Probing Purkinje neuron Intrinsic Dynamics using experimental and computational approaches

A Thesis

Submitted to the
Tata Institute of Fundamental Research, Mumbai
for the degree of Doctor of Philosophy in Subject Board of Biology

by

Aalok Varma

National Centre for Biological Sciences
Tata Institute of Fundamental Research
Bangalore, India

May, 2023
Final Version Submitted August, 2023

DECLARATION

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort has been made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

The work was done under the guidance of Dr. Vatsala Thirumalai, at the National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore.

Aalok Varma

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.

Dr Vatsala Thirumalai

Date: August 23rd, 2023

Acknowledgements

I must begin by thanking my supervisor, Dr Vatsala Thirumalai, whose guidance, expertise, and unwavering support have been invaluable throughout this journey. She gave me a lot of freedom to tinker and figure things out at my own pace, which I sincerely appreciate. She has been constructively critical, while also always being approachable, even pulling my leg every now and then! Most importantly, though, she has been extremely patient and has believed in me, especially during many moments of failure.

I next thank my thesis committee members, Prof MK Mathew and Prof Rohini Balakrishnan, for their advice and insight. In particular, this thesis would not have been possible if Rohini hadn't connected me to my collaborators - Dr Prasanta Kumar Ghosh and his student Mr Sathvik Udupa, both of whom I'm indebted to for their prompt and professional work.

I have spent the majority of my 20s at NCBS, and I am ever grateful to the entire community. The scientific exposure I got here has enabled me to not just have depth but also breadth in my knowledge of biology. I have also been fortunate to be a part of numerous extra-curricular activities on this campus, from Sympotein to the Dance Club. All these events have kept me going through tough times.

Perhaps the best thing I got from NCBS, though, has been a fantastic cohort of colleagues and lifelong friends. I'm immensely grateful to Mugdha for her tough love and for always being there and pushing me to be better. Through her I also became fast friends with Joseph, with whom I've taken fun trips and from whom I inherited a reliable mode of transport in Bangalore. I also must thank Rimple for being a fun collaborator and a very encouraging friend. Many other people have been incredible sources of support and/or fun over the years - Jagath, Dhanashree, Krishnan, Anzal and Thasneem.

A special thanks goes to Bhavya, Athulya, Jolly, Kaiwalya, and Batul. We went through our PhD journeys together and while we're all very different people, we somehow came together as a coherent group and I cherish all the shenanigans we've been up to together over the years.

I am lucky to have had passionate labmates and seniors who paved the way for me, all of whom have taught me something. In particular, I'd like to thank Mohini and Urvashi, who treated me as their equal despite the difference in our age and experience, and Meha and Shivangi, who have accepted me for who I am, always valued my opinion, and been there when the goings were tough. I am also thankful to Jagadeesh for keeping our fish facility running.

Lastly, I'd like to thank my family - my parents, my sister Anagha and brother-in-law Sandeep. They have supported all my endeavours over the years. I'm glad I lived in the same city as my aunt and my cousin Uddhav. They have been a critical link between me and the rest of society outside campus, and I don't know how I'd have managed to finish this degree without that.

I'm sure I've forgotten to mention many others - teachers, friends, counsellors, etc. - who have shaped my life in innumerable ways. I wouldn't have gotten this far without you. Thank you.

Table of Contents

List of figures

List of tables

Abstract

The cerebellum, an evolutionarily old and well-conserved vertebrate brain structure, has long fascinated neuroscientists for its involvement in motor adaptation. Despite its significance, much remains to be understood about the functioning of the cerebellar circuit. Traditionally, rodents and non-human primates have been utilised for cerebellar research. However, the zebrafish (*Danio rerio*) presents a promising alternative model system to investigate cerebellar function due to its simple yet analogous cerebellar architecture and the availability of advanced imaging and electrophysiology tools.

Purkinje neurons (PNs) are the central processing units of the cerebellum and play a crucial role in generating coordinated locomotion. These neurons generate two distinct electrophysiological events, namely simple spikes and complex spikes driven by climbing fiber inputs. Notably, in teleosts such as zebrafish complex spikes are replaced by giant all-or-none excitatory postsynaptic potentials (EPSPs) from climbing fibers. Additionally, PNs exhibit membrane potential bistability, existing in two distinct stable membrane potential states and displaying different firing patterns accordingly.

Understanding the function of bistability in the circuit has been a challenge, particularly at the population level. The widely used calcium imaging method has limitations in reporting specific electrophysiological events of PNs, hindering the study of bistability and its effects on circuit function and behaviour.

In this graduate thesis, I employed single and dual electrophysiological recordings to investigate the relationships between simple spikes and CF inputs within and between neurons. Furthermore, I mapped the correlation between a Purkinje neuron's electrophysiological activity and its calcium signal. Notably, in collaboration with a research group at the Indian Institute of Science, I developed a novel machine learning tool called CaMLsort that enables the classification of PN states solely from their calcium signals. This tool opens new avenues for studying the distribution and function of bistability.

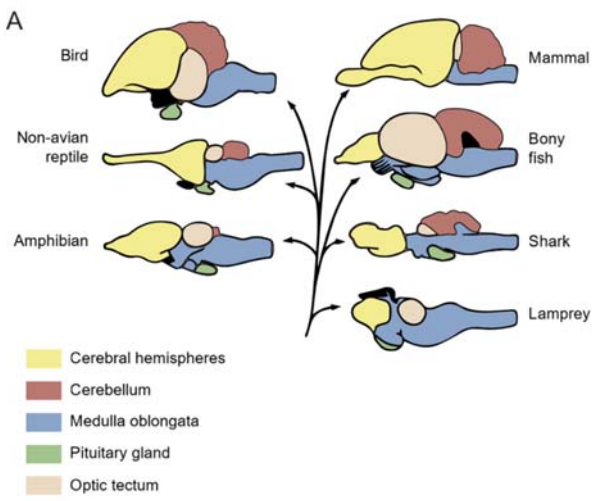
Leveraging the advantages of zebrafish as a model system, this thesis contributes to a deeper understanding of the cerebellar circuit's functionality through the exploration of bistability in Purkinje neurons. The findings presented here lay the groundwork for future investigations into the implications of bistability on neural circuits and behaviour, potentially leading to significant advancements in the field of neuroscience.

Chapter 1 - Introduction

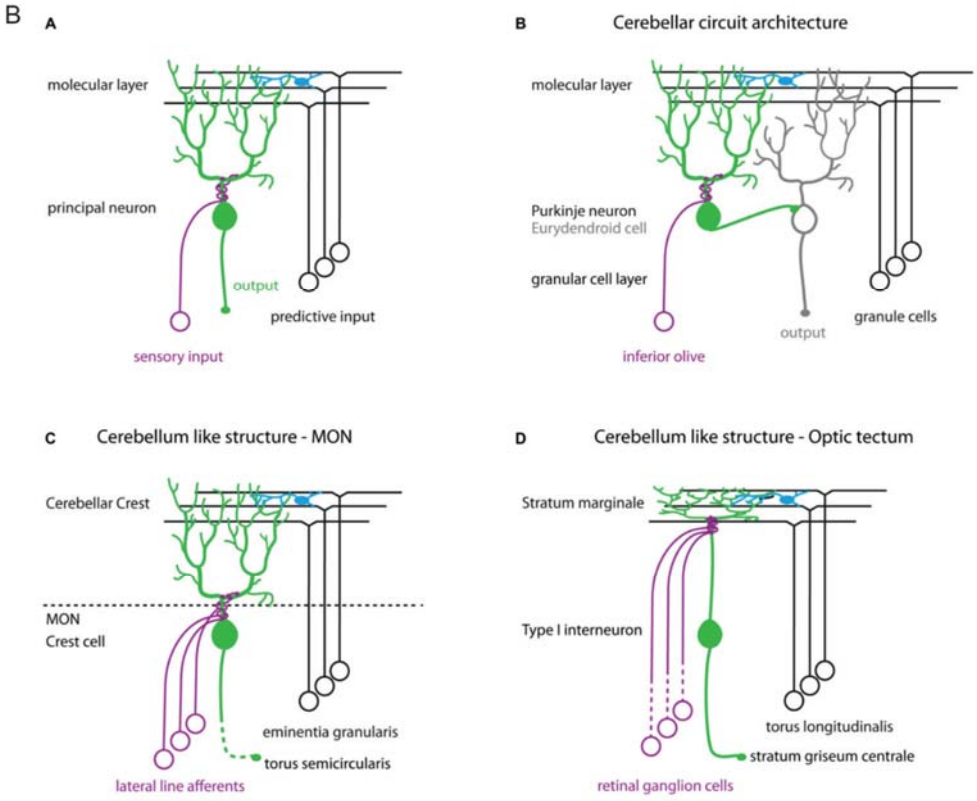
1.1 Cerebellar structure and evolution

The cerebellum is an evolutionarily old and well-conserved vertebrate brain structure (Fig 1.1A) (Bell, 2002) that has fascinated neuroscientists for over a century (Ramón y Cajal, 1911). A major feature of the hindbrain of vertebrates, the word cerebellum means “little brain”, attesting to the fact that it is usually smaller than the cerebrum, except in the case of certain bony fishes like mormyrid fish, where the cerebellum is very large, encompassing more than half the entire brain (Meek et al., 2008; Meek & Nieuwenhuys, 1991). The cerebellum has been thought to be primarily involved in the control of movement, as patients with cerebellar damage present with motor deficits (Kandel, 2013).

Besides the cerebellum, most vertebrates (with the exception of birds and reptiles) also have several cerebellum-like structures, which include the following regions - the MON and DON (medial or dorsal octavolateralis nucleus), the dorsal cochlear nucleus and the rostral nucleus of the thalamus (Bell et al., 2008). The key feature of all these regions is a molecular layer made up of parallel fibres, which originate from granule cells (Fig 1.1B). Most of these regions have a 3-layer structure consisting of the molecular layer, principal cell layer and a granule cell layer (although granule cells may be distributed elsewhere in certain cases). The granule cells receive peripheral sensory input (mechanosensory as well as electrosensory, in certain cases) and send out projections via parallel fibres to a neuropil layer called the molecular layer (Fig 1.1B). These parallel fibres and other peripheral input converge onto the principal cells of the structure (which are Purkinje neurons in the case of the cerebellum). The principal cells have spiny dendrites extending into the molecular layer and can be either excitatory or inhibitory, with excitatory principal cells often projecting to higher sensory centres in the brain and inhibitory principal cells connecting more locally and on efferent cells (Bell et al., 2008). The stereotyped circuit architecture is one of the reasons the cerebellum has garnered widespread interest for so long, with the idea being that its organisation might make it a relatively simpler region to study and understand.



Adapted from Fig 34.25, Principles of Life (David Hillis, 2 ed)



From Robra and Thirumalai (Frontiers in Neural Circuits, 2016)

Figure 1.1 The evolution and organisation of the cerebellum. (A) A schematic of the brains of several animal groups showing their phylogenetic position and marking homologous structures in the same colour. The cerebellum (copper red) is found in almost all vertebrates. **(B)** (A) Cerebellar and cerebellar-like circuits have a similar circuit architecture that is typically three layered - molecular layer (ML), principal cell layer (PCL) and granule cell layer (GCL). The granule cells (magenta) send out projections to the molecular layer that run in parallel (parallel fibers), which then impinge on the dendritic arbour of the principal neuron (green). (B-D) A comparison of the local circuits of the cerebellum and several cerebellum-like structures - the MON (medial octavolateralis nucleus) and optic tectum (Robra and Thirumalai, 2016)

Traditionally, rodents and non-human primates have been used to study the development and function of the cerebellum. Zebrafish (*Danio rerio*), however, are an excellent alternative model system to investigate nervous system function (Orger & de Polavieja, 2017), particularly that of the cerebellum. First, being vertebrates, their cerebellar architecture is very similar to that in higher vertebrates. It is much simpler, however, with about four orders of magnitude fewer (Purkinje) neurons - ~1 million in humans (Andersen et al., 2003), but only ~200-400 in larval zebrafish (Hamling et al., 2015). Second, mutant and transgenic lines allow cellular and circuit level analysis. Third, their eggs are fertilised and develop *ex utero*, allowing experimenters to watch the development of the brain in live animals. Moreover, they develop a large repertoire of motor behaviours several hours after fertilisation. Most importantly, with the latest imaging and electrophysiology tools, brain activity can be recorded *in vivo* during motor behaviours (Ahrens et al., 2013).

Zooming in from gross morphology and cell types to the level of molecules, the expression of several genes is also conserved across phylogeny. For instance, the expression of GluRδ2 is specific to the cerebellum. In mammals, it is found in Purkinje neurons (PNs) and principal neurons of the Deep Cerebellar Nuclei (DCN), whereas in zebrafish it is also found in the MON (medial octavolateralis nucleus) and OTML (marginal layer of the optic tectum), which are both cerebellar-like circuits (Bell et al., 2008). GluRδ2 is an interesting member of the glutamate receptor family, since it neither binds to glutamate nor does it conduct a current (Schmid & Hollmann, 2008). Yet we know that it is critical for cerebellar function because knocking it out or mutating it disrupts parallel fibre signalling to Purkinje neurons, leading to severe ataxia (Hashimoto et al., 2001; Kashiwabuchi et al., 1995).

Similarly, previous work from our lab has shown that the neuromodulatory signalling

molecule Darpp32 is also specifically expressed in Purkinje neurons (and principal neurons of other cerebellum-like circuits) of larval zebrafish (Robra & Thirumalai, 2016). Darpp32 is a regulatory subunit of protein phosphatase-1, which is downstream of G-protein receptors, such as dopaminergic receptors, and is expressed from the time of PN specification until adulthood (Robra & Thirumalai, 2016).

Another unique gene that is expressed only in Purkinje neurons of fish, birds and adult mammals alike is called aldolase c (shortened to aldoc and also known as zebrin II), which encodes a glycolytic enzyme. Bony fish (teleosts), which underwent a whole genome duplication event in their natural history, have two orthologs of this gene (*aldoca* and *aldocb*), of which *aldoca* is expressed strongly in all Purkinje neurons (Bae et al., 2009). Its promoter has been identified and cloned, too (Tanabe et al., 2010), a tool which will be used later in this thesis. Unlike fish, however, not all Purkinje neurons in the mammalian cerebellum express zebrin II (Bae et al., 2009). In fact, a striking pattern of alternating stripes of zebrin II positive and negative cell populations is found in 9 of the 10 lobules (Brochu et al., 1990; Zhou et al., 2014). While zebrin II is only expressed postnatally in rodents, it is expressed from the time of PN differentiation in zebrafish and is, therefore, an early-onset marker of PNs in zebrafish, like Darpp32 described above. Thus, over the course of evolution, cell populations have diversified and possibly become functionally specialised. Nevertheless, the conserved anatomical and gene expression profiles of the cerebellar circuitry points to a conserved function, which is easier to investigate in evolutionarily older and simpler organisms like zebrafish.

This thesis focuses on the developing cerebellum of larval zebrafish. In particular, it focuses on the physiology of Purkinje neurons, which I describe next.

1.2 Purkinje neurons - activity and function

Discovered in 1837 by the Czech physiologist Jan Evangelista Purkinje, Purkinje neurons (PNs) are the principal neurons of the cerebellum (Fig 1.2A). They have been the centre of attention in cerebellar research not only because of their striking, elaborate dendritic architecture (Fig 1.2B), but also because of their position as a key sensorimotor integrator and regulator of the cerebellar circuit.

Found in the eponymous Purkinje cell layer (PCL), these GABAergic neurons extend their dendrites into the molecular layer, where they receive two major sources of excitatory input - parallel fibre (PF) input from the molecular layer of the cerebellum, and climbing fibre (CF)

input from a region in the hindbrain called the inferior olive (Fig 1.2A). The former is a collection of peripheral sensory information, while the latter is thought to report motor error (Apps & Garwicz, 2005). They also receive inhibitory input from molecular layer interneurons. Thus, Purkinje neurons of the cerebellum are a sensorimotor integrating centre as both sensory and motor information modulates their activity.

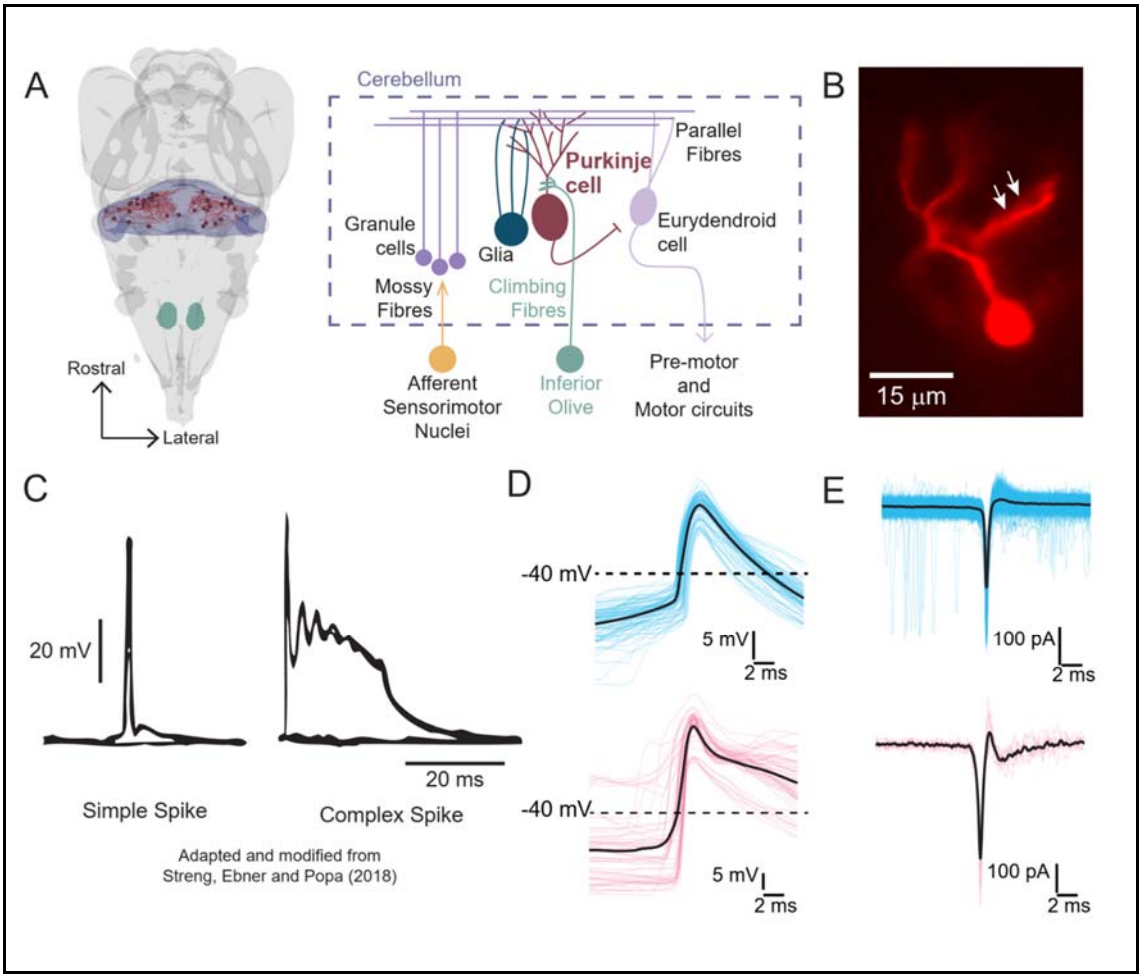


Figure 1.2 Location, morphology and electrophysiological activity of Purkinje neurons (A) *Left:* The location of the cerebellum (blue) in the larval zebrafish brain. Purkinje neurons are marked in dark red. This schematic was made using mapZebrain. *Right:* The circuit architecture of the zebrafish cerebellum showing Purkinje neurons (red) and the inferior olive (green). **(B)** A Purkinje neuron from a larval zebrafish filled with 5% sulforhodamine imaged after a whole-cell patch clamp recording. Note the elaborate dendritic arborisation of the neuron. Individual spines on the dendrite can also be seen (white arrows). **(C)** Electrophysiological signatures of Purkinje neuron activity in mammals - simple spikes are shown on the left, and complex spikes on the right. **(D)** Two kinds of electrical activity in Purkinje neurons of larval zebrafish, as observed using the whole-cell patch clamp recording configuration. Simple spikes are shown in cyan (top), and CF input-triggered EPSPs in magenta (bottom). The average event profiles are shown in black. **(E)** Simple spikes (cyan, top) and CF EPSPs (magenta, bottom), as in (D), but recorded in the loose patch configuration.

"Activity" in Purkinje neurons is reported as two separate electrical events. In mammals, these are called simple spikes (SS) and complex spikes (CS) (Eccles et al., 1967; Otis, 2016) (Fig 1.2C). Simple spikes occur spontaneously, even in the absence of synaptic input (Häusser & Clark, 1997), and can go up to high firing rates (~100 Hz), in response to modulation by parallel fibre input (Eccles et al., 1967). Complex spikes, on the other hand, are triggered by climbing fibre input and are characterised by a sharp initial discharge followed by a slow wave with several spikelets superimposed on it (Bell & Kawasaki, 1972; Eccles et al., 1966). Similar spiking properties have been reported in zebrafish (Sengupta & Thirumalai, 2015). However, unlike in mammals, zebrafish Purkinje neurons do not fire complex spikes. Instead, climbing fibre inputs lead to an all-or-none giant EPSP with a large amplitude (Fig 1.2D and 1.2E).

Research in primates, cats and rodents alike has shown that the simple spike firing rate in Purkinje neurons is correlated with kinematics, regardless of the task or effector in question (Popa et al., 2019). For example, simple spike (SS) firing rate is modulated by joint movement in cats (Thach, 1967) and correlates with arm movement velocity in monkeys performing a visually-guided reaching task (Marple-Horvat & Stein, 1987). Similarly, SS rate has been shown to encode eye movement kinematics (position, velocity and acceleration) in a smooth pursuit task, during saccades, and even during the vestibulo-ocular reflex (Popa et al., 2019; Shidara et al., 1993). Simple spiking also correlated with whisking in mice, with the majority of Purkinje neurons increasing their firing rate during whisking (S. Chen et al.,

2016). Thus, PN spiking does not unambiguously report any unique aspect of an animal's state, but instead correlates with many different sensorimotor parameters.

Complex spikes (CS), on the other hand, are thought to specifically report error - representing either sensorimotor error and/or prediction error (Herzfeld et al., 2018; Narayanan et al., 2021). During various ocular tasks, such as tracking, saccades and the vestibulo-ocular reflex, complex spikes occur during retinal slips (i.e. errors). CS modulation also occurs when the errors are not self-generated, but rather imposed externally, either by adding unexpected load, or with adaptation to visuomotor transformations (Gilbert & Thach, 1977; Ojakangas & Ebner, 1994).

Besides sensorimotor functions, the cerebellum also performs higher order functions, and has been known to be involved in conditioning, such as eyeblink conditioning behaviour (Mostofi et al., 2010). While the majority of our understanding of cerebellar function comes from work in rodents, we know that some of these functions are conserved in fish. For instance, cerebellar ablation leads to decreased task performance in associative behaviours in zebrafish (Harmon et al., 2017; Narayanan et al., 2021) and heterogeneous task-related activity was observed in Purkinje neurons during these tasks. How this multiplexing of motor and non-motor functions occurs is also poorly understood, although anatomical organisation of specialist cells could be a possible mechanism. The mammalian cerebellum can be partitioned into separate regions based on the expression of zebrin II in Purkinje neurons. This anatomical organisation appears to indicate functional segregation. Recent work showed that the zebrin negative (Z-) regions seemed to be involved in eyeblink conditioning (Mostofi et al., 2010; Wu et al., 2019), whereas zebrin positive (Z+) regions were involved in the vestibulo-ocular reflex (VOR) (Wu et al., 2019). Despite not having any anatomical segregation into domains marked by zebrin expression, the zebrafish cerebellum also has some degree of anatomical specialisation, as the regions involved in the optomotor response (OMR) and optokinetic response (OKR) are mutually exclusive (Matsui et al., 2014). One can only speculate how this specialisation comes to be, but one might expect that it is a result of a combination of a cell's synaptic connectivity and its intrinsic properties.

One such intrinsic physiological property of Purkinje neurons often gets overlooked: that they exhibit membrane potential bistability (described in the next section). It is possible that membrane potential bistability could confer a neuron the ability to be flexible in its function and thereby allow it to perform a variety of roles. This has been a particularly challenging domain of research, and while the question remains unanswered, this thesis aims at making such investigations feasible in the future.

1.3 Neuronal bistability and circuit function

The ability of neurons to exist in two stable states and exhibit distinct firing patterns/excitabilities in those states is called “bistability”. Bistability is a phenomenon that emerges from the nonlinear dynamical properties of a neuron and has been reported in a variety of systems across the animal kingdom.

1.3.1 Discovery and Types of Bistability

In 1977, Schwindt and Crill demonstrated that quiescent alpha motor neurons in anaesthetised cats showed sustained firing in response to a depolarising current injection, which continued well beyond the duration of the current pulse itself (Schwindt & Crill, 1980, 1977) (Fig 1.3A). This result was later reproduced by others in unanaesthetised decerebrate cats (Hounsgaard et al., 1984), which clarified that bistability wasn’t an effect of the anaesthetic itself, but rather a property of the neuron. Current injection didn’t always generate self-sustaining action potentials, but could instead make the neuron stay at a more depolarized, near-threshold membrane potential, called plateau potentials or the “up” state (Kiehn, 1991). The neuron’s membrane potential then spontaneously returns to baseline or it can be brought back to the “down” state with a short hyperpolarizing current pulse.

Unlike plateau neurons, whose states alternate between quiescence (subthreshold) and sustained, regenerative firing (depolarised), some neurons alternate between two distinct patterns of firing. The R15 neuron in the abdominal ganglion of the mollusc *Aplysia* was the first neuron shown to exhibit this kind of bistability (Lechner et al., 1996) (Fig 1.3B). The R15 neuron is an extensively characterised oscillatory neuron that is endogenously bursting. A brief current pulse triggers a prolonged period (on the order of minutes) of sustained firing, which the authors call “beating”. As with the plateau neurons, this neuron’s mode can then spontaneously switch back or be forcibly changed with a brief hyperpolarizing current pulse. Thus, these neurons have the potential to alternate between burst/phasic firing and beating/tonic firing.

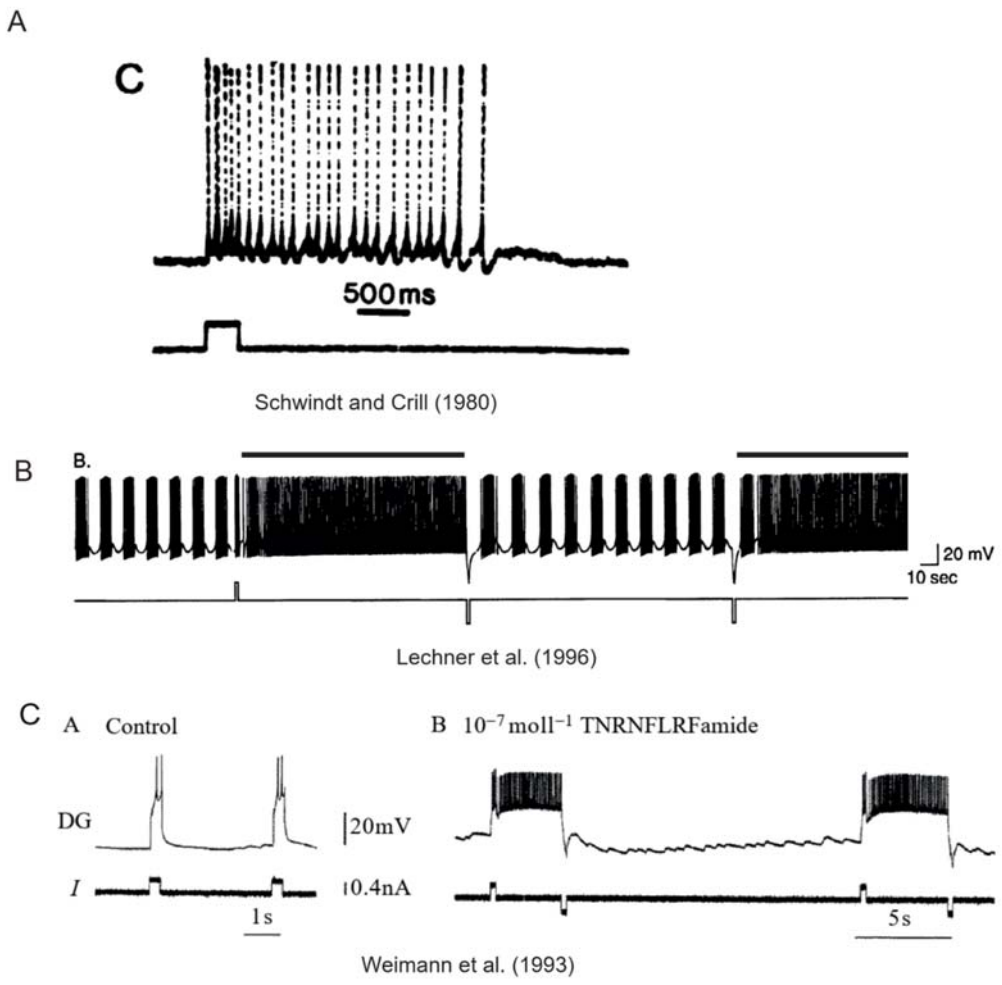


Figure 1.3 Bistable firing properties in different neurons. (A) Alpha-motorneurons in the spinal cord of anaesthetised cats fire action potentials (top) that outlast the duration of the current injection pulse (bottom). **(B)** The endogenously bursting R15 neuron of *Aplysia* switches to tonic firing upon current injection. The periods of tonic firing are highlighted with a bar above the trace. In these cells, the direction of current injected doesn't matter - both depolarising (first switch) and hyperpolarizing (second switch) currents trigger a switch to tonic firing. **(C)** Bistability can be induced in a neuron by neuromodulators. In control conditions (left), DG (dorsal gastric) neurons of the crab's stomatogastric nervous system are quiescent and fire (top trace) only in response to current injections (bottom trace). In the presence of the neuropeptide TNRNFLRFamide (right), these same neurons are able to fire for prolonged durations outlasting the current pulse.

Since the initial report of motor neuron bistability, several other cell types have been reported to be bistable, in both vertebrate and invertebrate systems alike. In the

stomatogastric nervous system (STG) of crustaceans, for example, the pyloric constrictor (PY) neurons were found to produce sustained plateaus of firing in response to a brief discharge of the anterior pyloric modulator (APM) neuron (Dickinson & Nagy, 1983; Marder et al., 1996). Similarly, the dorsal gastric (DG) neuron of the STG could alternate between quiescence and prolonged, self-sustained firing (Weimann et al., 1993). In the medicinal leech, the annulus erector (AE) motoneuron fired tonically and could be induced to switch to a burst firing mode (Puhl & Mesce, 2008). In vertebrates, examples of bistable neurons include thalamic reticular neurons (Fuentealba et al., 2005; Steriade et al., 1993), mitral cells of the olfactory bulb (Heyward et al., 2001), spiny neurons in the striatum (Misgeld et al., 1986; Wilson & Groves, 1981) and Purkinje neurons of the cerebellum (Loewenstein et al., 2005; Sengupta & Thirumalai, 2015; Williams et al., 2002). Bistability might be a more widespread phenomenon than currently reported given that much attention has not been devoted to it. For instance, given the homology between the cerebellum and cerebellum-like circuits, I propose that bistability may be a property of principal neurons in cerebellar-like circuits, too.

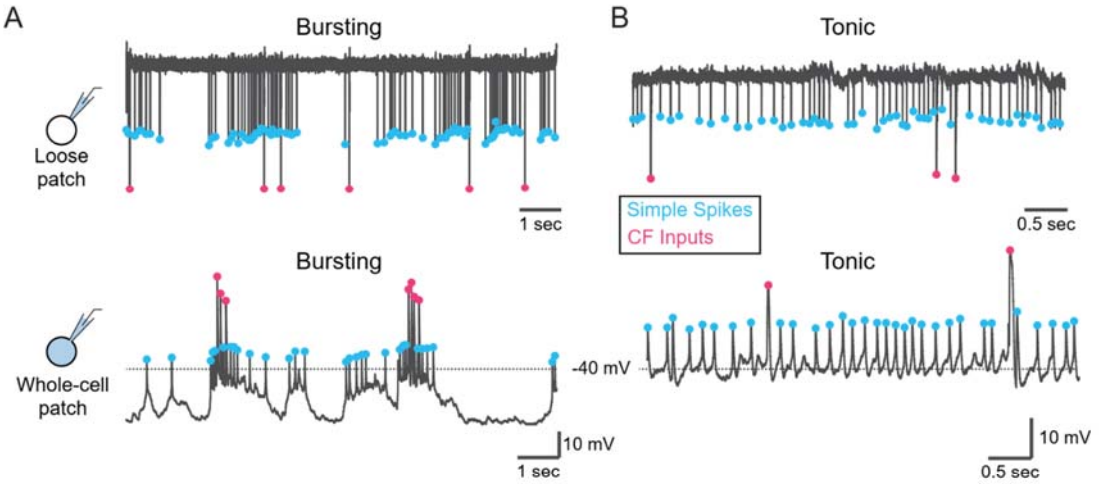
It is important to note that not all neurons of a kind are necessarily bistable and that bistability need not be a permanent property of a neuron, either. For instance, only ~20% of thalamic reticular neurons are bistable (Fuentealba et al., 2005), while about 67% of mitral cells are bistable (Heyward et al., 2001). Bistability can also be induced in response to neuromodulation, as in the case of motor neurons and DG neurons of the STG. Under control conditions, the DG neurons only fire for the duration of a current pulse. However, in the presence of the neuropeptide TNRNFLRFamide, they become bistable - alternating between quiescence and prolonged, self-sustained firing (Weimann et al., 1993) (Fig 1.3C). Similarly, serotonin has been shown to trigger bistability in various neurons - vertebrate motoneurons (Hounsgaard et al., 1988), R15 neurons of *Aplysia*, as well as Purkinje neurons (Williams et al., 2002). Serotonin activates a persistent inward current that is needed for sustained spontaneous firing (Raman & Bean, 1997; Schwindt & Crill, 1980). While the mechanism of action of serotonin has been uncovered, the ion channels that mediate a switch in cellular state in many other cases remain unknown.

1.3.2 Bistability in Purkinje neurons

Purkinje neurons of the cerebellum were first reported to be bistable in cerebellar slices in the 80s (Latham & Paul, 1971; Llinás & Sugimori, 1980). Since then, the existence of two different modes of PN firing has been shown in neuronal cultures (Gruol & Franklin, 1987) as well as *in vivo* (Loewenstein et al., 2005), although it was thought to be an artefact of anaesthesia (Schonewille et al., 2006). While PN bistability has been seemingly been

observed in awake monkeys, frogs and cats (Edgley & Lidierth, 1988; Loewenstein et al., 2005, 2006), the most unambiguous report of PN bistability comes from awake zebrafish (Sengupta & Thirumalai, 2015).

The bistability observed in these systems is not identical, though. In adult mammals, PNs are similar to plateau neurons, with a quiescent “down” state and an active “up” state. In larval zebrafish, however, the bistability observed is more similar to that of *Aplysia*’s R15 neurons, with the two modes being tonic firing and burst firing (Sengupta & Thirumalai, 2015) (Fig 1.4A and Fig 1.4B). In the tonic state, PNs rest at a relatively depolarized membrane potential of ~-40mV and continuously fire simple spikes, whereas in the burst state, neurons rest close to -65mV and have clusters of spiking separated by intervals of quiescence.


A
Bursting
Loose patch
1 sec
Bursting
Whole-cell patch
-40 mV
10 mV
1 sec
B
Tonic
Simple Spikes
CF Inputs
0.5 sec
Tonic
10 mV
0.5 sec

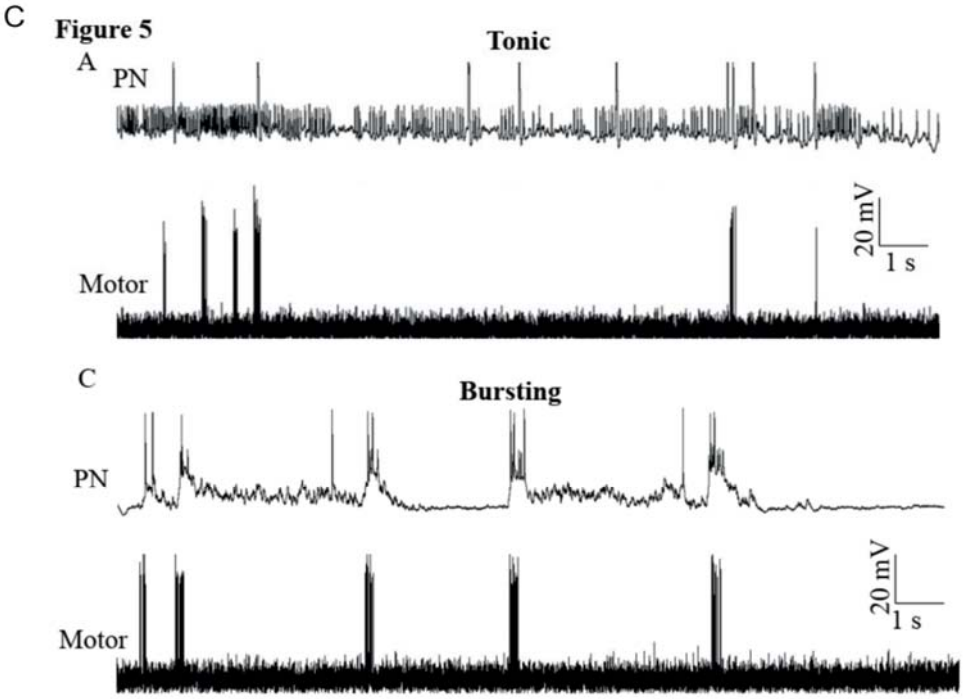

C
Figure 5
A
Tonic
PN
Motor
20 mV
1 s
C
Bursting
PN
Motor
20 mV
1 s

Figure 1.4 Bistability in Purkinje neurons of the larval zebrafish (A) Example recordings of zebrafish Purkinje neurons in the bursting state, obtained either extracellularly in the loose patch configuration (top) and intracellularly in the whole-cell configuration (bottom). Simple spikes (cyan) and CF EPSPs (magenta) can be distinguished from one another by their amplitudes. **(B)** Example recordings of zebrafish Purkinje neurons in the tonic state, obtained using loose patch (top) or whole-cell (bottome) recordings, organised as in (A). **(C)** Correlation of the activity of individual Purkinje neurons and simultaneously acquired fictive motor recordings. Note that there is no correlation between motor activity and simple spiking in tonic cells (top), but that bursts of activity in bursting cells (bottom) have a strong correlation with fictive swims. (Sengupta, 2016)

How Purkinje neurons switch their firing mode is not very clear, but it seems to be network-driven. In the complete absence of network activity, zebrafish PNs autonomously fire in the tonic state. This directly implicates the network in regulating bistability, and suggests that the complement of ion channels in a PN renders it depolarised and that hyperpolarisation from the network (presumably from the inhibitory input of local interneurons) is needed to make a cell quiescent/bursting. What's more, in rodents, the switch from the down state to the up state can be triggered by activation of the climbing fibre (Loewenstein et al., 2005). Other parts of the cerebellar network could also toggle a cell's state, such as molecular layer interneurons and Bergmann glia (Oldfield et al., 2010; Wang et al., 2012). Thus, in summary, network activity is critical to shape the bistable state of Purkinje neurons.

1.3.3 Possible functional consequences of bistability

It appears that phasic (burst) firing might have more biological significance than tonic firing across many systems. In larval zebrafish, for example, dopaminergic diencephalospinal neurons (DDNs) have been shown to fire both tonically and in bursts (Jay et al., 2015). The study also used paired recordings between DDNs and motor neurons/muscle fibers to show that periods of motor activity correlated with burst firing in DDNs, whereas periods of quiescence corresponded to periods of tonic activity. DDNs are homologs of midbrain dopaminergic neurons in adult mammals, which also exhibit mixed tonic and burst firing, although it is not clear whether both these states can exist in the same cell. A similar observation has also been made for Purkinje neurons, where bursts have been found to correlate with simultaneously-recorded fictive swims (Fig 1.4C). The utility of burst firing in these systems is yet to be understood. However, in the case of sensory processing, it is known that presynaptic burst firing increases the reliability of synaptic transmission at

unreliable synapses (Krahe & Gabbiani, 2004). Perhaps the same principle applies to DDNs and Purkinje neurons.

In certain cases, the bistability has been clearly associated with a function - such as in central pattern generation in the PY neuron of the STG (Dickinson & Nagy, 1983). The circuit consequences of Purkinje neuron bistability have not yet been understood, however. As hypothesised above, bistability could allow a cell to switch between performing alternative functions. It could also be a way for a cell to tune out of network activity, perhaps as a means to conserve energy.

It is interesting to note that many of the neurons reported to be bistable are involved in motor control, implying that bistability may be a feature that helps motor circuits perform their function smoothly. How this occurs remains to be seen.

1.4 Calcium imaging as a method to interrogate neural circuit function

Calcium imaging is a popular method of studying neuronal activity. A PubMed search with the terms “calcium imaging neuron” yielded 9,637 papers between 1975-2023 (as of March 29, 2023). The basis of this method is that action potentials in neurons activate voltage-gated calcium channels (VGCCs), thereby leading to calcium influx. The intracellular-extracellular concentration differential for calcium is very large, roughly 5 orders of magnitude ($[Ca^{2+}]_{in}$ ~ 50-100nM, whereas $[Ca^{2+}]_{out}$ ~ 2mM). So when VGCCs open, the intracellular calcium concentration can transiently rise to 10-100 times their baseline value (Grienberger & Konnerth, 2012), which makes it an excellent signal to detect. Imaging this signal, then, would be a way to unambiguously identify the source of neuronal activity in a population and monitor said activity at a high spatial resolution. This logic is what led to the development of calcium imaging.

The initial glimpses of the possibility of this method came around the 1970s, with the discovery of bioluminescent proteins like aequorin and obelin, whose luminescence depended on calcium concentration (Blinks et al., 1976; Moisescu et al., 1975; Robbins et al., 2021). Since these initial days, however, there have been a wide range of developments in (1) calcium sensors, (2) imaging systems (microscopy), (3) image analysis pipelines to extract calcium signals, and (4) algorithms to infer spiking activity from calcium signal time series. I summarise each of these below, to put into context work that will follow in this thesis.

1.4.1 Calcium sensors

Dyes or proteins whose fluorescence/bioluminescence is calcium-dependent have widely been adopted as sensors for calcium imaging. Bioluminescent proteins gave way to fluorescent dyes when BAPTA (an acronym for ***b***is(o-***a***mino***p***henoxy)ethane-N,N,N',N'-***t***etraacetic ***a***cid) and its quinoline derivatives were synthesised by Roger Tsien in the 1980s (Tsien, 1980). BAPTA is a very efficient, selective chelator of calcium, and its fluorescent derivatives Oregon Green BAPTA (OGB-1) and Calcium Green (CG1 and its dextran conjugates) are dyes which continue to be used today to monitor intracellular calcium dynamics. While OGB-1 and CG-1 are dyes whose fluorescence increases with increasing calcium concentration, thereby reporting relative changes in calcium concentration, dyes like Fura-2 and Indo-1 (Grynkiewicz et al., 1985) can report absolute concentration by the ratio of their fluorescence at two separate wavelengths (i.e. they are ratiometric indicators). Although these dyes continue to find uses in studying neurons in culture, they find limited use for *in vivo* functional imaging of neuronal populations during behaviour. This is because they are hard to deliver, require the use of two separate channels and rely on short-wavelength UV light that can lead to toxicity.

The arrival of fluorescent protein-based sensors which could be genetically encoded - genetically encoded calcium indicators (GECIs) - in the late 1990s led to a rapid growth in *in vivo* studies of neuronal function. The idea behind these sensors was to fuse the calcium-binding domain of the regulatory protein calmodulin (CaM) with various fluorescent proteins in configurations that suppressed their fluorescence under baseline conditions. When cytosolic calcium concentrations rise, however, the induced conformation change in the calcium-binding domain would then “release” the suppressed fluorescent signal (Fig 1.5A).

The first of these kinds of sensors were a family of FRET-based sensors called Cameleons (Miyawaki et al., 1997). Cameleons had two fluorescent proteins fused to them - one upstream and one downstream of the CaM domain. The key factor in the choice of the two proteins was that they were a FRET pair. In high calcium concentrations, the conformation change in the CaM domain would bring the FRET pair physically close to one another. (Fig 1.5Ai). The FRET-based design of the Cameleons limited experimental flexibility as two fluorescence channels needed to be monitored. This drawback was overcome with the arrival of GCaMP, an engineered protein that fused circularly permuted green fluorescence protein (cpGFP) with the CaM domain, which allowed the monitoring of calcium concentrations using a single fluorescence channel (Nakai et al., 2001) (Fig 1.5Aii). More than 20 years since the initial report of GCaMP, many generations of new GCaMPs continue

to be produced, with the latest in the series being the 8th generation (Zhang et al., 2023). Each version has been improved from its previous versions by iterative engineering to select for properties desirable for *in vivo* imaging - higher signal-to-noise ratios, faster reporter kinetics, high protein expression levels, linear calcium concentration-dependence, and so on (T.-W. Chen et al., 2013; Dana et al., 2019). The majority of these sensors continue to be primarily derivatives of GFP, and are hence green, but the colour palette of these sensors has been expanded to include more red-shifted reporters, like jRCaMP and jRGECO (Dana et al., 2016).

1.4.2 Imaging systems

Advances in microscopy methods have also propelled calcium imaging to becoming a staple of the neuroscientist's toolbox. In the initial days, conventional widefield epifluorescence microscopy was used to capture images of fluorescence signals from neuronal cultures or thin brain slices. In this method (Fig 1.5Bi), the sample is flooded with light of the excitation wavelength, and all the emission light obtained is collected and imaged using a camera. While this allows for fast imaging, images obtained have a lot of out-of-focus signals (Fig 1.5Bi). The development of optical sectioning methods like laser scanning confocal microscopy allowed experimenters to filter out-of-focus light and get clearer images with better resolution of depth. Simple confocal microscopy posed many problems for *in vivo* imaging, ranging from bleaching of the sensor due to overexposure and limited access to deeper tissues because of strong scattering of light in the visible wavelength range and phototoxicity. The advent of 2-photon microscopy solved several of these problems (Denk et al., 1990, 1994), which is why it became a very popular microscopy method among neuroscientists. In two-photon imaging, light of double the wavelength of the required excitation is used (Fig 1.5Bii, red area). At the plane of focus alone, a fluorophore can absorb two photons and thus get excited (Fig 1.5Bii, blue area). Thus, by stimulating regions point-by-point, phototoxicity is limited and the source of any emitted fluorescence signal (Fig 1.5Bii) that is obtained can be localised to that point. A major limitation of two-photon microscopy is that it uses point-by-point scanning, which restrains the maximum sampling speed. Selective plane illumination microscopy (SPIM), also known as light sheet microscopy (Huisken et al., 2004), was designed to overcome this problem. In this configuration, the excitation light (Fig 1.5Biii, blue) is delivered to the sample as a thin sheet, thereby stimulating many points in the focal plane simultaneously. The sheet can be made very thin (~2-10 μm width), which can allow imaging even at a single-cell resolution. The speed and resolution that light sheet microscopy provides has allowed researchers to successfully image the whole brain of larval zebrafish during behaviour (Ahrens et al., 2013; Vladimirov et al., 2014, 2018).

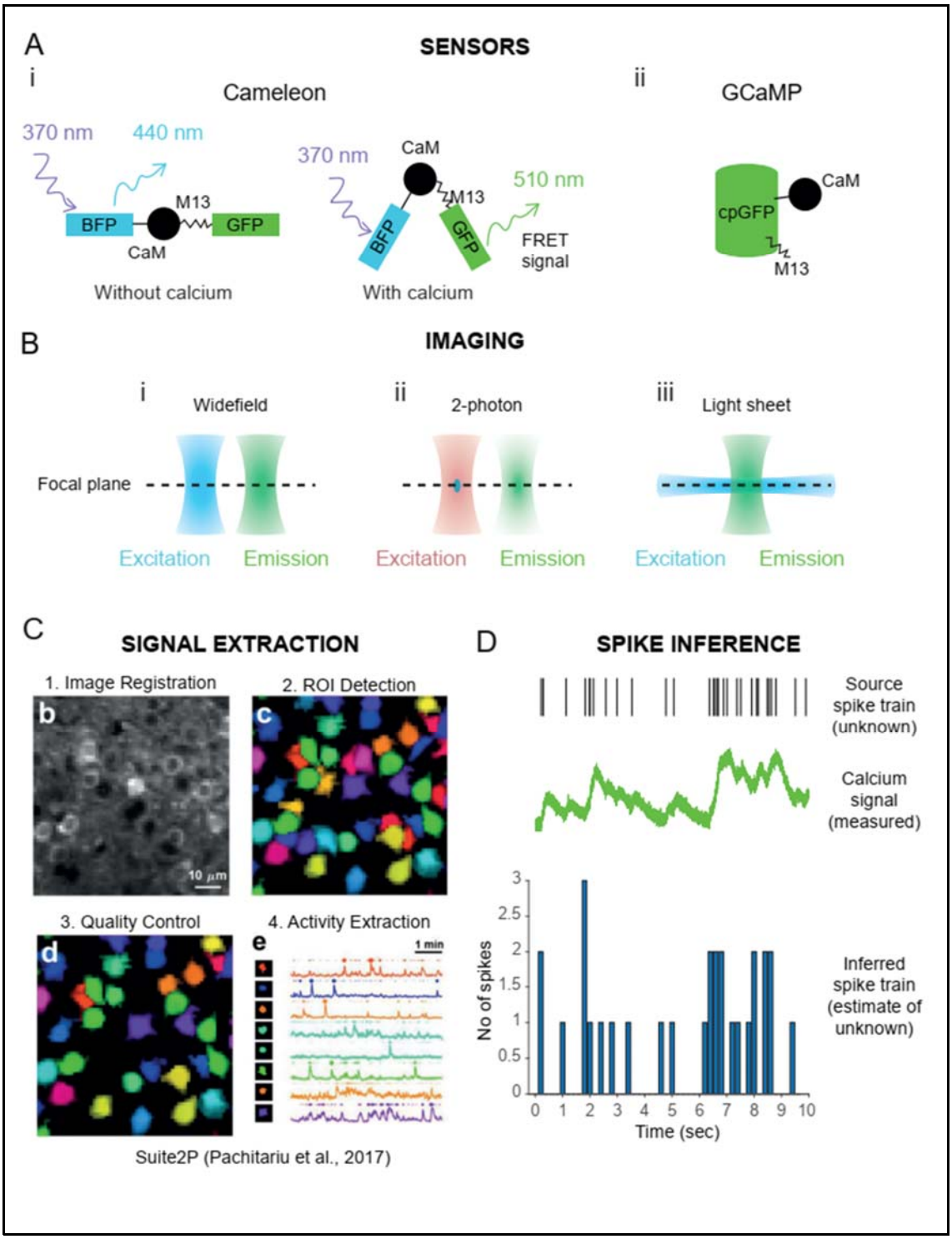

A
SENSORS
i
Cameleon
370 nm
440 nm
BFP
M13
GFP
CaM
Without calcium
370 nm
CaM
M13
510 nm
BFP
GFP
FRET
signal
With calcium
ii
GCaMP
CaM
cpGFP
M13
B
IMAGING
i
Widefield
Focal plane
Excitation
Emission
ii
2-photon
Excitation
Emission
iii
Light sheet
Excitation
Emission
C
SIGNAL EXTRACTION
1. Image Registration
2. ROI Detection
b
c
10 μm
3. Quality Control
4. Activity Extraction
d
e
1 min
Suite2P (Pachitariu et al., 2017)
D
SPIKE INFERENCE
Source
spike train
(unknown)
Calcium
signal
(measured)
No of spikes
Inferred
spike train
(estimate of
unknown)
Time (sec)

Figure 1.5 A summary of methods and developments in calcium imaging (A) Various genetically encoded calcium indicators (GECIs) have been developed to report intracellular calcium concentrations. **(i)** The first of these were the Cameleons (left), which are FRET-based reporters. In control conditions, the FRET pair fluorescent proteins upstream and downstream of the calcium-binding domain of calmodulin (CaM), are far apart. Calcium binding leads to a conformational change that brings the two fluorescent proteins close enough to yield a strong FRET signal. **(ii)** The latest calcium sensors mostly belong to the GCaMP family (right), which is derived from green fluorescent protein (circularly permuted GFP, cpGFP). Calcium binding leads to a conformational change in the sensor that increases its fluorescence. **(B)** The next step after expressing a calcium-sensitive sensor in cells is to image them. There are three main imaging methods popular in neuroscience - widefield imaging (left), two-photon imaging (middle) and light sheet imaging (right). The profiles of excitation and emission light are shown in this panel. In both widefield and two-photon imaging, the excitation and emission light paths are along the same axis, while for light-sheet imaging, they are orthogonal to one another. **(i)** Widefield imaging has a very broad cone of excitation (blue), and fluorophores that are outside the plane of focus are also exposed to excitation light. Thus, the emission light (green) that is collected also comes from fluorophores that are out of focus. **(ii)** To overcome the drawbacks of widefield imaging, two-photon microscopy was developed. In this method, light of double the required wavelength is used for excitation (red). Only when two photons are absorbed is the fluorophore excited (blue). This occurs only when enough light is concentrated, i.e. a point at the plane of focus. The emission signal (green) is collected through the same objective, but can be localised as having come only from that point in the plane of focus. **(iii)** In light-sheet imaging, a thin pencil of light is used as the source of excitation (blue). The fluorescence signal emitted from any excited fluorophore in this plane of excitation is collected using a separate set of optics placed orthogonal to the sheet of light (green). Thus, a localised field of excitation and emission are simultaneously collected, as opposed to the point-by-point excitation in two-photon microscopy. **(C)** After acquiring sequences of images, the next step is to identify cells and extract their calcium signals. This typically requires (1) image registration to correct for motion artefacts, (2) detection of regions of interest (ROIs) within the imaging frame, (3) verifying that the detected ROIs are, indeed, real cells, and (4) extracting the intensities from each ROI over time to obtain the final calcium activity profile for the cell. **(D)** The spike train of each source cell (black, top) is the unknown property of interest. The measured signal is the calcium sensor's fluorescence variation (green, middle). The goal of spike inference is to use the measured calcium signal to algorithmically estimate the

underlying spike train (blue histogram, bottom).

1.4.3 Signal extraction pipelines

Once data have been acquired, an experimenter must segment cells in the microscopy images and extract their fluorescence signals. This can be achieved in a number of ways, the simplest of which is to take a projection (either average or maximum) of the entire stack of images. Cells that were active at any time during the experiment would light up in this average image, and individual cells can be manually separated by drawing regions of interest. Alternatively, if possible, the experimenter can co-express a calcium insensitive fluorescent protein along with the calcium sensor. This fluorescence channel acts as a marker to segment cells, following which the calcium sensor's pixel intensities can be extracted. Other, more sophisticated, methods involving the use of pixel correlation analyses or matrix factorisation methods have also been developed to extract fluorescence signals from cells, as described below.

Given the challenges of using some of these methods *in vivo*, where it might be difficult to express multiple proteins simultaneously, or where an animal's movement can blur the image at the time of acquisition, work from several groups has led to advances in the engineering of algorithms for motion correction, image registration, cell segmentation and signal extraction. They have been put together into publicly accessible software packages, of which CaImAn (Giovannucci et al., 2019) and Suite2P (Pachitariu et al., 2017) (Fig 1.5C) are especially popular for use with large datasets acquired using 2-photon microscopy.

1.4.4 Spike inference methods

Neuronal activity in the form of action potential firing rates is what neuroscientists are typically trying to measure, for which the calcium signal is only a proxy. Thus, after acquiring a calcium signal time series, an experimenter, more often than not, would like to infer the underlying firing rate variations (Fig 1.5D). This process is called spike inference and it has been a major problem that neuroscientists are still trying to solve.

It is well-known that the calcium response of a neuron is a low-pass filtered version of its activity. In fact, this signal can be thought of as the spike events convolved with a calcium sensor kernel, with some added noise (Vogelstein et al., 2010). Since the calcium signal is produced by convolution of the spike train with a kernel, naturally one would expect that deconvolution of the calcium signal would yield back the original spike train. This reasoning (along with the added constraint that spike occurrence cannot take negative values) led to

the development of a fast, non-negative deconvolution (NND) strategy to recover the underlying spike train from a calcium signal time series (Vogelstein et al., 2010). While seemingly simple, the NND algorithm has been shown to be robust and work comparably well across datasets (Pachitariu et al., 2018). A version of this has been implemented by the CaImAn package (Giovannucci et al., 2019; Pnevmatikakis et al., 2016), and another fast implementation of the algorithm, called OASIS (Friedrich et al., 2017) has been adopted for use in the Suite2P package (Pachitariu et al., 2018).

Alternate methods of spike inference have also been developed, which implement methods ranging from Bayesian maximum likelihood estimates (MLspike) (Deneux et al., 2016) to supervised machine learning methods (CASCADE; SPIKEFINDER dataset; (Rupprecht et al., 2021; Theis et al., 2016) to generative models (Berens et al., 2018). Notably, a crowd-sourcing effort called *spikefinder* tried to develop and benchmark novel methods of spike inference, many of which used machine learning methods (Berens et al., 2018).

1.4.5 Limitations of calcium imaging and spike inference

Calcium imaging has several inherent limitations. For one, the variations in intracellular calcium concentration are only a proxy for neuronal activity. Furthermore, a calcium transient is obtained only if VGCCs hit their activation threshold, which is usually significantly higher than a cell's resting potential. This means that subthreshold activity (post-synaptic potentials, for instance) and inhibition are poorly reported by calcium imaging. Lastly, it is much slower than electrophysiology, which is another gold standard method of querying neural activity.

Spike inference algorithms also come with their own set of drawbacks and must be used with caution as they do not always yield consistent results. Different methods have varying performances, depending on the peculiarities of the dataset and the neuronal type being considered (Evans et al., 2020; Pachitariu et al., 2018). Furthermore, deconvolution only provides an approximation of the underlying firing rate by providing relative numbers of spikes. While this can be useful to capture stimulus-triggered variations, it misses out on other cell-intrinsic dynamics like membrane bistability. Lastly, spike inference algorithms assume that each action potential leads to a proportional, or linear, increase in sensor fluorescence. This might be a reasonable assumption, but doesn't always hold true, especially for high firing rates (T.-W. Chen et al., 2013). Non-linearities in the relationship between calcium concentration and GCaMP fluorescence mean that deconvolution, which is a linear method, isn't always directly interpretable.

These are limitations that can only be overcome by the development of voltage imaging, which might allow experimenters access to the membrane potential. Voltage imaging is still in its nascent stages, though, and the technology required for it hasn't yet advanced enough to make it easy to use across systems. Furthermore, the membrane potential is a signal across the surface of the plasma membrane (2D), whereas the rise in calcium concentration is a signal throughout the cytosol (3D). Therefore, the calcium signal will always be an easier signal to detect, and it is unlikely to become obsolete, despite its limitations.

1.5 Challenges in studying the function of bistability

Despite its occurrence in both vertebrates and invertebrates alike, bistability remains a poorly explored phenomenon. One major reason has been the choice of model system. Rodents and primates have extensively been used to study various bistable neurons. However, it is still difficult to acquire information from entire populations of neurons in these animals while they perform natural behaviours. Novel technologies like Neuropixels probes (Jun et al., 2017; Steinmetz et al., 2021) and advanced surgical methods for stable long-term recordings are trying to overcome this. Even so, the *ex utero* development of larval zebrafish and the fact that they are amenable to functional population imaging while awake and behaving means that they might be the ideal model system to explore the development and function of neuronal bistability *in vivo*.

Another major challenge has been the use of imaging based methods like functional calcium imaging in the study of bistability. Calcium imaging only reports relative changes in spiking frequency without directly reporting a neuron's membrane potential. And while a number of tools for spike inference have been developed and are currently being used, no tool exists for converting calcium imaging signals to maps of cellular state in bistable neurons.

This thesis aims to resolve both these problems by investigating the function of bistability in the cerebellar Purkinje neurons of the larval zebrafish. It also outlines the development of a machine learning tool that is able to classify calcium imaging traces as arising from either tonic or bursting cells. This tool offers a new way of analysing calcium imaging data from bistable neurons to understand how they participate in network computation and natural behaviours.

1.6 Machine learning tools in the biological sciences

1.6.1 Machine Learning Foundations

Machine learning is an interdisciplinary field that is an amalgamation of computer science and several disciplines of mathematics like statistics, linear algebra and optimisation. It has grown exponentially over the past decade and has had a significant impact on society in recent times. The goal of machine learning was to have a machine autonomously learn patterns from data without explicitly coding for said patterns. The origins of this idea go back to the 1950s, when Alan Turing asked whether machines could be “intelligent” and think like humans do. Since then, the field has made substantial progress in both technology and concepts (Schmidhuber, 2022).

There are four types of machine learning algorithms - unsupervised, supervised, reinforcement learning, and generative AI (*What Is Machine Learning?*, 2023). Unsupervised learning algorithms try to assign labels to unlabelled data and include methods like clustering, anomaly detection and dimensionality reduction. On the other hand, supervised learning algorithms are given training data with associated class labels, which are used iteratively to build a model of said data. Supervised learning includes a wide range of methods, such as regression, support vector machines and artificial neural networks. Reinforcement learning is similar to unsupervised learning in that no explicit labelled data is used. However, these models are not used to find underlying patterns, but instead given a target goal. They are then iteratively trained using “rewards” and “punishments” to achieve said goal with minimal error. Lastly, generative AI is a class of machine learning algorithms that use input from a user to produce novel content. Of these types of machine learning, supervised learning using neural networks is perhaps the most widely used.

Artificial neural networks (ANNs), often shortened to “neural networks” or “neural nets”, are a kind of machine learning model loosely inspired by how real neurons perform computations (Greener et al., 2022). Biological neurons typically receive many synaptic inputs of different strengths, which are added together (Fig 1.6A). If the sum of these signals reaches a threshold value, the neuron fires an action potential. In a similar fashion, the first ANN model - the perceptron (Rosenblatt, 1958) (Fig 1.6B) - was a binary classifier. A perceptron takes many numbers as an input, multiplies each of them by a unique weight, and then adds up all the resulting values. If this summed value is greater than a threshold, the perceptron outputs a 1, otherwise it outputs a 0. A perceptron could be made to produce continuous outputs by simply modifying this thresholding function. Since the “neuron” is thought to be “activated” if it hits threshold, the thresholding function is also called the “activation function”.

The “learning” part of the perceptron is as follows. First, all the weights for the perceptron are randomly initialised (or set to zero). For each example in the dataset, the input features and weights are multiplied and then added to produce an output value. Each output yielded in the dataset is compared to its corresponding ground truth value. The difference between the two is the error. Each input weight is then updated by adding the product of this error, the input and a scaling factor called the “learning rate” to the original weight. This is performed sequentially for every point in the dataset and the entire process iteratively performed until the model converges.

While initially promising, the perceptron was proven to be unable to learn patterns that were not linearly separable. This problem was solved by stacking several layers of perceptron units together (Rosenblatt, 1958) (Fig 1.6C), creating a “multilayer perceptron” with extra “hidden” layers of processing between the input and the output. Training these multilayer perceptrons needed the discovery of the backpropagation algorithm to transfer errors through the various layers and update their weights strategically (Linnainmaa, 1976). Multilayer perceptron models have since been extended to “deep” networks with many layers capable of not just producing binary classifications, but also complex, continuous outputs.

Their modular nature means that they can have any number of layers, and that individual units (“neurons”) in each layer can be modified to take inputs of different types (text, images, audio, etc.), and even have different activation functions. The diversity of neural network architectures is staggering, but certain patterns emerge. For example, convolutional neural networks (CNNs), i.e. neural networks which don’t just take weighted sums but rather perform convolutions of inputs to produce their outputs, are particularly good at solving problems related to image analysis. Similarly, recurrent neural networks, i.e. those with outputs from a neuron can be taken back as an input to the same neuron, or neurons in a “preceding” layer, are particularly good at solving problems related to pattern detection in time series (*What Are Recurrent Neural Networks?*, 2023).

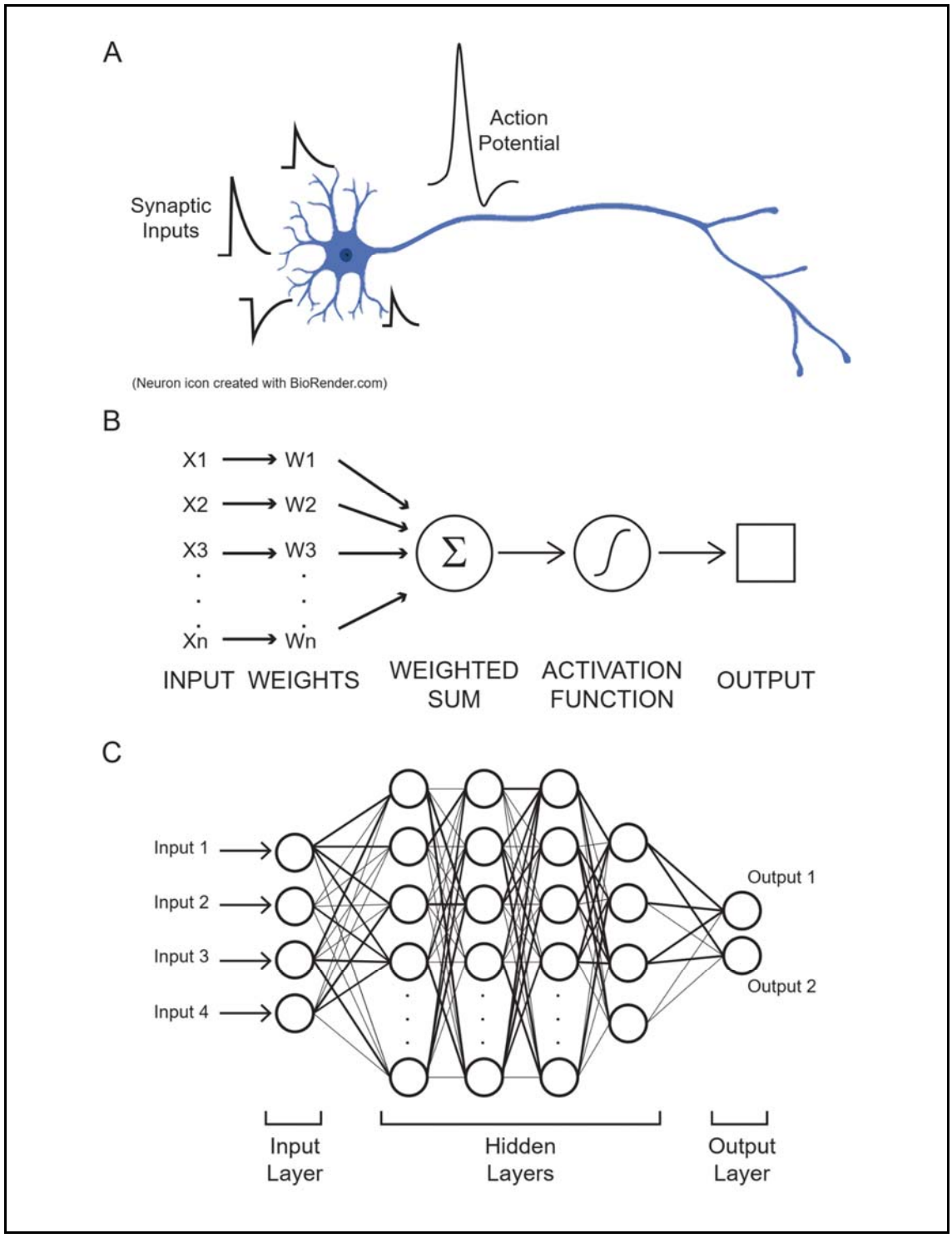

A
Action
Potential
Synaptic
Inputs
(Neuron icon created with BioRender.com)
B
X1
W1
X2
W2
X3
W3
Xn
Wn
Σ
INPUT
WEIGHTS
WEIGHTED
SUM
ACTIVATION
FUNCTION
OUTPUT
C
Input 1
Input 2
Input 3
Input 4
Output 1
Output 2
Input
Layer
Hidden
Layers
Output
Layer

Figure 1.6 Biology-inspired artificial neural networks (A) A schematic of a typical neuron that receives many synaptic inputs (black deflections on the left) - both excitatory (upward deflections) and inhibitory (downward deflections). If the sum of all these inputs hits threshold, the neuron fires an action potential. **(B)** The perceptron model takes many input numbers (X1, X2, …, Xn), each of which is multiplied by its unique weight (W1, W2, …, Wn) and then added up. The output of the perceptron is determined by the value of the "activation function" for the obtained weighted sum. In the simplest case, the activation function is a thresholding function that yields 0 for every input value below threshold and 1 for every input value above threshold. **(C)** A graph representation of an example multilayer perceptron, which is a classic feedforward neural network. Each layer is made of multiple nodes which are perceptron "neurons", i.e. they output a thresholded weighted sum of their inputs. The outputs of neurons in one layer get passed on as inputs to neurons in the next layer. Each such connection between neurons is represented as an edge between two nodes, the thickness of which corresponds to the magnitude of its weight. Between the "Input" and "Output" layers are several "Hidden" layers of processing. The network can yield one or more outputs, which can be interpreted by the end user in a meaningful way.

Multilayer perceptrons have successfully been used to solve many nonlinear problems. In fact, "deep networks", i.e. those with more than one hidden layer, can be trained to learn arbitrary associations of latent information in data. They have been used to accurately recognise handwritten digits in images (LeCun et al., 1989; Ciresan et al., 2010), for example, among many other achievements whose summary is beyond the scope of this text. One of the most outstanding achievements of trained deep networks that is worthy of note has been their ability to beat humans at complex strategy games like chess and Go (Silver et al., 2018). Also worth a special mention is the most recent deep network that has taken the world by storm - the language model ChatGPT/GPT-4 (OpenAI, 2023). Thus, machine learning/artificial intelligence is already showing great potential in solving many problems pressing humankind.

1.6.2 Machine learning in the life sciences

The size of data generated by life scientists is growing, and the power of machine learning is being leveraged to mine large data and facilitate discoveries in the life sciences. An exhaustive list of applications of ML in the life sciences is beyond the scope of this text, but I highlight certain examples here and in Table 1.1.

A landmark achievement in life science research came when Google released AlphaFold, a program that is able to accurately predict the 3D structure of a protein given its sequence of amino acids (Jumper et al., 2021). This tool has solved the protein folding problem to a great extent, ushering a new era of research in structural biology. Another powerful application of neural networks to find patterns in data invisible to the human eye is in the healthcare industry. Tools have been developed that could help with predicting cancer risk (Yala et al., 2021), to identify viable embryos from *in vitro* fertilisation (Bormann et al., 2020) and to screen for useful drugs (Ding & Zhang, 2021).

Machine learning has found widespread use in neuroscience, too, which has been a motivation behind this thesis. In ethology, various ML algorithms have been developed to segment videos to estimate the pose of a behaving animal, of which DeepLabCut is a notable example (Mathis et al., 2018). It is a very deep neural network, with 50 layers (or more), and is capable of estimating the pose of animals without the need for any reflective or high-contrast markers applied onto the animals themselves. Innovations in machine learning have also propelled systems neuroscience research. Programs like CaImAn and Suite2P have adopted neural networks to classify ROIs extracted from functional calcium imaging data as true cells or not (also see the “Calcium imaging as a method to interrogate neural circuit function” section above). Similarly, many of the top performing algorithms for spike train estimation from 2-photon calcium imaging were based on machine learning methods (Berens et al., 2018).

In summary, the number of applications of artificial intelligence in the life sciences is growing rapidly and enabling the solution of numerous difficult problems. The second half of this thesis poses one such challenging problem in the study of neuronal bistability and solves it using machine learning.

Table 1.1 - Examples of the applications of machine learning in life sciences research

Name of tool	**Type of Neural Network Used**	**Objective/Purpose**	**Reference**
Structural Biology			
AlphaFold	Deep “Self-attention” based neural network	Predicting protein structure based on the given amino acid	https://doi.org/10.1038/s41586-021-03819-2 and

		sequence	https://doi.org/10.1126/science.abj8754
Atomic Rotationally Equivariant Scorer (ARES)	Deep neural network	Predicting RNA secondary structures from its linear nucleotide sequence	https://doi.org/10.1126/science.abe5650
Medicine/Healthcare			
IDx-DR	k-Nearest Neighbour classification, followed by a multilayer convolutional neural network	Diagnose diabetic retinopathy from retinal images of patients	https://doi.org/10.1038/s41746-018-0040-6
-	Deep convolutional neural network (CNN)	Image classification for increasing IVF success rate	https://elifesciences.org/articles/55301
DeepBAR	deep generative models	calculates the binding affinities between drug candidates and their targets	https://doi.org/10.1021/acs.jpclett.1c00189
Neuroscience			
DeepLabCut	(Deep) Residual Networks (ResNet), CNN	Markerless pose estimation and tracking	https://www.nature.com/articles/s41593-018-0209-y
JAABA	GentleBoost	Classification of animal trajectories	https://www.nature.com/articles/nmeth.2281
CaImAn (submodule)	4-layer CNN	Used to classify true and false positives	https://elifesciences.org/articles/38173#s

		from segmented images from 2P imaging data	4
CASCADE	Deep CNN	Spike inference from calcium imaging data	https://doi.org/10.1038/s41593-021-00895-5

Chapter 2 - State-dependent relationships between electrophysiological events in Purkinje neurons

2.1 Introduction

The cerebellum is one of the oldest, most well-conserved regions of the vertebrate central nervous system. It has also received a lot of attention because of its stereotyped, repetitive architecture and because it houses what is arguably one of the most beautiful neurons to look at - the Purkinje neuron. These neurons have been studied in mammals for more than a century, dating back to Ramón y Cajal's work in the early 1900s.

Although these neurons have been well studied, the bulk of our understanding comes from *in vitro* studies, either from cultured neurons or from brain slices, and not always in physiologically relevant contexts. The larval zebrafish is an ideal model system to study these cells *in vivo* during development and during active behaviour. Larval zebrafish are mostly transparent when young, exhibit a diverse behavioural repertoire and have brains that are easily accessible as they are not hidden under a skull. Previous studies from our lab have investigated the development and physiology of these neurons. The study by Sengupta and Thirumalai (2015) was the first to report the electrophysiological properties of Purkinje neurons in larval zebrafish. It characterised the activity of zebrafish Purkinje neurons, showing that they were spontaneously active even in the absence of synaptic input and had two kinds of electrical activity - simple spikes and CF EPSPs, of distinct amplitudes. Furthermore, it showed that Purkinje neurons in awake larval zebrafish exhibit membrane potential bistability, either firing tonically or in bursts.

Despite recent progress in characterising the dynamics of these neurons, several questions remain unanswered. The main question that motivates further study in this area is whether changes in cellular state influence the temporal patterns of spiking activity, and how these patterns are modulated by the strong excitatory climbing fibre inputs. In order to address these questions, one needs to look deeper into the relationship (i.e. frequency of occurrence) between simple spikes and CF inputs. In order to calculate these relative statistics, many cells need to be looked at since there is considerable variability across cells. Therefore, I compiled all the recordings from individual Purkinje neurons that have been acquired in the lab into a single standardised dataset and then performed spike-train-level analyses on these data to work out the associations between the events in a Purkinje neuron.

2.2 Results

2.2.1 Compiling a dataset of zebrafish Purkinje neuron recordings

In order to analyse event statistics across Purkinje neurons as a function of their state, I first compiled electrophysiological recordings from various members of the lab into one large dataset. This dataset has a total of 173 recordings, both extracellular loose-patch recordings and intracellular whole-cell recordings (see Table 2.1 below), of Purkinje neurons from larvae between 6-8 dpf (days post fertilisation).

After compiling these recordings, I detected electrical events in each recording regardless of bistable state (Fig 2.1). There are two kinds of events in Purkinje neurons, which can be distinguished by their amplitude. Small amplitude events are simple spikes (Fig 2.1, cyan), and large amplitude events are climbing fibre (CF) inputs (Fig 2.1, magenta).

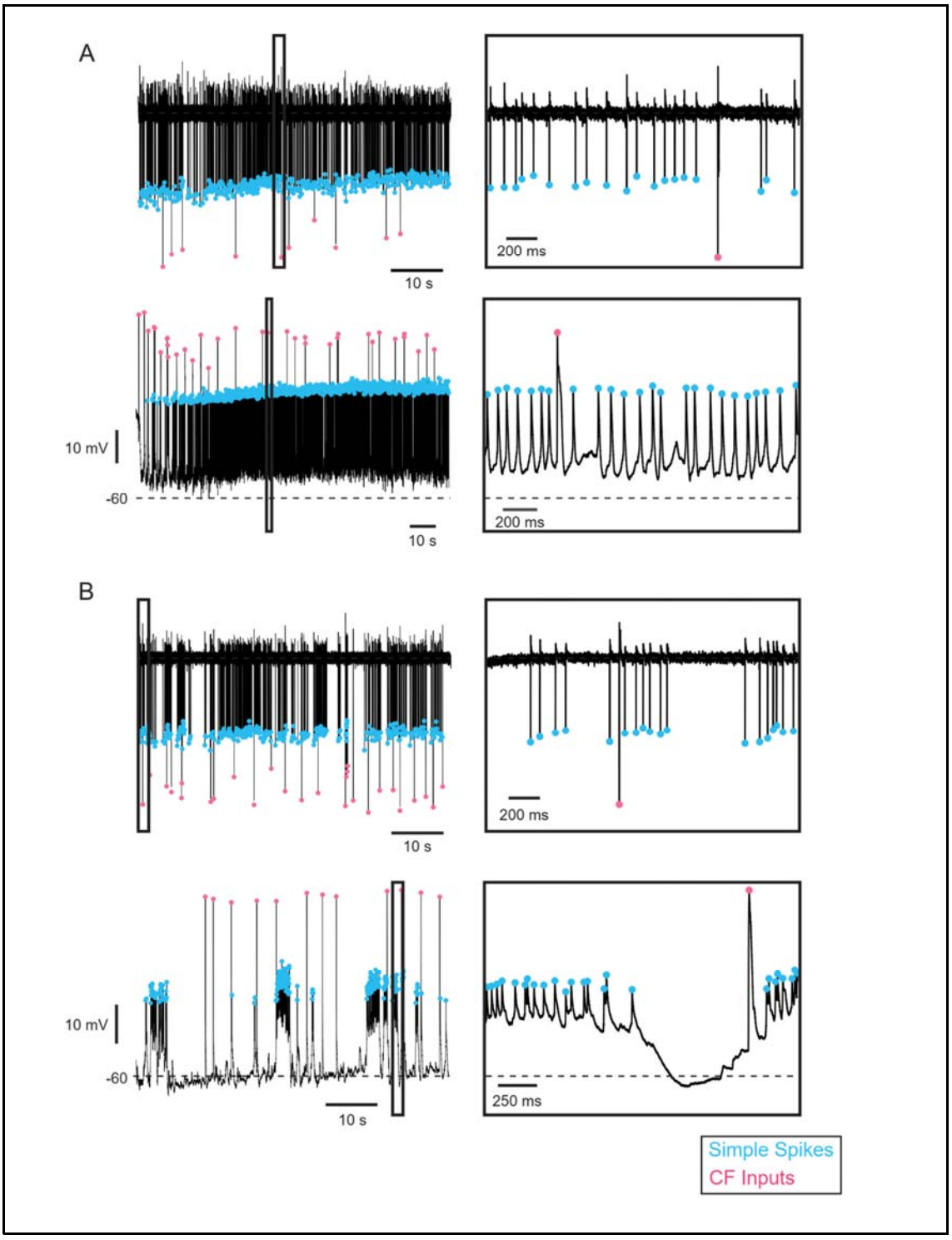

A
10 s
200 ms
10 mV
-60
10 s
200 ms
B
10 s
200 ms
10 mV
-60
10 s
250 ms
Simple Spikes
CF Inputs

Figure 2.1 - Detecting simple spikes and CF inputs from patch clamp recordings from Purkinje neurons in larval zebrafish (A) Event detection for an example Purkinje neuron in the tonic state - one from a juxtacellular recording (top) and one from a whole-cell recording (bottom). **(B)** Event detection for Purkinje neurons in the bursting state, from a juxtacellular recording (top) and a whole-cell recording (bottom), like in (A). For both (A) and (B), a small portion of the original recording is zoomed in (shown in the box on the right of each trace and marked on the trace on the left). Detected events are marked with circles - cyan for simple spikes and magenta for CF inputs.

Having identified the two kinds of events in the compiled dataset of recordings, they were then assigned a cellular state - either tonic, bursting, switching from one state to another, or ambiguous. Classification was first done by visual inspection (Fig 2.2A and B). These classifications were then further verified by inspecting the inter-event intervals of simple spikes. All inter-event intervals for tonic cells would form a single cluster, while bursting cells would have inter-event intervals that trailed to large values, which correspond to the inter-burst intervals (Fig 2.2C). Lastly, the coefficient of variation of the inter-spike intervals was used as an extra parameter of interest. This coefficient of variation is expected to be larger in bursting cells than the more regularly spiking tonic cells (Fig 2.2C and Fig 2.2D). Given that there is a lot of variability across cells and recordings, no single cutoff value could be determined for these parameters (like in Fig 2.2D). Instead, the three pieces of information were used in conjunction to yield the final state classification.

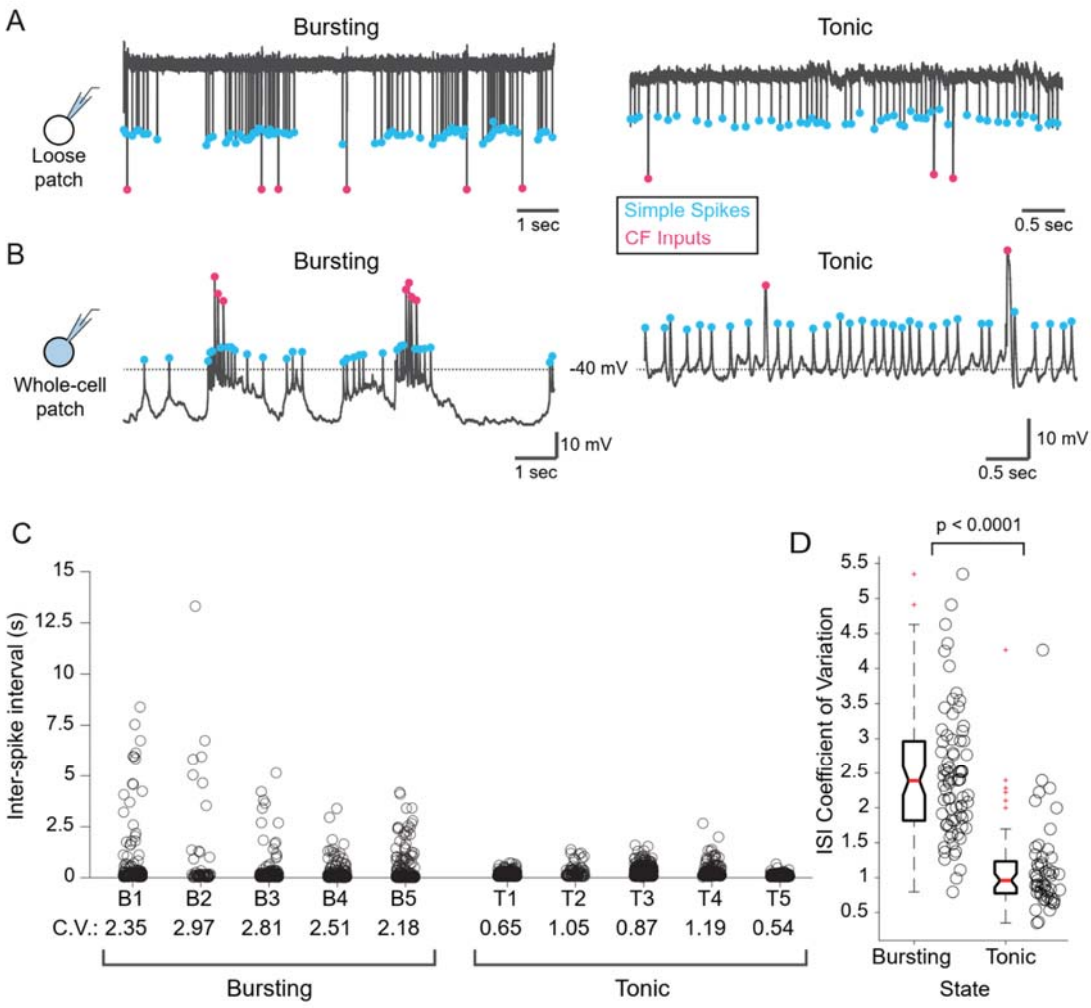


Figure 2.2 - Classification of larval zebrafish Purkinje neuron state as tonic or bursting (A) Representative patch clamp recordings in the loose patch configuration from a Purkinje neuron in the bursting state (left) and tonic state (right). Simple spikes are marked in cyan and CF inputs in magenta (also see Fig 2.2). **(B)** Representative whole-cell patch clamp recordings from Purkinje neurons in the bursting state (left) and the tonic state (right). Events are marked in the same colour scheme as in (A). **(C)** Example plots of simple spike inter-spike intervals from individual recordings. Five example cells are shown from each class (Bursting and Tonic). The coefficient of variation of this distribution is marked under each cell as “C.V.”. **(D)** The distribution of inter-spike interval coefficients of variation across all n=131 cells in the compiled dataset, sorted by state. The distributions are statistically significantly different, tested using a Mann-Whitney U-Test.

The types of recordings and their associated states are summarised in Table 2.1. Those recordings which had an ambiguous state (i.e. difficult to classify confidently as either tonic or bursting; n=25) were excluded from all further analyses. Similarly, those recordings which were of poor quality, i.e. where resting membrane potentials were higher than -40mV, or lower than -75mV or where events could not be identified cleanly, were excluded from further analysis (n=12). This left a total of 136 recordings for further analyses, of which 131 were either tonic (n=53) or bursting (n=78) alone, and 5 had a switch in state. These recordings were of quite a diverse range of durations, given that they were collected by different experimenters for different sets of experiments. The mean duration of the recordings was 72.2 seconds, with a standard deviation of 46.75 seconds, and the median duration of all the recordings was 60 seconds.

Table 2.1 - Summary of Purkinje neuron recordings in the compiled electrophysiology dataset, grouped by state (rows) and type of recording (columns)

	Extracellular	Intracellular	**Total**
Bursting	39	39	**78**
Tonic	11	42	**53**
Subtotal (included)	**50**	**81**	**131**
Switch	0	5	**5**
Ambiguous	0	25	**25**
Poor Quality	0	12	**12**
Grand Total	**50**	**123**	**173**

2.2.2 Quantifying Purkinje neuron event statistics and simulating realistic Purkinje neuron spike trains

The compilation of the electrophysiology dataset allowed me to inspect the relationships between the two electrical events within a Purkinje neuron. I first calculated event rates of simple spikes and CF inputs as a function of state, which are summarised in Table 2.2. I found that the average simple spike firing rates between tonic and bursting cells were statistically significantly different ($p=8.263 \times 10^{-4}$, *two-tailed* Mann-Whitney U-test). I noted that the average firing rate of bursting neurons was lower than that of tonic neurons. However, it was evident visually that the firing rates within bursts were higher than the typical tonic firing rate. Thus, I measured the average simple spike firing rate within bursts and compared it to the firing rate in tonic cells and found that in fact, the intraburst average simple spiking rate (12.46Hz) is almost double that of tonic cells (6.68Hz). This difference is statistically significant ($p=2.433 \times 10^{-11}$, *one-tailed* Mann-Whitney U-test). Unlike the simple spike firing rates, however, the average CF input rates between the two cellular states were not statistically significantly different (p=0.3702, *two-tailed* Mann-Whitney U-test). This implies that the inputs received by a neuron do not depend upon its state, consistent with the null expectation of the behaviour of synaptic input.

Table 2.2 - Average event rates as a function of cellular state

Average Simple Spike Rate (Hz)						
	Min	**Median**	**Max**	**Mean**	**SD**	**N**
All cells	0.23412	5.1714	17.969	6.1294	3.8777	131
Tonic	2.4009	6.6834	17.969	7.3628	3.5892	53
Bursting	0.23412	4.3767	15.765	5.2693	3.8608	78
Bursting (w/in bursts)	2.2145	11.491	133.33	13.5184	8.7113	1107 (bursts)
Bursting (w/in bursts)	3.9192	12.462	32.855	13.7262	5.9329	78 (cells)
Average CF Input Rate (Hz)						

All	0.086879	0.35764	15.987	0.5725	1.4136	131
Tonic	0.086879	0.32576	15.987	0.7306	2.2205	53
Bursting	0.11856	0.37726	2.5352	0.4712	0.3628	78

I next asked whether this information could be used to generate simulated spike trains that recapitulated the event statistics of biological Purkinje neurons. This is a useful exercise for several reasons. First, the ability to simulate spike trains of each state would allow us to better understand the dynamics of how each state encodes information. In addition, the ability to simulate spike trains with properties consistent with a cell's behvaiour in each state helps augments datasets that are difficult to acquire experimentally. These data are also useful for benchmarking novel algorithms, which will be useful later in this thesis.

One of the simplest models of neuronal spiking activity is to model it as a Poisson process (Gerstein & Mandelbrot, 1964). Although many more sophisticated models of spiking activity exist, I decided to first test the homogeneous Poisson process framework to model tonic firing in Purkinje neurons, as it is a model with very limited assumptions and only one parameter - the rate parameter, *lambda*. This parameter represents the average rate of events that the distribution tries to model.

To estimate the range of possible values for this parameter, I used the average simple spike firing rates reported above. The entire distribution of firing rates in tonic cells is shown in Fig 2.3A. Given that most of this distribution was between approximately 4Hz and 10Hz (10th percentile: 3.3974 and 90th percentile: 11.5498), I decided to use these as the limits of the parameter lambda for simple spikes.

The occurrence of CF inputs, being synaptic inputs, can also be considered to be a Poisson process. Similar to how it was done for simple spikes, I estimated the lambda for CF inputs from the distribution of average CF input rates (Fig 2.3B). For tonic cells, the 10th and 90th percentiles of average CF rate were 0.1850 Hz and 0.6383 Hz, respectively. Therefore, rounding off for convenience, I used 0.2 and 0.6 Hz as the limits for lambda.

To reconstruct SS and CF events based on their respective lambdas, I first needed to establish that the occurrence of these two events is independent. Thus, I tested whether there was any correlation between the average firing rates of simple spikes and CF inputs,

and found a weak, statistically insignificant negative correlation between the two (R = -0.1236, p = 0.4353) for tonic cells (Fig 2.3C). I obtained similar results for bursting cells (Fig 2.3F), which had a weak, statistically insignificant positive correlation (R = 0.0522, p = 0.6677). Hence, both the simple spike and CF input spike trains were treated as independent, and each of their lambda values were randomly picked from the aforementioned limits.

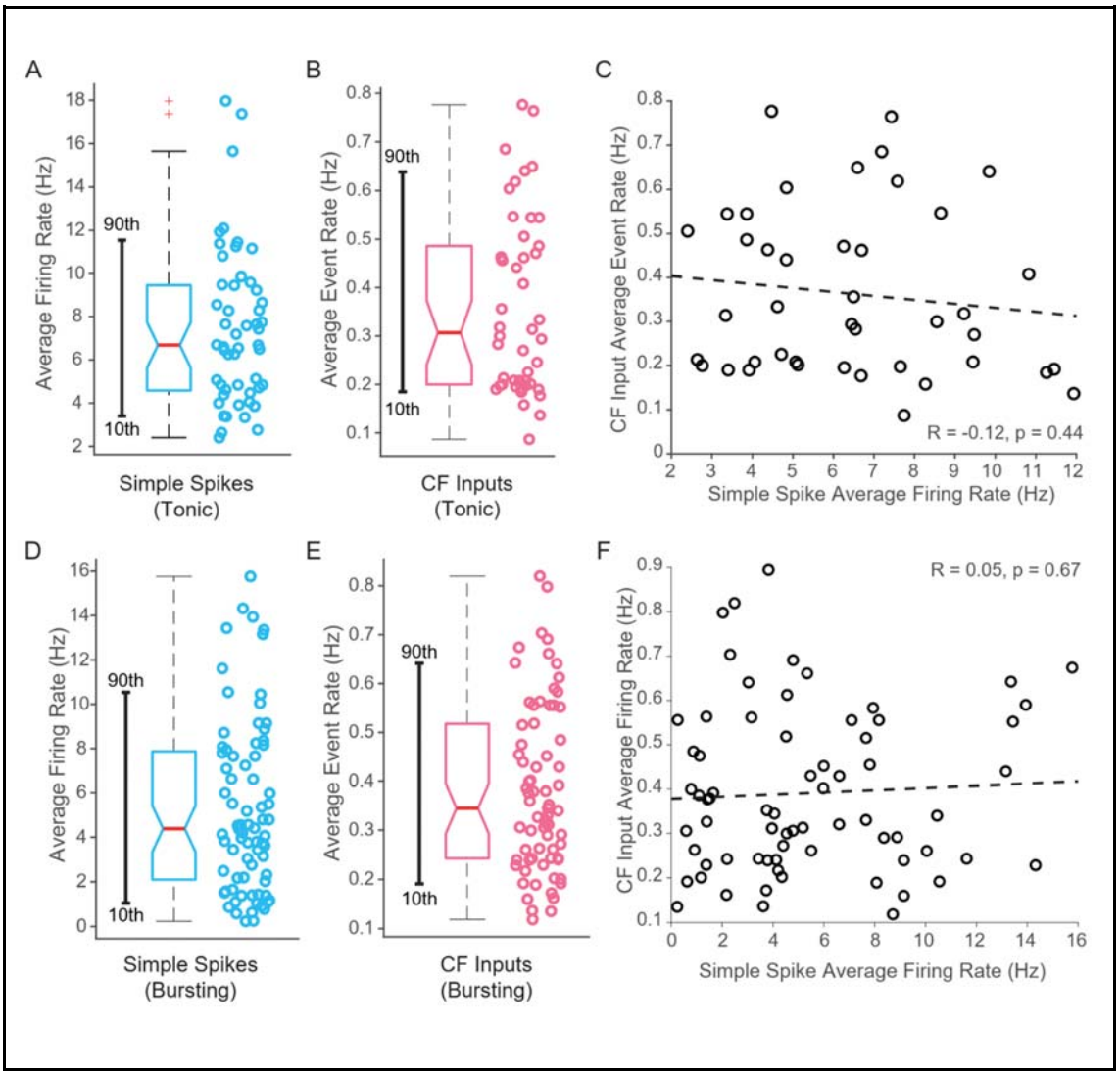


Figure 2.3 - **Average event rates of simple spikes and CF inputs in zebrafish Purkinje neurons. (A)** Boxplot showing the distribution of average simple spike firing rate for tonic cells. Individual data points are shown on the right. The 10th and 90th percentiles of the distribution are marked as a range on the left. **(B)** Boxplot showing the distribution of average CF input firing rate for all cells. Individual data points are shown on the right. The 10th and 90th percentiles of the distribution are marked as a range on the left. **(C)** Scatterplot showing the correlation between average simple spike firing rate and average CF input rate for tonic cells. The dashed line is a line of best fit for the data. **(D-F)** Average event rates for simple spikes (D) and CF inputs (E) and their correlation (F), as in A-C, but for neurons in the bursting state.

While this might work well in tonic cells, burst-like events violate the fundamental assumption underlying a Poisson process that each event be completely independent from the others. Hence, I simulated the simple spike train in bursting cells using a different strategy. The idea was to simulate spike trains such that each of their interevent distributions matched the distribution of real bursting neurons. To do so, I obtained the distribution of all simple spike interevent intervals across all bursting cells in our compiled electrophysiology dataset and drew random samples from this distribution to get the next event time. Similar to tonic cells, the CF input was simulated as a Poisson process with a random lambda value in the range of 0.2 and 0.6 Hz (10th percentile: 0.1910 and 90th percentile: 0.6480 Hz, Fig 2.3E).

To test whether these methods generated spike trains with event statistics similar to those of real cells, I compared the distributions of inter-spike intervals between real and simulated traces for both events using QQ plots. I explicitly did not use null hypothesis significance tests (NHSTs), like the Kolmogorov-Smirnov test, for these comparisons because the distributions had a large (>5000) and unequal number of data points, which drastically reduces the statistical power of these tests. The quantiles of inter-event intervals from real and simulated data for simple spikes are comparable in both tonic (Fig 2.4A, left) and bursting cells (Fig 2.4B, left), as the data points all lie very close to the line of equality. The same is observed for CF inputs in both tonic (Fig 2.4A, right) and bursting states (Fig 2.4B, right). Hence, I can conclude that the models used to generate synthetic spike trains of Purkinje neuron activity do so faithfully for both states.

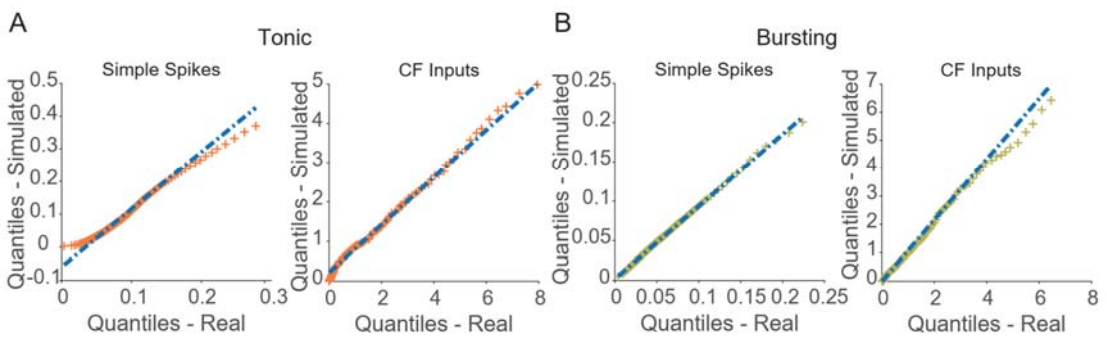


Figure 2.4 - **Comparing real and simulated spike trains. (A)** QQ plots comparing the distributions of real (x-axis) and simulated (y-axis) inter-event intervals of simple spikes (left) and CF inputs (right) in tonic cells. **(B)** QQ plots comparing inter-event intervals in real and simulated data for bursting cells.

The final test of whether or not these simple models were good approximations of real spike trains was performed with independent human reviewers. A mix of real and simulated spike trains for tonic and bursting neurons was presented to them, as shown in Fig 2.5. Independent reviewers were asked to classify these (unlabelled) spike trains as real or artificial and managed to do so only with an accuracy of ~42% on average, suggesting that this simple model captured the spike train trends in real Purkinje neurons and was good enough to confound experts.

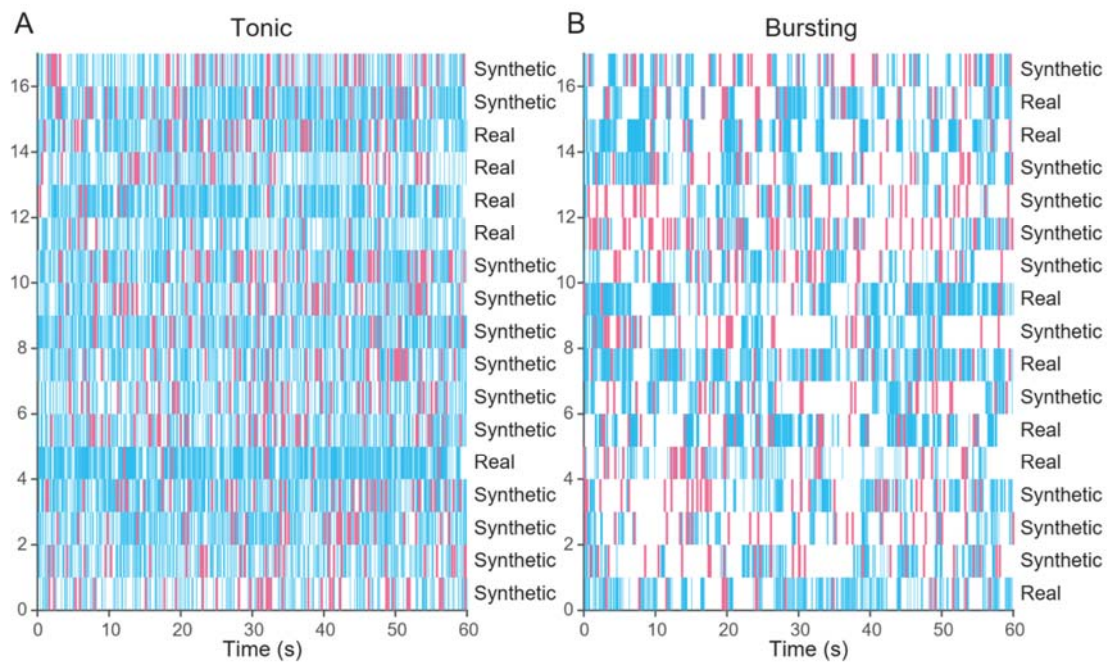


Figure 2.5 - Simulated spike train rasters for tonic and bursting Purkinje neurons. **(A)** Rasters of simple spikes (cyan) and CF inputs (magenta) for 'tonic' traces, from both real cells and artificially generated (synthetic) spike trains, as indicated by the label on the right. **(B)** Rasters of simple spikes and CF inputs for a mix of 'bursting' traces from real cells and artificially generated (synthetic) spike trains, as in (A).

2.2.3 Tonic and bursting cells have unique relationships between events

I next computed the autocorrelation functions for simple spikes (Fig 2.6A) and CF inputs (Fig 2.6B) in both bursting (Fig 2.6A,B *left*) and tonic states (Fig 2.6A,B *right*). This function (also see Methods) essentially measures the likelihood of finding another event of the same kind within a certain time interval, using time bins of reference that are suitable to the basal event rate (5ms for simple spike autocorrelation and 100ms for CF input autocorrelation).

I found that the valley of the simple spike autocorrelation function is broader for tonic cells (±65ms) than bursting cells (±45ms) (Fig 2.6A). This follows from the fact that tonic activity is periodic and that the firing rate within bursts (i.e. the period of activity in bursting neurons) is higher than the average firing rate during periods of activity in tonic cells.

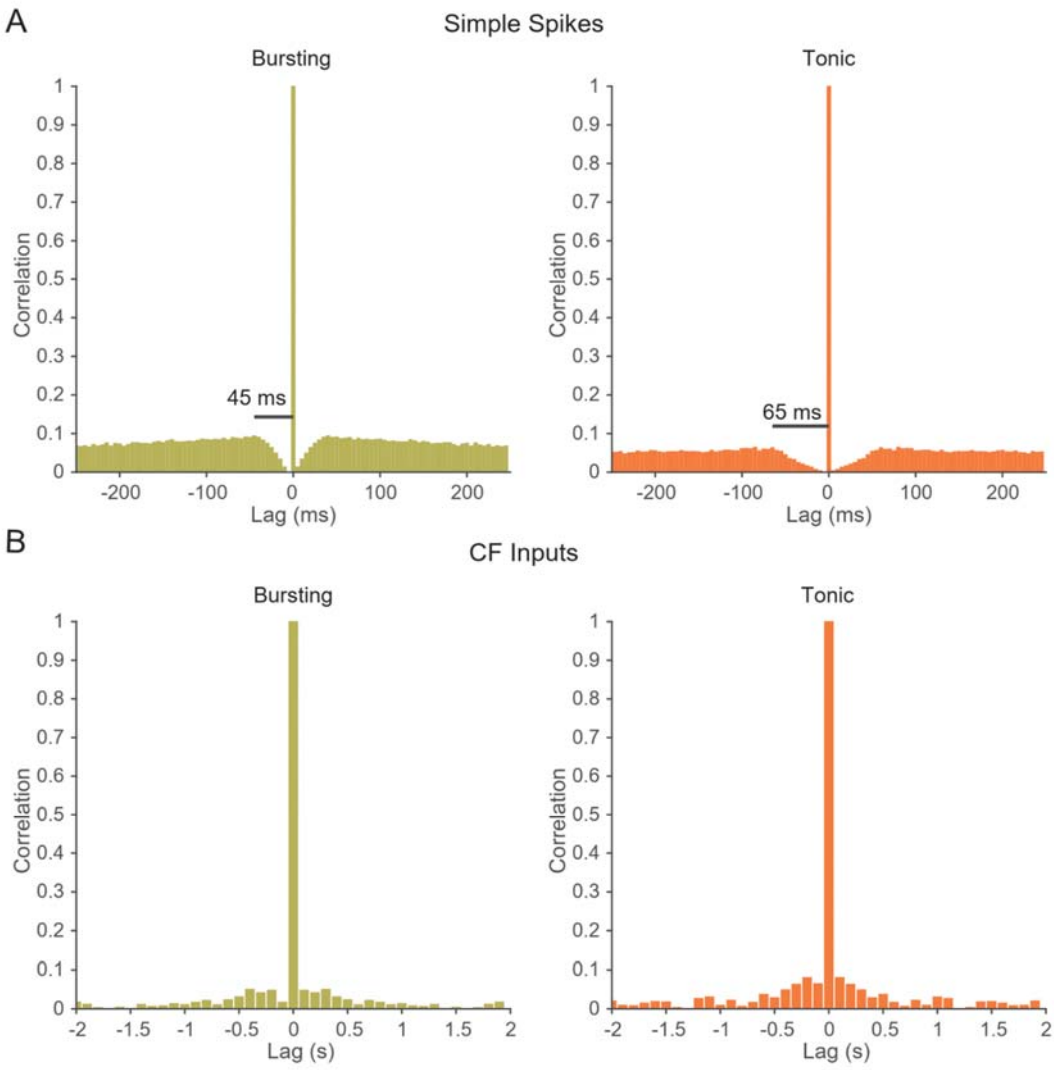


Figure 2.6 - Simple spike and CF input autocorrelation as a function of state. **(A)** Autocorrelation function of simple spikes in bursting cells (*left, green*) and tonic cells (*right, orange*). **(B)** Autocorrelation function of CF inputs in bursting cells (*left*) and tonic cells (*right*).

The autocorrelation functions for CF inputs, however, were similar between cells (Fig 2.6B). This is what one would expect, given that CF inputs are synaptic events whose frequency does not depend on the intrinsic state of the Purkinje neuron, and because the statistics of simple spikes are what distinguish tonic cells from bursting cells.

Though the statistics of the CF inputs themselves are not different between the two cellular states, I asked whether the relationship between simple spikes and CF inputs differed

between cells in the tonic and bursting modes. To address this, I calculated the cross-correlation function of simple spikes relative to CF inputs (Fig 2.7; see Materials and Methods for definition). Using this method, I found that there is an uptick in the simple spike event count ~50ms after a CF input in bursting cells (Fig 2.7A), whereas the simple spike count was lower after a CF input in tonic cells in this duration (Fig 2.7B). This suggests that CF inputs tend to lead to bursts in bursting cells, and a pause in tonic cells.

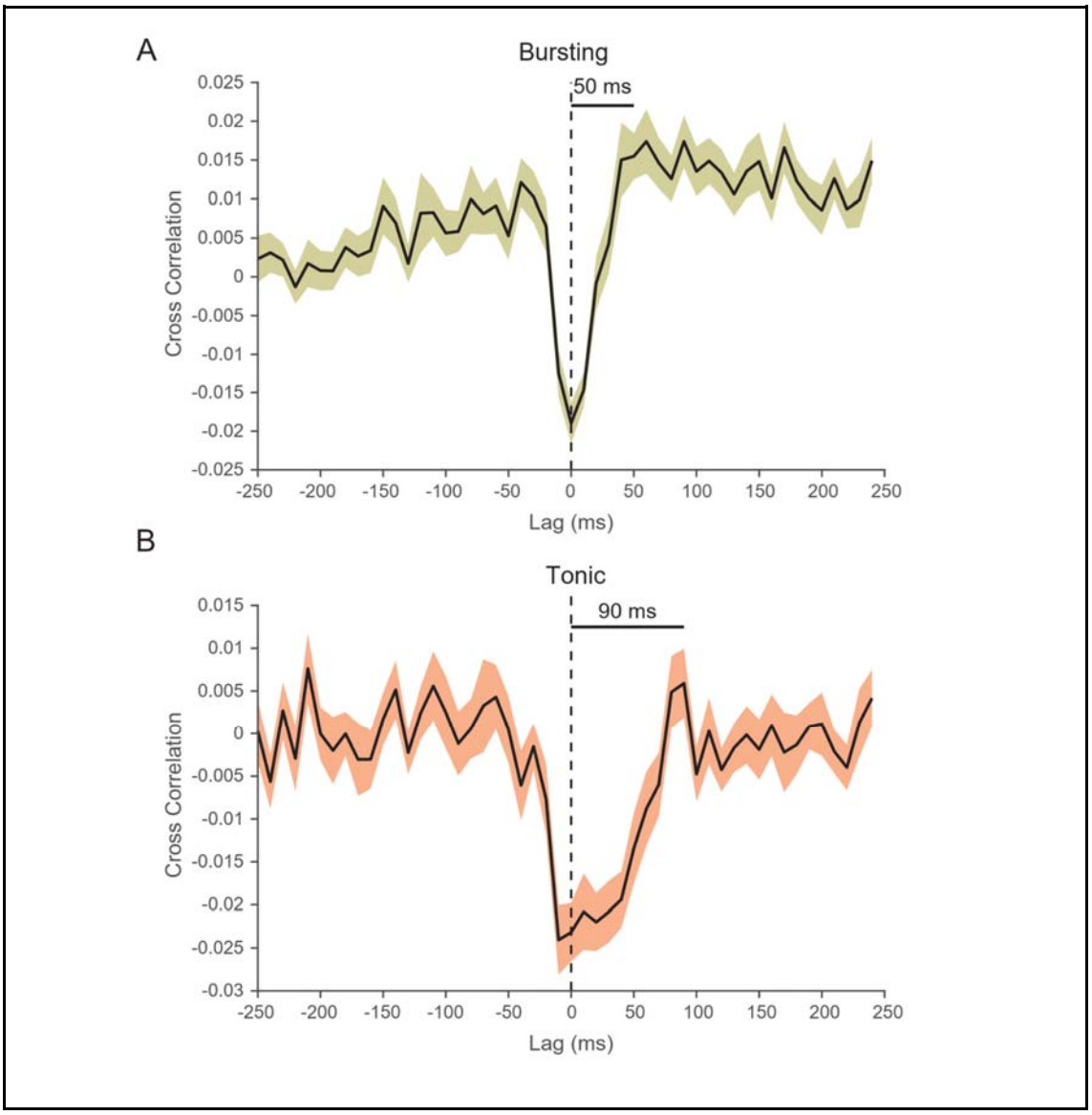


Figure 2.7 - Cross-correlation between simple spikes and CF inputs as a function of state. (A) Cross-correlation between simple spikes and CF inputs, averaged across all cells in the bursting mode. The dashed line represents the peak of the CF input, and the shaded region is the s.e.m. The time taken to return to the baseline correlation is marked with a horizontal line at the top of the plot. **(B)** Cross-correlation between simple spikes and CF inputs, like in (A), but for cells in the tonic mode.

2.2.4 Differential effect of CF inputs on simple spike firing as a function of state

In order to be certain that this effect is, indeed, a true one, I quantified the relationship between simple spikes and CF inputs in different ways. I first measured the inter-event interval between a CF input and the simple spikes immediately preceding it (pre-CF IEI, Fig 2.8A) and immediately following it (post-CF IEI, Fig 2.8A). If the null hypothesis - that the presence of a CF input has no effect on the succeeding simple spike - is true, then the scatter plot of the pre-CF IEIs and post-CF IEIs should show all the points distributed evenly around the line y = x. Similarly, the centroid of the cloud of (pre-CF IEI, post-CF IEI) pairs would lie close to the line y = x. While this seems to be true for bursting cells (Fig 2.8B, green dots and centroid on the top right), it doesn't seem to be the case for tonic cells (Fig 2.8B, orange dots and centroid on the bottom left), for which data points mostly lie above the line y = x, i.e. the post-CF IEI is higher than the pre-CF IEI, on average. This is consistent with the cross-correlation function, suggesting that CF inputs seem to lead to pauses in simple spiking activity in the tonic state.

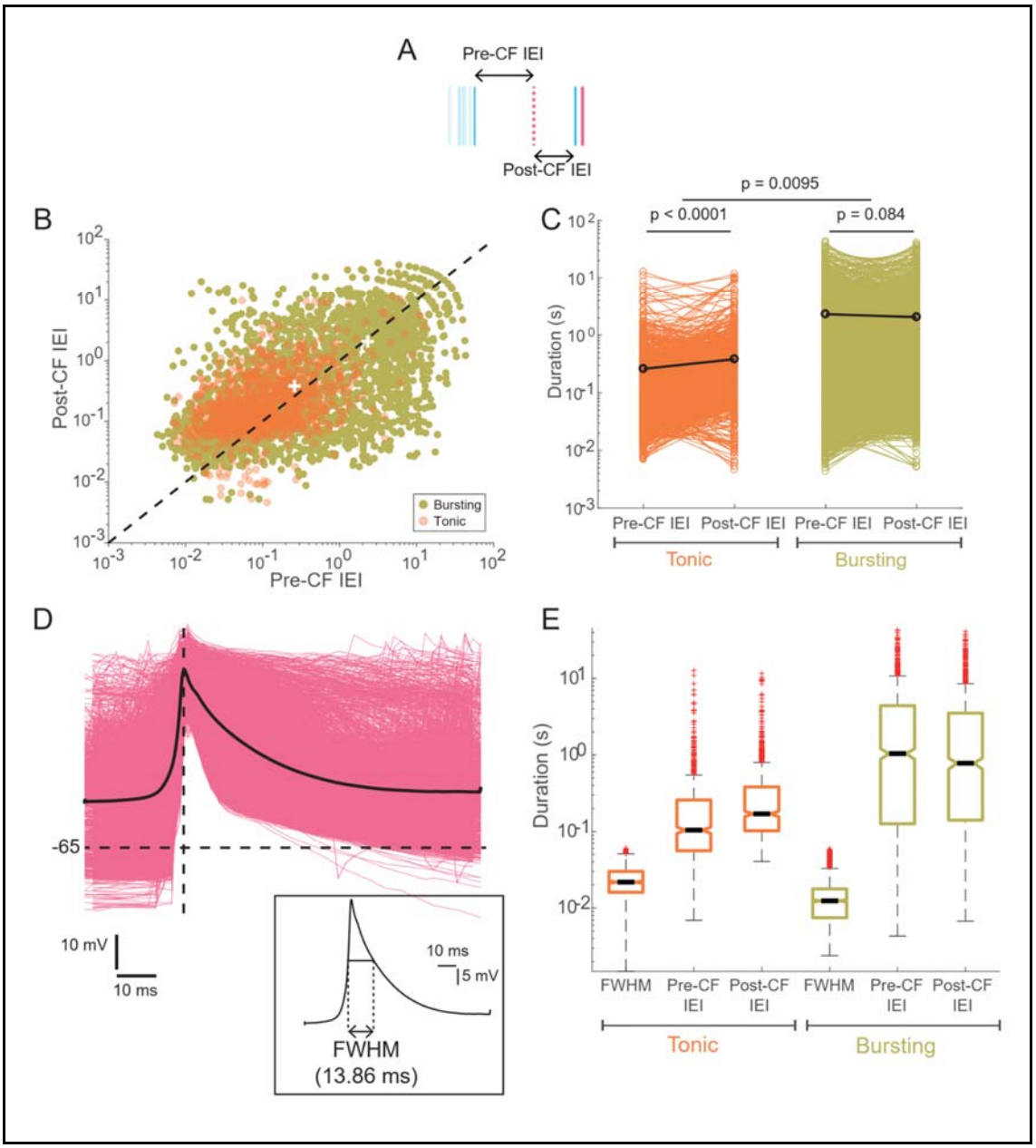

A
Pre-CF IEI
Post-CF IEI
B
Post-CF IEI
Pre-CF IEI
Bursting
Tonic
C
p = 0.0095
p < 0.0001
p = 0.084
Duration (s)
Pre-CF IEI
Post-CF IEI
Pre-CF IEI
Post-CF IEI
Tonic
Bursting
D
-65
10 mV
10 ms
10 ms
5 mV
FWHM
(13.86 ms)
E
Duration (s)
FWHM
Pre-CF IEI
Post-CF IEI
FWHM
Pre-CF IEI
Post-CF IEI
Tonic
Bursting

Figure 2.8 - Tonic cells have fewer simple spikes following the arrival of a CF input **(A)** Schematic representing the inter-event intervals relative to a CF input in a raster of electrophysiological events. Simple spikes are marked in cyan and CF inputs in magenta. The CF input in consideration is shown as a dashed line and the simple spikes immediately preceding and following it are in solid cyan lines. Those simple spikes that are farther away from the CF input are faded out. **(B)** Scatter plot of pre-CF and post-CF inter-event interval (IEI) colour-coded by state, with events from tonic cells in orange and those from bursting cells in green. The dashed line represents the line with a slope of 1, i.e. when the pre-CF and post-CF IEIs are equal. The black crosses mark the centroid of each cloud, bursting (top right cross) and tonic (bottom left cross). **(C)** Inter-event intervals before and after a CF input represented as paired data for tonic (orange) and bursting (green) cells. Each datapoint is an individual CF input. The black points and line represent the average across all events. p-values for each comparison were calculated using linear mixed-effects models (details of which are reported in Tables 2.3 and 2.4) are marked above each trace. **(D)** Superimposed individual CF inputs (magenta) showing the shape and duration of these events. The black line represents the average CF input, which is zoomed in in the inset below. The Full-width at half-maximum (FWHM) is represented on this trace. **(E)** Boxplots representing the pre- and post-CF IEIs with the FWHM (Full Width at Half-Maximum) of CF inputs alongside them for reference, classified by state.

I compared the difference between pre- and post-CF IEIs by plotting them pair-wise for every detected CF input (Fig 2.8C). The mean pre-CF IEI for every CF input across all tonic cells is 260.1ms and the mean post-CF IEI was 382.8ms. For bursting cells, the mean pre- and post-CF IEIs were 2.314s and 2.058s, respectively - approximately an order of magnitude higher than the same intervals for tonic cells. There are multiple cells in each state, and each cell has a different number (>1) of CF inputs. In other words, there is a nested structure within this dataset. Thus, I used a linear mixed-effects model to assess the statistical significance of the state-dependent difference between the pre-CF and post-CF intervals (Fig 2.8C and Table 2.3). I obtained a p-value of 0.0095, which is significant at the 5% significance level, which suggests that cells in the tonic and bursting states, indeed, have a differential response to CF inputs. I also tested whether the pre-CF and post-CF intervals were different *within* each state (Fig 2.8C and Table 2.4), also using a linear mixed-effect model, for reasons outlined above. This showed that the pre-CF and post-CF intervals were different for tonic cells ($p < 0.0001$), but not for bursting cells ($p = 0.084$).

Table 2.3 - *Statistical comparison of pre- and post-CF inter-event intervals as a function of state (related to Fig 2.8C)*

Statistical Model	**Interpretation**	**Model no**	**df**	**AIC**	**Comparison**	**p-value**	**Effect size**
IEI_Difference ~ 1	Baseline model (global mean)	1	2	20998.38			
IEI_Difference ~ 1|Cell_ID	Account for values from each cell	2	3	20999.99	1 vs 2	0.5313	
IEI_Difference ~ State|Cell_ID	Compare IEI differences between states, accounting for each cell's variability	3	4	20995.25	2 vs 3	**0.0095**	**0.3791***

*This is how much more the differences between post- and pre-CF ISIs in tonic cells are relative to what the differences between post- and pre-CF ISIs are in bursting cells.

Table 2.4 - *Statistical comparison of pre- and post-CF inter-event intervals for each state (related to Fig 2.8C)*

Statistical Model	**Interpretation**	**Model no**	**df**	**AIC**	**Comparison**	**p-value**	**Effect size**
Tonic cells							
Post_CF_SS ~ 1		1	2	3368.89			
Post_CF_SS ~ 1|Cell_ID		2	3	3174.242	1 vs 2	<.0001	

Post_CF_SS ~ Pre_CF_SS|Cell_ID		3	4	2937.678	2 vs 3	**<.0001**	**0.4901**
Bursting cells							
Post_CF_SS ~ 1		1	2	14134.42			
Post_CF_SS ~ 1|Cell_ID		2	3	13291.45		**<.0001**	
Post_CF_SS ~ Pre_CF_SS|Cell_ID		3	4	13290.46		**0.084**	**0.0345**

To rule out the possibility that the higher post-CF IEIs in tonic cells were observed because the CFs themselves are broad events and that the simple spike after a CF input was somehow drowned out or hidden in it, I calculated the full width at half maximum (FWHM) as a measure of the breadth of CF inputs in each state (Fig 2.8D). The median FWHM of CF inputs was 15.8ms, with an interquartile range of 12.1ms. This is at least one order of magnitude smaller than both the pre-CF IEI and the post-CF IEI (Fig 2.8E), in both tonic (orange boxplots) and bursting (green boxplots) cells. Thus, the ratio of the post- to pre-CF IEIs is very likely to be a true effect and not because the simple spike following a CF input could not be detected.

There are several limitations to looking only at the inter-event intervals, most importantly the fact that there is a lot of variability in the spiking activity of a Purkinje neuron. Thus, I decided to calculate an index that normalised for this variability - the asymmetry index (Fig 2.9A, left). This is defined as the ratio of the difference in intervals between consecutive events to the sum of those intervals (Hong et al., 2016). The asymmetry index (I_{Asym}) takes values between -1 and 1, with +1 indicating that there is a pause in activity, and -1 indicating that there is an increase in activity after the event in question. I measured 3 different AIs (Fig 2.9A, right): the SS-SS I_{Asym} (between simple spikes alone), CF-CF I_{Asym} (between CF inputs alone) and the CF-SS I_{Asym} (the asymmetry between a CF input and the simple spikes before and after it).

The distribution of the SS-SS I_{Asym} seems bell-shaped, centred around 0 (Fig 2.9B), suggesting that most spiking activity is symmetric/regular in both states. However, the distribution for bursting cells (green) have larger tails, which correspond to the asymmetries introduced by the start and end of bursts. The distributions of the CF-CF I_{Asym}, however, are uniform (Fig 2.9C) because they are sparse, aperiodic events. The important thing to note, though, is that the distributions for both tonic (orange) and bursting (green) cells are very similar, once again just as one would expect of synaptic events.

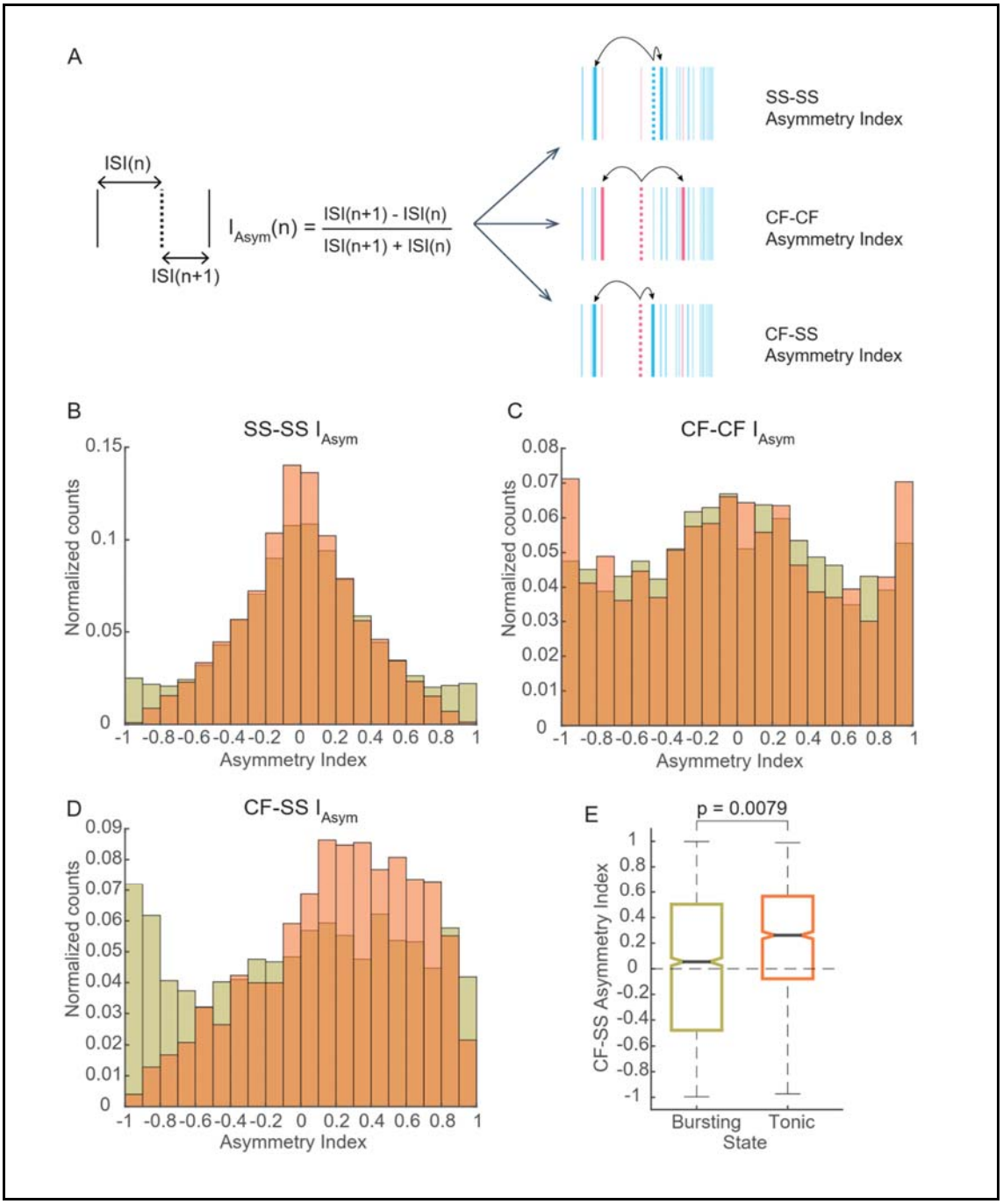


Figure 2.9 - **Asymmetry in inter-event intervals uncover unique state-dependent relationships between spiking events (A)** Schematic representing the asymmetry index (AI) calculation. Three kinds of asymmetry indices are calculated - SS-SS being the asymmetry index between a simple-spike and its two surrounding simple-spikes (top); CF-CF being the asymmetry index between a CF input and its two nearest CF inputs (middle); CF-SS being the asymmetry index between a CF input and its two nearest simple spikes (bottom). All simple spikes are represented in cyan and all CF inputs in magenta. The dashed line in the rasters represent the event being used as reference to calculate the asymmetry index. Events that are irrelevant to the calculation of the index are faded out. **(B)** Histogram of the SS-SS I_{Asym} for tonic (orange) and bursting (green) cells. **(C)** Histogram of the CF-CF I_{Asym} for tonic (orange) and bursting (green) cells. **(D)** Histogram of the CF-SS I_{Asym} for tonic (orange) and bursting (green) cells. **(E)** Boxplot of the CF-SS I_{Asym} in each state ($p = 0.0079$, linear mixed-effects model).

The most interesting of the asymmetry indices, however, is the CF-SS I_{Asym} (Fig 2.9D). This is the asymmetry of simple spikes around a CF input (i.e. the asymmetry between the interval between a CF input and the SS following it and the interval between the same CF input and SS preceding it - Fig 2.9A, right). It is immediately clear that the distributions for tonic and bursting cells are not identical. In fact, the distribution for tonic cells has mostly positive values, consistent with the induction of a pause in simple spike activity after a CF input in tonic cells. The distribution for bursting cells is more uniform, although there is an overrepresentation of more negative values, which suggest that many CF inputs are followed by a pause termination, which, in these cells, corresponds to a burst in activity. Both these distributions are statistically significantly different ($p=0.0079$, linear mixed-effects model - Table 2.5) with an effect size of 0.191.

Table 2.5 - *Statistical comparison of CF-SS asymmetry indices as a function of cellular state (related to Fig 2.9E)*

Statistical Model	Interpretation	Model no	df	AIC	Comparison	p-value	Effect size
Asymmetry Index ~ 1	Baseline model (global mean)	1	2	4407.051			

Asymmetry Index ~ 1|Cell_ID	Account for values from each cell	2	3	4273.297	1 vs 2	1.22E-30	
Asymmetry Index ~ State|Cell_ID	Compare IEI differences between states, accounting for each cell's variability	3	4	4268.251	2 vs 3	**0.00794**	**0.191**

Put together, these data suggest that although all Purkinje neurons receive similar synaptic input, their intrinsic state determines their response to this input (CF input, in particular), with tonic cells pausing their activity in response to the arrival of a CF input, and bursting cells responding neutrally. This seemed inconsistent with previous data that have shown that the arrival of CF inputs triggers a burst in Purkinje neurons in the bursting mode (Sengupta and Thirumalai, 2015). It is possible that we don't see a strong modulation in bursting cells, on average, because bursting neurons also receive numerous CF inputs in the intra-burst periods.

To check whether these intra-burst CFs have different effects from extra-burst CFs, I sorted CF inputs as those within a burst (Fig 2.10A, asterisk) and those outside (Fig 2.10A, hash). Across all bursting cells, there were 2461 CF inputs, of which 1524 (61.93%) were outside bursts and the remaining 937 (38.07%) were within bursts. I then measured the SS-CF asymmetry indices (Fig 2.10B and C) as well as the pre-CF and post-CF inter-event intervals (Fig 2.10D) for each of these groups. The distribution of asymmetry indices for those CF inputs which were outside bursts resembled that of tonic cells (Fig 2.10B, left), whereas the distribution of asymmetry indices for those CF inputs within bursts (Fig 2.10B, right) had an over-representation of negative asymmetry indices (i.e. they tended to terminate periods of quiescence). The median indices for CFs within and outside bursts were 0.1874 and -0.1148, respectively. This difference was statistically significant, as tested using a linear mixed-effects model ($p < 0.0001$ - Table 2.6).

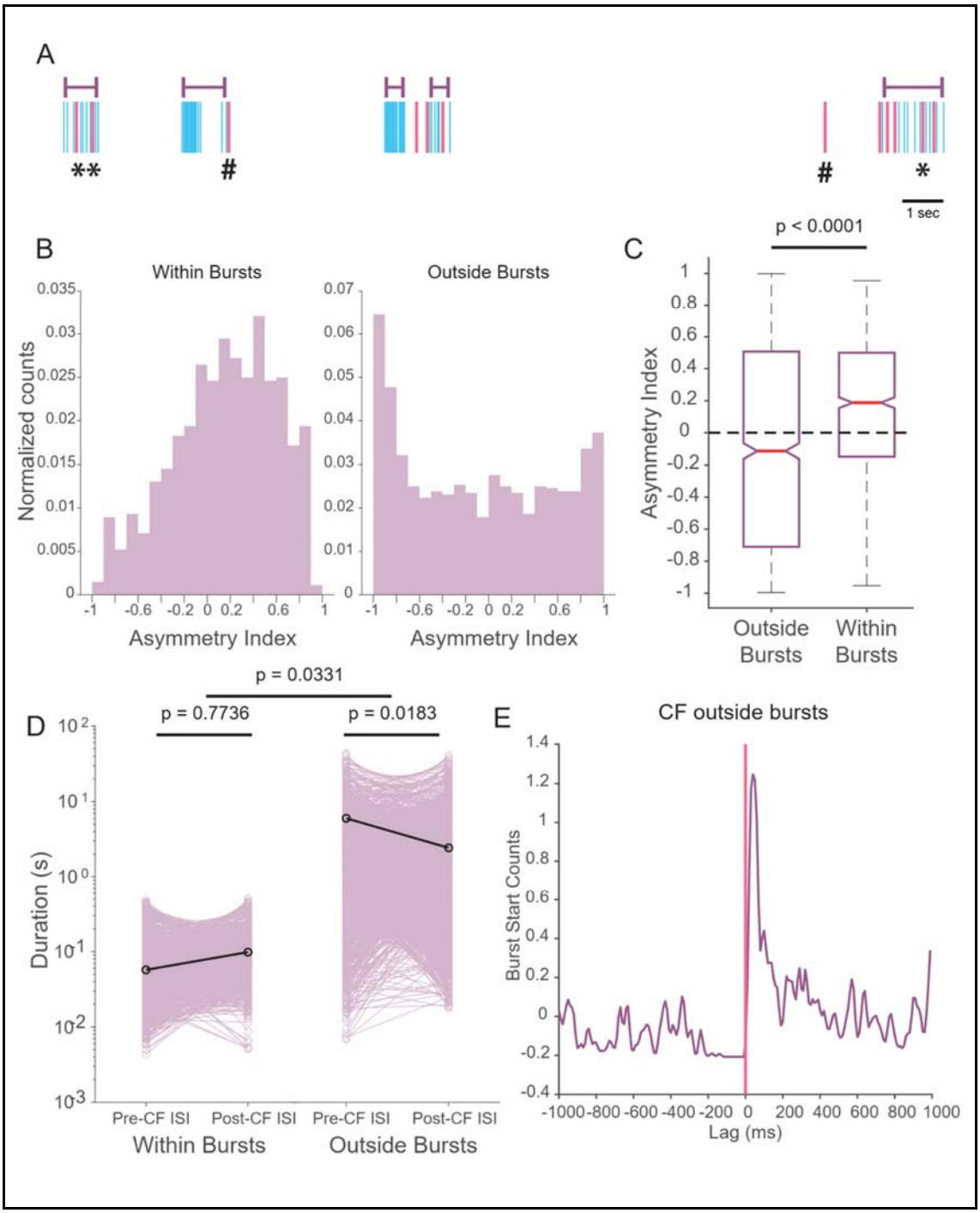

A
**
#
#
*
1 sec
B
Within Bursts
Outside Bursts
Normalized counts
Asymmetry Index
Asymmetry Index
C
p < 0.0001
Asymmetry Index
Outside Bursts
Within Bursts
D
p = 0.0331
p = 0.7736
p = 0.0183
Duration (s)
Pre-CF ISI
Post-CF ISI
Pre-CF ISI
Post-CF ISI
Within Bursts
Outside Bursts
E
CF outside bursts
Burst Start Counts
Lag (ms)

Figure 2.10 - **CF inputs within and outside bursts have differential effects. (A)** An example event raster from a Purkinje neuron in the bursting state, using the same colour scheme as in the rest of the thesis - i.e. simple spikes are marked in blue, and CF inputs in magenta. Bursts of simple spikes that are detected are marked above using purple bars. A few CF inputs that are within bursts are marked with asterisks, and a few CF inputs that are outside bursts are marked with hashes. **(B)** The distribution of simple spike asymmetry indices with respect to CF inputs (also see Fig 2.9, bottom), sorted with respect to whether the CF input was within a burst (left) or outside a burst (right). **(C)** The absolute durations of CF-SS interevent intervals before (pre-) and after (post-) a CF input, either within bursts (left) or outside bursts (right). **(D)** Perievent time histogram of burst starts (purple) relative to a CF input that is outside a burst (magenta vertical line).

The pre- and post-CF inter-event intervals were also significantly different between the within-burst and outside-burst CF inputs (Fig 2.10D and Table 2.7). When comparing them within each group, however, I found that the difference between these intervals is not significant for CF inputs within bursts ($p = 0.7736$, with an effect size of -0.00813). For the CF inputs outside bursts, however, the difference between these intervals is significant at the 5% level ($p = 0.0183$), although the effect size is quite small (-0.0611) (Fig 2.10D and Table 2.8).

I confirmed, however, that CF inputs tend to be followed immediately by bursts by plotting the frequency of burst start times around a CF input that is outside a burst (Fig 2.10E). In sum, I find that climbing fibre input can trigger an increase in activity in bursting neurons, contrary to its effect in neurons in the tonic state. When within a burst, however, they tend to lead to a pause, although the effect is much weaker than that in a tonic neuron.

Table 2.6 - *Statistical comparison of CF-SS asymmetry indices as a function of whether the CF was within or outside a burst (related to Fig 2.10C)*

Statistical Model	Interpretation	Model no	df	AIC	Comparison	p-value	Effect size
Asymmetry Index ~ 1	Baseline model (global mean)	1	2	4352.748			

Asymmetry Index ~ 1|Cell_ID	Account for values from each cell	2	3	4271.868	1 vs 2	<.0001	
Asymmetry Index ~ Within_Or_Outside_Burst|Cell_ID	Compare IEI differences between states, accounting for each cell's variability	3	4	4178.053	2 vs 3	**<.0001**	**0.2619 ***

*This indicates how much more positive the CF-SS asymmetry index is within bursts as compared to outside bursts.

Table 2.7 - *Statistical comparison of the difference in pre- and post-CF inter-event intervals as a function of state (related to Fig 2.10D)*

Statistical Model	**Interpretation**	**Model no**	**df**	**AIC**	**Comparison**	**p-value**	**Effect size**
IEI_Difference ~ 1	Baseline model (global mean)	1	2	14889.01			
IEI_Difference ~ 1|Cell_ID	Account for values from each cell	2	3	14891.01	1 vs 2	0.9986	
IEI_Difference ~ Within_Or_Outside_Burst|Cell_ID	Compare IEI differences between burst locations, accounting for each cell's variability	3	4	14888.47	2 vs 3	**0.0331**	**0.4402 ***

*This indicates how much larger the difference in inter-event intervals is within bursts relative to the same difference for CF events outside bursts.

Table 2.8 - *Statistical comparison of pre- and post-CF inter-event intervals for each state (related to Fig 2.10D)*

Statistical Model	Interpretation	Model no	df	AIC	Comparison	p-value	Effect size
CFs within bursts							
Post_CF_SS ~ 1	Baseline model (global mean)	1	2	-2025.374			
Post_CF_SS ~ 1|Cell_ID	Account for values from each cell	2	3	-2155.968	1 vs 2	<.0001	
Post_CF_SS ~ Pre_CF_SS|Cell_ID	Test if post-CF interval can be predicted based on pre-CF interval	3	4	-2154.051	2 vs 3	**0.7736**	**-0.00813 #**
CFs outside bursts							
Post_CF_SS ~ 1	Baseline model (global mean)	1	2	9274.197			
Post_CF_SS ~ 1|Cell_ID	Account for values from each cell	2	3	8845.739	1 vs 2	**<.0001**	
Post_CF_SS ~ Pre_CF_SS|Cell_ID	Test if post-CF interval can be predicted based on pre-CF interval	3	4	8842.175	2 vs 3	**0.0183**	**-0.06113 #**

These effect sizes indicate how different the intervals between SS and CF inputs are depending on whether the SS occurred before or after a CF input.

2.3 Discussion

2.3.1 A compiled dataset of Purkinje neuron recordings sorted by state

Previous work on Purkinje neuron bistability predominantly used whole-cell recordings to identify state (Sengupta & Thirumalai, 2015). This is facilitated by the fact that cells in the tonic state are at a more depolarized membrane potential (~-40mV), whereas cells in the bursting state are at a hyperpolarized membrane potential (~-65mV). Here, I used an alternative method of state identification based on simple spike firing, which is therefore independent of recording mode. I combined visual inspection of the raw electrophysiological trace and the inter-event interval distribution with said distribution's coefficient of variation. Together, they yielded a high degree of classification accuracy. It is important to note, however, that it is still difficult to assign a neuron's state completely unambiguously, and that there may be some cells in our compiled dataset whose state has been misidentified. The state of a cell is a dynamic property, and misclassifications could also occur if a cell is in the process of switching its state. However, the exercise of compiling this dataset proved useful in allowing me to investigate properties that are true properties of each state (since the occasional misclassification shouldn't dilute strong effects). It also allowed me to generate new spike trains for each state such that the statistics of real neurons are reasonably approximated.

2.3.2 State-dependent differential response to synaptic input

The fact that Purkinje neurons show bistability has been fairly well-established, yet its function isn't understood. Having recorded from zebrafish Purkinje neurons (PNs) over years, I had access to high temporal resolution data of their activity in both states. This data was compiled and analysed to extract information about the relationship between the two electrical events in these neurons, as a function of their state. These analyses included computing the auto- and cross-correlation functions and measuring various intervals between the events in a cell.

I show that Purkinje neurons differentially respond to the arrival of a CF input depending on their cellular state. In tonic cells, the arrival of a CF input led to a pause in simple spike activity. Unlike their tonic counterparts, bursting cells were neutral to the arrival of a CF input on average. Classifying CF inputs as those outside or within a burst revealed that the two types had opposing effects. Short post-CF pauses in simple spiking activity were observed within bursts, whereas CFs outside bursts tended to trigger a burst of simple spiking. Pauses in simple spiking after climbing fiber activity have been widely reported in mammals (Bloedel

& Roberts, 1971; Latham & Paul, 1971) as well as in mormyrid fish (Alviña & Sawtell, 2014). No state-dependent effect has been reported thus far, however.

Thus, I expect that the proportion of cells in the tonic or bursting state could strongly influence the output of the cerebellar circuit as a result. How this occurs, though, remains an open question that would be fascinating to address in the future.

This analysis is easy to perform for CF inputs, because they are easy to detect as a result of their large amplitude, so whether parallel fiber input (which is the other source of excitatory synaptic input to these neurons) also has similar effects is not clear. Since CF inputs are especially strong and lead to large calcium influxes, it is likely that this phenomenon is limited to CF inputs.

2.4 Materials and Methods

2.4.1 Animal care

Indian wild type (Ind WT) zebrafish (*Danio rerio*) were used for all experiments, unless otherwise stated, and the institutional animal ethics committee (IAEC) and institutional biosafety committee (IBSC) approved all experiments. Embryos were obtained by setting up an incross between adults raised and housed in a ZebTec multi linking system (Tecniplast, Italy) with a pH setting of 7.8 and conductivity of 1200µS. Larvae were kept in a MIR-154 incubator (Sanyo, Japan) with 14:10h light-dark cycle maintained at 28ºC in E3 medium (composition in mM: 5 NaCl, 0.17 KCl, 0.33 $CaCl_2$, and 0.33 $MgSO_4$, pH 7.8) in standard 90mm Petri dishes (Tarsons, Kolkata, India). The medium was routinely replaced, and starting at 5 dpf, larvae were fed Zeigler Larval diet AP100 (<100 microns) (Pentair AES, FL, United States). Larvae have not yet undergone sex specification by these stages, so experiments performed were agnostic to the sex of the animal.

2.4.2 Patch-clamp electrophysiology from Purkinje neurons

Dissection

Larvae were anaesthetised in 0.01% MS-222 (Sigma-Aldrich; Missouri, USA) and transferred to a custom-made acrylic recording chamber. Fine tungsten wire (California Fine Wire, CA, USA) was used to pin the larva onto a piece of Sylgard (Dow Corning, Midland, MI, United States) glued to the recording chamber. Two pins were put through the notochord at the level of the swim bladder and the tail, respectively. A third pin was then used to position the larval head dorsal-up. The MS222 was replaced by external solution

(composition in mM: 134 NaCl, 2.9 KCl, 1.2 $MgCl_2$, 10 HEPES, 10 glucose, 2.1 $CaCl_2$, 0.01 D-tubocurarine; pH 7.8; 290 mOsm) and skin on the dorsal surface of the larva was carefully removed using fine no. 5 forceps (Fine Science Tools, Foster City, USA) to expose the brain. This preparation was then taken to an electrophysiology rig with a Multiclamp 700B amplifier, Digidata 1440A digitiser and the pCLAMP software suite (Molecular Devices, Sunnyvale, CA, United States), and visualised using a microscope with water-immersion objectives (20x, 0.5NA and 60x, 1.0NA).

Juxtacellular patch-clamp

A Flaming-Brown P-97 pipette puller (Sutter Instruments, Novato, CA, United States) was used to pull patch pipettes made of standard-walled borosilicate capillaries with a filament (1.5 mm OD; 0.86 mm ID; Warner Instruments, Hamden, CT, United States), pulled to a final tip diameter of 1-1.5 µm. For loose patch recordings, pipettes were backfilled with external solution that was filtered through a 0.2µm filter (Millipore, Merck, Germany). These pipettes had a resistance between 8-12MΩ. Backfilled pipettes were loaded onto an ISO-S-1.5G electrode holder (G23 Instruments, London, UK) connected to a CV-7B headstage (Axon Instruments, Molecular Devices, Sunnyvale, CA, United States). The pipette was mounted onto a PatchStar micromanipulator (Scientifica Instruments, UK) which was used to focus near the sample. A target cell was then approached with mild positive pressure in the pipette. Once the pipette position relative to the cell was satisfactory, the positive pressure was released and mild suction was applied. Seals were formed which had resistances in the range of 100-250MΩ. If a gigaohm seal formed, positive pressure was applied to release the seal, the pipette repositioned and suction applied to obtain seals in the 100s of MΩs range. The amplifier gain (not feedback gain) was adjusted using the Multiclamp controller software to between 20-1000 so as to optimally use the dynamic range of the digitiser. Data was acquired at a sampling rate of 50kHz, low-pass filtered at 2kHz.

Whole-cell patch clamp

A similar procedure to the juxtacellular recordings was used to get whole-cell recordings. Patch pipettes were backfilled with filtered K-gluconate internal solution (composition in mM: 115 K gluconate, 15 KCl, 2 $MgCl_2$, 10 HEPES, 10 EGTA, 4 Mg-ATP; pH 7.2; 290 mOsm), however. Sulforhodamine (Sigma-Aldrich, St.Louis, MO, United States) was added to the patch internal solution at a final concentration of 5µg/mL to facilitate visualisation of cellular morphology post-recording.

The sample was approached with the pipette as done for obtaining juxtacellular recordings, with positive pressure, while monitoring the pipette resistance with square pulses of 10mV in

voltage clamp mode. Once the pipette was near the target cell, positive pressure was released and a mild negative pressure was applied to facilitate the formation of a gigaohm seal. Once the gigaohm seal formed, the pipette capacitance was compensated and holding potential set to -65mV before pulses of suction were applied to break into the cell. The break-in was successful if large capacitive transients were observed in the membrane test pulse response.

Recordings were obtained either in voltage clamp mode or current clamp mode, as needed. Depending on the experimenter and/or the experiment in question, data were acquired at one of two electrophysiology rigs. Both rigs had identical amplifiers and digitisers - Multiclamp 700B amplifier and Digidata 1440A digitiser, respectively (Molecular Devices, Sunnyvale, CA, United States). However, they had different microscopes and micromanipulators, either an Ni-E microscope (Nikon Corporation) with a PatchStar micromanipulator (Scientifica Instruments, UK), or an Olympus BX61WI microscope with an MP-285 micromanipulator (Sutter Instruments, Novato, CA, United States). Recordings were acquired at a 20-50kHz sampling rate, depending on the experimenter and the experiment in question, at an amplifier gain of 1 and after low-pass filtering at source at a cutoff frequency of 2kHz.

2.4.3 Compiling state-sorted electrophysiology traces

Intracellular and extracellular recordings from larval zebrafish Purkinje neurons acquired under baseline conditions (i.e. without the addition of small molecule drugs or injection of currents) during the period 2011-2020 by various members of the lab (Mohini Sengupta, Meha Jadhav, Shivangi Verma and myself (Aalok Varma)) were compiled. Good quality recordings were those in which the resting membrane potential of the cell was not below -80mV or not above -20mV, were uninterrupted gap-free recordings with no current injections and had two clearly distinguishable event types. Only those recordings which met the aforementioned criteria were retained. Besides poor quality recordings, those recordings with an ambiguous or indeterminate state were also excluded. The retained recordings were renumbered randomly, to remove grouping of data collected by the same individual. Recordings were then classified as tonic or bursting, first by visual inspection. These classifications were then further verified by inspecting the inter-event intervals of simple spikes and their coefficients of variation. If there was a switch in state within the recording, the recordings were labelled as “switch”, with the appropriate state switch timing and kind (tonic-to-bursting, or bursting-to-tonic) marked. A summary of the recording information is in Table 2.1.

2.4.4 Event detection

All electrophysiological analyses were done using custom scripts written in MATLAB, using functions defined in a publicly available repository (https://github.com/wagenadl/mbl-nsb-toolbox) to read electrophysiology data. For extracellular recordings, spikes were identified as events that showed large fluctuations relative to a 25ms rolling window baseline and the peak timing and amplitude were extracted. Events were sorted based on amplitude, with small amplitude events being classified as simple spikes and large amplitude events as climbing fiber (CF) inputs, as per Sengupta and Thirumalai (eLife, 2015). For intracellular recordings, peak detection algorithms inbuilt in MATLAB were used to identify spikes. Once again, events were sorted as simple spikes or CF inputs based on their amplitude relative to baseline.

2.4.5 Calculating autocorrelation and cross-correlation functions

The spike-train autocorrelation function was calculated as per (Dayan & Abbott, 2005). I start by first dividing time into bins of size Δt. For simple spikes, I used a value of 5ms, whereas for CF inputs, I used a bin size of 100ms. For integer values of m, the number of spike pairs which are separated by a time interval between $(m - ½)\Delta t$ and $(m + ½)\Delta t$ is calculated. In other words, each bin has the count of how many spike pairs are separated by a certain time lag. Note that this is done for all spike pairs, which includes the pairing of an event with itself. If all of the n^2 spike-pairs were uniformly distributed, then there would be $n^2\Delta t/T$ intervals in each bin, where n is the total number of spikes and T is the total duration of the recording. This value is subtracted from the autocorrelation histogram of all bins. Finally, the resulting values are divided by T, so that in the limit of $\Delta t \rightarrow 0$, the calculated values represent average firing rates.

Similarly, a cross-correlation function between spike trains from two different sources (either two different neurons or two different kinds of events in the same neurons) can be calculated as the histogram of spike time differences between all pairs of events being compared. I used a bin size of 10ms to calculate the cross-correlation function between simple spikes and CF inputs.

2.4.6 Generating artificial Purkinje neuron spike trains

Synthetic spike trains of the tonic class were generated using a Poisson statistical model. The only parameter needed for this is an average event rate, λ, which I took to be a random number between 4-10 Hz for simple spikes (also see Fig 2.9). Spike trains for bursting neurons do not follow Poisson statistics. Hence, I used an inverse sampling method to

generate spike trains of the bursting class. For this, I used the compiled electrophysiology data to obtain the real cumulative distribution of interevent intervals for simple spikes across all bursting cells. A random number, y, is generated from the uniform distribution and it is used to find a value x from the reference cumulative interspike interval (ISI) distribution such that P(ISI <= x) = y. Thus, by generating events sequentially with a given next interevent interval, I obtained artificial simple spike trains for the bursting class. A refractory period was taken into account such that if an event followed the preceding one in less than 3ms, it was removed from the spike train.

CF input event trains were generated using a Poisson statistics model with an average event rate between 0.2-1 Hz. The distributions of interevent intervals in the synthetic datasets were compared against the real distributions using QQ plots (Fig 2.9). Independent reviewers were asked to classify unlabelled spike trains as real or artificial and managed to do so with an accuracy of ~42% on average.

2.4.7 QQ Plots

Quantile-quantile (QQ) plots were used to compare two distributions, X and Y, for similarity. First, outliers, defined as those points which lie more than 1.5 interquartile ranges below or above the first (Q1) and third (Q3) quartiles, respectively, were removed from both distributions. Each of the two distributions was then sorted and their percentiles plotted against one another. This was performed using the inbuilt “qqplot” function in MATLAB. A line of reference was drawn between ($Q1_X$, $Q1_Y$) and ($Q3_X$, $Q3_Y$). Strong deviations from this line mean that X and Y weren’t sampled from the same distribution.

Chapter 3 - Relationships in electrophysiological activity between Purkinje neurons

3.1 Introduction

The compiled electrophysiological dataset described in the previous chapter allowed me to infer the properties of individual Purkinje neurons as a function of its state. Spike-train level analysis revealed that the firing mode of a Purkinje neuron impacts its response to synaptic input. However, these cells don't exist in isolation and are part of a network of neurons. The developing larval zebrafish has between 200-400 Purkinje neurons (Hamling et al., 2015), which likely increase even more until adulthood. Thus, features encoded by these neurons are possibly distributed across the population, with the state of the cell adding a layer of nuance to the code. If each of these neurons can independently exist in either one of the two bistable states, the space of combinations of states in this population is very large, and we are yet to understand its meaning, especially in how information gets transferred and processed further downstream from these neurons.

Thus, I next sought to ask how similar spiking and synaptic activity are between neurons as a function of their state. To address this question, I first decided to look at pairs of neurons and understand the relationships between their activity. To that end, I acquired paired electrophysiological recordings from neighbouring Purkinje neurons. I then performed spike-train-level analyses on these data to work out the associations between various parts of the cerebellar circuit, which is described in the following chapter of this thesis.

3.2 Results

3.2.1 Simple spike firing rates are more correlated in pairs of bursting neurons than tonic neurons

We know that in the absence of synaptic input, Purkinje neurons fire tonically (Häusser & Clark, 1997; Sengupta & Thirumalai, 2015), which implies that the tonic state could be the default state a neuron exists in, and that the bursting mode arises from a cell's integration into the network. Thus, I hypothesised that bursting cells might be more similar in activity to other bursting cells, as a result of being tuned into network activity.

To test this hypothesis, I performed simultaneous loose patch recordings on pairs of Purkinje neurons from the same lobe of the cerebellum (Fig 3.1A). I acquired 1 minute-long

recordings from a total of n=25 pairs of cells from N=13 fish (Table 3.1), of which 17 pairs were bursting-bursting pairs (Fig 3.1B), 3 were tonic-tonic pairs (Fig 3.1C) and 5 were tonic-bursting pairs (Fig 3.1D). The raster plots of all detected events from all pairs in this dataset are shown in Fig 3.2. Since I recorded from cells in their natural state in a minimally invasive way, and did not manipulate said state using current injections, this dataset is unevenly sampled and has an overrepresentation of bursting-bursting pairs. Nevertheless, I thought performing some preliminary analysis on these data would be useful to guide future work.

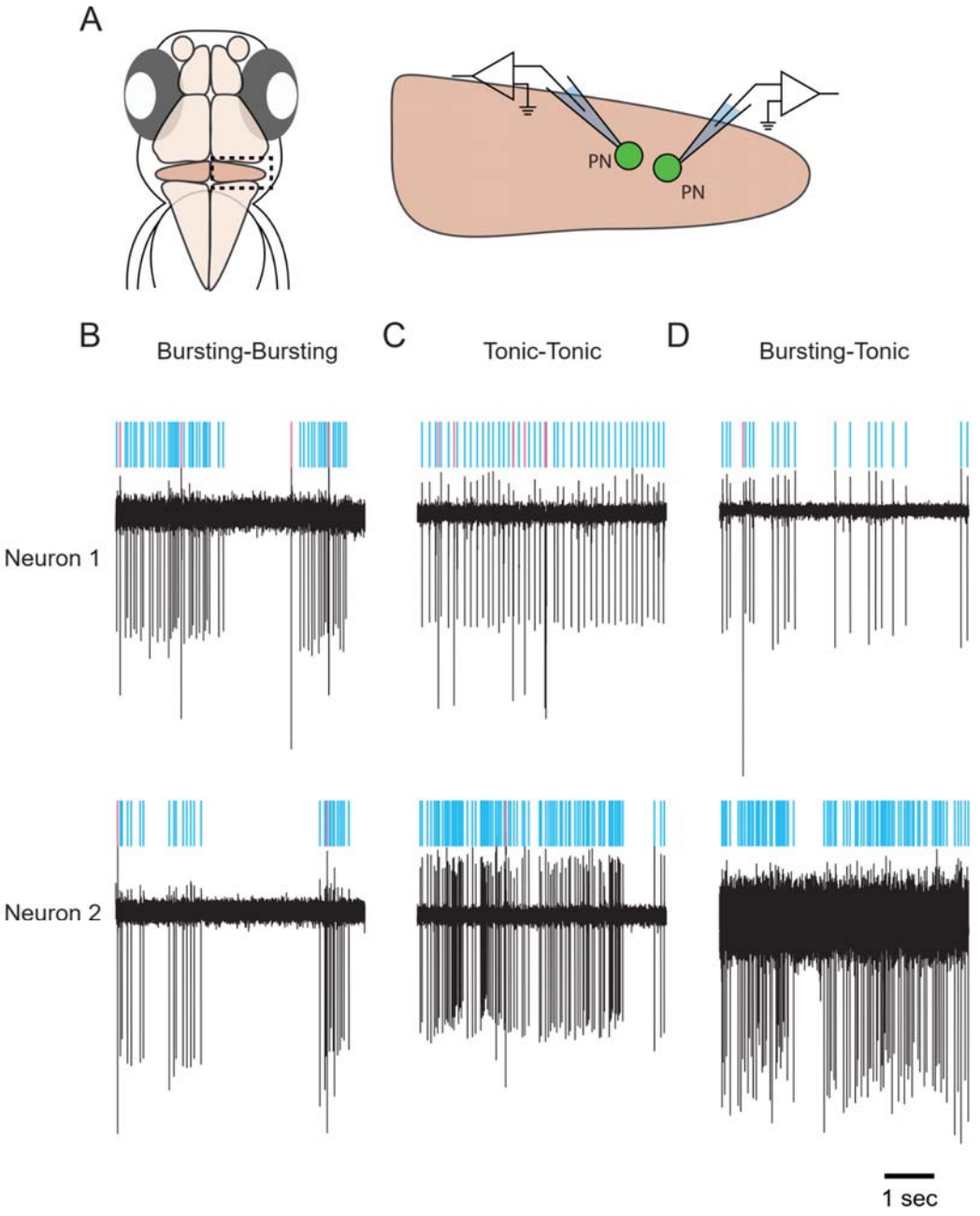


Figure 3.1 - Dual recordings to investigate the relationship between Purkinje neuron activity in different states. **(A)** Schematic of the dual recordings obtained, showing a single lobe of the cerebellum and two nearby Purkinje neurons (PN) which are recorded. **(B-D)** An example recording of each kind of state pair - Bursting-Bursting (B), Tonic-Tonic (C), and Tonic-Bursting (D).

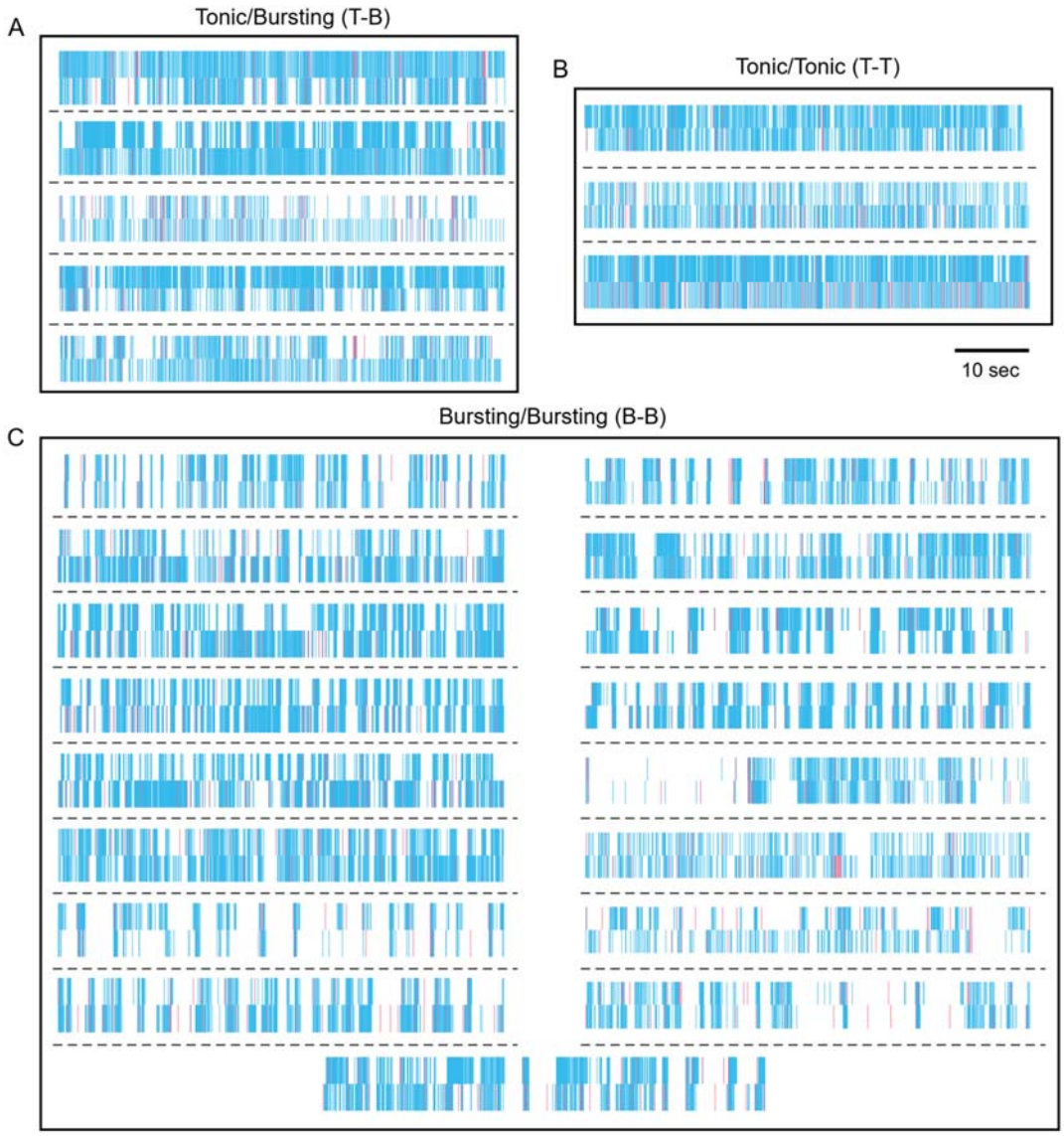


Figure 3.2 - **Event rasters for all paired PN-PN recordings in the dataset**. **(A)** Simple spike (cyan) and CF input (magenta) rasters for pairs of neurons of which one is in the tonic state and the other in the bursting state (labelled as Tonic/Bursting). **(A)** Event rasters as in (A), but for pairs of neurons that were both in the tonic state (Tonic/Tonic). **(C)** Events rasters as in (A), but for pairs of bursting neurons (Bursting/Bursting). In all panels, pairs of recordings are separated from each other by a dashed horizontal line. The scale bar in (B) applies for all recordings.

Table 3.1 *- Summary of number of pairs of PN-PN recordings by state*

	No. of pairs (n)	No. of fish (N)
Bursting-Bursting (B/B)	17	9
Bursting-Tonic (B/T)	5	5
Tonic-Tonic (T/T)	3	3
Total	**25**	**13**

Before performing correlation analyses between the spike trains of the recorded pairs, I wanted to establish that the baseline firing rates between the two neurons were not unusually different. Thus, I measured the average simple spike rate (Fig 3.3A) and the average CF input rate (Fig 3.3B) for each pair, as a function of their state combination. There did not seem to be any specific bias in the firing rates between pairs of neurons, so I proceeded with cross-correlation analysis.

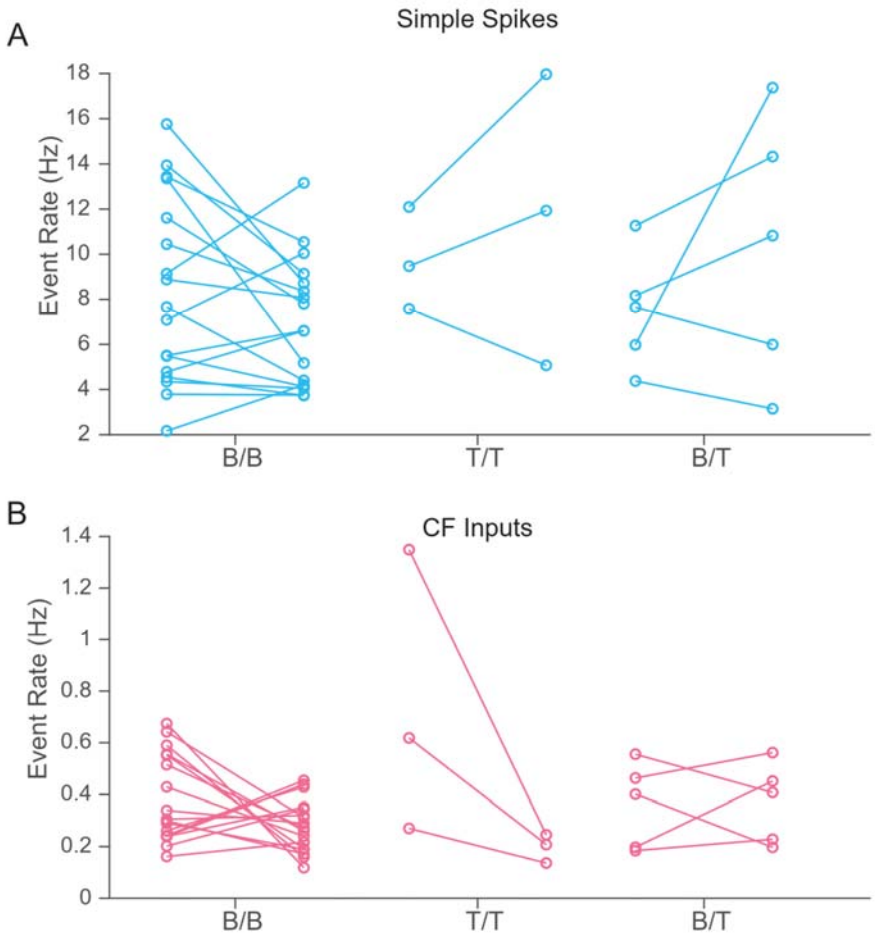


Figure 3.3 - **Baseline event rates for each pair of neurons recorded (A)** Average simple spike rate for each pair of neurons in the dataset. Each point is an individual neuron, and the point it is connected to is the other simultaneously recorded neuron. Pairs have been grouped by the combination of their cellular states (B/B = Bursting/Bursting; T/T = Tonic/Tonic; B/T = Bursting/Tonic). **(B)** Average CF input event rates for each pair of neurons in the dataset, as in (A).

I first looked at the firing rate profiles of simple spikes and CF inputs from each neuron (example pair shown in Fig 3.4A). I quantified the similarity between the two neurons' activity by calculating the Pearson's correlation coefficient between their firing rate profiles for both simple spikes (Fig 3.4B, top) and CF inputs (Fig 3.4B, bottom). From these data, it appears that the correlation in simple spike firing is highest in Bursting-Bursting pairs (median 0.5660) and lowest in Tonic-Tonic pairs (median 0.2912), with Tonic-Bursting pairs having intermediate correlations (median 0.3785). On the contrary, the CF input correlations were

low in Bursting-Bursting pairs (median 0.0896), with higher, similar correlations in both Tonic-Tonic and Tonic-Bursting pairs (medians 0.3352 and 0.3246, respectively).

The variability in the correlation between pairs of bursting neurons is rather high, with some pairs having close to zero correlation. However, all 3 pairs of tonic neurons I recorded from had a moderate correlation coefficient. It is possible that this could be just because of the high baseline firing rate of tonic cells. Similarly, it is rather surprising that the CF input correlations in Bursting-Bursting pairs are the lowest, given that they can trigger bursts of simple spikes, whose correlations are high. Thus, I wanted to verify that these obtained correlation values were, indeed, meaningful.

To do so, I obtained a significance value for each correlation by computing the distribution of values obtained by chance (by scrambling one of the cells' event times over 2000 different iterations) and calculating the p-value for the correlation, i.e. the probability of obtaining a correlation value greater than or equal to the true correlation coefficient by chance (i.e. from the random distribution). The random distribution is shown as the coloured boxplot in the upper plot in Fig 3.4C and Fig 3.4D, with the true value obtained marked as a circle. The associated p-values for each neuron pair is shown in the lower plot in Fig 3.4C and Fig 3.4D. From these, it is clear that the correlation coefficients for simple spikes are all rare to obtain randomly, whereas those for CF inputs were mostly easy to obtain by chance, suggesting that I should only interpret the distributions of simple spike correlations as meaningful.

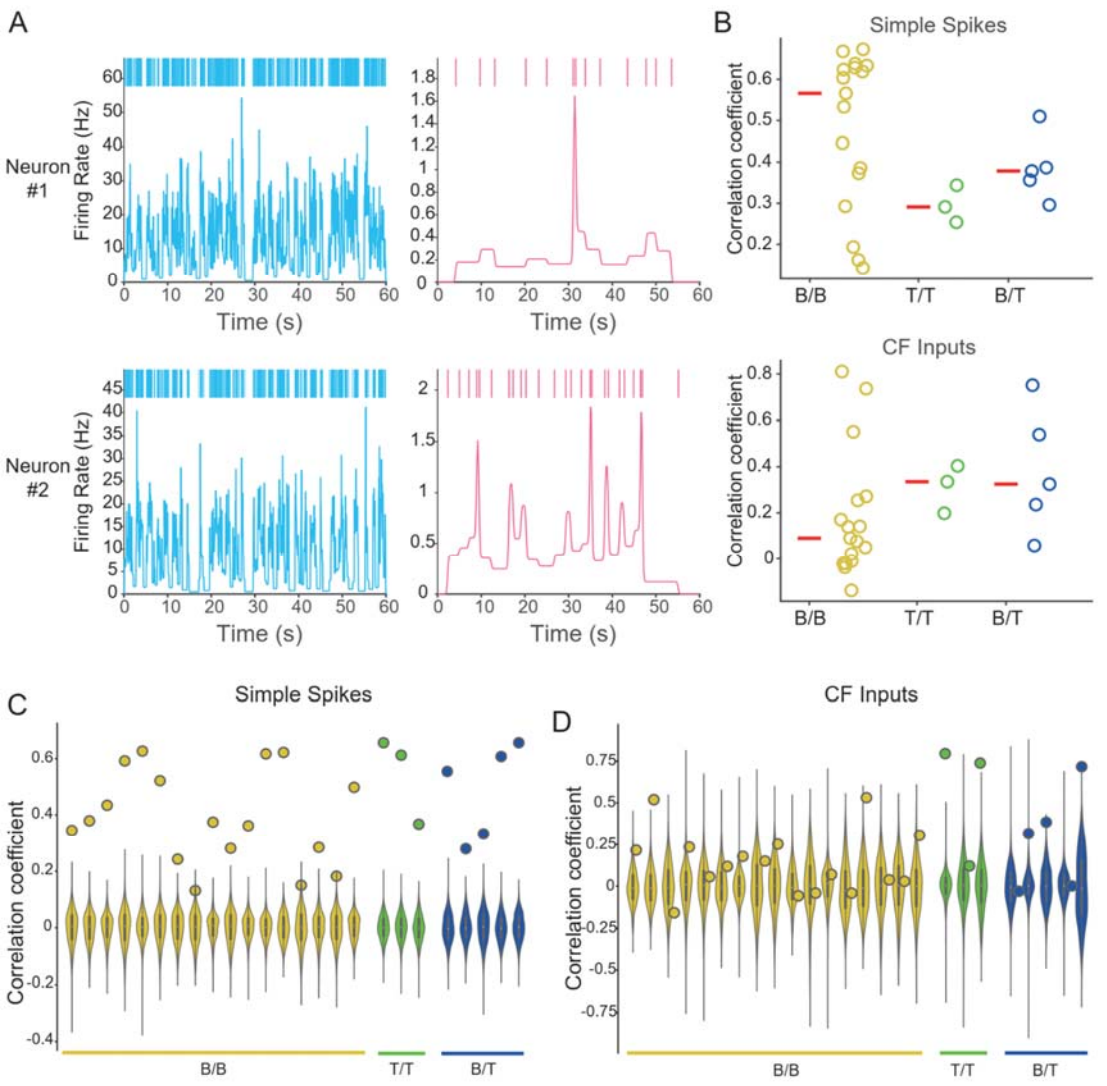


Figure 3.4 - Firing rate cross correlations by state (A) Simple spike (cyan) and CF input (magenta) spike train firing rates for an example pair (Bursting-Bursting) of cells. The event timings for each neuron are represented as rasters above their respective plots. **(B)** Pearson's correlation coefficient between the firing rates for simple spikes (above) and CF inputs (below) for each kind of state pair. B and T are short for Bursting and Tonic, respectively. The raw data are plotted alongside each boxplot. **(C)** True simple spike correlation coefficients (coloured dots) compared to the distribution of coefficients obtained by chance (violin plot), for each pair of neurons, sorted by their state combination. Correlations are significant if the true values lie outside the range of the violin plot. **(D)** Correlations and their chance distributions, as in (C), but for CF inputs.

I next asked whether the higher simple spike cross correlation between pairs of bursting neurons resulted in synchrony at the burst level. To quantify this, I decided to detect bursts of simple spikes in each electrophysiology trace and binarize them (Fig 3.5A). The degree of overlap between these binary time series was calculated as the ratio of their intersection to their union - a metric called the Jaccard Index (Chung & Edwards, 2019). The distribution of Jaccard Indices for all pairs of bursting neurons is shown in Fig 3.5B. 75% of the distribution lies above a JI of 0.5, meaning that more than half the bursts in each cell overlapped with one another.

I also measured the similarity between bursts in these pairs of cells at the burst level, i.e. for each burst in one channel, I asked what percentage of it overlapped with a burst in the other channel. This is shown in Fig 3.5C as a boxplot and Fig 3.5D as a histogram, both of which show that each burst has, on average, an 80% overlap with a burst in another nearby cell. Hence, pairs of nearby bursting neurons have a high degree of synchrony at the burst level.

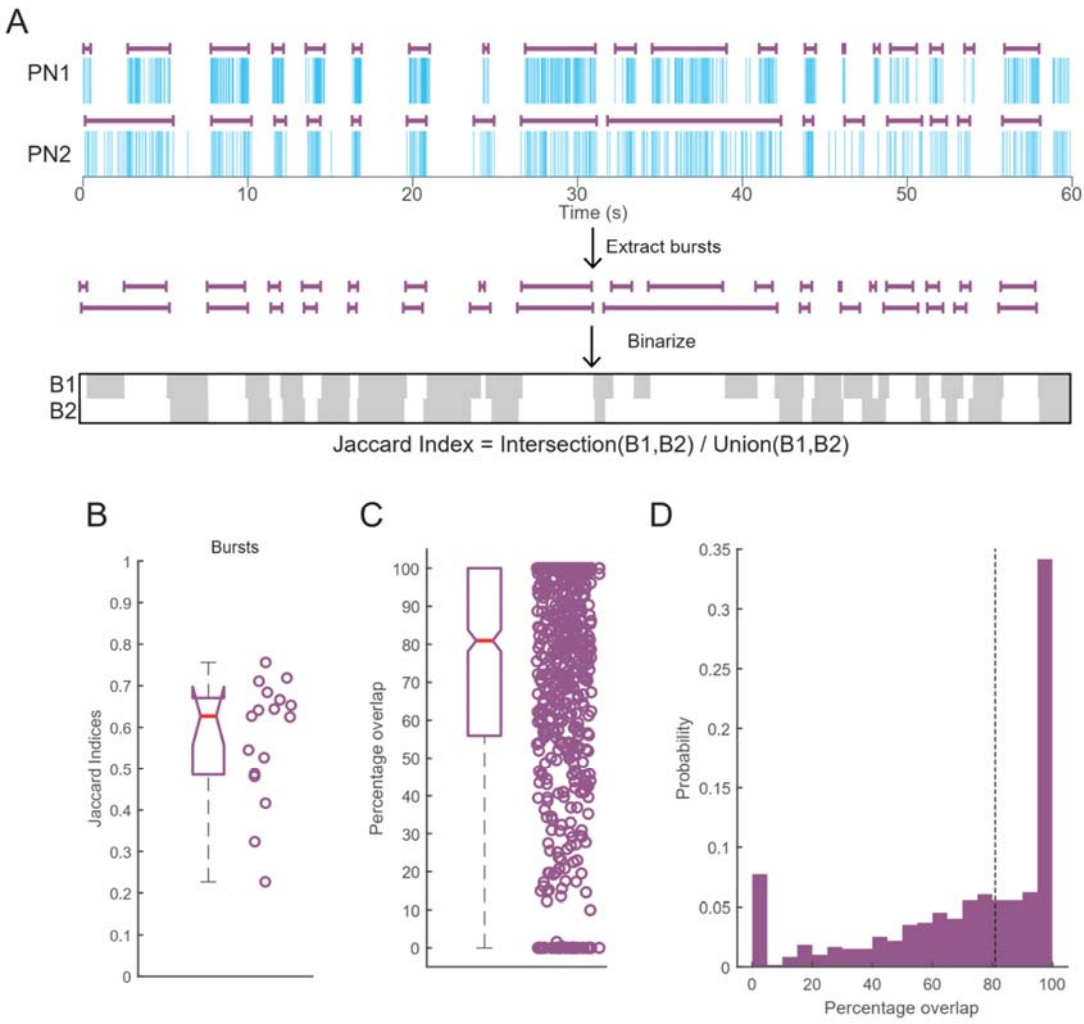


Figure 3.5 - **Overlap between bursts and CF inputs in bursting/bursting Purkinje neuron pairs (A)** Calculation of the Jaccard Index. Burst times for a pair of Purkinje neurons are first extracted. An example pair is shown, for which identified bursts are marked in purple above the corresponding simple spike raster. Burst times are then extracted and binarized such that time periods within a burst get the value 1 (white regions), while those periods outside the burst get a value of 0 (grey regions). Finally, the Jaccard Index is calculated for these binary time series as the ratio of their intersection to their union. **(B)** The distribution of Jaccard Indices for all Bursting-Bursting pairs of Purkinje neurons (n=17 pairs). **(C)** The distribution of percentage overlap of every burst with a burst in the other channel, also calculated from the binarized burst time series. The data points used to make the box plot are shown alongside it. **(D)** A histogram of the same data points as in (C). The vertical dashed line indicates the mean of this distribution.

3.2.2 Spike train distance metrics reveal that PNs with similar simple spike activity also have similar CF input activity

The analyses from firing rate profiles yielded meaningful information only for simple spikes. They also suffer from the disadvantage that these profiles are dependent on the time bins being considered (in other words, how much they have been smoothed). Furthemore, burst Jaccard indices could only be calculated for pairs of bursting neurons, which does not allow one to comment on tonic neurons. All these limitations led me to consider using less biassed methods of comparing two simultaneously-acquired spike trains: spike-train distance metrics.

In this study, I employed two distinct spike train metrics: the vanRossum distance and the SPIKE distance. The vanRossum distance relies on a time-scale parameter (τ) and convolves each spike train with a single exponential function. It calculates the sum of Euclidean distances between the two convolved traces, making it time-scale dependent (van Rossum, 2001). On the other hand, the SPIKE distance is time-scale independent and involves computing interval differences between each time point and the nearest preceding and succeeding spikes in both spike trains. These local time differences are then weighted, normalised, and averaged to obtain the SPIKE distance (Kreuz et al., 2013). Notably, the SPIKE distance yields values between 0 and 1, where 0 indicates perfect synchrony and 1 represents perfect asynchrony, unlike the vanRossum distance, which can take arbitrarily large values.

The vanRossum distances for simple spikes and CF inputs are shown in the top and bottom plots, respectively, of Fig 3.6A. All three kinds of pairs seem to have very similar vanRossum distances for both simple spikes (medians 112.9411, 122.3262 and 116.8576, respectively, for B-B, T-T and B-T pairs) and for CF inputs (medians 7.9472, 9.5929 and 7.2175, respectively, for B-B, T-T and B-T pairs). I found that there was a large correlation between the two vanRossum distances (Fig 3.6B, R = 0.60).

Like firing rate profiles, the vanRossum distances suffer from the use of a free time-scale parameter that must be chosen by the experimenter. Hence, I also quantified the time-scale independent SPIKE-distances of simple spikes and CF inputs between pairs of cells, by the type of state combination (Fig 3.6C). This metric shows that the distances between Bursting-Bursting pairs are lower (median 0.2483), in general, than those between Tonic-Tonic (median 0.2742) and Tonic-Bursting (median 0.3052) pairs. By contrast, the SPIKE distances for CF inputs are larger in Bursting-Bursting pairs (median 0.2717) than in Tonic-

Tonic (median 0.2526) and Tonic-Bursting (median 0.2216) pairs. This trend is somewhat similar to the one observed between the firing rate correlation coefficients (Fig 3.4). Moreover, just like in the case of the vanRossum distance, there exists a correlation between the simple spike SPIKE distance and the CF input SPIKE distance (Fig 3.6D), although it is more modest in comparison (R=0.41).

In sum, both the firing rate correlations and spike train distance measures suggest that pairs of bursting neurons have more similar simple spiking activity. There also exists a positive correlation between the simple-spike spike train distances and CF input spike train distances, which suggests that neurons receiving similar synaptic inputs have similar spike responses.

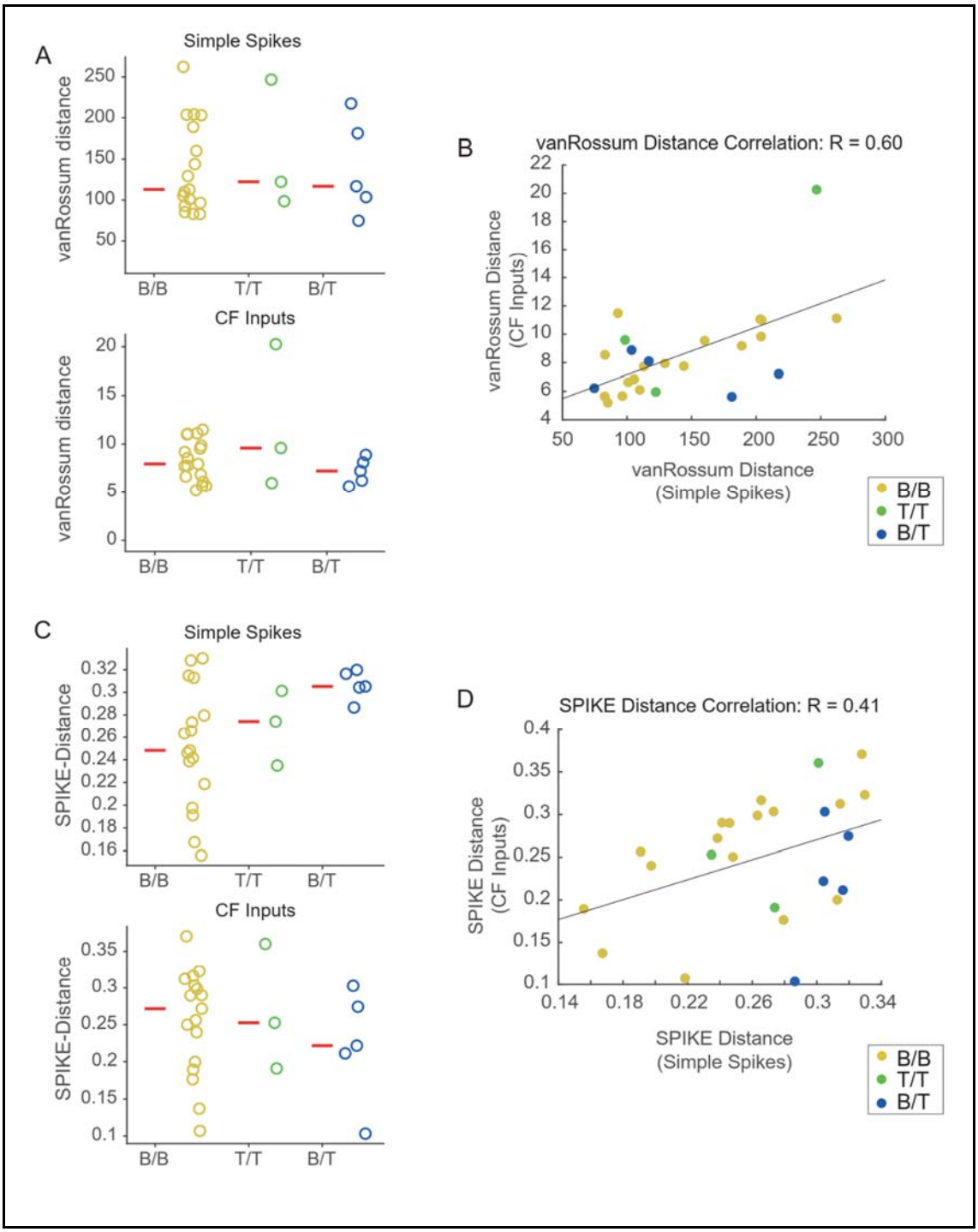

A
Simple Spikes
vanRossum distance
250
200
150
100
50
B/B
T/T
B/T
CF Inputs
vanRossum distance
20
15
10
5
0
B/B
T/T
B/T
B
vanRossum Distance Correlation: R = 0.60
vanRossum Distance (CF Inputs)
22
20
18
16
14
12
10
8
6
4
50
100
150
200
250
300
vanRossum Distance (Simple Spikes)
B/B
T/T
B/T
C
Simple Spikes
SPIKE-Distance
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
B/B
T/T
B/T
CF Inputs
SPIKE-Distance
0.35
0.3
0.25
0.2
0.15
0.1
B/B
T/T
B/T
D
SPIKE Distance Correlation: R = 0.41
SPIKE Distance (CF Inputs)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.14
0.18
0.22
0.26
0.3
0.34
SPIKE Distance (Simple Spikes)
B/B
T/T
B/T

Figure 3.6 - Measuring simple spike and CF input distances as a function of state (A) vanRossum distance between the two PN simple spike trains (top) and CF input trains (bottom), sorted by cellular states, Bursting-Bursting (B/B), Tonic-Tonic (T/T) and Bursting-Tonic (B/T). Each datapoint represents a single pair of neurons (n=25 pairs from N=13 fish), and the horizontal red line is the median of each distribution. **(B)** Correlation between the vanRossum distances for simple spikes and CF inputs. **(C-D)** Same as (A) and (B), but for the time-scale independent SPIKE distance.

3.2.3 Information theoretic measures support the hypothesis that pairs of bursting Purkinje neurons have more information shared between them

Information theoretic measures, such as mutual information, are superior to simple measures of correlation, since they capture non-linearities in the relationship between two variables and are measured on an absolute scale of bits, which facilitates comparison between different groups independent of the number of events or other differences. The drawback of these measures is the amount of data that is needed to measure them accurately (Strong et al., 1998; Victor, 2002). However, recently, a method of reliably estimating mutual information between two simultaneously-acquired spike trains was reported which doesn't require infinitely long recordings (Houghton, 2015, 2019).

I applied this method to our dataset to ask if there was more mutual information between spike trains from neurons in the same state versus those in different states. I found that the mutual information between the simple spike trains (Fig 3.7A, top) was highest in Bursting-Bursting pairs (median 1.7814 bits), least in Tonic-Tonic pairs (median 0.2513 bits) and of an intermediate value in Tonic-Bursting (median 0.7159 bits). In other words, there are almost 2 bits of information between two bursting spike trains, but less than 1 bit of information between two tonic spike trains or a tonic spike train and a bursting spike train.

On the other hand, the mutual information between CF input spike trains is comparable across all pairs of states (Fig 3.7A, bottom). The median CF spike train mutual information between Bursting-Bursting, Tonic-Tonic and Tonic-Bursting pairs are 0.8705, 0.9174 and 1.0438, respectively, meaning there is roughly 1 bit of information between CF input spike trains regardless of the state being considered. However, there was a lot of variability in the CF input mutual information between Bursting-Bursting pairs, which suggests that some of the sampled pairs could be in separate circuit sub-modules that receive dissimilar synaptic input.

Similar to the spike train distance measures, I found that there was a mild positive correlation (R = 0.27) between the simple spike and CF input mutual information values, i.e. those pairs which had a high simple spike mutual information also tended to have a higher value of CF input mutual information.

In sum, multiple kinds of analyses on these data all indicate that bursting cells tend to be more similar to other nearby bursting cells, although there is substantial variability between different B/B pairs. This variability probably comes from their position in distinct microcircuit modules. The relationship for tonic cells with other cells is harder to infer from this sample. Nevertheless, all methods suggest that cells with similar spiking activity also had more similar excitatory synaptic activity from CF inputs, further lending support to the idea that there may be several separate sub-circuit modules that receive correlated inputs.

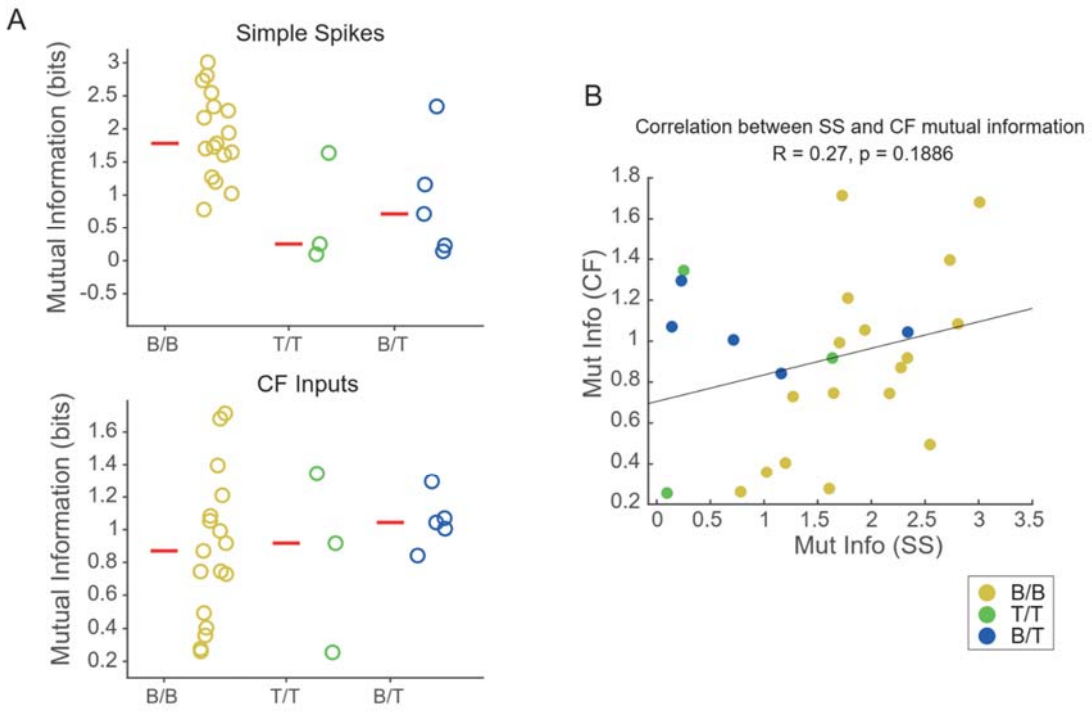


Figure 3.7 - Mutual Information between spike trains in pairs of Purkinje neurons (A) Boxplot of mutual information between two simple spike trains (top) and two CF input trains (bottom) that occur simultaneously, grouped by states - Bursting-Bursting (B/B), Tonic-Tonic (T/T) and Bursting-Tonic (B/T) - with individual data points alongside each boxplot coloured by the states. **(B)** Correlation in simple spike and CF input mutual information.

3.3 Discussion

3.3.1 Pairs of bursting neurons are more similar to each other than other pairs of states

I tried to get a preliminary handle on how the states of two neurons determine how their activities relate to one another. I did this by recording from pairs of Purkinje neurons in close proximity to one another. I consistently found, using different approaches of measurement, that a bursting cell had more similar simple spiking activity with another nearby bursting cell than a tonic cell had to another nearby tonic cell. Interestingly, the few pairs of proximal tonic cells recorded in this dataset seem to have more similar activity to one another than the bursting/tonic pairs. If this effect holds true even with a larger sample size, it could be explained by the differential synaptic response model. If tonic cells were weakly responsive to simple spike rate modulation by parallel fibers, while bursting cells were strongly responsive, for instance, there would be more common increases in spike rate between tonic and bursting neurons than between two tonic neurons, just by chance. Whether this is, indeed, what is happening in these neurons remains to be experimentally tested, since (1) we cannot know the state of a cell prior to recording and (2) because one cannot record weak parallel fiber synaptic currents while also recording spiking activity.

3.3.2 Major caveats of this approach

While electrophysiology is a high temporal-resolution way of looking at the activity of cells, the major caveat is that you can only record from so many cells simultaneously. Moreover, I have been recording from a neuron in its natural state, which is unknown *a priori.* Hence, these data are all undersampled - there are only 3 Tonic-Tonic pairs I recorded from, which is clearly not enough to make claims that can be generalised. Nevertheless, one must note that the methods of analysis described in the text above are promising and could be applied to a larger dataset of paired recordings.

Another limitation of this particular dataset is that I have recorded from proximal pairs of neurons. It would be interesting if one could collect paired recordings from distal pairs of neurons, maybe even neurons in different lobes. This would allow us to ask what similarities there are in synaptic and spiking activity across the cerebellum.

Lastly, electrophysiological recordings are quite invasive, which is why more non-invasive methods like calcium imaging are preferred to study population activity. However, how the bistable state of a neuron is reflected in its calcium signal is not self-evident. I explore this in

the next chapter.

3.4 Materials and Methods

3.4.1 Animal care

As described in Chapter 2.

3.4.2 Paired recordings from Purkinje neurons

The first recording of a pair of juxtacellular recordings was obtained using the same procedure as described in Chapter 2. After obtaining the first recording, the objective was lifted and a second electrode was brought into focus and manoeuvred to the sample. A second cell close to the first recording was targeted for recording, trying to minimise tissue disruption. Once both channels produced stable signals, 1 minute and 10 minute-long gap-free recordings were obtained before a different cell was targeted for recording.

3.4.3 Calculating firing rate functions

Instantaneous firing rate was calculated using the *instantfr* function in the MBL NSB Toolbox, at a sampling frequency of 100Hz for simple spikes and 5Hz for CF inputs. The resulting output was then filtered using low-pass Butterworth filters with a time constant of 0.1s and 1s for simple spikes and CF inputs, respectively. Cross-correlations between two firing rate functions were calculated as Pearson's product moment correlation coefficient between the traces, with zero lag. To test whether the true cross-correlation could be obtained by chance, events from one of the two channels were scrambled and the resulting firing rate function computed. The correlation coefficient between this scrambled firing rate function and the true firing rate function of the other cell was then computed. This entire process was repeated 2000 times to obtain the distribution of "chance" correlation coefficients. The p-value of the correlation coefficient reported is the proportion of "chance" correlation coefficients whose magnitude is larger than the true value obtained.

3.4.4 Calculating spike train distances

vanRossum distances were calculated using a custom function implemented in Python. For calculating this distance, a free parameter tau needs to be chosen, which was taken to be 0.01 for simple spikes and 1 for CF inputs. Bivariate SPIKE-distances were calculated using a publicly available function in Python (https://www.thomaskreuz.org/source-codes/isi-and-spike-distance). To get the distribution of metrics obtained by random chance for each cell, I computed the same distance metrics for 100 (vanRossum) or 500 (SPIKE) independent

spike trains whose event times were scrambled. p-values for each cell were then calculated as the fraction of the iterations that resulted in a metric value greater than or equal to the true metric obtained for the cell.

3.4.5 Calculating mutual information

The mutual information between two spike trains was calculated using a custom-script written in MATLAB (Mathworks, Natick, MA). It uses the Kozachenko-Leonenko approximation (Houghton, 2019) to get approximate values of mutual information between two spike trains. Let's call the first spike train ***u*** and the second spike train ***v.*** I first split the time into *N* bins of width Δt and then calculate a spike train distance metric (bivariate SPIKE distance, here) between every pair of bins within it. For each bin and different values of *h*, I find the *h-1* nearest bins in ***u*** and the *h-1* nearest bins in ***v*** (excluding the reference bin), and find the size of the overlap between the two groups. Let the mean size of this overlap be ***r.*** Then, the Kozachenko-Leonenko approximation of mutual information is $I_{KozLeo} = \log_2(N*r/h^2)$. Since this is just an estimator of the mutual information, it has a bias, which can be corrected for. In essence, the bias is the value of this estimate that would be obtained if both ***u*** and ***v*** came from independent distributions. This is calculated as the probability of randomly selecting *h-1* points out of *N-1* points from ***u*** and finding *r* of them in ***v***: $P(\text{int} = r) = {}^{h-1}C_{r-1} * {}^{n-h}C_{h-r} / {}^{n-1}C_{h-1}$. The overall bias in information is given by:

$I_0(N, h) = \sum_{r=1}^{h} P(\text{int} = r) * \log_2(N*r/h^2)$. This bias is subtracted from the Kozachenko-Leonenko approximation to yield the unbiased estimate of mutual information: $I = I_{KozLeo} - I_0(N, h)$. The entire calculation described above is performed for values of *h* ranging from 2 to *N* and the maximum value of I obtained is taken to be the best estimate of mutual information.

Chapter 4 - The representation of a Purkinje neuron's electrophysiological events in its calcium signal

4.1 Introduction

The previous chapter described methods used to assess the similarity/dissimilarity between pairs of Purkinje neurons. This was done by recording from pairs of neurons and quantifying the relationships between their electrophysiological activity at various timescales. While electrophysiology yields data at a high temporal resolution, there are limits to how many neurons can be recorded simultaneously, which in turn limits our ability to understand population-level activity at these resolutions. Neuroscience as a field has looked to imaging-based methods, particularly calcium imaging, to study the function of neuronal populations (also see Chapter 1). Calcium imaging, however, comes at the cost of temporal resolution, for various reasons. Firstly, cellular calcium dynamics in response to electrical signals (like action potentials) are slower than the signals themselves. Moreover, sensors used to report cytosolic calcium concentrations are also slow in their rise and decay kinetics. Lastly, there are limits to acquiring data from entire populations of cells at a high rate. Put together, the calcium signal reports only a low-pass filtered version of the underlying electrophysiological activity. Thus, this method has been of use predominantly to find cells that respond to stimuli and which are mostly silent otherwise.

Purkinje neurons are interesting because they are spontaneously active, have two electrical events (simple spikes and CF inputs) and exhibit membrane potential bistability. It is known that both simple spikes and complex spikes lead to increases in cytoplasmic calcium concentration, which means that a Purkinje neuron's calcium signal is not unambiguous. What effect a cell's state has on its calcium response, however, has not been explored. Here, I perform simultaneous electrophysiology and calcium imaging to model the relationship between a cell's electrical activity and its calcium signal as a function of its state. I then combine this model with the state-labelled electrophysiology dataset I compiled (Chapter 2) to obtain a state-labelled calcium signal dataset. The idea is to leverage this dataset to solve the inverse problem and identify a Purkinje cell's state from its calcium activity.

4.2 Results

4.2.1 Calcium imaging reports both simple spike bursts and individual CF inputs, independent of state

Given that PNs manifest two distinct activity patterns (tonic and bursting) and two types of electrical events (simple spike and CF events), I wondered if calcium imaging signals can reliably report these distinct patterns of activity. To do this, I performed simultaneous wide-field calcium imaging and targeted patch-clamp electrophysiology from zebrafish PNs. I used the carbonic anhydrase (Ca8) enhancer element and a minimal promoter (cfos) to drive expression of the genetically encoded calcium sensor GCaMP5G in PNs (Matsui et al., 2014). Embryos were microinjected at the 1-2 cell stage and those larvae that showed sparse labelling in PNs were selected for targeted patch-clamp electrophysiology and imaging (Fig 4.1A). I chose to perform juxtacellular recordings so as to be minimally invasive and to avoid the dialysis of cellular contents including GCaMP. I, along with Dr Mohini Sengupta, obtained such recordings from 15 cells in 12 larvae, out of which 9 cells were naturally in the tonic state, and 6 cells, in the bursting state.

Representative traces from this experiment for cells in the bursting and the tonic states are shown in Fig 4.1B and 4.1C, respectively. The first observation I made was that many of the calcium transients (marked with a hash symbol in Fig 4.1B and 4.1C) had an almost one-to-one correspondence with CF inputs (magenta lines in the rasters) in cells of both states. In tonic cells, most simple spikes are not reported as transients, as one would expect from cells that fire at high frequencies. However, in some cases, calcium transient peaks occurred in the absence of CF inputs, during times of simple spike firing (Fig 4.1C, see asterisk and cyan lines in raster). Similarly, in bursting cells, I found some calcium transients associated only with a burst of simple spikes (Fig 4.1B, asterisk). Naturally, then, one would expect to see calcium transients when CF inputs co-occurred with a burst of simple spikes, which is what I observed (marked with open circles in Fig 4.1B and 4.1C). Thus, regardless of the state of a cell, I found calcium transients that corresponded to individual CF inputs, bursts of simple spikes or a combination of both.

While this suggests that there is a causal link between the CF input's arrival and the measured calcium influx into the cell, it is not useful in the case of simple spikes. Firstly, there are multiple simple spikes per frame - approximately 2-3 spikes per frame when imaged at 30Hz. This is further worsened by the timescales of GCaMP5G's rise and decay, which occurs on the order of 100s of milliseconds (Akerboom et al., 2012; T.-W. Chen et al.,

2013). Moreover, one cannot separate out the individual effects of simple spikes and CF inputs because the two events are interspersed and thus, when calculating the triggered average response for one event, there is bleed-through from the other.

As an alternate measure, I decided to compare the number of events in the time leading up to the calcium peak (Fig 4.1D, region Ⅰ) and during the decay period (Fig 4.1D, region Ⅱ). The expectation was that if an event triggered calcium activity, there would be more of them in region Ⅰ, which includes the time of initiation of the calcium transient (Fig 4.1D), than in region Ⅱ, when the signal is decaying. GCaMP5G has a rise time constant of ~100ms, which means that the time it takes to reach a peak value from the baseline is ~201ms (modelled using a difference of single exponentials, see Methods). Hence, to be sure that the time of initiation was included in region Ⅰ, I considered a 300ms interval before the peak of the calcium signal. An interval of the same 300ms duration after the peak was taken to be region Ⅱ.

I first detected transient peaks in the dF/F signal (see Methods), of which there were a total of n=171 detected across the N=15 cells. The average simple spike counts in regions Ⅰ and Ⅱ for each cell are shown in Fig 4.1E and the average CF input counts in Fig 4.1F. On average, there were 2.4846 ± 1.1912 simple spikes in region Ⅰ and 2.6262 ± 1.2337 in region Ⅱ. As for CF inputs, there were 0.2120 ± 0.2590 events in region Ⅰ and 0.0974 ± 0.1693 in region Ⅱ. To test whether the difference between these were statistically significant, all 171 events were considered but grouped by their cell of origin, and a linear mixed-effects model was fit to the data (Table 4.1). This shows that after accounting for variation across cells, both simple spike and CF input counts between regions Ⅰ and Ⅱ were statistically significantly different ($p<0.0001$ and $p=0.0025$ for simple spikes and CF inputs, respectively). This implies that both simple spikes and CF inputs are associated with dF/F transients.

In summary, it is difficult to infer the underlying event train or even the cellular state of a Purkinje neuron directly from its calcium signal, as there is no one-to-one correspondence between calcium transients and their underlying electrophysiological events, as has also been reported previously (Knogler et al., 2017; Sengupta, 2016).

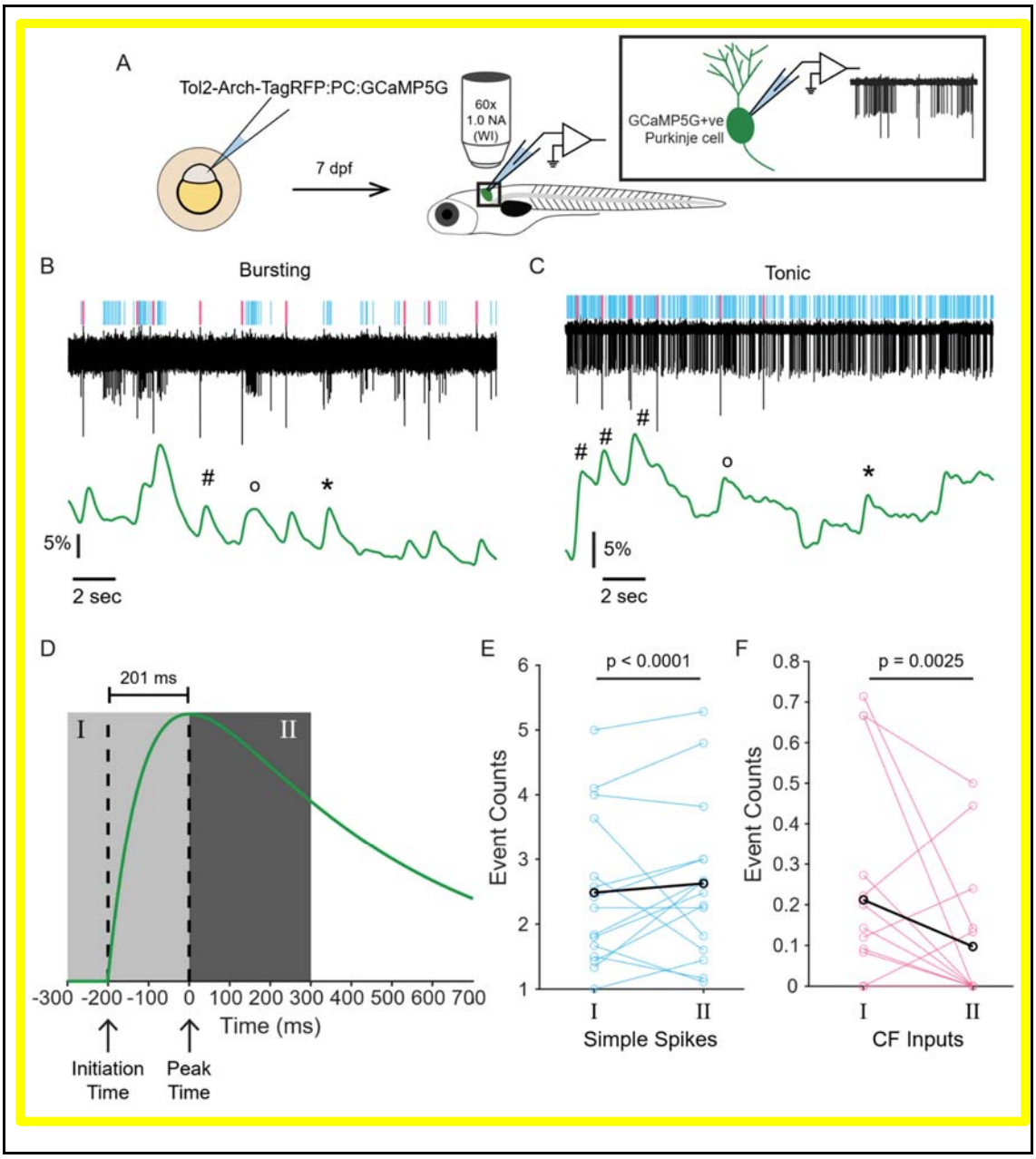

A
Tol2-Arch-TagRFP:PC:GCaMP5G
60x
1.0 NA
(WI)
7 dpf
GCaMP5G+ve
Purkinje cell
B
Bursting
C
Tonic
#
o
*
5%
2 sec
D
201 ms
I
II
-300 -200 -100 0 100 200 300 400 500 600 700
Time (ms)
Initiation
Time
Peak
Time
E
p < 0.0001
Event Counts
Simple Spikes
F
p = 0.0025
CF Inputs

Figure 4.1 - Both simple spikes and CF EPSPs contribute to the calcium signal in PNs irrespective of cellular state. (A) Schematic of the experimental design and setup of simultaneous calcium imaging and electrophysiology in larval zebrafish. Wild-type zebrafish embryos were injected at the 1-4 cell stage with the Tol2-ArchTagRFP:PC:GCaMP5G construct (left). At 7 days post fertilisation (dpf) (right), isolated PNs expressing GCaMP (right, inset) were targeted for simultaneous electrophysiology and imaging. **(B)** and **(C)** Representative traces obtained during simultaneous imaging and electrophysiology for cells in the bursting (B) and tonic (C) mode. Simple spikes and CF inputs in the trace are identified and marked in the raster above the trace in cyan and magenta, respectively. The simultaneously-recorded change in GCaMP fluorescence (dF/F) is shown below in green. Calcium transients are marked with an asterisk (*) if they correspond to only simple spikes in the electrophysiology channel. Similarly, they are marked with a hash symbol (#) if they correspond only to CF inputs, and with an open circle (o) if they correspond to both simple spikes and CF inputs. **(D)** The typical GCaMP5G response to a single action potential, assuming it follows the profile of a difference of single exponentials with a half-rise time of 100ms and half-decay time of 500ms. The vertical dashed lines indicate the time of initiation of the calcium transient and the time when the transient peaks, which corresponds to an interval of 201ms. Region Ⅰ is the period up to 300ms before the peak of the calcium signal, and region Ⅱ is the period up to 300 ms after the peak of the transient. There were a total of n=171 transient peaks in the dF/F signal detected across the N=15 cells. **(E)** Average simple spike counts in the periods corresponding to regions Ⅰ and Ⅱ as defined in (D). Each cyan line corresponds to an individual cell (N=15) and the average of all these lines is shown in black. **(F)** Average CF input event counts in the periods corresponding to regions Ⅰ and Ⅱ as defined in (D). Each magenta line corresponds to an individual cell (N=15) and the average of all these lines is shown in black.

Table 4.1 - *Linear mixed-effects model statistics for comparing event counts before and after a calcium peak (related to Fig 4.1E and Fig 4.1F)*

Model	Interpretation	DF	AIC	Test	p-value
SS_Counts_Ⅱ ~ 1	Baseline model - global mean for all	2	754.0859	-	-

	simple spike counts in region Ⅱ				
SS_Counts_Ⅱ ~ 1|Cell_ID	Model accounting for cell-wise variations in simple spike counts in region Ⅱ	3	745.6466	1 vs 2	**0.0012**
SS_Counts_Ⅱ ~ SS_Counts_Ⅰ|Cell_ID	Comparing spike counts between regions Ⅰ and Ⅱ having accounted for variation across cells	4	693.1542	2 vs 3	**<0.0001**
CF_Counts_Ⅱ ~ 1	Baseline model - global mean for all CF input counts in region Ⅱ	2	126.5696	-	-
CF_Counts_Ⅱ ~ 1|Cell_ID	Model accounting for cell-wise variations in CF input counts in region Ⅱ	3	123.932	1 vs 2	**0.0313**
SS_Counts_Ⅱ ~ CF_Counts_Ⅰ|Cell_ID	Comparing CF input counts between regions Ⅰ and Ⅱ having accounted for variation across cells	4	116.7789	2 vs 3	**0.0025**

4.2.2 The calcium signal for a Purkinje neuron can be reconstructed from its electrophysiological recording

Though the calcium signal in PNs could arise from more than one electrical event, I hypothesised that there might be sufficient information in a cell's calcium signal time series to allow us to determine whether it was firing tonically or in bursts. One possible approach to

this problem could be the application of machine learning algorithms to build a classifier that labels cells as tonic or bursting given a time series of calcium transients. For this, it was first necessary to generate a large ground truth dataset (100's of cells) comprising the electrical activity patterns and their corresponding calcium transients for training the classifier. Ideally, this dataset would take the form of what is shown in Fig 4.1B, where both electrophysiology and imaging data are simultaneously acquired. This ground truth dataset has only 15 cells, however, given that the simultaneous imaging and electrophysiology experiments are low yield and technically challenging. Hence, I needed a strategy to augment our dataset.

Mathematically, the calcium signal obtained by imaging can be thought of as the result of the convolution of spike trains with a calcium sensor kernel having a fast rise and a slow decay and described as a difference of exponentials. This transformation has been shown to work well in cells that are not spontaneously active, where every single action potential produces a large increase in intracellular calcium concentration (Rupprecht et al., 2021; Smetters et al., 1999; Vogelstein et al., 2010; Yaksi & Friedrich, 2006). I hypothesised that a similar transformation could be used to reconstruct the calcium signal profiles of PNs from their electrophysiological recordings alone. Such reconstructed calcium imaging traces with their corresponding electrical recording traces could then be applied to build the classifier.

However, given that these neurons have not one, but two electrophysiological events, each contributing differently to the calcium signal obtained from imaging, we must assign to each event a separate kernel (Fig 4.2), which I assumed could be considered independent of each other. This is a valid assumption, since simple spikes and CF events arise from distinct intrinsic mechanisms (Sengupta & Thirumalai, 2015) and because I found no correlation between average simple spike and CF input firing rates (also see Chapter 2, Fig 2.8). I used the simultaneous imaging and electrophysiology data to find optimal kernels for simple spikes and CF events. For each PN recording, I obtained simple spike and CF event rasters. Each event raster was then independently convolved with its own exponential response kernel, and the resulting "event-specific signals" were added to produce the final reconstruction (Fig 4.2). I then calculated the correlation coefficient and the mean squared error (MSE) between the reconstructed calcium signal time series and the experimentally obtained calcium imaging time series.

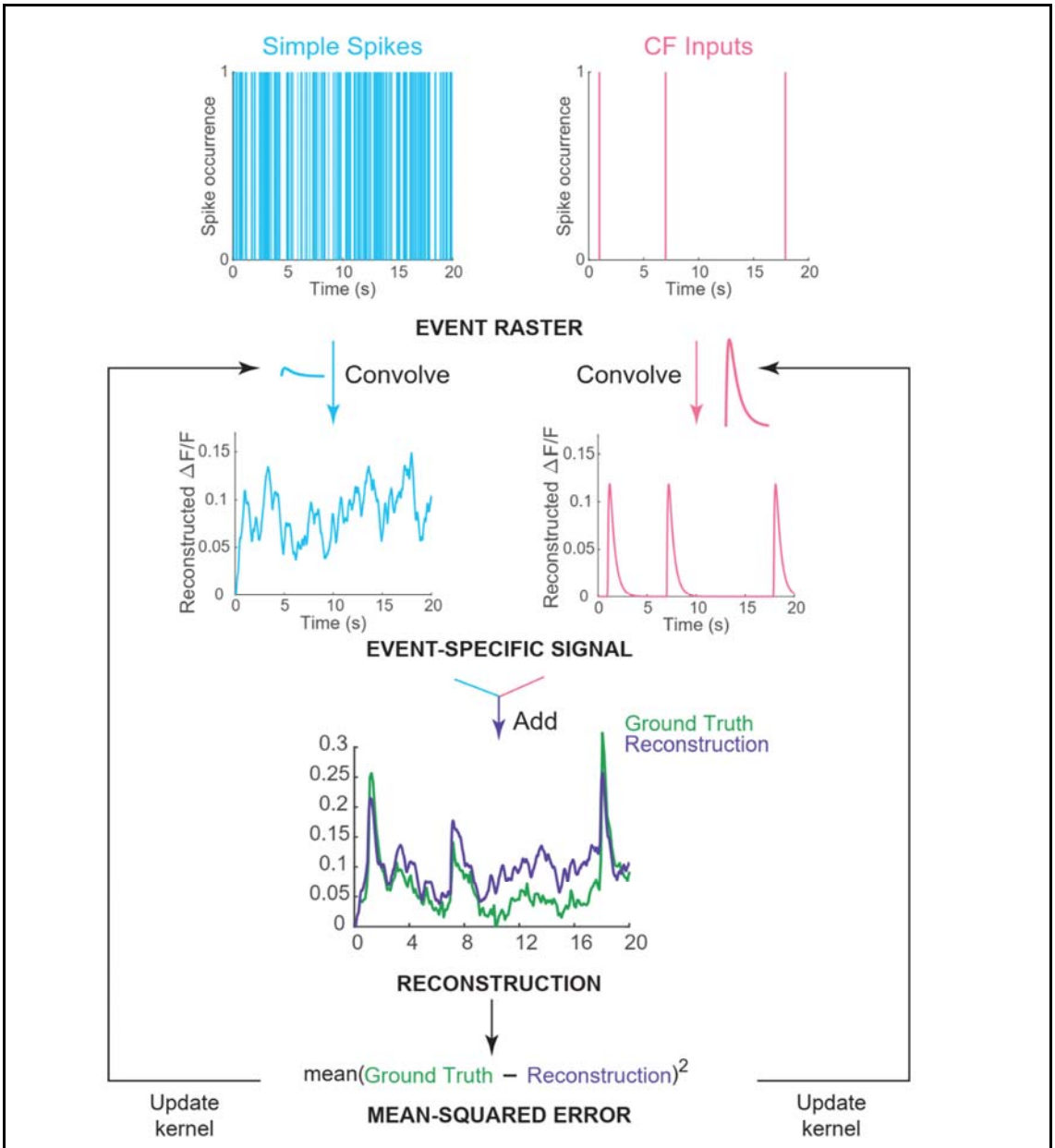


Figure 4.2 - Reconstructing the calcium signal from electrophysiology traces. Schematic showing the method used to reconstruct the calcium signal from an electrophysiological trace. Simple spike (cyan) and CF input (magenta) rasters (left) are independently convolved with calcium sensor kernels (cyan for simple spikes, magenta for CF inputs) to generate an event-specific calcium signal time series (middle). These are then simply added together to get the final reconstruction (right), shown in dark purple. This example is the same cell as shown in the rightmost plot in Fig 4.4B.

With this description in place, the problem then becomes one of finding optimal coefficients for each response kernel. This optimization was done by gradient descent to minimise the MSE (mean-squared error) between the ground truth and the reconstruction (Fig 4.2). The process was performed independently for each cell in the dataset. Then, these optimal kernel combinations were used to make reconstructions for other cells in the dataset, and the kernel that generalised best across all cells, as quantified by least total MSE (Fig 4.3A) and highest total cross-correlation (Fig 4.3B) between the ground truth and reconstruction, was taken to be the best kernel for the next steps. Kernel numbers 7 and 8 were very close to each other in both scores, so kernel 7 was chosen for further analysis.

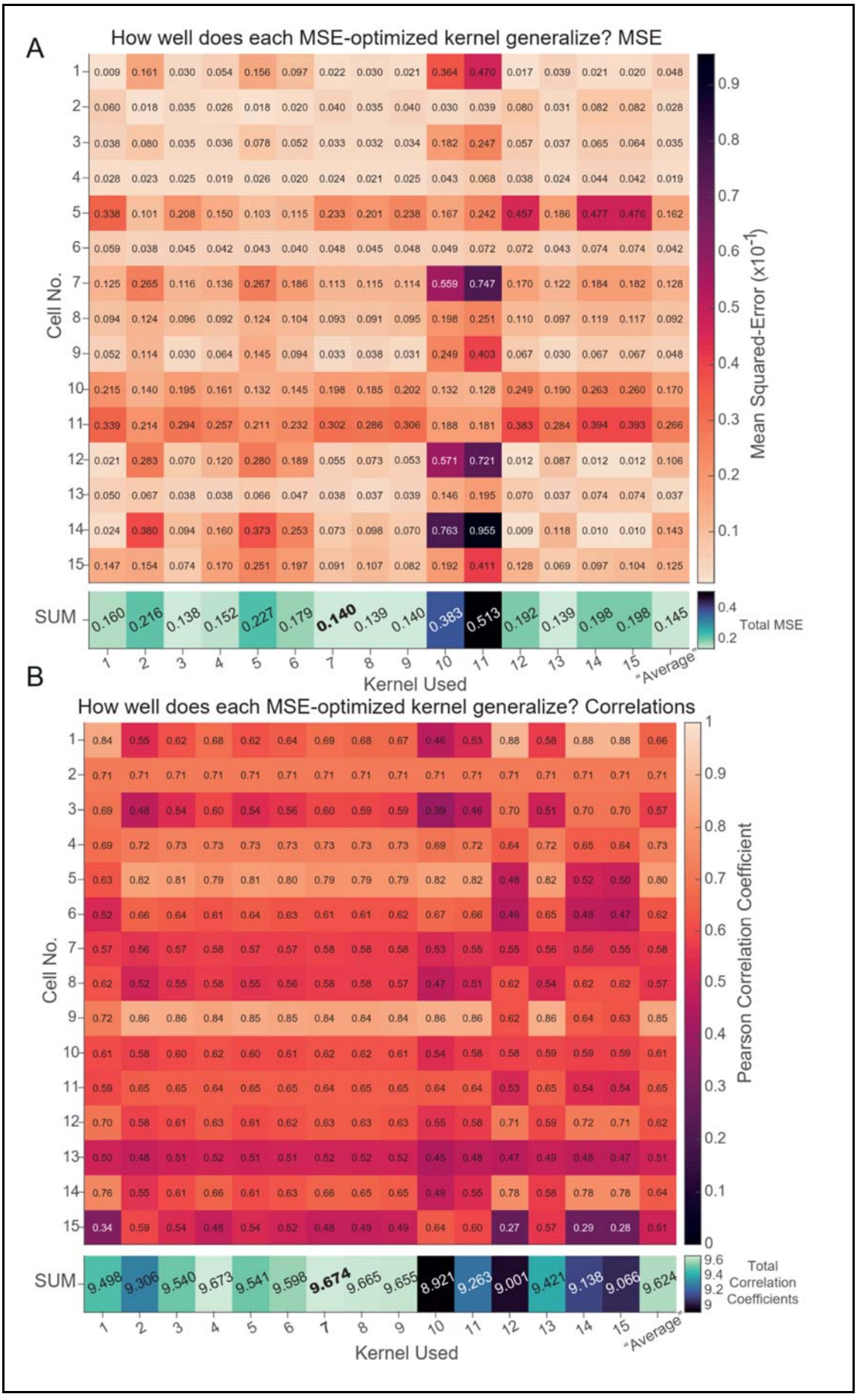

A
How well does each MSE-optimized kernel generalize? MSE
Cell No.
Mean Squared-Error (x10-1)
SUM
Total MSE
Kernel Used
B
How well does each MSE-optimized kernel generalize? Correlations
Cell No.
Pearson Correlation Coefficient
SUM
Total Correlation Coefficients
Kernel Used

Figure 4.3 - Finding the most generalizable GCaMP5G kernel (A) Mean-squared errors (MSEs) between the ground truth and reconstructed calcium signal for each cell and kernel combination used. Each row represents a single cell and each column, the kernels used. The “average” kernel’s coefficients were the mean of the coefficients of the kernels optimised for each cell. The last row, with a different colormap, represents the total MSE across all reconstructions for a kernel. The values corresponding to each colour are shown in the colour bars on the right. Bright colours represent lower MSEs. **(B)** Similar representation as above, but for Pearson’s correlation coefficients between the ground truth and reconstructions. Note that the colormaps are reversed as compared to (A), so that brighter colours represent higher correlations. In both (A) and (B), the kernel chosen for all future work (i.e. no 7) is highlighted in bold.

Although this is a simple forward model with calcium dynamics assumed to be linear and time-invariant, I found that this method to approximate the calcium signal works well, no matter the state of the neuron (Fig 4.4B). The quality of the reconstructions has been quantified by measuring the correlation coefficient between the ground truth and the reconstruction (Fig 4.4C). To benchmark the correlation coefficient values obtained against those obtained simply by chance, I also calculated the correlation coefficient between the ground truth and a scrambled version of the reconstruction, as well as that between the ground truth and a reconstruction generated after shuffling all the event times in the recording (Fig 4.4C). In both the latter cases, the correlation disappears, suggesting that the reconstructions generated by this forward model are, indeed, faithful reconstructions. I next calculated the residuals between ground truth and reconstructions, the distribution for which is shown in Fig 4.4D. It is not normally distributed, and has a positive skew, which means that these reconstructions tend to underestimate the true signal. Nevertheless, the small values of the residuals further lends support to the idea that the forward model of calcium signal reconstruction works well.

Given that only groups/bursts of simple spikes produce a calcium transient amplitude comparable to that generated by an individual CF input, one would expect a much smaller response kernel for simple spikes than for the CF input. Indeed, the optimal kernels found by this method show that the CF input response kernel is approximately 7.5 times larger than the simple spike response kernel (Fig 4.4E).

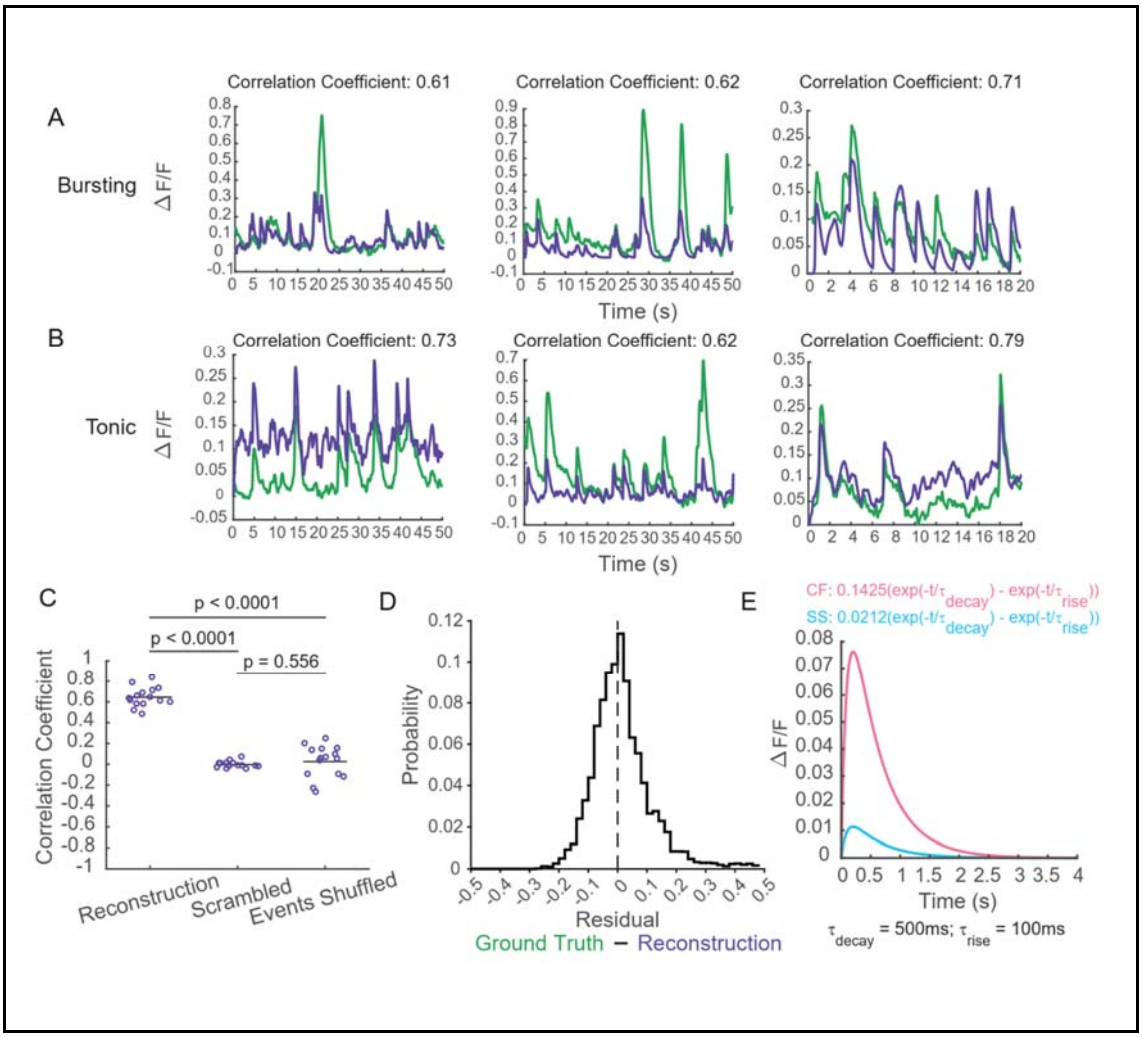

A
Bursting
ΔF/F
Correlation Coefficient: 0.61
Correlation Coefficient: 0.62
Correlation Coefficient: 0.71
Time (s)
B
Tonic
ΔF/F
Correlation Coefficient: 0.73
Correlation Coefficient: 0.62
Correlation Coefficient: 0.79
Time (s)
C
p < 0.0001
p < 0.0001
p = 0.556
Correlation Coefficient
Reconstruction
Scrambled
Events Shuffled
D
Probability
Residual
Ground Truth – Reconstruction
E
CF: 0.1425(exp(-t/τdecay) - exp(-t/τrise))
SS: 0.0212(exp(-t/τdecay) - exp(-t/τrise))
ΔF/F
Time (s)
τdecay = 500ms; τrise = 100ms

Figure 4.4 - Quality of calcium signal reconstructions obtained from electrophysiological data using a convolution-based forward model (A) Comparison of ground truth (green) and reconstructed (purple) calcium signals for 3 randomly chosen bursting cells. The Pearson's correlation coefficient between the two traces is marked above the plot for each pair. **(B)** Comparison of ground truth (green) and reconstructed (purple) calcium signals for 3 randomly chosen tonic cells, like in (B). **(C)** Quantification of the quality of reconstructions obtained using the optimal kernel combination. Correlation coefficients were calculated between the ground truth for each cell (n=15 cells; 6 bursting, 9 tonic) and either the true reconstruction, a scrambled version of the reconstruction, or the reconstruction for when simple spike and CF input event times were shuffled. A Kruskal-Wallis test yielded a p-value of 2.108 x 10^-7, which was followed by a post-hoc Dunn's test for individual comparisons, the latter of which is marked on the figure. **(D)** The distribution of the residuals between the ground truth and reconstructed calcium signal, pooled across all cells (n = 15 cells). **(E)** The optimal calcium sensor kernels for simple spikes (cyan) and CF inputs (magenta). The formula for each kernel is mentioned above the plots, and the values for rise and decay time constants are mentioned below the plot.

4.2.3 A state-labelled calcium signal dataset was generated by reconstructing the calcium signal for a compilation of electrophysiological traces

Having a method of converting any given PN electrophysiology trace to its corresponding calcium imaging trace allowed me to generate simulated calcium imaging traces for a compilation of all electrophysiological recordings acquired in our lab (Fig 4.5A). To do so, I went back to the compiled dataset of electrophysiological recordings (see Chapter 2) and applied the forward model of calcium signal reconstruction to it. This yielded a state-labelled reconstructed calcium signal dataset. The recordings used to generate this dataset were of different durations with a median duration of 60 seconds. In order to get multiple samples per recording, I decided to split the traces into 10- second-long non-overlapping chunks, representative samples from which are shown in Fig 4.5B. This produced a dataset of state-labelled calcium imaging traces (n=928 traces; n=580 bursting, n=348 tonic) that is much larger than could have been obtained by simultaneous imaging and electrophysiology of individual cells. The entire complement of non-overlapping samples is shown as a heatmap in Fig 4.5C.

To test whether there were any features of these traces that could be used to unambiguously classify their source state, I compared various properties between the two

classes - the number of dF/F peaks, average peak dF/F amplitude, area under the dF/F curve, mean dF/F, standard deviation of dF/F and the coefficient of variation of dF/F (Fig 4.5D). I found that most of the properties of the traces between the two classes were largely similar. While several of these differences are statistically significant (Table 4.2), they have small effect sizes, and the aim was not to just find properties that are statistically distinguishable from each other but to be able to classify a trace as having come from either a tonically firing cell or a bursting cell. To do so, I first tested whether a principal components analysis (PCA) on the dataset of trace features would be a useful strategy. There are large overlaps between tonic and bursting datapoints' principal components (Fig 4.5E). This implies that linear methods of separating the two classes weren't successful and I needed to consider more complex models to solve this classification problem.

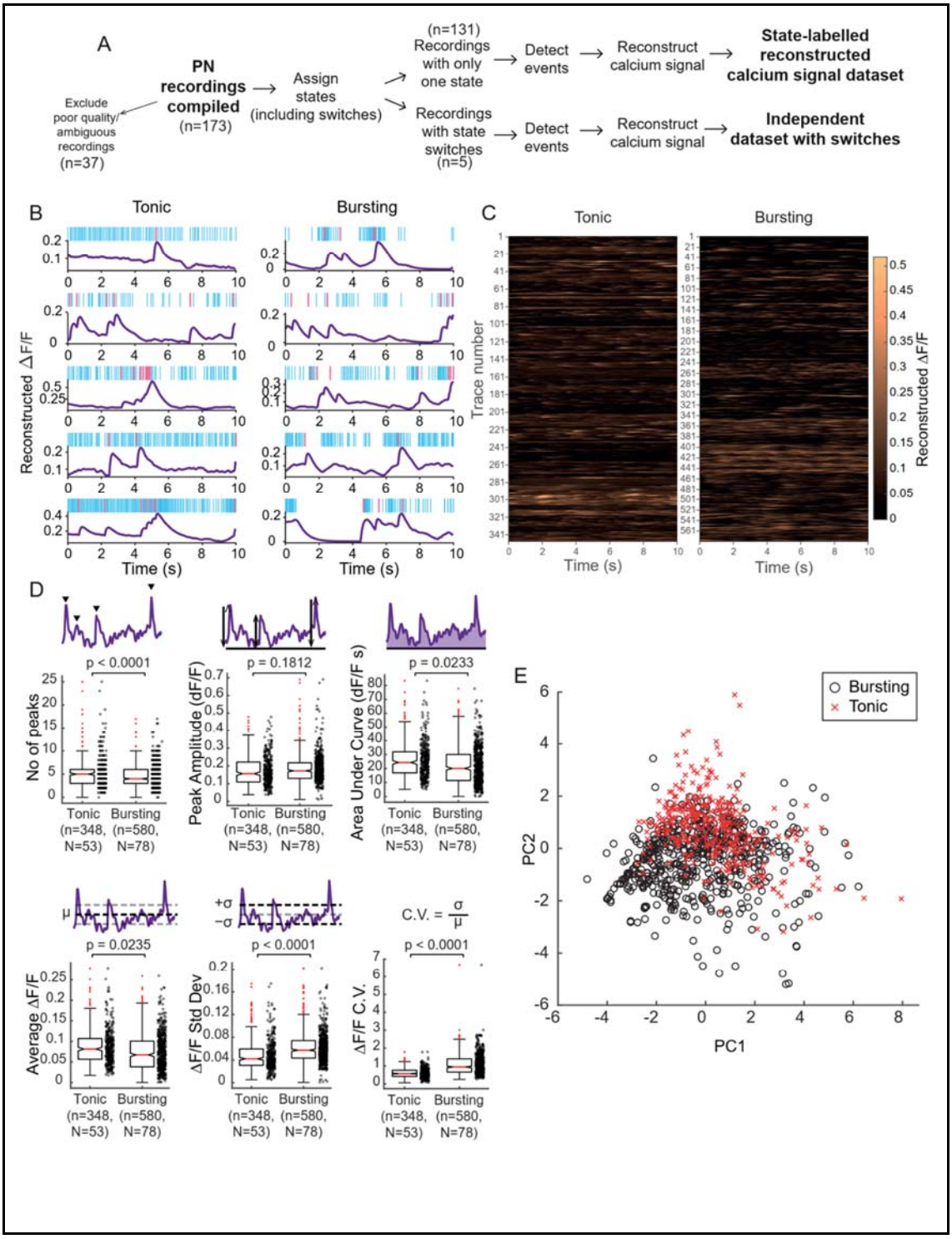

A
PN recordings compiled (n=173)
Exclude poor quality/ ambiguous recordings (n=37)
Assign states (including switches)
(n=131) Recordings with only one state
Recordings with state switches (n=5)
Detect events
Reconstruct calcium signal
State-labelled reconstructed calcium signal dataset
Independent dataset with switches
B
Tonic
Bursting
Reconstructed ΔF/F
Time (s)
C
Tonic
Bursting
Trace number
Time (s)
Reconstructed ΔF/F
D
p < 0.0001
p = 0.1812
p = 0.0233
No of peaks
Peak Amplitude (dF/F)
Area Under Curve (dF/F s)
Tonic (n=348, N=53)
Bursting (n=580, N=78)
p = 0.0235
p < 0.0001
p < 0.0001
Average ΔF/F
ΔF/F Std Dev
ΔF/F C.V.
C.V. = σ/μ
E
Bursting
Tonic
PC1
PC2

Figure 4.5 - Generating a labelled dataset of reconstructed calcium signal traces (A) Flowchart showing how electrophysiological recordings from PNs were processed to generate the state-labelled calcium signal database. **(B)** Representative randomly sampled traces from the state-labelled reconstructed calcium signal dataset. The left column has reconstructions from tonic cells and the right column has reconstructions from bursting cells. Above each trace is the raster showing the events in the source recording, with simple spikes in cyan, and CF inputs in magenta. **(C)** Heatmaps representing all the reconstructions in the state-labelled calcium signal dataset, for each state, after splitting all traces into 10 second-long non-overlapping chunks. (n=348 tonic traces; n=580 bursting traces). **(D)** Comparing the distribution of trace properties for both states. The top three plots show (from left to right) the number of peaks, peak amplitude and area under the curve of the reconstructions in the dataset. The bottom three plots show (from left to right) the mean, standard deviation and the coefficient of variation of dF/F values. n's are the number of 10s-long samples. N's are the number of recordings used to generate the samples. p-values were computed using linear mixed-effects models (see Table 4.2). **(E)** Principal Components Analysis of trace properties, with data points marked by state - tonic (red crosses) and bursting (black open circles).

Table 4.2 - *Linear mixed-effects model statistics for comparing the various trace properties (related to Fig 4.5D)*

Parameter	**Model #**	**Model**	**DF**	**AIC**	**Test**	**p-value**	**Effect Size (*b*) [Tonic relative to Bursting]**
No of peaks (NPeaks)	1	NPeaks ~ 1	2	5066.33	-	-	-
	2	NPeaks ~ 1|Cell	3	4550.27	1 vs 2	<0.0001	-
	3	NPeaks ~	4	4532.48	2 vs 3	**<0.0001**	**2.3575**

		State|Cell					
Average peak amplitude (Amp)	1	Amp ~ 1	2	-2549.08	-	-	-
	2	Amp ~ 1|Cell	3	-3000.83	1 vs 2	<0.0001	-
	3	Amp ~ State|Cell	4	-3000.62	2 vs 3	**0.1812**	**-0.0122**
Area Under the Curve (AUC)	1	AUC ~ 1	2	7192.146	-	-	-
	2	AUC ~ 1|Cell	3	6315.809	1 vs 2	<.0001	-
	3	AUC ~ State|Cell	4	6312.661	2 vs 3	**0.0233**	**4.4053**
Mean dF/F (Mean)	1	Mean ~ 1	2	-3365.760	-	-	-
	2	Mean ~ 1|Cell	3	-4244.373	1 vs 2	<.0001	-
	3	Mean ~ State|Cell	4	-4247.505	2 vs 3	**0.0235**	**0.0147**
Standard deviation of dF/F (Stdev)	1	Stdev ~ 1	2	-4566.303	-	-	-
	2	Stdev ~ 1|Cell	3	-5034.750	1 vs 2	<.0001	-
	3	Stdev ~ State|Cell	4	-5052.501	2 vs 3	**<.0001**	**-0.0135**
Coefficient of Variation	1	CV ~ 1	2	1378.348	-	-	-
	2	CV ~ 1|Cell	3	558.4086	1 vs 2	<.0001	-

of dF/F (CV)	3	CV ~ State|Cell	4	511.7434	2 vs 3	**<.0001**	**-0.4744**

Machine learning methods have successfully been used to solve such problems when the data points are not linearly separable (He et al., 2016). Thus, I turned to developing a machine learning model to identify the state of a cell from a calcium signal time series, which is discussed in the next chapter.

4.3 Discussion

4.3.1 Ambiguity of the calcium signal in zebrafish Purkinje neurons

Calcium imaging has been one of the core experimental techniques in neuroscience for more than three decades and has seen many major developments in sensor quality, microscopy and analytical algorithms and software (T.-W. Chen et al., 2013; Göbel & Helmchen, 2007; Peron et al., 2015; Renninger & Orger, 2013; Tian et al., 2009; Yang & Yuste, 2017). Yet, the technique is largely used to identify regions of the brain that respond to presented stimuli or correlate with behavioural output under well-defined experimental conditions. This is resultant from the fact that the predominant mode of transient calcium elevation occurs due to the opening of voltage-gated calcium channels during action potential firing (Borst & Helmchen, 1998; Göbel & Helmchen, 2007). Thus, these calcium transients are quite brief, typically lasting only a few milliseconds and report spiking activity of neuronal populations during sensory stimulation and/or motor output, which can be directly inferred using methods like nonnegative deconvolution (Vogelstein et al., 2010), maximum likelihood inference (Deneux et al., 2016) or supervised machine learning methods (Rupprecht et al., 2021).

In Purkinje neurons of adult mammals, too, calcium signals can be mapped back to their origin, because there is a spatial separation of calcium responses to simple and complex spiking. Only high-frequency bursting of simple spiking, but not calcium spiking, evokes a detectable calcium transient in the soma (Lev-Ram et al., 1992; Ramirez & Stell, 2016). On the other hand, calcium transients in the dendritic tree are produced in response to both spontaneous and evoked climbing fibre input (Kitamura & Häusser, 2011; Lev-Ram et al., 1992).

In contrast, however, I found that the calcium signals of zebrafish Purkinje neurons do not unambiguously report their activity. Somatic calcium transients can correspond either to bursts of simple spikes or to the arrival of climbing fibre (synaptic) input. The existence of

two spike trains, each of which contributes to the calcium signal, precludes the use of spike inference methods, which assume that a single spike train explains the origin of an observed calcium signal. The existence of two modes of firing - tonic and bursting - further renders the calcium signal hard to interpret directly.

It is known that these longer time scale dynamics can carry information to downstream neurons (Zeldenrust et al., 2018). However, as the currently available calcium sensors themselves are much slower compared to calcium dynamics, it is obvious that the recorded dF/F signals must manifest signatures of such long-term dynamics.

4.3.2 Data augmentation using forward generative models

Inferring spikes from a neuron's calcium signal isn't always straightforward. However, the converse is much easier. Since the early use of calcium indicator dyes in neurons, it has been known that the calcium signal of a neuron is a low-pass filtered version of its activity (Wei et al., 2020). The relationship between spiking activity and fluorescent calcium reporter traces has been modelled as a convolution of the spike train with an exponential kernel (Vogelstein et al., 2010). This method has even been previously used with Purkinje neurons in rats to demonstrate that only simple spikes are reported by the somatic calcium signal in these animals (Ramirez & Stell, 2016), unlike in the case of zebrafish.

I used this simple, linear forward model to convert electrophysiological event rasters to their corresponding calcium signal with considerable success. This allowed me to enlarge our dataset of labelled calcium signals, which was particularly useful since simultaneous imaging and electrophysiology is a low-throughput method of obtaining the same. The idea of using generative models to extend datasets is called "data augmentation" and is a powerful tool that finds frequent use in the field of machine learning. In these applications, artificial neural networks called Generative Adversarial Networks (GANs) are typically trained to simulate new data based on statistics learnt from real data. The training of these networks itself often relies on moderately sized datasets, though, which is why I did not pursue this strategy. The augmented data was fairly large enough to apply machine learning methods on it, which is what is described in the next chapter.

4.4 Materials and Methods

4.4.1 Simultaneous calcium imaging and electrophysiology

Indian WT larvae microinjected with the Tol2-tagRFP-T:PC:GCaMP5G construct (gift from Dr. Hideaki Matsui, Niigata University, Japan) and screened for sparse labelling were used

for this experiment. Purkinje neurons expressing GCaMP5G were targeted for the recordings. Widefield imaging was performed using a 60x water-immersion objective (1.0 NA) at an Olympus BX61WI microscope. Excitation light was provided using a pE300-ultra (CoolLED, Andover, England). Imaging was performed using an EVOLV EM-CCD camera (Photometrics, Tucson, AZ, United States) at the rig, with an exposure of 30 ms, and an on-chip gain of 170. An ROI was drawn around the targeted cell, so that only an isolated cell was imaged. Imaging was synchronised with electrophysiology by interfacing with the Digidata 1440A.

Image analysis was performed in Fiji/ImageJ (Schindelin et al., 2012). In order to identify cell boundaries, the averaged time series was used as a reference and an ROI drawn around the cell. Bleach correction was performed using the Exponential Fitting Method in the Bleach Correction plugin in Fiji. The average raw pixel intensity values within the ROI for each time point was extracted. The baseline intensity, F0, was taken to be the 5th percentile of these values. dF/F was then calculated for each frame using the formula (F - F0)/F0, where F represents the fluorescence in that frame and F0 the baseline fluorescence. Following this calculation, a Gaussian filter was applied to the time series in MATLAB R2019b (Mathworks, Natick, MA), using its function ‘smoothdata’, with a window size of 15.

To detect transients in the calcium signal, dF/F traces were first normalised to a range of [0, 1] using min-max feature scaling. Peaks in the normalised trace were then detected using the findpeaks function in MATLAB (Mathworks, Natick, MA), with a minimum peak height of 0.1, a minimum peak prominence of 0.05 and a minimum interval of 500ms between consecutive peaks.

4.4.2 Calcium imaging reconstruction from electrophysiology data

Simultaneous calcium imaging and extracellular electrophysiology data was used to infer optimal GCaMP5G kernels, which are of the general form

$$A_1 e^{-t/\tau_{decay}} - A_2 e^{-t/\tau_{rise}}$$

for both simple spikes and CF inputs of zebrafish PNs. Simple spike and CF input timings were first extracted from the electrophysiology data for a single cell. Simple spike timings were convolved with a simple spike kernel, with amplitudes A1s and A2s and CF input timings were convolved with a similar CF input kernel, with amplitudes A1c and A2c. The sum of both these convolutions was taken to be the putative calcium imaging reconstruction, downsampled to match the sampling rate of calcium imaging. The mean squared error (MSE) between the reconstruction and the true calcium imaging trace was calculated. The

values for A1s, A2s, A1c and A2c were obtained by minimising the MSE between the true calcium imaging trace and the reconstructed trace using gradient descent. To ensure that local minima were avoided, the gradient descent algorithm was repeated with randomised starting points. The kernels with these coefficients were taken to be the optimal simple spike and CF input kernels, respectively, for that cell.

The ‘optimal’ kernel would be one that generalised best. This was tested by using the kernels from one cell to reconstruct calcium imaging for other cells for which ground truth data was available. As before, the MSE and also the Pearson correlation coefficient between the reconstructed trace and the true trace were calculated. The final ‘optimal’ kernel was obtained by taking the kernel which generalised the best, i.e. had minimum MSE and maximum correlation across all cells. To test whether the correlation was real or not, the Pearson correlation coefficient between a scrambled reconstructed trace and the true trace was also calculated, for reference.

This process of convolving simple spike and CF timings with their optimal GCaMP5G kernels and summating the resultant traces was then used to generate a calcium imaging trace from both extracellular and intracellular Purkinje cell recordings.

4.4.3 Building a state-labelled calcium imaging dataset

Intracellular and extracellular recordings from larval zebrafish Purkinje neurons acquired during the period 2011-2020 by various members of the lab were compiled. Good quality recordings were those in which the resting membrane potential of the cell was not below -80mV or not above -20mV, were uninterrupted gap-free recordings with no current injections and had two clearly distinguishable event types. Only those recordings which met the aforementioned criteria were retained. Besides poor quality recordings, those recordings with an ambiguous, indeterminate state were also excluded. These recordings were renumbered randomly, to remove grouping of data collected by the same individual. Recordings were then classified as tonic or bursting. If there was a switch in state within the recording, the recordings were labelled as “switch”, with the appropriate state switch timing and kind (tonic-to-bursting, or bursting-to-tonic) marked. A summary of the recording information is in Table 2.1.

Here, data was split into two groups - one that contained traces only with a single state throughout the recording, and another that contained traces with switches in state. Once split, the spike detection protocol described above was applied to each group of recordings to detect the simple spikes and CF inputs. These event timings were then used to generate

simulated calcium imaging traces as described above. Combining the reconstruction with the assigned state resulted in the state-labelled calcium imaging dataset.

4.4.4 Principal Components Analysis

Principal Components Analysis was performed using the PCA module that is part of the sklearn package in Python. The input to the model was a vector with the trace features for each 10 second long trace in the reconstructions dataset (number of peaks, peak amplitude, area under the curve, mean, standard deviation and coefficient of variation). A 2-component PCA was performed, and the resulting fit (shown in Figure 4.5E) explained 88% of the variance in the data.

Chapter 5 - Developing machine learning models to identify bistable states from calcium imaging data

5.1 Introduction

The previous chapter described simultaneous electrophysiology and imaging from Purkinje neurons, which showed that the calcium signal in these neurons is ambiguous, reporting both simple spikes and CF inputs. Knowing this, I could use this data to develop a generative model of a cell's calcium signal given its electrophysiology signals. I combined this with the compiled dataset of Purkinje neuron recordings to develop a state-labelled reconstructed calcium signal database. I then found that the properties of a calcium trace from both tonic and bursting cells are not linearly separable. Nevertheless, I hypothesised that there could be latent information about long-term cellular dynamics in the calcium signal. Recent advances in the field of machine learning and artificial intelligence have shown promise in their ability to solve nonlinear problems. Thus, I decided to try applying machine learning methods to the problem of identifying the bistable state from which a given calcium signal is likely to have originated.

This chapter describes the process of identifying the best machine learning models that are not only capable of classifying bistable states in Purkinje neurons of larval zebrafish, but which also generalise well to other systems. This work was done in close collaboration with Dr Prasanta Kumar Ghosh and his student Mr Sathvik Udupa (Department of Electrical Engineering, Indian Institute of Science, Bangalore).

5.2 Results

5.2.1 Training neural network models to distinguish neuronal state from calcium imaging data

In order to find good machine learning models to solve this classification problem, we performed 5-fold cross validation, which is used to ensure that models generalise well and do not overfit the training data (Fig 5.1). On each fold, only 60% of the training traces were shown to the network in the training phase (Fig 5.1, white boxes), the network optimised on 20% of data used as a validation set (Fig 5.1, yellow boxes), and the generalizability of the trained network assessed using the remaining 20% (Fig 5.1, blue boxes). Each of the 5 folds had a different 60-20-20 split, such that the training and cross-validation datasets had a balance of traces from both classes.

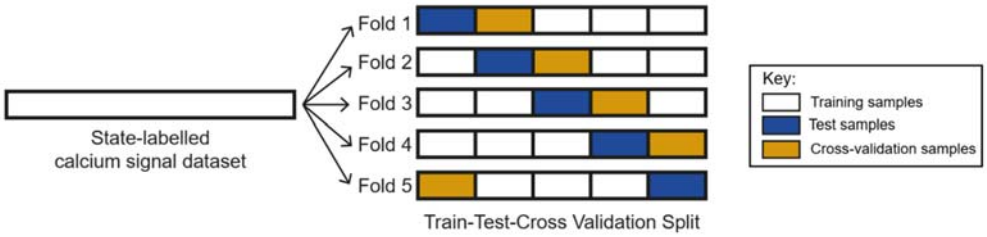


Figure 5.1 - Five-fold train/test/cross-validation split used for machine learning. Flowchart showing how the state-labelled calcium signal dataset was split into training and test datasets using a 5-fold 60-20-20 split into training (white boxes), test (blue boxes) and cross-validation (yellow boxes) samples, respectively. The networks that performed well in the training phase, as assessed by performance on the cross-validation samples, were then challenged with the unseen test data (blue boxes) in each fold. An independent dataset with state switches was also used to further validate the networks obtained (see Fig 5.5)

We used this strategy to test several different machine learning models on this classification problem - Support Vector Machines (SVMs), Deep Neural Networks (DNN), Convolutional Neural Networks (CNNs) and convolutional recurrent neural networks (CNN-LSTM, i.e. a Convolutional Neural Network with Long Short-Term Memory). The prediction accuracies and F1 scores (a metric that adjusts for class imbalances as well as whether a prediction was a true/false positive/negative, defined in the methods section) for the various models tested are shown in Fig 5.2A and Fig 5.2B, respectively.

Of these models, the 1D-CNN and CNN-LSTM models performed best across all folds. However, the former only predicts from snapshots of 10s-long snippets of the calcium trace, and the networks had more erroneous performance when predicting on the entire time series (data not shown). The CNN-LSTM model, however, also incorporates a recurrent neural network module called LSTM (Long Short-Term Memory) (Hochreiter & Schmidhuber, 1997), which has the ability to learn features across an entire sequence and not just across neighbouring positions in a trace sequence, like CNNs do.

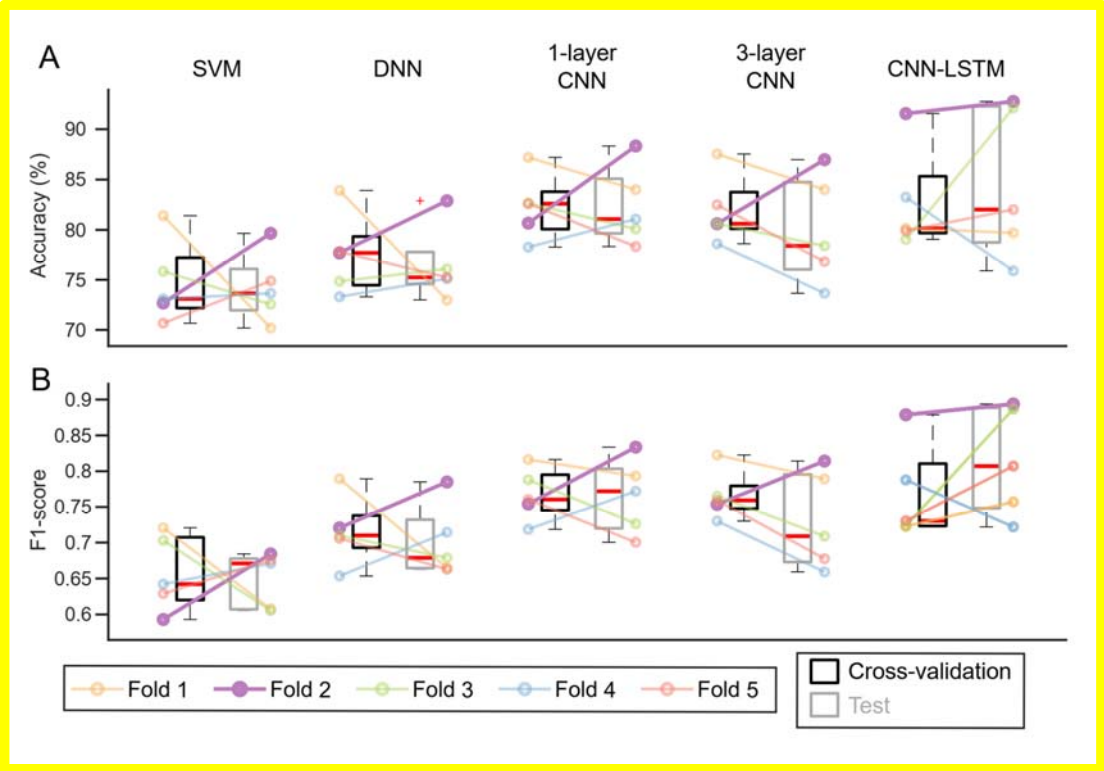


Figure 5.2 - Comparison of the performance of various machine learning models. (A) The percentage accuracy of the classifications made by five different network architectures - Support Vector Machines (SVM), Deep Neural Networks (DNN), 1- and 3-layer Convolutional Neural Networks (CNN) and a Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM). Each boxplot represents the distribution of the metric obtained after the training phase was completed. The black box plots show the classification percentage accuracy at the cross-validation phase, while the grey ones show the same with the unseen test data. Each connected pair of overlaid data points corresponds to one of the five different train/test/cross-validation folds, the colour code for which is shown in the legend below. **(B)** The same plots as (A), but showing the F1-score (see Methods), instead of percentage accuracy. **(C)** Area under the ROC curve for each fold of the trained CNN-LSTM networks for the cross-validation (black) and test (grey) cases.

The architecture of the network is shown in Fig 5.3. The input to the network is a trace sampled at 30Hz, which is some multiple of 10 seconds in duration. It is first remapped to values between 0 and 1 using min-max normalisation and then processed by a CNN, which acts as a local feature extractor that also down samples the input data, such that it produces 4 vector outputs at a 1Hz sampling frequency. These outputs are passed to a bidirectional

LSTM, one time-step at a time. A forward LSTM cell looks at trends in the forward direction (Fig 5.3, blue dotted box and arrow), while a reverse LSTM cell goes through the data in the reverse direction (Fig 5.3, red dotted box and arrow). Each of the two cells produces a "hidden state" as its output. Given that the LSTM is a recurrent neural network, this output is fed back into the network along with the next appropriate time step from the CNN's output, a process which is repeated 9 more times (for every 10 second trace). In essence, the LSTM is detecting trends across the entire sequence. The final hidden states from the forward and reverse LSTMs, produced after passing through the entire trace, are retained and concatenated. These outputs are combined by passing them through a Linear layer, which performs a matrix multiplication to produce an output that is 10 samples long. At this step, each sample represents the likelihood that the trace in question comes from a cell in the bursting mode, predicted at 1 second intervals. The likelihood of each time step having come from the tonic state can then be inferred (Fig 5.3, grayscale boxes on the right). The state with the larger likelihood score is taken to be the final classification at each time step.

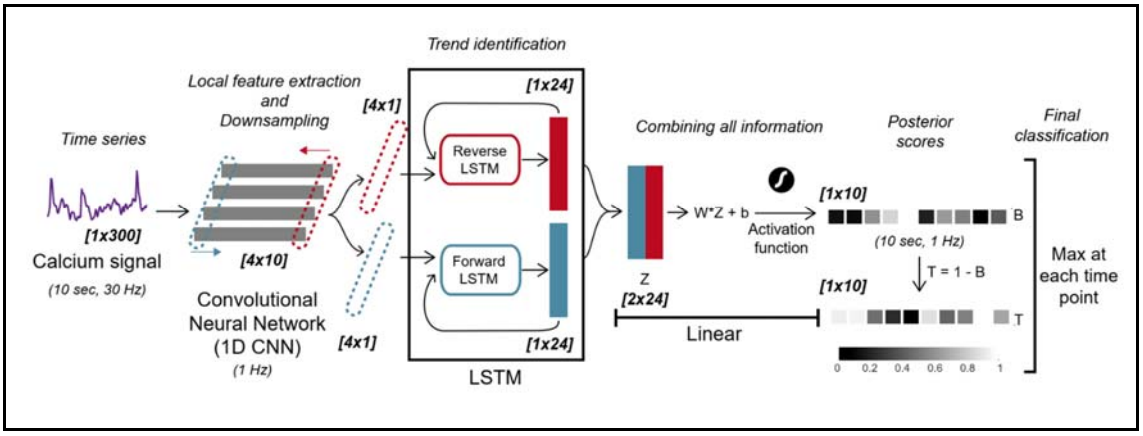


Figure 5.3 - Network architecture of the CNN-LSTM. The network takes an input time series trace sampled at 30Hz (purple) and that is a multiple of 10 seconds long in duration (here just 10 seconds long). This trace is first normalised using min-max normalisation, following which it is passed through a 1-dimensional convolutional neural network (1D CNN), which has 4 kernels with a step size and stride length of 30 each. The resulting 4 traces (grey boxes) are downsampled, local feature-extracted versions of the original trace. All 4 outputs are passed to the LSTM module (central black box) for trend identification. The LSTM is a bidirectional one, which processes traces in both the forward (blue arrow) and reverse (red arrow) directions. A single time step is taken from all the CNN outputs (red and blue dotted boxes) and passed to the respective LSTM cell, which then processes the input to produce a resultant "hidden state" output (red and blue filled boxes). This output is passed back into the LSTM cell along with the next sample from the CNN's output traces, until every time step from it has been processed (here 10 rounds). Once this is done, the final "hidden state" from each LSTM cell is retained and concatenated before their weighted average is taken by a Linear layer. The resulting average is converted to a "posterior score" using an activation function. These posterior scores represent the likelihood of the cell being in the bursting state at each time step. The likelihood of it being in the tonic state can be calculated from this. The final call is taken to be the state which has a higher posterior score. The text in italics represents the net effect of each phase of the neural network. The numbers in parentheses indicate the sampling frequency and/or the duration of the resultant vectors at various stages of processing by the neural network. Similarly, the numbers in bold and italics within square brackets indicate the size of the vectors/matrices at various stages of passing through the network.

Five networks were produced, one for each fold of cross-validation. The median accuracy was 80.15% for the cross-validation phase and 82.02% for the test phase (Fig 5.2A), whereas the median F1-scores were 0.7302 and 0.8071 for the cross-validation and test phases, respectively (Fig 5.2B). Next, we looked at the Receiver Operating Characteristic (ROC) curve for each network fold. This curve plots the false positive rate against the true positive rate of the classifier for various classification threshold values, and the area under this curve is an indicator of the quality of the classifier. The area under the ROC curve (AUC) for each of the five network folds is plotted in Fig 5.4A. The AUC values for all folds on both cross-validation and test data are much higher than 0.5, indicating that these networks have been trained well and make accurate predictions.

Of all the network folds, Fold 2 looked particularly promising, because it (1) consistently outperformed the other folds with respect to classification accuracy and F1-scores (Fig 5.2A, B, purple), and (2) had much higher AUC values than the other network folds (Fig 5.4A). The ROC curve for this network fold is shown in Fig 5.4B, which clearly shows that the network's true positive rate is much higher than its false positive rate at all classification threshold values. Thus, we decided to retain this network for future use, and have dubbed it CaMLsort (**Ca**lcium imaging and **M**achine **L**earning based tool to **sort** intracellular state).

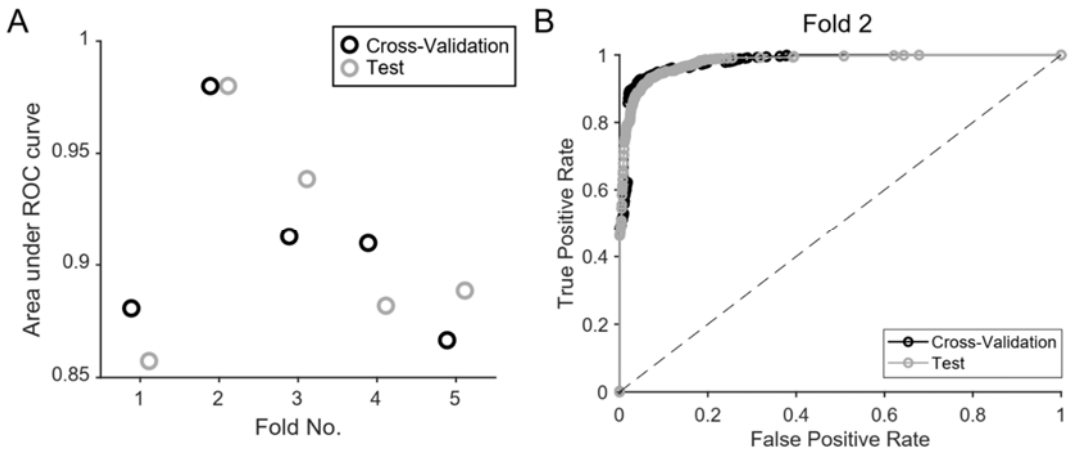


Figure 5.4 - Receiver Operating Characteristic (ROC) curves for trained CNN-LSTM networks. **(A)** Area under the ROC curve for each fold of the trained CNN-LSTM for both cross-validation (black) and test (grey) data. **(B)** ROC curve for Fold 2 of the CNN-LSTM for cross-validation (black) and test (grey) data.

5.2.2 Trained neural network models perform well on previously unseen data

To be sure that networks weren't performing well just because of peculiarities in the dataset used for training and testing, we presented an independent set of calcium signal reconstructions to the network that were not used in any phase of the training or testing. These reconstructions, of which there were n=19 in total, also included those from recordings with a state switch in them (also see Chapter 2). The raw predictions from the network are shown in Fig 5.5A (middle). Also shown at the top are the ground truth class labels.

At first glance, it appears that the network isn't making accurate calls, because each cell's predicted state alternates between both states. However, these fluctuations are transient and the overall F1-scores for the raw predictions are >0.5 for 13 out of the 19 cells (median

score 0.777). The distribution of F1-scores is shown in Fig 5.5C. On closer observation, it is clear that the 6 cells for which F1-scores were low (< 0.5) had a switch in state during the recording. The poor prediction accuracy for these can be explained by the fact that the training dataset didn't have any cases with state switches.

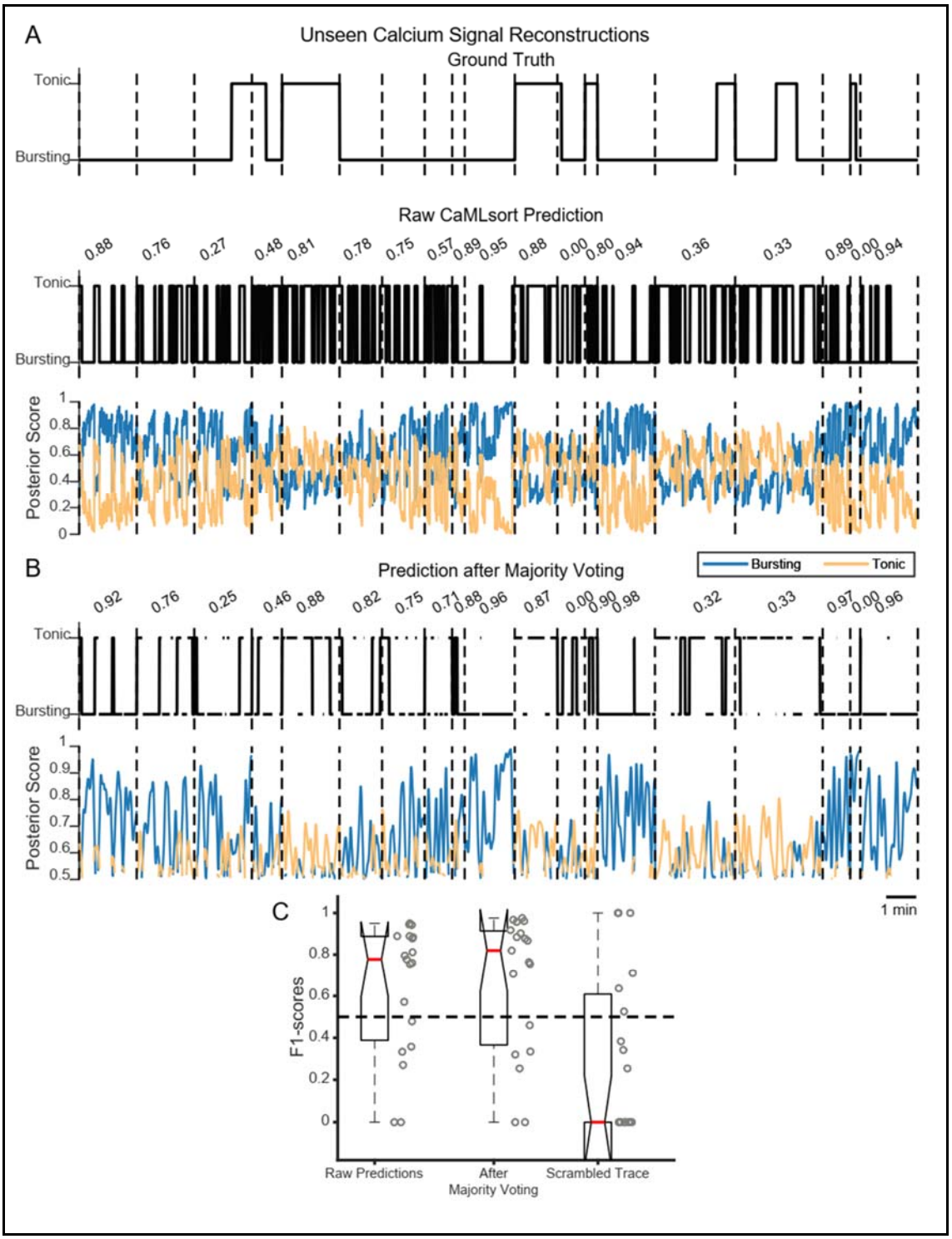


Figure 5.5 - CaMLsort predictions from previously unseen calcium signal reconstructions. (A) Raw CaMLsort class label predictions from previously unseen calcium signal reconstructions (n=19 recordings). Predictions from CaMLsort (middle) are compared against the ground truth state (top) at each time step. Predictions for each recording were made independently, but have been stitched here for the purposes of representation. Each recording is separated by a vertical dashed line. The number above each prediction is the F1-score for that recording. The bottommost plot has posterior scores for each time step, with those for the "bursting" class in blue, and those for the "tonic" class in light orange. **(B)** CaMLsort predictions (top) that remain after majority voting with a window length of 7 steps has been used to retain only confident classifications (see Results and Methods for details). The posterior score (bottom) for each class at each time step, taken as the average class posterior score across all 7-step windows that included said time step. As in (A), the posterior scores for the "bursting" class are in blue, whereas those for the "tonic" class are in light orange, and the numbers above each prediction represent the F1-scores for that recording. **(C)** The distribution of F1-scores between the ground truth state and various predictions from CaMLsort - either the raw predictions (left), predictions obtained after majority voting (middle), or the raw predictions obtained by passing a scrambled version of the calcium signal as an input.

Given our prior knowledge that the state of a neuron doesn't flicker rapidly between tonic and bursting, we decided to smoothen out the raw predictions using a majority voting strategy. The idea was to take a rolling window of length *N* around each time step of the prediction and ask what the majority of the network calls were across all the windows that included the time step. If the average posterior scores for the majority class were also higher than the average posterior scores for the minority class, we would confidently assign the majority class to that time step. If not, no confident call would be made for that time step. For example, to make a confident call at time t_0 with a window length of 5, we would first extract all the 5-step windows that include time t_0 (there will be 5 such windows, except at the start and end, where there will be fewer). Next, we would average the posterior scores within each of the windows and assign the 'average state' as the state with the higher posterior score. In this case, let's say these were 2:3 in favour of bursting. If the average bursting posterior score for the 3 "bursting" windows was greater than the average tonic posterior score for the 2 "tonic" windows, the state for time t_0 would be confidently assigned as "bursting". If not, the prediction at time t_0 would be left blank.

Using this majority voting strategy with a window length of 7, we were able to make confident calls for >90% of the duration for all 19 cells (the regions of the trace where no confident call could be made are marked as gaps). As seen in Fig 5.5B, the jitter in the raw CaMLsort predictions was smoothened by this process, and the F1-scores showed a marginal improvement as well, with a median F1-score of 0.8194 (Fig 5.5C). As before, the cases where F1-scores were <0.5 corresponded to recordings with state switches. As an extra control, we obtained CaMLsort predictions from signals that were scrambled. These predictions had much lower F1-scores (Fig 5.5C). In fact, when presented with scrambled traces, CaMLsort almost always (99.76% of the time) predicted the state to be tonic, with a very large, consistent differential between the posterior scores for the tonic (mean ± stdev: 0.7810 ± 0.0565) and bursting (mean ± stdev: 0.2190 ± 0.0565) classes. This implies that CaMLsort has, indeed, captured a latent trend present only in the true signal and not in a collection of random numbers of comparable magnitude.

To further challenge the trained networks, we decided to generate artificial spike trains that represent extreme cases of the tonic or bursting class (see thesis Chapter 2), reconstruct the calcium signal for these spike trains and ask the network to classify these completely synthetic data. The raw CaMLsort predictions obtained for these data (Fig 5.6A) were of high quality (median F1-score 0.8571, Fig 5.6C), although these, too, had several reversals in the state predictions. On applying the majority voting strategy described above, however, these reversals were smoothened (Fig 5.6B), yielding a median F1-score of 0.9189 (Fig 5.6C) and a confident prediction for at least 80% of the recording duration in every cell. As was observed in the previous case, scrambling the signal prior to prediction by the neural network yielded F1-scores which were either 1 (all right) or 0 (all wrong) (Fig 5.6C), as one would expect from predictions that were all of one class (tonic).

To sum up, we showed that CaMLsort has learnt to sort the cellular state of a Purkinje neuron based only on its calcium signal and is capable of producing confident calls of high quality.

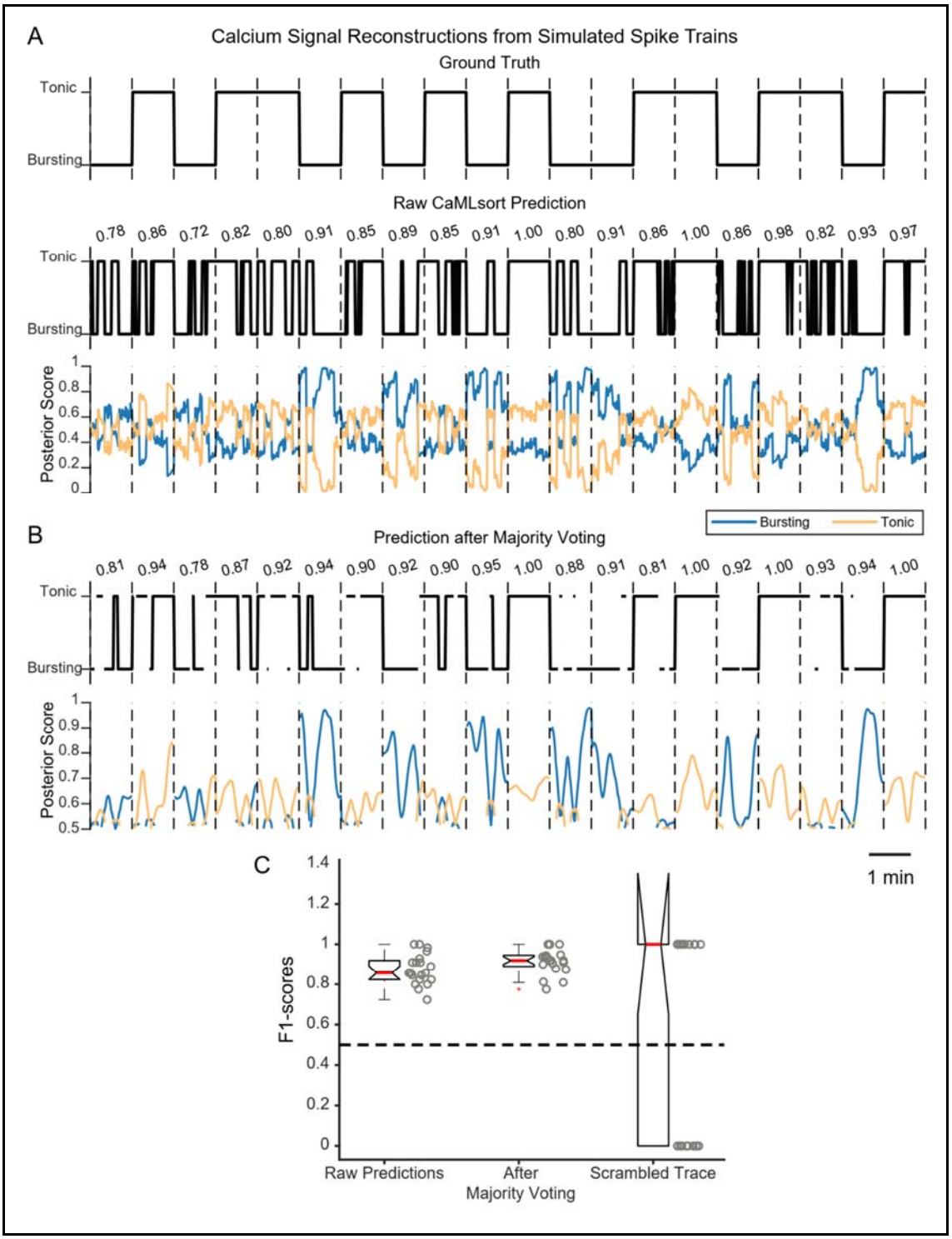

A
Calcium Signal Reconstructions from Simulated Spike Trains
Ground Truth
Tonic
Bursting
Raw CaMLsort Prediction
0.78 0.86 0.72 0.82 0.80 0.91 0.85 0.89 0.85 0.91 1.00 0.80 0.91 0.86 1.00 0.86 0.98 0.82 0.93 0.97
Tonic
Bursting
Posterior Score
1
0.8
0.6
0.4
0.2
0
Bursting
Tonic
B
Prediction after Majority Voting
0.81 0.94 0.78 0.87 0.92 0.94 0.90 0.92 0.90 0.95 1.00 0.88 0.91 0.81 1.00 0.92 1.00 0.93 0.94 1.00
Tonic
Bursting
Posterior Score
1
0.9
0.8
0.7
0.6
0.5
1 min
C
F1-scores
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Raw Predictions
After Majority Voting
Scrambled Trace

Figure 5.6 - CaMLsort predictions from calcium signal reconstructions from simulated spike trains. (A) Raw CaMLsort class label predictions from previously unseen calcium signal reconstructions that were generated from simulated spike trains (n=20 recordings). Predictions from CaMLsort (middle) are compared against the ground truth state (top) at each time step. Predictions for each recording were made independently, but have been stitched here for the purposes of representation. Each recording is separated by a vertical dashed line. The number above each prediction is the F1-score for that recording. The bottommost plot has posterior scores for each time step, with those for the "bursting" class in blue, and those for the "tonic" class in light orange. **(B)** CaMLsort predictions (top) that remain after majority voting with a window length of 7 steps has been used to retain only confident classifications (see Results and Methods for details). The posterior score (bottom) for each class at each time step, taken as the average class posterior score across all 7-step windows that included said time step. As in (B), the posterior scores for the "bursting" class are in blue, whereas those for the "tonic" class are in light orange, and the numbers above each prediction represent the F1-scores for that recording. **(C)** The distribution of F1-scores between the ground truth state and various predictions from CaMLsort - either the raw predictions (left), predictions obtained after majority voting (middle), or the raw predictions obtained by passing a scrambled version of the calcium signal as an input.

5.2.3 Neural networks trained only on simulated calcium signals correctly identify Purkinje neuron state from real data

Thus far, all challenges to the network were made using calcium signal reconstructions as the input. To test the applicability of these networks on data from experiments, we went back to our original simultaneous electrophysiology-imaging data and asked if the network could call the state of those cells correctly based on the dF/F traces alone.

We found that CaMLsort predicted the state of a cell from imaging data well (Fig 5.7A), with a median F1-score of 0.7654 across all cells (Fig 5.7C). Of the 15 cells in the dataset, 12 cells were predicted with an F1-score >0.6, while only 3 cells' states were predicted inaccurately. This outcome is interesting, because the networks have only been exposed to simulated calcium signals generated using a first-order generative model in the training and test phases, and yet are able to reasonably assess the state from real data, which have several challenging features - variable sensor expression levels, shot noise and signal bleaching, to name a few. Majority voting did not substantially improve the overall F1-score across all cells (median 0.75). However, by retaining only confident calls (at least 85% of the

duration in each recording), the flip-flopping of network predictions was markedly reduced (Fig 5.7B), as was observed before in the case of classifications using reconstructed calcium signals.

To benchmark the quality of these predictions, we obtained CaMLsort predictions from calcium signals reconstructed from the spike trains of these cells (since CaMLsort was trained only on simulated calcium signals). The prediction quality from reconstructions was better than what was obtained from true imaging data (median F1-score of 0.9011; Fig 5.7C). Therefore, while CaMLsort is clearly able to classify cellular state from experimentally-obtained imaging data, its performance has room for improvement. Also consistent with our previous results were the predictions from scrambled dF/F traces, which were almost all (96.83%) tonic, thereby yielding F1-scores that were mostly either 0 or 1 (Fig 5.7C).

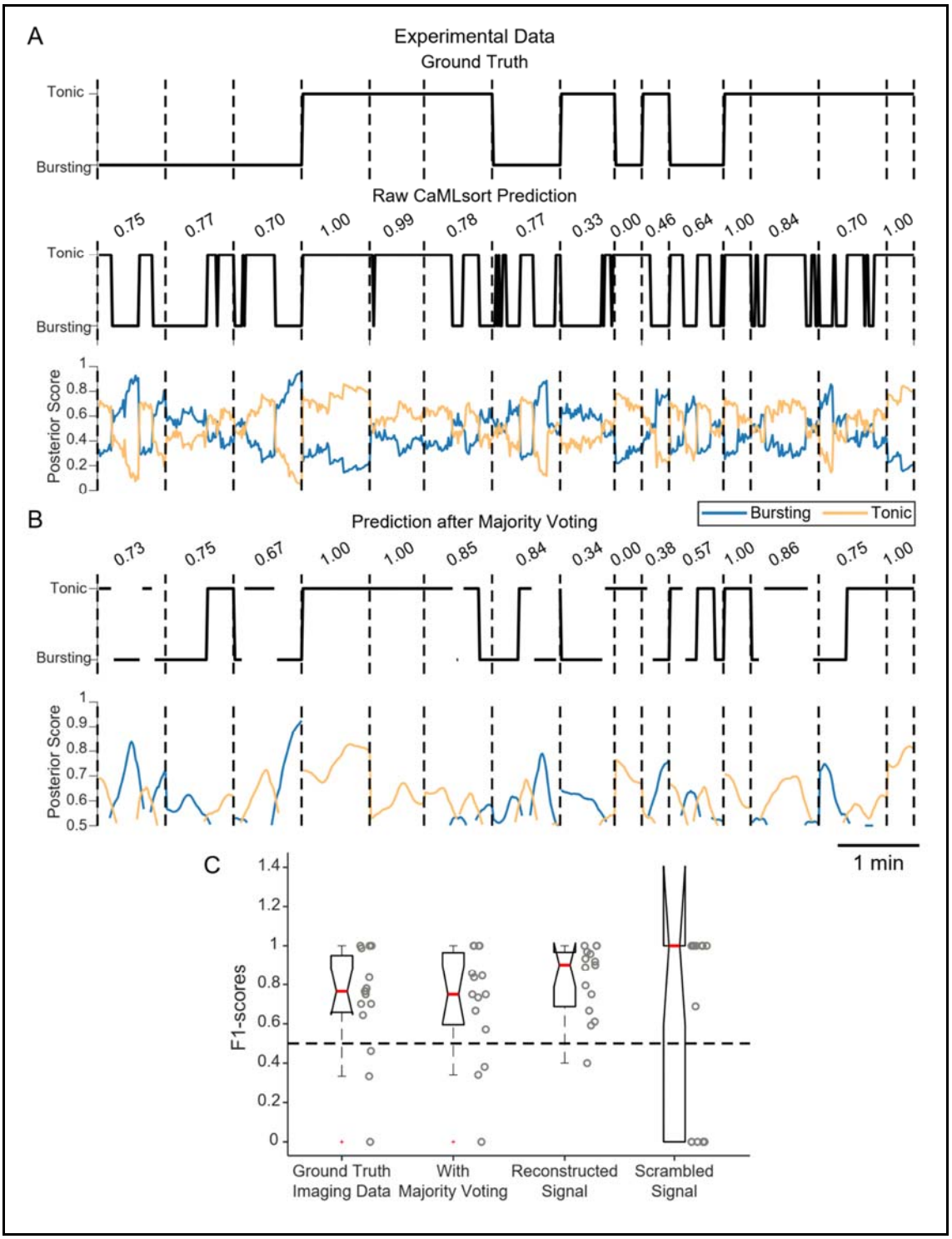

A
Experimental Data
Ground Truth
Tonic
Bursting
Raw CaMLsort Prediction
0.75 0.77 0.70 1.00 0.99 0.78 0.77 0.33 0.00 0.46 0.64 1.00 0.84 0.70 1.00
Tonic
Bursting
Posterior Score
1
0.8
0.6
0.4
0.2
0
Bursting
Tonic
B
Prediction after Majority Voting
0.73 0.75 0.67 1.00 1.00 0.85 0.84 0.34 0.00 0.38 0.57 1.00 0.86 0.75 1.00
Tonic
Bursting
Posterior Score
1
0.9
0.8
0.7
0.6
0.5
1 min
C
F1-scores
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Ground Truth Imaging Data
With Majority Voting
Reconstructed Signal
Scrambled Signal

Figure 5.7 - CaMLsort predictions on imaging data from zebrafish Purkinje neurons. **(A)** Raw CaMLsort predictions of cellular state inferred from experimental data acquired using widefield imaging (n=15; also see Chapter 4). Predictions from CaMLsort (middle) are compared against the ground truth state (top) at each time step. Each recording was treated independently, but predictions across cells have been stitched here for ease of visualisation. Each recording is separated by a vertical dashed line, and the number above each prediction is the F1-score for that recording. The bottommost plot has posterior scores for every time step, with those for the "bursting" class in blue, and those for the "tonic" class in light orange. **(B)** CaMLsort predictions (top) that remain after majority voting with a window length of 7 steps has been used to retain only confident classifications (see Results and Methods for details). The posterior score (bottom) for each class at each time step, taken as the average class posterior score across all 7-step windows that included said time step. As in (B), the posterior scores for the "bursting" class are in blue, whereas those for the "tonic" class are in light orange. **(C)** The distribution of F1-scores between the ground truth state and various predictions from CaMLsort - raw predictions from the imaging data (far left), predictions obtained after majority voting (second from left), raw predictions obtained from calcium signals reconstructed from spike times (third from left) and those predictions from imaging dF/F traces that were scrambled prior to prediction.

We wanted to ask whether there were any patterns in the incorrect predictions, so we examined CaMLsort predictions for each cell in detail by aligning them with the source event rasters and calcium signals, as shown in Fig 5.8. In general, it has been difficult to explain why network predictions were erroneous. Nevertheless, it is clear that "tonic" calls are more confident than "bursting" ones. Moreover, it appears that there is a tendency for the network to switch to "tonic" calls in between bursts (Fig 5.8A, C). However, with the knowledge that state transitions in Purkinje neurons are neither frequent nor of such short durations as predicted by the network, we think that an experimenter could overrule the brief transitions in (Fig 5.8A, C). In other words, the error rate of network predictions may be further reduced by human intervention.

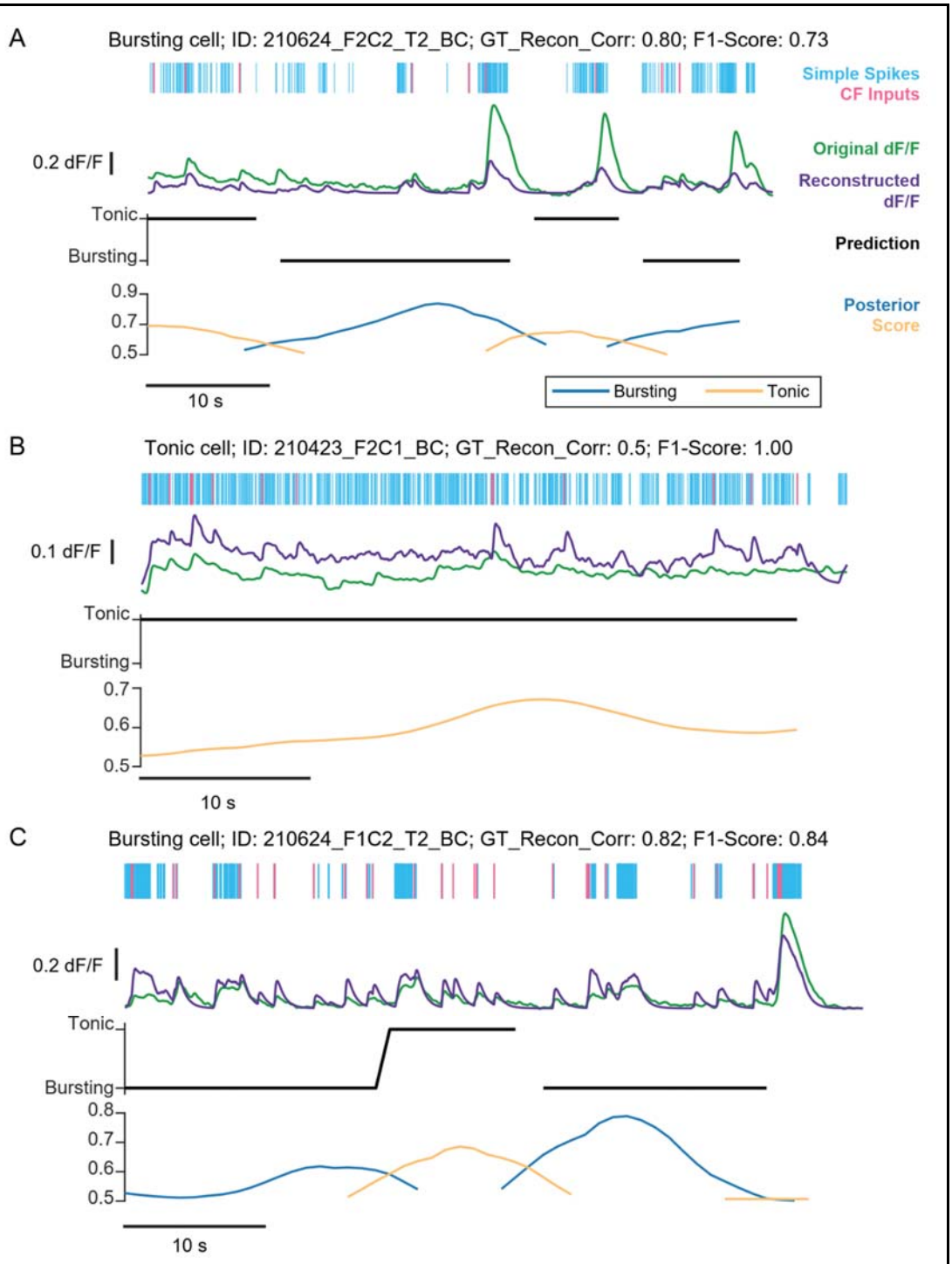


Figure 5.8 - Representative examples of CaMLsort state predictions aligned to the activity and calcium signals of their source cells. The electrophysiological activity of three representative Purkinje neurons **(A-C)** shown as rasters of their simple spikes (cyan) and CF inputs (magenta). Below each raster are the calcium signals of the same neuron, with the original dF/F signals obtained by imaging GCaMP5G-labelled neurons using

widefield microscopy in green, and the signal reconstructed from event timings in purple. The time-resolved state classification from CaMLsort is shown in the trace below in black, with average posterior scores for the bursting (blue) and tonic (light orange) classes at each time step shown in the lowest plot.

5.2.4 CaMLsort generalises well to calcium imaging data from other systems

Having developed a tool that can reasonably predict the bistable state of a Purkinje neuron given its calcium signal, we asked whether it could be broadly applicable to other neurons, too. I scanned the available literature and found a publicly available dataset of simultaneous electrophysiology and calcium imaging acquired from dopaminergic neurons of the ventral tegmental area (VTA) in mice (Fig 5.9A) (Fleming et al., 2021). These neurons also fire either tonically or phasically (Grace & Bunney, 1984b, 1984a). Imaging was performed using widefield fluorescence microscopy using either GCaMP7f or GCaMP7m. Thus, while the method of data acquisition was similar to ours, the sensors used were not.

I first compiled the dataset and sorted cells as tonic or bursting, using a combination of visual inspection of spike rasters (Fig 5.9E, F) and the coefficient of variation of inter-spike intervals (Fig 5.9B, C), as I did for Purkinje neurons. Three independent reviewers used this method to classify states, and a consensus was taken across reviewers, while also taking into consideration the state labels provided by the authors in their metadata. Data were included only if at least 2 out of the 4 labels (3 reviewers plus original metadata) agreed. Using this method, 78/84 cells (92.86%) of the original dataset were included for further analysis. Of these, 16 cells were in the Tonic state, while the remaining 62 cells were in the Bursting state. In order to test whether the event statistics of dopaminergic neurons of the VTA were similar to cerebellar Purkinje neurons in the corresponding state, I compared the cumulative distributions of ISIs of the two cell types with respect to their states (Fig 5.9D). From this, I could infer that interevent intervals in Purkinje neurons (Fig 5.9D, black lines) are generally shorter than in VTA neurons (Fig 5.9D, blue lines), with a much more pronounced difference in the tonic state (Fig 5.9D, solid lines).

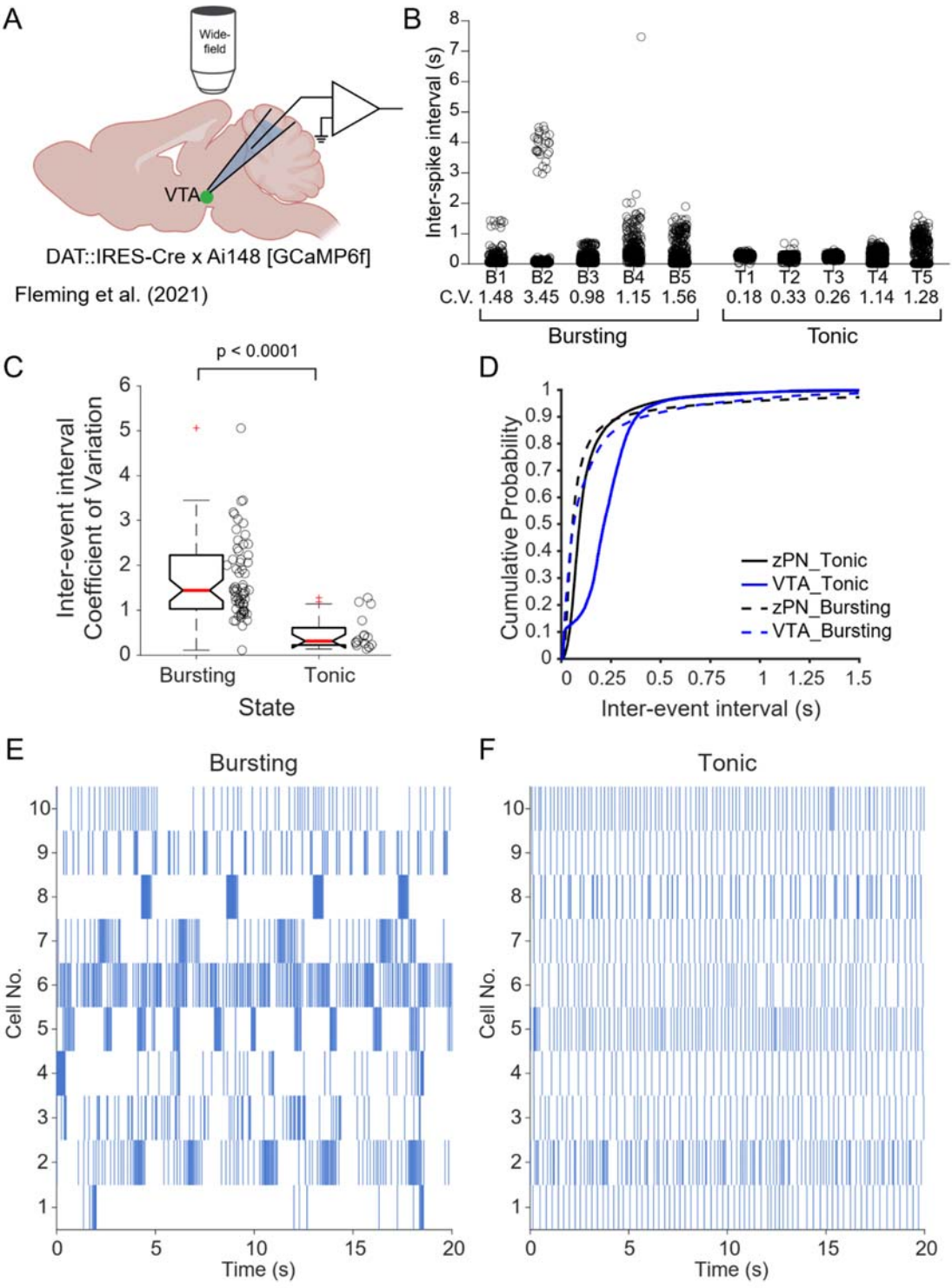


Figure 5.9 - Overview of publicly available simultaneous imaging and electrophysiology data from Dopaminergic (DA) neurons of the ventral tegmental area (VTA) in mice. (A) Schematic of the experimental paradigm used by Fleming et al. to acquire simultaneous widefield imaging and loose patch electrophysiology data from DA

neurons of the VTA in mice. Brain slices were obtained from adult transgenic mice resulting from a cross between the DAT::IRES-Cre and Ai-148 lines. VTA DA neurons expressed either GCaMP6f or GCaMP6m in these mice and could be targeted for simultaneous electrophysiology and imaging. **(B)** Individual interspike intervals from 10 neurons in this dataset, grouped by their firing mode (bursting or tonic). The coefficient of variation (C.V.) of the interspike interval distribution is indicated below each cell. **(C)** The distributions of C.V.s of interspike intervals for cells classified as either bursting (n=62) or tonic (n=16 cells). The p-value was calculated using a one-sided Mann-Whitney U-test. **(D)** Cumulative probability distributions of the inter-event intervals of simple spikes in zebrafish Purkinje neurons (black) and spikes in VTA DA neurons (blue). The distributions are grouped by cellular state - tonic (solid lines) or bursting (dashed lines). **(E-F)** Event rasters for 10 randomly selected VTA DA recordings in the bursting (E) or tonic (F) state.

Before trying to use CaMLsort on this data, I asked whether the convolution-based forward model of reconstructing calcium signals from electrophysiology would work on this dataset. Since there is only a single kind of spike in these neurons, I used only one kernel, the amplitude, half-rise time constant and half-decay time constant of which were taken to be the mean values for GCaMP6f or GCaMP5G (T.-W. Chen et al., 2013). No optimization of kernels was performed.

Visual inspection shows that reconstructions obtained from both GCaMP6f and GCaMP5G kernels were comparable to the signals imaged from VTA DA neurons, with the only difference between them being their amplitudes (Fig 5.10A, B). To quantify the similarity between the ground truth and the two simulated calcium signals, I used the Pearson correlation coefficient. The correlation coefficients between the ground truth and reconstructions from both GCaMP6f and GCaMP5G kernels were nearly identical, with a value of 0.6400 ± 0.3189, on average, for both (Fig 5.10C). Reconstructions that were either scrambled or generated after shuffling the DA neuron’s spike times had no correlation with the ground truth signal, as expected (Fig 5.10C).

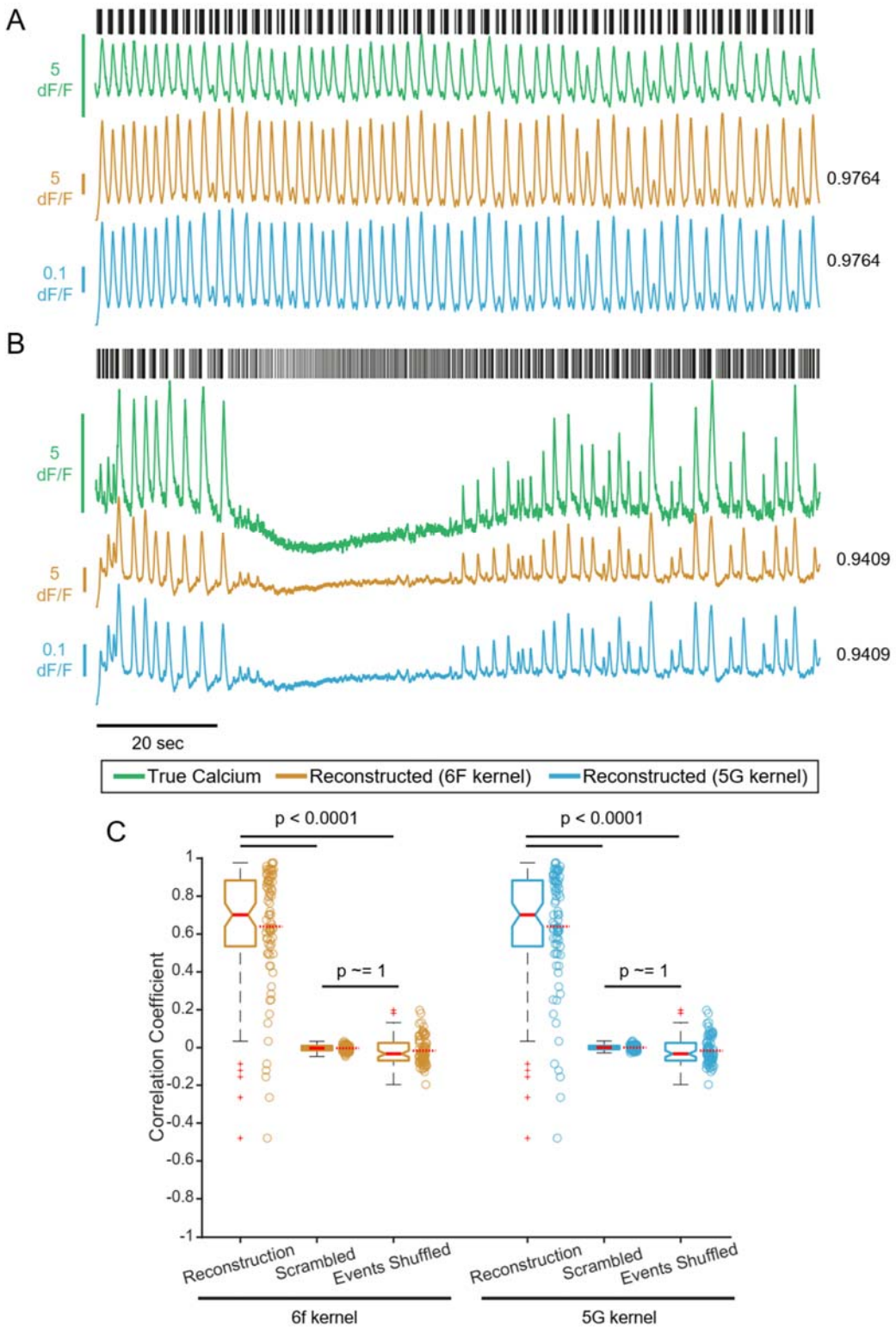


Figure 5.10 - Reconstructing calcium signals for VTA DA neurons using a convolution-based forward model. (A-B) Electrophysiological activity and associated

calcium signals from two representative VTA DA neurons. For each cell, spike timings are shown as rasters (black) and the raw simultaneously-acquired dF/F signals aligned below (green). Two calcium signal reconstructions were generated, using either a GCaMP6f kernel (gold) or a GCaMP5G kernel (blue). Numbers alongside each reconstruction indicate the Perason correlation coefficient between the ground truth and the corresponding calcium signal reconstruction. **(C)** Quantification of the reconstruction quality, as measured using the Pearson correlation coefficient between the ground truth and reconstructed signals ("Reconstruction"). For each cell (n=78), the correlation between ground truth and a scrambled version of the reconstructed signal ("Scrambled"), as well as the correlation between ground truth and a reconstruction generated after shuffling spike timings ("Events Shuffled") were measured. Individual data points are plotted alongside each boxplot. p-values were calculated using a Kruskal-Wallis test yielded a p-value of 4.6130×10^{-49}, and then using the post-hoc Dunn's test for individual comparisons, the latter of which are marked on the figure.

Having confirmed that the Fleming et al. dataset had both tonic and bursting data, and that the mapping from electrophysiology to imaging was also possible with this data, we were ready to test CaMLsort on it. As before, we first interpolated the calcium signals to get a sampling frequency of 30Hz before inputting them to CaMLsort. Raw predictions from CaMLsort and their corresponding ground truth labels for a subset of 12 of the 78 cells are shown in Fig 5.11A. It is apparent that predictions look similar to those obtained for Purkinje neurons. While the network appears to have a bias towards "tonic" class predictions, it is able to make accurate calls even for bursting neurons (Fig 5.11A, last cell, for example). The distribution of F1-scores obtained from all 78 cells, the median for which is 0.6841, is shown in Fig 5.11B.

We also calculated F1-scores for predictions made on calcium signals reconstructed from true electrophysiology data using the GCaMP5G kernel (Fig 5.11B). We decided to use the 5G kernel since the reconstructions that CaMLsort was trained on followed the kinetics of GCaMP5G. The predictions from signal reconstructions were much better than from raw imaging data, with a median of 0.8889. For comparison, predictions from calcium signals reconstructed after shuffling spike trains were poorer in quality, with a median of 0.5 (Fig 5.11B). Lastly, if the imaging signal was scrambled before prediction, CaMLsort predicted a default "tonic" state 89.03% of the time, yielding F1-scores that were mostly either 0 or 1 (Fig 5.11B).

In all, our results from VTA DA neurons in mice were similar to those from zebrafish Purkinje neurons, which means that CaMLsort generalises well and is likely to find broad use in classifying cellular state across systems. In order to facilitate this use, we will make CaMLsort publicly available as a GitHub package.

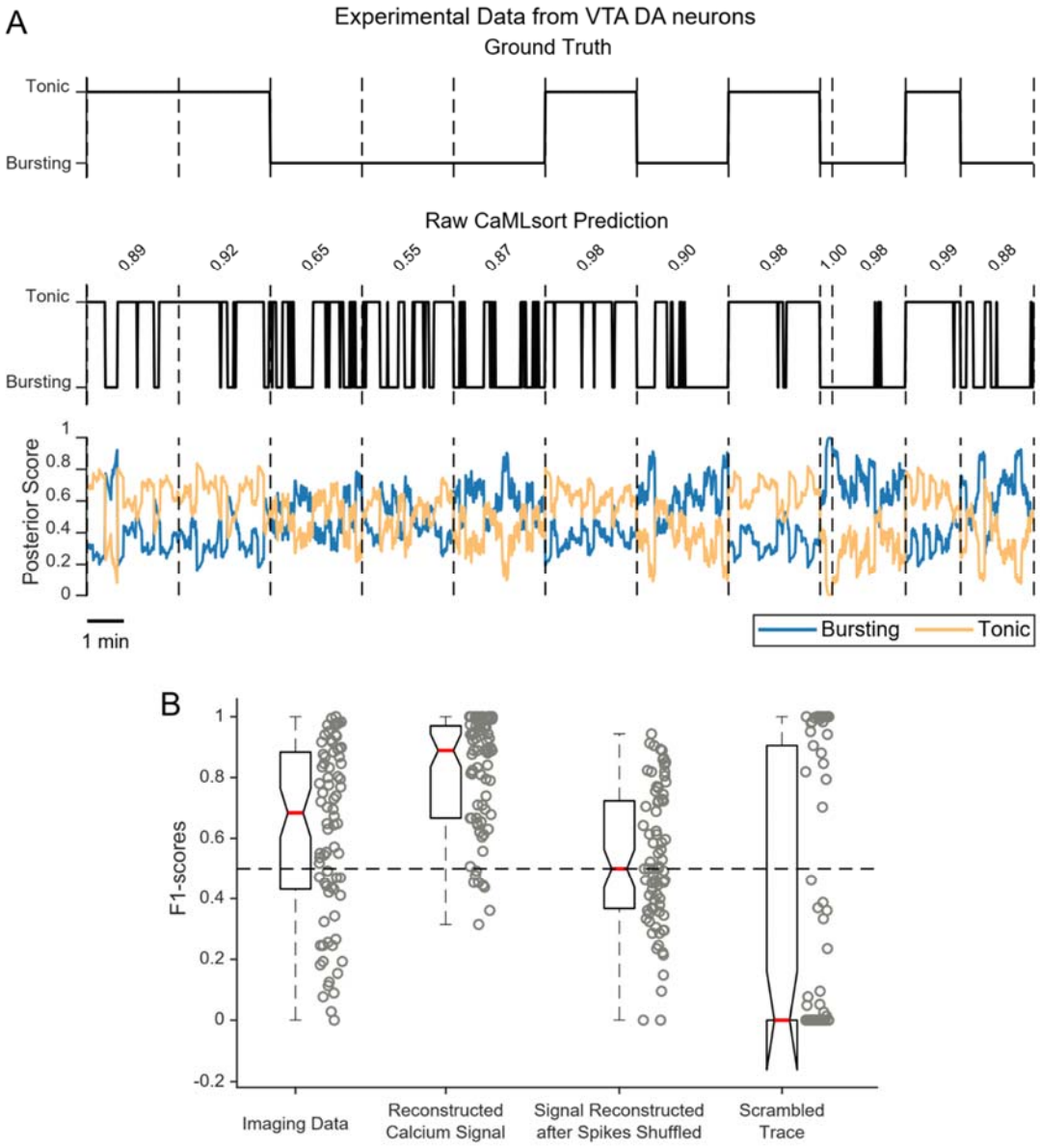


Figure 5.11 - CaMLsort predictions of cellular state from calcium signals of VTA DA neurons. (A) A subset of 12 (of 78) raw CaMLsort predictions from imaging data acquired from DA neurons of the VTA in mice. Each prediction was independently obtained but has been stitched together to produce a single time series with cells separated by vertical dashed lines. The ground truth state (top) is used as reference to assess the quality of CaMLsort's predictions (middle) using the F1-score, which is printed above the prediction for each cell. The posterior scores for each class - bursting (blue) and tonic (light orange) - are shown in the lowest plot. **(B)** The entire distribution of F1-scores (n=78) between the ground truth state and the binary classification from CaMLsort obtained from various inputs

- (*from left to right*) experimentally-acquired dF/F traces, calcium signals reconstructed from true spike trains using the GCaMP5G kernel, calcium signals reconstructed from shuffled spike trains using the GCaMP5G kernel, or a scrambled version of the raw dF/F data. Individual data points are shown alongside each boxplot. The horizontal dashed line corresponds to an F1-score of 0.5.

5.3 Discussion

5.3.1 Inferring long-term dynamics from calcium imaging data using CaMLsort

Calcium imaging, while a gold standard method in neuroscience, has thus far been used only to report activity or inactivity of neurons. This misses out many long-term dynamics of neurons, which can carry information. As the currently available calcium sensors themselves are much slower compared to the calcium transients, it is obvious that the recorded dF/F signals must manifest signatures of such long-term dynamics. CaMLsort addresses this issue by looking at longer term signatures in the calcium imaging time series data to identify cellular state.

The development of CaMLsort has opened up the possibility of a lot of future research in understanding PN bistability and its role in the function of the cerebellum as a whole. We could now experimentally address, for instance, what causes a switch in a PN's state by doing a high-throughput screen for chemical agents that affect bistability. We know that while bistability is a cell-intrinsic property, state switches could be a result of modulation of network activity. Neuromodulators are key candidates for this, as they can alter cell-intrinsic properties as well as network activity, and could thereby affect the bistability of a neuron, as has been demonstrated in multiple experimental systems (Abbinanti et al., 2012; Conway et al., 1988; Hounsgaard et al., 1988; Lechner et al., 1996; Williams et al., 2002). Another question we can now directly address is how the state of a cell changes over development. I hypothesise that the tonic state could be a developmentally early state, with the burst-like phenotype developing as they get integrated into the developing cerebellar circuit. This follows from the fact that PNs fire tonically in the presence of synaptic inhibitors. By birthdating PNs using a photoconvertible protein like Kaede and imaging their activity with GCaMP, we can use CaMLsort to ask whether younger cells tend to be more tonic or vice versa. While this question could potentially be addressed using electrophysiology, CaMLsort allows us to perform this experiment in a more high-throughput and minimally invasive fashion.

A number of cell types exhibit bistability including cortical neurons (Cowan and Wilson, 1994; Egorov et al., 2002), striatal neurons (Misgeld et al., 1986), thalamic neurons (Fuentealba et al., 2005) and spinal motoneurons (Hounsgaard et al., 1984). In all of these cases, cellular state dictates how synaptic inputs are integrated. Therefore, predicting cellular state is essential to understanding circuit computation and here CaMLsort will be immensely useful. It is natural to wonder whether CaMLsort will be equally effective in predicting cellular state in different cell types and in different model organisms. We have found that CaMLsort generalises well to a completely different system than it was trained on - neurons from a different brain region (ventral tegmental area) in a different model organism (mouse) imaged using a later generation of GCaMP (6th generation). We would caution against using CaMLsort right out of the box without validating its efficacy in the new system of interest using a small ground truth dataset.

Taken together, CaMLsort is a powerful tool that opens up new avenues of research into the functioning of any circuit with bistable neurons.

5.3.2 Limitations of CaMLsort

Having said the above, CaMLsort is not free of limitations. For instance, the generative model we use to train the neural network is a linear, time-invariant model and results may improve with the use of more accurate models, such as ones that estimate the true increase in intracellular calcium concentration resulting from simple spikes and CF inputs. Furthermore, it could be that CaMLsort may not work independent of the imaging method used. In both the cases we tested CaMLsort, widefield epifluorescence microscopy was used. The signal and noise characteristics of other, more modern and popular, imaging methods like multiphoton and light sheet microscopy are different from traditional widefield imaging. However, we posit that this should also be easy to resolve, either by the acquisition of a small ground truth dataset, or by comparing the traces obtained from imaging to those in the reconstructions dataset, or by denoising the data prior to prediction using CaMLsort.

5.4 Materials and Methods

5.4.1 Training and evaluating the CNN-LSTM model

The dataset of reconstructed calcium imaging time series was first split into 5 folds to perform cross validation (Fig 5.1). The experiments are repeated by considering each fold as a test set, with one of the folds being the validation set and remaining 3 folds to be training sets such that the data split for each experiment is as follows - 60% for training, 20% for

validation, 20% for testing. This splitting was done at the level of entire traces, such that the training, testing and cross-validation sets all had a balanced representation of traces from both the tonic and bursting classes. We chose to do a 5-fold cross-validation, i.e. the entire training dataset is split into 5 parts, each containing 20% of the data. Four of the folds are used for training and the remaining one is used for cross-validation.

The architecture of the network is shown in Fig 5.3. The input to the network is the entire trace sequence at 30Hz, normalised using min-max normalisation for every non-overlapping 10 second period, thus setting the minimum to 0, and maximum to 1. The first layer of processing is a CNN with kernel size 30 and stride size 30. The CNN downsamples the input trace from 30 Hz to 1Hz, and produces 4 vector outputs, which act as the input to the LSTM. The LSTM is bidirectional, so it looks at the sequence in both the forward and reverse order to produce an output. The LSTM has 24 hidden layers, so it produces 48 numbers per time step (Fig 5.3). The feature dimensions are collapsed by a Linear layer by taking their weighted average. This provides the classification logits at each position, which is passed through a softmax layer. The network is optimised using a cross entropy loss and the Adam optimizer (Kingma and Ba, 2017). The output of a trained neural network, then, is the probability of each timestep (at 1 Hz) belonging to the bursting class, from which the probability of it belonging to the tonic class can be derived. The class with the higher probability is assigned to each time step.

Since the model acts on the entire trace sequence, the amount of independent data for training the model is significantly reduced. To overcome this issue, we used a data augmentation strategy, which allows the network to learn features on arbitrary length sequences. For a sequence of length N, we determine a random segment size n, with n between [k, N], where k is the minimum sequence length. We randomly determine a starting position x, with a constraint that $x < N - n$. In our case, k=300
samples, and n varies in multiples of k (i.e. n = 300, 600, 900…). In each iteration of the network pass during the training phase, this random sequential segment was taken as the input trace to the model and gradient descent was performed to update weights. This was done for all the traces in the training set, so that the network saw entire traces at least once and in a non-overlapping fashion during the 500 epochs of training. The training is stopped before the full 500 epochs, based on early stopping criteria - if the validation loss does not improve after a few epochs. All models were trained using the PyTorch framework in Python 3 (Paszke et al., 2019).

For all evaluations - cross-validation, test and other independent validations, the network had to make predictions for the entire trace duration and not just the n-sample long chunks. Networks were evaluated using two metrics - classification accuracies and F1-scores. Classification accuracies are simply the fraction of predictions that were correct. F1-scores, on the other hand, are adjusted for class imbalances and account for the kind of errors made. By definition, an F1-score is the harmonic mean of the precision and recall metrics, where precision is the fraction of all predicted positives that were true positives, and recall is the fraction of all ground truth positives that were true positives.

To improve/smoothen predictions made by CaMLsort, we used a majority voting strategy. To do this, we re-sampled the time series using rolling window sampling with a window length N. For each time step of prediction, t_n, we found all the m windows (where $m \leq N$) that contained t_n (i.e. $[t_{n-N+1}, t_{n-N+2}, \ldots t_n, t_{n+1}]$, $[t_{n-N}, t_{n-N+1}, \ldots t_{n-1}, t_n]$, and so on until $[t_n, t_{n1}, \ldots t_{n+N-1}, t_{n+N}]$). For each of these windows, an average posterior score was calculated for each state, and the “average class” taken to be the state which had the higher score. The “majority class”, then, was the one which was the most frequent average class across all m windows. A “confident call” was made if and only if the average posterior scores for the majority class across all m windows was strictly higher than the average posterior scores for the minority class.

Chapter 6 - Conclusion

In this thesis, I have explored the intrinsic dynamics of Purkinje neurons in larval zebrafish, developing spike analysis pipelines and a machine learning tool to study membrane potential bistability. First, I showed that PN spike patterns are distinct between two states (Chapter 2). I also showed that the response of a cell to synaptic input was state-dependent, which means that the bistable state of a neuron has physiological relevance and should be further explored (Chapter 2). Next, I explored the relationship between pairs of neurons as a function of their state and found that pairs of bursting neurons have more similar activity than other state combinations (Chapter 3). This further lends support to the idea that neuronal state has implications for population coding. I then demonstrated that calcium imaging data can provide more information than just the presence or absence of spiking activity in PNs, and that it can be leveraged to provide information about more long-term dynamics like bistability (Chapter 4-5). I found that a limited amount of ground truth data can be used to develop generative models to enlarge datasets for use in machine learning, without losing out on its ability to generalise (Chapter 4-5). Finally, I show that tools developed for one kind of neuron (zebrafish Purkinje neurons) can generalise well to other cases, too, even when using newer versions of calcium sensors (Chapter 5).

Among the many systems in which neuronal bistability has been studied, Purkinje neurons of the cerebellum are particularly interesting because they are spontaneously active. Thus, they switch between firing either tonically or in bursts (Sengupta & Thirumalai, 2015), and such behaviour is not only unique to Purkinje neurons but has also been reported in many other cell types such as thalamic (Jahnsen & Llinás, 1984) and cortical neurons (McCormick et al., 1985). Though it is appreciated that bistability increases the information processing capacity of neurons, such as by allowing gating of synaptic inputs (Thirumalai & Jha, 2022), its implication for network computation or behavioural output is less well understood. This is true of Purkinje neurons as well - while it is established that these neurons are bistable and are important for motor behaviours (Engbers et al., 2013; Popa et al., 2019; Sengupta & Thirumalai, 2015), the role of bistability in Purkinje neuron computation or motor behaviour output has thus far not been clear.

I propose that having more than one stable state allows a cell to toggle in and out of participating in circuit computations. In the cerebellum, specifically, bistability may have significant consequences by modulating PN synchrony with other PNs. PNs make weak inhibitory synapses onto deep cerebellar nuclear (DCN) cells in mammals and the

eurydendroid cells in fish, with a high degree of convergence (Bae et al., 2009; De Zeeuw et al., 2011; Harmon et al., 2020; Ikenaga et al., 2006; Person & Raman, 2012). DCN cell spike times could be effectively controlled via inhibition from several PNs firing synchronously (Person & Raman, 2012). Since PNs fire tonically when synaptically-isolated and CF inputs can trigger bursts (Sengupta & Thirumalai, 2015), the bursting state is a network-embedded state while the tonic state could be thought of as an isolated state. By toggling between these two states PNs can increase or decrease the levels of population synchrony and thereby affect DCN firing differentially. Mapping state distribution across the PN population during various tasks will lead us to an understanding of how PNs synchronise or desynchronize during these tasks and how that in turn shapes cerebellar output and motor behaviours.

One obvious implication of bistability is that it allows for a large number of possible state combinations in a neuronal population (specifically 2^N, where N is the number of neurons). For an N of just 300 (a reasonable estimate for the number of PNs in larval zebrafish), this number is unimaginably large - larger than estimates of the total number of atoms in the observable universe! It is thus reasonable to imagine that only a small subset of these state combinations is accessed. Perhaps the solution to this problem of impossibly large combinations of states is a restricted spatial distribution of states in the population of Purkinje neurons *in vivo*. Both these can be addressed with the advent of CaMLsort.

Further, during development, the cerebellum undergoes drastic changes - new PNs are added (Hamling et al., 2015), PN dendritic arbours elaborate within the molecular layer, CFs undergo synapse elimination (Kano et al., 2018), and neuromodulatory fibers innervate the cerebellum (McLean & Fetcho, 2004). A developmental study of PN state distribution will allow us to ask whether in the face of such drastic circuit changes, PN state distributions are maintained or altered. It is now possible to investigate how these state changes affect PN activity during well-established motor behaviour paradigms**.** Thus, the immediate future goal would be to use CaMLsort to characterise the distribution of PN states in the developing cerebellum, both passively and during behaviour.

Voltage imaging is an exciting and promising technique that is considered to be the future of imaging-based reporting of neural activity. Unlike traditional calcium imaging, which measures changes in intracellular calcium levels, voltage imaging directly detects alterations in membrane potential, providing a more direct readout of neuronal electrical activity. This advantage makes voltage imaging a highly sought-after tool for understanding the dynamics of neural circuits with exceptional precision and sensitivity.

However, despite its immense potential, voltage imaging is not without its limitations. One of the most significant challenges is the need for bright, fast, and stable fluorescent indicators to accurately capture membrane potential changes. Developing and optimising these indicators remains a complex and ongoing task. Moreover, the phototoxicity associated with voltage imaging can lead to considerable damage to delicate neural tissues over prolonged imaging sessions, limiting its application in certain experimental settings.

On the other hand, calcium imaging, with its established and extensive usage in neuroscience research, continues to hold a significant place in the field. Over the years, calcium imaging has provided invaluable insights into neural activity patterns, population coding, and the dynamics of neural circuits. The wealth of existing calcium imaging data across various brain regions and species is a valuable resource for scientific investigations.

Although voltage imaging promises high-resolution insights into neural activity, the abundance of historical calcium imaging data cannot be overlooked. Researchers can leverage this vast pool of information to uncover new findings and revisit previous experiments using advanced analysis tools. Tools like CaMLsort have become essential for researchers to delve deeper into calcium imaging datasets, extract more nuanced information, and make new discoveries from data that have already been acquired.

In conclusion, voltage imaging indeed represents the future of imaging-based reporting of neural activity, offering unprecedented insights into the fine details of neural computations. Nevertheless, calcium imaging will continue to play a significant role in neuroscience research due to its well-established presence and extensive data availability. The inherent limitations of voltage imaging and the wealth of existing calcium imaging data necessitate the development and utilisation of sophisticated analysis tools like CaMLsort, allowing researchers to unlock the full potential of both imaging techniques and drive the progress of neuroscience into uncharted territories.

While the focus of this thesis was on larval zebrafish Purkinje neurons, it is worth exploring the generalisability of these findings to other neuron types. The demonstration that the tools developed for one type of neuron can be effectively applied to other cases, even when using different calcium sensors, suggests exciting possibilities for cross-species and cross-neuron comparisons. This could provide a more comprehensive understanding of bistability's role in neural computations across different brain circuits.

Finally, the exploration of the relationship between pairs of neurons in different states also paves the way for investigating functional connectivity within bistable networks. Understanding how the state-dependent responses of neurons influence their interactions

and information flow within the circuit could offer valuable insights into the emergent properties of neural networks and their role in shaping complex behaviours.

In conclusion, the thesis work on Purkinje neurons in larval zebrafish has profound implications for the study of neural dynamics and computation. By unravelling the intricacies of bistability and its effects on population coding, network computations, and motor behaviour, this research sets the stage for exciting future directions in neuroscience. From examining the impact of bistability during development to exploring its role in different brain regions and across species, the possibilities for further research are vast and promising.

References

Ahrens, M. B., Li, J. M., Orger, M. B., Robson, D. N., Schier, A. F., Engert, F., & Portugues, R. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. Nature, 485(7399), Article 7399. https://doi.org/10.1038/nature11057

Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M., & Keller, P. J. (2013). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nature Methods, 10(5), Article 5. https://doi.org/10.1038/nmeth.2434

Akerboom, J., Chen, T.-W., Wardill, T. J., Tian, L., Marvin, J. S., Mutlu, S., Calderón, N. C., Esposti, F., Borghuis, B. G., Sun, X. R., Gordus, A., Orger, M. B., Portugues, R., Engert, F., Macklin, J. J., Filosa, A., Aggarwal, A., Kerr, R. A., Takagi, R., … Looger, L. L. (2012). Optimization of a GCaMP Calcium Indicator for Neural Activity Imaging. Journal of Neuroscience, 32(40), 13819–13840. https://doi.org/10.1523/JNEUROSCI.2601-12.2012

Alcamí, P., & Pereda, A. E. (2019). Beyond plasticity: The dynamic impact of electrical synapses on neural circuits. Nature Reviews Neuroscience, 20(5), Article 5. https://doi.org/10.1038/s41583-019-0133-5

Alviña, K., & Sawtell, N. B. (2014). Sensory processing and corollary discharge effects in posterior caudal lobe Purkinje cells in a weakly electric mormyrid fish. Journal of Neurophysiology, 112(2), 328–339. https://doi.org/10.1152/jn.00016.2014

Andersen, B. B., Gundersen, H. J. G., & Pakkenberg, B. (2003). Aging of the human cerebellum: A stereological study. The Journal of Comparative Neurology, 466(3), 356–365. https://doi.org/10.1002/cne.10884

Apps, R., & Garwicz, M. (2005). Anatomical and physiological foundations of cerebellar information processing. Nature Reviews Neuroscience, 6(4), Article 4. https://doi.org/10.1038/nrn1646

Bae, Y.-K., Kani, S., Shimizu, T., Tanabe, K., Nojima, H., Kimura, Y., Higashijima, S., & Hibi, M. (2009). Anatomy of zebrafish cerebellum and screen for mutations affecting its development. Developmental Biology, 330(2), 406–426. https://doi.org/10.1016/j.ydbio.2009.04.013

Bagnall, M. W., Zingg, B., Sakatos, A., Moghadam, S. H., Zeilhofer, H. U., & du Lac, S. (2009). Glycinergic projection neurons of the cerebellum. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 29(32), 10104–10110. https://doi.org/10.1523/JNEUROSCI.2087-09.2009

Barnes-Davies, M., Barker, M. C., Osmani, F., & Forsythe, I. D. (2004). Kv1 currents mediate a gradient of principal neuron excitability across the tonotopic axis in the rat lateral superior olive. European Journal of Neuroscience, 19(2), 325–333.

https://doi.org/10.1111/j.0953-816X.2003.03133.x

Bell, C. C. (2002). Evolution of cerebellum-like structures. Brain, Behavior and Evolution, 59(5–6), 312–326. https://doi.org/10.1159/000063567

Bell, C. C., Han, V., & Sawtell, N. B. (2008). Cerebellum-Like Structures and Their Implications for Cerebellar Function. Annual Review of Neuroscience, 31(1), 1–24. https://doi.org/10.1146/annurev.neuro.30.051606.094225

Bell, C. C., & Kawasaki, T. (1972). Relations among climbing fiber responses of nearby Purkinje Cells. Journal of Neurophysiology, 35(2), 155–169. https://doi.org/10.1152/jn.1972.35.2.155

Berens, P., Freeman, J., Deneux, T., Chenkov, N., McColgan, T., Speiser, A., Macke, J. H., Turaga, S. C., Mineault, P., Rupprecht, P., Gerhard, S., Friedrich, R. W., Friedrich, J., Paninski, L., Pachitariu, M., Harris, K. D., Bolte, B., Machado, T. A., Ringach, D., … Bethge, M. (2018). Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. PLOS Computational Biology, 14(5), e1006157. https://doi.org/10.1371/journal.pcbi.1006157

Blinks, Prendergast, F. G., & Allen, D. G. (1976). Photoproteins as biological calcium indicators. Pharmacological Reviews, 28(1), 1–93.

Bloedel, J. R., & Roberts, W. J. (1971). Action of climbing fibers in cerebellar cortex of the cat. Journal of Neurophysiology, 34(1), 17–31. https://doi.org/10.1152/jn.1971.34.1.17

Bormann, C. L., Kanakasabapathy, M. K., Thirumalaraju, P., Gupta, R., Pooniwala, R., Kandula, H., Hariton, E., Souter, I., Dimitriadis, I., Ramirez, L. B., Curchoe, C. L., Swain, J., Boehnlein, L. M., & Shafiee, H. (2020). Performance of a deep learning based neural network in the selection of human blastocysts for implantation. ELife, 9, e55301. https://doi.org/10.7554/eLife.55301

Borst, J. G., & Helmchen, F. (1998). Calcium influx during an action potential. Methods in Enzymology, 293, 352–371. https://doi.org/10.1016/s0076-6879(98)93023-3

Brew, H. M., & Forsythe, I. D. (1995). Two voltage-dependent K+ conductances with complementary functions in postsynaptic integration at a central auditory synapse. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 15(12), 8011–8022. https://doi.org/10.1523/JNEUROSCI.15-12-08011.1995

Brewster, D. L., & Ali, D. W. (2013). Expression of the voltage-gated potassium channel subunit Kv1.1 in embryonic zebrafish Mauthner cells. Neuroscience Letters, 539, 54–59. https://doi.org/10.1016/j.neulet.2013.01.042

Brochu, G., Maler, L., & Hawkes, R. (1990). Zebrin II: A polypeptide antigen expressed selectively by purkinje cells reveals compartments in rat and fish cerebellum. Journal of Comparative Neurology, 291(4), 538–552. https://doi.org/10.1002/cne.902910405

Burgess, H. A., & Granato, M. (2007). Sensorimotor Gating in Larval Zebrafish. Journal of

Neuroscience, 27(18), 4984–4994. https://doi.org/10.1523/JNEUROSCI.0615-07.2007

Chen, S., Augustine, G. J., & Chadderton, P. (2016). The cerebellum linearly encodes whisker position during voluntary movement. ELife, 5, e10509. https://doi.org/10.7554/eLife.10509

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature, 499(7458), Article 7458. https://doi.org/10.1038/nature12354

Chung, B. P., & Edwards, D. H. (2019). Discrimination of bursts and tonic activity in multifunctional sensorimotor neural network using the extended hill-valley method. Journal of Neurophysiology, 122(3), 1073–1083. https://doi.org/10.1152/jn.00206.2018

Ciresan, D. C., Meier, U., Gambardella, L. M., & Schmidhuber, J. (2010). Deep Big Simple Neural Nets Excel on Handwritten Digit Recognition. Neural Computation, 22(12), 3207–3220. https://doi.org/10.1162/NECO_a_00052

Dana, H., Mohar, B., Sun, Y., Narayan, S., Gordus, A., Hasseman, J. P., Tsegaye, G., Holt, G. T., Hu, A., Walpita, D., Patel, R., Macklin, J. J., Bargmann, C. I., Ahrens, M. B., Schreiter, E. R., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2016). Sensitive red protein calcium indicators for imaging neural activity. ELife, 5, e12727. https://doi.org/10.7554/eLife.12727

Dana, H., Sun, Y., Mohar, B., Hulse, B. K., Kerlin, A. M., Hasseman, J. P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., Macklin, J. J., Chen, Y., Konnerth, A., Jayaraman, V., Looger, L. L., Schreiter, E. R., Svoboda, K., & Kim, D. S. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. Nature Methods, 16(7), Article 7. https://doi.org/10.1038/s41592-019-0435-6

Dayan, P., & Abbott, L. F. (2005). Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. MIT Press.

De Zeeuw, C. I., Hoebeek, F. E., Bosman, L. W. J., Schonewille, M., Witter, L., & Koekkoek, S. K. (2011). Spatiotemporal firing patterns in the cerebellum. Nature Reviews Neuroscience, 12(6), Article 6. https://doi.org/10.1038/nrn3011

Deneux, T., Kaszas, A., Szalay, G., Katona, G., Lakner, T., Grinvald, A., Rózsa, B., & Vanzetta, I. (2016). Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. Nature Communications, 7(1), Article 1. https://doi.org/10.1038/ncomms12190

Denk, W., Delaney, K. R., Gelperin, A., Kleinfeld, D., Strowbridge, B. W., Tank, D. W., & Yuste, R. (1994). Anatomical and functional imaging of neurons using 2-photon laser scanning microscopy. Journal of Neuroscience Methods, 54(2), 151–162. https://doi.org/10.1016/0165-0270(94)90189-9

Denk, W., Strickler, J. H., & Webb, W. W. (1990). Two-photon laser scanning fluorescence microscopy. Science (New York, N.Y.), 248(4951), 73–76. https://doi.org/10.1126/science.2321027

Dickinson, P. S., & Nagy, F. (1983). Control of A Central Pattern Generator by an Identified Modulatory Interneurone in Crustacea: II. Induction and Modification of Plateau Properties in Pyloric Neurones. Journal of Experimental Biology, 105(1), 59–82. https://doi.org/10.1242/jeb.105.1.59

Ding, X., & Zhang, B. (2021). DeepBAR: A Fast and Exact Method for Binding Free Energy Computation. The Journal of Physical Chemistry Letters, 12(10), 2509–2515. https://doi.org/10.1021/acs.jpclett.1c00189

Dodson, P. D., Barker, M. C., & Forsythe, I. D. (2002). Two Heteromeric Kv1 Potassium Channels Differentially Regulate Action Potential Firing. Journal of Neuroscience, 22(16), 6953–6961. https://doi.org/10.1523/JNEUROSCI.22-16-06953.2002

Eaton, R. C., & Nissanov, J. (1985). A review of Mauthner-initiated escape behavior and its possible role in hatching in the immature zebrafish, Brachydanio rerio. Environmental Biology of Fishes, 12(4), 265–279. https://doi.org/10.1007/BF00005457

Eccles, J. C., Ito, M., & Szentágothai, J. (1967). The Cerebellum as a Neuronal Machine. Springer. https://doi.org/10.1007/978-3-662-13147-3

Eccles, J. C., Llinás, R., & Sasaki, K. (1966). The excitatory synaptic action of climbing fibres on the Purkinje cells of the cerebellum. The Journal of Physiology, 182(2), 268–296. https://doi.org/10.1113/jphysiol.1966.sp007824

Edgley, S. A., & Lidierth, M. (1988). Step-related discharges of Purkinje cells in the paravermal cortex of the cerebellar anterior lobe in the cat. The Journal of Physiology, 401(1), 399–415. https://doi.org/10.1113/jphysiol.1988.sp017169

Engbers, J. D. T., Fernandez, F. R., & Turner, R. W. (2013). Bistability in Purkinje neurons: Ups and downs in cerebellar research. Neural Networks, 47, 18–31. https://doi.org/10.1016/j.neunet.2012.09.006

Evans, M. H., Petersen, R. S., & Humphries, M. D. (2020). On the use of calcium deconvolution algorithms in practical contexts (p. 871137). bioRxiv. https://doi.org/10.1101/871137

Fetcho, J. R. (1991). Spinal network of the Mauthner cell. Brain, Behavior and Evolution, 37(5), 298–316. https://doi.org/10.1159/000114367

Fleming, W., Jewell, S., Engelhard, B., Witten, D. M., & Witten, I. B. (2021). Inferring spikes from calcium imaging in dopamine neurons. PLOS ONE, 16(6), e0252345. https://doi.org/10.1371/journal.pone.0252345

Friedrich, J., Zhou, P., & Paninski, L. (2017). Fast online deconvolution of calcium imaging data. PLOS Computational Biology, 13(3), e1005423.

https://doi.org/10.1371/journal.pcbi.1005423

Fuentealba, P., Timofeev, I., Bazhenov, M., Sejnowski, T. J., & Steriade, M. (2005). Membrane Bistability in Thalamic Reticular Neurons During Spindle Oscillations. Journal of Neurophysiology, 93(1), 294–304. https://doi.org/10.1152/jn.00552.2004

Gerstein, G. L., & Mandelbrot, B. (1964). Random Walk Models for the Spike Activity of a Single Neuron. Biophysical Journal, 4(1, Part 1), 41–68. https://doi.org/10.1016/S0006-3495(64)86768-0

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods, 6(5), Article 5. https://doi.org/10.1038/nmeth.1318

Gilbert, P. F., & Thach, W. T. (1977). Purkinje cell activity during motor learning. Brain Research, 128(2), 309–328. https://doi.org/10.1016/0006-8993(77)90997-0

Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., Taxidis, J., Najafi, F., Gauthier, J. L., Zhou, P., Khakh, B. S., Tank, D. W., Chklovskii, D. B., & Pnevmatikakis, E. A. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. ELife, 8, e38173. https://doi.org/10.7554/eLife.38173

Göbel, W., & Helmchen, F. (2007). In Vivo Calcium Imaging of Neural Network Function. Physiology, 22(6), 358–365. https://doi.org/10.1152/physiol.00032.2007

Grace, A. A., & Bunney, B. S. (1984a). The control of firing pattern in nigral dopamine neurons: Burst firing. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 4(11), 2877–2890. https://doi.org/10.1523/JNEUROSCI.04-11-02877.1984

Grace, A. A., & Bunney, B. S. (1984b). The control of firing pattern in nigral dopamine neurons: Single spike firing. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 4(11), 2866–2876. https://doi.org/10.1523/JNEUROSCI.04-11-02866.1984

Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. Nature Reviews Molecular Cell Biology, 23(1), Article 1. https://doi.org/10.1038/s41580-021-00407-0

Grienberger, C., & Konnerth, A. (2012). Imaging Calcium in Neurons. Neuron, 73(5), 862–885. https://doi.org/10.1016/j.neuron.2012.02.011

Gruol, D. L., & Franklin, C. L. (1987). Morphological and physiological differentiation of Purkinje neurons in cultures of rat cerebellum. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 7(5), 1271–1293. https://doi.org/10.1523/JNEUROSCI.07-05-01271.1987

Grynkiewicz, G., Poenie, M., & Tsien, R. Y. (1985). A new generation of Ca2+ indicators with greatly improved fluorescence properties. Journal of Biological Chemistry, 260(6), 3440–

3450. https://doi.org/10.1016/S0021-9258(19)83641-4

Hamling, K. R., Tobias, Z. J. C., & Weissman, T. A. (2015). Mapping the development of cerebellar Purkinje cells in zebrafish. Developmental Neurobiology, 75(11), 1174–1188. https://doi.org/10.1002/dneu.22275

Harmon, T. C., Magaram, U., McLean, D. L., & Raman, I. M. (2017). Distinct responses of Purkinje neurons and roles of simple spikes during associative motor learning in larval zebrafish. ELife, 6, e22537. https://doi.org/10.7554/eLife.22537

Harmon, T. C., McLean, D. L., & Raman, I. M. (2020). Integration of Swimming-Related Synaptic Excitation and Inhibition by olig2+ Eurydendroid Neurons in Larval Zebrafish Cerebellum. Journal of Neuroscience, 40(15), 3063–3074. https://doi.org/10.1523/JNEUROSCI.2322-19.2020

Hashimoto, K., Ichikawa, R., Takechi, H., Inoue, Y., Aiba, A., Sakimura, K., Mishina, M., Hashikawa, T., Konnerth, A., Watanabe, M., & Kano, M. (2001). Roles of glutamate receptor delta 2 subunit (GluRdelta 2) and metabotropic glutamate receptor subtype 1 (mGluR1) in climbing fiber synapse elimination during postnatal cerebellar development. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 21(24), 9701–9712. https://doi.org/10.1523/JNEUROSCI.21-24-09701.2001

Häusser, M., & Clark, B. A. (1997). Tonic Synaptic Inhibition Modulates Neuronal Output Pattern and Spatiotemporal Synaptic Integration. Neuron, 19(3), 665–678. https://doi.org/10.1016/S0896-6273(00)80379-7

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. https://doi.org/10.1109/CVPR.2016.90

Herzfeld, D. J., Kojima, Y., Soetedjo, R., & Shadmehr, R. (2018). Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. Nature Neuroscience, 21(5), Article 5. https://doi.org/10.1038/s41593-018-0136-y

Heyward, P., Ennis, M., Keller, A., & Shipley, M. T. (2001). Membrane Bistability in Olfactory Bulb Mitral Cells. Journal of Neuroscience, 21(14), 5311–5320. https://doi.org/10.1523/JNEUROSCI.21-14-05311.2001

Hibi, M., & Shimizu, T. (2012). Development of the cerebellum and cerebellar neural circuits. Developmental Neurobiology, 72(3), 282–301. https://doi.org/10.1002/dneu.20875

Hines, M. L., & Carnevale, N. T. (2001). NEURON: A tool for neuroscientists. The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry, 7(2), 123–135. https://doi.org/10.1177/107385840100700207

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hong, S., Negrello, M., Junker, M., Smilgin, A., Thier, P., & De Schutter, E. (2016).

Multiplexed coding by cerebellar Purkinje neurons. ELife, 5, e13810. https://doi.org/10.7554/eLife.13810

Horn, K. M., Deep, A., & Gibson, A. R. (2013). Progressive limb ataxia following inferior olive lesions. The Journal of Physiology, 591(22), 5475–5489. https://doi.org/10.1113/jphysiol.2012.234898

Houghton, C. (2015). Calculating mutual information for spike trains and other data with distances but no coordinates. Royal Society Open Science, 2(5), 140391. https://doi.org/10.1098/rsos.140391

Houghton, C. (2019). Calculating the Mutual Information between Two Spike Trains. Neural Computation, 31(2), 330–343. https://doi.org/10.1162/neco_a_01155

Hounsgaard, J., Hultborn, H., Jespersen, B., & Kiehn, O. (1984). Intrinsic membrane properties causing a bistable behaviour of α-motoneurones. Experimental Brain Research, 55(2), 391–394. https://doi.org/10.1007/BF00237290

Hounsgaard, J., Hultborn, H., Jespersen, B., & Kiehn, O. (1988). Bistability of alpha-motoneurones in the decerebrate cat and in the acute spinal cat after intravenous 5-hydroxytryptophan. The Journal of Physiology, 405, 345–367. https://doi.org/10.1113/jphysiol.1988.sp017336

Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., & Stelzer, E. H. K. (2004). Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy. Science, 305(5686), 1007–1009. https://doi.org/10.1126/science.1100035

Ikenaga, T., Yoshida, M., & Uematsu, K. (2006). Cerebellar efferent neurons in teleost fish. The Cerebellum, 5(4), 268–274. https://doi.org/10.1080/14734220600930588

Jahnsen, H., & Llinás, R. (1984). Electrophysiological properties of guinea-pig thalamic neurones: An in vitro study. The Journal of Physiology, 349, 205–226.

Jay, M., De Faveri, F., & McDearmid, J. R. (2015). Firing Dynamics and Modulatory Actions of Supraspinal Dopaminergic Neurons during Zebrafish Locomotor Behavior. Current Biology, 25(4), 435–444. https://doi.org/10.1016/j.cub.2014.12.033

Johnson, B. R., Peck, J. H., & Harris-Warrick, R. M. (1993). Dopamine induces sign reversal at mixed chemical-electrical synapses. Brain Research, 625(1), 159–164. https://doi.org/10.1016/0006-8993(93)90149-H

Judd, E. N., Lewis, S. M., & Person, A. L. (2021). Diverse inhibitory projections from the cerebellar interposed nucleus. ELife, 10, e66231. https://doi.org/10.7554/eLife.66231

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), Article 7873. https://doi.org/10.1038/s41586-021-03819-2

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydın, Ç., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., … Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. Nature, 551(7679), Article 7679. https://doi.org/10.1038/nature24636

Kandel, E. R. (2013). Principles of Neural Science, Fifth Edition. McGraw Hill Professional.

Kani, S., Bae, Y. K., Shimizu, T., Tanabe, K., Satou, C., Parsons, M. J., Scott, E., Higashijima, S. I., & Hibi, M. (2010). Proneural gene-linked neurogenesis in zebrafish cerebellum. Developmental Biology, 343(1–2), 1–17. https://doi.org/10.1016/j.ydbio.2010.03.024

Kano, M., Watanabe, T., Uesaka, N., & Watanabe, M. (2018). Multiple Phases of Climbing Fiber Synapse Elimination in the Developing Cerebellum. Cerebellum (London, England), 17(6), 722–734. https://doi.org/10.1007/s12311-018-0964-z

Kashiwabuchi, N., Ikeda, K., Araki, K., Hirano, T., Shibuki, K., Takayama, C., Inoue, Y., Kutsuwada, T., Yagi, T., & Kang, Y. (1995). Impairment of motor coordination, Purkinje cell synapse formation, and cerebellar long-term depression in GluR delta 2 mutant mice. Cell, 81(2), 245–252. https://doi.org/10.1016/0092-8674(95)90334-8

Kemenes, I., Marra, V., Crossley, M., Samu, D., Staras, K., Kemenes, G., & Nowotny, T. (2011). Dynamic clamp with StdpC software. Nature Protocols, 6(3), Article 3. https://doi.org/10.1038/nprot.2010.200

Kiehn, O. (1991). Plateau potentials and active integration in the ‘final common pathway’ for motor behaviour. Trends in Neurosciences, 14(2), 68–73. https://doi.org/10.1016/0166-2236(91)90023-N

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. Developmental Dynamics: An Official Publication of the American Association of Anatomists, 203(3), 253–310. https://doi.org/10.1002/aja.1002030302

Kitamura, K., & Häusser, M. (2011). Dendritic Calcium Signaling Triggered by Spontaneous and Sensory-Evoked Climbing Fiber Input to Cerebellar Purkinje Cells In Vivo. Journal of Neuroscience, 31(30), 10847–10858. https://doi.org/10.1523/JNEUROSCI.2525-10.2011

Knogler, L. D., Markov, D. A., Dragomir, E. I., Štih, V., & Portugues, R. (2017). Sensorimotor Representations in Cerebellar Granule Cells in Larval Zebrafish Are Dense, Spatially Organized, and Non-temporally Patterned. Current Biology, 27(9), 1288–1302. https://doi.org/10.1016/j.cub.2017.03.029

Kohashi, T., & Oda, Y. (2008). Initiation of Mauthner- or Non-Mauthner-Mediated Fast Escape Evoked by Different Modes of Sensory Input. Journal of Neuroscience, 28(42), 10641–10653. https://doi.org/10.1523/JNEUROSCI.1435-08.2008

Korn, H., & Faber, D. S. (2005). The Mauthner cell half a century later: A neurobiological model for decision-making? Neuron, 47(1), 13–28. https://doi.org/10.1016/j.neuron.2005.05.019

Koyama, M., Minale, F., Shum, J., Nishimura, N., Schaffer, C. B., & Fetcho, J. R. (2016). A circuit motif in the zebrafish hindbrain for a two alternative behavioral choice to turn left or right. ELife, 5, e16808. https://doi.org/10.7554/eLife.16808

Krahe, R., & Gabbiani, F. (2004). Burst firing in sensory systems. Nature Reviews Neuroscience, 5(1), Article 1. https://doi.org/10.1038/nrn1296

Kreuz, T., Chicharro, D., Houghton, C., Andrzejak, R. G., & Mormann, F. (2013). Monitoring spike train synchrony. Journal of Neurophysiology, 109(5), 1457–1472. https://doi.org/10.1152/jn.00873.2012

Latham, A., & Paul, D. H. (1971). Spontaneous activity of cerebellar Purkinje cells and their responses to impulses in climbing fibres. The Journal of Physiology, 213(1), 135–156. https://doi.org/10.1113/jphysiol.1971.sp009373

Lechner, H. A., Baxter, D. A., Clark, J. W., & Byrne, J. H. (1996). Bistability and its regulation by serotonin in the endogenously bursting neuron R15 in Aplysia. Journal of Neurophysiology, 75(2), 957–962. https://doi.org/10.1152/jn.1996.75.2.957

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten Digit Recognition with a Back-Propagation Network. Advances in Neural Information Processing Systems, 2. https://proceedings.neurips.cc/paper_files/paper/1989/hash/53c3bce66e43be4f209556518c2fcb54-Abstract.html

Lev-Ram, V., Miyakawa, H., Lasser-Ross, N., & Ross, W. N. (1992). Calcium transients in cerebellar Purkinje neurons evoked by intracellular stimulation. Journal of Neurophysiology, 68(4), 1167–1177. https://doi.org/10.1152/jn.1992.68.4.1167

Lin, J. W., & Faber, D. S. (1988). Synaptic transmission mediated by single club endings on the goldfish Mauthner cell. I. Characteristics of electrotonic and chemical postsynaptic potentials. Journal of Neuroscience, 8(4), 1302–1312. https://doi.org/10.1523/JNEUROSCI.08-04-01302.1988

Linnainmaa, S. (1976). Taylor expansion of the accumulated rounding error. BIT Numerical Mathematics, 16(2), 146–160. https://doi.org/10.1007/BF01931367

Liu, L., & Chen, X. (2018). Intercellular and systemic trafficking of RNAs in plants. Nature Plants, 4(11), Article 11. https://doi.org/10.1038/s41477-018-0288-5

Liu, P., Chen, B., Mailler, R., & Wang, Z.-W. (2017). Antidromic-rectifying gap junctions amplify chemical transmission at functionally mixed electrical-chemical synapses. Nature Communications, 8(1), Article 1. https://doi.org/10.1038/ncomms14818

Llinás, R., & Sugimori, M. (1980). Electrophysiological properties of in vitro Purkinje cell

somata in mammalian cerebellar slices. The Journal of Physiology, 305(1), 171–195. https://doi.org/10.1113/jphysiol.1980.sp013357

Loewenstein, Y., Mahon, S., Chadderton, P., Kitamura, K., Sompolinsky, H., Yarom, Y., & Häusser, M. (2005). Bistability of cerebellar Purkinje cells modulated by sensory stimulation. Nature Neuroscience, 8(2), Article 2. https://doi.org/10.1038/nn1393

Loewenstein, Y., Mahon, S., Chadderton, P., Kitamura, K., Sompolinsky, H., Yarom, Y., & Häusser, M. (2006). Purkinje cells in awake behaving animals operate in stable upstate membrane potential. Nature Neuroscience, 9(4), Article 4. https://doi.org/10.1038/nn0406-461

Marder, E., Abbott, L. F., Turrigiano, G. G., Liu, Z., & Golowasch, J. (1996). Memory from the dynamics of intrinsic membrane currents. Proceedings of the National Academy of Sciences, 93(24), 13481–13486. https://doi.org/10.1073/pnas.93.24.13481

Marder, E., Gutierrez, G. J., & Nusbaum, M. P. (2017). Complicating connectomes: Electrical coupling creates parallel pathways and degenerate circuit mechanisms. Developmental Neurobiology, 77(5), 597–609. https://doi.org/10.1002/dneu.22410

Marple-Horvat, D. E., & Stein, J. F. (1987). Cerebellar neuronal activity related to arm movements in trained rhesus monkeys. The Journal of Physiology, 394(1), 351–366. https://doi.org/10.1113/jphysiol.1987.sp016874

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. Nature Neuroscience, 21(9), Article 9. https://doi.org/10.1038/s41593-018-0209-y

Matsui, H., Namikawa, K., Babaryka, A., & Köster, R. W. (2014). Functional regionalization of the teleost cerebellum analyzed in vivo. Proceedings of the National Academy of Sciences, 111(32), 11846–11851. https://doi.org/10.1073/pnas.1403105111

McCormick, D. A., Connors, B. W., Lighthall, J. W., & Prince, D. A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. Journal of Neurophysiology, 54(4), 782–806. https://doi.org/10.1152/jn.1985.54.4.782

McLean, D. L., & Fetcho, J. R. (2004). Ontogeny and innervation patterns of dopaminergic, noradrenergic, and serotonergic neurons in larval zebrafish. Journal of Comparative Neurology, 480(1), 38–56. https://doi.org/10.1002/cne.20280

Meek, J., & Nieuwenhuys, R. (1991). Palisade pattern of mormyrid Purkinje cells: A correlated light and electron microscopic study. Journal of Comparative Neurology, 306(1), 156–192. https://doi.org/10.1002/cne.903060111

Meek, J., Yang, J. Y., Han, V. Z., & Bell, C. C. (2008). Morphological analysis of the mormyrid cerebellum using immunohistochemistry, with emphasis on the unusual neuronal organization of the valvula. Journal of Comparative Neurology, 510(4), 396–

421. https://doi.org/10.1002/cne.21809

Meeker, N. D., Hutchinson, S. A., Ho, L., & Trede, N. S. (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. BioTechniques, 43(5), 610–614. https://doi.org/10.2144/000112619

Miller, A. C., Whitebirch, A. C., Shah, A. N., Marsden, K. C., Granato, M., O'Brien, J., & Moens, C. B. (2017). A genetic basis for molecular asymmetry at vertebrate electrical synapses. ELife, 6, e25364. https://doi.org/10.7554/eLife.25364

Misgeld, U., Calabresi, P., & Dodt, H. U. (1986). Muscarinic modulation of calcium dependent plateau potentials in rat neostriatal neurons. Pflugers Archiv: European Journal of Physiology, 407(5), 482–487. https://doi.org/10.1007/BF00657504

Miyawaki, A., Llopis, J., Heim, R., McCaffery, J. M., Adams, J. A., Ikura, M., & Tsien, R. Y. (1997). Fluorescent indicators for Ca2+based on green fluorescent proteins and calmodulin. Nature, 388(6645), Article 6645. https://doi.org/10.1038/42264

Moisescu, D. G., Ashley, C. C., & Campbell, A. K. (1975). Comparative aspects of the calcium-sensitive photoproteins aequorin and obelin. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 396(1), 133–140. https://doi.org/10.1016/0005-2728(75)90196-6

Mostofi, A., Holtzman, T., Grout, A. S., Yeo, C. H., & Edgley, S. A. (2010). Electrophysiological Localization of Eyeblink-Related Microzones in Rabbit Cerebellar Cortex. Journal of Neuroscience, 30(26), 8920–8934. https://doi.org/10.1523/JNEUROSCI.6117-09.2010

Nakai, J., Ohkura, M., & Imoto, K. (2001). A high signal-to-noise Ca2+ probe composed of a single green fluorescent protein. Nature Biotechnology, 19(2), Article 2. https://doi.org/10.1038/84397

Nakayama, H., & Oda, Y. (2004). Common Sensory Inputs and Differential Excitability of Segmentally Homologous Reticulospinal Neurons in the Hindbrain. Journal of Neuroscience, 24(13), 3199–3209. https://doi.org/10.1523/JNEUROSCI.4419-03.2004

Narayanan, S., Varma, A., & Thirumalai, V. (2021). Rapid acquisition of internal models of timing by Purkinje cells enables faster behavioral responses (p. 2021.04.28.441782). bioRxiv. https://doi.org/10.1101/2021.04.28.441782

Ojakangas, C. L., & Ebner, T. J. (1994). Purkinje cell complex spike activity during voluntary motor learning: Relationship to kinematics. Journal of Neurophysiology, 72(6), 2617–2630. https://doi.org/10.1152/jn.1994.72.6.2617

Oldfield, C. S., Marty, A., & Stell, B. M. (2010). Interneurons of the cerebellar cortex toggle Purkinje cells between up and down states. Proceedings of the National Academy of Sciences of the United States of America, 107(29), 13153–13158. https://doi.org/10.1073/pnas.1002082107

OpenAI. (2023). GPT-4 Technical Report (arXiv:2303.08774). arXiv.

https://doi.org/10.48550/arXiv.2303.08774

Orger, M. B., & de Polavieja, G. G. (2017). Zebrafish Behavior: Opportunities and Challenges. Annual Review of Neuroscience, 40(1), 125–147. https://doi.org/10.1146/annurev-neuro-071714-033857

Otis, T. S. (2016). Simple Spikes and Complex Spikes. In D. L. Gruol, N. Koibuchi, M. Manto, M. Molinari, J. D. Schmahmann, & Y. Shen (Eds.), Essentials of Cerebellum and Cerebellar Disorders: A Primer For Graduate Students (pp. 299–303). Springer International Publishing. https://doi.org/10.1007/978-3-319-24551-5_40

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L. F., Dalgleish, H., Carandini, M., & Harris, K. D. (2017). Suite2p: Beyond 10,000 neurons with standard two-photon microscopy (p. 061507). bioRxiv. https://doi.org/10.1101/061507

Pachitariu, M., Stringer, C., & Harris, K. D. (2018). Robustness of Spike Deconvolution for Neuronal Calcium Imaging. Journal of Neuroscience, 38(37), 7976–7985. https://doi.org/10.1523/JNEUROSCI.3339-17.2018

Patten, S. A., & Ali, D. W. (2007). AMPA receptors associated with zebrafish Mauthner cells switch subunits during development. The Journal of Physiology, 581(3), 1043–1056. https://doi.org/10.1113/jphysiol.2007.129999

Pereda, A. E. (2014). Electrical synapses and their functional interactions with chemical synapses. Nature Reviews Neuroscience, 15(4), Article 4. https://doi.org/10.1038/nrn3708

Pereda, A. E., & Faber, D. S. (1996). Activity-dependent short-term enhancement of intercellular coupling. Journal of Neuroscience, 16(3), 983–992. https://doi.org/10.1523/JNEUROSCI.16-03-00983.1996

Peron, S., Chen, T.-W., & Svoboda, K. (2015). Comprehensive imaging of cortical networks. Current Opinion in Neurobiology, 32, 115–123. https://doi.org/10.1016/j.conb.2015.03.016

Person, A. L., & Raman, I. M. (2012). Purkinje neuron synchrony elicits time-locked spiking in the cerebellar nuclei. Nature, 481(7382), Article 7382. https://doi.org/10.1038/nature10732

Pnevmatikakis, E. A., Soudry, D., Gao, Y., Machado, T. A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W., Ahrens, M., Bruno, R., Jessell, T. M., Peterka, D. S., Yuste, R., & Paninski, L. (2016). Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. Neuron, 89(2), 285–299. https://doi.org/10.1016/j.neuron.2015.11.037

Popa, L. S., Streng, M. L., & Ebner, T. J. (2019). Purkinje Cell Representations of Behavior: Diary of a Busy Neuron. The Neuroscientist, 25(3), 241–257. https://doi.org/10.1177/1073858418785628

Prekop, H.-T., Kroiss, A., Rook, V., Zagoraiou, L., Jessell, T. M., Fernandes, C., Delogu, A., & Wingate, R. J. T. (2018). Sox14 Is Required for a Specific Subset of Cerebello-Olivary Projections. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 38(44), 9539–9550. https://doi.org/10.1523/JNEUROSCI.1456-18.2018

Prinz, A. A., Abbott, L. F., & Marder, E. (2004). The dynamic clamp comes of age. Trends in Neurosciences, 27(4), 218–224. https://doi.org/10.1016/j.tins.2004.02.004

Puhl, J. G., & Mesce, K. A. (2008). Dopamine activates the motor pattern for crawling in the medicinal leech. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 28(16), 4192–4200. https://doi.org/10.1523/JNEUROSCI.0136-08.2008

Raman, I. M., & Bean, B. P. (1997). Resurgent sodium current and action potential formation in dissociated cerebellar Purkinje neurons. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 17(12), 4517–4526. https://doi.org/10.1523/JNEUROSCI.17-12-04517.1997

Ramirez, J. E., & Stell, B. M. (2016). Calcium Imaging Reveals Coordinated Simple Spike Pauses in Populations of Cerebellar Purkinje Cells. Cell Reports, 17(12), 3125–3132. https://doi.org/10.1016/j.celrep.2016.11.075

Ramón y Cajal, S. (1911). Histologie du système nerveux de l'homme et des vertébrés. https://gallica.bnf.fr/ark:/12148/bpt6k6213192g.texteImage

Rash, J. E., Dillman, R. K., Bilhartz, B. L., Duffy, H. S., Whalen, L. R., & Yasumura, T. (1996). Mixed synapses discovered and mapped throughout mammalian spinal cord. Proceedings of the National Academy of Sciences, 93(9), 4235–4239. https://doi.org/10.1073/pnas.93.9.4235

Renninger, S. L., & Orger, M. B. (2013). Two-photon imaging of neural population activity in zebrafish. Methods, 62(3), 255–267. https://doi.org/10.1016/j.ymeth.2013.05.016

Robbins, M., Christensen, C. N., Kaminski, C. F., & Zlatic, M. (2021). Calcium imaging analysis – how far have we come? (10:258). F1000Research. https://doi.org/10.12688/f1000research.51755.2

Robra, L., & Thirumalai, V. (2016). The Intracellular Signaling Molecule Darpp-32 Is a Marker for Principal Neurons in the Cerebellum and Cerebellum-Like Circuits of Zebrafish. Frontiers in Neuroanatomy, 10, 81. https://doi.org/10.3389/fnana.2016.00081

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65(6), 386–408. https://doi.org/10.1037/h0042519

Rupprecht, P., Carta, S., Hoffmann, A., Echizen, M., Blot, A., Kwan, A. C., Dan, Y., Hofer, S. B., Kitamura, K., Helmchen, F., & Friedrich, R. W. (2021). A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging. Nature Neuroscience, 24(9), Article 9. https://doi.org/10.1038/s41593-021-00895-5

Schmid, S. M., & Hollmann, M. (2008). To gate or not to gate: Are the delta subunits in the glutamate receptor family functional ion channels? Molecular Neurobiology, 37(2–3), 126–141. https://doi.org/10.1007/s12035-008-8025-0

Schmidhuber, J. (2022). Annotated History of Modern AI and Deep Learning (arXiv:2212.11279). arXiv. https://doi.org/10.48550/arXiv.2212.11279

Schonewille, M., Khosrovani, S., Winkelman, B. H. J., Hoebeek, F. E., De Jeu, M. T. G., Larsen, I. M., Van Der Burg, J., Schmolesky, M. T., Frens, M. A., & De Zeeuw, C. I. (2006). Purkinje cells in awake behaving animals operate at the upstate membrane potential. Nature Neuroscience, 9(4), Article 4. https://doi.org/10.1038/nn0406-459

Schwindt, P., & Crill, W. (1980). Role of a persistent inward current in motoneuron bursting during spinal seizures. Journal of Neurophysiology, 43(5), 1296–1318. https://doi.org/10.1152/jn.1980.43.5.1296

Schwindt, P., & Crill, W. E. (1977). A persistent negative resistance in cat lumbar motoneurons. Brain Research, 120(1), 173–178. https://doi.org/10.1016/0006-8993(77)90510-8

Sengupta, M. (2016). Purkinje neurons in action: From single cells to ensembles. Tata Institute of Fundamental Research.

Sengupta, M., & Thirumalai, V. (2015). AMPA receptor mediated synaptic excitation drives state-dependent bursting in Purkinje neurons of zebrafish larvae. ELife, 4, e09158. https://doi.org/10.7554/eLife.09158

Serrano-Velez, J. L., Rodriguez-Alvarado, M., Torres-Vazquez, I. I., Fraser, S. E., Yasumura, T., Vanderpool, K. G., Rash, J. E., & Rosa-Molinar, E. (2014). Abundance of gap junctions at glutamatergic mixed synapses in adult Mosquitofish spinal cord neurons. Frontiers in Neural Circuits, 8. https://www.frontiersin.org/article/10.3389/fncir.2014.00066

Sharp, A. A., O'Neil, M. B., Abbott, L. F., & Marder, E. (1993). The dynamic clamp: Artificial conductances in biological neurons. Trends in Neurosciences, 16(10), 389–394. https://doi.org/10.1016/0166-2236(93)90004-6

Shidara, M., Kawano, K., Gomi, H., & Kawato, M. (1993). Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. Nature, 365(6441), Article 6441. https://doi.org/10.1038/365050a0

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science, 362(6419), 1140–1144. https://doi.org/10.1126/science.aar6404

Smetters, D., Majewska, A., & Yuste, R. (1999). Detecting action potentials in neuronal populations with calcium imaging. Methods (San Diego, Calif.), 18(2), 215–221.

https://doi.org/10.1006/meth.1999.0774

Steinmetz, N. A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., Chen, S., Colonell, J., Gardner, R. J., Karsh, B., Kloosterman, F., Kostadinov, D., Mora-Lopez, C., O'Callaghan, J., Park, J., … Harris, T. D. (2021). Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. Science (New York, N.Y.), 372(6539), eabf4588. https://doi.org/10.1126/science.abf4588

Steriade, M., McCormick, D. A., & Sejnowski, T. J. (1993). Thalamocortical Oscillations in the Sleeping and Aroused Brain. Science, 262(5134), 679–685. https://doi.org/10.1126/science.8235588

Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R., & Bialek, W. (1998). Entropy and Information in Neural Spike Trains. Physical Review Letters, 80(1), 197–200. https://doi.org/10.1103/PhysRevLett.80.197

Takeuchi, M., Matsuda, K., Yamaguchi, S., Asakawa, K., Miyasaka, N., Lal, P., Yoshihara, Y., Koga, A., Kawakami, K., Shimizu, T., & Hibi, M. (2015). Establishment of Gal4 transgenic zebrafish lines for analysis of development of cerebellar neural circuitry. Developmental Biology, 397(1), 1–17. https://doi.org/10.1016/j.ydbio.2014.09.030

Takeuchi, M., Yamaguchi, S., Sakakibara, Y., Hayashi, T., Matsuda, K., Hara, Y., Tanegashima, C., Shimizu, T., Kuraku, S., & Hibi, M. (2017). Gene expression profiling of granule cells and Purkinje cells in the zebrafish cerebellum. Journal of Comparative Neurology, 525(7), 1558–1585. https://doi.org/10.1002/cne.24114

Tanabe, K., Kani, S., Shimizu, T., Bae, Y.-K., Abe, T., & Hibi, M. (2010). Atypical Protein Kinase C Regulates Primary Dendrite Specification of Cerebellar Purkinje Cells by Localizing Golgi Apparatus. Journal of Neuroscience, 30(50), 16983–16992. https://doi.org/10.1523/JNEUROSCI.3352-10.2010

Thach, W. T. (1967). Somatosensory receptive fields of single units in cat cerebellar cortex. Journal of Neurophysiology, 30(4), 675–696. https://doi.org/10.1152/jn.1967.30.4.675

Theis, L., Berens, P., Froudarakis, E., Reimer, J., Román Rosón, M., Baden, T., Euler, T., Tolias, A. S., & Bethge, M. (2016). Benchmarking Spike Rate Inference in Population Calcium Imaging. Neuron, 90(3), 471–482. https://doi.org/10.1016/j.neuron.2016.04.014

Thirumalai, V., & Jha, U. (2022). Recruitment of Motoneurons. Advances in Neurobiology, 28, 169–190. https://doi.org/10.1007/978-3-031-07167-6_8

Tian, L., Hires, S. A., Mao, T., Huber, D., Chiappe, M. E., Chalasani, S. H., Petreanu, L., Akerboom, J., McKinney, S. A., Schreiter, E. R., Bargmann, C. I., Jayaraman, V., Svoboda, K., & Looger, L. L. (2009). Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. Nature Methods, 6(12), Article 12. https://doi.org/10.1038/nmeth.1398

Truett, G. e., Heeger, P., Mynatt, R. l., Truett, A. a., Walker, J. a., & Warman, M. l. (2000). Preparation of PCR-Quality Mouse Genomic DNA with Hot Sodium Hydroxide and Tris (HotSHOT). BioTechniques, 29(1), 52–54. https://doi.org/10.2144/00291bm09

Tsien, R. Y. (1980). New calcium indicators and buffers with high selectivity against magnesium and protons: Design, synthesis, and properties of prototype structures. Biochemistry, 19(11), 2396–2404. https://doi.org/10.1021/bi00552a018

van Rossum, M. C. (2001). A novel spike distance. Neural Computation, 13(4), 751–763. https://doi.org/10.1162/089976601300014321

Victor, J. D. (2002). Binless strategies for estimation of information from neural data. Physical Review E, 66(5), 051903. https://doi.org/10.1103/PhysRevE.66.051903

Vladimirov, N., Mu, Y., Kawashima, T., Bennett, D. V., Yang, C.-T., Looger, L. L., Keller, P. J., Freeman, J., & Ahrens, M. B. (2014). Light-sheet functional imaging in fictively behaving zebrafish. Nature Methods, 11(9), Article 9. https://doi.org/10.1038/nmeth.3040

Vladimirov, N., Wang, C., Höckendorf, B., Pujala, A., Tanimoto, M., Mu, Y., Yang, C.-T., Wittenbach, J. D., Freeman, J., Preibisch, S., Koyama, M., Keller, P. J., & Ahrens, M. B. (2018). Brain-wide circuit interrogation at the cellular level guided by online analysis of neuronal function. Nature Methods, 15(12), Article 12. https://doi.org/10.1038/s41592-018-0221-x

Vogelstein, J. T., Packer, A. M., Machado, T. A., Sippy, T., Babadi, B., Yuste, R., & Paninski, L. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. Journal of Neurophysiology, 104(6), 3691–3704. https://doi.org/10.1152/jn.01073.2009

Wang, F., Xu, Q., Wang, W., Takano, T., & Nedergaard, M. (2012). Bergmann glia modulate cerebellar Purkinje cell bistability via Ca2+-dependent K+ uptake. Proceedings of the National Academy of Sciences, 109(20), 7911–7916. https://doi.org/10.1073/pnas.1120380109

Watanabe, T., Shimazaki, T., Mishiro, A., Suzuki, T., Hirata, H., Tanimoto, M., & Oda, Y. (2014). Coexpression of auxiliary Kvβ2 subunits with Kv1.1 channels is required for developmental acquisition of unique firing properties of zebrafish Mauthner cells. Journal of Neurophysiology, 111(6), 1153–1164. https://doi.org/10.1152/jn.00596.2013

Watanabe, T., Shimazaki, T., & Oda, Y. (2017). Coordinated Expression of Two Types of Low-Threshold K+ Channels Establishes Unique Single Spiking of Mauthner Cells among Segmentally Homologous Neurons in the Zebrafish Hindbrain. ENeuro, 4(5), ENEURO.0249-17.2017. https://doi.org/10.1523/ENEURO.0249-17.2017

Wei, Z., Lin, B.-J., Chen, T.-W., Daie, K., Svoboda, K., & Druckmann, S. (2020). A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology. PLOS Computational Biology, 16(9), e1008198.

https://doi.org/10.1371/journal.pcbi.1008198

Weimann, J. M., Marder, E., Evans, B., & Calabrese, R. L. (1993). The Effects of SDRNFLRFamide and TNRNFLRFamide on the Motor Patterns of the Stomatogastric Ganglion of the Crab Cancer Borealis. Journal of Experimental Biology, 181(1), 1–26. https://doi.org/10.1242/jeb.181.1.1

What are Recurrent Neural Networks? | IBM. (2023, August 1). https://www.ibm.com/topics/recurrent-neural-networks

What is Machine Learning? (2023, August 1). Google for Developers. https://developers.google.com/machine-learning/intro-to-ml/what-is-ml

Williams, S. R., Christensen, S. R., Stuart, G. J., & Häusser, M. (2002). Membrane potential bistability is controlled by the hyperpolarization-activated current IH in rat cerebellar Purkinje neurons in vitro. The Journal of Physiology, 539(2), 469–483. https://doi.org/10.1113/jphysiol.2001.013136

Wilson, C. J., & Groves, P. M. (1981). Spontaneous firing patterns of identified spiny neurons in the rat neostriatum. Brain Research, 220(1), 67–80. https://doi.org/10.1016/0006-8993(81)90211-0

Wu, B., Blot, F. G., Wong, A. B., Osório, C., Adolfs, Y., Pasterkamp, R. J., Hartmann, J., Becker, E. B., Boele, H.-J., De Zeeuw, C. I., & Schonewille, M. (2019). TRPC3 is a major contributor to functional heterogeneity of cerebellar Purkinje cells. ELife, 8, e45590. https://doi.org/10.7554/eLife.45590

Yaksi, E., & Friedrich, R. W. (2006). Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca2+ imaging. Nature Methods, 3(5), Article 5. https://doi.org/10.1038/nmeth874

Yala, A., Mikhael, P. G., Strand, F., Lin, G., Smith, K., Wan, Y.-L., Lamb, L., Hughes, K., Lehman, C., & Barzilay, R. (2021). Toward robust mammography-based models for breast cancer risk. Science Translational Medicine, 13(578), eaba4373. https://doi.org/10.1126/scitranslmed.aba4373

Yang, W., & Yuste, R. (2017). In vivo imaging of neural activity. Nature Methods, 14(4), Article 4. https://doi.org/10.1038/nmeth.4230

Yang, X.-D., Korn, H., & Faber, D. S. (1990). Long-term potentiation of electrotonic coupling at mixed synapses. Nature, 348(6301), Article 6301. https://doi.org/10.1038/348542a0

Zeldenrust, F., Wadman, W. J., & Englitz, B. (2018). Neural Coding With Bursts-Current State and Future Perspectives. Frontiers in Computational Neuroscience, 12, 48. https://doi.org/10.3389/fncom.2018.00048

Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G. J., Tsang, A., Tsegaye, G., Narayan, S., Obara, C. J., Lim, J.-X., Patel, R., Zhang, R., Ahrens, M. B., Turner, G. C., Wang, S. S.-H., Korff, W. L., … Looger, L. L. (2023). Fast

and sensitive GCaMP calcium indicators for imaging neural populations. Nature, 615(7954), Article 7954. https://doi.org/10.1038/s41586-023-05828-9

Zhou, H., Lin, Z., Voges, K., Ju, C., Gao, Z., Bosman, L. W., Ruigrok, T. J., Hoebeek, F. E., De Zeeuw, C. I., & Schonewille, M. (2014). Cerebellar modules operate at different frequencies. ELife, 3, e02536. https://doi.org/10.7554/eLife.02536

Appendices

Appendix 1 - Social behaviour in auts2a-/- mutants

A1.1 Introduction

The zebrafish, *Danio rerio*, a freshwater tropical teleost fish, has emerged as a powerful vertebrate model organism, particularly to study development and genetics. They are small, relatively easy to breed, and with the advent of multiple different technologies, are becoming easier to manipulate, both genetically and otherwise. More recently, there has been an increased interest in using them as models to study behavioural genetics. Adult zebrafish have been used to model complex behaviours, such as reward behaviour, aggression, and also social behaviour (Norton & Bally-Cuif, 2010).

Social behaviour studies in zebrafish typically involve a characterization of the group behaviour of these fish. Zebrafish form either a loose cluster of individuals (a shoal), or an aggregation that exhibits polarised, synchronised motion (a school) (Miller & Gerlai, 2012). Studies have shown that the social preference of the fish is dependent on the social exposure of the fish during its early life (Engeszer et al., 2004).
More recently, the development of social choice in zebrafish was studied, using a simple 2-choice assay, and it was reported that these fish develop a social preference by the age of 3 weeks post fertilisation. This study further reported that the behaviour could be elicited using vision alone (Dreosti et al., 2015).

Given that zebrafish are amenable to genetic perturbations and that they exhibit social behaviours, we proposed to use this organism as an animal model for a behavioural reverse genetic study. To that end, we identified an autism susceptibility candidate gene, called Auts2, which is the fish homolog of the human AUTS2 gene. A previous study from our lab (Kondrychyn et al., 2017) found that auts2 has four paralogs in zebrafish, namely, auts2a, auts2b, fibrosin (fbrs) and fibrosin-like 1 (fbrsl1). Of these, auts2a is the closest ortholog of the mammalian Auts2 (Jha et al., 2021; Kondrychyn et al., 2017) and both auts2a and auts2b are expressed in the zebrafish brain at various developmental timepoints. Mutants with knockouts of the Auts2a and Auts2b genes were generated using TALEN-based mutagenesis (Bedell et al., 2012; Joung & Sander, 2013). Although they show a normal developmental pattern, I wanted to investigate their ability to perform social choices using the assays developed for the same.

A1.2 Results

A1.2.1 A reverse genetics model of an autism susceptibility candidate gene, Auts2

The zebrafish *Danio rerio* has become a great model organism to make disease models and study the function of genes in the context of development and behaviour. We developed a reverse genetics model of the autism susceptibility candidate gene Auts2, which is the fish homolog of the human AUTS2 gene. Auts2 was first identified in monozygotic twins with autism (Sultana et al., 2002), and knocking out this gene in mouse models led to reduced righting reflexes and ultrasonic vocalisations (Gao et al., 2014), the latter of which are used by animals for intraspecific communication (Portfors & Perkel, 2014).

The main features of the Auts2 paralogs in zebrafish and the knockouts generated are summarised in Figure 1. auts2a is a large genomic locus on chromosome 10, 560 kb long, and produces multiple transcripts, having at least 5 independent transcription start sites (Fig A1.1A). auts2b, on the other hand, is much smaller, only 20kb long, and located on chromosome 15 (Fig A1.1B). In order to make mutants, the exons common to all the transcripts were targeted for TALEN-mediated mutagenesis - exon 8 for auts2a and exon 3 for auts2b (Fig A1.1).

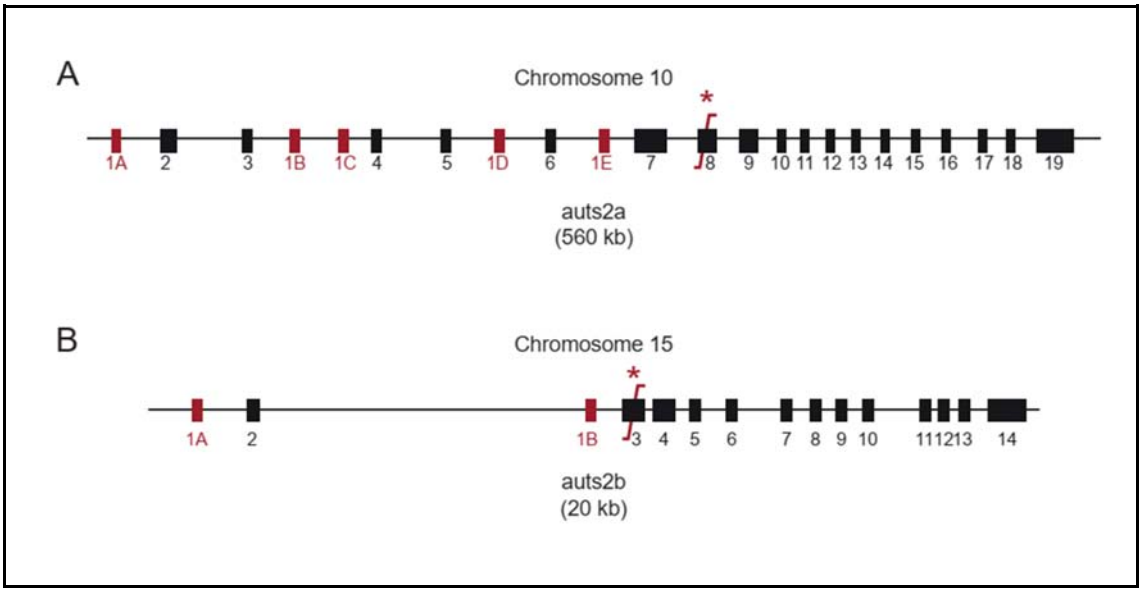


Figure A1.1 - Summary of gene structure and mutations of auts2 paralogs in the zebrafish *Danio rerio* (A) The gene structure of the auts2a gene, located on chromosome 10. Exons are marked as boxes, and introns as a line. Mutually exclusive transcription start sites giving rise to different first exons are marked in crimson and labelled as 1A-1E. Mutations were targeted to exon 8, marked with the crimson break mark and an asterisk. **(B)** The gene structure of the auts2b gene, located on chromosome 15, with features marked as in (A). Not to scale.

Given that the AUTS2 gene was identified in the context of autistic humans who have impaired social behaviour and because ultrasonic vocalisations were perturbed in Auts2 mutant mice, we wanted to ask whether defects in social behaviour are recapitulated in the zebrafish mutants, which could then potentially be used as a disease model.

A1.2.2 Adapting an existing assay to evaluate social choice in adult zebrafish

I first adapted an existing assay for social choice (Dreosti et al., 2015), identical in design except for the fact that compartment sizes were scaled up to suit zebrafish adults (Fig A1.2A). In this assay, a test fish is allowed to explore a compartment that is connected to two independent compartments that have cues present in them. First, the fish is acclimated to the tank for a period of 15 minutes, where no cue is presented in either compartment. Then, the social cue is introduced to one of the compartments and the behavioural response of the fish is measured for a period of 15 minutes. Two trials are performed per fish, the only difference being the side in which the cue was presented, in order to account for any side bias a fish might have.

The test compartment is separated from the cue compartments by a transparent barrier that allows the test fish to see the cues presented in the compartments. The entire setup is illuminated from below and imaged using a camera (Fig A1.2B, *left*). An image processing pipeline was written that segments the fish from the background and identifies the centroid of the fish (Fig A1.2B, *right*). These centroids are used to calculate a Social Preference Index (SPI) for each fish, which is taken to be the difference between the number of frames in which the fish was near the social cue compartment (Fig A1.2B, *left, green box*) and those in which the fish was near the no social cue compartment (Fig A1.2B, *left, magenta box*), divided by the total number of frames near either compartment (Fig A1.2B, *formula*). This index takes values between -1 (complete avoidance) to +1 (complete attraction).

For pilot experiments, I tested whether the response of fish in this assay is dependent on the time of day during which experiments were performed (Fig A1.2C). What I found was that,

regardless of the time of day, fish responded similarly. Initially, during the acclimatisation phase, most fish showed no particular preference for either compartment (the points are roughly symmetric around zero). However, in the test phase, fish seemed to have lower SPIs than during acclimatisation, suggesting that they tended to avoid the social cue stimulus presented to them. In order to test whether this response was specific to the social cue presented, or was independent of the kind of cue presented, I tested the same behaviour using a random object (a plastic plant, which is a naturalistic stimulus for fish) as cues instead. Fish seemed to not show a strongly aversive response to the random objects (Fig A1.2D). I went ahead and performed the experiment with wild-type (WT) and auts2b-/- fish. I saw wildly different SPIs for the same fish in the mutant, and quite a lot of variability even with WT fish in this assay (Fig A1.2E).

Furthermore, I also noticed that several fish froze in the corridor of the test compartment, not going towards either cue compartment (data not shown). Thus, I decided that while this assay may have been useful for young fish, it is not a feasible one for adult fish and has several design flaws that needed to be fixed before it could be used as a standard assay for social choice in adults.

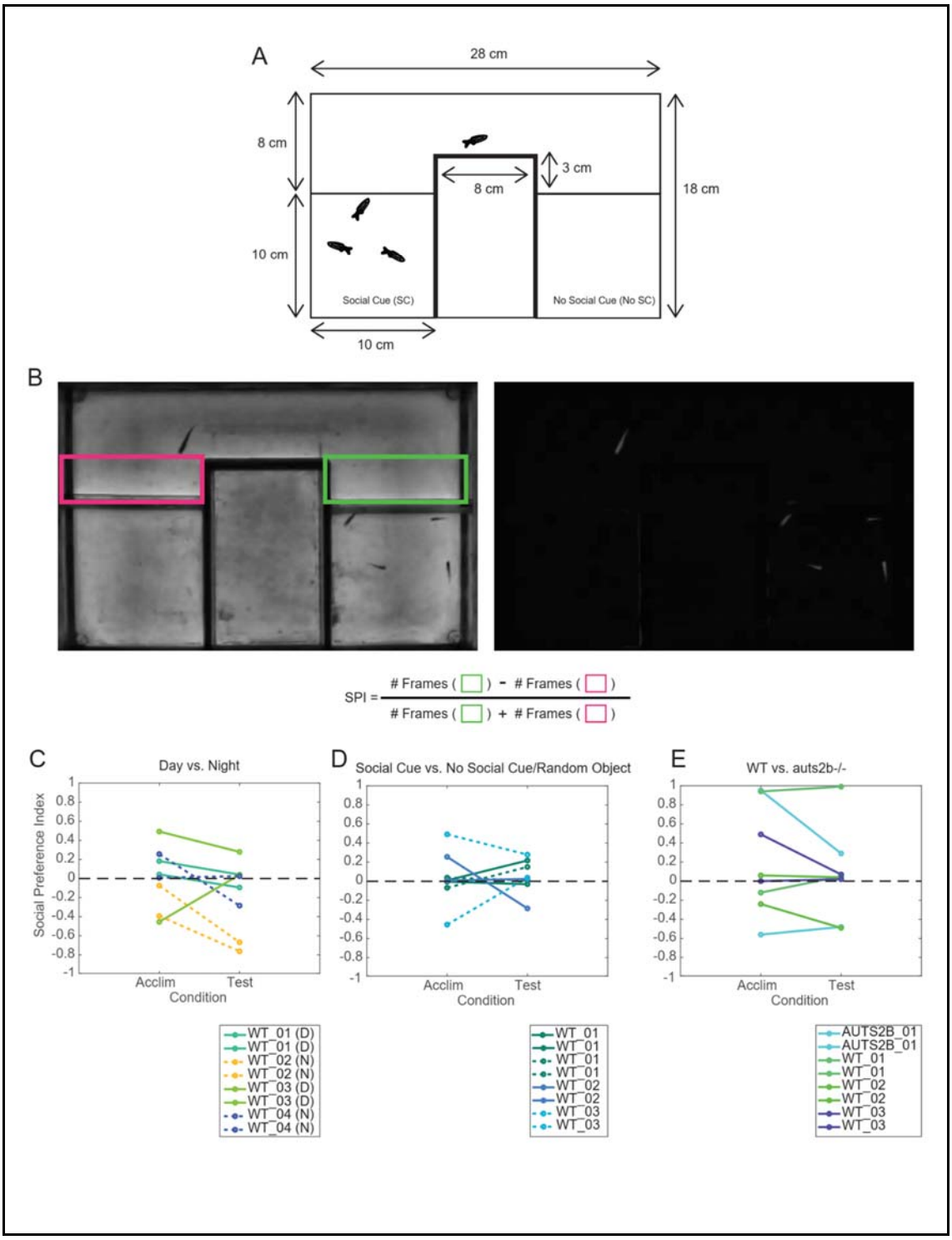

A
28 cm
8 cm
3 cm
8 cm
18 cm
10 cm
Social Cue (SC)
No Social Cue (No SC)
10 cm
B
SPI = # Frames ( ) - # Frames ( ) / # Frames ( ) + # Frames ( )
C
Day vs. Night
Social Preference Index
Acclim
Test
Condition
WT_01 (D)
WT_01 (D)
WT_02 (N)
WT_02 (N)
WT_03 (D)
WT_03 (D)
WT_04 (N)
WT_04 (N)
D
Social Cue vs. No Social Cue/Random Object
Acclim
Test
Condition
WT_01
WT_01
WT_01
WT_01
WT_02
WT_02
WT_03
WT_03
E
WT vs. auts2b-/-
Acclim
Test
Condition
AUTS2B_01
AUTS2B_01
WT_01
WT_01
WT_02
WT_02
WT_03
WT_03

Figure A1.2 - Adapting a behavioural setup to test social choice in adult zebrafish, *Danio rerio* (A) Schematic of the setup used to test social choice in adult zebrafish, with dimensions. Two square compartments (lower half) are for stimulus presentation. One compartment has fish (Social Cue compartment, SC), while the other is left empty (no SC compartment). The test fish is released in the third, U-shaped compartment and left free to explore both stimulus compartments, which it cannot enter but can see. **(B)** An example frame from a video recording of zebrafish behaviour. The raw image (*left*) was processed to segment the fish from the background (*right*) and extract its coordinates. The green and magenta boxes were manually drawn ROIs used to compute the Social Preference Index (SPI), using the formula shown below the images. **(C-E)** SPIs calculated for Indian Wild Type (WT) or Auts2b-/- (AUTS2B) fish under various conditions. Each animal was tested twice, with the side of the social cue presentation switched. Each test consists of a 15 minute long acclimatisation phase ("Acclim") when no stimulus was presented in either compartment followed by an equally long "Test" phase, where the cue compartments were appropriately filled. **(C)** The behaviour of WT fish, measured either during the day (solid lines and marked D in the legend) or during the night (dashed lines and marked N in the legend). **(D)** The behaviour of WT fish when presented with a social cue on one side, and either no cue (dashed lines) or a random object (solid lines and RO marked in the legend) on the other. **(E)** The behaviour of WT fish (solid lines) and Auts2b-/- fish (dashed lines).

A1.2.3 A T-Maze social choice assay could overcome many of the limitations of the previous setup

In order to overcome some of the design and practical flaws in the adapted setup, I decided to use a T-maze assay (Fig A1.3A), which has a large body of literature in ethological research (Darland & Dowling, 2001; Wendt et al., 2005). In this assay, the fish is left to acclimatise in one arm of the T-maze, and the two cues are in the remaining 2 arms of the maze. Once the animal has settled down and is ready to explore the remainder of the chamber it is housed in, it can exit the acclimatisation chamber and enter the test compartment, which is when the experiment begins. The advantage of this assay is that both cues are presented with equal likelihood, unlike in the previous assay, where a corridor reduces the likelihood that a fish will move from being near one cue compartment to exploring the other. Furthermore, this assay has the advantage that a trial begins only once the animal decides to actively explore the test compartment, and thus the acclimatisation phase would end at different times, unique to each fish.

Unfortunately, while the design of this assay was better than the previous one, most fish took very long times (> 30 minutes) to exit the acclimatisation chamber (Fig A1.3B), either due to a long period generally needed for acclimatisation or because the chamber was large enough that no exploration outside it was needed or because the opening to exit the acclimatisation compartment and enter the test compartment was too small.

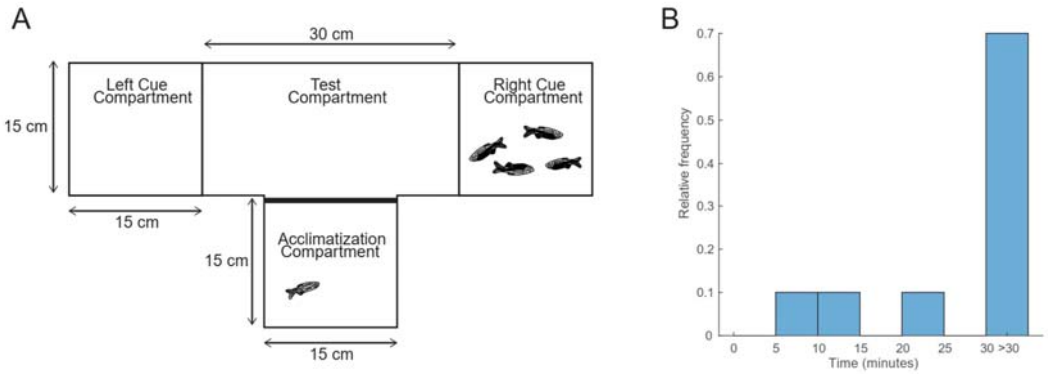


Figure A1.3 - An alternative, T-Maze setup used to test social choice in adult zebrafish (A) Schematic of the T-maze setup used to test social choice in adult zebrafish, with dimensions. In this setup, the acclimatisation compartment is separate from the test compartment. The cue compartments are first filled with the appropriate cue, and then a test fish placed in the acclimatisation compartment. Acclimatisation is complete when the test fish moves out of this compartment into the test compartment. The exit is such that it is equally likely for it to choose either of the cue compartments to go to first. **(B)** Histogram of exit times of N=10 fish in this assay.

A1.2.4 Both WT and auts2 mutant adult zebrafish adapt to a new environment in about 8-10 minutes

Given these results, I decided I would need to modify the social choice assay to ensure that the measurements of choice were not being affected by poor acclimatisation leading to anxiety or freezing behaviours. Thus, I used the novel tank test (Fig A1.4A), which is used to measure the time an animal takes to adapt to a new environment (Cachat et al., 2010). Adult zebrafish, when introduced into a new tank, first swim only at the bottom of the tank. Once they have adapted to the new environment, they begin to explore the surface of the water. The time taken for an animal to enter the top half of the tank and the number of times it crosses between the top and bottom halves are then taken as proxies of the anxiety levels of the animal.

I tested this in both Indian Wild-Type (WT) and auts2b mutant (hetero- or homozygous) animals. I found that the median time for WT fish to enter the top half of the tank was 500 s (~8.3 minutes). auts2b+/- and auts2b-/- fish took much shorter to explore the top half of the tank, with median times of 229 s and 47.5 s, respectively (Fig A1.4B). The median number of entries to the top half were 0, 4 and 9.5 for WT, auts2b+/- and auts2b-/- animals, respectively (Fig A1.4C).

Put together, this means that a 15 minute acclimatisation period should be more than sufficient to allow for animals of either genetic background to adapt to the new environment they are introduced to for the choice assay. Moreover, the results suggest that auts2b mutants have reduced anxiety levels than their wild-type counterparts.

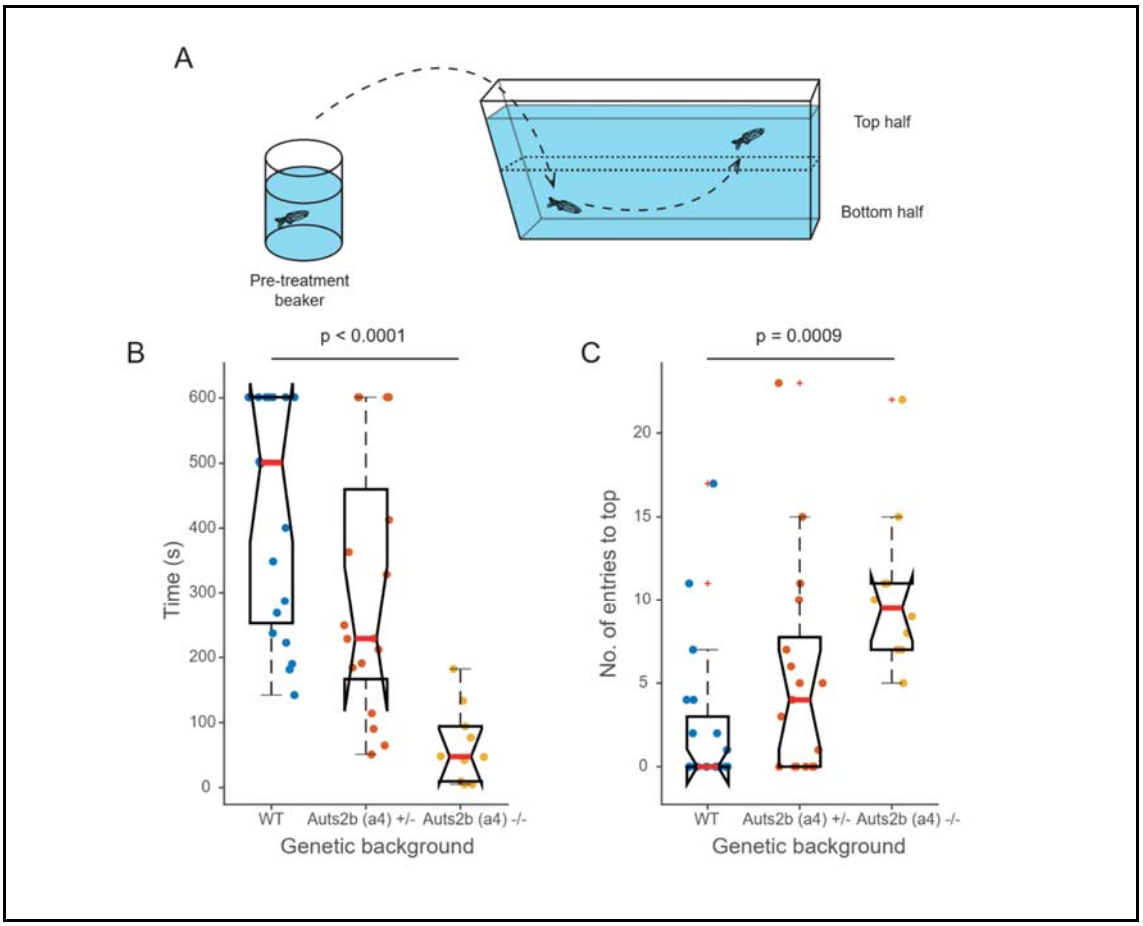


Figure A1.4 - Wild-type and auts2b mutant zebrafish acclimatise to a new environment in approximately 8 minutes (A) Schematic of the novel tank assay used to test the acclimatisation of adult zebrafish to a novel environment. Fish are tested individually. They are first kept in a pre-treatment beaker and then transferred to a new tank. The time taken for each animal to enter the top half of the tank **(B)**, and the number of entries to the top **(C)** are calculated. p-values were calculated using a generalised least-squares linear model predicting Latency~Genotype and No of entries to top~Genotype for (B) and (C), respectively.

A1.2.5 A three-compartment choice assay can be used to assess social preference

Taking into account the results from the novel tank test and the previous assays and their flaws, I modified the T-maze assay into a three-compartment choice assay (Fig A1.5A). In this assay, the test fish was directly introduced into the test compartment and left to acclimatise to its environment for a period of 15 minutes, following which the social cue was introduced into one of the two remaining compartments and was video recorded for a period of 15 minutes. Instead of calculating the SPI, which suffered from the drawback of having to manually draw regions of interest, I decided to calculate what I called the Social Interest Index (SII). This was simply taken to be the percentage of time spent in the half of the test compartment in which the cue was presented (Fig A1.5A). For the control condition, when no cue was present in either compartment, it was taken to be the percentage of time spent in the right half of the compartment.

I performed a pilot round of this assay with N=6 fish in five conditions - control (no cue in either compartment), a random object cue presented either in the right compartment or the left, and a social cue presented either in the right compartment or the left. As expected, I saw that the median control SIIs were close to 50% (49.33%), although 2/6 fish seemed to have a strong side bias (Fig A1.5B). The median SIIs for the random object stimuli was 54.67%, suggesting that fish had a slight preference for this object (Fig A1.5B). When presented with social cues, however, strong responses were elicited. The median SII in this condition was 61.39%, and almost all SIIs were higher than 50% (Fig A1.5B). In summary, this showed that the three-compartment choice assay was a robust assay to detect the social preference of adult zebrafish, and that this preference is, indeed, for a social cue and not the presence of a random object in one of the compartments.

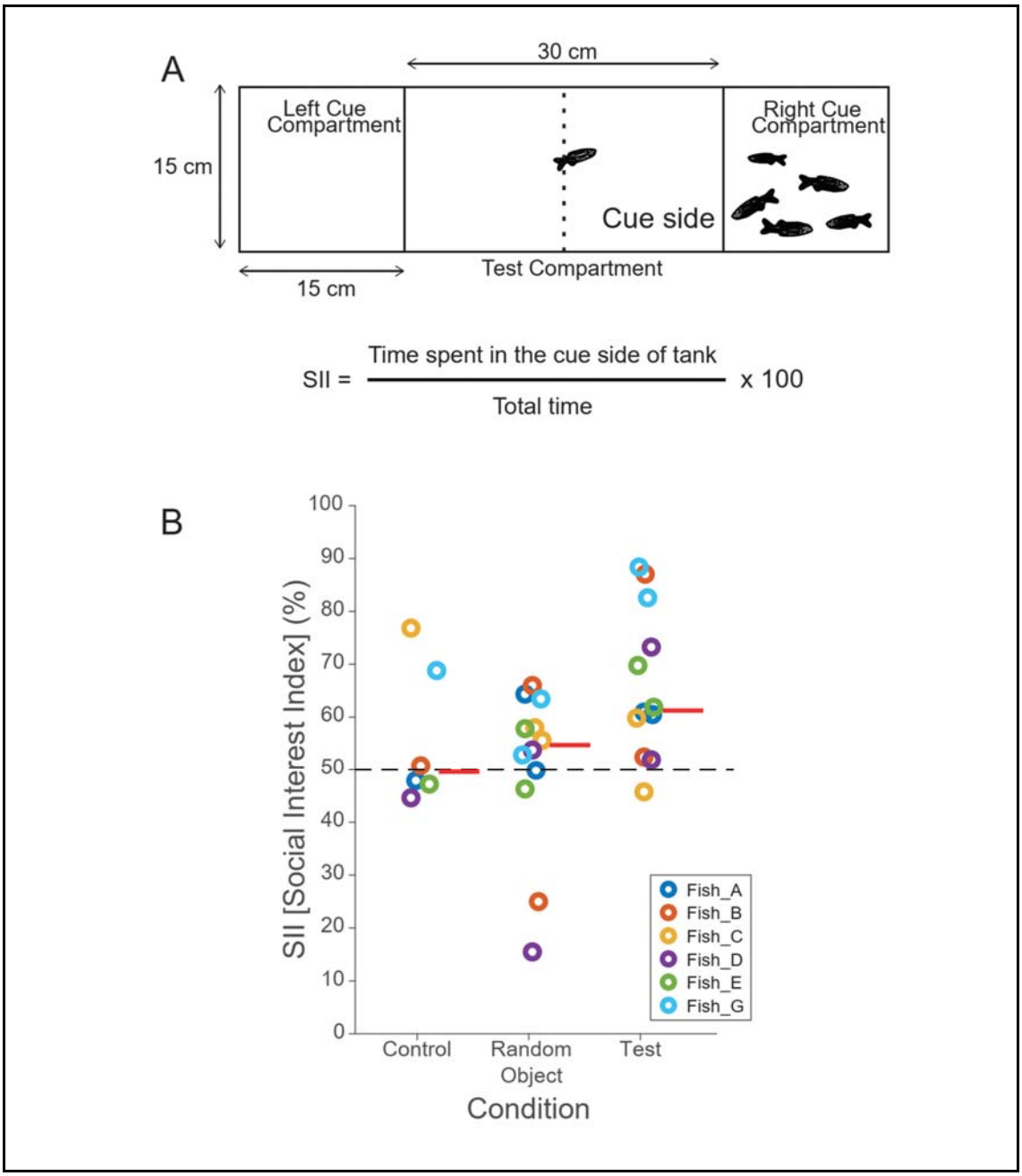

A
30 cm
Left Cue Compartment
Right Cue Compartment
15 cm
Cue side
Test Compartment
15 cm
SII = Time spent in the cue side of tank / Total time x 100
B
SII [Social Interest Index] (%)
100
90
80
70
60
50
40
30
20
10
0
Fish_A
Fish_B
Fish_C
Fish_D
Fish_E
Fish_G
Control
Random Object
Test
Condition

Figure A1.5 - The Three-Compartment Assay, a modified setup to test social choice in adult zebrafish (A) Schematic of the Three-Compartment Assay setup, modified from the T-Maze setup, used to test social choice in adult zebrafish, with dimensions. In this setup, the acclimatisation compartment and test compartment are the same. The test fish is first left to acclimatise in the Test Compartment (middle) and after a period, the cues were introduced by the experimenter into the appropriate Cue Compartments (left and right). A modified index, called the Social Interest Index (SII), was calculated as the fraction of time spent in the right side of the tank. **(B)** SIIs from N=6 fish, colour coded by fish identity. The horizontal dashed line marks an SII of 50%. Five conditions were presented to each fish - Control (no cue in either compartment), RO (random object vs no cue), and Test (no cue vs a social cue). The median for each group is marked with a red horizontal line.

A1.2.6 Adult wild-type and auts2 mutant zebrafish show a strong preference towards social cues

Before testing the social preference of animals, I first sought to further verify that an animal's control response was consistent over days. The reason for testing this was so that I could split the experiment over a two-day period and randomise the order of presentation of the control and the social cue, to prevent an animal's fatigue to the experiment while also further removing biases from the experimental protocol. I found that an animal's response is, indeed, very close to 50% in control conditions (Fig A1.6A) and uncorrelated between the two days (Fig A1.6B), which means that there are no inherent biases in the setup itself that would cause fish to prefer one side over another consistently across days.

Having established a setup that doesn't seem to have inherent biases in it and which also elicited a strong social choice behaviour in a subset of WT animals, I decided to use the three-compartment choice assay to test the social preference of adult zebrafish of both WT and auts2 mutant animals. To do this, I modified the SII index a bit to measure the time spent not in the right half of the setup, but the time spent in the social cue half of the tank. This allowed me to randomise the side in which the cue was presented, to reduce experimental bias.

I found a strong preference to the social cue in WT animals (Fig A1.6C). I tested a total of N=35 fish, of which N=14 were male and N=21 were female. I asked if there was any sexual dimorphism in the behavioural response, and found that there wasn't any - both males and females showed a strong preference to the social cue side (Fig A1.6D).

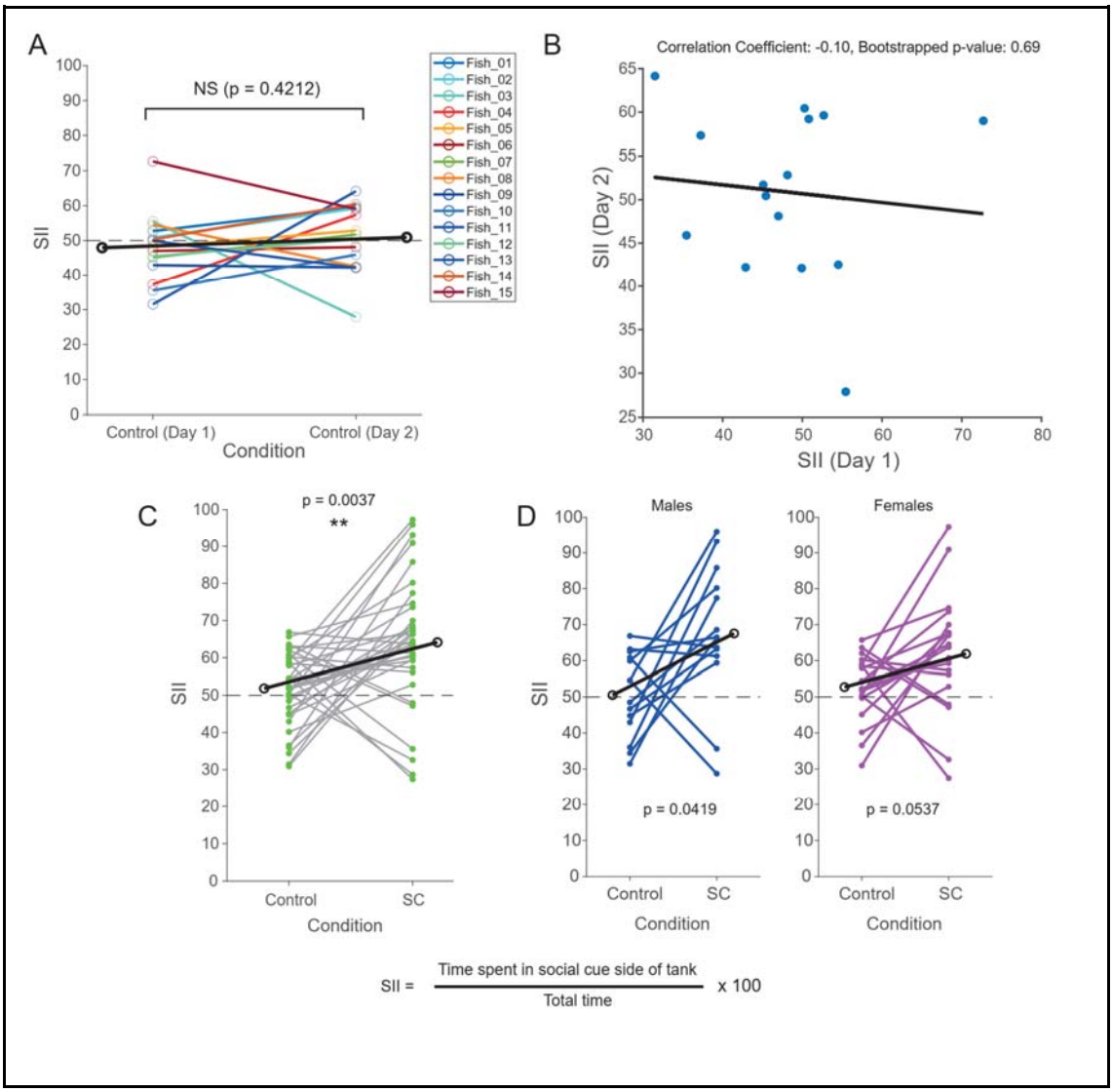

A
NS (p = 0.4212)
SII
Control (Day 1)
Control (Day 2)
Condition
Fish_01
Fish_02
Fish_03
Fish_04
Fish_05
Fish_06
Fish_07
Fish_08
Fish_09
Fish_10
Fish_11
Fish_12
Fish_13
Fish_14
Fish_15
B
Correlation Coefficient: -0.10, Bootstrapped p-value: 0.69
SII (Day 2)
SII (Day 1)
C
p = 0.0037
**
SII
Control
SC
Condition
D
Males
p = 0.0419
Females
p = 0.0537
Control
SC
Condition
SII = Time spent in social cue side of tank / Total time x 100

Figure A1.6 - Social choice in adult Indian WT zebrafish as assessed using the Three-Compartment Assay (A) Day-to-day variability of acclimatisation in N=15 Indian WT fish. Each pair of points connected by a line represents data from one fish, as shown in the legend alongside the plot. The mean values are marked by the black dots and the connecting line. The reported p-value was calculated using the WIlcoxon signed-rank test. **(B)** Cross correlation between the Social Interest Indices (SIIs) from Day 1 and Day 2. The Pearson's correlation coefficient and the significance of this correlation as calculated by bootstrapping are shown in the title. **(C)** Indian WT animals (N=35 fish) spend more time near a presented Social Cue (SC). The Control measurements are for the acclimatisation phase, when no cue is present in either compartment. The Social Interest Index (SII) is calculated here as time spent in the social cue side of the tank. **(D)** SIIs in (C) sorted by sex of the animal - males on the left (N=14 fish) and females on the right (N=21 fish). The mean for each condition in (C) and (D) is shown in the black dots connected by a black line, as in (A). p-values shown in (C) and (D) were calculated using the Wilcoxon signed-rank test.

I next tested the same behaviour in both auts2a-/- (Fig A1.7A) and auts2b-/- (Fig A1.7B) adult zebrafish, and found that animals of both these genotypes also had a strong preference for the presented social cue.

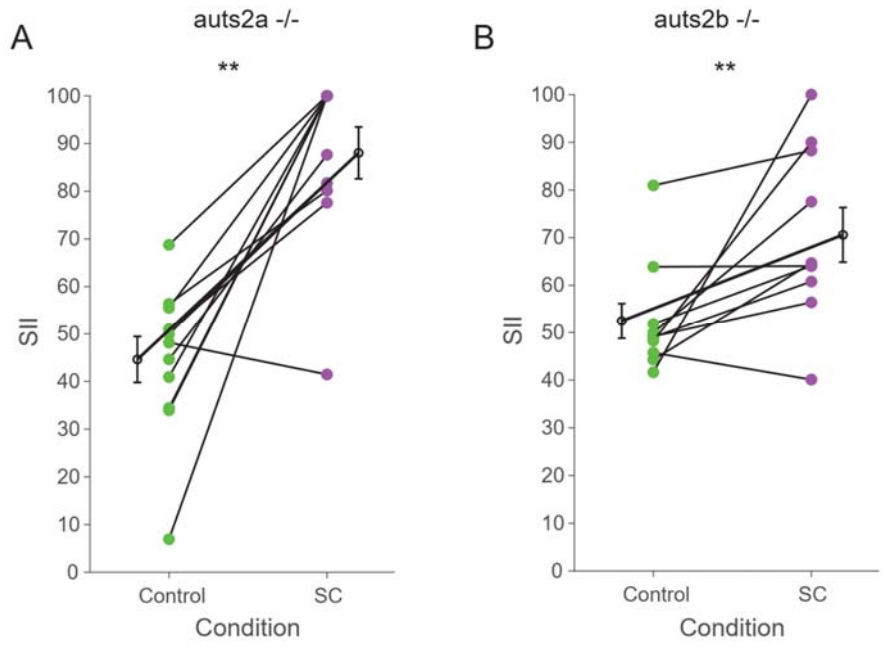


Figure A1.7 - Social choice is intact in both auts2a-/- and auts2b-/- mutants (A) Social choice in auts2a-/- mutants as measured by the SII in the acclimatisation phase (Control) and with the presentation of a social cue (SC) (N=11 fish, p=0.002, Wilcoxon signed-rank test). **(B)** Social choice in auts2b-/- mutants, similar to (A). N=10 fish, p=0.0059, Wilcoxon signed-rank test.

In summary, our reverse genetics study of auts2 showed that there was no detectable effect of auts2 knockout on social choice behaviour.

A1.3 Discussion

Auts2 is a protein that is part of the polycomb repressor complex (Gao et al., 2014) and is present in both the nucleus and the cytoplasm. Its function is to regulate gene expression, the disruption of which leads to autism-like symptoms. A single mutated copy is sufficient to lead to pathology, as there have been patients who present symptoms with a heterozygous genotype.

What I found from this study, however, was that there was no detectable effect of auts2a or auts2b knockout on social behaviour in adult zebrafish. It is possible that the presence of two paralogs of auts2 in zebrafish allows the loss of one to be compensated for by the other, which could be addressed by making double knockout animals and testing social choice in them. It may also be that the defect in social choice in mutants is a developmental one, i.e. the time course of the appearance of the circuits that control social choice is different in mutants. We know from an independent study from the lab (Jha et al., 2021) that auts2a knockout animals have defects in neuronal excitability in the escape circuitry early in their development. Furthermore, I haven't investigated defects in group behaviours like shoaling and schooling in the mutants.

Another interesting avenue to pursue from these results is the seemingly reduced anxiety in auts2b-/- animals. I didn't get to test this in the auts2a-/- animals, so I don't know if a similar effect is seen in them.

auts2 is a part of the polycomb repressor complex. Whether this molecule has an effect on the development of neuronal circuits that control social behaviour may be a more tractable problem in the future, once we have a better handle on which circuits control social behaviour and test whether double knockouts have defective social choice, and whether there are developmental defects, perhaps.

A1.4 Materials and Methods

A1.4.1 Animal care

All experiments were approved by the institutional animal ethics committee (IAEC) and institutional biosafety committee (IBSC). Experiments were performed on zebrafish (*Danio rerio*) of either the Indian wild type (Ind WT) background or the Auts2a or Auts2b mutant background. For this study, only adult animals (between 6 months-1.5 years old) were used, and their sex recorded prior to use for experimentation. Animals were obtained by setting up an incross between adults of the appropriate genetic background and raising them in a ZebTec multi linking system (Tecniplast, Italy) with a pH setting of 7.8, conductivity of 1200µS, and a 14:10h light-dark cycle. As larvae, they were kept in a MIR-154 incubator (Sanyo, Japan) at 28^{o}C in E3 medium (composition in mM: 5 NaCl, 0.17 KCl, 0.33 $CaCl_2$, and 0.33 $MgSO_4$, pH 7.8) with the same day-night cycle. Their medium was routinely replaced, and starting at 5 dpf, they were fed Zeigler Larval diet AP100 (<100 microns) (Pentair AES, FL, United States). From day 15 onwards, they were also fed freshly-hatched brine shrimp (*Artemia*) that is cultured in-house. Adults were also fed brine shrimp and their diet was supplemented with crushed Zeigler adult zebrafish food (Pentair AES, FL, USA). Animals taken for multi-day experiments were kept in separate breeding tanks outside of the multilinking system and were returned to the main system after experiments were complete.

A1.4.2 Social choice assay 1 (Adapted)

For this assay, I made a rectangular tank out of acrylic with dimensions shown in Fig A1.2A. All acrylic used was transparent, except for the thick edges in the schematic in Fig A1.2A, which were made using opaque black sheets of acrylic. Lighting was provided by placing an LCD screen under the tank at maximum brightness and displaying a blank white screen. Recordings were obtained using a PointGrey Flea3-USB3.0 camera (FLIR) at 24fps. First, the tank was filled with water taken from the multilinking system (henceforth referred to as system water), which was not changed between trials with different fish. A test fish was first transferred from its tank into the test compartment and left to acclimatise for a period of 5 minutes (‘Acclim’). After this time, the cues were introduced and the behaviour of the test fish was recorded for a period of 15 minutes (‘Test’). After this period, the test fish was transferred back into its source tank. The social cue provided was a group of 3 fish taken from the same tank, but not used as test fish. They were first kept in a 500ml beaker containing system water and then transferred to the relevant cue compartment using a net. Post-hoc video processing analysis was performed as described below.

A1.4.3 Social choice assay 2 (T-Maze)

For this assay, I made a T-maze tank out of transparent acrylic with dimensions shown in Fig A1.3A. The border between the acclimatisation and test compartments was a sheet of transparent acrylic that had a circular hole of ~3 cm diameter to allow the fish to go through it. A custom lighting panel was built using LED strips (2W) stuck to a cardboard sheet and placed under the tank. Recordings were obtained using a PointGrey Flea3-USB3.0 camera (FLIR) at 10fps. The tank was filled with system water at the start of the session and wasn't replaced between trials. The social cue (a group of 3 fish taken from the same tank but never used as test fish) was placed in the predetermined compartment. Next, a test fish was transferred from its tank into the acclimatisation compartment and left to acclimatise for a period of 5 minutes before video recording was started. Once the fish exited the acclimatisation period, a timer was started and fish were recorded for a duration of 15 minutes. If the test fish didn't exit the compartment, the trial for that fish was aborted. Post-hoc video processing analysis was performed as described below.

A1.4.4 Social choice assay 3 (Three-compartment assay)

This assay used the same setup as the T-maze, but only used the top part, with access between the test and acclimatisation compartments blocked. Video recording was performed in the same way as for the T-Maze assay. The tank was first filled with system water. For the pilot experiments, the test fish was first introduced into the test compartment. After 15 minutes of initial acclimatisation, a cue was introduced into one of the two side compartments and a video recording started for a period of 15 minutes. The cues were presented in the following order - no cue on either side ('Control'), fake plastic tree ('Random Object') in right cue compartment, fake plastic tree ('Random Object') in left cue compartment, 3 fish ('Social cue') in right cue compartment, and finally 3 fish ('Social Cue') in the left cue compartment. The test fish was left in place during the swapping out of cues, and no extra acclimatisation period was provided between trials.

Since fish tended to freeze during the experimental intervention, and to reduce fatigue, I altered the experimental protocol to only test one fish at a time on a day. Furthermore, since the system water wasn't replaced between trials, there could be chemical cues that linger after the initial experiments. To minimise the effects of these extraneous factors on the experimental results, I performed the experiment over two days. On day 1, it was first decided whether each animal would be presented with the 'Control' condition or the 'Social Cue' condition. If it was the 'Social Cue' condition, a random number generator was used to decide whether the cue was going to be in the right or the left cue compartment. After each trial, the test fish were transferred into a small chamber and labelled for use on day 2. On

day 2, the order of use of fish was first randomised, so that the order of introduction of the fish to the tank wasn't the same between days 1 and 2 (this reduces history-dependent cues). If a fish was presented with the 'Control' condition on day 1, it was presented with the 'Social Cue' condition on day 2, and vice versa. Each fish had a 15 minute period of acclimatisation followed by stimulus presentation and another 15 minutes of video recording. Post-hoc video analysis was done as described below.
Experiments with auts2a-/- and auts2b-/- adults were performed exactly as described above, except that the social cue presented were 3 Indian WT fish, since animals used as the social cue were never used as test animals and the number of mutant individuals was limited.

A1.4.5 Video Processing and Analysis

Video processing was done either in MATLAB or using Python's OpenCV module. In order to extract coordinates of the test fish in the social choice assays, background subtraction was performed, either using a blank reference image taken before each experiment began, or using a nearest neighbour background estimation method. After background subtraction, an auto-thresholding method was used to produce a binary image. The largest connected components in the image were taken to be the fish. Once segmented, the fish's centroid was extracted and used as the representative coordinate of the fish's location. Regions of interest (ROIs) were manually drawn on the reference image using the *selectROI* function (Python) or the *drawrectangle* function (MATLAB). For the initial choice assay, two ROIs were drawn per experiment, to define the 'Social Cue' (SC) region and the 'No Social Cue' (no SC) regions. The number of centroids that fell within these ROIs was calculated. For the three-compartment assay, the ROI drawn was that bounded the test compartment alone. The rectangle was split into two halves (right and left) and the number of centroids that fell within each ROI was calculated. Using these numbers, the Social Preference Index (SPI) or the Social Interest Index (SII) was calculated, as described in the corresponding Results section.

A1.4.6 Novel Tank Test

A trapezoidal tank from the multilinking system (Tecniplast, Italy) was used as the novel tank. It was filled to 80% of its maximum capacity and tape was used to mark half the height of the water level. Test animals were first acclimated to the experimental room, water temperature etc. in a 600ml 'pretreatment beaker' (Fig A1.4A). After a period of ~15 minutes, an animal was transferred to the novel tank using a net and the experiment immediately started. A two-camera setup was used to video record the behaviour. Logitech C170 cameras were used, which were triggered together using a custom workflow written in Bonsai (Lopes et al., 2015). Prior to each session, the cameras were calibrated by taking a photo of a calibration grid. This can be used to extract 3D coordinates of the fish during post-

hoc analysis. The fish's behaviour was also recorded manually for a period of 10 minutes. Two key parameters were measured - the time taken for the fish to enter the top half of the water level (latency to top) and the total number of crossings made from the bottom half to the top half in the first 6 minutes of this 10-minute assay.

A1.5 References

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug II, R. G., Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H., Fahrenkrug, S. C., Carlson, D. F., Voytas, D. F., Clark, K. J., Essner, J. J., & Ekker, S. C. (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature*, *491*(7422), Article 7422. https://doi.org/10.1038/nature11537

Cachat, J., Stewart, A., Grossman, L., Gaikwad, S., Kadri, F., Chung, K. M., Wu, N., Wong, K., Roy, S., Suciu, C., Goodspeed, J., Elegante, M., Bartels, B., Elkhayat, S., Tien, D., Tan, J., Denmark, A., Gilder, T., Kyzar, E., … Kalueff, A. V. (2010). Measuring behavioral and endocrine responses to novelty stress in adult zebrafish. *Nature Protocols*, *5*(11), Article 11. https://doi.org/10.1038/nprot.2010.140

Darland, T., & Dowling, J. E. (2001). Behavioral screening for cocaine sensitivity in mutagenized zebrafish. *Proceedings of the National Academy of Sciences*, *98*(20), 11691–11696. https://doi.org/10.1073/pnas.191380698

Dreosti, E., Lopes, G., Kampff, A., & Wilson, S. (2015). Development of social behavior in young zebrafish. *Frontiers in Neural Circuits*, *9*. https://www.frontiersin.org/article/10.3389/fncir.2015.00039

Engeszer, R. E., Ryan, M. J., & Parichy, D. M. (2004). Learned Social Preference in Zebrafish. *Current Biology*, *14*(10), 881–884. https://doi.org/10.1016/j.cub.2004.04.042

Gao, Z., Lee, P., Stafford, J. M., von Schimmelmann, M., Schaefer, A., & Reinberg, D. (2014). An AUTS2–Polycomb complex activates gene expression in the CNS. *Nature*, *516*(7531), Article 7531. https://doi.org/10.1038/nature13921

Jha, U., Kondrychyn, I., Korzh, V., & Thirumalai, V. (2021). High Behavioral Variability Mediated by Altered Neuronal Excitability in auts2 Mutant Zebrafish. *ENeuro*, *8*(5). https://doi.org/10.1523/ENEURO.0493-20.2021

Joung, J. K., & Sander, J. D. (2013). TALENs: A widely applicable technology for targeted genome editing. *Nature Reviews Molecular Cell Biology*, *14*(1), Article 1. https://doi.org/10.1038/nrm3486

Kondrychyn, I., Robra, L., & Thirumalai, V. (2017). Transcriptional Complexity and Distinct Expression Patterns of auts2 Paralogs in Danio rerio. *G3 Genes|Genomes|Genetics*,

7(8), 2577–2593. https://doi.org/10.1534/g3.117.042622

Lopes, G., Bonacchi, N., Frazão, J., Neto, J. P., Atallah, B. V., Soares, S., Moreira, L., Matias, S., Itskov, P. M., Correia, P. A., Medina, R. E., Calcaterra, L., Dreosti, E., Paton, J. J., & Kampff, A. R. (2015). Bonsai: An event-based framework for processing and controlling data streams. *Frontiers in Neuroinformatics*, *9*. https://www.frontiersin.org/articles/10.3389/fninf.2015.00007

Miller, N., & Gerlai, R. (2012). From Schooling to Shoaling: Patterns of Collective Motion in Zebrafish (Danio rerio). *PLOS ONE*, *7*(11), e48865. https://doi.org/10.1371/journal.pone.0048865

Norton, W., & Bally-Cuif, L. (2010). Adult zebrafish as a model organism for behavioural genetics. *BMC Neuroscience*, *11*(1), 90. https://doi.org/10.1186/1471-2202-11-90

Portfors, C. V., & Perkel, D. J. (2014). The role of ultrasonic vocalizations in mouse communication. *Current Opinion in Neurobiology*, *28*, 115–120. https://doi.org/10.1016/j.conb.2014.07.002

Sultana, R., Yu, C.-E., Yu, J., Munson, J., Chen, D., Hua, W., Estes, A., Cortes, F., de la Barra, F., Yu, D., Haider, S. T., Trask, B. J., Green, E. D., Raskind, W. H., Disteche, C. M., Wijsman, E., Dawson, G., Storm, D. R., Schellenberg, G. D., & Villacres, E. C. (2002). Identification of a Novel Gene on Chromosome 7q11.2 Interrupted by a Translocation Breakpoint in a Pair of Autistic Twins. *Genomics*, *80*(2), 129–134. https://doi.org/10.1006/geno.2002.6810

Wendt, K. D., Lei, B., Schachtman, T. R., Tullis, G. E., Ibe, M. E., & Katz, M. L. (2005). Behavioral assessment in mouse models of neuronal ceroid lipofuscinosis using a light-cued T-maze. *Behavioural Brain Research*, *161*(2), 175–182. https://doi.org/10.1016/j.bbr.2005.02.024

Appendix 2 - Functional Dissection of Mixed Synapses in the Escape Circuit of Larval Zebrafish

A2.1 Introduction

Neurons, the functional unit of electrical signalling in the brain, communicate with one another rapidly and with great precision. This communication occurs at specialised sites between two neurons, called synapses, which come in three varieties – electrical, chemical and mixed (Pereda, 2014). Electrical synapses are neuronal gap junctions, which form a cytoplasmic bridge between the connected neurons, allowing the free passage of ions and small molecules, such as secondary messengers. Chemical synapses, on the other hand, pass on signals using a chemical transmitter and a specialised postsynaptic detector. A third kind of synapse is the mixed synapse, which has both components between the same pair of neurons (Pereda, 2014). These connections are commonly found in many neural circuits, including (but not limited to) the stomatogastric nervous system of crustaceans (Johnson et al., 1993), motor circuits of *C. elegans* (Liu et al., 2017), and vertebrate spinal circuits (Rash et al., 1996; Serrano-Velez et al., 2014).

A particular circuit of interest where mixed synapses are found is the escape circuit of larval zebrafish. ‘Escape’ is a movement away from an imminent threat, and fast escape behaviour is essential for the survival of an organism. Larval zebrafish demonstrate a fast startle response to tactile, auditory or visual stimuli, with an initiation latency of 3-12 ms (Kohashi & Oda, 2008) (Fig A2.1A). This behavioural response is mediated by a set of neurons conserved across teleosts that originates in the hindbrain and sends long axonal projections to the spinal cord, called reticulospinal neurons (Burgess & Granato, 2007; Kohashi & Oda, 2008; Korn & Faber, 2005) (Fig A2.1B).

The main neurons of this circuit are a pair of large, bilateral neurons called the Mauthner neurons, often shortened to “M-cells” (Fig A2.1B and Fig A2.1C). These reticulospinal neurons, which are located in rhombomere 4 of the developing hindbrain, receive multimodal sensory input, and cross the midline to innervate motor neurons along the entire body axis of the larva (Fetcho, 1991). A single action potential fired in one of these neurons leads to the contraction of the contralateral muscles, leading to a C-shaped bend of the body away from the direction of the noxious stimulus, called a C-start (Eaton & Nissanov, 1985) (Fig A2.1A). Furthermore, Mauthner neurons receive mixed synaptic input from the 8th cranial nerve (Fig A2.1C), which transmits auditory information.

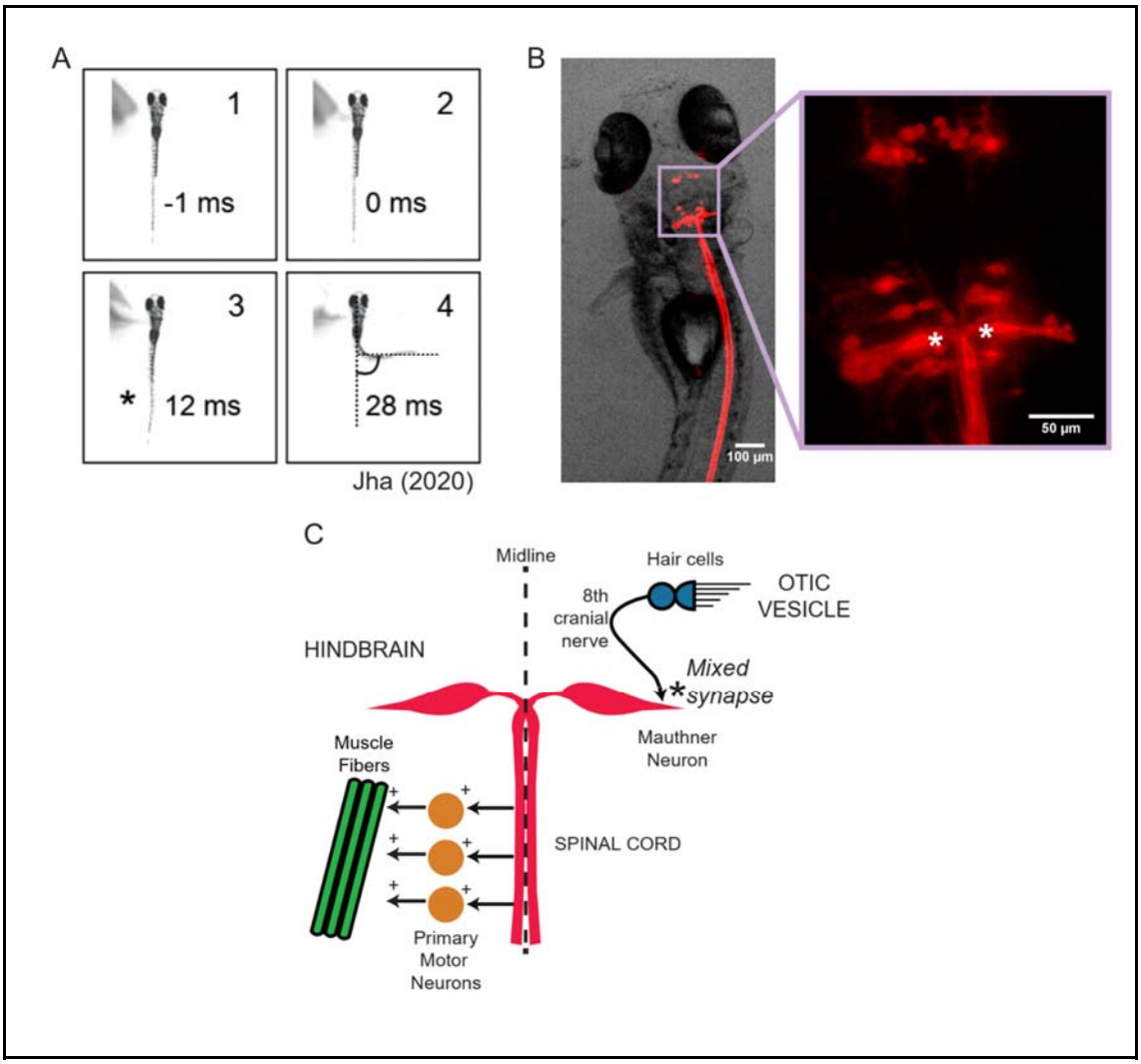

A
1
-1 ms
2
0 ms
3
*
12 ms
4
28 ms
Jha (2020)
B
100 µm
*
*
50 µm
C
Midline
Hair cells
OTIC
VESICLE
8th
cranial
nerve
HINDBRAIN
Mixed
*synapse
Mauthner
Neuron
Muscle
Fibers
SPINAL CORD
Primary
Motor
Neurons

Figure A2.1 - Escape behaviour in larval zebrafish and the circuit controlling it. (A) A montage from a high speed video recording of escape behaviour elicited in freely swimming larval zebrafish by a pulse of water delivered to the otic vesicle. The time stamps under each frame represent the time relative to the arrival of the water pulse at the otic vesicle. The frame with an asterisk marks the initiation of the escape response, which has occurred in 4ms in this case. The frame with the maximum bend angle is marked with a double asterisk. This montage was taken from a video obtained in the lab and reported in Jha (2020). **(B)** Reticulospinal neurons in the larval zebrafish, labelled by pressure injecting TMR-Dextran into the spinal cord. Note the long axons that span the length of the body of the animal. A region of the hindbrain is shown at a higher magnification in the inset to highlight the cell bodies of the reticulospinal neurons. Mauthner neurons are marked with a white asterisk. **(C)** Schematic of the neural circuit controlling the startle response. An auditory stimulus is detected by hair cells of the otic vesicle. This information is passed along the 8th cranial nerve to the Mauthner neuron via a mixed synapse, which is marked with an asterisk. If the Mauthner neuron hits threshold, an action potential is transmitted along its axon, which crosses the midline and activates primary motor neurons (and other interneurons not shown in the schematic) along the length of the animal. This, in turn, leads to contraction of muscle fibres on the side opposite to the stimulus, thereby leading to a turn away from the threat.

Despite their widespread occurrence, how electrical and chemical synapses interact at mixed synapses to produce function remains to be elucidated. Electrical synapses are thought to contribute speed, while chemical synapses are thought to contribute plasticity. However, this is too simplistic an explanation, as chemical synapses act fast relative to behavioural timescales, and because electrical synapses are also plastic (Pereda & Faber, 1996; Yang et al., 1990).

Thus, I decided to address the question of what each of the two parts of a mixed synapse does by asking how much of each type of synaptic input is needed to initiate escape behaviour. The escape circuit of the larval zebrafish is ideal to address this question because it is an ethologically vital circuit that has a mixed synapse at the rate limiting step of circuit activation.

The relative contribution of electrical and chemical synaptic conductances to M-cell firing was assessed using simulations. I also came up with a strategy to assess this *in vivo* using a technique called dynamic clamp, which allows the experimenter to artificially introduce ionic

conductances, such as synapses, of arbitrary strength and kinetics (Prinz et al., 2004; Sharp et al., 1993) (Fig A2.5).

A2.2 Results

A2.2.1 In silico modelling suggests that electrical synapses could be sufficient to make M-cells fire

One of the most flexible ways to study the effects of perturbations to individual units of a synapse is using *in silico* modelling. Hence, I first modelled the mixed synapse between the 8th cranial nerve and the M-cell (Fig A2.2A). I adapted a publicly available conductance-based model of zebrafish M-cells from ModelDB (Watanabe et al., 2017). This model itself was a modified version of a mouse cochlear neuron model, but accounted for various conductances that give M-cells their unique firing properties, such as the ability to only fire single action potentials in response to suprathreshold current injections (Watanabe et al., 2017). The model used was a single-compartment model, with parameters taken from (Koyama et al., 2016) and (Watanabe et al., 2017), wherever relevant. Table A2.1 summarises the parameters I used and also quotes values from both these papers for comparison.

Table A2.1 - Summary of parameters used for M-cell modelling.

Parameter [units]	**Mauthner Cell**			**8th cranial nerve**		
	My simulations	Koyama et al. (eLife, 2016)	Watanabe et al. (eNeuro, 2017)	My simulations	Koyama et al. (eLife, 2016)	Watanabe et al. (eNeuro, 2017)
Total Capacitance (pF)	~20	NA (LIF Model)	40	~20	20-200	NA
E_leak (mV)	-85		-85	-79	-79	
E_K (mV)	-90		-90	-90	-90	
E_Na (mV)	50*		30	50*	50*	
g_leak (S/cm^2)	12		12	0.0003 *	0.00005	
g_Na (S/cm^2)	6000		6000	0.120 *	0.1	
g_K (S/cm^2)	1500		1500	0.036 *	0.2	
g_K_A (S/cm^2)	300		300	NA	NA	

g_KCNQ (S/cm^2)	1400		1400			
g_{elec} (nS)	**1-1000**	15	NA			
g_{chem} (nS)	**1-100**	12				
t_{chem} (ms)	2	2				
Chemical Synaptic Delay (ms)	0.7	0.7				

* These are the default values for a new section in NEURON.

Parameters in bold were the ones varied over simulations.

Electrical synapses were modelled as a resistor connecting the two neuronal compartments, with a conductance of *g_el* (Fig A2.2A). Chemical synapses were modelled as a conditional conductance (i.e. one that is activated when a certain condition is met) in the postsynaptic compartment, using the ExpSyn module in NEURON. This module adds a conductance *g_chem* to the postsynapse a certain delay after the presynaptic neuron reaches threshold (Fig A2.2B), which decays back to zero with a certain decay time constant, tau (Fig A2.2B).

Three kinds of simulations were performed, corresponding to the three types of synapses - only electrical, only chemical, or both. For simulations with electrical synapses, the conductance *g_el* was varied. For simulations with chemical synapses, the delay and decay time constants were fixed (Table A2.2) and only the maximal conductance *g_chem* was taken as a free parameter. I sampled both conductances over several orders of magnitude (1-1000nS for *g_el* and 1-100nS for *g_chem*). Very low synaptic conductance values mean that the postsynaptic neuron is unable to fire (Fig A2.2C, left column, “Subthreshold”). Naturally, increasing the conductance meant the postsynaptic neuron could fire. The lowest

sampled conductance value at which the postsynaptic neuron fired was the “Threshold” (Fig A2.2C, middle column, “Around threshold”). Increasing conductance beyond this value also led to neuronal firing (Fig A2.2C, right column).

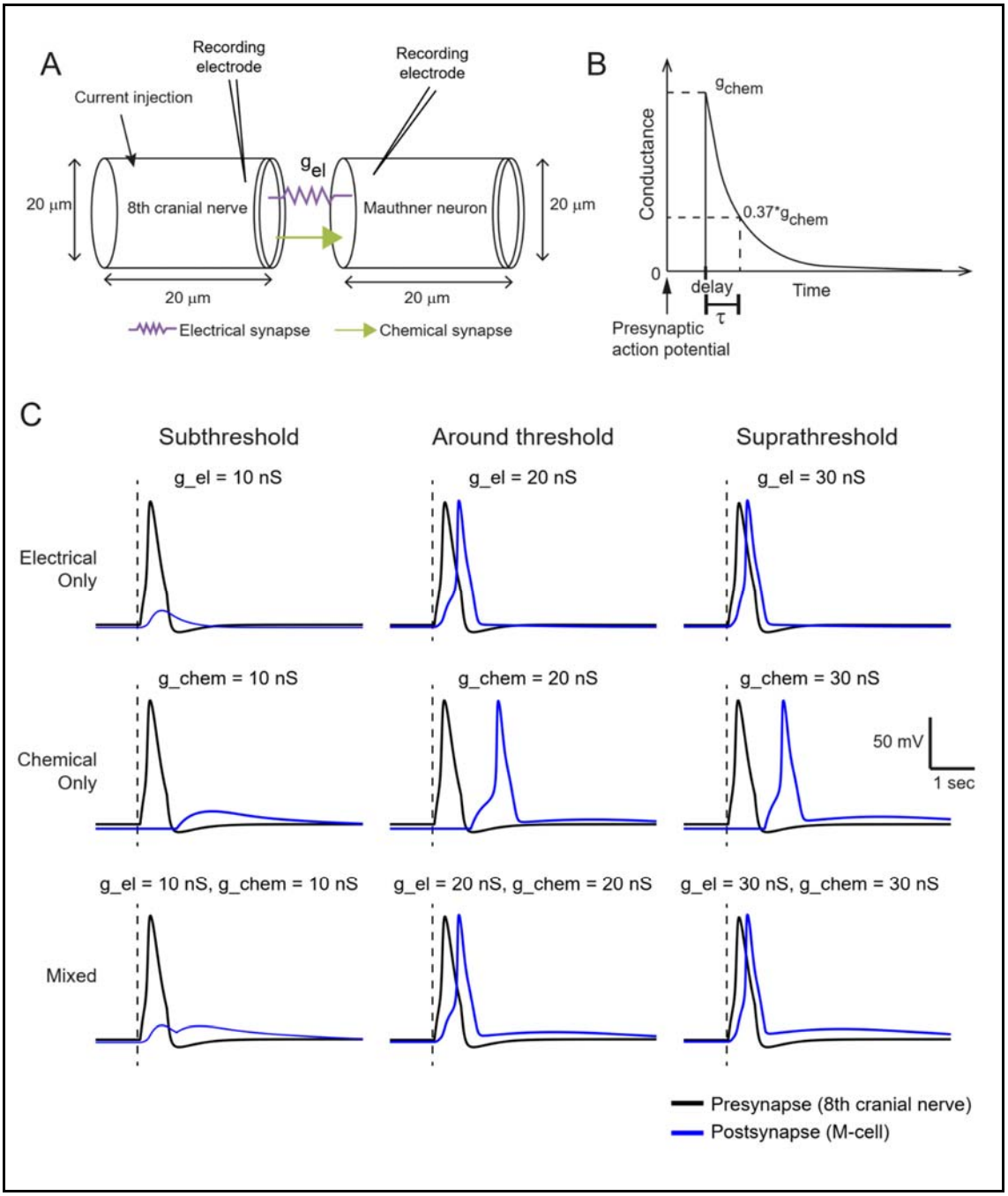


Figure A2.2 - *In silico* modelling of the effect of mixed synapses at the 8th cranial nerve-Mauthner neuron connection. (A) Schematic of the single compartment models of the 8th cranial nerve and the Mauthner neuron connected by a bidirectional electrical synapse and a chemical synapse. The electrical synapse is a resistor between the new neurons, with a conductance of g_el. The chemical synapse is a conditional conductance introduced into the postsynaptic neuron, the properties and profile of which are shown in (B). An ideal current injection electrode is placed in the 8th cranial nerve compartment, and an ideal recording electrode is placed in both compartments. **(B)** The conductance profile of an ExpSyn unit of a chemical synapse. It has three parameters - delay (the time between the presynaptic action potential and the activation of this conductance), g_chem (the maximal conductance) and τ (the conductance decay time constant). **(C)** Pre- (black) and postsynaptic (blue) voltage profiles for various combinations of electrical and chemical synaptic conductances. Each row represents a different combination of added synaptic conductances - only electrical synapses present (top), only chemical synapses present (middle), and both present (bottom). Each column represents a particular value of the synaptic conductance - those that didn't lead to M-cell firing ("Subthreshold", left), the lowest sampled value that led to M-cell firing ("Around threshold", middle) and values higher than threshold ("Suprathreshold", right). These three were calculated only for the "Electrical only" and "Chemical only" cases, and the traces shown for the "Mixed" case used the values obtained for each of these two cases. The vertical dashed line represents the time when a suprathreshold current was injected into the presynaptic neuron.

I found these three conductances - subthreshold, threshold and suprathreshold - only for simulations with one synaptic conductance. I then asked whether electrical and chemical synapses could work together to bring a cell to threshold even though individually both conductances were subthreshold. I found that this was not the case, and that the combination of two subthreshold synaptic conductances did not make the M-cell fire (Fig A2.2C, bottom row). Moreover, when electrical synaptic conductances were at threshold/suprathreshold values (Fig A2.2C, bottom row, middle and right columns), the latency to firing was shorter than for chemical synapses alone (Fig A2.2C, middle row, middle and right columns). This is also as one would expect, since chemical synapses have an inherent delay built into their mechanism of action.

Fig 2.3 shows traces from a range of conductance values sampled for both electrical and chemical synapses. These were sampled at a finer grain (2nS steps) than the initial broad sampling, which spanned several orders of magnitude. I found the threshold conductance

values for electrical and chemical synapses to be 16nS and 12nS, respectively. I compared the conductance values obtained in these three cases against reference values from literature (Koyama et al., 2016) (Table A2.2) and found that physiologically relevant electrical and chemical synaptic conductances (15nS and 12nS) were very close to threshold values in my simulations (16nS and 14nS).

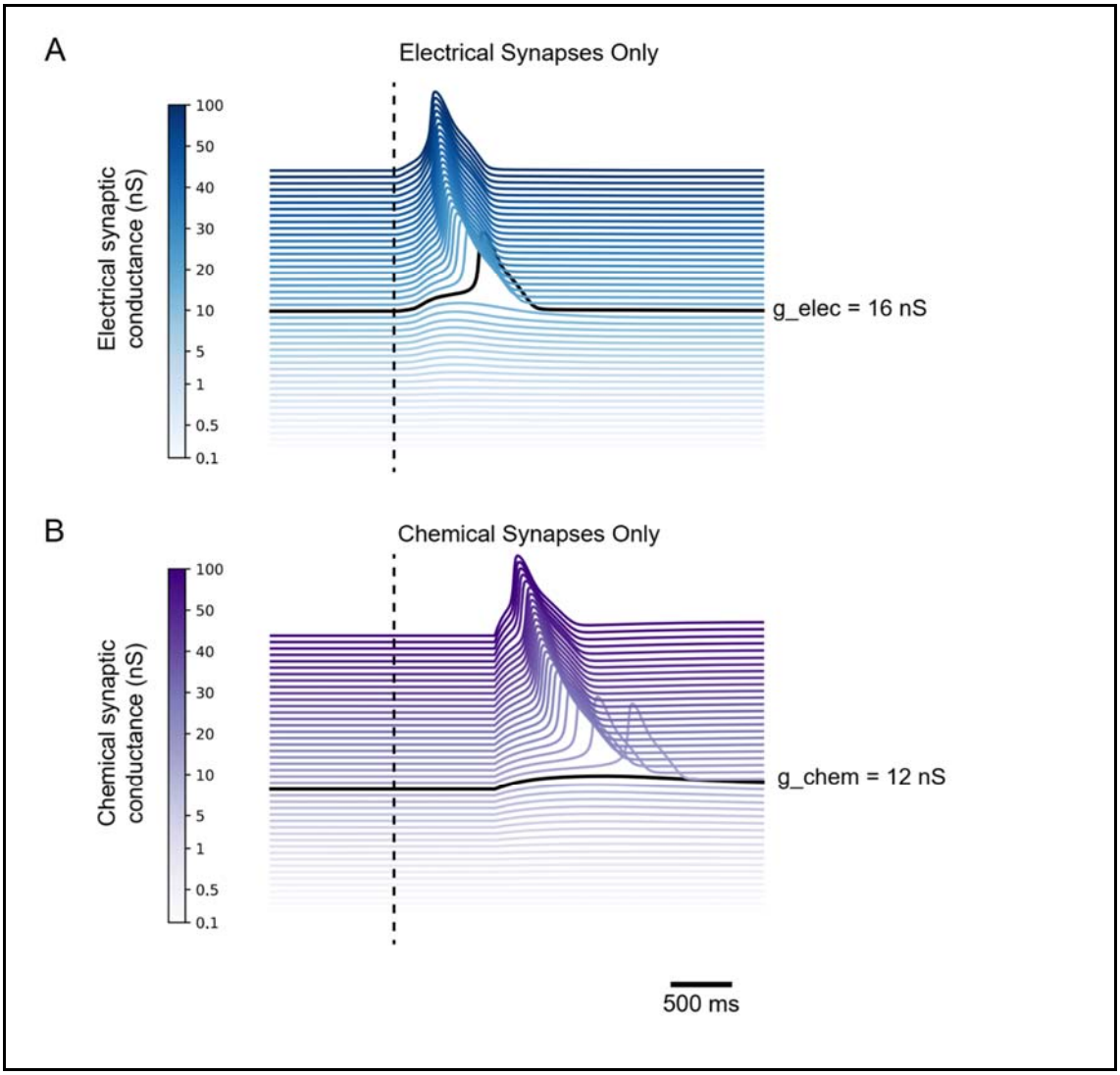


Figure A2.3 - Increasing synaptic conductances beyond threshold reduces postsynaptic firing latency. (A) Voltage traces from simulations where only electrical synapses were present between the two modelled compartments. Traces are sorted in order of synaptic strength and each trace is colour coded by the unique electrical synaptic conductance used for its simulation (see adjacent colour bar). All traces are aligned to the same time, with the vertical dashed line representing the time at which a suprathreshold current was injected into the presynaptic compartment. The trace marked in black represents the sampled conductance value that was closest to the reported physiological conductance. **(B)** Aligned voltage traces, as in (A), but for simulations with only chemical synapses. The black trace represents the chemical synaptic conductance closest to the reported physiological conductance.

In sum, *in silico* models showed that physiological levels of the electrical synaptic conductance were sufficient to get the M-cell to fire, and that in these cases, electrical synapses dominated over chemical synapses in terms of latency to M-cell firing. Thus, it seems that electrical synapses do, indeed, make the M-cell circuit respond fast. What the function of chemical synapses in this circuit, then, is unclear.

A2.2.2 Electrical synapses can slow down chemical synaptic transmission at a mixed synapse

Having looked at the firing profiles in the “electrical only” and “chemical only” cases, I decided to take a closer look at the M-cell firing profile for various combinations of electrical and chemical synaptic conductances (Fig A2.4A), in order to ask if the two were really completely independent of one another, or if they could interact. By and large, these simulations suggest that they act independent of each other. However, there is a small regime that is close to physiologically relevant values of these conductances that had an interesting effect.

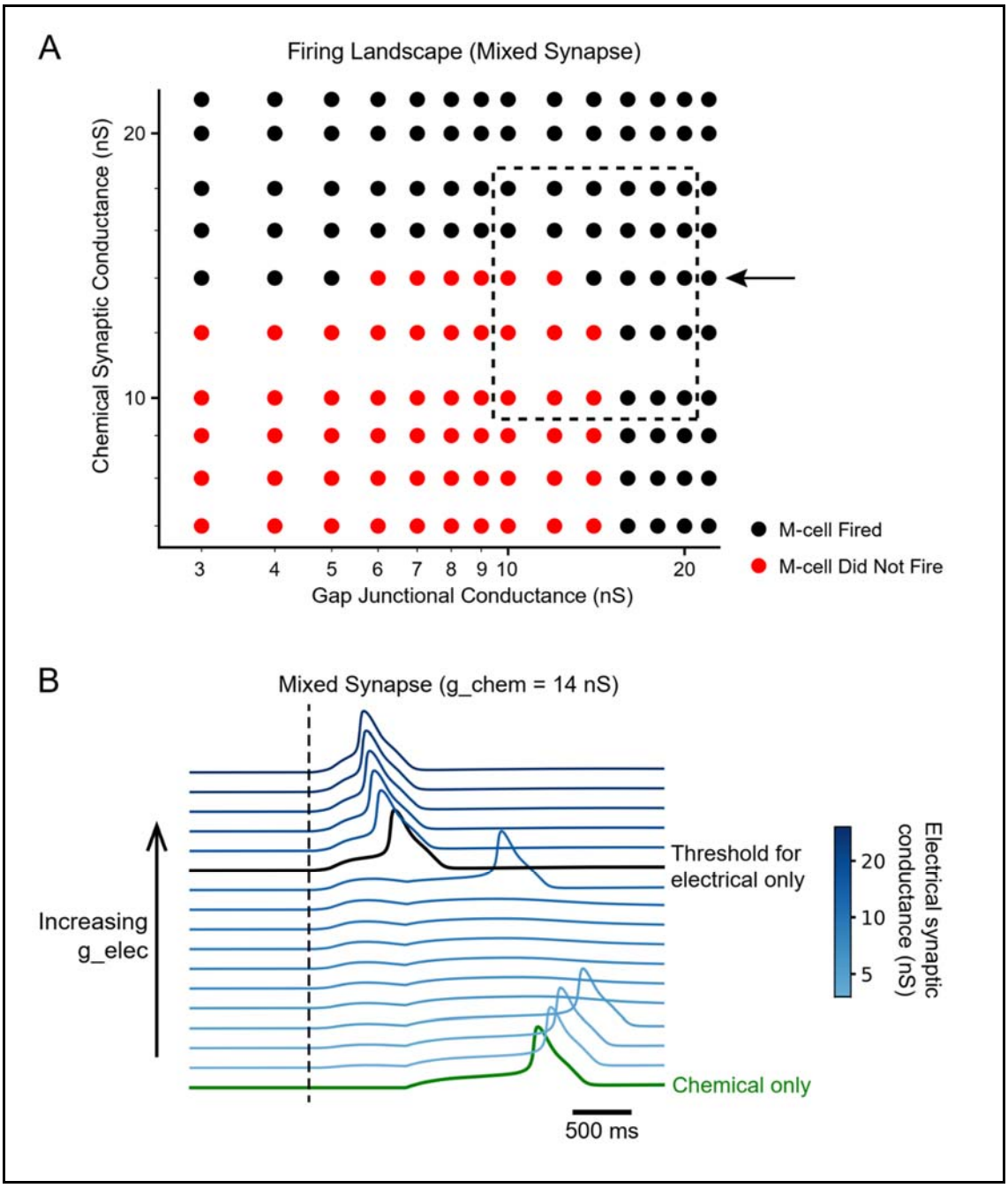

A
Firing Landscape (Mixed Synapse)
Chemical Synaptic Conductance (nS)
20
10
3
4
5
6
7
8
9
10
20
Gap Junctional Conductance (nS)
M-cell Fired
M-cell Did Not Fire
B
Mixed Synapse (g_chem = 14 nS)
Increasing
g_elec
Threshold for
electrical only
Chemical only
500 ms
Electrical synaptic
conductance (nS)
20
10
5

Figure A2.4 - Electrical synapses can slow down chemical synaptic activity in a physiologically relevant conductance regime. (A) Firing profile of M-cells in simulations with mixed synapses as a function of the combination of electrical (x-axis) and chemical (y-axis) synaptic conductances. Black dots mean that the M-cell fired an action potential in the simulation, whereas red dots indicate that it did not. The box with the dotted line marks physiologically relevant conductance regimes for both conductances. The arrow marks the row which is particularly interesting, corresponding to a maximal chemical synaptic conductance of 14 nS, and whose details are shown in (B). **(B)** Aligned voltage traces from M-cells with mixed synapses. In all these traces, the chemical synaptic conductance is fixed at 14 nS. The gap junctional conductance increases from bottom to top and traces are colour coded by this value (see adjacent colour bar). For reference, the bottommost (green) trace has no electrical synapses, and the black trace marks the threshold conductance value where electrical synapses alone led to M-cell firing. The vertical dashed line, as in Figs A2.3 and A2.4, represents the time at which a suprathreshold current was injected into the presynaptic neuron.

At near-threshold values of chemical synaptic conductances, I saw that subthreshold electrical synaptic conductances had conflicting effects. A value of *g_el* that was just under threshold sped up the action of the chemical synapse (Fig A2.4B, trace just below the black trace). If the value of *g_el* was lower, but non-zero, the action of the chemical synapse was slowed down or even completely inhibited (Fig A2.4B, traces between green and black traces). Hence, the action of the electrical synapse in the physiological range of conductances is not straightforward to predict. This is what made me want to ask what exactly happens *in vivo*.

A2.2.3 A dynamic clamp-based strategy to assess the interaction between electrical and chemical synaptic components of the mixed synapse *in vivo*

These *in silico* simulations have a range of limitations. Firstly, my model of the 8th cranial nerve is imperfect, as its structure isn't identical to that of the M-cell, and because I modelled it as having only Hodgkin-Huxley conductances, which may not be the case *in vivo*. Furthermore, I have only modelled this pair of connections, but the M-cells are actually part of a much larger network with many feedforward and feedback connections (Korn & Faber, 2005). The simulations I performed were also deterministic and accounted for no variation in synaptic transmission due to stochastic fluctuations or synaptic noise. Thus, any interaction between the two components of a mixed synapse needed to be addressed *in vivo*. To do so,

I came up with a dynamic clamp-based strategy to control each component of a mixed synapse (Fig A2.5A and Fig A2.5B).

Dynamic clamp is a powerful way of studying the responses of real biological samples to the addition/removal of conductances (ref). Briefly, a time- and voltage-dependent biophysical model of a conductance is computed. The amount of current that goes through it at any given point in time is computed from a real recording, and the appropriate amount is injected into the biological sample (Fig A2.5A). This cycle is repeated and is performed at a high rate. In essence, dynamic clamp is simply a current injection protocol with a profile that is not uniform, but rather dependent on the real-time voltage recorded from a neuron.

I decided to use this strategy to model electrical and chemical synaptic conductances and inject them into real Mauthner neurons in zebrafish to ask which combination of values leads to M-cell firing. The method would allow me to selectively perturb one (Fig A2.5B, middle trace) or both (Fig A2.5B, bottom trace) components of the mixed synaptic conductance in a live neuron. It would also allow me to make more graded perturbations, unlike what can be achieved using knockouts.

I established dynamic clamp using a NI-DAQ PCIe6361 card and the StdpC software (Fig A2.5A) (Kemenes et al., 2011). Using this hardware and software, I was able to inject Hodgkin-Huxley conductances for voltage-dependent Na^+ and K^+ channels into the PATCH 1-U Model Cell (Axon Instruments), which only has a passive RC circuit (Fig A2.5C, black trace). With dynamic clamp turned on, currents could be injected that led to action potential-like responses in this passive circuit (Fig A2.5C, blue trace). I also tried this with real neurons. I first obtained a whole-cell patch clamp recording from unidentified neurons in the hindbrain (n=11 neurons) in a ventral-up preparation. For a subset of these neurons, I injected a 100pA current with and without dynamic clamp turned on. Representative traces are shown in Fig A2.5D. Note that the waveform of the action potentials in the neuron are altered with dynamic clamp on (Fig A2.5D, blue trace) when compared to control (Fig A2.5D, black trace). In other words, I have a proof-of-principle showing that I could inject artificial conductances into real neurons. I could thus proceed to record from and perform the proposed dynamic clamp experiment in Mauthner neurons.

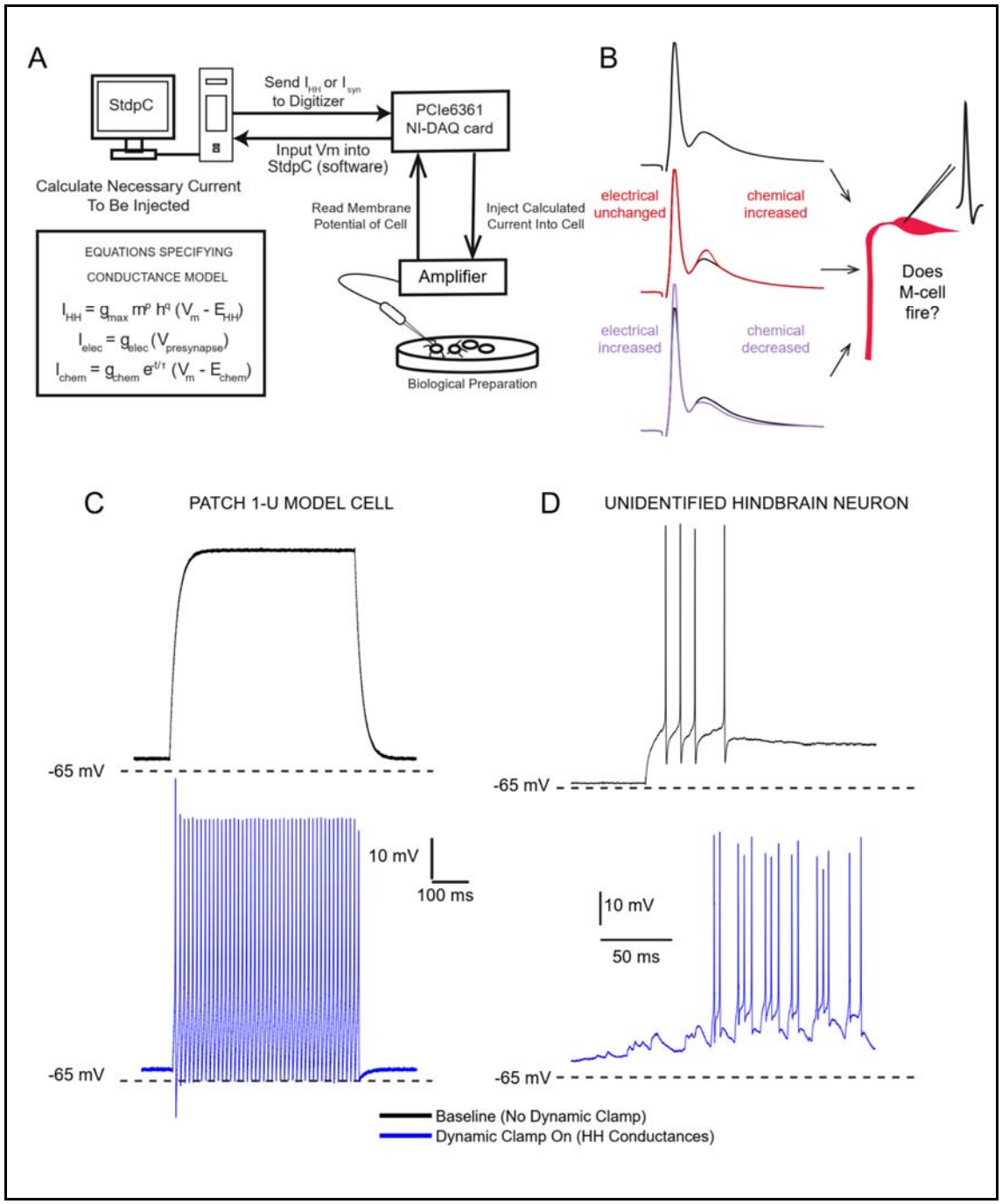

A
StdpC
Send I_{HH} or I_{syn} to Digitizer
PCIe6361 NI-DAQ card
Input Vm into StdpC (software)
Calculate Necessary Current To Be Injected
Read Membrane Potential of Cell
Inject Calculated Current Into Cell
Amplifier
Biological Preparation
EQUATIONS SPECIFYING CONDUCTANCE MODEL
$I_{HH} = g_{max} m^p h^q (V_m - E_{HH})$
$I_{elec} = g_{elec} (V_{presynapse})$
$I_{chem} = g_{chem} e^{t/\tau} (V_m - E_{chem})$
B
electrical unchanged
chemical increased
electrical increased
chemical decreased
Does M-cell fire?
C
PATCH 1-U MODEL CELL
-65 mV
10 mV
100 ms
-65 mV
D
UNIDENTIFIED HINDBRAIN NEURON
-65 mV
10 mV
50 ms
-65 mV
Baseline (No Dynamic Clamp)
Dynamic Clamp On (HH Conductances)

Figure A2.5 - A dynamic clamp-based strategy to test the relative contributions of each component of a mixed synapse. (A) Dynamic clamp set-up and calculations. In this method, the membrane potential of a biological preparation is constantly monitored and the currents that would flow through voltage-dependent conductances are calculated (*box*). These currents are then injected back into the sample to simulate their effects *in vivo*. **(B)** Idea behind specifically altering one or both components of a mixed synaptic response (*black trace, top*) - the first, faster component is obtained by electrical synaptic transmission, whereas the slower component is chemical synaptic transmission. Using dynamic clamp, individual synaptic conductances can be specifically altered. For example, the chemical synaptic conductance can be increased without altering the electrical synaptic conductance (*red trace*). Both components can also be altered simultaneously - for example, the electrical synaptic conductance can be increased while also reducing the chemical synaptic conductance (*purple trace*). In each of these cases, the readout would be whether or not the M-cell fired (*right*). **(C, D)** Proof of concept that the setup in (A) can be used to dynamic clamp voltage-dependent Hodgkin-Huxley conductances into (C) a passive electrical RC circuit - the PATCH 1-U model cell, and (D) a neuron from the hindbrain of larval zebrafish.

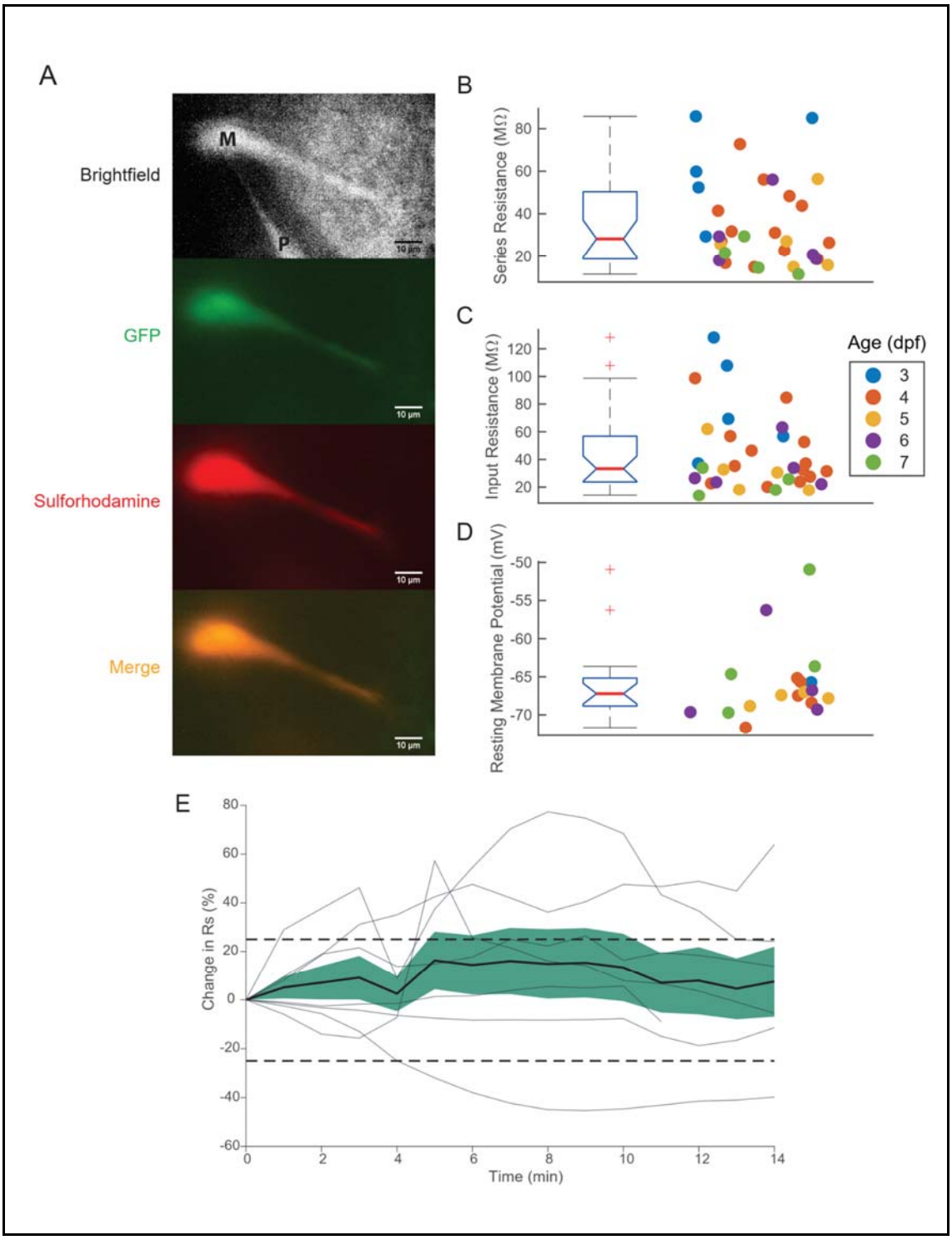

A
Brightfield
GFP
Sulforhodamine
Merge
M
P
10 µm
B
Series Resistance (MΩ)
C
Input Resistance (MΩ)
Age (dpf)
3
4
5
6
7
D
Resting Membrane Potential (mV)
E
Change in Rs (%)
Time (min)

Figure A2.6 - Verification of Mauthner neuron recordings from larval zebrafish. (A) Confirmation of whole cell patch clamp recordings from a GFP-expressing Mauthner neuron from a hspGFF62A larva. The brightfield image (top) is marked with the outline of the Mauthner cell (M) and the patch pipette (P). The remaining images show the individual channels - from top to bottom: transgenic GFP expression in the M-cell, sulforhodamine from the internal, and the merge showing the overlap of both channels. **(B-D)** Distribution of the various properties from the recorded cells, shown as a box plot with data from individual cells shown alongside it. The data points are colour coded by the age of the larva recorded from. Data for series and input resistance comes from n=32 cells, whereas resting membrane potential measurements come from n=18 cells **(E)** Change in series resistance over time, measured for a subset of cells. The mean change is marked in the black line, with the green shaded region marking the s.e.m. The grey lines represent data from individual cells. The horizontal dashed lines mark an absolute change in series resistance of 25%. Data from n=7 cells from N=5 larvae.

Table A2.2 - Summary of recordings from M-cells. Total number of cells recorded from, sorted by age group and genotype.

	Genotype		
Age (dpf)	**62A**	**Tol-056**	**Grand Total**
3	3	2	5
4	13	1	14
5	6	2	8
6	5	0	5
7	0	4	4
Grand Total	**27**	**9**	**36**

A2.2.4 Whole-cell patch clamp recordings from Mauthner neurons in larval zebrafish

I obtained whole-cell patch clamp recordings from Mauthner neurons at various developmental stages of the larval zebrafish (Fig A2.6A). Table A2.2 summarises the dataset of all recordings obtained from M-cells.

The critical factors for obtaining these recordings were as follows. First was the use of transgenic lines expressing GFP in Mauthner neurons (either hspGFF62A/Tol-026/Tol-056). Next was the development of a ventral-up preparation (as opposed to the dorsal-up preparation normally used). I also used thin-walled capillaries to make patch pipettes, whose wider inner channel allowed for lower resistances and a better tapered shape suitable for these large neurons. Lastly, I used a combination of visual guidance and a drop in open pipette resistance of 0.5Megohms to get higher gigaohm seal formation rates.

I calculated several intrinsic properties, such as membrane resistance (Fig A2.6C) and resting membrane potential (Fig A2.6D) for these cells, which matched the range of reported values in literature. I obtained slightly higher series resistances than typically reported in literature. Nevertheless, I could stably for a period of ~15 minutes per cell with changes in series resistance within 25%, on average (Fig A2.6E). I could also recapitulate the change in firing properties over development (Fig A2.7). Mauthner neurons in young animals (3-4dpf) fire multiple action potentials in response to current injections beyond the rheobase, and start firing fewer action potentials at later days, until they exclusively fire single action potentials (Watanabe et al., 2014). I saw the same trend in my data (Fig A2.8B). The rheobase of these neurons did not change much, though (Fig A2.8A).

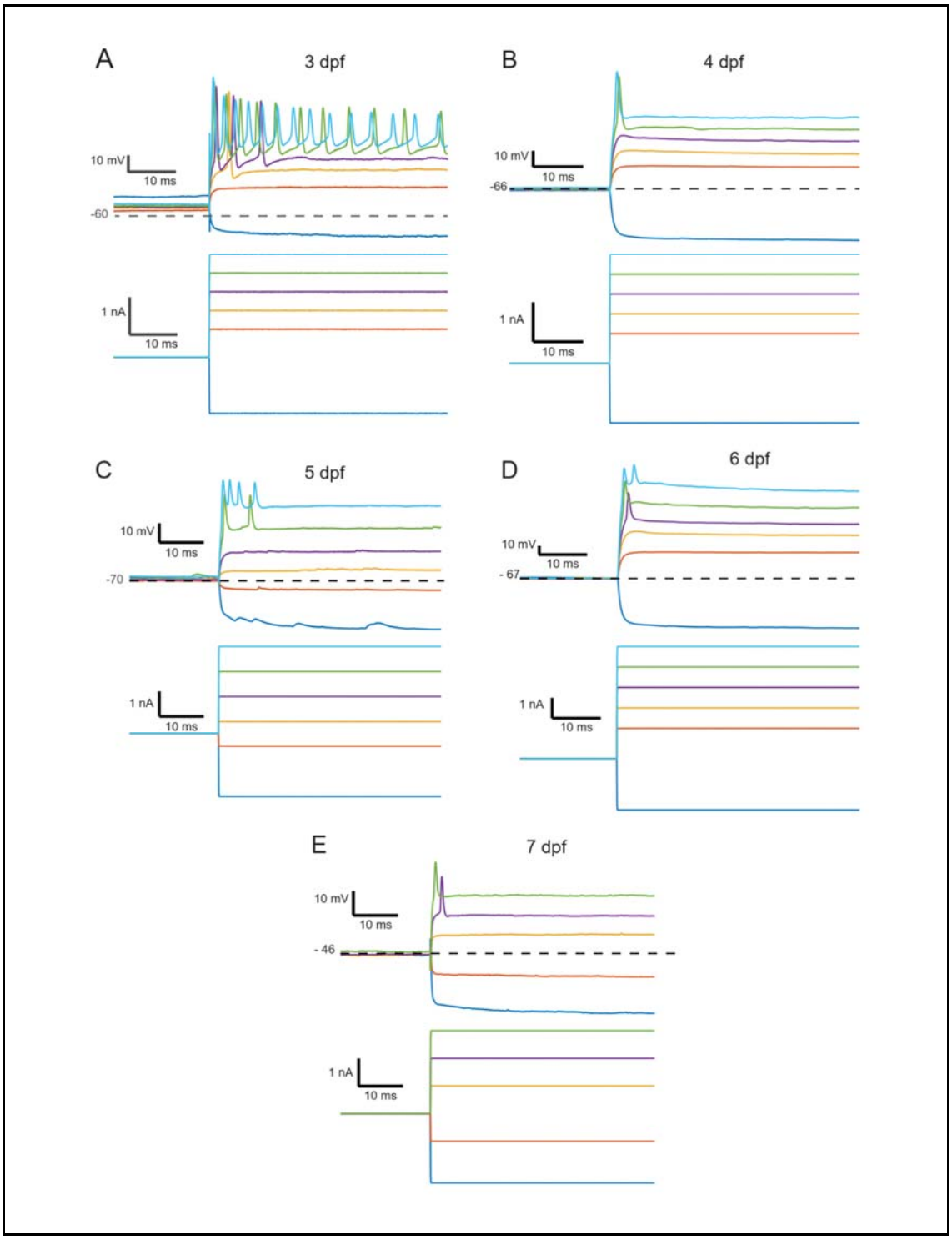

A
3 dpf
10 mV
10 ms
-60
1 nA
10 ms
B
4 dpf
10 mV
10 ms
-66
1 nA
10 ms
C
5 dpf
10 mV
10 ms
-70
1 nA
10 ms
D
6 dpf
10 mV
10 ms
- 67
1 nA
10 ms
E
7 dpf
10 mV
10 ms
- 46
1 nA
10 ms

Figure A2.7 - F/I curves from Mauthner neurons over development. (A-E) Representative current injection profiles from Mauthner neurons in zebrafish larvae from 3dpf-7dpf. In each panel, the plot at the bottom shows the current injection protocol and the correspondingly colour-coded plot on top represents the voltage recorded from the neuron in response to each applied step. Note that increasing current injection beyond the rheobase leads to multiple spikes in fish at 3dpf **(A)**, but only single action potentials in older fish, like in 7dpf **(E)**.

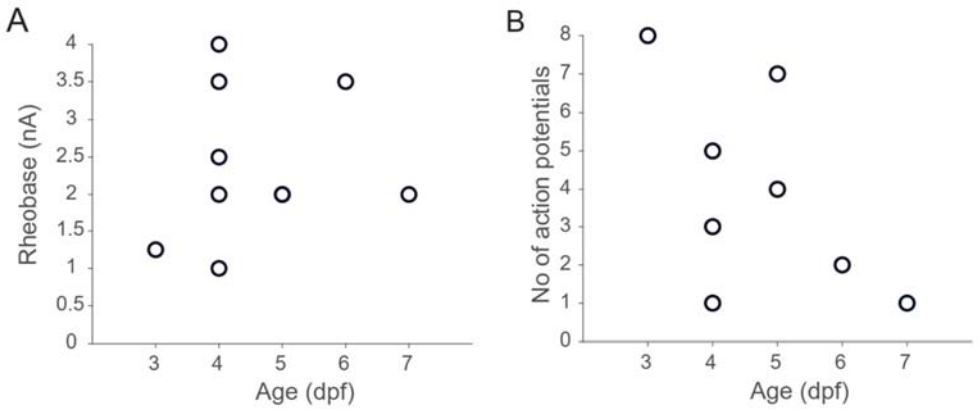


Figure A2.8 - F/I curves over development for all Mauthner neurons in the dataset. **(A)** Rheobase for all Mauthner neurons as a function of developmental age (n=10 cells). **(B)** Number of action potentials in a 100 ms-long window in response to a current injection of 1.5xRheobase (n = 8 cells).

A2.2.5 Spontaneous synaptic currents were observed in M-cells

In the process of establishing recordings from Mauthner neurons, I recorded spontaneous activity in these cells. Considering the function of the M-cells in escape responses, and the one-to-one correspondence between M-cell firing and behaviour, I expected these cells to be synaptically quiet when no stimulus was presented to them. However, to my surprise, I found that M-cells receive large, somewhat periodic synaptic input spontaneously, irrespective of developmental stage (Fig A2.9). Reports on the spontaneous synaptic activity in Mauthner neurons are not numerous. I only found a single study that reported the spontaneous synaptic activity in Mauthner neurons (Patten & Ali, 2007), although a few others have studied miniature EPSCs and IPSCs in these neurons (Brewster & Ali, 2013).

I also recorded spontaneous activity from these neurons in current clamp mode. I observed voltage fluctuations (Fig A2.9, brown traces) corresponding to excitatory synaptic currents recorded in voltage clamp mode (Fig A2.9, black traces). While these were repetitive, they were not strictly periodic and were all subthreshold and did not trigger action potential firing.

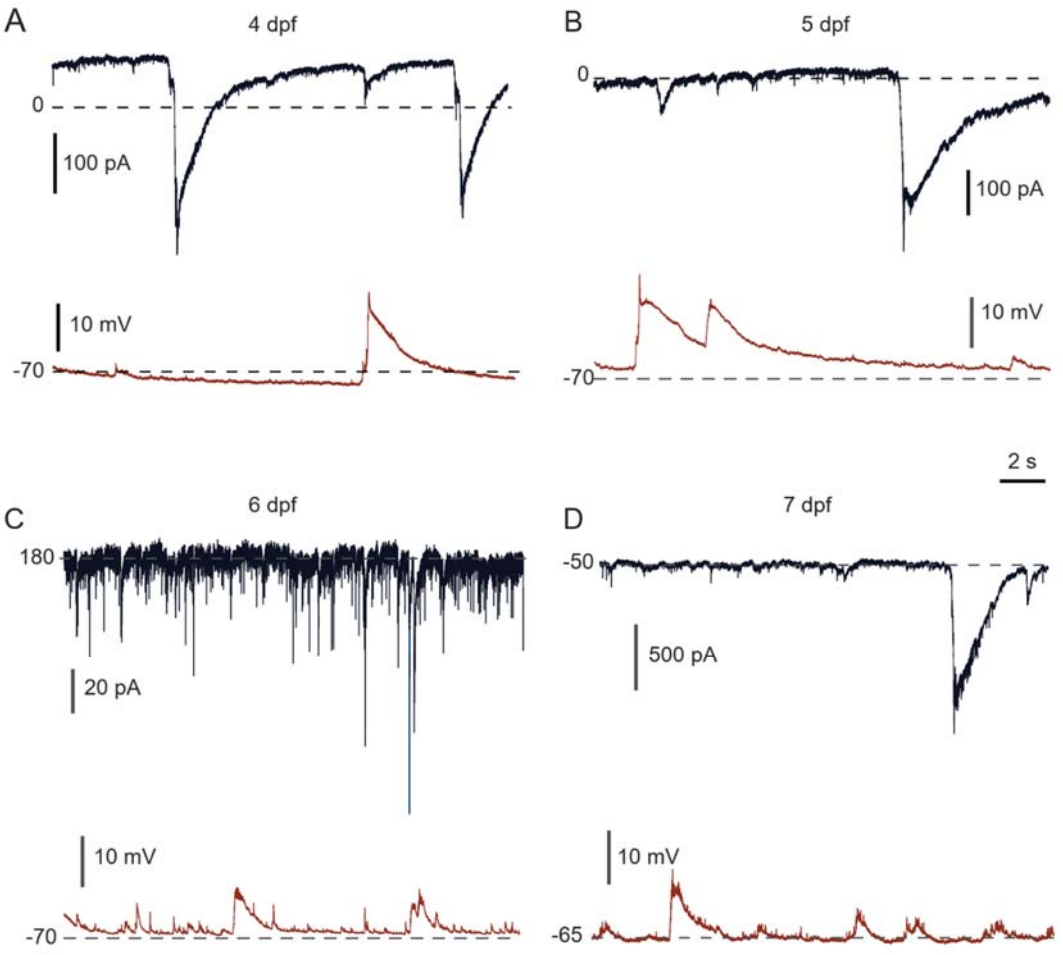


Figure A2.9 - Spontaneous subthreshold activity measured in Mauthner neurons over development. (A-D, top) Representative traces showing spontaneous inward currents in Mauthner neurons from 4-7 dpf, recorded in voltage clamp mode with a holding of -65mV. **(A-D, bottom)** Representative traces from the same cells as in (A-D, top), showing spontaneous depolarization in Mauthner neurons from 4-7 dpf, recorded in current clamp mode.

A2.2.6 Otic vesicle stimulation elicited complex post-synaptic potentials

Having established that I could record from Mauthner neurons for extended periods, I next wanted to record the mixed synaptic currents in these neurons, for two reasons. First, I wanted to ask how the strength of each of these synapses changed in response to increasing stimulus strength. Secondly, I needed to have some reference traces so that I could reproduce these conductances using dynamic clamp. Hence, I recorded the response of an M-cell to electrical stimulation of the otic vesicle with a bipolar electrode, using a modified ventral-up preparation that kept the otic vesicles intact (Fig A2.10A). I chose to stimulate the otic vesicle, and not just 8th cranial nerve fibres since I wanted to ask how the strength of each component of the mixed synapse changes in response to an ethologically relevant stimulus.

I found that otic vesicle stimulation elicited complex EPSCs and EPSPs, representative traces of which are shown in Fig A2.10B and Fig A2.10C. I performed otic vesicle stimulation in a total of n=6 cells from N=6 preparations, at 4 and 5 dpf, all of which yielded similar complex responses. These responses were quite variable across trials, and I was unable to isolate the mixed synaptic response from these traces, even in the average trial response (Fig A2.10B and C, black trace).

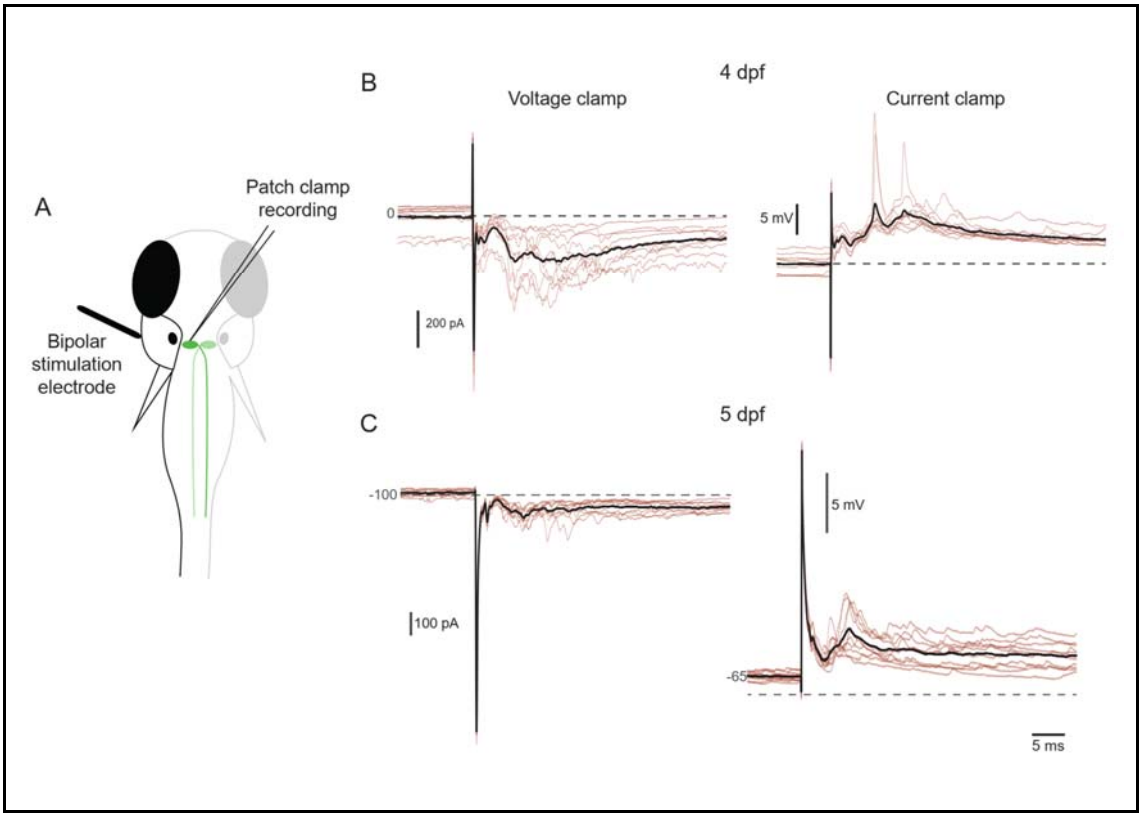


Figure A2.10 - Otic vesicle stimulation leads to a complex post-synaptic response. **(A)** Schematic of otic vesicle stimulation with simultaneous whole-cell recording from the ipsilateral M-cell. **(B)** Example M-cell response to otic vesicle stimulation from a larva at 4dpf, recorded in voltage clamp (left) and current clamp (right). Individual trials are shown in brown, and the average is shown in black. **(C)** Example M-cell response to otic vesicle stimulation, as in (B), but from a larva at 5dpf. The dotted lines in current clamp recordings represent -65 mV.

A2.3 Discussion

A2.3.1 Intrinsic properties and spontaneous synaptic activity in Mauthner neurons

I standardised a ventral-up preparation and was able to obtain stable recordings from Mauthner neurons in larval zebrafish, reproducing a key result, i.e. the developmental profile of single-spiking in the M-cell (Watanabe et al., 2014). Single-spiking is an interesting property of several auditory neurons in birds and mammals and is dependent on a dendrotoxin I (DTX)-sensitive low-threshold potassium current (Barnes-Davies et al., 2004; Brew & Forsythe, 1995; Dodson et al., 2002). In M-cells in goldfish and zebrafish, this single-spiking is dependent not only on a DTX-sensitive potassium current, but also recurrent inhibition (Brewster & Ali, 2013; Nakayama & Oda, 2004; Watanabe et al., 2014).

Another interesting property of M-cells that is poorly explored is their spontaneous subthreshold activity. I found that M-cells have spontaneous synaptic currents in voltage clamp mode, as has been reported previously (Patten & Ali, 2007). The existence of these currents could be attributed to poor space clamp associated with recording from such a large neuron. However, this hypothesis can be ruled out because corresponding subthreshold synaptic potentials were recorded in current clamp mode (Fig A2.9). The source and function of these semi-periodic excitatory currents is unclear. One possibility is that these subthreshold depolarisations could be priming the M-cell to be ready to fire faster when a noxious stimulus arrives, and may determine which of the two M-cells gets activated by a non-directional threatening stimulus. The expectation is that these currents would be out-of-phase in the two cells so that at least one M-cell is always closer to threshold at any given point in time. Simultaneously recording these currents from both M-cells could be used to verify this.

A2.3.2 The functions of electrical and chemical synapses at a mixed synapse

That electrical synapses contribute speed to a circuit is thought to be true because the postsynaptic membrane potential changes resulting from individual electrical synapses are faster than those resulting from chemical synapses. Moreover, experimental knockout of gap junctions leads to slower functional outcomes - knocking out Cx34.1 and/or Cx35.5 in larval zebrafish led to an ~2ms increase in median escape latency (Miller et al., 2017). However, the presence of electrical coupling in circuits often leads to unexpected circuit dynamics (Marder et al., 2017), and the interactions between the electrical and chemical synapses at a mixed synapse are not straightforward to predict. They can complement each other, such as in the backward locomotor circuit of *C. elegans* (Liu et al., 2017), but can also interfere with one another's function in various ways. Firstly, the addition of gap junctional conductances to a neuron reduces its input resistance, and, in turn, its excitability (Alcamí & Pereda, 2019). Secondly, gap junctional conductances can act as a leak of chemical synaptic currents, as I observed in my simulations of the M-cell mixed synapse (Fig A2.4). Lastly, connexin proteins are conduits for ions only when they assemble into 6-unit connexons, of which most are non-functional. For example, it has been estimated that only ~1-3% of channels in goldfish M-cell mixed synapses are conducting (Lin & Faber, 1988). Their role as signalling molecules is less appreciated, but may also be critical to circuit function.

In this study, I aimed at asking whether increasing gap junctional conductance correlated with increased escape speed, as measured using the M-cell firing latency as a proxy (since it is the rate-limiting step of this circuit). I wanted to then specifically perturb electrical and/or chemical synapse conductances *in vivo* by injecting simulated conductances into real, biological neurons using dynamic clamp. Unfortunately, I was unable to specifically measure the synaptic responses to realistic stimuli that I needed as an input for my dynamic clamp experiment. Thus, the question of what exactly each synapse type does at the M-cell mixed synapse remains open. Hopefully, developments which allow more specific, precise perturbations of electrical synapses might help answer this question in the future.

A2.4 Materials and Methods

A2.4.1 Simulations of mixed synapses in Mauthner neurons

Conductance-based models of the zebrafish Mauthner neuron were made using the NEURON programming language (Hines & Carnevale, 2001). The M-cell model was adapted from a model publicly available on ModelDB (Watanabe et al., 2017), which itself was modified from a mouse cochlear neuron model. The model is a single-compartment

model, with parameters as summarised in Table A2.2. Parameters used were taken from Koyama et al. (2016) and Watanabe et al., (2017), wherever relevant. Table A2.2 also quotes values from both these papers for comparison.

A presynaptic compartment was defined to represent the 8th cranial nerve that makes a mixed synapse with the M-cell. This compartment was also a single compartment, with only classical Hodgkin-Huxley conductances. Its capacitance was set such that the surface area matched the value specified by Koyama et al. (2016) [~2000 m^2].

The mixed synapse was modelled as a combination of the Gap point process mechanism, as defined in the NEURON book (Hines & Carnevale, 2001), and an ExpSyn conductance mechanism. For the gap junction, the free parameter was the gap junctional conductance (g_{elec}). The chemical synapse, however, has 3 free parameters - the conduction delay between the presynaptic action potential and postsynaptic response, the maximal conductance (g_{chem}), and the decay time constant (τ_{decay}) of the conductance. All these are schematically shown in Figure A2.2B.

The range of these values investigated was within physiological limits, as described in Table A2.2. Three types of simulations were performed – one with electrical synapses alone, one with chemical synapses alone, and the third with the mixed synapse. As a control, the key parameters of the Action Potential shape (spike threshold, spike amplitude, FWHM, and maximum rise slope) were measured by simulating independent current clamp electrodes into each compartment.

An ideal current clamp electrode was inserted into the dummy compartment to initiate a presynaptic action potential. Post-synaptic responses were recorded, and a range of parameters were analysed using custom scripts written in Python. Common parameters for all simulations included resting membrane potential, whether or not a postsynaptic action potential was observed, and if so, the spike latency. In the case when only gap junctions were there, the coupling coefficient was measured as a function of the gap junctional conductance. All simulations were performed at a temperature of 25^{o}C.

A2.4.2 Fish rearing and breeding

The use of all experimental animals has been approved by the Institutional Animal Ethics Committee (IAEC) and the Institutional Biosafety Committee (IBSC). Standard procedures were used to raise zebrafish (*Danio rerio*) of the following genotypes: Indian Wild-Type,

nacre -/- (albino), Tol-026, Tol-056 and hspGFF62A (henceforth referred to as 62A). In brief, embryos and larvae were grown in E3 medium (composition, in mM: 5 NaCl, 0.17 KCl, 0.33 CaCl2 and 0.33 MgSO4, pH 7.8). At about 15 days post fertilisation (dpf), larvae were transferred to a Zebtec standalone system, until about 2 months post fertilisation (mpf), when they were transferred to a Zebtec multilinking system. Young larvae, from 5-15 dpf, were fed Zeigler AP-100; 15dpf-1mpf were fed a combination of AP-100 and freshly hatched brine shrimp (*Artemia spp.*); 1mpf and above were fed Zeigler Adult Zebrafish Diet. Fish were fed at least 3 times per day, according to a standard schedule in the lab, and the system water in which they were raised was maintained at a temperature of ~28^{o}C, a pH of ~7.2, and a conductivity of ~1200mS.

In order to obtain embryos and larvae for experiments, a male and female are identified and placed overnight in a breeding tank with an internal perforated breeding net, filled with system water. Eggs were collected and raised as described above.

Tol-026 and Tol-056 lines were obtained from Prof. Hitoshi Okamoto's lab at the RIKEN Brain Science Institute, Saitama, Japan. The 62A line was obtained from Prof. Koichi Kawakami's lab at the National Institute of Genetics, Japan.

A2.4.3 Retrograde labelling of reticulospinal neurons

Larval zebrafish (4-6 dpf) were first anaesthetised in 0.01% MS-222 (Tricaine) and positioned on an agar dish. Excess water around the larva was drained using a tissue paper (Scott Multi-Fold Paper Towels, Kimberley-Clark). The larva was positioned such that its dorsal side would be approached by the injection needle. Injection needles were pulled from thin glass capillaries (OD 1.0mm and ID 0.56mm; Warner Instruments) using a Flaming-Brown P-97 pipette puller (Sutter Instruments, Novato, CA, United States). They were then backfilled with 25% TMR-dextran (1000 Da) in 10% Hank's Balanced Salt Solution (HBSS, composition: 137 mM NaCl, 5.4 mM KCl, 0.25 mM Na_2HPO_4, 0.44 mM KH_2PO_4, 1.3 mM $CaCl_2$, 1 mM $MgSO_4$, and 4.2 mM $NaHCO_3$). The needle was then connected to a Picospritzer III (Parker Hannifin Corporation, Hollis, NH, USA) - a pressure delivery system that allows the controlled ejection of picolitre volumes. The needle was then mounted onto a manual NM-3 micromanipulator on a GJ-1 magnetic stand (Narishige Instruments) and brought close to the larva under visual guidance using an Olympus SZ51 microscope. The tip of the pulled injection needle was first broken using fine forceps. The needle was pierced through the skin on the dorsal surface of the larva and slowly advanced until the tip touched the notochord. The needle was then retracted a bit before a pulse of dye was released into the spinal cord. Injected larvae were immediately transferred to freshly prepared HBSS and allowed to recover for a period of at least 12 hours prior to screening and experimentation. For screening, individual larvae were placed in separate droplets on the cover of a 90mm

Petri plate (Tarsons). Each larva was observed using the widefield epifluorescence module of an Olympus SZX16 dissection microscope and larvae with labelled reticulospinal neurons were separated for experiments.

A2.4.4 Patch-clamp electrophysiology from Mauthner neurons

Dissection

In order to get a preparation suitable for electrophysiological recordings, zebrafish larvae (3-7 dpf) of the 62A genotype were first anaesthetised in 0.01% MS-222 (Tricaine), positioned lateral-up in a recording chamber with a Sylgard (Dow Corning, Midland, MI, USA) block on it and pinned onto the block using fine tungsten wire (California Fine Wire, Grover Beach, CA, USA). The first pin was pushed through the tail, and second through the notochord of the larva, at the level of the yolk sac. The remaining procedures for larvae of 3-4dpf and for 5 dpf and above were slightly different, as described below.

3-4 dpf larvae were oriented dorsal-side up using a third pin through the jaw. The MS-222 was then replaced with external saline. Skin from the dorsal surface was carefully peeled off using a pair of fine No. 5 forceps (Fine Science Tools, British Columbia, Canada) to expose the brain. The eyes and jaws were then removed, and the brain positioned laterally. The notochord that lies just ventral to the brain was carefully separated from the brain using gentle strokes of the fine forceps. The brain is then gently rotated to a ventral-up position, and a pin is pushed through the rostral-most part of the forebrain, in order to stabilise the preparation.

5dpf and older larvae were oriented ventral-side up using a third pin through the jaw. The tricaine was then replaced with external saline. The yolk sac, air bladder, and lower jaw were pulled out using No. 5 forceps. Using a sharp tungsten pin, the region just caudal to the otic vesicles was gently scratched to release the upper jaw, and the entire bony tissue surrounding the brain was removed as one piece. The dorsal skin has pigments that can occlude vision at later stages, so this was gently removed from the under surface of the preparation using a tungsten wire, as needed. Finally, the preparation was stabilised by pushing a pin through the rostral-most part of the forebrain, as for the younger preparations.

Once the dissection was completed, debris from the dissection was cleared out by replacing the saline with fresh external saline. The recording chamber was then taken to the electrophysiology rig and the quality of the preparation assessed visually using an upright Nikon Ni-E microscope equipped with 20x (0.5NA) and 60x (1.0NA) water immersion objectives. The presence of both M-cells labelled with GFP, as seen by epifluorescence microscopy, was a key indicator of the quality of the preparation - even if one of the M-cells was not visible, the experiment was aborted and a fresh preparation was made.

Targeted Whole-Cell Patch Clamp Recording

Thin-walled borosilicate glass capillaries (OD 1.5mm; ID 1.17mm; Warner Instruments, Hamden, CT, USA) were pulled using a Flaming Brown P-97 pipette puller (Sutter Instruments, Novato, CA, USA) to produce patch micropipettes of tip diameter 1.2-2μm and resistance of 4-5MΩ. These were filled with a potassium gluconate or potassium methanesulfonate internal solution (Table A2.3) containing 5 μg/ml sulforhodamine or Alexa Fluor 488 hydrazide (Invitrogen), as needed (solution compositions detailed in Table 3). Before bringing the micropipette into the external bath solution, a mild positive pressure was applied to prevent pipette clogging. The patch electrode was manipulated using a PatchStar motorised micromanipulator (Scientifica Instruments, East Sussex, United Kingdom) and focused to the level of the M-cell in the preparation. Simultaneous digital GFP epifluorescence and low-intensity brightfield microscopy was performed to visualise the M-cell and the patch micropipette, using a CMOS camera (Flea3 USB 3.0, FLIR Systems, Inc.) attachment at the rig microscope. The M-cell was approached, until a slight dimple on the M-cell membrane was observed.

The positive pressure was released to form a seal only if both a dimple on the M-cell and an increase in open pipette resistance of at least 0.7 MΩ were observed. Seal formation (~1-5 GΩ) was assisted with the application of slight negative pressure, and the simultaneous slow decrease in holding potential from 0mV to -65mV. Pipette capacitance was then compensated for before entering the whole-cell mode. Attempts to enter the whole-cell configuration were made only if the seal resistance was in the gigohm range, by applying strong, sustained negative pressure via a 1ml syringe. A dramatic drop in resistance with a holding current less than 100pA in magnitude indicated successful entry into whole-cell mode. Recordings were then acquired in either Voltage Clamp mode or Current Clamp mode, as required, using a Multiclamp 700B amplifier, Digidata 1440A digitizer and pCLAMP software (Molecular Devices, Sunnyvale, CA, United States). Data was low-pass filtered using a 10 kHz Bessel filter and digitised at a rate of 50 kHz. The voltage clamp feedback resistor used was 5 GΩ, and the current clamp feedback resistor used is 50 MΩ. Whole-cell capacitance and series resistance compensation were not done in voltage clamp mode, although series resistance was monitored at the start and the end of each recording, and recordings were not included in analysis if series resistance changes by more than 25%. Bridge balance was auto-adjusted in current clamp mode. Liquid junction potential correction was not done, but has been calculated to be +20.2 mV for the potassium gluconate internal solution and +24 mV for the potassium methanesulfonate internal solution.

Successful whole-cell recordings were verified by colocalization in the GFP and RFP channels, as shown in Figure 3. Offline analysis was done using Clampfit software

(Molecular Devices, Sunnyvale, CA, United States), as well as custom-written Python and MATLAB scripts, wherever needed.

Otic vesicle stimulation

Otic vesicle stimulation was performed using concentric bipolar electrodes (FHC) and an ISO-Flex stimulus isolator (AMPI). The dissections were done as described above, ensuring that the otic vesicles were kept intact. After taking the preparation to the rig and focussing the sample, the bipolar electrode was positioned at the level of the otic vesicle. Once positioned, the Mauthner neuron on the side ipsilateral to the stimulation electrode was targeted for whole-cell patch clamp electrophysiology, performed as described above. The stimulus intensity was serially increased while the cell was recorded from in either voltage clamp or current clamp mode, as needed.

Table A2.3 - Composition of solutions used for electrophysiology of Mauthner neurons.

External Saline		**Internal (Potassium Gluconate)**		**Internal (Potassium Methanesulfonate)**	
Reagent	**Concentration (mM)**	**Reagent**	**Concentration (mM)**	**Reagent**	**Concentration (mM)**
NaCl	134	K gluconate	115	K methanesulfonate	115
KCl	2.9	KCl	15	HEPES	10
$MgCl_2.6H_2O$	1.2	$MgCl_2.6H_2O$	2	$MgCl_2.6H_2O$	2
HEPES	10	HEPES	10	EGTA	10

D-Glucose	10	EGTA	10	Na_2ATP	4
$CaCl_2$	2.1	MgATP	4	Na_2GTP	0.4
d-Tubocurarine	0.01				
pH 7.8 (adjusted using 3M NaOH); 290±10 mOsm		pH 7.2 (adjusted using 3M KOH); 290 mOsm		pH 7.2 (adjusted using 3M KOH); 274 mOsm	

A2.4.5 Dynamic Clamp

For dynamic clamping fast conductances, such as electrical synapses, I used a PCIe6361 Data Acquisition (NI-DAQ) card as the digitiser (National Instruments, Austin, TX, United States). This card was connected to the Multiclamp 700B using Analog Inputs and Outputs. Dynamic Clamp was implemented using the StdpC software (Kemenes et al., 2011). Preliminary experiments, demonstrating that dynamic clamp has been properly set up, were done using the PATCH 1-U Model Cell (Molecular Devices), as well as a hindbrain neuron patched from a ventral-up preparation, as shown in Figs A2.5C and D.

A2.5 References

Alcamí, P., & Pereda, A. E. (2019). Beyond plasticity: The dynamic impact of electrical synapses on neural circuits. *Nature Reviews Neuroscience*, *20*(5), Article 5. https://doi.org/10.1038/s41583-019-0133-5

Barnes-Davies, M., Barker, M. C., Osmani, F., & Forsythe, I. D. (2004). Kv1 currents mediate a gradient of principal neuron excitability across the tonotopic axis in the rat lateral superior olive. *European Journal of Neuroscience*, *19*(2), 325–333. https://doi.org/10.1111/j.0953-816X.2003.03133.x

Brew, H. M., & Forsythe, I. D. (1995). Two voltage-dependent K+ conductances with complementary functions in postsynaptic integration at a central auditory synapse. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *15*(12), 8011–8022. https://doi.org/10.1523/JNEUROSCI.15-12-08011.1995

Brewster, D. L., & Ali, D. W. (2013). Expression of the voltage-gated potassium channel subunit Kv1.1 in embryonic zebrafish Mauthner cells. *Neuroscience Letters*, *539*, 54–59. https://doi.org/10.1016/j.neulet.2013.01.042

Burgess, H. A., & Granato, M. (2007). Sensorimotor Gating in Larval Zebrafish. *Journal of Neuroscience*, *27*(18), 4984–4994. https://doi.org/10.1523/JNEUROSCI.0615-07.2007

Dodson, P. D., Barker, M. C., & Forsythe, I. D. (2002). Two Heteromeric Kv1 Potassium Channels Differentially Regulate Action Potential Firing. *Journal of Neuroscience*, *22*(16), 6953–6961. https://doi.org/10.1523/JNEUROSCI.22-16-06953.2002

Eaton, R. C., & Nissanov, J. (1985). A review of Mauthner-initiated escape behavior and its possible role in hatching in the immature zebrafish, Brachydanio rerio. *Environmental Biology of Fishes*, *12*(4), 265–279. https://doi.org/10.1007/BF00005457

Fetcho, J. R. (1991). Spinal network of the Mauthner cell. *Brain, Behavior and Evolution*, *37*(5), 298–316. https://doi.org/10.1159/000114367

Hines, M. L., & Carnevale, N. T. (2001). NEURON: A tool for neuroscientists. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *7*(2), 123–135. https://doi.org/10.1177/107385840100700207

Johnson, B. R., Peck, J. H., & Harris-Warrick, R. M. (1993). Dopamine induces sign reversal at mixed chemical-electrical synapses. *Brain Research*, *625*(1), 159–164. https://doi.org/10.1016/0006-8993(93)90149-H

Kemenes, I., Marra, V., Crossley, M., Samu, D., Staras, K., Kemenes, G., & Nowotny, T. (2011). Dynamic clamp with StdpC software. *Nature Protocols*, *6*(3), Article 3. https://doi.org/10.1038/nprot.2010.200

Kohashi, T., & Oda, Y. (2008). Initiation of Mauthner- or Non-Mauthner-Mediated Fast Escape Evoked by Different Modes of Sensory Input. *Journal of Neuroscience*, *28*(42), 10641–10653. https://doi.org/10.1523/JNEUROSCI.1435-08.2008

Korn, H., & Faber, D. S. (2005). The Mauthner cell half a century later: A neurobiological model for decision-making? *Neuron*, *47*(1), 13–28. https://doi.org/10.1016/j.neuron.2005.05.019

Koyama, M., Minale, F., Shum, J., Nishimura, N., Schaffer, C. B., & Fetcho, J. R. (2016). A circuit motif in the zebrafish hindbrain for a two alternative behavioral choice to turn left or right. *ELife*, *5*, e16808. https://doi.org/10.7554/eLife.16808

Lin, J. W., & Faber, D. S. (1988). Synaptic transmission mediated by single club endings on the goldfish Mauthner cell. I. Characteristics of electrotonic and chemical postsynaptic potentials. *Journal of Neuroscience*, *8*(4), 1302–1312. https://doi.org/10.1523/JNEUROSCI.08-04-01302.1988

Liu, P., Chen, B., Mailler, R., & Wang, Z.-W. (2017). Antidromic-rectifying gap junctions amplify chemical transmission at functionally mixed electrical-chemical synapses. *Nature*

Communications, *8*(1), Article 1. https://doi.org/10.1038/ncomms14818

Marder, E., Gutierrez, G. J., & Nusbaum, M. P. (2017). Complicating connectomes: Electrical coupling creates parallel pathways and degenerate circuit mechanisms. *Developmental Neurobiology*, *77*(5), 597–609. https://doi.org/10.1002/dneu.22410

Miller, A. C., Whitebirch, A. C., Shah, A. N., Marsden, K. C., Granato, M., O'Brien, J., & Moens, C. B. (2017). A genetic basis for molecular asymmetry at vertebrate electrical synapses. *ELife*, *6*, e25364. https://doi.org/10.7554/eLife.25364

Nakayama, H., & Oda, Y. (2004). Common Sensory Inputs and Differential Excitability of Segmentally Homologous Reticulospinal Neurons in the Hindbrain. *Journal of Neuroscience*, *24*(13), 3199–3209. https://doi.org/10.1523/JNEUROSCI.4419-03.2004

Patten, S. A., & Ali, D. W. (2007). AMPA receptors associated with zebrafish Mauthner cells switch subunits during development. *The Journal of Physiology*, *581*(3), 1043–1056. https://doi.org/10.1113/jphysiol.2007.129999

Pereda, A. E. (2014). Electrical synapses and their functional interactions with chemical synapses. *Nature Reviews Neuroscience*, *15*(4), Article 4. https://doi.org/10.1038/nrn3708

Pereda, A. E., & Faber, D. S. (1996). Activity-dependent short-term enhancement of intercellular coupling. *Journal of Neuroscience*, *16*(3), 983–992. https://doi.org/10.1523/JNEUROSCI.16-03-00983.1996

Prinz, A. A., Abbott, L. F., & Marder, E. (2004). The dynamic clamp comes of age. *Trends in Neurosciences*, *27*(4), 218–224. https://doi.org/10.1016/j.tins.2004.02.004

Rash, J. E., Dillman, R. K., Bilhartz, B. L., Duffy, H. S., Whalen, L. R., & Yasumura, T. (1996). Mixed synapses discovered and mapped throughout mammalian spinal cord. *Proceedings of the National Academy of Sciences*, *93*(9), 4235–4239. https://doi.org/10.1073/pnas.93.9.4235

Serrano-Velez, J. L., Rodriguez-Alvarado, M., Torres-Vazquez, I. I., Fraser, S. E., Yasumura, T., Vanderpool, K. G., Rash, J. E., & Rosa-Molinar, E. (2014). Abundance of gap junctions at glutamatergic mixed synapses in adult Mosquitofish spinal cord neurons. *Frontiers in Neural Circuits*, *8*. https://www.frontiersin.org/article/10.3389/fncir.2014.00066

Sharp, A. A., O'Neil, M. B., Abbott, L. F., & Marder, E. (1993). The dynamic clamp: Artificial conductances in biological neurons. *Trends in Neurosciences*, *16*(10), 389–394. https://doi.org/10.1016/0166-2236(93)90004-6

Watanabe, T., Shimazaki, T., Mishiro, A., Suzuki, T., Hirata, H., Tanimoto, M., & Oda, Y. (2014). Coexpression of auxiliary Kvβ2 subunits with Kv1.1 channels is required for developmental acquisition of unique firing properties of zebrafish Mauthner cells. *Journal of Neurophysiology*, *111*(6), 1153–1164. https://doi.org/10.1152/jn.00596.2013

Watanabe, T., Shimazaki, T., & Oda, Y. (2017). Coordinated Expression of Two Types of Low-Threshold K+ Channels Establishes Unique Single Spiking of Mauthner Cells among Segmentally Homologous Neurons in the Zebrafish Hindbrain. *ENeuro*, *4*(5), ENEURO.0249-17.2017. https://doi.org/10.1523/ENEURO.0249-17.2017

Yang, X.-D., Korn, H., & Faber, D. S. (1990). Long-term potentiation of electrotonic coupling at mixed synapses. *Nature*, *348*(6301), Article 6301. https://doi.org/10.1038/348542a0

Appendix 3 - Structure and function of the olivocerebellar loop in larval zebrafish

A3.1 Introduction

The olivo-cerebellar circuit, which includes the inferior olive and the cerebellum among other brain regions, is a well-conserved neuronal pathway across vertebrates that plays an important role in motor control. The inferior olivary nucleus, often shortened to the inferior olive, is a small cluster of cells that is part of the hindbrain. More interestingly, it is the sole source of climbing fibre input to the cerebellum (Eccles et al., 1966).

However, the role of the inferior olive is as yet unclear, although it is known to be involved in motor control. It is one of several brain regions that get affected in pathological conditions such as spinocerebellar ataxia (SCA), in which patients present with severe motor deficits. Similarly, lesioning the inferior olive in cats led to difficulties in movements such as reaching (Horn et al., 2013). In larval zebrafish, which is our model system of interest, it has been shown that inferior olivary neurons are active during a motor adaptation task, and that its ablation abrogates adaptation (Ahrens et al., 2012). Thus, while we know that the inferior olive is involved in motor control, its exact role has been hard to decipher.

A particular microcircuit module of interest in the olivocerebellar circuit is the olivocerebellar loop, which is a feedback loop between the inferior olive and the cerebellum. The existence of such a loop has been shown in adult mammals. Climbing fibres from the inferior olive split off and innervate both the Purkinje neurons (PNs) as well as the deep cerebellar nuclei (DCN) neurons, i.e. the target neurons of PNs (Fig A3.1A). There is also a reverse connection from the DCN back to the inferior olive (Fig A3.1A). The mouse deep cerebellar nuclei have both excitatory (glutamatergic) and inhibitory (GABAergic and glycinergic) neurons. The former are derived from atoh1+ progenitors in the upper rhombic lip, while the latter from Ptf1a+ progenitors in the ventricular zone (Hibi & Shimizu, 2012). Of the two kinds of inhibitory neurons, the GABAergic population projects predominantly to the inferior olive (Judd et al., 2021; Prekop et al., 2018), while the glycinergic population projects not to the inferior olive but to various premotor nuclei, such as those in the brainstem (Bagnall et al., 2009).

Unlike mammals, however, teleost fish don’t have deep cerebellar nuclei. Instead, they have projection neurons called eurydendroid cells, which are found in the Purkinje cell layer (to be exact, it is called the ganglionic cell layer, because it has both Purkinje cells as well as

efference cells (Meek et al., 2008). These cells are characterised as being glutamatergic neurons that typically have one to two dendrites that project into the molecular layer. A subset of them express olig2, while others are marked by calretinin (Bae et al., 2009; Harmon et al., 2020). None of the eurydendroid cells, however, are inhibitory, even though a fraction of them are derived from the ptf1a+ progenitors (Bae et al., 2009; Hibi & Shimizu, 2012; Ikenaga et al., 2006; Kani et al., 2010). Furthermore, no projection from the cerebellum to the inferior olive has been reported in fish thus far (Fig A3.1B). The proximity of eurydendroid cells and Purkinje neurons also confounds whether climbing fibre input goes to eurydendroid cells (Fig A3.1B). Whether this means that no olivocerebellar loop exists in primitive cerebelli is yet to be explored thoroughly. I therefore decided to undertake the task of studying the functional connectivity of the olivocerebellar loop in larval zebrafish, results for which are described in this chapter.

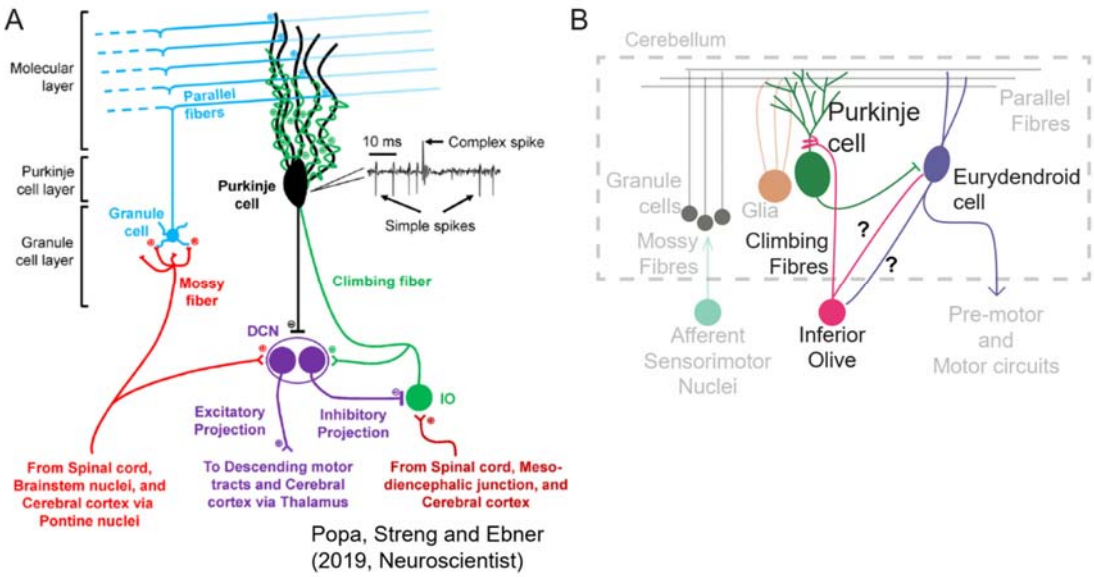


Figure A3.1 - **The olivocerebellar loop. (A)** Schematic of the olivocerebellar circuit in mice. Note the climbing fiber (green line) splitting off to provide excitatory input to the Deep Cerebellar Nuclei (DCN, purple). Note also the converse, inhibitory projection from the DCN back to the Inferior Olive (IO, green circle). **(B)** Circuit diagram of the cerebellum in larval zebrafish, with the olivocerebellar loop highlighted. Connections between the Inferior Olive (magenta) and Eurydendroid cells (purple) have not been shown, and are therefore marked with question marks.

A3.2 Results

A3.2.1 Shortlisting putative cell-type specific promoters for inferior olivary neurons

In order to be able to visualise and/or manipulate neurons of the inferior olive, I needed to first develop some tools. While there exists an enhancer trap line (hspGFFDMC28C, shortened to 28C) that drives GAL4 specifically in inferior olivary neurons (Takeuchi et al., 2015) (Fig A3.2A), we were unable to breed and maintain them successfully in the lab at that time. Fortunately, the group that generated the 28C line also made publicly available a transcriptome for these cells and various other cells of the olivocerebellar circuit (Takeuchi et al., 2017). I therefore decided to use this dataset to identify genes that were highly expressed specifically in the 28C line relative to the other lines. The idea, then, was that the promoters of these genes could be used to specifically express any agent of our interest in a cell-type specific manner without the use of an enhancer trap line.

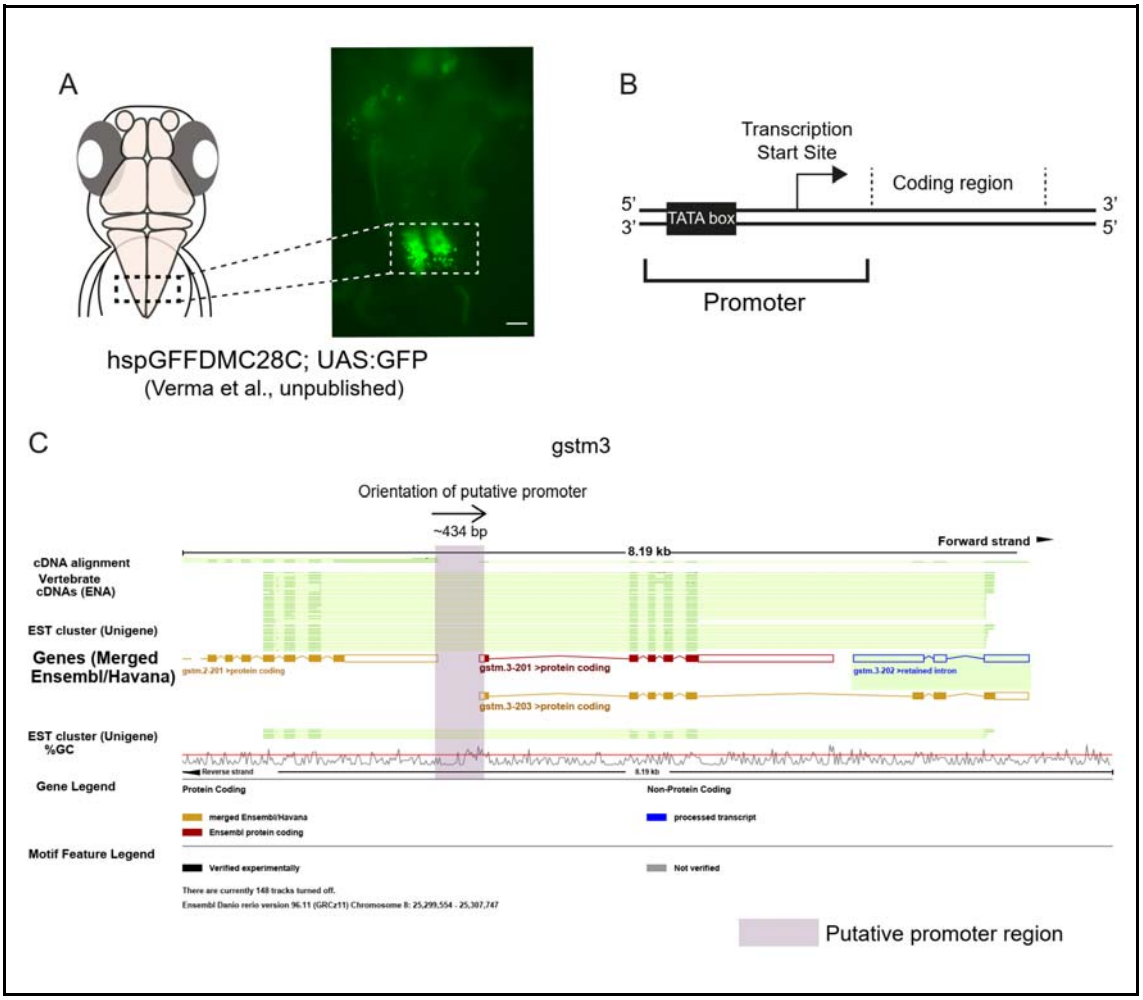


Figure A3.2 - Cell-type specific expression in the inferior olive in larval zebrafish. (A) UAS-driven expression of GFP in the hspGFFDMC28C GAL4 line in 5dpf larval zebrafish. Scale bars represent 100 µm. **(B)** The structure of a gene, highlighting features of the promoter region. **(C)** The organisation of a candidate inferior olive-specific gene (gstm3; also see Table A3.1) in the zebrafish reference genome, as visualised using the Ensembl Genome Browser. The putative promoter region of ~430 bp upstream of gstm3 is marked with a purple box.

I found 92 genes of interest that had at least a 20-fold higher expression in the 28C line than each of the other 4 olivocerebellar lines - aldoca (Purkinje neuron-specific), 152B (granule cell-specific), 156A (eurydendroid cell-specific) and 251A (Bergmann glia-specific). High relative expression cannot be the sole criterion to filter genes whose promoters could be useful cell-type specific markers. To be useful as a tool to drive gene expression, the promoter must yield transcripts at high absolute levels, so I further filtered through this list to identify genes that had an average FPKM value of >= 10 across all replicates. This shortened the list to 86 genes. Next, I filtered out genes that were not well annotated, which left 69 genes.

Finally, I manually selected 10 of these 69 genes to arrive at a final shortlist of genes of interest (Table A3.1). These genes spanned a wide range of functions - homeobox genes (*hoxb5a*), protease inhibitors (*itih6*, *serpinf1*), enzymes (*ampd1*, *gstm3*) to proteoglycans interacting with the extracellular matrix (*epyc*) and cytoskeleton-related proteins (*myhb* and *tnnc2*). I explored them using the Ensembl Genome Browser (Fig A3.2C) and identified regions of interest upstream of their coding sequence that could be their putative promoter. The putative promoter region (Fig A3.2B) includes the transcription start site, and must have at least one TATA box sequence (i.e. 5'-TATAA-3'). The shortlisted genes all had at least one such site up to 1kb from the transcription start site.

Table A3.1** - **Shortlist of candidate inferior olive-specific promoters.

Column labels:

A - Ensembl Gene ID

B - Gene symbol

C - Gene name

D - Is the gene on the forward or reverse strand of the reference genome?

E - Distance to the end of closest gene upstream (bp)

F - Closest TATA box in the correct orientation (bp)

G - Final amplicon sizes (bp) for two designed primer pairs

Genes whose putative promoter regions were successfully cloned are marked in **bold**.

A	B	C	D	E	F	G
ENSDARG00000033832	ampd1	adenosine monophosphate deaminase 1 (isoform M)	R	~2700	960	1461/1512
ENSDARG00000056950	**epyc**	**epiphycan**	**R**	**~3000**	**233**	**1003/1001**
ENSDARG00000078468	fap	fibroblast activation protein, alpha	F	~8350	670	1586/1466
ENSDARG00000088116	**gstm3**	**glutathione S-transferase mu tandem duplicate 3**	**F**	**~400**	**248**	**410**
ENSDARG00000013057	**hoxb5a**	**homeobox B5a**	**F**	**~2,000**	**417**	**1946/1896**
ENSDARG00000056037	**itih6**	**inter-alpha-trypsin inhibitor heavy chain family, member 6**	**F**	**~1570**	**583**	**1499/1485**
ENSDARG00000001993	**myhb**	**myosin, heavy chain b**	**R**	**~5,500**	**76**	**819/683**
ENSDARG00000069048	serpinf1	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1	F	~1550	1024	1112/1260
ENSDARG00000070835	**tnnc2**	**troponin C type 2 (fast)**	**F**	**~2400**	**1020**	**1178/1173**
ENSDARG00000001	TNXB	tenascin XBa	R	~15,50	863	1259/1043

760				0		

A3.2.2 Cloning and expression of putative inferior olive-specific promoters

Having shortlisted a set of high likelihood candidate inferior olive-specific genes, I decided to clone putative promoter regions from genomic DNA into an expression plasmid. I first tried using the Gateway cloning strategy that the Tol2 kit is based on. In this strategy, different parts of a sequence are put into separate vectors (called “entry” vectors) with unique att recombination sites. Multiple entry vectors with matching recombination sites are put together with a “destination” vector, which encodes a bacterial toxin, and the recombinase enzyme. Only one order of components will yield a successful recombination into the destination vector such that it replaces the toxin in the plasmid.

The promoter sequence was cloned into a 5’ entry vector - p5E-MCS (#228 in the Tol2 kit). This was done by PCR amplifying the putative regions using forward and reverse primers with overhangs for the restriction enzymes KpnI and XhoI, respectively. Two sets of primers were designed for each of the 10 shortlisted promoter regions (Table A3.1). Standard restriction digestion and ligation protocols were used to insert the putative promoter region into the 5’ entry vector. The use of different enzymes ensured that only the correct orientation of the promoter in the plasmid was obtained. The ligation mix was transformed into *E. coli* of the DH5alpha strain and plated onto media containing kanamycin. Colonies that were obtained were screened for the expression of the plasmid with the insert, which was then extracted using a standard plasmid isolation kit. The 5’ entry vector with the putative promoter was sequenced prior to use for downstream processing.

The reporter and polyA tail signal sequence have pre-existing plasmids that could be used out of the box. These are the pME-EGFP (“middle entry vector” - #383 in the Tol2 kit) and p3E-polyA (“3’ entry vector” - #302 in the Tol2 kit). The destination vector pDestTol2pA2 (#394 in the Tol2 kit) was used. All four components - the three entry vectors and one destination vector - were incubated with the LR recombinase enzyme. The reaction mixture was then transformed into *E. coli* of the DH5alpha strain and positive colonies were selected using ampicillin resistance as the marker.

Using this strategy, I successfully cloned 6 of the 10 shortlisted promoters of interest into the 5’ entry vector, which are highlighted in bold in Table A3.1. However, all attempts of the recombination reaction were unsuccessful and yielded no colonies. Thus, I switched cloning strategies and opted for Gibson Assembly (Gibson et al., 2009). I first designed primers for

PCR amplification such that they introduced extra overhangs with homology to upstream and downstream targets (Fig A3.3A) such that when stitched, the final plasmid would have the putative promoter followed by a marker gene (such as EGFP, for example) followed by the polyA sequence, in the correct order and orientation (Fig A3.3B). This destination vector also had Tol2 insertion sites that are needed for transgenesis in zebrafish (Fig A3.3B).

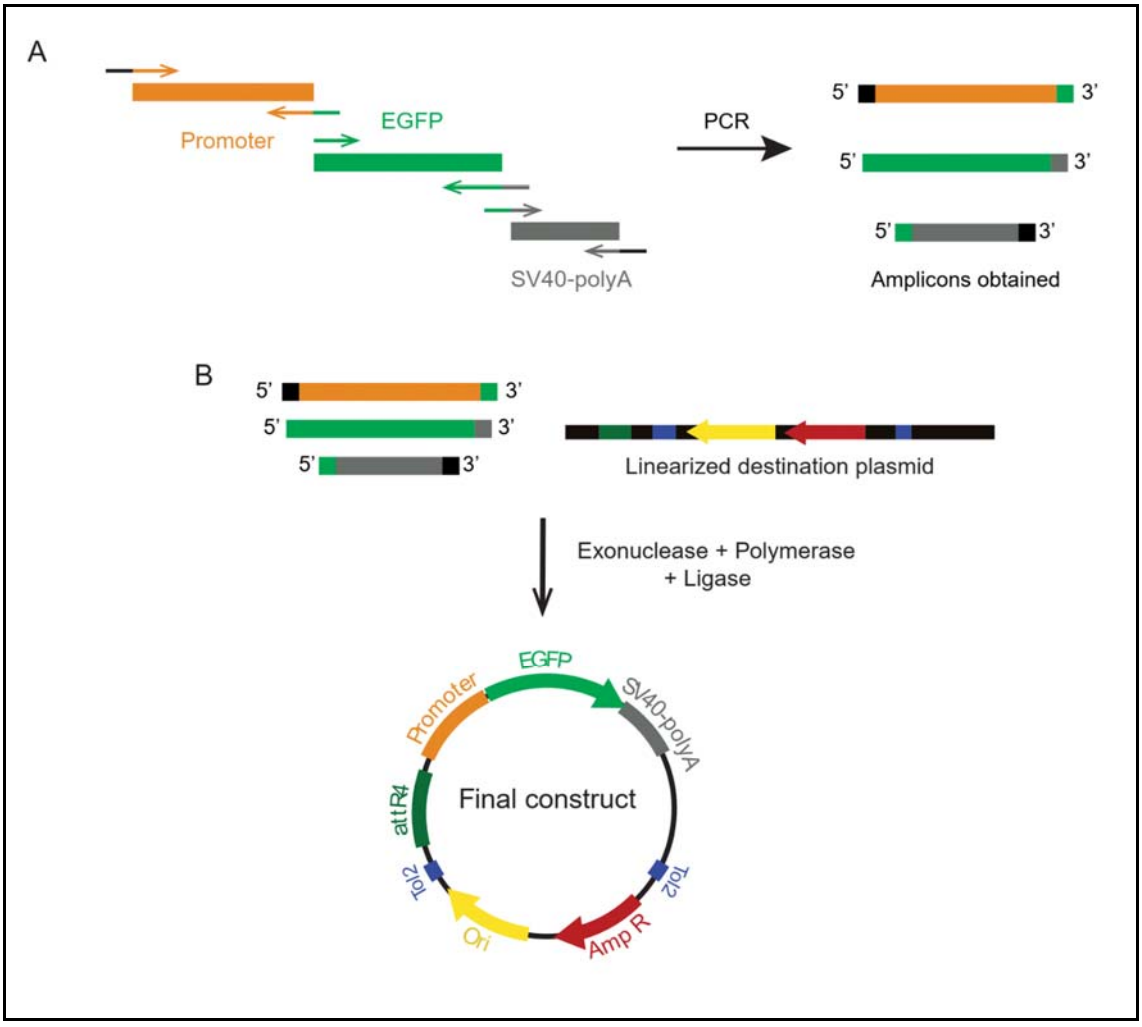


Figure A3.3 - **Promoter cloning and expression strategy. (A)** An initial PCR is performed to generate individual amplicons with overhangs that have homology with other parts that need to be stitched together. Sequence homology is colour-coded - black regions represent sequences from the destination plasmid; orange regions represent the putative promoter being cloned; green represents EGFP and grey represents the SV40 polyA signal needed for transcription. **(B)** Gibson Assembly-based strategy to stitch the amplicons from (A) to obtain final constructs containing Tol2 sequences (dark blue). Regions will be stitched together only because of homology, and thus only one possible construct can be obtained. attR4 (dark green) is a recombination site in the destination plasmid. Ori (yellow) is the origin of replication for the plasmid. AmpR (red) is the ampicillin resistance selection marker.

I next tested whether any of these could be used as cell type-specific drivers of gene expression. To do so, I microinjected the cloned construct along with Tol2 transposase mRNA and checked for expression at various developmental stages (Fig A3.4A) using green fluorescence as the reporter. To my disappointment, none of my constructs yielded expression in the hindbrain, where the inferior olive lies (Fig A3.4B, right). Interestingly, however, three of the constructs I injected (gstm3, hoxb5a and myhb) consistently showed green fluorescence in the heart (Fig A3.4B, right). The expression wasn't identical, with myhb constructs also showing strong expression in skeletal muscles (data not shown), which is consistent with what one might expect from a myosin heavy chain gene. Thus, while I seem to have successfully cloned putative promoters, I was unable to use it to drive expression specifically in the inferior olive.

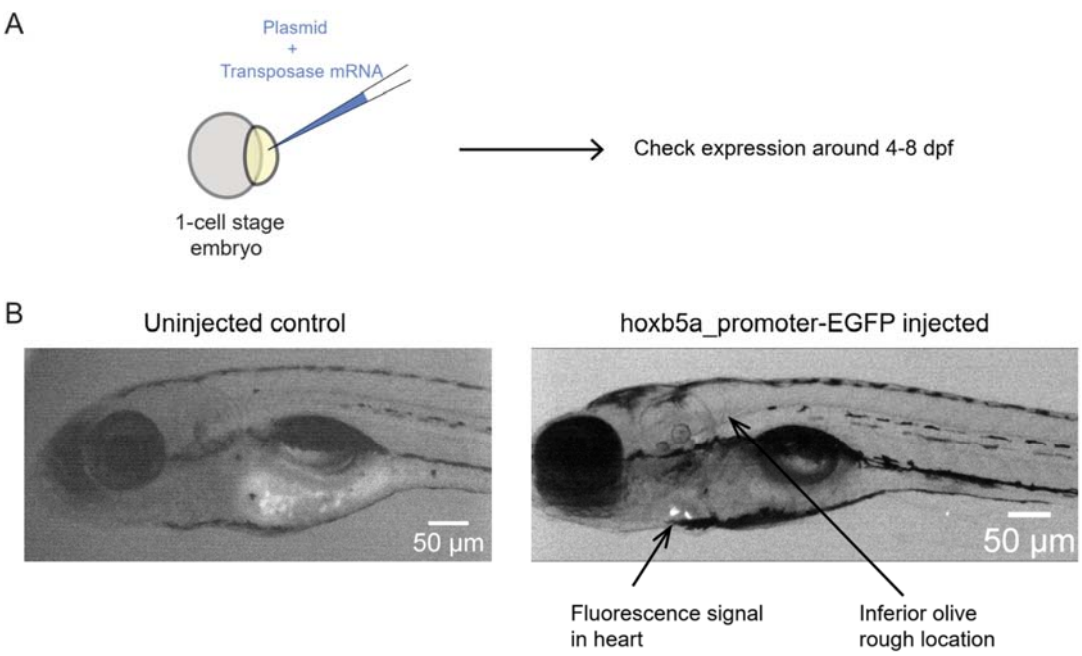


Figure A3.4 - Gene expression of putative cell-type specific promoter constructs. (A) Strategy of microinjection of plasmid constructs obtained (see Table A3.1 and Figure A3.3) **(B)** Overlaid green fluorescence and brightfield signals from uninjected control (left) and putative hoxb5a promoter injected animals (right) at 7dpf. Only autofluorescence in the gut was observed in uninjected animals. By contrast, a clear fluorescence signal was obtained in the heart in injected animals (right), marked with an arrow. However, no visible signal was observed in the inferior olive, the rough location of which in the hindbrain is marked with an arrow.

A3.2.3 There is no correlation between CF inputs and eurydendroid cell activity

While developing a tool to mark inferior olivary neurons, I also tried to address whether there was any functional connectivity between climbing fibers and eurydendroid cells. Since I could yet not directly stimulate CF inputs specifically, I thought of leveraging the large amplitude of CF inputs in Purkinje neurons as a fiduciary marker of inferior olive activity. Thus, I performed dual recordings of Purkinje neurons (PNs) and eurydendroid cells (ECs) (Fig A3.5A) to ask if there was any modulation of EC activity around CF inputs in PNs. I recorded from a total of 12 pairs of cells from 7 fish. This dataset was acquired from 9 unique ECs and 12 unique PNs, of which 2 PNs were in the tonic state and the remaining 10 were in the bursting state.

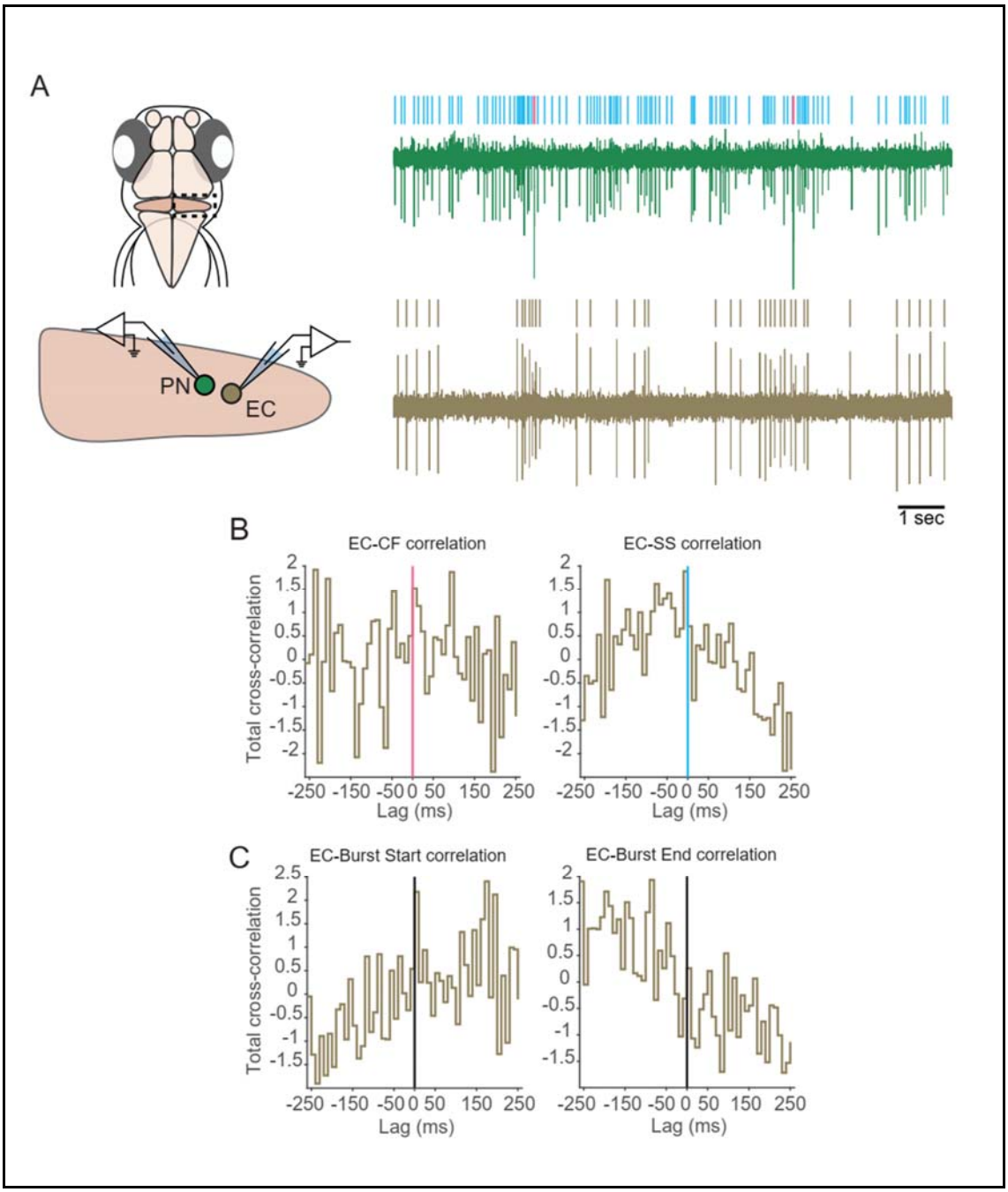

A
PN
EC
1 sec
B
EC-CF correlation
EC-SS correlation
Total cross-correlation
Lag (ms)
C
EC-Burst Start correlation
EC-Burst End correlation
Total cross-correlation
Lag (ms)

Figure A3.5 - **Testing functional olivocerebellar connectivity using paired Purkinje neuron-Eurydendroid cell recordings. (A)** Schematic of the recording configuration for dual PN-EC recordings. All recordings were in the loose patch mode, and an example dual recording is shown on the right, with the raw signal colour-coded by source (green for Purkinje neurons and brown for Eurydendroid cells). Detected events are shown as a raster above each raw trace, with PN simple spikes in cyan and CF inputs in magenta, and eurydendroid cell activity in brown. **(B)** Cross-correlation between EC activity relative to PN activity - CF inputs (left) and simple spikes (right). The start of CF inputs and simple spikes are shown as colour-coded vertical lines at 0. **(C)** Cross-correlation between EC activity relative to bursts of simple spikes in the simultaneously recorded Purkinje neuron in the bursting state, with correlation relative to the start of a burst on the left and the correlation relative to the end of a burst on the right. The reference point is marked as a black vertical line in each plot. Data for (B) and (C) come from n=12 pairs of cells from N=7 fish.

I first cross-correlated eurydendroid cell activity with CF inputs detected in Purkinje neurons. I found that there is no difference between EC activity before and after a CF input (Fig A3.5B, left). When cross-correlated with simple spiking activity in PNs, however, it was evident that EC activity is lower after a PN simple spike than before (Fig A3.5B, right). This is consistent with the fact that Purkinje neurons are inhibitory in nature. Inspecting raw data, it seemed that periods of PN and EC activity overlapped, so I looked deeper into the relationship between simple spikes and EC activity, by correlating PN bursts with EC activity. What I find is that, on average, EC activity is higher after the start of a burst (Fig A3.5C, left) and decreases after the end of a burst (Fig A3.5C, right).

Put together, what this suggests is that, on average, EC activity also occurs predominantly during bursts of simple spiking, although individual events are briefly suppressed when simple spiking occurs. Moreover, one can infer that there exists no direct functional connection between the inferior olive and eurydendroid cells in zebrafish at this developmental stage.

A3.3 Discussion

A3.3.1 Unsuccessful development of an inferior olive specific driver

Despite the strategic shortlisting of genes whose transcripts were highly expressed in cells in the 28C GAL4 line and the successful cloning of their putative promoter regions, I only

obtained expression of GFP in the heart. There could be several reasons for this result. Firstly, the RNASeq experiment performed for the various GAL4 lines used zebrafish at 14dpf, in order to get maximal RNA yield. I did not monitor the expression of GFP in microinjected animals beyond 8dpf. Thus, it is possible that expression in the inferior olive is visible only after this age. Another untested hypothesis is that the genes I shortlisted are transcribed in the inferior olive and then transported to the heart for protein expression. While it sounds unlikely, this kind of RNA transport to distant regions has been observed in plants (Liu & Chen, 2018). I would like to note that while I didn't get a marker for inferior olivary neurons, the constructs I generated could potentially be used to mark the heart - to monitor heart rate, or undertake other studies of cardiac function.

A3.3.2 Eurydendroid cell activity coincides with Purkinje neuron bursting, but is not modulated by CF inputs

Using paired recordings, I found that eurydendroid cell activity is modulated with respect to simple spikes, but that there is no modulation with respect to CF inputs. This suggests that there is no direct connection between the inferior olive and eurydendroid cells, unlike the case in mammals, where the inferior olive projects to both Purkinje neurons as well as deep cerebellar nuclear (DCN) neurons (Popa et al., 2019). This doesn't preclude the existence of an olivocerebellar loop, though, as the eurydendroid cells could still modulate the activity of inferior olivary neurons, which, in turn, affect the activity of Purkinje neurons. The existence of such a functional connection in fish hasn't yet been tested.

I would also like to acknowledge that this is a preliminary experiment and has a major drawback - that cells were chosen for recording only based on proximity, without information about whether or not they were functionally connected. I know that several Purkinje neurons converge onto a single eurydendroid cell. The most recent estimate of this convergence ratio suggests that 1-3 Purkinje neurons connect to an individual eurydendroid cell in larval zebrafish (Harmon et al., 2020). Given that the activity in ECs is reduced after individual simple spikes, consistent with the inhibitory nature of Purkinje neuron activity, I think that my samples do, indeed, include synaptic partners.

The absence of an olivocerebellar loop in zebrafish needs to be confirmed with more direct experiments with existing tools, such as (1) whole-cell voltage clamp recordings of eurydendroid cells with CF stimulation, or (2) recording the effect on CF inputs in PNs after ablation or optogenetic suppression of eurydendroid cell activity, which can be specifically achieved using the 156A GAL4 driver line (Takeuchi et al., 2015). The implications of the

absence of this olivocerebellar loop in zebrafish would be profound, as it would suggest that the feedback loop is an evolutionary novelty. It would then be interesting to ask what this feedback does for circuit function.

A3.4 Materials and Methods

A3.4.1 Genomic DNA extraction

Genomic DNA was extracted using the HotSHOT protocol (Meeker et al., 2007; Truett et al., 2000). First, 10 Indian WT zebrafish (*Danio rerio*) larvae at 9 dpf were put into a microfuge tube. Their water was drained and the larvae were washed with distilled water, which was also drained. 400 µl of alkaline lysis buffer (25 mM NaOH, 0.2 mM EDTA) was added to this tube, which was then incubated at 95ºC for 30 minutes and then cooled to 4ºC for 15 minutes. An equal volume (400 µl) of neutralisation buffer (40mM Tris-HCl, pH 5) was then added to this mixture. This crude DNA extract was incubated at 4ºC until use.

A3.4.2 Gateway cloning

For each of the 10 shortlisted genes of interest, two sets of primers were designed to capture slightly different regions that had all the features of a promoter. All the forward primers were designed with a KpnI restriction site added at the 5' end, and all the reverse primers were designed with an XhoI restriction site added at the 5' end. Table A3.2 lists all the primer sequences used. PCRs were performed using corresponding primer pairs (1F/1R or 2F/2R) only, using the high fidelity Phusion polymerase (NEB) to reduce the chances of mutations in the amplicons. 1 µl of the crude genomic DNA extract (~200 ng/µl) was used as the template.

Table A3.2 - List of primers used to clone putative promoters using restriction digestion and Gateway cloning. All forward primers have the suffix "F" and all reverse primers have the suffix "R". Restriction site recognition sites are highlighted in bold.

Primer name	Primer sequence (5' ➞ 3')
ampd1_AV_1F	TTAA**GGTACC**TTAAGCCATAACCGATTCGG
ampd1_AV_1R	TTAA**CTCGAG**GTAACTGGAGCCGTGTTC
ampd1_AV_2F	TTAA**GGTACC**CCCATTGTTGTTTTTATCATGG

ampd1_AV_2R	TTAA**CTCGAG**CTGGATATGAGGGGATTGGT
epyc_AV_1F	TTAA**GGTACC**TCAAGATGCAGATGAGAGG
epyc_AV_1R	TTAA**CTCGAG**ACCCAAACAAGTACCTGTTT
epyc_AV_2F	TTAA**GGTACC**AACAAGGGGTTACATGTGC
epyc_AV_2R	TTAA**CTCGAG**GGGATGTTTTTACAAGCTTCG
fap_AV_1F	TTAA**GGTACC**GACAGGTGTATGCCTTTT
fap_AV_1R	TTAA**CTCGAG**CTTCAGCTGCGTTTCTCT
fap_AV_2F	TTAA**GGTACC**TTACGTGTGATTTGTCAGC
fap_AV_2R	TTAA**CTCGAG**AGACGAGCGCTTATAACTCA
gstm3_AV_1F	TTAA**GGTACC**CTAAATAAAGAGCACAGATTTGTTCC
gstm3_AV_1R	TTAA**CTCGAG**GGAGATGAATGGGCGC
hoxb5a_AV_1F	TTAA**GGTACC**AAAAATTCAATTGTACGGAAAATG
hoxb5a_AV_1R	TTAA**CTCGAG**CTATGTATGCGGCCAAATAT
hoxb5a_AV_2F	TTAA**GGTACC**ATTGAGTCTGGGGGAAAAG
hoxb5a_AV_2R	TTAA**CTCGAG**TAGCAATGCACTGTACTTCG
itih6_AV_1F	TTAA**GGTACC**ATATTACAAAATGGAGGTTGG
itih6_AV_1R	TTAA**CTCGAG**CTGGGGAAAAAAACTGTTTT
itih6_AV_2F	TTAA**GGTACC**GTTGATAAGATGGTTACTTTGC
itih6_AV_2R	TTAA**CTCGAG**CCTAAAAATGATACCGAAGCC
myhb_AV_1F	TTAA**GGTACC**CTTGTTTCATCCCCTGGA
myhb_AV_1R	TTAA**CTCGAG**CTGTGTATAGCATAAACACTCA
myhb_AV_2F	TTAA**GGTACC**TGCAACTTTAGACAAGTCAC
myhb_AV_3F	TTAA**GGTACC**GAGTATTAGTAAACCCTCTGC

serpinf1_AV_1F	TTAA**GGTACC**CACAGTATTTACTTGTGGATG
serpinf1_AV_1R	TTAA**CTCGAG**CAAAGGATGGGTTAATAAGAAC
serpinf1_AV_2F	TTAA**GGTACC**ACCATAGAGAGTGGCAAC
serpinf1_AV_2R	TTAA**CTCGAG**CCTCCGCCAACTGTG
tnnc2_AV_1F	TTAA**GGTACC**CTGCTGGAAAGTATAGGGT
tnnc2_AV_1R	TTAA**CTCGAG**GGAGGAACTGATGAAAGACA
tnnc2_AV_2F	TTAA**GGTACC**GGGTTTCTTGGCTTGACT
tnnc2_AV_2R	TTAA**CTCGAG**CCCCAACAGAAAGACTAGAC
tnxb_AV_1F	TTAA**GGTACC**CATTTGTTTTGTTAGTTGTGG
tnxb_AV_1R	TTAA**CTCGAG**TTTGGAGCTGCTATCCTG
tnxb_AV_2F	TTAA**GGTACC**AACAGAATAAAACAAACCTCTG
tnxb_AV_2R	TTAA**CTCGAG**ACTTTTCAAGCAAAACAGAG

A3.4.3 Gibson Assembly

A3.4.3.1 Primer design - Putative Promoter

Primers were designed with overhangs as depicted in Fig A3.3. The target/destination vector used was the pDestTol2CG2 from the Tol2 kit. pDestTol2CG2 is a plasmid containing Tol2 sites, as well as the attR3 and attR4 recombination sites used for Gateway cloning. Between these two sites is the ccdB bacterial toxin gene. Double digestion with the SalI and KpnI enzymes excised this region (excluding the attR4 site) and also linearised the plasmid. The original plasmid was 7796bp long, while the linearised plasmid was 6198 bp long.

A 30 bp-long region from each end of the linearised plasmid was taken as the region of homology to be added to the appropriate primer. For the forward primer of the putative promoter, this sequence was 5'-TGTAAAACACAACATATCCAGTCACTATGG-3', which corresponded to the attR4 end of the linearised destination vector. The reverse primer of the promoter had the sequence 5'-GAACAGCTCCTCGCCCTTGCTCACCATGGT-3' appended to it, which is the reverse complement of the beginning of the EGFP coding sequence. The complete list of primers used for each putative promoter is shown in Table A3.3. Primers

were designed only for those regions that could be amplified using the restriction digestion strategy. The numbering of the Gibson Assembly primers in Table A3.3 matches the numbering of the Gateway cloning primers in Table A3.3, which indicates a mapping back to different regions of the genome.

Table A3.3 - List of primers used to clone putative promoters. Regions of homology are highlighted in bold - EGFP (green) and destination vector backbone (black).

Primer name	Primer sequence (5' → 3')
epyc_GA_AV_2F	**TGTAAAACACAACATATCCAGTCACTATGG**AACAAGGGGT TACATGTGC
epyc_GA_AV_2R_1	**GAACAGCTCCTCGCCCTTGCTCACCATGG**TGGGATGTTTT TACAAGCTTCG
gstm3_GA_AV_1F	**TGTAAAACACAACATATCCAGTCACTATGG**CTAAATAAAGA GCACAGATTTGTTCC
gstm3_GA_AV_1R_1	**GAACAGCTCCTCGCCCTTGCTCACCATGG**TGGAGATGAAT GGGCGC
hoxb5a_GA_AV_2F	**TGTAAAACACAACATATCCAGTCACTATGG**ATTGAGTCTG GGGGAAAAG
hoxb5a_GA_AV_2R_1	**GAACAGCTCCTCGCCCTTGCTCACCATGG**TTAGCAATGCA CTGTACTTCG
itih6_GA_AV_1F	**TGTAAAACACAACATATCCAGTCACTATGG**ATATTACAAAA TGGAGGTTGG
itih6_GA_AV_1R_1	**GAACAGCTCCTCGCCCTTGCTCACCATGG**TCTGGGGAAAA AAACTGTTTT
myhb_GA_AV_1F	**TGTAAAACACAACATATCCAGTCACTATGG**CTTGTTTCATC CCCTGGA
myhb_GA_AV_1R_1	**GAACAGCTCCTCGCCCTTGCTCACCATGG**TCTGTGTATAG CATAAACACTCA

tnnc2_GA_AV_1F	**TGTAAAACACAACATATCCAGTCACTATGG**CTGCTGGAAA GTATAGGGT
tnnc2_GA_AV_1R_1	**GAACAGCTCCTCGCCCTTGCTCACCATGG**TGGAGGAACTG ATGAAAGACA

A3.4.3.2 Primer design - EGFP and SV40pA

The two remaining parts of the complete insert are the reporter (EGFP) and the SV40 polyA tail signal, which is a necessary signal for transcription *in vivo*. These primers were common for all the promoter regions being cloned, and were designed to have regions of homology as shown in Fig A3.3. The primers had the following sequences: EGFP (forward) - 5'-ATGGTGAGCAAGGGCGAGGAG-3'. EGFP (reverse) 5'-CTCATCAATGTATCTTATCATGTCTGGATCCTATAGGGCTGCAGAATCTAGAGG-3'. The reverse primer has a 30 bp homology region with the SV40 polyA tail signal. The remaining part of this sequence is part of the pME_EGFP plasmid from the Tol2 kit, from which the EGFP region was amplified. The SV40pA forward primer was designed to have homology with the 3' end of the EGFP sequence and the SV40pA reverse primer was designed to have homology with the appropriate region of the linearized destination vector. The primer sequences are as follows: SV40pA (forward) - 5'-AGCCTCTAGATTCTGCAGCCCTATAGGGATCCAGACATGATAAGATACATT-3'. SV40pA (reverse) - 5'-GTATCTTATCATGTCTGGATCATCATCGATGTAACTTGTTTATTGCAGCTTATAATGG-3'.

A3.4.3.3 PCR Amplification and Component Purification

1 µl of crude genomic DNA extract (~200 ng/µl) or source plasmid (5 ng/µl) was used as the template. PCRs were performed using corresponding primer pairs, using the high fidelity Phusion polymerase (NEB). Gradient PCRs were performed to test annealing temperatures in the range of 50-74^{o}C. All reactions were run on a 1% agarose gel to check whether reactions worked and yielded specific amplicons. The band of interest was then cut and DNA was purified from the gel using the QIAGEN Gel Extraction Kit.

A3.4.3.4 Gibson Assembly

The four component parts - putative promoter, reporter gene (EGFP), polyA signal sequence and vector backbone - could not be stitched together in a single reaction. Thus, the assembly was done piecewise, by first stitching the reporter gene and SV40 polyA signal, which was the common factor across all target promoters to be cloned. This stitched sequence was then joined with the putative promoter sequence to obtain the full-length

insert. Finally, the insert and vector backbone were stitched together in the final assembly step. Each assembly step was performed by mixing each component (in a ~1:1 ratio) with the 2X Gibson Assembly master mix (NEB) to a final volume of 10µl. The reaction was incubated at 50ºC for at least 1 hour and the whole 10µl was transformed into competent *E. coli* DH5alpha cells, which were then plated onto media containing ampicillin at a final concentration of 100µg/ml and incubated at 37ºC overnight. Any colonies that were obtained were screened using colony PCR or by isolating plasmid and double digesting it with BamHI and PvuI. The BamHI site is introduced by Gibson Assembly (between the EGFP and SV40pA segments), so positive colonies would yield two bands, and negative colonies only a single band. Plasmid isolations were performed using the QIAGEN Miniprep Kit.

A3.4.3.5 Preparation of competent cells

A single colony of *E. coli* DH5alpha is inoculated into 50 ml of autoclaved LB media and incubated at 37ºC for 16 hours. 1ml of this primary culture is then inoculated into 100 ml of fresh autoclaved LB broth for 2.5-3 hours, until an OD of 0.35-0.4 has been reached. This culture is then centrifuged at 3000 rpm for 7 minutes at 4ºC. The supernatant is discarded and the culture is resuspended in ice-cold autoclaved 60mM $CaCl_2$. The cells are incubated at 4ºC for 20 minutes before being centrifuged at 2500 rpm for 5 minutes at 4ºC. Cells are then resuspended in 2 ml of sterile 60mM $CaCl_2$ with 15% glycerol. This mixture is then aliquoted (50µl) into autoclaved microfuge tubes and stored at -80ºC until use.

A3.4.3.6 Transformation into competent cells

An aliquot of competent cells is first thawed on ice. Cells are then incubated on ice with the DNA to be transformed for 20-30 minutes. A heat shock is delivered by incubating cells at 42ºC in a water bath for 45 seconds and then immediately transferring them back to ice for ~2 minutes. Cells are allowed to recover from this shock by incubating them at 37ºC in 250-1000µl of fresh antibiotic-free LB broth for 45 minutes to 1 hour. Cells are then plated onto LB agar containing the appropriate antibiotic needed for selection (here either 100µg/ml kanamycin or 100µg/ml ampicillin).

A3.4.4 Microinjection and screening

Adult male and female zebrafish are placed in a breeding chamber (Tecniplast, Italy) containing system water the night prior to microinjection. Males and females are separated by a barrier. On the morning of microinjection, breeding is timed by lifting the barrier. Zebrafish eggs are externally fertilised, and are collected at the bottom of the breeding chamber. The time to first division is ~45 minutes (Kimmel et al., 1995), which is typically the window of microinjection to obtain transgenic animals. The collected eggs are rinsed with E3

medium (composition in mM: 5 NaCl, 0.17 KCl, 0.33 $CaCl_2$, and 0.33 $MgSO_4$, pH 7.8) and then aligned against a glass slide in an orientation that allows them to be approached by a microinjection needle.

Microinjection needles are pulled from thin glass capillaries (OD 1.0mm and ID 0.56mm; Warner Instruments) using a Flaming-Brown P-97 pipette puller (Sutter Instruments, Novato, CA, United States). They are backfilled with a freshly-made mixture of Tol2 transposase mRNA (100 µg/ml final concentration) and the plasmid of interest (25-50 µg/ml final concentration). The needle is then connected to a Picospritzer III (Parker Hannifin Corporation, Hollis, NH, USA), which is a pressure delivery system that allows the controlled ejection of picolitre volumes. The tip of the pulled microinjection needle is sealed, and must be broken using forceps. The settings of the Picospritzer are then adjusted in order to deliver a droplet of liquid ~0.5nl in volume. The needle is mounted onto a manual NM-3 micromanipulator on a GJ-1 magnetic stand (Narishige Instruments), which is used to bring the needle close to the embryos, under visual guidance using an Olympus SZ51 microscope. The needle is then pierced through the chorion and brought into contact with the cell. A pulse of liquid is then delivered directly into the cell, after which the needle is retracted and the process repeated for the remaining embryos.

Post-injection, embryos are kept in E3 medium in an incubator at 28^{o}C. Their development is monitored over subsequent days and compared to uninjected controls of the same batch. Dead or unfertilised embryos are cleared out when needed. At the time of screening, individual larvae are placed in separate droplets on the cover of a 90mm Petri plate (Tarsons). Each larva is observed under an Olympus SZX16 dissection microscope with epifluorescence illumination, and are separated based on their expression pattern.

A3.4.5 Paired recordings from eurydendroid cells

The first recording of a pair of juxtacellular recordings was obtained using the procedure as described in Chapter 2. After obtaining the first recording, the objective was lifted and a second electrode was brought into focus and manoeuvred to the sample. A second cell close to the first recording was targeted for recording, trying to minimise tissue disruption. The identity of a cell as a eurydendroid cell was assigned based on its electrical activity. For a subset of recordings (in 156A transgenic animals), eurydendroid cell identity could be confirmed with GFP expression in the recorded cell. Once both channels produced stable signals, 1 minute and 10 minute-long gap-free recordings were obtained before a different cell was targeted for recording.

A3.5 References

Ahrens, M. B., Li, J. M., Orger, M. B., Robson, D. N., Schier, A. F., Engert, F., & Portugues, R. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature*, *485*(7399), Article 7399. https://doi.org/10.1038/nature11057

Bae, Y.-K., Kani, S., Shimizu, T., Tanabe, K., Nojima, H., Kimura, Y., Higashijima, S., & Hibi, M. (2009). Anatomy of zebrafish cerebellum and screen for mutations affecting its development. *Developmental Biology*, *330*(2), 406–426. https://doi.org/10.1016/j.ydbio.2009.04.013

Bagnall, M. W., Zingg, B., Sakatos, A., Moghadam, S. H., Zeilhofer, H. U., & du Lac, S. (2009). Glycinergic projection neurons of the cerebellum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(32), 10104–10110. https://doi.org/10.1523/JNEUROSCI.2087-09.2009

Eccles, J. C., Llinás, R., & Sasaki, K. (1966). The excitatory synaptic action of climbing fibres on the Purkinje cells of the cerebellum. *The Journal of Physiology*, *182*(2), 268–296. https://doi.org/10.1113/jphysiol.1966.sp007824

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), Article 5. https://doi.org/10.1038/nmeth.1318

Harmon, T. C., McLean, D. L., & Raman, I. M. (2020). Integration of Swimming-Related Synaptic Excitation and Inhibition by olig2+ Eurydendroid Neurons in Larval Zebrafish Cerebellum. *Journal of Neuroscience*, *40*(15), 3063–3074. https://doi.org/10.1523/JNEUROSCI.2322-19.2020

Hibi, M., & Shimizu, T. (2012). Development of the cerebellum and cerebellar neural circuits. *Developmental Neurobiology*, *72*(3), 282–301. https://doi.org/10.1002/dneu.20875

Horn, K. M., Deep, A., & Gibson, A. R. (2013). Progressive limb ataxia following inferior olive lesions. *The Journal of Physiology*, *591*(22), 5475–5489. https://doi.org/10.1113/jphysiol.2012.234898

Ikenaga, T., Yoshida, M., & Uematsu, K. (2006). Cerebellar efferent neurons in teleost fish. *The Cerebellum*, *5*(4), 268–274. https://doi.org/10.1080/14734220600930588

Judd, E. N., Lewis, S. M., & Person, A. L. (2021). Diverse inhibitory projections from the cerebellar interposed nucleus. *ELife*, *10*, e66231. https://doi.org/10.7554/eLife.66231

Kani, S., Bae, Y. K., Shimizu, T., Tanabe, K., Satou, C., Parsons, M. J., Scott, E., Higashijima, S. I., & Hibi, M. (2010). Proneural gene-linked neurogenesis in zebrafish cerebellum. *Developmental Biology*, *343*(1–2), 1–17. https://doi.org/10.1016/j.ydbio.2010.03.024

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages

of embryonic development of the zebrafish. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *203*(3), 253–310. https://doi.org/10.1002/aja.1002030302

Liu, L., & Chen, X. (2018). Intercellular and systemic trafficking of RNAs in plants. *Nature Plants*, *4*(11), Article 11. https://doi.org/10.1038/s41477-018-0288-5

Meek, J., Yang, J. Y., Han, V. Z., & Bell, C. C. (2008). Morphological analysis of the mormyrid cerebellum using immunohistochemistry, with emphasis on the unusual neuronal organization of the valvula. *Journal of Comparative Neurology*, *510*(4), 396–421. https://doi.org/10.1002/cne.21809

Meeker, N. D., Hutchinson, S. A., Ho, L., & Trede, N. S. (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *BioTechniques*, *43*(5), 610–614. https://doi.org/10.2144/000112619

Popa, L. S., Streng, M. L., & Ebner, T. J. (2019). Purkinje Cell Representations of Behavior: Diary of a Busy Neuron. *The Neuroscientist*, *25*(3), 241–257. https://doi.org/10.1177/1073858418785628

Prekop, H.-T., Kroiss, A., Rook, V., Zagoraiou, L., Jessell, T. M., Fernandes, C., Delogu, A., & Wingate, R. J. T. (2018). Sox14 Is Required for a Specific Subset of Cerebello-Olivary Projections. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(44), 9539–9550. https://doi.org/10.1523/JNEUROSCI.1456-18.2018

Takeuchi, M., Matsuda, K., Yamaguchi, S., Asakawa, K., Miyasaka, N., Lal, P., Yoshihara, Y., Koga, A., Kawakami, K., Shimizu, T., & Hibi, M. (2015). Establishment of Gal4 transgenic zebrafish lines for analysis of development of cerebellar neural circuitry. *Developmental Biology*, *397*(1), 1–17. https://doi.org/10.1016/j.ydbio.2014.09.030

Takeuchi, M., Yamaguchi, S., Sakakibara, Y., Hayashi, T., Matsuda, K., Hara, Y., Tanegashima, C., Shimizu, T., Kuraku, S., & Hibi, M. (2017). Gene expression profiling of granule cells and Purkinje cells in the zebrafish cerebellum. *Journal of Comparative Neurology*, *525*(7), 1558–1585. https://doi.org/10.1002/cne.24114

Truett, G. e., Heeger, P., Mynatt, R. l., Truett, A. a., Walker, J. a., & Warman, M. l. (2000). Preparation of PCR-Quality Mouse Genomic DNA with Hot Sodium Hydroxide and Tris (HotSHOT). *BioTechniques*, *29*(1), 52–54. https://doi.org/10.2144/00291bm09