that a subset of HTR6 antagonists cause eASRs in zebrafish.

IV. A neural substrate for paradoxical excitation

To determine which regions of the brain were active during eASRs, we visualized whole-brain activity patterns by pERK labeling [38]. In DMSO-treated control animals, pERK labeling showed broad patterns of activity in the optic tectum, telencephalon, and other brain regions (Fig. 5a, b). By contrast, in animals treated with etomidate or propofol, pERK labeling was broadly reduced (Fig. 5c-e; $P < 0.0005$, Mann-Whitney U test). Acoustic stimuli significantly activated a cluster of neurons in the caudal hindbrain at the base of the 4th ventricle near the auditory brainstem and the nucleus of the solitary tract (NST)[39] at the level of the area postrema (Fig. 5f, g; $P < 0.0005$, Mann-Whitney U test) [40], suggesting that this hindbrain neuron cluster represented a specific substrate of eASR behavior.

To determine if activity in this region specifically occurred during eASRs, we analyzed pERK labeling in this region during four other robust motor behaviors. First, in animals stimulated by optovin (a reversible photoactivatable TRPA1 ligand),[24] neuronal activity increased in many brain regions including the telencephalon and optic tectum but not in the hindbrain (Fig. 5j, k). Second, in DMSO-treated control animals stimulated by light, neuronal activity increased in the

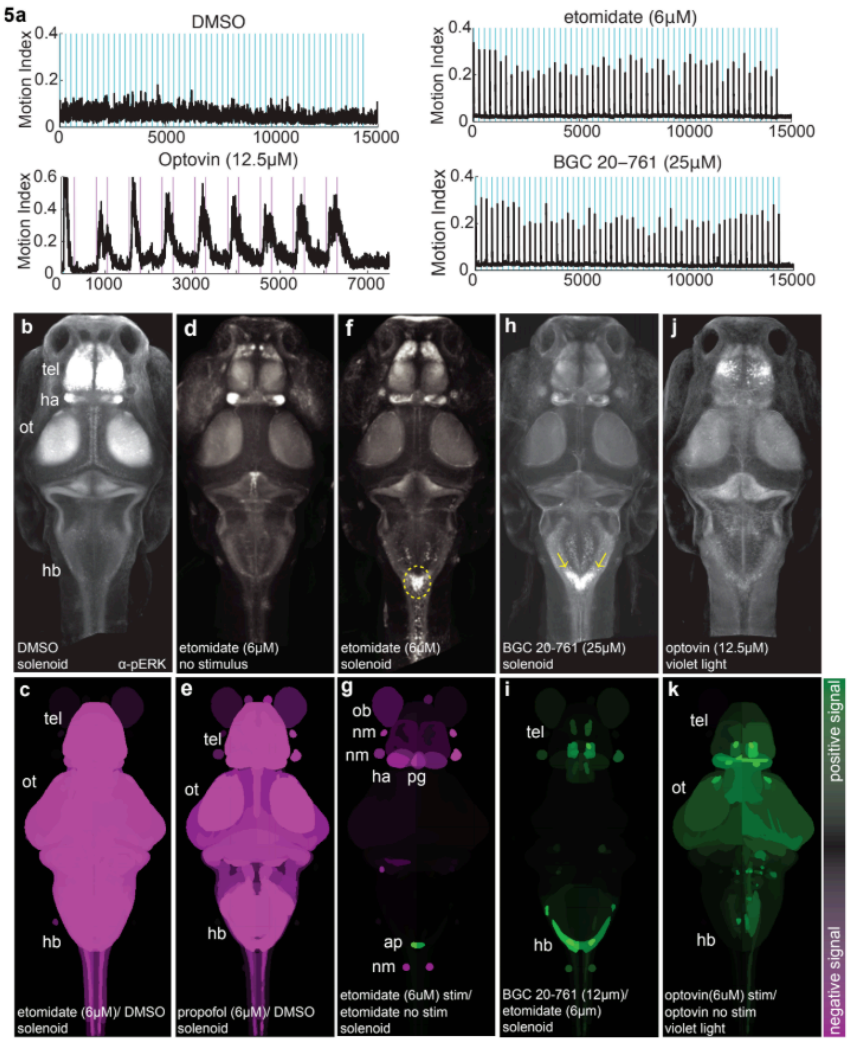


Fig 2.5 Hit compounds activate hindbrain neurons.

Animals were exposed to the indicated drugs and stimuli and analyzed for pERK levels as a readout of neuronal activity. (a) Plots showing motor activity (y-axis) over time (x-axis) for animals treated with the indicated compounds (n = 25-50 larvae) in response to the indicated acoustic (blue) or violet light (purple) stimuli. (b, d, f, h, j) Confocal projections showing the average fluorescent intensity of image registered larval brains stained with α-pERK (n = 10 larvae/condition). Larvae were treated with the indicated compounds and exposed to the low amplitude acoustic stimulus once every 10 seconds for 10 minutes, except for (b, no stimulus) and (f, violet light exposure). (c, e, g, i, k) Brain activity maps showing significant ΔpERK signals using the Z-brain online reference tool (n = 5-10 animals/condition). The heatmap indicates positive (green), negative (purple), and nonsignificant (black)
changes in pERK labeling ($P < 0.0005$, Mann-Whitney U test). All activity maps are comparisons between the indicated treatment conditions. Abbreviations: tel, telencephalon; ot, optic tectum; hb, hindbrain; ob, olfactory bulb; nm, neuromast;
ap, area postrema; pg, pineal gland.

V. Hit compounds produce distinct side effect profiles

To prioritize hit compounds for further development, we tested them for specific side effects. For example, a serious side effect of etomidate is that it suppresses corticosteroid synthesis due to off-target activity on 11β-hydroxylase, the enzyme that synthesizes cortisol in humans and zebrafish. To determine if any of the hit compounds suffered from similar liabilities, we measured their effects on cortisol levels. As a positive control, we used carboetomidate, a close structural analog of etomidate that was rationally designed to retain etomidate's activity on $GABA_AR$s, while disrupting its ability to suppress cortisol synthesis. Both etomidate and carboetomidate immobilized zebrafish and caused eASRs (Fig. 1g). As expected, etomidate lowered cortisol levels in zebrafish, whereas carboetomidate did not, suggesting that these compounds have similar side effects in humans and zebrafish (Fig. 6a). Next, we tested 12 hit compounds in the cortisol assay, including eight $GABA_AR$ ligands (found to be positive in the FLIPR assay), one compound predicted to target $GABA_AR$ by SEA (5951201), two HTR6 antagonists (6225936 and 6029941), and one mysterious compound with no target leads (5736224). None of these compounds reduced cortisol levels in zebrafish (Fig. 6a), indicating that these ligands cause sedation and paradoxical excitation without suppressing cortisol levels.

To determine if any of the hit compounds may be analgesic, we used optovin-induced motor activity as a potential analgesia-related assay. In humans, general anesthetics reduce perceptions of pain and suffering. Although it is unclear if fish feel pain, painful stimuli in humans also cause behavioral responses in zebrafish. For example, activation of the TRPA1 ion channel causes pain in humans [41], and optovin, a photoinducible TRPA1 ligand, induces strong behavioral responses in zebrafish [24]. As a positive control, we found that etomidate suppressed the optovin response at the same concentrations that caused eASRs (Fig. 6b). Similarly, we found that two GABAergic

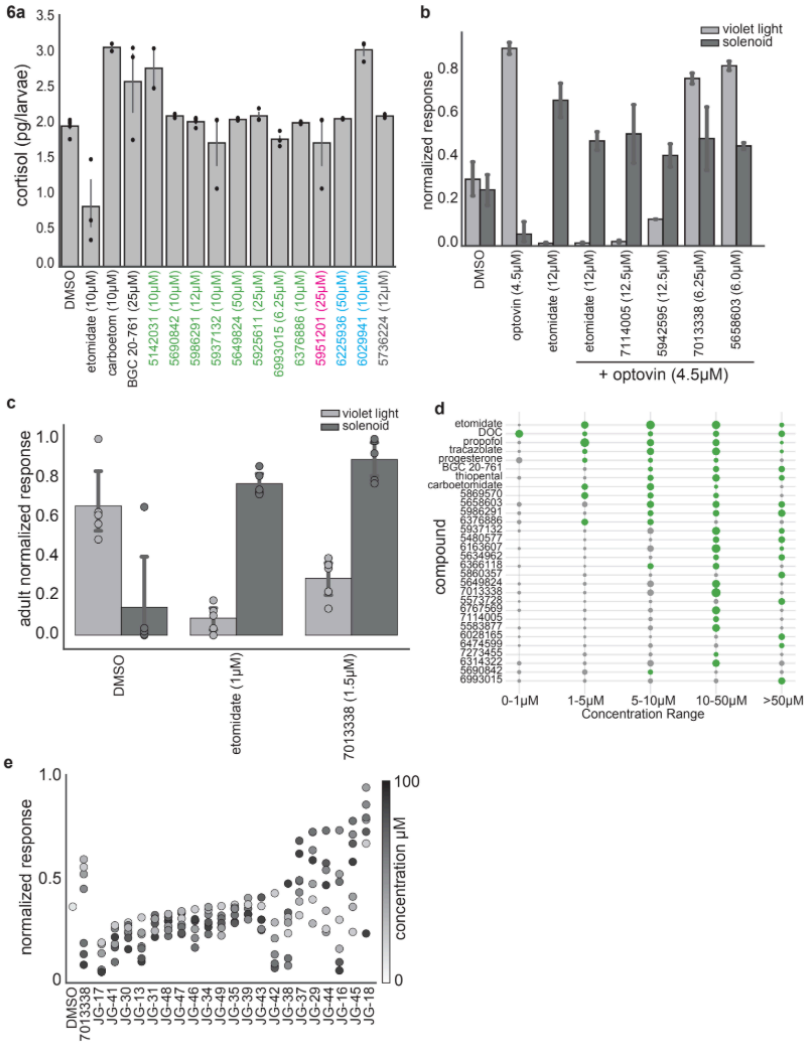


Fig 2.6 Hit compounds show diverse efficacy windows and side effect profiles.

(a) This bar plot shows cortisol levels (y-axis) in animals treated with the indicated compounds (x-axis) including FLIPR-positive GABAergics (green), SEA predicted GABAergics (magenta), serotonergics (blue), and a compound with undetermined targets grey) (n = 2-5 experiments, 15 animals/experiment, error bars: ± SEM). (b) This bar plot shows the normalized responses (y-axis) of animals treated with the indicated compounds (x-axis) in the pain-related optovin-response suppression assay. (c) This bar plot shows the magnitude of behavioral responses of adult zebrafish (y-axis) treated with with the 25 indicated compounds (x-axis). (d) Dot plot showing efficacy windows for the indicated compounds with strong (green) or weak (grey) phenocopy scores. Marker size represents the magnitude of the eASR response (n = 12 wells/condition). Compounds with broad efficacy windows have large green dots at multiple concentrations (x-axis). (e) This strip plot shows the normalized acoustic startle response (y-axis) of larvae treated with increasing concentrations (colorbar) of multiple analogs of the screening hit 7013338 (x-axis) (n = 4-6 wells/condition, 8 fish/well).

hit compounds, 7114005 and 5942595, also blocked optovin-induced motor activity at the same concentrations that caused eASRs (Fig. 6b). By contrast, compounds 5658603 and 7013338 did not

suppress the optovin response (Fig. 6b). The HTR6 antagonist BGC 20-761 also blocked the optovin response (Supplementary Figure 16), however serotonergic hit compounds 6225936, 6028165, and 6212662 only reduced the optovin response at concentrations that also reduced eASRs (Supplementary Figure 16). Together, these data suggest that the mechanisms controlling sedation and eASRs may be separable from analgesia, and that some eASR-causing compounds may cause analgesic-related effects in zebrafish.

To determine if eASRs also occur in adult zebrafish, we treated adult animals with etomidate and the hit compound 7013338, the most effective hit compound in the FLIPR assays (Fig. 2f). We found that both of these compounds also worked in adult animals, reducing the violet light response, while increasing acoustic startle (Fig. 6c). These data suggest that the mechanisms underlying eASR phenotypes are not limited to larvae but also exist in adult zebrafish.

In humans, therapeutic windows for many inhalational anesthetics are only 2-fold, while therapeutic indices for intravenous anesthetics are not much better [42]. Many of the hit compounds also had relatively narrow efficacy windows (Fig. 6d). Numerous analogs of key hit compounds including thiophenes, aryloxycarboxamides, quinazolines, and sulfonamides had lower activity than the original hits (Supplementary Figure 17), suggesting that substantial medicinal chemistry would be needed to increase the potency of the primary hit compounds.

Compound 7013338 activated human $GABA_A$Rs more than any other hit compound in the FLIPR assay (Fig. 2f). However, its efficacy window was relatively narrow (10 to 50 μM), raising questions about its structure activity relationship (SAR) (Fig. 6e). To analyze its SAR, we generated 21 analogs with different substituents on the A-, B-, and C-rings (Fig. 2d, Supplementary Figure 18, and Supplementary Table 10), and tested these analogs for biological activity *in vivo*. The most active analog, JG-18, increased eASR magnitude and widened the efficacy window from 1 to 50 μM (Fig. 6e). It had a chloro substituent on C2' of the B-ring, an ethyl substituent on C2 of the C-ring, and C-

6 propyl and C7 hydroxyl substituents on the A-ring (Supplementary Figure 18). In congruence with previous SAR analyses of isoflavones [34], JG-18 and other analogs with more lipophilic substituents on position C6 of the A-ring and position C2 of C-ring exhibited increased biological activity (Fig. 6e). By contrast, analogs with more dramatic enhancements in steric bulk and lipophilicity at these positions (i.e., phenethyl and propyl, respectively) exhibited reduced biological activity. Likewise, it was notable that capping the polar C7-hydroxy group of JG-18 with alkyl and acyl groups tended to lessen biological activity. Importantly, we found that the B-ring C2' chloro substituent was absolutely critical for biological activity, since analogs without it did not cause eASRs. Previously, it was reported that analogs with alkoxy or trifluoromethoxy substituents at multiple positions but especially at C3' on the B-ring were high affinity $GABA_AR$ binders *in vitro* [43]. Surprisingly, compound JG-17 (with a trifluoromethoxy substituent on C3' of the B-ring, an ethyl substituent on C2 of C-ring, and C6 propyl and C7 hydroxy substituents on the A-ring) had no biological activity *in vivo* (Fig. 6e). It is not clear why these ligands were not active in zebrafish. Perhaps, the anomaly could be ascribed to low penetrance *in vivo*, receptor subtype selectivity, and/or structural differences between the human and zebrafish $GABA_ARs$. Together, these data suggest that additional SAR analyses may yield analogs with greater potency and broader efficacy windows *in vivo*.

2.3 Discussion

These studies show that anesthetics and other $GABA_AR$ PAMs cause sedation and paradoxical excitation in zebrafish, and that this behavioral model has high predictive and construct validity for identifying modulators of human $GABA_ARs$. Indeed, these studies may have underestimated the number of hit compounds that targeted $GABA_ARs$ for several reasons. One reason is that the *in vitro* $GABA_AR$ FLIPR assay only tested a very small number of receptor subtypes and subunit isoforms ($\alpha_1\beta_2$ and $\alpha_1\beta_2\gamma_2$). As a result, these assays would have missed compounds that acted on other

$GABA_AR$ subtypes. A second reason is that some of the hit compounds may act on zebrafish-specific $GABA_ARs$. Finally, some hit compounds that caused eASRs in zebrafish may need to be bioactivated *in vivo*, and would therefore not be be active *in vitro*. Therefore, even more of the hit compounds may have targeted $GABA_ARs$.

These studies also suggest that non-$GABA_AR$ mechanisms may also affect paradoxical excitation, including HTR6 antagonism. For example, we found that HTR6 antagonists produced sedation and paradoxical excitation in zebrafish (Fig 4). These HTR6 antagonists likely reduce neuronal excitation via different mechanisms than $GABA_AR$ PAMs. $GABA_ARs$ are widely distributed in the CNS, suggesting that GABA ligands likely inhibit most neurons directly. By contrast, HTR6s are restricted to discrete neuronal populations [44], suggesting that their effects are likely propagated through indirect signaling networks. HTR6 antagonists can reduce 5-HT neuronal firing[45], presumably by blocking positive feedback control of raphe neurons that broadly project throughout the brain and spinal cord [21]. Researchers have made remarkable progress applying the principles of systems pharmacology to structure-based target predictions[47], computer assisted design of multi-target ligands[48-50], and the large-scale prediction of beneficial drug combinations[10,51]. Although we focused on predicting targets of compounds one at a time, in future studies it may be possible to calculate multi-target enrichment factors among the hit compounds from large-scale phenocopy screens and identify multi-target mechanisms.

The HTR6 antagonists identified in this study add to a growing body of evidence implicating various serotonin ligands and receptors in phenotypes related to neuronal inhibition and excitation. Our finding that HTR6 antagonists activate a region in the zebrafish NST (Fig. 5h, i), are consistent with previous work showing that HTR6 antagonists activate neurons in the mammalian NST [46]. In rodents, HTR6 antagonists promote sleep [47], reduce anxiety [48], and show anticonvulsant properties [49]. However, it is not clear if HTR6 causes these effects via specific neuronal circuits, or more generally

by coordinating nervous system tone and arousal. Furthermore, there are substantial differences in the central nervous system distribution and pharmacology of the mouse, rat, and human HTR6 receptors [50]. So, although HTR6 antagonists phenocopied etomidate in zebrafish, these effects may not translate to anesthetic activity in humans. Despite promising effects in rodents, several HTR6 antagonists failed in clinical trials as cognitive enhancers for the treatment of Alzheimer's disease [51], underscoring the caveats of generalizing between humans and model organisms.

These data suggest at least two possible models by which $GABA_AR$ PAMs could cause paradoxical excitation of the acoustic startle response. One possibility is that the ligands disinhibit the acoustic startle neurons. Alternatively, the ligands may excite specific neurons directly, due to conditions that reverse the chloride equilibrium potential, such as the tonic activation of $GABA_ARs$ [52]. Our observation that caudal hindbrain neurons were activated by acoustic stimuli in etomidate-treated zebrafish is not the first to link GABA signaling to auditory excitation. For example, in rodents, gaboxadol activates extrasynaptic $GABA_ARs$, hyperpolarizes resting membrane potential, and converts neurons in the auditory thalamus to burst mode [53]. In addition, etomidate causes purposeless muscle excitement that is exacerbated by acoustic stimuli in dogs [54]. In zebrafish, researchers have found that the offset of optogenetic-induced inhibition of caudal hindbrain neurons triggers swim responses [55]. In addition, zebrafish caudal hindbrain neurons have been shown to be activated during hunting behaviors, a behavior that requires strong inhibitory control [56]. However, exactly how these neurons impact motor activity, and why startle neurons remain active despite elevated inhibitory tone, remains unclear.

Although these studies show that $GABA_AR$ PAMs cause paradoxical excitation, pharmacological experiments to determine which $GABA_AR$ subtypes caused eASRs were ultimately inconclusive. While the majority of $GABA_ARs$ in the CNS are benzodiazepine-sensitive γ-containing subtypes, and multiple benzodiazepines did not cause strong eASRs (Fig. 1g), γ-containing subtypes

may still be very important for eASRs. One reason is that the benzodiazepines tested in this study only represent a very small subset of benzodiazepine analogs. Another reason is that diazepam produced intermediate eASR phenotypes (Fig. 1g, Supplementary Figure 2), suggesting that other benzodiazepines may cause even stronger eASR phenotypes. Although etomidate, propofol, neurosteroids, and other anesthetics are PAMs at δ-subunit containing $GABA_AR$ subtypes, these ligands also modulate γ-containing subtypes. Furthermore, although THIP and DS2 are reported to have preferential activity at δ-containing $GABA_ARs$, these compounds also modulate γ-containing receptors [57] , and they did not cause eASRs. One alternative explanation is that β-isoforms [58,59] could drive the presence or absence of eASRs. Another possible explanation is that whereas PAMs may produce immobilizing effects effects via some receptor subtypes, they may produce eASRs via other subtypes. In summary, although a subset of $GABA_AR$ PAMs caused eASRs, these compounds may do so via a variety of receptor subtypes. In future studies, it would be interesting to test additional benzodiazepines for such effects including midazolam, which causes paradoxical excitation in humans [60]. The specificity of currently available pharmacological tools may be insufficient to determine which $GABA_AR$ subtypes produce eASRs. Therefore, future studies may require targeted knockouts and other genetic tools to help identify the key receptor subtypes.

While these studies focused on behavioral profiling, other types of phenotypic profiling data may further improve the accuracy of neuroactive compound classification, including whole-brain imaging. Whole-brain imaging allows researchers to record real-time firing patterns will likely add massive amounts data to the behavioral pharmacology field [61,62]. For example, recent advances in high-throughput brain activity mapping for systems neuropharmacology illustrate how whole brain activity mapping can be used in primary screening for compounds that activate specific circuits, or allow researchers to discriminate between compounds with similar behavioral phenotypes but that work on different neuronal populations [62]. These approaches enable primary screening for

compounds that activate specific circuits and allow researchers to discriminate between compounds with similar behavioral phenotypes but that work on different neuronal populations.

In summary, we have shown that $GABA_AR$ PAMs cause sedation and paradoxical excitation in zebrafish. Whereas previous behavior-based chemical screens in zebrafish have identified neuroactive compounds related to behaviors including sleep [25], antipsychotics [23], learning [26], and appetite [63], we show here that behavioral profiling can also be used to rapidly identify compounds related to sedation and paradoxical excitation. Future studies will likely expand the utility of behavior-based chemical phenocopy screens to additional kinds of neuroactive ligands, targets, and pathways.

2.4 Methods

Fish maintenance, breeding, and compound treatments. We collected a large number of fertilized eggs (up to 10,000 embryos per day) from group matings of wild-type zebrafish (from Singapore). All embryos were raised on a 14/10-hour light/dark cycle at 28˚C until 7 dpf. Larvae were distributed 8 animals per well into square 96-well plates (GE Healthcare Life Sciences) with 300 μL of egg water. Compound stock solutions were applied directly to the egg water and larvae were incubated at room temp for 1 hour before behavioral analysis. To determine the impact of group size on this assay, we analyzed eASR behaviors from animals in different group sizes (1, 2, 4, 8, 12, 16, 32 animals per well). Although animals in all groups responded similarly to the stimulus (Supplementary Movie 5), the largest differences between treated and controls were seen in groups of 8 and 16 animals (Supplementary Figure 19). We therefore used 8 animals to balance small group size with a strong MI signal. All zebrafish procedures were and approved by the UCSF's Institutional Animal Care Use Committee (IACUC), and in accordance with the Guide to Care and

Use of Laboratory Animals (National Institutes of Health 1996) and conducted according to established protocols that complied with ethical regulations for animal testing and research.

Compounds and chemical libraries. All chemical libraries were dissolved in DMSO. The Chembridge library (Chembridge Corporation) contains 10,000 compounds at 1 mM. The Prestwick library (Prestwick Chemical) contains 1,280 approved drugs at 10 mM. All compounds were diluted in E3 buffer and screened at 10 μM final concentration in < 1% DMSO. Controls were treated with an equal volume of DMSO. Compounds were validated in 3-12 replicate wells, on 3 replicate plates. For dose response behavioral assays, compounds were tested at 7 concentrations that ranged from 0.1- 100μM, unless otherwise indicated.

Automated behavioral phenotyping assays. Digital video was captured at 25 frames per second using an AVT Pike digital camera (Allied Vision). Each assay duration was 30-120 seconds, and consisted of a combination of acoustic and light stimuli [23]. Low (70db) and high (100db) amplitude acoustic stimuli were delivered using push-style solenoids (12V) to tap a custom built acrylic stage where the 96-well plate was placed. Acoustic stimuli were recorded using a contact microphone (*Aquarian Audio Products*, model# H2a) and the freeware audio recording software Audacity (http://www.audacityteam.org). Digital acoustic stimulus was generated as a 70ms sine wave at various frequencies using Audacity. A computer was used to playback the audio stimulus as an mp3 using an APA150 150W powered amplifier (Dayton Audio) played through surface transducers adhered to the acrylic stage. Stimulus volume was measured using a BAFX 3608 digital sound level meter (BAFX Products). Light stimuli were delivered using high intensity LEDs (LEDENGIN) at violet (400nm, 11 $\mu W/mm^2$), blue (560nm, 18 $\mu W/mm^2$), and red (650nm, 11 $\mu W/mm^2$) wavelengths. Stimuli and digital recordings were applied to the entire 96-well plate simultaneously.

Instrument control and data acquisition were performed using custom scripts written in MATLAB and Python. The zebrafish motion index (MI) was calculated as follows: MI = sum(abs($frame_n$ – $frame_{n-1}$)). Normalized MI (nMI) was calculated as follows: nMI=(MI-min(MI))/max(MI). Startle magnitude was calculated using numerical integration via the trapezoidal method (Matlab function *trapz*) of MI values during stimulus.

Computing the phenoscore. To quantify distances between multi-dimensional behavioral profiles, we first defined a prototypic behavioral profile to compare everything else against. Etomidate's prototypical behavioral profile was determined from 36 replicates wells treated with etomidate (6.25 μM) on 3 different plates (12 replicates per plate). Using a simulated annealing procedure (described in the supplement) we identified 12 replicate profiles with the most consistent eASR response that was also most distant from the control (DMSO) wells. The reference profile was the average of these 12 profiles. Phenoscore distances were computed between each well and the reference profile by calculating the correlation distance (using the correlation distance module from the scipy package in python). The correlation distance (phenoscore) has a range from -1 to +1. Positive and negative values represent positive and negative correlation, respectively. Negative values represent anti-correlation. Experimentally, phenoscores tended to saturate at around 0.7, a value that represents substantial positive correlation given that the MI time series is a large vector with >10,000 values. Although etomidate and propofol are both anesthetic $GABA_AR$ PAMs with similar behavioral profiles in zebrafish, etomidate is more soluble than propofol, so we used etomidate as the archetypal positive control.

Ranking the screening hits. Phenoscores were computed to assign each compound in the screening library a rank order. Hit-compounds were defined as the top 125 scoring compounds from this ranked list

Calculating response magnitude Z-scores. Response magnitudes were calculated by averaging the maximum motion index value during 3 repeated violet stimuli or 6 repeated acoustic stimuli. These Motion index magnitudes were converted to Z-scores using the following equation: Z-score = (magnitude - mean)/*SD*. Z-scores were then normalized from 0-1 using the scikit function sklearn.preprocessing.normalize written for python.

***In vitro* receptor profiling**. *In vitro* binding assays and Ki data were generated by the National Institute of Mental Health's Psychoactive Drug Screening Program (PDSP), contract no. HHSN-271-2008-00025-C (NIMH PDSP), for assay details: http://pdsp.med.unc.edu/PDSP%20Protocols%20II%202013-03-28.pdf. Normalized Ki (npKi) values were generated as follows: npKi = 4 + (-log10 (Ki)) [64].

FLIPR. We used the FLIPR system (Molecular Devices) to quantify GABA-evoked activity of human GABAARs. We chose a membrane potential dye (Molecular Devices) to measure changes in membrane potentials and stably transfected HEK293 cells that expressed α_1, β_2 and γ_2. Since we observed an increase in GABA-evoked responses when transfected with γ_2 transiently, we describe the cells as having a low level of γ-subunit expression, indicating heterogeneity of $GABA_AR$ compositions in the cell ($\alpha_1\beta_2$ or $\alpha_1\beta_2\gamma_2$). To assay for direct agonists, fluorescence was subtracted pre- and post compound addition. To assay for PAMs, cells were treated with compound at 20 uM and then with 5 uM GABA.

Whole-brain activity mapping. Following behavioral experiments animals were immediately fixed in 4% paraformaldehyde in PBS and incubated overnight at 4˚C. Larvae were then washed in PBS + 0.25% Triton-X (PBT), incubated for 15 minutes at 70˚C in 150mM Tris-HCl, pH9, washed in PBT, pearmeablized in 0.05% Trypsin-EDTA for 30-45 minutes on ice and washed with PBT. Animals were then blocked for 1 hour at room temperature (PBT, 1% bovine serum albumin, 2% normal goat serum, 1% DMSO, and 0.02% sodium azide) [38]. The following primary antibodies were diluted into blocking buffer and incubated overnight at 4˚C: α-5HT (1:500, ImmunoStar), α-tERK (1:750, Cell Signaling), α-pERK (1:750, Cell Signaling). Secondary fluorescent antibodies (Life Technologies) were used at 1:500 and incubated in blocking buffer overnight at 4˚C in the dark. Whole-mount fluorescent images were obtained using a Leica SP8 confocal microscope. Image processing was performed in imageJ. Image registration was performed using the Computational Morphometry Toolkit (https://www.nitrc.org/projects/cmtk) and a user interface with the command string defined by Owen Randlett (-awr 010203 -T 8 -X 52 -C 8 -G 80 -R 3 -A '--accuracy 0.4' -W '--accuracy 1.6'). Multiple brains from each condition were then averaged using Matlab scripts to obtain a representative neural activity image. Brightness and contrast were adjusted using Fiji (imageJ). MAP-map calculations (whole brain ΔpERK significance heat maps) were performed using the analysis code for MAP-map which can be downloaded from the website (http://engertlab.fas.harvard.edu/Z-Brain/).

Cortisol detection assay. Cortisol levels were measured in zebrafish [65]. Briefly, 15, 7 day old larvae were treated with the indicated compounds for 1 hour. Larvae were anesthetized in ice-cold egg water and then snap-froze in an ethanol/dry ice bath. Larvae were then homogenized in 100μL of water. Cortisol was extracted from the homogenate with 1 mL of ethyl acetate, the resulting

supernatant was collected and the solvent allowed to evaporate. Cortisol was dissolved in 0.2% bovine serum albumin (A7030, Sigma) and frozen at -20°C. For cortisol ELISA experiments, 96-well plates (VWR International) were treated with cortisol antibody (P01-92-94M-P, EastCoast Bio; 1.6 g/mL in PBS) for 16 hours at 4°C, washed, and blocked with 0.1% BSA in PBS. Cortisol samples and cortisol-HRP (P91-92-91H, EastCoast Bio) were incubated at room temperature for 2 hours and washed extensively with PBS containing 0.05% Tween-20 (Invitrogen). Detection of HRP was performed using tetramethylbenzidine (TMB: 22166-1, Biomol) and Tetrabutylammonium borohydride (TBABH: 230170-10G, Sigma). Reaction was stopped using 1M H_2SO_4. Absorbance was read at 450 nm in an ELISA plate reader (SpectraMax MS, Molecular Devices).

Data Availability

The source data for all Figures and Supplementary Figures in the current study are available in the Zenodo repository, https://zenodo.org, DOI10.5281/zenodo.3336616

Software and code availability

Data acquisition and analysis were carried out using custom scripts in Matlab (MathWorks) and Python. Figures were prepared using Matlab, Python, ImageJ (NIH), Prism (GraphPad), and Adobe Illustrator. Scripts used for data acquisition and analysis are available from the corresponding author upon reasonable request. For Enrichment Factor calculations, code is available upon request.

2.5 Acknowledgements

We thank members of our research groups for helpful advice and critical reading of the manuscript. Dr. Owen Randlett for his assistance with the pERK immunostaining and whole brain imaging protocol. This work was supported by US National Institutes of Health (NIH) grants: R01AA022583 (DK), the Paul G. Allen Family Foundation (DK, MK), R01MH115705 (ST), and U01MH104984 (ST).

I. Contributions

MNM characterized the phenotype, designed and performed most of the experiments including behavioral assays, compound testing, cortisol detection, imaging, and immunohistochemistry and analyzed the data. LG analyzed the data and predicted targets using SEA. RK designed, collected, and analyzed the data related to most of the $GABA_AR$ reference set. JT, AC, and CR collected behavioral data. GB ran the initial compound screen with support from RTP. DMT and CH wrote code to collect and analyze the data. HJK performed the target binding assays and analyzed the data. ST performed the FLIPR assays. JHG and JKS synthesized the isoflavones. MJK and DK designed the experiments and wrote the paper with input from all the authors. All authors reviewed and edited the manuscript.

II. Competing Interests

The authors declare no competing financial interests.

III. Correspondence

Please address all correspondence to david.kokel@ucsf.edu or keiser@keiserlab.org

2.6 References

1 Jeong, S. *et al.* Increase of paradoxical excitement response during propofol-induced sedation in hazardous and harmful alcohol drinkers. *Br. J. Anaesth.* **107**, 930–933 (2011).

2. McCarthy, M. M., Brown, E. N. & Kopell, N. Potential network mechanisms mediating electroencephalographic beta rhythm changes during propofol-induced paradoxical excitation. *J. Neurosci.* **28**, 13488–13504 (2008).

3. Sneyd, J. R. Excitatory events associated with propofol anaesthesia: a review. *J. R. Soc. Med.* **85**, 288–291 (1992).

4. Fulton, S. A. & Mullen, K. D. Completion of upper endoscopic procedures despite paradoxical reaction to midazolam: a role for flumazenil? *Am. J. Gastroenterol.* **95**, 809–811 (2000).

5. Zuleta-Alarcon, A. *et al.* Anesthesia-related perioperative seizures: pathophysiology, predisposing factors and practical recommendations. *Austin Journal of Anesthesia and Analgesia* **2**, 1026 (2014).

6. Mancuso, C. E., Tanzi, M. G. & Gabay, M. Paradoxical reactions to benzodiazepines: literature review and treatment options. *Pharmacotherapy* **24**, 1177–1185 (2004).

7. Massanari, M., Novitsky, J. & Reinstein, L. J. Paradoxical reactions in children associated with midazolam use during endoscopy. *Clin. Pediatr.* **36**, 681–684 (1997).

8. Orser, B. A., Canning, K. J. & Macdonald, J. F. Mechanisms of general anesthesia. *Curr. Opin. Anaesthesiol.* **15**, 427–433 (2002).

9. Sachs, B. D. *et al.* The effects of brain serotonin deficiency on behavioural disinhibition and anxiety-like behaviour following mild early life stress. *Int. J. Neuropsychopharmacol.* **16**, 2081–2094 (2013).

10. Pothakos, K. *et al.* Decreased serotonin levels associated with behavioral disinhibition in tissue plasminogen activator deficient (tPA-/-) mice. *Brain Res.* **1326**, 135–142 (2010).

11. Fickbohm, D. J. & Katz, P. S. Paradoxical actions of the serotonin precursor 5-hydroxytryptophan on the activity of identified serotonergic neurons in a simple motor circuit. *J. Neurosci.* **20**, 1622–1634 (2000).

12. Yun, H.-M. & Rhim, H. The serotonin-6 receptor as a novel therapeutic target. *Exp. Neurobiol.* **20**, 159–168 (2011).

13. Griffin, A. *et al.* Clemizole and modulators of serotonin signalling suppress seizures in Dravet syndrome. *Brain* **140**, 669–683 (2017).

14. Dinday, M. T. & Baraban, S. C. Large-Scale Phenotype-Based Antiepileptic Drug Screening in a Zebrafish Model of Dravet Syndrome. *eNeuro* **2**, (2015).

15. Monti, J. M. Serotonin 5-HT(2A) receptor antagonists in the treatment of insomnia: present status and future prospects. *Drugs Today* **46**, 183–193 (2010).

16. Matsunaga, F. *et al.* Molecular interactions between general anesthetics and the 5HT2B receptor. *J. Biomol. Struct. Dyn.* **33**, 211–218 (2015).

17. James, R. & Glen, J. B. Synthesis, biological evaluation, and preliminary structure-activity considerations of a series of alkylphenols as intravenous anesthetic agents. *J. Med. Chem.* **23**, 1350–1357 (1980).

18. Mathur, P. & Guo, S. Use of zebrafish as a model to understand mechanisms of addiction and complex neurobehavioral phenotypes. *Neurobiol. Dis.* **40**, 66–72 (2010).

19. Cocco, A. *et al.* Characterization of the γ-aminobutyric acid signaling system in the zebrafish (Danio rerio Hamilton) central nervous system by reverse transcription-quantitative polymerase chain reaction. *Neuroscience* **343**, 300–321 (2017).

20. Monesson-Olson, B. *et al.* Expression of the eight GABAA receptor α subunits in the developing zebrafish central nervous system. *PLoS One* **13**, e0196083 (2018).

21. Panula, P. *et al.* The comparative neuroanatomy and neurochemistry of zebrafish CNS systems of relevance to human neuropsychiatric diseases. *Neurobiol. Dis.* **40**, 46–57 (2010).

22. Norton, W. H. J., Folchert, A. & Bally-Cuif, L. Comparative analysis of serotonin receptor (HTR1A/HTR1B families) and transporter (slc6a4a/b) gene expression in the zebrafish brain. *J. Comp. Neurol.* **511**, 521–542 (2008).

23. Bruni, G. *et al.* Zebrafish behavioral profiling identifies multitarget antipsychotic-like compounds. *Nat. Chem. Biol.* **12**, 559–566 (2016).

24. Kokel, D. *et al.* Photochemical activation of TRPA1 channels in neurons and animals. *Nat. Chem. Biol.* **9**, 257–263 (2013).

25. Rihel, J. *et al.* Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation. *Science* **327**, 348–351 (2010).

26. Wolman, M. A., Jain, R. A., Liss, L. & Granato, M. Chemical modulation of memory formation in larval zebrafish. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15468–15473 (2011).

27. Hill-Venning, C., Belelli, D., Peters, J. A. & Lambert, J. J. Subunit-dependent interaction of the general anaesthetic etomidate with the gamma-aminobutyric acid type A receptor. *Br. J. Pharmacol.* **120**, 749–756 (1997).

28. McCarthy, M. M., Brown, E. N. & Kopell, N. Potential network mechanisms mediating electroencephalographic beta rhythm changes during propofol-induced paradoxical excitation. *J. Neurosci.* **28**, 13488–13504 (2008).

29. Kornhuber, J. *et al.* [Neuronal potassium channel opening with flupirtine]. *Fortschr. Neurol. Psychiatr.* **67**, 466–475 (1999).

30. Popovici, F., Dorostkar, M. & Boehm, S. The non-opioid analgesic flupirtine is a modulator of GABAA receptors involved in pain sensation. *BMC Pharmacol.* **8**, A14 (2008).

31. Devulder, J. Flupirtine in pain management: pharmacological properties and clinical use. *CNS Drugs* **24**, 867–881 (2010).

32. Amato, G. *et al.* N-Pyridyl and Pyrimidine Benzamides as KCNQ2/Q3 Potassium Channel Openers for the Treatment of Epilepsy. *ACS Med. Chem. Lett.* **2**, 481–484 (2011).

33. Pérez, C., Limón, A., Vega, R. & Soto, E. The muscarinic inhibition of the potassium M-current modulates the action-potential discharge in the vestibular primary-afferent neurons of the rat. *Neuroscience* **158**, 1662–1674 (2009).

34. Luk, K. C., Stern, L., Weigele, M., O'Brien, R. A. & Spirt, N. Isolation and identification of 'diazepam-like' compounds from bovine urine. *J. Nat. Prod.* **46**, 852–861 (1983).

35. Lounkine, E. *et al.* Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **486**, 361–367 (2012).

36. Papadopoulos, V., Lecanu, L., Brown, R. C., Han, Z. & Yao, Z.-X. Peripheral-type benzodiazepine receptor in neurosteroid biosynthesis, neuropathology and neurological disorders. *Neuroscience* **138**, 749–756 (2006).

37. Papadopoulos, V., Mukhin, A. G., Costa, E. & Krueger, K. E. The peripheral-type benzodiazepine receptor is functionally linked to Leydig cell steroidogenesis. *J. Biol. Chem.* **265**, 3772–3779 (1990).

38. Randlett, O. *et al.* Whole-brain activity mapping onto a zebrafish brain atlas. *Nat. Methods* **12**, 1039–1046 (2015).

39. Biacabe, B., Chevallier, J. M., Avan, P. & Bonfils, P. Functional anatomy of auditory brainstem nuclei: application to the anatomical basis of brainstem auditory evoked potentials. *Auris Nasus Larynx* **28**, 85–94 (2001).

40. Longatti, P. *et al.* The human area postrema: clear-cut silhouette and variations shown in vivo. *J. Neurosurg.* **122**, 989–995 (2015).

41. Kremeyer, B. *et al.* A gain-of-function mutation in TRPA1 causes familial episodic pain syndrome. *Neuron* **66**, 671–680 (2010).

42. Kissin, I., McGee, T. & Smith, L. R. The indices of potency for intravenous anaesthetics. *Can. Anaesth. Soc. J.* **28**, 585–590 (1981).

43. Gavande, N., Karim, N., Johnston, G. A. R., Hanrahan, J. R. & Chebib, M. Identification of Benzopyran-4-one Derivatives (Isoflavones) as Positive Modulators of GABAA Receptors. *ChemMedChem* **6**, 1340–1346 (2011).

44. Ward, R. P. *et al.* Localization of serotonin subtype 6 receptor messenger RNA in the rat brain by in situ hybridization histochemistry. *Neuroscience* **64**, 1105–1111 (1995).

45. Brouard, J. T. *et al.* Pharmacological Evidence for 5-HT6 Receptor Modulation of 5-HT Neuron Firing in Vivo. *ACS Chem. Neurosci.* **6**, 1241–1247 (2015).

46. Garfield, A. S., Burke, L. K., Shaw, J., Evans, M. L. & Heisler, L. K. Distribution of cells responsive to 5-HT6 receptor antagonist-induced hypophagia. *Behav. Brain Res.* **266**, 201–206 (2014).

47. Morairty, S. R., Hedley, L., Flores, J., Martin, R. & Kilduff, T. S. Selective 5HT2A and 5HT6 receptor antagonists promote sleep in rats. *Sleep* **31**, 34–44 (2008).

48. Wesołowska, A. & Nikiforuk, A. Effects of the brain-penetrant and selective 5-HT6 receptor antagonist SB-399885 in animal models of anxiety and depression. *Neuropharmacology* **52**, 1274–1283 (2007).

49. Routledge, C. *et al.* Characterization of SB-271046: A potent, selective and orally active 5-HT6 receptor antagonist. *Br. J. Pharmacol.* **130**, 1606–1612 (2000).

50. Setola, V. & Roth, B. L. Why mice are neither miniature humans nor small rats: a cautionary tale involving 5-hydroxytryptamine-6 serotonin receptor species variants. *Mol. Pharmacol.* **64**, 1277–1278 (2003).

51. Atri, A. *et al.* Effect of Idalopirdine as Adjunct to Cholinesterase Inhibitors on Change in Cognition in Patients With Alzheimer Disease: Three Randomized Clinical Trials. *JAMA* **319**, 130–142 (2018).

52. Gulledge, A. T. & Stuart, G. J. Excitatory actions of GABA in the cortex. *Neuron* **37**, 299–309 (2003).

53. Richardson, B. D., Ling, L. L., Uteshev, V. V. & Caspary, D. M. Reduced GABA(A) receptor-mediated tonic inhibition in aged rat auditory thalamus. *J. Neurosci.* **33**, 1218–27a (2013).

54. Muir, W. W., 3rd & Mason, D. E. Side effects of etomidate in dogs. *J. Am. Vet. Med. Assoc.* **194**, 1430–1434 (1989).

55. Arrenberg, A. B., Del Bene, F. & Baier, H. Optical control of zebrafish behavior with halorhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17968–17973 (2009).

56. Del Bene, F. *et al.* Filtering of visual information in the tectum by an identified neural circuit. *Science* **330**, 669–673 (2010).

57. Ahring, P. K. *et al.* A pharmacological assessment of agonists and modulators at α4β2γ2 and α4β2δ GABAA receptors: The challenge in comparing apples with oranges. *Pharmacol. Res.* **111**, 563–576 (2016).

58. Zeller, A., Arras, M., Jurd, R. & Rudolph, U. Mapping the contribution of beta3-containing GABAA receptors to volatile and intravenous general anesthetic actions. *BMC Pharmacol.* **7**, 2 (2007).

59. Gee, K. W. *et al.* Limiting activity at beta1-subunit-containing GABAA receptor subtypes reduces ataxia. *J. Pharmacol. Exp. Ther.* **332**, 1040–1053 (2010).

60. Golparvar, M., Saghaei, M., Sajedi, P. & Razavi, S. S. Paradoxical reaction following intravenous midazolam premedication in pediatric patients--a randomized placebo controlled trial of ketamine for rapid tranquilization. *Pediatric Anesthesia* **14**, 924–930 (2004).

61. Ahrens, M. B. *et al.* Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature* **485**, 471–477 (2012).

62. Lin, X. *et al.* High-throughput brain activity mapping and machine learning as a foundation for systems neuropharmacology. *Nat. Commun.* **9**, 5142 (2018).

63. Jordi, J. *et al.* High-throughput screening for selective appetite modulators: A multibehavioral and translational drug discovery strategy. *Sci Adv* **4**, eaav1966 (2018).

64. Besnard, J. *et al.* Automated design of ligands to polypharmacological profiles. *Nature* **492**, 215–220 (2012).

65. Yeh, C.-M., Glöck, M. & Ryu, S. An Optimized Whole-Body Cortisol Quantification Method for Assessing Stress Levels in Larval Zebrafish. *PLoS One* **8**, e79406 (2013).

Preamble to Chapter 3

We have gotten this far using correlation distance (CD) – the previous chapter took the reader through our motion-index based discovery of the paradoxical excitation phenotype. But at some point, our goals broadened and we began asking bigger questions like "How do neuroactive drugs work?" and "Could we map out the entire Zebrafish behaviorome by phenotypically clustering entire libraries of drugs behaviorally profiled in fish*?"* In attempting to answer these penultimate questions, we ran into a bit of a roadblock. We didn't observe very strong phenotypic clusters using our traditional methods, and in the same vein, when we tried phenoblasting for novel compounds that elicited phenotypes similar to known drugs of interest such as stimulants, the top hits weren't very consistent and contained negative control wells. We reasoned there had to be some limitations in our methods.

These limitations could be related to either the representation used, the metric used to compare them, or both. Historically, we only considered learning a custom distance function to replace CD when we were faced with the challenge of comparing high-dimensional embeddings of Zebrafish videos (more on that in chapter 4). CD and other popular metrics operate on 1 dimensional vectors. This was a problem; we were now in a high-dimensional space. Should we just have concatenated all the dimensions into a 1D vector? This seemed like a fools-errand, since this procedure would basically eliminate most temporal relationships across features from different dimensions. Siamese neural networks (SNNs) seemed like a much more natural solution. These types of networks could be trained on a highly-replicated dataset to learn a custom distance metric for comparing embeddings in a space of arbitrary dimensionality. The network would be tasked with learning to distinguish replicates of same drugs from different drugs, and would in the process learn how to differentiate fundamental aspects of behavior. The advantage of SNNs is that the specific

architecture used for encoding input to output can be very flexible (see **Figure 3.1** for an illustration of this). There are deep neural network (DNN) architectures, like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), that can operate on multi-dimensional inputs. Indeed, many of the most popular CNNs, such as VGG[1] , Inception[2], and ResNets[3] operate on RGB images, where the inputs have 3-dimensions (red, green, and blue). These can be readily adapted to support arbitrary numbers of input dimensions. Further, in the time-series domains, 1D CNNs have demonstrated state-of-the-art performance in many time-series classification[4] and natural language processing tasks[5], and they too can be easily adapted to multi-dimensional inputs. Hence, we reasoned that perhaps Siamese 1D CNNs operating on multidimensional time-series embeddings of zebrafish videos had a lot of potential compared to other approaches.

In the process of developing and benchmarking these Siamese architectures on MI, we observed some promising results and learned a ton about our datasets and how to properly set them up for Siamese training; we present this story in the following chapter. We will then proceed to outline some of the progress we've made in computing more sophisticated embeddings of Zebrafish motion and some of the challenges that remain in learning distance metrics to compare them in chapter 4.

1.9 References

1. Simonyan, K. & Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv [cs.CV] (2014).
2. Szegedy, C. et al. Going Deeper with Convolutions. arXiv [cs.CV] (2014).
3. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. arXiv [cs.CV] (2015).
4. Peng, D., Liu, Z., Wang, H., Qin, Y. & Jia, L. A Novel Deeper One-Dimensional CNN With Residual Learning for Fault Diagnosis. IEEE Access 7, 10278–10293 (2019).
5. Yin, W., Kann, K., Yu, M. & Schütze, H. Comparative Study of CNN and RNN for Natural Language Processing. arXiv [cs.CL] (2017).

Chapter 3:

Deep phenotypic profiling of neuroactive drugs in larval zebrafish

Leo Gendelev[1], Douglas Myers-Turnbull[1], Jack Taylor[1], David Kokel[1,2*], Michael J. Keiser[1,3*]

[1] Institute for Neurodegenerative Diseases, University of California, San Francisco, California 94143, USA

[2] Department of Physiology, University of California, San Francisco, California 94158, USA

[3] Departments of Pharmaceutical Chemistry and of Bioengineering & Therapeutic Sciences, University of California, San Francisco, California 94158, USA

***Corresponding author. E-mail: keiser@keiserlab.org, david.kokel@ucsf.edu**

3.1 Abstract

High-throughput screening of neuroactive drugs in larval zebrafish has shown much promise in recent years for the discovery of new targets for known neuroactive drugs. Although the approach used in those studies – a combination of motion index and correlation distance - was highly successful for several interesting phenotypes including antipsychotics and paradoxical excitation, using it for characterization and discovery of more subtle phenotypes and behaviors has been elusive. Here, we propose to replace the correlation distance with a custom learned distance metric; using Siamese Neural Networks trained on a highly replicated screen of 650 known neuroactive drugs. We see a dramatic improvement in our ability to discern same-drug from different-drug replicates. The method generalizes to a different screen entirely, performed months before the highly replicated screen on which the model was trained. In that arena, the new distance metric gets higher classification accuracy on average, but also strikingly outperforms CD for 3 drugs with more subtle phenotypes. The Siamese distance metric can cluster neuroactive drugs into pharmacologically meaningful classes, and its ability to separate drugs from control-treated wells is significantly better than correlation distance.

3.2 Introduction

Approximately one in five adults in the US live with mental illness[1], and 970 million people around the world live with mental or substance abuse disorders[2,3]. Mental health disorders are the leading cause of disability adjusted life years in the US[4]. But treatment of mental health disorders, such as major depressive disorder (MDD), is critical; adult suicide rates have dramatically fallen since a significant increase in prescriptions of Fluoxetine and other serotonin re-uptake inhibitors SSRIs[5–7]. However, antidepressant use in adolescents - associated with increased suicide rates[9], is not clinically recommended[8], and in adults carries many commonly occurring but poorly understood side

effects, such as the alteration of sleep cycle[10] and sexual dysfunction[11] [12]. Antipsychotics such as Haloperidol are infamous for a plethora of severe and debilitating side effects, including extra-pyramidal side-effects (EPS)[13], and although second generation antipsychotics such as Clozapine manage to circumvent these to some extent, they have been shown to cause rare and fatal agranulocytosis and myocarditis[14], and have been implicated in hypertension and obesity[15].

To make matters worse, the genetics underlying Anxiety Disorder, MDD, and post-traumatic stress disorder (PTSD) are poorly understood[16], and antidepressants are the first-line treatment options given to patients in a "magic bullet"[17] approach[18,19] for these disparate mental conditions. But a reliable, consistent, and actionable mechanistic understanding of these disorders remains elusive. In patients with PTSD, Amygdala hyperactivity has been proposed based on magnetic resonance imaging (MRI) studies to play an important role[20], but it has also been implicated in other mood disorders including anxiety[21]. These observations highlight the challenges both in understanding the basic biology of these common mental disorders and the limits in our ability to effectively treat them.

To complicate things, most mental disorders are thought to be polypharmacological in nature[17], and antipsychotics such as haloperidol interact with an entire spectrum of various targets in the central nervous system (CNS)[22]. It is assumed its actions on Dopamine receptors[23] are crucial for its effectiveness in treating psychosis and schizophrenia, but various contributions from binding to the other targets likely contribute to both its desired and undesired effects[24]. For many mental illnesses, we are only beginning to unravel the complex and interconnected neuro-circuitry. Although most SSRI's work through serotonin transporter (SERT) inhibition[25], many other serotonin receptors, as well as the norepinephrine and neuropeptide systems, are thought to contribute to both their immediate and long-term effects[26]. For anxiety, while both the glutamatergic and GABAergic systems are thought to play key roles[19], SSRIs and SNRIs have effectively replaced

benzodiazepines as the preferred treatment options[27]. This exposes the gap between our already limited understanding of the polypharmacology underlying these complex and heterogeneous disorders and our ability to leverage it towards improving standard-of-care.

Perhaps instead of "magic bullets" we should be designing "magic shotguns"; drugs that selectively hit an entire profile of targets to elicit their effect[17]. In line with this philosophy, the Kokel Lab has been using phenotypic screens in larval zebrafish to discover novel compounds that elicit behavioral responses by potentially binding an entire gamut of receptors. A behavioral fingerprint for antipsychotics was established using known antipsychotics (such as Haloperidol), and using a "phenoblast" approach, compounds from a high-throughput screen (HTS) of a novel library were discovered[22]. These compounds were computationally predicted and validated to interact with a Haloperidol-like profile of targets in-vitro. Building on this initial success, a similar approach was taken for a phenotype caused by known anesthetics including etomidate and propofol (chapter 2). This phenotype was the first animal model for paradoxical excitation in the context of anesthetics and sedation in larval zebrafish, in a field where few reliable animal models exist. Here, novel compounds were discovered with chemical structures distinct from known anesthetics that were predicted and experimentally validated to bind both the expected target, the GABAA receptor, as well as a novel target entirely, the serotonin-6 receptor (HT6). Taken together, these two studies have demonstrated the ability of phenotypic screening in larval zebrafish to expose novel pharmacology for known neuroactive drugs, helping generate testable biological hypotheses and allowing for the discovery of novel chemical scaffolds acting through distinct pathways.

However, the phenoblasting approach used in these previous two studies - a combination of motion index time series (MI) and correlation distance (CD) to compare them - hasn't been readily applicable to other classes of neuroactive drugs, such as stimulants, anxiolytics, antidepressants, and convulsants. In this study we explore the notion that CD is not the ideal metric for comparing larval

fish swimming behavior. The issue with CD is that it presupposes perfect swim synchronization. An extreme but realistic example of this are misaligned videos, offset by some number of frames. In this situation, CD could in theory completely fail to identify the similarity between perfectly synchronized motions for identical behaviors. In a less extreme but likely scenario, fish in two different wells could respond strongly to a certain stimulus (for example UV-light) but not in an instantaneous or fully synchronized way. Again, correlation distance might severely penalize the similarity-score between the fish in these wells, even if they were treated by the same neuroactive drug at the same concentration.

In this study we propose to learn a custom distance function for computing similarity between zebrafish MI. For this goal, we perform a highly-replicated screen of neuroactive drugs on the Screen-Well Neurotransmitter Library[28] (from here on referred to as the Biomol library), which contains 661 CNS receptor ligands. From a highly-replicated screen, we can construct a dataset of "positive" pairs (drug-pairs that are replicates of each other), and "negative" pairs (drug-pairs from different plates that are not replicates of the same drugs). We will describe the data-preparation process in more detail shortly, but first we discuss some of the considerations that went into selecting which approach to use for distance learning.

We considered the following requirements for this: (1) the method should be able to learn which parts of the MI are important for distance learning and discard noisy or meaningless regions of the assays (2) it should be robust to poorly aligned videos and segments of video, as well as being robust to delayed fish responses and out-of-phase motions, and (3) it should be designed to take a pair of MI as input, and output a distance or similarity score. In principle, the Siamese Neural Network (SNN) is an architecture that seems to fit this set of requirements perfectly. SNNs were first developed at Bell Labs for signature validation[29] and personal identity fingerprints[30]; they were later applied towards facial recognition and "one-shot" learning, showcasing their ability to learn

even on smaller datasets[31]. In all of the above applications of SNNs, the goal was to learn a distance function rather than a classification label or regression score. The latter highlights the robustness of SNNs in the realm of real biological data, where we don't always have the luxury of "big-data" that is so critical towards successful training ultra-deep architectures; for example for the popular image-classification challenges like the ImageNet Large Scale Visual Recognition Challenge (ILSVRC)[32]. Additionally, SNNs are by design a very flexible architecture. The Siamese halves can be arbitrary neural network architectures (**Figure 3.1**). Indeed, there have recently been successful applications of a wide variety of SNN architectures in different domains, for example; convolution SNNs for object-tracking and image re-identification[31,33], multi-layer perceptron (MLP) SNNs[34] and convolutional SNNs[35] in the time-series and audio signals, and Long Short-Term Memory (LSTM) SNNs[36] for sentence matching.

Thus we are given a variety of options for encoding our motion-index time series data. As a baseline for performance we use a simple Siamese MLP (multilayer perceptron) architecture (from now on referred to as Sia-MLP). For a more advanced architecture inspired by recent state-of-the-art performance of CNNs in the time-series domain[37] we choose to modify a CNN architecture, the DenseNet, that has recently demonstrated state-of-the-art performance in the image-classification domain[38]. Since the zebrafish videos can be quite long, up to 20250 frames, an efficient encoding of the video would require a deep CNN architecture.

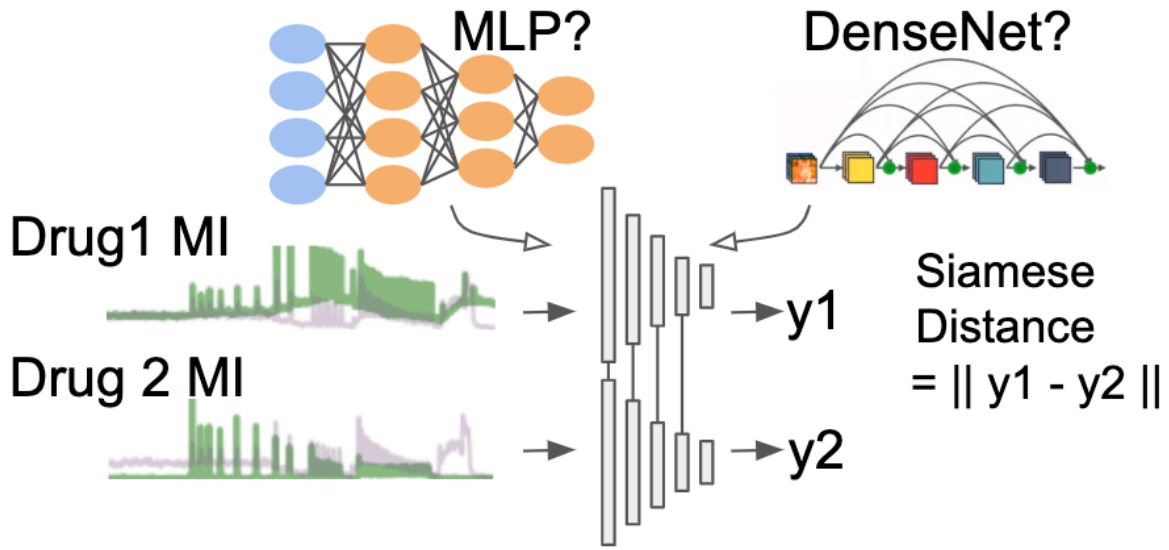


Figure 3.1 Siamese Neural Network Cartoon Illustration

Conceptual illustration of SNN in the context of learning a motion-index similarity function.
A pair of MI for two drugs are fed into the SNN. The SNN consists of two identical halves; the choice of architecture can be arbitrary. In our study, we use both MLP and DenseNet architectures for the encoding layers of the SNN. Once the network is trained, the distance can be computed as the Euclidean norm between the output vectors from each half of the network. For training (not shown), labels are also provided in the form of "0", or "positive", for same-drug replicate pairs and "1", or "negative", for different-drug pairs; the loss function is the contrastive loss function (see methods) that provides some measure of how different the Euclidean norm is from the actual label, clamping the maximum possible contribution from inputs that are drastically different with a "margin" term, and weights are updated accordingly.

The issue with using standard CNNs is that with deeper architectures, vanishing gradients can make training challenging or even impossible[39]. DenseNets circumvent this problem by allowing for a complete information flow across all the layers of the network by utilizing skip-connections (information from all previous layers gets added to the information from all subsequent layers)[40,41]. Doing a complete comparison of all possible choices of Siamese architectures is beyond the scope of this work; the goal of the current study is to demonstrate improvement over existing methods for neuroactive drug-classification and clustering in the context of larval zebrafish phenotypic screens.

We do not perform a full comparison of our performance against dynamic time warping (DTW)[42,43] because of the prohibitively slow nature of that algorithm; regardless of performance, one of our requirements is for rapid phenotypic profiling in high-throughput screens and clustering large

datasets of small molecules. In our hands, even the Python fastdtw package[44,45], which has a linearly increasing run time of O(n) as opposed to the $O(n^2)$ run time of traditional DTW, is prohibitively slow for computing distance matrices for our long time series of 20250 frames, with or without subsequence sampling. Due to the speed and parallelizable nature of Graphics Processing Units (GPUs), we can compute a Siamese distance (SD) trained on GPUs orders of magnitude faster than with DTW. This is another key advantage of our approach that can be adapted to other time-series classification and clustering domains.

Siamese dataset preparation

I. Splitting the highly-replicated screen into training and validation sets

One of the main requirements and applications we considered in deciding how to partition our dataset was that our new distance metric had to generalize to novel compounds from new screens. This would require a distance metric that has learned something fundamental about *phenotype*; not features specific to a certain set of drugs encountered during training. In theory, the most conservative approach would be a *by-plate* split, but in our specific situation this is undesirable since the Biomol library as supplied consists of activity-class based layouts (**Figure 3.2a**); such a split would omit entire pharmacological classes of drugs during training. The only way to fully address this limitation is by completely randomizing the drug-layouts; we have already completed this step and are currently in the process of finishing up collecting this new fully-randomized highly-replicated Biomol dataset (see methods, the computationally randomized plate layouts are presented in the appendix, **Supp Figure A.2.6)**; we will be validating results from this chapter on this dataset in the immediate future (see discussion and future directions).

The least conservative split is to randomly select a subset of the shuffled library for training and use the rest for testing. In our hands this leads to grotesque overfitting (not shown) – the SNNs

are likely able to memorize relationships between positive and negative pairs that are present in the validation set. This sort of combinatorial memorization during training can lead to exceedingly high validation performance but a meaningless distance metric that completely disregards any notion of *phenotype* underlying the positive and negative pair membership. With these considerations in mind, we use a *drug-split* approach to create our training and validation sets. This way, the validation set will never encounter a drug or any of its replicates that was made available to the training set, reducing the possibility of learning drug-specific effects or combinatorial patterns.

II. Plate-location effect considerations

As mentioned in the previous section, the Biomol library is supplied with activity-class based plate layouts (**Figure 3.2a**). Since positive pairs always come from the same plate-location (by definition), this could hypothetically introduce another kind of over-fitting; *plate-location* based effects. Here, instead of learning the actual phenotype, the SNNs might learn light-falloff patterns and edge-effects, and other location based artifacts common in microscopy[46,47]. To account for this possibility, we explore another way of circumventing plate-location effects without shifting away from the original dataset. We create a "*same-well*" version of our dataset where we keep the positive labels unchanged, but only allow negative pairs to come from the same physical locations on plates. That is, if two drugs come from the same well-location (for example, well A1), but from two different library plates (e.g. they are not replicates of the same drug), then we consider them negative-pairs (**Figure 3.2b**). This way, both positive-drug replicate pairs and negative pairs will always come from the same plate-location, so that the SNNs can't employ plate-location as a means of distinguishing positive from negative pairs. We call this dataset the *same-well* dataset and we call the original dataset where negative pairs can come from any wells the *all-well* dataset. In both datasets, we never allow

negatives to come from the same plate. Both versions of the dataset are used for training in the subsequent analyses; however, validation is always done on the *all-well* dataset.

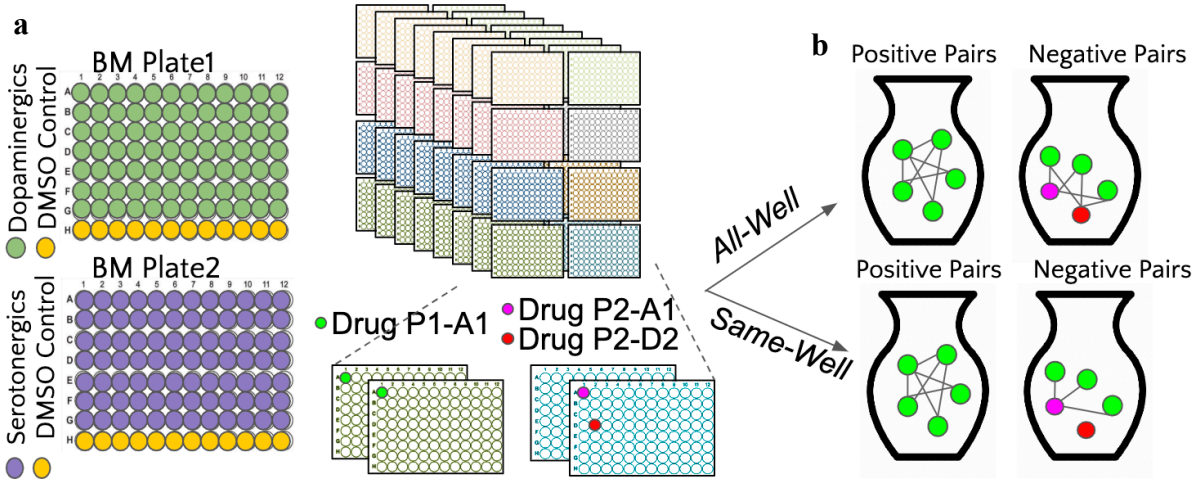


Figure 3.2 Illustration of Biomol highly-replicated screen and training dataset design

Our highly-replicated screen design a) 2 of 8 Biomol plates (cartoons); plates are supplied by activity-class. Some plates have multiple activity classes. Green circles might be dopaminergic drugs; purple circles might be serotonergic drugs. There are 12 DMSO Controls wells per plate (bottom rows, yellow) b) We replicate each copy of the library 7 times. Positive replicates are repeated drugs (for example, all combinations of well P1-A1 with itself; green with green). For the all-well negative dataset, negatives can come from different wells (P1-A1 with P2-D2 is allowed; green with red); for the same-well dataset, negatives can only come from same-well locations on different plates (P1-A1 with P2-A1; only green with magenta allowed).

3.4 Results

I. Initial benchmarks: Siamese distance vs correlation distance

Our first test is to compare both SNNs architectures (Sia-MLP and Sia-Dense) performance to CD and Euclidean distance (ED) at the task of separating out positive from negative pairs. For this analysis, we only consider the *all-well* dataset. Both the DenseNet and MLP distance metrics show drastically better performance, both in terms of positive vs negative distance (**Figure 3.3 a,b,d,e**), and AUC/PRC performance (**Figure 3.3 c,f)**.

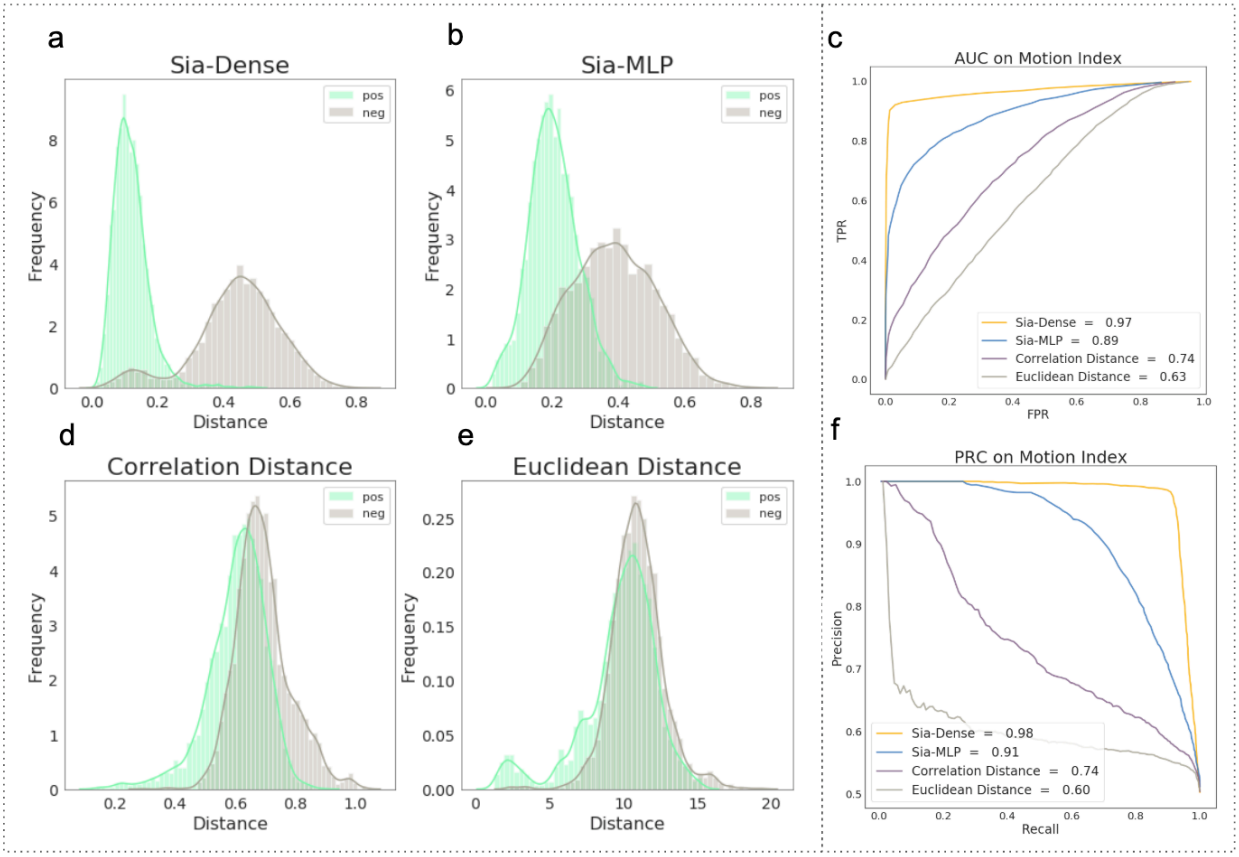


Figure 3.3 Motion index performance comparison; separation of same-drug replicate from different drug pairs

Here we present our trained SNNs ability to differentiate between pairs of drugs that have been replicated (pos labels, mint green) and different drug pairs (neg labels, gray) on a validation hold-out set. a) Sia-Dense b) Sia-MLP d) CD e) ED. c,f) AUC and PRC curves.

Sia-Dense in particular has somewhat outlandish performance; 0.97/0.98 AUC/PRC (compared to 0.74/0.74 for correlation distance), followed by the MLP, then CD, and finally ED. We notice that both SNN metrics and CD produce distances in the [0,1] range, unlike ED, as expected.

Another way of comparing distance metrics is using them to cluster data. We explore the ability of Sia-Dense to cluster our dataset based on the time-point wise means taken across all replicates of the drug compared to CD. In total, this *mean-MI-dataset* consists of 650 time-series, corresponding the 650 drugs from the Biomol library. We cluster this 650x20250 matrix (650 time series each with 20250 time-points) using the Python SciKit Learn[48] (sklearn) implementation of the

K-means algorithm[49], and we use the UMAP dimensionality reduction technique[50] to embed the data onto 2 dimension for visualization, plotting the data colored by K-means clusters. The Sia-Dense clusters (**Figure 3.4b)** are strikingly more aggregated and structured than the CD clusters (**Figure 3.4a)**. We decided to investigate whether the Sia-Dense network was overfitting our dataset as an explanation for such extreme performance.

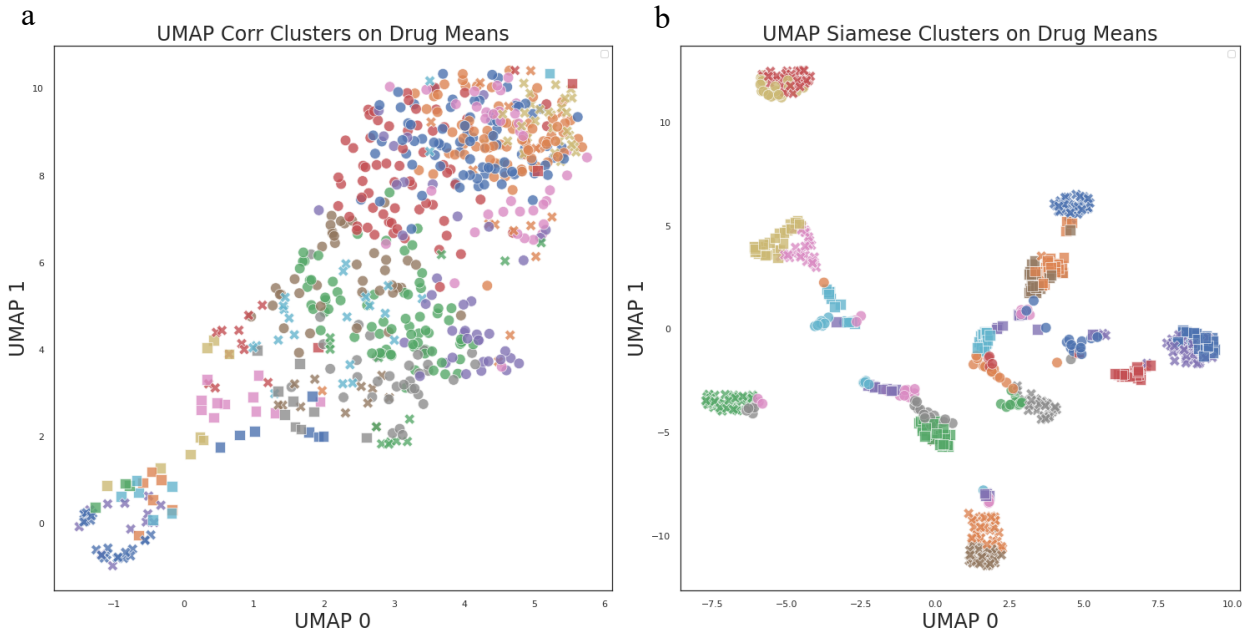


Figure 3.4 DenseNet clusters - potential overfitting?

UMAP[50] embeddings of Biomol drugs in terms of their mean-motion index time-series across all replicates. a) Using CD. The clusters are colored by cluster-membership based on a K-means clustering[49] of the time series means with 24 clusters. b) Same as a) but with Sia-Dense.

II. Investigating plate-location effects by comparing performance on the *same-well* dataset

We suspected that plate-location effects could have been employed by the network. Here we test the possibility that the extreme performance achieve by the Sia-Dense architecture could be explained by these plate-location effects by comparing performance versus the *same-well* dataset.

The AUC and PRC scores for replicate separation, for both the Sia-MLP and Sia-Dense architectures, are not strongly affected by dataset choice (**Table 3.1**, **Supplemental Figures A.2.1 and A.2.2**). For example, for Sia-Dense, the AUC scores are 0.97 for *all-well* and 0.96 for *same-well*. These results suggest that the potential overfitting observed with the Sia-Dense architecture can't be explained by plate-location effects alone.

Table 3.1 All-well vs same-well performance

Dataset	Arch	AUC Score	PRC Score
All-well	Sia-MLP	0.89	0.91
	Sia-Dense	**0.97**	0.98
Same-well	Sia-MLP	0.92	0.93
	Sia-Dense	**0.96**	0.97

III. Other possible explanations for over-fitting?

What other factors, aside from plate-location effects, could explain overfitting? Recent work in the time-series and signal processing domains with deep neural networks (DNNs) suggests that during training, they may fit the meaningful part of the signal first, but eventually overfit the high-frequency components[51]. With MI, factors like camera noise, water movement, and small vibrations of the plates could manifest themselves as high-frequency noise. We hypothesized that by removing this noise, we might reduce overfitting. To remove the high frequency components of the signal, we convolve a hanning smoothing function[52] over each MI time series (see methods). We explored how choice of window-size affected the result (**Supplemental Figure A.2.3**); by visual inspection, a window size of 11 filters out most of the high-frequency components while preserving the low-frequency, high amplitude motions. We create a smoothed version of the MI dataset using this window size.

Smoothing out the data drastically reduces performance for both Sia-MLP and Sia-Dense models (**Figure 3.5, Table 3.2**). The Sia-Dense in particular takes a huge hit to performance; decreasing from 0.97 AUC on raw MI to 0.78 on smooth MI. This indicates that a significant part of the high performance of the Sia-Dense raw model was based on the high- frequency components of the signal. This is also the case for the Sia-MLP model, but to a much lesser extent; its performance drops from 0.89 raw MI to 0.83 smooth MI.

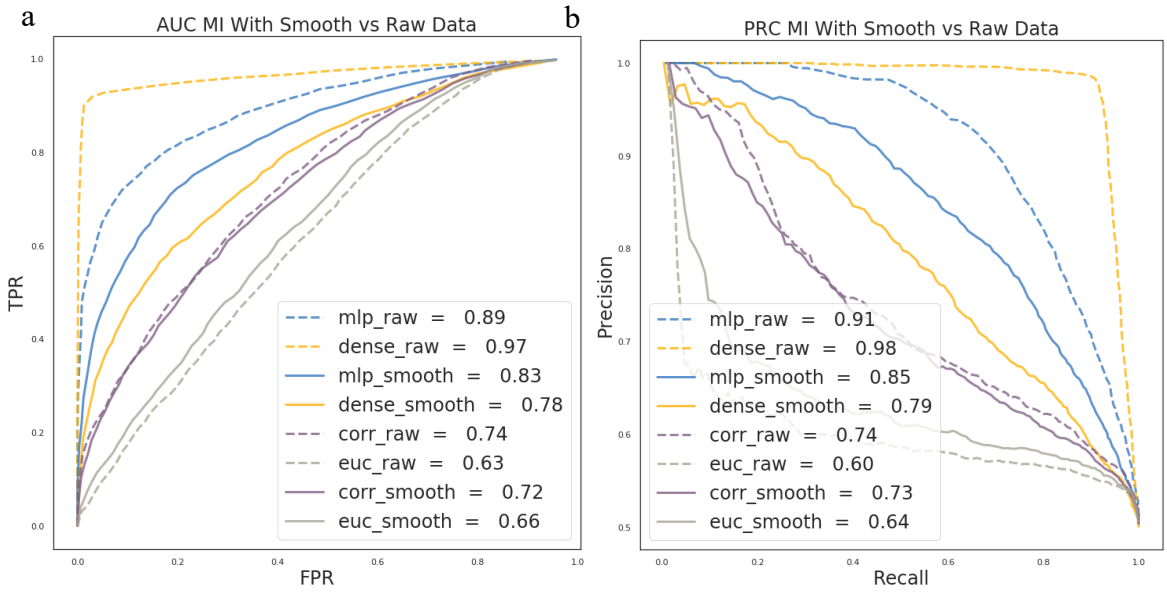


Figure 3.5 Smoothing out data lowers performance for both Siamese models

Smoothing out the motion index data with a hanning window lowers performance dramatically for both Sia-Dense and Sia-MLP models in terms of a) AUC and b) PRC. All smooth data is plotted in solid lines; raw data in dashed lines. CD and ED aren't strongly affected by data smoothing. All models benchmarked on the *all-well* dataset.

We we're faced with two conclusions from our experiments so far; either a) the high-frequency components were a real and meaningful feature of Zebrafish behavior, and we couldn't train effective distance metrics without them, or b) there were other reasons for the poor performance on the smoothed data. We decided to test the latter by considering additional potential underlying issues with our training datasets and models.

Table 3.2 Raw vs smooth performance

Data	Arch	auc_score	prc_score
Raw	Sia-MLP	0.89	0.91
	Sia-Dense	**0.97**	0.98
Smooth	Sia-MLP	0.83	0.85
	Sia-Dense	**0.78**	0.79

IV. Filtering out ineffective drugs from training

We realized another potential issue for SNN training; the positive and negative labels. They were based on the assumption that all the drugs actually induced a substantial effect on fish behavior. In reality, we knew this was unlikely the case. Some drugs may not induce behavioral change in Zebrafish given differences between human and Zebrafish orthologs, may require different pre-incubation periods, may not have been screened at their ideal therapeutic dose, or a may encounter a long-list of other potential issues including but not limited to drug solubility, drug degradation, experimental mixing and handling error, sick fish, and contaminants. So we set out to test whether removing ineffective drugs from the training dataset could improve the SNN performance on the smoothed dataset. We chose to use a random forest classifier (methods) for this simpler task of classifying wells as negative (ineffective drug or DMSO-control) or positive (effective drug) based on their MI.

We observe increasing AUC and PRC scores as we allow the classifier to use more of the time series (**Figure 3.5 b,c**). With the full time series, or 20250 frames (RF 100% in the legend), we observe an AUC of 0.88 and PRC of 0.72, which indicates reasonable performance at the task of distinguishing drug from DMSO-treated wells, even on a relatively small dataset.

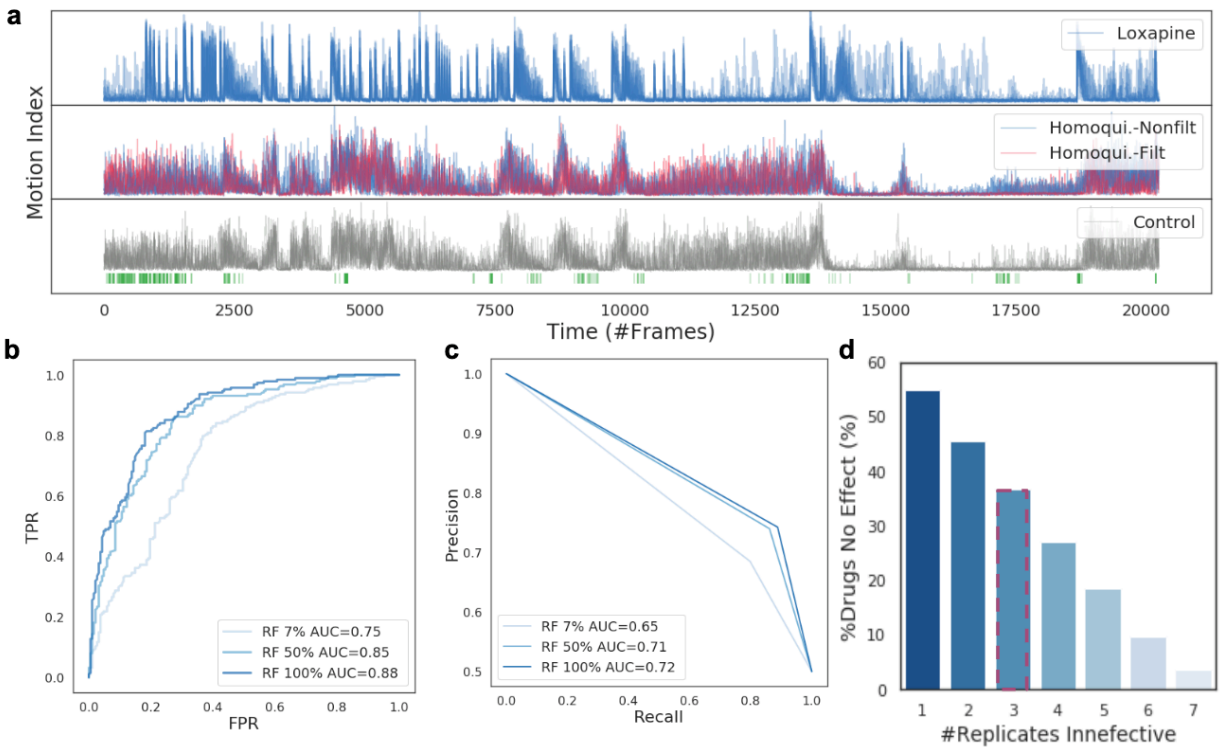


Figure 3.5 Random Forest Drug vs DMSO Control Classifier

Filtering out ineffective drugs using a random forest. a) The RF classifier doesn't flag any Loxapine replicates, but does flag 3 of 7 homoquinolonic replicates as indistinguishable from DMSO (red traces represent DMSO-classified replicates, blue represent non-DMSO classified traces). The top 500 ranked time points important to the random forest classifier are plotted in the green ticks below the DMSO traces. b,c) Area under the curve (AUC) and precision recall (PRC) performance of the DMSO-classifying RF. d) Barplot showing what percentage of drugs in the dataset (y-axis) we filter out depending on how stringent of a filter threshold we use. A cutoff of 3-ineffective replicates filters out 37% of the drugs (bar with red rectangle).

We showcase selected drug treatments where this classifier successfully filters out DMSO-like behaviors (**Figure 3.5a**). For loxapine, a typical antipsychotic (top row), which has a well defined phenotype, no replicates are filtered out (all effective time-series colored blue). For homoquinolonic, an excitotoxin, 3 of 7 replicates are filtered out (red time series, middle row), and its replicates at this concentration of 30μM look visually similar to DMSO control replicates (grey, bottom row). Extracting the most important features from the random forest classifier, we show that it utilizes a wide assays of time-points and assays to perform the classification (green marks, bottom row). We use the criterion that 3 or more of a drugs replicates must be classified as DMSO as grounds for

removal; this procedure removes 37% of the dataset (**Figure 3.5d**). We also filter out individual replicates predicted to be ineffective. The remaining drugs and replicates are used to construct the "RF-Filtered" Siamese training dataset. Note that we used this filtering procedure solely to filter confounding replicates from model training, so as to improve the clarity of the training signal. In the model's evaluation, however, we left all replicates in place, regardless of their quality, so as to avoid a self-referential and trivial performance boost pitfall.

Pre-filtering the dataset for ineffective drugs, we are able to "rescue" performance (**Figure 3.6a,b** dash-dotted lines) on smoothed data. This may be consistent with the idea that removing ineffective drugs improves the ability of our models to learn meaningful Siamese distance, without having to resort to exploiting high-frequency artifactual signal..

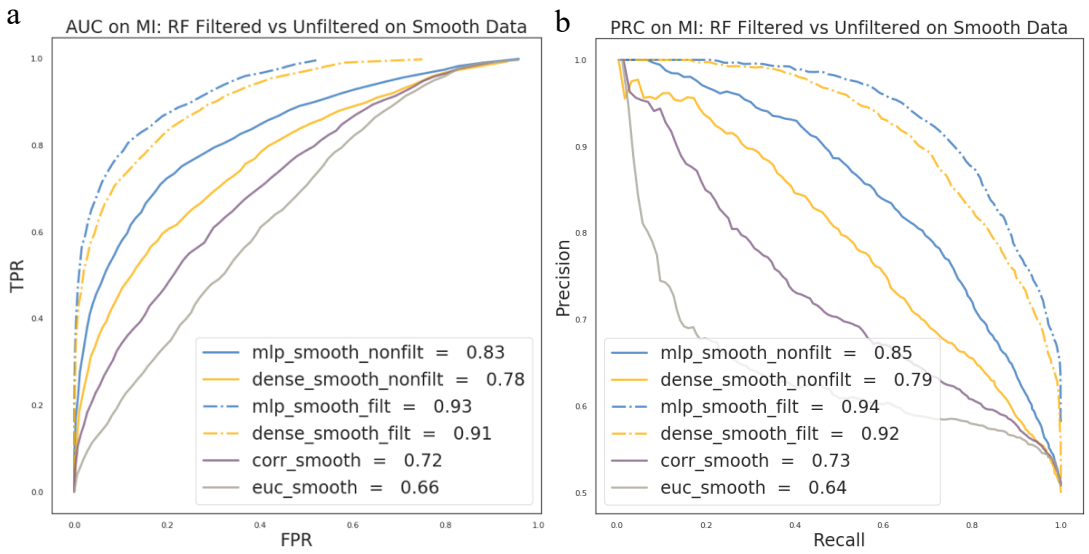


Figure 3.6 Filtering our ineffective drugs with RF classifier "rescues" performance

On the smoothed dataset, filtering out ineffective drugs with the RF classifier dramatically improves performance for both Sia-Dense and Sia-MLP models in terms of a) AUC and b) PRC. All RF-filtered data is plotted in dash-dotted lines; unfiltered data in smooth lines. All plots shown based on *all-well* dataset.

VI. Generalizability of models

Now that the Siamese models were yielding high performance on smoothed data, we wanted to see if they were capable of generalizing to a different dataset entirely. Fortunately the Kokel Lab

had completed a quality control (QC) screen (see methods) just months prior to the highly-replicated screen. This QC screen was done on a smaller set of 16 drugs distributed in different regions of the plates to account for potential plate-position effects. These drugs were replicated 108 times each. The average MI of those 16 drugs are presented in the appendix (see **Supp Figure A.2.4**) for reference.

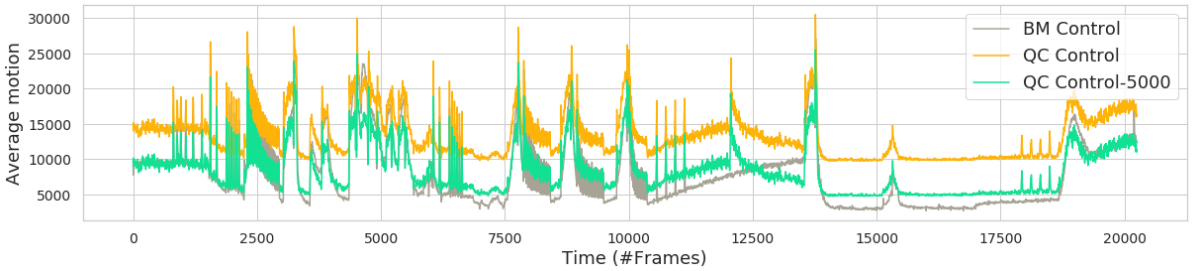


Figure 3.7 Quality control vs biomol screen

The mean motion index time series (average across all DMSO replicates) from both the Biomol highly-replicated screen (grey trace) and the QC control screen (yellow trace) are not the same. We subtract 5000 from the MI of the QC dataset to bring it closer to the Biomol dataset.

We noticed substantial differences in the motion index time series of DMSO control treated wells when comparing to the highly replicated screen (**Figure 3.7**). There could have been some differences between the camera settings or experimental setups between the two screens; after all, they were performed months apart. These kinds of experimental discrepancies are common in biological data. Carefully normalizing the two datasets is a highly non-trivial task and beyond the scope of this paper. One approach we've considered as a future direction for this is borrowing from the neural machine translation field and training neural network "transcoders" to translate between the two datasets. For now, in order to prevent the discrepancy from undermining the comparison, we base-subtract the QC data by 5000 – doing that seems to at least visually normalize the baselines between the two datasets, without over-engineering the problem and introducing overfitting. We note this is actually more of a problem for the SNN metric; CD doesn't depend on the absolute

magnitudes of the time-series vectors since it focuses on trends in the time-domain and is invariant to shifts across the y-axis. In a sense, CD might actually have an a priori advantage in this comparison over Siamese distance.

We use the Sklearn implementation of the K-Nearest-Neighbor classifier[53], supplying either the SD or CD as the metric by which the algorithm compares rows of the input data, to classify the QC screen replicates into the 16 possible labels. First we investigate the best choice for the number of neighbors for the kNN algorithm. We observe that the overall classification performance peaks around 15 neighbors for all kNN-Sia models; for the kNN-CD models, the highest performance is observed at n=1. This effect is similar for both RF-filtered (**Figure 3.8a**) and unfiltered (**Figure 3.8b**) datasets. Performance gradually declines with large neighbor number for all models – likely due to an averaging effect. For all subsequent analysis in this section, we set 15 as the choice of neighbors for all SNN models, and 1 for CD.

We find that multiple Siamese models outperform correlation distance in terms of overall accuracy at the task of classifying the 16 QC screen drugs (**Table 3.3**). CD classifies QC drugs at an overall accuracy of 59% while the highest scoring model - the Sia-MLP with the raw unfiltered all-well negative dataset reaches 68.75%. Consistent with the behavior we observed previously, the random forest filtered data outperforms unfiltered data on the smoothed dataset, same-well negatives don't change performance substantially, and the MLP networks outperform DenseNets across the board.

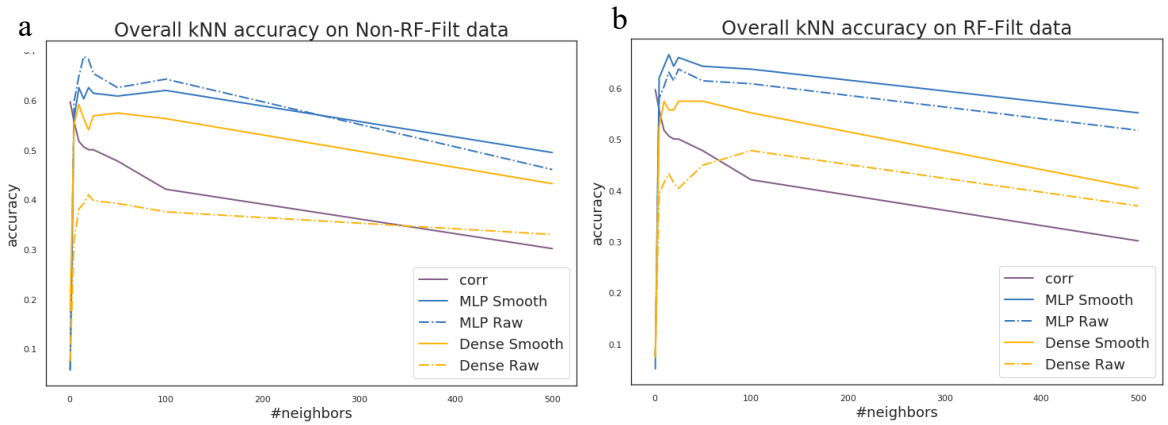


Figure 3.8 kNN accuracy as a function of number of neighbors

We investigate kNN accuracy for classifying the 16 QC screen drugs (y-axis) as a function of number of neighbors, for some of our architectures. a) Non RF-filtered data and b) RF-filtered data. All smooth dataset plotted with solid-lines; raw datasets with dash-dotted lines.

Our interpretation of these results is that the models indeed generalize to different datasets collected at different times, despite the substantial differences in the MI traces from these. The highest overall performance is achieved by a Sia-MLP on the raw data. This was rather surprising to us, since we thought the networks were overfitting the raw data, and didn't expect to see cross-screen generalization. This indicates that Sia-MLP performance on raw data could be largely based on real phenotypic similarity rather than similarity hidden within instrument or behavioral high-frequency components.

Table 3.3 Raw vs smooth performance

Arch	Data Type	RF filtered?	Well Type?	Accuracy
mlp	raw	nonfilt	allwells	68.75%
mlp	raw	nonfilt	samewell	66.48%
mlp	smooth	filt	allwells	66.48%
mlp	raw	filt	allwells	63.07%
mlp	smooth	nonfilt	samewell	60.23%
mlp	smooth	nonfilt	allwells	60.23%
corr	raw			59.66%
corr	raw			59.66%
mlp	smooth	filt	samewell	57.39%
dense	smooth	nonfilt	allwells	56.25%
dense	smooth	filt	allwells	55.68%
mlp	raw	filt	samewell	55.11%

Arch	Data Type	RF filtered?	Well Type?	Accuracy
dense	smooth	nonfilt	samewell	51.14%
dense	smooth	filt	samewell	47.73%
dense	raw	filt	allwells	43.18%
dense	raw	nonfilt	allwells	39.20%
dense	raw	filt	samewell	21.02%
dense	raw	nonfilt	samewell	9.66%

Next we examine the by-class performance of selected models. For the following analysis, we compare the two top performing models from the smooth and raw datasets; "mlp-raw-nonfilt-allwells" and "mlp-smooth-filt-allwells", against correlation distance (**Figure 3.9a)**. Both MLP models drastically outperform CD on 3 drugs in particular; Tracazolate, Tiagabine, and Lidocaine (starred on x-axis). The MI for these 3 starred drugs is shown for comparison in **Figure 3.9b** below.

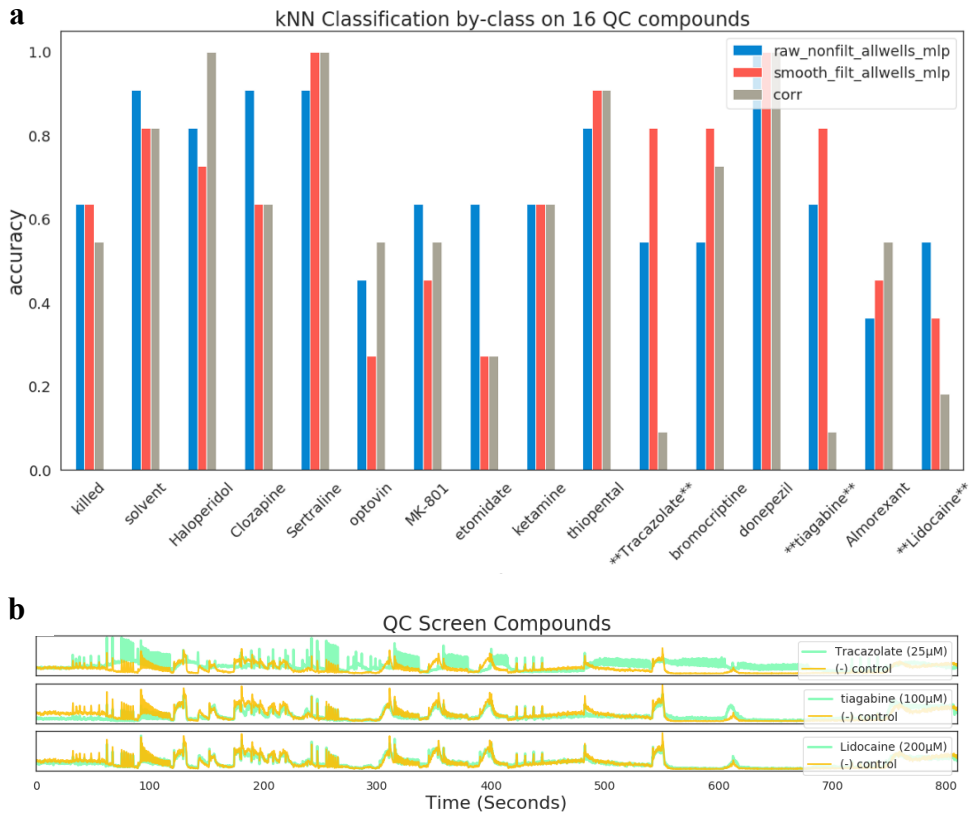


Figure 3.9 kNN by-class classification performance for top Siamese models

kNN performance by class is shown for QC screen for two of the top performing MLP Siamese models. a) barplot with 16 drugs on x-axis and accuracy on y-axis. b) The motion index averages for the 3 starred QC compounds is shown.

Although the raw MI model has the best overall accuracy by a small-margin, the smooth MI model has the best performance for Tracazolate and Tiagabine, 2 of the 3 drugs for which SD has a dramatic edge over CD. It is nice to see the SNN distances providing a significant edge over CD for drugs with potentially more challenging phenotypes. In the case of Tiagabine and Lidocaine, the phenotypes are quite subtle and tricky to discern from DMSO control by eye (**Figure 3.9b,** rows 2 and 3). Encouraged by the smooth Sia-MLP models generalizability and robust performance, we examine its ability to meaningfully cluster drugs.

VII. Neuroactive Clustering

Following the same procedure we used in section I (initial results), we cluster the mean MI of the 650 drugs from the Biomol highly replicated screen (**Figure 3.10a**) using the Kmeans algorithm with 48 clusters and embed the time-series matrix onto two components for visualization using the UMAP technique with all default parameters except for the neighbor number, which we set to 10, and visualize select clusters from distinct parts of the UMAP (**Figure 3.10b**). Mean traces for the 48 distinct clusters can be found in the appendix for reference (**Supp Figure A.2.5**). The most distinct region in the UMAP appears towards the middle-left part of the plot; this region consists of the "lethal" phenotype, and closely followed by sedation and the paradoxical excitation phenotype (cluster 31, red "x's", top 3 rows of panel b). This cluster contains MPEP, a specific inhibitor of the metabotropic glutamate receptor (mGLuR4), which we originally discovered as an inducer of the paradoxical excitation phenotype from chapter 2. This cluster also contains B-CCE and ICI-198,256, both of which have activity on Benzodiazepine receptors[54,55]. The Benzodiazepine association is quite interesting for this phenotype, since benzodiazepines induced partial paradoxical excitation in chapter 2.

The next cluster, 7, (**Figure 3.10b**, aqua, circles) is a cluster of antidepressants; we visualize the MIs for fluoxetine, nisoxetine, and fluvoxamine. Cluster 4 (green circles), has a strong adrenergic signal, with MIs for yohimbine, clenbuterol, and epinephrine visualized. Finally, we visualize cluster 13 (located in the middle of the large top-right lobe in the UMAP, magenta circles). This cluster is interesting because it is phenotypically similar to DMSO control. It does appear to have some opioid signal; nociceptin is the endogenous ligand for the nociceptin opioid receptor; endomorphin is an endogenous m-opioid receptor neurotransmitter; however muscimol is a potent $GABA_AR$ receptor agonist with hallucinogenic and sedative-hypnotic activity, so its membership in this cluster is grounds for further interrogation. It is possible this cluster is just one of several DMSO-like clusters, and is phenotypically non-distinct. Regardless, the Siamese distance metric does seem to do a reasonable job of placing drugs into phenotypic groups, and appears to be useful for neuroactive drug clustering.

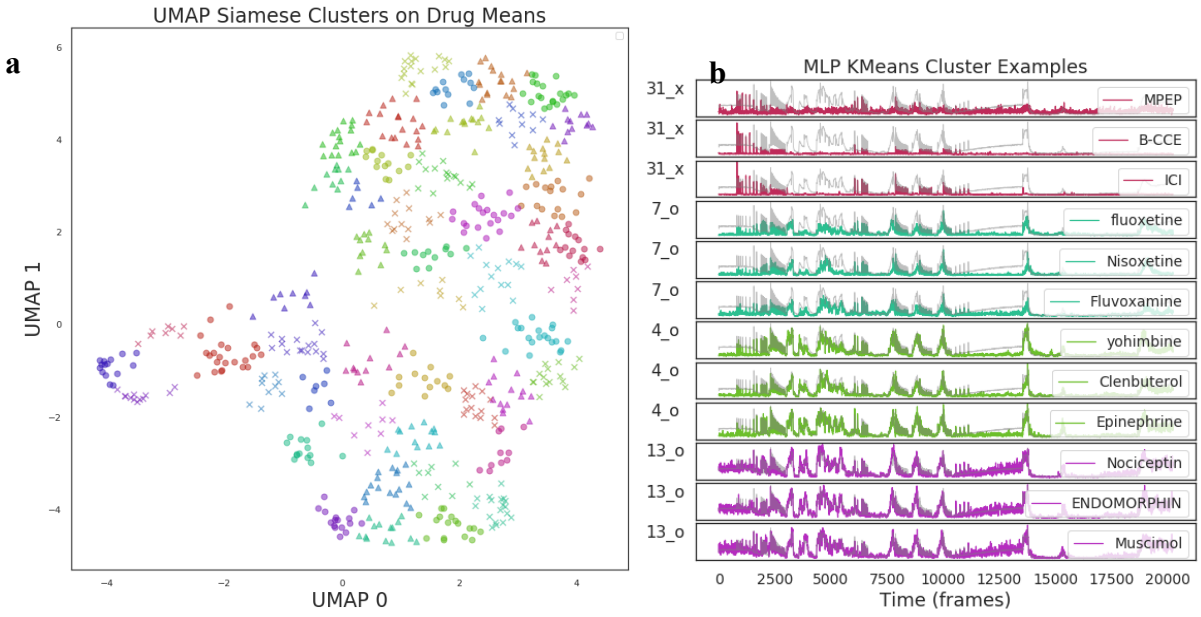


Figure 3.10 Siamese MLP clusters

UMAP embedding of Biomol drugs in terms of their mean-MI across all replicates. a) A UMAP two-component embedding with Siamese distance on the smooth data. The clusters are colored by the cluster-membership based on a K-means clustering. b) Visualization of select clusters. The row labels correspond to the cluster and marker style as they appear in the UMAP.

VIII. Metric sensitivity

A key performance criterion we'd like to satisfy with our new distance metric is sensitivity to weaker phenotypes. CD performs well for exceptionally strong, outlier phenotypes - such as paradoxical excitation - but it doesn't seem to robustly distinguish more subtle phenotypes. We compute the SD and CD between drug and DMSO for all 650 drugs and plot the result (**Figure 3.11a,b)**. In the left panel of the figure (a), the drugs on the x-axis are ordered by their SD from DMSO; in the right panel (b), by their CD from DMSO (in ascending order). SD from DMSO is greater than CD from DMSO for an overwhelming number of drugs; implying SD is more sensitive to a broader range of drugs and associated behavior.

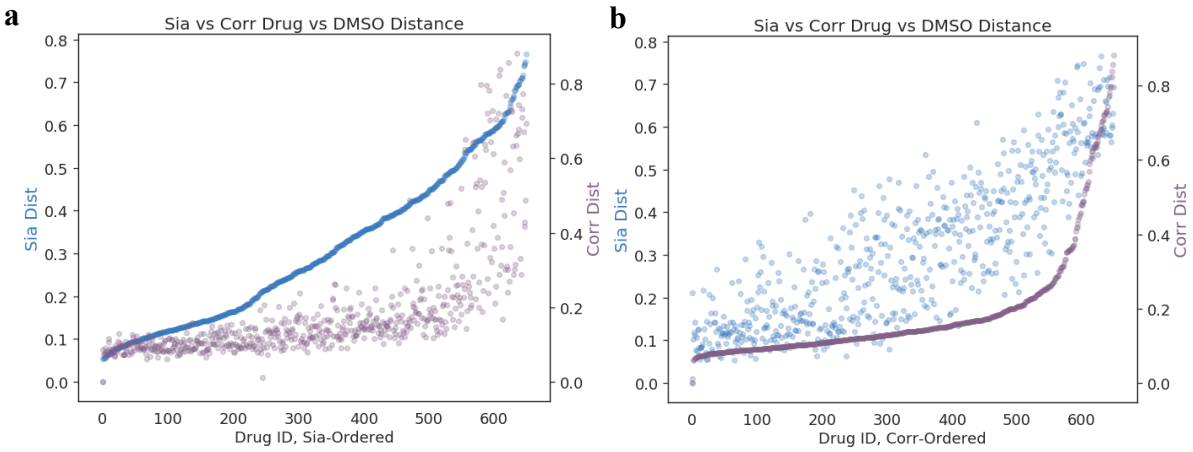


Figure 3.11 SD and CD vs DMSO for 650 Drugs

Sensitivity of Sia-MLP vs DMSO and CD vs DMSO for 650 drugs ordered by a) SD from DMSO, b) CD from DMSO. In both panels, blue markers are for SD, and purple markers are for CD.

We might have been worried if SD had little overlap with CD, but it was reassuring to see SD get the low-hanging fruit; drugs with the highest DMSO CDs also tend to get very high DMSO SDs. After all, we don't expect correlation distance to be completely "wrong"; it's probably just limited to the strongest phenotypes.

Discussion

Through a set of experiments performed in this study, we train distance metrics with fully-connected Siamese Neural Networks that, a) perform substantially better than correlation distance at the task of separating same-drug replicate pairs from different-drug pairs, b) are able to generalize to a completely different dataset, c) substantially outperform CD at classifying several drugs with more subtle phenotypes, d) produce neuroactive drug clusters that seems reasonable phenotypically and pharmacologically, and e) are more sensitive to a broader range of phenotypes. We believe these Siamese distance metrics should have an immediate impact and applicability in Zebrafish phenotypic drug discovery.

I. DenseNets vs MLPs

One of the most unexpected discoveries we make in the process is that the fully connected MLP networks outperform the more advanced and in-principle more powerful DenseNets. Further, we find that DenseNets are more prone to overfitting on high-frequency components of the time-series. Upon reflection, this is not so surprising. DenseNets learn hierarchical representations of the time-series with a deep stacks of convolutions; layers at various depths should be able to perceive different temporal features and connections between motions occurring at different length-scales. However, it is not clear that larval Zebrafish behavior actually exhibits a strong long-term memory component. It may not always be the case that fish experiencing one type of stimulus will respond differently to a future stimulus. Further, even if this effect does exist, its magnitude is uncertain; it is possible that for some stimuli, this "assay-memory" is not significant. Unfortunately, we currently lack strong evidence of this one way or another; future studies need to elucidate the prevalence of long-term assay effects on Zebrafish behavior.

For now, we suggest that one possible explanation for the poor DenseNet performance and overfitting stems from their detail-oriented view of motion index patterns at various length-scales, burdening their learning flux with high-frequency details and possibly weak long length-scale effects. In contrast, the fully-connected networks seem to accomplish the task of comparing motion index time-series effectively by learning all the possible ways time-points for two time-series can be compared to each other. By design, they should be able to accomplish what dynamic time warping sets out to do – allowing for shifts while comparing different combinations of time-points. In future studies, it should be insightful to test other architectures that incorporate the best from each approach. One such network might be an MLP/CNN hybrid[56], or a RNN/CNN hybrid[57,58], with 1D convolutions performed on shorter segments of time series (perhaps in the context of individual assays) stitched together by an outer fully-connected or recurrent layers operating on the CNN segment embeddings.

Another explanation entirely might be that the DenseNets - as the more complex architecture - require a larger dataset to effectively train; for this the solution is a performing a larger screen with more replicates (which can be prohibitively expensive), or generating more data with smart data augmentation. In this vein, recently there has been some promising work using Generative Adversarial Networks (GANs)[59] for generating new time-series data[60]. It would be ideal to be able to improve Siamese distance performance without having to generate new experimental data.

Yet another possibility entirely is that DenseNets are more likely to overfit MI data because of their higher complexity. With a fully randomized, highly replicated screen, there should be fewer opportunities for the DenseNets to overfit drug-layout and plate-layout-based effects, and they could play to their strengths better in such an arena.

II. Data smoothing, overfitting, and pre-filtering

In our experiments we observe that smoothing out MI with a Hann filter before training reduces performance, suggesting that overfitting occurs on high-frequency components. Pre-filtering ineffective drugs with a random forest DMSO classifier rescues performance on smoothed data; suggesting that data-smoothing and improving the quality of training labels is a key step. There is some support for this notion stemming from the observation that the smooth-MI MLP network performs better than its raw-MI counterpart on 3 of the 16 drugs for which correlation distance performance is particularly poor.

It is possible that the Hanning 11 filter is too aggressive for certain phenotypes; in particular ones where sharp peaky effects might be important. In this case a narrower window might yield better results. An important future direction is exploring how the choice of Hanning window size affects Siamese performance. It is possible that with a more optimized smoothing procedure, the smooth-MLP networks will outperform their raw-MLP counterparts on more drugs.

For the time being, we believe the smooth-MLP Siamese distance is the best choice for current phenotypic drug discovery efforts in larval Zebrafish. In future phenoblasting experiments on compounds from novel compound screens (see chapters 1 and 2), we believe the smooth-MLP distance will produce the most robust results, least prone to plate effects. We intend to test these notions once we have new Siamese distance metrics trained on the fully randomized screens.

We forgo drawing strong conclusions based on this work without training SNNs on a the fully well-location randomized screen. The SD metrics do generalize to the QC screen that was designed to control for plate-effects, giving us confidence that subsequent analysis on the fully randomized screen will validate our findings. We are also encouraged by the fact that we get good classification on the QC screen despite the underlying data incongruity between the two screens. In

fact we were quite surprised by how well the metrics performed in this task. With more uniform data across screens, we expect to achieve even higher classification performance.

Methods

Raw and smoothed motion index calculation

The raw motion index is calculated as follows: $MI = \sum_{n=0}^{N} abs(frame_n - frame_{n-1})$. The smoothed motion index time series is obtained by convolving a Hanning window function[52] of width 11 with the Python NumPy[61] convolve module. The effect of choosing various window sizes for a random segment of video of 750 frames is illustrated in (**Supp Figure A.2.3)**. We set the window size to be 11 because we observed this removed most of the high-frequency component of the signal while preserving most of the low-frequency motions.

Siamese networks

The Sia-MLP architecture (**Figure 3.12a**) consists of two MI inputs of 20250 frames, followed by a series of fully connected layers, batch-norm layers[62], and ReLU activations[63]. The hidden size of the siamese output layer is set to 10. The loss is computed using a contrastive loss function [64] commonly used in many SNN implementations, where we choose a margin of 0.5 on the negative labels. All MLP networks were run for 25 epochs using the learning rate of $1x10^{-4}$ with the Adam optimizer [66] and weight decay of $1x10^{-3}$.

The Sia-Dense architecture (**Figure 3.12b**) is an SNN with each identical half an adaptation of 2D DenseNets where we made the following modifications: all 2D convolutions are converted to 1D, instead of max-pooling we use convolutions with a stride of 2, and instead of the original large architecture (6, 12, 24, 16) dense blocks, we simplify to (3, 3, 3, 3), since our training dataset is

limited in size compared to the ImageNet dataset on which the original DenseNets were trained on. After the last convolution, the hidden size is 635. We added 2 fully connected layers to reduce the resulting hidden size to 10, to match Sia-MLP, adding batch-normalization and a ReLU activation in between them. Everything else is identical to the Sia-MLP implementation. Dense networks were run for 50 epochs using a learning rate of $1x10^{-3}$ with the Adam optimizer and weight decay of $1x10^{-3}$. Both architectures were designed and trained using the PyTorch deep learning software[65] (version 1.1.0), on in-house GPU clusters consisting of either Nvidia GTX 1080 or GTX 2080 GPUs.

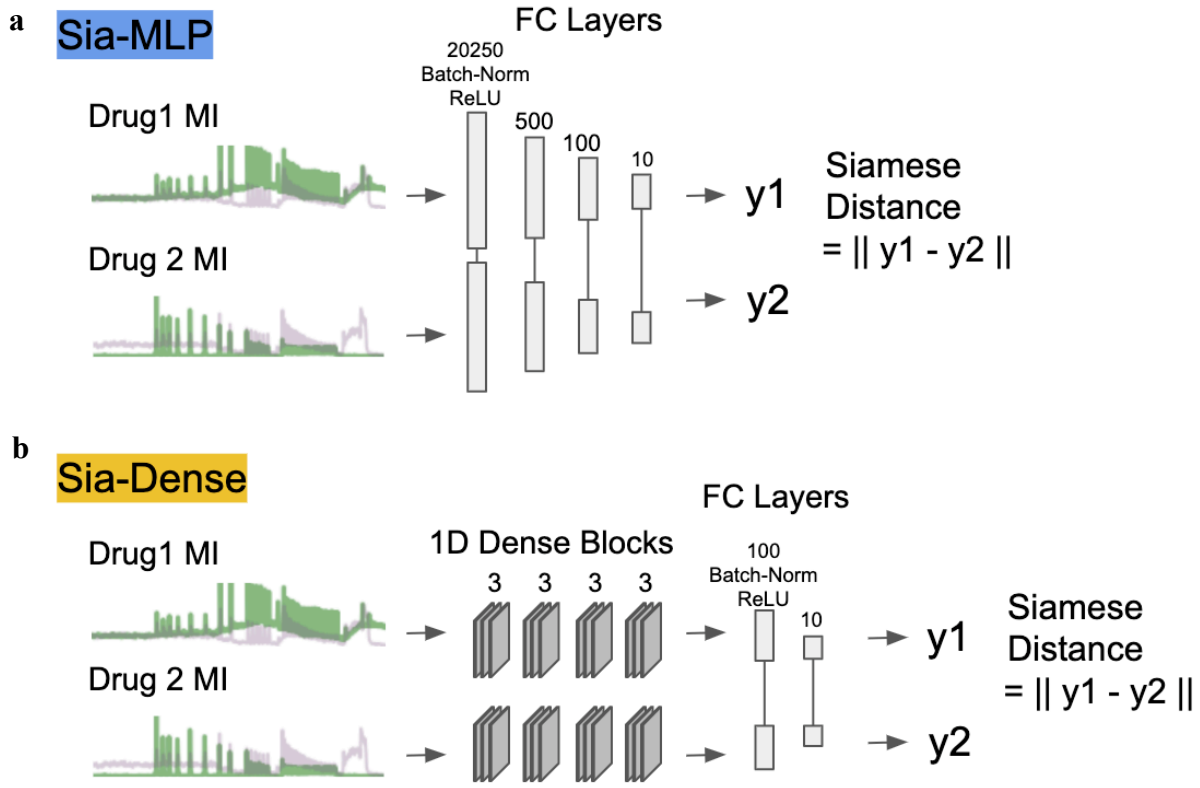


Figure 3.12 Sia-MLP and Sia-Dense Architectures

Sia-MLP vs Sia-MLP illustration. a) Sia-MLP consists of 4 fully-connected layers (20250, 500, 100, 10), with batch-norm and ReLU activations after each one except the output layer. b) Sia-Dense consists of 4 1D dense blocks (3, 3, 3, 3) followed by a fully-connected layer of 100, batch-norm and ReLU activations, and finally an output layer of size 10.

Random Forest DMSO Classifier

The dataset was constructed by enumerating all the DMSO replicates in the highly-replicated screen (a total of 748 unique traces), and under-sampling from the drug wells to match the number of DMSO for class balance. The drugs and DMSO replicates were split at random for the training and validation sets. The random forests were trained using the sklearn random forest package[67] using all default parameters except for the number of estimators. We explored this as a parameter and observed the best performance with 2000 estimators.

Performance evaluation

For all machine learning analysis, model performance was evaluated on the validation sets using the area under the curve (AUROC) and precision recall (PRC) functions from the Python sklearn package[48]. Correlation distance and Euclidean distance baselines are computed using the Python SciPy package[68].

Highly replicated screen and quality Control screens

The highly-replicated Biomol[28] library was ordered and layouts were unchanged. This library containing 661 known neuroactive drugs spread across 8 plates. DMSO control treated wells were designated as the last (bottom) row of all plates.

The layouts for the QC screen were designed to control for potential plate-location effects. 16 drugs, chosen based on diversity of phenotype, (see **Supp Figure A.2.4** for the drug names) were systematically spread out across 18 plates; each 96-well plate contained 6 copies of the 16 drugs at physically distinct locations. We refer the reader to chapter 2 (see methods) for all aspects of Zebrafish maintenance, breeding, and behavior phenotyping; they are analogous to the methods used for this chapter.

Fully randomized screen

This randomization involved a 3 step procedure. In the first step, we separated DMSO and Water dissolved drugs into separate 96 well plates. There were 76 water compounds, and they were all transferred robotically with the help from the Small Molecule Discovery Center at UCSF onto one 96-well plate, while all the DMSO-dissolved compounds were spread across the remaining 7 plates. The second step involved a "quad-map" transfer of compounds from the 7 DMSO compound plates onto 2 384-well plates, where the drugs from each 96 well plate are transferred to a unique quadrant of the 384-well plate. To account for higher water-evaporation rates, we transferred the water-dissolved compounds from their 96-well plate to a separate 384-well plate which was sealed off. For the final step, we computationally randomized layouts using Python that the robot used to transfer drugs from the 384-well plates to the target randomized 96-well plates. In total we prepared 80 randomized layouts (8 target plates with 10 replicates). For each library replicate, the DMSO-dissolved compounds were transferred first, followed by the water-dissolved compounds. The robot was programmed to open and close the seal on the Water-dissolved compounds before and after randomization, to minimize evaporation and make sure the final drug concentrations in the final randomized plates were as close to the final desired 30μM as possible. The final randomized plates were designed to have 73 drug wells, 8 DMSO wells, and 3 water wells, 2 eugenol (lethal positive control) wells, except for the last plate of each library copy, which was designed to have 72 drug wells and 7 water wells, and the same number of DMSO and eugenol wells.

3.8 Acknowledgments

We thank Steve Chen and Michelle Arkin at the UCSF Small Molecule Discovery Center for all their help on randomizing plate layouts.

3.9 Contributions

LG designed and performed the computational experiments and data analysis, designed fully randomized layouts, and wrote the chapter. JT performed all behavioral screens. DMT designed the QC dataset. MJK and DK designed the experiments.

3.9 References

1. NIMH » Mental Illness. Available at: https://www.nimh.nih.gov/health/statistics/mental-illness.shtml. (Accessed: 30th August 2019)
2. Ritchie, H. & Roser, M. Mental Health. *Our World in Data* (2018).
3. Global Burden of Disease (GBD). *Institute for Health Metrics and Evaluation* (2014). Available at: http://www.healthdata.org/gbd. (Accessed: 30th August 2019)
4. What do we know about the burden of disease in the U.S.? - Peterson-Kaiser Health System Tracker. *Peterson-Kaiser Health System Tracker* Available at: https://www.healthsystemtracker.org/chart-collection/know-burden-disease-u-s/. (Accessed: 30th August 2019)
5. McCain, J. A. Antidepressants and suicide in adolescents and adults: a public health experiment with unintended consequences? *P T* **34**, 355–378 (2009).
6. Isacsson, G., Rich, C. L., Jureidini, J. & Raven, M. The increased use of antidepressants has contributed to the worldwide reduction in suicide rates. *Br. J. Psychiatry* **196**, 429–433 (2010).
7. Grunebaum, M. F., Ellis, S. P., Li, S., Oquendo, M. A. & Mann, J. J. Antidepressants and suicide risk in the United States, 1985-1999. *J. Clin. Psychiatry* **65**, 1456–1462 (2004).
8. Center for Drug Evaluation & Research. Suicidality in Children and Adolescents Being Treated With Antidepressant Medications | FDA. *U.S. Food and Drug Administration* (2018). Available at: https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/suicidality-children-and-adolescents-being-treated-antidepressant-medications. (Accessed: 30th August 2019)
9. Bridge, J. A. *et al.* Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials. *JAMA* **297**, 1683–1696 (2007).
10. Wichniak, A., Wierzbicka, A., Walęcka, M. & Jernajczyk, W. Effects of Antidepressants on Sleep. *Curr. Psychiatry Rep.* **19**, 63 (2017).
11. Montejo, A. L., Montejo, L. & Navarro-Cremades, F. Sexual side-effects of antidepressant and antipsychotic drugs. *Curr. Opin. Psychiatry* **28**, 418–423 (2015).
12. Bala, A., Nguyen, H. M. T. & Hellstrom, W. J. G. Post-SSRI Sexual Dysfunction: A Literature Review. *Sex Med Rev* **6**, 29–34 (2018).
13. Sykes, D. A. *et al.* Extrapyramidal side effects of antipsychotics are linked to their association kinetics at dopamine D2 receptors. *Nat. Commun.* **8**, 763 (2017).

14. Lieberman, J. A. *et al.* Clozapine-induced agranulocytosis: non-cross-reactivity with other psychotropic drugs. *J. Clin. Psychiatry* **49**, 271–277 (1988).

15. Hirsch, L., Patten, S. B., Bresee, L., Jette, N. & Pringsheim, T. Second-generation antipsychotics and metabolic side-effects: Canadian population-based study. *BJPsych Open* **4**, 256–261 (2018).

16. Smoller, J. W. The Genetics of Stress-Related Disorders: PTSD, Depression, and Anxiety Disorders. *Neuropsychopharmacology* **41**, 297–319 (2016).

17. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature Reviews Drug Discovery* **3**, 353–359 (2004).

18. Cascade, E. F., Kalali, A. H. & Thase, M. E. Use of antidepressants: expansion beyond depression and anxiety. *Psychiatry* **4**, 25–28 (2007).

19. Cassano, G. B., Baldini Rossi, N. & Pini, S. Psychopharmacology of anxiety disorders. *Dialogues Clin. Neurosci.* **4**, 271–285 (2002).

20. Koenigs, M. & Grafman, J. Posttraumatic stress disorder: the role of medial prefrontal cortex and amygdala. *Neuroscientist* **15**, 540–548 (2009).

21. Nemeroff, C. B. & Marmar, C. *Post-Traumatic Stress Disorder*. (Oxford University Press, 2018).

22. Bruni, G. *et al.* Zebrafish behavioral profiling identifies multitarget antipsychotic-like compounds. *Nat. Chem. Biol.* **12**, 559–566 (2016).

23. Fox, C. A., Mansour, A. & Watson, S. J., Jr. The effects of haloperidol on dopamine receptor gene expression. *Exp. Neurol.* **130**, 288–303 (1994).

24. López-Muñoz, F. & Alamo, C. The consolidation of neuroleptic therapy: Janssen, the discovery of haloperidol and its introduction into clinical practice. *Brain Res. Bull.* **79**, 130–141 (2009).

25. Stahl, S. M. Mechanism of action of serotonin selective reuptake inhibitors. Serotonin receptors and pathways mediate therapeutic effects and side effects. *J. Affect. Disord.* **51**, 215–235 (1998).

26. Kent, J. M., Mathew, S. J. & Gorman, J. M. Molecular targets in the treatment of anxiety. *Biol. Psychiatry* **52**, 1008–1030 (2002).

27. Harvard Health Publishing. Benzodiazepines (and the alternatives) - Harvard Health. *Harvard Health* Available at: https://www.health.harvard.edu/mind-and-mood/benzodiazepines_and_the_alternatives. (Accessed: 30th August 2019)

28. SCREEN-WELL® Neurotransmitter library (10 plate set) - BML-2810 - Enzo Life Sciences. (2019). Available at: http://www.enzolifesciences.com/BML-2810/screen-well-neurotransmitter-library-10-plate-set/. (Accessed: 30th August 2019)

29. Bromley, J., Guyon, I., LeCun, Y., Säckinger, E. & Shah, R. Signature Verification using a 'Siamese' Time Delay Neural Network. in *Advances in Neural Information Processing Systems 6* (eds. Cowan, J. D., Tesauro, G. & Alspector, J.) 737–744 (Morgan-Kaufmann, 1994).

30. Baldi, P. & Chauvin, Y. Neural Networks for Fingerprint Recognition. *Neural Comput.* **5**, 402–418 (1993).

31. Koch, G., Zemel, R. & Salakhutdinov, R. Siamese neural networks for one-shot image recognition. in *ICML deep learning workshop* **2**, (2015).

32. Deng, J. *et al.* ImageNet: A large-scale hierarchical image database. in *2009 IEEE Conference on Computer Vision and Pattern Recognition* 248–255 (2009).

33. Bertinetto, L., Valmadre, J., Henriques, J. F., Vedaldi, A. & Torr, P. H. S. Fully-Convolutional Siamese Networks for Object Tracking. in *Computer Vision – ECCV 2016 Workshops* 850–865 (Springer International Publishing, 2016).

34. Zheng, Y., Liu, Q., Chen, E., Ge, Y. & Zhao, J. L. Exploiting multi-channels deep convolutional neural networks for multivariate time series classification. *Frontiers of Computer Science* **10**, 96–112 (2016).

35. Zhang, Y. & Duan, Z. Visualization and Interpretation of Siamese Style Convolutional Neural Networks for Sound Search by Vocal Imitation. in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 2406–2410 (IEEE, 2018).

36. Thyagarajan, A. Siamese Recurrent Architectures for Learning Sentence Similarity. in *AAAI-16* (2015).

37. van den Oord, A. *et al.* WaveNet: A Generative Model for Raw Audio. *arXiv [cs.SD]* (2016).

38. Huang, G., Liu, Z., van der Maaten, L. & Weinberger, K. Q. Densely Connected Convolutional Networks. *arXiv [cs.CV]* (2016).

39. Pascanu, R., Mikolov, T. & Bengio, Y. On the difficulty of training Recurrent Neural Networks. *arXiv [cs.LG]* (2012).

40. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. *arXiv [cs.CV]* (2015).

41. Srivastava, R. K., Greff, K. & Schmidhuber, J. Training Very Deep Networks. *arXiv [cs.LG]* (2015).

42. Senin, P. Dynamic Time Warping Algorithm Review. (2009).

43. Berndt, D. J. & Clifford, J. Using dynamic time warping to find patterns in time series. *KDD workshop* (1994).

44. fastdtw. *PyPI* Available at: https://pypi.org/project/fastdtw/. (Accessed: 30th August 2019)

45. Salvador, S. & Chan, P. Toward accurate dynamic time warping in linear time and

space. *Intelligent Data Analysis* **11**, 561–580 (2007).

46. Pluta, M. & Maksymilian, P. *Advanced light microscopy*. **1**, (Elsevier Amsterdam, 1988).

47. Nguyen, T. H. *et al.* Halo-free Phase Contrast Microscopy. *Sci. Rep.* **7**, 44034 (2017).

48. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

49. Hartigan, J. A. & Wong, M. A. Algorithm AS 136: A K-Means Clustering Algorithm. *J. R. Stat. Soc. Ser. C Appl. Stat.* **28**, 100–108 (1979).

50. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv [stat.ML]* (2018).

51. Xu, Z.-Q. J., Zhang, Y. & Xiao, Y. Training behavior of deep neural network in frequency domain. *arXiv [cs.LG]* (2018).

52. Blackman, R. B. & Tukey, J. W. The measurement of Power Spektra Doker, Publications. *Inc, New York* (1958).

53. Cover, T. & Hart, P. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* **13**, 21–27 (1967).

54. Vellucci, S. V. & Webster, R. A. Antagonism of the anticonflict effects of chlordiazepoxide by beta-carboline carboxylic acid ethyl ester, Ro 15-1788 and ACTH(4--10). *Psychopharmacology* **78**, 256–260 (1982).

55. Patel, J. B. *et al.* A novel potent non-benzodiazepine anxioselective agent with reduced dependence liability: ICI 198, 256. *Prog. Clin. Biol. Res.* **361**, 483–488 (1990).

56. Zhang, C. *et al.* A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS J. Photogramm. Remote Sens.* **140**, 133–144 (2018).

57. Sainath, T. N., Vinyals, O., Senior, A. & Sak, H. Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2015). doi:10.1109/icassp.2015.7178838

58. Donahue, J. *et al.* Long-term Recurrent Convolutional Networks for Visual Recognition and Description. *arXiv [cs.CV]* (2014).

59. Goodfellow, I. *et al.* Generative Adversarial Nets. in *Advances in Neural Information Processing Systems 27* (eds. Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D. & Weinberger, K. Q.) 2672–2680 (Curran Associates, Inc., 2014).

60. Antoniou, A., Storkey, A. & Edwards, H. Data Augmentation Generative Adversarial Networks. *arXiv [stat.ML]* (2017).

61. Oliphant, T. E. *A guide to NumPy*. **1**, (Trelgol Publishing USA, 2006).

62. Ioffe, S. & Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv [cs.LG]* (2015).

63. Agarap, A. F. Deep Learning using Rectified Linear Units (ReLU). *arXiv [cs.NE]* (2018).

64. Hadsell, R., Chopra, S. & LeCun, Y. Dimensionality Reduction by Learning an Invariant Mapping. in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* **2**, 1735–1742 (2006).

65. Adam, P. *et al.* Automatic differentiation in pytorch. in *Proceedings of Neural Information Processing Systems* (2017).

66. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. *arXiv [cs.LG]* (2014).

67. 3.2.4.3.1. sklearn.ensemble.RandomForestClassifier — scikit-learn 0.21.3 documentation. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html. (Accessed: 30th August 2019)

68. Jones, E., Oliphant, T. & Peterson, P. SciPy: Open Source Scientific Tools for Python. (2001).

Preamble to Chapter 4

Out of all the challenges I've come across during my PhD, none were greater than the DeepFish project. At first glance it feels like a huge disappointment to acknowledge that our approach remains as of yet unproven. But after some reflection I realize that only by setting sights high enough is real progress made. The idea behind the DeepFish project was simple in its formulation; construct a dictionary of Zebrafish behavior; where words are "motion primitives" from which "sentences" of behavior are formed. I understand now that I was thinking more like a reductionist and physicist and less like a biologist when I thought it would be simple to encode Zebrafish behavior.

Chapter 4 will go through some of our progress in forging this dictionary. But for now, I simply acknowledge that the promising results of chapter 3 owe their upbringing to the DeepFish project. We always suspected correlation distance was imperfect, but we may never have had enough motivation to embark on the endeavor of learning a custom distance metric (and performing highly replicated screens) if it hadn't become absolutely critical to do so in the context of comparing the high-dimensional encodings of Zebrafish behavior.

Even without reaching the summit, a mountain might reveal to its climbers glorious views and offer a new perspective of a complicated landscape extending past the horizon of its shadow. As of yet, that landscape hasn't fully revealed itself to us, with plenty of hills, forests, plate effects, and high-frequency components in the way. But we see a path forward, full of Siamese neural networks and fully randomized screens, and it may very well lead to lush and fertile land. I attempt to paint my best rendition of this vision in chapter 4.

Chapter 4:

Research in Progress: Deep Encoding of Larval Zebrafish Behavior

Leo Gendelev[1], Douglas Myers-Turnbull[1], Jack Taylor[1], David Kokel[1,2*], Michael J. Keiser[1,3*]

[1] Institute for Neurodegenerative Diseases, University of California, San Francisco, California 94143, USA

[2] Department of Physiology, University of California, San Francisco, California 94158, USA

[3] Departments of Pharmaceutical Chemistry and of Bioengineering & Therapeutic Sciences, University of California, San Francisco, California 94158, USA

***Corresponding author. E-mail: keiser@keiserlab.org, david.kokel@ucsf.edu**

4.1 Introduction

It's almost surprising how successful motion index has been despite its inherent limitations. It served us two bountiful phenotypes all while neglecting to encode information beyond bulk motion. Admittedly, our work in chapter 3 suggests that with a better distance metric, motion index might have even more riches in store for us. But from motion index's perspective, fish vigorously wiggling their tails in-place and fish swimming across the well might be indistinguishable. If there is even a remote chance that some of these potential behaviors and motions could be important drug-response signatures, it is worth developing a method that can effectively capture them.

We could have chosen to encode behavior based on existing work; the Zebrafish community has characterized many behavioral patterns and motions[1–5], that have been associated with neurological disease phenotypes; but these are likely not a complete repertoire. Even for just the well-studied phenotypes and behaviors, algorithmically encoding them in our 96-well, multi-animal setting, where the fish motion is not synchronized and somewhat stochastic is a daunting task. Further, the Kokel Lab has developed a robust sequence of in-house assays where the fish are stimulated in a myriad of ways. These assays have been pivotal for fingerprinting fish behavior because they cause the fish to respond in distinct and neuro-actively interesting ways[6–8]; but the full gamut of responses that the fish can potentially exhibit when stimulated by these unique assays might be larger than what has so far been illuminated by the Zebrafish community.

With these considerations, we take an unsupervised approach to learning a "dictionary" of zebrafish response "primitives", and fingerprinting the videos (or "stories") with these basic "words" of behavior. Our approach is visualized in **Figure 4.1**.

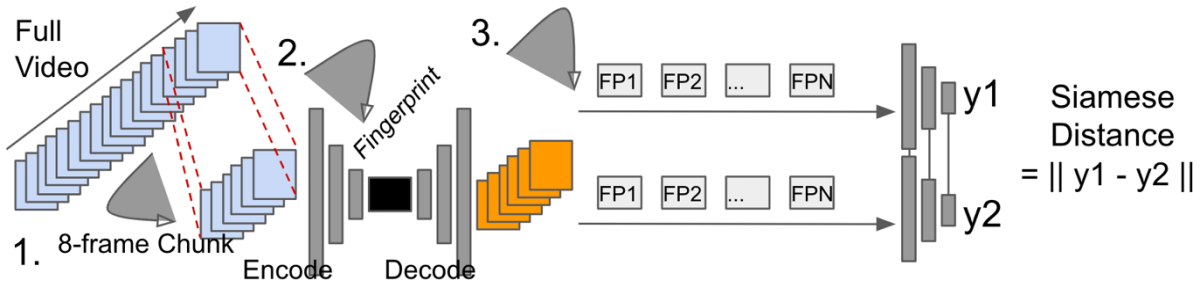


Figure 4.1 Unsupervised video encoding approach

A cartoon representation of our strategy for unsupervised video representation and subsequent Siamese distance metric training 1) A video of a drug-well treatment is broken up into 8-frame segments 2) We train an autoencoder on a large-dataset of segments to learn a hidden representation of 8-frame segments of video. 3) We concatenate all the "segment-fingerprints" from a video and train a Siamese neural network on our highly replicated dataset to learn a distance metric (analogous to Chapter 3).

Why autoencoders? Autoencoders are Neural Net (NN) architectures[9,10] used in many unsupervised learning applications[10–12], where instead of learning to *classify* data into bins, they instead learn an efficient *embedding* of the data. This kind of unsupervised pre-training[13] can be used for many subsequent purposes, including clustering or training a classifier. For example, in the popular Word2Vec[14] approach in the natural language processing domain, the first step is to embed the English vocabulary onto a space that captures some aspects of distance and relationships between words; subsequent training of LSTMs or other recurrent networks for text classification can be performed on this embedded text. This architecture isn't strictly an autoencoder, but the notion of learning an embedding of English vocabulary is somewhat analogous to our use of autoencoders in the context of zebrafish behavioral videos; we wish to embed behavior into a space that is meaningful in the context of fish behavior and can be used for subsequent phenotypic drug clustering and phenoblasting. In the text that follows, we will discuss each of the steps of our encoding strategy in more detail.

4.2 Training dataset for autoencoders

The first part of our pipeline (**Figure 4.1 step 1**), is to break up videos of drug-treated wells into 8-frame segments. This is done out of practical considerations; the full videos (20250 frames) are too long to use as inputs for Siamese distance training in their raw form. We break the video into segments so that we can embed each segment into a well-behaved latent space of motion "primitives". Any motion or behavior in any video should in principle be embeddable in this space. For constructing a training-dataset for the autoencoder, we take the highly-replicated Biomol screen (see chapter 3), and split it *by-plate* into training and validation sets, sampling 8-frame segments from these pools so as to balance segments with low versus high-motion content (**Figure 4.2**). We sample evenly across the 3 regions for a total of 1×10^6 segments for the training dataset, and 1×10^4 segments for the validation set.

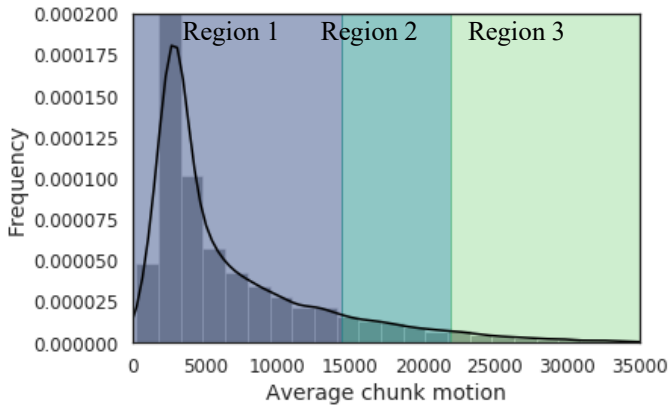


Figure 4.2 Sampling from a distribution of 8-frame segments

This is the distribution of 8-frame segments based on the average of motion across the 8 frames. We sample evenly from the three regions; low-motion, left (0-14413); average-motion, middle (14413-21913); and high-motion, right (14413-21913). These regions were calculated based on the mean and standard deviation of motion across the 8-frame segment dataset.

4.3 Removing noise from videos

In chapter 3, we observed that Siamese DenseNet architectures were able to easily overfit our datasets of raw motion index time-series (MI). By removing the high-frequency components of the signal, we were able to reduce overfitting and improve generalizability. With this insight, we looked for analogous ways of de-noising our video datasets. What spurious video artifacts might be the analogues of high-frequency components in MI? We thought the most likely culprits to be camera noise, water movement in wells, high-frequency movement of plates as a result of solenoid tapping, etc. Additionally, in the videos, the physical well-boundaries and plastic-artifacts from the 96-well plates are present; these artifacts can be exploited by our networks to learn features encoding for plate-location. To account for these potential artifacts, we sought to get rid of all the non-fish features from our videos. To do this, we used a semantic segmentation approach - the UNET architecture (**Figure 4.3a**)[15] - that has been applied with much success since its inception for various image segmentation applications across the biological sciences[16,17] To train UNETs, a labeled dataset of masks is required. We use the GIMP image editing software[18] to manually select fish from a hand-selected dataset of 100 frames with challenging fish occlusion situations and orientations from our training plates (**Figure 4.3b**). The result of this procedure are binary masks that are composed of values of "1" for regions of the image with fish, and values of "0" for background. When we first trained the UNETs, we observed that they sometimes confused circular well plastic-elements with rounded fish tails. We found that pre-training on a small set of 10 manually selected frames with these artifacts, followed by a main-training procedure on the rest of the 100 mask dataset (**Figure 4.3c**), helped to reduce this confusion and produced visually accurate whole-fish masks (**Figure 4.3d**).

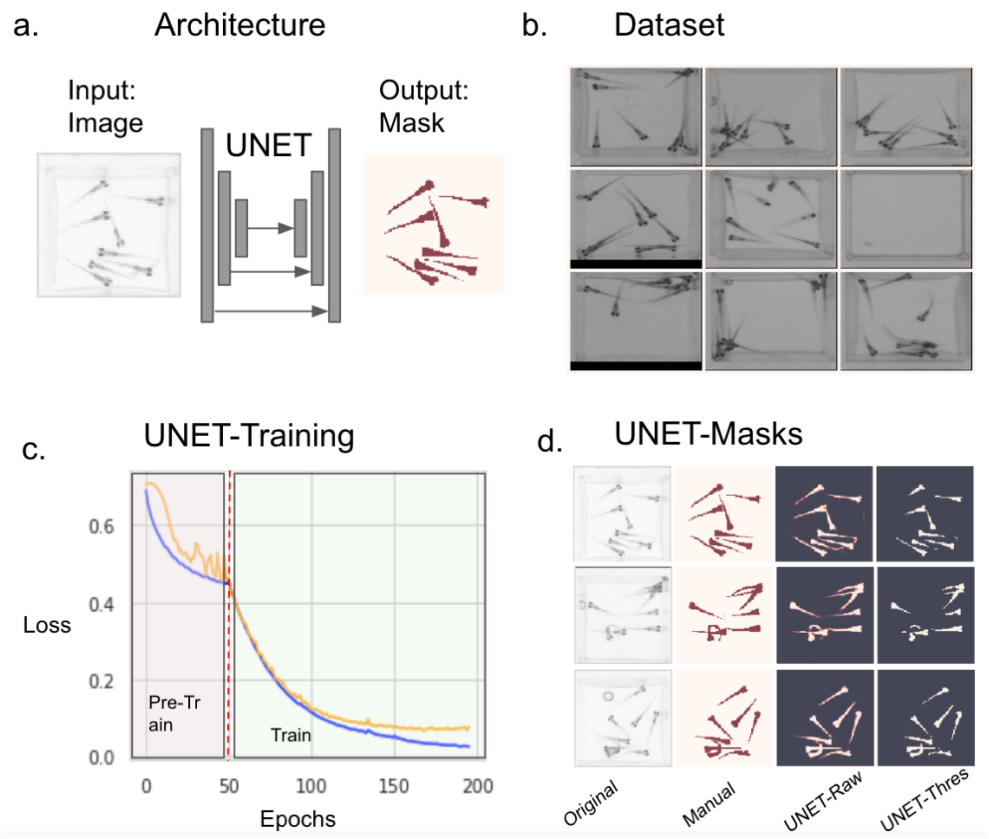


Figure 4.3 UNET Masking and Training

UNET-Mask dataset construction, training, and results. a) The UNET architecture[15] (cartoon representation), takes raw images as input, and is trained on a dataset of human-annotated masks to produce binary masks as output. b) Some examples shown from the manually curated dataset of challenging images. c) Pre-training and main-training loss curves (train loss; blue, validation loss; orange). The UNETs are first pre-trained on a set of 10 images and masks with plastic artifacts, then training is continued on the rest of the 100 masks. d) Some example masks from the validation set (original images; 1st column, human-annotated masks; 2nd-column, UNET-raw masks; 3rd column, UNET-thresholded at 0.5 masks; 4th column.

We explored other ways of removing camera artifacts. We realized that coordinate representations could be quite useful. They are the most parsimonious in terms of their feature space, and can be used with non-convolutional architectures, such as recurrent neural networks. Entire coordinate videos can even be fed directly into Siamese neural networks for end-to-end training. We wondered if a simple modification to the UNET architecture could be used for producing coordinate masks instead of whole-fish masks. To test this hypothesis, we manually annotated the same dataset of 100 images that we used for the whole-fish UNETs using GIMP, but

this time placing small circle masks around regions of interest in the image; specifically, the heads, bodies, tails, and tail-tips of the fish. It turned out that this UNET adaptation was quite robust and straightforward in terms of training and optimization (Figure 4.4).

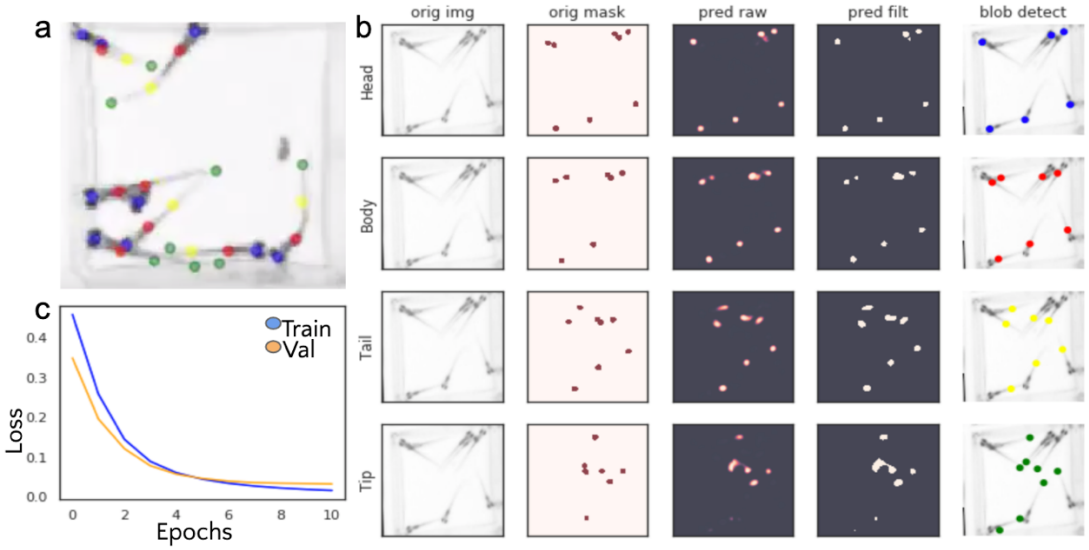


Figure 4.4 UNET Masking and Training

UNET-Coordinate mask dataset training and results a) Example from a fully trained model. b) Illustration of the dataset prep and performance of the UNET-Coordinate architecture. Original images (left column), are annotated with small circular masks (2nd from left column). The UNET is trained and raw masks are predicted (3rd from left column), and thresholded (2nd from right column). Open-cv blob-detection is applied to the thresholded masks to get the coordinates of the blobs for each of the possible fish-regions of interest (rightmost column, heads; blue, bodies; red, tails; yellow; tail-tips; green). c) UNET-Coordinate architecture training loss curves.

We kept the same architecture as our whole-fish UNETs; we only had to supply a different kind of mask (coordinate masks instead of whole-fish masks). Finally, we set the number of output channels as 4 (for the 4 possible regions of interest).

In principle, we could have used one of the many alternative approaches that already exist for coordinate extraction. There are deep neural net approaches such as DeepLabCut[19,20] for tracking multiple limbs of animals, idtracker[21] for tracking the center of mass of many animals, and YOLO[22,23] and Mask R-CNN[24] and Fast R-CNN[25] for tracking objects with bounding boxes. However, adapting any of these implementations for our high-throughput dataset of videos

composed of 8-fish per well where we'd like to track individual fish-body-parts would be quite challenging. Finally, we discovered it was straight-forward to adapt our UNETs to this task and it felt prudent to get an additional representation (coordinates) with recycled components.

4.4 Training autoencoders

I. Image-based Autoencoders

Our autoencoder architecture was inspired by a recent study in video-classification[26]. There, a 3D-convolutional architecture was designed to classify 16-frame clips from the UCF101 dataset[27] of 101 sports-action classes and was found to be more efficient and perform better or on-par with state-of-the-art LSTM[28] and LSTM-CNN Hybrid[29] networks. Here we adapt and modify their architecture for the purposes of auto-encoding video-fragments (**Figure 4.5**). We make some significant changes to the original architecture, with the biggest change being the addition of the decoding network to reconstruct video-fragments; other smaller changes are discussed in more detail in the methods section.

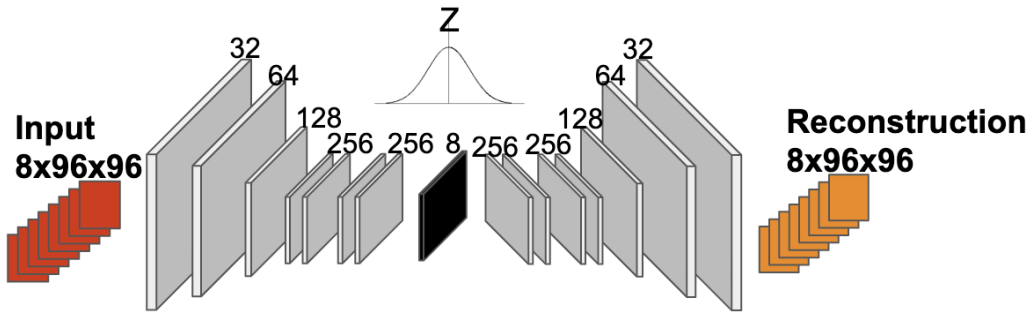


Figure 4.5 3D Convolutional Autoencoder Architecture

3D Convolutional autoencoder architecture inspired by recent work on video classification[26]. The time-dimension is not shown. An input video-segment of 8 frames is fed into the CNN, embedded into a hidden latent space of size 72, and training is done to minimize the loss between the input and reconstructed fragments. Time convolutions are performed only after the first 2 layers. The number of resulting channels after each convolution is displayed above the gray rectangle representing each layer.

Additionally, we considered the possibility that a better-behaved latent space could result in embeddings that perform better for subsequent tasks. With "vanilla" autoencoders (autoencoders that are trained simply by learning to minimize reconstruction loss), the learned features might be more useful for reconstructing images than for comparing behavior. In light of this we train variational autoencoders (VAEs)[30] as well, which have been successfully applied in many different domains and applications, including Generative Adversarial Nets (GANs)[31], generating images from text[32], and audio synthesis with generative WaveNets[33]. I will not attempt to reconstruct the beautiful mathematical underpinnings underlying VAEs here; I will simply state that in VAEs, a Kullback-Leibler (KL) Divergence term is added to the loss function, so that training is done not only to minimize the reconstruction error between input and output, but also such that the distribution of the embedded latent space follows a chosen distribution (often, and in our case, this is the Gaussian distribution). In practice, this insures that the hidden space is *well behaved*; it uses its neurons more evenly in the task of reconstructing input. The performance of our 3D-Convolutional Autoencoders both on raw-images and UNET masked images is visualized in **Figure 4.6**. We are able to get nice reconstructions (visually) with all architectures.

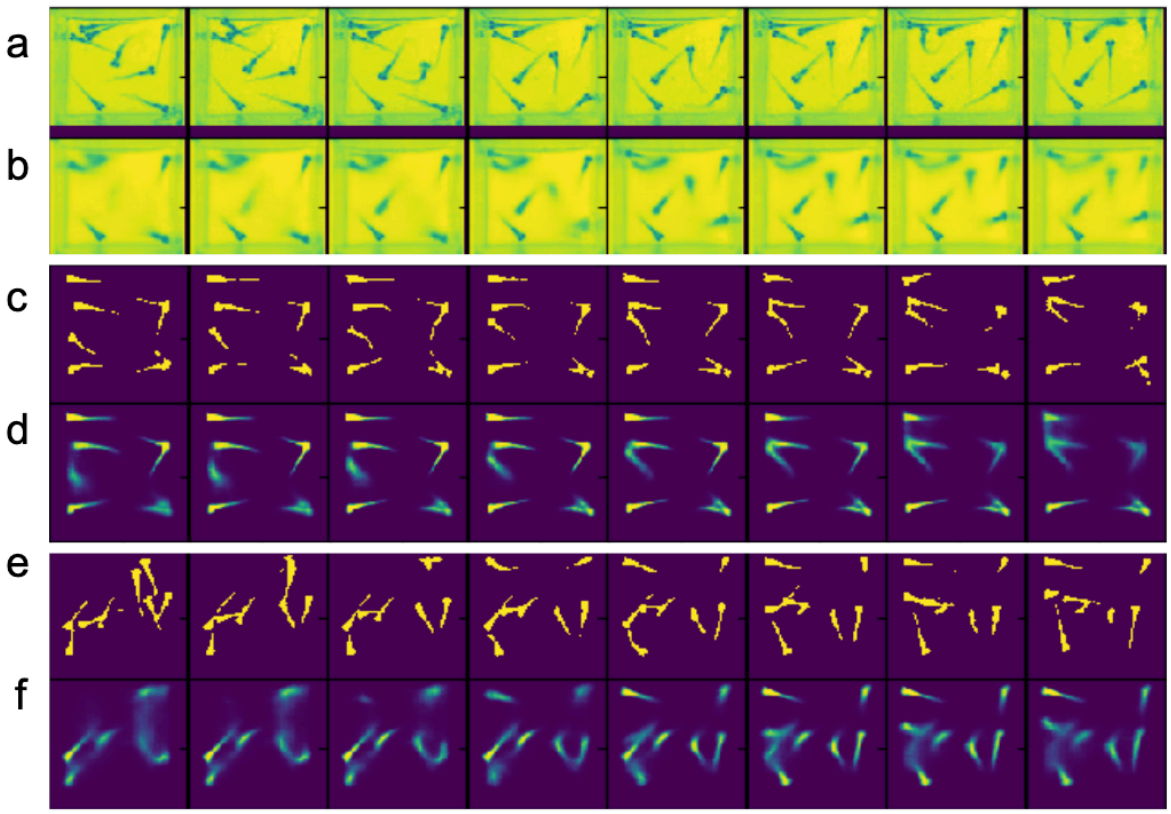


Figure 4.6 3D CNN Autoencoder Reconstructions

Original and reconstructed images are presented for 3 versions of the trained autoencoders. a,b) raw-image non-variational (original and reconstructed); c,d) UNET masked input non-variational; e,f) UNET masked variational.

II. Coordinate Autoencoders

Analogous to our work on segments of video, we explore the idea of embedding the UNET coordinate representations of fish video into a compressed, well-behaved space, using autoencoders. We use Gated Recurrent Units GRUs[34], a type of recurrent neural architecture with performance similar to Long Short-Term Memory Networks (LSTMs)[35,36], since they can accept multi-dimensional time-series inputs. We use a stacked GRU architecture to embed segments of 25 frames (corresponding to approximately 1 Zebrafish bout). We get good reconstructions (visually) with 256 hidden features (Figure 4.7). With the coordinates GRU Autoencoders, we can embed more frames at the cost of a larger hidden size.

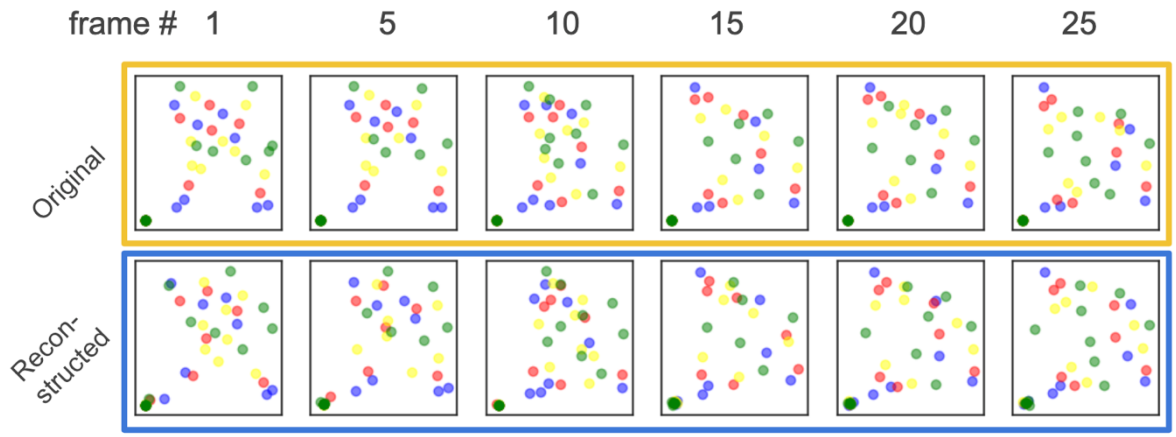


Figure 4.7 GRU Autoencoder Reconstructions

Original coordinate frames, plotted at an interval of each 5th frame from the 25-frame block, are encoded with a stacked-GRU autoencoder with 256 hidden neurons. The reconstructed coordinates are shown in the bottom panel.

4.5 Model Interpretability and visualization

Besides looking at reconstructions, one way to judge what an autoencoder is learning is visualizing the hidden layer. For our 3D CNN autoencoders, we can use the activation maximization[37–39] technique to visualize the input patterns that cause our 72 hidden neurons to maximally "fire". After 100 steps, we plot the results for each neuron (**Figure 4.7**), where we've taken the average over the 8 frames such that each of the 72 neurons is represented with its own individual 96x96 image. There is an obvious geometric pattern in the activations; each of the 8 rows has 9 possible spatial locations (4 corners, 4 edges, and 1 center-well); and each column seems to have one of 8 possible activation patterns. We believe the design choice of our architecture is responsible for these patterns. The last convolution right before the hidden layer uses a 3x3 filter, hence it learned to sense one of 9 possible spatial locations. It has 8 output channels, hence it learns 8 kinds of motions and fish-distribution patterns. In other words, the choice of filter and number of

channels right before the hidden layer has seemingly produced a latent space where neurons pick up on well-locations of fish with 8 different fish behavior "sensors".

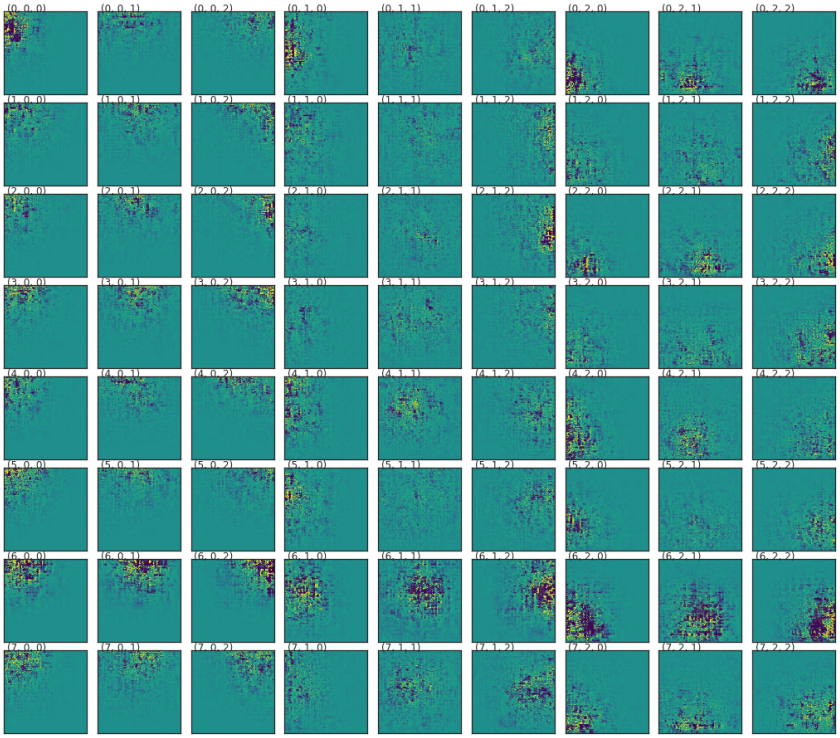


Figure 4.7 Activation maximization for 3D CNN (VAE)

Visualization of the hidden 72 neurons in the latent space of the fully trained VAE using the activation maximization technique. Each cell corresponds to a unique neuron. The rows appear to correspond to distinct motion types; the columns to one of 9 possible well locations.

A nice downstream result of performing this visualization was the discovery that the hidden neurons were not completely random; they were geometrically related. But we posit that for the purposes of encoding behavior, well-direction effects are undesirable. Preference of one corner over another, or one side over another, might not be a smart proxy for drug-induced behavior. For one, there could be physical directional preferences imposed by the experimental setup, which could lead to well-direction effects, making it all the more desirable to reduce them. By averaging over the corner, side, and middle-well features, we could bring the total number of hidden features down

from 72 to 24 (3 possible locations each with 8 possible behavior sensors). In the future, it would be interesting to explore a choice of autoencoder architecture that naturally performs such a directional "collapse". One such architecture might consist of polar-coordinate convolutions on polar-coordinate representations of video instead of cartesian; explicitly collapsing the possible range of angle to be between 0 and 45 degrees so as to remove directionality. One potential pitfall of collapsing directionality is if there are behaviorally-significant cases where fish flee to different corners of the well or purposefully group together in corners. Further studies need to be done to investigate how prevalent such behaviors are to inform this crucial choice of architecture or post-processing procedure.

We perform a case study with these 24-feature hidden embeddings, choosing to investigate an interesting behavior in Zebrafish related to fish moving in close contact with well walls; this phenomenon is known as "thigmotaxis" and is thought to be relevant for anxiety[40,41]. We find an example of such behavior in our dataset and visualize it (**Figure 4.8a**); a fish (circled) starts moving from the lower-right corner along the right wall towards the top-right corner of the well. We can track the activations of each of the 24 geometrically compressed neurons in the hidden layer as we feed in subsequent frames from this video (**Figure 4.8c).** As expected, the edge and corner neurons are activated more than the center-well neurons. There is some activation in one of the center-well channels likely because it catches a bit of fish tail. If we take the average across neurons, we get a trace that closely resembles the MI for this segment of video (**Figure 4.8b**). Together these observations highlight a key advantage of our autoencoders; they've learned an embedding of fish-behavior that can differentiate between distinct well-location movements. This virtue can be critical for successfully identify compounds that induce behaviors like Thigmotaxis.

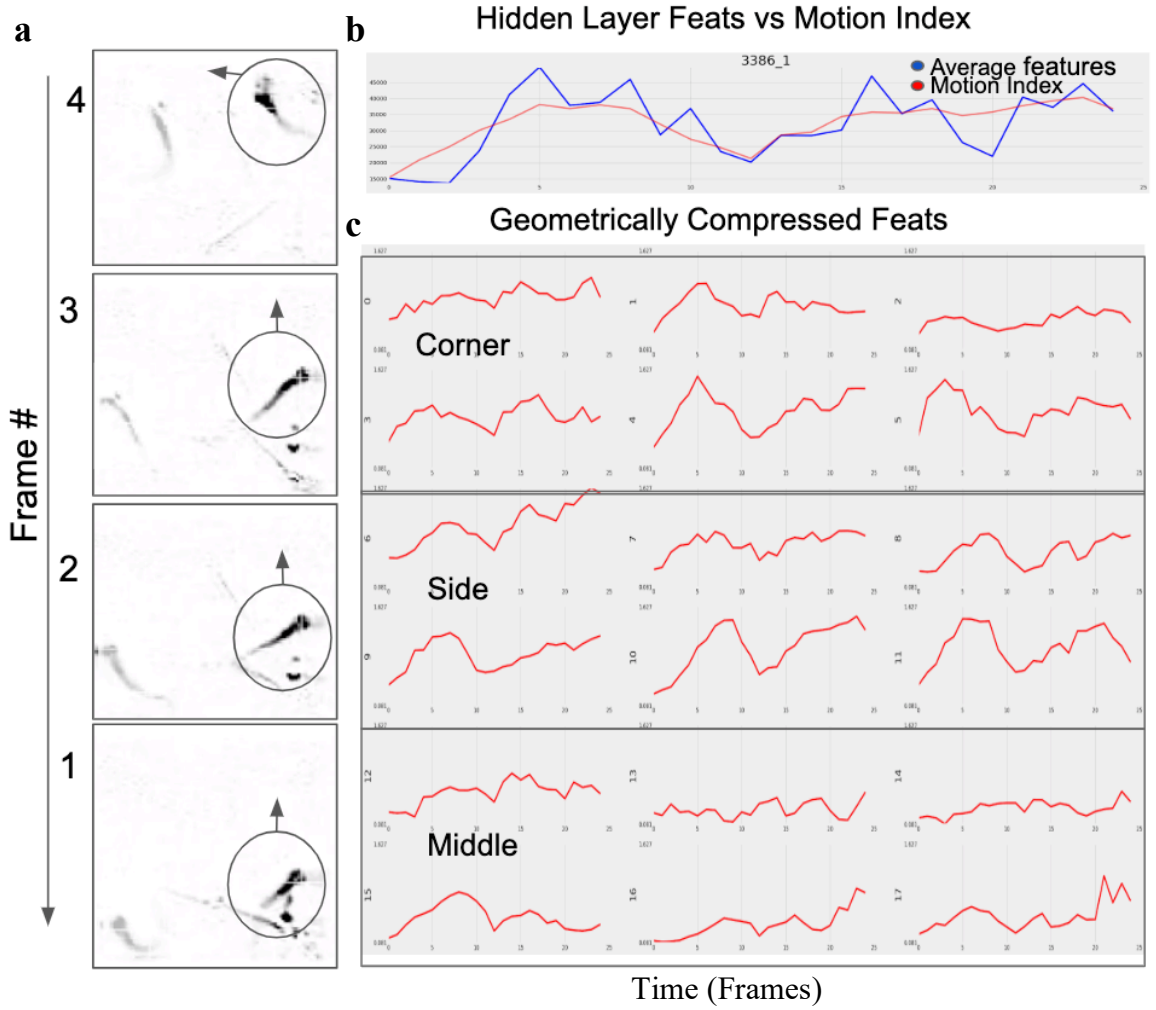


Figure 4.8 Edge-hugging (thigmotaxis) case-study

We investigate what geometrically-compressed hidden layer activation patterns can reveal in the context of edge hugging (thigmotaxis) behavior. a) Sequential frames selected from a ~30 second edge-hugging video (difference maps plotted to highlight moving fish only). The fish (circled) starts its motion from the lower-right corner and travels along the right wall towards the top-right corner of the well. b) Comparison of MI and average of hidden layer activations. c) Visualization of hidden layer activations by location.

4.6 Siamese distance learning with 3D encoded features

We were now at a point where our embeddings seemed to encode Zebrafish behavior in a meaningful and useable way, so we attempted to learn a new distance metric on them. Instead of encoding entire videos, which might contain a lot of redundancy, we sample frames only directly

around important stimuli during the assay. We include 39 stimuli events, and sample -24 to +48 (corresponding to roughly 1 bout before and 2 bouts after) frames around the stimulus (equivalent to 9 8-frame segments); this results in 351 8-frame segments of video. We use our trained 3D UNET Mask VAEs to embed these 351 8-frame segments into the 72 neuron hidden space; subsequently performing the geometric averaging to 24 distinct neurons per segment. The final fingerprints for 20250 long videos are 351 by 24 arrays. How do we use these arrays for Siamese training?

As a proof of concept, we start with a Sia-MLP, analogous to chapter 3. There, we used pairs of 20250 frame MI as inputs to the SNNs. Here, we project our 351x24 fingerprints into one dimension, resulting in 8424-long inputs. Thus the input layer of the Sia-MLP architecture has 8424 neurons instead of the 20250 neurons used in chapter 3. We train this Sia-3D network on the same dataset as our most promising model from chapter 3 (the smoothed, random forest filtered, all-well negative dataset). The Sia-3D does a reasonably good job of separating out positive and negative pairs (**Figure 4.9a**), and gets reasonable AUC/PRC performance, albeit substantially worse than Sia-MI from chapter 3, 0.86 vs 0.93, respectively (**Figure 4.9b,c**). Correlation distance and Euclidean distance calculated on the 3D embedded features perform atrociously. This highlights the importance of learning a custom distance metric, especially when dealing with high-dimensional, embedded representations of behavior.

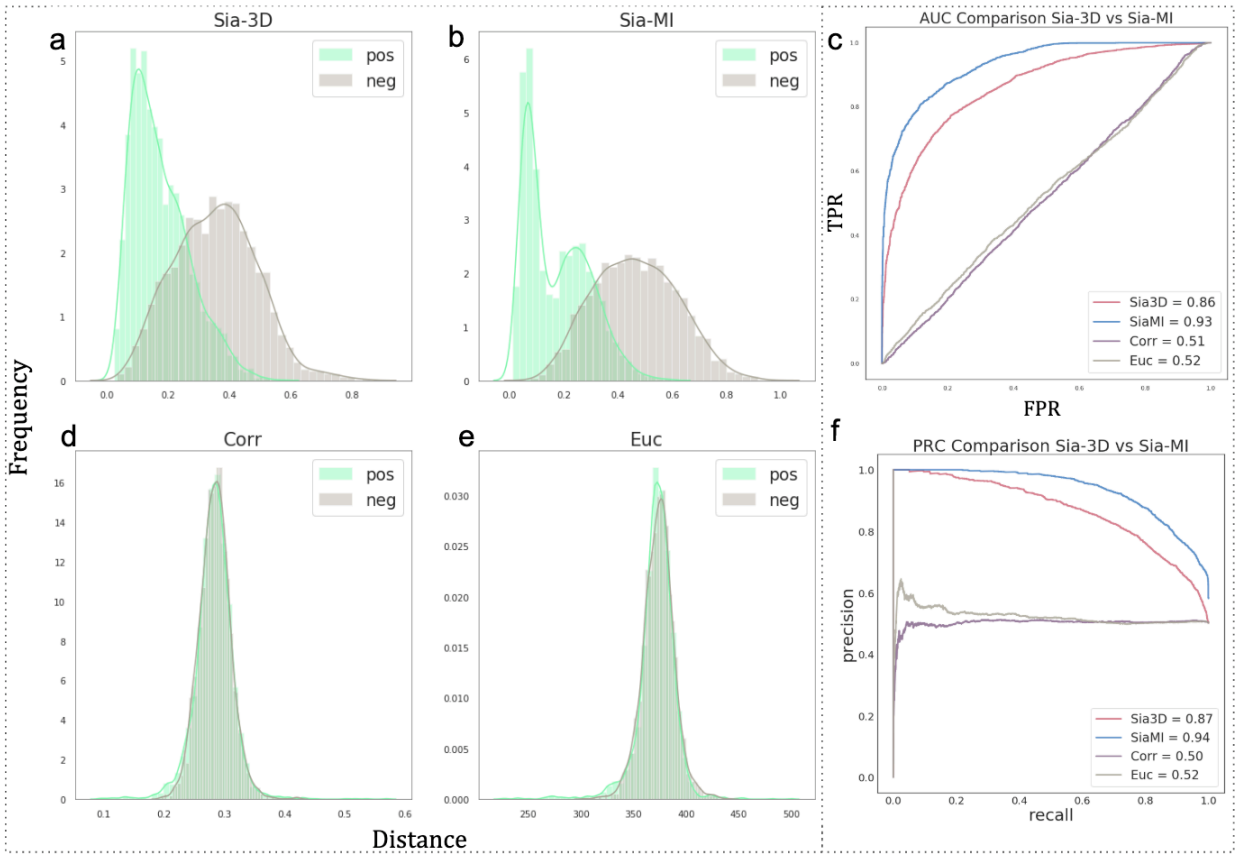


Figure 4.9 Sia-3D vs Sia-MI comparison

Sia-3D performance compared to Sia-MI, and correlation and Euclidean distance calculated on the 3D embeddings, for the highly-replicated, RF-filtered, all-well negatives dataset from chapter 3. a,b,d,e) positive-drug replicate pairs vs different drug pairs; c,f) AUC,PRC curves, respectively (red=3D, blue=MI, purple=Correlation, gray=Euclidean).

Next we investigate whether the Sia-3D distance can be used to cluster the high-replicated dataset (**Figure 4.10**) using the same methods (K-means to cluster, UMAP to embed onto 2 dimensions for visualization) as we used for visualizing the Sia-MLP clusters in chapter 3. We notice less pronounced structure in the clustering compared to Sia-MI (see **Figure 3.10a**). Since it's difficult to visualize the high-dimensional embeddings, we show the motion index time series of 10 drugs from cluster 15 (**Figure 4.10b**, red circles) for comparison. In terms of MI, the cluster is problematic; there are drugs that have the lethal phenotype as evidenced by flat-MI (GBR, DH, Loratadine, for example), but there are also several drugs with non-lethal phenotypes (Propranolol, B-CCE,

Clopidogrel, and Levallorphan). For a robust distance metric, we'd like to be able to discriminate between lethal and non-lethal drugs; there is certainly much room for improvement for our 3D distance metric.

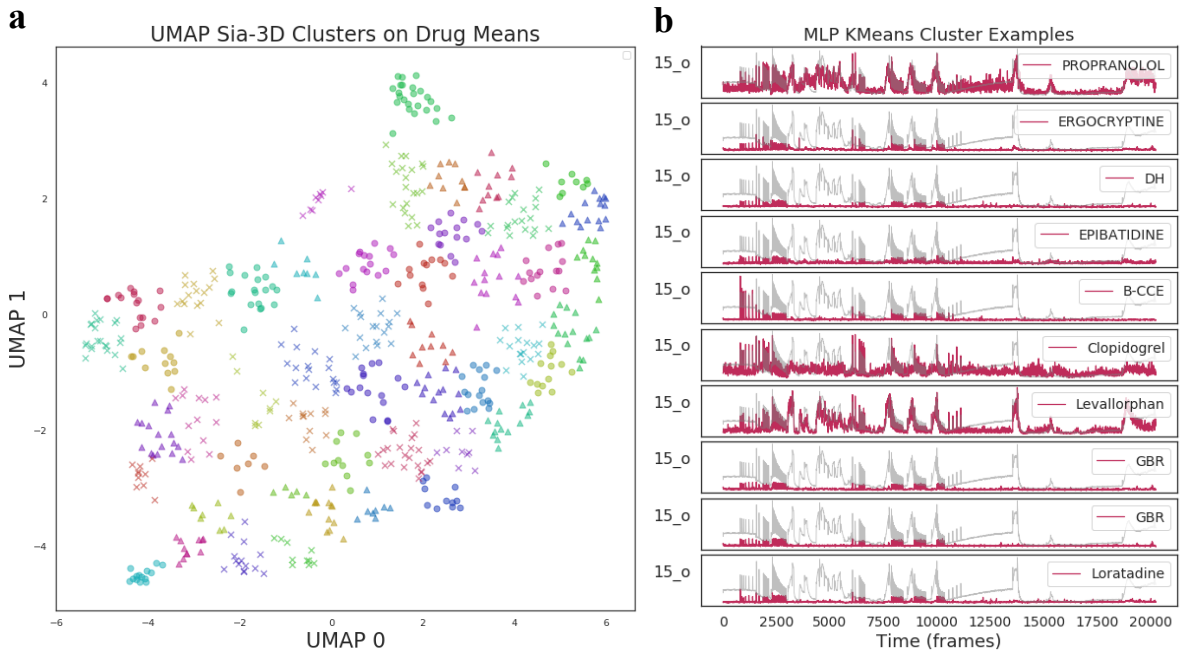


Figure 4.10 Sia-3D Clustering

Sia-3D clustering performed on the means across replicates from the highly-replicated screen on the embedded (3D-VAE UNET Masks fingerprints). a) UMAP embedding on fingerprints, colored and marked by the Kmeans clusters. b) For the drugs that have been grouped together into cluster 15, we visualize their time-series representations (red circles, middle-left part of the UMAP).

4.7 Discussion

We have in this chapter introduced an unsupervised approach for learning more advanced representations of Zebrafish behavior from videos. We are able to demonstrate we can do the following: a) denoise videos by extracting whole-fish masks with UNETs, b) extract body-part-specific coordinate masks for fish with UNETs, c) train 3D convolutional autoencoders to embed 8-frame segments of video, d) train GRU autoencoders to compress 25-frame segments of coordinates using 256 hidden layer features, e) provide evidence that the hidden activations from the 3D-CNN

autoencoders have learned to discern between corner, edge, and center-well positions and motions of fish, and show that averaging over the activations yields motion index, and finally f) show that we can modify the Sia-MLP architecture from chapter 3 to work on the high-dimensional embeddings.

Current limitations and future directions

Our first attempts at training a Sia-3D distance metric has produced mixed results; the AUC/PRC for separating same-drug replicates from different drugs is 0.86 and 0.87, respectively, which would indicate reasonable performance, but clustering performance seems relatively poor. Here we suggest several possible explanations for this.

I. Overfitting

In chapter 3 we saw that DenseNets were able to overfit on this dataset. We have already suggested that to fully control for overfitting, we need the fully randomized screen. Analogous to our smoothing procedure for motion index in chapter 3, we have attempted to use UNETs to denoise the images prior to training to circumvent this problem but it is possible that this hasn't removed all types of noise from the images. Perhaps there is still spurious high-frequency camera-noise or plate-vibrations that remain even after masking. It's likely the case that the only real solution for fully getting rid of overfitting is by training on the fully randomized screens.

II. Model selection

In our approach, we make many choices for architectures that might not be ideal and performance could be significantly improved with even the most obvious optimization. For instance, we attempt to encode 8-frame segments of video with the 3D-CNNs. Perhaps 8 frame segments of motion are just too short to encode behaviors that might actually occur on the scale of bouts (25-30 frames). Optimizing this critical choice of segment length should be a top priority for

subsequent work. The other critical choice is the Siamese architecture we use; the MLPs seem to work great on motion index, but they're probably not ideal for high-dimensional embedded video fingerprints. Here, we envision that a carefully designed convolutional or recurrent architecture should be able to greatly boost performance.

III. *Deeper* problems

It would be disappointing but not entirely surprising if the ways that fish distribute themselves around wells is not always significant from the behavioral and phenotypic standpoint. Our 3D-CNNs expend most of their learning-energy towards reconstructing these patterns. Perhaps, this distracts them too much from the really important behaviors and encourages them to overfit on potentially irrelevant fish-distribution patterns. An analogy is trying to track a specific car, driving erratically, in a sea of highway traffic during rush-hour. An autoencoder that takes in raw traffic camera footage and is tasked with reconstructing segments of traffic video might rarely encounter features that have any relationship at all with the swerving, speeding outlier, and might only learn general aspect of rush-hour traffic instead. A smarter approach might be to first learn what the "general traffic features" are, filter them out, and attempt to encode the remaining features - including the erratic car. In our realm of Zebrafish videos, this might be equivalent to first filtering out the static or motionless fish, and letting the 3D-CNNs focus on moving fish; or first learning features that are common to most videos and filtering those out, since they would likely not be useful for learning drug-specific encodings. In principle, neural networks should be able to perform feature selection automatically – that is, selecting drug-specific features and filtering out general ones – but this relies on having large enough datasets. With our limited datasets, this nice benefit of neural networks might be unrealized.

IV. Final thoughts

To conclude, we believe we've made important progress on finding improved representations of Zebrafish behavior, but much work still remains. At the least, we have an actionable pipeline which can be optimized further and benchmarked on the fully randomized screens. In chapter 3 we establish a context for learning custom distance functions for comparing MI; we should be able to extend it to this more challenging but potentially more-rewarding arena of high-dimensional embeddings.

We foresee the possibility that we might eventually get better classification and phenoblasting performance for specific *classes* of drugs with the Sia-3D distance functions. Perhaps certain classes of drugs and behaviors are just naturally predisposed for MI encoding; for sedating compounds this is a likely possibility. But for other classes of drugs – ones that cause the fish to behave in more unpredictable ways – learning representations that encode for more than just bulk motion seems prudent. For practical drug discovery purposes, we envision using the more advanced representations in conjunction with MI. With a combination of these representations we should be able to cast the widest net and provide the broadest coverage over the entire Zebrafish phenotypic and pharmacological landscape.

Acknowledgements

We thank members of the Keiser lab for insightful scientific discussion about various machine learning concepts that have helped inform key aspects of the research conducted in this chapter; this list includes but is not limited to: Kangway Chuang, Garrett Gaskins, Elena Caceres, Nick Mew, and Jacob Pfau. We thank the following members of the Kokel lab; Cole Helsell for help with troubleshooting experiments and datasets; Chris Ki for help with datasets; and Matt McCarroll for helpful scientific discussions.

References

1. Kalueff, A. V. *et al.* Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond. *Zebrafish* **10**, 70–86 (2013).
2. Cachat, J. *et al.* Three-dimensional neurophenotyping of adult zebrafish behavior. *PLoS One* **6**, e17597 (2011).
3. Miklósi, A. & Andrew, R. J. The zebrafish as a model for behavioral studies. *Zebrafish* **3**, 227–234 (2006).
4. Egan, R. J. *et al.* Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish. *Behav. Brain Res.* **205**, 38–44 (2009).
5. Green, J. *et al.* Automated high-throughput neurophenotyping of zebrafish social behavior. *J. Neurosci. Methods* **210**, 266–271 (2012).
6. Bruni, G. *et al.* Zebrafish behavioral profiling identifies multitarget antipsychotic-like compounds. *Nat. Chem. Biol.* **12**, 559–566 (2016).
7. Kokel, D. *et al.* Rapid behavior-based identification of neuroactive small molecules in the zebrafish. *Nat. Chem. Biol.* **6**, 231–237 (2010).
8. Kokel, D. & Peterson, R. T. Chapter 22 - Using the Zebrafish Photomotor Response for Psychotropic Drug Screening. in *Methods in Cell Biology* (eds. Detrich, H. W., Westerfield, M. & Zon, L. I.) **105**, 517–524 (Academic Press, 2011).
9. Baldi, P. Autoencoders, Unsupervised Learning, and Deep Architectures. in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning* (eds. Guyon, I., Dror, G., Lemaire, V., Taylor, G. & Silver, D.) **27**, 37–49 (PMLR, 2012).
10. Vincent, P., Larochelle, H., Bengio, Y. & Manzagol, P.-A. Extracting and composing robust features with denoising autoencoders. in *Proceedings of the 25th international conference on Machine learning* 1096–1103 (ACM, 2008).
11. Makhzani, A., Shlens, J., Jaitly, N., Goodfellow, I. & Frey, B. Adversarial Autoencoders. *arXiv [cs.LG]* (2015).
12. Masci, J., Meier, U., Cireşan, D. & Schmidhuber, J. Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction. in *Artificial Neural Networks and Machine Learning – ICANN 2011* 52–59 (Springer Berlin Heidelberg, 2011).
13. Erhan, D. *et al.* Why Does Unsupervised Pre-training Help Deep Learning? *J. Mach. Learn. Res.* **11**, 625–660 (2010).
14. Mikolov, T., Chen, K., Corrado, G. & Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv [cs.CL]* (2013).

15. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv [cs.CV]* (2015).

16. Li, X. *et al.* H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation From CT Volumes. *IEEE Trans. Med. Imaging* **37**, 2663–2674 (2018).

17. Oktay, O. *et al.* Attention U-Net: Learning Where to Look for the Pancreas. *arXiv [cs.CV]* (2018).

18. GIMP. *GIMP* Available at: https://www.gimp.org/. (Accessed: 31st August 2019)

19. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).

20. Welcome to idtracker.ai's documentation! — idtrackerai v3.0.17-alpha Manual. Available at: https://idtracker.ai/en/latest/. (Accessed: 31st August 2019)

21. Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S. & de Polavieja, G. G. idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nat. Methods* **11**, 743–748 (2014).

22. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: Unified, real-time object detection. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 779–788 (2016).

23. Redmon, J. & Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv [cs.CV]* (2018).

24. He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask r-cnn. in *Proceedings of the IEEE international conference on computer vision* 2961–2969 (2017).

25. Girshick, R. Fast r-cnn. in *Proceedings of the IEEE international conference on computer vision* 1440–1448 (2015).

26. Tran, D., Bourdev, L., Fergus, R., Torresani, L. & Paluri, M. Learning spatiotemporal features with 3d convolutional networks. in *Proceedings of the IEEE international conference on computer vision* 4489–4497 (2015).

27. Soomro, K., Zamir, A. R. & Shah, M. UCF101: A Dataset of 101 Human Actions Classes From Videos in The Wild. *arXiv [cs.CV]* (2012).

28. Srivastava, N., Mansimov, E. & Salakhudinov, R. Unsupervised Learning of Video Representations using LSTMs. in *International Conference on Machine Learning* 843–852 (2015).

29. Yue-Hei Ng, J. *et al.* Beyond short snippets: Deep networks for video classification. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 4694–4702 (2015).

30. Kingma, D. P. & Welling, M. Auto-Encoding Variational Bayes. *arXiv [stat.ML]* (2013).

31. Goodfellow, I. *et al.* Generative Adversarial Nets. in *Advances in Neural Information Processing Systems 27* (eds. Ghahramani, Z., Welling, M.,

Cortes, C., Lawrence, N. D. & Weinberger, K. Q.) 2672–2680 (Curran Associates, Inc., 2014).

32. Zhang, H. *et al.* Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. in *Proceedings of the IEEE International Conference on Computer Vision* 5907–5915 (2017).

33. Engel, J. *et al.* Neural audio synthesis of musical notes with WaveNet autoencoders. in *Proceedings of the 34th International Conference on Machine Learning - Volume 70* 1068–1077 (JMLR.org, 2017).

34. Chung, J., Gulcehre, C., Cho, K. & Bengio, Y. Gated Feedback Recurrent Neural Networks. in *Proceedings of the 32nd International Conference on Machine Learning* (eds. Bach, F. & Blei, D.) **37**, 2067–2075 (PMLR, 2015).

35. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).

36. Chung, J., Gulcehre, C., Cho, K. & Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv [cs.NE]* (2014).

37. Erhan, D., Bengio, Y., Courville, A. C. & Vincent, P. Visualizing Higher-Layer Features of a Deep Network. (2009).

38. Mahendran, A. & Vedaldi, A. Visualizing Deep Convolutional Neural Networks Using Natural Pre-Images. *arXiv [cs.CV]* (2015).

39. Yosinski, J., Clune, J., Nguyen, A., Fuchs, T. & Lipson, H. Understanding Neural Networks Through Deep Visualization. *arXiv [cs.CV]* (2015).

40. Basnet, R. M., Zizioli, D., Taweedet, S., Finazzi, D. & Memo, M. Zebrafish Larvae as a Behavioral Model in Neuropharmacology. *Biomedicines* **7**, (2019).

41. Schnörr, S. J., Steenbergen, P. J., Richardson, M. K. & Champagne, D. L. Measuring thigmotaxis in larval zebrafish. *Behav. Brain Res.* **228**, 367–374 (2012).

Appendix A:

A.1 Supplementary material for Chapter 2

Supplementary Note 1

$GABA_AR$ ligands produce paradoxical excitation in zebrafish. Compounds with weak phenoscores ($x < 0.51$) included one $GABA_B$ receptor agonist, one PAM of δ-subunit containing $GABA_A$Rs, two non-BZ-site ligands, three structurally-related $GABA_AR$ orthosteric agonists, and seven BZ-site $GABA_AR$ PAMs (**Fig. 1g**). For these compounds, the average phenoscores were significantly less than the positive controls ($P < 0.01$, Kolmogorov-Smirnov test, **Supplementary Figure 20a**), suggesting that these compounds did not phenocopy etomidate. For example, the highest scoring ocinaplon treatment produced a behavioral profile that resembled the negative controls (**Supplementary Figure 2**). These data suggest that a variety of GABAergic compounds do not cause sedation and paradoxical excitation.

Compounds with intermediate phenoscores ($0.51 < x < 0.71$) included several types of $GABA_AR$ PAMs including thiopental, carboetomidate, THDOC, alfaxalone, diazepam, and valerenic acid. The highest scoring profiles produced by some of these compounds (including alfaxalone, thiopental, and tracazolate) showed a barely detectable statistically significant difference compared to the positive controls ($0.01 < P < 0.05$). The highest scoring profiles of animals treated with diazepam, and valerenic acid were significantly lower than the positive controls ($P < 0.01$, Kolmogorov-Smirnov test, **Supplementary Figure 20a**), however these treatments produced interesting

intermediate effects on sedation and paradoxical excitation. For example, although the highest-scoring diazepam treatment was strongly sedating in most assays, it produced eASRs that were relatively weak and inconsistent (**Supplementary Figure 2**). These data suggest that a variety of PAMs have intermediate effects on sedation and paradoxical excitation.

Interestingly, although DOC and progesterone are neurosteroid precursors, they were among the most potent compounds tested (**Fig. 1g)**. As expected, progesterone's etomidate-like phenotype was suppressed by dutasteride, a 5-alpha-reductase inhibitor that blocks the metabolic conversion of progesterone to allopregnanolone, suggesting that these compounds were converted to active neurosteroids (**Supplementary Figure 5**).

Supplementary Note 2

Target prediction using SEA. We used the Similarity Ensemble Approach (SEA) to predict targets based on 'guilt-by-association' enrichment factor scores (EFs). These EFs were first developed for predicting adverse drug interactions [1], and balance the overall strength of a given target-to- compound-set association by correcting for the frequency that specific targets are predicted over random compounds sets in the screen [2]. Here, we used EFs to predict targets for the compounds that caused eASRs in the zebrafish. SEA identified 15 compounds with enriched target predictions for mGluRs (**Supplementary Table 9**, **Supplementary Figure 12a**). We chose eight of these compounds to reorder and retest and found that four of them reproducibly caused eASRs *in vivo* (**Supplementary Figure 12b, Supplementary Table 9**). Next, we tested five of these compounds as agonists and antagonists for activity at seven human mGluRs (mGluR1-6 and mGluR8). However, none of the compounds showed strong functional effects against mGluRs *in vitro* (**Supplementary Figure 12c**), suggesting that the 2compounds did not act via mGluRs *in vivo*.

To further test the mGluR hypothesis, we tried to phenocopy etomidate in dose-response experiments with a panel of structurally-diverse mGluR ligands and ligand combinations (**Supplementary Table 9**). Although MPEP, a mGluR5 antagonist, reproducibly caused eASRs, MPEP-induced eASRs were substantially lower in magnitude than etomidate-induced eASRs, and MPEP-induced eASRs only occurred in a narrow concentration range (**Supplementary Figure 12b, Supplementary Table 9**). Therefore, although MPEP weakly phenocopied etomidate, we found no further evidence that hit compounds targeted mGluRs, as predicted by SEA. SEA predicted that $GABA_AR$ was a target of four hit compounds (**Supplementary Figure 11**). We tested three of them (5658603, 5142031 and 7145248) and found that one (5658603) potentiated $GABA_AR$ *in vitro* (**Fig. 2f, red arrow)**. Curiously, we noted that SEA failed to predict $GABA_AR$ as a target for most hit compounds that tested positive in the $GABA_AR$ FLIPR assay (**Fig. 2f**), underscoring the value of behavior-based screens for identifying bioactive compounds with poorly annotated chemical structures.

Supplementary Note 3

$GABA_AR$ and HTR6 ligands likely converge on a common neural substrate. To determine if HTR6 antagonists activated the same neurons as GABAergic ligands, we took three approaches. First, we looked for overlap between 5HT immunohistochemistry and the eASR substrate neurons (Supplementary Figure 13f). Consistent with previous reports, we observed strong 5-HT staining in the telencephalon, pineal gland, hindbrain, and dorsal raphe nuclei [3]. In addition, we observed bilateral 5-HT staining in tracts that converged on the midline of the caudal hindbrain at the same location of the putative eASR substrate neurons in the caudal hindbrain (Supplementary Figure 13f). These tracts likely originated from the dorsal raphe, but we could not trace their origin definitively.

Second, we visualized HTR6 mRNA expression by RNAscope but could not detect reproducible expression patterns (Supplementary Figure 14), suggesting that HTR6 mRNA is not abundantly expressed. Finally, we tested for pharmacological interactions between GABAergic and serotonergic ligands. As expected, pretreatment with the $GABA_AR$ antagonist PTX rescued etomidate-treated animals, increasing and decreasing the magnitude of the violet light and eASR phenotypes, respectively (Supplementary Figure 15a). Similarly, PTX rescued the GABAergic compound 5658603, and partially rescued compounds 701338, and 5942595, albeit to a lesser extent than etomidate (Supplementary Figure 15a). PTX also partially rescued the behavioral phenotypes of several HTR6 antagonists including BGC 20-761, 6029941, 6028165, 6030006, and 6013263 (Supplementary Figure 15b). By contrast, EMDT oxalate, a HTR6 agonist, did not suppress eASRs caused by HTR6 antagonists (Supplementary Figure 15c), suggesting that the effects of HTR6 antagonists are not easily reversed. Together, these data suggest that $GABA_AR$ agonists and HTR6 antagonists likely cause eASR behaviors via different targets that converge on a common neural substrate in the zebrafish hindbrain.

Supplementary Discussion

Although etomidate and propofol are human anesthetics, the hit compounds identified in this study may not be useful as human anesthetics. One reason, is that the primary screen in zebrafish did not include behavioral correlates key anesthetic effects including analgesia and amnesia. Furthermore, only a minority of the hit compounds suppressed the TrpA1-induced pain-related assay (**Fig. 6b**), suggesting that many of the compounds may not cause the analgesic effects associated with human anesthesia. Another reason is hat paradoxical excitation is an unwanted side effect of anesthetic drugs. Even if the hit compounds translated to mammals (causing both sedation

and paradoxical excitation), additional studies would be necessary to determine if the paradoxical excitation phenotype could be overcome at higher concentrations or via medicinal chemistry. For example, the 21 analogs of compound 7013338 showed variable efficacies (**Fig. 6e**), suggesting that it may be possible to use medicinal chemistry to increase or decrease eASR activity. These shifts likely correspond specific effects on one or more molecular targets that would need to be identified with more sensitive functional assays, such as electrophysiological experiments, of recombinantly expressed receptor subtypes. Future studies may seek to identify different ligands that sedate zebrafish without causing eASRs, or eASRs may be used as a counter screen for other potential anesthetic lead compounds. Presumably, such compounds would work through different mechanisms than etomidate, propofol, and the other compounds identified in this study, and would further improve our understanding of GABAergic signaling, anesthesia, and paradoxical excitation.

Supplementary Methods

Behavioral assays for pERK. Animals were treated with DMSO or drug for one hour then exposed to a low-volume acoustic stimulus every 10 seconds for 10 minutes. Optovin-treated animals were stimulated with violet light for one second every 10 seconds for 10 minutes. Videos were recorded to measure response to the stimulus and quantified by motion index (MI). Animals were immediately fixed in 4% PFA in PBS.

High speed imaging. Digital video was recorded at 500-1000 frames per second using an Integrated Design Tools NX5-S1 digital camera. Assay duration was 500-1000 ms. Low amplitude tap stimuli was delivered as described above. Videos and still images were assembled using Fiji (imageJ) and Photoshop (Adobe).

RNAscope. Zebrafish (6 dpf) were fixed using 10% normal buffered formaldehyde overnight at 4˚C, dehydrated in 100% methanol and stored at -20˚C until the assay was performed. Fluorescent in situ hybridization was performed using the RNAScope kit (Advanced Cell Diagnostics)[4] with the following modifications: target retrieval was performed using 150 mM Tris pH 9.5 for 15 minutes at 70˚C, washes were performed using 0.2x SSCT (0.01% Tween-20, 3mM NaCl, 0.3mM TriNa-citrate, pH7), samples were stored overnight after probe hybridization and 2x 5 minute washes in 5x SSC (75 mM NaCl, 7.5 mM TriNa-citrate, pH7) at room temperature before amplification and detection steps were performed the following day.

Viability. Animals were scored manually to determine viability after anesthetic treatments. 100 animals were scored per condition at 10 minutes, 1 hour and 5 hours post treatment. Strong heart rate was observed under a dissecting scope (Leica M-80) for our viability score.

Structural Clustering. Structural clustering was performed on the top 125 hit compounds using the rdkit function FingerprintMols package written for python. Tanimoto similarity function was used with a threshold of 0.25 to define clusters and visualized using the scipy hierarchy dendrogram function.

SEA and EF calculations. Here we describe our computational pipeline: 1) Use the reference trace to discover the top 125 hit compounds (most similar phenotypically related to etomidate). 2) Organize hit compounds into hierarchical supersets of increasing numbers of hit compounds. Use SEA analysis to generate target predictions for each of the compounds in the sets. Perform enrichment factor calculations on the sets, which attempt to correct the occurrence of target predictions for a set of compounds by comparing to a background distribution.[2] To do so we generated 10,000 sets of 200 random screening compounds each, and applied the following formula to calculate the enrichment of target *y* for set *x*: E_xy = n*N / (A * T), where *n* is the number of times target *y* is predicted for set *x* compounds by SEA, A is the number of times any target shows

up for set *x*, T is the number of times target *y* shows up for any set, and N is a normalization factor equal to the product of all the targets and all the sets.

Determination of phenotypic thresholds and significance. For each ligand, we selected the dose that gave the highest average phenoscore, and for that dose, we performed a two-sample Kolmogorov-Smirnov (KS) test to calculate the KS statistic against the 12 positive control replicates of etomidate @ 6.25 μM using the scipy function ks_2samp from the scipy.stats package (**Supplementary Figure 20a**).

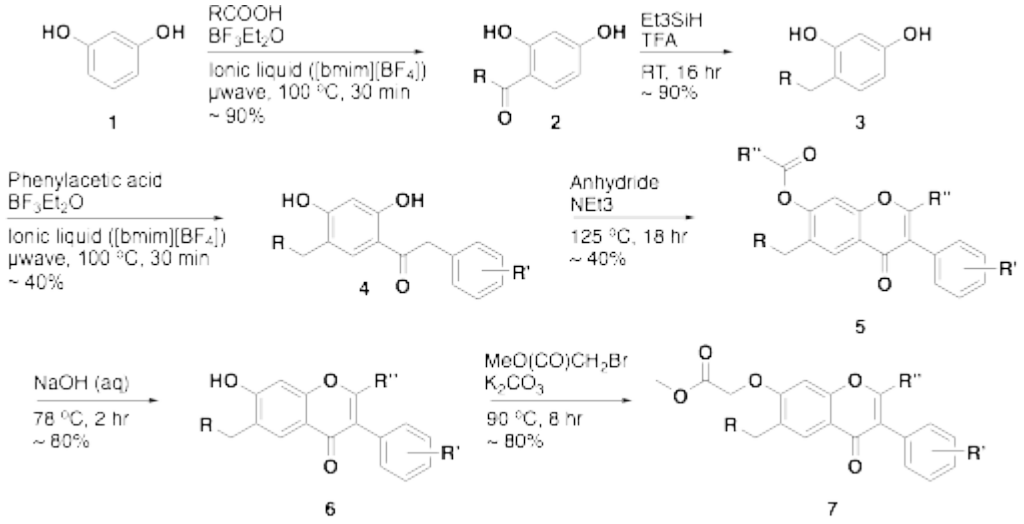
To calculate approximate thresholds of phenoscore significance, we performed a statistical simulation. For each score in the space of possible phenoscores (binned in 0.05 increments from 0 to 1), we sampled 12 replicates from a uniform distribution centered around the score ranging from -4σ to +4σ away from the mean, and calculated the KS statistic against the etomidate 6.25 μM replicates. We repeated this simulated procedure 100 times to get robust statistics, and took the average of these *P* values. However, we realized that the standard deviation of replicates across different $GABA_AR$ ligands was not a constant value. It tended to be low for extremely poor phenotypes, peaked for intermediate phenotypes, and decreased again for extremely strong phenotypes.

Therefore, we fit the standard deviations for $GABA_AR$ ligands as a function of phenoscore with a 10th order polynomial using the Polynomial package in numpy (**Supplementary Figure 20b**). Using this resulting polynomial, we calculated the KS *P* values from the simulated uniform distributions as we iteratively stepped along the y-axis ; these *P* values were smoothly distributed except for a discontinuity around phenoscore 0.5 due to rapidly increasing *P* values in this range (**Supplementary Figure 20c**). We derived the threshold phenoscores associated with these *P* values by fitting another polynomial to the resulting distribution in the smooth region (above phenoscore 0.5) (**Supplementary Figure 20d**) and calculating the roots of the function at those *P* values. The

resulting phenoscores corresponding to 0.01 and 0.05 *P* value thresholds were 0.51 and 0.71, respectively.

Z' and false positives and negatives calculation. To calculate the Z' (quality of screen coefficient), we use the formula $Z' = 1 - 3(\sigma p + \sigma n)/(\mu p - \mu n)$, expressed in terms of the mean and SD of the positive (p) and negative (n) controls. The false positive/negative rates were determined at a threshold of 3 SD. Any positive controls with a phenoscore 3 SD away from the positive control mean (μp) were labeled false positives. Likewise, any negatives controls 3 SD away from the negative control mean (μn) were labeled false negatives.

General Synthesis Scheme of Isoflavones



General Procedure for the Synthesis of 2. To a mixture of resorcinol (1 equiv) and carboxylic acid (1 equiv), BF3·Et2O (3 equiv) and ionic liquid ([bmim][BF4]) (3.5 equiv) was added. The reaction mixture was irradiated at 100 °C for 30 min in a microwave reactor at the maximum power of 300W. The solution was allowed to cool and poured into water. The reaction mixture was extracted with ethyl acetate, and the organic layer is separated and washed with brine, dried, and concentrated. The concentrate was purified by silica gel column chromatography to get ketone 2.

General Procedure for the Synthesis of 3. 2 was dissolved in trifluoroacetic acid (20 equiv) and triethylsilane (2.5 equiv) was added at room temperature. The resulting solution was stirred overnight, and the solvent was removed by flushing nitrogen gas in mild temperature. The residue was purified by silica gel column chromatography to get diol 3.

General Procedure for the Synthesis of 4. To a mixture of 3 and a phenylacetic acid (1 equiv), BF3·Et2O (3.5 equiv) and ionic liquid ([bmim][BF4]) (3 equiv) was added. The reaction mixture was irradiated at 100 °C for 30 min in a microwave reactor at the maximum power of 300W. The solution was allowed to cool and poured into water. The reaction mixture was extracted with ethyl acetate, and the organic layer was separated and washed with brine, dried and concentrated. The concentrate was purified by silica gel column chromatography to get ketone 4.

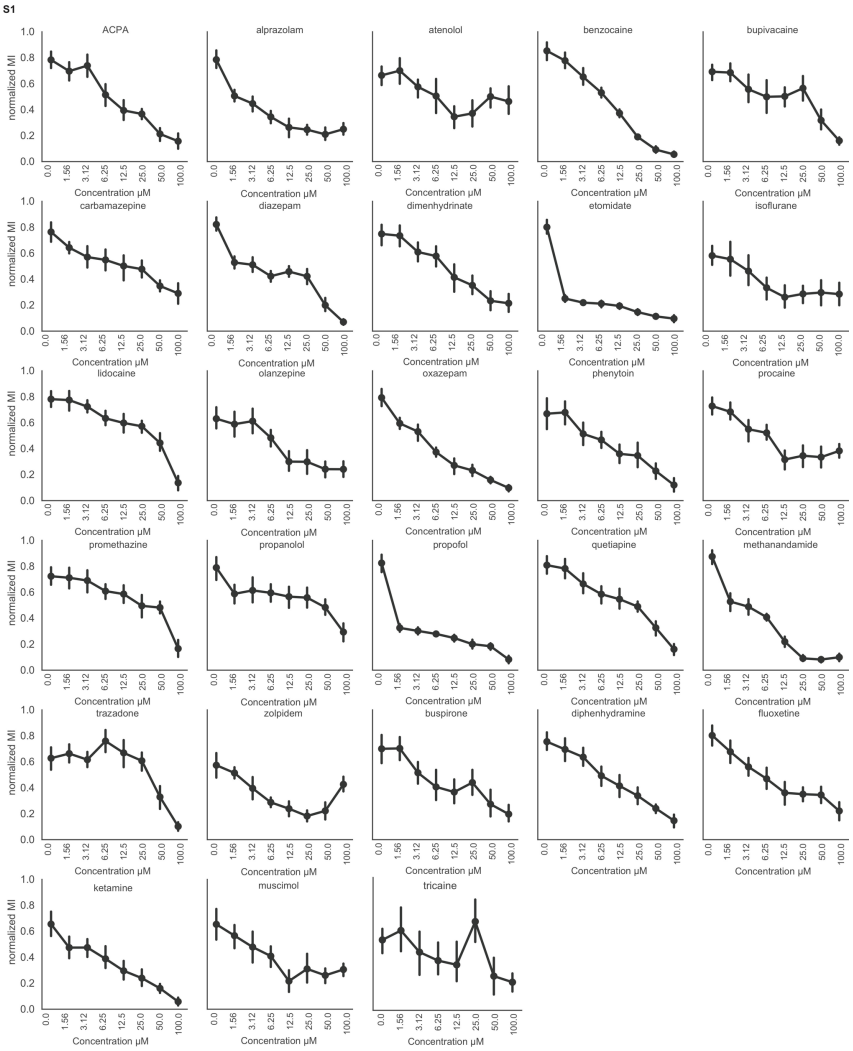
General Procedure for the Synthesis of 5. A mixture of 4, propionic anhydride (5 equiv) and triethylamine (4 equiv) was heated at 125 °C for 12 h. Then the reaction mixture was added to cold dilute 1M HCl solution. The mixture was extracted with ethyl acetate, and the organic layer was separated and washed with brine, dried and concentrated. The concentrate was purified by silica gel column chromatography to get isoflavone 5.

General Procedure for the Synthesis of 6. A solution of 5 in ethanol (0.2 M) containing 10% w/w NaOH was refluxed for 30 min. After 30 min, the same amount of water was added, and heating was continued for another 1.5 h. The reaction mixture was acidified with dilute hydrochloric acid and extracted with ethyl acetate, and the organic layer was separated and washed with brine, dried and concentrated. The concentrate was purified by silica gel column chromatography to get isoflavone 6.

General Procedure for the Synthesis of 7. To a solution of 6 in DMF (0.5 M), methyl bromoacetate (1.1 equiv) and K_2CO_3 (3 equiv) was added. The mixture was heated to 90 °C for 8 h. The reaction mixture was cooled to room temperature and extracted with ethyl acetate.

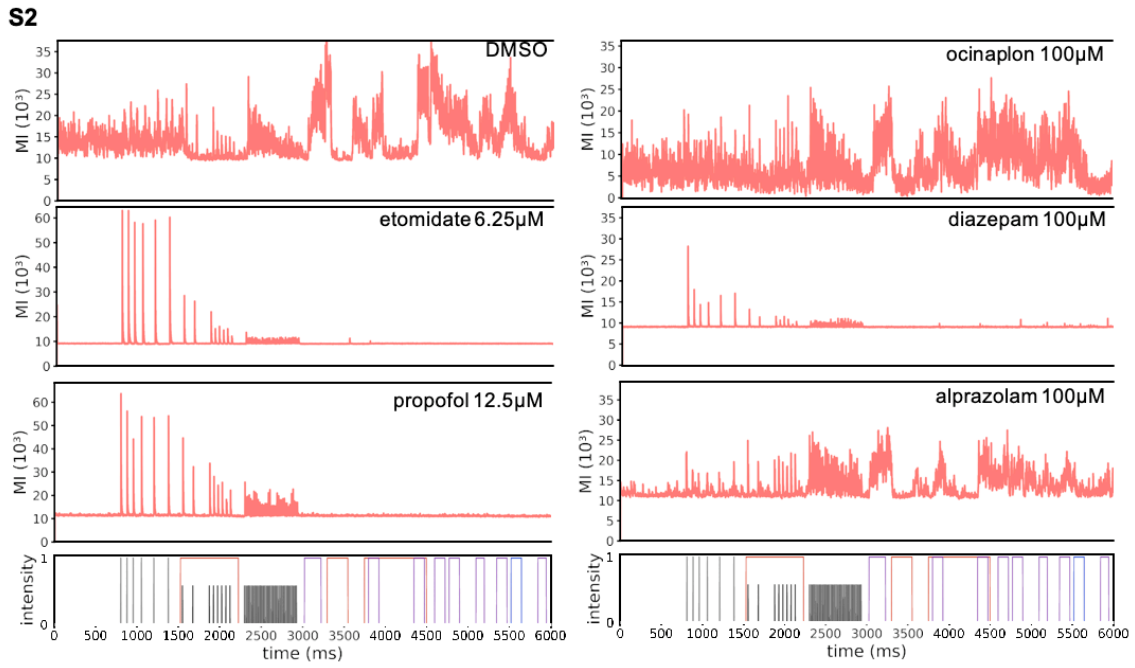
$GABA_AR$ terminology. Notation for GABA receptors conform to IUPHAR recommendations [5]. Receptor subunits are indicated by their greek symbols with subscripted numbers to indicate specific isoforms as in: "the α_1 subunit isoform". To refer to $GABA_A$ receptor ($GABA_AR$) subtypes, the term GABA is used to indicate the receptor type, and the subscript A is used to refer to all $GABA_ARs$. Subtypes comprised of specific subunit isoforms are indicated like: "the $\alpha_1\beta_2\gamma_2$ subtype".

Supplementary Figures



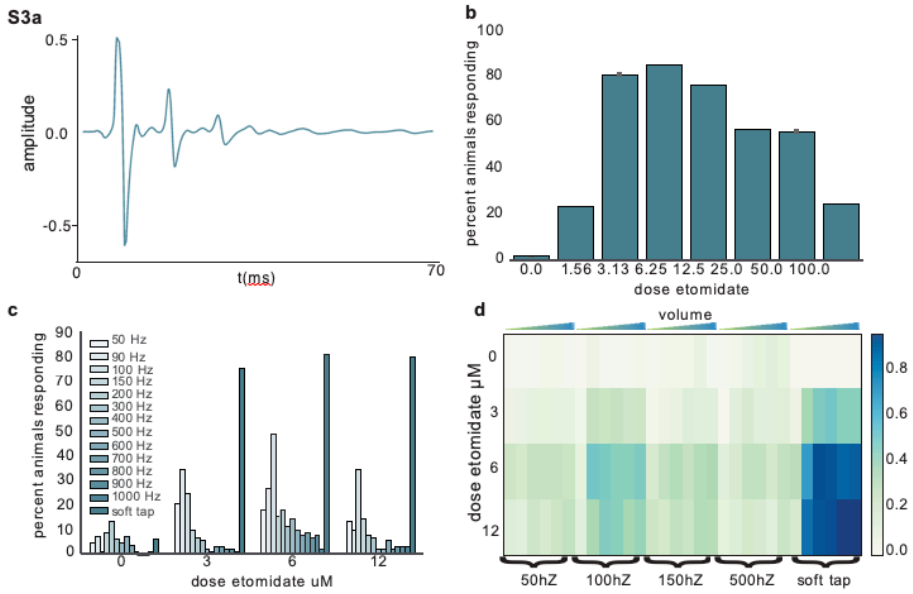
Supplementary Figure A.1.1.

Sedatives cause a dose dependent reduction in zebrafish motion. A panel of 30 known sedatives administered to 7dpf zebrafish larvae (n = 12 replicates; 96 fish/ condition) at a 2 fold dilution series. The y-axis represents motion index (MI) and the x-axis represents dose.



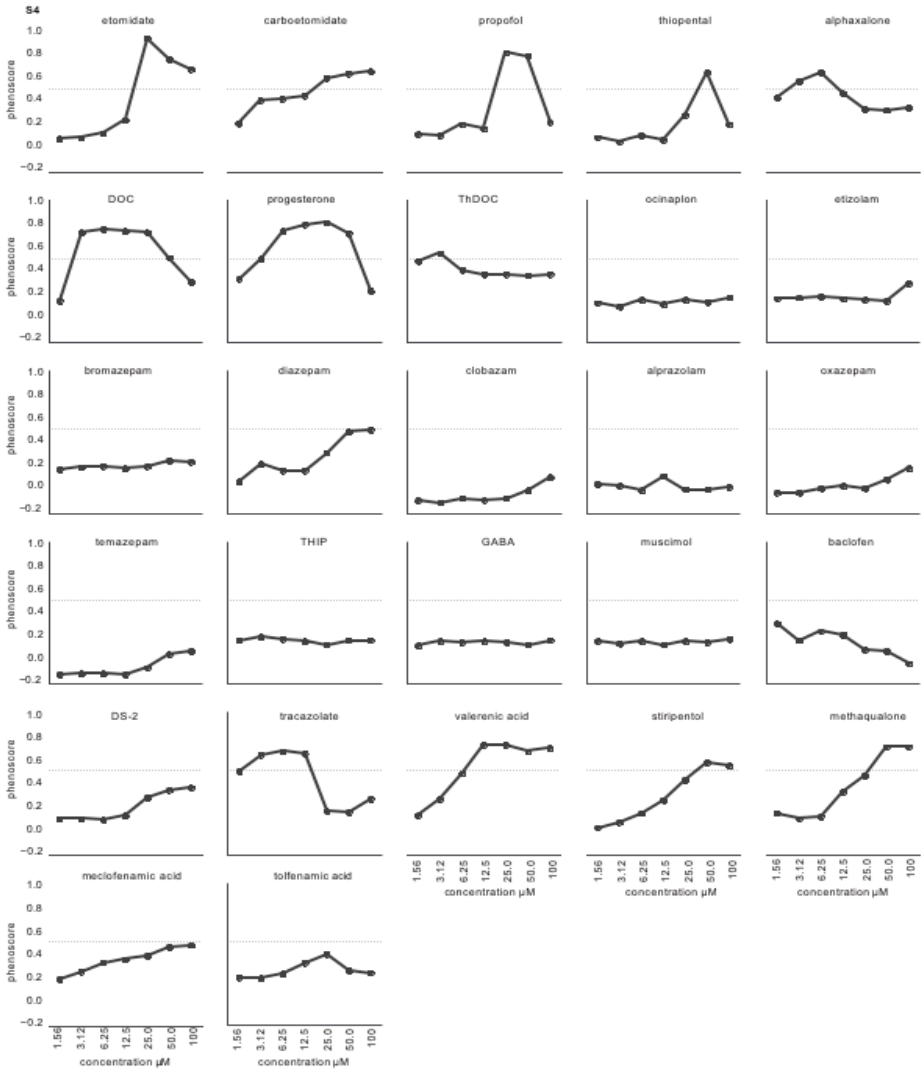
Supplementary Figure A.1.2.

Propofol and etomidate block light-induced behaviors, but enhance the acoustic startle response. The plots show motor activity (y-axis) of zebrafish treated with the indicated compounds (n = 12 wells). Colored bars above the x-axis represent the timing and duration of indicated stimuli.



Supplementary Figure A.1.3.

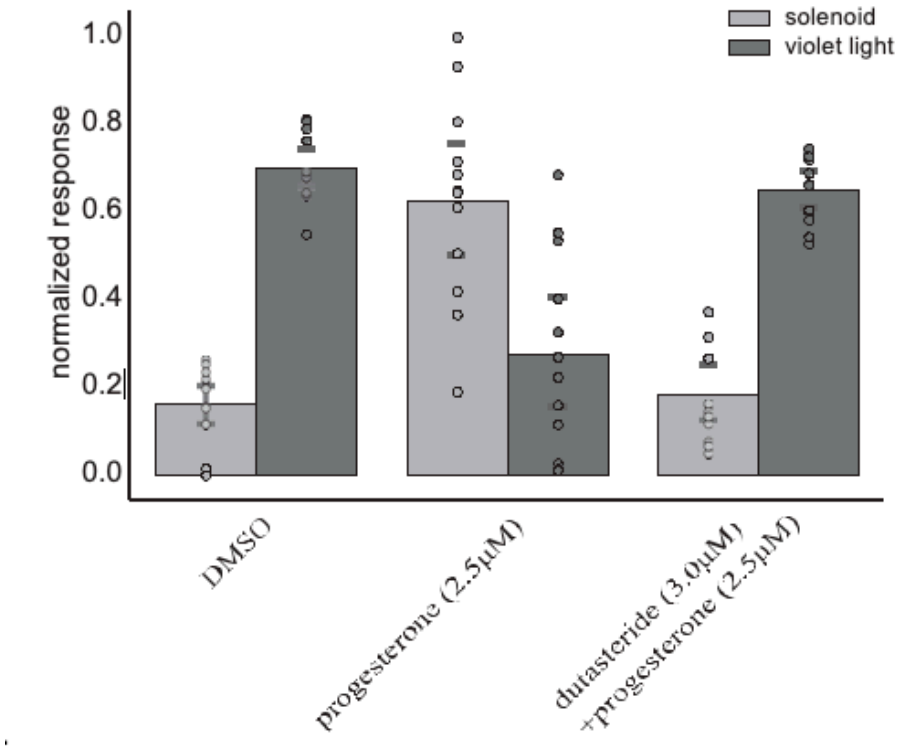
eASR stimulus characterization. We explored a range of digital, acoustic stimuli to understand which parameters were important for triggering eASRs. (a) Recorded waveform of the dampened solenoid. The original stimulus, generated by a dampened solenoid, approximated a 100 Hz inverse fading sine wave, with a 70 dB maximum volume and 70 ms duration. (b) Bar graph depicting startle frequency (y-axis) of 100 animals to the dampened solenoid stimulus at increasing concentrations of etomidate (x-axis), it elicited responses in 85% of etomidate-treated animals (6.25 µM) and in 2% of controls. (c) Startle frequency (y-axis) of 100 animals treated with indicated concentrations of etomidate (x-axis) in response to different frequencies of synthesized and dampened solenoid acoustic stimulus (colored bars). In frequency scans from 50-1000 Hz, the highest magnitude eASRs were elicited by 100 Hz stimuli. Interestingly, the most effective synthesized stimulus (a 100 Hz inverse fading sine wave; 70ms) was only 50% as effective as the original solenoid, suggesting that some unknown feature of the original solenoid-based stimulus was not captured by the synthesized waveform and/or the surface transducers. (d) Heat map of the startle frequency of 100 animals (color bar) in response to increasing volume (top y-axis) of different frequency synthesized acoustic stimulus and the solenoid stimulus (x-axis). Animals were treated with increasing concentrations of etomidate (y-axis). At 100Hz, all stimuli greater than 60 dB were effective, whereas those less than 55 dB were not.



Supplementary Figure A.1.4.

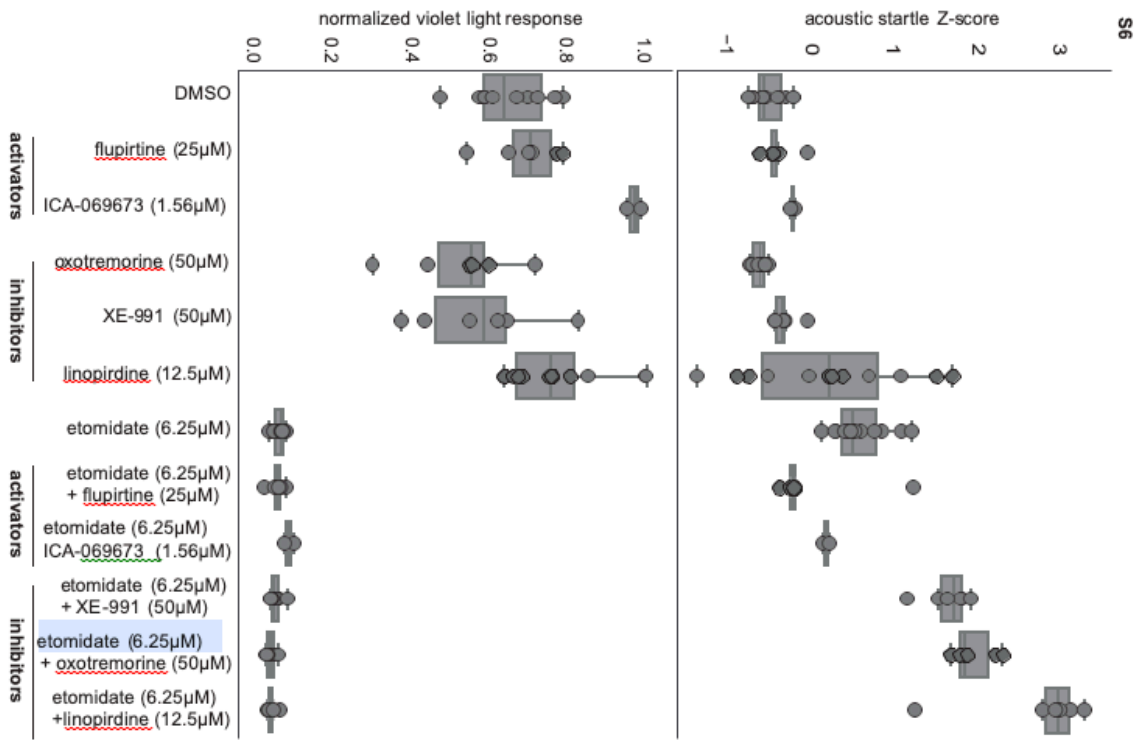
Dose response analysis of GABA reference compounds. Average phenoscores (y-axis) of zebrafish treated with the indicated compounds (n = 12 wells) at increasing concentrations (x-axis).

S5



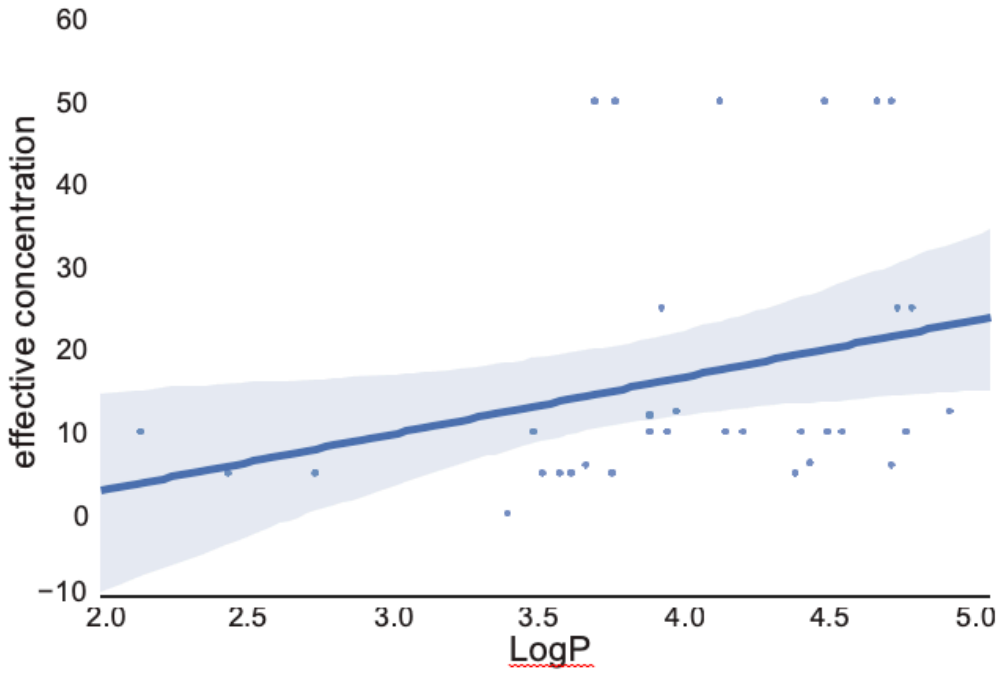
Supplementary Figure A.1.5. Dutasteride inhibits progesterone-induced eASRs.

The plots show the normalized behavioral responses (y-axis), to acoustic (grey) or light (black) stimuli, in animals treated with the indicated compounds (x-axis).



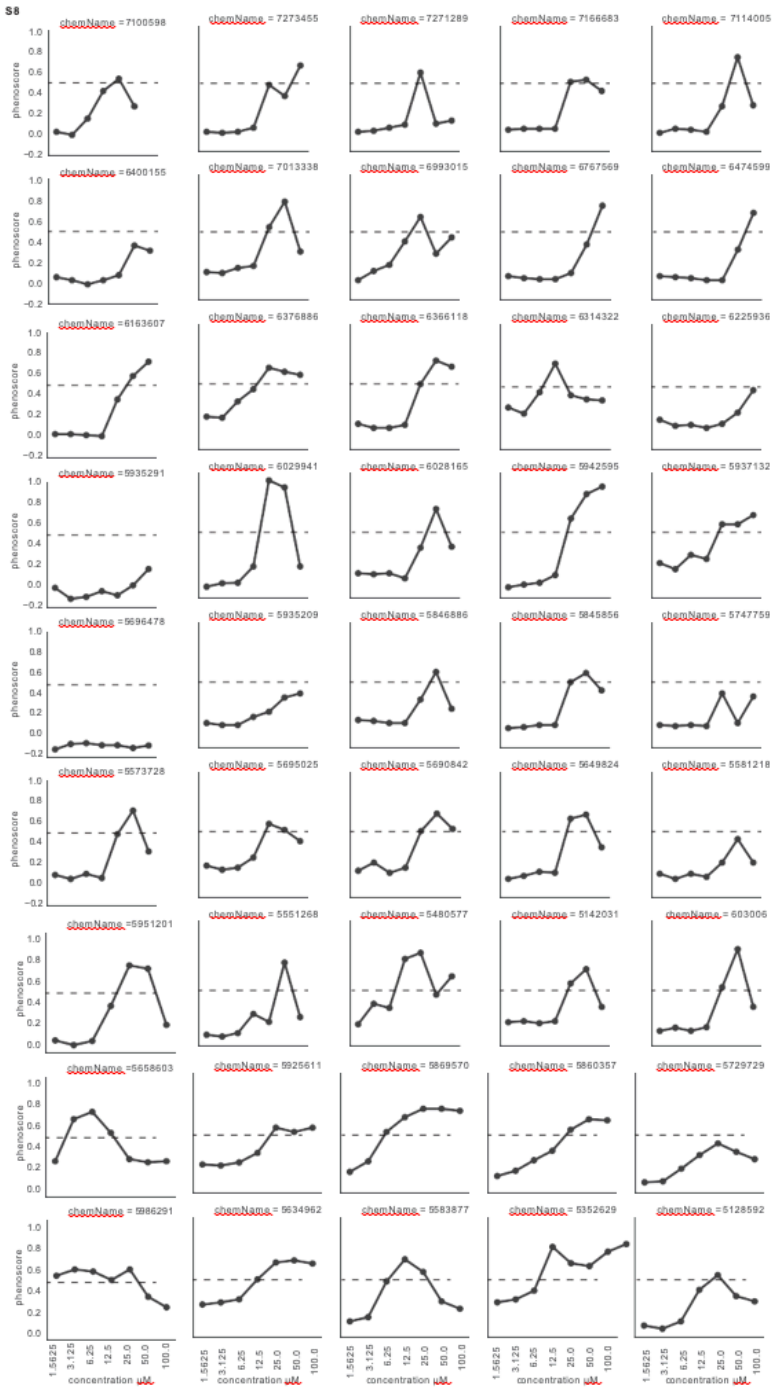
Supplementary Figure A.1.6.

M-current ligands modify eASRs. Boxplots depicting the motor activity (y-axis) of animals treated with the indicated compounds (x-axis) in response response to acoustic (top) or violet light stimuli (bottom). M-current activators and inhibitors were analyzed alone or combined with etomidate, at the indicated concentrations.



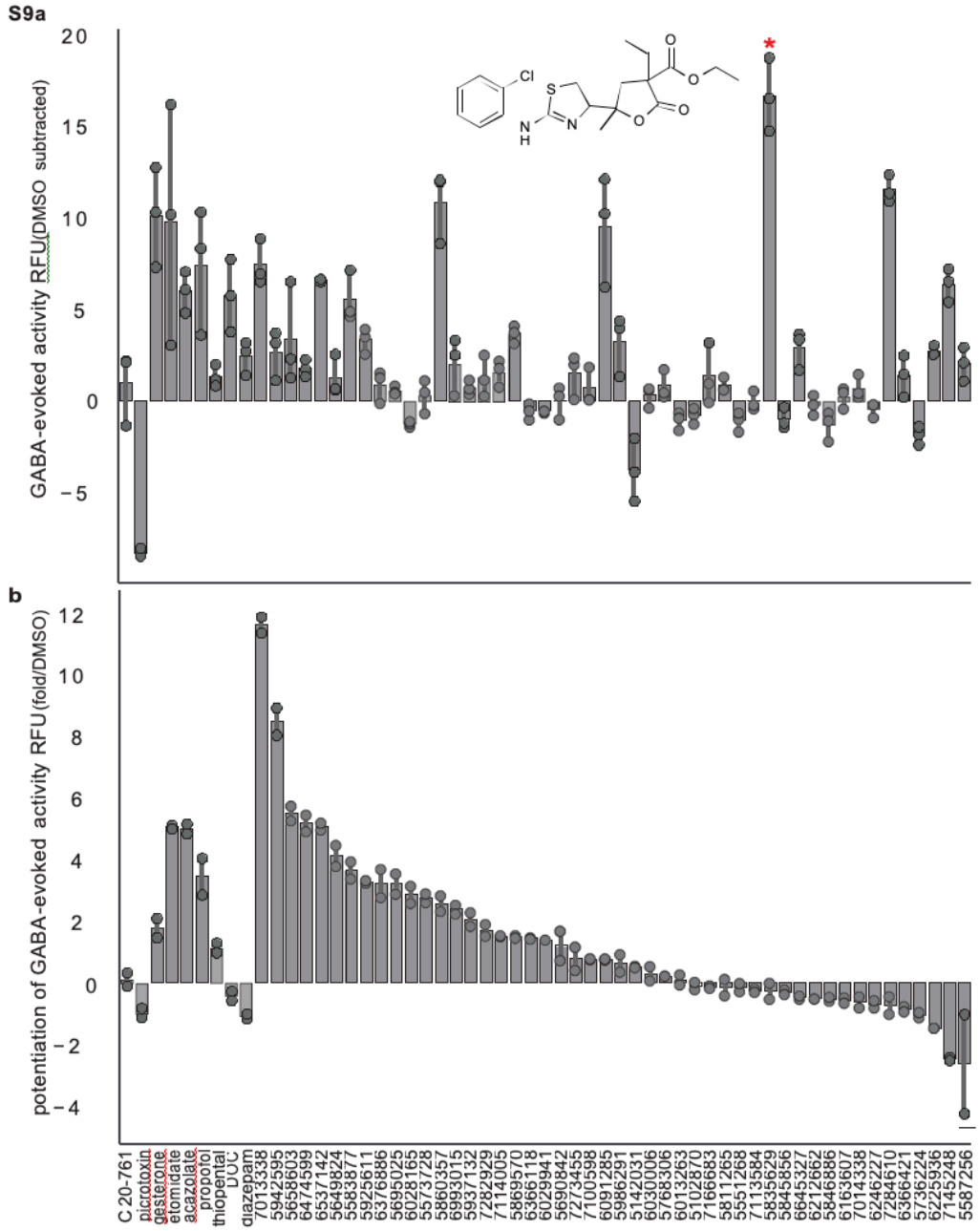
Supplementary Figure A.1.7.

Hit compound efficacy does not correlate with hydrophobicity. Animals were treated with hit compounds. For each compound, the cLogP (calculated partition coefficient) (x-axis) and minimum concentration required to cause the eASR phenotype, were plotted (y-axis). Unlike historical Myer-Overton analyses, the minimum effective concentration does not decrease with hydrophobicity. The best-fit line and shading represent the resulting regression line and a 95% confidence interval for that regression.



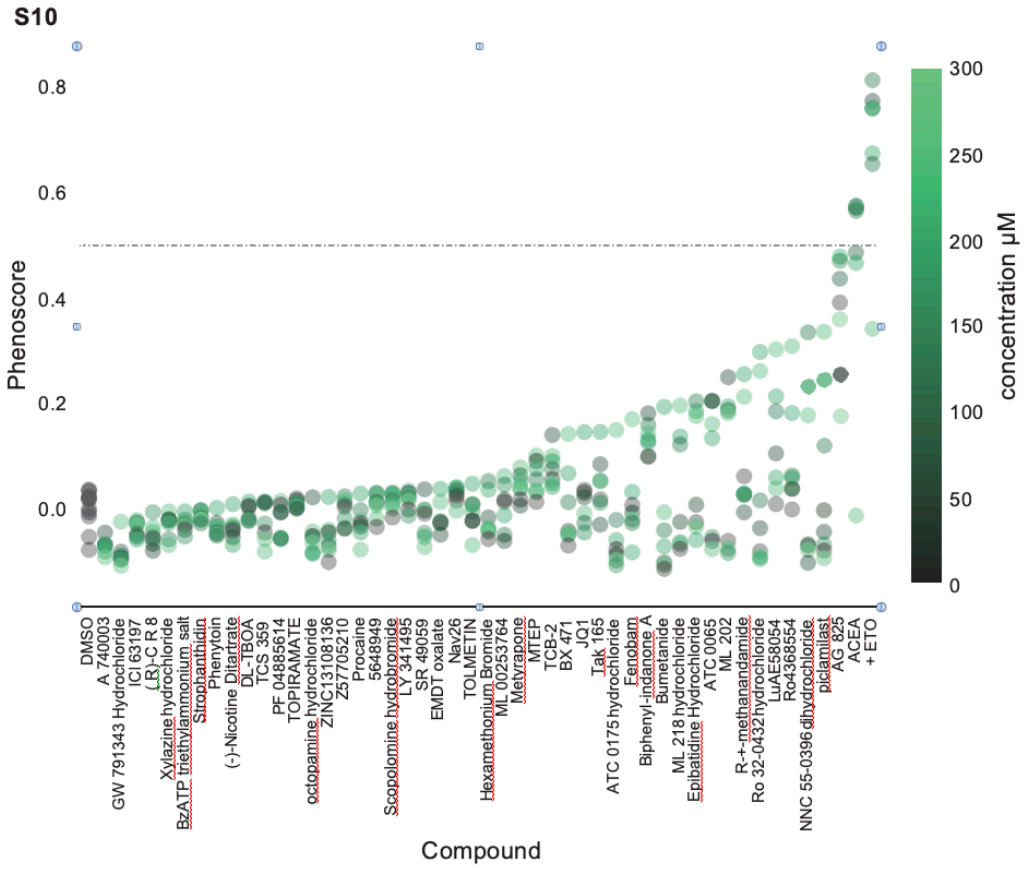
Supplemental Figure A.1.8.

Dose response retest of primary hit compounds. Average phenoscores (y-axis) of zebrafish treated with the indicated compounds (n = 12 wells) at the indicated concentrations (x-axis).



Supplementary Figure A.1.9.

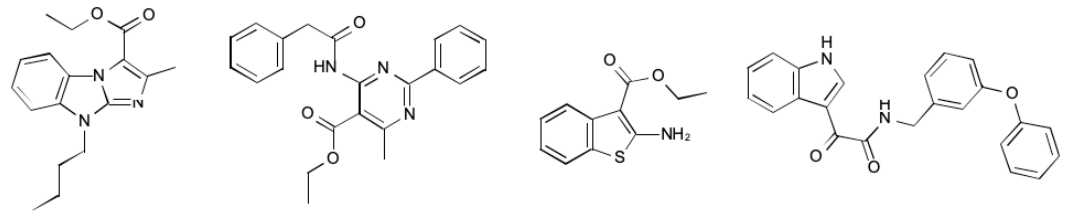
Hit compounds cause direct and indirect activation of GABARs. (a,b) Human GABAAR activation (*y*-axis) was measured by FLIPR analysis in random fluorescent units (RFUs). Direct **(a)** and indirect **(b)** activation was analyzed for the indicated hit compounds (x-axis, n= 2-4).



Supplementary Figure A.1.10.

Phenoscores of ligands at targets with low value EFs. Ligands for targets with low (left of dotted line) and high (right of dotted line) EF scores. The plot shows the phenoscore (y-axis) of the indicated compounds (x-axis). Color bar represents concentration in µM.

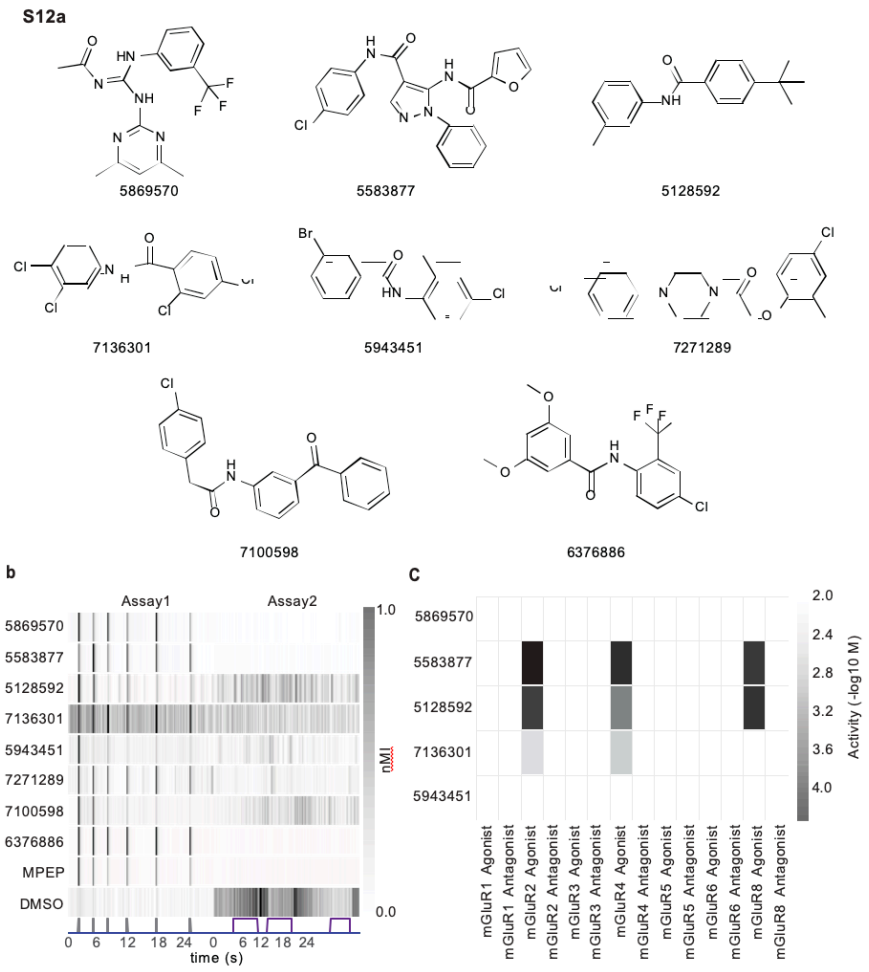
S11



5658603 5951201 5142031 7145248

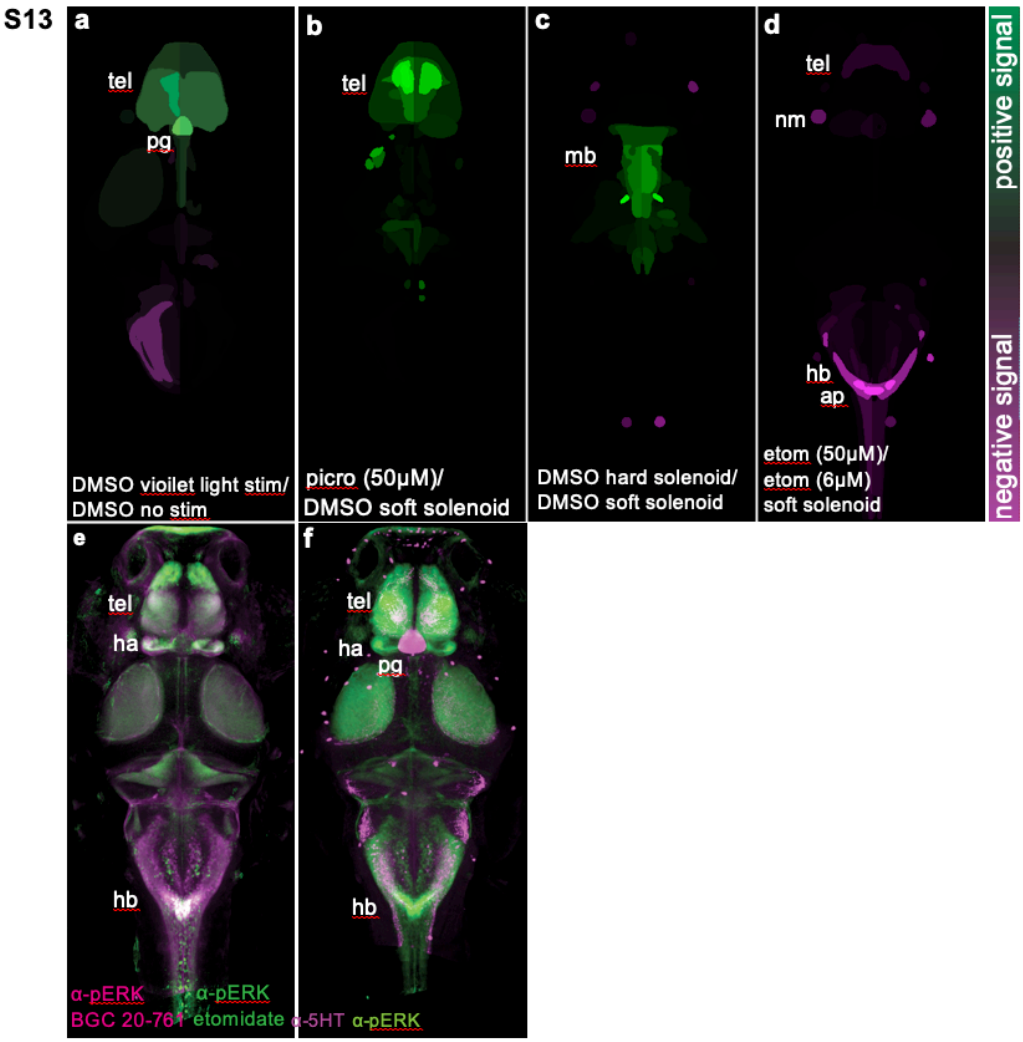
Supplementary Figure A.1.11.

Chemical structures of hit compounds predicted to target GABAARs by SEA.



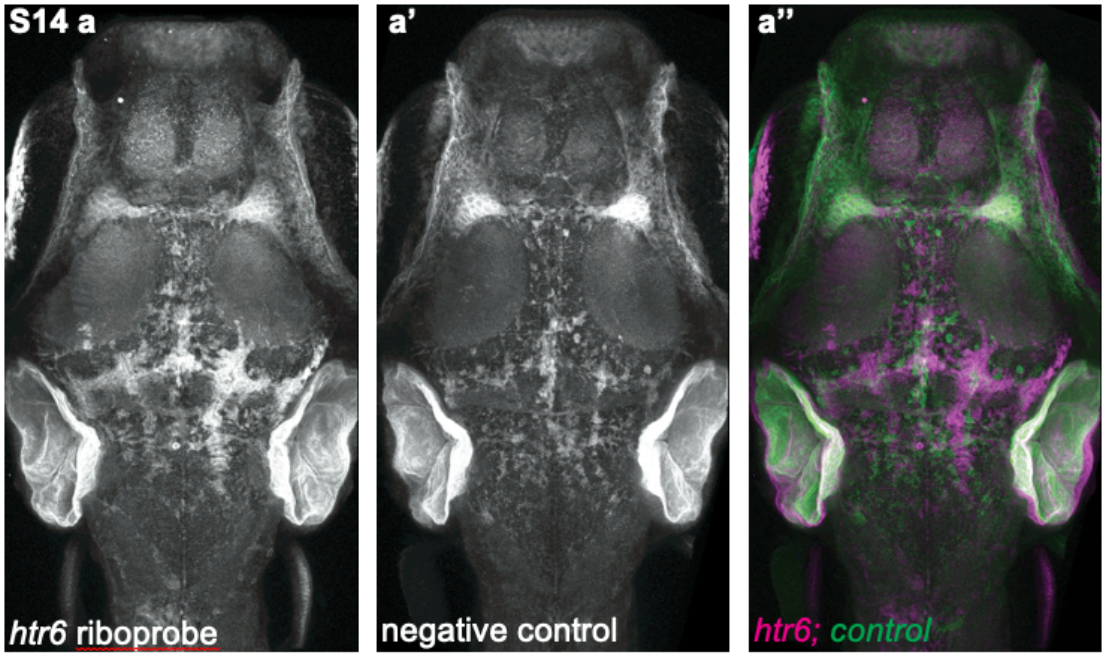
Supplementary Figure A.1.12. Characterization of hit compounds predicted to target mGluR by SEA.

(a) Chemical structures of 8 hit compounds predicted to target mGluR. (b) The heatmap represents the normalized motion index (nMI) of larvae treated with the indicated compounds. Assay 1 is composed of 6 low amplitude acoustic stimuli; Assay 2 is a series of 3 violet light pulses as indicated on the x-axis. MPEP is a known mGluR4/5 ligand. Compounds were tested for agonist and antagonist activity in Gq functional assays in-vitro. (c) The heat map represents the activity of 5 novel mGluR predicted compounds (y-axis) at the indicated receptor (x-axis). Low-level activation of mGlur2/4 was detected for compounds 5583877, 5128592, and 7136301 (46.13 μM to 2871 μM)



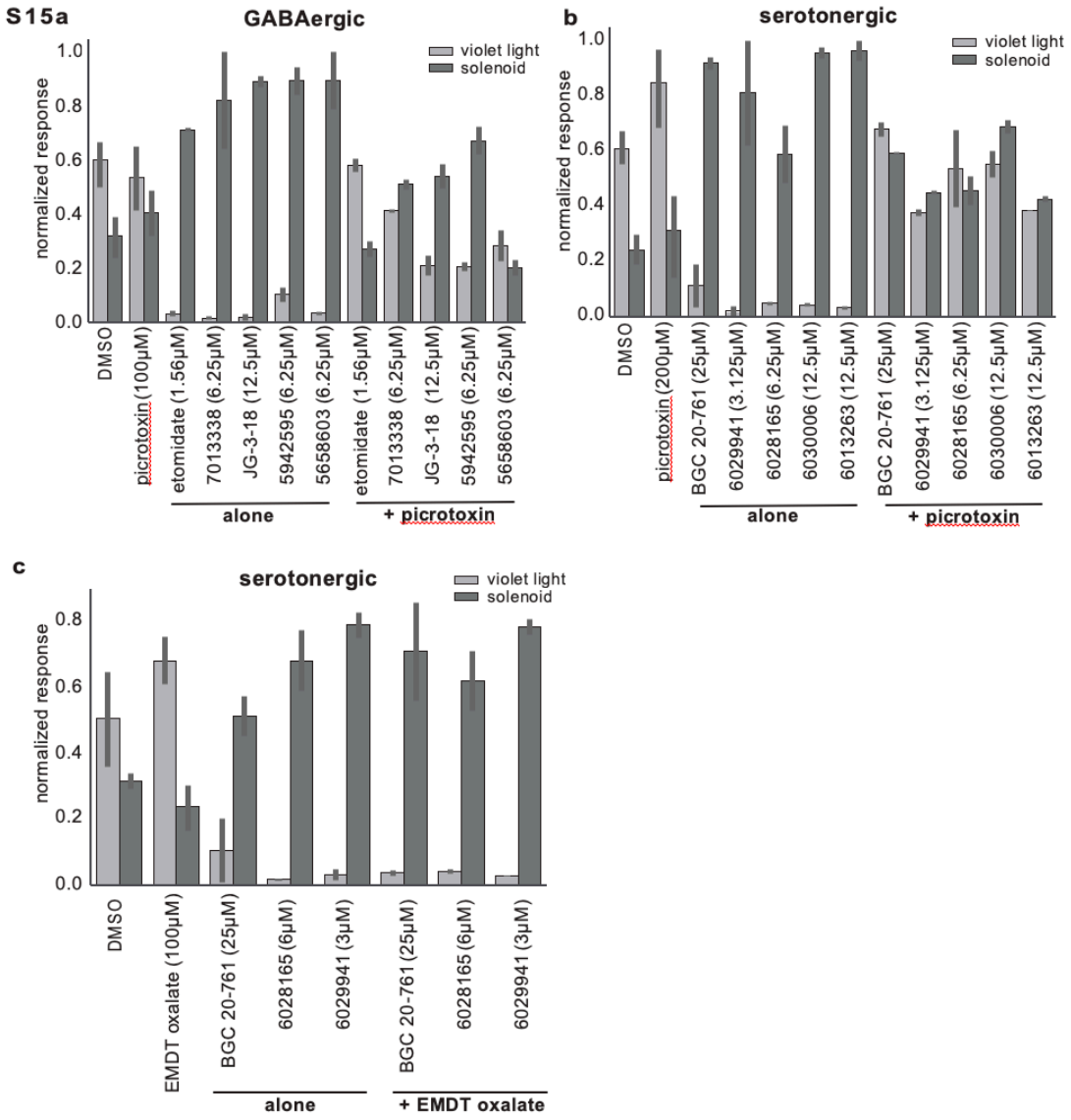
Supplementary Figure A.1.13. pERK whole brain neural activity maps in control assays and 5-HT immunohistochemistry.

(a-d) Brain activity maps showing significant ΔpERK signals using the Z-brain online reference tool (n = 5-10 animals/condition). Heatmaps indicate positive (green), negative (purple), and nonsignificant (black) changes in pERK labeling ($p < 0.0005$, Mann-Whitney U test). All activity maps are comparisons between the indicated treatment conditions. (**e**) Overlay of average α-pERK signal for BGC 20-761(magenta), and etomidate treated animals (green). (**f**) Overlay of α-5HT staining (magenta) and the average α-pERK staining (green) for BGC 20-761 treatment. Abbreviations: tel, telencephalon; mb, midbrain; ot, optic tectum; hb, hindbrain; ha, habenula; ob, olfactory bulb; nm, neuromast; ap, area postrema; pg, pineal gland.



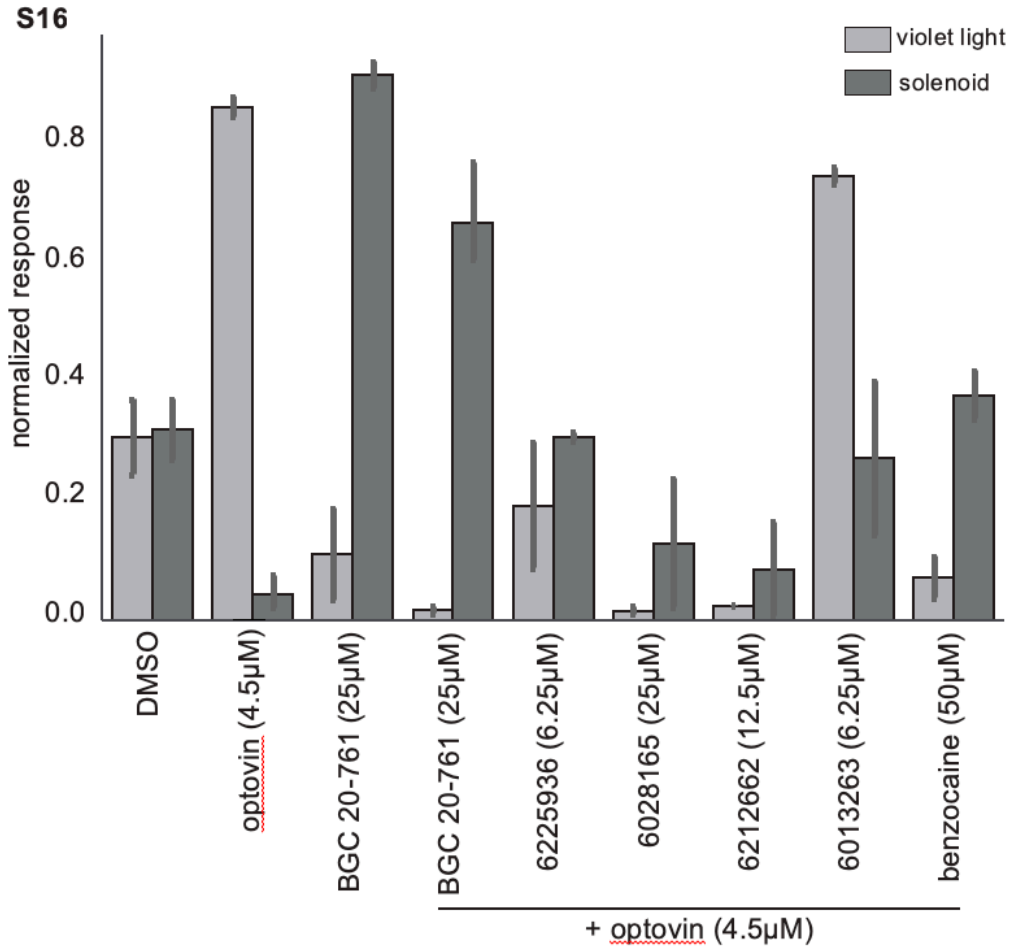
Supplementary Figure A.1.14. Fluorescent in situ hybridization of the zebrafish htr6 transcript shows low expression in the telencephalon.

(a-a") Confocal projections from image registered animals showing transcripts for htr6 (a) and non specific negative control antisense probe (a'), overlay in (a").



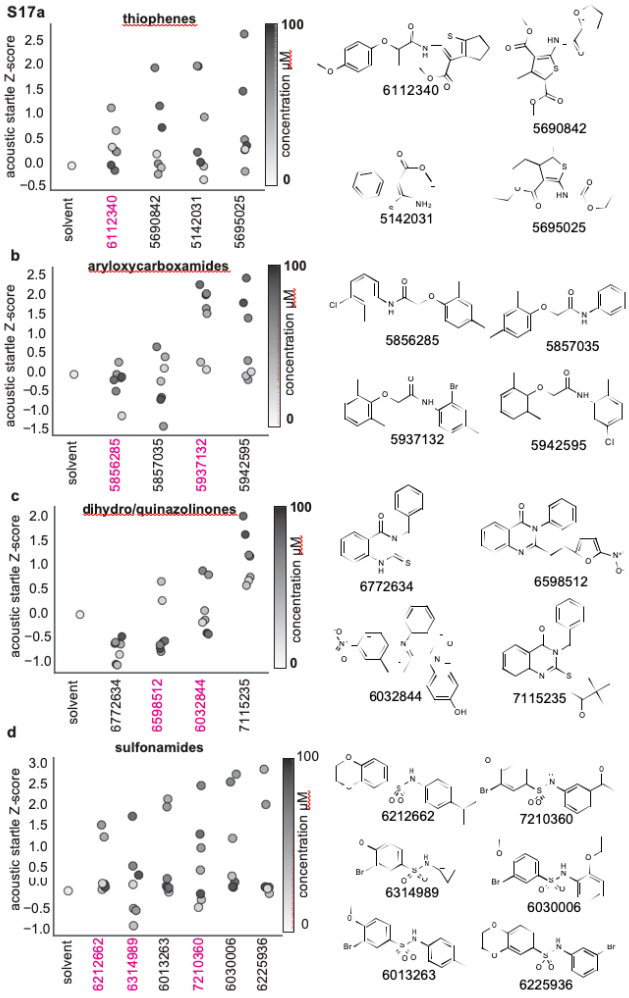
Supplementary Figure A.1.15.

The GABAergic antagonist picrotoxin reverses the eASRs-induced by some ligands, but the serotonergic agonist EMDT oxolate does not. (a-c) Normalized behavioral responses (y-axis) of animals treated with the indicated compounds (x-axis).



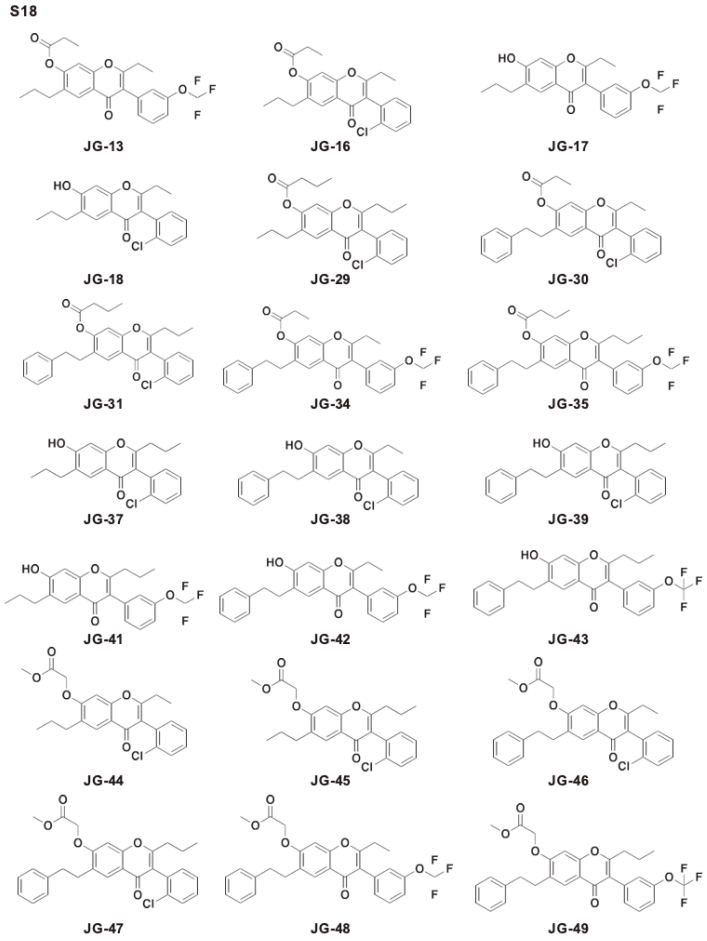
Supplementary Figure A.1.16.

Serotonergic hit compounds inhibit optovin response. Normalized behavioral response (y-axis) of animals treated with the indicated compounds (x-axis).



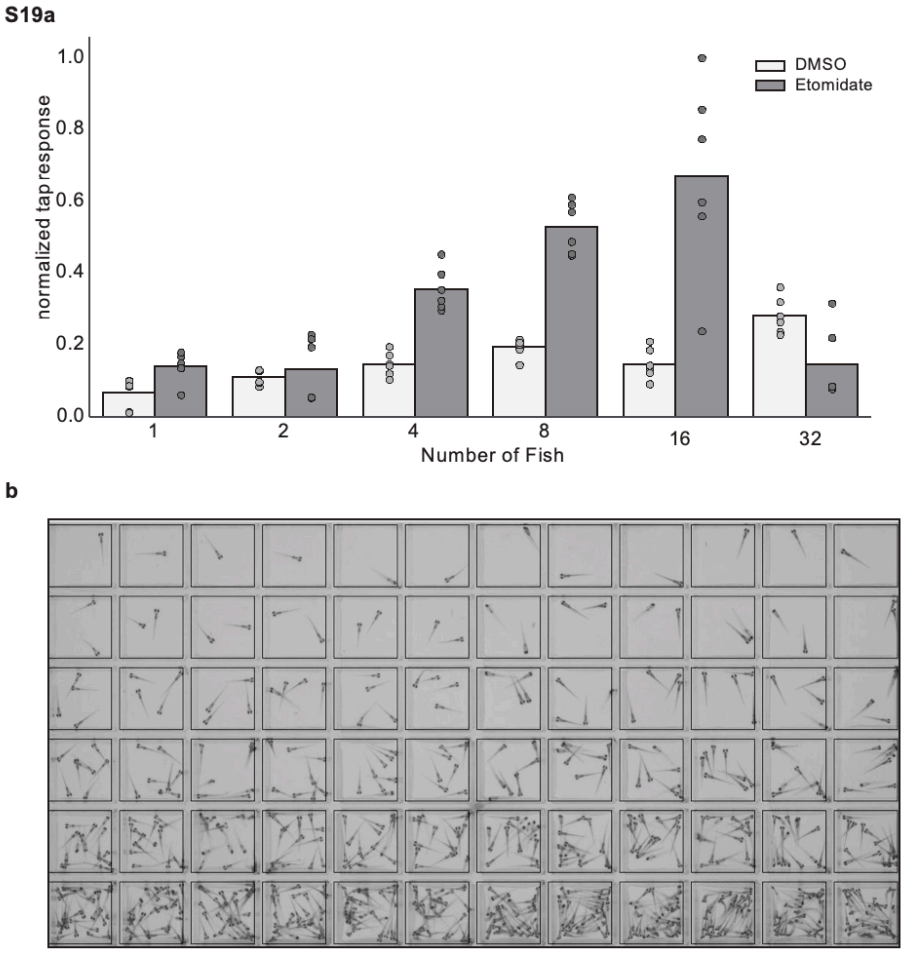
Supplementary Figure A.1.17. Preliminary SAR of key compound classes. (a-d)

The plots show the Z-score of the acoustic startle response (y-axis) in animals treated with the indicated compounds (x-axis). The compound structures in each class are shown to the right of each plot including the original hit compounds (black) and their analogs (red). Many analogs did not cause the eASR behaviors.



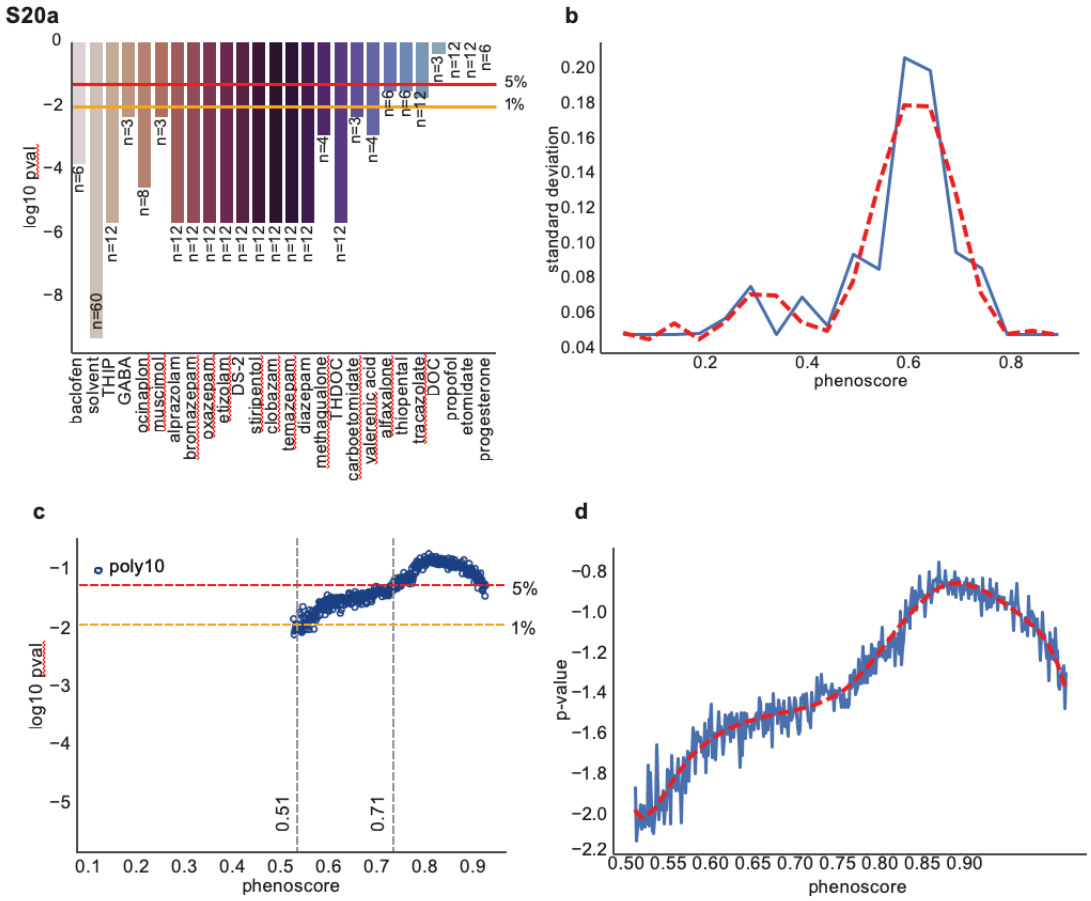
Supplementary Figure A.1.18.

Structures of Isoflavone analogs



Supplementary Figure A.1.19. Group size affects eASR quantification.

To determine the impact of group size on this assay, we analyzed eASR behaviors from animals in different group sizes (1, 2, 4, 8, 12, 16, 32 animals per well). Groups of 8 and 16 animals generated the most robust MI values. Here, we chose to use 8 animals per group because it balanced a small group size with high signal to noise using the MI metric. (a) Bar graph illustrating the average tap response (y-axis) per tap stimulus (each marker represents one of 6 total stimuli averaged over 6 replicate wells), for wells with the indicated number of fish larvae (x-axis) and treated with DMSO control or 6μM etomidate, as indicated. (b) Representative image of wells containing increasing numbers of animals.



Supplementary Figure A.1.20. Statistical analysis of phenotypic thresholds for $GABA_AR$ ligands.

(a) This plot shows the Kolmogorov-Smirnov (KS Test) statistic for the highest-scoring profiles produced by the indicated treatments (y-axis). On the x-axis, the ligands are sorted in order of ascending average phenoscore (left to right) from lowest to highest. Horizontal lines on the y-axis indicate the 1%(yellow) and 5% (red) P value significance thresholds (b) Plot showing the standard deviations for GABAAR ligands as a function of phenoscore with a 10th order polynomial. (c) Plot showing simulated P values as a function of phenoscore. Horizontal dashed lines indicate 1% and 5% P value thresholds, and vertical lines indicate the phenoscores at which these thresholds are met (0.51 and 0.71, respectively). (d) Plot showing a 10th order polynomial fit for the smooth region of the simulation where phenoscore > 0.5 in panel (c).

References

1. Lounkine, E. et al. Large-scale prediction and testing of drug activity on side-effect targets. Nature 486, 361–367 (2012).
2. Bruni, G. et al. Zebrafish behavioral profiling identifies multitarget antipsychotic-like compounds. Nat. Chem. Biol. 12, 559–566 (2016).
3. Gaspar, P. & Lillesaar, C. Probing the diversity of serotonin neurons. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367, 2382–2394 (2012).
4. Gross-Thebing, T., Paksa, A. & Raz, E. Simultaneous high-resolution detection of multiple transcripts combined with localization of proteins in whole-mount embryos. BMC Biol. 12, 55 (2014).
5. Barnard, E. A. et al. International Union of Pharmacology. XV. Subtypes of gamma-aminobutyric acidA receptors: classification on the basis of subunit structure and receptor function. Pharmacol. Rev. 50, 291–313 (1998).
6. Reis, G. M. L. & Duarte, I. D. G. Baclofen, an agonist at peripheral GABAB receptors, induces antinociception via activation of TEA-sensitive potassium channels. Br. J. Pharmacol. 149, 733– 739 (2006).
7. Masiulis, S. et al. GABAA receptor signalling mechanisms revealed by structural pharmacology.Nature 565, 454–459 (2019).
8. Johnston, G. A. R. Muscimol as an ionotropic GABA receptor agonist. Neurochem. Res. 39, 1942–1947 (2014).
9. Jensen, M. L. et al. A study of subunit selectivity, mechanism and site of action of the delta selective compound 2 (DS2) at human recombinant and rodent native GABA(A) receptors. Br. J. Pharmacol. 168, 1118–1132 (2013).
10. Braestrup, C. & Squires, R. F. Pharmacological characterization of benzodiazepine receptors in the brain. Eur. J. Pharmacol. 48, 263–270 (1978).
11. Pernea, M. & Sutcliffe, A. G. Clobazam and Its Use in Epilepsy. Pediatr. Rep. 8, 6516 (2016).
12. Heel, R. C., Brogden, R. N., Speight, T. M. & Avery, G. S. Temazepam: a review of its pharmacological properties and therapeutic efficacy as an hypnotic. Drugs 21, 321–340 (1981).
13. Lippa, A. et al. Selective anxiolysis produced by ocinaplon, a GABA(A) receptor modulator. Proc. Natl. Acad. Sci. U. S. A. 102, 7380–7385 (2005).
14. Sanna, E. et al. Molecular and neurochemical evaluation of the effects of etizolam on GABAA receptors under normal and stress conditions. Arzneimittelforschung 49, 88–95 (1999).
15. Thompson, S.-A., Wingrove, P. B., Connelly, L., Whiting, P. J. & Wafford, K. A. Tracazolate reveals a novel type of allosteric interaction with recombinant gamma-aminobutyric acid(A) receptors. Mol. Pharmacol. 61, 861–869 (2002).

16. Fisher, J. L. The effects of stiripentol on GABA(A) receptors. Epilepsia 52 Suppl 2, 76–78 (2011).
17. Hammer, H. et al. A Multifaceted GABAA Receptor Modulator: Functional Properties and Mechanism of Action of the Sedative-Hypnotic and Recreational Drug Methaqualone (Quaalude). Mol. Pharmacol. 88, 401–420 (2015).
18. Luger, D. et al. Identification of the putative binding pocket of valerenic acid on GABAA receptors using docking studies and site-directed mutagenesis. Br. J. Pharmacol. 172, 5403–5413 (2015).
19. Hosie, A. M., Wilkins, M. E. & Smart, T. G. Neurosteroid binding sites on GABA(A) receptors. Pharmacol. Ther. 116, 7–19 (2007).
20. Garcia, P. S., Kolesky, S. E. & Jenkins, A. General anesthetic actions on GABA(A) receptors. Curr. Neuropharmacol. 8, 2–9 (2010).
21. Troupin, A. S., Green, J. R. & Halpern, L. M. Carbamazepine (Tegretol) as an anticonvulsant. A controlled double-blind comparison with diphenylhydantoin (Dilantin). Acta Neurol. Scand. Suppl. 60, 13–26 (1975).
22. Yaari, Y., Selzer, M. E. & Pincus, J. H. Phenytoin: mechanisms of its anticonvulsant action. Ann.Neurol. 20, 171–184 (1986).
23. Wong, D. T., Bymaster, F. P. & Engleman, E. A. Prozac (fluoxetine, Lilly 110140), the first selective serotonin uptake inhibitor and an antidepressant drug: twenty years since its first publication. Life Sci. 57, 411–441 (1995).
24. Stahl, S. M. Mechanism of action of trazodone: a multifunctional drug. CNS Spectr. 14, 536–546 (2009).
25. Simons, F. E. R. & Simons, K. J. Histamine and H1-antihistamines: celebrating a century of progress. J. Allergy Clin. Immunol. 128, 1139–1150.e4 (2011).
26. Jann, M. W. Buspirone: an update on a unique anxiolytic agent. Pharmacotherapy 8, 100–116 (1988).
27. Maletzky, B. M. Anxiolytic efficacy of alprazolam compared to diazepam and placebo. J. Int. Med.Res. 8, 139–143 (1980).
28. Calcaterra, N. E. & Barrow, J. C. Classics in chemical neuroscience: diazepam (valium). ACS Chem. Neurosci. 5, 253–260 (2014).
29. Janecek, J., Vestre, N. D., Schiele, B. C. & Zimmermann, R. Oxazepam in the treatment of anxiety states: a controlled study. J. Psychiatr. Res. 4, 199–206 (1966).
30. Misra, L. K., Erpenbach, J. E., Hamlyn, H. & Fuller, W. C. Quetiapine: a new atypical antipsychotic. S. D. J. Med. 51, 189–193 (1998).
31. Bymaster, F. P. et al. Radioreceptor binding profile of the atypical antipsychotic olanzapine. Neuropsychopharmacology 14, 87–96 (1996).
32. Waal-Manning, H. J. Atenolol and three nonselective beta-blockers in hypertension. Clin.Pharmacol. Ther. 25, 8–18 (1979).
33. Stapleton, M. P. Sir James Black and propranolol. The role of the basic sciences in the history of cardiovascular pharmacology. Tex. Heart Inst. J. 24, 336–342 (1997).

34. Hillard, C. J. et al. Synthesis and characterization of potent and selective agonists of the neuronal cannabinoid receptor (CB1). J. Pharmacol. Exp. Ther. 289, 1427–1433 (1999).
35. Abadji, V. et al. (R)-methanandamide: a chiral novel anandamide possessing higher potency and metabolic stability. J. Med. Chem. 37, 1889–1893 (1994).
36. Sanger, D. J. & Depoortere, H. The Pharmacology and Mechanism of Action of Zolpidem. CNS Drug Rev. 4, 323–340 (1998).
37. Terrell, R. C. The invention and development of enflurane, isoflurane, sevoflurane, and desflurane. Anesthesiology 108, 531–533 (2008).
38. Chidambaran, V., Costandi, A. & D'Mello, A. Propofol: a review of its role in pediatric anesthesia and sedation. CNS Drugs 29, 543–563 (2015).
39. Forman, S. A. Clinical and molecular pharmacology of etomidate. Anesthesiology 114, 695–707 (2011).
40. Kurdi, M. S., Theerth, K. A. & Deva, R. S. Ketamine: Current applications in anesthesia, pain, and critical care. Anesth Essays Res 8, 283–290 (2014).
41. Fozzard, H. A., Sheets, M. F. & Hanck, D. A. The sodium channel as a target for local anesthetic drugs. Front. Pharmacol. 2, 68 (2011).
42. Clarkson, C. W. & Hondeghem, L. M. Mechanism for bupivacaine depression of cardiac conduction: fast block of sodium channels during the action potential with slow recovery from block during diastole. Anesthesiology 62, 396–405 (1985).
43. Carter, K. M., Woodley, C. M. & Brown, R. S. A review of tricaine methanesulfonate for anesthesia of fish. Rev. Fish Biol. Fish. 21, 51–59 (2011).

Supplementary Tables

Supplementary Table A.1.1. CNS depressants characterized on zebrafish larval behavior.

class ID	chemical name	class	citation
1	carbamazepine	anticonvulsant	19
1	phenytoin	anticonvulsant	20
2	fluoxetine	antidepressant	21
2	trazodone	antidepressant	22
3	diphenhydramine	antihistamine	23
3	dimenhydrinate	antihistamine	23
3	promethazine	antihistamine	23
4	buspirone	anxiolytic	24
4	alprazolam	anxiolytic	25
4	diazepam	anxiolytic	26
4	oxazepam	anxiolytic	27
5	quetiapine	atypical antipsychotic	28
5	olanzapine	atypical antipsychotic	29
6	atenolol	beta blocker	30
6	propranolol	beta blocker	31
7	ACPA	cannabinoid	32
7	methanandamide	cannabinoid	33
8	zolpidem	hypnotic	34
9	benzocaine	local anesthetic	35
9	lidocaine	local anesthetic	39
9	bupivacaine	local anesthetic	40
9	tricaine	local anesthetic	41
9	procaine	local anesthetic	42
10	ketamine	intravenous anesthetic	38
10	isoflurane	inhalational anesthetic	35
10	propofol	intravenous anesthetic	36
10	etomidate	intravenous anesthetic	37

Supplemental Table A.1.2. Viability of anesthetic treated animals

treatment	concentration	alive/total 10 min	alive/total 1 hour	alive/total 5 hours
DMSO	0 µM	100/100	100/100	100/100
etomidate	3 µM	100/100	100/100	100/100
etomidate	6 µM	100/100	100/100	100/100
etomidate	12 µM	100/100	100/100	98/100
propofol	3 µM	100/100	100/100	100/100
propofol	6 µM	100/100	100/100	100/100
propofol	12 µM	100/100	100/100	99/100

Supplementary Table A.1.3. GABAAR ligand reference set

Chemical Name	Class	Reference
dmso	vehicle control	NA
baclofen	$GABA_B$ agonist	4
GABA	orthosteric $GABA_AR$ agonist	5
muscimol	orthosteric $GABA_AR$ agonist	6
gaboxadol (THIP)	$GABA_AR$ delta subtype preferring PAM	7
DS-2	$GABA_AR$ delta subtype preferring PAM	7
ocinaplon	$GABA_AR$ BZ-site PAM	11
bromazepam	$GABA_AR$ BZ-site PAM	8
etizolam	$GABA_AR$ BZ-site PAM	12
alprazolam	$GABA_AR$ BZ-site PAM	5
oxazepam	$GABA_AR$ BZ-site PAM	8
clobazam	$GABA_AR$ BZ-site PAM	9
temazepam	$GABA_AR$ BZ-site PAM	10
diazepam	$GABA_AR$ BZ-site PAM	5
stripentol	$GABA_AR$ non BZ-site PAM	14
methaqualone	$GABA_AR$ non BZ-site PAM	15
valerinic acid	$GABA_AR$ non BZ-site PAM	16
thiopental	$GABA_AR$ non BZ-site PAM	18
tracazolate	$GABA_AR$ non BZ-site PAM	13
carboetomidate	$GABA_AR$ anesthetic PAM	18
propofol	$GABA_AR$ anesthetic PAM	18
etomidate	$GABA_AR$ anesthetic PAM	18

Chemical Name	Class	Reference
tetrahydrodeoxycorticosterone (THDOC)	$GABA_AR$ neurosteroid PAM	17
alphaxalone	$GABA_AR$ neurosteroid PAM	17
progesterone	$GABA_AR$ neurosteroid PAM	17
deoxycorticosterone (DOC)	$GABA_AR$ neurosteroid PAM	17

Supplementary Table A.1.4. Chemical names and SMILES of the top 125 hit compounds from a zebrafish behavioral drug screen.

Chemical Name	SMILES	Retested
progesterone	CC(=O)C1CCC2C3CCC4=CC(=O)CCC4(C)C3CCC12C	yes
alfaxalone	CC(=O)C1CCC2C3CCC4CC(O)CCC4(C)C3C(=O)CC12C	yes
DOC	CC12CCC(=O)C=C1CCC1C2CCC2(C)C(C(=O)CO)CCC12	yes
7166683	Cc1cc(C)n2c(SCc3ccc(C(=O)c4ccccc4)cc3)nnc2n1	yes
etomidate	CCOC(=O)c1cncn1C(C)c1ccccc1	yes
6587027	CC(=O)c1cccc(NC(=O)c2ccc(-c3ccc(Cl)cc3)o2)c1	no
etomidate	CCOC(=O)c1cncn1C(C)c1ccccc1	yes
6858658	O=C(c1cc(Cl)ccc1Cl)N1CCN(c2cccc(C(F)(F)F)c2)CC1	no
5846886	Nc1ccc(Oc2ccc(Cl)c3cccnc23)c(Cl)c1	yes
6767569	Cc1nc2ccccc2n1C(=O)N(c1ccccc1)c1ccccc1	yes
6762995	Fc1ccccc1OCCn1c(S)nc2ccccc21	yes
alfaxalone	CC(=O)C1CCC2C3CCC4CC(O)CCC4(C)C3C(=O)CC12C	yes
7013338	CCc1cc2c(=O)c(-c3ccccc3Cl)c(C)oc2cc1OC	yes
7010474	CCc1cc2c(=O)c(-c3ccccc3Cl)coc2cc1OC	no
6376886	COc1cc(OC)cc(C(O)=Nc2ccc(Cl)cc2C(F)(F)F)c1	yes
7100598	O=C(c1ccccc1)c1cccc(N=C(O)Cc2ccc(Cl)cc2)c1	yes
7113584	Nc1ccc(OC(F)(F)F)cc1C(=O)c1ccccc1	yes
7114005	CCOC(=O)c1c(C)n(Cc2ccco2)c2ccc(OC)cc12	yes
6576466	O=[N+]([O-])c1cccc(C(O)=Nc2ccc(Cl)cc2C(F)(F)F)c1	no
6029941	COc1ccc(OC)c(NS(=O)(=O)c2ccc(OC)c(Br)c2)c1	yes
7285168	OC(=Nc1ccc(Cl)cc1F)c1cccc2ccccc12	no
7136301	OC(=Nc1ccc(Cl)c(Cl)c1)c1ccc(Cl)cc1Cl	yes
6890102	CCOc1ccc2c(=O)c(-c3ccccc3Cl)c(C)oc2c1	no
6225936	O=S(=O)(Nc1cccc(Br)c1)c1ccc2c(c1)OCCO2	no
6525784	O=c1c2ccccc2nc(C=Cc2ccc([N+](=O)[O-])cc2)n1-c1ccccc1Cl	no
7273455	Cc1cccc(OCCn2c(NC(=O)c3ccco3)nc3ccccc32)c1	yes
6314322	CC(C)(C)N=C(O)COc1ccc(C(C)(C)c2ccccc2)cc1	Yes

Chemical Name	SMILES	Retested
6474599	CCCN(CCC)S(=O)(=O)c1ccc(Cl)c(Cl)c1OC	yes
6682129	C=CCn1c(SCC(=O)c2cccs2)nc2sc(CC)cc2c1=O	no
ivermectin	CCC(C)C1OC2(CCC1C)CC1CC(CC=C(C)C(OC3CC(OC)C(OC4CC(OC)C(O)	yes
6028165	C(C)O4)C(C)O3)C(C)C=CC=C3COC4C(O)C(C)=CC(C(=O)O1C34O)O2	
	COc1ccc(S(=O)(=O)NC2CCCC2)cc1Br	yes
7271289	Cc1cc(Cl)ccc1OCC(=O)N1CCN(c2ccc(Cl)cc2)CC1	yes
5584178	CCc1nc2c(cnn2-c2ccccc2)c(=O)n1-c1ccc(C)cc1	no
5735460	COc1ccccc1CC(O)=Nc1ccc(Br)cc1F	no
7305598	OC(=Nc1cccc(N=C(O)c2cc(Cl)ccc2Cl)c1)c1cc(Cl)ccc1Cl	no
6993015	CC(C)CC(O)=NC(S)=NCc1ccccc1	yes
progesterone	CC(=O)C1CCC2C3CCC4=CC(=O)CCC4(C)C3CCC12C	yes
6645327	COc1ccccc1-c1coc2cc(OC(=O)c3cccs3)ccc2c1=O	yes
6772634	O=c1c2ccccc2nc(S)n1Cc1ccccc1	no
6366118	COc1ccc(S(=O)(=O)N2CCC(C)CC2)cc1Br	yes
7282929	CCOC(=O)c1c(N=C(O)C2CCCC2)sc2c1CCCC2	yes
6193422	O=S(=O)(Nc1cccc(Cl)c1)c1ccc2c(c1)OCCO2	no
5694163	Cc1csc(N=C(O)C2c3ccccc3Oc3ccccc32)n1	no
5754452	COC(=O)c1cccc(N=C(O)COc2ccc(C(C)(C)C)cc2)c1	yes
6565557	CCCCOc1ccc(C(=O)NC(C)C2COc3ccccc3O2)cc1	yes
5587256	O=C(Nc1c(C(=O)Nc2ccc(Cl)cc2)cnn1-c1ccccc1)c1ccccc1F	yes
5551268	Clc1ccc(CSc2nc3ccccc3[nH]2)cc1	yes
5149665	Brc1ccc(CSc2nc3ccccc3[nH]2)cc1	no
5869570	CC(O)=NC(=Nc1cccc(C(F)(F)F)c1)Nc1nc(C)cc(C)n1	yes
5729729	O=C(CC(c1ccccc1)c1ccccc1)N1CCN(c2ccc(F)cc2)CC1	yes
6877352	CC(c1ccccc1)n1c(S)nc2ccccc21	no
6595002	CCc1ccc(O)c(N=C(S)N=C(O)c2ccccc2I)c1	no
5352629	COc1ccccc1N1CCN(C(=O)c2ccc(C(=O)c3c(C)cc(C)cc3C)cc2)CC1	yes
7184284	CCc1cccc(C)c1N=C(O)COc1ccc(Cl)cc1Cl	no
7114335	Nc1ccc(SC(F)(F)F)cc1C(=O)c1ccccc1	yes
7014338	CCc1cc2c(=O)c(-c3ccccc3Cl)coc2cc1O	yes
DOC	CC12CCC(=O)C=C1CCC1C2CCC2(C)C(C(=O)CO)CCC12	yes
5142031	CCOC(=O)c1c(N)sc2ccccc12	yes
5695025	CCOC(=O)c1c(N=C(O)OCC)sc(C)c1CC	yes
5935291	Cc1ccc(N2C(=O)CC(N3CCN(c4cccc(Cl)c4)CC3)C2=O)cc1Cl	yes
6625147	CC1CC(OCC(O)CN2CCN(c3ccccc3F)CC2)CC(C)(C)C1.Cl.Cl	no
6340625	O=C(c1cc2ccccc2o1)N1CCCC(c2ccccc2)C1	no
6383686	C=C(C)Cn1c(-c2ccc(OC)c(OC)c2)nc2ccccc21	no

Chemical Name	SMILES	Retested
5813444	CCC(O)=Nc1nc(-c2ccccc2)nc(SC)c1C(C)=O	no
6570890	CCOc1ccc(NCc2ccccc2NS(=O)(=O)c2ccc(C)cc2)cc1	no
6400155	O=C(c1cc2ccccc2o1)N1CCN(c2ccc(C(F)(F)F)cc2[N+](=O)[O-])CC1	yes
5951201	CCOC(=O)c1c(C)nc(-c2ccccc2)nc1N=C(O)Cc1ccccc1	yes
riluzole	N=c1[nH]c2ccc(OC(F)(F)F)cc2s1	yes
5583877	O=C(Nc1ccc(Cl)cc1)c1cnn(-c2ccccc2)c1N=C(O)c1ccco1	yes
6642835	Cc1c(O)ccc(C(=O)Cc2ccccc2)c1O	no
5573728	CC(=O)c1c(C)nc2n(Cc3ccccc3)c3ccccc3n12	yes
5658603	CCCCn1c2ccccc2n2c(C(=O)OCC)c(C)nc12	yes
6013263	COc1ccc(S(=O)(=O)Nc2ccc(C)cc2)cc1Br	yes
riluzole	N=c1[nH]c2ccc(OC(F)(F)F)cc2s1	yes
6652383	CC(C)(C)C(=O)Oc1ccc2c(=O)c(-c3ccccc3Cl)coc2c1	no
6522346	CC(=O)Oc1ccc(OC(C)=O)c(S(=O)(=O)c2ccc(Cl)c(Cl)c2)c1	no
6271180	Cc1cccc(OCC(=O)N2CCN(c3ccc(C(F)(F)F)cc3[N+](=O)[O-])CC2)c1C	yes
5480577	CCOC(=O)c1cc2n(c1N=CN(C)C)-c1ccccc1C2=O	yes
6204912	CCC(C)c1ccc(NC(=O)CC(C)(C)C)cc1	no
6386892	Cc1cc(C)c(N=C(O)Cc2cccs2)c(C)c1	no
5846693	OC(=NCc1ccco1)c1cc2nc(-c3ccc(F)cc3)cc(C(F)(F)F)n2n1	no
6030006	CCOc1ccccc1NS(=O)(=O)c1ccc(OC)c(Br)c1	yes
6756477	COc1cccc(-n2nnnc2SCC(O)=Nc2c(C)cc(C)cc2C)c1	no
6353053	CC(C)(c1ccccc1)c1ccc(OCC(O)=Nc2ccccc2C(=N)O)cc1	no
6667020	CCCOC(O)=Nc1nc(-c2ccc(C)cc2)c(C)s1	no
5978667	OC1=Nc2c(ccc3ccccc23)C(c2ccc(C(F)(F)F)cc2)C1	no
5649824	CCC(=O)c1c(C)nc2n(Cc3ccccc3)c3ccccc3n12	yes
7211089	O=[N+]([O-])c1cccc(N=C(O)c2ccc(Cl)cc2Cl)c1	no
5577990	O=C(c1c(N=C(O)c2ccccc2Cl)sc2c1CCCCC2)N1CCCCC1	no
6805976	CCC(C)c1ccccc1N=C(O)c1ccccc1I	no
5799128	CCCCOc1nnc(-c2ccccc2)c2ccccc12	no
5795075	Clc1ccc(OCc2nc(-c3ccccn3)no2)c(Br)c1	no
7115521	O=C1CCCc2c1[nH]c1ccc(C(F)(F)F)cc21	no
7145248	O=C(NCc1cccc(Oc2ccccc2)c1)C(=O)c1c[nH]c2ccccc12	yes
7015047	Cc1ccc(C(C)C)c(OCCNCc2ccccc2)c1	no
5802987	Clc1ccc(OCc2nc(-c3cccnc3)no2)c(Br)c1	no
5982161	COC(=O)c1c(N=C(O)CCN2C(=O)c3ccccc3C2=O)sc2c1CCC2	no
6647373	CC(=O)Oc1ccc2c(=O)c(Oc3ccc(F)cc3)c(C(F)(F)F)oc2c1C	no
7112518	CCOC(=O)COc1ccc2c(c1)c(C(C)=O)c(C)n2Cc1ccccc1	No

Chemical Name	SMILES	Retested
6653154	CC(=O)Oc1ccc2c(=O)c(Oc3ccccc3)c(C(F)(F)F)oc2c1C	no
6689594	CCc1cc2c(=O)c(Oc3ccccc3F)c(C)oc2cc1OC(C)=O	no
6793728	Cc1cc(=Nc2cccc3ccccc23)c2cc(Cl)ccc2[nH]1.Cl	no
5846232	CC(=O)c1cc(F)c(N2CCN(C(=O)c3ccc(Cl)cc3Cl)CC2)cc1C	yes
6639477	COc1ccc2c(=O)c(-c3ccccc3Cl)c(C)oc2c1	no
5128592	Cc1cccc(NC(=O)c2ccc(C(C)(C)C)cc2)c1	yes
6216715	Cc1cccc(C)c1N=C(O)CSCc1cccc(Br)c1	no
6053334	O=C(Cc1cccc(Cl)c1)Nc1ccc(Cl)cc1	no
6646535	COc1ccc2c(=O)c(-c3ccccc3)c(C)oc2c1C	no
6367818	CCC(=O)Nc1ccc(Oc2ccc(Cl)cc2)cc1	no
7115235	CC(C)(C)C(=O)CSc1nc2ccccc2c(=O)n1Cc1ccccc1	no
promazine	CN(C)CCCN1c2ccccc2Sc2ccccc21	no
6661919	CCCCN=C(S)N=C(O)c1ccc(Cl)cc1Cl	yes
6163607	O=S(=O)(Nc1ccc(OCc2ccccc2)cc1)c1ccc2c(c1)OCCO2	yes
5690842	COC(=O)c1sc(N=C(O)C2CCCO2)c(C(=O)OC)c1C	yes
6673619	Cc1ccccc1CC(=O)c1ccc(O)c(C)c1O	no
7011253	COc1ccccc1-c1nnc(SCc2ccccc2F)n1-c1ccccc1	no
7150160	O=C(CC(CC(=O)c1ccccc1)c1ccccc1)c1ccccc1	no
5867832	COc1ccc(N=C(O)c2ccccc2Oc2ccccc2)c(OC)c1	no
5942595	Cc1ccc(Cl)cc1N=C(O)COc1c(C)cccc1C	yes
6150813	Cc1noc(C)c1CSc1nc2sc3c(c2c(=O)n1-c1ccccc1)CCC3	no
6165550	Cc1cccc(-n2c(C=Cc3cccc([N+](=O)[O-])c3)nc3ccccc3c2=O)c1	no
6678692	CCc1cc2c(=O)c(Oc3ccc(F)cc3)coc2cc1OC(C)=O	no
7237541	Cc1ccc2c3c1C(=O)C(=O)N3C(C)(C)CC2(C)c1ccccc1	yes

Supplementary Table A.1.5. Compound names and summary of results from *in vitro* FLIPR experiments for GABAA.

Compound Name	FLIPR Results	Ave Norm FLIPR Score	SD
DMSO	Negative	1	0.992
BGC 20-761	Negative	1.188647747	3.043
picrotoxin	Negative	-9.759599332	2.087
progesterone	Positive	17.80300501	4.5
etomidate	Positive	50.67111853	0.892
tracazolate	Positive	50.13856428	2.255
propofol	Positive	34.52337229	8.439
thiopental	Positive	11.39232053	2.274
DOC	Negative	-4.308013356	2.091
diazepam	Negative	-10.89565944	1.052
7013338	Positive	116.3580968	3.651
5942595	Positive	84.9148581	6.339
5658603	Positive	55.23539232	3.39
6474599	Positive	52.02420702	3.831
6537142	Positive	50.97579299	1.64
5649824	Positive	41.41402337	4.859
5583877	Positive	36.70784641	3.964
5925611	Positive	32.94657763	0.645
6376886	Positive	32.44240401	6.486
5695025	Positive	32.2721202	4.675
6028165	Positive	28.75792989	3.955
5573728	Positive	27.62103506	2.061
5860357	Positive	25.8706177	3.621
6993015	Positive	23.92654424	2.021
5937132	Positive	20.70033389	3.122
7282929	Positive	17.15317195	2.717
7114005	Positive	15.23205343	0.272
5869570	Positive	15.03672788	0.574
6366118	Positive	14.5033389	0.336
6029941	Positive	14.00584307	NA
5690842	Positive	12.09432387	6.775
7273455	Positive	7.892320535	5.349
7100598	Positive	7.736227046	0.217
6091285	Positive	7.664440735	0.253
5986291	Positive	6.420701169	4.018
5142031	Positive	5.23706177	0.349
6030006	Negative	3.04590985	3.264
5768306	Negative	2.243739566	0.059
6013263	Negative	1.035058431	2.342

Compound Name	FLIPR Results	Ave Norm FLIPR	SD
5102870	Negative	-0.944908181	1.695
7166683	Negative	-1.302170284	0.581
5811265	Negative	-1.4557596	3.792
5551268	Negative	-1.463272121	1.876
7113584	Negative	-2.454924875	0.79
5835629	Negative	-2.682804675	3.707
5845856	Negative	-2.947412354	1.101
6645327	Negative	-4.733722872	0.776
6212662	Negative	-5.178631052	0.047
5846886	Negative	-5.25542571	0.94
6163607	Negative	-5.679465776	1.265
7014338	Negative	-6.270450752	2.581
6246227	Negative	-6.910684474	1.718
7284610	Negative	-7.29966611	4.074
6366421	Negative	-8.515859767	1.221
5736224	Negative	-10.45158598	1.238
6225936	Negative	-14.696995	0.004
7145248	Negative	-24.78631052	0.673
5587256	Negative	-26.36894825	22.82

Supplementary Table A.1.6. SEA predictions prioritized by EF from the top 1000 hit compounds.

CHEMBL ID	EF	p-Value	q-Value	Description
CHEMBL5469	10.832	2.82E-153	9.38E-148	Protein tyrosine kinase 2 beta
CHEMBL2094122	9.2246	3.67E-20	1.22E-14	GABA-A receptor; alpha-5/beta-3/gamma-2
CHEMBL3746	8.2498	5.69E-90	1.89E-84	11-beta-hydroxysteroid dehydrogenase 2
CHEMBL3012	6.1087	6.66E-54	2.22E-48	Phosphodiesterase 7A
CHEMBL4501	5.2561	9.62E-22	3.20E-16	Ribosomal protein S6 kinase 1
CHEMBL5936	5.2175	8.15E-29	2.71E-23	Toll-like receptor 7
CHEMBL1907607	5.16	1.14E-53	3.79E-48	GABA-A receptor; anion channel
CHEMBL4409	4.6661	1.65E-09	0.000549	Phosphodiesterase 10A
CHEMBL2835	4.3081	2.33E-56	7.74E-51	Tyrosine-protein kinase JAK1
CHEMBL1787	4.3081	4.35E-30	1.45E-24	Steroid 5-alpha-reductase 1
CHEMBL2095227	4.2252	3.99E-178	1.33E-172	Vascular endothelial growth factor receptor
CHEMBL4296	4.2035	1.41E-35	4.69E-30	Sodium channel protein type IX alpha subunit
CHEMBL3055	4.1484	9.87E-10	0.000328	Cyclin-dependent kinase 7
CHEMBL4975	4.0012	8.05E-14	2.68E-08	Adenosine A1 receptor
CHEMBL2034	3.827	6.36E-34	2.12E-28	Glucocorticoid receptor
CHEMBL1856	3.5805	9.87E-79	3.28E-73	Steroid 5-alpha-reductase 2

CHEMBL4977	3.5709	8.07E-63	2.69E-57	Proto-oncogene c-JUN
CHEMBL ID	**EF**	**p-Value**	**q-Value**	**Description**
CHEMBL2488	3.5637	3.50E-43	1.16E-37	Prostanoid EP2 receptor
CHEMBL4040	3.4143	7.31E-47	2.43E-41	MAP kinase ERK2
CHEMBL259	3.3269	5.86E-17	1.95E-11	Melanocortin receptor 4
CHEMBL3072	3.2333	7.50E-24	2.49E-18	Androgen Receptor
CHEMBL1907605	3.1718	1.64E-15	5.46E-10	Cyclin-dependent kinase 2/cyclin E1
CHEMBL1918	3.0062	8.12E-184	2.70E-178	Glutamate receptor ionotropic kainate 1
CHEMBL2337	2.8332	9.25E-110	3.08E-104	Glycine transporter 1
CHEMBL3687	2.7944	2.78E-49	9.26E-44	Arachidonate 12-lipoxygenase
CHEMBL4430	2.7913	4.86E-43	1.62E-37	Cytochrome P450 17A1
CHEMBL330	2.7913	3.71E-67	1.24E-61	Glutamate (NMDA) receptor subunit zeta 1
CHEMBL3514	2.7732	4.06E-30	1.35E-24	LDL-associated phospholipase A2
CHEMBL5652	2.7499	2.49E-110	8.27E-105	Glucose-dependent insulinotropic receptor
CHEMBL2569	2.7405	3.98E-110	1.32E-104	Microsomal triglyceride transfer protein large subunit
CHEMBL2971	2.6595	2.83E-47	9.41E-42	Tyrosine-protein kinase JAK2
CHEMBL5658	2.6443	6.50E-66	2.16E-60	Prostaglandin E synthase
CHEMBL1293255	2.6013	8.60E-140	2.86E-134	15-hydroxyprostaglandin dehydrogenase [NAD+]
CHEMBL1889	2.5984	4.47E-68	1.49E-62	Vasopressin V1a receptor
CHEMBL2568	2.5325	5.69E-92	1.89E-86	Liver glycogen phosphorylase
CHEMBL3230	2.4991	3.40E-103	1.13E-97	Sphingosine 1-phosphate receptor Edg-6
CHEMBL4036	2.454	2.29E-26	7.61E-21	Cyclin-dependent kinase 5
CHEMBL2903	2.4465	7.17E-17	2.39E-11	Arachidonate 15-lipoxygenase
CHEMBL3227	2.4305	6.64E-166	2.21E-160	Metabotropic glutamate receptor 5
CHEMBL1790	2.3894	1.62E-28	5.39E-23	Vasopressin V2 receptor
CHEMBL275	2.3347	2.22E-16	7.40E-11	Phosphodiesterase 4B
CHEMBL4652	2.3307	3.35E-40	1.11E-34	Somatostatin receptor 1
CHEMBL3371	2.2818	1.39E-12	4.62E-07	Serotonin 6 (5-HT6) receptor
CHEMBL5409	2.2229	7.73E-26	2.57E-20	G-protein coupled bile acid receptor 1
CHEMBL3351	2.1837	2.62E-12	8.72E-07	Acetyl-CoA carboxylase 1
CHEMBL4336	2.1425	6.56E-15	2.18E-09	Prostanoid EP3 receptor
CHEMBL235	2.1375	3.50E-48	1.16E-42	Peroxisome proliferator-activated receptor gamma
CHEMBL2095160	2.1118	2.42E-09	0.000807	Leukotriene B4 receptor
CHEMBL5071	2.0202	2.18E-57	7.25E-52	G protein-coupled receptor 44
CHEMBL1811	1.9836	1.35E-19	4.51E-14	Prostanoid EP1 receptor
CHEMBL4315	1.9642	1.48E-22	4.93E-17	Purinergic receptor P2Y1

CHEMBL ID	EF	p-Value	q-Value	Description
CHEMBL1906	1.9542	1.78E-33	5.91E-28	Serine/threonine-protein kinase RAF
CHEMBL2001	1.9359	1.07E-72	3.55E-67	Purinergic receptor P2Y12
CHEMBL4478	1.9234	9.22E-19	3.07E-13	Voltage-gated N-type calcium channel alpha-1B subunit
CHEMBL3338	1.9062	1.44E-28	4.79E-23	Squalene synthetase
CHEMBL1966	1.8407	7.61E-12	2.53E-06	Dihydroorotate dehydrogenase
CHEMBL2993	1.8048	1.81E-19	6.01E-14	Monoamine oxidase B
CHEMBL4051	1.7585	8.50E-33	2.83E-27	Cystic fibrosis transmembrane conductance regulator
CHEMBL3974	1.7366	3.74E-29	1.24E-23	Proteinase-activated receptor 1
CHEMBL2868	1.7226	1.58E-63	5.24E-58	Vasopressin V1a receptor
CHEMBL1868	1.694	3.93E-17	1.31E-11	Vascular endothelial growth factor receptor 1
CHEMBL2736	1.6741	1.24E-32	4.14E-27	Metabotropic glutamate receptor 4
CHEMBL244	1.6674	2.31E-22	7.70E-17	Coagulation factor X
CHEMBL3553	1.6663	4.38E-44	1.46E-38	Tyrosine-protein kinase TYK2
CHEMBL285	1.6627	2.46E-10	8.20E-05	Acyl coenzyme A:cholesterol acyltransferase 1
CHEMBL1901	1.6571	2.88E-28	9.58E-23	Cholecystokinin A receptor
CHEMBL1741186	1.6522	2.90E-36	9.66E-31	Nuclear receptor ROR-gamma
CHEMBL311	1.6361	5.80E-22	1.93E-16	Glutamate [NMDA] receptor subunit epsilon 2
CHEMBL3969	1.6281	2.92E-10	9.72E-05	Carbonic anhydrase VB
CHEMBL4892	1.6142	6.42E-10	0.000213	Alpha-1a adrenergic receptor
CHEMBL3238	1.6083	4.80E-75	1.60E-69	Carnitine palmitoyltransferase 2
CHEMBL249	1.6064	1.82E-46	6.06E-41	Neurokinin 1 receptor
CHEMBL3766	1.5872	3.26E-14	1.09E-08	Vasopressin V2 receptor
CHEMBL2413	1.5732	1.63E-13	5.43E-08	C-C chemokine receptor type 1
CHEMBL2047	1.5349	1.17E-10	3.88E-05	Bile acid receptor FXR
CHEMBL3858	1.5283	7.25E-56	2.41E-50	Carnitine palmitoyltransferase 1A
CHEMBL2564	1.5215	5.48E-46	1.82E-40	Metabotropic glutamate receptor 5
CHEMBL1913	1.4884	1.72E-11	5.72E-06	Platelet-derived growth factor receptor beta
CHEMBL1293194	1.4697	1.28E-46	4.25E-41	Carnitine O-palmitoyltransferase 1 liver isoform
CHEMBL2095150	1.4663	4.96E-23	1.65E-17	Phosphodiesterase 1
CHEMBL3156	1.457	8.36E-13	2.78E-07	Thromboxane A2 receptor
CHEMBL4018	1.4434	3.32E-17	1.10E-11	Neuropeptide Y receptor type 2
CHEMBL1844	1.3767	3.24E-17	1.08E-11	Macrophage colony stimulating factor receptor
CHEMBL2216739	1.3745	2.01E-23	6.68E-18	Carnitine O-palmitoyltransferase 1 muscle isoform
CHEMBL4722	1.3586	2.22E-13	7.40E-08	Serine/threonine-protein kinase Aurora-A

CHEMBL ID	EF	p-Value	q-Value	Description
CHEMBL2093869	1.3358	4.11E-15	1.37E-09	Integrin alpha-IIb/beta-3
CHEMBL2093866	1.3354	6.69E-25	2.23E-19	Estrogen receptor
CHEMBL5669	1.2425	3.27E-24	1.09E-18	Epoxide hydrolase 2
CHEMBL4140	1.2125	3.30E-37	1.10E-31	Epoxide hydratase
CHEMBL256	1.1181	3.59E-12	1.19E-06	Adenosine A3 receptor
CHEMBL3572	1.0034	8.31E-15	2.77E-09	Cholesteryl ester transfer protein
CHEMBL3105	0.89889	1.50E-20	4.99E-15	Poly [ADP-ribose] polymerase-1
CHEMBL1937	0.86861	2.11E-25	7.01E-20	Histone deacetylase 2
CHEMBL1951	0.8684	3.01E-09	0.001	Monoamine oxidase A
CHEMBL255	0.85701	2.34E-19	7.79E-14	Adenosine A2b receptor
CHEMBL2111429	0.8541	7.13E-15	2.37E-09	Histone deacetylase (HDAC1 and HDAC2)
CHEMBL3254	0.82456	9.65E-20	3.21E-14	Monoamine oxidase A
CHEMBL273	0.82284	3.24E-17	1.08E-11	Serotonin 1a (5-HT1a) receptor
CHEMBL2095189	0.78671	5.92E-11	1.97E-05	Platelet-derived growth factor receptor
CHEMBL251	0.78053	3.07E-14	1.02E-08	Adenosine A2a receptor
CHEMBL321	0.77988	1.51E-14	5.02E-09	Matrix metalloproteinase 9
CHEMBL6009	0.76082	2.88E-15	9.60E-10	Diacylglycerol O-acyltransferase 1
CHEMBL4191	0.74784	1.01E-11	3.36E-06	Monoglyceride lipase
CHEMBL205	0.73443	1.20E-10	3.99E-05	Carbonic anhydrase II
CHEMBL3571	0.73416	2.94E-14	9.79E-09	Cannabinoid CB1 receptor
CHEMBL3192	0.72428	1.63E-38	5.42E-33	Histone deacetylase 8
CHEMBL5393	0.71706	5.11E-59	1.70E-53	ATP-binding cassette sub-family G member 2
CHEMBL219	0.68396	5.99E-16	1.99E-10	Dopamine D4 receptor
CHEMBL3594	0.68047	7.19E-21	2.39E-15	Carbonic anhydrase IX
CHEMBL281	0.67296	3.75E-19	1.25E-13	Carbonic anhydrase IV
CHEMBL1945	0.66941	1.07E-24	3.57E-19	Melatonin receptor 1A
CHEMBL1980	0.6652	2.77E-10	9.22E-05	Sodium channel protein type V alpha subunit
CHEMBL332	0.65466	2.02E-22	6.71E-17	Matrix metalloproteinase-1
CHEMBL4302	0.65163	1.11E-12	3.71E-07	P-glycoprotein 1
CHEMBL3138	0.64379	1.19E-20	3.98E-15	Dopamine D3 receptor
CHEMBL246	0.62469	5.68E-10	0.000189	Beta-3 adrenergic receptor
CHEMBL261	0.62233	6.21E-24	2.07E-18	Carbonic anhydrase I
CHEMBL2095171	0.58411	9.03E-10	0.0003	Sodium channel alpha subunits; brain (Types I II III)
CHEMBL3180	0.58231	1.61E-12	5.36E-07	Carboxylesterase 2
CHEMBL2093865	0.57073	9.07E-25	3.02E-19	Histone deacetylase
CHEMBL4792	0.55951	8.64E-24	2.87E-18	Orexin receptor 2

CHEMBL ID	EF	p-Value	q-Value	Description
CHEMBL2093870	0.55882	4.04E-20	1.34E-14	Serotonin 2 (5-HT2) receptor
CHEMBL260	0.55449	2.22E-15	7.37E-10	MAP kinase p38 alpha
CHEMBL3898	0.55037	1.05E-20	3.50E-15	Bone morphogenetic protein 1
CHEMBL1878	0.54726	1.45E-14	4.82E-09	Calcium sensing receptor
CHEMBL2096671	0.53752	1.76E-11	5.86E-06	Serotonin 2 (5-HT2) receptor
CHEMBL3229	0.53436	5.63E-21	1.87E-15	Anandamide amidohydrolase
CHEMBL3473	0.53004	2.87E-34	9.54E-29	C-C chemokine receptor type 3
CHEMBL234	0.52084	4.19E-46	1.39E-40	Dopamine D3 receptor
CHEMBL1946	0.51445	3.27E-42	1.09E-36	Melatonin receptor 1B
CHEMBL214	0.51376	1.16E-28	3.84E-23	Serotonin 1a (5-HT1a) receptor
CHEMBL3455	0.5133	1.69E-28	5.62E-23	Anandamide amidohydrolase
CHEMBL2094268	0.50345	2.06E-31	6.84E-26	Melatonin receptor
CHEMBL283	0.50011	9.33E-46	3.10E-40	Matrix metalloproteinase 3
CHEMBL3361	0.49714	4.58E-39	1.52E-33	Dopamine D4 receptor
CHEMBL3427	0.49437	2.44E-14	8.12E-09	Dopamine D2 receptor
CHEMBL264	0.49369	5.49E-16	1.83E-10	Histamine H3 receptor
CHEMBL217	0.48105	2.23E-52	7.43E-47	Dopamine D2 receptor
CHEMBL3465	0.48013	9.07E-15	3.02E-09	Sigma opioid receptor
CHEMBL220	0.47337	8.15E-17	2.71E-11	Acetylcholinesterase
CHEMBL2409	0.44961	2.31E-28	7.68E-23	Epoxide hydratase
CHEMBL4588	0.4342	2.36E-49	7.85E-44	Matrix metalloproteinase 8
CHEMBL1873	0.42696	3.40E-20	1.13E-14	Tissue-type plasminogen activator
CHEMBL325	0.42455	3.61E-39	1.20E-33	Histone deacetylase 1
CHEMBL3286	0.41945	7.47E-29	2.49E-23	Urokinase-type plasminogen activator
CHEMBL287	0.39167	2.26E-39	7.51E-34	Sigma opioid receptor
CHEMBL3223	0.37809	4.88E-32	1.62E-26	Serotonin 7 (5-HT7) receptor
CHEMBL3199	0.36708	2.07E-29	6.88E-24	Acetylcholinesterase
CHEMBL3602	0.33883	6.17E-42	2.05E-36	Sigma opioid receptor
CHEMBL3198	0.3105	2.27E-38	7.55E-33	Acetylcholinesterase

Supplementary Table A.1.7. SEA predictions prioritized by EF from the top 30 hit compounds.

CHEMBL ID	EF	p-Value	q-Value	description
CHEMBL3746	53.234	0	0	11-beta-hydroxysteroid dehydrogenase 2
CHEMBL1787	27.799	7.30E-267	2.43E-261	Steroid 5-alpha-reductase 1
CHEMBL3072	20.864	1.26E-225	4.18E-220	Androgen Receptor
CHEMBL1856	19.253	5.42E-146	1.80E-140	Steroid 5-alpha-reductase 2
CHEMBL1907607	14.798	6.99E-127	2.33E-121	GABA-A receptor; anion channel
CHEMBL2096664	8.055	2.14E-54	7.12E-49	Steroid 5-alpha-reductase
CHEMBL230	7.6613	1.04E-24	3.47E-19	Cyclooxygenase-2
CHEMBL2903	5.92	1.39E-48	4.61E-43	Arachidonate 15-lipoxygenase
CHEMBL2095227	4.8113	2.18E-125	7.25E-120	Vascular endothelial growth factor receptor
CHEMBL1889	4.5727	3.51E-99	1.17E-93	Vasopressin V1a receptor
CHEMBL1918	4.4766	5.57E-125	1.85E-119	Glutamate receptor ionotropic kainate 1
CHEMBL3371	4.4172	2.91E-50	9.70E-45	Serotonin 6 (5-HT6) receptor
CHEMBL2564	4.0909	1.99E-49	6.60E-44	Metabotropic glutamate receptor 5
CHEMBL2337	3.6564	5.47E-147	1.82E-141	Glycine transporter 1
CHEMBL2568	3.4403	2.69E-29	8.97E-24	Liver glycogen phosphorylase
CHEMBL244	3.3105	3.62E-30	1.20E-24	Coagulation factor X
CHEMBL2993	2.9115	2.77E-20	9.22E-15	Monoamine oxidase B
CHEMBL3230	2.8458	8.52E-70	2.83E-64	Sphingosine 1-phosphate receptor Edg-6
CHEMBL1966	2.6394	3.89E-14	1.29E-08	Dihydroorotate dehydrogenase
CHEMBL2868	2.6154	1.21E-59	4.04E-54	Vasopressin V1a receptor
CHEMBL3238	2.5944	7.37E-27	2.45E-21	Carnitine palmitoyltransferase 2
CHEMBL3766	2.5604	2.80E-50	9.32E-45	Vasopressin V2 receptor
CHEMBL2216739	2.2174	5.63E-11	1.87E-05	Carnitine O-palmitoyltransferase 1 muscle isoform
CHEMBL4681	2.0292	2.43E-10	8.09E-05	Aldo-keto-reductase family 1 member C3

Supplementary Table A.1.8. Reference compounds targeting receptors with relatively low EF scores

Compound	Target	InChiKey
SR 49059	V1 antagonist	CEBYCSRFKCEUSW-NAYZPBBASA-N
TCS 359	FLT3 RTK antagonist	FSPQCTGGIANIJZ-UHFFFAOYSA-N
Piclamilast	PDEantagonist	RRRUXBQSQLKHEL-UHFFFAOYSA-N
(R)-C R 8	CDK antagonist	ORYSYXHQFOWNDK-RGFWRHHQSA-N
DL-TBOA	EAAT antagonist	BYOBCYXURWDEDS-IUCAKERBSA-N
ML 218 hydrochloride	CAC1G antagonist	IDCVEUISZZKMKJ-ZXVFAPHLSA-N
AG 825	Erbb2 antagonist	KXDONFLNGBQLTN-WUXMJOGZSA-N
Tak 165	Erbb2 antagonist	ZTFBIUXIQYRUNT-MDWZMJQESA-N
Z57705210	PPBT antagonist	OTFKFIWLDAHEKF-UHFFFAOYSA-N
BX 471	CCR1 antagonist	XQYASZNUFDVMFH-CQSZACIVSA-N
ML 00253764	MC4R antagonist	KZUMGPQDDCBFBF-UHFFFAOYSA-N
Nav26	SCN9A antagonist	ICGMZCVSHDKQTE-UHFFFAOYSA-N
ATC 0065	MCHR1 antagonist	BPGUWYBAINNZQH-LFOVFOEYSA-N
5648949	PD2R2 antagonist	WVBVIRRWTMFJAK-AWQFTUOYSA-N
JQ1	BRD3 antagonist	DNVXATUJJDPFDM-KRWDZBQOSA-N
R-+-methanandamide	CNR1 agonist	SQKRUBZPTNJQEM-FQPARAGTSA-N
TOPIRAMATE	GluR5 antagonist	KJADKKWYZYXHBB-XBWDGYHZSA-N
ML 202	PKM2 antagonist	MORBXZMIXGYQDB-UHFFFAOYSA-N
PF 04885614	SCNAA antagonist	AGORGFNWYAUYSU-UHFFFAOYSA-N
LY 341495	Group II mGluR antagonist	VLZBRVJVCCNPRJ-KPHUOKFYSA-N
ICI 63197	PDE4 antagonist	UQDVRVNMIJAGRK-UHFFFAOYSA-N
GW 791343 Hydrochloride	P2X7 allosteric modulator	WSBRAHWNJBXXJM-UHFFFAOYSA-N
A 740003	P2X7 antagonist	PUHSRMSFDASMAE-UHFFFAOYSA-N
TOLMETIN	IL-8 antagonist	UPSPUYADGBWSHF-UHFFFAOYSA-N
EMDT oxalate	5HT6 Agonist	IFGWAHGHGDZBEH-UHFFFAOYSA-N
Metyrapone	11β-hydroxylase antagonist	FJLBFSROUSIWMA-UHFFFAOYSA-N
ACEA	CB1 receptor agonist	SCJNCDSAIRBRIA-DOFZRALJSA-N
Strophanthidin	Na+ /K+-ATPase antagonist	ODJLBQGVINUMMR-HZXDTFASSA-N
Procaine	NaV general antagonist	MFDFERRIHVXMIY-UHFFFAOYSA-N
ZINC13108136	BACE2 antagonist	SVHVIRBGQRDCIB-UHFFFAOYSA-N
ATC 0175 hydrochloride	MCH1 antagonist	FAIMGWSOSCFGRU-UHFFFAOYSA-N
Fenobam	mGlu5 agonist	DWPQODZAOSWNHB-UHFFFAOYSA-N
NNC 55-0396 dihydrochlorid	NaV general antagonist	BCCQNBXHUMKLFW-HNQRYHMESA-N
Xylazine hydrochloride	α2-adronergic agonist	DIIBRMSCONGGIN-UHFFFAOYSA-N
Bumetanide	NKCC cotransporter antagonist	MAEIEVLCKWDQJH-UHFFFAOYSA-N
Epibatidine Hydrochloride	nicotinic agonist	NLPRAJRHRHZCQQ-UTLUCORTSA-N
(-)-Nicotine Ditartrate	Nicotinic acetylcholine receptor agonis	RFEJUZJILGIRHQ-UHFFFAOYSA-N
Hexamethonium Bromide	nicotinic receptor antagonist	FAPSXSAPXXJTOU-UHFFFAOYSA-L
Scopolomine hydrobromide	muscarinic antagonist	STECJAGHUSJQJN-FWXGHANASA-N
octopamine hydrochloride	β3 adrenoceptor agonist	QHGUCRYDKWKLMG-UHFFFAOYSA-N
Ro 32-0432 hydrochloride	protein kinase C antagonist	HSPRASOZRZDELU-LMOVPXPDSA-N
BzATP triethylammonium sal	P2X7 receptor agonist	HVOVBTNCGADRTH-WBLDMZOZSA-N
Phenytoin	NaV general antagonist	CXOFVDLJLONNDW-UHFFFAOYSA-N
TCB-2	5HT2AR agonist	TYMMXVZAUGQKRF-UHFFFAOYSA-N
MTEP	mGluR5 antagonist	NRBNGHCYDWUVLC-UHFFFAOYSA-N
Ro4368554	5HT6R antagonist	AOPYPEADLGTXRA-UHFFFAOYSA-N
Biphenyl-indanone A	mGluR2 allosteric modulator	KMKBEESNZAPKMP-UHFFFAOYSA-N
LuAE58054	5HT6R antagonist	YBAWYTYNMZWMMJ-UHFFFAOYSA-N

Supplementary Table A.1.9. Compound names and descriptions for annotated and novel small molecules that interact with 3 of the primary SEA predicted targets (mGluR, GABAA, and HTR6).

mGluR Compounds tested	SMILES	Mechanism Notes	Phenoscore
Fenobam	CN1CC(=O)N=C1NC(=O)Nc2cccc(c2)Cl	mGlu5 agonist	0.3
STK234931	c1ccc(cc1)C#Cc2ccc(cc2)C(=O)N3CCC(CC3)O	MGluR antagonist	0.4
MPEP	CC1=NC(=CC=C1)C#CC2=CC=CC=C2.Cl	mGlu5 antagonist	0.73
TOPIRAMATE	CC1(O[C@@H]2CO[C@@]3([C@H]([C@@H]2O1)OC(O3)(C)C)COS(=O)(=O)N)C	mGluR5 antagonist	0.1
UBP 302	c1ccc(c(c1)Cn2c(=O)ccn(c2=O)C[C@@H](C(=O)O)N)C(=O)O	mGluR5 antagonist	0.17
VU 0361737	COc1cc(ccc1Cl)NC(=O)c2ccccn2	mGluR4 positive allosteric modulator	0.28
VU 0155041 NaSalt	c1c(cc(cc1Cl)Cl)NC(=O)[C@H]2CCCC[C@H]2C(=O)O	mGluR4 positive allosteric modulator	0.11
Cinnabarinic acid	c1cc(c2c(c1)oc-3cc(=O)c(c(c3n2)C(=O)O)N)C(=O)O	mGluR4 agonist	0.18
L-AP4	[H][C@](N)(CCP(O)(O)=O)C(O)=O	mGluR4/6/7/8 agonist	0.1
LY-354740	C1C[C@]([C@H]2[C@@H]1[C@@H]2C(=O)O)(C(=O)O)N	mGluR2/3 agonist	0.1
Biphenylindanone A	CC1=C(C=C2CC(C(=O)C2=C1C)C3CCCC3)OCC4=CC(=CC=C4)C5=CC=C(C=C5)C(=O)O	mGluR2 positive allosteric modulator	0.11
Predicted mGluR			
7285168	OC(=Nc1ccc(Cl)cc1F)c1cccc2ccccc12	predicted mGluR	0.54
7211089	O=[N+]([O-])c1cccc(N=C(O)c2ccc(Cl)cc2Cl)c1	predicted mGluR	0.5
5128592	Cc1cccc(c1)NC(=O)c2ccc(cc2)C(C)(C)C	predicted mGluR	0.52
6576466	O=[N+]([O-])c1cccc(C(O)=Nc2ccc(Cl)cc2C(F)(F)F)c1	predicted mGluR	0.57
5795075	Clc1ccc(OCc2nc(-c3ccccn3)no2)c(Br)c1	predicted mGluR	0.39
7136301	O=C(N[C@H](c1occc1)C)COC(=O)c3cc([N+]([O-])=O)ccc3N2CCCC2	predicted mGluR	0.42
6587027	CC(=O)c1cccc(NC(=O)c2ccc(-c3ccc(Cl)cc3)o2)c1	predicted mGluR	0.57
7271289	Cc1cc(Cl)ccc1OCC(=O)N1CCN(c2ccc(Cl)cc2)CC1	predicted mGluR	0.56
7305598	OC(=Nc1cccc(N=C(O)c2cc(Cl)ccc2Cl)c1)c1cc(Cl)ccc1Cl	predicted mGluR	0.51
7100598	O=C(c1ccccc1)c1cccc(N=C(O)Cc2ccc(Cl)cc2)c1	predicted mGluR	0.54
6053334	O=C(Cc1cccc(Cl)c1)Nc1ccc(Cl)cc1	predicted mGluR	0.34
6376886	COc1cc(OC)cc(C(O)=Nc2ccc(Cl)cc2C(F)(F)F)c1	predicted mGluR	0.59
5869570	Cc1cc(nc(n1)N/C(=N\C(=O)C)/Nc2cccc(c2)C(F)(F)F)C	predicted mGluR	0.68
5583877	c1ccc(cc1)n2c(c(cn2)C(=O)Nc3ccc(cc3)Cl)NC(=O)c4ccco4	predicted mGluR	0.66
5943451	CCOC(=O)c1c(C)n(Cc2ccco2)c2ccc(OC)cc21	predicted mGluR	0.37
GABA Compounds Tested			
diazepam	CN1c2ccc(cc2C(=NCC1=O)c3ccccc3)Cl	GABAA agonist	0.68
propofol	Oc1c(cccc1C(C)C)C(C)C	GABAA agonist	0.78
DMCM	CCc1c(ncc2[nH]c3cc(OC)c(OC)cc3c12)C(=O)OC	GABAA negative allosteric modulator	0.54
β-CCE	CCOC(=O)c1cc2c3ccccc3[nH]c2cn1	GABAA inverse agonist	0.37
flumazenil	CCOC(=O)c1ncn2c1CN(C)C(=O)c1cc(F)ccc1-2	GABAA antagonist	0.1
TB 21007	CC1(CC2=C(SC(=C2C(=O)C1)SCCO)C3=NC=CS3)C	GABAA inverse agonist	0.12
bicuculline (-)	CN1CCc2cc3c(cc2[C@H]1[C@H]4c5ccc6c(c5C(=O)O4)OCO6)OCO3	GABAA antagonist	0.05
picrotoxin	C[C@@]12[C@H]3[C@H]4[C@H]([C@@H]([C@@]1(C[C@@H]5[C@]2(O5)C(=O)O3)O)C(=O)O4)C(C)(C)O	GABAA antagonist	0.03
bicuculline (+)	CN1CCc2cc3c(cc2[C@H]1[C@H]4c5ccc6c(c5C(=O)O4)OCO6)OCO3	GABAA antagonist	0.23
etomidate	N1(C=NC=C1C(=O)OCC)C(C)c2ccccc2	GABAA agonist	0.81
Alphaxalone	CC(=O)[C@H]1CC[C@@H]2[C@@]1(CC(=O)[C@H]3[C@H]2CC[C@@H]4[C@@]3(CC[C@H](C4)O)C)C	GABAA agonist	0.58
thipoental	CCCC(C)C1(C(=O)NC(=S)NC1=O)CC	GABAA agonist	0.51
SR-95531	[H]/N=c/1\ccc(nn1CCCC(=O)O)c2ccc(cc2)OC.Br	GABAA antagonist	0.21
muscimol	NCc1cc(O)no1	GABAA agonist	0.21
GABA	NCCCC(=O)O	GABA agonist	0.21
tracazolate	CCCCNc1c2cnn(c2nc(c1C(=O)OCC)C)CC.Cl	GABAA agonist	0.8
clonazepam	c1ccc(c(c1)C2=NCC(=O)Nc3c2cc(cc3)[N+](=O)[O-])Cl	GABAA agonist	0.23
zolpidem	Cc1ccc(cc1)c2c(n3cc(ccc3n2)C)CC(=O)N(C)C	GABAA agonist	0.23
21-Hydroxyprogesterone	C[C@]12CC[C@H]3[C@H]([C@@H]1CC[C@@H]2C(=O)CO)CCC4=CC(=O)CC[C@]34C	GABAA agonist	0.79
SCS	C1=CC=C(C(=C1)/C=N/NC(=O)C2=CC=CC=C2O)O	GABAA antagonist	0.04
Isoflurane	C(C(F)(F)F)(OC(F)F)Cl	GABAA agonist	0.03
TACA	C(/C=C/C(=O)O)N	GABAA agonist	0.07
R-Baclofen	c1cc(ccc1C(CC(=O)O)CN)Cl	GABAB Agonist	0.26
Progesterone	CC(=O)[C@H]1CCC2C3CCC4=CC(=O)CC[C@]4(C)C3CC[C@]12C	GABAA agonist	0.68
Indiplon	CC(=O)N(C)C1=CC=CC(=C1)C2=CC=NC3=C(C=NN23)C(=O)C4=CC=CS4	GABA agonist	-0.06
Ocinaplon	C1=CC=NC(=C1)C(=O)C2=C3N=CC=C(N3N=C2)C4=CC=NC=C4	GABA agonist	-0.01
L-655,708	CCOC(=O)c1c2n(cn1)-c3ccc(cc3C(=O)N4[C@H]2CCC4)OC	GABAA antagonist	0.51
Gabapentin	C1CCC(CC1)(CC(=O)O)CN	increases GABA biosynthesis	0.15
Valproic Acid	CCCC(CCC)C(=O)O	GABA agonist	0.75
Predicted GABA			
5658603	CCCCn1c2ccccc2n2c(C(=O)OCC)c(C)nc12	predicted GABAA	0.74
5951201	CCOC(=O)c1c(C)nc(-c2ccccc2)nc1N=C(O)Cc1ccccc1	predicted GABAA	0.48
5142031	CCOC(=O)c1c(N)sc2ccccc12	predicted GABAA	0.54
7145248	c1ccc(cc1)Oc2cccc(c2)CNC(=O)C(=O)c3c[nH]c4c3cccc4	predicted GABAA	0.5
Serotonin-6 Compounds Tested			
BGC 20-761	CN(C)CCC1=C(NC2=C1C=C(C=C2)OC)C3=CC=CC=C3	5-HT6 antagonist	0.74
Idalopirdine	C1=CC(=CC(=C1)OCC(C(F)F)(F)F)CNCCC2=CNC3=C2C=CC(=C3)F	5-HT6 antagonist	0.55
R1485 DI HCL	c1ccc(c(c1)F)S(=O)(=O)N2CCOc3c2cccc3N4CCNCC4.Cl.Cl	5-HT6 antagonist	0.07
MS 245 OXALATE	CN(C)CCC1=CN(C2=C1C=C(C=C2)OC)S(=O)(=O)C3=CC=CC=C3.C(=O)(C(=O)O)O	5-HT6 antagonist	0.09
SB 399885 hydrochloride	COC1=C(C=C(C=C1)S(=O)(=O)NC2=C(C(=CC(=C2)Cl)Cl)OC)N3CCNCC3	5-HT6 antagonist	0.16
Ro 04-6790	CNc1cc(NS(=O)(=O)c2ccc(N)cc2)nc(NC)n1	5-HT6 antagonist	0.11
WAY 208466	CN(C)CCN1C=C(C2=C1N=CC=C2)S(=O)(=O)C3=CC=CC(=C3)F.Cl.Cl	5-HT6 agonist	0.21
ST 1936 OXALATE	Cc1c(c2cc(ccc2[nH]1)Cl)CCN(C)C.C(=O)(C(=O)O)O	5-HT6 agonist	0.09
EMD 386088 HYDROCHLORIDE	Cc1c(c2cc(ccc2[nH]1)Cl)C3=CCNCC3.Cl	5-HT6 agonist	0.15
EMDT oxalate	CCc1c(c2cc(ccc2[nH]1)OC)CCN(C)C.C(=O)(C(=O)O)O	5-HT6 agonist	-0.03

Predicted Serotoni-6	SMILES	Mechanism Notes	Pheno score
6028165	COc1ccc(S(=O)(=O)NC2CCCC2)cc1Br	predicted 5-HT6	0.66
6366118	COc1ccc(S(=O)(=O)N2CCC(C)CC2)cc1Br	predicted 5-HT6	0.66
6029941	COc1ccc(OC)c(NS(=O)(=O)c2ccc(OC)c(Br)c2)c1	predicted 5-HT6	0.61
6030057	CCc1ccccc1NS(=O)(=O)c1ccc(OC)c(Br)c1	predicted 5-HT6	0.36
6193422	O=S(=O)(Nc1cccc(Cl)c1)c1ccc2c(c1)OCCO2	predicted 5-HT6	0.46
6030006	CCOc1ccccc1NS(=O)(=O)c1ccc(OC)c(Br)c1	predicted 5-HT6	0.52
6013263	COc1ccc(S(=O)(=O)Nc2ccc(C)cc2)cc1Br	predicted 5-HT6	0.61
5352629	COc1ccccc1N1CCN(C(=O)c2ccc(C(=O)c3c(C)cc(C)cc3C)cc2)CC1	predicted 5-HT6	0.48

Other Serotonin Compounds Tested			
ALMOTRIPTAN MALATE	CN(C)CCc1c[nH]c2c1cc(cc2)CS(=O)(=O)N3CCCC3.C(C(C(=O)O)O)C(=O)O	5-H1B/1D agonist	0.18
GR 55562 DIHYDROCHLORIDE	CN(C)CCCc1cc(ccc1O)C(=O)Nc2ccc(cc2)c3ccncc3.Cl.Cl	5-HT1B silent antagonist	0.24
LISURIDE MALEATE	CCN(CC)C(=O)N[C@@H]1CN([C@@H]2Cc3c[nH]c4c3c(ccc4)C2=C1)C.C(=C\C(=O)O)\C(=O)O	5-HT2b antagonist	0.11
MESULERGINE HYDROCHLORIDE	Cn1cc2c3c1cccc3[C@H]4C[C@@H](CN([C@@H]4C2)C)NS(=O)(=O)N(C)C.Cl	5-HT2A/2C2C antagonist	0.4
Rizatriptan Benzoate	CN(C)CCc1c[nH]c2c1cc(cc2)Cn3cncn3.c1ccc(cc1)C(=O)O	5-HT1B/1C agonist	-0.01
Duloxetine Hydrochloride	CNCCC(c1cccs1)Oc2cccc3c2cccc3.Cl	SNRI	0.27

Other Serotonin Compounds Tested	SMILES	Mechanism Notes	Phen oscor e
TCB-2	COc1cc(c(c2c1C(C2)CN)OC)Br.Br	5-HT2a agonist	0.1
Fluoxetine Hydrochloride	CNCCC(c1ccccc1)Oc2ccc(cc2)C(F)(F)F.Cl	SSRI	0.46
Ondansetron Hydrochloride	N2(c1c(cccc1)C3=C2CCC(C3=O)CN4C(=NC=C4)C)C	5-HT3 antagonist	0.13
SB-216641	Cc1cc(ccc1c2ccc(cc2)C(=O)Nc3ccc(c(c3)OCCN(C)C)OC)c4nc(on4)C	5-HT1B antagonist	0.11
SDZ-205557	CCN(CC)CCOC(=O)c1cc(c(cc1OC)N)Cl	5-HT3/4 antagonist	0.07
sumatriptan	CNS(=O)(=O)Cc1ccc2c(c1)c(c[nH]2)CCN(C)C	5-HT1 agonist	0.07
BW 723C86	CC(N)Cc1c[nH]c2ccc(OCc3cccs3)cc12	5-HT2B agonist	0.1
DOI HYDROCHLORIDE	CC(Cc1cc(c(cc1OC)I)OC)N.Cl	5-HT2A/2C agonist	0.1
Cisapride	Fc1ccc(cc1)OCCCN2CC(C(CC2)NC(=O)c3c(cc(c(c3)Cl)N)OC)OC	5-HT4 agonist	0.08
Sertraline	ClC1=CC=C([C@H]2C3=C([C@H](CC2)NC)C=CC=C3)C=C1Cl	SSRI	0.54
Fluvoxamine maleate	FC(F)(F)c1ccc(cc1)/C(=N/OCCN)/CCCCOC	SSRI	0.23
Paroxetine Hydrochloride	Cl.Fc1ccc(cc1)[C@@H]1CCNC[C@H]1COc1ccc2OCOc2c1	SSRI	0.2
8-HYDROXY-DPAT	CCCN(CCC)C1CCc2cccc(O)c2C1	5-HT1A agonist	0.49
alpha-METHYLSEROTONIN	CC(Cc1c[nH]c2c1cc(cc2)O)N	5-HT agonist	0.18
Trazodone hydrochloride	Cl.Clc1cccc(c1)N1CCN(CCCn2nc3ccccn3c2=O)CC1	SSRI	0.22
1-(3-CHLOROPHENYL)BIGUANIDE	c1cc(cc(c1)Cl)NC(=N)NC(=N)N.Cl	5-HT3 agonist	0.18
Quipazine maleate salt	c1ccc2c(c1)ccc(n2)N3CCNCC3.C(=C\C(=O)O)\C(=O)O	5-HT agonist	0.16
Zimelidine dihydrochloride monohydrate	Brc1ccc(cc1)/C(=C/CN(C)C)/c2cnccc2	SSRI	0.16
Chlorpheniramine maleate	Clc1ccc(cc1)C(CCN(C)C)c2ncccc2	SNRI	0.05
Buspirone hydrochloride	N1(CCN(CC1)c2ncccn2)CCCCN4C(=O)CC3(CCCC3)CC4=O	5-HT1A agonist	0.08
FENFLURAMINE	CCNC(C)Cc1cccc(c1)C(F)(F)F.Cl	SSRI and 5-HT release stimulator	0.13
57-DIHYDROXYTRYPTAMINE	c1c(cc(c2c1c(c[nH]2)CCN)O)O	5-HT neurontoxn	0.05
serotonin	NCCc1c[nH]c2ccc(O)cc12	5-HT agonist	0.19
Desipramine hydrochloride	CNCCCN1c2ccccc2CCc3c1cccc3.Cl	SNRI	0.49
clomipramine	CN(C)CCCN1c2ccccc2CCc2ccc(Cl)cc12	SNRI	0.13

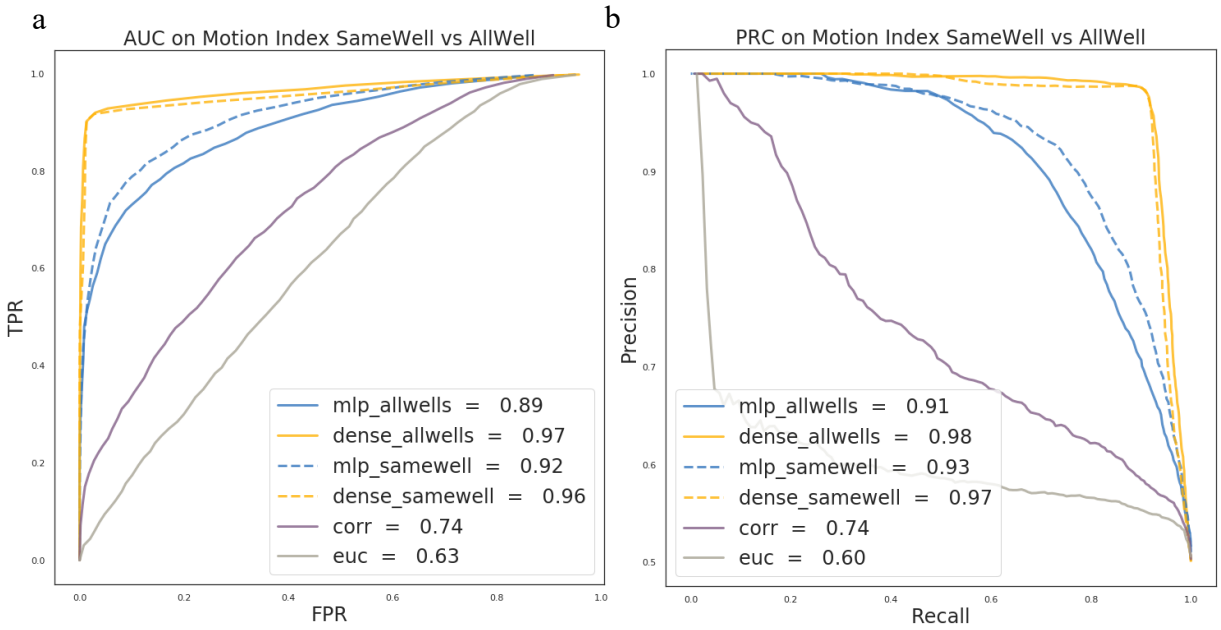
Supplementary Table A.1.10. Isoflavone analogs of hit 7013338

Cpnd	C2	C2'	C3'	C6'	C6	C7	phenocopy
7013338	CH_3	Cl	H	H	C_2H_5	CH_3	positive
JG-17	CH_3	H	CF_3O	Cl	C_3H_7	H	negative
JG-41	C_3H_7	H	CF_3O	H	C_3H_7	H	negative
JG-30	C_2H_5	H	H	Cl	C_8H_9	C_3H_5O	negative
JG-13	C_2H_5	H	CF_3O	H	C_3H_7	C_3H_5O	negative
JG-31	C_3H_7	H	H	Cl	C_8H_9	C_4H_7O	negative
JG-48	C_2H_5	H	CF_3O	H	C_8H_9	$C_3H_5O_2$	negative
JG-47	C_3H_7	H	H	Cl	C_8H_9	$C_3H_5O_2$	negative
JG-46	C_2H_5	H	H	Cl	C_8H_9	$C_3H_5O_2$	negative
JG-34	C_2H_5	H	CF_3O	H	C_8H_9	C_3H_5O	negative
JG-35	C_3H_7	H	CF_3O	H	C_8H_9	$C_3H_5O_2$	negative
JG-49	C_3H_7	H	CF_3O	H	C_8H_9	C_4H_7O	negative
JG-39	C_3H_7	H	H	Cl	C_8H_9	H	negative
JG-43	C_3H_7	H	CF_3O	H	C_8H_9	H	negative
JG-42	C_2H_5	H	CF_3O	H	C_8H_9	H	negative
JG-38	C_2H_5	H	H	Cl	C_8H_9	H	negative
JG-37	C_3H_7	H	H	Cl	C_3H_7	H	positive
JG-29	C_3H_7	H	H	Cl	C_3H_7	C_4H_7O	positive
JG-44	C_2H_5	H	H	Cl	C_3H_7	$C_3H_5O_2$	positive
JG-16	C_2H_5	H	H	Cl	C_3H_7	C_3H_5O	positive
JG-45	C_3H_7	H	H	Cl	C_3H_7	$C_3H_5O_2$	positive
JG-18	C_2H_5	H	H	Cl	C_3H_7	H	positive

Spectral Analysis of Isoflavone Analogs Omitted

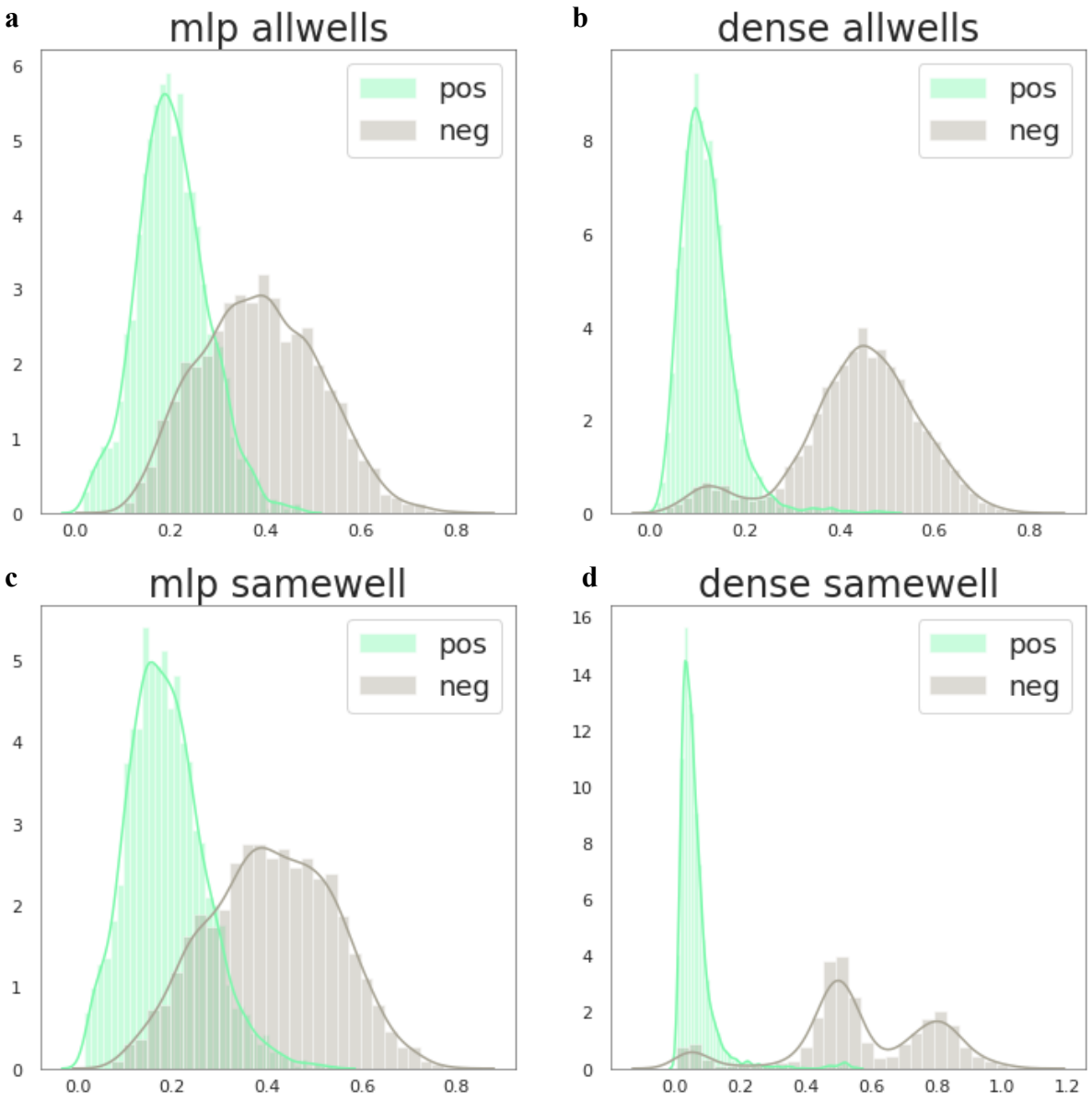
A.2 Supplementary material for Chapter 3

Supplemental Figures



Supplementary Fig A.2.1 Samewell vs Allwell AUC performance

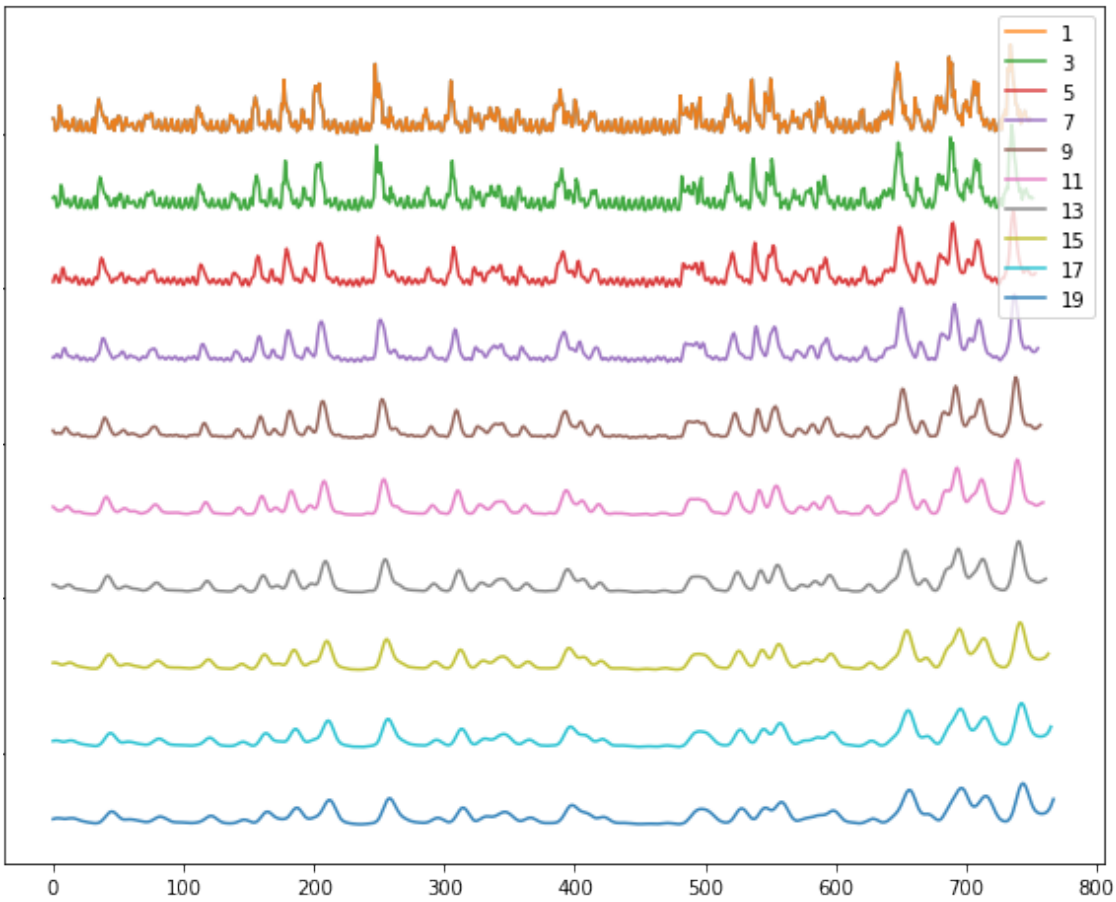
Same-well (dashed) vs *All-well* (dotted) negatives dataset performance comparison. a) Area under the curve plots. b) Precision recall curves.



Supplementary Fig A.2.2 Samewell vs Allwell positive vs negative separation

---

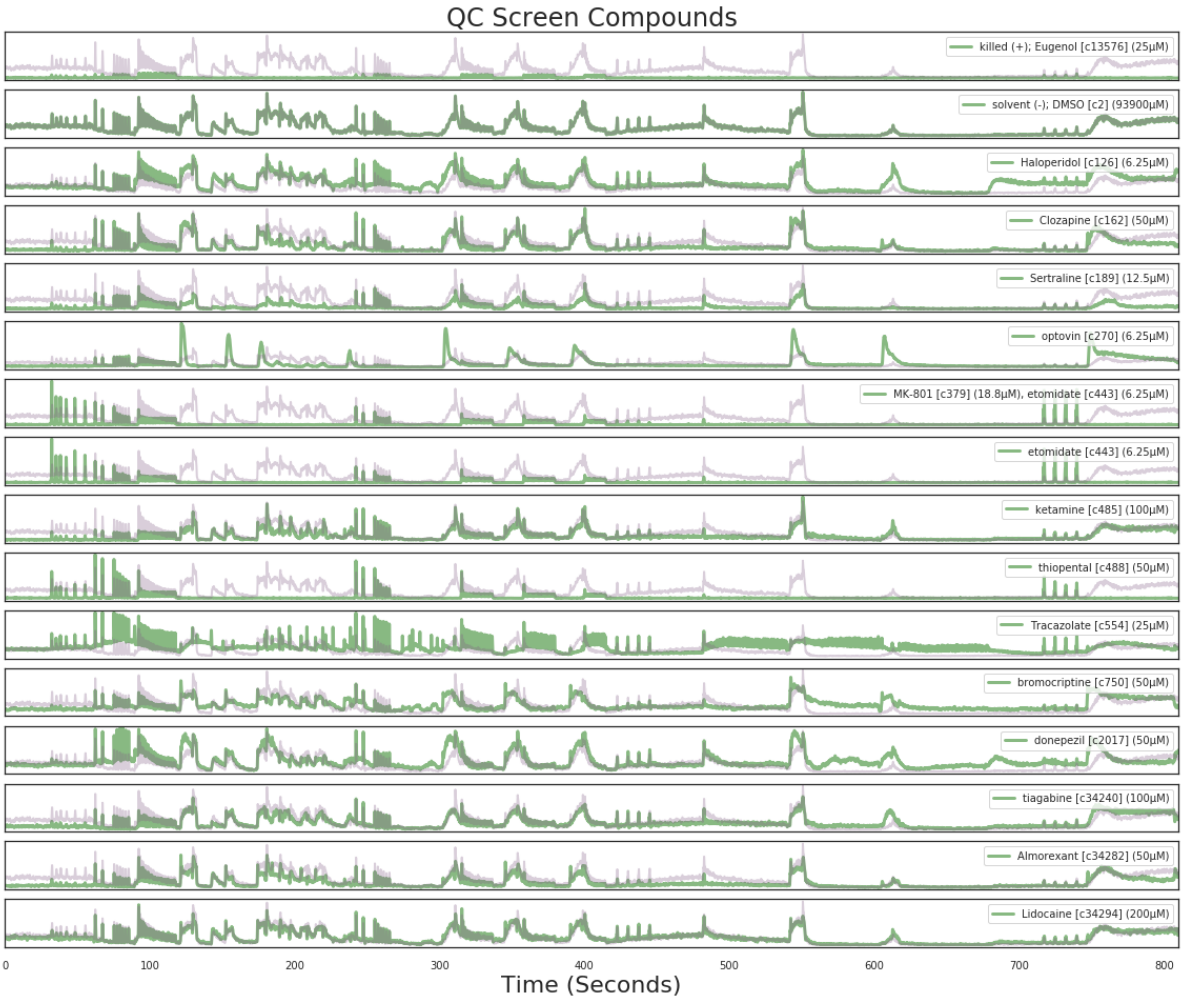
Same-well vs *All-well* negatives for positive vs negative pair separation. a) mlp all-well b) dense all-well c) mlp same-well d) dense same-well



Supplementary Fig A.2.3 Hanning-smoothing motion index time-series

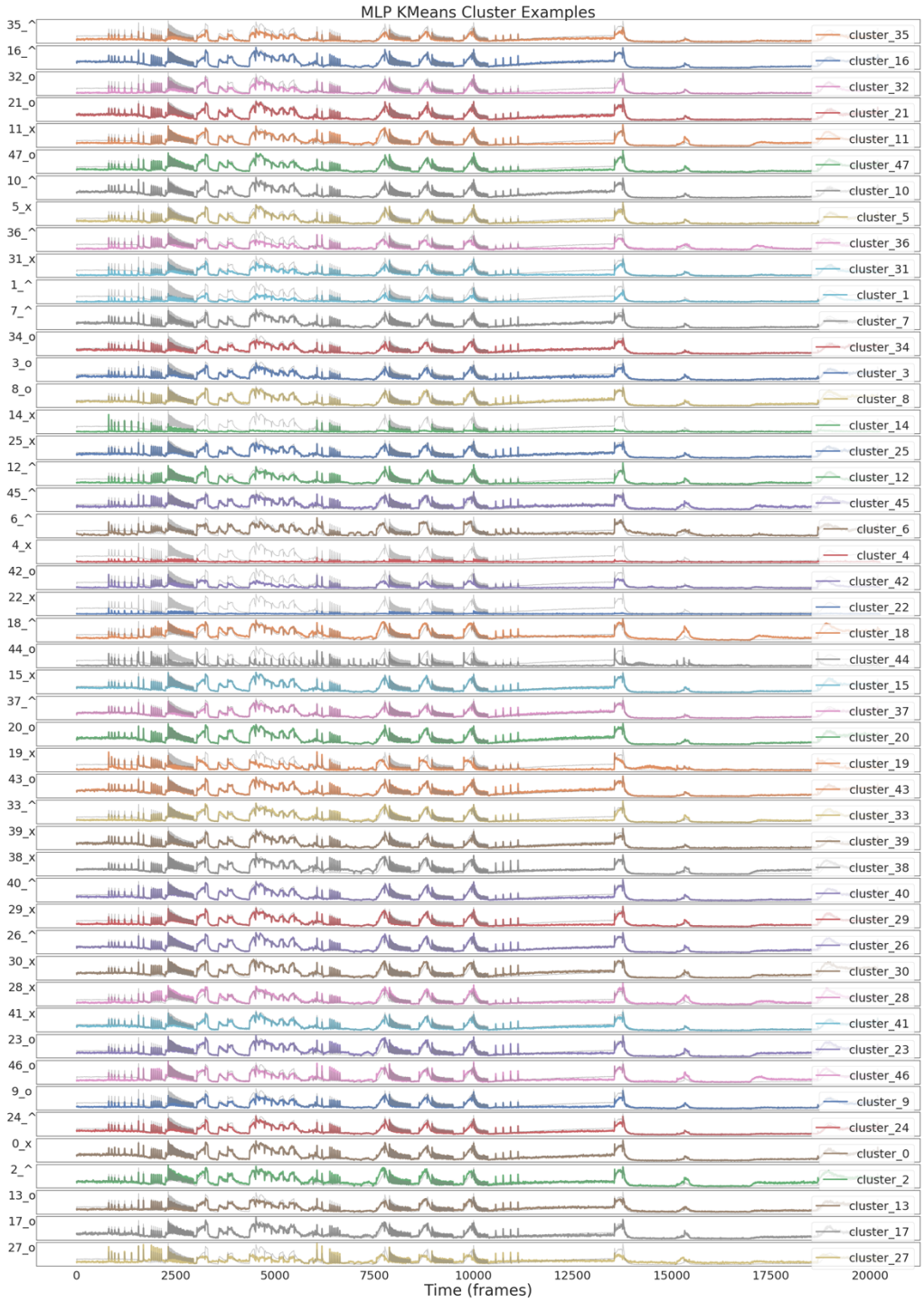
---

We examine how the choice of window-size for the hanning windowing function affects the resulting smoothed time-series. We decide to use the hanning 11 window because it seems to reduce most of the high-frequency components without sacrificing too much of the low-frequency, high amplitude motions.



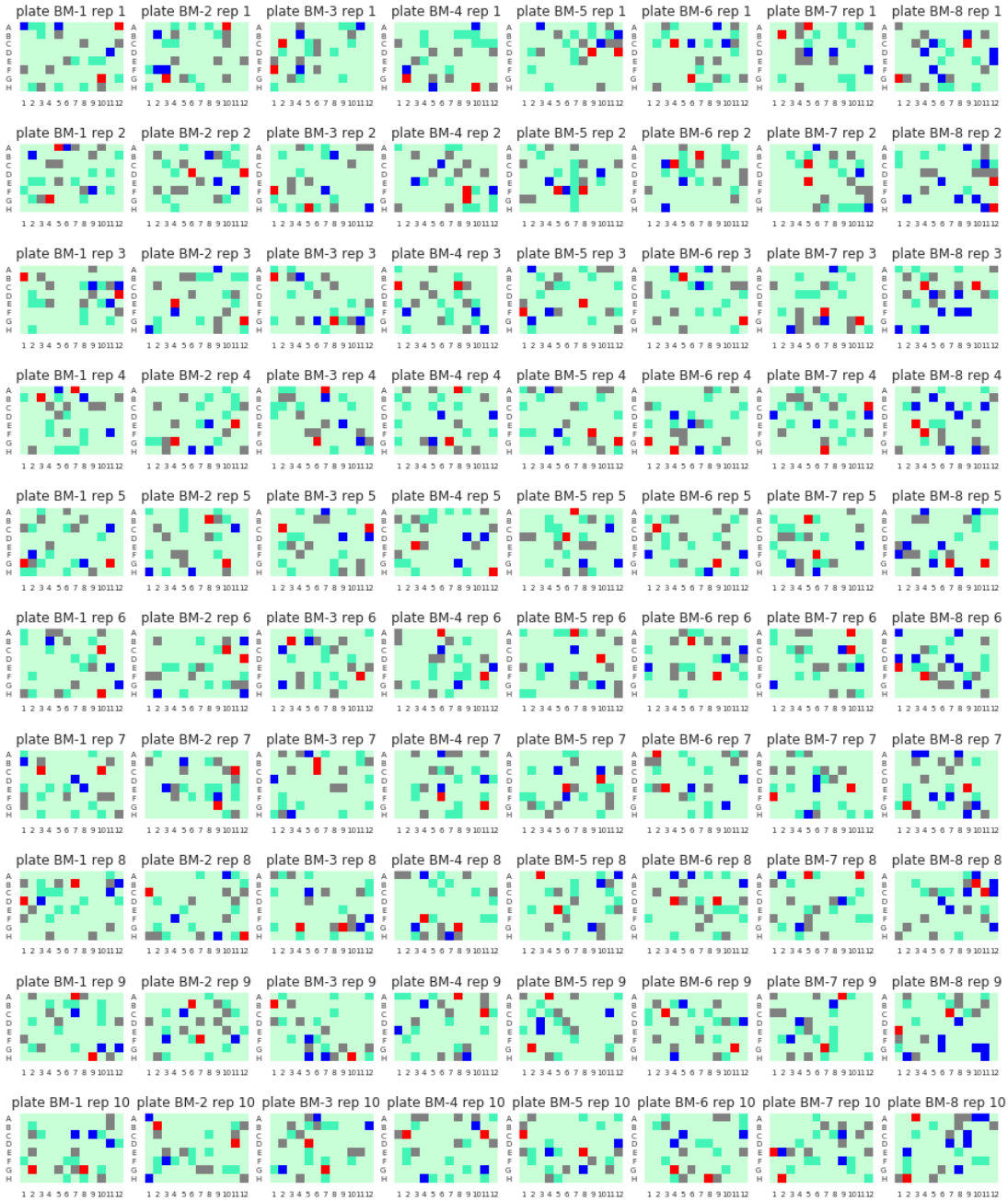
Supplementary Fig A.2.4 Quality control compounds

Mean motion-index time-series plots for each of the 16 quality control compounds. The mean DMSO control motion index is plotted in grey in all rows.



Supplementary Fig A.2.5 MLP Time-series Clusters

Mean motion-index time-series plots for the mean of each of the 48 Kmeans MLP smooth clusters. The mean DMSO control motion index is plotted in grey in all rows.



Supplementary Fig A.2.6 Computational randomized plate-layouts for new randomized screen

---

We computationally randomized the Biomol screen. Each row is based on one full library replicate (10 replicates total), and each column is a library plate (one of 8 possible plates). For each plate, we show the controls (Grey=DMSO (-), Red=Eugenol (+lethal), Blue = Water (-), light green = DMSO-dissolved drugs, darker green = Water-dissolved drugs). These layouts are saved into CSV files and read in by the randomization robot.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

9/5/19

Date