Concordant spatio-temporal patterns of brain activation in zebrafish exposed to compounds with similar pharmacodynamics or with similar seizurogenic potential.

Submitted by Joseph Pinion to the University of Exeter

As a thesis for the degree of

Masters by Research in Biological Sciences

In September 2018

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: …………………………………………………...

Abstract

Drug development is a highly resource intensive process that uses large numbers of animals for assessing the safety and efficacy of drugs prior to clinical testing. Improving the efficiency of drug development in terms of financial expenditure and number of animals used is therefore of utmost concern, not only to industry, but also to animal welfare organisations such as the NC3Rs. Poor efficiency in drug development largely stems from drug attrition, particularly attrition in the latter stages of the testing due to the large amount of resources expended at the point of failure. It is therefore imperative that deleterious off-target effects are identified as early as possible. However, typically, identification of seizure as a side-effect of drugs is performed in the later stages of development due to the highly intensive and low-throughput nature of seizure assays. At which point, if a compound fails, a large amount of resources have been squandered. There therefore exists a need for high-throughput and relatively inexpensive seizure liability assays that can be used early in drug development to prevent compounds destined for failure undergoing unnecessary resource intensive testing.

In this thesis we propose a refined approach using non-invasive imaging techniques in non-protected life stage zebrafish as a method for the detection of seizurogenic compounds early in drug development. In addition, we highlight its utility for elucidating the pharmacodynamics of compounds. In this study, a transgenic zebrafish line containing a GCaMP6s calcium sensor under the control of the pan-neuronal promoter *elavl3* was used for functional profiling of compounds with varied pharmacologies. Light sheet microscopy was used to record fluorescent activity in three spatial dimensions over time (4-dimensions) from the zebrafish brain after exposure to forty-three different compounds with varied pharmacodynamics and seizure liability profiles. Hierarchical clustering was employed in order to assess if compounds with seizurogenic activity or similar pharmacodynamics elicited specific functional brain activity. It was found that compounds with dopaminergic and serotonergic mechanisms of action elicited highly specific and similar brain activity patterns and that non-seizurogenic drugs also clustered separately from seizurogenic ones. Subsequent analyses, focussed on the utilisation of machine learning techniques, developing a model that could be used to discriminate between

compounds with and without potentially seizurogenic effects. It is clear, from the analyses presented here, that drugs do in fact elicit specific brain patterns in zebrafish and that these brain patterns are effectively detected using light sheet microscopy. This system is highly applicable for use within the drug industry and even in its relatively preliminary stages provided an accurate method of discriminating between compounds based on their physiological effects in zebrafish.

Chapter One: Introduction

The drug industry currently relies heavily on mammalian models for neurotropic drug profiling. Drug profiling is used to better understand drug effects as well as for assessing safety. The main approaches used for assessing the effects of compounds on the brain *in vivo* are pharmaco-EEG studies in mammals, which tend to have poor coverage of the brain and involve stressful and invasive procedures on rodents (1). In this thesis I outline the utility of zebrafish larvae as an alternative approach for profiling compounds for both assessing drug effects, in terms of pharmacodynamics, and also for safety testing. Zebrafish larvae have a number of advantages over mammalians models, including high reproductive rate and a transparent epidermis making them uniquely suited for high throughput imaging assays (2). Zebrafish are a quick and easy organism to cultivate and relatively easy to genetically engineer, compared with to their mammalian counterparts (3). Moreover, due to their small brain volume, zebrafish larvae are particularly suited to whole brain functional imaging making them an ideal model organism for studying functional connectivity, as the entire brain can imaged in a relatively short timeframe (~1.8 seconds using light sheet microscopy (4).

In terms of safety assessment, this thesis focuses on identifying compounds with potentially seizurogenic effects, a highly dangerous side effect of a number of neurotropic drugs. Currently seizure liability is performed late in drug development because seizure liability assays are highly resource intensive. As such, if compounds fail seizure liability testing, a large amount of resources are wasted, as said compound has already undergone a large battery of tests. It is therefore imperative that a relatively quick and inexpensive assay is developed so that seizurogenic compounds can be identified earlier on in drug development. The advantages conferred by using larval zebrafish make them an ideal candidates for just such a low resources and high throughput assay.

 It should be noted, however, that the techniques described could be applied in a variety of drug profiling scenarios and are not necessarily solely applicable to the seizurogenic potential of compounds.

In order to understand the potential contribution of zebrafish as a model organism, to the area of neurotropic drug screening, it is necessary to start by giving an appraisal of the use of mammalian systems for drug profiling and discovery.

Mammalian models for classifying the effects of pharmacological agents on human neurophysiology.

Traditional approaches to assessing the effects of drugs on human brain function are based, to a large extent, around performing pharmaco-electroencephalography (p-EEG) on mammals. It is clear from the numerous studies that have used this approach, that drugs can elicit specific and measurable changes in the electrical activity of the brain. As such, it is possible to use EEG data to elucidate the properties of drugs (5). In addition to the classification of drugs, p-EEG has been used to better understand the relationship between pharmacokinetics and pharmacodynamics. For example, by looking at the effect of a range of benzodiazepines on EEG power spectra it was observed that there is a strong correlation between $GABA_A$ receptor affinity (as measured *in vitro*) and changes in the amplitude of β frequency bands *in vivo* (6).

Multi-electrode EEGs are a commonly used tool and allow recording from multiple brain areas simultaneously (7). The information provided using these techniques can be used to assess the effect of drugs on brain activity extracted from different brain regions. For example, multi-electrode epicranial EEG recordings taken from freely behaving mice have demonstrated that cognition enhancing compounds, with varied pharmacologies, have distinctive effects on brain activity. When cognition enhancing compounds were clustered alongside a plethora of pharmacologically diverse compounds, cognition enhancing drugs were shown to cluster very tightly with one another relative to other compounds (see figure 1) (8).

Figure 1: Hierarchical clustering of compounds using EEG parameters taken from rats instrumented with epidural electrodes in different cortical areas (5).

In addition to multi-electrode EEGs, there are number of other useful tools available. Transgenic mouse models are highly useful for the discovery and development of therapeutic drugs (9). In fact, phenotypic drug discovery, such as the use of transgenic mouse models of disease, was responsible for development of 56% of first in class molecules (molecules with novel mechanisms of action) between 1999 and 2008 (10).

The combination of EEG and transgenic mouse models provides a powerful tool for assessing the therapeutic potential of drugs for brain disorders. For example, Levetiracetam, a drug traditionally used to treat epilepsy sufferers,

has been shown to have additional therapeutic potential after partially restoring cognitive deficits in a mouse model of Alzheimer's disease (11).

Due to the success of mammalian research in elucidating the mechanisms of action of compounds and relative lack of viable alternatives, the number of animals used in pharmaceutical research, is extremely high. Over two hundred thousand mammals were used, for a variety of purposes, in pharmaceutical research and development in 2013, eight and a half thousand of which were used for surveying drug pharmacodynamics and pharmacokinetics (12).

However, although the use of mammals in pharmacological research is very popular, there is a high financial cost. This is partially due to the high standards of care and regulatory requirements for housing and working with animals. Animal units must be temperature controlled, have air filtered ventilation, and be highly secure among a host of other specific requirements. In addition, mice take between forty and sixty days to reach sexual maturity with a gestation period of a further three weeks (13). These limitations make breeding mice for animal experiments a relatively time consuming and expensive process.

Moreover, in the case of EEG brain recordings, the experiments themselves are highly invasive and stressful procedures. Even epicranial EEG insertion requires surgery under general anaesthetic and runs the risk of infection. Moreover, due to tissue injury, related to the invasiveness of electrode implantation, inflammation is likely to occur (14). Inflammation is problematic because it has been implicated in epileptogenesis, thus furthering the suffering experienced by the animal while making them unsuitable for research (15). In addition, multi-electrode EEG recordings have relatively poor spatial resolution as the number of possible electrodes is limited.

All of these factors suggest there is a strong need for alternative models for assessing CNS effects of drugs. One such alterative gaining traction is the zebrafish (16).

Zebrafish EEG models for screening pharmacological compounds and the utility of zebrafish as tools for assessing neuropharmacological agents.

The zebrafish has a number of advantages that make it an attractive animal model. Zebrafish females take around three months to mature, spawn around a hundred eggs at a time and spawn every few days (17). Extremely high rate of reproduction and the fact that zebrafish embryos are not a protected species until four days post fertilisation makes them an ideal candidate for high-throughput pharmacological testing (18).

The utility of zebrafish for observing drug effects is already well established. Zebrafish have been used in high throughput behaviour based screens in attempts to identify CNS active compounds with novel structures, targets and functions (2,16,19–21). However, behavioural readouts, while a useful marker of brain activity, often diminish complex and varied brain function into basic metrics that do not fully represent the underlying brain mechanisms. Direct recordings of brain activity patterns have been utilised in order to better assess drug activity and side effects. Local field potential recordings have high temporal resolution and provide a good summation of fast brain electrical activity but only of a relatively small area of brain tissue (22). The relative ease with which genetic modification can be undertaken on zebrafish means they are ideal organisms on which to perform phenotype based screening. One such study, utilised a *snc1lab* mutation, to produce a transgenic zebrafish with photo-sensitive seizure-genesis. Using a seizure modelling algorithm, the ability of a wide variety of compounds to prevent or reduce seizure perpetuation could be realised. A measure of local field potential (LFP) complexity was found to negatively correlate with the presence of the *snc1lab* mutation. Interestingly the drug clemizole did not improve LFP complexity scores and was therefore not considered significantly efficacious in the LFP screening, despite being shown to reduce the number of spontaneous seizures (23). Conversely, clemizole has been identified as efficacious in a number of different screens using similar model organisms but with different approaches in terms of analysis and methodology (24). It is possible these discrepancies in results, with regard to clemizole, occur as a result of the poor coverage allowed of local field potential recordings; seizures may not be properly identified due to recordings only being taken from a very specific area.

Furthermore, zebrafish have been used in combination with multi-channel EEG recordings to study Pentylenetetrazole (PTZ) induced seizures. It was found

that seizures were largely propagated through the telencephalon. However as the electrode array only recorded from the midbrain and telencephalon, propagation through other regions cannot be excluded (25). This highlights the limitations of electrophysiological experiments, as even multi-electrode arrays have relatively poor coverage of all brain areas.

A similar study demonstrated the viability of machine learning approaches for automatic detection of seizures in zebrafish (26). The study in question utilised multiple models of seizures including pharmacologically induced seizures via PTZ and seizures in genetically modified models. Local field potential recordings and inducible seizures in combination with automated seizure detection could provide a powerful screening model of anti-epileptic drugs (26). Moreover, hyperthermia induced seizures have been shown to be dependent on TRPV4 channels as demonstrated by utilising TRPV4 antagonist RN-1734. In further analyses it was shown that seizures could also be attenuated by exposing fish to post-synaptic glutamate antagonists, thus supporting the theory that hypothermia induced seizures are partly mediated by NMDA receptor dependent glutamate transmission (27).

It's clear that EEGs can be used in conjunction with zebrafish to great effect and have already demonstrated their power in pharmacological research, particularly, in the field of epilepsy. However, there are limitations with regard to EEGs. In particular, EEGs and LFPs merely provide a summation of brain activity and even with the use of multi electrode arrays it is difficult to localise brain activity to specific areas. Being able to identify multiple specific areas of activity would allow for a better understanding of network connectivity and whole brain dynamics. In contrast with EEG, brain imaging techniques have been shown to be well adapted for these purposes and zebrafish, with their transparent epidermis, are perfectly suited for use in conjunction with imaging techniques (28).

Calcium imaging of neuronal activity.

The ability to record activity from multiple brain regions simultaneously and specifically is imperative in order to better understand the functioning of neuronal networks. Calcium imaging is one approach that can be used for imaging large populations of neurons and is a useful tool for assessing functional networks (29). Calcium transients are excellent markers of neuronal activity due to the strong relationship between intracellular Ca^{2+} ions in the presynaptic terminal and neurotransmitter release (see figure 2). The extent of action potentials controls the amount of Ca^{2+} that enters the presynaptic terminal which in turn effects the amount of neurotransmitter released. Moreover, action potentials usually produce a very brief rise in Ca^{2+} ions. This is due to a combination of Ca^{2+} channels being open a very short amount of time and Ca^{2+} ions being rapidly buffered by calcium binding proteins (30). These properties make Ca^{2+} an ideal marker of neural activity because it can represent neurotransmitter release at high spatial resolution.

Due to the effectiveness of using Ca^{2+} ions as a measure of brain activity a number of genetically
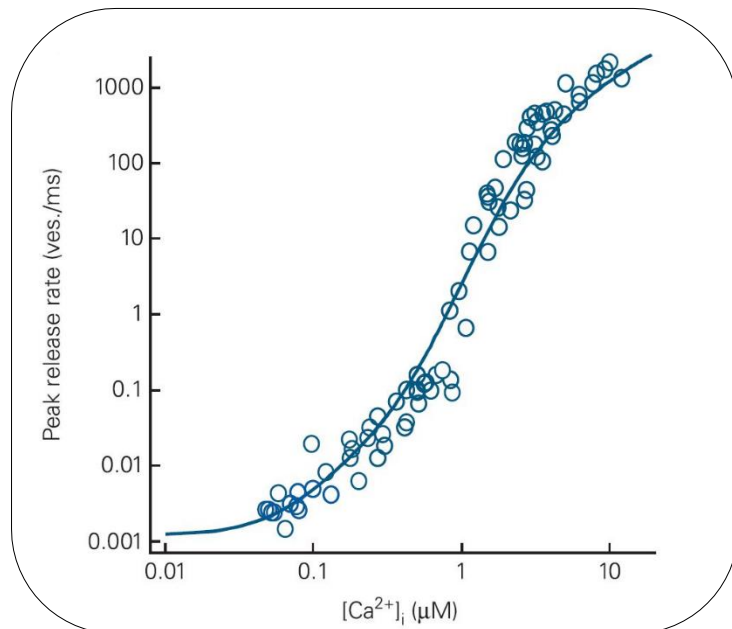
Figure 2: The relationship between intracellular Ca^{2+} and neurotransmitter release rate (27).

encoded calcium indicators have been developed; prominent among which is the GCaMP family of sensors. GCaMPs have been a popular approach for measuring calcium activity and have been continually updated and improved over the course of the last two decades (31–33). One of the more recent members of this family, GCaMP6, has been shown to reliably detect single action potentials in the mouse visual cortex when comparing fluorescent activity to simultaneous *in vivo* electrophysiology. Indeed, it was evident that GCaMP6 detected three-fold more neuronal responses than its predecessor, GCaMP5G, and could even detect calcium transients in dendritic spines. Moreover,

because GCaMPs are genetically encoded, it is possible to express them in specific cell subpopulations (34).

In zebrafish, GCaMPs can be genetically encoded under the pan-neuronal *elavl3* promotor, allowing constitutive expression throughout the brain. Larval zebrafish provide an ideal framework for calcium imaging of this type due to the relative ease with which genetic modification can be done and their largely transparent epidermis (16).
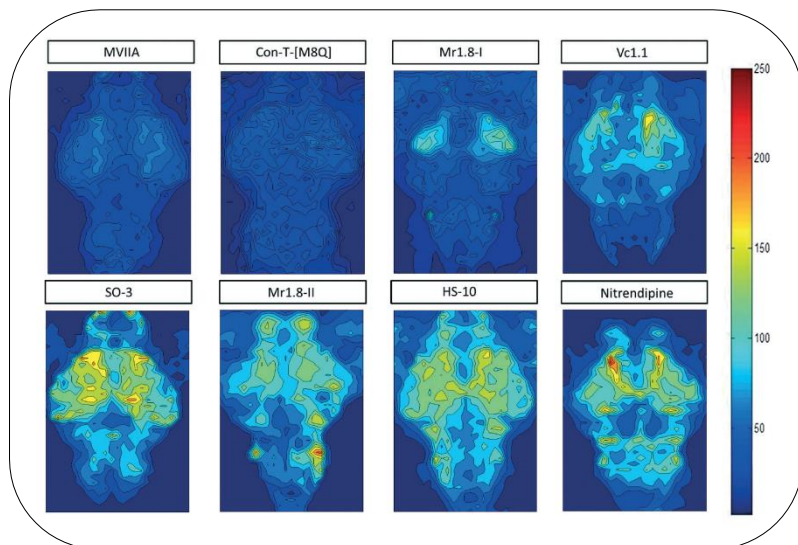
Figure 3: Neural activity density map of a representative elavl3:GCaMP5 zebrafish larva under treatment of different neurotoxin peptides (33).

The combination of advanced microscopy techniques, GCAMP calcium sensors and larval zebrafish provides a powerful approach for brain network analysis due to the exceptionally high spatial resolution available. Xudong Lin *at al* utilised GCaMP sensors to record brain wide activity in zebrafish larva at single cell resolution using confocal microscopy. Zebrafish were exposed to a variety of different neurotoxins and their brain wide activity was recorded. This experimental technique allowed for the detection of the distinct spatio-temporal patterns of activation induced by different neurotoxins, as illustrated in figure 3 (35).

A similar experiment setup, combining transgenic GCaMP6s zebrafish and wide-field fluorescence microscopy has been used for mapping of neuronal activity during pentyltetrazole (PTZ) induced seizures. This study demonstrated the concentration-dependent effect of PTZ on zebrafish brain dynamics in terms of signal amplitude and frequency. Moreover, a PTZ concentration-dependent increase in functional connectivity was also shown to be present (36). The use of a wide-field fluorescent microscope, in this study, meant that image data will only be of super-imposed fluorescent activity across the zebrafish brain. This is

because wide field fluorescent microscopes can only image in 3-dimensions (2 spatial dimensions over time). As such, region of interest selection can only be of the most rudimentary nature and thus cannot provide a comprehensive perspective whole brain activity.

Alternative approaches to widefield fluorescence microscopy include imaging techniques such as two photon (37), confocal (38) and light-sheet microscopy (LSM) (39). However, while these techniques allow three-dimensional imaging of tissue at very high spatial resolution, temporal resolution is sacrificed relative to widefield fluorescence microscopy. Two photon and confocal microscopy are particularly slow when performing volumetric imaging because both techniques are based on scanning and therefore only image a small area at a time. In addition, confocal microscopy is limited to a depth of around 50µm and two photon microscopy, due to its use of longer wave length light, comes with the risk of localised heating and the resultant damage (39). As such neither approach is ideal for imaging the whole zebrafish brain. Conversely LSM illuminates a single optical plane at a time, collecting fluorescence from all excited molecules within the field of view. LSM, therefore, can image 10 optical sections, enough to encompass an entire zebrafish larvae, within 1.9 seconds (4). Light sheet microscopy, therefore, provides respectable temporal resolution, while still offering excellent volumetric coverage. It is for this reason that it is the imaging approach of choice for this study.

Analytical approaches to 4-dimensional brain activity data and how it has been used in the past to model brain function.

The combination of LSM, GCaMP calcium sensors and zebrafish allows for imaging of four-dimensional brain activity (three spatial dimensions over time). Imaging over time, which is known as functional imaging, generates large data sets. Consequently, it is often necessary to condense some of this data, either to make it more understandable or less computationally intensive. A common way of doing this is by combining individual voxels from the 3d images into regions of interest (ROIs).

ROIs can be assigned to regions based on established cytoarchitecture or some other defined anatomical architecture. Atlases for cytoarchitecture exist for a number of species including the mouse (40) and macaque (41). Brain atlases have also been developed for use in zebrafish. One such atlas uses a variety of techniques to identify regions related to specific behaviours and stimuli in larval zebrafish between 5 and 7 days post fertilisation (DPF) (42).

Mapping ROIs to an atlas is advantageous as it allows inferences to be made about why specific regions behave in a certain way as ROIs are delineated based, to a certain degree, on the current understanding of their functional capacity. However, this approach is not without issues: for example it does not account for significant amounts of variability between individuals but rather assumes a certain degree of uniformity. This downside could be particularly relevant in the case of larval zebrafish because small differences in the number of DPF could have significant effects on brain anatomical architecture (43). Moreover, it is possible that larger anatomical regions could contain functionally discrete populations of neurons, meaning the true properties of the network are distorted due to mixing distinct signals of activity (44).

A viable alternative to defining ROIs according to predefined anatomical areas, is to randomly aggregate voxels into ROIs of equal size (45). This approach allows for a certain degree of choice in terms of node size and number. In the case of zebrafish this is useful because, with such high spatial resolution, the size of selected ROIs can be chosen based on the smallest possible functionally distinct region. In this case it is often necessary to repeat analyses with a different random parcellation for each, in order to ensure consistent findings.

In addition to random parcellation ROIs can be selected in a data driven manner using clustering algorithms. One such algorithm combined clustering of voxels based on their functional similarity with a spatial constraint function. This allowed the creation of similar sized ROIs, each containing functionally similar voxels. However, it has been demonstrated that while this approach was an improvement over selecting ROIs using an anatomical atlas it was not a significant improvement over random parcellation approaches (46).

Once ROIs have been selected there are plethora of analytical approaches that can be utilised to make sense of the generated data, the starting point for a number of analytical approaches is calculating the functional connectivity. Functional connectivity is an approach whereby the correlation co-efficient or another method of time dependent similarity is calculated between all pairs of ROIs in a brain network. This quantifies the level of similarity over time between all the different brain regions. How brain regions relate to one another is directly related to brain function. Functional connectivity has been used to identify differences in resting state fMRI between healthy controls and people with internet addiction tendency (47), schizophrenia (48) and borderline personality disorder (49) as well as being used a predictor of treatment response to antipsychotic medication (50).

Functional connectivity is therefore a very well-studied marker of brain function. How functional connectivity is optimally compared between experimental groups is dependent on the amount of available data, number of experimental groups and aims of individual researchers. A commonly used approach is network based statistics (NBS). NBS utilises permutation testing to statistically test the overall differences in the functional connectivity networks of two different groups. NBS is a useful tool as it provides a statistical significance cut-off for comparisons of experimental groups, however it cannot be used to make statistical inference about specific region-region functional connectivity values (49). Moreover, it is not useful for comparing large numbers of experimental groups to one another as it is both computationally intensive and would need to be corrected for multiple comparisons.

Instead of attempting to make elucidations about whole networks, many experimenters prefer to focus on specific region to region changes in functional connectivity. For example, linear regression analyses have been used to identify relationships between functional connectivity in specific cortical areas and the progression of Parkinson's disease (51). Making inferences about functional connectivity between specific brain regions is easier to put into the context of wider literature and opens up discussion about specific neuro-anatomical mechanisms that may be driving findings.

Functional connectivity analyses, however, do not provide a full picture of brain activity. Functional connectivity is an analysis method focused largely on

assessing the similarity between regions of interest and overlooks information pertaining to variability in localised temporal brain dynamics. Temporal information is manifest in individual time series. One such study, utilising GCaMP6s zebrafish used characteristics of temporal dynamics, in the form of parameters such as peak amplitude and frequency, to assess the effect of PTZ on brain activity (36). While in this particular case the imaging data were 3-dimensional (two spatial dimensions and time) this approach easily translates to 4-dimensions. Simple peak-based measures provide variables relating to temporal dynamics that are not measured explicitly in functional connectivity analyses. Peak parameters extracted from 4-dimensional brain activity data have already been utilised in the context of convulsant-treated larval zebrafish to highlight differences in temporal dynamics between pharmacologically distinct pro-convulsant compounds (4). Similar approaches have been used in the study described here, however, in this context they are paired with multivariate clustering approaches similar to those utilised during clustering of spectral indices extracted from EEG data in mice (8).

The overall aim of this thesis was to highlight the utility of 4-dimensional functional imaging of elavl3:GCaMP6s zebrafish for the assessment of drug function and the identification of unwanted side effects.

Objectives

To address the overall aim, this thesis addressed the following main objectives:

1. The first objective was to identify if compounds with specific pharmacologies or potentially seizurogenic properties induce common spatio-temporal patterns of activation in the zebrafish brain (chapter 2). This was achieved by exposing zebrafish, with genetically encoded calcium sensors, to a variety of compounds with varying pharmacologies and seizurogenic profiles. Four-dimensional activity data were extracted, from which parameters were calculated that reflected the spatio-temporal patterns of activation induced by the different compounds. Clustering was then performed on these parameters to identify if similar compounds with similar pharmacological or seizurogenic profiles cluster together based on these parameters.

2. Moreover, I wished to identify whether classification algorithms were able to discriminate between drugs, with and without seizurogenic properties, based on their spatio-temporal dynamics in the zebrafish brain (chapter 3). A variety of classification learning algorithms were generated and tested using the same parameters used for the clustering analyses.

3. In addition, I sought to train and test classifiers using hold out data (data not used during the generation of the algorithm), in order to obtain information about generalisability (chapter 3).

Chapter Two: Data acquisition and Clustering

Introduction

All analyses in this thesis were performed using 4-dimensional brain activity imaged from zebrafish larvae. As described in chapter one zebrafish larvae have a plethora of advantages over their mammalian counterparts, not least, being a more ethical and cost effective alternative. Also highlighted was the utility of a combination of light-sheet microscopy, GCaMP calcium sensors and transgenic zebrafish for providing a real-time model of whole brain activity. The work described in this thesis uses such an experimental model for the generation of 4-dimensional brain activity data.

This chapter of the thesis will discuss the methods used to generate the 4-dimensional brain activity data in addition to a number of analyses performed. These analyses are concerned with identifying whether compounds with specific pharmacological mechanisms or seizurogenic profiles induce specific patterns of activity in the brain. This question is of the utmost importance with regards to demonstrating the utility of zebrafish, as a viable alternative to mammals, in the area of drug safety screening and pharmacological profiling.

The ability to effectively determine if drugs are potentially seizurogenic is paramount to assessing the safety of neurotropic drugs (52). This can be a complex task, because drugs have varying degrees of seizurogenicity and there is a large variety of mechanism through which compounds can elicit a seizurogenic effect (53–55). As such, approaches used to assess such

compounds must be able to consider a spectra of drug effects. The main advantage of the experimental approach used in this thesis is the ability to record activity from the whole brain, meaning activity indicative of seizurogenesis occurring in specific brain regions will likely be detected. If known seizurogenic compounds elucidate a similar and specific pattern of activation in the zebrafish brain it validates this experimental setup as an effective screening approach.

In addition to identifying potentially seizurogenic compounds, I wished to see if there were patterns of activation indicative of specific pharmacologies. Being able to make inferences about the pharmacology of compounds based on zebrafish brain activity data would provide information about the applications of compounds before clinical trials are conducted. Moreover, it has the potential to provide insight into new applications of compounds currently used for different therapeutic purposes. Ketamine for example has come to the forefront in recent years as a highly effective antidepressant, despite not traditionally being used for such purposes (56). In addition, it is imperative that novel ways of assessing neurotropic drugs are developed, as current approaches have failed to produce effective therapies for large proportions of both epilepsy sufferers (57) and people with depression (58).

In this thesis, a number of parameters were measured that quantify the spatio-temporal properties of the 4-dimensional zebrafish brain activity. These parameters were analysed using univariate and multivariate approaches in order to assess whether common spatio-temporal patterns of activation exist between drugs with similar seizurogenic or pharmacological properties.

Methods

Methods Section 1: Data acquisition

The experimental approach and data acquisition, including the extraction of timeseries from the 4-dimensional image data and the extraction of the peak parameters was performed as previously (4). Below is a brief outline of the process, more detail is available in the original paper (See (4) appendix A).

Model organism

Transgenic *elavl3:GCaMP6s* zebrafish with an unpigmented epidermis. Fish were used four days post fertilisation (4dpf). These fish express GCaMP6s in the cytoplasm under control of the pan-neuronal *elavl3* promotor meaning that the Ca2+ sensor is expressed in virtually every neuron of the CNS (59). This model allows us to visualise neural activity, as changes in fluorescence intensity, across the whole CNS in response to treatment with neuroactive drugs.
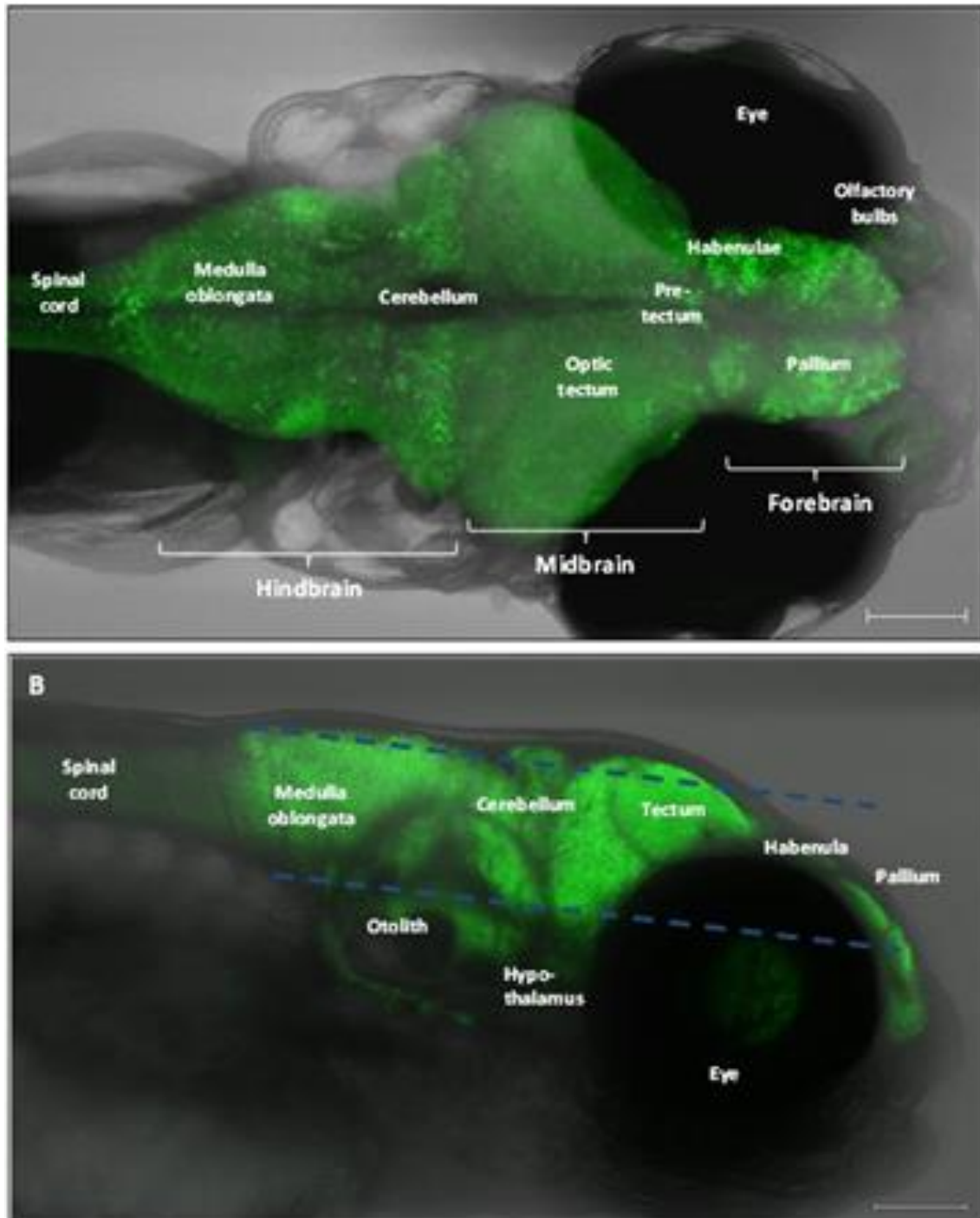

Figure 4: Confocal maximum intensity projection image of a 4dpf elavl3:GCaMP6s larva (10x magnification). Scale bar: 100µM.

<u>Test Compounds</u>

Forty-three compounds were selected for study which included multiple representatives of GABA$_A$ receptor antagonists, antidepressants, muscarinic agonists, acetylcholine esterase inhibitors, δ-opioid agonists and glutamate subtype agonists (details on drug names, uptakes and specific mechanisms of action are included in table 1). Appropriate exposure times and exposure concentrations were determined based on previous data, literature information and pilot work.

In addition, information on the bioavailability of the compounds tested (percentage uptake into the body) were based on previously conducted studies with zebrafish embryo-larvae (publication in progress).

<u>Compound Exposure and LSM</u>

Drug exposure and LSM was performed as described previously in (4). Briefly, larvae were exposed to each drug in a single well of a twenty-four well plate. Following exposure, larvae were subsequently exposed to the neuromuscular blocker tubocurarine (4mM) until cessation of movement. Each drug was tested at three concentrations (outlined in Table 1) alongside an unexposed control group. Eight larvae were used per concentration, although in some cases this number varied due to loss of animals, failure of data sets to correctly process experimental error with regard to the registration of ROIs to the 4-dimensional image data. As ROIs are based on an anatomical atlas and there is a degree of heterogeneity between larvae there is, consequently, a degree of ROI registration error. In addition, eight control experiments were run alongside each compound whereby fish were exposed to embryo culture water only. After exposure to the test compound, fish were transferred into 1.4% low melting point agarose containing the test compound and tubocurarine. Optical sectioning was executed in the horizontal plane from the dorsal to the ventral surface of each larva. Optical sectioning was repeated for six minutes generating two hundred cycles, each comprising ten z-stacks. Each z-stack took approximately 1.875 seconds to capture.

<u>Image processing and analysis</u>

Image processing was performed using a custom python image processing pipeline as previously described (4). Briefly, a labelled reference brain was aligned to each image and fifty regions of interest (ROIs) were selected in order to provide representatives of all the major brain regions (see table 2) (60). The mean fluorescent intensity across all the voxels within each ROI was calculated generating a time series of neural activity for each ROI (see figure 4). Peak analysis was then performed on each timeseries, producing nine peak parameters for each ROI within each animal: mean peak height, peak width, peak separation in frames, area under the curve and number of peaks. For this all the time series were filtered using Gaussian filtering (sigma = 1.5) to remove high frequency noise. The baseline of the smoothed signal was subtracted and a threshold of two times the standard deviation was used to identify peaks.
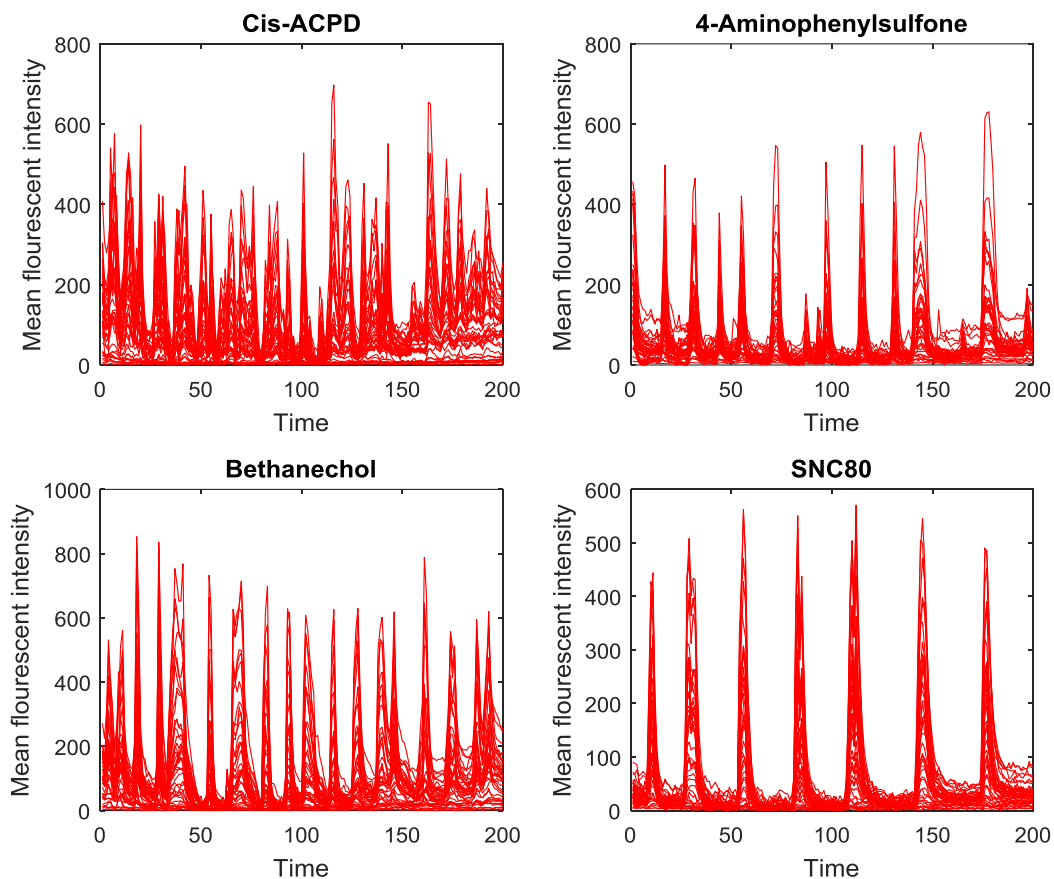
Figure 5: Example timeseries from larvae exposed to four pharmacologically distinct drugs. Graphs depict mean fluorescent intensity for each ROI (represented by distinct lines) on the y-axis and time in number of frames on the x-axis. Note the coordinated activity response profiles across multiple ROIs for each drug. Each graph is a single representative experiment. All the timeseries have their linear trend removed in order to provide a better perspective of the peak heights.

Compound	CAS No.	Mode of Action	Exposure conditions		% uptake across
			Concentrations	Duration	
Donepezil	884740-09-4	ACh esterase inhibitor (61)	125, 62.5, 31.25µM	30mins	349.65
Physostigmine	57-64-7	ACh esterase inhibitor (62)	500, 250, 125µM	20mins	43.21
Tacrine	1684-40-8	ACh esterase inhibitor (61)	1, 0.5, 0.25mM	20mins	16.44
Clozapine	5786-21-0	Atypical antipsych (polypharm) (63)	100, 50, 25µM	20mins	958.70
Olanzapine	132539-06-1	Atypical antipsych (polypharm) (64)	250, 125, 62.5µM	20mins	88.43
Bicuculline	40709-69-1	GABAA antagonist (65)	250, 125, 62.5µM	30mins	1.00
Pentetrazol (PTZ)	54-95-5	GABAA antagonist (66)	5, 2.5, 1.25mM	20mins	8.00
Picrotoxin	124-87-8	GABAA antagonist (67)	60, 30, 15µM	20mins	2.00
Bemegride	64-65-3	GABAA antagonist (68)	1, 0.5, 0.25mM	20mins	N/A
Gabazine	104104-50-9	GABAA antagonist (69)	1, 0.5, 0.25mM	20mins	1.03
Kainic acid	58002-62-3	Glutamate subtype agonist (Kainate) (70)	1, 0.5, 0.25mM	30mins	<LOQ
Cis-ACPD	477331-06-9	Glutamate subtype agonist (NMDA) (71)	2, 1, 0.5mM	30mins	N/A
NMDA	6384-92-5	Glutamate subtype agonist (NMDA) (72)	2, 1, 0.5mM	30mins	N/A
Strychnine	60-41-3	Glycine antagonist (73)	200, 100, 50µM	20mins	36.00
4-aminopyridine	504-24-5	Non-selective K+ blocker (74,75)	1, 0.5, 0.25mM	10mins	9.00
Tetraethylammonium Cl	56-34-8	Non-selective K+ blocker (76)	1, 0.5, 0.25mM	1 hour	N/A
Bupropion	31677-93-7	Mixed mechanism antidepressant (77)	60, 30, 15µM	30mins	63.00
Maprotiline	10347-81-6	Mixed mechanism antidepressant (78)	31.25, 15.6, 7.8µM	20mins	9.00
Amoxapine	14028-44-5	Mixed mechanism antidepressant (79)	62.5, 31.25, 15.6µM	20mins	34.00
Amitriptyline	549-18-8	Mixed mechanism antidepressant (80)	125, 62.5, 31.25µM	20mins	863.78
Pilocarpine	54-71-7	Muscarinic agonist (81)	10, 5, 2.5mM	30mins	7.64
Muscarine	2936-25-6	Muscarinic agonist (82)	1, 0.5, 0.25mM	1 hour	0.93
Bethanechol	590-63-6	Muscarinic agonist (83)	1, 0.5, 0.25mM	1 hour	0.50
Aminophylline	317-34-0	Phosphodiesterase and Adenosine (84)	2, 1, 0.5mM	30mins	2.00
Theophylline	58-55-9	Phosphodiesterase and Adenosine (85,86)	1, 0.5, 0.25mM	20mins	N/A
Caffeine	58-08-2	Phosphodiesterase and Adenosine (87)	1, 0.5, 0.25mM	20mins	19.65
SNC80	156727-74-1	δ-opioid agonist (88)	100, 50, 25µM	20mins	224.67
SB205607 dihydrobromide	1217628-73-3	δ-opioid agonist (89)	500, 250, 125µM	20mins	24.33
Cocaine	53-21-4	Dopamine transporter inhibitor (90)	250, 125, 62.5µM	20mins	150.55
Amphetamine	51-63-8	Dopamine, 5-HT and noradrenaline (91)	1, 0.5, 0.25mM	20mins	76.14
4-Aminophenylsulfone	80-08-0	Antibiotic (sulphonamide-like) (92)	1, 0.5, 0.25mM	20mins	7.78
Clonidine	4205-91-8	a2 agonist (in CNS) (93)	1, 0.5, 0.25mM	20mins	25.50
Quinine HCl	6119-47-7	Alkaloid antimalarial (94)	1, 0.5, 0.25mM	20mins	11.28
Yohimbine	65-19-0	Alkaloid a2 antagonist (95)	1, 0.5, 0.25mM	20mins	49.17
Mizolastine	108612-45-9	H1 antihistamine (96)	100, 50, 25µM	20mins	185.60
Emetine	316-42-7	Alkaloid (97)	500, 250, 125µM	1hour	33.77
Apomorphine	41372-20-7	Dopamine agonist, 5HT and a2 antagonist (95)	100, 75, 50, 25µM	1hour	77.77
Cisplatin	15663-27-1	DNA cross linking chemotherapeutic (98)	500, 250, 125µM	2hours	3.00
Morphine	52-26-6	µ-opioid receptor activator (99)	1, 0.5, 0.25mM	2hours	2.60
Rolipram	61413-54-5	Phosphodiesterase-4 inhibitor (100)	300, 150, 75µM	24hours	24.39
Ketamine	1867-66-9	Glutamate subtype antagonist (NMDA) (101)	62.5, 31.25, 15.6µM	20mins	214.23
Nomifensine	32795-47-4	Noradrenalin and dopamine uptake inhibitor	100, 50, 25µM	20mins	255.51
Ketoconazole	65277-42-1	CYP17c inhibitor (104)	100, 50, 25µM	20mins	567.05

Methods Section 2: Data analysis

Overview of hierarchical clustering

Hierarchical clustering was performed using parameters that quantified the effect of each compound on spatio-temporal dynamics. I extracted two distinct but complementary sets of parameters from the mean fluorescent intensity data that are indicative of the spatio-temporal patterns of activity. One set was derived from analysis of peaks of time series in each region of interest, and the other was derived from functional connectivity analysis. Subsequent analysis, including hierarchical clustering, was performed separately on each set of parameters. For each set of parameters, I formed a vector representation of each compound and the euclidean distances between each of the vectors was calculated and used for hierarchical clustering (105–108). The hierarchical clustering, for each of the two sets of parameters, produced a dendrogram, which represents each of the compounds similarity to one another in terms of either their peak parameters or functional connectivity.

Data used in subsequent advanced statistical analysis (input data)

The experimental unit was considered to be the data obtained, as described in 'methods section one', from one individual larva imaged for approximately 200 time-cycles, each cycle comprising 1 z-stack of 10 z-slices (to be referred to herein as a "frame"). Experiments below one hundred and ninety-eight frames in length were removed from analyses. This was done in order to ensure all timeseries were equally representative and to make sure parameters were not biased by the length of experiment.

As previously stated, the image processing pipeline generated mean fluorescent intensity data for 50 ROIs comprising all major brain regions. It was imperative that all experiments used for analyses contained the same number of ROIs, this

was necessary in order to ensure consistency. As a result eight ROIs were excluded from all analyses. ROIs included and excluded are outlined in table 2.

ROIs were excluded if they were missing more than fifteen times across all experiments for all compounds. Such ROIs were considered to be frequently missing. ROIs were marked as missing if the mean intensity signal was zero across all the frames recorded in a single experiment. The reason for the absence of such frames was a failure of the image processing pipeline to correctly register that region, typically as the region consistently lay at the extreme of the imaging range (i.e. at the very top or bottom of the image z-stack such as the pituitary).

After removing all ROIs deemed to be missing frequently, individual experiments were removed from analyses if they had non-frequently missing ROIs. That is to say, experiments were removed from analyses if they contained a ROI that was missing in less than 14 other experiments.

As stated above ROIs were excluded if missing more than fifteen times. The number fifteen was chosen in order to strike a balance between the number of ROIs excluded from analyses and the number of experiments excluded from analyses. This is because if a large number of ROIs are excluded then coverage of the zebrafish brain is lost and if a large number of experiments are lost n-numbers for compounds go down.

There were forty-four different compounds including the control group. The number of experiments for each compound, not including control, varied between six and twenty-seven, with a mean of twenty-two.

The total number of control experiments was three hundred and thirty-three. The large amount of variation in number of experiments per compound was due to some compounds having poor registration for ROIs that were not frequently missing across all experiments. This meant a large number of experiments from these compounds had to be excluded from analyses.

In addition to the mean intensity data and peak analysis parameters provided by the image processing pipeline, for each experiment a functional connectivity matrix (109) was generated by calculating the Pearson correlation coefficient between the time series of all possible pairs of ROIs. Several alternative measures were considered, including cross correlation and coherence, the merits

and disadvantages of which are discussed in later chapters. Pearson correlation coefficient was chosen because it is simple to understand and implement while considering the amplitude and phase of any underlying oscillatory activity.

Connectivity matrices are designed to give an indication of the level of functional similarity present between different ROIs and thus provide information on how functionally connected each ROI is to every other ROI. A schematic providing an overview of initial data input generation process is provided in Figure 5.

The correlation coefficient between two time series was calculated as (eqn 1)

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}, (eqn\ 1)$$

Here, n is the number of time points in each time series, x and y the time series vectors of fluorescent intensity over time, $\bar{x}$ and $\bar{y}$ are the mean of the corresponding timeseries and $r_{x,y}$ is the Pearson's correlation coefficient between x and y:

To generate a functional connectivity matrix, each ROI was compared to every other ROI, using the correlation coefficient (eqn 1). Each row and column of the resulting matrix represents a different ROI and the matrix element positioned at the intersection of each row and column is the correlation coefficient between the two corresponding ROIs (see Figure 5).

Here, X is a $m - by - o$ matrix of time series such that m is the number of time points and o is the number of ROIs. Here, R is a connectivity matrix of X such that R is given by:

$$R = \begin{bmatrix} r_{x1,x1} & \cdots & r_{xo,x1} \\ \vdots & \ddots & \vdots \\ r_{x1,xo} & \cdots & r_{xi,xo} \end{bmatrix}, (eqn\ 2)$$

The connectivity matrix is symmetric along the diagonal, with all the values along the diagonal being equal to one. As a result of this, only the values in the upper triangle of the connectivity matrix R, excluding the diagonal, were selected for further analyses. This was done to remove repetitions of correlation coefficients and comparisons of ROIs to themselves.

Region of interest used in analyses	ROIs excluded from analyses
Diencephalon	Pituitary
Dorsal thalamus	Torus lateralis
Eminentia thalami	
Habenulae	
Pineal	
Posterior tuberculum	
Preoptic area	
Pretectum	
Ventral thalamus	
Mesencephalon	
Tectum stratum periventriculare	Caudal Hypothalamus
Tectum neurophil	Intermediate Hypothalamus
Tegmentum	Rostral Hypothalamus
Torus longitudinaris	
Torus semicircularis	
Rhombencephalon	
Area postrema	
Cerebellum	
Corpus cerebelli	
Eminentia granularis	
Inferior olive	
Interpendnuclar nucleus	
Lateral reticular nucleus	
Lobus caudalis cerebelli	
Locus coeruleus	
Mauthner	
Medial Vestibular nucleus	
Noradrendergic neurons	
Raphe Inferior	
Raphe Superior	
Tangential vestibular nucleus	
Valvula cerebelli	
Telencephalon	
Anterior commissure	Postoptic commissure
Olfactory bulb	Optic commissure
Pallium	
Subpallium	
Eyes	
Olfactory epithelium	
Vagal ganglia	
Spinal Cord	
Spinal cord neuropil region	

Table 2: Colour coding represents anatomical categorisation as subdivisions of the diencephalon, mesencephalon, rhombencephalon and telencephalon, along with ganglia respectively from top to bottom.
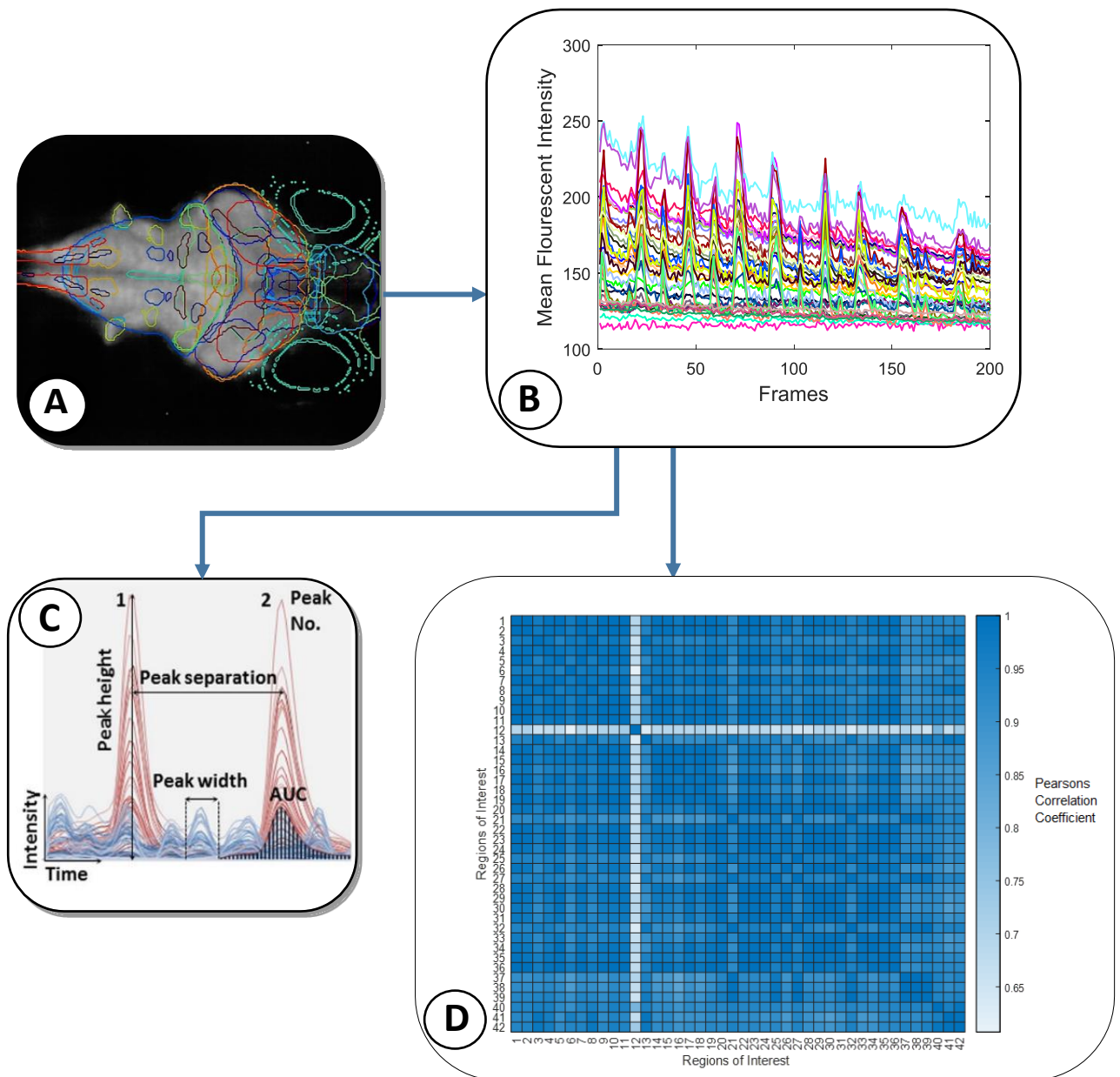
Figure 5: A schematic showing the approach used for the advanced analysis applied to the experimental datasets A: ROIs outlined in colours over a maximum intensity projection taken from a single 4dpf larval zebrafish brain. B: A single 'experiments' time series depicting the mean fluorescent intensity of the different ROIs over time taken from a single experiment. C: image showing the definition of each of the peak parameters extracted from each time series. D: Heat map of connectivity matrix comparing forty-two ROIs. Each element represents a correlation coefficient with dark blue elements representing strong correlation and light blue representing negative correlation.
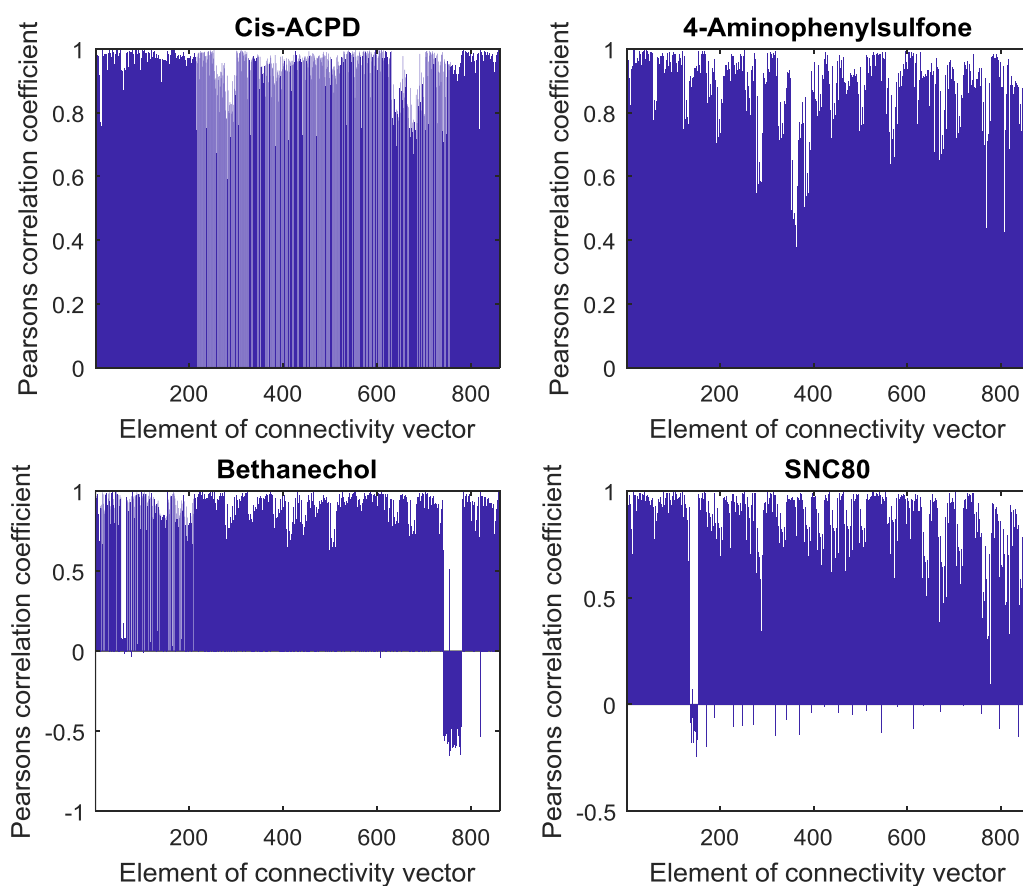
Figure 6: Bar graphs depicting the functional connectivity vectors of the time series in figure 4. The x-axis is each element of the functional connectivity vector and the y-axis is the Pearsons correlation coefficient value (measure of functional connectivity) for that element. Each element of the connectivity vector represents a comparison between two different ROIs

The selected values of the connectivity matrix were then reshaped into a vector r, which consisted of eight hundred and sixty-one unique values henceforth referred to as the connectivity vector. A connectivity vector was generated for all the experiments (see figure 6). This was done in order to generate parameters reflective of how the brain dynamics of regions change in relation to one another in response to certain compounds. The correlations were vectorised so that experiments could be indexed by row upon data matrix generation, as subsequently described.

Next a data matrix was generated in which each row is a connectivity vector from a single experiment. This data matrix D is an m-by-t matrix where m is equal to the total number of experiments across all conditions and t is equal to the number of connectivity parameters. Data were organised as such, so that principal component analysis could be performed across all experiments and over all the parameters.

In addition, and separately to the connectivity analyses, several peak parameters were calculated for each of the ROIs and for each experiment. These parameters consisted of the peak height, peak width, peak separation, number of peaks and the area under the curve. These were calculated as described previously (4). The mean and standard deviations of each peak parameter were calculated for each ROI for the time span of the experiment. There were therefore nine peak parameters, including peak number, calculated for forty-two ROIs totalling three hundred and seventy-eight different values for each experiment. A data matrix was generated and organised in the same way as for the connectivity analyses. Thus, a second $m - by - p$ data matrix D was generated where m was equal to the number of experiments and p was equal to the number of peak parameters for each experiment (p=333).

Next, each of the three hundred and thirty-three peak parameters were normalised such that the maximum for each parameter across all experiments was equal to 1, and the minimum equal to -1. If * and / represent pairwise multiplication and pairwise division of matrices respectively then the normalised peak parameters are calculated using equation 3, where $D_{i,max}$ denotes the largest value in the ith column of data matrix D, $D_{i,min}$ denotes the smallest value in the ith column of data matrix D and D_i denotes all the elements of the ith columns of data matrix D:

$$\left(D_i - D_{i,min}\right) * \left(\frac{2}{D_{i,max} - D_{i,min}}\right), (eqn\ 3)$$

Principal Component Analysis

The data matrices of the peak parameters and connectivity parameters were analysed separately but identically, thus in all future descriptions analyses were undertaken on both sets of parameters independently of one another. Principal component analysis (PCA) was used to reduce the dimensionality of each of the two data sets while retaining the majority of the variability present across compounds. The principal components were computed as follows:

First, the covariance was calculated between all possible pairs of parameters using equation 4, where: x is a vector of experiments of a single parameter (e.g.

peak number for the diencephalon) and y a vector of experiments of a different individual parameter (e.g. mean peak height for the diencephalon). Here, m is equal to the total number of experiments. The covariance between x and y, $c_{x,y}$ was calculated as

$$c_{x,y} = \frac{\sum_{i-1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{m-1}, (eqn\ 4)$$

Where p is the total number of parameters. The covariance was calculated between all the different pairs of parameters across all experiments to produce a $p - by - p$ covariance matrix C. Each row and column of the resulting matrix represented a different parameter and the matrix element positioned at the intersection of each row and column represented the covariance between the two corresponding parameters. C was therefore a square matrix that was symmetrical down the diagonal such that the diagonal was the variance of each of parameter:

$$C = \begin{bmatrix} c_{1,1} & \cdots & c_{1,p} \\ \vdots & \ddots & \vdots \\ c_{p,1} & \cdots & c_{p,p} \end{bmatrix}, (eqn\ 5)$$

Subsequently, the n eigenvalues and n eigenvectors of the covariance matrix were calculated. Mathematically the eigenvalues and eigenvectors satisfy:

$$Cv_i = \lambda_i v_i \ (eqn\ 6)$$

The eigenvalues and corresponding eigenvectors were then sorted by descending size.

The eigenvectors of the covariance matrix C represent "directions" in the coordinate space in which the variability across experiments was organised. The amount of variability accounted for by each eigenvector decreased with the size of their corresponding eigenvalue. Therefore, to reduce the size of the data, but retain variability, the eigenvectors that contained 99% of the variability present in the data, were selected.

This amount of variability accounted for by each eigenvector was directly proportional to the size of their corresponding eigenvalue. As such, the percentage of variability accounted for by each eigenvector was calculated by dividing the eigenvectors corresponding eigenvalue by the sum of all the

eigenvalues and multiplying the answer by one hundred. The number of eigenvectors that accounted for 99% of the variability was 127 and 211 for the connectivity and peak parameters respectively.

The original data for each set of parameters was then projected into this principal component space by multiplying the vector of values of each experiment by the eigenvectors with either the 127 or 211 largest eigenvalues (110). The data transformed into the principal component space is hereafter referred to as the principal component scores.

Outlier detection

Outlier detection was performed on all the experiments representing an individual compound. Outlier detection was performed by first calculating the euclidean distances between the vectors of principal component scores of the experiments for each individual compound as follows. The euclidean distance between two experiments for which p and q are their corresponding vectors of principal component scores, denoted as $d_{p,q}$, is given by:

$$d_{p,q} = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \quad (eqn\ 7)$$

Here, n is the number of principal component scores. Distances were then calculated between all the experiments of each compound. The resulting matrix of distances between all the experiments of that compound, denoted D, is given by:

$$D = \begin{bmatrix} d_{1,1} & \cdots & d_{1,e} \\ \vdots & \ddots & \vdots \\ d_{e,1} & \cdots & d_{e,e} \end{bmatrix}, (eqn\ 8)$$

Here e is the number of experiments of a compound. The mean of each of the columns of the $e-by-e$ distance matrix D was then calculated to find the mean distance each experiment was from every other experiment of the same

compound, hereafter referred to as the mean within compound distance. This was done to identify experiments that are significantly different from other experiments of the same compound.

The mean within compound distances were then put into size order and the interquartile range (IQR) and upper quartile (UQ) were calculated. Experiments were excluded from further analyses if their mean within compound distance was more than the UQ plus one and half times the IQR of the within compound distances (equation 14). Thus, the experiments whose corresponding within class distance satisfies eqn 13 were removed from analyses.

$$d_i > (IQR \times 1.5 + UQ), (eqn\ 9)$$

Hierarchical Clustering

To simplify the analysis, I sought to utilise a single representative vector for each compound. In order to achieve this the mean principal component scores for all the experiments (after exclusion of outliers) representing each compound were then calculated, producing a vector mean of principal component scores for each compound. Each element of the resulting vectors was the mean of a set of principal components scores across all the experiments representing a compound. Subsequently the euclidean distance was calculated between the representative vectors of all of the possible pairs of compounds. The euclidean distance was used to assess the level of similarity between pairs of compounds. These similarities can be visualised by a hierarchical clustering dendrogram.

A hierarchical clustering algorithm was employed to cluster compounds into a dendrogram based on their euclidean distances. Hierarchical clustering was performed using a simple agglomerative method known as 'unweighted pair group method with arithmetic mean' (UPGMA) (111). UPGMA works by first clustering the two compounds that are the shortest distance apart. Once the first cluster is generated it recalculates the distances to find the average distance between the newly formed cluster and the remaining compounds. Let a, b and c be vectors representing individual compounds and $d_{a,b}$ represent the euclidean distance between a and b. Assuming a and b are clustered together to form the cluster ab, the distance between the cluster ab and the compound c is given by:

$$\frac{d_{a,c} + d_{b,c}}{2}, (eqn\ 10)$$

Once the distances between the newly formed cluster and the remaining compounds have been generated, the smallest distance is again selected. If that distance is between the newly formed cluster and another compound, that compound is added to the cluster. If this is not the case a new and separate cluster is formed between the two compounds that are the shortest distance apart. The algorithm repeats itself, finding the smallest distance, and connecting the relevant compound or cluster into a new higher order cluster. The distances between two clusters are calculated by simply finding the mean of the pairwise distances between all the compounds within each cluster.

The UPGMA algorithm was applied to the principal components scores of each the compounds to group compounds together based on how similar their functional connectivity or peak parameters were. The UPGMA algorithm was chosen because all the distances between compounds contribute equally to the separation of each cluster.

Mean functional connectivity

In addition to the hierarchical clustering algorithm, the mean functional connectivity was calculated across all ROIs for all the connectivity values of each compound. This was achieved by selecting all the connectivity vectors representing each experiment of a compound, excluding outliers. An average functional connectivity vector was then calculated as the element-wise mean across all the experiments of a compound. Subsequently the mean of all the elements in the resulting vectors was calculated to generate a single mean connectivity , an approach that has been applied to evaluating the progressions of Parkinson's disease (112). This was repeated for all the different compounds (see table 3).

Mean peak parameters

In addition to the hierarchical clustering algorithm, the mean of each of the peak parameters was calculated across all ROIs for all the input data of each compound. This was achieved by selecting all the peak parameters representing each experiment of a compound, excluding outliers. The mean peak parameters were then calculated by calculating the mean across all the experiments.

Subsequently the mean of all the peak parameters across all the ROIs was calculated. This was repeated for all compounds (see table 6).

Functional connectivity between areas connected by established neurochemical pathways

Established neurochemical pathways connecting ROIs were identified from the literature. Pathways between the relevant ROIs were listed alongside the neurotransmitter present in each pathway (see table 4). Subsequently, the elements of the connectivity matrices that represented the connectivity between the two ROIs connected by an identified neurochemical pathway were extracted from each experiment. These values were then separated into their corresponding compounds.

For each compound the connectivity values representing each identified neurochemical pathway were compared with the corresponding connectivity values for water using a Kruskal and Wallis one-way analysis of variance test. Compounds with a significantly lower or significantly higher ($p < 0.05$) mean connectivity values are included in table 4.

Results

Mean functional connectivity

The mean functional connectivity across all ROIs was generated in order to assess if generalised functional connectivity information provides insight into pharmacology or seizurogenic potential. This was done in an effort to identify if spatial information is an important indicator of either pharmacology or seizurogenesis. When the mean functional connectivity scores were generated, and put in order of magnitude, non-seizurogenic compounds were randomly distributed throughout table 4 and there were no obvious similarities relating to their pharmacology.

As such it seems that non-spatially specific functional connectivity information does not provide much insight into pharmacology or seizurogenesis. It therefore seems that spatially specific information is a necessary component of the analysis, with regards to functional connectivity.

With this in mind, specific functional connectivity scores between ROIs with known neurochemical connections, were assessed for viability as biomarkers of drug action.

Compound	Mean Functional connectivity (≥ Control)	Compound	Mean Functional connectivity (< Control)
Amitriptyline	0.891	Bupropion	0.744
Muscarine	0.827	Nomifensine	0.743
NMDA	0.812	Olanzapine	0.741
Bethanechol	0.805	Clozapine	0.740
Amphetamine	0.802	Morphine	0.732
Tetraethylammonium	0.799	Cocaine	0.731
Yohimbine	0.794	Bemegride	0.731
CisACPD	0.792	SB205607	0.729
Donepezil	0.792	Apomorphine	0.728
Cisplatin	0.786	Clonidine	0.725
Ketamine	0.769	Caffeine	0.720
Physostigmine	0.761	Pilocarpine	0.720
Ketoconozole	0.759	Rolipram	0.716
Kainic	0.757	Quinine	0.711
Tacrine	0.756	Aminophylline	0.711
4-Aminopyridine	0.752	Amoxipine	0.711
Emetine	0.751	Maprotiline	0.708
Water - (Control)	0.744	SNC80	0.708
		Theophylline	0.697
		4-Aminophenylsulfone	0.692
		Strychnine	0.684
		Mizolastine	0.682
		Pentylenetetrazole	0.677
		Picrotoxin	0.649
		Bicuculine	0.646
		Gabazine	0.644

Table3: Mean functional connectivity data across all ROI pairs. Left two columns contain compounds whose mean functional connectivity is equal to or greater than the mean functional connectivity of water. Right two columns are the compounds whose mean functional connectivity is less than water.

Functional connectivity between areas connected by established neurochemical pathways

This approach identified neurochemical pathways containing specific neurotransmitters from the literature. Subsequently the functional connectivity between the two ROIs connected by the identified pathways was calculated. Statistical analyses was performed to identify if compounds caused increased functional connectivity between the two selected ROIs, relative to control.

As an intuitive example of the reasoning behind this process, one would expect that the functional connectivity between ROIs connected by neurotransmitters

such as serotonin would be increased by a serotonin reuptake inhibitor. Generally speaking this proved not to be the case.

Ref	Neurotransmitter	Pathways	Drugs with increased functional connectivity	Drugs with decreased functional connectivity
(113)	Serotonin	Raphe Inferior - Telencephalon	Cocaine, Muscarine, Physostigmine, SB205607, NMDA, Amitriptyline, Amphetamine, Bethanechol	Mizolastine, Picrotoxin, SNC80, Amoxipine, Strychnine, Bemegride, Clonidine, Gabazine
(113)	Serotonin	Telencephalon Pallium	Ketamine, Ketoconazole, Muscarine, Physostigmine, Tacrine, Donepezil, NMDA, Amitriptyline, Amphetamine, Bethanechol, Cis-ACPD	Morphine, Amoxipine, Strychnine, Gabazine, Maprotiline
(113)	Serotonin	Raphe Inferior - Diencephalon	Rolipram, Amitriptyline, Bethanechol, SB205607,	PTZ, Apomorphine, Amoxipine, Cocaine, Ketamine, Ketoconazole, Tacrine, NMDA, Pilocarpine, Strychnine, Aminophylline, Bupropion, Donepezil, 4-Aminophenylsulfone, Bicuculine, Clonidine, Gabazine, Mizolastine, Picrotoxin, SNC80, Bemegride
(113)	Serotonin	Raphe Superior - Diencephalon	Kainic, Rolipram, Amitriptyline, Muscarine, Nomifensine,	PTZ, Pilocarpine, Strychnine, Bemegride, Bicuculine, Gabazine, Picrotoxin, Amoxipine,
(113)	Serotonin	Raphe Inferior - Preoptic Area	Amitriptyline, Bethanechol,	Bicuculine, Clonidine, Gabazine, Picrotoxin, Quinine, SNC80, Theophylline, Strychnine, Aminophylline, Donepezil, Bemegride,
(113)	Serotonin	Raphe Superior - Preoptic Area	Bethanechol, Cocaine, Ketamine, Muscarine, NMDA, Cisplatin, Amitriptyline, Amphetamine	Strychnine, Bicuculine, Gabazine, Picrotoxin,
(113)	Serotonin	Raphe Inferior - Ventral Thalamus	Bethanechol, Muscarine, SB205607, Tetraethylammonium, Apomorphine, Morphine, Amitriptyline, Amphetamine	Gabazine, Picrotoxin, SNC80, Theophylline, Emetine, Rolipram, Aminophylline, 4-Aminophenysulfone, Bemegride, Bicuculine, Clonidine
(113)	Serotonin	Raphe Superior - Ventral Thalamus	Amitriptyline, Amphetamine, Bethanechol, Cocaine, Muscarine, Tetraethylammonium,	Strychnine, Aminophylline, Bicuculine, Picrotoxin, Quinine, SNC80, Theophylline,
(113)	Serotonin	Raphe Inferior - Posterior Tuberculum	Amitriptyline, Amphetamine, Bethanechol, Cis-ACPD, Muscarine, Tetraethylammonium	Bicuculine, Clonidine, Gabazine, Picrotoxin, Quinine, SNC80, Theophylline, Apomorphine, PTZ, Pilocarpine, Strychnine, Aminophylline, Donepezil, 4-Aminophenylsulfone, Bemegride
(113)	Serotonin	Raphe Superior - Posterior Tuberculum	NMDA, Kainic, Cis-ACPD, Muscarine	Amoxipine
(113)	Serotonin	Raphe Inferior - Habenulae	Clonidine, Kainic, Apomorphine, Rolipram, Aminophylline, Donepezil, NMDA	SB205607
(113)	Serotonin	Raphe Superior - Habenulae	Donepezil, NMDA, Kainic, Amitriptyline, Nomifensine,	PTZ, Bicuculine, Gabazine, Picrotoxin, Quinine, Caffeine, Cis-ACPD,
(113)	Serotonin	Raphe Inferior - Mesencephalon	Bethanechol, Cis-ACPD, Muscarine, Physostigmine, SB205607, Tacrine, NMDA, Clozapine, Yohimbine, Amitriptyline, Amphetamine	4-Aminophenylsulfone, Gabazine, SNC80, Theophylline, Apomorphine, Rolipram,
(113)	Serotonin	Raphe Superior - Mesencephalon	Cis-ACPD, Muscarine, Physostigmine, SB205607, Tacrine, NMDA, Amitriptyline, Amoxipine, Amphetamine, Bethanechol	Theophylline, Apomorphine, Emetine, Morphine, Rolipram, 4-Aminophenylsulfone, Bicuculine, Clonidine, Gabazine, Quinine, SNC80
(113)	Serotonin	Raphe Inferior - Torus semicircularis	Bethanechol, Cocaine, Muscarine, Tetraethylammonium, Bupropion, NMDA, Yohimbine, Amitriptyline, Amoxipine	Bicuculine, Picrotoxin, Rolipram, Caffeine,
(113)	Serotonin	Raphe Superior - Torus semicircularis	Bethanechol, Cocaine, Muscarine, Tetraethylammonium, Bupropion, NMDA, Yohimbine, Amitriptyline	Bicuculine, Picrotoxin, Caffeine,
(114)	Serotonin	Raphe Inferior - Spinal Cord	Physostigmine, SB205607, Tacrine, Tetraethylammonium, NMDA, Amitriptyline, Bethanechol, Muscarine	Gabazine, Picrotoxin, SNC80, Theophylline, Strychnine, Bemegride, Bicuculine, Clonidine
(114)	Serotonin	Raphe Superior - Spinal Cord	Cis-ACPD, Cocaine, Ketamine, Ketoconazole, Tetraethylammonium, Muscarine, Nomifensine, Olanzapine, SB205607, Tacrine, Caffeine, Amitriptyline, Amoxipine, Amphetamine, Bethanechol, Physostigmine	Clozapine, Gabazine, Kainic, Apomorphine, Cisplatin, Emetine, Morphine, Rolipram, 4AP, PTZ, Pilocarpine, Strychnine, Aminophylline, Bupropion, 4-Aminophenylsulfone, Bemegride, Bicuculine, Maprotiline, Mizolastine, Picrotoxin, Quinine, SNC80, Theophylline, Yohimbine, Clonidine
(114)	Serotonin	Raphe Inferior - Olfactory bulb	Rolipram, Amitriptyline, Bethanechol, Muscarine,	Mizolastine, Picrotoxin, Quinine, SNC80, Theophylline, Apomorphine, Amoxipine, Cocaine, Nomifensine, Physostigmine, Tacrine,

	Neurotransmitter	ROIs	Higher connectivity	Lower connectivity
				PTZ, Pilocarpine, Strychnine, Aminophylline, Donepezil, 4-Aminophenylsulfone, Bemegride, Bicuculine, Clonidine, Clozapine, Gabazine, Maprotiline,
(114)	Serotonin	Raphe Superior - Olfactory bulb	Amitriptyline,	Cocaine, Ketamine, Ketoconazole, Nomifensine, Olanzapine, Physostigmine, Tacrine, Strychnine, Bupropion, 4-Aminophenylsulfone, Bicuculine, Mizolastine, Quinine, SNC80, Caffeine
(115)	Dopamine	Diencephalon - Telencephalon	Rolipram, Amitriptyline,	PTZ, Pilocarpine, Strychnine, Aminophylline, Bicuculine, Caffeine, Cis-ACPD, Olanzapine, Physostigmine, SB205607, Tacrine, Tetraethylammonium
(115)	Dopamine	Diencephalon - Pretectum	Bicuculine, Clozapine, Kainic, Apomorphine, Cisplatin, Emetine, Amitriptyline, Muscarine,	Amoxipine, Caffeine, Ketoconazole, Nomifensine, Olanzapine, Tetraethylammonium, Strychnine, Bemegride, Gabazine, Maprotiline, Quinine, Rolipram
(115)	Dopamine	Diencephalon - Subpallium	Clozapine, Kainic, Yohimbine, Apomorphine, Cisplatin, Emetine, Rolipram, Amitriptyline, Muscarine,	PTZ, Strychnine, Bemegride, Maprotiline, Quinine, Amoxipine, Caffeine, Nomifensine, Tetraethylammonium,
(115)	Dopamine	Ventral Thalamus - Subpallium	Amitriptyline, Amphetamine, Bethanechol, Caffeine, Cis-ACPD, Ketoconazole, Muscarine, SB205607,	PTZ, Strychnine, 4-Aminophenylsulfone, Clonidine, Gabazine, Kainic, Maprotiline, Mizolastine, Quinine, SNC80, Apomorphine, Cisplatin, Emetine, Morphine,
(115)	Dopamine	Posterior Tuberculum - Subpallium	Donepezil, Cis-ACPD, Cocaine, Muscarine, Olanzapine, Physostigmine, Tetraethylammonium,	Aminophylline, Theophylline, Ketoconazole,
(116)	Dopamine	Pretectum - Dorsal Thalamus	Kainic, Cisplatin, Emetine, Rolipram,	PTZ, Pilocarpine, Strychnine, Bupropion, Bicuculine, Cocaine,
(116)	Dopamine/ Serotonin	Posterior Tuberculum - Dorsal Thalamus	Amitriptyline,	NMDA, Bicuculine, Maprotiline, Amoxipine, Caffeine, Cis-ACPD, Cocaine, Olanzapine, Tacrine, 4AP, PTZ, Pilocarpine, Strychnine, Aminophylline, Donepezil
(117)	Acetylcholine	Habenulae - Interpendnuclar nucleus	Aminophylline, NMDA, Kainic, Apomorphine, Emetine, Rolipram, Amitriptyline, Muscarine,	Strychnine, Bicuculine, Picrotoxin, Quinine, Amoxipine, Caffeine, Cis-ACPD, Nomifensine, SB205607,
(118)	Acetylcholine	Tegmentum - Diencephalon	Aminophylline, NMDA, Kainic, Apomorphine, Rolipram, Amitriptyline,	PTZ, Strychnine, Bicuculine, Picrotoxin, Amoxipine, Bethanechol, Caffeine, Cis-ACPD, Olanzapine, Physostigmine, SB205607,
(118)	Acetylcholine	Torus semicircularis - Pretectum	Amitriptyline, Amphetamine, Bethanechol,	Strychnine, Bicuculine, Quinine, Theophylline, Caffeine, Muscarine, Tacrine,
(118)	Acetylcholine	Diencephalon - Rhombencephalon	Kainic, Rolipram, Amitriptyline,	Strychnine, Bupropion, Donepezil, Bemegride, Bicuculine, Picrotoxin, Amoxipine, Caffeine, Ketoconazole, Muscarine,
(116)	Noradrenaline	Cerebellum - Locus coeruleus	, Bethanechol, Caffeine, Cis-ACPD, Cocaine, Ketamine, Ketoconazole, Muscarine, Nomifensine, Physostigmine, Donepezil, NMDA, Bemegride, Clozapine, Picrotoxin, Amitriptyline, Amoxipine, Amphetamine	Kainic, Theophylline, Apomorphine, Morphine,
(116)	Noradrenaline/ Dopamine	Area postrema - Spinal Cord	Yohimbine, Apomorphine, Cisplatin, Emetine, Morphine, Rolipram, Amitriptyline, Amphetamine, Bethanechol, Cocaine, Muscarine, Physostigmine, 4AP, Aminophylline, Bupropion, Donepezil, NMDA, Bemegride, Clozapine, Kainic, Theophylline,	

Table 4: from left to right: Column 2: The neurotransmitter present in the identified neurochemical pathway. Column 3: The two ROIs the neurochemical pathway connects. Column 4: Compounds that have a signficantly higher mean connectivity value between the two corresponding ROIs than control. Column 5: Compounds that have a signifcantly lower mean connectivity value between the two corresponding ROIs than water.

Amitriptyline features heavily as amitriptyline causes a large increase in functional connectivity generally (see table 3). Moreover, dopamine agonists

such as cocaine and amphetamine do not appear to increase functional connectivity between the identified ROIs for dopamine. The only exception being between the posterior tuberculum and the Subpallium. Similarly, acetylcholine upregulators such as donepezil, tacrine, and physostigmine noticeably do not increase functional connectivity between ROIs connected by alleged acetylcholine pathways.

Connections between the raphe superior and the spinal cord appear to have some sensitivity for serotonergic compounds. Cocaine, ketamine, nomifensine, amoxipine and amphetamine all increase functional connectivity between the regions and all are agonists of serotonin. However, the functional connectivity of this area was increased for a number of other compounds also.

Drugs with effects on noradrenaline seemed to consistently increase functional connectivity in both the noradrenaline pathways with compounds such as bupropion, amphetamine and amoxipine. However, again relevant compounds appear beside a wide array of other compounds.

Generally speaking this particular approach did not show that there were any specific functional connectivity values that were indicative of specific pharmacologies. It is on this basis that multivariate techniques were used, in to take into account all the parameters simultaneously, as opposed to the univariate approaches described above.

Hierarchical clustering of Principal component scores

PCA product

I used a hierarchical clustering algorithm to produce a dendrogram, clustering together drugs with similar functional connectivity values or peak profile parameters in the zebrafish brain, based on their respective principal component scores (figure 8). To demonstrate that the principal component scores were accurately represented by the dendrogram and to highlight the way in which principal component scores represent compounds, the principal component scores of drugs that were closely linked in the dendrogram are shown in figure 7.

As an example, CisACPD and bethanechol on the left-hand side of figure 7 are closely associated with one another in the dendrogram, (see figure 8). Similarly,

4-aminophenylsulfone and SNC80, on the right hand-side of figure 7 are closely associated, but the pair on the left and the pair on the right hand-sides are distinct from one another. These similarities and distinctions in figure 8 are mirrored in figure 7.
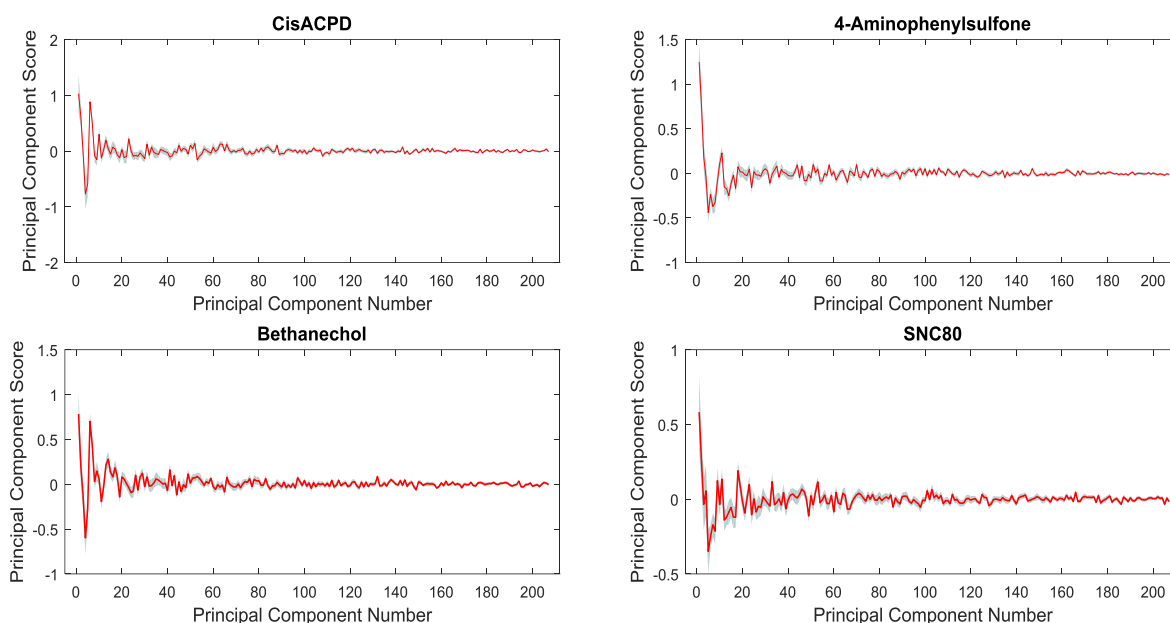

Figure 7: Line graphs of the mean principal component scores. Grey outline is standard error of the mean of the principal component scores across all experiments of the respective class. The two compounds on the left-hand side cluster together strongly and have highly similar mean principal component scores, as do the two drugs on the right. Comparatively comparing the left and right-hand sides there is relatively little similarity as evidenced by the difference in profiles exhibited.

Hierarchical Clustering of Functional Connectivity: Relation to Known Pharmacology

4-Aminophenylsulfone and SNC80 were the two compounds deemed most similar in functional connectivity by the clustering algorithm (see figure 7). 4-Aminophenylsulfone is an antibiotic with no known effect on brain function, conversely SNC80 is a delta opioid agonist. However, there is evidence to suggest that SNC80 has poor affinity to zebrafish delta opioid receptors and it seems in this case to have little effect in the zebrafish brain (119), thus explaining its close proximity to 4-aminophenylsulfone in this dendrogram. The antihistamine mizolastine is also clustered closely to both SNC80 and 4-aminophenylsulfone. This is unsurprising as mizolastine has been shown to be free of the sedative effects often associated with H-1-receptor antagonists (120).

The next two closely associated compounds were apomorphine and emetine, both drugs have emetic properties which may explain their close proximity (121,122). Cisplatin and kainic acid form a close-knit cluster with emetine and apomorphine. Cisplatin is also a powerful emetic which strengthens the suggestion that emetic drugs induce distinct functional connectivity patterns. (123). Conversely kainic acid is a glutamate agonist (124), however in the larval cultures for exposure via the medium it does not appear to be bioavailable and so is unlikely to elicit any pharmacological effect (see table 1).

The medications amoxipine and olanzapine also have very similar functional connectivity patterns to one another. Amoxipine and olanzapine are used for treating depression and schizophrenia, respectively, and both compounds have effects on dopamine, serotonin and noradrenaline. Interestingly, while amoxipine is a serotonin and noradrenaline reuptake inhibitor (79), olanzapine is a serotonin and noradrenaline antagonist (64). In this context, it is counter intuitive that these two seemingly contradictory compounds associate so closely. However, amoxipine does also have some antagonistic effects on serotonin and noradrenaline receptors and it may be the case that these effects are accentuated in zebrafish (79).

The illicit drug cocaine is the next most similar compound, in terms of functional connectivity, to amoxipine and olanzapine. Cocaine prevents reuptake of serotonin, noradrenaline and dopamine by inhibiting their associated transporters (125). Cocaine therefore shares some of its primary pharmacological actions with amoxipine which may explain the similar functional connectivity patterns. The cluster most closely associated with amoxipine olanzapine and cocaine includes the three compounds nomifensine, ketamine and ketoconazole. Much like cocaine, nomifensine is a dopamine reuptake inhibitor (102). Nomifensine is most closely associated with ketamine in the dendrogram. Ketamine is a drug with a large repertoire of pharmacological effects including N-methyl D-aspartate (NMDA) receptor inhibition (101), dopamine receptor agonism (126) and serotonin receptor agonism (127). The latter two effects may explain ketamine's similarity in functional connectivity with drugs such as amoxipine, nomifensine and cocaine. Ketoconazole is the compound most closely linked to nomifensine and ketamine, despite being an antifungal with no recorded effects on dopamine, serotonin or noradrenaline (104).
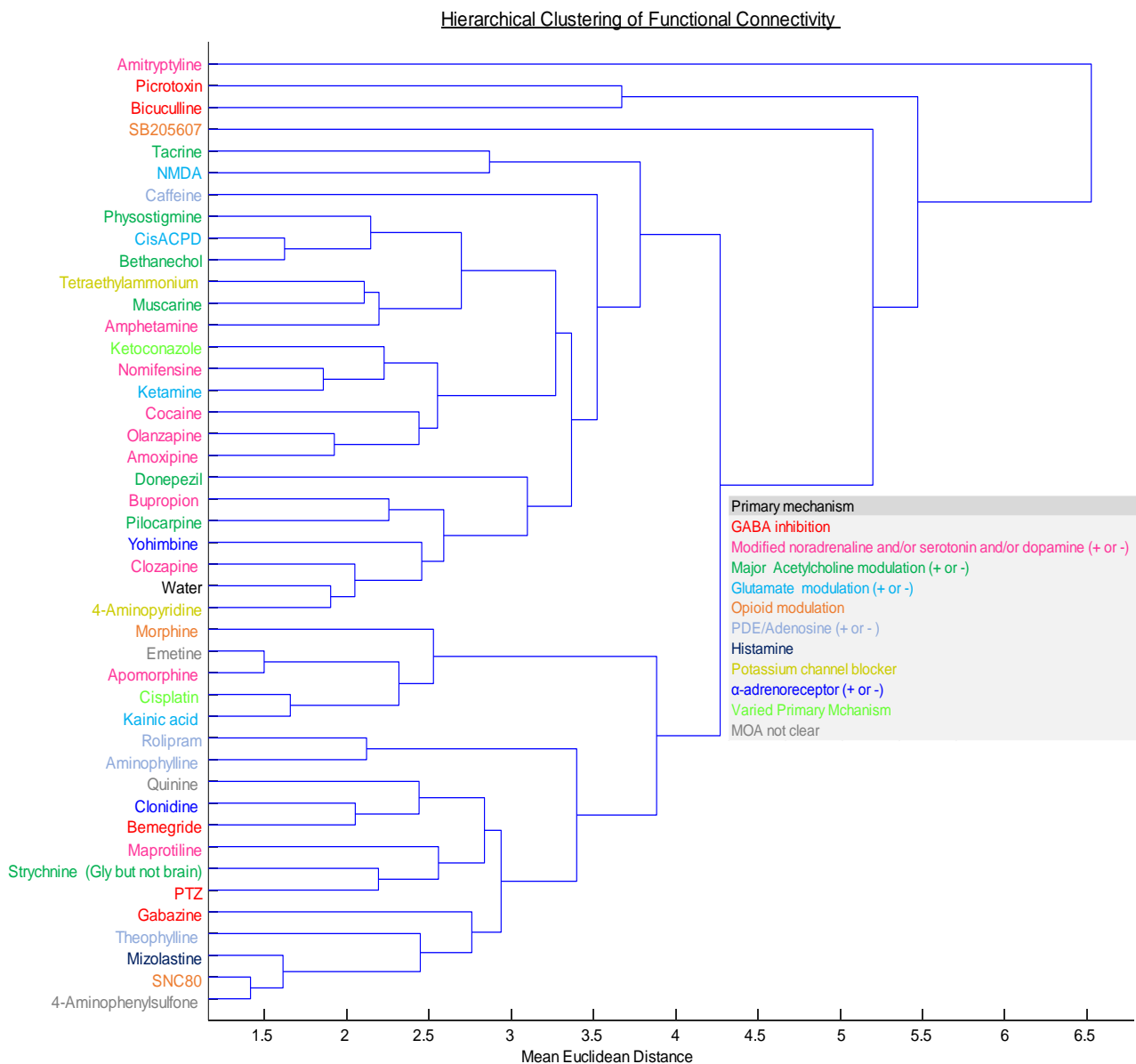
Figure 8: Cluster dendrogram output of hierarchical clustering algorithm performed on functional connectivity data. Colours indicate broader pharmacological classes.

Higher up in the dendrogram from the aforementioned clusters, is the pairing of cis-ACPD and bethanechol. Bethanechol is a parasympathomimetic which is pharmacologically and structurally related to acetylcholine (83). In contrast, cis-ACPD is a potent NMDA receptor agonist (71). Both compounds induce a generalised increase in functional connectivity across the zebrafish brain (see table 3) but do so using different mechanisms of action. The next most closely associated drug to this pair is physostigmine, an acetylcholinesterase inhibitor. Much like bethanechol, physostigmine is a parasympathomimetic, and is

responsible for the stimulation of muscarinic acetylcholine receptors (62). The cluster most closely associated to physostigmine, bethanechol and cis-ACPD also contains a muscarinic agonist, muscarine, suggesting there is some clustering of parasympathomimetic compounds (82). Muscarine is partnered most closely with tetraethylammonium, a compound that increases the length of action potentials by blocking potassium channels (128). Tetraethylammonium may have other properties that explain its similarity to muscarine, as tetraethylammonium is not a well understood compound. Directly adjacent to this pair is the illicit drug amphetamine. Amphetamine has a wide range of pharmacological effects including dopamine reuptake inhibition (129). Amphetamines dopaminergic activity may explain its similarity to muscarine, as stimulation of muscarinic receptors can initiate dopamine release (130).

Interestingly, tacrine and NMDA for a cluster. This cluster is markedly distinct from all other clusters however this pair is mirrored in the cis-ACPD and bethanechol pairing. NMDA, like cis-ACPD, is a NMDA receptor agonist (131) and tacrine, like bethanechol, is an acetylcholine agonist (132). The pairing of acetylcholine agonists and NMDA receptor agonists is unexpected as they are two distinct pharmacological mechanisms however it may be explained as a reflection of the seizurogenic properties of the compounds. Drugs that increase acetylcholine activity have been implicated in seizurogenesis in rats (133), similarly NMDA agonism is known to be effective in generating seizures (134).

Peak parameter clustering: relation to known pharmacology

When clustering the peak parameters, the two most closely associated compounds were the water and cisplatin (see figure 9). This suggests cisplatin has little effect on zebrafish brain dynamics. Given that cisplatin is not typically considered a neurotropic drug it's similarity to water conforms is what would be expected pharmacologically (135). The next two most closely associated drugs with cisplatin and water are kainic acid and emetine. Kainic acid, as mentioned above, does not appear in detectable quantities in the zebrafish brain (see table 1), using the culture approach adopted, thus, its similarity with water is unsurprising. Emetine is closely paired with kainic acid and like cisplatin isn't typically associated with any cognitive effects (97). The next most closely associated compound with the aforementioned four compounds is 4-aminophenylsulfone, another drug without established effects on the central

nervous system (92). The next most closely allied pair contains the compounds ketoconazole and clonidine. Ketoconazole is an anti-fungal drug with no known brain effects, as such it would not be expected to induce changes in the zebrafish brain, thus explaining its relatively close localisation with water (14). Conversely, clonidine is an alpha-adrenergic agent that has well established effects on the brain and is known to localise in the CNS. By binding to a variety of alpha-adrenergic receptors it has a plethora of effects including pain relief and reduced
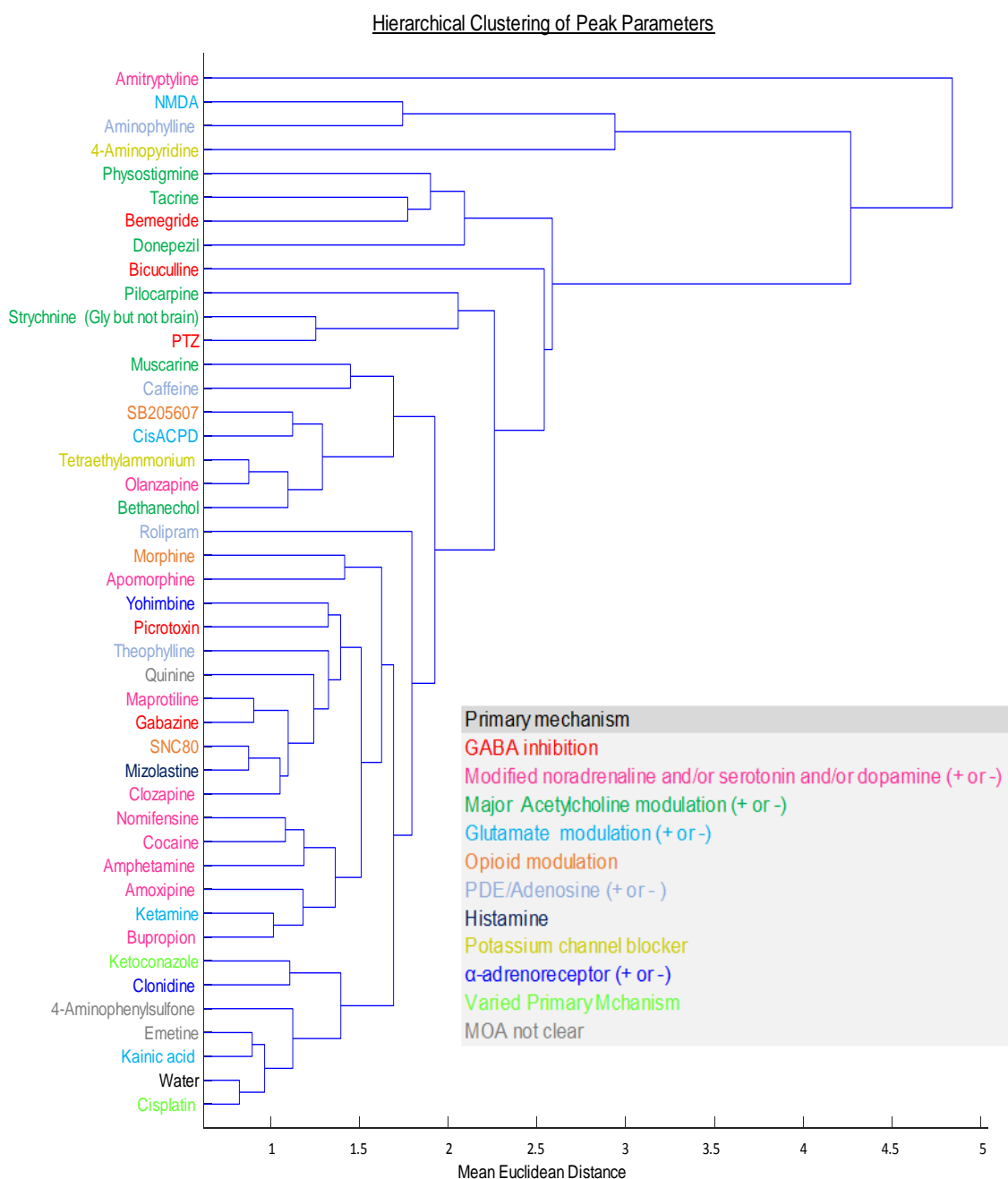

Figure 9: Cluster dendrogram output of hierarchical clustering algorithm performed on peak parameters. Colours indicate broader pharmacological classes.

blood pressure (93). Clonidine may represent the boundary after which brain active drugs start to be present in the dendrogram.

The next cluster positioned higher up the dendrogram from water and contains an assortment of serotonergic and dopaminergic active compounds. Of these, the pair of compounds with most similar peak parameters are ketamine and bupropion. Ketamine, among a range of other effects enhances descending inhibiting serotonergic pathways and inhibits reuptake of serotonin, noradrenaline and dopamine (136). Similarly, bupropion has inhibiting effects on serotonin, noradrenaline and dopamine reuptake (77). The next most similar compound is

Compound	Mean AUC	Peak Number	Mean Peak height	Mean Peak width	Mean Peak Seperation
Aminophylline	605.76	3.98	11063.30	17.48	21.31
NMDA	588.79	3.19	10726.20	18.45	21.01
Amitriptyline	43.68	4.01	1406.69	35.73	34.18
Tacrine	286.85	4.78	3903.53	12.06	32.91
Kainic	318.20	6.34	5069.05	11.86	22.23
Donepezil	295.31	5.69	1793.80	8.44	32.93
Bicuculine	213.74	8.04	1502.78	10.05	24.83
Amoxipine	190.07	6.67	1070.26	8.06	33.93
Apomorphine	153.70	5.05	1386.41	11.97	38.49
Bupropion	132.28	7.15	961.52	10.27	29.82
Nomifensine	162.90	5.90	1095.01	9.66	33.87
Cocaine	139.18	6.08	1107.05	11.19	32.57
Bemegride	227.41	4.76	2833.84	10.95	31.30
Rolipram	108.16	5.94	1045.88	12.38	31.85
Cisplatin	100.13	8.13	764.55	9.55	25.31
Ketamine	160.63	6.90	1150.27	9.47	27.36
Caffeine	162.85	6.19	1269.79	10.33	28.95
Morphine	137.05	6.43	1023.21	9.77	29.63
Quinine	124.05	7.14	929.89	9.11	27.47
4AP	320.29	4.68	5378.09	14.93	17.28
Clozapine	113.02	5.79	828.60	10.57	32.30
Tetraethylammonium	121.82	6.93	842.45	10.14	27.33
SNC80	124.78	6.48	908.01	9.62	29.31
Mizolastine	126.06	6.51	991.01	10.51	28.22
Emetine	105.66	7.37	843.92	10.68	25.41
Theophylline	99.83	5.78	929.72	11.58	31.54
Water - (Control)	111.14	7.41	943.92	10.37	25.00
Olanzapine	119.04	7.19	789.63	10.10	25.78
Ketoconozole	102.84	7.53	702.29	9.09	25.53
4-Aminophenylsulfone	111.84	7.31	768.02	8.78	26.20
Clonidine	97.70	7.19	688.04	8.94	27.02
Pilocarpine	113.27	6.59	1520.32	12.12	24.94
Picrotoxin	109.05	5.72	880.96	9.63	31.95
Amphetamine	135.04	5.88	1119.62	10.40	29.19
Physostigmine	177.49	4.41	1282.08	10.54	33.44
Gabazine	166.17	5.94	2039.30	10.73	25.24
Muscarine	120.60	5.80	1184.65	13.27	27.32
Maprotiline	82.19	6.76	630.31	8.83	28.44
SB205607	97.59	7.56	644.51	9.90	23.22
Strychnine	122.12	6.61	698.90	7.68	27.90
Cis-ACPD	183.08	7.25	1397.33	9.41	19.22
Bethanechol	97.44	6.85	738.20	10.88	23.23
Yohimbine	76.74	5.95	629.12	11.45	26.56
Pentylenetetrazole	40.89	6.42	248.04	8.10	28.20

Table 3: from left to tight: compounds used in experiment, mean area under the curve, mean peak number, mean peak height, mean peak width, and mean peak seperation. All veargaes were calculated across all ROIs and all experiments of each class.

amoxipine, a tricyclic antidepressant that induces noradrenaline and serotonin reuptake inhibition (137).

Adjacent to this cluster is a trio of drugs including: nomifensine, amphetamine and cocaine. All three of these compounds are involved in either dopamine reuptake inhibition and or direct dopamine agonism (103,138,139). Amphetamine and cocaine both also increase synaptic neurotransmission of serotonin and noradrenaline (138,139). Notably, nomifensine and apomorphine occupy different parts of the dendrogram, despite their similar chemical structure. However, this is perhaps not unexpected as while both compounds upregulate dopaminergic neurotransmission, they have contrasting effects on the adrenergic system. Nomifensine is a noradrenaline reuptake inhibitor (140) and thus is responsible for increasing adrenergic synaptic transmission while apomorphine is thought to have antagonistic effects on adrenergic receptors (141).

Of the five-remaining serotonergic, dopaminergic or noradrenergic compounds apomorphine, clozapine and olanzapine all have some form of serotonergic, adrenergic or dopaminergic antagonism and thus their absence, from the aforementioned cluster of six mood enhancing compounds, makes pharmacological sense (95,142–144).

 Moreover, maprotiline, which is also absent, has only weak dopaminergic and serotonergic activity (78). Amitriptyline is the only remaining drug that should potentially cohabit this cluster due to its serotonin and noradrenaline reuptake inhibition (80). Given that amitriptyline is not remotely similar, in terms of its peak profiles or indeed its functional connectivity values, to any other compound it could be argued that it is an anomaly. Adjacent to the cluster of mood elevating compounds, described above is another tightly knit group of compounds consisting of clozapine, mizolastine, SNC80, gabazine and maprotiline. This cluster contains a diverse array of pharmacologies. Clozapine is an atypical antipsychotic medication which is an antagonist of dopamine, serotonin and adrenergic receptors (143). SNC80 and mizolastine, clozapine's closest allies, are a μ–δ opioid receptor agonist and antihistamine respectively (96,145). Maprotiline and gabazine also form a tight pairing and are an adrenergic reuptake inhibitor with weak action on serotonin and dopamine transporters and a $GABA_A$ antagonist respectively (78,146). While they do have

discordant pharmacologies gabazine and maprotiline both have relatively poor bioavailability in embryo-larval cultures (see table 1) and thus in the zebrafish brain which may explain their pairing.

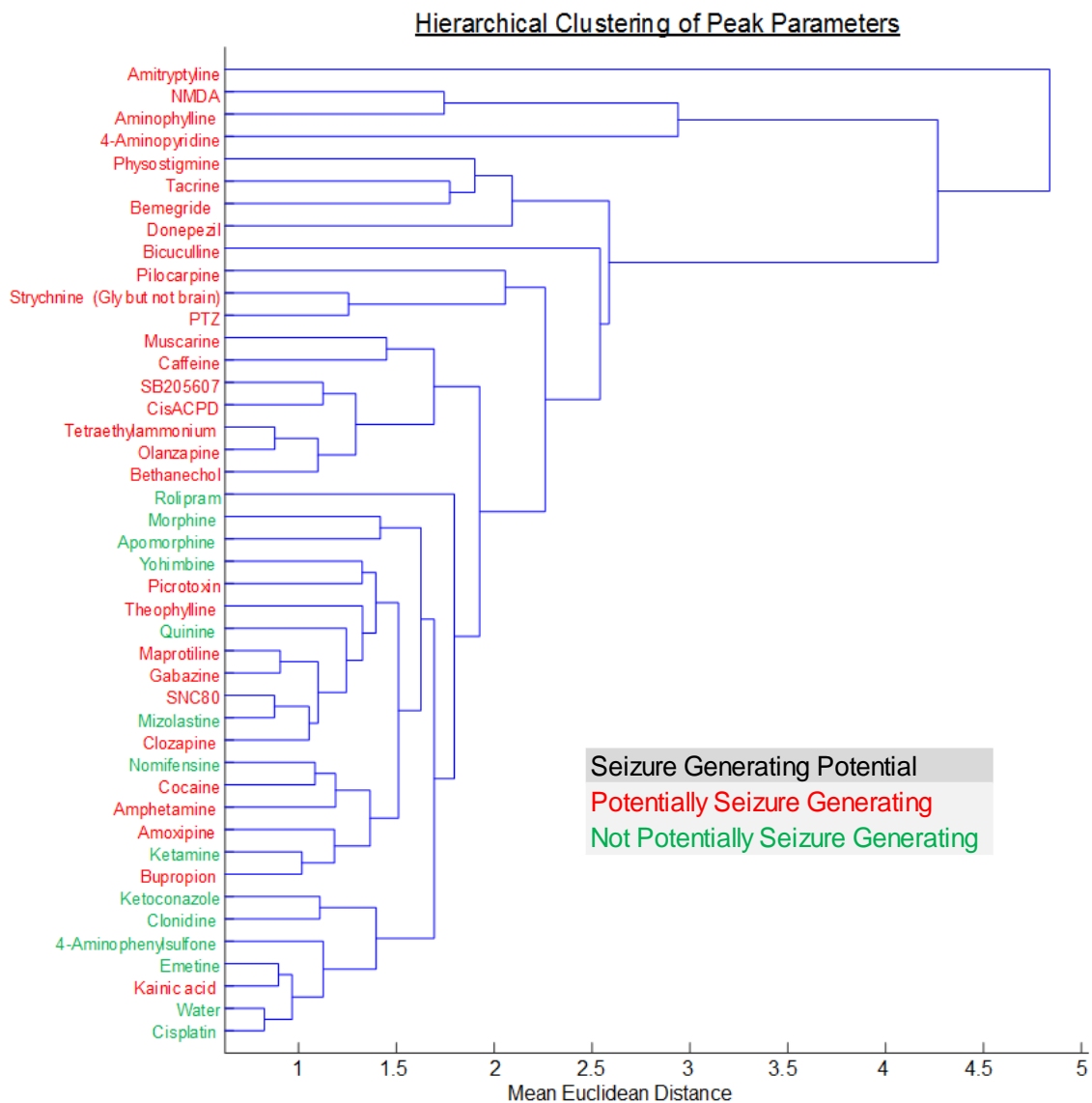Peak parameter clustering results in relation to seizurogenic properties

Figure 10: Cluster dendrogram output of hierarchical clustering algorithm performed on peak parameters.

Moving up the dendrogram from maprotiline, the following 14 compounds have little pharmacologically validatable clustering. At the point of PTZ, however, there again appears to be some clustering of acetylcholine modulating compounds. Strychnine, PTZ and pilocarpine form a cluster however a reasonably well spread one. PTZ and strychnine are both seizure inducing compounds however via different modes of action. PTZ inhibits GABA$_A$ to reduce inhibitory

neurotransmission and initiate seizures (147). Conversely strychnine induces seizure via glycine inhibition in the spinal cord (148). Pilocarpine, in contrast to strychnine, is a muscarinic agonist (81), rather than a acetylcholine antagonist (149). Sitting above pilocarpine in the dendrogram is a quartet of compounds: donepezil, bemegride, tacrine and physostigmine forming a cluster, however each member of the cluster is relatively far away in euclidean space. Of these four compounds, all except bemegride are acetylcholinesterase inhibitors (150–152).

At the very top of the dendrogram a single cluster of three compounds reside: NMDA, aminophylline and 4-Aminopyridine. All three of these compounds have distinct pharmacologies and indeed the dendrogram places them reasonably separate from each other in euclidean space. While they are not similar pharmacologically they all have well established seizurogenic properties (84,153–155).

Using the peak hierarchical clustering approach, drugs with seizure generating potential were highlighted to assess if seizurogenic compounds showed some clustering. Upon observation it seemed that they were, at least to some extent. All the non-seizurogenic drugs occupied a single large cluster at the bottom of the dendrogram (figure 10). All the sub-clusters within this large cluster were less than two units in euclidean space away from each other (see figure 10). However, there are several seizurogenic compounds within this cluster. Of these some have low uptake into the zebrafish CNS including picrotoxin, maprotiline, gabazine and kainic acid (Table 1). Additionally as SNC80 has poor affinity for zebrafish delta opioid receptors (119) it is unlikely to demonstrate its seizurogenic potential (156). Moreover, there are no uptake data for theophylline and thus it is also possible that low bioavailability is an issue. Aside from theophylline, all the remaining seizurogenic drugs that reside within this cluster are serotonin, dopamine and noradrenaline inhibitors. These mood elevating compounds cluster together regardless of seizurogenic potential and seem to elicit spatially specific peak profiles.

Functional connectivity clustering results in relation to seizurogenic properties

Clustering of seizurogenic compounds was less specific when clustering was performed on the connectivity parameters. There was some clustering of

seizurogenic drugs at the top of the dendrogram (figure 11), including amphetamine and Picrotoxin, two compounds that clustered with non-seizurogenic compounds in the peak profile dendrogram. However, generally speaking, non-seizurogenic drugs were spaced fairly evenly throughout. It is clear that while functional connectivity shows some clustering of seizurogenic compounds it is, generally speaking, an inferior method for assessing the seizurogenic properties of compounds relative to assessing the peak profiles.
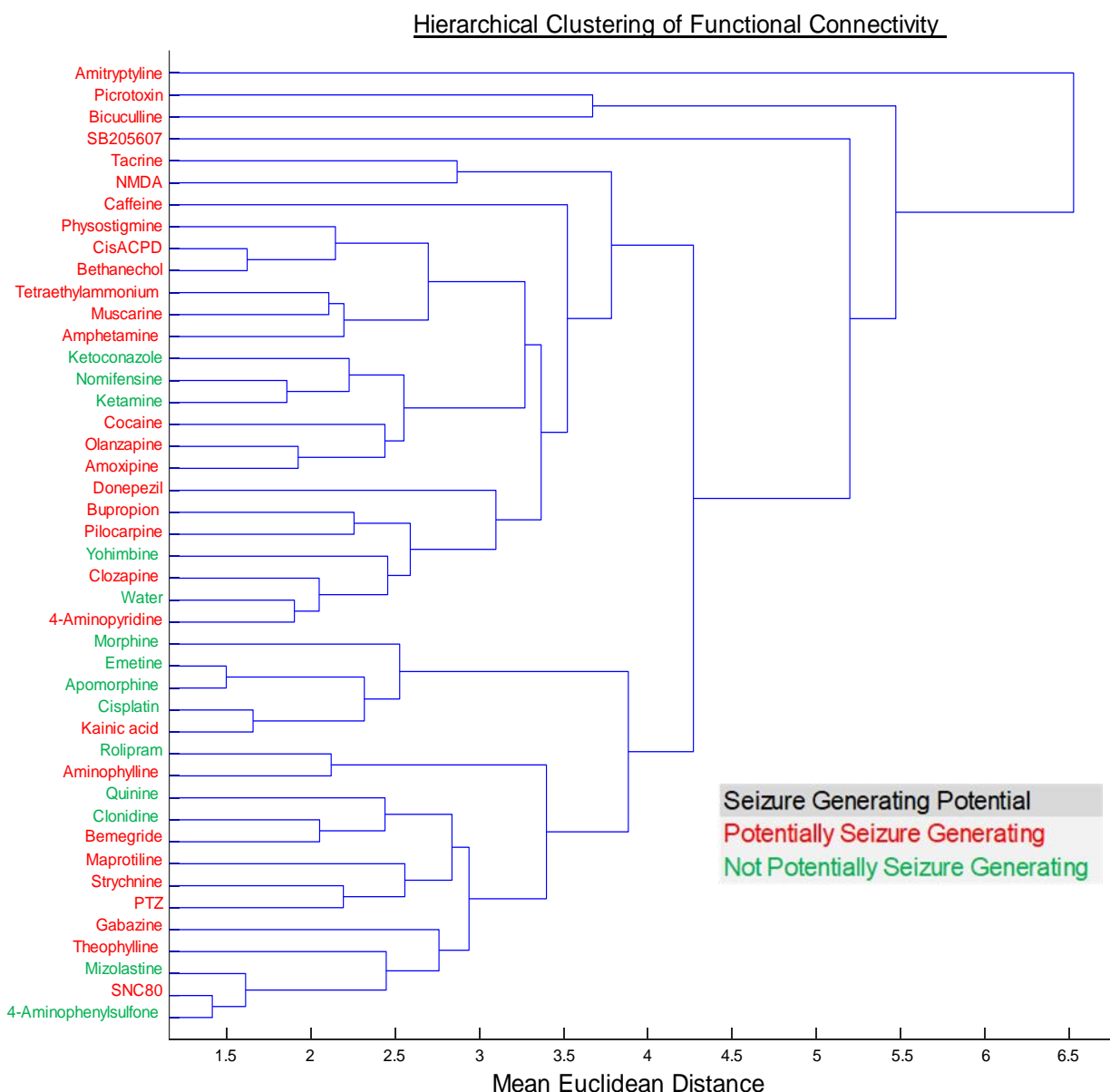

Figure 11: Cluster dendrogram output of hierarchical clustering algorithm performed on connectivity parameters. The colours of compounds indicate seizurogenic potential, with compounds in red being potentially seizurogenic, and those in green being non-seizurogenic

Differences between peak and functional connectivity analyses

Water and 4-Aminopyridine cluster tightly together when analysed based on their respective functional connectivity values. However, when clustered based on their peak profiles the two compounds were distinct from one another (see figure 10 and 11). Similarly, water and cisplatin clustered closely based on their peak parameters but were distinct from one another when compared using their functional connectivity values (see figure 10 and 11). Figure 12 (top of next page) highlights the distinction between utilising peak parameters extracted from the time series and functional connectivity. While some compounds may have closely related functional connectivity this does not necessarily translate to closely related functional connectivity and vice versa. These significant distinctions between the two sets of parameters underlies the motivation for analysing them separately.
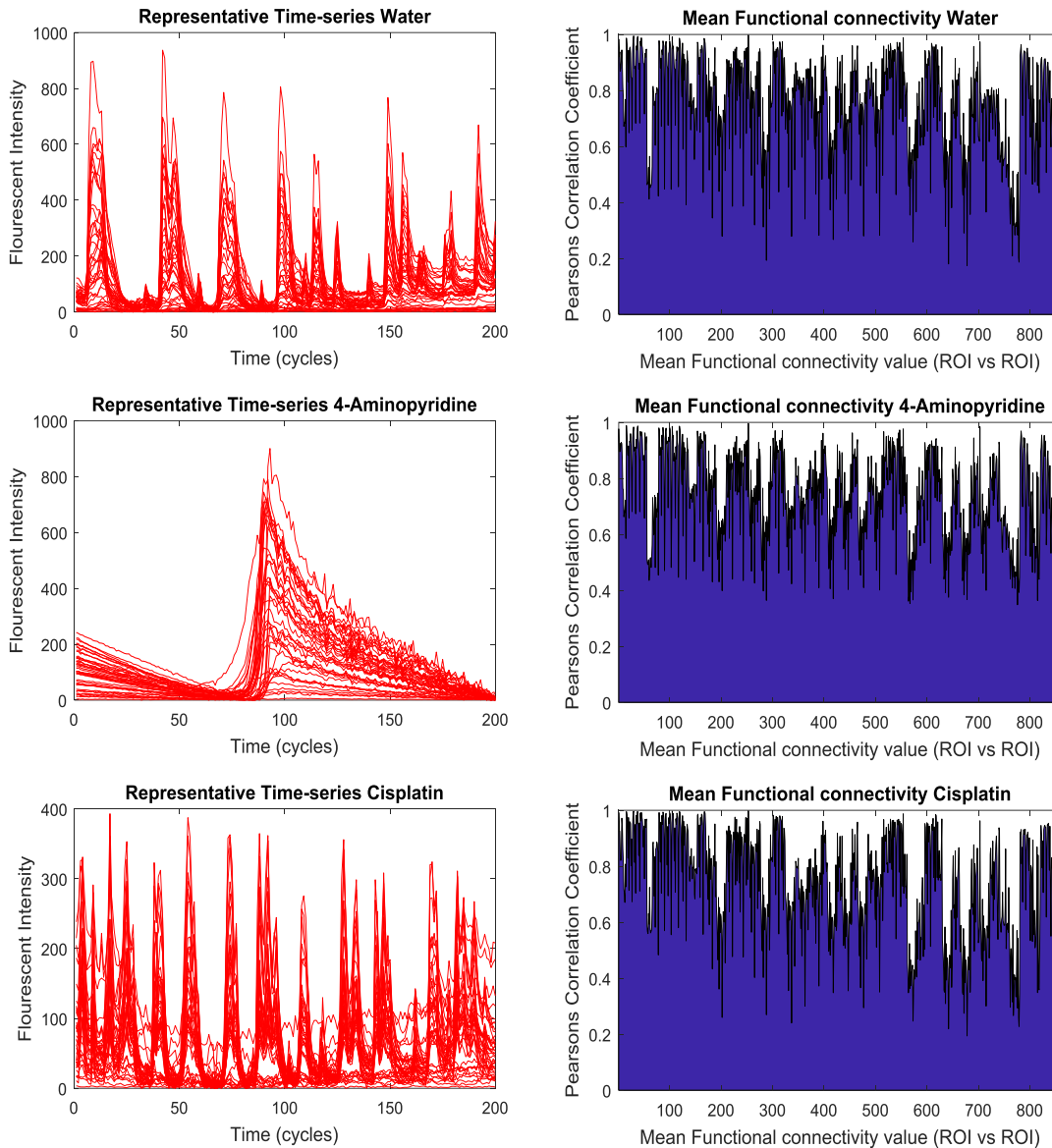
Figure 12: On the left-hand side: Representative time-series of all 42 ROIs. ON the right-hand side: Mean functional connectivity values across experiments of individual classes (with outliers removed). The x-axis here represents every possible pair of ROIs.

Discussion

Of the analytical techniques used the most effective was the clustering approach. Similar multivariate clustering approaches have been used previously to successfully cluster cognition enhancing compounds based on parameters extracted from cortical EEG oscillations (8). The evidence from EEGs in mice suggests certain drugs elicit distinct temporal activation patterns and that analysis of brain dynamics over time could provide useful biomarkers for specific pharmacodynamics. The study describe here utilises calcium imaging techniques

for detection of brain activity and thus has a far greater emphasis on spatial information than temporal. Comparison of the spatially specific parameter clustering with the mean parameter values suggest that spatially specific temporal information is highly informative for profiling drugs and that spatial element is of particular importance.

Of the two sets of parameters, peak profiles and functional connectivity data, peak profiles seemed to provide the most useful data, both in terms of identifying seizurogenic compounds and identifying pharmacology. In terms of the functional connectivity data, it may be that for some compounds the pharmacological signatures were overridden due to hyperexcitation, induced by the high concentrations used for some compounds. Despite this the functional connectivity data were not totally unsuccessful; there was some clustering of both antidepressant and seizurogenic compounds.

During analysis of the peak profiles, five of the six drugs most closely associated with water would be considered unlikely to induce any significant cognitive effect based on their respective pharmacologies, suggesting that non-neurologically active compounds could be consistently identified using this approach. Moreover, the entire upper half of the dendrogram, contained a cluster exclusively populated with seizurogenic drugs.

Aside from an aptitude for identifying non-seizurogenic compounds, the peak profiles displayed an efficacy for clustering compounds based on their pharmacologies. An interesting example of this is the clustering of bupropion and ketamine. Bupropion is a well-established antidepressant drug and ketamine has come to the forefront in more recent years as an alternative antidepressant that is effective in treating patients with major depressive disorder, including those resistant to other medications (56). Including this pair, there is a cluster of six drugs that have a mixture of dopaminergic, noradrenergic and serotonergic agonistic effects, all of which could be classified as mood elevating compounds.

The functional connectivity analyses contained a similar cluster containing a number of the same compounds, however it was not as extensive nor as tightly knit as in the case of the peak profiles.

It seems, therefore, that mood elevating compounds, such as recreational and antidepressant drugs, elicit very specific and similar spatio-temporal patterns in

the zebrafish CNS. The approaches detailed here could therefore have some implication in the categorisation and identification of drugs for treating depression, a disease for which current therapeutics have a number of concerns in terms of efficacy and toxicity (157).

Peak profile parameters, unlike functional connectivity, provide a measure of the magnitude of activity. This is of particular relevance in the case of identifying seizurogenic compounds, as seizures very much derive from hyperexcitation (158), thus a measure of magnitude should confer greater differentiation between seizurogenic and non-seizurogenic compounds. Moreover, differences in magnitude are translated well in the analyses as euclidean distance is sensitive to these differences. In terms of pharmacology magnitude is also relevant as some pharmacologies will likely increase the number of action potentials in specific areas, therefore spatially specific increases in magnitude likely confer useful information for modelling of pharmacodynamics.

While the lack of information, in terms of magnitude, conferred by functional connectivity could be perceived as a flaw, it does allow for analyses to be performed without concerns that slight variations present in the model organism may have a detrimental effect on the results. This is of particular relevance, in this case, due to the lack of a within experiment control that could account for variation between animals or indeed batches of animals. As such, variations in terms of baseline fluorescent intensity are more difficult to account for as removing the mean linear trend doesn't fully mitigate the heterogeneity between animals. As functional connectivity analyses compare different ROIs from the same animal to one another, differences in overall magnitude between animals is mitigated.

In addition, there are distinction in terms of the spatial information provided by the two sets of parameters. The peak parameters consider the spatial element of the data but using summarised representations of the time series. Conversely functional connectivity considers the whole timeseries point by point and consequently provides a more complete representation of the spatio-temporal patterns, however does not provide direct information about the precise temporal dynamics, for example the size and amplitude of spikes.

Moreover, it is possible that functional connectivity could perform better in terms of representing pharmacology were hyperactivity not an issue and concentrations were kept at effective but non-hyperactive levels. It should be noted that for this animal model within experiment controls are not always possible, as drug exposure times can be long, and fish may not stay alive and healthy in agar for long enough for experiments to be performed. In this respect functional connectivity is advantageous and while it performed poorly in this study it should not be written off on this basis.

Chapter Three: Classification learners.

Introduction

In the previous chapter it was identified that potentially seizurogenic compounds cluster towards each other when compared based on their peak profiles. Therefore, in this chapter I set out to identify if a classification model could be trained that recognises seizurogenic compounds based on their peak profiles, and thus build a framework that could be used to separate seizurogenic from non-seizurogenic compounds. The overall aim is to create a model that can accurately identify seizurogenic compounds based on the peak profiles of unseen data. That is to say, were identical experiments performed on an unknown or novel compound and identical parameters extracted from said experiments, the data could be fed into the model and the model would provide a reliable assessment of whether that compound has the potential to be seizurogenic or not.

In order to do this, I utilised the MATLAB built in classification learner tool box (159) to generate and test a number of different machine learning algorithms. Machine learning algorithms vary greatly in terms of the kind of data they are adept at classifying, because different approaches depend on different assumptions of the data (160) (see below for more detail). Given the lack of information on the problem at hand I decided to test multiple classifiers to see how they performed against each other. Three different discriminative models were chosen for this task: a shallow neural network (NN), a support vector machine (SVM) and a logistic regression. The models were chosen because they are three models widely used in science and because they each have very different ways of addressing the classification problem.

Classification algorithms can be highly accurate predictive models. In past studies, utilising data generated from structural MRI scans, SVMs have been able predict the clinical response of patients to antidepressant drugs with up to 89% accuracy, prior to the initiation of treatment (161). Similarly, EEG data in conjunction with machine learning techniques, has been used to predict responses to clozapine therapy with up to 85% accuracy in an independent test group (162).

In addition to drug response prediction, machine learning approaches have also been used in the area of drug discovery. Novel compounds can be compared to known compounds and biological targets, based on their structural similarity, using parameters such as molecular properties, topological descriptors and pharmacophore fingerprints. Using machine learning techniques in conjunction with these measures can provide valuable insight into the applications of novel compounds (163). Such approaches provide information about compounds very early on in the drug creation process and are therefore an invaluable tool.

Later on, in the drug development process, machine learning can provide insight into the effects of drugs *in vivo*. Machine learning based analysis of brain networks has proved effective in differentiating drugs based on their *in vivo* effects. In one such study, fMRI scans were performed on healthy volunteers exposed to two compounds with distinct pharmacologies. The image data obtained was analysed using machine learning approaches and the resulting model could sensitively differentiate between the two compounds and a placebo (164). Machine learning algorithms have also been applied to behavioural assays in zebrafish as a predictor of drugs pharmacological profiles (165).

Given the efficacy of machine learning approaches on similar kinds of data sets, in the area of drug discovery, and the ease with which such tools can be applied to new data, machine learning algorithms were applied to differentiate between the seizurogenic and non-seizurogenic compounds in our data set.

The following sections give an overview of the principals necessary to understand machine learning. In addition, I describe intuitively how each of the algorithms used in this thesis function. This is not only to provide the reader with understanding but also to highlight how varied machine learning algorithms are and in what contexts particular types of algorithm flourish.

Training and test data

Classification learners are trained using training data. Training data is given to a classification learner in conjunction with information regarding how the data should be classified (i.e. the ground truth). Unseen test data, with known classes, is subsequently used to test the classifier to see how often the classifier classifies correctly (166).

For example, to generate a model that predicts whether or not a student was going to pass an exam a previous cohort of students that sat the exam could be analysed, with variables such as number of hours studied, previous test scores in other subjects, numbers of hours of sleep a night etc. extracted. A parameter, in this context, is any information that can be extracted from a data set that is potentially predictive of the outcome. The cohort could then be split into training and test sets, with half the students that passed the exam and half the students that failed the exam contained in each set.

The extracted parameters would be run through a training algorithm for the model in conjunction with the known output: whether or not that student passed the exam. This is what is known as supervised learning. That is to say, a classification learner is provided with the data to be classified and provided with the correct classification for each piece of input data.

Once the classifier has been 'trained' it is necessary to 'test' the classifier to see if it can effectively classify new data. This simply means providing a classifier with the test set. The classifier is then used to predict the classification of each student in the test set: whether or not they pass or fail the exam. Because the exam scores of the students are known the accuracy of the classification learner can be calculated.

Support Vector Machine

Extending upon the exam pass example in the context of SVM, for this example, I envisage two parameters: the test scores from two previous exams. First the data would be split into test and training sets. Then each of the individuals, from the training data, are mapped onto data points in a feature

space in this case points on a two-dimensional plot (one dimension for each parameter). Subsequently, a dividing line is drawn, in two dimensions, between individuals that fall into the different classifications; in this case individuals that pass an exam and individuals that fail.

The position of this dividing line is decided as follows: lines are drawn along the data points from each classification that are closest to the data points from the opposing class, these are known as the support vectors. A line is then drawn that is equidistant from each support vector and that maximally separates the classifications. Considering the application of this algorithm to higher dimensional problem, the line is called a hyperplane (167).
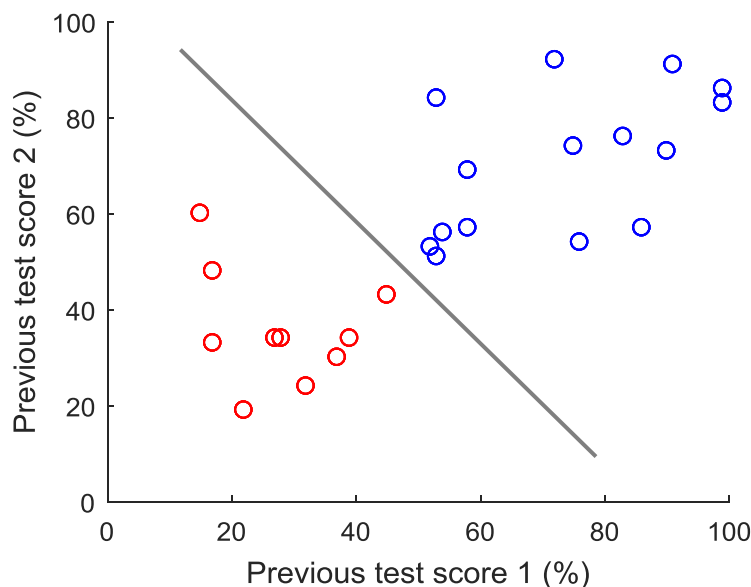
Figure 63: The blue circles represent students who pass the test. The red circles represent students who fail the test. The x and y axes are the parameters, in this case previous test scores. The line separating them is the hyperplane. When a new student's test scores are input into this model and are above the hyperplane they will be classified as a student who passes the exam.

When the test data is input into the constructed support vector machine it too is mapped into the feature space and the side of the line it falls on dictates its classification.

For the example above, two parameters were used for ease of visualisation. However this approach can be applied to data sets with a huge number of

parameters. The number of dimensions in the feature space is simply increased to be equal to (or higher than – in the case of kernel functions) the number of parameters. The parameters then act as co-ordinates for each of the data points in the multidimensional feature space.

The advantage of SVMs is they can be applied to very high dimensional data and don't necessarily require a large number of data points. In fact, the only data points that influence the outcome in an SVM are those that influence the positions of the support vectors.

Now the hyperplane is easy to draw in the case of linearly separable data such as in figure 13, because the data are completely separated in their feature space. Thus, the position of the hyperplane is easily dictated. However, if this is not the case, such as in figure 14, then generating the hyperplane becomes more complex.
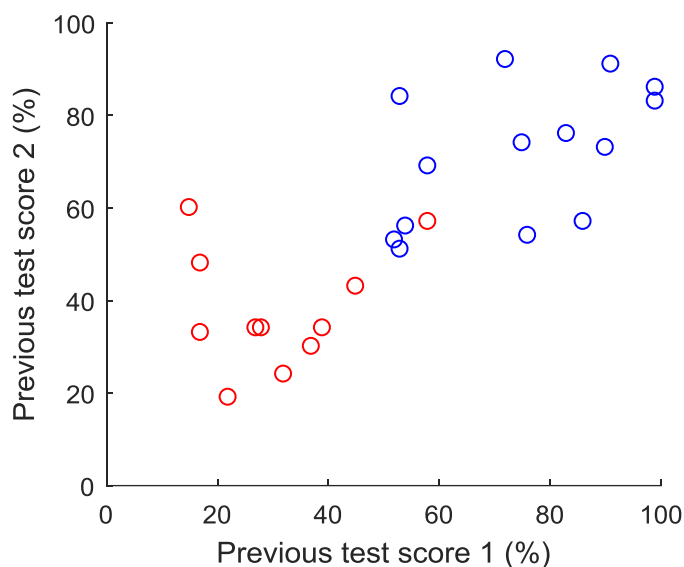

Figure 14: The blue circles represent students who pass the test. The red circles represent students who fail the test. The x and y axes are the parameters, in this case previous test scores. There is overlap between the blue and red circles, thus these data are not linearly separable.

Given that I don't expect most real-world data to be perfectly linearly separable it is necessary to employ kernel functions. Kernel functions represent the distance between data points in higher dimensional space. They essentially take the parameters used to map the data points and combine them to form additional parameters. The goal of this process is to take the data points that aren't linearly separable and map them into a higher dimensional space where

they are (167). This allows the computation of a hyperplane in data sets that are not linearly separable.

<u>Neural Networks</u>

Neural networks come in many shapes and sizes, the one used in this study is a simple single layer feed forward neural network (168). Neural networks have input, hidden and output layers. In each layer there are a number of 'neurons' each of these neurons has a connection to every single one of the neurons in the subsequent and previous layer. The number of neurons in the input layer is equal to the number of parameters in the data, the hidden layer has as many neurons as the experimenter chooses, and the output layer has as many neurons as there are classifications. In the test pass example, this would mean that both the input layer and output layer would have 2 neurons, due to there being two parameters (two past test results) and two classifications (pass or fail).

The strength of each of the connections between neurons are decided by the 'weights' and 'biases'. Essentially, the weights and biases are just co-efficients in a regression equation. The weights and bias initially are randomly selected, but as training data is fed into the neural network the weight and biases are changed as more information is made available and the network learns. This is achieved using an algorithm called back-propagation. Back-propagation works by changing the weights and bias in response to how correctly the neural network classifies the training data. The training data is fed into the neural network multiple times and on each iteration the weights and biases are modified by back propagation to increase the accuracy of the output (169). Unlike SVMs, NNs use all of the data provided to inform the weights and biases. NNs function best when the n-number is significantly higher than the number of parameters. In the test pass example this would mean a significantly higher number of students relative to past test results. Conversely SVMs work equally well, regardless of this ratio, for this reason SVMs can perform significantly better in some contexts.

When test data is fed into the network, the size of the output of each of the input neurons is influenced by the size of each of the parameters of the test data. The output of each of the input neurons is fed forward into every neuron in the

hidden layer, which subsequently is fed forward into the output layer. The output provided by the output layer is calculated based on the input received from the hidden layer. The input received by the output layer is therefore influenced by the weights and biases, in the hidden layer, calculated by the training algorithm and the parameters of the test data.

The output layer provides a number between 1 and 0 for each classification, and the sum of all the outputs is equal to one. The values provided by the output neurons give the relative probability that each output is the correct classification based on the training data provided. In the test pass scenario, it would provide two numbers, one giving the probability of a pass, and one giving the probability of a fail. The sum of these two numbers would be equal to one.

<u>Binary Logistic regression</u>

The binary logistic regression provides probabilistic prediction of categories. It does this by generating an s-shaped graph for each individual parameter. Training data for each of the parameters are placed on a graph based on the category they represent and the value they have for the corresponding parameter. Their position on the graph dictates the shape of the s-shaped curve which is subsequently used to confer the probability of the unknown test data belonging to one of two output categories.

The probabilistic outputs for each individual parameter are combined to generate an overall probability of the outcome for each test example. Logistic regressions are often inaccurate in situations where there is a large ratio of predictor parameters to observations and if parameters are highly correlated (170).
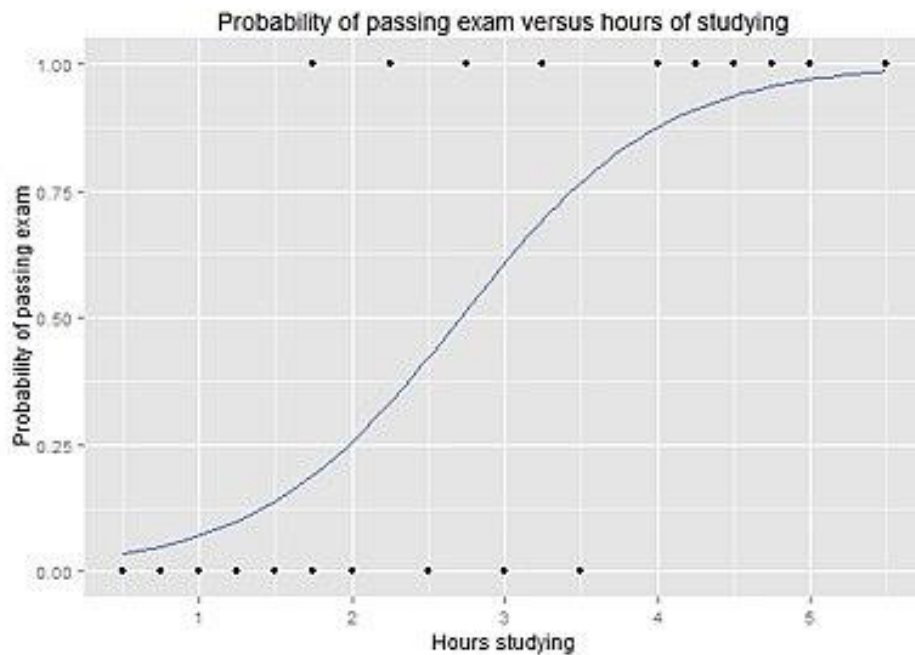
Figure 15: An example of what a logistic regression curve for an individual parameter looks like. Each of the points represent a different observation. Observations of students who passed the exam are placed at the top and observations of students who did not pass are placed at the bottom. Their position along the x-axis is relative to their value for the corresponding parameter. Test data have a probabilistic output given by the y-axis dependent on their parameter value. In this example 3 hours of studying would relate to an approximately 59% pass probability (171).

Methods

Data input

Data acquisition, image processing and peak parameter extraction were performed as detailed in chapter two. Again, input data were limited to experiments of more than one hundred and ninety-seven frames in length. Kainic acid was excluded from analyses due to not being present in measurable quantities within the animals (see table 1). For these analyses PCA was not performed in order to keep the number of parameters identical in both the training and test sets. Outlier removal was performed in an identical manner to that describe in chapter two with the exception that it was performed directly on the vectors of peak parameters as opposed to the principal component scores. The input data were therefore a vector of peak parameters for all the experiments remaining after outlier detection was performed. Each experiment is therefore a vector of three hundred and seventy-eight peak parameter

parameters consisting of the standard deviation and mean of each peak
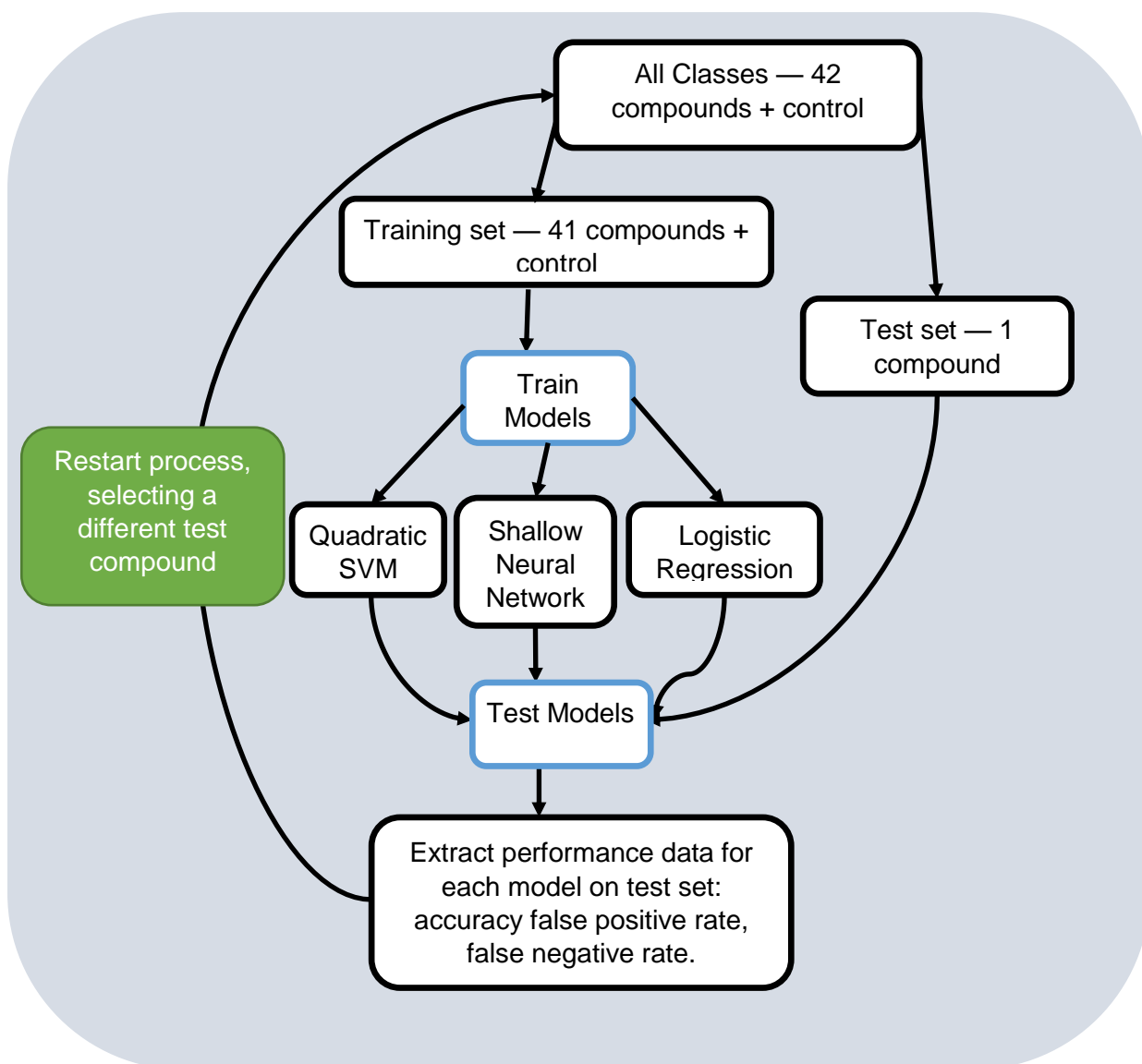parameter for each ROI.

*Figure 16: Schematic outlining the leave one out approach used for testing each model's
performance. Allows all of the classes to be tested on each type of model.*

Training and test data – For informing model selection.

Initially, training and test data input was selected in a leave one out training and
testing algorithm (see figure 16). Each model was trained on all of the data
except one compound and then subsequently each of the trained models was
tested using the compound that was left out of the original training data set. This
was done in order to assess how well each trained model performs on each
compound, when the model has been trained entirely independently of that

compound. Once a model has been trained and tested using the 'left out' compound, the model is discarded. A new model is then trained using a setup in which a different compound is left out and subsequently used to test the trained model. The performance information generated for each model, in the form of its sensitivity, specificity and accuracy was used to gain a preliminary understanding of which model would be the most appropriate choice for the data. Sensitivity was calculated as the percentage of seizurogenic compounds classified as seizurogenic, specificity was the percentage of non-seizurogenic compounds classified as non-seizurogenic and accuracy was the overall proportion of compounds correctly classified.

Given the large number of classification models available, this approach was utilised to inform the choice of model to use on randomised training and test sets. Moreover, the 'leave one out' approach allowed machine learning algorithms to be tested in the situation where a maximum amount of training data is available. Note that I expect to significantly expand on the amount of data available in the future.

Training and test data - For model generation.

After using the leave one out approach, to compare the efficacy of the different models for categorising the data, a randomised approach was utilised in order to provide an indication of the possible generalisability of the selected model. In this stage, the same data were split in two. Non-seizurogenic compounds were split: half into a training set and half into a test set. Seizurogenic compounds were also split: half into a training set and half into a test set.

Training model support vector machines

One of the models trained and tested was a support vector machine. This model was built using the default settings in MATLABs inbuilt classification learner toolbox. A quadratic support vector machine was trained by first mapping the experiments into multidimensional feature space based on their parameter values. A line, or 'hyperplane' was then placed between the experiments, in the multidimensional space, that fell in the space that maximised the distance between the two classifications: seizurogenic and non-seizurogenic. Hyperplane generation was achieved using a second polynomial kernel function. As mentioned previously kernel functions are used to help find

the hyperplane in data that isn't easily separable. In the internal model building algorithm, the SVM was tested for overfitting using a k-fold cross-validation algorithm with its own, internal hold out group. During the cross validation process data was randomly partitioned into five equal sized sub samples, each of which was used as the validation set for each iteration of the cross-validation process. This allowed the hyperplane to be optimised, ensuring that the SVM will perform accurately on new data sets.

Training model neural network

A two-layer feed-forward neural network was trained, with sigmoid hidden and softmax output neurons. The sigmoid function in the hidden layer neurons takes the weight and bias produced when training a neural network and coverts their output into a continuous spectrum between 1 and 0. The sigmoid function therefore provides a continuous measure of output for each of the hidden layer neurons. Moreover, the sigmoid function has an easily calculable derivative that can be used by the back-propagation algorithm to inform the changes in the weights and biases of the neural network that ultimately dictates its outputs.

The softmax function receives outputs from all the neurons in the hidden layer. It then takes the output of the hidden layer of the neural network and puts it into a vector of values between zero and one, the sum of which add up to one. The idea behind this is that each element of the vector represents a corresponding output and the size of the number is relative to the probability that the output is true. The softmax output neurons are therefore used to represent a categorical distribution of the output of the neural network.

The model was trained with scaled conjugate backpropagation in a supervised learning process. There were five hundred neurons in the hidden layer. Although this large number of neurons could potentially lead to overfitting, the cross-validation steps in the training algorithm are designed to counter this. As such some neurons will have little to no input on the output the model but as this classification procedure is not computationally heavy it will have little impact on the overall procedure.

Overfitting was prevented in the training algorithm by randomly splitting the training data into training, test and validation sets. The training algorithm used the validation set to test the performance of the model against unseen data to

ensure the models generalisability to new data. When the validation set error begins to rise, indicating overfitting, the weights and biases of the model were saved for when the validation set error is at its minimum. The test set error was not used during training however it was used to scrutinize varied models. The training, test and validation sets were selected by randomly splitting the training data in a ratio of 0.7:0.15:0.15 respectively. The performance of the network was tested using cross-entropy.

Training Logistic regression model

A generalised linear model was fitted to all of the training data. A 'logit' link function was used in order to generate probabilities for each of the classifications. Similar to the support vector machine the training algorithm involved k-fold cross validation, partitioning the training data into five equally sized random 'k-folds' which were used to help prevent the model from being subject to overfitting.

Calculating performance of models on independent test sets

Once each model was trained, the test data were run through each model. The following applies to both the leave one out training/test data and the randomly selected training/test data.

Once constructed, the models provided an output for each experiment or 'observation' provided in the test data. For the SVM the output was binary, i.e. for each experiment there was either an output of 'yes this experiment is seizurogenic' or 'no this experiment is not seizurogenic'. In the case of the NNs and logistic regression the output was probabilistic. The NN gives two values between zero and one for how likely each experiment is from each class – seizurogenic or non-seizurogenic. In the case of logistic regression, it provides a single value for how likely the test data were seizurogenic. For both sets of outputs, in order to quantify the performance of the classifier a threshold of a half was applied to decide which output the classifier had chosen. That is to say, if the probability output for an experiment to be seizurogenic was more than a half then that experiment was considered seizurogenic.

The performance of the models was calculated with respect to how effectively they could detect seizurogenic compounds. Compounds which had more than

half of their corresponding experiments classified as seizurogenic were considered to have been classified as seizurogenic overall. Compounds which had more than half of their corresponding experiments classified as non-seizurogenic were considered to have been classified as non-seizurogenic. Compounds which had exactly half of the corresponding experiments classified as (non)seizurogenic were considered to be undecided, and therefore were considered misclassified overall.

Seizurogenic compounds were considered 'positive' test results, thus true positives were seizurogenic compounds that were classified as seizurogenic. The sensitivity was calculated as the true positive rate and the specificity was calculated as the true negative rate. Accuracy was calculated as the percentage of compounds correctly classified overall.

<u>Leave one out vs random selection</u>

It is worth noting that the sensitivity, specificity and accuracy figures attained during the leave one out testing protocol are not for any specific model but rather an assessment of the utility of the algorithm. As previously mentioned, during the leave one out testing protocol multiple models are trained for each machine learning algorithm, this was done in order to assess how machine learning algorithms performed when provided with the maximum amount of training data. However, this protocol can give no information on the generalisability of each of the models as the amount of test data used on each model is insufficient.

In order to generate and better interrogate a usable model it was necessary to randomly split data in half. The specificity, sensitivity and accuracy figures generated using this approach were specific to a single model for each of the algorithms used.

Results

SVM performance

Using the leave-one out testing protocol the trained SVM classified most non-seizurogenic compounds reasonably well with 85% of the compounds that are non-seizurogenic having the majority of their experiments classified as non-seizurogenic (see table 6). There were two non-seizurogenic compounds that were misclassified as seizurogenic: nomifensine and ketoconazole.

Of the seizurogenic compounds used as test data approximately 30% were misclassified as non-seizurogenic. Of these picrotoxin, bicuculline and gabazine had poor uptake into the zebrafish CNS – with 1, 2 and 1.03% uptake respectively. Clozapine, maprotiline, SNC80, theophylline, amoxipine, cocaine and tetraethylammonium were also misclassified by their corresponding SVMs. Of those with good or unknown uptake into the zebrafish larvae: theophylline, maprotiline, SNC80, clozapine, cocaine, and amoxipine also clustered with non seizurogenic compounds during hierarchical clustering (see figure 10). Of the seizurogenic compounds that clustered with non-seizurogenic compounds during hierarchical clustering of peak profiles, amphetamine and bupropion are the only two that have been correctly classified as seizurogenic by their corresponding SVMs. There is a noticeable similarity between the hierarchical clustering outputs and the outputs of the SVMs which suggests the SVMs are

Support Vector Machines - Leave one out testing algorithm						
Correctly classified Seizurogenic Compounds	Correctly classified Non-seizurogenic compounds	Incorrectly classified Seizurogenic compounds	Incorrectly classified Non-seizurogenic compounds	Sensitivity	Specificity	Accuracy
4-Aminopyridine, Pentylenetetrazol, Strychnine, Bupropion, Donepezil, Bemegride, Bicuculine, Amitriptyline, Amphetamine, Bethanechol, Caffeine, CisACPD, Muscarine, Olanzapine, Physostigmine SB205607, Tacrine	4-Aminophenylsulfone, Clonidine, Mizolastine, Quinine, Yohimbine, Apomorphine, Cisplatin, Emetine, Morphine, Rolipram, Ketamine	Clozapine, Gabazine, Maprotiline, Picrotoxin, SNC80, Theophylline, Amoxipine, Cocaine Tetraethylammonium,	Ketoconozole, Nomifensine	69%	85%	74%

Table 6: Each of the classes named were used to test an SVM that was trained on all the remaining classes. From left to right column one: Seizurogenic compounds classified as seizurogenic. Column two: Non-seizurogenic compounds that were classified as non-seizurogenic. Column three: seizurogenic compounds that were classified as non-seizurogenic. Column four: non-seizurogenic compounds classified as seizurogenic. Column five: the sensitivity of this modelling approach for seizurogenic compounds. Column six: the specificity. Column seven: the overall accuracy.

effectively classifying the data and that perhaps it is the data driving the misclassification, due to issues with things like bioavailability and receptor

affinity, as opposed to the choice of machine learning algorithm. It is worth noting that misclassifications are based on if compounds are considered seizurogenic based on the literature from humans and mammalian models as opposed to what is known in fish. In this sense misclassifications rates are effected to some extent by limitations relating to the experimental model, as well as the algorithm in question.

Neural network performance

The trained NN classified most non-seizurogenic compounds reasonably well with 85% of the compounds that are non-seizurogenic having the majority of their experiments classified as such (see table 7). The two compounds that were misclassified were nomifensine and mizolastine. Nomifensine was also misclassified by the SVM classifier.

Of the seizurogenic compounds used as test data approximately 66% were classified correctly by their corresponding NN. NNs were therefore marginally less sensitive for detecting seizurogenic compounds, than SVMs, when tested using the leave one out approach. Of the compounds misclassified by the NNs, strychnine, muscarine, bupropion and bemegride were correctly classified by SVMs. However, the NN approach correctly classified tetraethylammonium, clozapine and amoxipine where the SVM did not. Both sets of models

Neural Network Performance - Leave one out testing algorithm						
Correctly classified Seizurogenic Compounds	Correctly classified Non-seizurogenic compounds	Incorrectly classified Seizurogenic compounds	Incorrectly classified Non-seizurogenic compounds	Sensitivity	Specificity	Accuracy
4-Aminopyridine, Pentylenetetrazol, Pilocarpine, Aminophylline, Donepezil, NMDA, Bicuculine, Clozapine, Amitriptyline, Amoxipine, Amphetamine, Bethanechol, Caffeine, CisACPD, Olanzapine, Physostigmine SB205607, Tacrine, Tetraethylammonium	4-Aminophenylsulfone, Clonidine, Quinine, Yohimbine, Apomorphine, Cisplatin, Emetine, Morphine, Rolipram, Ketamine, Ketoconozole	Strychnine, Bupropion, Bemegride, Gabazine, Maprotiline, Picrotoxin, SNC80, Theophylline, Cocaine, Muscarine	Mizolastine, Nomifensine	66%	85%	71%

Table 7: Each of the classes named were used to test an NN that was trained on all the remaining classes. From left to right column one: Seizurogenic compounds classified as seizurogenic. Column two: Non-seizurogenic compounds that were classified as non-seizurogenic. Column three: seizurogenic compounds that were classified as non-seizurogenic. Column four: non-seizurogenic compounds classified as seizurogenic. Column five: the sensitivity of this modelling approach for seizurogenic compounds. Column six: the specificity. Column seven: the overall accuracy.

misclassified gabazine, maprotiline, picrotoxin, SNC80 and theophylline.

The NNs had a classification accuracy of 71% when using the leave one out approach to test the efficacy of the model, marginally less than the SVM classifier at 74%.

Logistic regression performance

The logistic regression classifier was the poorest performing classifier on non-seizurogenic compounds, classifying 77% of non-seizurogenic compounds correctly (see table 8). This equates to one additional misclassification relative to the SVM and NN approaches. Again, nomifensine is misclassified, as is mizolastine. 4-aminophenylsulfone is also misclassified despite being correctly classified by the other two approaches.

Of the seizurogenic compounds approximately 55% were classified as seizurogenic by their corresponding logistic regression. This is a significant decrease in the classification rate of seizurogenic compounds compared to the other two approaches. Of the compounds misclassified pilocarpine, bicuculline, amitriptyline and bethanechol were only misclassified by logistic regression and were correctly classified by the other classifiers.

Logistic Regression - Leave one out testing algorithm						
Correctly classified Seizurogenic Compounds	Correctly classified Non-seizurogenic compounds	Incorrectly classified Seizurogenic compounds	Incorrectly classified Non-seizurogenic compounds	Sensitivity	Specificity	Accuracy
4-Aminopyridine, Pentylenetetrazol, Strychnine, Aminophylline, Donepezil, NMDA, Amoxipine, Amphetamine, Caffeine, CisACPD, Muscarine, Olanzapine, Physostigmine SB205607, Tacrine, Tetraethylammonium,	Clonidine, Quinine, Yohimbine, Apomorphine, Cisplatin, Emetine, Morphine, Rolipram, Ketamine, Ketoconozole,	Pilocarpine, Bupropion, Bemegride, Bicuculine, Clozapine, Gabazine, Maprotiline, Picrotoxin, SNC80, Theophylline, Amitriptyline, Bethanechol, Cocaine,	4-Aminophenylsulfone, Mizolastine, Nomifensine,	55%	77%	62%

Table 8: Each of the classes named were used to test a logistic regression that was trained on all the remaining classes. From left to right column one: Seizurogenic compounds classified as seizurogenic. Column two: Non-seizurogenic compounds that were classified as non-seizurogenic. Column three: seizurogenic compounds that were classified as non-seizurogenic. Column four: non-seizurogenic compounds classified as seizurogenic. Column five: the sensitivity of this modelling approach for seizurogenic compounds. Column six: the specificity. Column seven: the overall accuracy.

The overall accuracy of the logistic regressions was significantly worse than the other two classifiers at 62%, a 9% drop in accuracy compared to NNs and 11% relative to SVMs.

Support Vector Machine - Randomly Assigned Test Set						
Correctly classified Seizurogenic compounds	Correctly classified Non-seizurogenic compounds	Incorrectly classified Seizurogenic compounds	Incorrectly classified Non-seizurogenic compounds	Sensitivity	Specificity	Accuracy
Aminophylline, Bupropion, Donepezil, NMDA, Bicuculine, Clozapine, Picrotoxin, SNC80, Cocaine, Tacrine	Cisplatin, Morphine, Quinine, Apomorphine, Ketamine	Strychnine, Bethanechol, CisACPD, Tetraethylammonium	Rolipram	67%	83%	75%

Neural Network Performance - Randomly Assigned Test Set						
Correctly classified Seizurogenic Compounds	Correctly classified Non-seizurogenic compounds	Incorrectly classified Seizurogenic compounds	Incorrectly classified Non-seizurogenic compounds	Sensitivity	Specificity	Accuracy
Strychnine, Aminophylline, Bupropion, Donepezil, NMDA, Bicuculine, Clozapine, Picrotoxin, SNC80, Bethanechol, CisACPD, Cocaine, Tacrine, Tetraethylammonium	Quinine, Apomorphine, Ketamine		Cisplatin, Morphine, Rolipram	100%	50%	85%

Table 9: Top: Classifications of randomly selected test compounds by SVM. Bottom: Classification of randomly selected test compounds by NN. From left to right column one: Seizurogenic compounds classified as seizurogenic. Column two: Non-seizurogenic compounds that were classified as non-seizurogenic. Column three: seizurogenic compounds that were classified as non-seizurogenic. Column four: non-seizurogenic compounds classified as seizurogenic. Column five: the sensitivity of this modelling approach for seizurogenic compounds. Column six: the specificity. Column seven: the overall accuracy.

SVM and NN performance – Using Randomly selected test and training sets

SVMs and NNs had similar performance when tested using the leave one out approach. As such both were subsequently tested using randomly assigned test and training sets. Both algorithms performed well overall with the SVM and NN achieving 75% and 85% accuracy respectively.

The NN appeared to be the most sensitive approach, correctly classifying all seizurogenic compounds correctly. However, it performed poorly on non-seizurogenic compounds achieving only 50% specificity. This means the model could be expected to classify non-seizurogenic compounds as seizurogenic 50% of the time. It is worth noting that although it achieved a very high accuracy of 85% this is not truly reflective of the model's performance as there is a

greater number of seizurogenic compounds than non-seizurogenic compounds. Were there as many non-seizurogenic compounds in the test set as seizurogenic compounds, and the sensitivity and specificity remained the same, the accuracy would drop due to its poor specificity.

The SVM achieved a sensitivity of 67% and a specificity of 83%, and thus had much higher specificity than the neural network while sacrificing a significant amount of sensitivity.

<center>Discussion</center>

Nomifensine was consistently misclassified across all classifiers tested, when using the leave one out approach. Nomifensine misclassification was unsurprising given that Nomifensine is similar in pharmacology to a host of compounds, such as cocaine, amphetamine and bupropion that are classified as 'seizurogenic' in the training data. Moreover, Nomifensines peak profile clusters well with these compounds during hierarchical clustering (see figure 10). Interestingly, when training and testing the SVM with randomly selected compounds, nomifensine was correctly classified, despite being incorrectly classified during leave one out. One possible explanation for this, is that cocaine wasn't used in the training set, as it was randomly selected for the test set. Given Nomifensines, close association with cocaine during hierarchical clustering (see figure 10) it is possible cocaine's presence in the training data was driving Nomifensines misclassification. This highlights the importance of having representative training data, as even a small amount of data can change how the model classifies. This is of particular relevance in the case of SVMs, because even a single data point can change the position of a hyperplane. Comparatively, NNs should be more resilient to small changes in training data.

Seizurogenic compounds that were misclassified across all the models, during the leave one out testing algorithm, included cocaine, gabazine, maprotiline, picrotoxin, SNC80 and theophylline. Of these gabazine and picrotoxin have extremely low uptake with both averaging less than 2.1% uptake (see table 1). It is therefore likely that these compounds don't reach the concentrations required to have substantive seizurogenic effects. Similarly, maprotilines uptake is low, especially in the context of the other mixed mechanism antidepressants and thus it is possible that maprotiline doesn't reach adequate concentrations to

induce seizure. In the case of SNC80, as previously discussed, it has been establish that SNC80 has poor affinity for zebrafish delta opioid receptors (119). As such its seizurogenic mechanism is unlikely to be reflected in the animal model these analyses are based on.

All of the compounds consistently misclassified, across the different models tested, clustered with non-seizurogenic compounds in the hierarchical clustering analyses (see figure 10). It seems that a number of the compounds understood to be seizurogenic in the literature are not representative of seizurogenesis in this model. This could be due to either: not being present in high enough concentrations in zebrafish to elicit a seizurogenic profile or having poor ligand receptor affinity. Better training data selection could therefore improve the specificity of this approach, but all testing must be used in conjunction with compound uptake data which appears to be a specific issue in this experimental method. Experimental, modifications could be made such as injecting compounds directly into the zebrafish, a subject discussed in greater detail in the subsequent chapter.

Notably, the experimental approach used to generate the data used a variety of different concentrations of compounds. It is therefore possible that one or two of the concentrations used were not high enough to induce seizure. For example, some compounds may be seizurogenic only at the highest concentration and, despite correct classification of individual experiments, be considered misclassified overall.

For the classification purposes described, SVMs and NNs perform almost identically but both performed better on these data than a logistic regression. Logistic regression is ill-suited to the data because a lot of the parameters are likely to be highly correlated as some regions of interest can behave very similarly in terms of their peak dynamics. In addition, it is suggested that logistic regression is only effective when there is a high ratio of experiments to parameters and in this case there is not (170).

During the leave one out test protocol, the SVM and NN algorithms correctly classified seizurogenic compounds 70% of the time. Moreover, both algorithms correctly classified non-seizurogenic compounds 85% of the time. Given that

both algorithms were similarly effective both were selected for further testing using randomly selected data sets.

When trained and tested using the random selection protocol the SVM accurately identified 83% of the non-seizurogenic compounds it was tested with, however managed just over 67% sensitivity for seizurogenic compounds.

Despite the relatively poor performance of the SVM in terms of accurately identifying seizurogenic compounds it did achieve very high specificity. Given that the specificity was high, the final model could to some extent detect seizurogenic compounds based on their peak parameters albeit with a relatively high false negative rate.

The NN achieved significantly higher overall accuracy than the SVM due to its very high sensitivity. However, it had low specificity which must be taken to account if using this model for identification of seizurogenic compounds. The low specificity and very high sensitivity are likely symptoms of overfitting possibly occurring due to the relatively small amount of training data. Neural networks can suffer from overfitting when used on lower sample size, higher dimensional data (172).

Moreover, because there was a smaller amount of non-seizurogenic compounds available the high overall accuracy is perhaps not reflective of the true efficacy of the models. The imbalance in the test data causes the accuracy to be weighted toward the sensitivity of the model and less affected by the specificity. Were these analyses to be repeated it would be important to have an equal number of seizurogenic and non-seizurogenic compounds in order for the models to learn features of both seizurogenic and non-seizurogenic compounds.

Given that the experimental and analytical approaches used here are still very preliminary stages there is a huge amount of progress that could still be made. It is clear, based on these analyses, that machine learning approaches could play a key role in its future development.

Improvements could made on a number of fronts, including data acquisition, training data selection and fine tuning of modelling algorithms. Here I have used the inbuilt matlab classification learner tools box to highlight the potential

efficacy of these kind of approaches. However, purpose built, and custom designed algorithms could yield far superior results. Techniques such as unsupervised pre-training of neural networks have been shown to significantly improve neural network performance (173). Techniques, such as unsupervised learning of deep belief networks, can be used as feature detectors. Feature detectors can be used to initialise the weights and biases of neural networks, prior to supervised training, to provide powerful classification algorithms. This approach has already been utilised in the area of drug discovery to greatly improve compound selectivity prediction (174).

Moreover, machine learning techniques have been utilised in the clinic for epilepsy surgery outcome prediction (175) in addition to seizure prediction and identification (176). Efforts to create seizure prediction models based on using electroencephalography combined with PCA and SVMs resulted in an algorithm with 100% sensitivity and over 80% specificity (176). A similar study using PCA and SVM achieved similarly impressive results (177). In these cases, EEGs were recorded in human subjects and all signals were either from patients known to be suffering a seizure or from healthy controls (176,177). By comparison the study described here could not guarantee seizure phenotypes were present in the fish representing the 'seizurogenic' training and test groups. Moreover, EEGs have the advantage of extremely high temporal resolution, future work on detection of seizurogenic compounds using zebrafish could benefit from a comparison of local field potential and calcium imaging peak analyses as input data for machine learning algorithms.

Previous studies have utilised local field potentials recorded in transgenic zebrafish with seizure phenotypes to identify anti-epileptic drugs with high accuracy (23). Similar analyses could be repeated in zebrafish in addition to calcium imaging and both data sets could be combined with the aim of identifying potentially seizurogenic compounds.

Chapter 4: Overall discussion and conclusion

The overall aim of this thesis was to highlight the utility of 4-dimensional functional profiling of drugs in zebrafish. It is clear from the analyses performed that drugs elicit spatio-temporal patterns of activation in zebrafish specific to

their pharmacological properties. It is also clear that these specific spatio-temporal patterns can be exploited to make elucidations about the properties of drugs and that machine learning approaches are well suited for such purposes.

The clustering analyses showed clear congregations of compounds with similar pharmacodynamics, in particular with regards to dopamine and serotonin active compounds. This is of particular relevance to anti-depressant drug discovery, an area in which novel therapeutics are desperately needed (58). On this basis, it seems that compounds with specific pharmacodynamics do seem to elicit specific patterns of activation in the zebrafish brain. However, not all compounds clustered in a pharmacologically meaningful way, either due to bioavailability issues, lack of receptor affinity, or perhaps simply because their particular pharmacodynamics did not elicit distinctive patterns. Moreover, it should be noted that the cluster analyses do not provide a metric that states whether certain compounds are significantly different or not, due to not using any form of statistical hypothesis testing. This is due in part to the fact that statistical hypothesis testing on data with a large number of experimental groups, such as the data in this thesis, encounters problems related to multiple comparisons (178).

As a result it could be argued that it is difficult to draw firm conclusions from the clustering dendrogram with regards to the significance of the similarities and differences between compounds. Moreover, as clustering was performed on the average principal components across the experiments of a compound, the clustering doesn't account for variability between experiments of the same compound, which was at times significant.

Despite the limitations related to clustering, this approach allows users to gain an understanding of the data set and illustrates how the different experimental groups relate to one another. This is highly useful for step and in this case showed that there were some strong similarities between compounds with similar pharmacologies, relative to other pharmacologically distinct compounds. And it is this concept that underpinned the rationale behind the application of machine learning to the data set.

It is clear that machine learning techniques can be applied effectively to 4-dimensional zebrafish brain activity data for the identification of potentially

seizurogenic compounds. The tests showed fairly high accuracy, sensitivity and specificity in the classifiers chosen for testing. The analyses showed that either SVMs or NNs could prove useful for compound classification purposes, however neural networks could benefit from a larger data set. Classifiers, such as the machine learning techniques described here, have a huge amount of potential utility in the pharmaceutical industry due to their ease of application to new data sets and new classification problems. It is clear from the analyses presented here, that machine learning tools provide a powerful and replicable approach for making elucidations about drug properties.

However, in the context of this study, there are a number of limitations with regards to the machine learning approaches that should be reiterated. SVMs are sensitive to anomalous data, so if one compound classified as seizurogenic in the training data is actually non-seizurogenic, due to experimental constraints, this is likely to have a significant effect on the how the algorithm classifies test data. Moreover, neural networks are likely to perform poorly if not enough data is available because the ratio of the number of experiments relative to the number of parameters needs to be reasonably high (179).

When designing the analytical approaches used during this project there were a number of decisions to make. These different analytical choices all have different advantages and disadvantages worth noting. As discussed, the spatial resolution of light sheet microscopy is very high, to reduce the computational intensity of the analyses and to put analyses into a more biologically relevant context, ROIs were defined using a standardised 3D anatomical map of the larval zebrafish brain (4,42). The atlas in question utilised 6dpf zebrafish larvae to generate the three-dimensional anatomical atlas utilised for ROI selection. The zebrafish larva used for this study were 4dpf and as such could have a considerably different anatomical make up to the zebrafish used to generate this map. This could cause voxels from distinct functional areas to be averaged together with the results being an output of time series that misses large chunks of informative data.

Alternatives, such as spatially constrained spectral clustering, an algorithm developed for use in functional MRI, utilises the functional similarities of voxels to generate ROIs, thus generating ROIs based on the functional make up of individual fish (46). However, such data driven ROI selection approaches also

have problems, such as lack of consistency in ROI selection making between subject valid comparisons of ROIs difficult.

Another alternative could simply be to use individual voxels however the varying sizes of fish may cause the number of voxels present within the animal itself to vary, again making between subject comparisons difficult. This is therefore an ongoing challenge.

An additional important analytical decision worth noting was the choice to use Pearson's correlation coefficient for pairwise comparisons of ROIs during functional connectivity analysis. As mentioned Pearsons correlation coefficient was chosen because it is simple to understand and implement while considering the amplitude and phase of any underlying oscillatory activity.

However other approaches, such as cross correlation, coherence and grangers causality, may be more suitable in future if experimental approaches or analytical approaches are modified.

Coherence is a measure of the level of similarity in frequency content and has been utilised in electrophysiology and magnetoencephalography due to its capability to deal with very high temporal frequency data (180). This particular approach is not applicable to this study as the temporal resolution was simply not high enough. Were there a change in experiment approaches that caused a significant increase in temporal resolution coherence would be a potentially useful tool.

Lagged cross correlation is also a useful measure of similarity as it can be utilised to take into the 'lag' between signals. This measure is useful in situations where there is a slight time delay between signals as it can be used to account for this lag when performing analyses. This approach was tested during the analyses conducted however it was deemed that the measurements of calcium transients in this study were too slow for lag correction to have be of any real utility. As such Pearson's correlation coefficient was deemed a preferable option. Were a different transgenic fish line used, with faster measures of calcium transients, cross correlation may provide a superior measure of functional connectivity. The calcium transient utilised for this study, GCaMP6s is considered the most sensitive for changes in calcium levels however GCaMP6f has a considerably higher rise and decay time allowing for

the detection of calcium transients with higher temporal resolution (34). Given that the temporal resolution of LSM is low it made sense in this context to utilise the slower more sensitive sensor GCaMP6s.

The combination of LSM and zebrafish with genetically encoded calcium indicators is relatively new experimental approach. The advantages this method allows experimenters, wishing to study whole brain network dynamics, are significant. LSM allows imaging of the entire zebrafish brain at sub-single cell resolution and with significantly improved temporal resolution over equivalent scanning microscopes (181). However, despite this advantage, there are a number of limitations to be considered. Light sheet microscopes still have low temporal resolution, relative to electrophysiological recordings, particularly when imaging a large number of z-planes, typically taking just under two seconds for ten z-stacks (4). This temporal resolution is significantly lower than the potential temporal resolution available using GCaMP sensors, which are capable of temporal resolution in the millisecond range (182). However the calcium sensor for this study (GCaMP6s) is a slower acting calcium sensor due to the slow rise and fall rate of the dye (183), which means LSM has more than adequate temporal resolution.

Therefore, a potential alteration to the experimental setup, worth investigating is the temporal resolution of LSM. As mentioned, when utilising LSM it is a trade-off between the volumetric coverage achieved and the temporal resolution. In future, studies could focus on more specific brain regions at significantly higher temporal resolution, utilising GCaMP6f sensors and measuring functional connectivity with cross correlation to provide more complete temporal coverage of specific regions.

While the temporal resolution of LSM is relatively low, it has extremely high spatial resolution and can been used for imaging at cellular resolution (181). High spatial resolution has the potential to be particularly useful when attempting to classify neuroactive compounds based on brain activity because often specific neurotransmitters are found in specific brain areas (113,116,184,185). This seems to be particularly true of the serotonergic and dopaminergic systems and may explain why serotonergic (113,116) and dopaminergic compounds (116) clustered in close proximity during analyses (see figures 8 and 9).

In this study there were a number of issues pertaining to bioavailability. This is because, fish were exposed to compounds dissolved in water and in order for compounds to make it into the zebrafish CNS they need to be absorbed through external tissues which may be impossible or difficult in the case of certain compounds. This meant a number of compounds being absorbed only in small quantities (see table 1). In addition to absorption issues, it is clear drugs that are insoluble are going to be very difficult to test using this approach. It is therefore necessary, if using this approach as an assay, to perform bioavailability assays to ensure compound performance is related to its pharmacological properties, as opposed to its absorption into the zebrafish CNS. Alternatively, it is possible to inject compounds into the cardinal vein of zebrafish embryos which guarantees uptake into the zebrafish body and thus maybe be a preferable technique for this assay in future.

Among the considerable advantages, of using zebrafish, is that they have a blood brain barrier homologous with higher invertebrates (186). This poses a significant benefit over *in vitro* models and *in vivo* models where drugs are injected though a cannula directly into brain tissue, because it provides a model that gives insight into the relevant pharmacokinetics of compounds in addition to the pharmacodynamics. While worth mentioning, this is not hugely relevant to this study as the zebrafish blood brain barrier isn't fully developed at 4dpf and therefore isn't likely to provide useful insight into pharmacokinetics at this developmental stage (187). It is possible future studies could use zebrafish at later developmental stages that do take advantage of this useful aspect of zebrafish physiology. However there are a number of problems associated with using zebrafish at latter stages chief among which are the ethical considerations of using neuromuscular blocker in protected life stage fish. There are number of alternative methods for 'trapping' zebrafish for imaging including using hydrodynamic force to hold zebrafish in specifically designed microfluidic channels (35). This approach confers a number of other advantages including the ability to keep zebrafish in water (as opposed to agarose) and the ability to perform 'wash on wash off' compound exposure.

Another potential problem associated with using later stage zebrafish is the increased brain volume may cause issues with brain coverage. The obvious solution is to reduce magnification resulting in reduced resolution. This would be

of little consequence, in the *x* and *y* planes set up described here, as the brain is divided into anatomically defined regions during image analysis, as such the full resolution isn't utilised. However as the brain gets larger it requires more z-stacks (that is to say images taken in different z-planes) to provide adequate coverage and this reduces temporal resolution. Loss of temporal resolution would potentially cause recordings to miss some dynamic brain activity, as such it is unlikely that functional full brain coverage would be ideal past 7 day post fertilisation.

A notable issue, relevant to this study, is the poor affinity of SNC80 for zebrafish delta opioid receptors (119). This issue is chiefly notable because it provides an example of where brain receptors in zebrafish are no longer adequately homologous with humans to bind compounds with acute effects on human tissues. This would be of particular concern, in this case, as SNC80 is a compound with well characterised seizurogenic properties (156). Given that SNC80 was among the compounds most consistently misclassified or misplaced during analyses, it is plausible other compounds may produce false negative results for seizurogenic properties, not due to flaws in the analytical or experimental method, but due to the physiology of the zebrafish itself.

However, it should be acknowledged that SNC80 is in contradiction to the broader trend. Zebrafish are generally considered to be excellent models for studying the central opioid and endocannabinoid systems and have just over 60% pairwise alignment homology scores for both the opioid and endocannabinoid systems when compared with humans (188). Moreover, zebrafish have shown to be highly sensitive to antidepressants (189), a fact corroborated by the analyses describe here (see figures 8 and 9). More generally speaking, zebrafish have orthologous genes for approximately 70% of the human genome; zebrafish also have equivalent genes for 84% of human genes associated with disease (190).

The high level of homology with higher invertebrates and genetic malleability of zebrafish means there are many different transgenic zebrafish models of human disease including: depression (21), anxiety (191), epilepsy (192) and serotonin syndrome (193). Future work could utilise transgenic zebrafish models with spontaneous seizure phenotypes to train classification learners to recognise seizure and thus identify seizurogenic compounds. A similar approach could be

utilised using serotonin syndrome mutants to identify antidepressant compounds. The zebrafish as a model organism provides a huge number of potential future directions for this project.

Among the more prominent experimental limitations present in this study is the lack of a within experiment control. The zebrafish analysed here are relatively heterogeneous and the lack of drug free baseline period for each fish makes normalisation of experiments difficult. Lack of an adequate normalisation process makes between subject comparisons in heterogeneous data unreliable. This is of particular relevance with regards to differences in overall brain activity induced by compounds. Because control fish can vary significantly in their overall fluorescent intensity, identifying if compounds are increasing overall fluorescent intensity and thus brain activity is very difficult as there is no drug free baseline period for each fish with which to generate a relative measure of fluorescent intensity (and thus brain activity). The introduction of a drug free baseline period for each experiment would therefore allow better characterisation of the changes in overall brain activity induced by particular drug classes and would therefore add a whole new dimension of informative data into the analytical process. This could prove to be of particular relevance to the seizurogenic potential of compounds as PTZ, a well categorised seizurogenic compound (194), has been demonstrated to rapidly and significantly increase baseline activity (36).

Introducing a drug free baseline period has potential downsides. Some compounds utilised for these analyse require long exposure times. Long exposure times would require fish to be imaged for a very long time, extended period in agar tend to cause the health of fish to deteriorate which would introduce an unwanted confound into the analyses. As such it would be logistically difficult to perform analyses on compounds with long exposure times such as Rolipram, which requires a 24-hour exposure regimen (see table 1).

Despite the limitations, it is clear from the analyses that the experimental and analytical approaches described here are highly useful for profiling drugs. As discussed previously, there are a plethora of advantages to using zebrafish and as such the methodologies described here are highly applicable for use in industry. There are also a number of methodological alterations that have the potential to improve these approaches, both experimentally and analytically.

Overall, it is clear from the works described here, that the future is bright for the zebrafish in pharmaceutical research.

Conclusions

Transgenic zebrafish are a powerful tool for assessing the effects of drugs on the central nervous system. When combined with genetically encoded calcium sensors and light sheet microscopy, zebrafish provide a 4-dimensional model of brain networks with extremely high spatial resolution and modest temporal resolution.

Clustering analysis revealed that spatial patterns of activity as reflected by their spatially specific peak parameters and functional connectivity values are specific to the pharmacodynamics of certain compounds. It is clear that dopamine and serotonin active compounds in particular elicit specific spatio-temporal patterns of activity. Further analysis also showed strong clustering of seizurogenic compounds, in particular when clustered based ion their peak parameters. Further analysis showed that peak parameters can be used to train classification learners such as SVMs and NNs to produce models that can accurately discriminate between seizurogenic and non-seizurogenic compounds.

A number of experimental modifications could be utilised in future work, such as using GCaMP sensors with faster acting dyes (e.g. GCaMP6f), introducing within experiment controls and reducing the number of z-slices to increase temporal resolution (at the expense of brain coverage).

Future directions, in terms of analysis, could focus on utilising different ROI selection approaches, such as spatially constrained spectral clustering and using unsupervised learning algorithms, such as deep belief networks, for feature selection prior to neural network modelling. In addition, when choosing compounds for the training and testing of classification learners stringent thresholds for bioavailability and receptor affinity could be implemented that improve the performance of models.

Bibliography

1. Drinkenburg WHIM, Ruigt GSF, Ahnaou A. Pharmaco-EEG Studies in Animals: An Overview of Contemporary Translational Applications. Neuropsychobiology. 2016;72(3–4):151–64.

2. Kalueff A V., Stewart AM, Gerlai R. Zebrafish as an emerging model for studying complex brain disorders. Trends Pharmacol Sci [Internet]. 2014;35(2):63–75. Available from: http://dx.doi.org/10.1016/j.tips.2013.12.002

3. Stewart AM, Braubach O, Spitsbergen J, Gerlai R, Kalueff A V. Zebrafish models for translational neuroscience research: From tank to bedside. Trends Neurosci [Internet]. 2014;37(5):264–78. Available from: http://dx.doi.org/10.1016/j.tins.2014.02.011

4. Winter MJ, Windell D, Metz J, Matthews P, Pinion J, Brown JT, et al. 4-Dimensional Functional Profiling in the Convulsant-Treated Larval

Zebrafish Brain. Sci Rep [Internet]. 2017;7(1):1–16. Available from: http://dx.doi.org/10.1038/s41598-017-06646-6

5. Krijzer F, Koopman P, Olivier B. Classification of psychotropic drugs based on pharmaco- electrocorticographic studies in vigilance-controlled rats. Neuropsychobiology. 1993;28(3):122–37.

6. Mandema JW. Electroencephalogram effect measures and relationships between pharmacokinetics and pharmacodynamics of centrally acting drugs. LACDR, devision Pharmacol. 1991;23(3):191–215.

7. Mégevand P, Quairiaux C, Lascano AM, Kiss JZ, Michel CM. A mouse model for studying large-scale neuronal networks using EEG mapping techniques. Neuroimage. 2008;42(2):591–602.

8. Ahnaou A, Huysmans H, Jacobs T, Drinkenburg WHIM. Cortical EEG oscillations and network connectivity as efficacy indices for assessing drugs with cognition enhancing potential. Neuropharmacology [Internet]. 2014;86:362–77. Available from: http://dx.doi.org/10.1016/j.neuropharm.2014.08.015

9. Lee H. Genetically Engineered Mouse Models for Drug Development and Preclinical Trials. Biomol Ther (Seoul) [Internet]. 2014;22(4):267–74. Available from: http://www.biomolther.org/journal/DOIx.php?id=10.4062/biomolther.2014.074

10. Swinney DC, Anthony J. How were new medicines discovered? Nat Rev Drug Discov. 2011;10(7):507–19.

11. Sanchez PE, Zhu L, Verret L, Vossel KA, Orr AG, Cirrito JR, et al. Levetiracetam suppresses neuronal network dysfunction and reverses synaptic and cognitive deficits in an Alzheimer's disease model. Proc Natl Acad Sci [Internet]. 2012;109(42):E2895–903. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1121081109

12. UK HO-NS. Annual Statistics of Scientific Procedures on Living Animals [Internet]. UK, Home Office - National Statistics. 2012. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/212610/spanimals12.pdf

13. Kentucky U of. Reproductive Data for Mice. Presentation. 2015;

14. Balzekas I, Hernandez J, White J, Koh S. Confounding effect of EEG implantation surgery: Inadequacy of surgical control in a two hit model of temporal lobe epilepsy. Neurosci Lett [Internet]. 2016;622:30–6. Available from: http://dx.doi.org/10.1016/j.neulet.2016.04.033

15. Vezzani A, Granata T. Brain inflammation in epilepsy: experimental and clinical evidence. Epilepsia [Internet]. 2005;46(11):1724–43. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16302852

16. Zon LI, Peterson RT. In vivo drug discovery in the zebrafish. Nat Rev Drug Discov. 2005;4(1):35–44.

17. Reed B, Jennings M. Guidance on the housing and care of Zebrafish Danio rerio. Res Anim Dep Sci Group, RSPCA. 2011;1–27.

18. Strähle U, Scholz S, Geisler R, Greiner P, Hollert H, Rastegar S, et al. Zebrafish embryos as an alternative to animal experiments-A commentary on the definition of the onset of protected life stages in animal welfare regulations. Reprod Toxicol [Internet]. 2012;33(2):128–32. Available from: http://dx.doi.org/10.1016/j.reprotox.2011.06.121

19. Rogério dos Santos Alves; Alex Soares de Souza et all. Perspectives of zebrafish models of epilepsy: what, how and where next? Igarss 2014. 2014;(1):1–5.

20. Meshalkina DA, Kysil E V., Warnick JE, Demin KA, Kalueff A V. Adult zebrafish in CNS disease modeling: A tank that's half-full, not half-empty, and still filling. Lab Anim (NY) [Internet]. 2017;46(10):378–87. Available from: http://dx.doi.org/10.1038/laban.1345

21. Norton W, Bally-Cuif L. Adult zebrafish as a model organism for behavioural genetics. BMC Neurosci. 2010;11:90.

22. Baraban SC. Forebrain Electrophysiological Recording in Larval Zebrafish. 2013;(January):2–5.

23. Eimon PM, Ghannad-Rezaie M, De Rienzo G, Allalou A, Wu Y, Gao M, et

al. Brain activity patterns in high-throughput electrophysiology screen predict both drug efficacies and side effects. Nat Commun [Internet]. 2018;9(1). Available from: http://dx.doi.org/10.1038/s41467-017-02404-4

24.   Griffin A, Hamling KR, Knupp K, Hong SG, Lee LP, Baraban SC. Clemizole and modulators of serotonin signalling suppress seizures in Dravet syndrome. Brain. 2017;140(3):669–83.

25.   Cho S-J, Byun D, Nam T-S, Choi S-Y, Lee B-G, Kim M-K, et al. Zebrafish as an animal model in epilepsy studies with multichannel EEG recordings. Sci Rep [Internet]. 2017;7(1):3099. Available from: http://www.nature.com/articles/s41598-017-03482-6

26.   Hunyadi B, Siekierska A, Sourbron J, Copmans D, de Witte PAM. Automated analysis of brain activity for seizure detection in zebrafish models of epilepsy. J Neurosci Methods [Internet]. 2017;287:13–24. Available from: http://dx.doi.org/10.1016/j.jneumeth.2017.05.024

27.   Hunt RF, Hortopan GA, Gillespie A, Baraban SC. A novel zebrafish model of hyperthermia-induced seizures reveals a role for TRPV4 channels and NMDA-type glutamate receptors. Exp Neurol [Internet]. 2012;237(1):199–206. Available from: http://dx.doi.org/10.1016/j.expneurol.2012.06.013

28.   Feierstein CE, Portugues R, Orger MB. Seeing the whole picture: A comprehensive imaging approach to functional mapping of circuits in behaving zebrafish. Neuroscience [Internet]. 2015;296:26–38. Available from: http://dx.doi.org/10.1016/j.neuroscience.2014.11.046

29.   Yuste R, Katz LC. Control of postsynaptic Ca2+ influx in developing neocortex by excitatory and inhibitory neurotransmitters. Neuron. 1991;6(3):333–44.

30.   Kandel ER, Schwartz JH, Jessell TM. Principles of neural science [Internet]. Neurology. 1990. Available from: http://ftp.kermit-project.org/itc/gsas/g9600/2003/HenReadings/KandellCh3.pdf

31.   Nakai J, Ohkura M, Imoto K. A high signal-to-noise ca2+ probe composed of a single green fluorescent protein. Nat Biotechnol. 2001;19(2):137–41.

32.   Tian L, Hires SA, Mao T, Huber D, Chiappe ME, Chalasani SH, et al.

Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. Nat Methods [Internet]. 2009;6(12):875–81. Available from: http://dx.doi.org/10.1038/nmeth.1398

33.   Akerboom J, Chen T-W, Wardill TJ, Tian L, Marvin JS, Mutlu S, et al. Optimization of a GCaMP Calcium Indicator for Neural Activity Imaging. J Neurosci [Internet]. 2012;32(40):13819–40. Available from: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2601-12.2012

34.   Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature [Internet]. 2013;499(7458):295–300. Available from: http://www.nature.com/doifinder/10.1038/nature12354

35.   Lin X, Wang S, Yu X, Liu Z, Wang F, Li WT, et al. High-throughput mapping of brain-wide activity in awake and drug-responsive vertebrates. Lab Chip [Internet]. 2015;15(3):680–9. Available from: http://xlink.rsc.org/?DOI=C4LC01186D

36.   Turrini L, Fornetto C, Marchetto G, Müllenbroich MC, Tiso N, Vettori A, et al. Optical mapping of neuronal activity during seizures in zebrafish. Sci Rep. 2017;7(1):1–12.

37.   Nemoto T, Kawakami R, Hibi T, Iijima K, Otomo K. Two-photon excitation fluorescence microscopy and its application in functional connectomics. Microscopy [Internet]. 2015 Feb [cited 2018 Aug 25];64(1):9–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25540030

38.   Prilloff S, Fan J, Henrich-Noack P, Sabel BA. In vivo confocal neuroimaging (ICON): Non-invasive, functional imaging of the mammalian CNS with cellular resolution. Eur J Neurosci [Internet]. 2010 Feb [cited 2018 Aug 25];31(3):521–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20105236

39.   Girkin JM, Carvalho MT. The Light sheet microscopy revolution. J Opt. 2018;

40.   Vogt BA, Paxinos G. Cytoarchitecture of mouse and rat cingulate cortex with human homologies. Brain Struct Funct. 2014;219(1):185–92.

41.   Walker AE. A cytoarchitectural study of the prefrontal area of the macaque monkey. J Comp Neurol. 1940;73(1):59–86.

42.   Randlett O, Wee CL, Naumann EA, Nnaemeka O, Schoppik D, Fitzgerald JE, et al. Whole-brain activity mapping onto a zebrafish brain atlas. Nat Methods. 2015;12(11):1039–46.

43.   Katherine J. Turner, Thomas G. Bracewell TAH. Brain Development. 2013. 197-214 p.

44.   Butts CT. Revisiting the Foundations of Network Analysis. Science (80-) [Internet]. 2009;325(5939):414–6. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1171022

45.   Hagmann P, Kurant M, Gigandet X, Thiran P, Wedeen VJ, Meuli R, et al. Mapping human whole-brain structural networks with diffusion MRI. PLoS One. 2007;2(7).

46.   Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp. 2012;33(8):1914–28.

47.   Wen T, Hsieh S. Network-Based Analysis Reveals Functional Connectivity Related to Internet Addiction Tendency. Front Hum Neurosci [Internet]. 2016;10(February):1–13. Available from: http://journal.frontiersin.org/article/10.3389/fnhum.2016.00006

48.   Zalesky A, Fornito A, Bullmore ET. Network-based statistic: Identifying differences in brain networks. Neuroimage [Internet]. 2010;53(4):1197–207. Available from: http://dx.doi.org/10.1016/j.neuroimage.2010.06.041

49.   Xu T, Cullen KR, Mueller B, Schreiner MW, Lim KO, Schulz SC, et al. Network analysis of functional brain connectivity in borderline personality disorder using resting-state fMRI. NeuroImage Clin [Internet]. 2016;11:302–15. Available from: http://dx.doi.org/10.1016/j.nicl.2016.02.006

50.   Ph D, Karlsgodt KH, Ph D, John M, Ph D, Weissman N. Baseline Striatal Functional Connectivity as a Predictor of Response to Antipsychotic Drug Treatment. Am J Psychiatry. 2017;173(1):69–77.

51. Stoffers D, Bosboom JLW, Deijen JB, Wolters EC, Stam CJ, Berendse HW. Increased cortico-cortical functional connectivity in early-stage Parkinson's disease: An MEG study. Neuroimage. 2008;41(2):212–22.

52. Markowitz JC, Brown RP. Seizures with neuroleptics and antidepressants. Gen Hosp Psychiatry. 1987;9(2):135–41.

53. Wang CY, Hung CH, Lin CS, Lee HH, Yang CH, Jong YJ, et al. Differential alterations of GABAA receptor (α1, β2, γ2 subunit) expression and increased seizure susceptibility in rat offspring from morphine-addicted mothers: Beneficial effect of dextromethorphan. Neurosci Lett [Internet]. 2011;489(1):5–9. Available from: http://dx.doi.org/10.1016/j.neulet.2010.11.055

54. Carter Snead O. Noradrenergic mechanisms in γ-hydroxybutyrate-induced seizure activity. Eur J Pharmacol. 1987;136(1):103–8.

55. Barker-Haliski M, White HS. Glutamatergic mechanimms associated with seizures and epilepsy. Cold Spring Harb Perspect Med. 2015;5(8):a022863.

56. Fond G, Loundou A, Rabu C, Macgregor A, Lançon C, Brittner M, et al. Ketamine administration in depressive disorders: A systematic review and meta-analysis. Psychopharmacology (Berl). 2014;231(18):3663–76.

57. Laxer KD, Trinka E, Hirsch LJ, Cendes F, Lang J, Delanty N, et al. Epilepsy & Behavior The consequences of refractory epilepsy and its treatment. 2014;37:59–70.

58. Al-Harbi KS. Treatment-resistant depression: Therapeutic trends, challenges, and future directions. Patient Prefer Adherence. 2012;6:369–88.

59. Fig S, Table S, Note S. correspondEnce Light-sheet functional imaging in fictively behaving zebrafish Guided visual exploration of genomic stratifications in cancer. 2014;11(9):883–4.

60. Z Brain atlas [Internet]. [cited 2018 Jul 16]. Available from: https://engertlab.fas.harvard.edu/Z-Brain/#/home/

61. Piecoro LT, Wermeling DP, Schmitt FA, Ashford JW. Seizures in patients

receiving concomitant antimuscarinics and acetylcholinesterase inhibitor. Pharmacotherapy. 1998;18(5):1129–32.

62. Physostigmine | C15H21N3O2 - PubChem [Internet]. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/physostigmine#section=Pharmacology-and-Biochemistry

63. PubChem. Clozapine | C18H19ClN4 - PubChem [Internet]. 2018 [cited 2018 Jul 14]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/clozapine#section=Pharmacology-and-Biochemistry

64. Olanzapine [Internet]. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/4585#section=Pharmacology-and-Biochemistry

65. (+)-Bicuculline | C20H17NO6 - PubChem [Internet]. [cited 2018 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/bicuculline

66. Pentetrazol | C6H10N4 - PubChem [Internet]. [cited 2018 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/pentylenetetrazole#section=Top

67. Law RJ, Lightstone FC. Gaba receptor insecticide non-competitive antagonists may bind at allosteric modulator sites. Int J Neurosci [Internet]. 2008 Jan 7 [cited 2018 Jul 16];118(5):705–34. Available from: http://www.tandfonline.com/doi/full/10.1080/00207450701750216

68. Tutughamiarso M, Bolte M. 4-Ethyl-4-methylpiperidine-2,6-dione. Acta Crystallogr Sect E Struct Reports Online [Internet]. 2007 Dec 15 [cited 2018 Jul 16];63(12):o4743–o4743. Available from: http://scripts.iucr.org/cgi-bin/paper?S1600536807058436

69. Bemegride [Internet]. [cited 2018 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/2310

70. Ben-Ari Y. Kainate and Temporal Lobe Epilepsies: 3 decades of progress. Jasper's Basic Mech Epilepsies [Internet]. 2012 [cited 2018 Jul

16];(Md):1–35. Available from:
https://www.ncbi.nlm.nih.gov/books/NBK98166/pdf/Bookshelf_NBK98166.
pdf

71. Conti P, De Amici M, Brauner-Osborne H, Madsen U, Toma L, De Micheli
C. Synthesis and pharmacology of 3-hydroxy-delta2-isoxazoline-
cyclopentane analogues of glutamic acid. Farmaco [Internet]. 2002 Nov
[cited 2018 Jul 10];57(11):889–95. Available from:
http://www.ncbi.nlm.nih.gov/pubmed/12484537

72. N-Methyl-D-aspartic acid | C5H9NO4 - PubChem [Internet]. [cited 2018
Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/N-
methyl-D-aspartic_acid

73. Komatsu H, Furuya Y, Sawada K, Asada T. Involvement of the
strychnine-sensitive glycine receptor in the anxiolytic effects of GlyT1
inhibitors on maternal separation-induced ultrasonic vocalization in rat
pups. Eur J Pharmacol [Internet]. 2015 Jan [cited 2018 Jul 16];746:252–7.
Available from:
http://linkinghub.elsevier.com/retrieve/pii/S001429991400805X

74. Tseng G. Different State Dependencies of 4-Aminopyridine Binding to
rKv1 . 4 and rKv4 . 2 : Role of the Cytoplasmic Halves of the Fifth and
Sixth Transmembrane Segments 1. 1999;290(2):569–77.

75. Bouchard Ron, David Fedida. Closed- and Open state binding of 4-
Aminopyridine to the cloned human potassium channel Kv1.5. J
Pharmacol Exp Ther. 1995;275:864–76.

76. Tetraethylammonium chloride | C8H20ClN - PubChem [Internet]. [cited
2018 Jul 16]. Available from:
https://pubchem.ncbi.nlm.nih.gov/compound/5946

77. bupropion | C13H18ClNO - PubChem [Internet]. [cited 2018 Jul 13].
Available from:
https://pubchem.ncbi.nlm.nih.gov/compound/bupropion#section=Pharmac
ology-and-Biochemistry

78. National Center for Biotechnology Information. PubChem Compound
Database. Maprotiline [Internet]. [cited 2018 Jul 13]. Available from:

https://pubchem.ncbi.nlm.nih.gov/compound/4011

79. Amoxapine | C17H16ClN3O - PubChem [Internet]. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/amoxapine#section=Pharmacology-and-Biochemistry

80. Database NC for BIPC. Amitryptiline [Internet]. [cited 2018 Jul 13]. Available from: ttps://pubchem.ncbi.nlm.nih.gov/compound/2160

81. Figueroa KW, Griffin MT, Ehlert FJ. Selectivity of Agonists for the Active State of M1 to M4 Muscarinic Receptor Subtypes. J Pharmacol Exp Ther [Internet]. 2009 Jan 1 [cited 2018 Jul 14];328(1):331–42. Available from: http://jpet.aspetjournals.org/cgi/doi/10.1124/jpet.108.145219

82. Muscarine | C9H20NO2+ - PubChem [Internet]. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/muscarine#section=Pharmacology-and-Biochemistry

83. Bethanechol | C7H17N2O2+ - PubChem [Internet]. National Center for Biotechnology Information. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/bethanechol#section=Pharmacology-and-Biochemistry

84. Ray A, Gulati K, Anand S, Vijayan VK. Pharmacological studies on mechanisms of aminophylline-induced seizures in rats. 2005;43(10):849–53.

85. Deree J, Martins JO, Melbostad H, Loomis WH, Coimbra R. Insights into the regulation of TNF-a production in human mononuclear cells: the effects of non-specific phosphodiesterase inhibition. Clinics [Internet]. 2008 [cited 2018 Jul 16];63(3):321–8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2664230/pdf/cln63_3p0321.pdf

86. Daly JW, Jacobson KA, Ukena D. Adenosine receptors: development of selective agonists and antagonists. Prog Clin Biol Res [Internet]. 1987 [cited 2018 Jul 16];230:41–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3588607

87. Ribeiro JA, Sebastio AM. Caffeine and adenosine. In: Cunha RA, de Mendonça A, editors. Journal of Alzheimer's Disease [Internet]. 2010 [cited 2018 Jul 16]. p. S3–15. Available from: http://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JAD-2010-1379

88. Metcalf MD, Yekkirala AS, Powers MD, Kitto KF, Fairbanks CA, Wilcox GL, et al. The δ opioid receptor agonist SNC80 selectively activates heteromeric μ-δ Opioid receptors. ACS Chem Neurosci [Internet]. 2012 Jul 18 [cited 2018 Jul 16];3(7):505–9. Available from: http://pubs.acs.org/doi/10.1021/cn3000394

89. GWXFBFMLKRAWEU-YJKXCHRFSA-N | C23H26Br2N2O - PubChem [Internet]. [cited 2018 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/56972161#section=Top

90. Pomara C, Cassano T, D'Errico S, Bello S, Romano AD, Riezzo I, et al. Data Available on the Extent of Cocaine Use and Dependence: Biochemistry, Pharmacologic Effects and Global Burden of Disease of Cocaine Abusers. Curr Med Chem [Internet]. 2012 Nov 14 [cited 2018 Jul 16];19(33):5647–57. Available from: http://www.eurekaselect.com/openurl/content.php?genre=article&doi=10.2174/092986712803988811

91. Heal DJ, Smith SL, Gosden J, Nutt DJ. Amphetamine, past and present - A pharmacological and clinical perspective [Internet]. Vol. 27, Journal of Psychopharmacology. 2013 [cited 2018 Jul 10]. p. 479–96. Available from: http://journals.sagepub.com/doi/10.1177/0269881113482532

92. Gengenbacher M, Xu T, Niyomrattanakit P, Spraggon G, Dick T. Biochemical and structural characterization of the putative dihydropteroate synthase ortholog Rv1207 of Mycobacterium tuberculosis. FEMS Microbiol Lett [Internet]. 2008 Oct [cited 2018 Jul 13];287(1):128–35. Available from: https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.2008.01302.x

93. Clonidine | C9H9Cl2N3 - PubChem [Internet]. [cited 2018 Jul 13].

Available from:
https://pubchem.ncbi.nlm.nih.gov/compound/2803#section=Pharmacology-and-Biochemistry

94. Qualaquin (quinine) dosing, indications, interactions, adverse effects, and more [Internet]. [cited 2018 Jul 16]. Available from: https://reference.medscape.com/drug/qualaquin-quinine-342696#showall

95. Newman-Tancredi A, Cussac D, Quentric Y, Touzard M, Verriele L, Carpentier N, et al. Differential actions of antiparkinson agents at multiple classes of monoaminergic receptor. III. Agonist and antagonist properties at serotonin, 5-HT 1 and 5-HT 2 , receptor subtypes. JPharmacolExpTher [Internet]. 2002 [cited 2018 Jul 13];303(2):815–22. Available from: http://jpet.aspetjournals.org.

96. Mizolastine | C24H25FN6O - PubChem [Internet]. [cited 2018 Jul 14]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Mizolastine#section=Pharmacology-and-Biochemistry

97. EMETINE | C29H40N2O4 - PubChem [Internet]. [cited 2018 Jul 13]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/emetine

98. Casini A, Sun RWY, Ott I. Medicinal chemistry of gold anticancer metallodrugs. In: Sigel A, Sigel H, Freisinger E, Sigel RKO, editors. Metallo-Drugs: Development and Action of Anticancer Agents [Internet]. Berlin, Boston: De Gruyter; 2018 [cited 2018 Jul 16]. p. 199–217. Available from: http://www.degruyter.com/view/books/9783110470734/9783110470734-007/9783110470734-007.xml

99. Yekkirala AS, Kalyuzhny AE, Portoghese PS. Standard opioid agonists activate heteromeric opioid receptors: Evidence for morphine and [d -Ala2-MePhe4-Glyol 5]enkephalin as selective ??-delta; Agonists. ACS Chem Neurosci. 2010;1(2):146–54.

100. Rolipram | C16H21NO3 - PubChem [Internet]. [cited 2018 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/rolipram

101. Morris PJ, Moaddel R, Zanos P, Moore CE, Gould T, Zarate CA, et al.

Synthesis and N -Methyl- d -aspartate (NMDA) Receptor Activity of Ketamine Metabolites. Org Lett [Internet]. 2017 [cited 2018 Jul 10];19(17):4572–5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5641405/pdf/nihms909228.pdf

102. Nomifensine | C16H18N2 - PubChem [Internet]. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/4528#section=Pharmacology-and-Biochemistry

103. Brogden RN, Heel RC, Speight TM, Avery GS. Nomifensine: A Review of its Pharmacological Properties and Therapeutic Efficacy in Depressive Illness. Drugs [Internet]. 1979 Jul [cited 2018 Jul 13];18(1):1–24. Available from: http://link.springer.com/10.2165/00003495-197918010-00001

104. Loose DS, Kan PB, Hirst MA, Marcus RA, Feldman D. Ketoconazole blocks adrenal steroidogenesis by inhibiting cytochrome P450-dependent enzymes. J Clin Invest [Internet]. 1983 [cited 2018 Jul 10];71(5):1495–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC437014/pdf/jcinvest00154-0455.pdf

105. Dendrogram plot - MATLAB dendrogram - MathWorks United Kingdom [Internet]. [cited 2018 Aug 31]. Available from: https://uk.mathworks.com/help/stats/dendrogram.html

106. Agglomerative hierarchical cluster tree - MATLAB linkage - MathWorks United Kingdom [Internet]. [cited 2018 Aug 31]. Available from: https://uk.mathworks.com/help/stats/linkage.html

107. Pairwise distance between pairs of observations - MATLAB pdist - MathWorks United Kingdom [Internet]. [cited 2018 Aug 31]. Available from: https://uk.mathworks.com/help/stats/pdist.html

108. Hierarchical Clustering - MATLAB [Internet]. [cited 2018 Aug 31]. Available from: https://uk.mathworks.com/help/stats/hierarchical-clustering.html

109. Bastos AM, Schoffelen J-M. A Tutorial Review of Functional Connectivity

Analysis Methods and Their Interpretational Pitfalls. Front Syst Neurosci [Internet]. 2016;9(January):1–23. Available from: http://journal.frontiersin.org/Article/10.3389/fnsys.2015.00175/abstract

110. Smith LI. A tutorial on Principal Components Analysis Introduction. Statistics (Ber) [Internet]. 2002;51:52. Available from: http://www.mendeley.com/research/computational-genome-analysis-an-introduction-statistics-for-biology-and-health/

111. Sokal RR, Michener CD. A Statistical Method for Evaluating Systematic Relationships [Internet]. Vol. 38, The University of Kansas Science Bulletin. 1958. 1409-1438 p. Available from: http://ci.nii.ac.jp/naid/10011579647/en/

112. Olde Dubbelink KTE, Schoonheim MM, Eijen JB, Twisk JR, Barkhof F, Berendse HW. Functional connectivity and cognitive decline over 3 years in Parkinson disease. Neurology. 2014;83:2046–53.

113. Lillesaar C, Stigloher C, Tannhäuser B, Wullimann MF, Bally-Cuif L. Axonal projections originating from raphe serotonergic neurons in the developing and adult Zebrafish, Danio Rerio, using transgenics to visualize Raphe-specific pet1 expression. J Comp Neurol. 2009;512(2):158–82.

114. Gaspar P, Lillesaar C. Probing the diversity of serotonin neurons. Philos Trans R Soc B Biol Sci. 2012;367(1601):2382–94.

115. Du Y, Guo Q, Shan M, Wu Y, Huang S, Zhao H, et al. Spatial and Temporal Distribution of Dopaminergic Neurons during Development in Zebrafish. Front Neuroanat [Internet]. 2016;10(November):1–7. Available from: http://journal.frontiersin.org/article/10.3389/fnana.2016.00115/full

116. McLean DL, Fetcho JR. Ontogeny and innervation patterns of dopaminergic, noradrenergic, and serotonergic neurons in larval zebrafish. J Comp Neurol. 2004;480(1):38–56.

117. Hong E, Santhakumar K, Akitake CA, Ahn SJ, Thisse C, Thisse B, et al. Cholinergic left-right asymmetry in the habenulo-interpeduncular pathway. Proc Natl Acad Sci [Internet]. 2013;110(52):21171–6. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1319566110

118. Clemente D, Porteros Á, Weruaga E, Alonso JR, Arenzana FJ, Aijón J, et al. Cholinergic elements in the zebrafish central nervous system: Histochemical and immunohistochemical analysis. J Comp Neurol. 2004;474(1):75–107.

119. Sanchez-Simon FM, Ledo AS, Arevalo R, Rodriguez RE. New insights into opioid regulatory pathways: Influence of opioids on Wnt1 expression in zebrafish embryos. Neuroscience [Internet]. 2012;200:237–47. Available from: http://dx.doi.org/10.1016/j.neuroscience.2011.10.026

120. Rosenzweig P, Patat A. Lack of behavioural toxicity of mizolastine: a review of the clinical pharmacology studies. Clin Exp Allergy [Internet]. 1999;29(S3):156–62. Available from: http://dx.doi.org/10.1046/j.1365-2222.1999.0290s3156.x

121. Eggleston C, Hatcher RA. The seat of the emetic action of apomorphine. J Pharmacol Exp Ther. 1912;3:551–80.

122. Wu W, Zhou HR, Bursian SJ, Link JE, Pestka JJ. Emetic responses to T-2 toxin, HT-2 toxin and emetine correspond to plasma elevations of peptide YY3–36and 5-hydroxytryptamine. Arch Toxicol [Internet]. 2016;90(4):997–1007. Available from: http://dx.doi.org/10.1007/s00204-015-1508-7

123. Callaham EM, Alphin R. The emetic activity of centrally administered cisplatin in cats and its antagonism by zacopride. 1988;126(1986):142–3.

124. Olney JW, Rhee V, Ho OL. Kainic acid: a powerful neurotoxic analogue of glutamate. Brain Res. 1974;77(3):507–12.

125. Cocaine | C17H21NO4 - PubChem [Internet]. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/cocaine

126. Kapur S, Seeman P. NMDA receptor antagonists ketamine and PCP have direct effects on the dopamine D2 and serotonin 5-HT2receptors— implications for models of schizophrenia. Mol Psychiatry [Internet]. 2002 [cited 2018 Jul 10];7(8):837–44. Available from: www.nature.com/mp

127. Roth BL, Gibbons S, Arunotayanun W, Huang XP, Setola V, Treble R, et al. The Ketamine Analogue Methoxetamine and 3- and 4-Methoxy

Analogues of Phencyclidine Are High Affinity and Selective Ligands for the Glutamate NMDA Receptor. PLoS One [Internet]. 2013 [cited 2018 Jul 10];8(3):59334. Available from: www.plosone.org

128. Bisset D, Chung SH. Efficacy of external tetraethylammonium block of the KcsA potassium channel: Molecular and Brownian dynamics studies. Biochim Biophys Acta - Biomembr [Internet]. 2008 Oct [cited 2018 Jul 10];1778(10):2273–82. Available from: http://linkinghub.elsevier.com/retrieve/pii/S000527360800165X

129. Amphetamine [Internet]. [cited 2018 Jul 10]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/3007#section=Pharmacology-and-Biochemistry

130. James MK, Cubeddu LX. Frequency-dependent muscarinic receptor modulation of acetylcholine and dopamine release from rabbit striatum. J Pharmacol Exp Ther. 1984;229(1):98–104.

131. N-Methyl-D-aspartic acid | C5H9NO4 - PubChem [Internet]. [cited 2018 Jul 12]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/N-methyl-D-aspartic_acid#section=Pharmacology-and-Biochemistry

132. Tacrine | C13H14N2 - PubChem [Internet]. [cited 2018 Jul 12]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Tacrine

133. Zimmerman G, Njunting M, Ivens S, Tolner E, Behrens CJ, Gross M, et al. Acetylcholine-induced seizure-like activity and modified cholinergic gene expression in chronically epileptic rats. Eur J Neurosci. 2008;27(4):965–75.

134. Trommer BL, Pasternak JF. NMDA receptor antagonists inhibit kindling epileptogenesis and seizure expression in developing rats. Dev Brain Res. 1990;53(2):248–52.

135. MedScape. Cisplatin dosing, indications, interactions, adverse effects, and more [Internet]. [cited 2018 Jul 16]. Available from: https://reference.medscape.com/drug/qualaquin-quinine-342696#showall

136. Kohrs R, Durieux ME. Ketamine: Teaching an Old Drug New Tricks Rainer. Anesth Analg. 1998;87(5):1186–93.

137. Amoxapine | C17H16ClN3O - PubChem [Internet]. [cited 2018 Jul 13]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/2170#section=Drugs-at-PubMed-Health

138. Miller G. The emerging role of trace amine-associated receptor 1 in the functional regulation of monoamine transporters and dopaminergic activity. J Neurochem [Internet]. 2011;116(2):164–76. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1471-4159.2010.07109.x/full

139. Rothman RB, Baumann MH, Dersch CM, Romero D V., Rice KC, Carroll FI, et al. Amphetamine-type central nervous system stimulants release norepinephrine more potently than they release dopamine and serotonin. Synapse [Internet]. 2001 Jan 1 [cited 2018 Jul 13];39(1):32–41. Available from: http://doi.wiley.com/10.1002/1098-2396%2820010101%2939%3A1%3C32%3A%3AAID-SYN5%3E3.0.CO%3B2-3

140. Nomifensine | C16H18N2 - PubChem [Internet]. [cited 2019 Jan 7]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/4528

141. Apomorphine | C17H17NO2 - PubChem [Internet]. [cited 2019 Jan 7]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/6005

142. Seeman P, Tallerico T. Antipsychotic drugs which elicit little or no Parkinsonism bind more loosely than dopamine to brain D2receptors, yet occupy high levels of these receptors. Mol Psychiatry. 1998;3(2):123–34.

143. PubChem. Clozapine | C18H19ClN4 - PubChem [Internet]. 2018 [cited 2018 Jul 13]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/2818#section=Pharmacology-and-Biochemistry

144. Newman-Tancredi A, Cussac D, Quentric Y, Touzard M, Verri`ele L, Verri`ele V, et al. Differential Actions of Antiparkinson Agents at Multiple Classes of Monoaminergic Receptor. III. Agonist and Antagonist Properties at Serotonin, 5-HT 1 and 5-HT 2 , Receptor Subtypes. 2002 [cited 2018 Jul 13]; Available from: http://jpet.aspetjournals.org.

145. Metcalf MD, Yekkirala AS, Powers MD, Kitto KF, Fairbanks CA, Wilcox

GL, et al. The δ opioid receptor agonist SNC80 selectively activates heteromeric μ-δ Opioid receptors. ACS Chem Neurosci [Internet]. 2012 [cited 2018 Jul 14];3(7):505–9. Available from: http://pubs.acs.org.

146. Gabazine free base | C15H17N3O3 - PubChem [Internet]. [cited 2018 Jul 14]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Gabazine#section=Pharmacology-and-Biochemistry

147. Pentetrazol | C6H10N4 - PubChem [Internet]. [cited 2018 Jul 14]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/pentylenetetrazole#section=Top

148. O'Connor V, Phelan PP, Fry JP. Interactions of glycine and strychnine with their receptor recognition sites in mouse spinal cord. Neurochem Int [Internet]. 1996 Oct 1 [cited 2018 Jul 14];29(4):423–34. Available from: https://www.sciencedirect.com/science/article/pii/0197018695001603?via%3Dihub

149. Matsubayashi H, Alkondon M, Pereira EF, Swanson KL, Albuquerque EX. Strychnine: a potent competitive antagonist of alpha-bungarotoxin-sensitive nicotinic acetylcholine receptors in rat hippocampal neurons. J Pharmacol Exp Ther [Internet]. 1998 [cited 2018 Jul 14];284(3):904–13. Available from: http://www.jpet.org

150. Qizilbash N, Whitehead A, Higgins J, Wilcock G, Schneider L, Farlow M. Cholinesterase inhibition for Alzheimer disease: a meta-analysis of the tacrine trials. Dementia Trialists' Collaboration. Jama [Internet]. 1998;280(20):1777–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9842955

151. Birks J. HR. Donepezil for dementia due to Alzheimer's disease. Cochrane Libr. 2009;(1):1–3.

152. Coelho Filho J, Birks J. Physostigmine for dementia due to Alzheimer's disease (Review). Cochrane Database Syst Rev. 2008;(2):1–82.

153. Pe�a F, Tapia R. Seizures and neurodegeneration induced by 4-aminopyridine in rat hippocampus _in vivo_: Role of glutamate- and

GABA- mediated neurotransmission and of ion channels. Neuroscience. 2000;101(3):547–61.

154. Stafstrom CE, Sasaki-Adams DM. NMDA-induced seizures in developing rats cause long-term learning impairment and increased seizure susceptibility. Epilepsy Res. 2003;53(1–2):129–37.

155. Reichert C, Reichert P, Monnet-Tschudi F, Kupferschmidt H, Ceschi A, Rauber-Lüthy C. Seizures after single-agent overdose with pharmaceutical drugs: Analysis of cases reported to a poison center. Clin Toxicol. 2014;52(6):629–34.

156. Chu Sin Chung P, Boehrer A, Stephan A, Matifas A, Scherrer G, Darcq E, et al. Delta opioid receptors expressed in forebrain GABAergic neurons are responsible for SNC80-induced seizures. Behav Brain Res [Internet]. 2015 Feb [cited 2018 Jul 15];278:429–34. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0166432814006949

157. Henry A, Kisicki MD, Varley C. Efficacy and safety of antidepressant drug treatment in children and adolescents. Mol Psychiatry [Internet]. 2012;17(12):1186–93. Available from: http://dx.doi.org/10.1038/mp.2011.150

158. Zhang W, Peterson M, Beyer B, Frankel WN, Zhang Z -w. Loss of MeCP2 From Forebrain Excitatory Neurons Leads to Cortical Hyperexcitation and Seizures. J Neurosci [Internet]. 2014;34(7):2754–63. Available from: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4900-12.2014

159. Train models to classify data using supervised machine learning - MATLAB - MathWorks United Kingdom [Internet]. [cited 2018 Aug 15]. Available from: https://uk.mathworks.com/help/stats/classificationlearner-app.html

160. Rogers S, Girolami M. A first course in machine learning [Internet]. Finance. 2011. p. 306. Available from: http://books.google.com/books?hl=en&lr=&id=rdQ1daD8BH8C&oi=fnd&pg=PP1&dq=A+First+Course+in+Machine+Learning&ots=aYmoPlDLa8&sig=9cOppAxCXfcrM2lg-2lah8YeTwA

161. Doyle OM, Mehta MA, Brammer MJ. The role of machine learning in

neuroimaging for drug discovery and development. Psychopharmacology (Berl). 2015;232(21–22):4179–89.

162. Khodayari-Rostamabad A, Hasey GM, MacCrimmon DJ, Reilly JP, Bruin H de. A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy. Clin Neurophysiol [Internet]. 2010;121(12):1998–2006. Available from: http://dx.doi.org/10.1016/j.clinph.2010.05.009

163. Cereto-massagué A, José M, Valls C, Mulero M, Garcia-vallvé S, Pujadas G. Molecular fingerprint similarity search in virtual screening. Methods. 2015;71:58–63.

164. Marquand AF, Daly OGO, Simoni S De, Alsop DC, Maguire RP, Williams SCR, et al. Dissociable effects of methylphenidate, atomoxetine and placebo on regional cerebral blood flow in healthy volunteers at rest : A multi-class pattern recognition approach. Neuroimage [Internet]. 2012;60(2):1015–24. Available from: http://dx.doi.org/10.1016/j.neuroimage.2012.01.058

165. Stewart AM, Gerlai R, Kalueff A V, Brennan CH, Bonan CD. Developing highER-throughput zebrafish screens for in-vivo CNS drug discovery. 2015;9(February):1–8.

166. Training and Test Sets: Splitting Data | Machine Learning Crash Course | Google Developers [Internet]. [cited 2018 Aug 21]. Available from: https://developers.google.com/machine-learning/crash-course/training-and-test-sets/splitting-data

167. Steinwart I CA. Support vector machines [Internet]. Springer; 2008 [cited 2018 Aug 21]. 601 p. Available from: https://books.google.co.uk/books?id=HUnqnrpYt4IC&printsec=frontcover&dq=support+vector+machines&hl=en&sa=X&ved=0ahUKEwjH8-2Dh_7cAhWKL8AKHVO6BMsQ6AEIKTAA#v=onepage&q&f=false

168. Classify Patterns with a Shallow Neural Network - MATLAB & Simulink - MathWorks United Kingdom [Internet]. [cited 2018 Aug 21]. Available from: https://uk.mathworks.com/help/nnet/gs/classify-patterns-

with-a-neural-network.html

169. Krose B, Smagt P van der. An introduction to neural networks. 1993 [cited 2018 Aug 21];234. Available from: https://books.google.co.uk/books?id=sn6oBHq8qQQC&printsec=frontcover&dq=neural+networks&hl=en&sa=X&ved=0ahUKEwiZ1K3jiv7cAhWlJ8AKHSL2DX8Q6AEIKTAA#v=onepage&q=neural networks&f=false

170. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Logistic regression. Perspect Clin Res [Internet]. 2017 [cited 2018 Aug 13];8(3):148–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28828311

171. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nat Publ Gr [Internet]. 2015;16(6):321–32. Available from: http://dx.doi.org/10.1038/nrg3920

172. Liu B, Wei Y, Zhang Y, Yang Q. Deep Neural Networks for High Dimension, Low Sample Size Data. Proc Twenty-Sixth Int Jt Conf Artif Intell [Internet]. 2017;2287–93. Available from: https://doi.org/10.24963/ijcai.2017/318

173. Erhan D, Courville A, Vincent P. Why Does Unsupervised Pre-training Help Deep Learning ? J Mach Learn Res [Internet]. 2010;11:625–60. Available from: http://portal.acm.org/citation.cfm?id=1756025

174. Zhang R, Li J, Lu J, Hu R, Yuan Y, Zhao Z. Using Deep Learning for Compound Selectivity Prediction. Curr Comput Aided Drug Des. 2016;12(1):5–14.

175. Senders JT, Staples PC, Karhade A V., Zaki MM, Gormley WB, Broekman MLD, et al. Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. World Neurosurg [Internet]. 2018;109:476–486.e1. Available from: https://doi.org/10.1016/j.wneu.2017.09.149

176. Ghaderyan P, Abbasi A, Sedaaghi MH. An efficient seizure prediction method using KNN-based undersampling and linear frequency measures. J Neurosci Methods [Internet]. 2014;232:134–42. Available from: http://dx.doi.org/10.1016/j.jneumeth.2014.05.019

177. Jaiswal AK, Banka H. Epileptic seizure detection in EEG signal using machine learning techniques. Australas Phys Eng Sci Med [Internet]. 2018;41(1):81–94. Available from: http://link.springer.com/10.1007/s13246-017-0610-y

178. J H. Multiple comparisons: theory and methods. Chapman and Hall/CRC. 1996.

179. Brandon Rohrer. How to choose machine learning algorithms | Microsoft Docs [Internet]. 2016 [cited 2018 Aug 13]. Available from: https://docs.microsoft.com/en-us/azure/machine-learning/studio/algorithm-choice

180. Bowyer SM. Coherence a measure of the brain networks: past and present. Neuropsychiatr Electrophysiol [Internet]. 2016;2(1):1. Available from: http://npepjournal.biomedcentral.com/articles/10.1186/s40810-015-0015-7

181. Ahrens MB, Orger MB, Robson DN, Li JM, Keller PJ. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nat Methods [Internet]. 2013;10(5):413–20. Available from: http://www.nature.com/doifinder/10.1038/nmeth.2434

182. Russell JT. Imaging calcium signals in vivo: A powerful tool in physiology and pharmacology. Br J Pharmacol. 2011;163(8):1605–25.

183. Chen J, Xia L, Bruchas MR, Solnica-Krezel L. Imaging early embryonic calcium activity with GCaMP6s transgenic zebrafish. Dev Biol [Internet]. 2017;430(2):385–96. Available from: http://dx.doi.org/10.1016/j.ydbio.2017.03.010

184. Filippi A, Mueller T, Driever W. vglut2 and gad expression reveal distinct patterns of dual GABAergic versus glutamatergic cotransmitter phenotypes of dopaminergic and noradrenergic neurons in the zebrafish brain. J Comp Neurol. 2014;522(9):2019–37.

185. Wullimann MF, Rupp B, Reichert H. Neuroanatomy of the Zebrafish Brain [Internet]. 1996. Available from: http://link.springer.com/10.1007/978-3-0348-8979-7

186. Jeong JY, Kwon HB, Ahn JC, Kang D, Kwon SH, Park JA, et al. Functional and developmental analysis of the blood-brain barrier in zebrafish. Brain Res Bull. 2008;75(5):619–28.

187. Umans RA, Taylor MR. Zebrafish as a model to study drug transporters at the bloodbrain barrier. Clin Pharmacol Ther. 2012;92(5):567–70.

188. Demin KA, Meshalkina DA, Kysil E V., Antonova KA, Volgin AD, Yakovlev OA, et al. Zebrafish models relevant to studying central opioid and endocannabinoid systems. Prog Neuro-Psychopharmacology Biol Psychiatry. 2018;86(January):301–12.

189. Demin KA, Kolesnikova TO, Khatsko SL, Meshalkina DA, Efimova E V., Morzherin YY, et al. Acute effects of amitriptyline on adult zebrafish: Potential relevance to antidepressant drug screening and modeling human toxidromes. Neurotoxicol Teratol [Internet]. 2017;62:27–33. Available from: http://dx.doi.org/10.1016/j.ntt.2017.04.002

190. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. The zebrafish reference genome sequence and its relationship to the human genome. Nature. 2013;496(7446):498–503.

191. Stewart A, Gaikwad S, Kyzar E, Green J, Roth A, Kalueff A V. Modeling anxiety using adult zebrafish: A conceptual review. Neuropharmacology [Internet]. 2012;62(1):135–43. Available from: http://dx.doi.org/10.1016/j.neuropharm.2011.07.037

192. Stewart AM, Desmond D, Kyzar E, Gaikwad S, Roth A, Riehl R, et al. Perspectives of zebrafish models of epilepsy: What, how and where next? Brain Res Bull [Internet]. 2012;87(2–3):135–43. Available from: http://dx.doi.org/10.1016/j.brainresbull.2011.11.020

193. Stewart AM, Cachat J, Gaikwad S, Robinson KSL, Gebhardt M, Kalueff A V. Perspectives on experimental models of serotonin syndrome in zebrafish. Neurochem Int [Internet]. 2013;62(6):893–902. Available from: http://dx.doi.org/10.1016/j.neuint.2013.02.018

194. Loscher W. Critical review of current animal models of seizures and epilepsy used in the discovery and development of new antiepileptic drugs. Seizure Eur J Epilepsy [Internet]. 2011;20(5):359–68. Available

from: http://dx.doi.org/10.1016/j.seizure.2011.01.003

Appendix

A:

SCIENTIFIC REP♥RTS

OPEN

4-dimensional functional profiling in the convulsant-treated larval zebrafish brain

Matthew J. Winter[1], Dylan Windell[1], Jeremy Metz[1], Peter Matthews[2], Joe Pinion[2], Jonathan T. Brown[2], Malcolm J. Hetheridge[1], Jonathan S. Ball[1], Stewart F. Owen[3], Will S. Redfern[4], Julian Moger[5], Andrew D. Randall[2] & Charles R. Tyler[1]

Functional neuroimaging, using genetically-encoded Ca^{2+} sensors in larval zebrafish, offers a powerful combination of high spatiotemporal resolution and higher vertebrate relevance for quantitative neuropharmacological profiling. Here we use zebrafish larvae with pan-neuronal expression of GCaMP6s, combined with light sheet microscopy and a novel image processing pipeline, for the 4D profiling of chemoconvulsant action in multiple brain regions. In untreated larvae, regions associated with autonomic functionality, sensory processing and stress-responsiveness, consistently exhibited elevated spontaneous activity. The application of drugs targeting different convulsant mechanisms (4-Aminopyridine, Pentylenetetrazole, Pilocarpine and Strychnine) resulted in distinct spatiotemporal patterns of activity. These activity patterns showed some interesting parallels with what is known of the distribution of their respective molecular targets, but crucially also revealed system-wide neural circuit responses to stimulation or suppression. Drug concentration-response curves of neural activity were identified in a number of anatomically-defined zebrafish brain regions, and *in vivo* larval electrophysiology, also conducted in 4dpf larvae, provided additional measures of neural activity. Our quantification of network-wide chemoconvulsant drug activity in the whole zebrafish brain illustrates the power of this approach for neuropharmacological profiling in applications ranging from accelerating studies of drug safety and efficacy, to identifying pharmacologically-altered networks in zebrafish models of human neurological disorders.

Real-time cell-level imaging of neural activity is a powerful means to study the activity of neuronal assemblies at scales ranging from the sub-cellular to entire nervous systems. The most widely employed imaging-based read-outs for this are fluorescent reporters of intracellular Ca^{2+} concentration. These effectively reflect neural activity as voltage changes in neurons are usually coupled closely with Ca^{2+} transients that arise through the gating of voltage-sensitive Ca^{2+} channels. Of the fluorophores available, GCaMPs, which can be genetically encoded, are amongst the most sensitive[1]; Ca^{2+} binding to the GCaMP molecule results in conformational changes that provide a fluorescent indication of increased neural activity[2]. GCaMPs have been employed in a range of model organisms[3]; amongst these, larval zebrafish offer an unparalleled combination of higher vertebrate (including human) relevance, optical transparency, and a small brain size (comprising approximately 100000 neurons)[4] allowing simultaneous spatial coverage of multiple brain regions within a fully intact connectome. Moreover, the availability of transgenic zebrafish, in which genetically-encoded GCaMPs are under the control of the pan-neuronal *elavl3* promotor have greatly facilitated the potential for assessing neural network dynamics (e.g. refs 5–7). When coupled with imaging modalities with high spatiotemporal resolution, such as light sheet microscopy (LSM)[8], such models offer a uniquely powerful experimental system with which to functionally characterise large-scale, integrated neuronal assemblies at rest[4], under conditions of sensory stimulation[9, 10] or, as in the case with the current study, following treatment with neuroactive drugs.

[1]Biosciences, College of Life and Environmental Sciences, Exeter, Devon, EX4 4QD, United Kingdom. [2]Medical School, University of Exeter, Exeter, Devon, EX4 4PS, United Kingdom. [3]AstraZeneca, Global Compliance, Alderley Park, Macclesfield, Cheshire, SK10 4TF, United Kingdom. [4]AstraZeneca R&D Innovative Medicines, Drug Safety & Metabolism, Babraham Research Campus, Cambridge, CB22 3AT, United Kingdom. [5]Physics and Medical Imaging, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, Devon, EX4 4QL, United Kingdom. Correspondence and requests for materials should be addressed to M.J.W. (email: m.winter@exeter.ac.uk)

Compound	CAS No.	Drug concentrations in external medium	Pre-imaging exposure period
4-Aminopyridine (4AP)	504-24-5	1, 0.5 and 0.25 mM	15 minutes
Pentylenetetrazole (PTZ)	54-95-5	5, 2.5 and 1.25 mM	20 minutes
Pilocarpine hydrochloride	54-71-7	5, 2.5 and 1.25 mM	30 minutes
Strychnine hemisulphate (STR)	60-41-3	200, 100 and 50 μM	20 minutes

Table 1. Compounds and respective exposure regimes used for imaging neural activity responses in 4dpf *elavl3:GCaMP6s* zebrafish larvae. All sourced from Sigma-Aldrich, Poole, UK.

Using four exemplar drugs known to induce neural network hyperactivity by different molecular mechanisms in both mammals and zebrafish, we utilise a combination of LSM and a transgenic zebrafish (*elavl3: GCaMP6s*)[11] to quantify functional brain responses in 4 dimensions. Our approach addresses certain key challenges associated with profiling drug action in a living vertebrate, such as the contextualised quantification of pharmacologically-modified neural function within a 3D landscape of brain anatomy[4]. Accordingly, we present a computational image analysis framework that enables 3D anatomical registration against a standardised larval zebrafish brain, thus providing quantitative functional data for defined regions across the whole larval brain. Using this methodology we have quantified several measures of neural activity in multiple brain regions, revealing differential spatiotemporal patterns associated with differing molecular mechanisms of drug action. The use of *in vivo* extracellular local field potential recording from the optic tectum of comparable animals provides additional information on the temporal profile characteristics of drug-induced neural activity to aid interpretation of the GCaMP imaging data.

Here we present novel data that validate the use of this experimental system for the routine 4D functional profiling of drug action in the brain of a highly relevant vertebrate model. Furthermore our approach has many potential applications ranging from studies of CNS drug safety and efficacy, and, when combined with the relative simplicity of gene editing in the zebrafish, to the identification of functionally-altered network activity in various human neurological disorders.

Materials and Methods

Experimental animals. Transgenic *elavl3:GCaMP6s*, in an unpigmented Casper background, were the kind gift of Misha B. Ahrens (Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia, USA) and were used at 4 days post fertilization (dpf). This line incorporates stable cytoplasmic expression of GCaMP6s under the control of the *elavl3* promotor[11]. All animal work was approved by the University of Exeter's Animal Welfare and Ethical Review Body, and undertaken under project and personnel licences granted by the UK Home Office under the United Kingdom Animals (Scientific Procedures) Act, and in accordance with The University of Exeter's ethical policies.

Test compounds. The drugs selected for study (Table 1) were: pentylenetetrazole (PTZ), a GABA_A receptor antagonist[12]; pilocarpine, a muscarinic acetylcholine receptor (mAChR) agonist[13]; strychnine, a chemoconvulsant that acts predominantly via glycine receptor (GlyR) antagonism[14]; and 4-Aminopyridine (4AP), a K$^+$ channel blocker that enhances neuronal excitability and consequently the release of various neurotransmitters including glutamate[15, 16].

Drug Treatment and Light Sheet Microscopy. A schematic of the basic experimental process is presented in Fig. 1. Larvae were pre-exposed to each drug in a well of a 24 well plate (Table 1), and then transferred to another well containing drug plus the anti-nicotinic neuromuscular blocker tubocurarine (4 mM), until muscle tone was lost. Each drug was tested at 3 concentrations alongside an untreated control group (8 larvae per treatment). The concentrations and exposure periods were selected based on our previous studies in which they were observed to induce convulsions, but were well below lethal concentrations in larval zebrafish[17] (Supplementary Table 1), and also from range-finding work undertaken using wide-field fluorescence microscopy (see below). Following drug exposure, individuals were transferred to drug plus tubocurarine in 1.4% Low Melting Point (LMP) agarose, and drawn into a clear borosilicate glass capillary (940 μM internal diameter) plugged with 1.5% LMP agarose. Control larvae were treated identically without the presence of the chemoconvulsant. Each mounted fish was positioned vertically (Head down), and optical sectioning was undertaken repeatedly in the horizontal plane from the dorsal to ventral surface of the larva. This was repeated for 6 mins generating 200 × 10-Z-plane image stacks, each spanning a total depth of 220 μm. At the end of each 6 min run, the larva was checked for a normal heart rate and blood flow by microscopy to ensure viability. Standard wide-field imaging (see below) suggested minimal photobleaching (<10%/hour) supporting the ability to undertake longer term assessment. However, the image capture parameters employed were considered appropriate for obtaining a representative period of drug-induced neural activity across all major brain structures, whilst yielding datasets suitable for routine image processing and analysis techniques.

The LSM (See Supplementary Figure 1) was custom built using the Open Access platform for SPIM[8, 18]. The system comprised a 488 nm Argon laser (Melles Griot, Didam, Netherlands) for illumination, a 20x/0.5 NA objective lens (Olympus, Southend-on-Sea, UK) with an intermediary magnification of 1x, followed by a 525/50 nm emission filter and 495 nm bandpass filter (Chroma, Olching, Germany) to collect emitted light, and a 5.5MP Zyla sCMOS camera (30FPS, 640 × 540 pixels, 4 × 4 binning, 40 ms exposure, Andor, Belfast, Northern Ireland) to capture images. The camera and rotational axis (Picard Instruments, Albion, USA) were controlled though μManager[19] and the OpenSPIM plugin[18]. The Point Spread Function (PSF) of our system was determined by

Page | 105
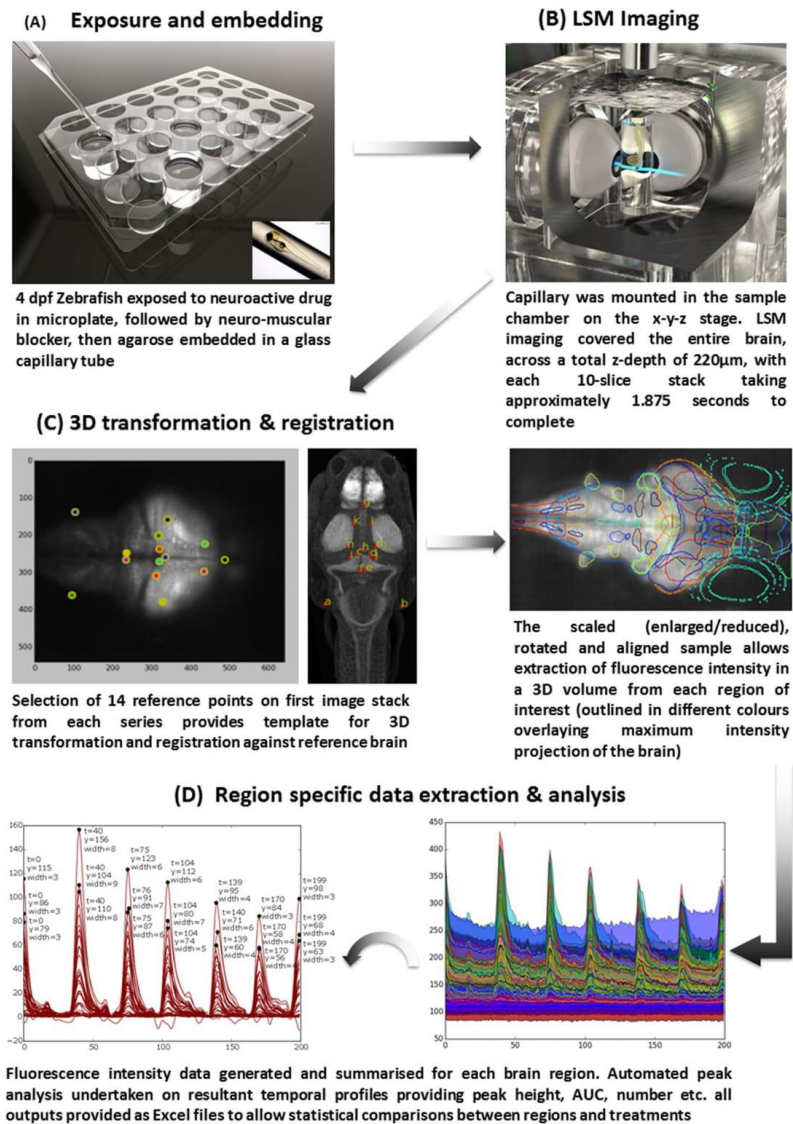
Figure 1. Schematic of the basic experimental process used to obtain 4D whole brain neural activity data from drug treated larval *elavl3:GCaMP6s* zebrafish (**A**) Larvae are cultured to 4 days post fertilisation (dpf), exposed to each drug, and neuromuscular blocker and embedded in agarose in a capillary tube; (**B**) larvae are then imaged on the LSM to capture the whole brain volume; (**C**) Resultant images are processed using a custom Python script in which 3D anatomical registration against a standard brain allows identification of regions of interest (ROIs); (**D**) fluorescence intensity data are summarised for ROI and automated peak analysis performed on the resultant temporal profiles (example plots that are automatically output are shown for illustrative purposes).
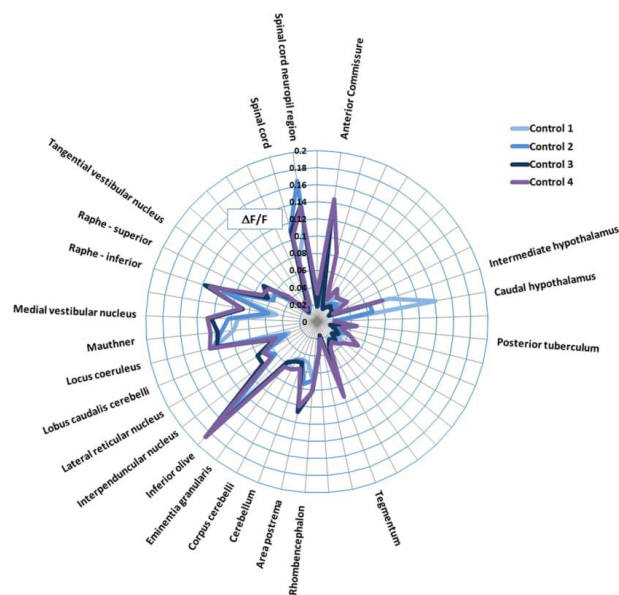
Figure 2. Summary of the control fish data across 4 experiments. Data are shown as the average of $\Delta F/F$ ($=(F_1-F_0)/F_0*100$, where $F_1=$ peak fluorescence intensity, and $F_0=$ baseline fluorescence intensity) of the fish within each control group. Only those regions showing activity greater than the median of all areas (>0.0386) are labelled. Note the consistency in brain regions showing measurable baseline activity. In particular, activity was most pronounced in the anterior commissure, hypothalamus, tegmentum, rhombencephalon, inferior olive, locus coeruleus, mauthner cells, medial vestibular nucleus, superior and inferior raphe and spinal cord neuropil region.

performing Z-stacks of $50 \times 0.5\,\mu m$ fluorescent beads, embedded in 1% agarose within a glass capillary. Images were acquired at z-intervals of $1.5\,\mu m$ using the same system settings employed during data acquisition (detailed above), and the PSF calculated using the MOSAICsuite PSF Tool (http://imagej.net/MOSAICsuite). The spatial resolution of our system was measured using a segment from a multi-grid stage micrometer (Edmund Optics, UK) with $100\,\mu m$ divisions. The grid was imaged in the sample chamber using the same optics and medium to determine the pixel length of each $100\,\mu m$ division. Each $100\,\mu m$ division at 1×1 binning was 304 pixels, and at 4×4 binning was 76 pixels.

Image processing and analysis. Data processing was performed using a custom Python image processing pipeline incorporating functions from the Scikit-image[20], Scipy[21] and Scikit-Learn[22] libraries. Each 3D image was first shift-corrected using 3D registration and then down-sampled by averaging across $3 \times 3 \times 1$ voxel blocks. The data were then baseline corrected by subtracting each voxel's baseline, estimated using a sliding minimum filter. Due to the absence of an appropriate control with which to normalise to account for regional variations in the expression of *elavl3:GCaMP6s*, the data presented in Fig. 2 (see later) were then divided by the same baseline. Next, a labelled reference brain (http://engertlab.fas.harvard.edu/Z-Brain/downloads.html)[23] was aligned with each image in 3D by using key-point registration allowing full 3D affine transforms to optimise the quality of the alignment. Key-points comprised 14 predetermined anatomical reference points that were user-selected on a representative stack from each series. The optimal affine transformation for aligning the reference points to the expert-selected points was then determined using a standard optimisation algorithm (Umeyama). The whole series was then aligned to the labelled reference brain using this optimal transform, and registration was undertaken for 45 anatomical Regions of Interest (ROIs) selected to encompass major brain structures (see Supplementary Table 2). Each dataset was visually inspected for accuracy of the registration process using the reference brain overlay image shown in Fig. 1, panel C, and any showing poor alignment were reprocessed following adjustment of the user defined key-points. The error of the registration process was estimated by calculating the median distance between the aligned reference key-points and the user selected key-points for each dataset.

For each ROI, fluorescence intensity data were provided as the time-averaged mean, median and standard deviation of the incorporated voxels. The median is a robust measure of the central tendency of the fluorescence signal within each ROI as it is less influenced by small variations in the spatial accuracy of the registration process. In addition, temporal profile peak analyses were undertaken for each ROI. For this, all activity profile

baselines underwent fine-scale smoothing using Gaussian filtering (sigma = 1.5) to preserve the signal features while smoothing high-frequency noise in the signal. For peak analysis, the baseline from the smoothed signal was subtracted (as above), and then a threshold applied of twice the standard deviation to identify significant peak regions. Any peaks spanning 2 or more consecutive imaging cycles (>1.875 secs, and thus considered non artefactual) were automatically logged and analysed, and for each animal mean peak height, peak width, peak separation in frames, Area Under the Curve (AUC) from summed intensities, and number of peaks were provided in an output file.

The use of an independent sample experimental design meant that the most appropriate representation of untreated baseline activity was the average of the median intensity values from the water-exposed animals, within each region. Consequently, statistical analysis of concentration-response relationships within each ROI was undertaken using median fluorescence intensity data comparing treated with untreated groups. This was achieved using Kruskal-Wallis tests coupled with Dunn's post hoc comparisons of treated with corresponding non-treated control fish data. The four control groups were also compared within each ROI using the same method.

For graphical representation of relative induction or suppression of neural activity between compounds and brain regions, data are expressed as the % increase over the same region from the mean of the corresponding untreated control group according to the formula: $\Delta F/F = (F_1 - F_0)/F_0 * 100$, where F_1 = treated group fluorescence intensity, and F_0 = control group fluorescence intensity.

Wide-field fluorescence microscopy for range finding. Full details of the wide-field fluorescence microscopy are contained within the Supplementary Materials and Methods section. Briefly, larvae were immobilized with tubocurarine, mounted in custom designed microchambers in LMP agarose, and transferred onto a standard inverted microscope equipped with a 10x objective, fast switching fluorescent light source and sensitive high speed CCD camera. Excitation was provided by constant 490 nm illumination, and images collected and assessed. In each case, a drug-free baseline recording of 10 mins was undertaken prior to the addition of the test compound to the microchamber, and subsequent recording was continued for a further 70 mins.

***In vivo* electrophysiology.** Full details are contained in the Supplementary Materials and Methods. Briefly, larvae were immobilised with tubocurarine and then fixed and positioned dorsal side up in the recording chamber in LMP agarose containing extracellular solution (ECS). Under 4x magnification, a glass ECS-filled microelectrode (resistance of 3–5 MΩ) was inserted into the optic tectum (Supplementary Figure 2) to record extracellular local field potential from small networks of neurons. Larvae were equilibrated for 600 secs prior to data acquisition to establish a baseline response. Local field potential (LFP) recording protocols allowed for a 300secs baseline period before drug addition to the ECS and experimental recording lasted for a total of 4200 secs. During all experimentation on larval zebrafish, the heart rate was monitored visually to confirm viability. Comparisons of the frequency (Hz) and power (mV^2) of neuronal network events were made before (0–900 secs) and after drug addition (3300–4200 secs). Student's t-tests were performed on paired data with statistical significance defined as $P < 0.05$.

Results

Estimates of registration error, and spatial and temporal resolution.
The average of the median error of the automated 3D affine registration processes was calculated as 27.56 ± 0.43 µm (median, ± SEM, n = 128 datasets), a figure in line with the error reported for an alternative registration process published by Ahrens *et al.*[24].

The spatial resolution of the system is summarised in Supplementary Figure 3. The theoretical point spread function (PSF), calculated from the average of 50 × 0.5 µm fluorescent beads and expressed as the width at half maximum intensity, was 0.751 µm. The measured lateral resolution was 3.04 pixels per µm, or 0.76 pixels per µm when employing the 4 × 4 binning used to achieve faster acquisition rates for registration of broad brain regions.

The temporal resolution of our system was 1.875 secs for each full z-series (including stage movement and image acquisition). This cycle time was suitable for capturing dynamic neural events at a level of spatial resolution appropriate for our intended aim of mapping whole brain functional responses to neuroactive drug treatment.

Mapping whole brain neural activity in untreated larvae.
Control larvae exhibited a relatively low level of neural activity compared with drug-treated animals; however, there was consistent above-baseline activity in certain brain regions (Fig. 2). Active regions were mainly rhombencephalic in origin and included the area postrema (sensory circumventricular organ involved in multiple autonomic processes)[25], inferior olive (predominantly associated with adaptive motor control)[24], locus coeruleus (involved in autonomic arousal associated with stress)[26, 27], Mauthner cells (control of the escape response)[28], raphe (serotonergic nuclei with functions in nociception, sleep-wake cycles and alertness)[29], and medial vestibular nucleus (vestibular signal processing)[30], in addition to the telencephalic anterior commissure (interhemispheric functional connectivity)[31], diencephalic caudal hypothalamus (multisystem homeostasis)[32], and mesencephalic tegmentum (motor control)[33, 34].

Mapping whole brain neural activity during drug treatment - time averaged intensity data.
Treatment with the 4 drugs confirmed distinct spatiotemporal patterns of neural activity initially identified using standard wide field fluorescence imaging (Supplementary Figure 4). Using LSM, this differential patterning was evident both in the anatomical ROIs being activated or supressed, as well as in the temporal profile characteristics exhibited. Figure 3 exemplifies the different patterns of neural activity shown for each drug, compared with that exhibited by a typical untreated control animal. From the temporal profiles shown, one can see the pronounced, relatively high amplitude Ca^{2+} transients in the drug-treated larvae compared with the low level undulation of the untreated state. Furthermore, the waveform, peak separation, peak height and number of peaks

Page | 108
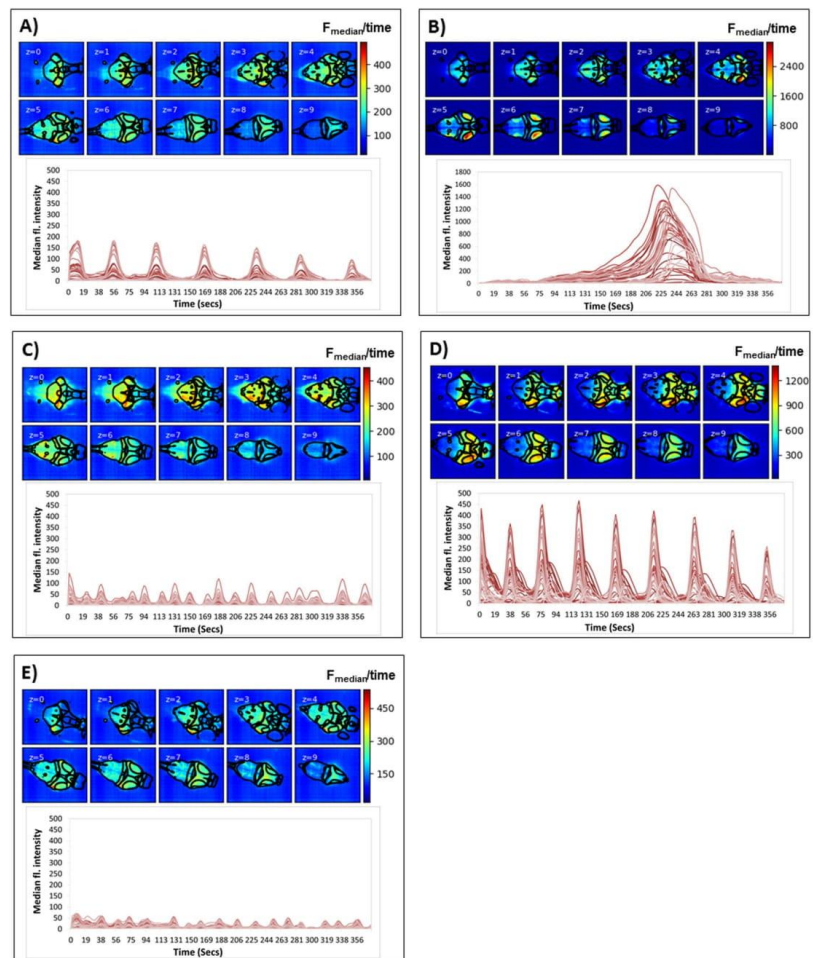
Figure 3. Example larvae exposed to each of the 4 neuroactive drugs plus an untreated control. Images are median intensity projections across time, for each z-plane (0–9 ventral to dorsal surface) and the corresponding temporal profile with each line representing the median intensity of the voxels within each of the 45 3D-registered anatomical regions. Each panel represents an example larva treated with: (**A**) PTZ, (**B**) 4-AP, (**C**) Pilocarpine (**D**) Strychnine and (**E**) untreated control. Superimposed on each image are outlines of the regions registered within that z-plane. From each montage, one can clearly see differences in the GCaMP fluorescence intensity within specific brain regions, and the different characteristics of the temporal profiles generated within these regions of interest. Accompanying colour scales show the fluorescence intensity range for each montage as the median fluorescence intensity value averaged over the experimental duration.

clearly varies between drugs, in addition to the regions affected, as shown in the pseudo-coloured fluorescence intensity images that accompany these profiles.

For each drug, Fig. 4 summarises the median GCaMP6s fluorescence intensity for each ROI as a % increase compared with the corresponding control group value, thus allowing visual comparison of the relative induction or suppression of each ROI compared to each another, and between drugs. 4AP-treated animals exhibited a widespread and considerable increase in activity across the majority of ROIs. None of these exhibited a reduction in activity and a number (e.g. pallium, habenulae, various tectal sub-regions, and the torus longitudinalis and semicircularis) exhibited average levels of more than 7 times that shown in the same ROI from the mean of the untreated control larvae. Although some of these areas also showed considerable inter-animal variation,

Median data	Control	% Change, control versus treated			
Region of Interest	mean	4AP	PTZ	Pilocarpine	Strychnine
Telencephalon	4.23	790.98	58.99	35.82	51.40
Anterior commissure	23.34	336.60	14.01	-30.74	-7.59
Olfactory bulb	3.54	572.67	12.11	-26.34	-35.24
Pallium	3.33	1034.41	107.74	26.38	75.02
Subpallium	0.32	494.19	125.10	217.42	33.69
Diencephalon	10.45	508.20	29.65	1.36	113.46
Dorsal thalamus	14.79	500.17	31.33	-1.60	136.97
Eminentia thalami	10.06	442.46	-23.59	7.40	35.73
Habenulae	3.65	748.30	94.28	18.60	47.21
Intermediate hypothalamus	26.00	408.71	17.96	8.39	57.45
Caudal hypothalamus	30.23	344.74	n/a	n/a	n/a
Pineal	2.21	329.47	25.87	6.06	77.45
Posterior tuberculum	10.82	409.52	23.27	-10.58	111.93
Preoptic area	5.36	391.36	-5.97	-7.79	75.84
Pretectum	15.73	574.66	61.26	-10.05	118.77
Rostral hypothalamus	7.36	542.64	-3.77	-45.49	-42.72
Ventral thalamus	13.62	363.66	30.96	-9.92	148.14
Mesencephalon	19.21	752.84	32.27	-7.19	57.22
Tectum stratum periventriculare	16.42	700.23	35.52	-13.35	48.85
Tectum neuropil	18.93	1170.10	19.04	-6.54	54.91
Tegmentum	34.65	387.56	43.20	5.99	76.86
Torus longitudinaris	3.10	809.06	45.79	-0.53	53.85
Torus semicircularis	21.28	762.88	14.31	0.96	76.23
Rhombencephalon	24.64	381.24	20.23	-15.60	49.30
Area postrema	13.91	312.06	-0.02	-45.18	6.87
Cerebellum	18.59	531.37	38.58	-22.08	55.61
Corpus cerebelli	20.33	527.22	42.98	-20.40	57.73
Eminentia granularis	19.82	510.37	-5.49	-9.94	30.11
Inferior olive	35.64	216.91	-18.11	-4.33	-2.20
Interpenduncular nucleus	34.07	326.64	28.96	22.04	45.00
Lateral reticular nucleus	16.82	586.10	-16.12	-21.08	3.63
Lobus caudalis cerebelli	9.85	528.72	67.48	-36.16	32.01
Locus coeruleus	41.48	269.94	28.48	-5.61	73.24
Mauthner	35.67	319.86	25.33	-11.18	70.15
Medial vestibular nucleus	35.44	389.25	28.90	-27.81	48.52
Noradrendergic neurons	14.59	541.97	19.63	-61.34	-13.24
Raphe − inferior	34.90	246.85	-3.67	-14.13	46.73
Raphe − superior	34.99	331.15	40.87	-1.59	79.19
Tangential vestibular nucleus	21.45	436.13	18.63	-25.38	31.60
Valvula cerebelli	15.60	579.89	60.25	-25.30	79.80
Eyes	1.20	86.68	-4.32	-7.94	19.80
Ofactory epithelium	1.19	191.42	27.16	-19.21	29.92
Vagal ganglia	6.34	271.56	-9.78	-13.39	13.69
Spinal cord	12.52	188.50	-28.14	-31.52	-17.12
Spinal cord neuropil region	21.34	194.17	-31.28	-22.90	-16.39

Figure 4. Summary of the voxel GCaMP6s fluorescence intensity obtained per anatomical region for each of the model chemoconvulsant compounds applied. Data are expressed as the time-averaged median fluorescence intensity values, averaged across all fish in that treatment group, expressed the % change versus the corresponding control fish group (second left column presented for brevity as the mean across the 4 control groups): $\Delta F/F = (F_1 - F_0)/F_0 * 100$ (where F_1 = treated group fluorescence intensity, F_0 = control group fluorescence intensity). Colour coding represents the degree of activation (shades of red), or suppression (shades of blue) versus that region in the control larvae, relative to other brain regions within that treatment group. The treatment group shown is that at which the highest activity was observed, which was the top concentration for 4AP, pilocarpine and strychnine, but 2.5 mM (the second highest) for PTZ due to a slight dip in neural activity at the highest treatment level. Colour coding in the left most column represents anatomical categorisation as subdivisions of the telencephalon, diencephalon, mesencephalon and rhombencephalon, along with ganglia (green) respectively from top to bottom. n/a – values not obtained due to a failure in registration probably due to the high z-depth of this region. N = 8 larvae were imaged per group. For corresponding SEM values, see Supplementary Table 3.

(Supplementary Table 3) several ROIs showed a significant and concentration-dependent increase in median fluorescence intensity (Supplementary Table 4), namely the torus semicircularis, Mauthner cells, and the medial and tangential vestibular nuclei (all $P < 0.05$).

Larvae treated with PTZ showed less widespread and extensive activation compared with the 4AP treated larvae, but with specific ROIs exhibiting a range of increased and decreased % activity compared with the corresponding control animals (Fig. 4). Most notably, there was a greater than 15% reduction in GCaMP fluorescence in the eminentia thalami, inferior olive, lateral reticular nucleus, and the spinal cord and associated neuropil region. Again, inter-animal variability in some cases was high (Supplementary Table 3), and consequently only the eminentia thalami, preoptic area and spinal cord exhibited a significant reduction in activity compared with the corresponding untreated larval data (all $P < 0.01$). This variability was most evident in terms of the areas showing apparent stimulation of activity of activation as a number of ROIS exhibited a $>50\%$ increase in activity compared with control larvae (e.g. pallium and subpallium, habenulae, pretectum, lobus caudalis cerebelli, and valvula cerebelli), although only one achieved statistical significance (Olfactory bulb, $P < 0.01$, Supplementary Table 4).

When the time averaged GCaMP fluorescence intensity data were considered, pilocarpine treatment resulted in predominantly suppressed neural activity compared with control larvae (Fig. 4). Most notably, $>30\%$ reductions in fluorescence intensity compared with the corresponding control data were observed in the anterior commissure, rostral hypothalamus, area postrema, lobus caudalis cerebelli, and the noradrendergic neurons of the interfascicular and vagal areas. As occurred with the other drugs, there was some inter-animal variability (Supplementary Table 3) and consequently only the area postrema ($P < 0.05$), and the noradrendergic neurons of the interfascicular and vagal areas ($P < 0.01$) were significantly lower than in controls, in addition to the inferior raphe ($P < 0.05$) and spinal cord ($P < 0.01$, Supplementary Table 4). Conspicuous by their increased levels of activity ($>18\%$ of controls) were the pallium, subpallium, habenulae, and interpenduncular nucleus, although none achieved statistical significance (Supplementary Table 4).

In strychnine-exposed larvae, there was an apparent altered activity in multiple ROIs compared with the untreated control larvae (Fig. 4), most notably $>100\%$ increase over controls in multiple sub-regions of the diencephalon (e.g. the dorsal and ventral thalamus, posterior tuberculum, and pretectum), and $>30\%$ decrease in activity in the olfactory bulb and rostral hypothalamus. Despite the widespread trends for altered GCaMP fluorescence in many brain regions, however, there were no statistically-significant changes in any ROI (Supplementary Table 4) perhaps, again, due to high inter-individual variability (Supplementary Table 3).

Collectively, from these data it was clear that averaging fluorescence intensity across the whole experimental period resulted in the loss of resolution and statistical power, with respect to relative regional changes in neural activation or suppression following drug treatment. Consequently, the detailed analysis of resultant waveforms within each region was required in order to reveal the more subtle effects of treatment with these 4 pharmacologically-distinct chemoconvulsant drugs.

Mapping whole brain neural activity during drug treatment - temporal profile waveform analysis.
A summary of the data obtained from the temporal waveform profile analysis is provided within Supplementary Figure 5, and graphically presented in Fig. 5.

Compared with untreated larvae, those exposed to 4AP exhibited an overall trend for highly increased AUC, peak heights (Fig. 5, note this is divided by 10 to allow plotting on the same scale) and peak width. Corresponding with this was a widespread reduction in peak number and increased separation suggesting that 4AP-exposed animals showed larger and more intense (area, duration and intensity), but less frequent, bursts of neural activity in multiple brain regions compared with those in control animals. These trends were further supported by a number of statistically-significant, concentration-related changes within specific brain regions (Supplementary Table 5). Significant concentration-related increases in AUC and peak height were observed most consistently in mesencephalic and rhombencephalic regions, and were especially centred on tectal, cerebellar and vestibular centres as well as within the vagal ganglia (example concentration response curves shown in Fig. 6). Similar significant changes were observed with respect to peak width and number further supporting the induction of fewer, larger spikes of activity occurring in a concentration-dependent manner (although no significant effects in peak separation were detected in any brain region assessed, see Supplementary Table 5).

Larvae exposed to PTZ (Fig. 5) generally exhibited increases in AUC and peak height, together with an increase in peak separation and decreased peak number (note peak activity was at 2.5 mM rather than 5 mM). This again suggests less frequent bursts of higher intensity activity in particular regions compared with untreated larvae. The areas exhibiting the most consistently significant elevations in peak intensity and AUC (Supplementary Table 5) were centred on the rhombencephalon, particularly in sub-regions of the cerebellum, in addition to some regions of the mesencephalon (example concentration-response curves shown in Fig. 6). However, few significant changes in peak width were observed. Conspicuously, only 2 regions exhibited significant reductions in activity (the preoptic area and the ementenia thalami), both of which were in the diencephalon (Fig. 6 and Supplementary Table 5). As stated, increased intensity was mirrored with significant reductions in peak number and increased peak separation across many of the same areas (Supplementary Table 5).

Following exposure to pilocarpine, larvae generally exhibited decreased peak height and AUC, along with decreased peak number and increased separation, but with a concomitant increase in peak width (Fig. 5). This overall pattern suggested a reduction in event intensity, magnitude and frequency in pilocarpine-treated larvae, alongside a slight increase in event duration. These changes were observed predominantly in the rhombencephalon, but also in the tegmentum, spinal cord neuropil region and the anterior commissure. These trends were supported by statistically-significant responses in a number of defined brain regions (Supplementary Table 5 with example curves in Fig. 6). In most cases, however, these effects were significant only at the lower or intermediate treatment concentrations.
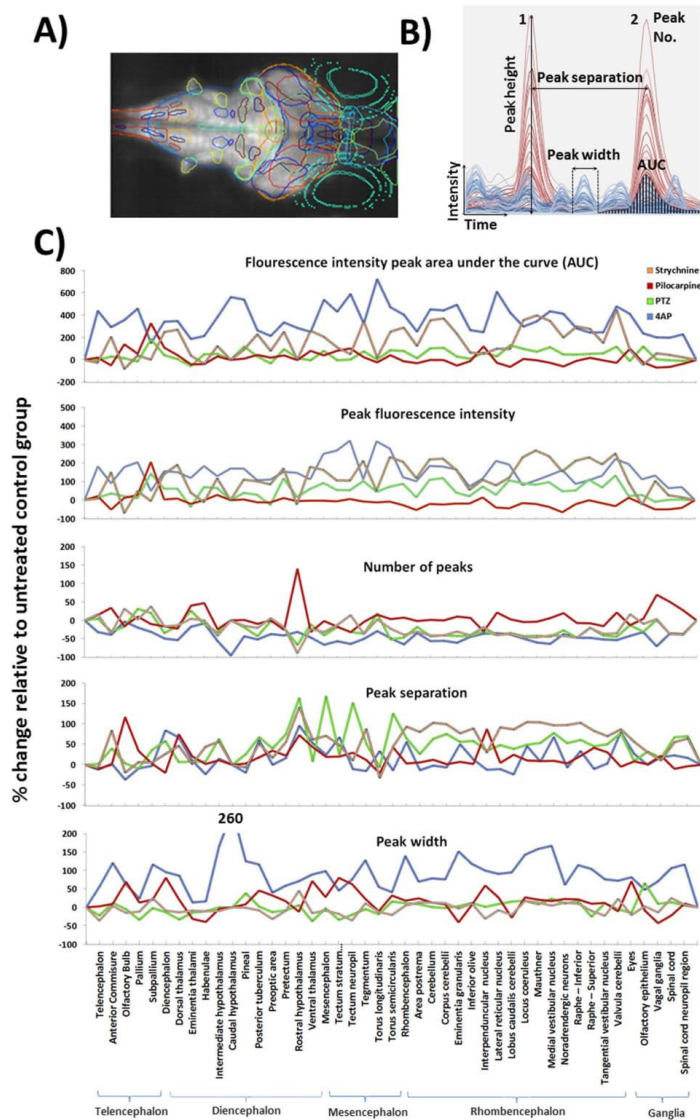
Figure 5. Results of the profile peak analysis. (**A**) Maximum intensity projection through a larva with overlays showing the positions of the 45 registered anatomical regions. (**B**) The various peak analysis parameters automatically quantified for each registered anatomical region, within each fish, are defined in panel B. For illustrative purposes only partial example profiles for a PTZ-exposed (red) and control fish (blue) are shown and the various parameters highlighted on these traces. (**C**) Peak profile analysis data for each region across treatments. The treatment group shown is that at which the highest activity was observed (Top concentrations for 4AP, pilocarpine and strychnine, and 2.5 mM for PTZ due to a slight dip in neural activity at the highest treatment level) Each data point is the mean across all fish/treatment group expressed as the % change versus the mean of the corresponding control group $(F_1 - F_0)/F_0 * 100$, where F_1 = treated group value, and F_0 = control group value. Note that for the AUC and the peak height profiles, 4AP data were divided by 10 to allow plotting on the same axis. Data also summarised in Supplementary Figure 5.
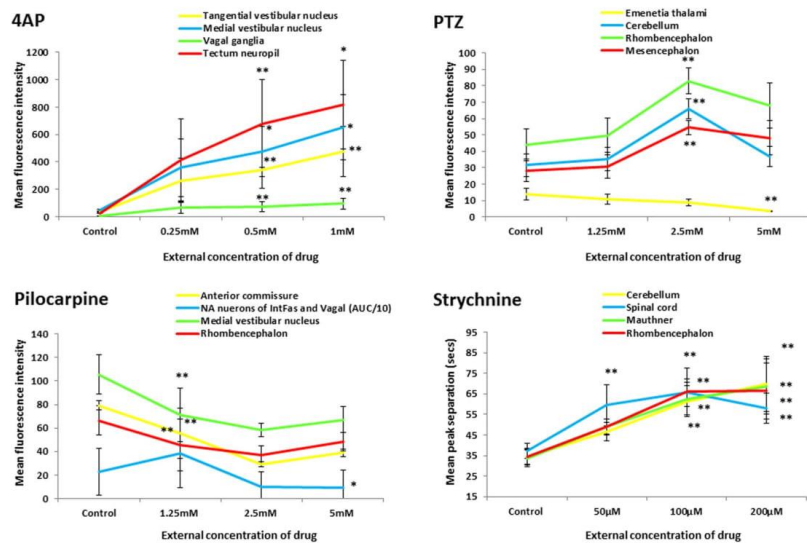
Figure 6. Example concentration response curves generated for selected brain regions of interest in fish exposed to each of the 4 drugs. To provide representative curves, all data shown are those derived from measurements of mean peak fluorescence intensity across all larvae within that treatment group, except pilocarpine noradrendergic (NA) neurons of the interfascicular and vagal areas for which the AUC is presented (divided by 10 to allow plotting on the same axis) and all of the strychnine data (peak separation data shown). Data are shown as the average of the median fluorescence intensity measures obtained for each fish in that treatment group, ±SEM (n = 7–8). All statistical analyses were undertaken using a Kruskal Wallis analysis across treatment groups, followed by a Dunn's post-hoc test in which each drug-treated group was compared with the corresponding control group. *Denotes significance at the p < 0.05, and ** at the P < 0.01 level in order of lines on graph. Strychnine significance levels are shown below each point in order of the legend. For full concentration-response datasets please see Supplementary Table 5.

Larvae exposed to strychnine exhibited a general trend for elevated AUC and peak height (Fig. 5 and Supplementary Figure 5), however, none of the ROIs exhibited a statistically-significant change in activity when compared with the corresponding untreated control larva data (Supplementary Table 5). In addition, there was a concomitant trend for reduced peak number and increased separation and variable changes in peak width in strychnine-treated animals. A number of ROIs exhibited a significant concentration-related decrease in event number and increase in separation (Supplementary Table 5) compared with untreated larvae. Many of the concentration-related changes in peak number and separation were confined to sub-regions of the rhomben-cephalon and spinal cord whereas areas showing a significant reduction in peak width were conspicuous by the dominance of the mesencephalon (example curves shown in Fig. 6).

***In vivo* electrophysiology.** A summary of the data obtained from electrophysiological recordings in drug-treated 4dpf *elavl3:GCaMP6s* larvae are presented in Fig. 7, with full results contained within Supplementary Figures 6–10. Following application of 1 mM 4AP, the frequency of neuronal network activity was significantly increased (P < 0.05), however there was no significant effect on power. Interestingly, the significant increase in tectal region AUCs observed in our functional imaging may reflect this elevated spiking across large neuronal populations resulting in large scale GCaMP fluorescence increases, with little time for recovery in between. Exposure to 5 mM PTZ resulted in reduced event frequency (P < 0.05) with no effect on power. The former mirrored the observed reduction in tectal fluorescence intensity, peak number and frequency detailed above. In contrast, pilocarpine exposure (1 mM) resulted in an increase in event power (P < 0.05), but not in frequency, and exposure to 200 μM strychnine resulted in no significant changes in either parameter. The imaging data for pilocarpine suggested relatively widespread reduction in activity, however this was predominantly observed in the rhombencephalon. Similarly, the absence of any effect of strychnine exposure perhaps reflects the relatively low occurrence of GlyR in mid- and forebrain regions (see Discussion), and both of these points illustrate the issue of relatively low spatial coverage when using electrophysiological recording approaches to gather information on the large-scale modulation of neural networks.

Page | 113
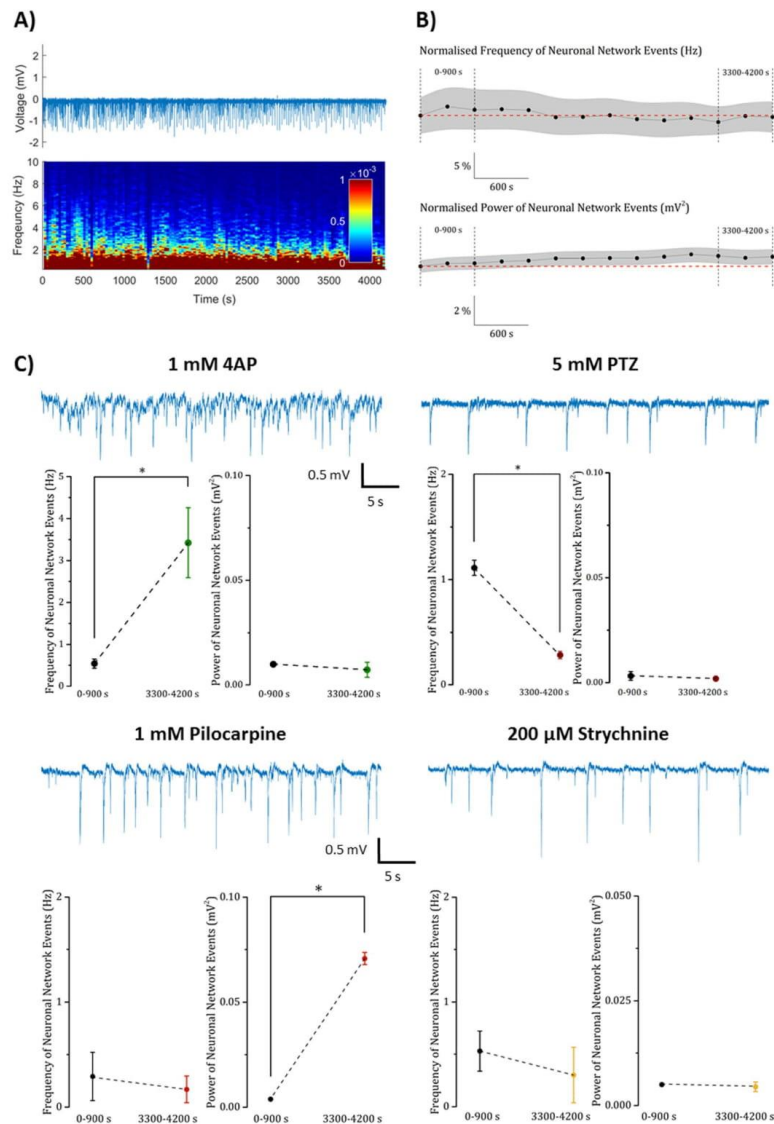
Figure 7. Summary of the electrophysiology data obtained after exposure to the 4 drugs. Panel (A) Experimental trace and supporting spectrogram of basal waveform activity recorded from the optic tectum of untreated control zebrafish larvae. (**B**) The baseline normalised frequency and power of neuronal network events in untreated control larvae does not change over the time course of the experiment (n = 6 zebrafish). (**C**) Representative traces obtained from larvae exposed to the 4 exemplar neuroactive drugs and below each trace, graphs showing control versus treated fish neural event frequency and power. In each case, data are shown as the Mean ± SEM (n = 3 larvae, except pilocarpine = 2). Baseline and post-treatment traces were compared using paired Student's t-test from which * denotes a significant difference at the P < 0.05 level.

Discussion

The power of transgenic zebrafish and LSM for functional 4D drug profiling in the vertebrate brain.

Here we show that the combined use of a pan-neuronal, genetically-encoded Ca^{2+} sensor in transparent larval zebrafish with LSM affords an extremely powerful approach for the quantitative functional analysis of vertebrate neural networks, following drug treatment.

A key limiting factor in interpreting LSM-generated neural function data has been the efficient transformation and 3D anatomical registration of the many resultant images. This has posed a significant challenge for accurately identifying ROIs and associated activity profiles when quantifying treatment-related effects[4]. We have overcome this by developing and applying a novel Python image processing pipeline in which we are able to automatically superimpose captured image voxel data onto a standardised 3D anatomical map of the larval zebrafish brain[23] and extract spatiotemporal profile data for each registered brain region. Our methodology allows us to quantify treatment-related changes in neural function within and between specific ROIs, and to compare neuropharmacological profiles between different treatment regimes. We have demonstrated, using selection of chemoconvulsant drugs, that this type of approach is suitable for the identification and quantification of activity across large-scale, integrated neuronal assemblies, and greatly expedites the time and region-directed analysis of network activation after drug treatment.

Control animal activity is associated with arousal, alertness, and escape responsiveness.

In untreated control larvae, there was a high degree of consistency in brain regions showing low level elevation of GCaMP fluorescence, compared with the corresponding baseline intensity. Most of these regions were within the rhombencephalon, most prominently the inferior olive, locus coeruleus, Mauthner cells, raphe and medial vestibular nucleus. Consistent with the nature of our conscious, but immobilised, preparation, these regions are notable for their functional roles in sensory processing, arousal, and the escape response[24, 26–30, 35]. Some this of this low level activity observed could indeed reflect network-wide responses to accelerational, photic and or auditory stimulation that may be present due to movement of the stage and/or laser illumination during imaging. Further supporting this, the tegmentum was the only mesencephalic region showing consistently high activity in control larvae, having a key role in the coordination of sensory inputs and motor control[33, 34], and particularly high activity was also noted in the anterior commissure in the telencephalon, and the caudal hypothalamus in the diencephalon which are involved in functional connectivity[31] and autonomic process control[32] respectively. The inferior olive, which is thought to be essential for adaptive motor control[24], also showed consistent above-baseline levels of spontaneous neural activity in control animals. Interestingly, a number of previous authors[9, 10, 36] have reported the existence of a 'hindbrain oscillator' adjacent to the inferior olive that exhibited 'slow lateralized oscillations' in larvae in the absence of any exogenous sensory stimulation. In support, the coefficient of variation of peak frequency in the inferior olive, across all control larvae, was 33% (n = 32), suggesting a relatively consistent temporal activity profile in this region. Furthermore, the mean frequency, in seconds (38, n = 32) is broadly in line with the oscillatory period of 20–30 s reported by Ahrens *et al.*[9]. Collectively these data support the suitability of our approach with respect to the identification and quantification of neural signals within discrete brain structures.

Different pharmacological mechanisms elicit distinct spatiotemporal neural activity patterns.

Exposure to the K^+ channel blocker 4AP resulted in considerable and widespread GCaMP activation, particularly in tectal and cerebellar nuclei. Very little is documented regarding the distribution of Kv1 channels and associated neural circuitry, or the activity of K^+-blockers in the zebrafish CNS, however, Fosque *et al.*[37] recently published a confocal image of CNS activity following exposure to 800 μM 4AP, from transgenic zebrafish larvae expressing the novel photoconvertible Ca^{2+} sensor CaMPARI. CaMPARI expression was evident throughout the brain and in particular the hindbrain and spinal cord following 4AP treatment, although, due to the nature of the CaMPARI reporter, the temporal profile and regional quantification of activity were not available for close comparison with our data. In rats injected with 4AP or picrotoxin (a PTZ-like $GABA_A$ antagonist), Salami *et al.*[38] reported distinct EEG patterns: $GABA_A$ antagonism resulted in rapid focal onset and dissipation, whereas 4AP administration resulted in more diffuse activity. Furthermore, it was noted that high frequency oscillations differed between 4AP and picrotoxin: fast ripples were characteristic of picrotoxin exposure (thought to reflect hypersynchronous firing of principle glutamatergic neurons); and slow ripples characteristic of 4AP exposure. These descriptions resonate with the evolution of network hyperactivity seen here (see also Supplementary Videos 2 and 4), in which 4AP treatment resulted in a relatively slow progression of GCaMP activation across the entire brain, whereas PTZ treatment resulted in sharp rhythmic spiking within discrete brain regions from the onset. Distinct *in vivo* electrophysiological profiles were also obtained following 4AP and PTZ treatment: PTZ-induced network activity was regular, consistent and discrete, whereas 4AP-induced network activity was more frequent and variable in terms of magnitude. When considering these profiles alongside the functional imaging output, in most cases, the 4AP imaging data revealed a small number of large magnitude increases in GCaMP fluorescence across multiple brain regions (see also Supplementary Video 2) perhaps suggesting that the temporal resolution of the imaging was insufficient to capture the faster temporal components seen with *in vivo* electrophysiology; rather it manifested as a single global elevation in GCaMP fluorescence. This possibility highlights the relative advantages and disadvantages of each approach, namely the superior spatial resolution juxtaposed with inferior temporal resolution of the functional imaging versus *in vivo* electrophysiology. It is also important to reiterate that, although both methods report neural activity, electrical recordings and GCaMP-based imaging do not measure identical phenomena. Pivotally, extracellular electrical recording can effectively report sub-threshold electrical activities, for example those arising from synaptic activity, and in particular synchronised synaptic activity; whereas GCaMP signals will predominantly arise from suprathreshold membrane potential changes and the voltage-gated Ca^{2+}

Page | 115

entry they drive. Also the influence of the kinetics of Ca^{2+} binding and particularly unbinding to/from GCaMP determine the nature of how it can report neural activity.

GABA-immunoreactive cells are found in the CNS of zebrafish early in development, and by 3dpf their distribution is similar to that of comparatively developed mammals[39, 40]. The reported distribution shows some similarities with our temporal analysis-derived peak height and AUC data, which suggest removal of inhibitory synaptic drive triggering neural hyperactivity in a number of prominent regions. In support, Mueller *et al.*[40] reported many GABA-immunoreactive (-IR) cells in the olfactory bulb and subpallium, with some penetration of the pallium; all regions showing elevated GCaMP activity here following PTZ-treatment. These authors also reported GABA-IR cells across the diencephalon, but noted relative absences in the anterior preoptic area and emenetia thalami, both areas of which showed conspicuous reductions in peak GCaMP fluorescence after PTZ treatment. These authors also reported detailed thalamic expression of GABA-IR: higher levels were observed in posterior parts of the dorsal and ventral thalamus into the prectectum (which is the only one of these regions showing a significant increase in peak height in our data), compared with a relative absence in the anterior portions of these regions. Although we are able to provide region-specific fluorescence intensity measures, our registration method does not differentiate between posterior and anterior sections of such regions. At present we can only quantify median intensity across the whole region which may mask sub-regional variations in expression revealed using GABA-IR. Rhombencephalic regions reported as having notable GABA-IR cell populations by Mueller *et al.*[40] centred on the cerebellum and sub regions, such as the valvula and corpus cerebelli. Our data also showed elevated GCaMP fluorescence in these regions which was consistently higher than control levels, with respect to both peak height and AUC from the temporal profile peak analysis. A key consideration to bear in mind when comparing the output from our functional imaging approach, versus descriptions of the distribution of neurotransmitters or their target receptors, is that our approach provides quantitative data on whole network activation and suppression. For example, each neural target possesses a network of stimulatory or inhibitory inter-neuronal connections with other regions, which may be functionally activated regardless of the presence or absence of target receptors. Similarly each of these has multiple downstream regional activation profiles, which may be targeted by the neural circuit under investigation, as well as other circuits and inputs independent of the initial target network. Consequently, the presence or absence of immunoreactivity in a specific region is not necessarily a predictor of activation after pharmacological treatment. As such our method's key strength is in providing a greater understanding of network structure and regional cross-talk than is achievable using non-functional imaging-based approaches. This allows us to understand which neural circuits are affected by drug treatment, rather than which act as the initiator or initial propagator of such activity.

No comparable zebrafish functional imaging data are available. However, Van Camp *et al.*[41] employed functional Magnetic Resonance Imaging (MRI) in rats dosed with PTZ and reported enhanced activity in cortical regions, the hypothalamus and brain stem. They also identified that the neocortex (analogous to the dorsal pallium in zebrafish) was strongly involved in PTZ-induced hyperactivation. Although we observed some enhanced activity in the telencephalon, there was a relative absence of significant-concentration-related effects on measures of GCaMP fluorescence intensity. One exception was the olfactory bulbs, which exhibited a highly significant elevation in peak height, but only at the lowest PTZ concentration. It could be postulated that this reflects an initial olfactory area response, subsequently superseded by global network hyperactivity at higher PTZ concentrations. As stated, Van Camp *et al.*[41] also reported elevated MRI signals in the hypothalamus, thalamic regions and 'brain stem' although the specific regions were not revealed. Here, (hypo)thalamic regions by and large exhibited modest, non-significant changes, but a number of mid- and hindbrain structures showed significant concentration-related increases in GCaMP fluorescence after exposure to PTZ, particularly focussed on the tectum and cerebellum.

Pilocarpine exposure typically resulted in a trend for reduced GCaMP fluorescence, particularly in the rhombencephalon. No comparable zebrafish functional imaging data are available and data on the occurrence and distribution of mAChRs in zebrafish are limited. Park *et al.*[42], however, showed that the mixed mAChR/nicotinic receptor agonist carbachol reversibly suppressed electrically-evoked field potentials in the isolated telencephalon of adult zebrafish in a dose-dependent manner. These authors noted that, in the rat telencephalon, mAChR activation has opposing effects due to its inhibition of K^+ conductance resulting in neuronal excitation, and simultaneous activation of GABAergic inhibition resulting in neuronal depression. Furthermore, this effect is carbachol-concentration dependent in rats, with enhancement of evoked activity predominant at low concentrations and the opposite true at higher concentrations. It is possible, therefore, that our observations of reductions in GCaMP fluorescence across multiple regions reflect inhibitory dominance due to the relatively high concentrations of pilocarpine used. In contrast with the predominant inhibition of GCaMP fluorescence across mesencephalic and rhombencephalic regions, some forebrain regions exhibited a trend for modest (non-significant) elevations in GCaMP fluorescence. Interestingly, the *in vivo* electrophysiology exhibited a significant increase in event power in the optic tectum following pilocarpine treatment. Although perhaps surprising given the modest elevation in GCaMP activity observed in the forebrain, as discussed previously, the two methods do not provide identical readouts. In addition, electrophysiological recordings sample small sub-populations of cells within the wider region, whilst GCaMP fluorescence data reflect the median intensity value for that whole region within which isolated populations of cells may show much higher localised activity in response to drug treatment (for example as observed in the subpallium). This further highlights the difficulty in interpreting electrophysiological data, with its low spatial resolution and coverage, in the context of whole, or even regional, brain functionality.

Strychnine exposure resulted in a general trend for elevated time-averaged GCaMP activity. However, despite the apparent widespread activation across multiple regions, the responses exhibited between animals were variable, and no significant effects on peak height or AUC were observed. Assessment of the significant effects of strychnine treatment on other peak parameters revealed that, in the vast majority of cases, these were confined to the hindbrain and spinal cord. This pattern is consistent with the role of glycine as the major inhibitory

neurotransmitter of the caudal CNS, reflected in the distribution of the GlyR which shows a clear rostral to caudal increase in expression in both mammals[43] and zebrafish[44, 45]. Despite this, however, there were still no significant increases in GCaMP activity (median intensity, AUC or peak height) within hindbrain and spinal cord regions, perhaps due to the relatively high variability in responsiveness shown between individual animals. A lack of a significant effect, on either event power or frequency, was also evident from the *in vivo* electrophysiology following strychnine treatment. This however, is more easily explained due to the likely low density of GlyR in the optic tectum of these animals, and the predominant role of glycine in hindbrain and spinal cord neurotransmission. Interestingly, Malosio *et al.*[46] observed that rat GlyR expression was abundant amongst cell populations known also to express $GABA_A$ receptors, and speculated that coexpression may allow more precise control of inhibitory inputs. Our findings may also support this at a functional network level. When considering the time-averaged median intensity data, the spatial distribution of regions showing altered neural activity after exposure to both PTZ and strychnine were broadly similar, but with the generally lower activity evident in strychnine-exposed larvae perhaps reflecting the lower density of GlyR compared with $GABA_A$ receptors in midbrain structures.

Two regions that were consistently (significantly) affected across all drug-treated groups were the Mauthner cells and the medial vestibular nucleus. Mauthner cells in the Goldfish have been shown to be sensitive to the effects of strychnine, the $GABA_A$ antagonist bicuculline, as well as the voltage-gated potassium channel blocker dendrotoxin I[47]. This is of interest as exposure here, to representatives of both of these pharmacological classes, resulted in evidence of elevated GCaMP fluorescence and thus neural activity in the same discrete region. The fact that both the Mauthner cells and vestibular nuclei appear to be consistently affected (either positively or negatively) is perhaps unsurprising given the nature of our conscious larval preparation, and that multiple auditory and vestibular afferents are known to innervate Mauthner cells in larval zebrafish[48].

The future of functional neuropharmacological profiling in zebrafish.

Despite the small scale of the zebrafish brain, there is still currently a trade-off between spatial and temporal resolution for applications focussed on whole brain functional profiling[49]. Our approach was appropriate for the capture of neural activity within specific brain regions following treatment with our selected drugs, however, faster acquisition rates would allow greater temporal resolution as well as higher spatial resolution at any given acquisition rate. In this respect, the use of piezoelectric drives to move illumination and detection objectives relative to a static sample holder to minimise accelerational forces, in combination with a faster camera, would allow more rapid cycle times. Another consideration is the differential expression of GCaMP across different anatomical regions. Arrenberg and Driever[4] pointed out that the *HuC/elavl3* promoter induces relatively low levels of expression in the diencephalon at 5dpf, and that GCaMP expression is cytosolic rather than nuclear. Although the use of untreated versus treated groups, and baseline-corrected control animal data allowed comparison of relative activity levels within regions here, diencephalic regions do appear underrepresented in terms of significant concentration-related effects, especially when assessing the peak analysis data. Analysis of the % fluorescence intensity (% of the average across the whole brain) exhibited by each broad brain region compared with the overall average for all regions, reveals that the diencephalon, mesencephalon and rhombencephalon exhibit 75%, 111% and 127% of the whole brain mean fluorescence respectively, whereas the telencephalon and ganglia exhibit just 39% and 52% respectively (Supplementary Table 6). It is not possible to say whether these regional differences reflect low inherent expression or low activity. The occurrence of significantly elevated activity within specific structures from all of these broad regions following drug treatment, however, suggests that any differential expression did not hamper our ability to detect the effects of the drugs that were applied. Nevertheless, such approaches would benefit greatly from the identification of other promotors[4], particularly those allowing fluorophore-based differentiation between brain regions. The cytosolic expression of GCaMP6s also has the potential to confound precise cellular location of neural activity[4]. For our neuropharmacological profiling approach, which focusses on broad scale expression patterns and the analysis of dynamic neural network activation and suppression, this factor is less important. However, single cell, and even subcellular resolution is clearly achievable using multi-photon based systems (see Supplementary Figure 11 and Video 5), and therefore for experimental paradigms that require cellular detailed assessment, nuclear and other compartmentalised sensors would be desirable.

In summary we have demonstrated the power of using a pan-neuronal genetically-encoded Ca^{2+} sensor in transparent larval zebrafish, along with LSM, for profiling integrated system-wide neural circuit responses to treatment with CNS-active drugs in a highly relevant vertebrate model. Our approach has clear potential for application to many areas of CNS-focussed research, particularly when combined with emerging technologies that have greatly simplified *in vivo* gene editing in the zebrafish[50]. In the current context of compounds resulting in neural network hyperexcitability, this includes: the identification of convulsive properties and their respective mechanisms as side effects of new drugs, by allowing comparison of resultant neural activity patterns against a library of known pharmacology-linked phenotypes; for testing the efficacy and identifying novel mechanisms of action of anticonvulsive drugs; and for aiding in the clarification of poorly understood molecular mechanisms. More generally this approach has considerable potential in research encompassing the assessment of spatiotemporal neural network functionality in the normal and diseased vertebrate brain.

References

1. Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature.* **499**, 295–300 (2013).
2. Akerboom, J. *et al.* Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32**, 13819–13840 (2012).
3. Grienberger, C. & Konnerth, A. Imaging calcium in neurons. *Neuron.* **73**, 862–885 (2012).
4. Arrenberg, A. B. & Driever, W. Integrating anatomy and function for zebrafish circuit analysis. *Front. Neural Circuits.* **7**, 74, doi:10.3389/fncir.2013.00074 (2013).
5. Muto, A. *et al.* Genetic visualization with an improved GCaMP calcium indicator reveals spatiotemporal activation of the spinal motor neurons in zebrafish. *Proc. Natl. Acad. Sci. USA* **108**, 5425–5430 (2011).

6. Muto, A., Ohkura, M., Abe, G., Nakai, J. & Kawakami, K. Real-time visualization of neuronal activity during perception. *Curr. Biol.* **23**, 307–311 (2013).
7. Riley, E., Kopotiyenko, K. & Zhdanova, I. Prenatal and acute cocaine exposure affects neural responses and habituation to visual stimuli. *Front. Neural Circuits.* **9**, 41, doi:10.3389/fncir.2015.00041 (2015).
8. Huisken, J. Slicing embryos gently with laser light sheets. *Bioessays.* **34**, 406–411 (2012).
9. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M. & Keller, P. J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods.* **10**, 413–420 (2013).
10. Dunn, T. W. *et al.* Brain-wide mapping of neural activity controlling zebrafish exploratory locomotion. *eLife* **5**, e12741, doi:10.7554/eLife.12741 (2016).
11. Vladimirov, N. *et al.* Light-sheet functional imaging in fictively behaving zebrafish. *Nat. Methods.* **11**, 883–884 (2014).
12. Brevard, M. E., Kulkarni, P., King, J. A. & Ferris, C. F. Imaging the neural substrates involved in the genesis of pentylenetetrazol-induced seizures. *Epilepsia.* **47**, 745–754 (2006).
13. Hoff, E. I. *et al.* Pharmacological magnetic resonance imaging of muscarinic acetylcholine receptor activation in rat brain. *Neuropharmacol.* **58**, 1252–1257 (2010).
14. Marrs, T. C. & Maynard, R. L. Neurotransmission systems as targets for toxicants: A review. *Cell. Biol. Toxicol.* **29**, 381–396 (2013).
15. Pena, F. & Tapia, R. Seizures and neurodegeneration induced by 4-aminopyridine in rat hippocampus *in vivo*: role of glutamate- and GABA-mediated neurotransmission and of ion channels. *Neuroscience.* **101**, 547–561 (2000).
16. Wahab, A., Albus, K., Gabriel, S. & Heinemann, U. In search of models of pharmacoresistant epilepsy. *Epilepsia.* **51**(Suppl 3), 154–159 (2010).
17. Winter, M. J. *et al.* Validation of a larval zebrafish locomotor assay for assessing the seizure liability of early-stage development drugs. *J. Pharmacol. Toxicol. Methods.* **57**, 176–187 (2008).
18. Pitrone, P. G. *et al.* OpenSPIM: an open-access light-sheet microscopy platform. *Nat. Methods.* **10**, 598–599 (2013).
19. Edelstein, A. D. *et al.* Advanced methods of microscope control using µManager software. *J. Biol. Methods.* **1**(2), pii: e10 (2014).
20. van der Walt, S. *et al.* Scikit-image: image processing in Python. *PeerJ.* **2**, e453, doi:10.7717/peerj.453 (2014).
21. Van Der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).
22. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn Res.* **12**, 2825–2830 (2011).
23. Randlett, O. *et al.* Whole-brain activity mapping onto a zebrafish brain atlas. *Nat. Methods.* **12**, 1039–1046 (2015).
24. Ahrens, M. B. *et al.* Brain-wide neuronal dynamics during directed behavior. *Nature.* **485**, 471–477 (2012).
25. Price, C. J., Hoyda, T. D. & Ferguson, A. V. The area postrema: a brain monitor and integrator of systemic autonomic state. *Neuroscientist.* **14**, 182–194 (2008).
26. Ma, P. M. Catecholaminergic systems in the zebrafish. I. Number, morphology, and histochemical characteristics of neurons in the locus coeruleus. *J. Comp. Neurol.* **344**, 242–255 (1994).
27. Samuels, E. R. & Szabadi, E. Functional Neuroanatomy of the Noradrenergic Locus Coeruleus: Its Roles in the Regulation of Arousal and Autonomic Function Part II: Physiological and Pharmacological Manipulations and Pathological Alterations of Locus Coeruleus Activity in Humans. *Curr. Neuropharmacol.* **6**, 254–285 (2008).
28. Sillar, K. T. Mauthner cells. *Curr. Biol.* **19**, R353–355 (2009).
29. Yokogawa, T., Hannan, M. C. & Burgess, H. A. The dorsal raphe modulates sensory responsiveness during arousal in zebrafish. *J. Neurosci.* **32**, 15205–15215 (2012).
30. Moorman, S. J. Development of sensory systems in zebrafish (*Danio rerio*). *ILAR J.* **42**, 292–298 (2001).
31. Cho, K. H., Murakami, G. & Rodriguez-Vazquez, J. F. Early fetal development of the anterior commissure. *Pediatr. Neurol.* **48**, 56–58 (2013).
32. Biran, J., Tahor, M., Wircer, E. & Levkowitz, G. Role of developmental factors in hypothalamic function. *Front. Neurosci.* **9**, 47, doi:10.3389/fnana.2015.00047 (2015).
33. Mark, R. F., Peer, O. & Steiner, J. Integrative functions in the midbrain commissures in fish. *Exp. Neurol.* **39**, 140–156 (1973).
34. Wullimann, M. F., Rupp, B. & Reichert, H. In *Neuroanatomy of the Zebrafish Brain: A Topological Atlas*, 89–101 (Birkhäuser Basel, 1996).
35. Yokogawa, T., Hannan, M. C. & Burgess, H. A. The Dorsal Raphe Modulates Sensory Responsiveness during Arousal in Zebrafish. *J. Neurosci.* **32**, 15205–15215 (2012).
36. Freeman, J. *et al.* Mapping brain activity at scale with cluster computing. *Nat. Methods.* **11**, 941–950 (2014).
37. Fosque, B. F. *et al.* Neural circuits. *Labeling of active neural circuits in vivo with designed calcium integrators. Science.* **347**, 755–760 (2015).
38. Salami, P., Levesque, M., Gotman, J. & Avoli, M. Distinct EEG seizure patterns reflect different seizure generation mechanisms. *J. Neurophysiol.* **113**, 2840–2844 (2015).
39. Stewart, A. M. *et al.* Molecular psychiatry of zebrafish. *Mol Psychiatry.* **20**, 2–17 (2015).
40. Mueller, T., Vernier, P. & Wullimann, M. F. A phylotypic stage in vertebrate brain development: GABA cell patterns in zebrafish compared with mouse. *J. Com. Neurol.* **494**, 620–634 (2006).
41. Van Camp, N. *et al.* Simultaneous electroencephalographic recording and functional magnetic resonance imaging during pentylenetetrazol-induced seizures in rat. *Neuroimage.* **19**, 627–636 (2003).
42. Park, E., Lee, Y., Kim, Y. & Lee, C.-J. Cholinergic modulation of neural activity in the telencephalon of the zebrafish. *Neurosci. Lett.* **439**, 79–83 (2008).
43. Betz, H. & Laube, B. Glycine receptors: recent insights into their structural organization and functional diversity. *J. Neurochem.* **97**, 1600–1610 (2006).
44. Hirata, H., Takahashi, M., Yamada, K. & Ogino, K. The biological role of the glycinergic synapse in early zebrafish motility. *Neurosci. Res.* **71**, 1–11 (2011).
45. Imboden, M., Devignot, V., Korn, H. & Goblet, C. Regional distribution of glycine receptor messenger RNA in the central nervous system of zebrafish. *Neuroscience.* **103**, 811–830 (2001).
46. Malosio, M. L., Marqueze-Pouey, B., Kuhse, J. & Betz, H. Widespread expression of glycine receptor subunit mRNAs in the adult and developing rat brain. *EMBO J.* **10**, 2401–2409 (1991).
47. Nakayama, H. & Oda, Y. Common Sensory Inputs and Differential Excitability of Segmentally Homologous Reticulospinal Neurons in the Hindbrain. *J. Neurosci.* **24**, 3199–3209 (2004).
48. Lacoste, A. M. B. *et al.* A convergent and essential interneuron pathway for Mauthner cell mediated escapes. *Curr. Biol.* **25**, 1526–1534 (2015).
49. Keller, P. J., Ahrens, M. B. & Freeman, J. Light-sheet imaging for systems neuroscience. *Nat. Methods.* **12**, 27–29 (2015).
50. Li, M., Zhao, L., Page-McCaw, P. S. & Chen, W. Zebrafish Genome Engineering Using the CRISPR–Cas9 System. *Trends Genet.* **32**, 815–827 (2016).

Acknowledgements

Page | 118

Author Contributions

M.J.W. conceived the project, developed the overall concept and designed the LSM studies. D.W. and J.M.o. designed and built the L.S.M., M.J.W. and D.W. conducted the LSM experiments and P.M., J.P. and A.D.R. designed and conducted the electrophysiology experiments. J.M.e., M.J.W. and D.W. developed the Image processing pipeline and J.M.e. coded the script. J.T.B. designed and coded the electrophysiology analysis pipeline. M.J.W., D.W., J.M., P.M., J.P., J.T.B., M.J.H., J.S.B., S.F.O., W.S.R., J.M., A.D.R., and C.R.T. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-06646-6

Competing Interests: S.F.O and W.S.R. are employees of AstraZeneca P.L.C. and W.S.R. and M.J.H. are shareholders of AstraZeneca P.L.C. The authors declare no other competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Page | 119