IMMEDIATE-EARLY GENE HOMER1A INTRANUCLEAR TRANSCRIPTION FOCUS INTENSITY AS A MEASURE OF RELATIVE NEURAL ACTIVATION

WING KAR LI WITHARANA
Bachelor of Science, University of Alberta, 2009
Master of Science, University of Lethbridge, 2012

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfilment of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

Department of Neuroscience
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

IMMEDIATE-EARLY GENE HOMER1A INTRANUCLEAR TRANSCRIPTION FOCUS INTENSITY AS A MEASURE OF RELATIVE NEURAL ACTIVATION

WING KAR LI WITHARANA

Date of Defence: November 8, 2018

Dr. Bruce L. McNaughton Supervisor	Professor	Ph.D.
Dr. Robert J. Sutherland Thesis Examination Committee Member	Professor	Ph.D.
Dr. Ute Wieden-Kothe Thesis Examination Committee Member	Professor	Ph.D.
Dr. Majid H. Mohajerani Internal External Examiner	Associate Professor	Ph.D.
Dr. Diano F. Marrone External Examiner Wilfred Laurier University Waterloo, Ontario	Associate Professor	Ph.D.
Dr. Artur Luczak Chair, Thesis Examination Committee	Professor	Ph.D.

DEDICATION

For my children: I hope you grow up to be open, critical, and deep thinkers. I hope you learn to value intellect and scientific reason. Above all, I hope you understand that the gift of education is not available to everyone in the world and to never take it for granted.

For mothers everywhere: May you never lose sight of your own dreams, especially during the grueling seasons of motherhood.

ABSTRACT

Although immediate-early gene expression analyses using fluorescent *in situ* hybridization is an effective method to identify recently activated neurons; and non-Boolean variations in transcription foci have been documented, it remains unclear whether there is a systematic relationship between magnitude of neural activation and corresponding RNA signal. Here, we quantified the *Homer1a* response of hippocampal neurons in rats that ran laps on a familiar track to induce consistent cell firing. A strong linear trend ($r^2 > 0.9$) in INF intensity (brightness) was observed between 1 and 25 laps, after which INF signal dispersed within the nucleus. When the integrated intranuclear fluorescence was considered instead, the linear relationship extended to 50 laps. But there was only an approximate doubling of *Homer1a* RNA detected for this 50-fold variation in total spiking. Thus, this low-gain dynamic range likely precludes INF intensity as a precise quantitative readout of neural activation, albeit a useful qualitative tool.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Bruce McNaughton, for over 9 years of academic mentorship and scientific training. Dr. McNaughton's consistently high standard of excellence has been valuable in propelling me to strive for quality and thoroughness in my scientific endeavours. I am deeply grateful and honoured to have received such prestigious guidance. Also, I am indebted to my supervisory committee members: Dr. Robert Sutherland, Dr. Robert McDonald, and Dr. Ute Wieden-Kothe for their time and energy in providing insightful evaluations of my thesis work. Thank you to Dr. Ben Clark who was an important source of encouragement and mentorship in our joint efforts on this project. I would also like to acknowledge Jeanne Xie and Sutherland Dubé for their valuable moral support and friendship through the years of rigour. Thank you to Valérie Lapointe and Aubrey Demchuk for their technical expertise with tissue processing, imaging, and seemingly endless trouble-shooting. In terms of logistics, this project would not have been possible without the conscientious and excellent administrative management provided by Amanda Mauthe-Kaddoura and Naomi Cramer who were consistently reliable and dependable. I have been influenced by and received assistance from many, many people at the University of Lethbridge and the Canadian Centre for Behavioural Neuroscience, and for this I will be eternally grateful.

On a personal level, I want to thank my husband, Dr. Dinesh Witharana, for constantly pushing me to pursue my goals even when it meant we had to make substantial sacrifices, or be apart for long stretches of time. I am forever grateful for my parents, Yuen Leung and Michael Li, for bringing me to Canada for a better life, and now helping

raise my children so I can still pursue education. Lastly, thank you to my friends and family who never understood what I was studying but still cared and supported me anyways.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	2-dimensional
3D	3-dimensional
AC	adenylyl cyclase
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
Arc	Activity-regulated cytoskeleton-associated protein
ATP	adenosine triphosphate
B_i	minimum blue intensity
C	Celsius
CA1	Cornu Ammonis 1
CA3	Cornu Ammonis 2
CaM	calmodulin, calcium-modulated protein
CaMK-II	calcium/calmodulin-dependent protein kinase type II
CaMKIV	calcium/calmodulin-dependent protein kinase type IV
cAMP	cyclic adenosine monophosphate
catFISH	compartment analysis of temporal activity for fluorescent *in situ* hybridization
CC	coiled-coil domain
CCD	charge-coupled device
CDK9	cyclin-dependent kinase 9
CRE	cAMP response element
CREB	cyclic AMP response element-binding protein
DAPI	4',6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
DRG	dorsal root ganglion
EGFP	enhanced green fluorescent protein
ERK	extracellular-signal regulated kinase
EVH1	Enabled/Vasp homology
FISH	fluorescent *in situ* hybridization
FITC	fluorescein isothiocyanate
GPCR	G-protein coupled receptor
GKAP	guanylate kinase-associated protein
H1a	*Homer1a*
HRP	horseradish peroxidase
IEG	immediate-early gene
INF	intranuclear foci
IP3R	inositol-1,4,5-triphosphate receptor
JPEG	joint photographic experts group
kb	kilobase
kDa	kilodalton
LTP	long-term potentiation
MAPK	mitogen-activated protein kinase
MECS	maximal electroconvulsive shock
MEF2	myocyte enhancer factor-2
MEK	mitogen-activated protein kinase

mGluR	metabotropic glutamate receptor
μOhms	microOhms
MSK	mitogen and stress-activated kinase
NDPI	NanoZoomer Digital Pathology Image
NMDA	N-methyl-D-aspartate
Npas4	Neuronal PAS 4
NR2	N-methyl-D-aspartate receptor 2
PAS	Per-Arnt-Sim
PCR	polymerase chain reaction
PDZ	post-synaptic density/*Drosophila* disc large tumor suppressor/ zonula occludens-1
P_i	peak intensity
PKA	protein kinase A
PMT	photomultiplier
PSD95	post-synaptic density protein 95
px	pixels
RGB	red, green, blue
RNA	ribonucleic acid
RNAP	ribonucleic acid polymerase
RNAPII	ribonucleic acid polymerase II
ROI	region of interest
RTK	receptor tyrosine kinase
SARE	synaptic activity-responsive element
SD	standard deviation
SEM	standard error of the mean
SRF	serum response factor
SSC	saline sodium citrate
TDI	time-delayed integration
T_i	minimum green intensity (threshold)
TIFF	Tagged Image Format File
TSA	tyramide signal amplification
UTR	untranslated region
VASP	vasodilator-stimulated phosphoprotein
VGCC	voltage-gated calcium channels

IMMEDIATE-EARLY GENE HOMER1A INTRANUCLEAR TRANSCRIPTION FOCUS INTENSITY AS A MEASURE OF RELATIVE NEURAL ACTIVATION

1 INTRODUCTION

1.1 Immediate-early gene expression analyses can identify recently activated neurons

In the brain, immediate-early genes (IEGs) are characterized by their rapid and transient expression which can be induced almost instantaneously by neuronal activation (Link et al., 1995; Flavell & Greenberg, 2008). IEGs can be functionally classified as regulatory transcription factors or effector IEGs (Kubik, Miyashita, & Guzowski, 2007). Regulatory transcription factors such as *c-fos*, *c-jun*, and *zif268* encode proteins that control transcription of other downstream genes (Clayton, 2000; Guzowski, 2002). Effector IEGs such as *Arc* (activity-regulated cytoskeleton-associated protein, also known as *Arg 3.1*), and *Homer1a* (also known as *H1a* or *vesl-1s*), encode proteins that directly regulate cellular functions and contribute to synaptic plasticity (Lyford et al., 1995; Brakeman et al., 1997).The IEGs *Arc* & *Homer1a* exhibit upregulated transcription in neurons in direct response to neuronal activation associated with behavioural tasks such as exploring a novel environment (Guzowski, McNaughton, Barnes, & Worley, 1999; Witharana et al., 2016; Chawla, Sutherland, Olson, McNaughton, & Barnes, 2018), contextual fear conditioning (Inoue et al., 2009; Minatohara, Akiyoshi, & Okuno, 2016), operant-response learning (Kelly & Deadwyler, 2003), or running laps around a track (Miyashita, Kubik, Hahighi, Steward, & Guzowski, 2009; Burke et al., 2005). This behaviourally-relevant IEG upregulation occurs in principal cells in many brain regions, including the cortex, amygdala and hippocampal subregions (Burke et al., 2005;

Imamura, Nonaka, Yamamoto, Matsuki, & Nomura, 2011). The proteins encoded by these effector IEGs may be implicated in establishing long-lasting cellular and synaptic changes that are associated with memory formation (Flavell & Greenberg, 2008; Adams & Dudek, 2005; Lanahan & Worley, 1998). IEG expression has also been shown to persist during post-experience reactivation of neuronal populations, suggestive of their role in consolidation of new memories (Marrone, Schaner, McNaughton, Worley, & Barnes, 2008).

By analyzing the RNA transcription of these activity-regulated IEGs through fluorescent *in situ* hybridization (FISH), neuroscientists have developed a useful method to map behaviourally-relevant neuronal populations. Experiments can be designed such that animals are sacrificed at a specific time-point after behaviourally-associated IEG transcription. For example, brief and ongoing transcription of *Arc* RNA induced by the behaviour immediately preceding the animal's death (about 2- 5 min) can be detected as bright probe-labeled foci within the neuronal nuclei (*Arc*-INF) of activated cells, while activation induced 20-25 min prior to sacrifice can be observed as diffuse labeling outside the nucleus, typically in the cytoplasm (*Arc*-cyto). This method of detecting activated neuronal populations by measuring anatomically distinct RNA signal is known as cellular compartment analysis of temporal activity by *in situ* hybridization, or catFISH, and was initially applied to monitor hippocampal populations in response to novel environmental exposures by John Guzowski and colleagues (Guzowski et al., 1999). In this original paradigm, rats were exposed to two behavioural epochs (duration of 5 min each), separated by a 20 min rest period. The population of neurons activated in the earlier epoch (activated 20 – 25 min prior to sacrifice) would be labeled with *Arc*-cyto and the

population of neurons activated in the second epoch (activated 2-5 min prior to sacrifice) would be labeled with *Arc*-INF (see also Chawla et al., 2005).

One variant of catFISH only detects intranuclear RNA signal but differentiates populations based on the transcription of two different IEGs, *Arc* and *Homer1a.* It has been shown that both *Arc* & *Homer1a* are co-expressed in similar populations of activated cells in the hippocampus and neocortex, and their expression are both induced by similar stimuli, such as novel context experiences (Vazdarjanova, McNaughton, Barnes, Worley, & Guzowski, 2002). *Homer1a (H1a)* mRNA is generated from a relatively long transcript (~45 kb, Bottai et al., 2002), whereas the *Arc* transcript is much shorter (~3.5 kb, Lyford et al., 1995). This length discrepancy between the mRNA transcripts of these two genes permits the targeting of the 3'-untranslated region (UTR) of the *H1a* RNA to identify neuronal populations activated 25-40 min before sacrifice in conjunction with the targeting of *Arc* to identify neurons that were activated between 2-5 min prior to sacrifice (Vazdarjanova et al., 2002; Vazdarjanova & Guzowski, 2004). Another adaptation of catFISH that permits the detection of cells activated in a third time-point, about 60 min prior to sacrifice, adds the tracking of cytoplasmic *H1a*, the RNA of which takes about 60 min to be export out of the nucleus (Marrone et al., 2008; Bottai et al., 2002).

Through these precisely timed protocols, IEG image analyses provide striking temporal and spatial resolution for the detection of activated neurons in large populations. Effectively, fluorescent labeling of RNA transcribed from induced IEGs can generate a whole-brain snapshot of time-stamped neural activation (for further reveiws: Guzowski,

McNaughton, Barnes, & Worley, 2001; Guzowski, Timlin, Roysam, McNaughton, Worley, & Barnes, 2005; Kubik et al., 2007).

After fluorescent *in situ* hybridization (FISH), the labeled RNA can appear as two bright intranuclear transcription foci (INF), or as diffuse cytoplasmic staining (Guzowski et al., 1999). Notably, some cells express only one focus as a result of mono-allelic expression, although this stochastic process only occurs in as low as 0.5% of autosomal genes (Eckersley-Maskin & Spector, 2014) and bi-allelic expression is typically the default process for IEGs based on empirical observations. Transcription foci detected in the nucleus represent the first compartment in which ongoing rounds of RNA transcription are concentrated in a small cellular region, and can typically appear as compact and tightly bound puncta of fluorescent signal (Figure 1.1). INFs may correspond to 45-100 nm diameter "transcription factories" composed of RNA polymerases and transcription factors (Mitchell & Fraser, 2008). It has been proposed that these "factories" are maintained as stable compartments for long periods of time (Iborra, Pombo, Jackson, & Cook, 1996; Wansink, Schul, Van Der Kraan, Van Steensel, Van Driel, & De Jong, 1993; Ghamari et al., 2013), or these factories may correspond to temporary clusters of RNA polymerase in which transcription occurs (Cisse et al., 2013; Zhao et al., 2014). The current conflicting views regarding the intranuclear spatial distribution and clustering dynamics of RNA polymerase and associated transcription factors will be discussed further in section 1.7. However, in any of the proposed models of RNA polymerase distribution, it is reasonable to assume that while IEGs undergo repeated rounds of transcription, the RNA molecules produced on each round can accumulate as conglomerates of transcripts before exportation to the cytoplasm. The

number of RNA copies would therefore affect the intensity of the RNA-based fluorescent signal. Furthermore, the number of RNA copies generated may be regulated by the intensity of the transcription induction signal conveyed by the intensity of neural activation based on stoichiometric properties of electro-transcriptional coupling, which will be examined in detail in section 1.5. Therefore, it would be useful to examine the strength of the relationship between spiking and INF intensity, since this would determine whether IEG fluorescence can be used as a histological readout of relative magnitudes of activation within a neural population.

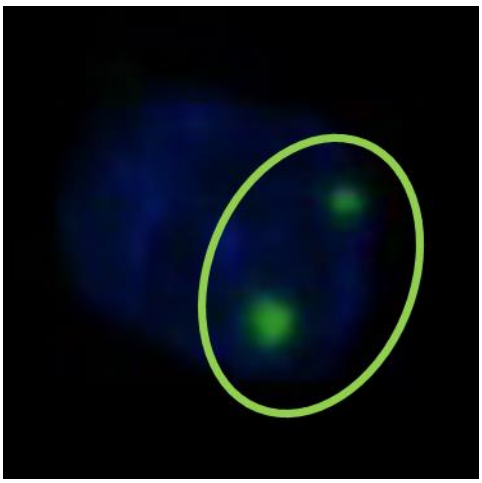


Figure 1.1 Sample fluorescent image depicting a single neuronal nucleus and two intranuclear *Homer1a* transcription foci as green puncta. This is a DAPI-stained nucleus (blue circular object) in a neuron of the CA1 subregion of the rat hippocampus. DAPI binds to eukaryotic chromatin and as such the neuronal soma and other cell parts are not stained (see section 2.5). The two green puncta correspond to the RNA that has been labeled with a fluorescent probe interlaced with FITC (green) targeting the 3'-untranslated region (UTR) of *Homer1a* transcript. There are two foci because there are two alleles of the *Homer1a* gene, one on either chromosome in the rat.

1.2 Non-Boolean foci analysis as alternative for all-or-nothing quantification

Many archetypal IEG imaging studies adopted a Boolean assumption in their transcription foci quantification protocols. That is, quantification was typically based on whether neurons were IEG-RNA-positive (signal present +) or IEG-RNA-negative (signal absent -) (see Table 1.1 for a brief overview of these studies).

Perhaps these previous studies implicitly assumed that either FISH amplification processes or the resolution limitations of optical imaging eliminated quantifiable variations in intranuclear transcription foci (which could be affected by relative copy numbers of transcribed RNA). As a result, previous Boolean approaches to IEG analyses, although effective for determining *which* cells fired *when*, considered transcription foci as all-or-nothing entities, without overt consideration of differences in foci characteristics such as foci intensity (brightness), size, or volume. The common use of laser confocal microscopy combined with stereological approaches in IEG detection also required the implementation of relatively high intensity thresholds or image adjustment for INF signal clarity, which would presumably alter INF characteristics to permit detection by the human eye (Guzowski et al., 1999; Vazdarjanova et al., 2002). If INF intensity and size truly did vary on a non-Boolean spectrum, then confocal thresholding would inadvertently eliminate the lower portions of these intensity/size distributions.

In our own investigations, the serendipitous observation of non-Boolean variations of *Homer1a* INF intensity and size in neurons of rats exposed to either one or five consecutive environments led us to think at length about the implications of these variable INF characteristics. A comprehensive report of non-Boolean INF characteristics in four hippocampal subregions was provided in thesis form (Witharana, 2011) and then

published in a meta-analysis of seven IEG studies (Witharana et al., 2016). These studies showed that transcription foci intensities were log-normally distributed within the population of activated cells, and the average INF integrated intensities increased with cumulative environmental exposure, or extensive surface area exploration. Our interpretation suggested the following: if electro-transcriptional coupling could regulate the number of transcribed RNA copies transcribed from IEGs based on the magnitude of neural activation (encoded by rate of action potentials), it is plausible that not all IEG transcription foci are uniform in size and intensity if different neurons were activated with variable magnitudes. Therefore, Boolean quantification may preclude the possibility that IEG images could provide firing rate information (*how much* these cells fired) in relation to neural coding, precluding the ability to detect any possible relationships between electrical neural activation and the corresponding transcription response. Furthermore, perhaps a non-Boolean approach, wherein the intensity of intranuclear transcription foci was also quantified, could offer a new type of information previously overlooked by Boolean characterization.

After this discovery, we searched through the literature in depth and found that supplementary to their main Boolean foci quantification methods, several publications provided secondary notes about non-Boolean changes in *Arc* INF intensity. One such report was presented by Miyashita and associates (2009) when they demonstrated that the fluorescence intensities of *Arc* transcription foci in CA1 differed as a result of cumulative track-running. They trained rats to run laps within a rectangular box for one lap a day, four laps a day, one lap for four days in a row, or four laps a day for four days, and reported on the effects of extended experience on the extent of new neuronal recruitment.

Despite relatively high confocal thresholding, Miyashita *et al.* (2009) clearly showed *Arc*-INF pixel intensity variations increased with massed running trials. Later, Penner *et al.* (2011) also reported significant variations in integrated intensity levels in *Arc* transcription between groups of young versus aged rats that were tested on spatial exploration paradigms. Both Penner *et al.* (2011) and Miyashita *et al.* (2009) showed corroborating PCR quantification data of *Arc* mRNA levels declining as a result of natural aging or overtraining in the hippocampus. Penner's group reported significant decline in behaviourally-induced *Arc* mRNA levels as a result of aging, which also corresponded to decreased average integrated intensities of *Arc* foci in CA1 and dentate gyrus. They suggested that this decline in mRNA expression paralleled the degradation of place cell expansion in aged rats, and therefore aged neurons fire fewer spikes (Shen, Barnes, McNaughton, Skaggs, & Weaver, 1997; Burke, Maurer, Yang, Navratilova, & Barnes, 2008). Miyashita's group showed increased levels of *Arc* mRNA that correlated with increased *Arc* foci intensity as a result of running more laps, and presumably in neurons firing more rounds of spikes than lower lap groups. Both studies provide support for the proposal that binary IEG foci quantification alone can omit information regarding IEG expression levels and corresponding neural activation magnitudes; whereas non-Boolean quantification might provide more insight into these other aspects of neuronal population activation.

To reiterate, non-Boolean variations in transcription foci observed in these studies (Witharana et al., 2016; Miyashita et al., 2009; Penner et al., 2011) might arise from electro-transcriptional coupling, wherein the intensity or frequency of neural activity can directly increase or decrease expression levels of activity-dependent immediate-early

gene transcription, resulting in variable amounts of transcribed RNA. Before we can elaborate on this concept, it is important to examine the specific functional and biomolecular bases of activity-dependent IEG expression.

Table 1.1 Examples of published studies that used IEG-RNA labelling to identify recently activated neuronal populations with a Boolean approach to INF quantification. Although this is a non-comprehensive list of references, this list demonstrates the versatility of INF characterization in examining a diverse range of questions related to neural activation. Note the ability to differentiate two or three separate neural populations active in distinct behavioural epochs by using single or double-IEG catFISH with one or two IEGs.

Authors	**Title**	**Paradigm**	**catFISH variant**
Guzowski *et al.* 1999	Environment-specific expression of the immediate-early gene *Arc* in hippocampal neuronal ensembles	A,immediate/ A,delay A/A A/B	*Arc*-INF *Arc*-Cyto
Vazdarjanova *et al.* 2002	Experience-coincident expression of the effector immediate-early genes *Arc* and *Homer1a* in hippocampal neocortical neuronal networks	A, A delay at various time-points	*Arc*-INF *Arc*-Cyto *Homer1a*-INF
Vazdarjanova & Guzowski, 2004	Differences in hippocampal neuronal population responses to modifications of an environmental context: Evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles	A / A object / A configuration A/B B/B	*Arc*-INF *Homer1a*-INF
Burke *et al.*, 2005	Differential encoding of behavior and spatial context in deep and superficial layers of neocortex	Same room/ different turn; Different room/ same turn	*Arc*-INF *Arc*-Cyto
Marrone *et al.*, 2008	Immediate-early gene expression at rest recapitulates recent experience	Explore-rest-explore / Rest-explore-explore / Rest-rest-rest	*Homer1a*-Cyto *Arc*-Cyto *Homer1a*-INF *Arc*-INF

1.3 IEGs encode proteins that contribute to long-lasting neuronal changes

Several effector immediate-early genes regulate neuronal plasticity processes by encoding proteins that bind to downstream molecular complexes and cascades involved in the generation of new synapses, trafficking of receptors, calcium signaling, and receptor scaffold targeting (Tully, 1997; Lanahan & Worley, 1998). As such, the transient and immediate transcription of synaptic-plasticity related IEGs is an important component of long-lasting memory in the nervous system, especially in the regulation of long-term synaptic potentiation, or LTP (Bliss & Collingridge, 1993; Flavell & Greenberg, 2008; Kandel, 2001). For example, Arc (Arg 3.1) proteins are transported to dendritic targets and bind to cytoskeletal complexes for AMPA receptor induction (Lyford et al., 1995). Down-regulating or knocking out the *Arc* gene inhibits the sustenance of long-term potentiation in hippocampal networks and impairs spatial memory consolidation (Guzowski et al., 2000). Similarly, the IEG *c-fos* appears to play a key role in the facilitation of fear memory formation and retrieval (Garner et al., 2012; Ramirez et al., 2013). Neurons that express *c-fos* in the dentate gyrus could be optogenetically re-activated through the activation of channelrhodopsin-2 by shining light neurons that express this type of light-gated receptor. After the training and cell-identification phase, freezing behaviour in trained rats could be induced even in a neutral, unpaired context (Ramirez et al., 2013). Thus, *c-fos* likely plays an important role in the facilitation and maintenance of fear memory processes (Morgan, Cohen, Hempstead, & Curran, 1987).

Both *Arc* and *c-fos* transcription foci appear in neuronal nuclei after ~2-5 min post-activation, but their intranuclear RNA load sharply declines after about 10 – 12 min

from activation. However, *Homer1a* is a relatively long (~45 kb) isoform of the *Homer* family (Bottai et al., 2002) and its nuclear signal has been shown to persist beyond 25 min after activation (when the 3' UTR of the *Homer1a* RNA strand is targeted with fluorescently-labeled antisense riboprobe) (Bottai et al., 2002). It is also possible to label *Homer1a* RNA persisting within the nucleus at substantially delayed time-points of 40 to 50 min after initial expression, likely due to the stability of its long transcript (Dubé et al., 2012; Dubé, 2016;). Figure 1.2 (adapted with permission from Figure 1.2 from Dubé's Master's thesis [2016]) depicts persisting INFs (labeled as "blobs" in y-axes) over the time course of 2 min to 50 min (x-axes) directly after the administration of maximal electroconvulsive shock (MECS). Briefly, MECS was administered to rats through ear-clips which were connected to a current-generating machine. The electroconvulsive shock delivery robustly excited hippocampal neurons, induced a visible seizure, and induced *Homer1a* transcription (Brakeman et al., 1997). The number of identifiable INFs is maintained above a quantifiable range even up to 50 min post-shock. This persistence of *Homer1a* mRNA permits this IEG as a reliable marker of neural activity while animals are subjected to long durations of behavioural experimentation.

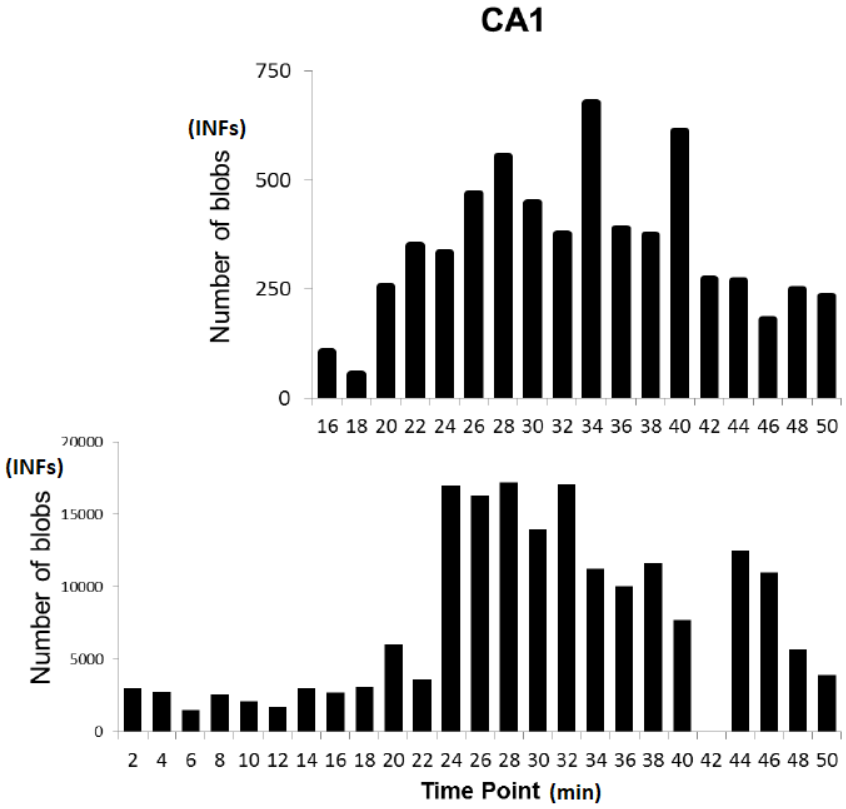


Figure 1.2 Post-maximal-electroconvulsive-shock (MECS) time-course showing persistence of *Homer1a* mRNA foci (blobs) up to 50 min after shock administration. The relative number of INFs (blobs) were quantified at 2 min intervals. Top panel shows laser confocal microscope image analysis of INF quantification in CA1 from 16 min to 50 min post-MECS administration. Bottom panel shows NanoZoomer image analysis of INF quantification from 2 min to 50 min after MECS administration. Adapted with permission from Dubé (2016).

1.4 Homer1a regulates synaptic plasticity in the nervous system

As previously mentioned, *Homer1a* is an effector IEG whose expression has been used extensively as a marker of recent neuronal activation (Vazdarjanova et al., 2002; Vazdarjanova & Guzowski, 2004; Marrone et al., 2008; Witharana, 2011; Witharana et al., 2016). *Homer1a* RNA transcript spans ~45kb (Bottai et al., 2002), and because of its sequence length, the transcription of its 3'UTR takes about 20-25 min, which affords a long time window for behavioural testing. In this section, the biological functionality of *Homer1a* will be discussed in the context of other members of the *Homer1* family, particularly regarding its role in regulating synaptic plasticity in the nervous system.

In the mammalian genome, three *Homer* genes (*Homer1*, *Homer2*, and *Homer3*) have been identified to date, and each gene seems to be expressed in different tissue groups (Xiao et al., 1998). *Homer* genes encode 47-48 kDa proteins which possess: 1) an amino (N)-terminal *Drosophila* Enabled/Vasp homology (EVH1) domain (Gertler, Niebuhr, Reinhard, Wehland, & Soriano, 1996) belonging to the Ena/VASP (vasodilator-stimulated phosphoprotein) group (Ponting & Phillips, 1997); and 2) a carboxy (C)-terminal coiled-coil (CC) domain and leucine zipper which enable oligomerization (Kato et al., 1998). The *Homer1* gene is mainly expressed in the brain, spans ~100kb, consists of 10 exons, and is the only *Homer* gene to have evolved bimodal expression of both IEG and constitutive isoforms (Bottai et al., 2002). Homer1 proteins are abundantly localized at the post-synaptic density, which is a protein-rich specialization region attached to the postsynaptic membrane of excitatory neurons consisting of a dense concentration of membrane receptors (Sheng & Hoogenraad, 2007). Homer1a (186 amino acids) was the

first isoform of the Homer1 family to be identified based on its significant mRNA upregulation in neurons of seizure-induced rat hippocampi (Brakeman et al., 1997; Kato et al., 1998). It was originally named just Homer, or vesl-1S (for VASP/Ena-related gene up-regulated during seizure and LTP), until the other members of the extended Homer family were also identified (Kato et al., 1998; Xiao et al., 2000). Homer1b/c are derived from alternatively spliced primary transcripts of Homer1 that are constitutively expressed, which means they are continuously expressed at a low rate that does not seem to increase substantially with neuronal activation (Wang, Chikina, Pincas, & Sealfon, 2014; Bottai et al., 2002). *Ania3* is the other IEG short-length variant of Homer1. Both *Homer1b* and *Homer1c* transcripts contain exons 1-10 of the parent *Homer1* gene, but *Homer1a* and *Ania3* mRNA only contain exons 1-5 as a result of synaptic-activity induced premature termination of transcription in the large central intron 5 of *Homer1* (Bottai et al., 2002). Although both Homer1a and Ania3 expression are upregulated by synaptic activity (Brakeman et al., 1997; Bottai et al., 2002; Kato et al, 1998), they differ from each other by several C-terminal amino acids, and their RNA transcripts contain unique 3'UTRs (Brakeman et al., 1997; Kato et al., 1997). Homer1a and Ania3 proteins are about half the size of the constitutive proteins, both contain the EVH1 domain but lack the C-terminal oligomerization component (Bottai et al., 2002). *Homer1a* is rapidly and transiently expressed in response to neuronal activation in excitatory neurons in the hippocampus, neocortex, and the basolateral amygdala, and in inhibitory neurons in the dorsal striatum and central amygdala (Imamura, Nonaka, Yamamoto, Matsuki, & Nomura, 2011).

The absence of the C-terminal coiled-coil (CC) domain in Homer1a proteins has been proposed as an important property that permits their endogenous negative inhibition

of their constitutive Homer1 counterparts, which, by contrast, all possess the CC domain. The constitutively expressed Homer1b/c will be denoted as a group by the term "CC-Homer1" from this point forward as we discuss the antagonistic interactions between the constitutive versus activity-driven Homer1 isoforms (Naisbitt et al., 1999, Xiao et al., 2000).

The PDZ domain is an important part of the conserved EVH1 homologous region in Homer1 proteins (Kato et al., 1998). A common structural component made up of approximately 80 to 90 amino acids, the PDZ domain contributes to protein-protein interactions, especially for anchoring to membrane receptors (Cho, Hunt, & Kennedy, 1992; Lee & Zheng, 2010). The acronym PDZ is derived from the first letters of the original three proteins found to share this domain: post-synaptic density protein (PSD95), *Drosophila* disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1) (Kennedy, 1995). In Homer1 proteins, the PDZ region interacts with group 1 metabotropic receptors: mGluR1 and mGluR5. Homer1a was one of the first PDZ-domain containing proteins that was also up-regulated by LTP, implicating H1a as a modulator of mGluR signaling and trafficking (Brakeman et al., 1997). The identification of a Homer-binding site on inositol-1,4,5-triphosphate receptors (IP3R) and ryanodine receptors, corroborated with evidence of their co-immunoprecipitation, suggested that CC-Homer1 couples mGluRs and IP3Rs. In contrast, H1a likely uncouples the mGluR-CC-Homer1-IP3R scaffold since it lacks the CC domain needed for self-association and cross-linkage. In support of this decoupling mechanism, the expression of a *Homer1a* transgene in Purkinje neurons decreased mGluR-induced calcium release, but *Homer1b* transgene did not (Tu et al., 1998). CC-Homer1 proteins have also been implicated in the

linkage of NMDA receptors (an ionotropic glutamate receptor) with mGluRs at the post-synaptic density via another family of scaffold proteins called the Shank family. The Shank group consists of proteins that interact with guanylate kinase-associated protein (GKAP) and PSD-95. PSD-95 possesses the PDZ domain, and anchors NMDARs to the postsynaptic membrane by interacting with the cytoplasmic C-termini of these receptors' NR2 subunits (Sheng, 2001). Shank proteins also interact with the EVH1 region of Homer, and are crucial in contributing to the quaternary structure of the Homer-Shank-GKAP-PSD-95 assembly. The Shank-CC-Homer1 interaction also causes mGluR clustering by linking mGluRs with the NMDAR assembly (Sheng and Kim, 2000). It has been proposed that the monovalent Homer1a protein uncouples proteins by targeting Shank and competing for the linkage positions, thereby disassembling the signaling complex, and reducing glutamate-dependent release of Ca^{2+} (Xiao et al., 2000).

Although the evidence suggest that Homer1a regulates synaptic plasticity by down-regulating neural excitability by competing with long-form CC-Homer1 proteins' interactions with membrane receptors, H1a has several other modulatory roles. H1a has also been shown to facilitate calcium spikes via L-type voltage-dependent calcium channels, which can be interpreted as a mechanism of bi-directional (either enhancement or down-regulation of) synaptic plasticity (Yamamoto et al., 2005). When H1a was infused by patched pipette into brain slices of rat neocortical pyramidal cells, Yamamoto and colleagues observed that spike-induced calcium release actually increased and this influx required concurrent mGluR-IP3 signalling. Furthermore, H1a can also act as an endogenous agonist of mGluRs, such that H1a can stimulate group 1 mGluRs in the absence of its natural ligand glutamate (Ango, Prezeau, Muller, Tu, Xiao, & Worley,

2001; Chung & Kim, 2017). Together, these results suggest that H1a can both negatively and positively affect neural excitability and its diverse functions likely contribute to a tightly controlled feedback loop for optimizing or downregulating specific neuronal responses.

The behavioural impacts of Homer1a expression parallel this protein's bi-directional regulation of neural plasticity. Overexpression of H1a in the lateral and basal regions of the amygdala impairs auditory fear conditioning and interferes with normal social cognition (Banerjee, Luong, Ho, Saib, & Ploski, 2016). In rats that experienced amygdala kindling (the gradual intensification of brain excitation via repetitive administration of electrical stimulation until a seizure is induced), *H1a* was the most strongly-induced gene in the kindled hippocampus (Potschka et al., 2002). Transgenic mice that permanently overexpressed H1a showed a decreased seizure susceptibility after kindling when compared to wild-type mice, suggesting that H1a may act as an endogenous antiepileptogenic or anticonvulsant agent (Potschka et al., 2002) by downregulating NMDAR and mGluR interactions. Based on these findings, H1a likely participates in a feedback loop, induced by intense neural activation followed by subsequent downregulation to modulate synaptic strength.

1.5 Mechanisms of electro-transcriptional coupling of IEGs: the role of calcium

Next, we will return to a more general discussion of how most immediate-early gene transcription can be induced by electrical neural activation through the process of "electro-transcriptional coupling."

High frequency neuronal activation can induce long-lasting potentiation (LTP) or enhancement, which is an enduring increase in the strength of the synaptic response of a neuron (Bliss & Lomo, 1973). A neuron that displays LTP has a higher probability of firing again when activated by the same pre-synaptic neuron that triggered this LTP (Barnes, Jung, McNaughton, Korol, Andreasson, & Worley; McNaughton, 1982). As such, LTP is the primary model of the cellular basis of memory in the mammalian brain. LTP can be divided into distinct stages based on the longevity of the potentiation (Teyler & DiScenn, 1987). The brief initial stage can last minutes, while early phase LTP lasts multiple hours, and late-phase LTP can be sustained between several hours to weeks. Only the maintenance of late-phase LTP requires gene transcription and subsequent protein synthesis (Kandel, 2001). Synaptic activity-driven immediate-early gene (IEG) expression can occur rapidly, transiently, and in the absence of *de novo* protein synthesis, and therefore IEGs play an important role in encoding proteins that regulate long-lasting synaptic plasticity (Flavell & Greenberg, 2008). Electro-transcriptional coupling (ETC) is the process in which neuronal activation (membrane depolarization) signals a rapid molecular cascade triggering transcription of IEGs (Figure 1.3). One model of ETC suggests that action potentials can induce gene transcription without direct transport of molecules from the synapses, which are typically located far from the nucleus. Instead, action potentials can signal the translocation of molecules already in the cell body into the nucleus or by direct activation of nuclear transcription through the use of calcium ions (Ca^{2+}), which can enter the cell through voltage-gated membrane channels or released from intracellular stores (Adams & Dudek, 2005).

Although there are several signal transduction pathways through which action potentials can signal gene expression in the nucleus, Ca^{2+} seems to be a common essential messenger among these diverse pathways. During neuronal membrane depolarization, there are several types of membrane receptors that can be triggered to permit the subsequent influx of Ca^{2+}. The pre-synaptic release of glutamate can activate post-synaptic NMDA receptors (NMDAR) which opens a Ca^{2+} channel, allowing the flow of Ca^{2+} into the cell. This NMDAR activation has been shown to directly upregulate the transcription of *Homer1a* (Sato, Suzuki, & Nakanishi, 2001), *Arc* (Lyford et al., 1995), and other IEGs such as *zif268* (Cole, Saffen, Baraban, & Worley, 1989). Calcium can also enter the cell via L-type voltage-gated calcium channels (VGCCs), which are opened when the membrane is depolarized (Yamamoto et al., 2005).

Activity-dependent increases of Ca^{2+} in the cell can have widespread convergent interactions on downstream pathways triggered by the activation of other membrane proteins. For example, Ca^{2+} can bind to calmodulin (CaM), also known as calcium binding protein or calcium-modulated protein, which has interactions with divergent pathways as well. In relation to gene expression, CaM can activate calcium/calmodulin-dependent kinase type IV (CAMKIV), or indirectly activate protein kinase A (PKA), both of which phosphorylate transcription factors in the nucleus, which then bind to gene sequences to initiate transcription (Adams & Dudek, 2005). In the indirect transcriptional signal pathway of CaM, adenylyl cyclase (AC) activation also requires the concurrent binding of an activated subunit of G-protein coupled receptor (GPCR). Then, adenylyl cyclase catalyzes the conversion of adenosine triphosphate (ATP) into cyclic adenosine monophosphate (cAMP) (Van Nguyen, Kobierski, Comb, & Hyman, 1990) which then

goes on to activate PKA. In both the direct and indirect pathways, Ca^{2+} and cAMP function as second messengers that mediate action potential-induced gene transcription by directly or indirectly activating kinases that phosphorylate transcription factors. Calcium can also regulate kinases in a third pathway, the extracellular-signal regulated kinase/mitogen-activated protein kinase (ERK/MAPK) system (Chuderland & Seger, 2008), which is triggered when receptor tyrosine kinases (RTKs) at the membrane are bound by extracellular proteins such as growth factors or hormones. The Ras protein family and mitogen-activated protein kinase kinase (MEK) are intermediary enzymes that, along with Ca^{2+}, contribute to the activation of ERK/MAPKs (Figure 1.3). The ERK/MAPK pathway serves as a secondary route for the activity-dependent phosphorylation of transcription factors in the nucleus, and also contributes to the regulation of IEG transcription (Xia, Dudek, Miranti, & Greenberg, 1996; Wiegert & Bading, 2011). For example, *Arc* mRNA is upregulated by mGluR1/5 activation as a result of concurrent Ca^{2+} activity on both CaM and ERK1/2 (Wang, Zheng, Zhou, Sun, & Wang, 2009). In addition, glutamate stimulation of NMDARs increased *Homer1a* mRNA levels in cerebellar granule cells, but the induction of transcription also required the presence of MAPK as a downstream mediator, most likely in its role of activating *H1a*-related transcription factors (Sato et al., 2001).

In the nucleus, several transcription factors are required for the successful and rapid recruitment of RNA polymerase for transcribing RNA from the DNA sequence. Originally, the primary candidate for mediating IEG transcription was Ca^{2+}/cAMP response element binding protein (CREB) (Kandel, 2001). The simple CREB-dependent transcription model suggested that neuronal membrane depolarization led to subsequent

increases in calcium influx through the opening of voltage-gated calcium channels (Hardingham, Chawla, Johnson, & Bading, 1997). Calcium cAMP activate protein kinases, either directly or through intermediary molecules, which in turn phosphorylate CREB, which then binds to the cAMP response element (CRE) in the promoter or enhancer regions upstream of IEGs, and thereby initiates the transcription of IEGs (Sheng, McFadden, & Greenberg, 1990; Sheng & Greenberg, 1990) (Figure 1.3). In cell cultures, Greenberg, Thompson and Sheng (1992) delineated a model in which voltage-gated calcium channels were activated by membrane depolarization, which led to the activation of CaM kinase, followed by CREB phosphorylation, which then activated *c-fos* transcription.

However, it has since been shown that CREB is not the only transcription factor that is required for IEG transcription. Furthermore, the CRE gene element only constitutes one part of a longer, unique enhancer region, called the synaptic activity-responsive element (SARE) that is located upstream of several IEGs such as *Arc* and *Homer1a* (Rodriguez-Tornos, Aniceto, Cubelos, & Nieto, 2013). Other transcription factors include serum response factor (SRF) (Ramanan et al., 2005), myocyte enhancement factor (MEF2) (Flavell et al., 2006), and neuronal PAS (Per-Arnt-Sim) domain binding protein (NPas4) (Lin et al., 2008), which, along with CREB, can all bind to SARE to initiate transcription in a conjunctive manner. More importantly, all these transcription factors implicated in IEG expression require calcium-dependent activation (Miranti, Ginty, Huang, Chatila, & Greenberg, 1995; Flavell et al., 2006; Sun & Lin, 2016).

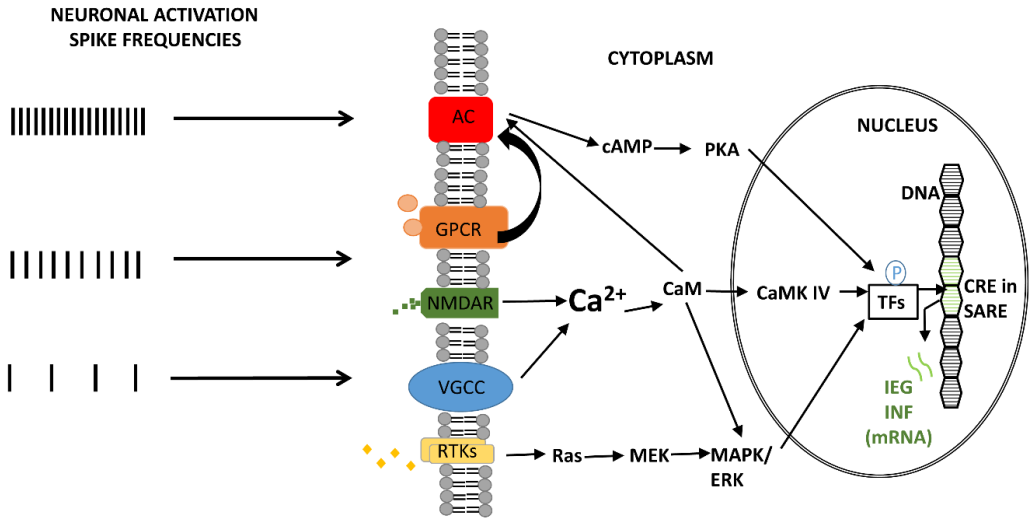


Figure 1.3 Schematic of several calcium-dependent signal-transduction pathways that can mediate the conversion of neuronal activation (spike trains or series of action potentials) into cascades that initiate rapid transcription of immediate-early genes in the neuronal nucleus. Several convergent calcium (Ca^{2+})-dependent pathways can lead to the phosphorylation of transcription factors (TFs), which binds to upstream synaptic-activity response elements (SARE) on mammalian DNA, signalling transcription of immediate-early genes (IEGs). The binding of glutamate to NMDA receptors (NMDAR) in the cellular membrane opens the ion channel which permits the influx of Ca^{2+} into the cell. Membrane depolarization can also lead to the opening of voltage-gated Ca^{2+} channels (VGCC), another source of extracellular Ca^{2+}. In another pathway, G-protein coupled receptors (GPCR) activate adenylyl cyclase (AC) with the concurrent binding of calcium-modulated protein, or calmodulin (CaM). This synergistic interaction generates cyclic adenosine monophosphate (cAMP) which phosphorylates protein kinase A (PKA). CaM itself can directly activate calcium/calmodulin-dependent kinase IV (CaMKIV) which also phosphorylates transcription factors. Lastly, CaM can interact with kinases in the ERK/MAPK (extracellular signal-regulated kinase/mitogen activated kinase) pathway, which is triggered by the binding of endogenous growth factors or hormones to receptor tyrosine kinases (RTKs) at the membrane. The intermediary proteins Ras and mitogen-activated protein kinase kinase (MEK) contribute to the activation of ERK/MAPKs which then phosphorylate transcription factors. Some transcription factors that have been linked to IEG-specific transcription include CREB (cAMP responsive element binding protein), serum response factor (SRF), Npas4, and myelin elongation factor (MEF2). The identification of CREB's affinity for the cAMP response element (CRE) upstream of IEGs was the original primary model for electro-transcriptional coupling, but has since been modified to include the other important transcription factors that bind to other parts of the longer synaptic activity-responsive element (SARE). Therefore, the transcription of mRNA of immediate-early genes relies on widespread calcium signalling. Finally, transcription and accumulation of mRNA at transcription sites generate intranuclear transcription foci which are the target of probes used in fluorescent *in situ* hybridization.

1.6 Electro-transcriptional coupling may be linearly co-regulated

In light of the above discussion highlighting the importance of calcium ions as a second messenger in signalling IEG transcription, it has been proposed that the intensity of neural activation can affect the level of calcium increases, and subsequently influence the level of gene expression. The stoichiometric property of Ca^{2+} influx suggests the possibility of a direct proportionality between firing rates and transcription cycles, particularly for IEGs like *Homer1a* and *Arc*. For example, it is likely that spiking frequency or intensity can affect the amount of available calcium ions, which in turn signal and amplify the level of recruitment of transcription factors within the nuclear transcription sites. In this regard, the magnitude of the electrical activated is converted into a chemical signal (based on level of calcium increase) that translates into a continuous transcriptional signal. The number of cycles of IEG transcription may be controlled by the rate of recruitment of transcription factors (how many are phosphorylated or stay phosphorylated); thereby affecting how many RNA transcripts are produced based on the magnitude of the initial neuronal activation.

Indirect evidence supporting this proportional relationship was reported *in vitro* by Fields and associates (1997). Mouse dorsal root ganglion (DRG) cells were electrically stimulated with 540 action potentials administered in four variable patterns for 30 min. Transient intracellular calcium signals were measured via fluorescent imaging with Ca^{2+}-indicators while *c-fos* expression was measured with semiquantitative PCR (polymerase chain reaction). Immunocytochemical staining was also used to quantify CREB phosphorylation (based on quantifying intensities within cell nuclei). The authors observed that specific temporal features of action potentials (i.e. shorter inter-spike

duration) influenced Ca^{2+} transients, and correlated with increases in *c-fos* expression and CREB phosphorylation levels (Fields, Eshete, Stevens, & Itoh, 1997). These findings suggest that temporal dynamics of neuronal activation or membrane depolarization lead to changes in calcium ion influx, which can transiently coordinate the transcription rates of immediate-early genes via specific signal cascades with variable kinetics.

As a starting point, it could be proposed that a discrete train of action potentials (i.e. variable magnitude of electrical signal) could be bio-mechanistically converted into variable repetitions of IEG transcription, which would lead to the synthesis of different quantities of RNA transcripts, therefore resulting in non-Boolean variations in riboprobe-labeled transcription foci intensity.

Optimistically, the electro-transcriptional relationship might be linear:

$$M(n) = kn + M_0$$

where M is the number of RNA transcripts, n is the number of spikes (or possibly bursts of spikes), M_0 is the pre-stimulation number (possibly resulting from previous activity), and k is a constant reflecting the probability of a new cycle of transcription following a spike (or burst). Ideally, for use as a robust proxy for neural activation, there would be a high gain relationship between spikes and cycles of transcription. Of course, many factors, such as changes in transcription factor availability, saturation of available RNA polymerases, diffusion and post-transcriptional modification of primary transcripts, stearic hindrances within transcription complexes, the number of RNA polymerases and

state of polymerase readiness (more discussion in the next section) might affect both the gain and linearity of the spike vs. transcription function. Therefore, one cannot currently estimate *a priori* the extent to which IEG transcription vs. spiking rate stoichiometry might be limited but the overall linear relationship might be conserved until this biomolecular availability limit is reached.

Given that this proposal suggests a direct relationship between incoming signal strength (neural activation via calcium influx) which we have discussed at length, and corresponding transcription foci signal (amount of RNA); it would now be helpful to discuss the biological implications of the resulting variable RNA transcription foci.

1.7 Punctate transcription foci correspond to clusters of RNA

After fluorescent *in situ* hybridization, the bright punctate intranuclear foci labeled by fluorescent riboprobes correspond to the spatial location of tight clusters of nascent RNA transcripts. One spatial model of transcription suggests that RNA is synthesized within stable and self-contained "transcription factories" (Jackson, Hassan, Errington, & Cook, 1993) that are sustained as distinct domains throughout the nucleus (Wansink, Schul, van der Kraan, van Steensel, van Driel, & de Jong, 1993). A prevailing assumption is that the spatially compact feature of transcription foci is the result of synergistic labeling of tightly bound clusters of RNA copies contained within these "transcription factories." Electron microscopy imaging suggested these discrete intranuclear factories are between 45-100 nm in diameter (Iborra et al., 1996; Martin & Pombo, 2003), and are present in consistent numbers within similar cell types (Iborra et al., 1996). Each factory houses between 4-30 RNA polymerases, multiple transcription factors, and many other molecules associated with transcriptional activation and mRNA processing (Carter,

Eskiw, & Cook, 2008; Martin & Pombo, 2003; Rieder, Trajanoski, & McNally, 2012; Grande, van der Kraan, de Jong, & van Driel, 1997). Also, these factories are maintained in the absence of transcription (Mitchell & Fraser, 2008) and even after removal of chromatin (Jackson et al., 1993). Associated with this model is the prediction of "immobilized polymerase" wherein DNA containing the target immediate-early gene sequences is mobilized and sequestered to stable transcription factories at the initiation of transcription (Osborne et al., 2007). During transcription, DNA is looped by the stationary polymerase enzymes and the subsequent elongation of nascent RNA strands from the same gene is restricted to small areas (about 50 nm in diameter) corresponding to the location of the transcription factories (Martin & Pombo, 2003). The "immobilized polymerase" theory is contrary to the model of free-polymerase moving down immobilized DNA during transcription (Osborne et al., 2007). Several studies also provide evidence for co-localization of genetic transcription activity (initiation and elongation) across multiple genes such that several activated genes have been shown to share the same transcription factories (Mitchell & Fraser, 2008; Osborne et al., 2004). The densely packed polymerase enzymes anchored within transcription factories may permit the concurrent activity of two or more RNA polymerases at a specific locus, providing a mechanism for rapid re-initiation of transcription and accumulation of multiple RNA copies in a short period of time (Jackson, Iborra, Manders, & Cook, 1998). Although much of the earlier evidence for transcription factories was observed in fixed cells, more recent live-image tracking of an RNAPII-associated factor, cyclin-dependent kinase 9 (CDK9), corroborated the presence of transcription factories as stable compartments in unfixed mouse erythroid leukemia cells (Ghamari et al., 2013). It is plausible that, because of this stable spatial concentration of RNA transcripts at discrete

sites of synthesis, fluorescent intranuclear transcription foci (INFs) may increase in intensity and size in response to increasing neuronal stimulation (cumulative action potentials) (Miyashita et al., 2009; Penner et al., 2011; Witharana, 2011; Witharana et al., 2016). In effect, INFs may expand because of the accumulation of newly synthesized RNA in these factories, and signal dissipation may be the result of eventual post-processing export to the cytoplasm or mRNA degradation.

In contrast to the model of stable transcription factories, evidence has emerged suggesting that RNA polymerases do not stay in pre-assembled long-lasting factories. On the contrary, live imaging experiments reported that clusters of RNAPII are only transiently assembled, and these temporary RNAPII clusters were only maintained briefly for an average of 5.1 ($\pm$0.4) s in human osteosarcoma cells (Cisse et al., 2013). A new quantitative technique in giant human salivary glands has also shown the brief assembly and subsequent dissipation of "clusters" of only a minute fraction of all cellular transcription enzymes in real-time, and the majority of RNAPII do not form permanent clusters on a global scale (Zhao et al., 2014). However, a recent experiment showed that although the basal cluster duration of RNAPII in mouse embryonic fibroblasts only lasts ~8 s, serum stimulation actually extended RNAPII cluster lifetime up to 20 min and the duration of clustering was predictive of corresponding levels of β-actin mRNA (Cho et al., 2016). Therefore, to corroborate these three studies, although there may be a low rate of stochastic RNAPII clustering at basal, unstimulated levels (Cisse et al., 2013), it is possible that stimulation or neuronal activation substantially extends RNAPII cluster time to transcribe small population of IEGs (Cho et al., 2016), which would naturally constitute a small portion of all available RNAPII in the nucleus (Zhao et al., 2014). In

light of the "transient RNAPII cluster" model, the number of RNA copies that can be synthesized may be temporally regulated by stimulation-based extended duration clustering, providing an upper limit to the linearity of the electro-transcriptional relationship. However, despite this possible upper limit, the spatially-concentrated accumulation of nascent RNA transcripts in this limited time window could still be affected by the magnitude of neural activation if the activation-dependent polymerase cluster time was increased to accommodate variable rounds of transcription, still producing quantifiable changes in the amount of RNA in this short time frame (Cho et al., 2016).

Another recent finding pertinent to RNA polymerase localization is the characterization of "poised" RNA polymerase near the transcription start sites in certain genes, especially in immediate-early genes (Saha et al., 2011). Also known as RNA polymerase "stalling," this mechanism permits the near-instantaneous induction of transcription triggered by neuronal activation since the polymerase has already been recruited and targeted to the appropriate target gene, both processes that is typically considered rate-limiting steps in transcription initiation (Lis, 1998). Saha and colleagues (2011) discovered RNAPII substantially concentrated near the *Arc* transcription start site which showed classic traits of pre-initiation such as phosphorylation of serine 5 residues and histone enrichment of the promoter region. This evidence of pre-initiation by stalled polymerase can fit into either the stable "transcription factory" model, or the "transient polymerase cluster" model. In the case of the first model, the stalled and immobilized polymerase could already be poised on the gene locus, and this stalling is maintained within the transcription factory. In the case of the second model, the transient clustering

of the stalled polymerases on their specific gene loci could involve the translocation of the DNA-RNAPII complex into these RNAP groups and continue transcription within the RNAPII clustering period. Moreover, since stalled polymerase contributes to the retention of permissive epigenetic markers on promoter sequences, rapid and successive rounds of polymerase recruitment to the accessible domain of DNA could result in robust transcriptional output and increased RNA copy numbers of IEGs (Saha & Dudek, 2013).

1.8 Variable levels of RNA signal can be maintained during FISH amplification

Despite the proposed proportional relationship in electro-transcriptional coupling, another reason for continued Boolean quantification of IEG images might be the implicit assumption that FISH amplification steps are not sensitive enough to detect minute differences in the local density of target RNA molecules. In the FISH protocol that has been widely adopted for IEG processing, a hapten-labeled antisense, single-stranded oligonucleotide probe is hybridized to the primary RNA transcript of the target IEG (Guzowski et al., 1999). The hapten molecules, typically small ligand dye molecules (e.g. fluorescein or digoxigenin), usually benefit from signal amplification to generate a larger, quantifiable optical signal. An antibody that binds to the hapten is conjugated to an HRP (horseradish peroxidase) enzyme that catalyzes and forms bonds with a signal amplifying dye (such as tyramide). This amplification exponentially increases the detectable RNA signal by recruiting many fluorescent dye molecules to bond to single hapten molecules (see Figure 1 of Witharana, 2011), and it is this aggregate of fluorochromes that produces an enhanced fluorescent foci (reduced signal to noise ratio, higher resolution at lower magnifications). However, if the factor of amplification across behavioural conditions is

monotonic, then even this exponential amplification ratio may still produce detectable differences in the intensity and size of transcription foci, granted image acquisition parameters are consistent across behavioural conditions as in the Miyashita *et al.* (2009) study. Thus, if the current FISH protocol is carried out uniformly across behavioural conditions (such as combining test groups on a single microscope slide to prevent technical biases from tissue processing), then variable levels of RNA signal may still potentially be quantified on the resulting images, through non-Boolean INF intensities and sizes. This type of quantification can potentially be accomplished using populations of neurons whose physiological response function to given stimuli are already known, for example, visual cortical responses to variably repeated oriented bar stimuli, or, in the case of the hippocampus, repetitive place cell activation (O'Keefe & Dostrovsky, 1971) over repeated traversals of a fixed region of space.

1.9 Systematic quantification of *Homer1a* RNA expression

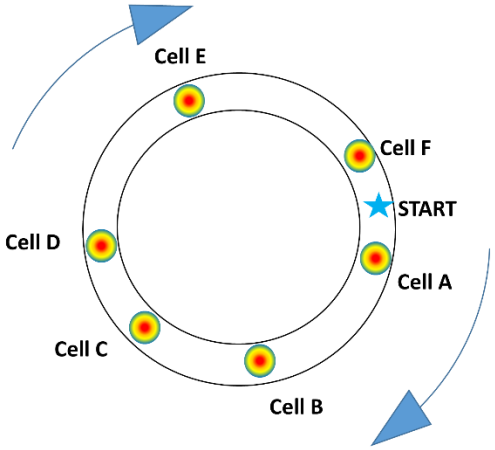
The possibility that static fluorescent IEG images could convey not only the identity of activated cells in a large population, but also the approximate magnitudes of neuronal activation (i.e. firing rates) has important implications for brain imaging applications. At this time, IEG images can provide a sufficient readout of *which* cells fired *when*. However, if there is a robust relationship between activation magnitude and RNA signal intensity, then it might be possible to use IEG images as a proxy for firing rate analysis. That is, we could potentially also identify *how much* those cells fired.

Previous reports (Miyashita et al., 2009; Penner et al., 2011; Witharana et al., 2016) have shown non-Boolean variations in the intensity and size of IEG transcription foci as a function of lap running, aging, or cumulative environmental exposure. However,

these studies did not explicitly investigate the transfer function (linear or nonlinear) between electro-transcriptional co-regulation. In the present study, animals were subjected to systematic track running in order to delineate changes in *Homer1a* transcription levels as lap numbers increased. The integrated fluorescence intensities of individual *Homer1a* transcription foci were characterized to determine whether repetitive epochs (such as place cell re-activation) would result in a linear increase in foci intensity. On a broader scope, this body of work focused on whether a linear relationship between foci fluorescence signal could be directly translated into a surrogate indication of that particular neuron's firing rate. Previous experiments (Penner et al., 2011; Witharana et al., 2016) showed that INF intensities are variable and log-normally distributed within populations of hippocampal neurons, and INF intensities of a fraction of hippocampal neurons increased with cumulative environmental exposure, or extensive area exploration. Log-normality of INF intensity might reflect log-normal distributions of firing rates during spatial behaviour (Barnes et al., 1990; Mizuseki & Buzsaki, 2013). Our goal in this thesis was to quantify the stoichiometry of IEG RNA fluorescence and the total number of spikes within a brief time window.

Place cells in the hippocampus fire when an animal occupies a specific region in space (O'Keefe & Dostrovsky, 1971; Wilson & McNaughton, 1993). Based on robust electrophysiological recordings of neurons, repetitive track running in a familiar environment elicits repeated and reasonably consistent activation of many of the same hippocampal place cells on each lap (explained in Figure 1.4) (Lee, Yoganarasimha, Rao, & Knierim, 2004; Maurer, Cowen, Burke, Barnes, & McNaughton, 2006). Moreover, there is a robust correspondence between place cell activity and *H1a* transcriptional

activity (Vazdarjanova et al., 2002), making this a reasonable behavioural paradigm to theoretically induce linearly increasing numbers of total spiking events to compare with the total IEG fluorescence. In addition to its strong relationship with place cell firing, we also chose to examine *Homer1a* expression because of its longer transcript length (Bottai et al., 2002), in conjunction with the persistence of intranuclear RNA signal up to 50 min (Dubé, 2016). Both properties permit delayed appearance of its 3'UTR, and latent signal decay, which would minimize loss of accumulated signal over the time (about 10 min) needed to run 50 laps. Portions of the experimental protocol and some of the results outlined in this thesis have been published as a peer-reviewed article (Witharana, Clark, Trivedi, Mesina, & McNaughton, 2018).



Lap	Firing sequence
1	ABCDEF
2	ABCDEF
3	ABCDEF
4	ABCDEF
5	ABCDEF
6	ABCDEF

Cell	# theoretical spike trains after 6 laps
A	IIIII IIIII IIIII IIIII IIIII IIIII
B	IIIII IIIII IIIII IIIII IIIII IIIII
C	IIIII IIIII IIIII IIIII IIIII IIIII
D	IIIII IIIII IIIII IIIII IIIII IIIII
E	IIIII IIIII IIIII IIIII IIIII IIIII
F	IIIII IIIII IIIII IIIII IIIII IIIII

Figure 1.4 Theoretical activation pattern of a sequence of place cells in the hippocampus during unidirectional lap running on a circle track.
Electrophysiological recordings show that many of the same place cells representing areas on a track are repeatedly activated when an animal runs through the same zone in the same direction over many laps. This diagram shows theoretical place cells that fire in the same sequence over six laps. These place cells would show a burst or packet of firing during each traversal of that position on the track. After 6 unidirectional continuous laps, cumulative firing for each cell would be the sum of all spikes fired over each consecutive lap. Furthermore, the six cumulative firing events would result in cumulative transcription of IEG RNA induced by each re-activation on each lap.

2 METHODS & MATERIALS

2.1 Subjects

Adult male Long-Evans rats (n=31) between the ages of 3-6 months, were handled daily and food-restricted to 85% of *ad libitum* body weight. Animal handling and care procedures were approved by the University of Lethbridge Animal Welfare Committee, as outlined by the Canadian Council on Animal Care. Subjects were housed in a shared colony with other rats with the lights on during the day, and lights off during the evening. Training and testing were performed during the lights on portion of the day.

2.2 Experimental design

Rats were trained to run in one direction (clockwise) for food reward (crushed Froot Loops™ and Cheetos™) at a fixed location on a 90 cm-diameter circular track placed on the floor (Figure 2.1). The training room contained various distal cues including posters, an electrophysiology rig, and other visible cues that were not changed during training or testing. At the start of each training session, rats were transported from their colony housing rooms (in temporary housing cages identical to their colony home-cages) to a darkened room next to the experimental room, where they spent 2-3 h in quiet acclimatization. Each rat was trained with the lights on in the experimental room for about 10 min/day over consecutive days until their running speed reached a criterion of 20 laps or more in under 5 min. On test day, rats were moved to the familiar darkened antechamber as was done during training, and again rested for 2-3 h prior to lap running, in order to minimize the residual *H1a* mRNA induced by prior behaviour and transportation in the home-cage. Rats were randomly pre-assigned a running condition between 1 to 25 laps. Four rats that demonstrated exceptionally faster running speeds

were pre-assigned to run 30 or 50 laps (two per group). Previous studies (Ekstrom et al., 2001) show that variations in running speed do not markedly affect the total number of spikes emitted in the place field. Once the rat completed the designated number of laps [1 lap (n =4 rats), 5 laps (n=4), 10 laps (n=3), 20 laps (n=3), 25 laps (n=5), 30 laps (n=2), 50 laps (n=2)], they were returned to the darkened antechamber and rested for ~25 min, followed by immediate sacrifice. Caged controls (n=4) were sacrificed immediately after the 2-3 h period in the dark chamber, and positive controls (n=4) (maximal electroconvulsive shock, MECS) were given a single electroconvulsive shock treatment and sacrificed ~29 min later (Figure 2.1). MECS was administered via ear-clips attached to a current generator (UgoBasile, Italy), with the following parameters: frequency of 100 pulses/s; pulse width of 0.5 ms; shock duration of 1.1 s; current at 85 mA. MECS-induced expression is typically assumed to reflect the maximal possible expression within the population since it robustly upregulates *H1a* mRNA transcription (Brakeman et al., 1997; Bottai et al., 2002) but this is currently an untested assumption.

Following *in situ* hybridization, image acquisition, and image analysis of fluorescence signal of *Homer1a* mRNA transcripts (detailed below), transcription foci integrated intensity values were collected by automated quantification software. Automated analyses permitted blind collection of intensity values such that subjects could be pooled within a lap group, or analyzed as separate subjects.

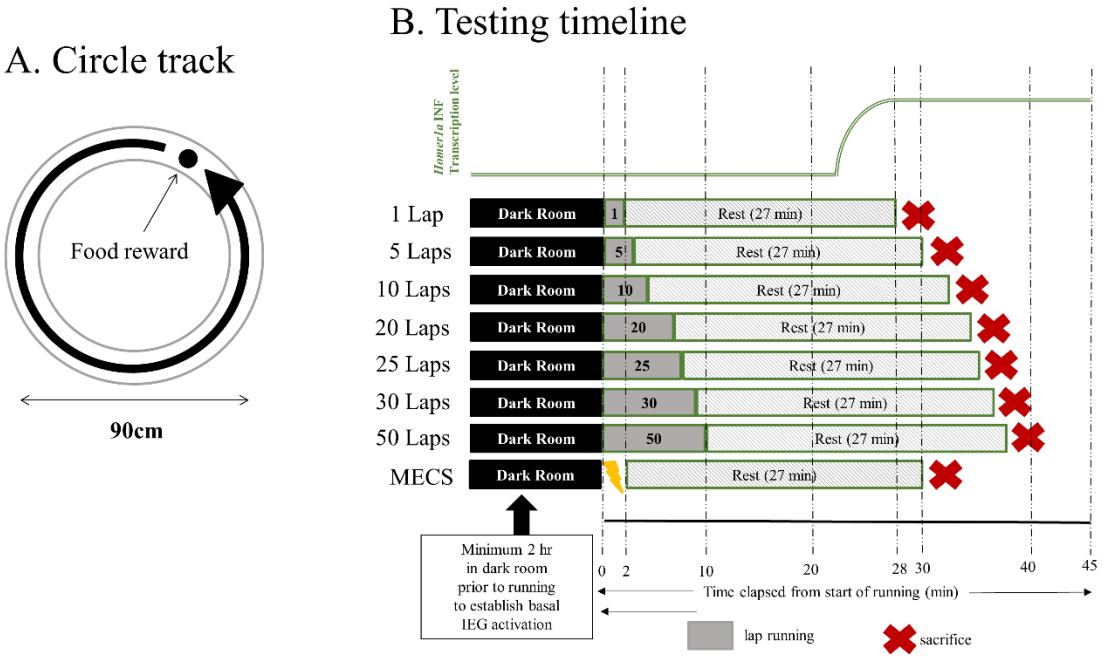


Figure 2.1 Schematic of running track and time-course of events for each group on testing day. A) Circle track used for behavioural training and testing. Track measured 90 cm in diameter and a fixed reward site was replenished with a mixture of crushed Froot Loops™ and Cheetos™ after each visit to encourage continuous lap running. Distal room cues were visible throughout all trials, including test day. **B)** Timeline of experimental procedure on testing day. Each rat rested in a dark antechamber for a minimum of 2 h prior to first exposure to the familiar running track in a separate, well-lit room. Each rat ran a fixed number of laps (randomly assigned except for the 30 and 50 lap groups which were selected based on their higher speeds to ensure fast and consistent track running) and rested for a minimum of 27 min in the same dark antechamber before rapid sacrifice and brain extraction. Upper green line is the theoretical INF signal level of *H1a*-3'UTR in correspondence with timing of the experiment. Exploratory study of post-MECS signal time-course showed that *H1a*-3'UTR fluorescence is sustained between ~25-40 min from the start of induction (see Figure 1.2). Therefore the variation in total time due to differences in number of laps is expected to have minimal effect on the results.
Figure also published in Witharana *et al.*, 2018 (Figure 1).

2.3 Sacrifice and brain extractions

After the post-running rest period, rats were transported to a wet lab for quick sacrifice. Rats were placed in a tightly-sealed chamber containing 5% isoflurane for a minimum of 40 s until fully anaesthetized. Between 0.4-0.8 mL of sodium pentobarbital (Euthansol) was injected intracardially to arrest cardiac activity and rats were then decapitated with a guillotine. Following rapid brain extraction and skull removal (the entire procedure occurred in less than 1 min for each subject), the brains were submerged for 2 min in a metal container of -50^{o}C liquid 2-methylbutane surrounded by a slurry of dry ice in 70% ethanol. Frozen brains were then wrapped in aluminum foil, labeled and stored in Falcon tubes at -80^{o}C until cryosectioning.

2.4 Cryosectioning of brains

Brains were removed from -80^{o}C storage, and hemispheres were separated with a razor blade after removing the cerebellum. The hemi-brains were "blocked" with TissueTek^{TM} Optimal Cutting Temperature medium (VWR, Radnor, PA) such that all experimental groups and controls from a cohort were represented on the same slide to minimize procedural variabilities that can arise during tissue handling or FISH. Blocking means that hemi-brains were placed vertically on their flat cerebellar ends, arranged in an array in a rectangular mold, and layers of OCT were gradually poured around the array to create a solid block of embedded brains. Cryosectioning was performed on a cryostat (LEICA Microsystems, model CM1900, Concord, Ontario, Canada). SuperFrost^{TM} charged slides (ThermoFisher Scientific, Waltham, MA) were used to capture 20 μm thick serial coronal sections of the brain blocks. Slides were dried at room temperature for

a maximum of 30 min to prevent ice crystal formation within the brain tissue, frozen at -20^{o}C for 2 h, and then stored at -80^{o}C.

2.5 Fluorescent *in situ* hybridization (FISH)

Every tenth 20μm thick coronal brain section representing about -2.64 mm to -4.68 mm from bregma (based on the rat brain atlas, Paxinos & Watson, 2007) was thawed in preparation for fluorescent *in situ* hybridization to label *Homer1a* RNA. The main steps of the hybridization protocol are described below and additional information can also be found in previous publications (Montes-Rodriguez, Lapointe, Trivedi, Lu, Demchuk, & McNaughton, 2013; Witharana et al., 2016).

Antisense riboprobes targeting the 3'UTR (untranslated region) of the *Homer1a* RNA transcript were synthesized from *Homer1a* DNA template with Maxscript RNA Synthesis Kit (Ambion, Austin, TX) and fluorescein-isothiocyanate-labeled UTP (Roche Diagnostics; Indianapolis, IN) to conjugate uracil bases on the probe with fluorescein isothiocyanate (FITC). FITC is a derivative of fluorescein dye molecule which has been chemically optimized for biological labeling. The 3'UTR of *Homer1a* was chosen as the probe target because it takes about 20 min to be transcribed in neuronal nuclei after the start of transcription by synaptic activity, and this sequence also differs from the 3'UTR of the other IEG form of *Homer1,* known as *Ania3* (Bottai et al., 2002). *H1a* mRNA has been shown to persist in the nucleus up to about 40 – 50 min (Figure 1.2) and is not exported to the cytoplasm until 60 min after the first transcription event (Bottai et al., 2002). It should also be noted that portions of intron 5 of *Homer 1* is converted into the 3'UTR of *Homer1a* (Bottai et al., 2002) but is maintained as an intron in *Homer1b/c* so it possible that the probe would also bind to constitutively low levels of *Homer1b/c* in the

nucleus prior to intron 5 being spliced out (Figure 2.2). However, our lab performed verification experiments to ensure that the amount of *Homer1b/c* that bound by the probe did not significantly contribute to the overall *Homer1a* signal. These experiments confirmed that the constitutive expression of *H1b/c* in CA3 and CA1 was extremely low, resulting in very few numbers of RNA transcripts, and the contamination of the total *H1a* signal by *H1b/c* was only about 0.5% $\pm$ 0.19% (personal communication from Aubrey Demchuk, 2013).

The sequence of the *Homer1a* 3'UTR probe was as follows:

```
CATTGAAGGTTGTTTTTGTATGCCAACAGGAGGAAAGCTTGAGTTGCTGCTGATTCTTAAAAGAATTCTG
TATTCTAAAAGATACACATCATGTTCTAAATGCATTTTAAACTAGTGACATTAGTTATTGGGCATACTGT
GGTATTACTAGACTACAAAGAGGAATATGAAGTGGCACCATTGAAAGTATTTTTTTAAAAAGCCTGTCTA
CCTTAACACTAATTTTTACCCTTATTTAAATGCTTTTTACTAAACAGTTTTAGGTAAAATTAAGAAAACA
GTTTTGTTGACTGCACATCTTTTAGAAGGACCAACTTTTAGAGAATTACATTCTTTGACAGATTAAAAAT
TGCAAAGTGAGATATTTCAAACTCTTAAGTGAGTTTTATTGCCGTTGGACTGCATTAATACGGACATACG
ATTAAACTTAGTAGACCAACACTGAGGGATCTCCTTACCAGGCTGCAGAACAAGGAAATTAAGCAATAAA
TGGGACTTGTGAATGGAAGGACACTCTACTGCTAGTGCTAGTAATTCTGCATAAGATGGTATACATTTTG
            AAGAAAGCTGCTTTTAATTACTTTTAATAATGATTTTAATTACTCT
```

All solutions and buffers were prepared with distilled water filtered through a Nanopure water purification system (Barnstead, ThermoFisher Scientific, Waltham, MA) set at 18.2 μOhms to remove DNAses and RNAses. Slides were removed from -80^{o}C storage, thawed for 30 min at room temperature (~21^{o}C), and then fixed in 4% paraformaldehyde for 7 min at 4^{o}C. Following a 2X saline sodium citrate buffer (SSC) wash, slides were then treated with acetic anhydride in triethanolamine buffer for 10 min to lower background signal by binding to polar groups that bind to the probe, followed by a 5 min 1:1 acetone-methanol treatment (to perforate nuclear envelope to enable probe penetration). Sections were then pre-hybridized with hybridization buffer for a minimum of 1 h to prevent background staining, and then incubated overnight (16 h) at 56^{o}C in the

hybridization oven with the *Homer1a* 3'UTR labeled riboprobes. Slides were then cooled for 15 min, incubated with RNAse A for 30 min to digest free single-stranded RNA that did not bind to the probe and then washed in a succession of buffers in increasing stringency. To quench the endogenous peroxidases that would bind to the antibody, slides were washed in 2% hydrogen peroxide for 15 min, and washed in buffered solution. TSA blocking buffer with 5% normal sheep serum was pipetted onto each slide to block all non-specific binding sites for anti-FITC to reduce the background and incubated for 40 min at room temperature. The antibody anti-FITC (Jackson Immuno Research, Cedarlane, Burlington, NC) was added, then the slides were incubated at 4^{o}C for 18 h. After three more washes in buffered solution, slides were then incubated with 1:100 FITC-tyramide (PerkinElmer, Waltham, MA) for 30 min, washed in buffer, and counter-stained with DAPI (Sigma-Aldrich, St. Louis, MI). DAPI, or 4',6-diamidino-2-phenylindole, is a fluorescent dye that binds selectively to adenine-thymine rich regions in DNA (Kapuscinski, 1995). DAPI absorbs light maximally at 340nm and emits light maximally at 453nm, which is in the blue range of fluorescence emission (Peters, 1979), and it has been commonly applied in RNA expression analyses as a counterstain to identify neuronal nuclei (Guzowski et al., 1999; Guzowski et al., 2006; Montes-Rodriguez et al, 2013; Witharana et al., 2016). It should be noted that DAPI has also been shown to bind to adenine-uracil regions of RNA, although with lower affinity than to DNA (Manzini, Xodo, Barcellona, & Quadrifoglio, 1985; Tanious, Veal, Buczak, Ratmeyer, & Wilson, 1992). DAPI-RNA interactions show significantly lower fluorescence (20% of DAPI-DNA signal), and emission of DAPI-RNA shifts to about 500 nm (Krishan & Dandekar, 2005; Kapuscinski, 1995). Therefore, it is possible that the DAPI counterstain contributes slightly to the *Homer1a*-fluorescein emission signal, although this would be a minor

contribution and should the contamination should be uniform across test groups, likely cancelling out in the cross-group comparisons.

Finally, slides were protected with glass coverslips after applying VectaShield® Antifade Mounting Media for Fluorescence (Vector Labs, Burlingame, CA) to preserve the slides, and then sealed on the sides with clear nail polish.

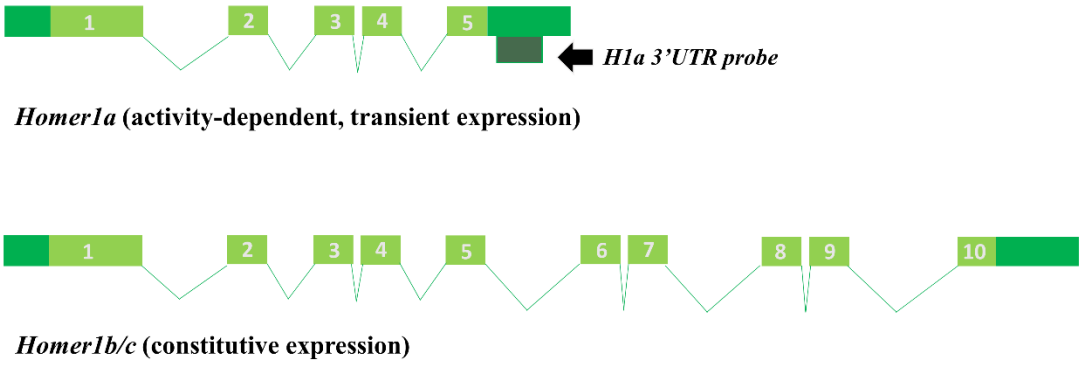


Figure 2.2 Comparison of RNA transcript profiles of *Homer1a* and *Homer1b/c*. *Homer1a* is a short-form transcript of the *Homer1* parent gene and only contains exons 1 through 5; while *Homer1b/c* are long-form transcripts and contain all exons 1-10. *Homer1a* is expressed in a synaptic-activity dependent manner; and is upregulated transiently but then its transcription subsequently declines. *Homer1b/c* are constitutively expressed and their RNA is expressed in a low basal rate in neurons at all times. Solid bars represent exonic sequences, and the thin connector lines represent introns. The 3'UTR probe target is shown with an arrow on the *Homer1a* transcript.

2.6 Image acquisition

2.6.1 NanoZoomer whole slide scanning

To capture large brain areas, we first used a digital slide scanner, NanoZoomer Digital Pathology RS (Model C10730, Hamamatsu Photonics K.K., Japan), equipped with a specialized fluorescence illumination optics module to scan all processed slides. The NanoZoomer Digital Pathology scanner is a high-speed and high-resolution digital imaging system that captures whole-slide digital images by employing TDI (time-delayed-integration) technology. This TDI technology permits the synchronization of the sensor signal with the movement of a microscopic that is being scanned, and uses a triple-chip TDI camera that accurately reproduces sample colors to produce a single RGB (red, green and blue) formatted image. In addition, the NanoZoomer can scan in tri-color fluorescence using three CCD sensors with a full multiband filter system, that covers the emission spectra of green (FITC/530 nm), blue (DAPI/457 nm), and red (628 nm, not applicable here). This fluorescence imaging module uses a light source (halogen lamp) and automatic focus by dark field illumination to detect fluorescence sample locations. The NanoZoomer can also capture weak fluorescence signals, and reduces bleaching, since the excitation light is not targeted onto unwanted regions of the sample, which is an important advantage to prevent significant bleaching of the FITC dye.

Every tenth coronal brain section representing about -2.64 mm to -4.68 mm from bregma (according to the rat brain atlas; Paxinos & Watson, 2007) was scanned to obtain fluorescent images of dorsal CA3 and CA1 of the hippocampus. Each slide consisted of multiple hemi-brains, each representing the different lap conditions within a cohort. Selection of ROIs for scanning was blind to condition as the slide orientations were

unmarked on the ROI review screen. Single median z-planes of 2 μm depth were scanned at 40X magnification, with numerical aperture of 0.75, at a scan speed of ~6 min per slide (2 cm X 2 cm area) at a resolution of 0.23 μm/pixel. Whole slide images were acquired at 4X exposure for all three color channels (approximately 18 ms effective photon collection) and 40X magnification.

Regions of interest (ROIs) from CA3 and CA1 were manually cropped into sub-images by outlining with NDPToolkit image cropping software (Hamamatsu Photonics K.K., Japan), with reference to the Rat Brain Atlas (Paxinos & Watson, 2007). ROIs of CA1 were selected from about the proximal two-thirds of CA1 along the proximal-distal axis in relation to CA3. The entire CA3 present on each hemi-brain on each sampled slide was imaged. Uncompressed images were saved in the Hamamatsu proprietary format, NanoZoomer Digital Pathology Image (NDPI).

2.6.2 Laser confocal microscopy

After acquisition of digital images by NanoZoomer scanning, a sample of z-stacks of the same slides were acquired using a laser scanning confocal microscope (Fluoview FV1000, Olympus America Inc.) with 40X oil immersion objective lens (numerical aperture 1.30) equipped with diode lasers for fluorescence detection of DAPI (blue, 458 nm) and FITC (green, 515 nm). Each image in the stack represented 1μm depth of tissue, so each image stack consisted of between 15 to 18 images. Although the brain sections started as 20 μm thick before FISH, slight tissue shrinkage is common especially during the fixation and dehydration steps. The confocal acquisition method permitted more reliable INF size and intensity measurements by substantially eliminating out-of-focus fluorescent noise (Pawley, 2006). Despite the high scanning speed and resolution, and

larger scanned areas, NanoZoomer images showed a high amount of such noise in the form of blurry FISH signal and out-of-plane fluorescent rings. Confocal z-stacks also permitted accurate nuclear segmentation, which was necessary for some of the analysis (section 2.8). Accurate automated nuclear segmentation was not achievable from the NanoZoomer images because they were only single plane images. For confocal images, the imaging software provided an intensity histogram so that the experimenter could adjust the laser intensity in the appropriate channel to give about 5% saturation of intranuclear foci in the MECS sample, thus covering as much of the dynamic range as possible. The DAPI (blue) laser power was maintained at 5.1%, with photomultiplier (PMT) gain of 460-475. The FITC (green) laser power ranged from 4.7% - 5%, with the PMT gain of 460-520. After laser power and PMT gain were determined, all acquisition parameters were kept constant for all lap groups represented on the slide. Each 24-bit TIFF image in a stack measured 1024 X 1024 pixels, with both horizontal and vertical resolutions of 0.23 μm/pixel.

2.7 Image analysis (Automated INF-boundary-dependent characterization)

2.7.1 NanoZoomer image analysis

Single-plane NDPI images (digital images acquired from the NanoZoomer Digital Pathology scanner) were processed through automated 2D transcription foci intensity analyses using custom Java codes for the open source program ImageJ (NIH, Bethesda, MD). This program first converted the uncompressed NDPI images into TIFF format compatible with ImageJ. These TIFF images were then pre-processed to correct blue/green channel bleed-through since the emission spectra of fluorescein (green, peak at 520 nm, PerkinElmer, Waltham, MA) and DAPI (blue, peak at 461 nm, Sigma-Aldrich,

St. Louis, MI) overlap (see Appendix B for details on bleed-through correction algorithm). Finally, for INF characterization, the automated quantification program assumed that each fluorescein-tagged FISH signal (intranuclear focus labelling for *Homer1a* RNA) possessed a central, bright peak of intensity (P_i) as a single pixel or group of adjacent pixels. Since all digital pixels possess a brightness value (intensity) between 0 – 255, this P_i value cannot exceed 255 intensity units. Also, this local intensity maximum and its adjacently connected pixels in 2D space must meet or exceed two intensity threshold criteria: minimum green intensity (T_i) and the minimum blue (DAPI) background intensity (B_i) (Du & Zhang, 2011). T_i was set at a low value (35) to capture as many INFs as possible, even the faint objects, and were verified to be consistent with human evaluations during prior threshold testing. If the local maximum or adjacently connected pixels fell below T_i, then this object was considered noise. The B_i threshold ensured that the putative INF signal was co-localized with DAPI staining, and thus was actually located within a neuronal nucleus. For each processed image, a corresponding Excel (Microsoft) file was generated, which listed the location of each detected intranuclear focus according to the x- and y- coordinates of the local maximum (intensity peak). For each INF, corresponding maximum intensity values (P_i), total INF size (sum of pixels included in the INF area), and integrated intensities (sum of all pixel intensities within that INF) were also computed.

For NanoZoomer images, the following thresholds were used and held constant across all images: minimum size = 25 pixels at 0.23 μm/pixel, minimum green intensity (T_i) = 35, minimum blue intensity (B_i) = 20. The same parameters were maintained in both CA1 and CA3 stacks.

INF statistics from each image were sorted based on region and lap group, and then either pooled or separated by individual animal. Distributions of intensity values were generated as frequency histograms after uniform binning of intensity ranges. Group medians were calculated from each distribution. Linear regression statistical testing was performed to examine whether there was a relationship between INF intensity values with number of laps.

2.7.2 Confocal image stack analysis

Optical z-stacks of confocal microscope acquired images were saved as 24-bit TIFF stacks using ImageJ software (NIH, Bethesda, MD). Image stacks were then processed through automated 3D INF quantification software written in Java for ImageJ. The automated quantification program assumed that each intranuclear FISH signal (intranuclear focus) possessed a central, bright peak of intensity (P_i) as a single pixel or group of adjacent pixels. This local maximum and its adjacently connected pixels in 3D space must meet or exceed two intensity threshold criteria: minimum green intensity (T_i) and the minimum blue (DAPI) background intensity (B_i) (Du and Zhang, 2011). T_i set to a low value (35) to capture many INFs as possible, even the faint objects, and were verified to be consistent with human evaluations during prior threshold testing. If the local maximum and adjacently connected pixels were below T_i, then this object was discarded as noise. The B_i value ensured that the putative INF signal was co-localized with DAPI staining, and was indeed within a neuronal nucleus. For each z-stack, a corresponding Excel file was generated, which listed the location of each detected intranuclear focus according to the x-, y-, and z- coordinates of the local maximum (intensity peak). For each INF, corresponding maximum intensity values (P_i), total INF

size (sum of pixels included in the INF volume), and integrated intensities (sum of all pixel intensities within that INF) were also computed.

For confocal images, the following thresholds were used and held constant across all z-stacks analyzed: minimum volume = 15 px at 0.32 μm/pixel, minimum green intensity (T_i) = 35, minimum blue intensity (B_i) = 15. The same thresholds were maintained for both CA1 and CA3 stacks.

As with NanoZoomer images, INF statistics from each image were sorted based on region and lap group, and then either pooled or separated by individual animal. Distributions of intensity values were generated as frequency histograms after uniform binning of intensity ranges. Group medians were calculated from each distribution. Linear regression statistical testing was performed to examine whether there was a relationship between INF intensity values with number of laps.

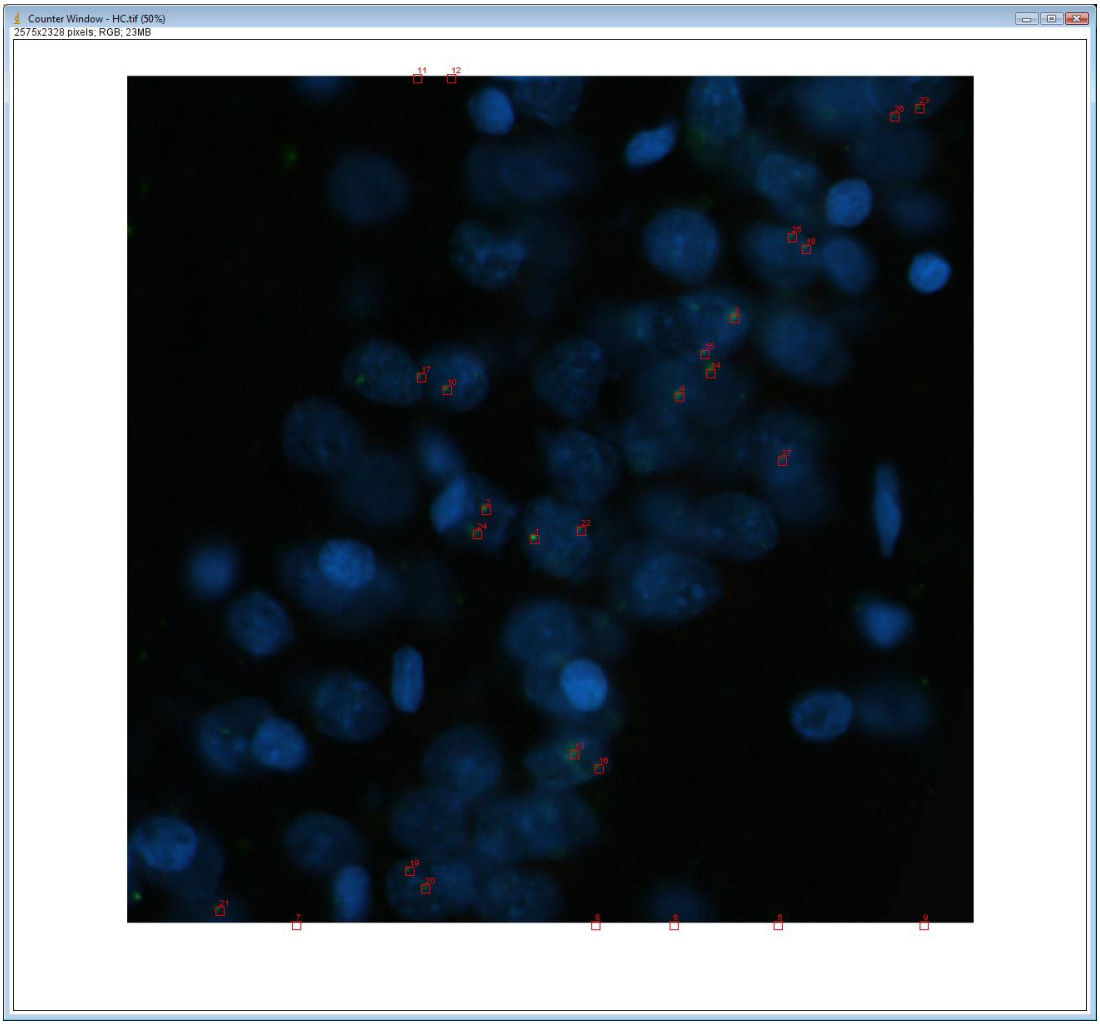


Figure 2.3 Screenshot of custom INF-boundary-dependent software designed for multi-layer z-stack analysis of confocal laser microscope acquired images of fluorescently-labeled *Homer1a* INFs. Green (FITC-labeled) foci were detected by the program based on user-defined thresholds set for pixel intensity, shape, and pixel adjacencies. Red boxes at the bottom of the screen indicate origins of foci detected on a different layer that is either above or below current layer in view.

2.8 Image analysis with segmentation (Automated INF-boundary-independent characterization)

Secondary pixel intensity analysis was performed on the confocal images to compare INF characterization with within-nuclear-boundary (INF-independent) intensity values. INF-independent analysis required the demarcation of DAPI-stained neuronal nuclear boundaries, and total green pixel (*H1a*) intensity within this boundary was cumulatively measured. At the time this study was performed, our software required z-stack (multi-plane) data for accurate segmentation, so only confocal images were included in the INF-independent analyses since our NanoZoomer images were single-plane images. As outlined the previous section, putative INFs had to consist of adjacent pixels and the connected pixels had to adhere to strict shape requirements (i.e. the entire pixel group's outer boundary had to form a near-circular shape). Also, all pixels considered as part of the INF must exceed both T_i and B_i. However, we observed that for higher lap groups, the INFs seemed to diffuse from the peak intensity origin and no longer met the threshold or pixel connectivity requirements of the INF-boundary-dependent algorithm.

The first step of the INF-boundary-independent approach was the application of object segmentation to delineate nuclear (DAPI-stained) boundaries (Figure 2.4). Color TIFF images were separated into blue and green channels. The blue channel image was used for nuclear segmentation via FARSIGHT (Al-Kofahi et al., 2011; Bjornsson et al., 2008; Mesina et al., 2016; Roysam et al., 2008) to segregate all individual neurons in CA1 and CA3. After segmentation, integrated intensity of *Homer1a* signal within segmented nuclear boundaries was quantified by cross-referencing the marked nuclear

boundaries to the corresponding green channel image. The cumulative intensity values were calculated and collated with Matlab (Mathworks, Natick, MA). After image acquisition, fluorescent image files were compiled by referencing from their blind numerical files back to their original image file names that indicated region (CA3 or CA1) and lap group. Image statistics were sorted based on region and lap group and then either pooled, or separated by individual animal. Distributions of intensity values were generated with frequency histograms after equal binning of intensity ranges. Group means, medians and modes were calculated. Linear regression statistics were performed to determine the relationship of intensity values with increasing laps.

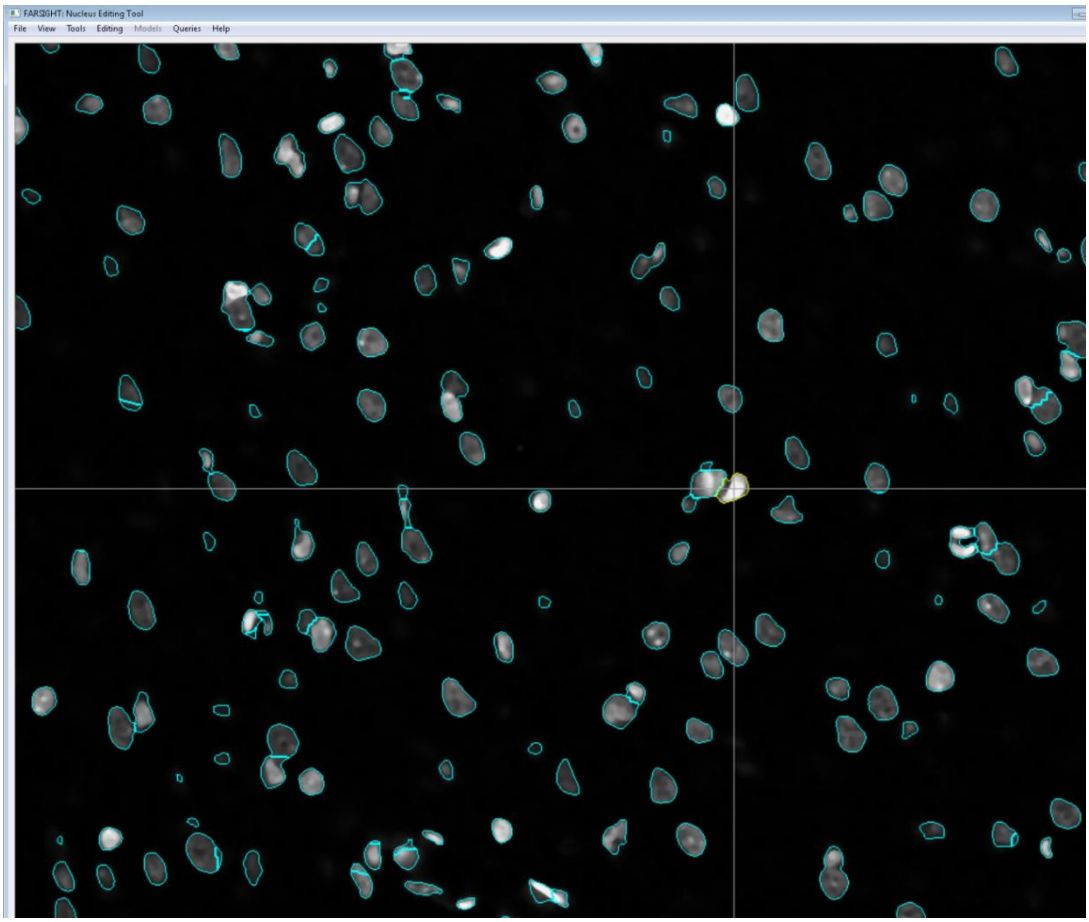


Figure 2.4 Screenshot of an output image generated by neuronal segmentation software used for INF-boundary-independent integrated intensity analysis (FARSIGHT). This ImageJ-based software was customized to identify all individual neurons in CA1 and CA3. After segmentation of nuclear boundaries, integrated intensity of *Homer1a* transcription foci (green) within segmented boundaries was quantified. Over longer time lapses, INFs become diffuse within the nuclear space so INF-boundary-dependent analysis could not accurately quantify mRNA signal in higher lap groups.

3 RESULTS

3.1 Automated INF-boundary-dependent characterization

"Integrated intensity" refers to the sum of the brightness values of all pixels included in a detected intranuclear transcription focus (INF; total signal within a single FITC-labeled RNA focus). Each pixel holds a value between 0 – 255. After intensity data were collected with the INF-boundary-dependent characterization software written for ImageJ (output screen shown in Figure 2.3, described in section 2.7.1), the median integrated intensity for each animal was calculated ("un-pooled"), and also the group median integrated intensity for all INFs belonging to a lap condition regardless of animal ("pooled") was calculated. Both the un-pooled and pooled medians were calculated for comparison since the between-animal variation was high. Then, linear regression tests were performed on both the average un-pooled and pooled median values for CA1 and CA3. The integrated intensity analyses were performed independently on NanoZoomer images and then on confocal images. It should be noted that the sample size for confocal images is substantially smaller than the NanoZoomer analysis due to the comparatively longer time required for confocal scanning. As a result of visually observed INF diffusion (Figure 3.3, I and J), the 30 and 50 lap groups were not included in the linear regression tests, but their un-pooled group average medians are shown for comparison in Figures 3.1 and Figure 3.4.

3.1.1 NanoZoomer images

NanoZoomer image files consisted of single focal plane sub-images cropped from uncompressed NDPI files, which were then converted to TIFF format, and then processed through a custom ImageJ software which identified putative INFs based on user-defined

thresholds and other criteria such as INF area in pixels, as described in 2.7.1. These putative INFs and corresponding pixel intensity and size data were then compiled and re-assigned back to their original lap groups and subregions on a running database.

Un-pooled INF data (Figure 3.1): For each animal, the INF data from all its CA1 images were combined to generate a single distribution, and a single median integrated intensity value was calculated from this distribution. This was also done for the animal's CA3 images, generating another median integrated intensity value. Within a lap group and subregion, each animal's median values was averaged to generate a within-lap group average, and these averages were used to perform linear regression statistics across laps 1 through 25. The un-pooled average median regressions are shown in Figure 3.1. Average group medians for home-cage and MECS were not included in the regression analyses. Significant linear trends were observed in the un-pooled group comparisons for only CA1 (r^2=0.9125, $p<0.05$), but not significant for CA3 (r^2=0.421, p=0.236).

Regression statistical testing for the un-pooled NanoZoomer average lap group medians (divided by 1000 for concise graphing) yielded the following parameters:
CA1: y=0.078x + 4.1, r^2=0.9125; F(1,3)=31.266; p=0.0113; 95% confidence interval (CI) of intercept: 4.1$\pm$0.068; 95% CI slope: 0.08$\pm$0.045. Home-cage average = 2.596 and MECS average = 8.284.
CA3: y=0.011x+3.57, r^2=0.421; F(1,3)=2.188; p=0.236; 95% CI intercept: 3.57 $\pm$0.372; 95% CI slope: 0.011$\pm$0.025. Home-cage average = 2.830, MECS average = 7.141.

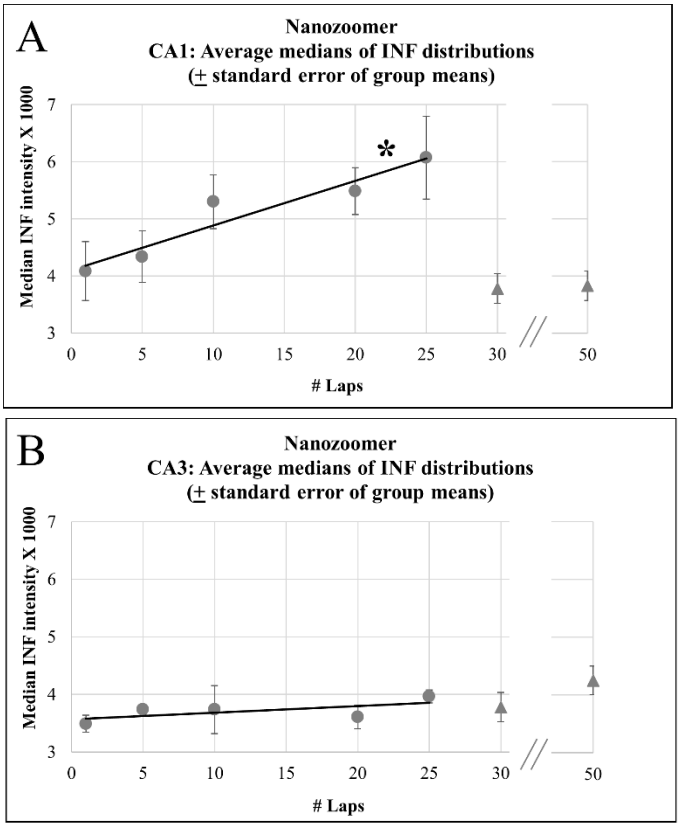


Figure 3.1 Un-pooled, per-animal comparisons; Linear regression statistics on average median INF integrated intensity (summed brightness values X 1000) of each subject within a lap group based on NanoZoomer image analysis with INF-boundary-dependent characterization. Dots represent the group average of all median INF intensities per lap group. Triangles represent averages from 30 and 50 lap groups, neither of which were included in the regression calculations due to INF diffusion. Integrated intensity (y-axis) reported in brightness values (digital pixel brightness can measure 0 – 255 per pixel). Error bars indicate standard error of the group mean between rats within a lap group: 1 lap (n =4 rats), 5 laps (n=4 rats), 10 laps (n=3 rats), 20 laps (n=3 rats), 25 laps (n=5 rats), 30 laps (n=2 rats), 50 laps (n=2 rats), caged controls (n=4 rats), MECS control (n=4 rats). **(A)**: NanoZoomer CA1 average median INF intensity. Regression calculated on laps 1 – 25. $y=0.078x + 4.1$, $r^2=0.9125$; $F(1,3)=31.266$; $p=0.0113$; 95% CI intercept: 4.1 ± 0.068; 95% CI slope: 0.08 ± 0.045. Home-cage average = 2.596 and MECS average = 8.284 (not pictured). **(B)**: NanoZoomer CA3 average median INF intensity. Regression calculated on laps 1 – 25. $y=0.011x+3.57$, $r^2=0.421$; $F(1,3)=2.188$; $p=0.236$; 95% CI intercept: 3.57 ± 0.372; 95% CI slope: 0.011 ± 0.025. Home-cage average = 2.830, MECS average = 7.141 (not pictured). Figure also published in Witharana *et al.*, 2018 (Figure 3a/b).

Pooled INF data (linear regression not displayed in figure form): Since the between-animal variation was relatively high (as observed by the standard error of the group means, especially in CA1), all INFs of animals within a lap group were also pooled and regression statistics were performed on the entire distribution of pooled intensities for the NanoZoomer image data to compare with un-pooled animal medians (Figure 3.1). Integrated intensity values of all INFs within a lap group (all animals' data collapsed) followed a log-normal distribution (Figure 3.2). Both raw and log-transformed data are shown in Figure 3.2 to demonstrate the lognormality and range of INF integrated intensities. The medians of the pooled log integrated intensities were also calculated for each group (shown as vertical red lines in Figure 3.2). Linear regression analyses were performed on the pooled medians for 1 to 25 lap groups. Medians for home-cage and MECS were not included in the regression analyses. Significant linear trends were observed in the pooled distributions for both CA1 (r^2=0.907, $p<0.05$) and CA3 (r^2=0.864, $p<0.05$).

Regression statistical testing for the pooled NanoZoomer INF values (divided by 1000 for concise graphing) yielded the following parameters:

CA1: $y=0.061x + 4.480$; $F(1,3)=29.29$; $r^2=0.907$; $p=0.012$; 95% confidence interval (CI) of intercept 4.48 $\pm$ 0.548; 95% CI of slope 0.061$\pm$0.036. Home-cage median: 2.505, MECS median: 8.992.

CA3: $y=0.019x + 3.559$; $F(1,3)=19.02$, $r^2=0.864$, $p=0.022$, 95% CI of intercept 3.559$\pm$0.217; 95% CI of slope 0.019$\pm$0.014. Home-cage median: 2.527, MECS median: 8.277.

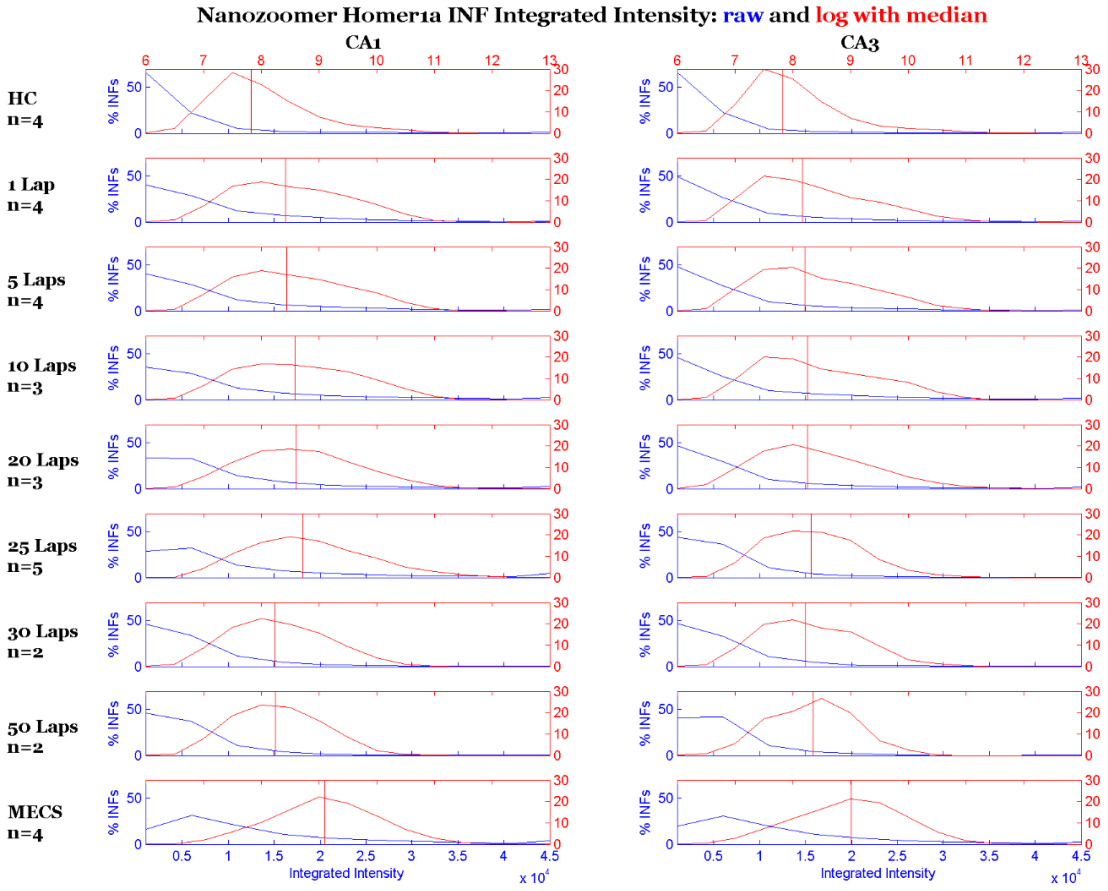


Figure 3.2 Pooled distributions of integrated intensity values of all detected *Homer1a* INFs within a lap group (all animals' INFs within a test group pooled into a single distribution) derived from INF-boundary dependent analysis of NanoZoomer images. Raw distributions of Homer integrated intensities pooled within the different lap groups in blue. Log distributions shown in red (medians shown as vertical red lines). X-axes are integrated intensity values of INFs (x 10^4) and y-axes depict the % of the total INF population in that test group with the corresponding integrated intensity. Each pixel included in an INF can have an intensity (brightness) value between 0-255. The integrated intensity is the sum of all pixels included within that particular INF based on the shape and threshold boundaries outlined in section 2.7.1.

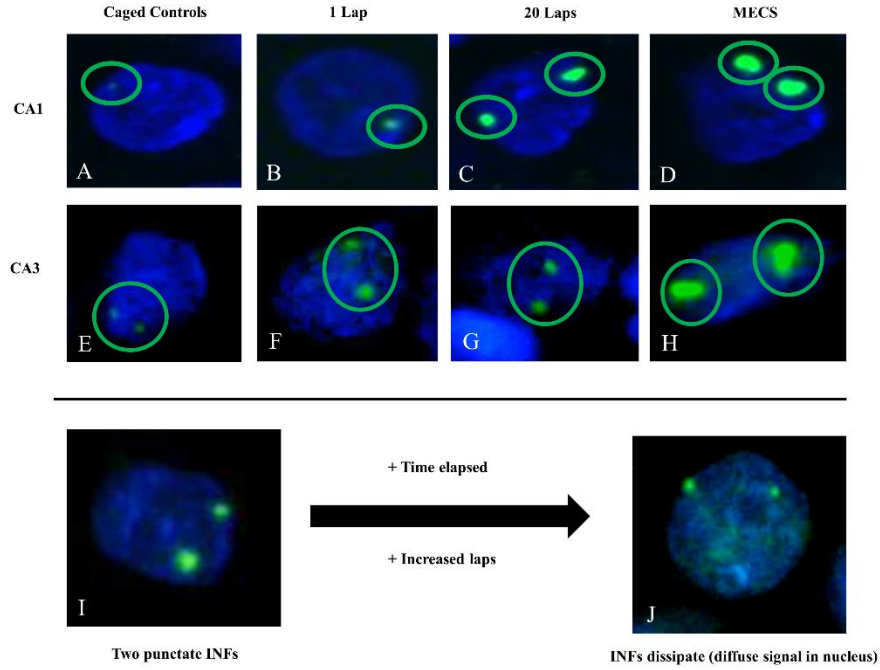


Figure 3.3 Sample images of neuronal nuclei and intranuclear transcription foci of *Homer1a* mRNA tagged with fluorescent label. Top panel: Single-plane NanoZoomer images of DAPI-stained single neuronal nuclei containing fluorescently-labeled *Homer1a* intranuclear transcription foci from CA1 and CA3 in rats that ran 0 (A & E); 1 (B & F); 20 (C & G) laps, or were administered MECS (D & H). Some cells may only express one focus (stochastic phenomenon of mono-allelic expression) or the second focus may be present in another image plane above or below the z-plane in view. The second focus may be detected only in the z-axis in another image plane. **Bottom Panel:** I) A standard neuronal nucleus with two green punctate INFs with tight boundaries (from an image of a subject that ran 20 laps, CA1). These punctate INFs correspond to the fluorescent-labeled transcription loci of the two *Homer1a* alleles on either chromosome. J) An example nucleus (subject ran 30 laps, CA1) with two smaller punctate INFs near the top, but also visible green signal throughout the nuclear area. This diffuse signal is still detectable through the automated algorithm above threshold but cannot contribute to accurate integrated intensity values through INF-boundary-dependent characterization protocol as this signal does not reside within the boundaries of a clean INF packet.

Figure also published in Witharana *et al.*, 2018 (Figure 3).

Estimation of activation proportions: For the un-pooled NanoZoomer data set, total INFs detected were summed and then divided by an estimate of the total neurons sampled for each animal and for both CA1 and CA3 (data shown in Table 3.1). The total neuronal count was estimated by summing all blue pixels within a specific intensity range that should eliminate glial or background DAPI stain (see Chapter 6: Appendix A for details on this pixel-based neuronal count estimation). Presumably, most neurons expressed two transcription puncta, one for each chromosome, so the total INF counts were divided by two to determine the approximate neuron count. It should be noted that although most hippocampal neurons express bi-allelic transcription of *Homer1a*, stochastic mono-allelic expression has been documented in some cells (personal communication from Aubrey Demchuk, 2016). However, we also presume this random mono-allelic expression occurs at an equal rate across all lap groups, so the under-estimation error in dividing the total INF counts by two should cancel out in the cross-group comparison. In any event, this estimated activation proportion should not change dramatically with increasing laps since this measure should reflect that the same place cells were activated on each traversal, leading to accumulation of transcribed RNA. To analyze whether the average activation proportions changed across laps or stayed the same, a linear regression test was applied to values in CA1 across laps; and also in CA3. In CA1, regression statistics yielded these parameters: $y=0.09x + 26.85$. $r^2=0.06$, $F(1,3) = 0.1989$, $p=0.686$. In CA3, regression statistics yielded these parameters: $y=-0.15x +20.43$, $r^2=0.17$, $F(1,3)=0.608$, $p=0.493$. In both CA1 and CA3, changes in average activation proportions were not significant, confirming that the number of place cells activated stayed relatively consistent across laps.

Table 3.1 Estimated percent activation in CA1 and CA3 across laps from NanoZoomer image sets. MECS estimation in CA1 is over 100% due to minor error in the gross estimation method by simply counting blue pixels within neuronal intensity range and dividing total INFs by two. Presumably, this error transfers across all lap groups and both regions, to serve as a gross estimate that cells were not substantially recruited across laps (same place cells activated on each traversal). Regression statistics did not show a significant effect of laps on percent of cells activated (statistics in text) and thus, approximately the same number of cells were activated per lap. Table also published in Witharana *et al.*, 2018 (Table 1).

Number of laps	**CA1 raw %**	**CA3 raw %**
Home-cage	8.50	8.77
1	30.48	20.00
5	23.60	18.78
10	28.18	18.68
20	25.16	22.82
25	32.29	12.47
MECS	118.11	71.77

3.1.2 Confocal image stacks

Confocal image stacks consisted of multi-plane TIFF files with z-axis data. Image stacks were analyzed with custom ImageJ software which identified putative INFs based on user-defined thresholds and other criteria, including minimum INF volume in pixels as

outlined in section. These putative INFs and corresponding pixel intensity and size data were then compiled and re-assigned back to their original lap groups and subregions on a running database.

Un-pooled INF data (Figure 3.4): For each animal, the INF data from all its CA1 image stacks were combined to generate a single distribution, and a single median integrated intensity value was calculated from this distribution. This was also done for the animal's CA3 image stacks, generating another median integrated intensity value. Within a lap group and subregion, each animal's median values was averaged to generate a within-lap group average, and these averages were used to perform linear regression statistics across laps 1 through 25. The un-pooled average median regressions are shown in Figure 3.4. Average group medians for home-cage and MECS were not included in the regression analyses. Significant linear relationships were observed across laps in both CA1 (r^2= 0.90, $p<0.05$) and CA3 (r^2=0.937, $p<0.05$).

Linear regression testing on laps 1 to 25 yielded these parameters:

CA1: y= 0.16x + 3.948; r^2= 0.90; F(1,3)=27.78; p=0.013; 95% confidence interval (CI) of intercept: 3.948 $\pm$ 1.503; 95% CI slope: 0.164 $\pm$ 0.099. Home-cage average = 4.09, MECS average = 12.313.

CA3: y=0.13x + 3.978; r^2=0.937; F(1,3)=44.401; p=0.007; 95% CI intercept: 3.978 $\pm$ 0.937; 95% CI slope: 0.129 $\pm$ 0.062. Home-cage average = 3.455, MECS average = 10.600.

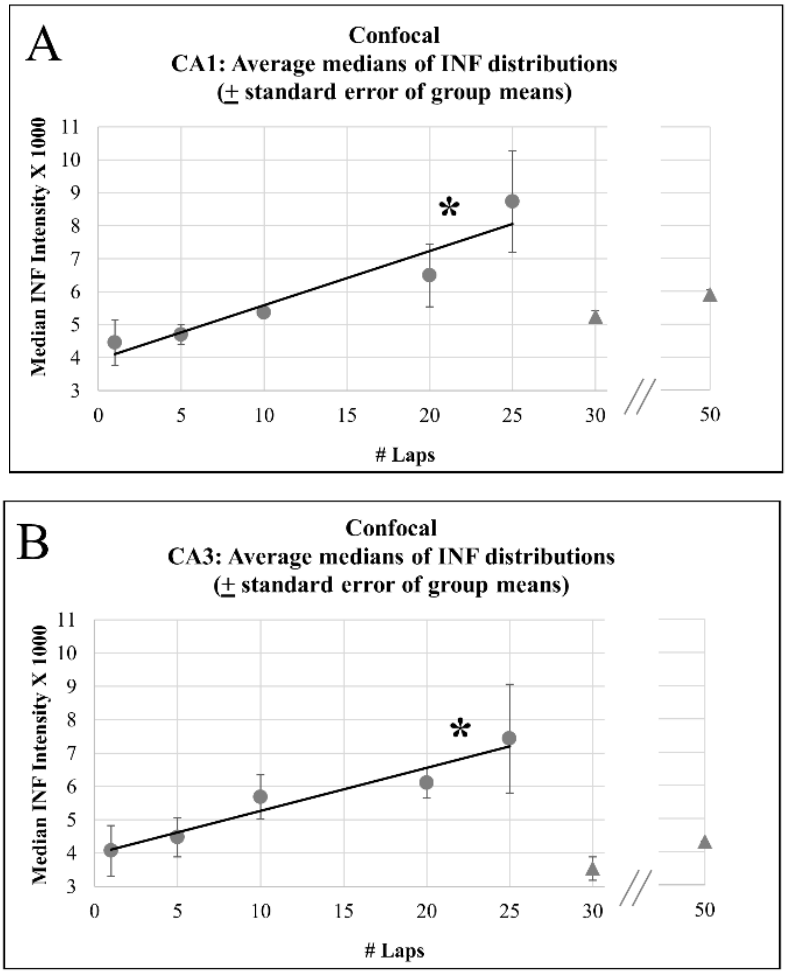


Figure 3.4 Un-pooled, per-animal comparisons: Linear regression statistics on average median INF integrated intensity (summed brightness values X 1000) of each subject within a lap group based on confocal image analysis with INF-boundary-dependent characterization. Dots represent the group average of all median INF intensities per lap group. Triangles represent averages from 30 and 50 lap groups, neither of which were included in the regression calculations due to INF diffusion. Integrated intensity (y-axis) reported in brightness values (digital pixel brightness can measure 0 – 255 per pixel). Intensities divided by 1000 to highlight comparison of changes of transcription focus intensity across lap groups. Error bars indicate standard error of the group mean between rats within a lap group: 1 lap (n =4 rats), 5 laps (n=4 rats), 10 laps (n=3 rats), 20 laps (n=3 rats), 25 laps (n=5 rats), 30 laps (n=2 rats), 50 laps (n=2 rats), caged controls (n=4 rats), MECS control (n=4 rats). **(A)**: Confocal CA1 average median INF intensity. Regression calculated on laps 1- 25. $y = 0.16x + 3.948$; $r^2 = 0.90$; $F(1,3)=27.78$; $p=0.013$; 95% CI intercept: 3.948 ± 1.503; 95% CI slope: 0.164 ± 0.099. Home-cage average = 4.09, MECS average = 12.313 (not pictured). **(B)**: Confocal CA3 average median INF intensity. Regression calculated on laps 1- 25. $y=0.13x + 3.978$; $r^2=0.937$; $F(1,3)=44.401$; $p=0.007$; 95% CI intercept: 3.978 ± 0.937; 95% CI slope: 0.129 ± 0.062. Home-cage average = 3.455, MECS average = 10.600 (not pictured).
Figure also published in Witharana *et al.*, 2018 (Figure 3c/d).

Pooled INF data (linear regression not displayed in figure form): Between-animal variation was also high in the confocal imaging analysis (as observed by the standard error of the group means in both CA1 and CA3). In addition, the confocal image stack sample size was lower than compared to the large amount of images available in the NanoZoomer analysis, so again all INFs of animals within a lap group were pooled and regression statistics were performed on the entire distribution of pooled intensities for the confocal image data to compare with un-pooled animal medians. Integrated intensity values of all INFs within a lap group (all animals' data collapsed) followed a log-normal distribution (Figure 3.5). Both raw and log-transformed data are shown in Figure 3.5 to demonstrate the lognormality and range of INF integrated intensities. The medians of the pooled log integrated intensities were also calculated for each group (shown as vertical red lines in Figure 3.5). Linear regression analyses were performed on the pooled medians for 1 to 25 lap groups. Medians for home-cage and MECS were not included in the regression analyses. Significant linear trends were observed in the pooled distributions for both CA1 (r^2=0.961, $p<0.005$) and CA3 (r^2=0.983.864, $p<0.005$).

Linear regression statistical testing for the pooled confocal INF values (divided by 1000 for concise graphing) yielded the following parameters:

CA1: $y= 0.118x + 4.57$, r^2=0.961; $F(1,3)=74.72$; $p<0.005$; 95% confidence interval (CI) of intercept: 4.57 ± 0.662; 95% CI slope: 0.118 ± 0.044. Home-cage median = 3.580, MECS median = 11.867.

CA3:$y=0.058x + 4.562$; r^2=0.983; $F(1,3)=170.05$; $p< 0.005$; 95% CI intercept: 4.562 ± 0.216; 95% CI slope: 0.058 ± 0.015. Home-cage median = 3.481, MECS median = 9.209.

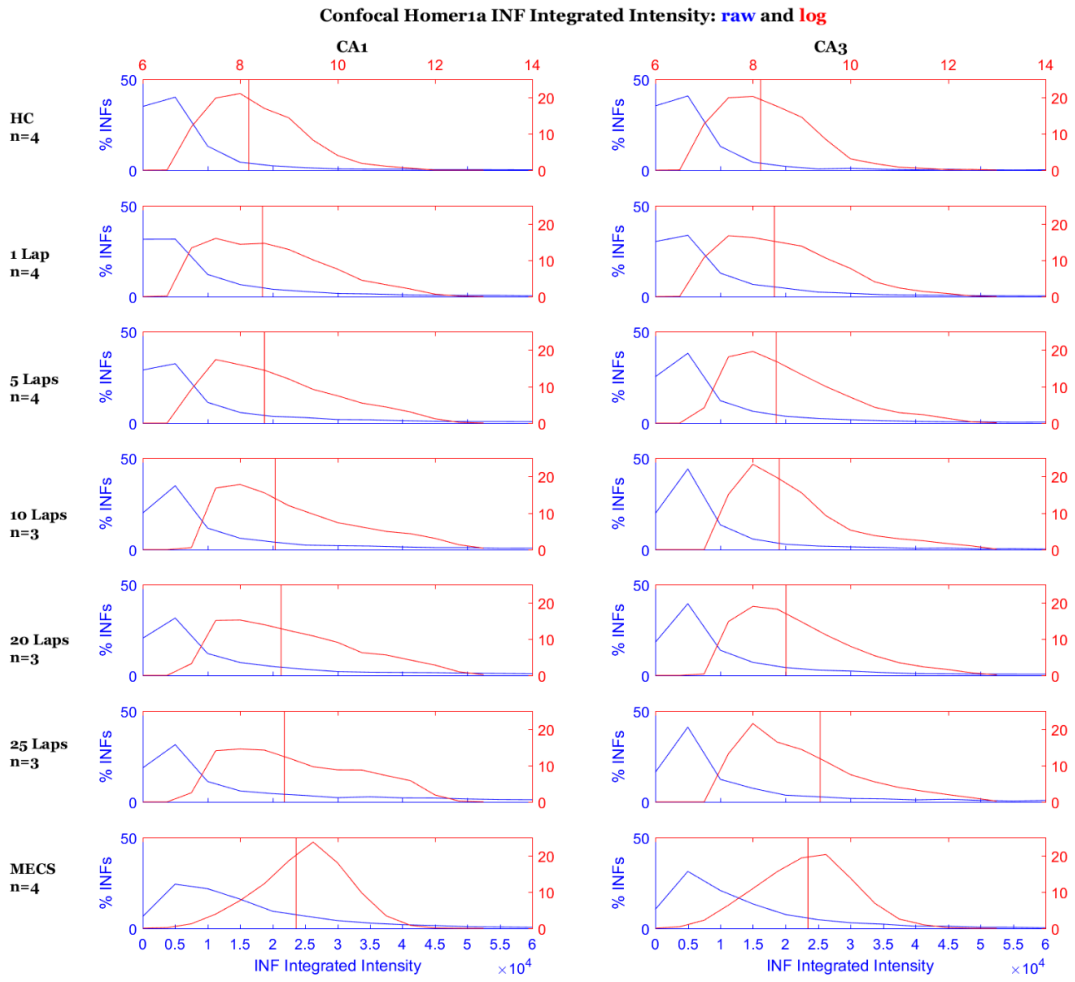


Figure 3.5 Pooled distributions of integrated intensity values of all detected *Homer1a* INFs within a lap group (all animals' INFs within a test group pooled into a single distribution) derived INF-boundary-dependent analysis of confocal images. Raw distributions of Homer integrated intensities pooled within the different lap groups in blue. Log distributions shown in red (median shown in vertical lines on log scale). X-axes are integrated intensity values of INFs (x 10^4) and y-axes depict the % of the total INF population in that test group with the corresponding integrated intensity. Each pixel included in an INF can have an intensity (brightness) value between 0-255. The integrated intensity is the sum of all pixels included within that particular INF based on the shape and threshold boundaries (in the x-,y-, and z-axes) outlined in section 2.7.2.

3.2 Automated INF-boundary-independent characterization (nuclear segmentation)

Although INF-boundary-dependent characterization was useful for analyzing the integrated intensity of transcription foci in lap groups 1 to 25, visual examination of both NanoZoomer and confocal images showed that the INF boundaries were not as well defined in animals that ran 30 or 50 laps when compared to the punctate foci from animals that ran fewer laps (Figure 3.3, I & J). These diffuse signals from the 30- and 50-lap subjects were not accurately characterized by the INF-boundary-dependent method, because of the nature of the boundary detection algorithm described in section 2.7.1 and 2.7.2. Thus, samples of confocal images were analyzed using an INF-boundary-independent analysis (FARSIGHT) which required the detection and segmentation of neuronal nuclear boundaries (marked by DAPI signal), and then the intensities of all green pixels within these nuclear zones (whether diffuse or bound to punctate INFs) were summed (described in detail in section 2.8). Each neuronal nucleus that could be delineated by the segmentation showed a corresponding integrated green value which represented the amount of detectable *Homer1a* RNA signal within that particular nuclear boundary regardless of whether it was contained in a distinct focus. Subjects were pooled within lap groups and the corresponding sums of green pixels were then analyzed and median nuclear signals were calculated. Linear regression tests were performed on the distribution medians (Figure 3.6), but now included 30 and 50 lap animals (whereas these two lap groups were excluded from regression tests in the INF-boundary-dependent analyses). In this INF-boundary-independent intensity analysis, both CA1 and CA3 median integrated intensities increased linearly up to 50 laps. Linear regression test

results are reported in the legend of Figure 3.6. The total number of neurons sampled is listed in Table 3.2.

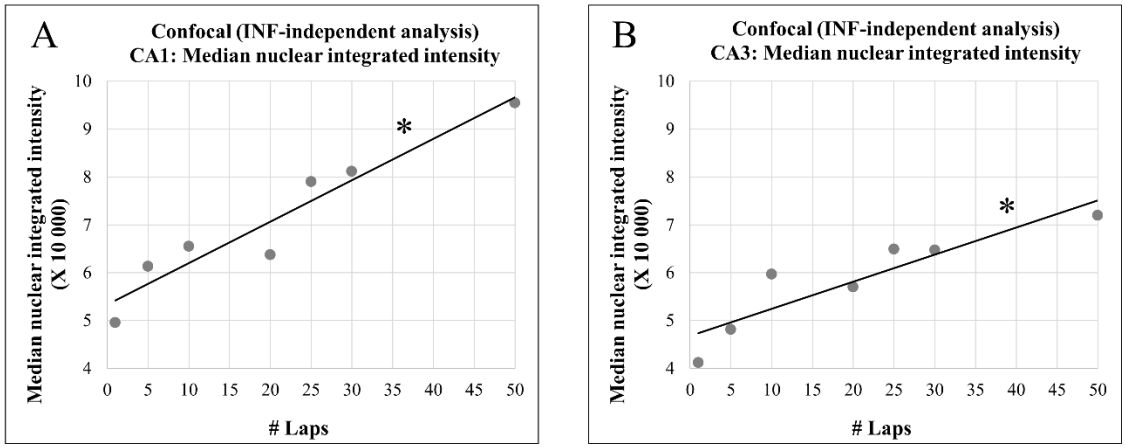


Figure 3.6 Linear regression performed on average median nuclear integrated intensity distributions from confocal images across lap groups in (A) CA1 and (B) CA3 following INF-boundary-independent analysis (FARSIGHT). Since this INF-boundary-independent (nuclear segmentation) analysis includes green signal from diffuse RNA signal outside of clear INF boundaries, the inclusion of 30 and 50 lap group values contributed to significant linear increases in transcription signal across laps. Integrated intensity (y-axis) measured in brightness values (digital pixel brightness can measure 0 – 255 per pixel). **(A)** CA1: Regression statistics: $y= 0.087x + 5.336$, $r^2=0.917$, $F(1,5)=55.09$, $p=0.0007$, 95% CI intercept: 5.336 ± 0.765, 95% CI slope: 0.087 ± 0.023. Home-cage median: 3.989, MECS median: 12.597 (not pictured). **(B)** CA3: Regression statistics: $y=0.057x + 4.680$, $r^2=0.820$, $F(1,5)=22.814$, $p=0.005$, 95% CI intercept: 4.680 ± 0.777, 95% CI slope: 0.057 ± 0.030. Home-cage median: 4.404, MECS median: 17.79 (not pictured). Figure also published in Witharana *et al.*, 2018 (Figure 5).

Table 3.2 Total number of neurons analyzed for INF-boundary-independent integrated intensity based on nuclear segmentation. A cross-section of the available data was analyzed for intranuclear integrated intensity and due to low numbers per animal, median nuclear integrated intensities were pooled within a lap group. Table also published in Witharana *et al.*, 2018 (Table 2).

Number of laps	**CA1 Total neurons sampled**	**CA3 Total neurons sampled**
Home-cage	1179	783
1	550	1830
5	1961	1045
10	1593	1953
20	861	1972
25	1218	1589
30	663	768
50	663	768
MECS	1718	2024

4 DISCUSSION & CONCLUSIONS

4.1 *Homer1a* transcription foci increased with laps within a narrow dynamic range

To estimate the relationship between neuronal spiking and expression of the synaptic plasticity-related immediate-early gene *Homer1a (H1a)*, rats ran 0, 1, 5, 10, 20, 25, 30 or 50 unidirectional laps on a circular track for food reward and then their brains were analyzed for *H1a* RNA expression. Automated fluorescent image analyses revealed systematic increases in integrated intensity of *H1a* intranuclear transcription foci as laps increased in both CA1 and CA3 subregions of the hippocampus, with no change in the percentage of cells passing the $H1a^{+}$ threshold (estimated neuronal activation proportions remained relatively constant across laps). It should be emphasized that the methods for estimating total neuronal numbers are approximate. The intensity ranges of neurons and glia overlap to some degree, and it may be the case that some neurons exhibit only a single intranuclear transcription focus due to mono-allelic expression (Eckersley-Maslin & Spector, 2014); however, neither error would affect the overall conclusion concerning the observed increase in expression over laps, since these errors would apply equally to all groups.

Therefore, the intensity data support the hypothesis that there is a linear, but low-gain, relationship between number of spikes and, presumably, the average amount of transcribed *H1a* RNA signal as a result of direct electro-transcriptional coupling, over a 50-fold range of spiking activity. The data also suggest dynamic properties of electro-transcriptional regulation wherein the average number of IEG transcription cycles within the overall population increases monotonically with the number of times that the cells

have been activated within a fixed time window, at least over a certain range, in hippocampal subfields CA1 and CA3, even after two weeks of experience in a familiar environment. Since this direct proportionality appears to be preserved in both CA1 and CA3 (over a certain behavioural range), it is conceivable that at least some part of the complex biochemical cascade occurring during electro-transcriptional coupling (Fields et al., 1997; Flavell & Greenberg, 2008; Link et al., 1995) may manifest as visible and quantifiable variations in INF characteristics in certain behavioural paradigms. However, the data do not address how such changes might be manifested at the single neuron level since these results are based on averages over large populations of cells.

This study used two optical image acquisition systems and several analytical methods to characterize and verify that, on average, the fluorescence intensity increased as subjects ran consecutive laps. Parallel analysis streams were applied for verification in image acquisition (NanoZoomer scanner versus laser confocal microscope); and also for fluorescence intensity detection (INF-boundary-dependent versus INF-boundary-independent analyses). In the INF-boundary-dependent protocol using NanoZoomer images, only single focal plane z-stack analyses were performed, which would have led to a consistent degree of measurement error since the data only show INFs in the x-y plane. For example, the inclusion of partial INFs would underestimate certain data points, or the inclusion of out-of-focus INFs from below or above the focal plane would have inflated actual values. Presumably this error is constant across lap-groups. Furthermore, the same images were also sampled on a confocal laser microscope with z-stack information and similar trends of fluorescence intensity increase were confirmed; although the INF sample is much smaller because of the time involved in confocal z-stack acquisition. In the

future, if confocal acquisition times can be dramatically reduced from a technical standpoint, it would be advantageous to replicate this analysis with large-scale confocal imaging including the z-axis.

The eventual disbanding of transcription puncta into blurred and diffuse signal within the neuronal nucleus at the longest time-points after initiation of transcription (i.e. at 30 and 50 laps) necessitated the use of an alternate signal characterization protocol (INF-boundary-independent nuclear segmentation based approach). INF-boundary-independent quantification of transcription foci intensity showed that electro-transcriptional coupling is maintained within the specific behaviour up to 50 laps or approximately 10 minutes of consecutive lap-running. As such, while both methods are useful for intensity quantization, the INF-boundary-independent method may be generally preferable and computationally more efficient. Furthermore, the INF-boundary-independent method is likely less biased since it employs less user-defined thresholding and criteria. Presumably, the initiation of transcription occurs at localized nuclear zones and as nascent RNA molecules are transcribed, these transcription puncta or factories gradually disband as individual RNA molecules separate from the conglomerate of newly synthesized transcripts and diffuse to other nuclear zones *en route* to exportation to the cytoplasm. While the diffusion of *Homer1a* RNA from the transcription origin seems to begin between 25 to 30 laps in this current experimental paradigm, it is conceivable that time-courses of diffusion patterns vary depending on the immediate-early gene (*Arc, zif268*) or behavioural requirements (running on a narrow track versus free exploration, for example).

4.2 Electro-transcriptional coupling captured by IEG fluorescence analysis

Although a linear relationship exists in INF fluorescence across laps, is this empirical relationship a true representation of the amount of *Homer1a* RNA, or is the relationship merely a reflection of a stochastic fluorescence amplification process? Presumably, even if there is a stochastic process of random non-uniform process of fluorescent probe binding or dye amplification throughout the sample, we would assume this random process would cancel out across test groups since it would occur at similar rates in all test groups. Also, all test groups were represented on each processed slide, mainly to eliminate these technical biases. Therefore, the continuous variation in intensity and size of transcriptional foci most likely indicates variable numbers of RNA transcripts generated from repeated transcription cycles of RNA polymerase activity at the IEG loci. Furthermore, if the data is a true representation of *Homer1a* RNA quantities, does the absolute mRNA amount coincide in perfect relationship with exact spike rates, and if so, is there a minimal "spike number" threshold to trigger detectable transcription in the form of fluorescent transcription foci? For example, is a single action potential responsible for triggering a set number of transcription cycles (*t*), such that traversal of the same place field *n* times in a behavioural epoch, results in a number of IEG transcripts that coincides with *n* X *t*? This remains to be investigated systematically, for example, through the use of direct quantification of transcribed RNA during controlled place field firing similar to the PCR quantification performed by Miyashita *et al.* (2009) and Penner *et al.* (2011).

In addition, biomolecular techniques (e.g. Northern blot quantification of transcription factors) could further characterize differential upregulation during electro-transcriptional coupling. Ideally, this quantification would occur in tissue preserved

immediately following electrophysiological recordings to confirm the direct coupling between electrical activity and the onset of genetic transcription. Perhaps future experiments could also capitalize on single-molecule FISH procedures which provide accurate integer counts of mRNA copy numbers in individual activated neurons (Raj, Van Den Bogaard, Rifkin, Van Oudenaarden, & Tyagi, 2008). With recent technical advances in Ca^{2+} imaging in real-time, it could be possible to measure neuronal activity by proxy of calcium influxes and simultaneously visualize live mRNA transcription with the use of conjugate IEG-promoters and fluorescent markers to quantify the exact number of IEG transcripts directly transcribed by a set number of action potentials. Given the complex and dynamic quantitative molecular processes involved in the electro-transcriptional conversion action potentials into packets of RNA, an important follow-up study would be one that observed the real-time output of RNA in live cells when IEG expression is triggered by neuronal electrical activation. With the advent of complex genetic techniques and high-resolution optical imaging, the relationship between the intensity of neural activation and the corresponding magnitude of the IEG transcriptional response at the single neuron level could be investigated in real time with *in vivo* imaging through the use of transgenic animals expressing IEG-promoter driven fluorescent effectors such as *Arc*-enhanced green fluorescent protein (EGFP), akin to the imaging done by Na and colleagues (2016) although they monitored the translation of *Arc* in dendrites caused by glutamate stimulation.

4.3 Possible explanations for the low dynamic range of *H1a* signal

Although there were increases in transcription foci intensity as the number of laps also increased, the largest rate of change in signal intensity was typically between the

home-cage condition and the first lap, which coincides with previous reports that one lap is sufficient to induce distinct IEG activation (Miyashita et al., 2009). We compared the slope between home-cage and 1 lap to the slope over all lap groups (without including the home-cage). On average, over all analysis methods and regions, the ratio of slopes was 54 (+/- 43 SD). This may reflect a condition in which ongoing *H1a* transcription in the population is very low during rest, but, once initiated following the first lap, which is relatively novel, even on a familiar track as evidenced by plasticity effects in CA1 over multiple spatial learning epochs (Lee et al., 2004), there is very little remaining capacity for new initiation. Since ongoing place cell firing rates are consistent over multiple laps (Maurer et al., 2006; Ekstrom et al., 2001), the non-persistence of this initial burst in *H1a* expression (i.e. the rate of change in expression decreases after lap 1) suggests a mechanism of electro-transcriptional decoupling in which ongoing neural activation ceases to induce further immediate-early gene expression.

According to the "transcription factory" model, each factory contains only a small number (4-30) of RNA polymerase complexes (Rieder et al., 2012), some of which will likely be occupied transcribing other genes. This would provide a signal-transduction based upper limit on the number of copies of *H1a* mRNA that could be produced. Therefore, whether a burst of spikes initiates a cycle of transcription depends entirely on the probability that there is an available RNA polymerase, and whether multiple cycles of transcription can be initiated simultaneously on the *H1a* gene. Apparently, in the case of *H1a,* this probability is about 2/50 = 0.04. In the case of the “transient RNAP cluster” paradigm (Cisse et al., 2013; Cho et al., 2016) the level of permissible transcriptional cycles might be limited by the cluster time of RNA polymerases. This assumes that

transcription stops when RNAPII de-cluster, which would also lead to the ungrouping of RNA transcript copies. In either scenario, increase in median *H1a* transcriptional activity per lap might depend on the limited pool of available or stalled polymerases near the gene promoter (Saha et al., 2011), which might logistically limit the number of cycles of transcription and subsequent re-initiation. Thus, even the most simplified model of the dynamics of the transcriptional process in a population of neurons would predict an asymptotic response as the number of spikes within the relevant time window increases. Such a response could appear relatively linear over some range, but its first derivative would steadily decrease. Therefore, the biological limitations of enzyme availability and temporal requirements for enzymatic activation may control the responsive window of IEG transcription. These cellular constraints might have important implications for electro-transcriptional coupling wherein only the initial activity "burst" can be encoded.

In regards to behavioural context and learning, perhaps *Homer1a* expression levels are regulated by motivation since the first lap at the beginning of the test sessions should be when animals are most hungry and most interested in the food reward. Behavioural studies implicate an important role for *H1a* in regulating motivated learning in rats (Aragona & Carelli, 2006), and since food deprivation is a highly motivating condition, perhaps the initial bursts of *Homer1a* expression are important for encoding the salience of the food reward based on starting motivational levels. However, as the animals become more satiated with each subsequent lap, their motivation could be decreasing and as such the need to encode the action/reward association of lap running in the same place over and over again also decreases. Continuous and repetitive lap running could also introduce as aspect of familiarity, in addition to the previous training on the

same track and room over many days. The redundancy or lack of novel salient features over many laps could cause electro-transcriptional decoupling since no new learning is occurring, and the unchanging context is already becoming a stable representation. In addition, on test day, rats were completing each lap within 30 – 40 s, and this speed stayed relatively constant over subsequent laps. Although speed does not affect the number of spikes within a place field (Ekstrom et al., 2001), it is possible that the consistent speed also decreased the salience of the running epochs since the behaviour became habitual after the first lap, thus also likely contributing to electro-transcriptional decoupling. It would be interesting to alter the test day paradigm so that the rats had to take a pause after a few laps, then resumed running, to see whether the initial burst could be reinstated by renewing the transcriptional response through reactivation of electro-transcriptional coupling. This pause would be introduced only on test day so it served as a novel component of a familiar context, which should be encoded as new information in the presumably stable representation.

4.4 Is IEG fluorescence a reliable indicator of firing rates?

One application-related purpose of the present study was to determine whether IEG fluorescence can be used as a reliable proxy for individual neuronal spikes (or spike bursts) at the level of the individual neurons, which would make it an invaluable tool for the study of experience-dependent changes in neural coding dynamics on a large scale. We conclude that, at least for *H1a*, this is generally not the case when the spike-inducing stimulation or behaviour is repeated multiple times within a short time window. On average, there was only a 3-fold change in fluorescent signal between the home-cage and MECS conditions, and at best a twofold increase in median fluorescence signal (by any

method of analysis) over a 50-fold repetition of the spike inducing behaviour. Given the degree of error expected in the quantification process, this very low signal to noise ratio appears to preclude accurate estimates of spike rate differences within a given population of neurons or between experimental conditions, unless the sample size and/or the spike rate differences are very large.

Also, it was apparent that there were high degrees of between-animal variation in the median integrated intensity in both CA1 and CA3. Qualitatively, the data suggested higher variance in CA1 specifically, based on the higher standard errors of the group averages of median INF intensities. There could be several possibilities for these variations: intrinsic biological variance in gene expression, or systematic bias in the FISH protocol and subsequent sampling. The second explanation could probably be ruled out since the different lap groups were blocked onto the same slides to ensure all brains underwent identical processing and therefore technical variations would apply to all subjects uniformly. During training, some rats seemed to show more motivation or inclination to reach criterion faster than others, so it seems there are natural underlying inter-subject variations in behaviour, which could also indicate intrinsic difference between each individual's learning and memory processing. Indeed, Long-Evans rats show natural individual differences in motivation to stay in darkness or to seek food rewards, and these initial inclinations could reliably predict their level of success in subsequent goal-oriented task learning (Franks, Higgins, & Champagne, 2014). Extra-species evidence supporting the possibility of natural differences in motivation can be seen in the African cichlid fish. Some fish could be trained easily on a spatial task (learners), while some fish would attempt the task but could not succeed (non-learners),

and some fish never attempted the training task at all (non-attempters) (Wood, Desjardins, & Fernald, 2012). In any event, as a result of these underlying inter-subject variations at least in *Homer1a* gene expression, it would be difficult to definitively use IEG images as an absolute readout for firing rates.

In summary, the low gain of the transcriptional response in the present study may reflect a dramatic decrease in electro-transcriptional coupling as the number of exposures of the animal to a cell's place field increases. Such a decrease has been shown previously over a longer time scale (Guzowski et al., 2006). The cause of this decrease is unknown. It could be intrinsic to the transcription induction process itself, or it might reflect rapid habituation of neuronal modulatory processes which might play a role in coupling transcription to spiking. If such a hypothetical fast habituation of permissive modulatory responses is the explanation, then it remains possible that the log-normal distribution of IEG integrated intensities may indeed reflect the log-normal distribution of spiking activity in the hippocampal population on a single trial. This question can only be addressed by directly correlating differential spiking rates in a population of neurons with the corresponding level of IEG expression, which would be technically challenging, but not impossible.

5 REFERENCES

Adams, J.P., & Dudek, S.M. (2005). Late-phase long-term potentiation: getting to the nucleus. *Nature Reviews Neuroscience, 6*:737-743.

Al-Kofahi, Y., Lassoued, W., Grama, K., Nath, S.K., Zhu, J., Oueslati, R.,…Roysam, B. (2011). Cell-based quantification of molecular biomarkers in histopathology specimens. *Histopathology, 59*(1):40-54.

Ango, F., Prezeau, L., Muller, T., Tu, J.C., Xiao, B., & Worley, P.F. (2001) Agonist-indepdent activation of metabotropic glutamate receptors by the intracellular protein Homer. *Nature, 411*:962-965.

Aragona, B.J., & Carelli, R.M. (2006). Dynamic neuroplasticity and the automation of motivated behavior. *Learning & Memory, 13*:558-559.

Banerjee, A., Luong, J.A., Ho, A., Saib, A.O., & Ploski J.E. (2016). Overexpression of *Homer1a* in the basal and lateral amygdala impairs fear conditioning and induces an autism-like social impairment. *Molecular Autism, 7*(1):16. doi:10.1186/s13229-016-0077-9.

Barnes, C.A., Jung. M., McNaughton, B.L., Korol, D., Andreasson, K., & Worley, P.F. (1994). LTP saturation and spatial learning disruption: effects of task variables and saturation levels. *Journal of Neuroscience, 14*(10):5793-5806.

Barnes, C.A., McNaughton, B.L., Mizumori, S.J.Y., Leonard, B.W., & Lin, L.H. (1990). Comparison of spatial and temporal characteristics of neuronal activity in sequential stages of hippocampal processing. *Progress in Brain Research, 83*:287-300.

Bjornsson, C.S., Lin, G., Al-Kofahi, Y., Narayanaswamy, A., Smith, K.L., Shain W., & Roysam B. (2008). Associative image analysis: a method for automated quantification of 3D multi-parameter images of brain tissue. *Journal of Neuroscience Methods, 170*(1):165-178.

Bliss, T.V., & Collingridge, G.L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. *Nature, 361*(6407):31-39.

Bottai, D., Guzowski, J.F., Schwarz, M.K., Kang, S.H., Xiao, B., Lanahan, A.,…Seeburg, P.H. (2002). Synpatic activity-induced conversion of intronic to exonic sequence in Homer 1 immediate early gene expression. *Journal of Neuroscience, 22*(1):167-175.

Bliss, T.V. & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology, 232*(2):331:356.

Brakeman, P.R., Lanahan, A., O'Brien, R., Roche, K., Barnes, C.A., Huganir, R., & Worley, P.F. (1997). Homer: a protein that selectively binds metabotropic glutamate receptors. *Nature, 386*(6622):284-288.

Burke, S.N., Chawla, M.K., Penner, M.R., Crowell, B.E., Worley, P.F., Barnes, C.A., & McNaughton, B.L. (2005). Differential encoding of behavior and spatial context in deep and superficial layers of the neocortex. *Neuron, 45*(5):667-674.

Burke, S.N., Maurer, A.P., Yang, Z., Navratilova, Z., & Barnes, C.A. (2008). Glutamate receptor-mediated restoration of experience-dependent place field expansion plasticity in aged rats. *Behavioral Neuroscience, 122*:535-548.

Carter, D.R., Eskiw, C., & Cook, P.R. (2008). Transcription factories. *Biochemical Society Transactions, 36*(4):585-589. doi:10.1042/BST0360585.

Chawla, M.K., Guzowski, J.F., Ramirez-Amaya, V., Lipa, P., Hoffman, K.L., Marriott, L.K.,… Barnes, C.A. (2005). Sparse, environmentally selective expression of *Arc* RNA in the upper blade of the rodent fascia dentata by brief spatial experience. *Hippocampus, 15*(5):579-86.

Chawla, M.K., Sutherland, V.L., Olson, K., McNaughton, B.L., & Barnes, C.A. (2018). Behavior-driven *arc* expression is reduced in all ventral hippocampal subfields compared to CA1, CA3, and dentate gyrus in rat dorsal hippocampus. *Hippocampus, 28*(2):178-185.

Cho, K.O., Hunt, C.A., & Kennedy, M.B. (1992). The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumor suppresor protein. *Neuron, 9*:929-942.

Cho, W.K., Jayanth, N., English, B.P., Inoue, T., Andrews, J.O., Conway, W.,…Cisse, I.I. (2016). RNA polymerase II cluster dynamics predict mRNA output in living cells. *eLife, 5*:e13617. doi: 10.7554/eLife.13617.

Chuderland, D., & Seger, R. (2008). Calcium regulates ERK signaling by modulating its protein-protein interactions. *Communicative & Integrative Biology, 1(1)*: 4-5.

Chung, G., & Kim, S.J. (2017). Sustained activity of metabotropic glutamate receptor: Homer, Arrestin, and Beyond. *Neural Plasticity, 2017*:5125624. doi:10.1155/5125624.

Cisse, I.I., Izeddin, I., Causse, S.Z., Boudarene, L., Senecal, A., Muresan, L.,…Darzacq, X. (2013). Real-time dynamics of RNA polymerase II clustering in live human cells. *Science, 341*(6146):664-667.

Clayton, D.F. (2000). The genomic action potential. *Neurobiology of Learning and Memory, 74*:185-216.

Cole, A.J., Saffen, D.W., Baraban, J.M., & Worley, J.F. (1989). Rapid increase of an immediate early gene messenger RNA in hippocampal neurons by synaptic NMDA receptor activation. *Nature, 340*:474-476.

Du, B., & Zhang, L. (2011). Random-selection-based anomaly detector for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(5):1578-1589.

Dubé, S. (2016). *Mapping behaviourally induced immediate-early gene expression along the septo-temporal axis of CA1* (Master's thesis, University of Lethbridge, Lethbridge, Canada). Retrieved from http://hdl.handle.net/10133/4789.

Dubé, S., Lapointe, V., Trivedi, V., & McNaughton, B.L. (2012, October). *Persistent transcription of immediate-early gene Homer1a in the dentate gyrus after MECS*. Poster session presented at Society for Neuroscience Meeting, New Orleans, LA.

Eckersley-Maslin, M.A. & Spector, D.L. (2014). Random monoallelic expression: regulating
gene expression one allele at a time. *Trends in Genetics, 30*(6): 237-244.

Ekstrom, A., Meltzer, J., McNaughton, B.L., & Barnes, C.A. (2001). NMDA receptor antagonism blocks experience-dependent expansion of hippocampal place fields. *Neuron, 31*(4):631-638.

Fields, R.D., Eshete, F., Stevens, B., & Itoh, K. (1997). Action potential-dependent regulation of gene expression: temporal specificity in Ca^{2+}, cAMP-responsive element binding proteins, and mitogen-activated protein kinase signaling. *Journal of Neuroscience, 17*(19):7252-7266.

Flavell, S.W., Cowan, S.W., Kim, T.K., Greer, P.L., Lin, Y., Paradis, S., …Greenberg, M.E. (2006). Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number. *Science, 311*(5763):1008-1012.

Flavell, S.W., & Greenberg, M.E. (2008). Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. *Annual Reviews Neuroscience, 31*:563-590.

Franks, B., Higgins, E.T., & Champagne, F.A. (2014). A theoretically based model of rat personality with implications for welfare. *PLOS ONE 9*(4): e95135. doi:10.1371/journal.pone.0095135.

Garner, A.R., Rowland, D.C., Hwang, S.Y., Baumgaertel, K., Roth, B.L., Kentros, C., & Mayford, M. (2012). Generation of a synthetic memory trace. *Science, 335*(6075):1513-1516.

Gertler, F.B., Niebuhr, K., Reinhard, M., Wehland, J., & Soriano, P. (1996). Mena, a relative of VASP and Drosophila *Enabled*, is implicated in the control of microfilament dynamics. *Cell, 87*:227-239.

Ghamari, A., van de Corput, M.P., Thongjuea, S., van Cappellen, W.A., van Ijcken, W., van Haren, J., …Grosveld, F.G. (2013). In vivo live imaging of RNA polymerase II transcription factories in primary cells. *Genes & Development, 27*(7):767-777.

Grande, M.A., van der Kraan, I., de Jong, L., & van Driel, R. (1997). Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. *Journal of Cell Science, 110*(15):1781-1791.

Greenberg, M., Thompson, M., & Sheng, M. (1992). Calcium regulation of immediate early gene transcription. *Journal of Physiology-Paris, 86*(1-3):99-108.

Guzowski, J.F. (2002). Insights into immediate-early gene function in hippocampal memory consolidation using antisense oligonucleotide and fluorescent imaging approaches. *Hippocampus, 12*:86-104.

Guzowski, J.F., Lyford, G.L., Stevenson, G.D., Houston, F.P., McGaugh, J.L., Worley, P.F., & Barnes C.A. (2000). Inhibition of activity-dependent arc protein expression in the rat hippocampus impairs the maintenance of long-term potentiation and the consolidation of long-term memory. *Journal of Neuroscience, 20*(11):3993-4001.

Guzowski, J.F., McNaughton, B.L., Barnes, C.A., & Worley, P.F. (1999). Environment-specific expression of the immediate-early gene *Arc* in hippocampal neuronal ensembles. *Nature Neuroscience, 2*(12):1120-1124.

Guzowski, J.F., McNaughton, B.L., Barnes, C.A., & Worley, P.F. (2001). Imaging neural activity with temporal and cellular resolution using FISH. *Current Opinion in Neurobiology, 11*(5):579-584.

Guzowski, J.F., Miyashita, T., Chawla, M.K., Sanderson, J., Maes, L.I., Houston, F.P.,…Barnes, C.A. (2006). Recent behavioral history modifies coupling between cell activity and *Arc* gene transcription in hippocampal CA1 neurons. *Proceedings of the National Academy of Sciences, 103*(4):1077-1082.

Guzowski, J.F., Timlin, J.A., Roysam, B., McNaughton, B.L., Worley, P.F., & Barnes, C.A. (2005). Mapping behaviorally relevant neural circuits with immediate-early gene expression. *Current Opinion in Neurobiology, 15*(5):599-606.

Hardingham, G.E., Chawla, S., Johnson, C.M., & Bading, H. (1997). Distinct functions of nuclear and cytoplasmic calcium in the control of gene expression. *Nature, 385*(6613):260-265.

Iborra, F.J., Pombo, A., Jackson, D.A., & Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription “factories' in human nuclei. *Journal of Cell Science, 109*(6):1427-1436.

Imamura, N., Nonaka, A., Yamamoto, H., Matsuki, N., & Nomura, H. (2011). Experience-dependent *Homer1a* expression in excitatory and inhibitory neurons. *NeuroReport, 22*(7):353-357.

Inoue, N., Nakao, H., Migishima, R., Hino, T., Matsui, M., Hayashi, F.,… Inokuchi, K. (2009). Requirement of the immediate early gene vesl-1S/homer-1a for fear memory formation. *Molecular Brain, 2*(1):7. doi:10.1186/1756-6606-2-7.

Jackson, D.A., Hassan, A.B., Errington, R.J., & Cook, P.R. (1993). Visualization of focal sites of transcription within human nuclei. *The EMBO Journal, 12*(3):1059-1065.

Jackson, D.A., Iborra, F.J., Manders, E.M., & Cook, P.R. (1998). Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. *Molecular Biology of the Cell, 9*:1523-1536.

Kandel, E.R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science, 294*(5544):1030-1038.

Kapuscinski, J. (1995). DAPI: a DNA-specific fluorescent probe. *Biotechnic & Histochemistry, 70*(5), 220-233.

Kato, A., Ozawa, F., Saitoh Y., Hirai, K., & Inokuchi, K. (1997). vesl, a gene encoding VASP/Ena family related protein, is upregulated during seizure, long-term potentiation, and synaptogenesis. *Federation of Biochemical Societies Letters, 412*:183-189.

Kato, A., Ozawa, F., Saitoh, Y., Fukazawa, Y., Sugiyama, H., & Inokuchi, K. (1998). Novel members of the Vesl/Homer family of PDZ proteins that bind metabotropic glutamate receptors. *Journal of Biological Chemistry, 273*(37):23969-23975.

Kelly, M.P., & Deadwyler, S.A. (2003). Experience-dependent regulation of the immediate-early gene *arc* differs across brain regions. *Journal of Neuroscience, 23*(16):6443-6451.

Kennedy, M.B. (1995). Origin of PDZ (DHR, GLGF) domains. *Trends in Biochemical Sciences, 20*(9):350.

Krishan, A. & Dandekar, P.D. (2005). DAPI fluorescence in nuclei isolated from tumors. *Journal of Histochemistry and Cytochemistry, 53(8)*:1033-1036.

Kubik, S., Miyashita, T., & Guzowski, J.F. (2007). Using immediate-early genes to map hippocampal subregional functions. *Learning & Memory, 14*(11):758-70.

Lanahan, A., & Worley, P.F. (1998). Immediate-early genes and synaptic function. *Neurobiology of Learning and Memory, 70*:37-43.

Lee, H-J., & Zheng, J.J. (2010). PDZ domains and their binding partners: structure, specificity, and modification. *Cell Communication and Signaling, 8*:8. doi:10.1186/1478-811X-8-8.

Lee, I., Yoganarasimha, D., Rao, G., & Knierim, J.J. (2004). Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature, 430*(6998):456-459.

Lin, Y., Bloodgood, B.L., Hauser, J.L., Lapan, A.D., Koon, A.C., Kim, T.K., …Greenberg, M.E. (2008). Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature, 455*(7217):1198-1204.

Link, W., Konietzko, U., Kauselmann, G., Krug, M., Schwanke, B., Frey, U., & Kuhl, D. (1995). Somatodendritic expression of an immediate early gene is regulated by synaptic activity. *Proceedings of the National Academy of Sciences of the United States of America, 92*(12):5734-5738.

Lis, J. (1998). Promoter-associated pausing in promoter architecture and postinitiation transcriptional regulation. *Cold Spring Harbor Symposia on Quantitative Biology, 63*:347-356.

Lyford, G.L., Yamagata, K., Kaufmann, W.E., Barnes, C.A., Sanders, L.K., Copeland, N.G., …Worley, P.F. (1995). Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron, 14*(2):433-445.

Manzini, G., Xodo, L., Barcellona, M.L., & Quadrifoglio, F. (1985). Interaction of 4'6-diamidino-2-phenylindole 2HCl with synthetic and natural deoxy-and ribonucleic acids. *Journal of Biosciences, 8*:699-711

Marrone, D.F., Schaner, M.J., McNaughton, B.L., Worley, P.F., & Barnes, C.A. (2008). Immediate-early gene expression at rest recapitulates recent experience. *Journal of Neuroscience, 28*(5):1030-1033.

Martin, S., & Pombo, A. (2003). Transcription factories: quantitative studies of nanostructures in the mammalian nucleus. *Chromosome Research, 11*(5):461-470.

Maurer, A.P., Cowen, S.L., Burke, S.N., Barnes, C.A., & McNaughton, B.L. (2006). Organization of hippocampal cell assemblies based on theta phase precession. *Hippocampus, 16*(9):785-794.

McNaughton, B.L. (1982). Long-term synaptic enhancement and short-term potentiation in rat fascia dentata act through different mechanisms. *Journal of Physiology 324*(1):249-262.

Mesina, L., Wilber, A.A., Clark, B.J., Dubé, S., Demecha, A.J., Stark, C.E., & McNaughton, B.L. (2016). A methodological pipeline for serial-section imaging and tissue realignment for whole-brain functional and connectivity assessment. *Journal of Neuroscience Methods, 266:*151-160.

Minatohara, K., Akiyoshi, M., & Okuno, H. (2016). Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. *Frontiers in Molecular Neuroscience, 8*:78. doi:10.3389/fnmol.2015.00078.

Miranti, C.K., Ginty, D.D., Huang, G., Chatila, T., & Greenberg, M.E. (1995). Calcium activates serum response factor-dependent transcription by a Ras and Elk-1 independent mechanism that involves a Ca^{2+}/Calmodulin-dependent kinase. *Molecular and Cellular Biology*, *15*:3672-3684.

Mitchell, J.A., & Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes & Development, 22*(1):20-25.

Miyashita, T., Kubik, S., Haghighi, N., Steward, O., & Guzowski, J.F. (2009). Rapid activation of plasticity-associated gene transcription in hippocampal neurons provides a mechanism for encoding of one-trial experience. *Journal of Neuroscience*, *29*(4):898-906.

Mizuseki, K., & Buzsáki, G. (2013). Preconfigured, skewed distribution of firing rates in the hippocampus and entorhinal cortex. *Cell Reports, 4*(5):1010-1021.

Montes-Rodriguez, C.J., Lapointe, V., Trivedi, V., Lu, Q., Demchuk, A.M., & McNaughton, B.L. (2013). Postnatal development of *Homer1a* in the rat hippocampus. *Hippocampus 23*:890-902.

Morgan, J.I., Cohen, D.R., Hempstead, J.L., & Curran, T. (1987). Mapping patterns of *c-fos* expression in the central nervous system after seizure. *Science 237*(4811):192-197.

Morgan, J.I., & Curran, T. (1989). Stimulus-transcription coupling in neurons: role of cellular immediate-early genes. *Trends in Neurosciences, 12*(11):459-462.

Na, Y., Park, S., Lee, C., Kim, D.K., Park, J.M., Sockanathan, S., …& Worley, P.F. (2016). Real-time imaging reveals properties of glutamate-induced Arc/Arg 3.1 translation in neuronal dendrites. *Neuron, 91*(3):561-573.

Naisbitt, S., Kim, E., Tu, J.C., Xiao, B., Sala, C., Valtschanoff, J.,…Sheng, M. (1999). Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. *Neuron, 23*(3):569-582.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely moving rat. *Brain Research, 34*:17175.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E.,… Reik, W. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature Genetics, 36*(10):1065-1071.

Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J.,… Fraser, P. 2007. *Myc* dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biology, 5*(8):1763-1772. doi:10.1371/journal.pbio.0050192.

Pawley J.B. (2006). Fundamental limits in confocal microscopy in J.B. Pawley (Ed.) *Handbook of Biological Confocal Microscopy* (pp. 20-42). Boston, MA: Springer. doi:10.1007/978-0-387-45524-2.

Paxinos G., & Watson C. (2007). *The rat brain in stereotaxic coordinates*. Cambridge, MA: Academic Press.

Penner M., Roth T., Chawla M., Hoang L., Roth E., Lubin F., ..., Barnes C.A. (2011). Age-related changes in *Arc* transcription and DNA methylation within the hippocampus. *Neurobiology of Aging, 32*(12):2198-2210.

Peters, D.C. (1979). A comparison of mercury arc lamp and laser illumination for flow cytometers. *Journal of Histochemistry and Cytochemistry, 27*:241-245.

Ponting, C.P., & Phillips, C. (1997). Identification of homer as a homologue of the Wiskott-Aldrich syndrome protein suggests a receptor-binding function for WH1 domains. *Journal of Molecular Medicine, 75*:769-771.

Potschka, H., Krupp, E., Ebert, U., Gumbel, C., Leichtlein, C., Lorch, B.,…Hiemisch, H. (2002). Kindling-induced overexpression of Homer 1A and its functional implications for epileptogenesis. *European Journal of Neuroscience, 16*(11):2157-2165.

Raj, A., Van Den Bogaard, P., Rifkin, S.A., Van Oudenaarden, A., & Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods, 5*(10):877-879.

Ramanan, N., Shen, Y., Sarsfield, S., Lemberger, T., Schutz, G., Linden, D.J., & Ginty, D.D. (2005). SRF mediates activity-induced gene expression and synaptic plasticity but not neuronal viability. *Nature Neuroscience, 8*(6):759-767.

Ramirez, S., Liu, X., Lin, P.-A., Suh, J., Pignatelli, M., Redondo, R.L.,… Tonegawa, S. (2013). Creating a false memory in the hippocampus. *Science, 341*(6144):387-391.

Rieder, D., Trajanoski, Z., & McNally, J. (2012). Transcription factories. *Frontiers in Genetics 3*:221. doi:10.3389/fgene.2012.00221.

Rodriguez-Tornos, F.M., Aniceto, I.S., Cubelos, B., & Nieto, M. (2013). Enrichment of conserved synaptic-activity responsive element in neuronal genes predicts a coodrinated response of MEF2, CREB and SRF. *PLOS ONE, 8*(1):e53848. doi:10.1371/journal.pone.0053848

Roysam, B., Shain, W., Barnes, C.A., Mohler, W., Lin, G., Bjornsson, C.,…Narayanaswamy, A. (2008). FARSIGHT: A divide and conquer methodology for analyzing complex and dynamic biological microenvironments. In J. Rittscher *et al.* (Eds.), *Microscropic image analysis for life science applications* (pp.115-151). Norwood, MA: Artech House.

Saha, R.N., Wissink, E.M., Bailey, E.R., Zhao, M., Fargo, D.C., Hwang, J.Y.,… Dudek, S.M. (2011). Rapid activity-induced transcription of *Arc* and other IEGs relies on poised RNA polymerase II. *Nature Neuroscience, 14*(7):848-56.

Saha, R.N., & Dudek, S.M. (2013). Splitting hares and tortoises: a classification of neuronal immediate early gene transcription based on poised RNA polymerase II. *Neuroscience, 247*:175-181.

Sato, M., Suzuki, K., & Nakanishi, S. (2001). NMDA receptor stimulation and brain-derived neurotrophic factor upregulate *homer 1a* mRNA via the mitogen-activated protein kinase cascade in cultured cerebellar granule cells. *Journal of Neuroscience, 21*(11):3797-3805.

Shen, J., Barnes, C.A., McNaughton, B.L., Skaggs, W.E., & Weaver, K.L. (1997). The effect of aging on experience-dependent plasticity of hippocampl place cells. *Journal of Neuroscience, 17*:6769-6782.

Sheng, M. (2001). The postsynaptic NMDA-receptor-PSD-95 signaling complex in excitatory synapses in the brain. *Journal of Cell Science, 114*:1251.

Sheng, M., & Greenberg, M.E. (1990). The regulation and function of *c-fos* and other immediate early genes in the nervous system. *Neuron, 4*(4):477-485.

Sheng, M. & Hoogenraad, C.C. (2007). The postsynaptic architecture of excitatory synapses: a more quantitative view. *Annual Reviews in Biochemistry*, 76:823-847.

Sheng, M., & Kim, E. (2000). The Shank family of scaffold proteins. *Journal of Cell Science. 113*(11):1851-1856.

Sheng, M., McFadden, G., & Greenberg, M.E. (1990). Membrane depolarization and calcium induce *c-fos* transcription via phosphorylation of transcription factor CREB. *Neuron, 4*(4):571-582.

Sun, X., & Lin, Y. (2016). Npas4: Linking neuronal activity to memory. *Trends in Neurosciences, 39(*4):264-275.

Sutherland, H., & Bickmore, W.A. (2009). Transcription factories: gene expression in unions? *Nature Reviews Genetics, 10*(7):457-466.

Tanious, F.A., Veal, J.M., Buczak, H., Ratmeyer, L.S., & Wilson, W.D. (1992). DAPI (4',6-diamidino-2-phenylindole) binds differently to DNA and RNA: minor-groove binding at AT sites and intercalation at AU sites. *Biochemistry, 31*(12):3103-3112.

Teyler, T.J., & DiScenn, P. (1987). Long-term potentiation. *Annual Review ofg Neuroscience, 10*:131-161.

Tu, J.C., Xiao, B., Yuan, J.P., Lanahan, A.A., Leoffert, K., Li, M.,…Worley, P.F. (1998.) Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. *Neuron, 21*(4):717-726.

Tully, T. (1997). Regulation of gene expression and its role in long-term memory and synaptic plasticity. *Proceedings of the National Academy of Sciences, 94*(9):4239-4241.

Van Nguyen, T., Kobierski, L., Comb, M., & Hyman, S. (1990). The effect of depolarization on expression of the human proenkephalin gene is synergistic with cAMP and dependent upon a cAMP-inducible enhancer. *Journal of Neuroscience, 10*(8):2825-2833.

Vazdarjanova, A., & Guzowski, J.F. (2004). Differences in hippocampal neuronal population responses to modifications of an environmental context: evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles. *Journal of Neuroscience, 24*(29):6489-6496.

Vazdarjanova, A., McNaughton, B.L., Barnes, C.A., Worley, P.F., & Guzowski, J.F. (2002). Experience-dependent coincident expression of the effector immediate-early genes *Arc* and *Homer1a* in hippocampal and neocortical neuronal networks. *Journal of Neuroscience, 22*:10067-10071.

Vazdarjanova, A., Ramirez-Amaya, V., Insel, N., Plummer, T.K., Rosi, S., Chowdhury, S.,… Barnes C.A. (2006). Spatial exploration induces ARC, a plasticity-related immediate-early gene, only in calcium/calmodulin-dependent protein kinase II-

positive principal excitatory and inhibitory neurons of the rat forebrain. *Journal of Comparative Neurology, 498*(3):317-329.

Wang, Y., Zheng, F., Zhou, X., Sun, Z., & Wang, H. (2009). Converging signal on ERK1/2 activity regulates group 1 mGluR-mediated Arc transcription. *Neuroscience Letters, 460*(1):36-40.

Wang, Q., Chikina, M.D., Pincas, H., & Sealfon, S.C. (2014). Homer1 alternative splicing is regulated by gonadotropin-releasing hormone and modulates gonadotropin gene expression. *Molecular Cell Biology, 34*(10):1747-1756.

Wansink, D.G., Schul, W., Van Der Kraan, I., Van Steensel, B., Van Driel, R., & De Jong L. (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *Journal of Cell Biology, 122*:283-283.

Wiegert, J.S. & Bading, H. (2011). Activity-dependent calcium signaling and ERK-MAP kinases in neurons: a link to structural plasticity of the nucleus and gene transcription regulation. *Cell Calcium, 49*(5):296-305.

Wilson, M.A., & McNaughton, B.L. (1993). Dyanmics of the hippocampal ensemble code for space. *Science, 261*(5124):1055-1058.

Witharana, W.K.L. (2011). *Non-Boolean characterization of Homer1a intranuclear transcription foci* (Master's thesis, University of Lethbridge, Lethbridge, Canada). Retrieved from http://hdl.handle.net/10133/3402.

Witharana, W.K.L., Cardiff, J., Chawla, M.K., Xie, J.Y., Alme, C.B., Eckert, M., …McNaughton, B.L. (2016). Nonuniform allocation of hippocampal neurons to place fields across all hippocampal subfields. *Hippocampus, 26*(10):1328-1344.

Witharana, W.K.L., Clark, B.J., Trivedi, V., Mesina, L., & McNaughton, B.L. (2018). Immediate-early gene *Homer1a* intranuclear transcription focus as a measure of relative neural activation. *Hippocampus*. doi:10.1002/hipo.23036

Wood, L.S., Desjardins, J.K., & Fernald, R.D. (2012). Effects of stress and motivation on performing a spatial task. *Neurobiology of Learning and Memory, 95*(3):277-85.

Xia, Z., Dudek, H., Miranti, C.K., & Greenberg, M.E. (1996). Calcium influx via the NMDA receptor induces immediate early gene transcription by a MAP kinase/ERK-dependent mechanism. *Journal of Neuroscience, 16*(17):5425-5436.

Xiao, B., Tu, J.C., Petralia, R.S., Yuan, J.P., Doan, A. Breder, C.D.,…Worley, P.F. (1998). Homer regulates the association of group 1 metabotropic glutamate receptors with multivalent complexes of Homer-related, synaptic proteins. *Neuron, 21*:707-716.

Xiao, B., Tu, J.C., & Worley, P.F. (2000). Homer: a link between neural activity and glutamate receptor function. *Current Opinion in Neurobiology, 10*(3):370-374.

Yamamoto, K., Sakagami, Y., Sugiura, S., Inokuchi, K., Shimohama, S., & Kato, N. (2005). Homer 1a enhances spike-induced calcium influx via L-type calcium channels in neocortex pyramidal cells. *European Journal of Neuroscience, 22*(6):1338-1348.

Zhao, Z.W., Roy, R., Gebhardt, J.C., Suter, D.M., Chapman. A.R., & Xie, X.S. (2014). Spatial organization of RNA polymerase II inside a mammalian cell nucleus revealed by reflected light-sheet superresolution microscopy. *Proceedings of the National Academy of Sciences of the United States of America, 111*(2):681-686.

6 APPENDIX A: Estimation of proportion of activated neurons in NanoZoomer images (pixel-based quantification)

Based on many neurophysiological studies, the estimated activation proportion was not expected to increase or decrease dramatically over laps and in our data there was indeed no systematic trend. Although many neurophysiological studies have shown that the number of active hippocampal pyramidal cells remains relatively constant across repeated laps on track mazes such as the one used here, we felt it necessary to confirm this result for the present experiment, since the dependent measure was IEG INF intensity and not spikes *per se*. Unfortunately, automated segmentation of neuronal nuclei was not possible from single plane NanoZoomer images. To assess stability of the active fraction of cells, it would have been sufficient to estimate the number of INFs per unit area of the pyramidal cell layer ROIs; however, to bring the results more into alignment with physiology, we adopted the following method for estimating neuronal numbers, which relies on a pixel-based cell count method. To determine the neuronal pixel intensity range, and also to reduce glial contamination, a sampling exercise was conducted. Blind counters (blind to animal or test group) manually segmented and classified 347 neuronal nuclei and 34 glial nuclei (i.e., the glial fraction from the stratum pyramidale ROIs was approximately 0.10) from randomly selected CA1 and CA3 sub-images. We then computed the distributions of blue (DAPI-stained) pixel intensities within nuclei of either cell class (Figure A1). We also estimated from these manually segmented nuclei the mean nuclear image area in pixels (3013.58). Based on these sample distributions of pixel intensities, we then set upper and lower boundaries for included blue pixels within each ROI (blue{15,85}). This would eliminate ~57% of glial pixels and ~6% of neuronal pixels, resulting in a glial contamination of less than 5%. No further correction was made. We then estimated the number of neurons included by dividing the number of included pixels by the mean nuclear image area (in number of pixels). The proportion of activated neurons was estimated by dividing the INF count by 2 (assuming 2 foci per activated neuron) and then dividing the result by the estimated number of neurons. The resulting percent of *H1a+* neurons (Table 3.1) is somewhat lower than previously published neurophysiological and IEG activation data (e.g., Maurer et al. 2006; Guzowski et al., 1999), possibly as a consequence of assuming 2 foci per activated neuron (there may be mono-allelic expression and often only one of 2 foci appear in a given single plane image).

This method of approximation, however, is subject to two sources of potential error (see methods). Presumably though, this error would apply to all groups analyzed. The purpose of the analysis was to determine whether the number of activated neurons changed across lap numbers. Since it did not, it is of no consequence to the overall conclusion whether there was a relatively small error in the estimates of absolute numbers of neurons.

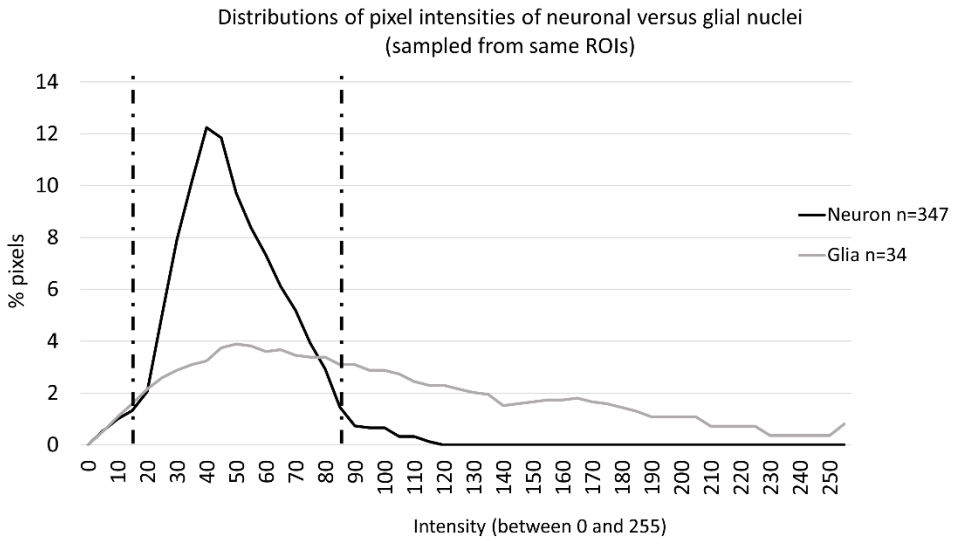


Figure A1. Distributions of pixel intensities in neuronal versus glial nuclei in the blue channel (corresponding to the DAPI counterstain). Based on the distributions, a threshold was set at minimum of 15 and maximum of 85, which accounted for 94% of the neuronal pixels, but eliminated 57% of glial pixels sampled, leaving a potential glia contamination of < 5%. Figure also published in Witharana *et al.*, 2018 (Figure 2).

7 APPENDIX B: Blue/green bleed-through correction for NanoZoomer images

As described in section 2.7.1, prior to INF detection and characterization, uncompressed NanoZoomer TIFF images were pre-processed to minimize blue/green channel bleed-through correction since the emission spectra of fluorescein (green, peak at 520 nm, PerkinElmer, Waltham, MA) and DAPI (blue, peak at 461 nm, Sigma-Aldrich, St. Louis, MI) overlap (Figure B1).

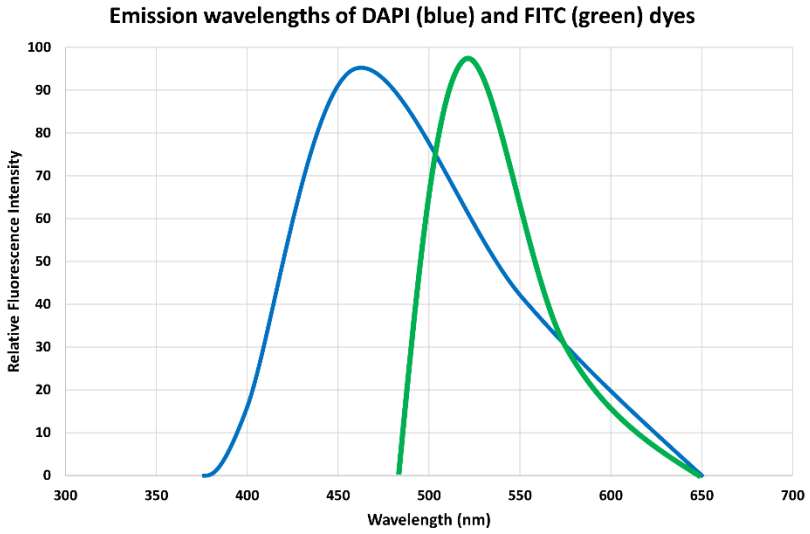


Figure B1. Emission spectra of DAPI (blue) and FITC (green) overlap so the emitted light requires post-acquisition correction to eliminate as much signal bleed-through as possible.

The custom ImageJ software was written with a special bleed-through correction algorithm based on calculated subtraction of cross-over intensity values. First, the program plotted blue intensities versus green intensities for a single image. For each interval of blue intensities, the program plotted the green intensity versus the minimum blue intensity (brightness values of 0 to 255) in each interval (Figure B2). A linear regression was then performed on these values and the regression line was subtracted from the green intensity values. After this subtraction, a new filtered image was created and this image was then run through the INF-analysis step of the analysis.

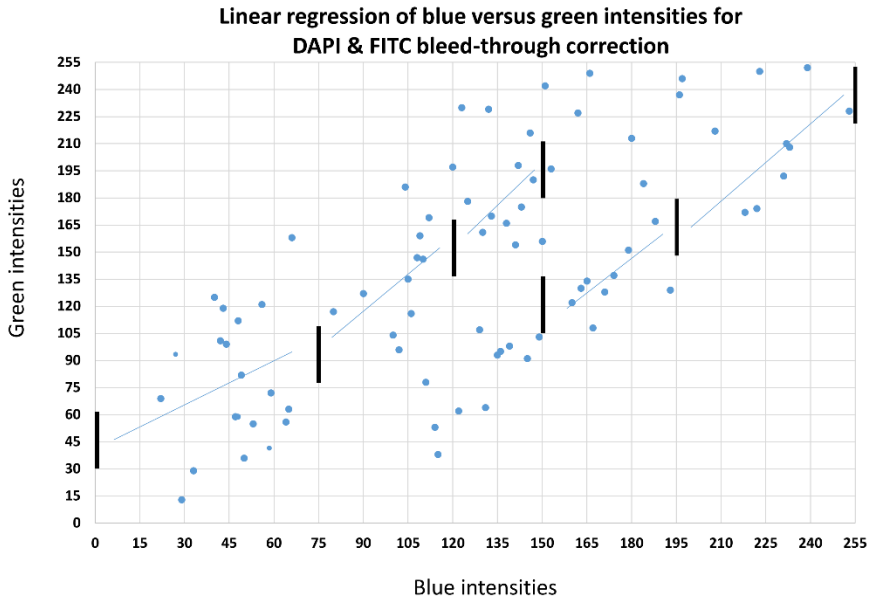


Figure B2. Scatterplot of green versus blue intensities of pixels from a NanoZoomer image used to compute linear regressions that were subtracted from each interval as bleed-through correction of blue and green signal. Custom software program generated intervals of blue intensities and plotted the green versus blue intensities for each pixel on the image. For each interval of blue, a linear regression was computed between the green and blue relationship. This linear regression line was then subtracted from the green intensity as a correction to eliminate blue signal of DAPI emission overlapping into the true FITC (green) signal.