MOLECULAR BIOMARKERS AND NETWORK ANALYSES OF NEURON-TYPE CONNECTIVITY REVEAL CLOAKED INSIGHTS INTO HIPPOCAMPAL FUNCTION

by

Christopher Loren Rees
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Bioinformatics and Computational Biology

Committee:

__________________________________  Dr. Giorgio A. Ascoli, Dissertation Director
__________________________________  Dr. Iosif I. Vaisman, Committee Chair
__________________________________  Dr. Dmitri K. Klimov, Committee Member
__________________________________  Dr. Iosif I. Vaisman, Acting Director,
  School of Systems Biology
__________________________________  Dr. Donna M. Fox, Associate Dean, Office
  of Student Affairs & Special Programs,
  College of Science
__________________________________  Dr. Peggy Agouris, Dean, College of Science

Date: ____________________________  Fall Semester 2016
George Mason University
Fairfax, VA
Molecular Biomarkers and Network Analyses of Neuron-type Connectivity Reveal Cloaked Insights into Hippocampal Function

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Christopher Loren Rees
Master of Science
The George Washington University, 2009
Bachelor of Science
Virginia Polytechnic Institute and State University, 2005

Director: Giorgio A. Ascoli, Professor
Department of Bioengineering and Neuroscience Program

Fall Semester 2016
George Mason University
Fairfax, VA
DEDICATION

To my amazing wife Rachel, and to all those people in my life who have made it past the title page and are still reading. Your encouragement, support, prayers, and love have made this possible.
ACKNOWLEDGEMENTS

I owe many thanks. This dissertation was many years in the making, and, at times, I know it felt like a lifetime.

First to Rachel, for your patience with me through this process and in all things. It is with great joy that I report to you that, after 7+ years, I am finally finished. Our journey together led us across state and national boundaries; we were often unsure of the next step but never the destination, and always hand-in-hand. Next to you is where I belong: it’s where I’m most comfortable and most happy. You sustain me and surprise me, and you make my life fun. I “loves” you, and I look forward to our next adventure together. (I promise: no more school.)

To my Mom and Dad: thank you both for your unwavering help, encouragement, belief, pride, and love. You have always freely and joyfully given your time and expertise, financial support, delicious meals, and daily prayers on my behalf. I learned priorities (and so much else) from your example. You were my first and will always be my favorite teachers. Without your help, I would not have survived 5th grade, high school, or college. Without your love, I would not have thrived.

Thanks and love in great heaps are also due to my mother-in-law, “Mom,” for your love and for always keeping our freezer stocked to bursting with great food. And thank you for a tireless stream of emails of cheer and inspiration. But most of all, thank you (and George) for Rachel.

To Giorgio, my advisor and mentor: I am deeply indebted to you for all that you have done for me throughout my time as a graduate student and in launching my career as a published researcher. You have taught me not just about the brain and research methodology, but you have shown me, through your example, what it means to manage diverse personalities, treat everyone with respect, work honorably and ethically, relish learning, and show up every day with a smile. For this I am forever grateful. And to the rest of the Hippocampome team, including Diek, Charise, Alexander, David, Keivan, Siva, and Sean: thank you for making work fun. I imagine I am one of the few Ph.D. students who truly enjoyed his project.

As for my brother, confidant, best friend, and FIFA teammate, Matthew: though our gaming skills leave something to be desired (someday we’ll win Division 1!), talking to you and playing soccer is a highlight of every week. You are a person of great
intelligence, principle, commitment, joy, celebration, and loyalty. I hope that we can collaborate in research someday. And eternal thanks to my friend and sister-in-law, Laura, who (like Rachel) has the unenviable task of putting up with not just one but two Rees brothers on soccer nights.

I am forever grateful to my grandparents for your unending generosity, selflessness, prayerfulness, and for the lifetime of memories you made possible. The financial and spiritual investments you have made in me, and each of your grandchildren, are without parallel. I wish that Grandpa Rees and Gima could be here today, but I look forward to seeing them again someday – there will be much to celebrate! I truly believe our extended family is exceptional in its love, care, and genuine interest in the success of each other. Thank you to Uncle Tim, Aunt Grace, April, Alyssa, Michael, and (little) Rachel, and to Uncle Jeff, Aunt Lisa, and Elijah Bleu for your prayers, love, and inexplicable curiosity (and great questions) about my research!

Special appreciation also goes out to the nighttime cleaning staff at Mason and the employees of Chick-fil-A, who saw a lot of me during my late nights and early mornings writing. Thank you for being quick with a smile, a hello, or a much-needed frosted coffee refill.

Thank you also to the handsome, talented kitty, Spock, that I am lucky enough to share with Rachel. Your love and happiness at seeing me walk through the door is unconditional and uplifting.

But above all, thank you to my God and Savior for the gifts listed here and so many others. May I always seek to serve You, and may I find in You my hope and my future.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>List of Equations</td>
<td>xi</td>
</tr>
<tr>
<td>Abstract</td>
<td>xii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Conceptualizing and building the Hippocampome knowledge base</td>
<td>2</td>
</tr>
<tr>
<td>Dissertation outline</td>
<td>6</td>
</tr>
<tr>
<td>Chapter One</td>
<td>11</td>
</tr>
<tr>
<td>Abstract</td>
<td>11</td>
</tr>
<tr>
<td>Introduction</td>
<td>13</td>
</tr>
<tr>
<td>Types of biomarkers in the brain</td>
<td>15</td>
</tr>
<tr>
<td>Names, synonyms, and identifying features of major biomarkers</td>
<td>20</td>
</tr>
<tr>
<td>Membrane transporters</td>
<td>22</td>
</tr>
<tr>
<td>Calcium-binding proteins</td>
<td>22</td>
</tr>
<tr>
<td>Neuropeptides</td>
<td>23</td>
</tr>
<tr>
<td>Neurotransmitter receptors</td>
<td>24</td>
</tr>
<tr>
<td>Extracellular matrix proteins</td>
<td>26</td>
</tr>
<tr>
<td>Functional roles of biomarkers</td>
<td>26</td>
</tr>
<tr>
<td>Membrane transporters</td>
<td>26</td>
</tr>
<tr>
<td>Calcium-binding proteins</td>
<td>31</td>
</tr>
<tr>
<td>Neuropeptides</td>
<td>34</td>
</tr>
<tr>
<td>Neurotransmitter receptors</td>
<td>38</td>
</tr>
<tr>
<td>Extracellular matrix proteins</td>
<td>43</td>
</tr>
<tr>
<td>Pharmacological implications</td>
<td>46</td>
</tr>
<tr>
<td>Relationship between biomarkers and neuron types</td>
<td>48</td>
</tr>
<tr>
<td>Conclusions</td>
<td>51</td>
</tr>
</tbody>
</table>
Chapter Two......................................................................................................................... 52
Abstract ............................................................................................................................ 52
Big data connectomes and the usage of potential-synapse proxies............................. 53
Peters’ rule: its original intent and current tripartite usage .................................... 54
  Level 1 – assessment of potential connectivity between neuron types .............. 55
  Level 2 – assessment of potential connectivity between individual neurons .... 59
  Level 3 – assessment of synapse numbers between individual neurons .......... 60
Testing the predictive accuracy of Peters’ rule (Level 1) using Hippocampome.org... 60
Reconciling opposing views in the community ......................................................... 63
Chapter Three...................................................................................................................... 65
Abstract ............................................................................................................................ 65
Significance statement.................................................................................................... 66
Introduction ....................................................................................................................... 67
Materials and methods ............................................................................................... 68
  Identification of neuron types .................................................................................. 68
  Culling of known connectivity information ........................................................... 71
  Computation of potential connectivity ................................................................. 71
  A web-accessible resource for connectome visualization .................................... 74
  Graph theory analyses .............................................................................................. 74
Results ............................................................................................................................. 82
  A highly specialized topology .................................................................................. 83
  Significant community structure .......................................................................... 84
  Degree distribution and hubs .................................................................................. 87
  Rich and ultra-rich clubs ......................................................................................... 90
  Robustness to random failures .............................................................................. 91
  Characteristic connectivity superpattern profile ................................................ 94
  Weighted pattern profile and neuron type fingerprint analysis ........................... 96
  Motifs and anti-motifs ............................................................................................ 99
  Pairwise correlations .............................................................................................. 102
  Sensitivity to future additions or subtractions of neuron types ......................... 104
Discussion ...................................................................................................................... 105
Appendix 1 ...................................................................................................................... 113
Appendix 2 ............................................................................................................. 142
References ........................................................................................................... 155
LIST OF TABLES

Table                                      Page
Table 1 Nomenclature and structural features for 15 biomarkers in mice......................... 21
Table 2 Key roles for the selected biomarkers (continues next page)................................. 28
Table 3 Selected drugs that target biomarker receptors....................................................... 47
Table 4 Recent applications of Peters’ rule .............................................................................. 57
Table 5 Meta-analysis of Peters’ rule on the Hippocampome.org knowledge base.............. 62
Table 6 Neuron type glossary .................................................................................................. 70
Table 7 Identification of hubs and anti-hubs by TD............................................................... 89
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Neuron types and web portal functionality</td>
<td>4</td>
</tr>
<tr>
<td>Figure 2</td>
<td>The Hippocampome project and my roles</td>
<td>7</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Overview of biomarker categories and functionality</td>
<td>18</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Peripheral tissue expression of brain biomarkers</td>
<td>27</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Biomarker expression for 43 major neuron types</td>
<td>50</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Three interpretation levels of resolution for Peters’ rule</td>
<td>55</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Increased usage and concomitant confusion of Peters’ rule over time</td>
<td>55</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Potential connectivity of neuron types</td>
<td>72</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Topological comparison of the HC to well-known random network types</td>
<td>84</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Modular structure of the potential hippocampal connectome</td>
<td>85</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Degree distribution analysis to isolate types with unusual connectivity</td>
<td>88</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Nested rich clubs within the HC</td>
<td>91</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Alternate pathways between types afford resilience to the network</td>
<td>93</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Superpatterns and HC usage</td>
<td>95</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Weighted trimers analysis based on E/I neuron type distinction</td>
<td>97</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Utilization of feedforward versus feedback loops in the HC</td>
<td>100</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Motifs and anti-motifs are largely determined by superpattern topology</td>
<td>101</td>
</tr>
</tbody>
</table>
LIST OF EQUATIONS

Equation                                                                 Page
Equation 1 Computation of potential connectivity ...................................... 73
Equation 2 Characteristic path length of the network .................................... 76
Equation 3 Communication cost ......................................................................... 77
ABSTRACT

MOLECULAR BIOMARKERS AND NETWORK ANALYSES OF NEURON-TYPE CONNECTIVITY REVEAL CLOAKED INSIGHTS INTO HIPPOCAMPAL FUNCTION

Christopher Loren Rees, M.S.
George Mason University, 2016
Dissertation Director: Dr. Giorgio A. Ascoli

To launch Hippocampome.org, a framework for classifying neuron types in the rodent hippocampus, we assembled >14,000 pieces of experimental evidence supporting knowledge of approximately 3,700 properties for 122 neuron types. This resource, the first comprehensive machine-readable neuron inventory for a mammalian cortical region, provides an opportunity to analyze the hippocampus in novel ways and garner unique insights into its intricate structure and enigmatic function. Here, I describe my two-pronged attack on the sub-cellular and cellular fronts.

First, although chemical messages are disseminated throughout the hippocampus and beyond, their individual and collective purpose as it relates to overall functionality remains shrouded. Thus, in a systematic review of a convoluted body of literature, I surveyed 15 molecular biomarkers present in the brain as well as peripheral tissues. By
tracing structural information alongside functional, developmental, pathological, and pharmacological roles, I began to characterize the molecular-level blueprint of the brain.

At the cellular level, knowledge of synaptic connectivity among neuron types is sparse. To fill the gaps, I employed Peters’ rule to determine sites of potential connectivity. In testing the principle by comparing these predictions with data for known connections, I uncovered gross inconsistencies in the community as to the interpretation and usage of the rule. This afforded the opportunity to reconcile opposing views and propose a more-nuanced version of Peters’ rule, which, based on the data at hand, is highly accurate, at least in the hippocampus. Emboldened, I combined known connectivity data with the potential connections and assembled a full connectome for the hippocampal formation at the neuron-type level. Using a full complement of graph theory analytic techniques, I investigated the relationship between the connectivity properties and functions of this circuit and unearthed a goldmine of information. Nuggets included, but were not limited to, a highly specialized topology built upon the tri-synaptic loop, a core set of neuron types that manage and control much of the network traffic, an innate resilience to random perturbations, and a set of simple connectivity patterns that combine to produce complex global functionality.
The hippocampus plays a crucial role in learning and memory. In rodents, the functionality is manifested in both spatial and emotional retention and recall. Although the cellular organization of this structure has been extensively studied, including more than 100 years of detailed and well-documented research using traditional anatomical methods (e.g. Golgi staining and light microscopy), understanding of the mechanisms underlying hippocampal functionality remains elusive. Though recent advances in technology and methodology mean that the diversity of neuron types comprising the circuit can now be probed with modern molecular, genetic, anatomical, and physiological approaches, the literature in the field is vast, scattered, and employs inconsistent terminology.

In a project we call the “Hippocampome,” our lab has produced the first comprehensive accounting of all the known component parts of the hippocampus at the neuron-type scale. As a founding, core member of this team, I have helped to relate the key anatomical, molecular, electrical, and synaptic features of these types, and have aided in producing an online web portal for browsing and searching the knowledge base. But beyond merely cataloging information and illuminating gaps in the knowledge, we have afforded the community (and have ourselves capitalized on) an opportunity to analyze and model the hippocampus in novel ways. This entire endeavor makes up the nucleus of
my dissertation and serves as a critical stepping stone toward our long-term goal of understanding how the menagerie of neuron types interrelates to produce sophisticated functions such as spatial maps and memories.

**Conceptualizing and building the Hippocampome knowledge base**

Two post-doctorate fellows and I mounted a massive literature search to identify all known neuron types in the rodent hippocampus. Starting from seminal work by prominent experts in the past 20 years, we expanded the search radially, such that version 1.0 included usable information from 484 sources. This data was gathered, collated, and cataloged based on the following principles.

First, while no two neurons are identical, neurons of the same type (e.g. dentate gyrus granule cells) are much more similar to each other than to neurons of other types (e.g. dentate gyrus basket cells or CA1 pyramidal cells). By comprehending and exploiting the key distinction between the diversity among neuron types and the variability within one, we seized the opportunity to reduce the problem complexity by orders of magnitude; a circuit defined by millions of neurons can be expressed as one comprising hundreds of neuron types.

Secondly, it is currently unfeasible to fully quantify the exact parameters of neurons in the hippocampus. Such a characterization would necessarily entail capturing morphometric properties (e.g. neurite densities, bifurcations, surface area, branching angles), molecular properties (e.g. full gene and protein expression profiles), and electrophysiological properties (e.g. passive parameters and spike responses to a variety
of applied currents), to name a few. However, by identifying the essential features and extracting them from a statistically representative sample, we were able to build a scheme for classification. Namely, we first considered the primary neurotransmitter and distinguished glutamatergic from GABAergic neurons. Next, we formulated a basic morphological characterization of the neurites: in each of the 26 commonly recognized cytoarchitectural layers of the major hippocampal formation regions, we answered, at a binary level (for now), the question of significant presence or absence of the dendrites and axons. In addition, in some cases, thorough and convincing evidence for differentiable targeting, molecular expression, or electrophysiological parameters allowed for the creation of multiple neuron types with a common morphology.

Thirdly, in performing literature searches and assembling the knowledge base, it was necessary to ignore the author-bestowed names of neurons. Unfortunately, we found that the same name is frequently employed in the literature to describe neurons with different morphological patterns. Moreover, in a separate issue, varying emphases of individual researchers tend to result in different names being ascribed to neurons with very similar morphologies. In combination, these factors result in a chaotic body of literature that confounded the assembly and interrelation of information. Applying the philosophies mentioned above, we were able to identify 122 neuron types in the hippocampal formation. The morphological definitions of these types are presented in Figure 1A (types are listed in rows; columns represent the 26 anatomically defined sub-region parcels in the hippocampal formation; filled red boxes with horizontal lines indicate the presence of axons, blue boxes with vertical lines show dendrites, and purple
Figure 1 Neuron types and web portal functionality
boxes with a ‘+’ show both axons and dendrites; black circles show possible soma locations; abbreviations: DG=dentate gyrus, SUB=subiculum, EC=entorhinal cortex, SMo=stratum moleculare (outer), SMi=stratum moleculare (inner), SG=stratum granulosum, H=hilus, SLM=stratum lacunosum moleculare, SR=stratum radiatum, SL=stratum lucidum, SP=stratum pyramidale, SO=stratum oriens, SM=stratum moleculare, PL=polymorphic layer). Upon this framework, for each type, we layered knowledge regarding approximately 3,700 properties based on more than 14,000 pieces of experimental evidence. We initially added available data regarding the expression or lack of expression of 93 molecular biomarkers and values for 10 electrophysiological passive, spike, and other response parameters.

To display and share this information, we built a web portal. I taught myself PHP and SQL and was a major contributor in bringing to fruition what is a vast resource that is now publically available at Hippocampome.org. Figure 1B displays screenshots from the web portal to illustrate browsing (here, for molecular marker information) and searching functionality (in this case, for the subset of neuron types that are positive for cholecystokinin (CCK) with an input resistance < 867.6 MΩ) capabilities. This, and the knowledge behind it was the product of more than five years of work, and it was described and published in eLife (Wheeler et al., 2015; see Appendix 1) to shining reviews. One reviewer expressed a desire to “applaud and thank the authors for undertaking this endeavor,” noting that “it should be helpful in bringing experimental and modeling communities closer together and to shorten literature exploration time.” To that end, we next described and addressed the problem of inconsistent neuron naming and
terminology in a 2016 article, on which I am a co-author, that was accepted in the journal Brain Informatics (Hamilton et al., 2016; see Appendix 2).

Dissertation outline
The knowledge base described above was foundational and stemmed from my collaborations as part of a small team. But Hippocampome.org is ripe for analysis along multiple dimensions. Therefore, individually, I took the lead on several of these aspects, which are summarized briefly below and documented in detail in the chapters that follow. In addition, I assisted in several other facets that resulted in substantial scientific output, which are currently in preparation for publishing. Figure 2 illustrates the scope of the project (Foundation; capital letters), my various roles, and their relationship with this dissertation (dissertation chapters indicated as numerals 1-3).

Chapter 1 concerns biomarkers. Though a profusion of molecules has been assayed for various neurons, little is known about their importance or function. In fact, most expression data only see the light of day in published literature because their detection was undemanding and/or previously reported. Further, the markers, like the neurons themselves, are entrapped in a maze of confusing nomenclature. These issues extend beyond the hippocampus and impact both the brain at large and peripheral tissues. Thus, I engaged in a unique, systematic review of the most studied biomarkers across categories that included membrane transporters, calcium-binding proteins, neuropeptides, receptors, and extracellular matrix proteins. I amassed and delineated information on the structure, function, and developmental and pathological roles of each, which led to a
The Hippocampome project and my roles

Figure 2 The Hippocampome project and my roles

discussion on the potential benefits for pharmacology and neuron type identification. This chapter will be submitted for publication in Current Medicinal Chemistry.

The latter two-thirds of the dissertation relate to the connectivity among neuron types. Since knowledge about synapsing is quite sparse (i.e. detailed information exists for a handful of connections but most are lightly treated or still unidentified), I supplemented known connections by identifying “potential connections” at axon-dendrite
juxtapositions. This principle is known as Peters’ rule, and it has been misapplied just as frequently as it has been applied, leading to widespread misconceptions and debate over its correctness. Thus, the purpose of Chapter 2, which was invited for submission in Trends in Neurosciences (manuscript in preparation), was two-fold. First, I conducted a conceptual analysis of the three distinct levels at which Peters’ rule has been employed in the literature: one, the originally intended level, involves inferring synapsing among neuron types, and the other two, respectively, predict synapsing probabilities and estimate numbers of synapses between individual cells. All three are discrete questions that intermingle in a common pool of results. Secondly, I critically evaluated the accuracy of type-level connectivity predictions derived from Peters’ rule by comparing Hippocampome known and potential connectivity data. This led me to propose a new, more-nuanced version of Peters’ rule that allows for both proximity-based and highly specific affinity for post-synaptic types.

In the final chapter, which was accepted for publication in the open-access journal eNeuro, I combined the known connectivity and potential connectivity data into a single network with ~3,200 connections among the 122 neuron types. This assemblage represents the full type-level connectome of the hippocampus. I then applied a suite of mathematically-based graph theory analytic techniques to the network, which led to a wealth of insights on the organization and communication of the hippocampus. For example, I unearthed a non-random, modular structure within the tangle of cabling; I discerned the neuron types that serve as critical (but vulnerable) hubs of communication, as well as those that are resilient to degenerative change; and I analyzed the motif
building blocks of local circuitry that help to characterize the function of each neuron type.

It is also worth mentioning that I played notable, secondary roles in three additional branches of analysis pertaining to biomarker and electrophysiological data. In the first, we were able to supplement the Hippocampome molecular marker information we derived from the literature with data from Allen Brain Atlas (brain-map.org). By capitalizing on their ~20,000-gene dataset obtained from 27,000 in situ hybridization experiments, and linking them to hippocampal neuron types whenever possible, we augmented our molecular marker information by a factor of 5. This work was invited for submission and is currently under review in the Journal of Pharmaceutical and Biomedical Analysis. Secondly, when molecular knowledge is incomplete for a given neuron type, I helped to develop an inference engine by which we can sometimes make an expression prediction (with calculable confidence) based on the expression of other molecules and those molecules’ known co-expression levels with each other in other types. With this methodology, we approximately double the amount of knowledge available. This boost allows us (1) to perform clustering of the neuron types along the molecular dimension and identify families, and (2) to extract previously hidden pairwise correlations between markers and other properties (e.g. morphology, connectivity, and electrophysiology). This manuscript is in preparation. Thirdly, I developed a tool that was used in the systematic classification of neuron responses to long-lasting supra-threshold depolarizing current injections. Data were assembled from the literature, and I designed and built software to calculate firing pattern parameters (e.g. inter-spike intervals,
adaptation index, relative deviation, action potential half-widths, fast and slow afterhyperpolarization amplitudes, and sag ratio) from digitally recovered spike-train plots. The information was subsequently used to automatically classify the firing patterns and firing pattern elements of hippocampal neuron types. This manuscript, too, is in final preparation. These secondary roles are not discussed further in this dissertation.
CHAPTER ONE

Title: Molecular biomarkers of mammalian neurons: structure, function, and pharmacological relevance

Authors: Christopher L. Rees, Charise White, Giorgio A. Ascoli

Abstract
Knowledge of molecular biomarker expression in neural systems can provide insight to the chemical blueprint of signal processing and transmission, assist in tracking developmental or pathological progressions, and yield key information regarding potential medicinal targets. These markers are particularly relevant in the mammalian brain in light of its unsurpassed cellular diversity. Accordingly, molecular expression profiling is rapidly becoming a major approach to classify neuron types. Despite a profusion of research, however, the biological interactions of common neuronal biomarkers remain incompletely understood. Furthermore, most brain biomarkers are also present in other organs, therefore complicating considerations of their potential pharmacological effects. Here, we survey 15 prominent neural markers from five categories, namely membrane transporters, calcium-binding proteins, neuropeptides, receptors, and extracellular matrix proteins, explaining their relation and relevance to
synaptic communication. For each marker, we summarize fundamental structural features, cellular functionality, distributions within and outside the brain, as well as known drug effectors and mechanisms of action. This essential primer therefore links the cellular complexity of the brain, the chemical properties of key molecular players in neurotransmission, and open biomedical opportunities.
Introduction

Biological markers, or biomarkers, are genes and proteins that are differentially expressed throughout the body. Though such data is far from complete, rudimentary comprehension and exploitation of this remarkable molecular diversity has already opened many promising research avenues. The expression (or lack thereof) of biomarkers can yield information about cellular populations and unlock our understanding of both normal biological processes and pathogenic events. More specifically, biomarkers have been used to (1) simultaneously classify and sub-classify assemblages of cells while garnering informative signatures about their functional mechanisms (Sloviter, 1989; DeFelipe, 1993; Sakai et al., 1995; Marc and Jones, 2002; Hevner et al., 2003; Markram et al., 2004; Somogyi and Klausberger, 2005; Houser, 2007; Karnas et al., 2013), (2) estimate direct and indirect counts of cellular populations (Jinno and Kosaka, 2006; Neddens and Buonanno, 2009), and (3) detect and diagnose abnormal cells (e.g. cancerous tumors) (Walther et al., 2009; Winter et al., 2013; Nalajska, 2014). In the mammalian brain, where the relationships between molecular structures and higher-level cognitive functions, as well as complex diseases, are still not well understood, such insights are especially critical.

Biomarker expression may be detected through various methods that depend on the type of biomarker. For proteins, immunohistochemistry is widely used, a technique in which antibodies designed to interact with certain antigens are introduced into the tissue and protein-binding is detected through fluorescent or color-staining reactions. At the pre-translation level, in situ hybridization is a prevalent method utilizing a labeled complementary DNA or RNA probe to pinpoint specific sequences in the tissue. Myriad
studies that characterize and, sometimes, relate these molecular expressions with morphological, electrophysiological, and synaptic properties in individual neurons or in a slice of brain tissue are assembled in a vast wealth of literature. Recently, large-scale studies (Mikula et al., 2007; Colantuoni et al., 2011) have begun amassing datasets containing mRNA and protein localization information across the entire brain. In a particularly notable repository built from more than 25,000 experiments, the Allen Brain Atlas collected in situ hybridization gene expression data for approximately 20,000 genes in the mouse brain (Lein et al., 2007).

Although molecular signals are continuously passed around the brain and other body systems, and expression data is accumulating, their function is rarely well-characterized. In fact, the literature is rife with expression reports for biomarkers that simply had readily available detection reagents, were easy to detect, and/or were previously cited. Worse, biomarkers are often referred to by numerous variations of names, synonyms, and abbreviations.

Here, to ease the bewilderment, we assemble a systematic review of fifteen of the most studied markers in the brain, restricting our discussion to mice. We begin with an overview of neurotransmitters and the categories of supporting molecules that make viable biomarker candidates. Then, for biomarkers from each category, we investigate the nomenclature, sequences, and structures, before delving into function and pathology, and, finally, pharmacological applications. We close with an overview of the major, known relationships between biomarkers and cell types in the brain. Together, this information on various membrane transporters, calcium-binding proteins (CaBP), neuropeptides,
receptors, and extracellular matrix proteins will serve as a primer for researchers, illuminate gaps in the knowledge, and collate biomarker-pathology linkages that may be useful in drug targeting.

Types of biomarkers in the brain
The primary chemical messengers in the nervous system are neurotransmitters (NT), which transfer signals across synapses from pre-synaptic neurons to post-synaptic neurons, muscles, and other targets. NTs may be classified according to their function (i.e. excitation or inhibition), action (i.e. direct or neuromodulatory), or, more discriminately, molecular structure (e.g. amino acids, monoamines, and peptides, among others).

Glutamate and gamma-aminobutyric acid (GABA), two amino acids, are the most prevalent NTs in the brain. Glutamate, which promotes excitatory effects by increasing the probability that the target cell will fire an action potential, is utilized in more than 90% of synapses (Abeles, 1991; Braitenberg and Schüz, 1998; Attwell and Laughlin, 2001). It has a major role in synaptic plasticity and, thus, is implicated in cognitive functions such learning and memory (McEntee and Crook, 1993). GABA, a non-standard and non-proteinogenic amino acid, is the next most prevalent NT and is inhibitory in most cases.

Though other NTs are less prevalent in terms of overall usage in the brain, they serve critical functions. One category, biogenic monoamines, includes histamine, serotonin, and the catecholamines (i.e. epinephrine, also known as adrenaline, norepinephrine or noradrenaline, and dopamine). Histamine is synthesized from the
amino acid histadine, serotonin is derived from tryptophan, and the catecholamines are obtained from phenylalanine and tyrosine. These biogenic amines, which may be more familiar for their roles in other body systems, are distributed in the brain, where they play a role in regulating emotional behavior. Histamine, known primarily for its part in immune response, is also involved in arousal, perception of anxiety, the stress-related release of hormones from the pituitary, and water retention and the suppression of eating (Brown et al., 2001). Serotonin is synthesized both in the gut, where it controls gastrointestinal motility (Yano et al., 2015), and in the brain to help regulate mood, appetite, and sleep (Jouvet, 1969; Young and Leyton, 2002; Curzon, 1990). The catecholamines act as hormones that are present in the bloodstream; high levels indicate stress and, in particular, epinephrine and norepinephrine are released by the adrenal medulla as part of the fight-or-flight response (Haller et al., 1998). Dopamine acts within the brain to regulate motor control and reward-motivated behavior systems; the latter function is played upon by addictive drugs such as cocaine and amphetamines (Kelley and Berridge, 2002). Non-monoamine NTs include gasotransmitters (e.g. nitric oxide, carbon monoxide, and hydrogen sulfide), purines (e.g. ATP and adenosine), various peptides, and acetylcholine, which has a primary role at neuromuscular junctions (review in Hall and Sanes, 1993).

Naturally, biomarkers have maximal utility when they are differentially expressed across brain regions and neuron types. Thus, while glutamate and GABA are useful in determining the basic excitatory or inhibitory operation of cells, their widespread usage renders them less practical in identification roles. Neuropeptides, a special class of NT
that do not undergo cellular reuptake after secretion, are in many cases more suitable as biomarkers. Examples of neuropeptide biomarkers, including cholecystokinin (CCK), enkephalin (ENK), neuropeptide Y (NPY), somatostatin (SOM), and vasoactive intestinal polypeptide (VIP) will be detailed in subsequent sections. Neurons often produce both a conventional NT and one or more neuropeptides. The peptides are slower-acting than the classical NTs and are involved in supporting a fine-tuning of neuronal signaling. Though they may be packaged together with traditional NTs, peptides are generally enveloped in large, dense-core vesicles found throughout a neuron (i.e. including the soma, dendrites, axonal varicosities), and classical NT in small, clear synaptic vesicles found in clusters at presynaptic locations (Zhang et al., 2011). Release of the large vesicles and the small vesicles is independently regulated.

Indeed, NTs are not the only viable biomarkers in the brain, as molecules that are differentially expressed from cell to cell play many supporting roles leading up to and following NT release (see Figure 3 for an overview; markers reviewed here shown in purple). Prior to the arrival of an electrical action potential at a synapse, vesicular transporters regulate the passage of NTs across the vesicle via active transport coupled with the antiport (i.e. facilitated diffusion) of $H^+$ ions out of the vesicle. These transporter proteins are specific to the NT and include vesicular glutamate transporters 1-3 (i.e. vGluT1, vGluT2, and vGluT3) and vesicular GABA transporter (vGAT). vGluT3 will be described in detail in this report. Subsequently, an action potential depolarizes the cell membrane and leads to an influx of $Ca^{2+}$ ions through calcium channels into the pre-synaptic cell. At this point, CaBP, such as potential biomarkers calbindin (CB), calretinin
(CR), or parvalbumin (PV) (all of which are detailed below) bind to the ions and regulate calcium levels. When high concentrations of intracellular calcium are reached, the vesicles fuse with the synaptic membrane and their contents are expelled into the synaptic cleft via exocytosis.

Figure 3 Overview of biomarker categories and functionality

After crossing the synaptic membrane, NTs, including neuropeptides, seek to bind specific, differently expressed receptors on post-synaptic cells. In general, receptors are one of two types: ionotropic (or ligand-gated) and metabolic (or G protein-coupled)
receptors. The former variety are a group of transmembrane ion channels that open or close in response to the binding of a NT and are usually very selective to certain cations or anions. Conversely, metabotropic receptors are a subtype of neuroreceptors that use signal transduction (i.e. second messenger) mechanisms, often G proteins, to activate a series of intracellular events. Ionotropic receptors tend to mediate fast synaptic transmission, while metabotropic receptors are associated with a more prolonged stimulus due to their cascading nature. Receptors for both glutamate and GABA consist of both types. For glutamate, AMPA, NMDA, and kainate are ionotropic receptors that allow the passage of Na$^{+}$, K$^{+}$, and/or Ca$^{2+}$. Metabotropic glutamate receptors are classified into eight types (mGluR1-8) belonging to three families according to their location, preferred agonists and activators, and function (Ferraguti and Shigemoto, 2006). Group 1, which includes mGluR1 and mGluR5, primarily operate on the post-synaptic side of the cleft, whereas receptors belonging to groups 2 and 3 are primarily pre-synaptic (Shigemoto et al., 1997). The splice variant known as mGluR1a is described below.

GABA receptors are divided into two classes. GABA$_A$ (along with the GABA$_A$-$\rho$ subclass, formerly known as GABA$_C$) are ionotropic receptors with fast kinetics that allow anions (i.e. usually Cl$^-$) to pass through the membrane and hyperpolarize the post-synaptic cell. The primary type of GABA$_A$ receptors (i.e. the “non-rho” subclass of receptors) are usually composed of various alpha, beta, and gamma subunits, though delta, epsilon, theta, and rho subunits may also be present (Barnard et al., 1998). The various GABA$_A$ subunits are themselves differentially expressed, and GABA$_A$$\alpha_1$, a popular biomarker, is detailed below. The second type of GABA receptor, GABA$_B$, is
metabotropic. Also note that, in general, the affinity of receptors for NTs (i.e. the synaptic plasticity) may be modified by extracellular matrix proteins, including reelin (RLN), a biomarker detailed below.

Finally, after the action potential, membrane-bound transporters participate in the reuptake of recycled or cytosolically synthesized NTs back into the pre-synaptic axon. These proteins, like the vesicular transporters, are specialized to carry specific NTs, including glutamate and aspartate (excitatory amino acid transporters, or EAAT1-5), GABA (GAT1-3), and monoamines such as dopamine (DAT), norepinephrine (NET), and serotonin (SERT).

**Names, synonyms, and identifying features of major biomarkers**

In the following sections, we review fifteen commonly used biomarkers from the categories mentioned above: membrane transporters (i.e. vGluT3), CaBP (i.e. CB, CR, and PV), neuropeptides (i.e. CCK, ENK, NPY, SOM, and VIP), NT receptors (i.e. mGlR1a, GABA	extsubscript{Aa1}, serotonin receptor 3 or “5-HT3”, cannabinoid receptor 1, and substance P receptor), and extracellular matrix proteins (i.e. RLN). The various names and synonyms of these biomarkers, combined with basic structural and chemical properties, are encapsulated in Table 1 and discussed below. Structural information was obtained from the UniProt knowledge base (www.uniprot.org) and the RCSB protein data bank (www.rcsb.org). Hydrophathy scores are calculated based on the “GRAVY”
### Table 1 Nomenclature and structural features for 15 biomarkers in mice

<table>
<thead>
<tr>
<th>Category</th>
<th>Marker (UniProt accession #)</th>
<th>Synonyms, related names, and abbreviations</th>
<th>Subcellular localization(s)</th>
<th>Sequence length (aa)</th>
<th>Molecular weight (kDa)</th>
<th>Isoelectric point</th>
<th>Amino acid composition (GlyCAVY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transporters</td>
<td>vGluT3 (Q88FU8)</td>
<td>Solute carrier family 17 (sodium phosphate), member 3; Solute carrier family 17 (vesicular glutamate transporter), member 3</td>
<td>Cytoplasmic vesicle membrane</td>
<td>601</td>
<td>66.148</td>
<td>6.05</td>
<td>0.186</td>
</tr>
<tr>
<td>CB (P12658)</td>
<td>Brain-2; CaBP; Calb; Calb-1; Calb1; Calbindin 1 (28kD); Calbindin D28 k; Calbindin D28K; Calbindin-28K; Calbindin-D28k; PCD-29; Spot 35 protein; Vitamin D-dependent calcium-binding protein, avian-type</td>
<td>Cytosol, nucleus, axon, dendrite, synapse</td>
<td>261</td>
<td>29.996</td>
<td>4.71</td>
<td>-0.484</td>
<td></td>
</tr>
<tr>
<td>CR (Q08331)</td>
<td>CAL2; Calb2; Calbindin 2 (29kD)</td>
<td>Cytosol, nucleus, terminal bouton, synapse</td>
<td>271</td>
<td>31.374</td>
<td>4.94</td>
<td>-0.655</td>
<td></td>
</tr>
<tr>
<td>PV (P32848)</td>
<td>Parvalbumin alpha</td>
<td>Cytosol, nucleus, axon, terminal bouton</td>
<td>110</td>
<td>11.931</td>
<td>5.02</td>
<td>-0.355</td>
<td></td>
</tr>
<tr>
<td>Neuropeptides</td>
<td>CK (P09240)</td>
<td>--</td>
<td>Nucleus, axon, axon hillock, axon initial segment, dendrite</td>
<td>115</td>
<td>12.770</td>
<td>9.44</td>
<td>-0.381</td>
</tr>
<tr>
<td>ENK (P22005)</td>
<td>ENK; Penk; Penk1; PPA; preproenkephalin 1; Preproenkephalin; Proenkephalin-A</td>
<td>Nucleus, axon, dendrite, plasma membrane</td>
<td>268</td>
<td>31.004</td>
<td>5.55</td>
<td>-0.821</td>
<td></td>
</tr>
<tr>
<td>NPY (P77774)</td>
<td>C-terminal peptide of NPY; CPON; Neuropeptide tyrosine; Neuropeptide Y; NPY</td>
<td>Nucleus, axon, cytosol</td>
<td>97</td>
<td>10.874</td>
<td>6.56</td>
<td>-0.405</td>
<td></td>
</tr>
<tr>
<td>SOM (P06041)</td>
<td>Antrin; preprosomatostatin; SMS; Smst; SOM; Somatostatin; Somatostatin-14; Somatostatin-28; SRIF; SS; Sst</td>
<td>Nucleus, extracellular matrix</td>
<td>116</td>
<td>12.746</td>
<td>5.47</td>
<td>-0.290</td>
<td></td>
</tr>
<tr>
<td>VIP (P32648)</td>
<td>Intestinal peptide PHI-27; Intestinal peptide PHI-42; MGC107202; Peptide histidine isoleucineamide 27; PHI, peptide histidine isoleucine; Vasactive intestinal polypeptide; VIP; VIP peptides</td>
<td>Nucleus, extracellular matrix</td>
<td>170</td>
<td>19.049</td>
<td>6.13</td>
<td>-0.416</td>
<td></td>
</tr>
<tr>
<td>Receptors</td>
<td>mGluR1a (P97772)</td>
<td>isoform 1, a, or alpha of the following: G protein-coupled receptor, family C, group 1, member A; Glutamate receptor, metabotropic 1; Glutamate receptor, metabotropic, type 1; Gprc1a; Grm1; GRM1; mGlu1 receptor; mGluR1</td>
<td>Post-synaptic cell membrane</td>
<td>1,199</td>
<td>133.212</td>
<td>6.44</td>
<td>-0.122</td>
</tr>
<tr>
<td>GABAa (P62812)</td>
<td>GABA A receptor alpha 1 subunit; GABA(A) receptor alpha1; GABA(A) receptor subunit alpha1; GABA(A)/Ralph1(1) subunit; GABA-A receptor subtype alpha 1; GABAA receptor alpha1 subunit; GABAA receptor subunit alpha 1; GABA-1; GABA1; gamma-Aminobutyric acid (GABA) A receptor alpha 1 subunit; gamma-Aminobutyric acid type A receptor alpha 1 subunit (GABRA1)</td>
<td>Post-synaptic cell membrane</td>
<td>455</td>
<td>51.754</td>
<td>9.34</td>
<td>-0.234</td>
<td></td>
</tr>
<tr>
<td>5-HT3 (P23979)</td>
<td>5-HT-3; 5-HT3; 5-HT3 receptor; 5-HT3-A; 5-HT3A; 5-HT3R; 5-hydroxytryptamine (serotonin) receptor 3A; 5-hydroxytryptamine receptor 3; 5-hydroxytryptamine receptor 3A; 5-hydroxytryptamine receptor 3A; 5-hydroxytryptamine receptor 3A; 5-hydroxytryptamine receptor 3A; 5-hydroxytryptamine receptor 3A; Serotonin receptor 3A; Serotonin-gated ion channel receptor</td>
<td>Post-synaptic cell membrane</td>
<td>487</td>
<td>56.056</td>
<td>6.15</td>
<td>0.091</td>
<td></td>
</tr>
<tr>
<td>CB1 (P47746)</td>
<td>Brain-type cannabinoid receptor; Cannabinoid receptor 1 (brain); Cannabinoid receptor CB1; CB-1; CB1 receptor; CB1R; CNNR; CNNR1; Cnr1</td>
<td>Post-synaptic cell membrane</td>
<td>473</td>
<td>52.831</td>
<td>6.58</td>
<td>0.333</td>
<td></td>
</tr>
<tr>
<td>sub P rec (P30548)</td>
<td>Neurontin-1 receptor; NK1-receptor; NK1 (substance P) receptor; NK1-R; NK1R; SPR; Tachykinin NK1 receptor; Tacr1</td>
<td>Post-synaptic cell membrane</td>
<td>407</td>
<td>46.322</td>
<td>7.56</td>
<td>0.322</td>
<td></td>
</tr>
<tr>
<td>Mek/ERK protein</td>
<td>RLN (Q60841)</td>
<td>CR-50 antigen; extracellular matrix serine protease; reeler; Reeler protein; Retin; Rl</td>
<td>Extracellular matrix, dendrite, cytosol</td>
<td>3,461</td>
<td>387.495</td>
<td>5.44</td>
<td>-0.264</td>
</tr>
</tbody>
</table>
methodology, wherein the hydrophathy values of each residue are added and the sum is divided by the length of the sequence; scores below 0 are more likely hydrophilic, while scores above 0 are more likely hydrophobic (Kyte and Doolittle, 1982).

**Membrane transporters**

vGluT3, also commonly known as solute carrier family 17 member 8, is encoded by the SLC17A8 gene. It is a synaptic vesicle membrane protein with both the amino- and carboxy-termini located in the cytoplasm, surrounding 12 transmembrane domains that alternately cross into and out of the vesicle. This results in a net hydrophobic protein, but the cytosolic and vesicular domains are generally hydrophilic. The chain consists of 601 amino acids (aa) with a total mass of more than 66.1 kDa.

**Calcium-binding proteins**

CaBP contain EF-hand domains with a high affinity for Ca\(^{2+}\) ions and, accordingly, undertake a crucial role in intracellular calcium homeostasis. Calbindin D28K and calretinin, which is confusingly sometimes designated as 29 kDa calbindin, are two closely related proteins belonging to the Troponin-C superfamily. The two sequences are quite similar; BLAST reveals that calretinin shares 60% identical residues with CB (with 76% positives, or similar residues, and only 2% gaps). Calretinin is the slightly longer (271 aa versus 261) and slightly heavier of the two; it contains 6 EF-hand domains that are home to 6 Ca\(^{2+}\)-binding sites, while CB contains 6 EF-hand domains and only 4 binding sites, as potential binding sites II and VI have lost their affinity for calcium. PV, formally known as parvalbumin alpha, is less than half the length of the other two (i.e. 110 aa) and only contains two EF-hand domains and Ca\(^{2+}\)-binding sites. All three have
isoelectric points near a pH of 5, above which they carry a net negative charge and below which they carry a net positive charge.

**Neuropeptides**

Peptides are cleaved from protein precursors, or pro-peptides, as post-translational modifications; in some cases, both the precursor and the peptide may be used as a biomarker. Table 1 lists data for the pro-peptide form of the molecules for comparison purposes. These peptides are generally produced in the nucleus of the cell, and sometimes the neurites, and are hydrophilic.

Pro-CCK, a 115-residue chain with a high isoelectric point, is cleaved into three CCK peptides with 32, 12, and 8 residues, respectively, each of which is hydrophilic. These peptides act on two receptors (i.e. CCK-A and CCK-B), both of which are expressed in the brain (Lay et al., 2000). The enkephalin precursor is formally known as proenkephalin, proenkephalin-1, or proenkephalin-A. With 268 aa, it is relatively long and heavy, and is cleaved into 8 active peptides, including two pentapeptides known as met-enkephalin (Tyr-Gly-Gly-Phe-Met) and leu-enkephalin (Tyr-Gly-Gly-Phe-Leu). Upon cleavage, proenkephalin-A generates four copies of met-enkephalin (from positions 100–104; 107–111; 136–140; and 210–214) and a single copy of leu-enkephalin (231–235). Both met- and leu-enkephalin bind to opioid receptors. Pro-somatostatin is a 116-residue molecule that is processed into three chains: antrin, somatostatin-28, and somatostatin-14, the latter two of which coincide with the C-terminus of the precursor. The actions of SOM are galvanized by binding to five different receptor subtypes (Patel, 1999). Pro-neuropeptide Y, or pro-neuropeptide tyrosine, is the shortest and lightest of
the biomarkers reviewed here (97 aa). It is processed into two chains: the 36-residue length NPY and the C-flanking peptide of neuropeptide Y, which is 30 aa long. NPY receptors are metabotropic and have five subtypes (Michel et al., 1998). Finally, the “VIP-peptides” encoding protein is 170 residues long and is processed into three chains: intestinal peptide PHI-42 and intestinal peptide PHI-27, which overlap, and VIP (28 aa). Either of two seven-transmembrane G protein-coupled receptors bind VIP in the brain.

**Neurotransmitter receptors**

Metabolic glutamate receptor 1 is a G protein-coupled receptor for glutamate that is alternatively spliced into three isoforms: mGluR1a, mGluR1b, and mGluR1c. mGluR1a is the canonical form, which is described here. 1,199 aa in length, the membrane-bound protein has an extracellular domain beginning at the N-terminus that stretches 572 residues; this portion contains the five glutamate binding sites. Seven helical transmembrane domains are then followed by a long, cytoplasmic C-terminal domain. Both the extracellular and the cytoplasmic domains are hydrophilic, but the membrane-bound portions lead to a slightly hydrophobic GRAVY score.

GABA$_A$ is an ionotropic heteropentamer receptor for GABA formed usually by one type of alpha and beta subunit, along with a single gamma polypeptide, in a ratio of 2:2:1 (Möhler et al., 2001), though in some cases other subunits such as epsilon or delta may replace gamma. The subunits arrange themselves around a central chloride anion conduction pore. The most common GABA$_A$ receptor is the $\alpha_1\beta_2\gamma_2$ subtype, which accounts for 60% of all GABA$_A$ receptors, followed by $\alpha_2\beta_3\gamma_2$ (15-20%) and $\alpha_3\beta\gamma_2$ (10-15%) (Möhler, 2006). The $\alpha_1$ subunit described here is a major biomarker in the
mammalian brain. It has a 223 aa-long extracellular ligand-binding domain near the N-terminus, three hydrophobic transmembrane helical regions that form the ionic channel, a cytoplasmic domain, and a fourth transmembrane helical region at the C-terminal of the sequence.

Receptors for three other NTs, serotonin, endocannabinoids, and substance P, are also reviewed. 5-HT3 is one of several receptors for 5-hydroxytryptamine, also known as serotonin. It is a ligand-gated ion channel that allows various cations to pass into a post-synaptic cell, affecting fast, depolarizing responses in neurons. Isoform A, the canonical sequence, is 487 aa long. Much like GABA_A1, 5-HT3A has a long extracellular region for NT binding, three transmembrane helices, a post-synaptic cell cytoplasmic domain, and a fourth transmembrane region. However, this serotonin receptor has a lower isoelectric point and is, on the whole, hydrophobic. Cannabinoid receptor 1 is a metabotropic receptor containing seven transmembrane-spanning regions that render it a net hydrophobic molecule. It is activated by natural, plant-based (e.g. by way of the psychoactive drug, cannabis), or synthetic cannabinoids (Felder et al., 1995; Ben-Shabat et al., 1998). Finally, the substance P receptor is also known as the neurokinin NK1 receptor. (NK2 and NK3 receptors bind preferentially to substance K and neuromedin K, respectively.) Of the receptors reviewed here, the substance P receptor (sub P rec) has the shortest extracellular domain (i.e. only 32 aa) and is the only one with a longer cytoplasmic domain (i.e. 74 residues); in between, it has seven transmembrane regions.
**Extracellular matrix proteins**
RLN, which is sometimes also referred to as the reeler protein, is the longest and heaviest biomarker reviewed here by nearly a factor of three. It is encoded by the Reln gene, which is disrupted in reeler mice, giving them a distinctive gait and the protein its name. Structurally, the N-terminal end has a 27-aa signaling peptide, a “reeler” domain with unknown function, and eight “reelin repeats” that are 300-350 aa long. These repeats have an epidermal growth-factor motif at their center that divides each repeat into two subrepeats, referred to as A (the BNR/Asp-box repeat) and B (the EGF-like domain). Despite this interlude, the two subdomains make direct contact, resulting in a compact, but massive, overall structure.

**Functional roles of biomarkers**
We now turn our attention to the various physiological roles for these biomarkers, both inside of and outside of the nervous system. Note that in peripheral tissues, many markers are also implicated in the digestive and/or endocrine systems (see Fig. 4 for an overview). We also discuss the function of the markers during development and their part in various pathologies. This information is summarized in Table 2 for each of the biomarkers under consideration here.

**Membrane transporters**
In contrast with vGluT1 and vGluT2, vGluT3 is expressed at rather low levels in the brain. In fact, in situ hybridization signals for vGluT3 in the medial septum and basal forebrain are absent (Gras et al., 2002; Schafer et al, 2002). Thus, the other two transporter isoforms shoulder the primarily responsibility for packaging glutamate in the
brain. However, the transporter is expressed by a specific subset of cells in multiple brain regions, including serotonergic neurons in the dorsal raphe, cholinergic interneurons of the striatum, and certain GABAergic interneurons of the hippocampus and the cortex (Fremeau et al., 2002; Gras et al., 2002; Schafer et al, 2002; Takamori et al., 2002). In the dorsal raphe, both in situ hybridization and immunohistochemistry show partial overlap (i.e. at least 70% co-expression) between the serotonin-positive populations and vGluT3-
<table>
<thead>
<tr>
<th>Category</th>
<th>Marker</th>
<th>Functional role</th>
<th>Developmental role</th>
<th>Pathological role</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Neurological</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>Trans-</td>
<td>vGlut3</td>
<td>1. Transport and packing of glutamate into vesicles; may be co-released with GABA or serotonin 2. Buffer of cytoplasmic glutamate</td>
<td>1. Unknown. mRNA expressed in liver and kidney [Fremeau et al., 2002]</td>
<td>1. Released transiently in selected cells; role unclear 2. Nonsyndromic hearing impairment</td>
</tr>
<tr>
<td></td>
<td>CR</td>
<td>1. (See CB) 2. Expressed in retina and somatosensory pathways (e.g. cochlear nuclei and olfactory bulb) 3. Also induces LTP</td>
<td>1. Expressed in lung mesothelial 2. Detected in Leydig cells of the testis, theca lutein and theca interna cells of the ovary 3. Expressed in in mast cells and mast cell lesions of the skin</td>
<td>1. (See CB) 2. CR absent from nerves in the bowel in Hirschsprung disease 3. In mesothelioma, CR differentially expressed in malignant and benign lung tumors 3. Decreased CR expression in hippocampus linked with temporal lobe epilepsy</td>
</tr>
<tr>
<td></td>
<td>PV</td>
<td>1. (See CB)</td>
<td>1. Involved in muscle relaxation after contraction (shuttles Ca⁺⁺ from cytosol to intracellular stores to accelerate relaxation of fast-twitch fibers)</td>
<td>1. (See CB) 2. Decreased PV expression in schizophrenia; disruption may affect synchronization of cortical circuits</td>
</tr>
<tr>
<td>Neuroptides</td>
<td>CCK</td>
<td>1. Modulates the effects of glutamate, GABA, dopamine, and serotonin 2. Increased CCK activity on exposure to stress, suggesting it might participate in stress response 3. Role in memory</td>
<td>1. Induces gall bladder contraction and the release of pancreatic enzymes in the gut 2. Hunger suppressant</td>
<td>1. Unclear, but CCK is detected in the nervous system as early as embryonic day 8 (E8) and in the digestive system by E15 2. Cause of visual hallucinations in Parkinson’s disease 3. Colorectal carcinomas produce CCK</td>
</tr>
<tr>
<td></td>
<td>ENK</td>
<td>1. Pain perception and analgesia 2. Respond to stress</td>
<td>1. Present in the digestive system, but role unclear</td>
<td>1. Help to regulate cell proliferation 2. Role in addiction and reward systems 3. Shown to induce seizures</td>
</tr>
<tr>
<td></td>
<td>NPY</td>
<td>1. Regulates food intake and storage of energy as fat</td>
<td>1. Induces vasoconstriction in cardiovascular system</td>
<td>1. During development, NPY expression corresponds to levels of maternal food provision [Grove et al., 2001] 2. Increases in NPY mRNA and NPY release linked to obesity, anorexia, and bulimia 3. Related to alcoholism</td>
</tr>
<tr>
<td>Receptor</td>
<td>Effect</td>
<td>Function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-HT3</td>
<td>Triggers neurotransmitter serotonin</td>
<td>1. Mediates fast excitatory transmission in the neocortical interneurons, amygdala, and hippocampus, and in the visual cortex.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Activates fast excitatory transmission in the hippocampus and visual cortex.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Inhibits GABAergic hippocampal neurons.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Regulates spatial memory.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Modulates pain perception, inflammation, and adaptive stress response.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CB1</td>
<td></td>
<td>1. Expression detected in heart, liver, kidney, spleen, small intestine, and gonads.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Present as early as E11 [Buckley et al., 1997].</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sub P</td>
<td>Binds substance P</td>
<td>1. High of substance P expression before birth decreases to adult levels by P14.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rec</td>
<td></td>
<td>2. Linked to chronic pain.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Role in development of obesity.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Neuropeptides (continued)

<table>
<thead>
<tr>
<th>Neuropeptide</th>
<th>Effect</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIP</td>
<td></td>
<td>1. Local energy metabolism and usage via glycolysis.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Neuroprotection.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Circadian rhythm regulation: synchronizes the suprachiasmatic nucleus function with the environmental light-dark cycle.</td>
</tr>
<tr>
<td>mGluR1a</td>
<td></td>
<td>1. Binds glutamate and initiates a host of electrical and chemical signaling pathways.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Regulates ion channels and cell excitability.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Auto-regulates synaptic transmission by reducing pre-synaptic release of glutamate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Implicated in LTP and LTD.</td>
</tr>
<tr>
<td>GABA&lt;sub&gt;a1&lt;/sub&gt;</td>
<td></td>
<td>1. Identified in adrenal gland, gonads, and small intestine.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. Minimal role during embryonic and prenatal development.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. Links to several neurological and psychiatric diseases, including epilepsy, Huntington's disease, anxiety disorders, alcoholism, and schizophrenia.</td>
</tr>
<tr>
<td>5-HT3</td>
<td></td>
<td>1. Mediation of signals of nausea, pain, and bloating from gut to brain.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. Serotonin present, but receptor roles not well documented [Gaspar et al., 2003].</td>
</tr>
<tr>
<td>CB1</td>
<td></td>
<td>1. Regulates the release of GABAergic neurotransmission.</td>
</tr>
<tr>
<td>sub P</td>
<td>Binds substance P</td>
<td>1. High of substance P expression before birth decreases to adult levels by P14.</td>
</tr>
<tr>
<td>rec</td>
<td></td>
<td>2. Linked to chronic pain.</td>
</tr>
<tr>
<td>RIN</td>
<td></td>
<td>1. Involved in development and cellular migration of small intestine.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Linked to bone and tooth formation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Also expressed in liver, blood, plasma, and reproductive organs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. Regulates processes of neuronal migration and positioning.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Plays a role in layering of neurons in the cortex, hippocampus, and cerebellum.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. Dysregulation of RIN gene linked with varicose cancers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Reduced expression detected in schizophrenia and bipolar disorder.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Also implicated in autism and Alzheimer's disease.</td>
</tr>
</tbody>
</table>
positive cells. Similarly, co-expression between these molecules is observed in the hippocampus near the border between stratum radiatum and stratum lacunosum-moleculare (Schafer et al., 2002; Somogyi et al., 2004). In other hippocampal neurons, such as some perisomatic-targeting, CCK+ basket cells (which are also negative for CR, PV, SOM, and VIP), vGluT3 is co-expressed with GABA (Somogyi et al., 2004), suggesting a co-release of GABA and glutamate into the synapse. It is not known whether GABA and glutamate are packaged in the same vesicle or whether glutamate is stored separately. Regardless, the abundant expression of this transporter in certain interneurons predicts an important role for glutamate in their signaling.

vGluT3 is also released transiently at certain locations during postnatal development, including in selected migrating cells and in Purkinje cells in the cerebellum (Boulland et al., 2004). In the rat, between P1 and P15, vGluT3 is expressed in striatum, accumbens, hippocampus, as well as in certain caudal brain structures, including the colliculi, pons and cerebellum. During a second phase extending from P15 to adulthood, the labeling in the caudal brain fades away.

Clinically, the gene that encodes for vGluT3 has been linked with nonsyndromic, slowly progressive, hearing impairment. Though the gene was originally identified in a human molecular genetics study (Green et al., 2001), the effect has been studied with a mouse model. Seal et al. (2008) generated a Slc17a8-null mice lacking the VgluT3 protein and found them to be deaf due to the lack of glutamate release from hair cells at the first synapse of the auditory pathway.
Calcium-binding proteins

Although there are many CaBP in the brain, the three mentioned here are particularly useful as biomarkers due to their abundancy and specific distribution. In general, CB, CR, and PV are segregated and rarely co-expressed with each other, though they do frequently co-express with GABA (Baimbridge et al., 1992). Thus, certain neuron types, like the CA3 pyramidal cells of the hippocampus, do not contain any of these CaBP (Sloviter, 1989; Toth and Freund, 1992).

Although there are many CaBP in the brain, the three mentioned here are particularly useful as biomarkers due to their abundancy and specific distribution. In general, CB, CR, and PV are segregated and rarely co-expressed with each other, though they do frequently co-express with GABA (Baimbridge et al., 1992). Thus, certain neuron types, like the CA3 pyramidal cells of the hippocampus, do not contain any of these CaBP (Sloviter, 1989; Toth and Freund, 1992).

The primary neurological roles of calbindin, calretinin, and parvalbumin are in binding to calcium ions (CB holds four, CR six, and PV two) and regulating cytosolic calcium levels, a responsibility which requires interplay with calcium ion pumps in the plasma membrane. Because high Ca$^{2+}$ levels trigger exocytosis, calcium-binding proteins play a major part in modulating and mediating NT release into the synaptic cleft. The effects could be manifested as modulation of action potential duration, promotion of neuronal “bursting” (i.e. by inhibiting potassium currents that rely on Ca$^{2+}$), or as protection against an excessive calcium influx that could damage the cell during prolonged periods of high activity (Baimbridge et al., 1992; Kojetin et al., 2006). In
addition, CR has been shown to modulate neuronal excitability by inducing long-term potentiation (Schurmans, 1997).

Outside of the brain, CB was first discovered in cytosolic fractions of the chicken intestine (Wasserman and Taylor, 1966; Wasserman et al., 1969). It was originally described as vitamin D-dependent because levels were diminished in vitamin D-deficient animals, but expression could be induced by treatment with vitamin D metabolites such as calcitriol (Wasserman and Taylor, 1966). In the brain, however, its synthesis is independent of vitamin D. CB has also been found in the mammalian pancreas (Johnson et al., 1994), kidney (Clemens et al., 1989; Bindels et al, 1991; Li and Christakos, 1991), and osteoblasts, where it regulates apoptosis and serves as a critical determinant in the rate of bone formation (Bellido et al., 2000). On the other hand, while CR is expressed in other parts of the body, including the lung mesothelium (Lugli et al., 2003), it is primarily and abundantly expressed in the brain, including in the retina, which gave it its name (Rogers, 1987), and in various sensory pathways (e.g. certain cells and fibers of the cochlear nuclei and olfactory bulb) (Résibois and Rogers, 1992). PV is localized at high levels in fast-contracting muscles, where it assists in the contraction-regulation cycle by playing a role in Ca\(^{2+}\) exchange between the sarcoplasmic reticulum and the myofibrils (Celio and Heizmann, 1982).

During development, these CaBPs play a role in regulating Ca\(^{2+}\) pools critical for synaptic plasticity, and the transient expression of all three in the brain seems to be stage-dependent. CB expression is detected before the other two, shortly after cessation of mitosis when neurons begin migration and differentiation, followed by PV, which is
expressed in parallel with an increase of neuronal activity, and CR is generally detected later and in specific sub-populations (Permyakov, 2007).

Links between these three CaBPs and various pathological conditions are also well established. First, CB has been associated with Alzheimer's disease. In addition to reports of reduced CB expression in the brains of mice and humans with Alzheimer’s (Heizmann et al., 1992), Kook and colleagues showed experimentally that removal of CB from amyloid-precursor-protein transgenic mice aggravates Alzheimer’s dysfunction (Kook et al., 2014). Similarly, in Huntington’s disease, immunohistochemistry showed a substantial loss from the neostriatum of neurons containing CB. This population of neurons is particularly damaged in this pathology, suggesting that a failure of calcium buffering or homeostasis may contribute to cell death in Huntington’s disease (Seto-Ohshima et al., 1988). Thirdly, CB may confer some protection to substantia nigra dopaminergic neurons against the pathological process that is responsible for Parkinson's disease (Yamada et al., 1990). CR, too, is a diagnostic marker for diseases, including Hirschsprung disease, where CR is absent from nerve trunks in the bowel (Alexandrescu et al., 2013), and mesothelioma, where it can help differentiate benign and malignant lung tumors (Marchevsky, 2008). In addition, loss of expression of CR in hippocampal interneurons has been implicated in temporal lobe epilepsy (Tóth and Maglóczky, 2014). Lastly, normal PV expression was found to be diminished in interneurons of individuals with schizophrenia (Nakazawa et al., 2012). These interneurons, many of which act via fast-spiking firing patterns, are critical for tight temporal control of cortical inhibition, so
their disruption may affect synchronization of disparate cortical circuits and result in schizophrenia symptomology.

**Neuropeptides**

CCK plays roles in both the nervous and digestive systems. The physiologic role of CCK in the central nervous system is not well understood, although CCK has been shown to have NT-like properties. It often acts as a modulator of the effects of classical neurotransmitters, including glutamate (e.g. in inhibitory hippocampal basket cells, where, in coordination with vGluT3, it stimulates exocytosis of glutamate (Williams, 1982; Breukel et al., 1997; Somogyi et al., 2004)), GABA (e.g. CCK has been shown to increase GABA release in the rat cortex and neostriatum (Sheehan and de Beller, 1983; Peréz de la Mora et al., 1993; Rakovska, 1995; Lanza and Makovec, 2000)), dopamine (Crawley, 1991), and serotonin (Woodruff and Hughes, 1991). In addition, CCK may regulate behaviors such as anxiety (Harro et al., 1993; Daugé and Léna, 1998) and memory (Flood et al., 1987). In the gut, in response to a meal, CCK induces gall bladder contraction and the release of pancreatic enzymes and gastric acid, which catalyze the digestion of fat, protein, and carbohydrates (Liddle, 1995). The peptide also plays a role in hunger suppression (Shillabeer and Davison, 1987). Pathologically, CCK has been implicated as a cause of visual hallucinations in Parkinson’s disease (Lenka et al., 2016), and, in the digestive system, colorectal carcinomas have been found to produce CCK and CCK receptors (Baldwin and Shulkes, 1998).

Enkephalins (i.e. both met- and leu- forms) are one of three families of endogenous opioid peptides, along with endorphins and dynorphins. Met-enkephalin acts
through μ- and δ-opioid receptors, and leu-enkephalin acts solely through δ-opioid receptors, which are expressed differentially throughout the brain (see Mansour et al., 1995). In either case, by competing with and mimicking the effects of opiate drugs (e.g. morphine), the primary function of enkephalins is to modulate information related to pain perception, analgesia, and pleasure, along with responses to stress, aggression, and dominance (Rosenfeld, 1994; Mansour et al., 1995). Though δ-opioid receptors are restricted to the brain, μ-opioid receptors are also found in the digestive tract, and enkephalin immunoreactivity was observed in many parts of the gut where its role is unclear (Schultzberg et al., 1980). In development, the endogenous opioids have been shown to serve as inhibitory growth factors limiting cell proliferation (i.e. specifically, neurogenesis and gliogenesis in the cerebellum) (Zagon and McLaughlin, 1991).

Clinically, enkephalin is a factor in drug addiction and in stimulation of the reward system, so much that rats have been shown to work for enkephalin injections delivered directly into their brains (Stein and Belluzzi, 1979). Linkage has also been shown to epileptic seizures (Frenk et al., 1978).

NPY, which was first isolated in the pig, has two intestinal counterparts known as peptide YY and pancreatic polypeptide, both of which are also 36-residue peptides with a similar structure (Tatemoto et al., 1982). Because their amino acid sequences differ, these relatives are not reviewed here. NPY, is localized to the nervous system, including the peripheral nervous system, where immunoreactivity has been detected in nerves of the heart, spleen, kidney, respiratory and urogenitary tracts, around blood vessels, and within visceral smooth muscle (Tatemoto, 2004). In the central nervous system, NPY acts
primarily as a regulator of energy homeostasis (Chronwall et al., 1985). Increased NPY activity has been directly tied to food intake, and it is a very potent appetite stimulant. Specifically, administration of NPY agonists increases food intake, and blocking NPY receptors inhibits activity of the peptide and decreases intake (Clark et al., 1984; Levine and Morley, 1984; Stanley and Leibowitz, 1984). Further, NPY concentrations were found to be altered in both cases of anorexia and bulimia (Kaye et al., 1990), as well as in obesity (Dryden et al., 1995), suggesting a key role in eating disorders. Additional evidence in mice suggests NPY might protect against alcoholism (Thiele et al., 2002; review in Ciccocioppo et al., 2009).

SOM is a neuropeptide with wide distribution in the body that was first identified in the hypothalamus. It is known to regulate growth hormone secretion from cells of the anterior pituitary gland (Brazeau et al., 1973). In the central nervous system, it is particularly expressed in the hippocampal formation and the neocortex, where it has been implicated in a variety of functions, such as pain perception (Kuriashi et al., 1985) and memory formation (Matsuoka et al., 1994) via changes in neuroplasticity (Kullman et al., 2002; Scheyltjens and Arckens, 2016; see Liguz-Lecznar et al., 2016 for review). In the hippocampus, SOM is co-released with GABA from interneurons (Freund and Buzsaki, 1996; Somogyi and Klausberger, 2005), and it has been found to both excite and inhibit cortical neurons in tissue culture (Delfs and Dichter, 1983). In peripheral tissues, it regulates endocrine and exocrine secretion and acts as a modulator of motor activity in the gastrointestinal tract (Reichlin, 1983). Altered levels of SOM are observed in several brain dysfunctions, such as Alzheimer’s disease (Ferrier et al., 1983; Terry and Katzman,
Parkinson’s disease (Jiménez-Jiménez et al., 2014), multiple sclerosis (Roca et al., 1999), and temporal lobe epilepsy (Buckmaster et al., 2002), where a near complete loss of SOM-containing neurons was observed (Sloviter, 1987).

As its full name implies, VIP was discovered in the digestive system. There, it induces smooth muscle relaxation to increase motility in the esophagus, stomach, gallbladder, and intestinal lumen, where it also regulates secretion of gastric acid, water, and electrolytes (Bitar and Makhlouf, 1982; Racusen and Binder, 1977). In addition, VIP is widely expressed in the cardiovascular system, respiratory system, kidney, reproductive system, and blood cells (see Table 2 for details; see Said et al., 2013 for review). In the brain, VIP and its receptors are expressed in those regions thought to be involved in learned behaviors, including the hippocampus, cortex, amygdala and hypothalamus. There, VIP is commonly expressed in GABAergic interneurons (Dietl et al., 1990; Somogyi and Klausberger, 2005), where it serves a number of functions. It stimulates local energy metabolism through glycogenolysis, the breakdown of glycogen into glucose, which results in increased glucose availability for the generation of phosphate-bound energy (Magistretti et al., 1981). Secondly, VIP appears to provide neuroprotection against cell death and diseases such as Alzheimer’s disease and Parkinson’s disease (Offen et al., 2000; Brinneman, 2007). Thirdly, VIP neurons regulate circadian rhythms in the suprachiasmatic nucleus by producing higher firing rates during the day relative to night (Hermanstyne et al., 2016). Throughout early nervous system development, VIP can stimulate neurogenesis, modulate the development of the neural
tube, and affect proliferation of precursors of the cerebral cortex (DiCicco-Bloom, 1996). In fact, defects in the VIP regulation system in mice have been shown to cause failure of neural tube closure and death by embryonic day 9 (E9) (Pinhasov et al., 2003). VIP has also been shown to have proliferative activity on neural precursor cells during mice embryonic development, as it increases the rates of neural cell proliferation by shortening the G1 and S phases of the cell cycle (Gressens et al., 1998). During embryogenesis, unregulated VIP may have major consequences and has been linked to disorders of neurodevelopment, including autism, Down syndrome, and fetal alcohol syndrome (Hill, 2007). Lastly, VIP also appears to be involved in disease states like epilepsy, in which neurogenesis is permanently altered (Parent et al., 2002).

**Neurotransmitter receptors**

mGluR1a is a receptor that is distributed widely throughout the brain, though, unlike its ligand, glutamate, not ubiquitously. mGluR1 is one of eight receptor variants that are divided into three groups based on sequence homology. Members of the same group (e.g. mGluR1 and mGluR5) share 70% sequence identity, while between-group identities drop to 45% (Conn and Pin, 1997). As a metabotropic receptor, mGluR1 exerts its effects through secondary messenger systems and is therefore linked to a litany of signaling pathways that are beyond the scope of this review. Its function was first discovered in the hippocampus (Charpak et al., 1990). There, high-frequency firing was induced in CA3 pyramidal cells, leading to an overabundance of intracellular Ca\(^{2+}\), which in turn resulted in the activation of K\(^+\) channels to hyperpolarize the cell and inhibit further action potentials. However, agonist-binding to Group 1 receptors was found to
suppress this adaptive behavior, meaning that action potentials were continuously generated throughout the depolarizing input. Thus, mGluR1 helps to regulate ion channels and cell excitability (Gerber and Gähwiler, 1994). In addition, on the pre-synaptic side of the synapse, these receptors have been shown to act as autoreceptors, meaning that they regulate synaptic transmission and inhibit further glutamate release (reviewed in Glaum and Miller, 1994). Finally, mGluRs are also implicated in the mechanisms leading to long-term changes in synaptic transmission (i.e. long-term potentiation and long-term depression) in the cerebellar cortex (Kano and Kato, 1987), hippocampus (Bortolotto and Collingridge, 1993; Bashir et al., 1993), striatum (Calabresi et al., 1992), and visual cortex (Kato, 1993). These alterations have been suggested to form the basis for memory and learning.

mGluR1 receptors also have been found to have widespread distribution outside the brain, including in the gastrointestinal tract, peripheral sensory cells, and heart (see Table 2 and Julio-Pieper et al., 2011 for review). Though mGluR1 mRNA expression is present and increasing during postnatal life, data imply that, unlike mGluR3 and mGluR5, mGluR1 plays a minimal role during synaptogenesis but a large role in mature synaptic transmission (Catania et al., 1994). Pathologically, in the brain, mGluR1 has been linked to Huntington’s disease (Cha et al., 1998) and multiple sclerosis (Geurts et al., 2003; Newcombe et al., 2008). In the periphery, dysfunction of the receptor has a role in ulcer formation (Akiba et al., 2009) and the pathogenesis of melanoma (Shin et al., 2008).
GABA receptors are present on the majority of interneurons in the brain, holding the primary responsibility of binding GABA and converting this chemical signal into an electrical mediation of the fast inhibitory neurotransmission (Fritschy and Brünig, 2003). GABA_\textsubscript{A} is one of three receptors, and it is composed from a selection of subunits; this heterogeneity appears to be a key factor in fine-tuning of inhibitory transmission under physiological and pathophysiological conditions. Specifically, GABA_\textsubscript{A}α1 subunit agonists potentiate IPSC amplitude to various extents depending on the degree of receptor occupancy (Hajos et al., 2000). The GABA_\textsubscript{A} α1β2γ2 subtype is localized in the cerebral cortex, hippocampus, cerebellum, and amygdala, among other regions, though it appears to be absent from the olfactory bulb, striatum, thalamic reticular nucleus, inferior olive, and motoneurons (Fritschy and Brünig, 2003). Expression in the hippocampus is especially interesting, where two subtypes of basket cells (i.e. those that express PV and are fast-spiking, and those that express CCK and are regular spiking (Pawelzik et al., 2002)) are also separable based on their usage of the α1 (for PV) and α2 (for CCK) receptor subtypes in synapses onto post-synaptic CA1 pyramidal cells (Nyíri et al., 2002; Klausberger et al., 2002).

The GABA_\textsubscript{A} α1 receptor has been identified in peripheral tissues such as the adrenal gland, gonads, and small intestine (Erdö and Wolff, 1990; Akinci and Schofield, 1999). Unfortunately, though GABA has also been detected in several additional systems and organs, much of the evidence is not specific to the subunits of GABA_\textsubscript{A}. In development, though the α1 subunit displays weak expression in the cortex, thalamus, hippocampus, and cerebellum as soon as embryonic day 18, it appears that other receptor
subunits play a larger role. Clinically, in the adult, functions in multiple neurological and psychiatric diseases have been reported, including epilepsy, anxiety disorders, alcoholism, Huntington's disease, Angelman syndrome, and schizophrenia (Möhler, 2000).

In binding serotonin, 5-HT3 receptors regulate the release and effects of other neurotransmitters. These receptors are found both pre- and post-synaptically (Nichols and Nichols, 2008), with highest expression in the brainstem (Gehlert et al., 1993). The receptor also has been detected at lower levels in several areas of the forebrain, including the hippocampus, nucleus accumbens, putamen, caudate nucleus, amygdala, and visual cortex, where it mediates fast excitatory synaptic transmission (Sugita et al., 1992; Bufton et al., 1993; Roerig et al., 1997; Férézou et al., 2002). In addition, in the hippocampus, blockage of 5-HT3 receptors (i.e. by receptor antagonists) was shown to reduce hyperpolarization and thereby enhance both the frequency of the naturally occurring theta rhythm and the induction of long-term potentiation; concomitantly, retention was observed to be improved in both an odor-matching problem and a spatial task (Stäubli and Xu, 1995). Similarly, 5-HT3 receptors modulate dopamine: agonists were found to increase dopamine release in the striatum (Blandina et al., 1988; Blandina et al., 1989) and antagonists had anti-dopaminergic effects (Costall et al., 1987). Antagonists also temper the psychomotor stimulant effects of drugs such as cocaine, amphetamines, nicotine, and morphine (Grant et al., 1995). Peripheral roles for the receptor are less well known but, within the digestive system, 5-HT3 receptors mediate
transmission of signals from the gut to the brain, including messages of nausea, pain, and bloating (Gershon, 2004).

Cannabinoid receptors (CB1 and CB2) are metabotropic receptors that bind to three major groups of ligands: endo-, plant, and synthetic cannabinoids. CB1 is found mainly in the brain, with high levels occurring in the basal ganglia, hippocampus, cerebellum, and cerebral cortex, with more moderate levels in the hypothalamus (Herkenham et al., 1990; Jansen et al., 1992). This receptor is often localized to axon terminals (e.g. in CCK-positive interneurons of the hippocampus and amygdala), and its activation leads to inhibition of NT release (Katona et al., 1999; Katona et al., 2001; Szabo and Schlicker, 2005). The pre-synaptic nature of action also indicates a role as a retrograde messenger, meaning that when endocannabinoids are released by post-synaptic principal cells, these cells are then able to modulate their own GABAergic inputs (i.e., an effect known as depolarization-induced suppression of inhibition (Wilson and Nicoll, 2001; Wilson et al., 2001)). This mechanism may explain some of the emotionally relevant behavioral effects of cannabinoid exposure. The CB1 receptor is also present in some non-neuronal cells and tissues, including the heart, liver, kidney, spleen, small intestine, and gonads (Pertwee, 1997). Pathologically, CB1 is considered to mediate the majority of the psychoactive properties of cannabis (Ameri, 1999), and increased binding levels have been detected in schizophrenia (Dalton et al., 2011) and Parkinson’s disease (Lastres-Becker et al., 2001), though the effects of drug treatments could not be ruled out.

In the guinea pig brain, sub P rec is expressed to some degree in superficial layers of the cortex, thalamus, hypothalamus, CA1-3 and the dentate gyrus of the hippocampus,
and various mid- and hind-brain structures (Yip and Chall, 2000). Generally, as substance P is widely implicated in pain perception, inflammatory response, and stress response, antagonists that act on its receptors negate these effects (Henry, 1992; De Felipe et al., 1998; Nichols et al., 1999; Zubrzycka and Janecka, 2000; Duric and McCarson, 2005; Ebner and Singewald; 2006). In peripheral tissues, sub P rec are found in the digestive system (Sternini et al., 1995) and endocrine glands, where they reduce levels of hormones produced in response to stress upon binding of substance P (Jessop et al., 1999).

Interestingly, during development, the density of sub P rec is maximal one day before birth. These high levels gradually decrease until they reach adult levels at approximately 2 weeks after birth (P14), suggesting an important but as-yet-uncharacterized role for substance P in the early organization of the central nervous system (Quirion and Dam, 1986). Pathologically, sub P rec plays a chief role in chronic pain perception (Nichols et al., 1999). In addition, administration of a specific antagonist was shown to lead to reduced weight gain and circulating levels of insulin and leptin after a high-fat diet, pointing to a role for sub P rec in obesity (Karagiannides et al., 2011).

**Extracellular matrix proteins**

In the nervous system, reelin plays a pivotal role in deploying and positioning neurons during brain development, especially in laminated brain regions such as the cerebral cortex, hippocampus, and cerebellum. After a multi-step synthesis involving the rough endoplasmic reticulum and Golgi vesicles, reelin is secreted through porosomes into the extracellular matrix. As young neurons migrate from the site of their origin to the final destinations, where they adopt their final morphological patterns, reelin binds to
their transmembrane lipoprotein receptors (D’Arcangelo et al., 1999; Hiesberger et al., 1999). This triggers an involved intracellular signaling cascade through the cytoplasm and nucleus that instructs the neurons to occupy their proper locations in the developing brain (Rice and Curran, 2001).

In the cortex, transient reelin-expressing cells are primarily found in the marginal zone and in the temporary sub-pial granular layer (Meyer and Goffinet, 1998). In the hippocampus, Cajal-Retzius cells secrete reelin into the extra-cellular matrix of the stratum lacunosum-moleculare and the outer molecular layer of the dentate gyrus (Del Rio et al., 1997; Borrell et al., 1999). Reelin is also expressed in the external granule cell layer of the cerebellum, where granule cell migration to the internal granule cell layer is initiated (Schiffmann et al., 1997).

Synthesis of reelin decreases postnatally, becoming more diffuse compared with the distinctly laminar expression in the developing brain. However, in the adult, the protein still plays a role at active neurogenesis sites, including the subventricular zone and the dentate gyrus (Frotcher et al., 2003), and it is secreted in adulthood by GABAergic neurons that originate in the caudal ganglionic eminence (Tricoire et al., 2011). In addition, in the mature brain, the reelin receptor Apoer2 modulates synaptic plasticity to maintain long-term potentiation and favor memory formation (D’Arcangelo, 2005).

Reelin expression in the brain is well documented, but a large body of evidence now suggests that reelin is involved in the cellular migration and proliferation of non-neuronal tissues (see Khialeeva and Carpenter, 2016 for a full review). In brief, reelin
mRNA is found in the small intestine, where the epithelium has a high turnover rate and is subject to rapid renewal (Garcia-Miranda et al., 2010), and reelin is involved in the processes of osteogenesis and dentinogenesis (i.e. bone and tooth formation) (Maurin et al., 2004; Rawlinson et al., 2009). Reelin is also expressed in hepatic stellate cells of the liver, an organ that is self-regenerative (Smalheiser et al., 2000; Kobold et al., 2002), along with the blood (Smalheiser et al., 2000; Chu et al., 2014), mammary gland (Khialeeva et al., 2011), and the bovine and chicken endometrium and ovarian follicle (Fayad et al., 2007; Eresheim et al., 2014). Though the exact mechanism of actions in these widespread systems are unclear, the pervasive nature of reelin, in spite of its large size and the energy required to secrete it, points to a broad but significant role in the development of a diverse set of organs and tissues.

Finally, regarding the various pathological roles of reelin, the processes involved in tumorigenesis often employ the same pathways that are required for normal development. Thus, it is unsurprising that reelin up- or down-regulation has been implicated in many cancers. Reelin expression is reduced in hepatocellular carcinoma and breast, gastric, and pancreatic cancers (Hong et al., 2008; Dohi et al., 2010; Stein et al., 2010; Okamura et al., 2011). In contrast, upregulation is observed in esophageal carcinoma, prostate cancer, and retinoblastoma (Wang et al., 2002; Perrone et al, 2007; Seigel et al., 2007). In addition, reelin has been linked with several brain disorders, including schizophrenia (Impagnatiello et al., 1998; Eastwood et al., 2003; Fatemi et al., 2005), bipolar disorder (Guidotti et al., 2000), and Alzheimer’s disease (Wirths et al., 2001; Botella-López et al., 2006; Seripa et al., 2008).
**Pharmacological implications**

The sobering assortment of documented pathological roles for the biomarkers reviewed here, coupled with rich compendia of Cre transgenic mouse lines for drug testing, has opened the door to opportunities for pharmacological intervention.

However, for our fifteen markers, current prospects appear to be largely restricted to the receptor biomarkers and the receptors of the neuropeptides. In fact, according to the drug-gene interaction database (DGI-DB; [http://www.dgi-db.org](http://www.dgi-db.org)) (Griffith et al., 2013; Wagner et al., 2016), neither vGluT3 nor any of the CaBP are known to be targeted or affected by any drug presently used by humans. Moreover, current drug development strategies do not currently focus on intervention at the calcium-regulation level. In the case of CaBP, though it has been reported that CB can protect cells against amyloid-β peptide toxicity (Mattson et al., 1997), the use of agents affecting intracellular calcium was found to be of limited value in the treatment of Alzheimer’s and dementia (Fritze and Walden, 1994; De Vry et al., 1997).

Similarly, drugs are not used to directly target any of the neuropeptides discussed here. However, an impressive barrage of exogenous agonists and antagonists are lobbed at the receptors for many of these neuropeptides. For each biomarker in this category, Table 3 lists the human genes that encode its various receptors. Queries leveraging DGI-DB were then run, and the number of hits for each gene was noted. From the results that we deemed to be of “high-quality” (i.e. meaning they had DGI-DB scores ≥5, which are calculated based on number of source databases with information and the amount of supporting evidence from PubMed), a sampling of drugs was included in Table 3. They
Table 3 Selected drugs that target biomarker receptors

<table>
<thead>
<tr>
<th>Marker</th>
<th>Gene for receptor</th>
<th>DGI-DB hits (score &gt;8)</th>
<th>Sampling of drugs</th>
<th>Drug indication</th>
<th>Interaction type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCK</td>
<td>CCKBR</td>
<td>12 (1)</td>
<td>Pentagastrin</td>
<td>Diagnostic aid for evaluation of gastric acid secretory function</td>
<td>Agonist</td>
<td>(Radu et al., 2002)</td>
</tr>
<tr>
<td>ENK</td>
<td>OPRD1</td>
<td>53 (16)</td>
<td>Butorphanol</td>
<td>Moderate to severe pain</td>
<td>Agonist</td>
<td>(Chen et al., 1993)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Morphine</td>
<td>Relief and treatment of severe pain</td>
<td>Agonist</td>
<td>(Abdelhamid et al., 1991)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Naloxone</td>
<td>Reversal of narcotic depression induced by opioids and opioid overdose; to increase blood pressure in management of septic shock</td>
<td>Antagonist</td>
<td>(Strang et al., 1999)</td>
</tr>
<tr>
<td>OPRM1</td>
<td>84 (32)</td>
<td>Fentanyl</td>
<td>Treatment of cancer patients with severe pain</td>
<td>Agonist</td>
<td>(Dosen-Miovic et al., 2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Methadone</td>
<td>Pain; drug withdrawal syndrome; opioid type drug dependence; dry cough</td>
<td>Agonist</td>
<td>(Shi et al., 2002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sufentanil</td>
<td>General anesthetia</td>
<td>Agonist</td>
<td>(Leysern and Gommeren, 1982)</td>
<td></td>
</tr>
<tr>
<td>NPY</td>
<td>NPY1R</td>
<td>2 (0)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>NPY5R</td>
<td>4 (1)</td>
<td>Velpeptin</td>
<td>Obesity</td>
<td>Antagonist</td>
<td>(Powell et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>NPY2R</td>
<td>2 (0)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>SOM</td>
<td>SST1</td>
<td>10 (2)</td>
<td>Octreotide</td>
<td>Acromegaly; side effects from cancer chemotherapy</td>
<td>Binder</td>
<td>(Matrone et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>SST2</td>
<td>15 (2)</td>
<td>Pasireotide</td>
<td>Cushing’s syndrome</td>
<td>Binder</td>
<td>(Zafetti et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>SST3</td>
<td>15 (1)</td>
<td>[Pasireotide &amp; Pasireotide]</td>
<td>[see above]</td>
<td>[see above]</td>
<td>[see above]</td>
</tr>
<tr>
<td></td>
<td>SST4</td>
<td>7 (0)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>SST5</td>
<td>15 (2)</td>
<td>[Octreotide &amp; Pasireotide]</td>
<td>[see above]</td>
<td>[see above]</td>
<td>[see above]</td>
</tr>
<tr>
<td>VIP</td>
<td>VIPR1</td>
<td>2 (0)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>VIPR2</td>
<td>2 (0)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>mGluR1a</td>
<td>GRM1</td>
<td>10 (0)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>GABAa</td>
<td>GABRA1</td>
<td>84 (41)</td>
<td>Pentobarbital</td>
<td>Short-term treatment of insomnia</td>
<td>Potentiator</td>
<td>(Steinbach and Aik, 2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Thiamylal</td>
<td>General anesthetia; inducing a hypnotic state</td>
<td>Agonist</td>
<td>(Sugimura et al., 2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Phenobarbital</td>
<td>Seizures</td>
<td>Potentiator</td>
<td>(Macdonald and McLean, 1986)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ethanol</td>
<td>Chronic pain in such conditions as inoperable cancer and trigeminal neuralgia</td>
<td>Agonist</td>
<td>(Santhakumar et al., 2007)</td>
</tr>
<tr>
<td>S-HT3</td>
<td>HTR3A</td>
<td>54 (8)</td>
<td>Ondansetron</td>
<td>Alcohol use disorders; chemotherapy-induced nausea and vomiting</td>
<td>Agonist</td>
<td>(Müller et al., 1998; Litten et al., 2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Granisetron</td>
<td>Nausea and vomiting</td>
<td>Antagonist</td>
<td>(Rubenstein et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Clecapeptide</td>
<td>Symptomatic treatment of nocturnal heartburn due to gastroesophageal reflux disease</td>
<td>Agonist</td>
<td>(Talley, 1992)</td>
</tr>
<tr>
<td>CB1</td>
<td>CNR1</td>
<td>10 (3)</td>
<td>Dronabinol</td>
<td>Anorexia; chemotherapy-induced nausea and vomiting; disturbed behavior in Alzheimer’s</td>
<td>Agonist</td>
<td>(Voilier et al., 1997)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Nablimine</td>
<td>Chemotherapy-induced nausea and vomiting</td>
<td>Partial agonist</td>
<td>(Davis et al., 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rimonabant</td>
<td>Obesity</td>
<td>Antagonist</td>
<td>(Shire et al., 1996)</td>
</tr>
<tr>
<td>sub P rec</td>
<td>TACR1</td>
<td>30 (1)</td>
<td>Aprapant</td>
<td>Chemotherapy-induced nausea and vomiting</td>
<td>Antagonist</td>
<td>--</td>
</tr>
</tbody>
</table>

were chosen to represent the range of drug indications and interaction types (e.g. agonist, antagonist, etc.) based on available information in the PubChem database. Not surprisingly, the δ- and μ-opioid receptors were the most popular targets by a considerable margin. These receptors, to which ENK also binds, play a major role in pain perception and treatment. Receptors for other biomarkers appear currently to be in
limited pharmacological use and, when they are utilized, it is in treatment relating to digestive distress or obesity.

Additionally, 4 out of the 5 NT receptors we reviewed are major pharmacological targets, with the exception mGluR1a. For the others, DGI-DB searches reveal a plethora of chemicals that can be aimed at these biomarkers to treat a variety of conditions. In particular, GABA\textsubscript{A} receptors represent a major site of action for clinically important drugs, including benzodiazepines, barbiturates, and some general anesthetics, as well as drugs of abuse such as ethanol (see Table 4 for references).

However, for all of the biomarker targets that were analyzed, current use in the brain was found to be notably limited in favor of the peripheral tissues overviewed in Figure 4. Dronabinol, a psychoactive compound extracted from Cannabis plants that binds to CB1 in the central nervous system, which has been investigated for use in Alzheimer’s patients who refuse food and exhibit “disturbed behavior,” was the exception. Thus, though the viability of many of these markers as drug targets remains to be seen, the field would seem to be quite open.

**Relationship between biomarkers and neuron types**

Although biomarkers are often first thought of as potential drug targets, in theory, they also may be used to identify neuron types. In practice, however, this is complicated by the resolution of still-improving techniques, the condition-dependence of certain markers, and the large number of biomarkers that often work in concert.

Although no specific type of neuron can be fully defined by a single marker, certain interneurons reliably express a given molecule across regions of the brain. For
example, some perisomatic-targeting cells, like axo-axonic cells and fast-spiking basket cells, are consistently PV-positive, regular spiking basket cells are CCK-positive, neurogliaform cells usually express NPY, and Cajal-Retzius cells are positive for reelin.

Further, some types express specific combinations of different biomarkers that can aid in their identification. For example, though interneuron-specific interneurons in CA1 of the hippocampus come in a variety of morphological flavors, so as to inhibit distinct sets of interneurons in the CA1 subregion, they resolutely express CR and VIP (Acsády et al., 1996; Gulyás et al., 1996). Therefore, we combed the literature for well-supported and agreed-upon evidence of expression for prominent neuron types across the brain. Figure 5 shows the consensus for expression (green squares) or lack of expression (blue squares) of our 15 markers of interest across 43 types from 7 regions (abbreviations: HF – hippocampal formation, DG – dentate gyrus).

In the amygdala, in addition to confirming the perisomatic-targeting type expression data mentioned above (Capogna, 2014), neurogliaform cells were found to express NPY, as well as CB and SOM (Mańko et al., 2012). Purkinje cells in the cerebellum are known to co-express two CaBPs, CB and PV (Weyer and Schilling, 2003), and data from other GABAergic types in the region encompass bipolar cells, Golgi cells, Lugaro cells, and stellate cells (D’Angelo et al., 2013; Cauli et al., 1997; Celio and Heizmann, 1991; Mertz et al., 2000; Sahin and Hochfield, 1990). Expression information from the hippocampus was obtained from www.hippocampome.org, a rich knowledge base of neuron types and their literature-ascribed properties, covering the dentate gyrus, CA1-3, the subiculum, and the enthorinal cortex. In the neocortex, unique
and well-known types with expression information include, among others, double bouquet cells, which are positive for CB, CR, CCK, and VIP, and Martinotti cells, which,
also, are positive for CB, CR, and CCK in addition to famously being positive for SOM and RLN (DeFelipe, 1997; Douglas and Martin, 2004; Kawaguchi and Kondo, 2002; Kubota et al., 2002; Lee et al., 2010; Markram et al., 2004; Pesold et al., 1999; Porter et al., 1998; Wang et al., 2002). Finally, striatal data on basket cells, medium spiny neurons, and neurogliaform cells is also included (Muñoz-Manchado et al., 2016). Biomarker data for the olfactory bulb and retina were considered but are largely not localized to cell types.

As such information becomes more complete (i.e. across both dimensions), one can begin to see how biomarkers can be used to efficiently locate, label, and study neuron populations across the brain.

Conclusions
A vast array of biomarkers and corresponding expression information can be used in brain research. In a systematic review of 15 of the most widely used (though not well understood) markers, we have disentangled a web of synonyms and abbreviations, explored their structural features, and shed light on their operation both in the brain and in peripheral tissues. We also investigated developmental and pathological roles, cataloged information of pharmacological relevance to potential medicinal targets, and amassed expression data for well-known neuron types. Together, this information helps to ease the gridlock in our understanding of the molecular-level traffic in the brain, and it serves as a critical guidepost along the pathway between obtaining massive, molecular-expression datasets and using them to our advantage to improve our knowledge and our medicine.
CHAPTER TWO

Title: Weighing Peters’ rule: Does neuronal morphology predict connectivity?

Authors: Christopher L. Rees, Keivan Moradi, Giorgio A. Ascoli

Abstract
   In anxious anticipation of a full-brain connectome, assessment of potential connectivity via axon-dendrite juxtapositions has been employed as a surrogate. However, in recent years, this principle has been hastily or imprecisely applied, fueling passionate debates regarding its utility. We critically review the literature of what has come to be known as Peters’ rule and identify three distinct conceptual applications that have become muddled over time: inferring synapsing among neuron types, predicting synapsing between individual cells, and estimating synapse numbers. Paradoxically, at the originally proposed cell-type level, Peters’ rule remains largely untested. Through efforts like Hippocampome.org, we have an opportunity to validate and re-characterize a more concentrated rule to aid in the interpretation of future, massive synaptic datasets.
**Big data connectomes and the usage of potential-synapse proxies**

Substantial, global research endeavors such as the BRAIN Initiative (Insel et al., 2013) and the Human Brain Project (Markram et al., 2015) aspire to experimentally map the human brain connectome with detailed representations of individual cells and their spatial and temporal interactions. However, the existing experimental approaches (e.g. electron microscopy of a volume of tissue, electrophysiological paired recordings, and locating marked appositions under light or electron microscopes) are cumbersome. Thus, the challenge of collecting, integrating, and interpreting data for approximately $10^{11}$ neurons (Williams and Herrup, 1988; Azevedo et al., 2009; Herculano-Houzel, 2009) and $10^{15}$ synapses (Braitenberg and Schüz, 1998) is quite daunting. Even disregarding inter-individual diversity or temporal changes and focusing on small volumes of tissue in simpler species, the undertaking remains formidable.

While awaiting these promised data, which may still be a decade or more hence, many researchers have turned to Peters’ rule (Peters and Feldman, 1976) to answer the question of where synapses *might be*. Although certain connections may be genetically or developmentally specified, an as-yet-undetermined proportion of synapses appear to form between crossing axonal and dendritic arbors, mainly based on activity-dependent structural plasticity. Accordingly, Peters’ rule operates by identifying all spatial appositions between axons and dendrites as potential connections. The result is that readily available morphological information can be used to provide a statistical summary of the connectome.

The principle, little-heralded at inception, went unnamed and lightly cited for decades (Braitenberg and Schüz, 1998), but has recently and rapidly gained traction...
attributable to the approaching Big Data deluge. But even with a relatively brief history, the state of the Peters’ rule literature is untidy: confusion abounds regarding the original intent, scope, and even the proper spelling and punctuation of the name. Anecdotally, Google Scholar searches for the incorrectly apostrophized “Peter’s rule,” “Peters rule,” “Peters’s rule,” and the preferred variation, “Peters’ rule,” all returned neuroscience-relevant results that intermingled with hits pertaining to (1) the reign of Tsar Peter the Great, (2) Peter’s half-slope method (i.e. a way to find the depth of buried magnetic sources), and (3) the Peter principle (i.e. a management concept postulating that employees will continue to be promoted until they rise to the level of their incompetence). Mispunctuations aside, even after filtering to the connectivity literature, misunderstandings and misapplications of the rule are commonplace, and they have incited fervent discussion on both sides regarding its correctness and utility.

**Peters’ rule: its original intent and current tripartite usage**

We have objectively and critically reviewed the neuroscience literature and, for the first time, identified three conceptually distinct resolution levels at which the principle has been applied. These levels are discussed in detail below and summarized in Figure 6. Briefly, Peters’ rule has been used to determine potential connectivity between neuron types, potential connectivity between individual neurons, and the number of synapses between individual neurons. Worryingly, as exploitation of the concept has proliferated over time, the usage of the term as applied to each of the three levels has grown nearly proportionally (see Figure 7). This has led to a muddy pool of literature comprised of results that are not directly comparable.
Level 1 – assessment of potential connectivity between neuron types

Peters’ rule can be applied to identify the potential for connectivity between morphologically-defined neuron types with proximal axon and dendrite arbors. Naturally, several factors affect whether these connections are realized as one or more synapses,
including the numerosity of each type, the three-dimensional neurogeometries (Stepanyants and Chklovskii, 2005) and other branch morphometrics (Binzegger et al., 2004; Jefferis et al., 2007), the number of junctures at given touch distances (Stepanyants and Chklovskii, 2005; Kasthuri et al., 2015; Reimann et al., 2015), and the synaptic transmission method (i.e. chemical, electrical, or volume transmission). However, in general, the co-location of similar axonal arbors from multiple neurons with multiple dendritic trees reasonably forecasts at least some level of connectivity between the two types.

This is the resolution level upon which Peters’ rule was conceptually founded. In the late 1970s, Peters and Feldman reported that thalamic axon terminals projecting to layer IV of the rat visual cortex formed synapses in a haphazard fashion that was attributable largely to the distributions of the spines, shafts, and somata of the neuron types that happened to be available (Peters and Feldman, 1976; Peters et al., 1979; Peters; 1979). Two decades later, the notion of the lack of synaptic selectivity of neuron types was extended to the cortex at large, and the term “Peters’ rule” was coined (Braitenberg and Schüz, 1998). Subsequently, this idea has been applied in diverse species and neurological systems, including the hatchling frog tadpole spinal cord (Li et al., 2007), cat primary visual cortex (Binzegger et al., 2004; Potjans and Diesmann; 2014), rodent hippocampus (Wheeler et al., 2015), mouse retina (Morgan et al., 2011; Helmstaedter et al., 2013; Kim et al., 2014) and neocortex (Packer et al., 2013), and large-scale studies of the rat somatosensory cortex (Reimann et al., 2015) and fly photocerebral bridge (Lin et al., 2013) (see Table 4 for details of selected publications from the past five years). In
certain cases in the literature, the actual synaptic connectivity was determined and found to confirm the statistical impartiality of neuron types and, thus, Peters’ rule (green shading in Table 4) (Binzegger et al., 2013; Packer et al., 2013). Other reports challenged the hypothesis by relating exceptional connectivity biases among certain types (red shading in Table 4) (famously White and Keller, 1987, known as White’s Exceptions; also, Li et al., 2007; Morgan et al., 2011; Helmstaedter et al., 2013; Kim et al., 2014).

Table 4 Recent applications of Peters’ rule

<table>
<thead>
<tr>
<th>Reference</th>
<th>Level</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Briggman et al., 2011</td>
<td>2</td>
<td>Using electron microscopy and two-photon calcium imaging, reconstructed 24 cells of mouse starburst amacrine cells and 6 direction-selective retina ganglion cells. Found the types to be connected, but connections between individual cells were highly specific depending on the preferred direction of the ganglion cell.</td>
</tr>
<tr>
<td>Fino et al., 2011</td>
<td>2</td>
<td>Utilized a two-photon photostimulation technique to systematically map connections with single-cell resolution and found nearly full connectivity between somatostatin-positive GABAergic interneurons and pyramidal cells in mouse frontal cortex. Do not speculate as to whether dense connectivity is due to Peters’-rule overlap or due to more selective mechanisms.</td>
</tr>
<tr>
<td>Morgan et al., 2011</td>
<td>1</td>
<td>Upon examining synaptic development of 3 glutamatergic bipolar cell types converging onto a common postsynaptic retinal ganglion cell, found that, during early developmental periods, synaptogenesis may be predicted by neurogeometry. However, with maturation, synaptic specificity was generated and Peter’s rule broke down.</td>
</tr>
<tr>
<td>Ramaswamy et al., 2011</td>
<td>2</td>
<td>Identified axo-dendritic appositions among thick-tufted layer 5 pyramidal neurons packed and modeled in a volume of rat somatosensory cortex and analyzed the resulting microcircuit.</td>
</tr>
<tr>
<td>Ropireddy et al., 2011</td>
<td>2</td>
<td>Digitally embedded 3D reconstructions from NeuroMorpho.Org (Ascoli et al., 2007) and reported on the potential connectivity from various pre-synaptic principal cells and interneurons across rat hippocampal sub-regions onto CA1 pyramidal cell dendrites.</td>
</tr>
<tr>
<td>Czajkowski et al., 2013</td>
<td>2</td>
<td>Found that the density of retrosplenial cortex axonal fibers was linearly proportional to the number of their synaptic contacts on medial entorhinal cortex layer V cell dendrites.</td>
</tr>
<tr>
<td>Helmstaedter et al., 2013</td>
<td>1</td>
<td>In mouse retina, presented the dense reconstruction of 950 neurons and their mutual contacts, as determined by crowd-sourced manual annotation and machine-learning from electron-microscopic data. Stated that the absence of contacts between certain cell types with extensively co-mingled neurites routinely violates Peters’ rule.</td>
</tr>
<tr>
<td>Lin et al., 2013</td>
<td>1</td>
<td>Examined 3D reconstructions and proposed a comprehensive wiring diagram based on the intersections among 194 neuron types in the Drosophila protocerebral bridge.</td>
</tr>
</tbody>
</table>
(Table 4 continues from previous page)

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packer et al., 2013</td>
<td>1</td>
<td>Investigated whether Peters’ rule could explain the non-specific connectivity between somatostatin-positive neurons and parvalbumin-positive neurons to pyramidal cells in neocortex. Determined that the spatial profile of the connectivity maps and even the postsynaptic position of interneuron contacts could result from axo-dendritic overlap of their arbor.</td>
</tr>
<tr>
<td>van Pelt et al., 2013</td>
<td>3</td>
<td>Derived an expression for the expected number of contacts between two neurons based on the overlap of their axonal and dendritic density fields, which accurately estimated the number of synapses between neuron pairs.</td>
</tr>
<tr>
<td>Kim et al., 2014</td>
<td>1</td>
<td>Using crowd-sourced reconstructions, identified contact areas between starburst amacrine cell types and bipolar cell types. Showed that fairly subtle violations of Peters’ Rule may be important for visual function.</td>
</tr>
<tr>
<td>Merchán-Pérez et al., 2014</td>
<td>3</td>
<td>In neuropil samples of layer III of the young rat somatosensory cortex, fully reconstructed ~1,700 synaptic junctions, while determining synapse position and analyzing spatial distributions. Results indicated that the distribution of synaptic junctions was nearly random.</td>
</tr>
<tr>
<td>Potjans et al., 2014</td>
<td>1</td>
<td>Derived a modified version of earlier connectivity maps from (Binzegger et al., 2004) and (Thomson et al., 2002) in order to correct for the excitatory/inhibitory target specificity of certain cell types in a cortical network model.</td>
</tr>
<tr>
<td>Rieubland et al., 2014</td>
<td>2</td>
<td>Investigated the connectivity of the interneuron network in the molecular layer of the cerebellum and used graph theoretic techniques to determine that connectivity among cell pairs is random but connectivity among triplets is structured.</td>
</tr>
<tr>
<td>Kasturi et al., 2015</td>
<td>3</td>
<td>Used Peters’ rule to explore whether redundant synapses form among random appositions between excitatory axons and dendritic spines. Refuted the idea that physical proximity was sufficient to predict the formation of a synapse.</td>
</tr>
<tr>
<td>Krishnaswamy et al., 2015</td>
<td>2</td>
<td>In mouse retina, detected strong and weak connectivity between types (i.e. as determined by optogenetic stimulation) that was partially attributable to the fraction of individual cells connected and was not explainable by the overlap of arbors.</td>
</tr>
<tr>
<td>Reimann et al., 2015</td>
<td>1</td>
<td>Applied Peters’ Rule extensively to predict connectivity from morphology and digitally reconstructed the complete synaptic connectivity between neuron types in a small, well-defined volume of tissue.</td>
</tr>
<tr>
<td>Takemura et al., 2015</td>
<td>3</td>
<td>In the <em>Drosophila</em> visual system, determined that synapse numbers did not correlate with the area of contact between pairs of neurons.</td>
</tr>
<tr>
<td>Wheeler et al., 2015</td>
<td>1</td>
<td>Constructed a knowledge base for neuron types and their properties in the rodent hippocampal formation. In the absence of known connectivity information between types, inferred potential connectivity by identifying axon-dendrite co-localizations in 26 anatomically defined parcels.</td>
</tr>
<tr>
<td>Lee et al., 2016</td>
<td>2 &amp; 3</td>
<td>In the rodent primary visual cortex, found that functionally specific connections do not result from the spatial arrangement of the neuropil and that multiple synapses between pairs of neurons occurred far above chance levels.</td>
</tr>
</tbody>
</table>
**Level 2 – assessment of potential connectivity between individual neurons**

Analogously, Peters’ rule has also been used to unearth potential connections between single neurons. As a baseline, in the cellular-level human brain connectome, the connection density is quite low. Comparing the $10^{15}$ synapses present (Braitenberg and Schüz, 1998) with the number of possible neural connections (i.e. $10^{22}$, the number of neurons squared) results in a probability of roughly 1 synapse for every 10 million neuron pairs; this ratio is even more unbalanced when accounting for multi-synaptic connectivity. Peters’ rule is quite useful here in ruling out the possibility of connectivity between neurons that are spatially remote. However, even granting proximity, the co-location of axons and dendrites is a necessary but insufficient condition for synapsing. As in Level 1, neurogeometrical properties and the selective and dynamic nature of synaptogenesis (Reid, 2012) are key predictive factors. Yet realizing potential connectivity at this resolution is further complicated by the simple fact that fewer neurites (and thus neurite juxtapositions) are involved when dealing with individual cells rather than with neuron types.

This variation of the principle has been applied frequently and, surprisingly often, under the same name as in Level 1. In point of fact, the widespread extension of the rule to individual cell pairs (Fino and Yuste, 2011; Ropireddy and Ascoli, 2011; Ramaswamy et al., 2012), with both positive (Czajkowski et al., 2013) (i.e. cases with high filling fractions (Stepanyants et al., 2002)) and, more frequently, negative results (Hamos et al., 1987; Stepanyants et al., 2002; Shepherd, 2005; Mishchenko et al., 2010; Briggman et al., 2011; Rieubland et al., 2014; Krishnaswamy et al., 2015; Lee et al., 2016) represents an
interesting and even critical pursuit, but it is a conceptual departure from the intended
application of Peters’ rule. (See again Table 4 for details of recent usage at this level.)

Level 3 – assessment of synapse numbers between individual neurons
Peters’ rule has even been invoked with sub-cellular scope, again in cases that
support (Stepanyants et al., 2008; van Pelt and van Ooyen, 2013; Merchán-Pérez et al.,
2014) and refute (Kasthuri et al., 2015; Takemura et al., 2015; Lee et al., 2016) the
predictive accuracy of the hypothesis (details in Table 4). Though this resolution, too,
represents a worthy scientific inquiry, it is another case of a fundamentally distinct
principle masquerading under borrowed terminology. Studies employing the rule in this
way utilize single-neuron reconstructions of three-dimensional axonal and dendritic
arbors in an attempt to correlate the expected number of synapses with the branch
overlaps within a given touch distance. Even if touch distances are allowed to vary
according to the reach of the potential post-synaptic elements (i.e. spines, shafts, gap
junctions, and terminal boutons), the cause of synapse formation is not well understood,
and the occurrence of monosynaptic or multi-synaptic connectivity (or even a lack of
connectivity) between neurons is highly condition-dependent.

Testing the predictive accuracy of Peters’ rule (Level 1) using
Hippocampome.org
Paradoxically, at the originally intended neuron-type level, Peters’ rule remains
largely untested for whole brain regions. We take advantage of the opportunity to
evaluate available data and leverage Hippocampome.org, a freshly compiled open-access
catalogue of all known hippocampal neuron types (Wheeler et al., 2015), annotated from
~14,000 pieces of experimental evidence through a multi-year literature-mining effort.
This resource tracks the axonal and dendritic location of 122 neuron types across 26 sub-regions and layers of the rodent hippocampal formation, including the dentate gyrus (DG), CA1-3, subiculum (SUB), and the entorhinal cortex (EC) (see Hippocampome.org/morphology). Applying Peters’ rule and identifying intra-parcel colocations of neurites for the 122×122 pairs of neuron types in this dataset yields a potential connectome of >3,300 potential synapses. In other words, Peters’ rule is useful in refuting approximately 80% of all-to-all type-level connectivity. For the remainder, nearly 10% of the potential connections have been experimentally examined in peer-reviewed publications and confirmed or refuted (e.g. via electron microscopy or dual electrophysiological recordings). Notably, this data is not restricted to connections to or from principal cells, as 25% of all confirmed potential connections are between interneurons (see Hippocampome.org/connectivity; view evidence for known connections by clicking on the green ‘+’ icons and for refuted connections by clicking on a red ‘x’).

This sample size is sufficient to assess the efficacy of Peters’ rule for predicting connectivity among cell types in the hippocampal formation.

The results from the meta-analysis, overall and broken down by sub-region, are displayed in Table 5. We find that, at least in the hippocampal formation, the original version of Peters’ rule is correct in 73.9% of investigated cases (204 confirmed connections relative to 72 refuted), with the lowest accuracy in CA3 (57.4%) and the highest for those connections that project across sub-region boundaries. Moreover, when discounting the handful of interneuron types with well-known and exceptional targeting specificities, and applying a more nuanced version of Peters’ rule that accounts for the
connectivity preferences of certain cell types, the predictive accuracy jumps to 99% (204 corroborations relative to 2 incorrect cases). Specifically, axo-axonic cells (Li et al., 1992; Bernard and Wheal, 1994; Freund and Buzsáki, 1996; Klausberger and Somogyi; 2008), which have been shown to avoid interneurons and connect exclusively to the axon initial segment of pyramidal cells and granule cells, and interneuron-specific cells (Acsády et al., 1996; Freund and Buzsáki, 1996; Gulyás et al., 1996; Klausberger and Somogyi; 2008), which synapse preferentially with other interneurons, clearly are not subject to Peters’ rule and thus skew the numbers.

<table>
<thead>
<tr>
<th>Connections</th>
<th>HC</th>
<th>DG</th>
<th>CA3</th>
<th>CA2</th>
<th>CA1</th>
<th>SUB</th>
<th>EC</th>
<th>Inter-subregional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential, by Peters' rule</td>
<td>302</td>
<td>263</td>
<td>451</td>
<td>25</td>
<td>1075</td>
<td>8</td>
<td>719</td>
<td>758</td>
</tr>
<tr>
<td>Experimentally confirmed</td>
<td>204</td>
<td>34</td>
<td>27</td>
<td>0</td>
<td>77</td>
<td>4</td>
<td>6</td>
<td>56</td>
</tr>
<tr>
<td>Experimentally refuted</td>
<td>72</td>
<td>11</td>
<td>20</td>
<td>0</td>
<td>35</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>- Axo-axonic</td>
<td>60</td>
<td>11</td>
<td>20</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>- Interneuron-specific</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>- Other</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Current known accuracy of original Peters' rule</td>
<td>73.9%</td>
<td>75.6%</td>
<td>57.4%</td>
<td>Unk.</td>
<td>68.8%</td>
<td>100%</td>
<td>85.7%</td>
<td>91.8%</td>
</tr>
<tr>
<td>Current known accuracy of nuanced Peters' rule</td>
<td>99.0%</td>
<td>100%</td>
<td>100%</td>
<td>Unk.</td>
<td>100%</td>
<td>100%</td>
<td>85.7%</td>
<td>98.2%</td>
</tr>
<tr>
<td>Sample size</td>
<td>8.4%</td>
<td>17.1%</td>
<td>10.4%</td>
<td>0%</td>
<td>10.4%</td>
<td>50.0%</td>
<td>1.0%</td>
<td>8.0%</td>
</tr>
</tbody>
</table>

In full disclosure, it is worth noting that the connectivity literature itself is inclined towards reports of confirmed synapses over refuted synapses. This is due both to a positive-results publication bias and the fact that it is quite difficult to disprove connectivity with existing electrophysiological methodology. Nevertheless, the data at
hand do suggest that most cell types may connect as needed with a broad selection of their potential partners.

**Reconciling opposing views in the community**

Progress in connectomics has been fettered not only by neural complexity and technological hurdles, but also by the lingering confusion surrounding a key, four-decade-old hypothesis. Here, we first provided a primer illuminating the distinct conceptual interpretations of Peters’ rule: (1) at the level of neuron types, (2) between individual neurons, and (3) to estimate the number of synapses between neurons. As we have shown, a large portion of the polarizing discourse on Peters’ rule (Helmstaedter et al., 2013; Packer et al., 2013; Kasthuri et al., 2015; Markram et al., 2006) can be ascribed to confusion and disagreement over the various levels at which to apply it, and many results that seem at first blush to be disparate, may be “explained away” by recognizing and clarifying the multiple notions behind the single name.

Regardless of the application level, however, Table 4 shows that Peters’ rule works in some instances and not in others. Clearly, more nuanced connectivity rules are needed. Reid referred to three abstract orders of specificity (Reid 2012). First, certain connections are *topographic*, or, in our terminology, based on the coincidence of neurite proximity. We have demonstrated using data from Hippocampome.org that axonal-dendritic patterns are quite serviceable as first-order proxies for neuron-type connectivity. Such wiring is prevalent within the brain and may provide the network with the capability for rapid signal propagation or fundamental, local processing unit blocks. Secondly, *cell-type* specific connections (e.g. axo-axonic or interneuron-specific cells in the
hippocampus), in violation of Level 1 of Peters’ rule, have a statistically-biased aversion to certain other types. These connections have highly precise functional ramifications, including direct inhibition or disinhibition of key players within the larger network. Thirdly, some connectivity is likely functionally specific, where synapsing (or the lack thereof) cannot be attributed to the morphology or genetics of neuron type classes but, rather, may arise because of physiological circumstances or the temporal nature of synapse formation.

Thus, as far as we know, in a well-defined set of cases with modest and known corrections, Peters’ rule is quite predictive and imminently testable. As information grows about exceptions to topographic connectivity, we can refine the various tiers of rules for synaptic specificity and improve the accuracy of potential connectivity predictions. Though the original intent of Peters’ rule has been blurred by a variety of applications over the ensuing decades, the time is right to re-clarify and even re-implement a well-defined, more nuanced rule at the mesoscopic resolution. In so doing, reconciling opposing views in the community may pave the path towards deriving a ‘draft’ connectome at the neuron-type level from neuronal morphology, to help guide and interpret forthcoming synaptic data.
CHAPTER THREE

Title: Graph theoretic and motif analyses of the hippocampal neuron type potential connectome

Authors: Christopher L. Rees, Diek W. Wheeler, David J. Hamilton, Charise M. White, Alexander O. Komendantov, Giorgio A. Ascoli

Abstract
We computed the potential connectivity map of all known neuron types in the rodent hippocampal formation by supplementing scantly available synaptic data with spatial distributions of axons and dendrites from the open-access knowledge base Hippocampome.org. The network that results from this endeavor, the broadest and most complete for a mammalian cortical region at the neuron-type level to date, contains more than 3,200 connections among 122 neuron types across six sub-regions. Analyses of these data employing graph theory metrics unveil the fundamental architectural principles of the hippocampal circuit. Globally, we identify a highly specialized topology minimizing communication cost; a modular structure underscoring the prominence of the tri-synaptic loop; a core set of neuron types serving as information processing hubs as well as a distinct group of particular anti-hub neurons; a nested, two-tier rich club managing much
of the network traffic; and an innate resilience to random perturbations. At the local level, we uncover the basic building blocks, or connectivity patterns, that combine to produce complex global functionality, and we benchmark their utilization in the circuit relative to random networks. Taken together, these results provide a comprehensive connectivity profile of the hippocampus, yielding novel insights on its functional operations at the computationally crucial level of neuron types.

**Significance statement**

Brain connectomes are being constructed at two disjointed levels. Microscopically, the wiring of individual neurons is being accumulated into cumbersome synaptomes; macroscopically, region-to-region projectomes obscure important circuit details. Neuron types provide a fertile middle ground. Using the 122 hippocampal formation types from Hippocampome.org, we augmented sparse connectivity knowledge with morphological evidence to obtain a full “potential connectome.” Though this network contains >3,200 connections that are not easily amenable to intuitive hypothesis generation and testing, such complexity can be tackled using graph theory analysis, whereby we investigate the relationship between the circuit’s connectivity properties and functions. As type-level data grows, the array of analyses detailed here can be extended to rapidly supplement our understanding of the computational operation of the hippocampus.
**Introduction**

The rodent hippocampus encompasses millions of neurons (West et al., 1991; Hosseini-Sharifabad and Nyengaard, 2007; Bandeira et al., 2009; Fu et al., 2013), each synapsing with tens of thousands of others (Gulyás et al., 1999; Megías et al., 2001). Examination and quantitative analysis of anatomical connectivity constitute a critical step towards understanding the circuit function (Sporns et al., 2005).

Major community efforts like the BRAIN Initiative (Insel et al., 2013) and the Human Brain Project (Markram et al., 2015) are currently attempting to reconstruct the entire synaptic connectivity of each individual neuron. These undertakings produce massive datasets, but their necessary focus on extremely contained anatomical domains cannot comprehensively reveal long-range circuit architecture (see e.g. Mishchenko et al., 2010; Kasthuri et al., 2015). At the other extreme, approaches such as the Human Connectome Project (van Essen et al., 2013) and the Allen Mouse Brain Connectivity Atlas (Oh et al., 2014) use diffusion tensor imaging or anterograde/retrograde tractography to map brain-wide regional connectivity (see also Mitra, 2014; Zingg et al., 2014). However, the limited spatial resolution and lack of cellular specificity restrict the utility of these data to inform our understanding of neuronal computation.

In between these popular “synaptome” (DeFelipe, 2010) and “projectome” (Kasthuri and Lichtman, 2007) levels lies an arguably more immediately fertile neuron-type circuitry approach that intuitively harmonizes well with Cajal’s “Neuron Doctrine” (Shepherd, 1991). Though neurons are indeed unique cellular units, they may be readily grouped according to sets of properties that cluster along a continuum. Over the past six years, we mounted a massive literature search to catalog all known neuron types in the
rodent hippocampal formation based on their main neurotransmitter, axonal-dendritic morphologies, somatic location, molecular expression, and electrophysiological parameters (Wheeler et al., 2015). All the properties (and underlying experimental evidence) of the resulting 122 neuron types are collated in a publicly available, highly curated knowledge base (Hippocampome.org) that is ripe for analysis along multiple dimensions.

Here, because knowledge about synaptic connectivity among the types is sparse, we fill the gaps by exploiting Peters’ rule (Braitenberg and Schüz, 1998), which recognizes axon-dendrite juxtapositions among the types as potential connections. We then quantitatively examine the resulting complex network using graph theory (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Wig et al., 2011; Binicewicz et al., 2015). Through a suite of analyses, we investigate global degree distribution, circuit modularity, rich club coefficients, absorption and driftness, as well as local motif composition, in order to foster intuition on how the functionality of the hippocampus relates to its fundamental architectural properties (Sporns et al., 2000). We also present an interactive, online, open-source toolbox for exploring the potential neuron-type connectome in the rodent hippocampal formation.

Materials and methods

Identification of neuron types

This work focuses on the rodent (mouse and rat, of either sex) hippocampal formation, defined as the dentate gyrus (DG), CA3, CA2, CA1, subiculum, and entorhinal cortex (EC). Each of these sub-regions is divided in layers (e.g. CA3 oriens,
pyramidale, lucidum, radiatum, and lacunosum-moleculare; or EC L1-L6) giving rise to a total of 26 anatomical parcels. Over a period of several years, we amassed information on hippocampal formation neuron types from the century-deep and information-rich body of literature. However, because neurons are often named on an *ad hoc* basis without full mappings to previous names and descriptors (Hamilton et al., 2016), author-provided names of types were treated warily. Instead, neuron types were identified chiefly based on their primary neurotransmitter (i.e. glutamate or GABA) and for having a unique binary pattern of axonal and dendritic presence or absence across the 26 parcels (see Wheeler et al., 2015 for details). In rare cases (e.g. fast-spiking/parvalbumin-positive and regular-spiking/cholecystokinin-positive basket cells, or ivy and bistratified cells), aligned molecular marker and electrophysiological evidence was sufficiently different to support the creation of two distinct types out of neurons with the same morphological pattern and primary neurotransmitter. Type names were then selected, differentiated, combined, or created anew in order to minimize confusion with the existing literature and fully mapped to their synonyms (see Hamilton et al., 2016 for explanation). The complete set of terms, definitions, data, and supporting experimental evidence collectively underlying the identification of the resulting 122 hippocampal neuron types is publicly available in open access form at Hippocampome.org (RRID: SCR_009023). Table 6 provides a glossary of neuron types to facilitate identification in figures throughout the paper. Types are ordered first by sub-region, then by primary neurotransmitter, then alphabetically. Asterisks indicate types that are either not well known or contain relatively little molecular marker and electrophysiological evidence.
### Table 6 Neuron type glossary

<table>
<thead>
<tr>
<th>Neuron Type</th>
<th>CA3 R-LM</th>
<th>CA1 R-Recv Apical-Targeting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Granule</td>
<td>[1]</td>
<td>[62]</td>
</tr>
<tr>
<td>Hilir Ectopic Granule *</td>
<td>[2]</td>
<td>Schaffer Collateral-Associated</td>
</tr>
<tr>
<td>Semilunar Granule</td>
<td>[3]</td>
<td>SCR R-Targeting *</td>
</tr>
<tr>
<td>Mossy</td>
<td>[4]</td>
<td>CA1 SO-SO (orients-oriens) *</td>
</tr>
<tr>
<td>Mossey MOLDEN *</td>
<td>[5]</td>
<td>CA1 Hip-Subiculum Proj ENK+ *</td>
</tr>
<tr>
<td>AIPRIM (Aspiny int w/ proj. to SM)</td>
<td>[8]</td>
<td>CA1 Trilaminar</td>
</tr>
<tr>
<td>DG Axo-Axonic</td>
<td>[7]</td>
<td>CA1 Radial Trilaminar</td>
</tr>
<tr>
<td>DG Basket</td>
<td>[9]</td>
<td>SUB EC-Projecting Pyramidal</td>
</tr>
<tr>
<td>DG Basket CCK+</td>
<td>[10]</td>
<td>SUB CA1-Projecting Pyramidal</td>
</tr>
<tr>
<td>HIPPP</td>
<td>[12]</td>
<td>CA1 Radiatum Giant</td>
</tr>
<tr>
<td>HiPRom</td>
<td>[13]</td>
<td>CA1 Axo-Axonic</td>
</tr>
<tr>
<td>MOGAP (Molecular Commisural-Associated Pathway related) *</td>
<td>[14]</td>
<td>MEC LIi-III Pyramidal-Fan</td>
</tr>
<tr>
<td>MOLAX</td>
<td>[15]</td>
<td>CA1 Basket CCK+</td>
</tr>
<tr>
<td>MOPP</td>
<td>[16]</td>
<td>CA1 Basket CCK+</td>
</tr>
<tr>
<td>DG Neurogliaform</td>
<td>[17]</td>
<td>CA1 Horizontal Basket</td>
</tr>
<tr>
<td>Outer Molecular Layer *</td>
<td>[18]</td>
<td>CA1 Neuralgiaform</td>
</tr>
<tr>
<td>Total Molecular Layer</td>
<td>[19]</td>
<td>CA1 Int-spec-O-Target QuadD</td>
</tr>
<tr>
<td>CA3 Pyramidal</td>
<td>[20]</td>
<td>CA1 Int-specific LMO-O *</td>
</tr>
<tr>
<td>CA3c Pyramidal</td>
<td>[21]</td>
<td>CA1 Int-specific LM-R</td>
</tr>
<tr>
<td>CA3 Giant</td>
<td>[22]</td>
<td>CA1 Int-specific LMR-R</td>
</tr>
<tr>
<td>CA3 Granule</td>
<td>[23]</td>
<td>CA3a Int-specific O-R *</td>
</tr>
<tr>
<td>CA3 Axo-Axonic</td>
<td>[24]</td>
<td>CA2 Int-specific RO-O *</td>
</tr>
<tr>
<td>CA3 Horizontal Axo-Axonic *</td>
<td>[25]</td>
<td>CA3 Basket CCK+</td>
</tr>
<tr>
<td>CA3 Basket</td>
<td>[26]</td>
<td>CA3 Basket CCK+</td>
</tr>
<tr>
<td>CA3 Basket CCK+</td>
<td>[27]</td>
<td>CA3 Basket CCK+</td>
</tr>
<tr>
<td>CA3 Bistratified</td>
<td>[28]</td>
<td>CA3 Basket CCK+</td>
</tr>
<tr>
<td>CA3 Interneuron-specific Orients</td>
<td>[29]</td>
<td>CA3 Interneuron-specific Quad *</td>
</tr>
<tr>
<td>CA3 Ivy</td>
<td>[30]</td>
<td>CA1 O-LM</td>
</tr>
<tr>
<td>CA3 LMR-Targeting</td>
<td>[31]</td>
<td>CA1 Recurrent O-LM</td>
</tr>
<tr>
<td>Lucidum LAX (lucidum axons)</td>
<td>[32]</td>
<td>CA1 C-LMR</td>
</tr>
<tr>
<td>Lucidum ORAX (orians axons)</td>
<td>[33]</td>
<td>CA1 O-LM</td>
</tr>
<tr>
<td>Lucidum-Radiatum *</td>
<td>[34]</td>
<td>CA1 O-LM</td>
</tr>
<tr>
<td>Spiny Lucidum</td>
<td>[35]</td>
<td>CA1 O-LM</td>
</tr>
<tr>
<td>Mossy Fiber-Associated (MFA)</td>
<td>[36]</td>
<td>CA1 OR-LM *</td>
</tr>
<tr>
<td>MFA ORDEN (orians-dendrites)</td>
<td>[37]</td>
<td>CA1 Perforant Path-Associated</td>
</tr>
<tr>
<td>CA3 O-LM</td>
<td>[38]</td>
<td>CA1 Perforant PathQuadD</td>
</tr>
<tr>
<td>CA3 QuadD-LM</td>
<td>[39]</td>
<td>CA1 Quadrilaminar</td>
</tr>
<tr>
<td>CA3 Radiatum *</td>
<td>[40]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA3 QuadD-LM</td>
<td>[41]</td>
<td>CA1 Quadri-Targeting</td>
</tr>
<tr>
<td>CA3 R-LM</td>
<td>[42]</td>
<td>CA3 R-LM</td>
</tr>
<tr>
<td>CA3 SO-SO (orients-oriens) *</td>
<td>[43]</td>
<td>CA3 Trilaminar</td>
</tr>
<tr>
<td>CA2 Pyramidal</td>
<td>[44]</td>
<td>CA2 SP-SR</td>
</tr>
<tr>
<td>CA1 Trilaminar</td>
<td>[45]</td>
<td>CA1 Pyramidal</td>
</tr>
<tr>
<td>CA1 SO-SO (orients-oriens) *</td>
<td>[46]</td>
<td>CA1 Pyramidal</td>
</tr>
<tr>
<td>CA2 Basket</td>
<td>[47]</td>
<td>CA2 Bistratified</td>
</tr>
<tr>
<td>CA2 Wide-Arbor Basket</td>
<td>[48]</td>
<td>CA2 SP-SR</td>
</tr>
<tr>
<td>CA1 Pyramidal</td>
<td>[49]</td>
<td>CA1 Basket</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[50]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[51]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[52]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[53]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[54]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[55]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[56]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[57]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[58]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[59]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[60]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[61]</td>
<td>CA1 Radiatum</td>
</tr>
<tr>
<td>CA1 Radiatum</td>
<td>[62]</td>
<td>CA1 Radiatum</td>
</tr>
</tbody>
</table>
Culling of known connectivity information

All 484 peer-reviewed literature references comprising v1.0 of the
Hippocampome.org knowledge base were mined in a first-pass attempt to determine
which of the 14,884 (122^2) directed pairs of neuron types are known to synapse, or
known not to synapse. Information verified by various methods (e.g. electron microscopy
or electrophysiological paired recordings) was annotated and relevant quotes and figures
were extracted. Future versions of Hippocampome.org will additionally examine sources
that cite and those that are cited by the original references, as well as search for specific
peer-reviewed articles with neuron-type connectivity information.

Computation of potential connectivity

In the absence of literature evidence for known connections or non-connections,
information on potential connectivity between types was exploited in order to achieve a
full hippocampal connectome (simply referred to as HC in the remainder of the paper).
The co-existence of the axons of one type with the dendrites of another within any
hippocampal formation parcel indicates relative spatial proximity and a potential for
synapsing. The rows in Figure 8A show a subset of neuron types and their defining axo-
dendritic patterns (type names in bold are glutamatergic and in gray are GABAergic; red
boxes with horizontal lines correspond to axons; blue with vertical lines to dendrites;
purple with horizontal and vertical lines to both axons and dendrites; black circles
indicate soma locations). Potential connections of Granule cells are spotlighted with red
arrows). For example, the axons of Granule cells are present in the DG hilus (H), CA3
stratum lucidum (SL) and stratum pyramidale (SP), and CA2 SP. Therefore, any neuron
type with dendrites in any one or more of these parcels, including Mossy-Fiber-
Associated Oriens-Dendrite (MFA ORDEN) cells (type 37 in Table 6; Fig. 8B left), is a
potential target of the Granule cell axons (Fig. 8B right; for both neurons: axons in red;
dendrites in black; morphological reconstruction of the granule cell downloaded from
NeuroMorpho.Org (Ascoli et al., 2007), with layers drawn in, from a tracing originally presented in Bausch et al. (2006); permission to reprint the MFA ORDEN cell (Vida and Frotscher, 2000) granted by PNAS (Copyright 2000 National Academy of Sciences, U.S.A.)). Further, types that do not have dendrites in the aforementioned parcels are excluded as potential Granule cell targets because of the lack of neurite overlap. This approach was extended to account for axo-somatic and axo-axonic connections of basket and chandelier cells, respectively. Mathematically, 26-dimensional binary vectors were utilized to encode the presence or absence across hippocampal parcels of the axon of each potential pre-synaptic type and of the dendrites (or soma or axonal initial segment) of each potential post-synaptic type. Potential connectivity was then calculated (Equation 1) as the dot-product of these vectors: a non-zero result indicated a potential connection whereas a zero-value dot-product denoted that connectivity was not possible between the types in question.

\[
\text{Equation 1 Computation of potential connectivity}
\]

Axons of type \( \vec{A} = (a_1, a_2, ..., a_{26}) \);
Dendrites of type \( \vec{D} = (d_1, d_2, ..., d_{26}) \);
\[
c = \vec{A} \cdot \vec{D}
\]

Many neuron types (101/122, not including Granule cells) have axons and dendrites co-located within one or more parcels, indicating the potential for within-type connectivity; such self-connections are not necessarily indications of single-neuron autapses.
A web-accessible resource for connectome visualization
A Java-based online toolbox was developed and deployed
(Hippocampome.org/connectivity; click “Launch” button or link to “Potential Connectivity Map”) to assist in the visualization and exploration of the HC.

Glutamatergic (excitatory) and GABAergic (inhibitory) neuron types are represented as black and gray circles, respectively, and are placed randomly within the parcel (or along the parcel boundary) where their soma is most commonly located. Hovering over a type reveals its name; clicking on a type displays all of the connections that may be received by its dendrites (lines with arrows in) or sent by its axons (arrows out). A snapshot of the toolbox, taken after the selection ofGranule cells, is shown in Fig. 8C. Toggles provide the ability to show or hide additional information, including connections made by the other (i.e. unselected) types and schematic illustrations of many of the major cell types and pathways in the hippocampal circuit.

Graph theory analyses
The Brain Connectivity Toolbox (BCT: Rubinov and Sporns, 2010; brain-connectivity-toolbox.net) was utilized to compute many graph theory measures for the HC. In certain cases, the MATLAB code was modified slightly to allow (or correct) for the possibility of self-connections of neuron types along the main diagonal of the connectivity matrix. In addition, some topological metrics were measured on the unweighted network while others mandated connections weighted by the sign of the primary neurotransmitter of the pre-synaptic type: +1 for glutamatergic and -1 for GABAergic. To study the robustness of our results, we also examined a version of the
network wherein connections of the most numerous principal cell (PC) types, namely DG Granule cells, CA3 Pyramidal cells, and CA1 Pyramidal cells, were weighted as +10.

**Clustering coefficient, characteristic path length, and degree**

Certain standard measures, including clustering coefficient (CC) and characteristic path length (CPL), are used to encapsulate the topology of the graph and are thus computed on a static, binary version of the network that disregards excitation and inhibition. Briefly, clustering coefficient is the fraction of connections among the immediate neighbors of a node (i.e. the set of neuron types that may receive information directly from that node) relative to the number of possible connections (Fagiolo, 2007). For example, Granule cells have 33 immediate neighbors that are interconnected with 476 (out of a possible $33^2=1089$) edges; $CC_{\text{Granule}} = \frac{476}{1089} = 0.437$. This quantity, computed for each node, is then averaged over all neuron types to yield a single global value, $CC_{\text{HC}}$.

Characteristic path length is defined as the mean of the shortest directed (i.e. axon to dendrite) path from a node to every other neuron type in the network. For example, Granule cells and CA1 Pyramidal cells are not in direct contact, so communication requires at least one intermediary; in fact, there are five two-step pathways (via types [19], [20], [44], [46], or [47]). Determining analogous distances from Granule cells to the other types in the network and averaging gives $CPL_{\text{Granule}}=2.11$, meaning that they can send information anywhere in the network in an average of just over two steps. Then, averaging this quantity over all 122 types yields a single global value, $CPL_{\text{HC}}$.

Mathematically, Equation 2 gives:
where $n$ is the number of nodes in the network, $i$ is the set of presynaptic types, and $j$ is the set of postsynaptic types. Neuron types that have axons and dendrites co-located within at least one parcel are self-connected and have a shortest path length to themselves of zero (e.g. the shortest path from CA1 Pyramidal cells to CA1 Pyramidal cells is 0); non-self-connected types require multi-step paths to communicate with themselves (e.g. traveling from Granule cells to Granule cells requires two steps).

Node degree is the number of connections made by a node (out-degree; OD), to a node (in-degree; ID), or the sum of these quantities (total-degree, also called degree centrality; TD). Again, Granule cells have 33 immediate neighbors (OD\text{Granule}=33), and they are immediate neighbors to 26 other types (ID\text{Granule}=26); TD\text{Granule}=33+26=59. Self-connected neuron types thus contribute two connections to their total-degree. A related measure, polarity, is defined as (ID–OD)/TD (Shih et al., 2015).

**Topology comparison analysis**

For six well-known network types, we generated 1,000 random networks identical in size to the HC and compared their CC and CPL. The two metrics were then combined to measure the overall (i.e. global and local) “communication cost.” Specifically, the cost was computed as follows (Equation 3):
Equation 3 Communication cost

\[ \text{Communication cost} = -\log_{10}(CC) + \log_{10}(CPL) \]

For each network type, the resulting cost was linearly scaled so that the reference network (HC) was given unitary value.

The algorithms to produce the Erdős-Rényi (ER), lattice, ring, Watts-Strogatz (WS), Barabási-Albert (BA), and Klemm-Eguílez (KE) networks were also implemented in MATLAB (open-source code: [github.com/Hippocampome-Org/GraphTheory](https://github.com/Hippocampome-Org/GraphTheory)) using published pseudo-code (Prettejohn et al., 2011). Briefly, an ER network (Erdős and Rényi, 1960) is constrained only by its number of nodes and its connection density; we used HC network values of 122 and 21.7%. These graphs were constructed by considering all possible connections among the nodes and inserting them with probability equal to the connection density. A square lattice network, in contrast, is heavily constrained by the number of nodes, edges, and the fact that each node must be connected to its K nearest neighbors (where K is the ratio between HC edges and nodes: 3,236/122=26.5). A ring network, a one-dimensional string of nodes “bent” into circular form by joining the ends, is similarly constrained. Starting from a highly clustered ring graph, WS (Watts and Strogatz, 1998) networks were created by considering each connection for random rewiring with constant probability (\( p_{\text{rewiring}} = 0.4 \)) in order to introduce long-distance (i.e. cross-network) edges. For BA scale-free networks (Barabási and Albert, 1999), we started from an initial size of ten fully connected nodes and serially attached the remaining 112 nodes to pre-existing nodes chosen with probability proportional to their OD in the growing network. This preferential addition of new nodes
to higher-degree nodes yields the desired power law distribution for the final network degree, with the vast majority of nodes having very small OD and a select few types having large OD. Finally, KE networks (Klemm and Eguiluz, 2002) are generated to obtain high CC and low CPL (like WS networks), along with a scale-free OD distribution (like BA networks). The algorithm is similar to that used for BA networks, but attachment of new nodes is preferentially biased toward high-degree, highly clustered “active” nodes (Prettejohn et al., 2011).

Modularity
The HC modular, or community, structure was bared by computationally assigning neuron types into non-overlapping groups to maximize within-community connectivity and minimize extra-modular cabling. Community assignments are evaluated by a modularity score, Q, which quantifies the fraction of connections in a module relative to those expected by chance (Newman, 2004). Practically, we use BCT code based on a spectral algorithm that optimizes Q over possible HC divisions (Leicht and Newman, 2008). The algorithm was run 100 times and the detected communities did not change.

Rich club analysis
Rich club (RC) analysis utilized a modified version of BCT code to identify cores of nodes that are more highly connected to each other than expected by chance (Zhou and Mondragón, 2004; Colizza et al., 2006; McAuley et al., 2007; van den Heuvel and Sporns, 2011). First, a connectivity fraction (C_t) is computed for each degree level k from 1 to the maximum TD in the network (i.e. 114, for CA3c Pyramidal cells) as the proportion of edges that connect nodes of degree>\(k\) relative to the maximum number of
edges that such nodes might share (Colizza et al., 2006). These $C_f$ values are then normalized relative to the average for a given $k$ in a population of 1,000 random networks synthetically generated to have fixed OD and ID distributions matching the HC. Raw $p$ values were calculated at each $k$ based on the $C_f$ percentile rank of HC within the population of 1,000 random networks. Normalized $C_f$ values that were significantly greater than 1 over a range of $k$’s with $p$ values smaller than 0.05 after “false discovery rate” (FDR) multiple testing correction (Storey, 2002) were designated as members of RC tier I. The cutoff for inclusion in RC tier II was selected based on the relatively large $C_f$ increase in HC between $k=77$ ($C_f=0.630$) and $k=78$ ($C_f=0.766$).

Absorption and driftness analyses

Shortest path lengths between neuron types were again measured using BCT. The number of paths of length $y$ between all pairs of types may be found simultaneously by multiplying the unweighted connectivity matrix, $M$, by itself $y$ times (e.g. the matrix entries obtained by calculating $M^3$ are the number of paths of exactly three steps between types). Absorption and driftness (Costa et al., 2011) values were also computed in MATLAB. The absorption metric simulates average random walks as a surrogate for dynamic activity in the network. In a given random walk from a starting neuron type to a destination, the walk progresses with equal probability to any of the connected types and continues until the target is reached. Averaging a large number of independent random walks mimics parallel propagation of activity over all possible paths connecting two neuron types. Driftness is calculated as the absorption value divided by the CPL for each pair of neuron types (Costa et al., 2011).
**Connectivity superpattern and pattern profiles**
At a local level, we investigated the configurations of connectivity (or lack thereof) for all groupings of three neuron types. In a circuit with 122 elements chosen three at a time without regard to order, this equates to a total of 295,240 combinatorial relationships. In one analysis, we examined connectivity “superpatterns” without distinguishing excitatory and inhibitory connections (i.e. the network was considered directed but unweighted); in a second, we studied the directed and weighted network “patterns.” For the sake of interpretability, self-connections among types were not considered as differentiators in this analysis. Superpattern and pattern libraries and detection algorithms were built from scratch using MATLAB (github.com/Hippocampome-Org/GraphTheory).

Excitability scores (ES) quantify the net counterbalance of excitation versus inhibition occurring within a triad of neuron types. These scores are computed at each node in the pattern, then summed over the three nodes. If a node does not receive a connection from either of the two other nodes, its score is equal to its sign (+1 if excitatory node, -1 if inhibitory): this node is not amplified or dampened by the rest of the pattern. If a node receives a connection from one or both other node(s) in the pattern, its score equals its sign multiplied by 1.1 for each incoming excitatory connection, and by 0.9 for each incoming inhibitory connection. Explicit examples of this computation are included in the Results.

Patterns may or may not have a unique ES. Thus, for each ES, we also quantified the relative prevalence within the detected modules, so as to determine whether the underlying communities tended to utilize repeatedly certain sets of excitatory/inhibitory
configurations. The relative importance of these interactions to a module is computed based on the number of times an ES appears within that module relative to the overall network.

*Detection of motifs and anti-motifs*

Analogously to the rich club analysis, counts of HC superpatterns and patterns were compared to a population of 1,000 random networks to find those that were significantly over- or under-utilized relative to expectancy, respectively called motifs and anti-motifs (Milo et al., 2002; Milo et al., 2004). These random networks were generated in parallel by selective edge swaps chosen stringently and conservatively so as to maintain the underlying spectrum of two-node (i.e. dimer) superpatterns and patterns. Specifically, the random networks preserved the HC number of excitatory (E) and inhibitory (I) nodes and connections; the number of E to E, E to I, I to E, and I to I connections; and the OD and ID of each node. Accordingly, each edge in the graph had a limited number of valid swap partners from which a suitable mate was randomly chosen. Fifty swapping passes were made over all edges in order to sufficiently scramble the original network.

The statistics employed for the motif/anti-motif analysis were similar to the RC analysis: raw p values for each pattern were based on the percentile rank of the HC count within the random population. Patterns with percentile ranks >95 (meaning that the pattern appeared in the HC more than in 95% of the 1,000 random networks) underwent multiple testing corrections to determine whether they constituted statistically significant motifs. Similarly, patterns with percentile ranks <5 underwent testing to identify anti-motifs. Adjusted p-values were calculated with the step-down “min P” procedure.
(Westfall and Young, 1993), and patterns with corrected values < 0.05 were deemed significant.

**Pairwise correlation analysis**

Pairwise correlation were evaluated among 315 properties across the 88 neuron types in the DG, CA3, CA2, and CA1 sub-regions. In addition to connectivity properties detailed herein (degree, strength, polarity, and usage of superpatterns/patterns), we examined morphological features (e.g. somatic, axonal, and dendritic locations, as well as the projecting or local nature of axons), molecular markers (e.g. expression, or lack thereof, of various calcium-binding proteins, neuropeptides, or receptors), and assorted passive and spiking electrophysiological parameters (e.g. input resistance, fast or slow membrane time constants, action potential width). Neuron types from subiculum and EC were excluded from this analysis due to the scarcity of available molecular and electrophysiological information. Direct and inverse relationships between properties were detected using 2x2 contingency matrices, and p-values were calculated with Barnard’s exact test (Lydersen et al., 2009).

**Results**

In building the neuron type connectome of the hippocampal formation, we extracted information for 167 known connections and 68 non-connections from the literature. For the remaining 14,649 type pairs (i.e. $122^2 - 235$ known connections or non-connections), we calculated the spatially-based potential connectivity, which excluded the possibility of connections for 11,580 pairs of neuron types. Consequently, 3,069 potential connections were combined with 167 known connections to obtain the HC network explored here: a graph of 38 excitatory and 84 inhibitory neuron types (nodes).
interlinked by 3,236 edges (1,216 excitatory and 2,020 inhibitory; full connectivity data may be downloaded from Hippocampome.org/netlist).

**A highly specialized topology**

We first compared the HC clustering coefficient and characteristic path length to those of six identically-sized, well-known network types (ER, BA, WS, KE, rings, and lattices). Because of the relatively small size of the graphs, CC and CPL showed little variance over the 1,000 randomly generated variants of each network type (Fig. 9A; standard deviations illustrated by the diameter of the data points). CC is indicative of the tendency of nodes to gather in tightly knit groups that may correspond to functional processing units, while CPL reveals the relative expanse of the network. Together, these metrics characterize network topology in terms of communication cost: large-world networks (Boccaletti et al., 2006) contain densely connected groupings of nearby nodes, but remote nodes are only reachable by paths with many steps (dark green background shading in Fig. 9A). At the opposite extreme, uniform random networks have low CC and CPL because of the arbitrary placement of their edges (dark gray shading). Scale-free networks (gold shading) and small-world networks (blue shading) represent two popular mixed cases, with low CC/high CPL and high CC/low CPL respectively. HC displays both high CC, analogous to rings and lattices, and low CPL comparable to ER random networks. This suggests that hippocampal neuron types rapidly combine information across short path lengths into targeted areas, where specialized processing occur within tightly interconnected circuits. In fact, not only is HC classifiable as a small-world network, but its combined global and local communication cost is lower than any of the
other tested networks (Fig. 9B). Moreover, when PC connections were weighted ten times more heavily than other edges, both CPL and the overall communication cost further decreased by 25%.

Figure 9 Topological comparison of the HC to well-known random network types

**Significant community structure**

The organization of HC connectivity can be visually inspected on a circular graph (Fig. 10A; chord diagram produced with Circos software: Krzywinski et al., 2009). In the graph, types are identifiable both by numbers in brackets and the corresponding names provided in Table 6 and by axon-dendrite patterning within the sub-region of their soma location (colored box convention and layer ordering from inside-out as in Fig. 8A); shaded bars in the innermost ring show the total number of signed connections made by
Figure 10 Modular structure of the potential hippocampal connectome
that type; excitatory types have outward-facing black bars and inhibitory types inward-facing gray bars. The innate community structure is identified by grouping the 122 neuron types to maximize intra-modular wiring and minimize inter-modular wiring. The modularity score Q measures the effectiveness of the resulting grouping, with values of Q=0 indicating randomness (i.e. the groupings are equally good, or poor) and, in practice, Q>0.3 pointing to noteworthy community structure (Newman, 2004). The HC network is optimally subdivided into four modules with Q=0.53. Connections between neuron types within one of these communities account for 81% (2,622/3,236) of all graph edges, and the average connection density of the four modules is 0.675, dwarfing the between-module connection density of 0.041 (614/3,236). Interestingly, the communities do not themselves partition into smaller submodules as the average Q score for each of the modules is 0.09 (Fig. 10B). Thus, the four detected communities are the major, high-level processing units of the network.

Even though axons of neuron types frequently cross sub-region boundaries to form connections (i.e. 33/122 types project to different sub-regions from their soma location), the detected communities closely aligned with DG, CA3, CA1, and EC (the first three are shown in Fig. 10C), the most-highly-studied sub-regions of the hippocampal formation. These sub-regions are also the major players in the trisynaptic loop (TSL) relay (highlighted as thick, brightly colored chords in Fig. 10A; perforant pathway lines: dark green; temporoammonic path: light green; mossy fibers: red; Schaffer collaterals: blue; projection from CA1 to EC layer V: orange; other connections are colored randomly to optimize visibility). The “DG module” identified by this analysis
contained all 18 DG types, along with one of the CA1 types that projects to DG (i.e. CA1 Neurogliaform Projecting). The “CA3 module” included all 25 types from CA3 and four out of five types from less-researched CA2. The exception, CA2 Bistratified cells, belonged to the “CA1 module,” along with the remaining 39 CA1 types and SUB CA1-Projecting Pyramidal cells. Finally, the “EC module” contained all 31 EC types and the other two subicular neuron types. Notably, this core modular structure was revealed even without differentially weighting the PC connections.

**Degree distribution and hubs**

The numbers of connections made and received by a neuron type respectively correspond to out- and in-degrees (Fig. 11A), and types with a relatively high total-degree may be generally considered network hubs (van den Heuvel and Sporns, 2013). The HC OD distribution (thin red columns) has both a more asymmetric and a more heavily tailed spread than the ID distribution (thick blue columns), as quantified respectively by skewness and kurtosis values. While both distributions have positive skewness, indicating a right-shifted distribution attributable to the presence of network hubs, the skewness of the OD distribution is more than four times larger. More strikingly, the OD kurtosis (2.87) denotes a much heavier tail than is found with the in-degree distribution, whose negative value close to zero indicates near-normal, if not thinner-than-normal, tails (Pearson, 1905; Westfall, 2014).

Altogether, this evidence points to an axonal architecture that is both anomalous and nonrandom in contrast to a relatively ordinary dendritic architecture. The peculiarity in the axonal distribution is accentuated by breaking down the data by neuron types that
project to another sub-region (Hippocampome.org/morphology) versus those with only local axons (Fig. 11B). The projecting types (thick green columns; n=33) show a heavy, right tail versus the light, left tail of the local types (thin gray columns; n=89). These tails indicate, separately, the presence of neuron types that serve as highly connected hubs and types that are decidedly particular in the connections they form.

Figure 11 Degree distribution analysis to isolate types with unusual connectivity
The top and bottom neuron types by TD may be respectively considered global hubs and anti-hubs (Table 7; type names in black are excitatory; gray inhibitory). The list signals the importance of the CA3 module and highlights its central role in the TSL circuit: pyramidal cells from CA3, CA3c, and CA2 are three of the four most connected neuron types in the network. The other, CA1 Back-Projection cells, is an interneuron type in CA1 with axons that project upstream to CA3 (and DG), opposite to the TSL flow. Notably, Granule cells are not a global hub based on pure topology (i.e. they are only the 43rd most connected neuron type), but they become the third most critical type in the 10x PC weighted network. Thus, Granule cells do not influence a large number of neuron types; rather, their importance in the HC is largely due to their relative abundance.

<table>
<thead>
<tr>
<th>Module</th>
<th>Neuron type</th>
<th>Out-deg</th>
<th>In-deg</th>
<th>Total-deg</th>
<th>Polarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA3</td>
<td>CA3c Pyramidal</td>
<td>84</td>
<td>30</td>
<td>114</td>
<td>-0.47</td>
</tr>
<tr>
<td>CA1</td>
<td>CA1 Back-Projection</td>
<td>79</td>
<td>25</td>
<td>104</td>
<td>-0.52</td>
</tr>
<tr>
<td>CA3</td>
<td>CA3 Pyramidal</td>
<td>65</td>
<td>33</td>
<td>98</td>
<td>-0.33</td>
</tr>
<tr>
<td>CA3</td>
<td>CA2 Pyramidal</td>
<td>79</td>
<td>9</td>
<td>88</td>
<td>-0.80</td>
</tr>
<tr>
<td>CA1</td>
<td>CA1 Pyramidal</td>
<td>41</td>
<td>46</td>
<td>87</td>
<td>0.06</td>
</tr>
<tr>
<td>CA1</td>
<td>CA1 Quadrilaminar</td>
<td>41</td>
<td>43</td>
<td>84</td>
<td>0.02</td>
</tr>
<tr>
<td>CA1</td>
<td>CA1 Radial Trilaminar</td>
<td>37</td>
<td>47</td>
<td>84</td>
<td>0.12</td>
</tr>
<tr>
<td>CA1</td>
<td>CA1 LMR Projecting</td>
<td>45</td>
<td>35</td>
<td>80</td>
<td>-0.13</td>
</tr>
<tr>
<td>EC</td>
<td>MEC LV Pyramidal</td>
<td>51</td>
<td>27</td>
<td>78</td>
<td>-0.31</td>
</tr>
<tr>
<td>CA3</td>
<td>CA3 Trilaminar</td>
<td>62</td>
<td>13</td>
<td>75</td>
<td>-0.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Module</th>
<th>Neuron type</th>
<th>Out-deg</th>
<th>In-deg</th>
<th>Total-deg</th>
<th>Polarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA1</td>
<td>SUB Axo-axonic</td>
<td>2</td>
<td>15</td>
<td>17</td>
<td>0.76</td>
</tr>
<tr>
<td>CA3</td>
<td>CA3 Horizontal AA</td>
<td>2</td>
<td>13</td>
<td>15</td>
<td>0.73</td>
</tr>
<tr>
<td>CA3</td>
<td>CA2 Basket</td>
<td>5</td>
<td>9</td>
<td>14</td>
<td>0.29</td>
</tr>
<tr>
<td>CA3</td>
<td>CA2 SP-SR</td>
<td>5</td>
<td>8</td>
<td>13</td>
<td>0.23</td>
</tr>
<tr>
<td>DG</td>
<td>MOLAX</td>
<td>17</td>
<td>32</td>
<td>49</td>
<td>0.31</td>
</tr>
<tr>
<td>CA3</td>
<td>CA3 Bistratified</td>
<td>26</td>
<td>28</td>
<td>54</td>
<td>0.04</td>
</tr>
<tr>
<td>CA3</td>
<td>CA3 Ivy</td>
<td>26</td>
<td>25</td>
<td>51</td>
<td>-0.02</td>
</tr>
<tr>
<td>CA1</td>
<td>CA1 Oriens-Bistratified</td>
<td>36</td>
<td>24</td>
<td>60</td>
<td>-0.20</td>
</tr>
</tbody>
</table>
Certain neuron types are global hubs due to a high OD even with a low ID (i.e. they have a dominant axonal architecture); others are more balanced. The top four global hubs have OD>>ID (i.e. a large, negative polarity) and constitute network divergence points. Global anti-hubs, with positive polarities, like axo-axonic cells, basket cells, and interneuron specific cells, indicate selected targets of information convergence. In contrast, provincial hubs, which by definition do not project outside of their module, are highly connected within the module and serve as critical traffic directors (van den Heuvel and Sporns, 2013). The top provincial hubs in the DG, CA3, and CA1 modules (bottom portion of Table 7) are again not restricted to a certain polarity. In DG, MOLAX cells tend to funnel information to specific points within the DG, but CA1 Oriens-Bistratified cells distribute information widely to 36 out of 41 CA1-module types; CA3 Bistratified and Ivy cells have relatively balanced, neutral polarities.

Rich and ultra-rich clubs
Rich club analysis showed that the global hubs are significantly more connected amongst each other than could be expected by chance. In fact, beyond just the top global hubs, all nodes with TD≥55 (see Fig. 12A top), have statistically higher interconnectivity than in equivalent random networks (Fig. 12A; normalized data>1). Although nearly half (56/122; 45.9%) of the neuron types belong to this RC I (both light- and dark-purple shaded regions of Fig. 12A), the top 8 hubs from Table 7 are also members of a tighter “richest of the rich” club within (RC II; dark-purple shading only), boasting the highest edge density of 76% (Fig. 12B). Members of RC I consisted of types from all four modules (Fig. 12C), but a disproportionate number (42/56) came from CA1 and EC, as
these modules are densely connected and have the most nodes. Fig. 12D shows the 56 neuron types of RC I (top) and the subset constituting RC II (bottom). Interestingly, all RC II types were located in CA3 and CA1.

**Figure 12** Nested rich clubs within the HC

**Robustness to random failures**

Any two neuron types within a network can typically connect through multiple alternative pathways, providing redundancy for information flow. The maximum shortest
pathway length in HC, five steps, was only found between certain DG and EC types, a route that requires upstream travel against TSL current (Fig. 13A; see Table 6 for type names and ordering; color gradient key: yellow=direct connection; orange, red, and dark red indicate 2, 3, and 4 steps, respectively; black=highest path length: 5 steps).

Approximately two-thirds of all pairs of neuron types are connected by two steps or fewer, and nearly 95% by three (Fig. 13B; orange labels indicate the percentage of type pairs that can be bridged by a path of a given length, k, shown for k≤5). Moreover, increasing the length of possible paths by just a single step raises the number of available alternates by successive orders of magnitude (Fig. 13B; blue columns show the average number of available conduits across all pairs of types). For example, there are on average 5.81 available two-step pathways in HC between two neuron types. While certain pairs have no such pathways (e.g. from DG to EC), others have many possible two-step routes at their disposal (Fig. 13C).

The absorption value for a pair of neuron types is the average length of all paths (the length of a “random walk”) between them (Fig. 13D; color intensity gradient in panel A). The overall average absorption of HC is 230.5. While the out-absorption vectors (i.e. the rows in the matrix) are relatively similar across all nodes, the in-absorption vectors (columns) are highly specific to a given node. Therefore, neuron types are activated with just as much ease or difficulty from any part of the network, depending primarily on the dendritic architecture of the type and its close neighbors. CA2 and subiculum, being hard to reach with few connections arriving at a select few types, and EC, due to overall unidirectional information flow through the TSL, have high in-
absorptions. Driftness corresponds to the absorption normalized by the shortest path lengths between the types (Fig. 13E; color intensity gradient in panel A). Intra-CA2 and intra-EC values are again high, but those for most other type pairs are low, indicating the existence of multiple pathways of similar scale to the shortest path. This feature suggests that the HC network could continue to operate at near-optimal levels after insult to random neuron types and connections. Notably, however, while the absorption values are unchanged when accounting for 10x PC weighting, the overall average driftness increases from 111.8 to 170.0. Thus, the “shortest” path between many types (i.e. through the PCs) is both unique and irreplaceable, and the PCs represent points of vulnerability.

Figure 13 Alternate pathways between types afford resilience to the network
**Characteristic connectivity superpattern profile**

To examine the local interactions of neuron types, we consider all 16 possible ways in which three unweighted nodes can interrelate: 13 where all three nodes directly participate in at least one connection (Fig. 14A, graphs labeled A through M) and 3 cases with at least one node disconnected from the others (labeled -A, -B, and -C). The frequency of occurrence for these 16 building blocks constitutes the HC connectivity superpattern profile (Fig. 14B). An absolute majority of the 295,240 trimer occurrences in HC (248,118 or 84%) involve at least one disconnected type. Of the fully connected trimers, the superpattern consisting of a single uplinked mutual dyad (superpattern F; name modified from Milo et al., 2002) and the single input module (superpattern C; named after Zaslaver et al., 2002; Alon, 2007) represent nearly a collective one third (15,443/47,122; 32.8%). Most interactions take place intermodularly (Fig. 14C) involving projecting neuron types. In both F and C, a single node disperses information to two neuron types that do not interact with each other, with superpattern F receiving direct feedback from one partner and C containing no feedback from either partner. In contrast, superpattern I, a similar structure containing a chain of two mutual dyads, wherein both receivers of the dispersed signal provide direct feedback, is one of the least frequent. These observations point to a strong, net-unidirectional information flow between modules, often in the TSL direction. Furthermore, as the internal connection densities of trimer superpatterns increase (i.e. across vertical, black dotted lines in Fig. 14C), with the exception of F and C, intramodular utilization also increases, underscoring the importance of signal fine-tuning for local microcircuit interactions.
Next we analyzed the breakdown of superpatterns in the CA1 module as utilized by Pyramidal cells versus interneurons (Fig. 14D, blue line). Though many CA1 interneurons interact primarily with pyramidal neurons, many others, including calretinin-positive (CR+) (Gulyás et al., 1996) and vasoactive intestinal peptide-positive (VIP+) (Acsády et al., 1996) cells do not, and the interactions of many other interneuron types are unknown. Pyramidal cells dominate the usage of all superpatterns except C, E (a feedforward loop), and K (a double uplinked mutual dyad). Interestingly, in CA1, such
diminution is specifically attributable to an elevated employment of these trimers by perisomatic-targeting (PST) interneurons, namely CA1 Axo-axonic cells, Horizontal Axo-axonic cells, Basket cells, Basket CCK+ cells, and Horizontal Basket cells (Fig. 14D, red dotted line).

**Weighted pattern profile and neuron type fingerprint analysis**

The number of distinct connectivity trimers grows substantially when distinguishing excitatory and inhibitory nodes. After considering rotational symmetry, 104 possible interaction patterns exist between three nodes: 86 fully connected and 18 with at least one disconnected node. For example, eight connectivity patterns correspond to superpattern F, the single uplinked mutual dyad (Fig. 15A; black lines and nodes are excitatory, gray lines and nodes are inhibitory, blue lines indicate reciprocal connections that are excitatory in one direction and inhibitory in the other). The excitability score ES captures the overall excitatory or inhibitory nature of a pattern by accounting for the net amplification and dampening of each node by their connected partners. Consider, for instance, pattern F4: a mutually interacting pair of excitatory and inhibitory types, with the former being “uplinked” to a second inhibitory type. Every node in F4 receives a connection from another node and is scored independently. The score of the inhibitory node on the right of the pattern (when rotated as in Fig. 15A) is equal to -1.1 because the original value, -1, is amplified (i.e. multiplied) by an inbound excitatory signal. The score of the node at the bottom of the pattern, another inhibitory type that receives excitation, is identical. At the top of the pattern, the original value of this excitatory type (+1) is dampened by an inbound inhibitory edge from the bottom node, yielding a value of 0.9.
Summing the scores from all three types gives an overall $ES = (-1.1) + (-1.1) + (0.9) = -1.3$ for this pattern. ES values for the eight patterns of superpattern F are shown in boxes in Fig. 15A.

Figure 15 Weighted trimers analysis based on E/I neuron type distinction

The ES distribution for all 104 trimer patterns has four narrow peaks determined by the number of excitatory or inhibitory nodes therein; accordingly, trimer neuron type patterns may be categorized as strongly inhibitory, mildly inhibitory (e.g. pattern F4,
which has two inhibitory types), mildly excitatory, and strongly excitatory. The abundance of inhibitory types over excitatory types in the DG, CA3, and CA1 modules yields higher usages of strongly and mildly inhibitory patterns, consistent with the known diversity of GABAergic interneurons (Fig. 15B). Two patterns, L8 and F4, are particularly prominent in certain modules. L8, the all-inhibitory version of superpattern L (refer back to Fig. 14A) consisting of a feedback loop between two mutual dyads, is used ubiquitously in CA1 to fine-tune inhibition. In contrast, the pervasiveness of F4 in CA3 (and, in fact, the whole HC) is almost entirely due to CA3 Pyramidal cells (Fig. 15C), which heavily employ this pattern to disperse information simultaneously to both CA3 interneurons (many of which supply direct feedback) and CA1 interneurons via the Schaffer collaterals (no direct feedback). In fact, pattern ‘fingerprint’ profiling reveals that F4 is by far the dominant class of interactions for CA3 Pyramidal cells (Fig. 15D, brown data series; note log radial scale). The overall HC usage of all 86 connected patterns (blue shading) is shown for comparison.

Moreover, HC conspicuously underutilizes superpattern G, along with each of the patterns G1-G4 (45 total interactions out of 295,240). Indeed, these patterns, corresponding to a unidirectional feedback loop with no reciprocal connections, are avoided in favor of other structural blocks. When two unidirectional connections transmit signals forward along a chain in HC, nodes 1 and 3 are rarely connected by unidirectional feedback as in G (0.5% of such cases); instead, they are either unconnected in a three-node chain (superpattern B; 38.5%), connected by a feedforward link (superpattern E; 49.8%), or connected by a reciprocal edge that serves both feedback and feedforward
purposes (superpattern J; 11.2%). This strikingly uneven distribution is consistent with forward-directional circuitry where feedback tends to be relatively immediate and curbed (e.g. in the form of a reciprocal edge) or else drawn out over a more global scale. Comparing the counts and circuit locations of feedforward and feedback loops (superpatterns E and G, respectively), along with specific examples and tabulated interpretations, clearly illustrates this trend (Fig. 16). Note that, while only four feedback patterns exist due to rotational symmetry, there are 8 feedforward patterns. Black dots and arrows indicate excitatory types and connections; gray signifies inhibitory types and connections; blue dots, located at the output of the loop, are excitatory in one pattern combination and inhibitory in the other. Total network occurrences for each pattern are shown in square brackets.

Motifs and anti-motifs

To identify significantly over- or under-represented subcircuits, we benchmarked the HC networks to random graphs in which global topology was obliterated, but the underlying composition of all dimers was preserved. Surprisingly, superpattern topology was the most important factor in determining whether a pattern was a motif or an anti-motif (Fig. 17A). In other words, most superpatterns (13/16) are either motifs or anti-motifs, independent of the excitatory/inhibitory make-up of their nodes. Only the three-node chain, the single downlink to a mutual dyad, and the single uplinked mutual dyad (superpatterns B, D, and F) contained a mixture of motifs and anti-motifs (stacked green and red columns). For the others, the connectivity itself was either over- or under-employed relative to the expectation based on HC dimer distribution. Patterns belonging
Figure 16 Utilization of feedforward versus feedback loops in the HC

to D (single downlink to a mutual dyad), G (feedback loop), and I (chain of two mutual dyads) were severely underutilized, but patterns -B (disconnected single edge), -A
Figure 17 Motifs and anti-motifs are largely determined by superpattern topology
(disconnected mutual dyad), E (feedforward loop), H (double downlink to a mutual dyad), K (double uplinked mutual dyad), L (feedback with two mutual dyads), and M (fully connected triad) were all strong motifs. These results were robust to PC weighting. Additionally, motifs and anti-motifs were module-specific (Fig. 17B-D). The DG module contained a mixture of motif and anti-motifs for many superpatterns. Less-densely connected superpatterns, including simple regulation (name from Alon, 2007), three-node chain, single input module, and single downlink to a mutual dyad (superpatterns A through D) were underutilized in CA3; instead, these superpatterns tended to be slightly over-utilized in CA1, where F and G were anti-motifs.

**Pairwise correlations**

For the 88 neuron types not located in EC or subiculum, we tested the interactions among 315 connectivity, morphological, molecular, and electrophysiological properties and detected 14,217 (14.3%) significant correlations (p<0.05). These results fell across a spectrum of novelty, and the more interesting outcomes are presented here.

Like in the motif analysis, sub-regional differences in the usage of superpatterns and patterns were revealed. DG types have high participation (i.e. relative to other sub-regions) in three-node chains (superpattern B; p<0.02) but shun dense, highly connected superpatterns, including I (chain of two mutual dyads; p<0.002), J (single point feedforward and feedback loops; p<0.006), K (double uplinked mutual dyad; p<0.00002), L (feedback with two mutual dyads; p<0.00006), and M (fully connected triad; p<0.0005). Interestingly, CA3 and CA1 do not share parallel high or low participation in any superpattern or pattern. Instead, one pattern is highly used in CA1 but avoided in
CA3 and one pattern displays the opposite trend. Pattern H6, where an interneuron acts as a single input module dispersing information to two interneuron recipients with reciprocal feedback between them, is highly utilized in CA1 and under-employed in CA3. Contrarily, pattern L3, a reciprocal edge between two excitatory types, one of which is connected reciprocally with an interneuron and the other of which receives unidirectional information from that interneuron, is avoided in CA1 but highly used in CA3. This connectivity pattern is especially utilized by CA3 and CA3c Pyramidal cells to communicate with each other and with a third (interneuron) partner.

Pairwise contingency analysis also detected a set of characteristics differentiating projecting from local neurons. Projecting types participate highly in superpatterns where information converges to a single point onward through the TSL (p<0.0003): superpattern A (simple regulation) and superpattern E (feedforward loops). Conversely, projecting types use sparingly superpattern C (single input module) and particularly pattern C5 (p<0.0003), which is disperse excitation to two inhibitory nodes. CA1 Back-Projection cells, a major GABAergic projecting type, are the exception. This neuron type primarily interacts with other hippocampal interneurons and makes use of patterns C6 (inhibitory dispersal to two other GABAergic types) and F8 (Fig. 15A) more than any other neuron type.

Lastly, connectivity was clearly correlated with molecular expression (Hippocampome.org/markers). While the correlation between subcircuits involving only inhibitory types (e.g. from superpatterns A, D, I, and L) with expression of VIP+ and CR+ (two markers of interneuron-specific interneurons) was expected, other observations
were not. For example, even though somatostatin is not associated with interneuron-specific interneurons, somatostatin-positive (SOM+) cell types also tend to interact in groupings with two other GABAergic types (sometimes, but not always, with interneuron-specific types). More specifically, SOM+ types are among the top users of the all-inhibitory versions of superpatterns D, E, F, H, J, L, and M. In addition, neurons expressing parvalbumin (including perisomatic-targeting basket and axo-axonic cells) participate copiously in superpattern C (single input module), but sparingly in superpatterns B (three-node chain), D (single downlink to a mutual dyad), J (a single point feedforward and feedback loop), and M (fully connected triad).

**Sensitivity to future additions or subtractions of neuron types**

Finally, to examine the robustness of our results to reasonable changes in network size and composition, we reran all analyses on two modified networks. First, we eliminated 26 of the 122 neuron types (asterisks in Table 6), along with their connections, that were either not well known (e.g. described by a single peer-reviewed publication) or contained no or sparse molecular marker and electrophysiological evidence. In the second network, we added 23 new neuron types that are currently being annotated for inclusion in future versions of Hippocampome.org. Remarkably, analysis of both networks yielded results very similar to those reported here for the network of 122 types. Specifically, connection density, CPL, CC, and scaled communication cost were all within 4% of their HC values. More complex analyses (e.g. rich club and motifs) were similarly dependable.
Discussion

Knowledge about synaptic connectivity between identified neuron types in the hippocampal formation is currently quite scant: empirical information on the presence or absence of synapses is available for less than 1.6% of all possible neuron type pairs. These limited data, however, can be supplemented by leveraging spatially co-aligned axonal and dendritic patterns based on the evidence annotated in Hippocampome.org. Although axonal-dendritic co-location does not guarantee synaptic presence, applying the original, neuron-type version of Peters’ rule at least reveals the potential connectivity of the full hippocampal circuitry. While the concept of potential connectivity is extensible in other parts of the brain, it is particularly pertinent in the hippocampus due to its superior structural plasticity (Leuner and Gould, 2010). In particular, in this region, the lack of synapses between neurons at any given moment may not necessarily foreshadow the absence of connection at a different time.

This level of description of the rodent cerebro-hippocampal cortex complements (and fills a gap between) previous large-scale syntheses of tract-tracing studies (Burns and Young, 2000; van Strien et al., 2009) and sparse synaptic sampling (Druckmann et al., 2014). In fact, this effort represents the first comprehensive, literature-based neuron-type circuitry inventory for a mammalian cortical region. Thus, we began to unravel the structural complexity of the hippocampal network through graph theory analyses, shedding light on the functional roles of the component neuron types.

Although networks are quantitatively differentiable according to myriad metrics, two of the most topologically illustrative are CC and CPL (Watts and Strogatz, 1998). We first identified and quantified the specialized topology that brings about higher
efficiency and lower overall communication cost in HC than in any equivalent, well-studied network type. Small-world networks, in particular, have been researched and applied fashionably to brain networks for decades (Hilgetag and Goulas, 2016), but we detected significantly higher CC than in equivalent WS networks. The in-built capacity for rapid response times and precise processing, common elsewhere in the brain (Latora and Marchiori, 2001; Bassett and Bullmore, 2006; Rubinov and Sporns, 2010; Mišić et al., 2014), might be especially relevant to the demands of the hippocampus, where the tasks of memory consolidation, retrieval, cognitive navigation, and path-finding have inherent temporal and spatial constraints.

Next, we exposed a significant modular substructure comprising four densely intra-connected communities. It is worth noting that CA2 and the subiculum, the two hippocampal formation sub-regions with the fewest known neuron types (five and three, respectively), are currently subsumed into communities dominated by other sub-regions. CA2 types are split into the CA3 and CA1 modules; subicular types are divided among CA1 and EC. As future knowledge in these areas increases, presumably hailing a proliferation in interneuron diversity, one or both of these sub-regions might become independent modules.

Regardless, the major high-traffic links between the sub-regionally-based communities recapitulate the TSL and various shortcuts through it. This excitatory relay includes the perforant pathway (PP; grid and border cells from EC layer II to DG and CA3), the temporoammonic pathway (head direction and border cells from EC layer III to CA1), mossy fibers (from DG Granule cells to CA3), Schaffer collaterals (from CA3
Pyramidal cells to CA1), and the nameless projection from CA1 to EC layer V (Amaral and Lavenex, 2007; van Strien et al., 2009). Although the functional ramifications of these individual conduits are not yet fully understood, in a loop-heavy network that lacks discrete beginning and endpoints, the detected modules likely act as processing stations regulated by dense intra-modular connections (both excitatory and inhibitory). For example, though most studies of the PP focus on the well-known glutamatergic-to-glutamatergic connections onto Granule cell dendrites, feedforward inhibition also plays a major role in controlling information processing in DG (Ferrante et al., 2009). As we have shown, the PP can also affect DG interneurons such as MOPP and Neurogliaform cells (pattern E2 in Fig. 16). Under physiological conditions, these parallel routes might selectively respond to particular oscillatory input frequencies from EC reflecting different behavioral states (Tateno et al., 2007; Akam and Kullmann, 2010; Ewell and Jones, 2010; Jones and McHugh, 2011). Novelty, for example, induces a slight decrease in Granule cell firing rates concomitant with increased DG interneuron activity (Nitz and McNaughton, 2004).

Analysis of out-degree and in-degree distributions revealed that the peculiar HC topology was largely due to its axonal architecture while the dendritic circuitry was fairly unremarkable. This result is consistent with the recent finding that the computational load of neurons is unrelated to their in-degree (Timme et al., 2016); instead, neurons that process large amounts of information tend to receive connections from high OD neurons. The axonal distribution further pointed to anti-hub and hub neuron types that utilized highly specific or largely blind targeting, respectively. Hubs are notable because, by
directly connecting to many neuron types that are themselves neighbors, they violate
typical tenets of wiring minimization (Chklovskii et al., 2002; Chklovskii and Koulakov,
2004; Chen et al., 2006; Bullmore and Sporns, 2012). In fact, both the construction of
these superfluous, often long-distance connections and the regular handling of a
disproportionately large volume of traffic come at a high cost of energy. At the same
time, these nodes facilitate the integration of distributed neural activity and are well
situated to integrate the network modules. Such double-edged nature justifies the “high
cost, high value” characterization of these circuit elements (van den Heuvel and Sporns,
2013).

While quantification of modularity yielded informative but non-overlapping
groupings of HC neuron types, the rich club analysis produced a hierarchical ordering of
importance of each type to the network. Rich clubs also have been detected in other parts
of the brain across several species (van den Heuvel and Sporns, 2011; Harriger et al.,
2012; Shanahan et al., 2013; Binicewicz et al., 2015). We identified two nested rich clubs
that are likely to route much of the network traffic (Mišić et al., 2014) within the
hippocampal formation. Like hubs, this feature accentuates a departure from the
parsimonious wiring typically observed in neural systems, but the paths between these
critical types provide a highly efficient network core with built-in protection against
neurodegeneration. Members of these rich clubs, including the global hubs, are
potentially vulnerable to targeted attacks: damaging all neurons within one of these types
could lower network efficiency and increase processing times, possibly impairing storage
and retrieval functionality. Interestingly, these same types tend to be particularly
abundant in terms of cell numerosity, with principal cells present in quantities up to 10 times higher than other neuron types (Bezaire and Soltesz, 2013), thus providing a certain level of resistance against random neurodegeneration. More generally, we showed that the plethora of alternate pathways available in the circuit serves as a second countermeasure.

Finally, we analyzed the superpattern and pattern building blocks responsible for the local interactions that enable global functionality. Three-node subgraphs have attracted considerable attention for their role in complex networks across disciplines (Milo et al., 2002; Shen-Orr et al., 2002; Milo et al., 2004), including neuroscience (Sporns and Kötter, 2004; Song et al., 2005; Santana et al., 2011; Binicewicz et al., 2015). They have been specifically studied in the hippocampus with focus directed at DG Mossy cells and unidentified hilar interneurons (Larimer and Strowbridge, 2008) and among recurrent connections of CA3 pyramidal cells (Guzman et al., 2016). Building on this well-defined framework, we added excitatory/inhibitory weights and identified connectivity pattern relations among HC neuron types. In truth, the empirical characterization of even simple (e.g. two-node) interactions between excitatory and inhibitory cells is still vastly incomplete. Recent recordings from more than 500 pyramidal cells and 1,500 GABAergic neurons in the mouse neocortex delineated fifteen interneuron types that could be grouped based on broad connectivity preferences (Jiang et al., 2015). One group preferentially formed synapses with pyramidal neurons; another, referred to as “master regulators,” connected nonspecifically to all types in proximity of their axons; two additional groups contained interneuron specific (IS) cells that synapsed
primarily with other interneurons of the same or of different types, respectively. When these interactions are extended to include a third party (i.e. trimers), the functional implications are more complex. IS cells, in particular, have recently been the subject of much study for their role in disinhibition. Specifically, IS cells can influence principal neurons by inhibiting other GABAergic interneurons (Pi et al., 2013; Jiang et al, 2015). In the cortex, this type of circuit control has been linked to enhanced plasticity (Fu et al., 2015) and shown to affect social behavior (Yizhar et al., 2011), sensorimotor integration (Lee et al., 2013), attention (Vogels and Abbott, 2009; Sridharan and Knudsen, 2014; Zhang et al., 2014), and associative learning and memory (Letzkus et al., 2011; Letzkus et al., 2015). The specific involvement of the hippocampus in many of the above functions makes these connectivity patterns particularly worthy of study.

On a related note, the excitability scores computed for each patterns only capture the overall excitatory/inhibitory nature of a structurally defined trimer. In actuality, each trimer can produce multiple functional states that are affected by the degree of activation, delays in signal propagation, the surrounding neural context, and the behavioral state of the organism (Sporns and Kötter, 2004). Those various functional states of patterns and superpatterns are not analyzed here.

Further meaningful interpretation of our results was hindered by two main factors, both imputable to data incompleteness. First, the neuron types identified in HC are limited to the information available in the literature. While certain hippocampal areas are well studied (e.g. CA1), other domains, including CA2, subiculum, and entorhinal GABAergic neurons are still under-researched. Other parts of the subicular complex,
including the prosubiculum, presubiculum, postsubiculum, and parasubiculum, are not tracked in version 1.0 of Hippocampome.org. Though morphologically based neuronal-type information was recently reported for the presubiculum (Nassar et al., 2015), on the whole, the breadth of knowledge within these additional areas is relatively narrow. As the scientific community overcomes these shortcomings, the published evidence and Hippocampome.org will grow, which will result in additions and alterations to the connectivity. This accumulating knowledge could impact some of the HC circuit properties described herein. To assess and mitigate this issue, we repeated the analyses on modified (reduced and expanded) networks, and concluded that the main network properties of HC are innate to the well-known, constituent neuron types and unshaken by reasonable additions or deletions.

The second hindrance is the lack of connection weights beyond the binary differentiation of glutamatergic and GABAergic types. A comprehensive solution of this problem requires quantitatively estimating both the counts for each neuron type and the corresponding axonal and dendritic length distributions. Ideally, cell counts would be determined from absolute, stereology-derived numbers for each morphologically defined neuron type, but relative ratios of molecularly-defined subpopulations across anatomical parcels could already be useful. These challenging experimental tasks are complicated further by variation across rodent species, strains, ages, sexes, and anatomical axes. Nevertheless, it is generally assumed that principal cells dominate the relative abundances of other neuron types by an order of magnitude (Bezaire and Soltesz, 2013), with experimental observations ranging from 89% of the hippocampal neuron population
as a whole (Woodson et al., 1989) to 93% within CA1 (Aika et al., 1994). Accordingly, we also carried out the graph theory analyses with principal cell connections weighted as +10. Remarkably, our conclusions were largely unchanged and, in many cases detailed herein, their significance was even amplified. In addition to obtaining counts for the neuron types, weighting for connections should also be based on measuring the three-dimensional overlap of the neurite trees of each type. However, this approach is currently unfeasible as three-dimensional reconstructions (e.g. from NeuroMorpho.Org) are only available for a small fraction of neuron types. Lastly, the expression levels of the primary neurotransmitters, as well as the prevalence of membrane receptor proteins in distinct post-synaptic cell types, also play a role in weighting the connections. Though this information, too, is currently lacking, appropriate data can be included in the future to extend potential connectivity analyses beyond binary values.

With the set of tools deployed in this work, future updates to the connectome (through addition, merging, splitting, and weighting of nodes and edges, or through augmentation of known connectivity) can be analyzed in relatively short order. Furthermore, as information accumulates about aging and disease states, the analyses can be repeated with a comparative bent. Extending the foundational results presented here with the expected continuous growth of data will progressively improve our understanding of how network architecture mediates hippocampal function and dysfunction.
Hippocampome.org: a knowledge base of neuron types in the rodent hippocampus

Diek W Wheeler, Charise M White, Christopher L Rees, Alexander O Komendantov, David J Hamilton, Giorgio A Ascoli*

Krasnow Institute for Advanced Study, George Mason University, Fairfax, United States

Abstract Hippocampome.org is a comprehensive knowledge base of neuron types in the rodent hippocampal formation (dentate gyrus, CA3, CA2, CA1, subiculum, and entorhinal cortex). Although the hippocampal literature is remarkably information-rich, neuron properties are often reported with incompletely defined and notoriously inconsistent terminology, creating a formidable challenge for data integration. Our extensive literature mining and data reconciliation identified 122 neuron types based on neurotransmitter, axonal and dendritic patterns, synaptic specificity, electrophysiology, and molecular biomarkers. All ~3700 annotated properties are individually supported by specific evidence (~14,000 pieces) in peer-reviewed publications. Systematic analysis of this unprecedented amount of machine-readable information reveals novel correlations among neuron types and properties, the potential connectivity of the full hippocampal circuitry, and outstanding knowledge gaps. User-friendly browsing and online querying of Hippocampome.org may aid design and interpretation of both experiments and simulations. This powerful, simple, and extensible neuron classification endeavor is unique in its detail, utility, and completeness.

DOI: 10.7554/eLife.09960.001

Introduction

A century after the advent of the neuron doctrine (Jones, 2006), neuroscience is challenged with data on neuronal characteristics spread throughout thousands of articles in hundreds of journals, growing at a rate of dozens of new papers each day. The sweeping array of research designs, data formats, and presentation styles, reflecting the tremendous diversity in neuronal structure and functions, makes it extremely difficult to assess the available information. Overcoming the hindrances of this variety is problematic because of limited consensus on which properties to use for neuron classification and how to define them effectively (Petilla Interneuron Nomenclature Group et al., 2008; Hamilton et al., 2012). The lack of a complete accounting of neuron types is recognized as a critical omission for experimental and computational progress alike (Lichtman and Denk, 2011).

As a case in point, the rodent hippocampus is among the most intensively studied neural systems. Numerous seminal efforts have summarized the enormous amount of information on the morphology, connectivity, biochemistry, and electrophysiology of hippocampal neurons (Bernard and Wheel, 1994; Patton and McNaughton, 1995; Freund and Buzsáki, 1996; Parra et al., 1998; McBain and Fisahn, 2001; Maccari and Lacaille, 2003; Amaral and Lavenex, 2007; Canto et al., 2008; Klausberger and Somogyi, 2008; Klausberger, 2009; Somogyi, 2010; Bezaire and Sotetsu, 2013). However, these notable advances have yet to translate into an integrated understanding of corresponding functions. The hippocampus plays a critical role in the consolidation and retrieval of episodic memory (Nadel and Moscovitch, 1997; Wang and Morris, 2010) as well as in spatial representation and navigation (Foster and Krierm, 2012; Moser et al., 2015). Many theories have been formulated to connect these important cognitive functions to hippocampal architecture (Clark and Squire, 2010; Eichenbaum and Cohen, 2014), rhythmic activity (Burgess and O’Keefe, 2011;
**eLife digest** The hippocampus is a seahorse-shaped region of the brain that is responsible for learning, emotions, and memory. Like other regions of the brain, it contains many types of neurons that send information to each other by releasing chemicals called neurotransmitters across junctions known as synapses. Identifying all the different neuron types in the hippocampus is an important step towards understanding in detail how this brain region works.

Thousands of articles have been published that attempt to characterize the neurons in the hippocampus, but many of these studies report only some of the properties of a new neuron type. It is also often difficult to compare the results of different studies, as many different approaches have been used to investigate neuron types, and different studies often use different terms to describe similar features.

Wheeler et al. have now created a resource called Hippocampome.org that combines approximately 14,000 pieces of experimental evidence about neuron types in the rat hippocampus into a unified database. Analyzing these data has revealed about 3700 different neuron properties. By primarily considering the pattern formed by the branched axons and dendrites, the outputs and inputs that extend out of a neuron, Wheeler et al. have identified over a hundred different neuron types. This classification system also considers how selectively the neuron forms synapses with other cells and the identity of the neurotransmitter released by the neuron. In the future, other features of the neurons will also be incorporated into the system to help refine the classifications.

All of this information is online and freely available at Hippocampome.org. This resource is expected to provide a solid basis for analyzing how the hippocampus works, by helping researchers to design and interpret experiments and simulations.

DOI: 10.7554/eLife.09960.002

---

Hasselmo and Stern, 2014), and synaptic plasticity (Rolls and Treves, 1994; Bliss and Collingridge, 2013). However, current models are far from comprehensively accounting for the available experimental data on all neuron types and relevant properties.

As a stepping stone towards filling this gap, we mounted a large-scale literature mining effort to assemble a knowledge base of neuron types in the rodent hippocampal formation: dentate gyrus (DG), CA3, CA2, CA1, subiculum (Sub), and entorhinal cortex (EC). Our data-driven analysis of thousands of peer-reviewed publications identified a specific set of neuron properties suitable to define a basic classification scheme for collating, organizing, and integrating available knowledge. A cornerstone of this approach is the distributions of axons and dendrites across the distinct areas and layers of the hippocampal formation, such as CA1 stratum radiatum (SR) or EC layer II. Axonal and dendritic arbors perennially have been central to the experimental identification of neurons (Jones, 2006; DeFelipe et al., 2013), as they underlie network connectivity and profoundly influence information processing. Systematic mining of published data on axonal and dendritic profiles, augmented with information on neurotransmitter and synaptic specificity, led to the tentative definition of over 100 distinct neuron types across the hippocampal formation. Importantly, this initial identification of ‘morphological’ types enabled the unambiguous incorporation of existing molecular and electrophysiological data into the knowledge base. Other key features, such as developmental origin, firing dynamics, synaptic properties, neuron counts, and connectivity ratios, are being progressively integrated into the same classification framework.

Essential to this effort is the digestion of information from original publications into both human- and machine-readable forms. We present Hippocampome.org, a publicly accessible web portal to browse and query data with direct links to specific evidence in the scientific literature. Information summaries are available for each neuron type, anatomical parcel, molecular marker, and cited author. Searching for combined morphological, molecular, and electrophysiological properties returns lists of all known neuron types with those features. These functionalities, coupled with the knowledge base’s comprehensiveness and inter-relatedness, reveal novel insights on neuron types, their properties, and circuit connectivity, which are impractical or impossible to derive from traditional literature searches. Moreover, the collation of data enables quantification of available information, unearthing residual knowledge gaps.

The ‘Description of resource’ section of this article aims to offer a clear description of both the conceptual foundation and practical utility of Hippocampome.org according to the following
organization. First, we explain the neuron type classification criteria, including the definition of morphological patterns and the distinctions by main neurotransmitter, synaptic specificity, biomarkers, and electrophysiology. We then provide a summary of the knowledge base content and describe its online accessibility. Next, we illustrate the usefulness and possible applications of Hippocampome.org through analyses of pairwise correlations among neuron properties, potential circuit connectivity, and use case scenarios in experimental and computational investigations. A discussion follows the 'Description of resource' section. In the final 'Materials and methods' section, we provide practical details of how the Hippocampome.org knowledge base was assembled. We define parcels and neuron types, explain how biomarker and electrophysiological data are linked to morphological data, describe how names are assigned to neuron types, and expand upon how the knowledge base will be maintained going forward. For optimal comprehension of this resource, we recommend online consultation of Hippocampome.org while reading this article.

Description of resource

Identifying neuron types by axonal and dendritic patterns

As a first step in the knowledge-base design, we sought the optimal level of description to capture the largest possible extent of available information. For example, just distinguishing neurons into projection cells and local interneurons is too coarse to reflect the known variety of hippocampal neuron types. In contrast, demanding a detailed quantification of axonal and dendritic morphology and connectivity excludes the majority of available scientific reports that only depict neuronal arbors qualitatively. An intermediate approach is to describe axons and dendrites based on the specific areas and layers they invade. The orderly anatomical organization of the hippocampal formation is commonly delineated into 26 cytoarchitectural parcels (Figure 1A,B): DG outer stratum (s.) moleculare (SMo), inner s. moleculare (SMi), s. granulosum (SG), hilus (H); CA3 s. lacunosum-moleculare (SLM), s. radiatum (SR), s. lucidum (SL), s. pyramidal (SP), s. oriens (SO); 4 each in CA2 and CA1 (same as CA3 except SL); Sub s. moleculare (SM), SP, polymorphic layer (PL); and EC layers I–VI. Most publications that report morphological information on hippocampal neurons include evidence of axonal and dendritic presence in at least a subset of these parcels in the form of reconstructions, tracings, microscopic images, schematics, or text descriptions (Figure 1C–I).

This relatively simple description is surprisingly effective to map neuronal diversity in the hippocampus. Specifically, the binary representation (present or not) of both axons and dendrites across the 26 hippocampal parcels (see ‘Materials and methods’) is sufficient to identify >100 unique morphological profiles based on existing literature (selection in Figure 1F; complete data: hippocampome.org/morphology). A key assumption of this framework is that neurons differing in axonal or dendritic patterns are different types. The same morphological pattern is further separated into distinct types if the neurons can be discriminated by neurotransmitter, synaptic selectivity, or consistent molecular and electrophysiological differences (see below), giving rise to 122 neuron types distinguished so far (18 in DG, 23 in CA3, 5 in CA2, 40 in CA1, 3 in Sub, and 31 in EC). The soma layer location is annotated for all neuron types, but neurons are not considered of separate types solely due to this feature if their other main characteristics are the same. Several neuron types have somata distributed across both sides of, or just along, a layer boundary, such as CA1 quadrilaminar cells (Pawelzik et al., 2002) or DG hilar commissural-associational pathway (HICAP) cells (Mott et al., 1997).

If only dendrites or axons are observed (but not both), or a single instance of a neuron morphology is described (n = 1), the characterization is deemed insufficient to include a neuron type in the knowledge base. We often attempt to contact authors to ascertain whether additional data may exist. If no further evidence is available, the existing information is extracted but not integrated with the rest of the knowledge base, and the possible neuron type is placed 'on hold' until more data are published. Currently, 151 potential neuron types are in this state (hippocampome.org/on-hold).

Hippocampome.org links information from various publications to specific neuron types primarily based on parcel-delineated neurite morphology and does not rely on author given names for neuron classification. The reason is that the same name is frequently employed in the literature to describe neurons with different morphological patterns. For example, the term ‘bistratified’ has been used to describe different morphologies corresponding to Hippocampome cell types CA1 Bistratified (Freund and Buzsáki, 1996), CA1 Radial Trilaminar (Daw et al., 2009), and Schaffer Collateral-receiving Radial-targeting (Leão et al., 2012). Furthermore, neurons of the same type are often
Figure 1. Defining neuron types with anatomical parcels and morphological patterns. (A, B) Nissl staining of a P56 mouse (coronal sections 74 and 85 from the Allen Brain Atlas) overlaid by color-coded parcels (in accord with Temporal-Lobe.com) of (A) the hippocampus proper and (B) the rest of the hippocampal formation. See main text for abbreviations. (C–E) Example morphological reconstructions from NeuroMorpho.org (left) with red axons and blue dendrites; their estimated axonal and dendritic breakdown by layers (middle); and Hippocampome.org representation (right) with blue square and Figure 1. continued on next page.
given different names due to varying emphases of researchers. For instance, CA1 Radiatum cells (Cope et al., 2002) have also been called CaBP-positive non-pyramidal cells (Tooth and Freund, 1992) and Schaffer collateral associated cells (Vida, 2010). The authors’ original names are always included as synonyms of the Hippocampome.org-assigned identifiers (see ‘Materials and methods’) to facilitate tracing information provenance and for ease of communication.

An advantage of defining neuron types by their axonal and dendritic patterns is the practical cross-species invariance of these features, at least within rodents. Hippocampome.org collates data from all healthy young or adult rats, mice, and guinea pigs, but all evidence is meta-annotated with species and age information to allow separate tracking. The few cross-species morphological differences identified so far, such as the extent of CA3c infrapyramidal mossy fibers (Blaabjerg and Zimmer, 2007), are immaterial to the current granularity of the knowledge base.

Major neurotransmitter distinctions

Hippocampal neurons mainly release glutamate or gamma-aminobutyric acid (GABA) (Kullmann, 2007). When the literature does not present explicit evidence for determining the neurotransmitter, a neuron type is deemed putatively glutamatergic or GABAergic depending on ancillary characteristics including (asymmetry of synapses, excitatory or inhibitory effect, relative somatic abundance within an area, presence of dendritic spines, and local or projecting nature of the axons. Although exceptions exist for all these criteria, taken together, they are sufficiently indicative to enable the tentative inference of the main neurotransmitter for all identified neuron types based on published information (Figure 1F: black, presumed glutamatergic; gray, presumed GABAergic).

Neurons with the same axonal and dendritic patterns, but different neurotransmitters, belong to different types. Interestingly, only one such case is known: in CA1, Cajal–Retzius cells, which were recently characterized as glutamatergic and more abundant than previously assumed in adult rats (Quattrocchio and Maccapani, 2014), and (GABAergic) Neurogliaform cells (Price et al., 2003) have both axons and dendrites confined to SLM, though with extremely different arbor densities and shapes. Thus, putatively excitatory and inhibitory neurons in the hippocampus tend to have completely distinct morphologies.

Although most excitatory neurons in all hippocampal areas have long-range axons, Hippocampome.org also includes several local glutamatergic types in DG, CA3, and CA1. Similarly, a substantial minority of GABAergic types in these areas (16/71) have projecting axons. Knowledge is sparser in CA2, Sub, and EC. The axons of several presumed glutamatergic types in EC have not been reconstructed beyond layer VI or the angular bundle, so their participation in long-range pathways remains largely unknown (Cantó and Witter, 2012a, 2012b).

Synaptic specificity

Certain interneurons selectively discriminate between GABAergic and glutamatergic post-synaptic partners. Like the neurotransmitter, preferential connectivity profoundly affects circuit function. This characteristic is thus essential to distinguish neuron types with the same axonal–dendritic pattern. For example, A xo-axonic and Basket cells in CA3 and CA1 have dendrites spanning all layers and axons limited to SP (e.g., Freund and Buzsáki, 1996; Klausberger and Somogyi, 2008). A xo-axonic cells, however, contact exclusively Pyramidal cells on the axon initial segments (Li et al., 1992; Ganter et al., 2004), whereas Basket cells synapse perisomatically on both principal neurons and interneurons (Buhl et al., 1994; Vida, 2010). The separation of A xo-axonic and Basket cells yields two additional types with the same axonal–dendritic patterns.
Several ‘interneuron-specific’ neuron types preferentially target other GABAergic neurons over glutamatergic types. Since all known cases correspond to unique morphological patterns, such selectivity does not separate further neuron types. However, this core property is indicated in their adopted names, as in CA1 Interneuron-specific O-R cells (cf. Gulyás et al., 1996).

**Other properties distinguishing neuron types**

The hippocampal literature offers a wealth of biochemical and electrophysiological data (see below) that are linked in the knowledge base to specific morphologically defined neuron types (‘Materials and methods’). Differences in individual molecular or biophysical features, such as the presence/absence of a single biomarker or the high/low value in one membrane property, are frequently reported for a given axonal-dendritic pattern. These cases are consistently annotated in Hipcampom.org as indicating potential sub-types, but do not establish full neuron types.

In contrast, when opposite expression of multiple markers and large discrepancies in several membrane properties strongly support the existence of distinct neuron types by converging molecular and electrophysiological evidence, the same morphological pattern is divided into two types. For example, one type of CA1 Basket cells expresses parvalbumin (PV) and µ-opioid receptor, but not cholecystokinin (CCK), cannabinoid receptor 1 (CB1), and substance P receptor (sub P rec), while another type displays the opposite expression profile (Pawelski et al., 2002); moreover, PV+ cells have significantly lower input resistance ($R_i$), faster membrane time constant ($\tau_m$), narrower action potential width ($AP_{peak}$) and smaller slow after-hyperpolarization (sAHF) than CCK+ cells (Bartos and Elgula, 2012).

Similarly, both CA1 Bistratified and Ivy cells have axons and dendrites extending from SO to SR, but the former are positive for PV and somatostatin (SOM) and negative for neuronal nitric oxide synthase (nNOS), while the opposite holds for the latter; furthermore, relative to Ivy cells, Bistratified cells have lower firing threshold ($V_{thresh}$) and maximum firing rates (maxFR), larger action potential amplitude ($AP_{peak}$), and narrower $AP_{peak}$ (Fuentealba et al., 2006; Tricoire et al., 2011). Equally convincing data support the distinction of CA3 Ivy and Bistratified cells as well as the separation of basket cells in DG and CA3. In all, five additional types with non-unique axonal-dendritic patterns are established on the basis of clear and substantial molecular and electrophysiological information.

**Molecular markers**

Expression data are obtained from studies detecting proteins (immunohistochemistry) or mRNA (in situ hybridization, promoter constructs, or single-cell RT-PCR). The knowledge base has information on 96 biomarkers, and positive or negative expression of at least one biomarker is known for two-thirds (81/122) of the neuron types (Figure 2; extended listing: hippocampom.org/markers). The most cited biomarker is parvalbumin (PV), appearing in ~25% of the biomarker sources in Hipcampom.org, and whose expression is known for 56 neuron types.

In 79 of the 56 cases with available expression data for the 20 most cited biomarkers (non-gray squares in Figure 2 and hippocampom.org/markers), evidence exists for a neuron type to both express and not express a given biomarker. In 13 of these cases, the incongruity can be attributed to species, technique, or sub-cellular localization differences (Figure 2, orange, down-facing flags). For example, mossy cells express calretinin (CR) in mice (Blasco-Ibáñez and Freund, 1997), but not in rats (Freund et al., 1997). In another case, in situ hybridization showed GABA_A-α1 expression in rat CA1 Pyramidal cells (Miralles et al., 1994), yet somatic immunohistochemistry of these neurons is negative; however, their dendrites show positive expression (Lopez-Tellez et al., 2004).

Another 58 mixed cases are designated as possible subtypes (Figure 2, blue and green flags in the same square), when in a population of cells with equivalent morphology, from the same experiment, a sizeable proportion expresses a biomarker and a sizeable proportion does not. The interpretation of potential subtypes is sometimes supported by distinct somatic locations. For example, CA3 Pyramidal cells are positive for chicken ovalbumin upstream promoter transcription factor II (COUP-TFI) in the temporal, but not septal hippocampus (Fuentealba et al., 2010). Similarly, superficial CA1 Pyramidal cells express calbindin (CB) while those deep in CA1 SP do not (Jinno et al., 1999; Sadowski et al., 2002).

In the remaining eight cases of mixed expression (red, upward-facing flags), data come from multiple sources that use the same species and technique (e.g., rat immunohistochemistry); however, non-identical experimental details (e.g., the antibodies used) prevent conclusive interpretation.
Figure 2. Expression of 20 common biomarkers for 100 representative neuron types (full matrix: hippocampome.org/markers or hippocampome.org/php/images/marker/Marker_Matrix.jpg; complete list of abbreviations: hippocampome.org/help). Positive expression: left green flags; negative expression: right blue; mixed expression (possible subtypes): left/right green blue; mixed expression (different experimental protocols, species, or conditions): mixed green blue. Figure 2 continued on next page.
Figure 2. Continued

Sub-cellular localization: top orange; unresolved mixed expression: bottom red; empty gray boxes indicate that morphologically linkable information was not found. The right summary column reports the number of additional biomarkers with known expression for each neuron type. Bottom values are counts of neuron types with available information for each biomarker.

DOI: 10.7554/eLife.09960.004

Electrophysiology

Electrophysiological characteristics vary extensively across hippocampal neurons. Passive, spike, and other response parameters are commonly recorded in the slice preparation. Slice recording is highly dependent on experimental method (patch clamp or sharp microelectrode), animal species and age, bathing and intra-electrode solution, and temperature. Thus, summary values (Table 1 for 50 representative neuron types; complete listing: hippocampome.org/electrophysiology) are not averaged across all studies, but rather reflect data from a single report under ‘preferred’ conditions (operationally defined as rat patch clamp at body temperature) with the largest number of measurements (n). These settings represent 61% of the data in Table 1. If these conditions are unavailable, the precedence order is patch clamp over microelectrodes, body temperature over room temperature, and rats over mice over guinea pigs. However, all collected parameter values (preferred or otherwise) for every neuron type are accessible at Hippocampome.org, each complete with metadata, available statistics, and direct links to underlying experimental evidence. Three-quarters of the neuron types (93/122) have measurements reported for four or more parameters, the most common being AP_{max}, fast afterhyperpolarization amplitude (fAHP), and V_{thres}.

Summary of design criteria

Although Hippocampome.org primarily categorizes neuron types according to their morphologies (locations of axons and dendrites across the 26 parcels of the hippocampal formation), several other properties are also considered (Figure 3). When evaluating publications describing a neuron type, we first apply a set of interpretative rules to the morphological information (see ‘Materials and methods’ and hippocampome.org/full-interp). We then follow a series of systematic criteria that can result in one of three end points: the mined information is consistent with (and supplementary of) a currently active, fully defined neuron type (hippocampome.org/supplemental); the information is distinct enough to establish a new active neuron type; or the information is insufficient, thus generating an ‘on hold’ neuron type (hippocampome.org/on-hold).

A digital storehouse of explicit knowledge

Hippocampome.org provides free, user-friendly online access to the entire information content. The annotated evidence supports 3697 distinct ‘pieces of knowledge’ (PoK) regarding the presence or absence of axons or dendrites within any of the 26 parcels (e.g., Granule cells axons are found in CA3 SL represents one PoK, the expression or non-expression of a biomarker (e.g., Granule cells are positive for CB, and individual electrophysiological properties (e.g., the AP_{max} of Granule cells is 1.71 ± 0.58 ms, mean ± standard deviation). This knowledge exceeds by 2–3 orders of magnitude (Figure 4A); the average numbers of PoK collated from relevant articles (<5) and reviews (<27). Integration of published information from multiple sources, by linking together molecular, electrophysiological, and morphological data, produces an explicit knowledge web of references, properties, and neuron types interlaced in a many-to-many fashion (Figure 4B). Of equal importance, this resource explicitly highlights the non-uniform distribution of knowledge among neuron types and across morphological, molecular, and electrophysiological properties (Figure 4C). The approximate ratio of morphological, molecular, and electrophysiological PoK is 5:3:2. While CA1 is the most studied hippocampal area (1321 PoK), outstanding knowledge gaps remain on interneuron diversity in subiculum (only Avo-axonic cells have enough information for inclusion as an active neuron type in Hippocampome.org) and EC (no GABAergic cells characterized with either soma or axons in deep layers).

The massive amount of knowledge and online availability of Hippocampome.org enable both broad-scope analytics and quick-use information checks. All information can be browsed or searched.
### Table 1. Electrophysiological properties for 50 representative neuron types (full table: hippocampome.org/electrophysiology/Electrophysiology_Table.pdf)

<table>
<thead>
<tr>
<th>Neuron Type</th>
<th>$V_{rest}$ (mV)</th>
<th>$R_{n}$ (MΩ)</th>
<th>$\tau_\alpha$ (ms)</th>
<th>$V_{threshold}$ (mV)</th>
<th>fAHP (mV)</th>
<th>AP amplitude (mV)</th>
<th>maxFR (Hz)</th>
<th>saHP (mV)</th>
<th>Sag ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Granule</td>
<td>−75 ± 2(16)</td>
<td>228 ± 79.1</td>
<td>26.9 ± 6.7</td>
<td>35 ± 6</td>
<td>11.7 ± 1</td>
<td>91 ± 5</td>
<td>0.87 ± 0.06</td>
<td>72 ± 8(16)</td>
<td>0.6 ± 0.8</td>
</tr>
<tr>
<td>Mosey</td>
<td>−62 ± 19(9)</td>
<td>199 ± 19(9)</td>
<td>41 ± 3(8)</td>
<td>23.7</td>
<td>6.2 ± 0.09</td>
<td>62.5</td>
<td>0.78 ± 0.04(9)</td>
<td>50 ± 6(9)</td>
<td>2.8 ± 0.7</td>
</tr>
<tr>
<td>AIPRIM</td>
<td>−64 ± 2(4)</td>
<td>363 ± 62(4)</td>
<td>30 ± 6(4)</td>
<td>16</td>
<td>12.3 ± 2</td>
<td>3(3)</td>
<td>0.5 ± 0.02(4)</td>
<td>81 ± 9(4)</td>
<td>3 ± 0.7</td>
</tr>
<tr>
<td>DG Axo-axonic</td>
<td>−65.1 ± 3(14)*</td>
<td>73.9 ± 18.6(14)*</td>
<td>7.7 ± 3.8(14)*</td>
<td>7.7 ± 2</td>
<td>7, 8 ± 1</td>
<td>86</td>
<td>0.42 ± 0.2(2)</td>
<td>85 ± 5(4)</td>
<td>3 ± 0.8</td>
</tr>
<tr>
<td>DG Basket</td>
<td>−62 ± 3(3)</td>
<td>43 ± 3(3)</td>
<td>10 ± 3(3)</td>
<td>17.6 ± 19.0</td>
<td>20 ± 2.3(3)</td>
<td>3(3)</td>
<td>0.25 ± 0.04(3)</td>
<td>230 ± 15</td>
<td>7 ± 1.2</td>
</tr>
<tr>
<td>HIPROM</td>
<td>−65 ± 6(3)</td>
<td>371 ± 47(3)</td>
<td>35</td>
<td>25.1 ± 27.3</td>
<td>13.1 ± 3</td>
<td>80.8</td>
<td>0.72 ± 0.9(3)</td>
<td>47 ± 3(3)</td>
<td>3.1 ± 1</td>
</tr>
<tr>
<td>MOLAX</td>
<td>−54 ± 1.9</td>
<td>198.2 ± 23.8</td>
<td>18.4 ± 1.1(13)</td>
<td>15.2 ± 2</td>
<td>11 ± 2</td>
<td>432 ± 44.4</td>
<td>1.12 ± 0.6(3)</td>
<td>50 ± 4(3)</td>
<td>3.5 ± 0.9</td>
</tr>
<tr>
<td>Total Molecular Layer</td>
<td>−54 ± 1.9</td>
<td>198.2 ± 23.8</td>
<td>18.4 ± 1.1(13)</td>
<td>15.2 ± 2</td>
<td>11 ± 2</td>
<td>432 ± 44.4</td>
<td>1.12 ± 0.6(3)</td>
<td>50 ± 4(3)</td>
<td>4.5 ± 0.9</td>
</tr>
<tr>
<td>CA3 (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA3 Pyramidal</td>
<td>−60.5 ± 5.4</td>
<td>126 ± 8(3)</td>
<td>61 ± 24(3)</td>
<td>13 ± 5</td>
<td>10.2 ± 0.5</td>
<td>97.6 ± 1.9</td>
<td>1 ± 0.1(4)</td>
<td>40 ± 8(3)</td>
<td>7.5m ± 3</td>
</tr>
<tr>
<td>CA3 Giant</td>
<td>−57 ± 1.2(28)</td>
<td>229 ± 22(28)</td>
<td>22</td>
<td>34.8 ± 2</td>
<td>74 ± 9.2(28)</td>
<td>1.1 ± 0.1(28)</td>
<td>1.2 ± 0.0(8)</td>
<td>50 ± 5(13)</td>
<td>7.1 ± 0.9</td>
</tr>
<tr>
<td>CA3 Granule</td>
<td>−78 ± 0.15</td>
<td>139 ± 11(5)</td>
<td>17.1 ± 8.1(15)</td>
<td>27.4</td>
<td>7.9 ± 2</td>
<td>104.3</td>
<td>0.82 ± 1.0(2)</td>
<td>&gt;100 ± 1.5</td>
<td>1.75 ± 0.9</td>
</tr>
<tr>
<td>CA3 Basket</td>
<td>−58.5 ± 2.8</td>
<td>122.9 ± 26.7</td>
<td>11.2 ± 2.9</td>
<td>25.6m</td>
<td>35m</td>
<td>77.1m</td>
<td>0.54 ± 0.1(8)m</td>
<td>33.6 ± 6.4</td>
<td>3m ± 0.93m</td>
</tr>
<tr>
<td>Lucidum ORAX</td>
<td>−61 ± 6(9)</td>
<td>284 ± 16(9)</td>
<td>42 ± 17(8)</td>
<td>16.3 ± 2</td>
<td>15 ± 6.6(9)</td>
<td>77.7 ± 12</td>
<td>0.52 ± 0.2(9)</td>
<td>75 ± 8(9)</td>
<td>6.6 ± 3.9</td>
</tr>
<tr>
<td>MFA ORDEN</td>
<td>−57 ± 5(13)</td>
<td>225 ± 193(13)</td>
<td>29.1 ± 14.6(13)</td>
<td>20</td>
<td>13.1 ± 1.91(13)</td>
<td>74</td>
<td>0.72 ± 0.15(13)</td>
<td>73 ± 1.6(13)</td>
<td>6.5 ± 1</td>
</tr>
<tr>
<td>CA3 O-LM</td>
<td>−60 ± 12</td>
<td>315.1 ± 161.15(15m)</td>
<td>33.3 ± 5.4</td>
<td>1.16 ± 0.37(15m)</td>
<td>34.8m</td>
<td>109m</td>
<td>0.84 ± 0.2 (15m)</td>
<td>0.82 ± 1.0(15m)</td>
<td>82 ± 1.1</td>
</tr>
<tr>
<td>CA3 Trilaminar</td>
<td>−61.2 ± 13</td>
<td>167.3 ± 59.1</td>
<td>16.9 ± 8.8 (8m)</td>
<td>12m</td>
<td>30m</td>
<td>101.5 ± 62.28(15m)</td>
<td>0.57 ± 0.8 (8m)</td>
<td>&gt;180m</td>
<td>3m ± 0.99</td>
</tr>
<tr>
<td>CA2 (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA2 Basket</td>
<td>−71.2 ± 4(68)</td>
<td>77 ± 19.36(68)</td>
<td>8.2 ± 3.2(68)</td>
<td>17, 27(1)</td>
<td>22 ± 5.16(68)</td>
<td>62.7 ± 9.6(68)</td>
<td>0.5 ± 0.18(68)</td>
<td>&gt;180(1)</td>
<td>3m ± 0.99</td>
</tr>
<tr>
<td>CA2 Wide-</td>
<td>−74.9 ± 5.6(10)</td>
<td>118.8 ± 36.7(10)</td>
<td>12.6 ± 4.8(10)</td>
<td>[26, 34]</td>
<td>19.5 ± 9(10)</td>
<td>65.5 ± 7.1(10)</td>
<td>0.6 ± 0.10(10)</td>
<td>&gt;125(1)</td>
<td>7m ± 0.55</td>
</tr>
<tr>
<td>Abo BC</td>
<td>−72.7 ± 1.13(13)</td>
<td>83.3 ± 16.7(13)</td>
<td>13.7 ± 4.2(13)</td>
<td>42, 56(13)</td>
<td>14.4 ± 11.8(3)</td>
<td>65.2 ± 7.8(3)</td>
<td>0.4 ± 0.13(3)</td>
<td>–</td>
<td>3.2m ± 0.99</td>
</tr>
<tr>
<td>CA2 SP-SR</td>
<td>−71.1 ± 4.3(68)</td>
<td>82.6 ± 24.4(68)</td>
<td>12.7 ± 9.3(68)</td>
<td>19.9 ± 11.5(68)</td>
<td>67 ± 6.7(68)</td>
<td>&gt;160(1)</td>
<td>3m ± 0.18(68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA1 (14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA1 Pyramidal</td>
<td>−62.4 ± 2.4</td>
<td>65.6 ± 4.4</td>
<td>22.4 ± 1.5</td>
<td>16.6 ± 7</td>
<td>6.8 ± 2</td>
<td>90 ± 4</td>
<td>1 ± 3</td>
<td>&gt;30 ± 3</td>
<td>1.7 ± 0.7</td>
</tr>
<tr>
<td>CA1 Radiusiam Giant</td>
<td>−66 ± 3.6</td>
<td>56 ± 14(25)</td>
<td>50 ± 7(9)</td>
<td>31 ± 2</td>
<td>8 ± 2</td>
<td>70 ± 2</td>
<td>1.9 ± 0.2(7)</td>
<td>&gt;26 ± 8</td>
<td>8.3 ± 0.86</td>
</tr>
<tr>
<td>CA1 Horizontal AA</td>
<td>−57 ± 5</td>
<td>185, 252</td>
<td>[16, 32]</td>
<td>25m</td>
<td>6.0, 13.7,</td>
<td>[71.4, 90.2]</td>
<td>0.7, 11, 87</td>
<td>6.1m</td>
<td>0.74m</td>
</tr>
<tr>
<td>CA1 Basket</td>
<td>−57 ± 5(15)m</td>
<td>116 ± 63(15)</td>
<td>13 ± 8(15)m</td>
<td>19 ± 3</td>
<td>62 ± 3</td>
<td>&gt;60</td>
<td>2 ± 3</td>
<td>0.84 ± 0.06 (15m)</td>
<td></td>
</tr>
<tr>
<td>CA1 BC CCK+</td>
<td>−61.4 ± 3.2(5)</td>
<td>281.48 ± 79.7(5)</td>
<td>25.07 ± 5.6(5)</td>
<td>21.5 ± 3.5</td>
<td>15.17 ± 3.4</td>
<td>76.92 ± 11.75(3)</td>
<td>0.84 ± 0.15</td>
<td>&gt;30</td>
<td>7 ± 2.825</td>
</tr>
</tbody>
</table>

Table 1. Continued on next page
<table>
<thead>
<tr>
<th>EC</th>
<th>[14]</th>
<th>Table 1. Continued</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vmax (mV)</td>
<td>Rm (MΩ)</td>
<td>tmax (ms)</td>
</tr>
<tr>
<td>CA1 Horizontal BC</td>
<td>–55.4 ± 9.5 (17)</td>
<td>116 (19)</td>
</tr>
<tr>
<td>CA1 Ivy</td>
<td>–71</td>
<td>72.8 ± 53.6 (5)</td>
</tr>
<tr>
<td>CA1 LMR</td>
<td>–53.1 ± 4.0 (4)</td>
<td>49 (3)</td>
</tr>
<tr>
<td>CA1 Neurogliaform</td>
<td>–63.1 ± 5.6 (33)</td>
<td>32 (2)</td>
</tr>
<tr>
<td>CA1 NGF Projecting</td>
<td>–63.1 ± 5.6 (33)</td>
<td>21.35 ± 92.8 (32)</td>
</tr>
<tr>
<td>CA1 Recurrent O-LM</td>
<td>–85, –65 (235)</td>
<td>70 ± 13.7 (235)</td>
</tr>
<tr>
<td>Schaffer Collateral-Associated</td>
<td>–55.8 ± 2.8 (100)</td>
<td>96.3 ± 36.0 (100)</td>
</tr>
<tr>
<td>CA1 SO-SO</td>
<td>–59 ± 10 (19)</td>
<td>401 ± 212 (19)</td>
</tr>
<tr>
<td>CA1 Trilaminar</td>
<td>–64 ± 7 (235)</td>
<td>216 ± 124 (235)</td>
</tr>
<tr>
<td>CA1 Radial Trilaminar</td>
<td>–57 ± 5 (15)</td>
<td>116 ± 63 (13)</td>
</tr>
<tr>
<td>L-II Multipolar-Pyramidal</td>
<td>–70, –56</td>
<td>430 ± 121.7 (37)</td>
</tr>
<tr>
<td>L-II Pyramidal-Fan</td>
<td>–62, –59</td>
<td>400 ± 98.9 (90)</td>
</tr>
<tr>
<td>MEC LII PC-Multiform</td>
<td>–70, –56</td>
<td>430 ± 121.3 (37)</td>
</tr>
<tr>
<td>MEC LII Oblique Pyramidal</td>
<td>–62, –59</td>
<td>400 ± 98.9 (90)</td>
</tr>
<tr>
<td>MEC LII Stellate</td>
<td>–62, –59</td>
<td>400 ± 98.9 (90)</td>
</tr>
<tr>
<td>LII-III Pyramidal-Triple</td>
<td>–60, –56</td>
<td>460 ± 78.27 (27)</td>
</tr>
<tr>
<td>LEC ULI Multipolar Principal</td>
<td>–68, –65</td>
<td>450 ± 78.27 (27)</td>
</tr>
<tr>
<td>LII Pyramidal-Stellate</td>
<td>–68, –65</td>
<td>450 ± 78.27 (27)</td>
</tr>
<tr>
<td>ULI-I Bipolar Pyramidal</td>
<td>–69, –67</td>
<td>490 ± 79.28 (28)</td>
</tr>
<tr>
<td>LV-VI Deep Multipolar</td>
<td>–65.02 ± 3.810 (2)</td>
<td>272.3 ± 105.88 (4)</td>
</tr>
<tr>
<td>LV Deep Pyramidal</td>
<td>–65.02 ± 3.810 (2)</td>
<td>75.67 ± 28.710 (2)</td>
</tr>
<tr>
<td>MEC LV-VI PC-Polymorph</td>
<td>–63</td>
<td>480 ± 94 (14)</td>
</tr>
<tr>
<td>LEC LVII Multipolar-PC</td>
<td>–64</td>
<td>450 ± 90 (13)</td>
</tr>
</tbody>
</table>
Browsing starts with summaries of morphological (hippocampome.org/morphology), molecular (hippocampome.org/markers), or electrophysiological (hippocampome.org/electrophysiology) data, similar to hyperlinked versions of Figures 1F, 2, and Table 1, respectively. Searching for neuron types is accomplished through selections from dynamically updated pull-down menus of any combination of properties, such as ‘axons located in DG’, ‘PV-negative’, and ‘V_{mem} > 20 mV’ (Figure 5A). Compound queries with AND & OR Boolean connectors can uncover unexpected results even for the most experienced hippocampal researchers. Searches can also be conducted by PubMed identifiers or author names (Figure 5B). Returned results contain full bibliographic information, a link to the article, and a list of neuron types with information from the article.

Both browsing and searching lead to summaries of all information associated with a given neuron type (Figure 5C): synonyms, morphology, electrophysiology, biomarkers, a representative figure, and known pre- and post-synaptic connectivity (see below). Every property on each neuron page or browse summary links to an evidence page that lists all supporting bibliographic citations complete with extracted quotes, figures, tables, and exact pointers to the relevant pages and paragraphs.

Figure 3. Flow chart of inclusion criteria for neuron types. Beginning with a reconstruction, schematic, or text description of a neuron morphology, the flow chart ends with either a new ‘on hold’ neuron type, supplemental information for an existing active neuron type, or a new active neuron type. Intermediate decision points evaluate the presence of both axons and dendrites, the determination of the main neurotransmitter, the uniqueness of the new type, and whether information is sufficient to create a new active type.

DOI: 10.7554/eLife.09960.006
Figure 4. Quantifying knowledge in Hippocampome.org about morphology, biomarkers, and electrophysiology of hippocampal neuron types. (A) Histograms comparing the sum of pieces of knowledge (PoK) in relevant journal articles or book chapters, in reviews, and in Hippocampome.org. (B) Interconnected knowledge graph of neuron type properties mined from two typical journal articles. (C) Balloon plot of collated knowledge for a majority of GABAergic neuron types. The balloon size indicates the sum of PoK for that type across all three dimensions; balloon color denotes the subregion (as in Figure 1). Note the dearth of biomarker information in entorhinal cortex (EC) and the uneven distribution of data between CA3 and CA1. DOI: 10.7554/eLife.09960.007

(Hippocampome.org contains 13,888 pieces of evidence, including all known sources for many neuron types and properties, but only an adequate number of sources to firmly support established knowledge (e.g., we have only annotated a fraction of all published evidence that granule cells extend axons in CA3 SL).

**Pairwise correlations**

Knowledge integration facilitates the discovery of relations between neuronal properties that would ordinarily remain hidden in the scattered literature. Most PoK in Hippocampome.org, such as axonal presence in a layer or expression of a biomarker, are categorical, and their statistical co-occurrence can be analyzed with contingency tables (‘Materials and methods’). To allow comparison across experiments, electrophysiological parameters are converted (for this analysis only) from continuous to categorical variables by labeling values in the top and bottom one-third of the range, respectively, as high and low. This approach reveals several interesting relationships (Box 1): for example, positive or
negative expression of neuropeptide Y (NPY) tends to co-occur, respectively, with high or low fAHP 
(p < 0.001 with Barnard’s exact test; post hoc t-test p < 0.02 with real values from all n = 42 pieces of
evidence); vasoactive intestinal polypeptide (VIP) is mutually exclusive with CB (p < 0.04); and of
Box 1. Trends assessed from data collated in the Hippocampome.

The trend across clear expression patterns of calbindin (CB), calretinin (CR), and parvalbumin (PV) supports the general idea that only one of these calcium binding proteins is expressed at a time within a cell. While CB is never co-expressed with either CR (p < 0.05, n = 19) or PV (NS, n = 31), PV and CR do co-occur in 2 out of 12 neuron types known to express at least one of them (p > 0.49, n = 31). A less known but more stringent relationship exists between CB and vasoactive intestinal peptide (VIP): among the 24 neuron types for which the expression information on both of these markers is known, 14 are positive for either CB or VIP, but none express both (p < 0.04, n = 24).

Clearly positive expression of neuropeptide Y (NPY) tends to coincide with very high values of the fast afterhyperpolarization (fAHP) (p < 0.001, n = 23) and membrane firing threshold (Vfiring) (p < 0.005, n = 21), and clearly negative expression tends to be associated with very low values of both parameters. fAHP is similarly related also to expression of chicken ovalbumin upstream promoter transcription factor II (CoupTF II) (p < 0.005, n = 14), which tends to clearly co-express either both positively or both negatively with NPY (p < 0.005, n = 18).

Confirming a well-known relationship, of 45 cell types with extreme values of input resistance (Ri) and membrane time constant (τm) (p < 0.001, n = 45), 43 are positively correlated (R² = 0.70). Likewise, high values of the spike width (APWidth) are associated with high values of Ri (p < 0.001, n = 45), and vice versa for the low values of both (R² = 0.36).

Of the 69 cell types with axons in more than one layer, only DG Granule and CA1 O-LMR do not have overlapping axons and dendrites (p < 0.001, n = 122). All of the 43 cell types that have axons in three or more layers have axons that overlap with their dendrites (p < 0.001, n = 122). Furthermore, cell types with axons in three or more layers tend to have very high Ri (p < 0.001, n = 58) and τm (p < 0.001, n = 59).

Clear expression of cannabinoid receptor 1 (CB1) implies (in 5/5 cases) clear expression of cholecystokinin (CCK) (if known), and clear absence of CCK implies (in 8/8 cases) clear absence of CB1 (if known) (p < 0.005, n = 15). Of the two neuron types expressing CCK and not CB1, one expresses VIP; however, independent of CCK expression, CB1 and VIP are not co-expressed in any of the eight types that are positive for either biomarker (NS, n = 17). Overall, these patterns suggest that CCK+VIP+ and CCK+CB1+ neurons constitute completely separate groups.

The p values and sample sizes (n) pertain to Bernard’s exact test on 2 x 2 contingency tables (see Materials and methods).

DOI: 10.7554/eLife.09960.009

31 entorhinal neuron types, only Oblique Pyramidal cells in LII of medial EC have no layer overlap between axons and dendrites.

Many observed trends also reinforce expected associations and serve as validation for the dataset. For example, 96% (43/45) of neuron types with high or low input resistance have correspondingly high or low τm, consistent with RC-circuit theory. The two outliers correspond to data from a paper on CA1 Radium Giant cells (Kirson and Yase, 2000), for which the particularly high τm/Ri ratio is explained by the oversized soma and consequently large capacitance.

Potential connectivity

Hippocampome.org contains known connectivity information (synapses or lack thereof, established e.g., by electron microscopy or paired electrophysiological recordings) among its neuron types. Such data, however, are only available for 202 out of 14,884 possible pairs (<2%). In the remaining cases, co-existence of axons and dendrites across parcels allows inference of ‘potential connectivity’.
between cell types (Braitenberg, 1991). Specifically, potential connectivity is computed as the dot-product of the 26-dimensional binary vectors encoding the presence/absence of axons for one neuron and dendrites for another across hippocampal parcels. The possibility of connections can be tentatively excluded for 11,668 pairs of neuron types (78.4%) based on non-overlapping distributions of their respective axons and dendrites. However, incomplete axons and the binary thresholding of the morphological encoding (see ‘Materials and methods’) might yield false negatives. In 3014 pairs (20%) the axons and dendrites of the two neuron types share at least one parcel. Although the connection probability may not be estimated in the absence of experimental evidence, even the opportunity to make synapses is computationally relevant due to the superior structural plasticity of the hippocampal formation throughout the lifespan (Leuner and Gould, 2010).

All (known and potential) connectivity data are summarized in a matrix with rows as pre-synaptic and columns as post-synaptic neuron types (100 neuron types in Figure 6A; complete data: hipcampomem.org/connectivity). Filled boxes along the main diagonal, corresponding to neurons with axons and dendrites co-located in any parcel (purple boxes in Figure 1F), indicate within-type connectivity (not necessarily autapses), and are more frequent among excitatory (87%) than inhibitory (67%) types.

Neuron type connectivity as depicted in Figure 6A constitutes an intermediate level of description between the neuron-by-neuron connectome (‘synaptome’) and the region-by-region connectome (‘projectome’). In order for two regions to be connected, there must be at least one neuron type in one region connecting to one neuron type in the other (Martone and Ascoli, 2013). Within the hippocampal formation, this relationship is illustrated explicitly in Figure 6B: all six areas have potential excitatory, and with the exception of Sub and EC, inhibitory connections to at least one other area, most of which are not considered in the canonical trisynaptic loop (cf. Bums and Young, 2000; van Strien et al., 2009). The only potential monosynaptic connection between DG and CA1 is inhibitory. While this regional wiring diagram appears simpler than the corresponding seminal illustration for the visual cortex (Felleman and Van Essen, 1991), the underlying neuron type circuitry is highly complex. In fact, the explicit wiring diagram of all neuron type potential connections is impractical to render on a single page. Selecting only 15 representative neuron types in the DG entails 253 connections among 43 somatic and dendritic compartments (Figure 7A). This compartmental representation of network connectivity captures the computationally distinct subunits that emerge from layer-specific axonal targeting, a key component of mesoscopic neuron type circuitry.

Use cases

Hippocampomem.org can deliver valuable information in many day-to-day research scenarios (Box 2) and at multiple stages of a neuroscience study (hipcampomem.org/usage). They include determining whether a hippocampal neuron with a particular morphological pattern is known, obtaining a list of candidate neuron types based on partial reconstructions, and finding biomarker and electrophysiological properties for most neuron types. In all highlighted use cases, Hippocampomem.org delivers the needed information with very few mouse clicks, in a matter of seconds. Even queries that are nearly impossible with literature search engines (e.g., finding an unnamed neuron by its axonal pattern) allow, at Hippocampomem.org, straightforward retrieval of the original evidence.

Discussion

Systematic organization of present knowledge on neuron types might revolutionize neuroscience akin to the impact the Periodic Table of the Elements had on chemistry 150 years ago. Key to this endeavor is the judicious selection of pivotal variables with the most discriminant and explanatory power. Notably, the initial choice may prove to be a useful proxy even if later recognized as unprincipled:
**Figure 6.** Neuron type connectivity (area color coding and numbering as in Figure 3). (A) Known and potential connections for 100 neuron types (full matrix: hippocampus.org/connectivity or hippocampus.org/php/images/connectivity/Connectivity_Matrix.jpg), with pre-synaptic types in rows and post-synaptic types in columns. Black squares indicate potential glutamatergic connections and gray squares GABAergic. Pairs of neuron types with experimentally established and refuted synapses are shown using green and red squares, respectively. Colored boundaries demarcate intra-area

Wheeler et al. eLife 2015;4:e09960. DOI: 10.7554/eLife.09960
Mendeley’s first attempt to order chemical elements by atomic mass subsequently led to predicting their correct atomic numbers four decades before the discovery of the proton. The distribution of axons and dendrites across anatomical areas and layers is known for the majority of hippocampal neurons identified to date and effectively discriminates among them. Among the features commonly used to describe cortical neurons, morphology is also considerably robust to experimental conditions. Furthermore, just like the positions of the elements in the Periodic Table indicate their potential to combine into molecules, axonal and dendritic patterns provide the blueprint for the potential network connectivity.

Our classification scheme links biochemical, physiological, and synaptic data to structural knowledge of neuron types by methodical literature mining. The cost of this integration is the need to interpret published information to abstract specific evidence from neuron instances into general properties of neuron types. The resulting knowledge base constitutes the first comprehensive machine-readable neuron inventory for a mammalian cortical region. Parallel efforts are underway for the retina (Siegert et al., 2009), somatosensory (Markram, 2006) and visual cortex (celltypes.brain-map.org), and drosophila (dos Santos et al., 2015). Complementary synergies in the rat hippocampal formation include Temporal-lobe.com (van Strien et al., 2009), Rat Hippocampus Atlas (Kjönigsen et al., 2011), and Hippocampus 3D (Ropreaddy et al., 2012).

The vision driving Hippocampome.org is a real-scale computer model of the entire hippocampal formation. Accurate simulations require knowledge of the component parts, their locations, numbers, properties, and connectivity. We have started with an accounting of the neuron types defined by their most essential morphological, molecular, and electrophysiological features. Further works-in-progress include neuron counts (Bezaire and Soltész, 2013), firing patterns, synaptic profiles, developmental origins (Tricoire et al., 2011), and rhythmic phase-locking (Somogyi and Klausberger, 2005). Moreover, the binary assignment of axonal and dendritic distribution across parcels will gradually be substituted by quantitative morphological estimates (Ropreaddy and Ascoli, 2011).

Neuronal classification will likely benefit by advances in optogenetics, sequencing, and machine learning (Fenno et al., 2014; Kohara et al., 2014; Roux et al., 2014; Armalíanzas and Ascoli, 2015; Zeisel et al., 2015). Rapid growth is expected particularly in the knowledge of molecular markers. Thousands of the ~20,000 genes mapped in the Allen Mouse Brain Atlas (Lein et al., 2007) are expressed in the hippocampus. Connecting such large-scale information to the morphology and physiology of particular neuron types will likely answer many open questions while raising new ones.

By integrating available information, Hippocampome.org aims to accelerate discovery. Neuro-Morpho.Org (Ascoli et al., 2007), a digital archive of neuronal reconstructions, has enabled numerous secondary findings from amounts of data far outweighing the collection means of individual labs (Parekh and Ascoli, 2014). While this article offered selected examples, further in-depth analyses of neuron type connectivity and pairwise property correlations are ongoing.

Consolidation of knowledge also allows assessing the available information to prioritize missing data in the hippocampal formation (cf. Olshausen and Field, 2005 for V1). The limited interneuron diversity in CA3, CA2, Sub, and EC compared to CA1 suggests several to-be-discovered neurons. Many cell types found in CA1 (e.g., Basket, O-LM, Ivy, Neurogliiform, Axo-axonic, Trilaminar, and Bistratified) have counterparts in some other area(s), but they are probably present in more. As experimental techniques advance, several new neuron types will likely be reported. The Hippocampome.org framework also sets a standard for the minimal information that must be included in publications describing hippocampal neuron types in order to link the resulting data to the existing body of knowledge.

Initially enforcing strict information granularity has proven critical for the launch of this resource by preventing unending mining of published neuronal properties with clear-cut delimiters. The germinal selection of information to include in Hippocampome.org was pragmatically guided by the density of available published data. For example, in vivo neuronal reconstructions and recordings, although
Figure 7. Neuron type circuitry. (A) Circuit diagram of selected neuron types in DG (full diagram: hippocampome.org/php/images/connectivity/DG_Circuit_Diagram.png) or hippocampome.org/php/images/connectivity/DG_Circuit_Diagram.graphviz.zip). Axonal connections from pre-synaptic somata (orange glutamatergic, green GABAergic) to post-synaptic somata or dendrites (blue) have unique line colors for each pre-synaptic type for clarity. Lines ending in arrows and open circles indicate, respectively, glutamatergic and GABAergic connections; 22 known (thick lines) and 231 potential (thin) connections.

Figure 7 continued on next page.
connections are depicted. 1: Granule, 2: Semilunar Granule, 3: Mospy, 4: AIPRM, 5: A xo-axonic, 6: Basket, 7: Basket CCK+, 8: HICAP, 9: HIP, 10: HIPROM, 11: MOCAP, 12: MOLAX, 13: MOPP, 14: Neurogliafilm, 15: Outer Molecular Layer. (B) Pre- and post-synaptic connections for DG Granule and CA3 Basket cells. Red lines ending in arrows indicate glutamatergic connections, and blue lines ending in open circles indicate GABAergic connections. Thick and thin lines indicate, respectively, known and potential connections. Neuron types are color coded by area of origin (as in Figure 1).

DOI: 10.7554/eLife.09960.011

extremely valuable, are too sparse to be adopted as defining dimensions in a comprehensive census of neuron types. Similarly, axonal and dendritic patterns are more commonly identified across layers than in the longitudinal (septo-temporal) and transverse (e.g., CA3a,b,c) axes. In limited cases, information is sufficient to distinguish specific neuron types based on finer anatomical sub-divisions. For instance, CA3c Pyramidal cells are differentiated from other CA3 Pyramidal cells, because of their axonal back-projection to DG and lack of dendrites in SLM. Some neuron types are similarly separated between medial and lateral EC.

At the same time, Hippocampe.org is well equipped to manage the expected ‘data deluge’ from ongoing big science projects (Kandel et al., 2013). With accelerating knowledge expansion, our linking protocol to interrelate each additional piece of data with pre-existing information will ensure continuous integration, facilitating user-friendly analysis and modeling. Moreover, Hippocampe.org infrastructure is designed to accommodate additions along many dimensions: new areas (e.g., septum), finer parcels (medial vs lateral EC), additional neuron types (hippocampe.org future), and graded reporting of neurite densities. The open-source code of this resource also encourages community-led extension to other brain regions, inter-operability with related initiatives, and progressive adoption of unique identifiers for neuron types and properties.

Materials and methods

Anatomical parcels and neuron type identification

Several attempts are underway to establish definitive layer and area boundaries through molecular expression (Thompson et al., 2008; Boccara et al., 2014). Until formal definitions are in place, generally agreed upon anatomical and histochemical characteristics provide for the delineation of the hippocampal formation into six areas (DG, CA3, CA1, Sub, and EC) and their respective layers (e.g., CA1 SO, SR, etc.; Figure 1A,8). Most of these parcels are standard and can be found in all reviews of the hippocampal formation.

Part of the DG literature only defines a single s. molecule (SM) between the fissure and SG. Hippocampe.org divides SM into the outer two-thirds (SMo) and the inner one-third (SMi), because SMo receives perforant path input from EC and SMi receives hilar collaterals from mossy cells (Amaral et al., 2007).

In Sub, Hippocampe.org lists a single SM superficial to SP and a PL below SP (van Strien et al., 2009), differing from the Allen Mouse Brain Atlas (Lein et al., 2007) delineation of two parcels above SP and none below.

A neuron type must satisfy four criteria to be included in Hippocampe.org. (1) The soma must be located in the hippocampal formation. (2) The major neurotransmitter must be at least tentatively specifiable; at present, only glutamatergic and GABAergic types have qualified for inclusion in Hippocampe.org (see below for cholinergic). (3) The locations of both axons and dendrites within any of the 26 hippocampal parcels (Figure 1A,8) must be clearly presented, either textually or graphically. (4) Experimental evidence of more than a single neuron of the type must be reported.

Hippocampe.org neuron types are primarily distinguished by criterion 3 (axonal-dendritic patterning) independent of author provided names or grouping, but are further differentiated by their synaptic specificity (see Figure 3). Neurite densities within parcels, beyond a minimal threshold (see below), are not currently a factor, as most reconstructions are incomplete; this first approximation is likely to change when more quantitative data become available for most neuron types. All evidence collated in Hippocampe.org comes from peer-reviewed articles or book chapters (hippocampe.org/bibliography) and pertains to healthy rodents, predominantly rats and mice ≥13 days old.

Wheeler et al. eLife 2015;4:e09960. DOI: 10.7554/eLife.09960 19 of 28
Box 2. Representative Hippocampome.org use case scenarios.

Property-based recognition of known neuron types

While patching in DG stratum granulosum (SG), a researcher encounters several interneurons with high input resistance ($R_m$), membrane time constant ($\tau_m$), and spike amplitude ($A_{\text{pm}}$). Biotin filling reveals both axonal and dendritic presences in the inner stratum moleculare (SM) but only axons in the hilus (H). Is this a newly discovered neuron type?

Hippocampome.org

These characteristics are consistent with those reported for MOCAP neurons, first described by Markwardt et al. (2011) as non-ly/NG cells.

Comprehensive listing of potential pre-synaptic sources

In the presence of outward channel blockers, CA1 Pyramidal cells display rebound spikes upon repeated GABA puffing on their distal apical dendrites. What neurons besides O-LM and Neuroglialform cells might trigger such post-inhibitory firing?

Hippocampome.org

14 different interneurons have axons in stratum lacunosum-moleculare (SLM), including CA1 Quadrlaminar, Back-projection, Radiatum-receiving Apical-targeting, Perforant Path-associated, Orients-Bistratified Projecting, and LMR cells.

Comprehensive listing of potential post-synaptic targets

Glutamatergic Cajal–Retzius neurons have recently been observed in the adult rodent hippocampus in larger numbers than previously assumed. What interneurons could they excite in CA1?

Hippocampome.org

Although the axons of CA1 Cajal–Retzius are confined to SLM, they could nonetheless contact no fewer than 17 types of distinct GABAergic cells, such as Basket, Axo-axonic, Radial Trilaminar, Orients/Alveus, Schaffer Collaterals-associated, as well as several interneuron-specific interneurons.

Discrimination of perisomatic neuron types by electrophysiological measures

CA3 Basket and Axo-axonic neurons have similar somatic locations, dendritic tree shapes (invading all layers from stratum orenis (SO) to SLM), and overall axonal patterns. Short of the labor-intensive determination of their post-synaptic targets, could they be tentatively distinguished by patch-clamp recording?

Hippocampome.org

While CA3 Axo-axonic and Basket cells have practically indistinguishable resting potential ($V_{\text{rest}}$), $\tau_m$, and $A_{\text{pm}}$, the former tend to display lower firing threshold ($V_{\text{threshold}}$) as well as greater slow afterhyperpolarizations ($sAHP$) and $R_m$.

Positive identification of neuronal phenotypes

Two varieties of CA2 basket cells have been described: a classic type with dendrites confined within CA2 and a ‘wide-arbor’ type, whose dendrites enter CA1 and CA3. How can the borders between CA2 and adjacent subregions be reliably demarcated?

Hippocampome.org

CA2 pyramidal neurons, unlike their CA1 and CA3 counterparts, are positive for Purkinje cell protein 4 (PCP4). Conversely, pyramidal neurons in CA1 and CA3 express $\alpha$-mannosidase-I.
(Man1a), while those in CA2 do not. Immunostaining with either of these markers can thus delineate CA2 boundaries.

**Constraining simulation parameters**

To build a circuit model of grid cell activity, a computational neuroscientist is searching for plausible values for $V_{\text{rest}}$, $V_{\text{thresh}}$, fast afterhyperpolarizations (fAHP), and $AP_{\text{ampl}}$ for the principal cells and main interneurons above the lamina dissecans (LIV) of the medial entorhinal cortex (MEC).

**Hippocampome.org**

These data are available for MEC I/II Stelelate cells and EC I/II Pyramidal cells (the two major glutamatergic neurons), for EC I/II Basket-Multipolar and MEC I/II Superficial Trilayered interneurons (the major perisomatic and dendritic-targeting GABAergic cells, respectively), as well as for eight additional, if less prominent, neuron types (5 excitatory and 3 inhibitory) of layers I-III.

**Distilling information relevant to specific hypotheses**

A novel theory of hippocampal function requires direct feedback inhibition from CA3 and CA1 to DG Granule cells, contrary to the canonical triynaptic loop and the common assumption of non-projecting GABAergic cells. Does the literature provide any experimental evidence to support the new assumption?

**Hippocampome.org**

Based on axonal–dendritic overlap, both the Granule cell page and the connectivity matrix indicate as (potential) sources of monosynaptic input one CA3 (Spiny Lucidium) and three CA1 (LMR Projecting, Perforant Path-associated, and Neurogliaform Projecting) neurons. Of these, Perforant Path-associated neurons have already been shown to form synapses onto DG Granule cells (Vida et al., 1998).

See [hippocampome.org/usage](https://hippocampome.org/usage) for an extended example.

DOI: 10.7554/eLife.09960.012

When incomplete information prevents inclusion in Hippocampome.org, neuron types are placed ‘on hold’ (hippocampome.org/on-hold). For example, none of the references providing evidence for cholinergic neurons in CA1 clearly describe the axonal and dendritic arbors of individual neurons. On-hold types also include neurons from animals younger than P13 (e.g., DG Cajal–Retzius cells).

**Encoding of neuronal morphology**

Hippocampome.org binarizes the locations of axons and dendrites in the 26 parcels: neurites either have ‘sufficient’ presence in a given parcel or they do not. Lexical and visual thresholds, briefly described here, specify sufficiency criteria (full explanation and examples: hippocampome.org/full-interp). For neuronal reconstructions and schematics, a layer must contain either $\geq$15% of the overall arbor or at least half the amount included in the most abundant layer. Moreover, the neurite must penetrate $\geq$15% of the layer depth. In particular, the axonal tree of the CA1 basket cell can spill over from SP into SO and SR, but the penetration threshold is not crossed (cf. Figure 3E from Pawelzik et al., 2002). Thus, if the layers are not delineated, the figure may be unusable for neuron typing.

For text descriptions, categorical statements such as ‘The dendrites (of CA1 Horizontal Basket cells) are restricted to stratum oriens’ (Maccari et al., 2000), are straightforward. However, quotes that are ambiguous in terms of neurite quantity within a parcel require an interpretive threshold. The term ‘most’ referring to branches is considered as evidence for presence within that parcel, while the term ‘some’ is not, although the quote is still included in Hippocampome.org as evidential information. Interpretation of other equivocal statements is carefully annotated.
Although the authors of an article may group a set of neurons together, differing axonal and dendritic patterns can lead to multiple Hippocampome neuron types. This occurs if the number of neurons in each subset is greater than the square root of the total number of neurons described (e.g., if 4 neurons out of 15 have distinct neurite distributions, they would be split into a different cell type because $4 > \sqrt{15}$). For instance, DG neuroglialform cells (Armstrong et al., 2011) are divided into two separate neuron types: those that extend their axons into Sub (keeping the name 'DG Neuroglialform') and those that remain local within DG SMo. This latter type is merged with MOPP cells, which have the same axo-dendritic pattern in DG and no projection. Future accumulation of neurite density, molecular expression, and electrophysiology data may result in separating local DG neuroglialform and MOPP cells.

Conversely, neurons assigned to different groups in a paper might belong to the same Hippocampome type, such as R-LM and P-LM cells (Oliva et al., 2000), that differing only in somatic location, are merged into a single CA1 OR-LM type in Hippocampome.org.

Nomenclature

Hippocampome.org neuron names were assigned progressing from the most to the least prominent types. When a single name is used in the literature, that name is adopted into the Hippocampome.org neuron name. For example, DG granule cells in the knowledge base are referred to as DG (e)2201p-CA3_00110 Granule, a formal name encoding the pattern of axons and dendrites (hippocampome.org/brief-interp), and Granule, a common name.

The formal name is composed of multiple parts (hippocampome.org/formal-name), the first of which (e.g., DG (e)2201p) encodes information for the area where the neuron resides: DG designates the home area of the soma; (e) informs about the putative major neurotransmitter being glutamatergic; 2201 is the axo–dendritic pattern in the home area ordered from the most superficial to the deepest layer (e.g., dendrites in SMo and SMi, no neurites in SG, and axons in H); and the p indicates that the axons and/or dendrites project out of the home area. The second part of the formal name encoding (e.g., CA3_00110) describes the pattern of the projecting neurites (e.g., axons in SL and SP of CA3). The final portion of the formal name is the same as the common name.

Common neuron names in Hippocampome.org are assigned a prefix denoting their hippocampal area if the same name is used for neurons in different areas (e.g., CA3 Bistratified and CA1 Bistratified), but not if the name is unique (as in the case of HIPP cells, which only exist in DG). Furthermore, the prefix MEC is assigned to neurons that have been characterized solely (or are known to exist predominantly) in the medial EC, such as MEC Layer II Stellate cells. Similarly, the prefix LEC is assigned to lateral entorhinal neurons. In contrast, neurons that are believed to exist in both the medial and lateral ECs are generically referred to with the prefix EC.

When multiple names appear in the literature for the same neuron type, the most frequent or best-known is selected, such as CA1 O-LM over CA1 oriens interneurone of the second type (McBain et al., 1994) or CA1 horizontal oriens-rectus interneurone (All and Thomson, 1998). If no literature name clearly emerges, or if all suitable names are already adopted for more prominent neuron types, we adopt or combine multiple author-originated names, as in EC II Pyramidai-Fan (Germroth et al., 1991; Lengenhjelm and Finch, 1991; Empson et al., 1999; Tabeshdari and Alonso, 2000). All terms used in publications are always reported as synonyms in Hippocampome.org.

Linking molecular and electrophysiological data to neuronal morphology

To search for molecular and electrophysiological data, the articles defining the axonal–dendritic pattern of each Hippocampome.org neuron type are mined first. All references citing (webmknownedge.com) or cited by these sources are mined next. Lastly, full-text searches using all known synonyms of the target property and neuron type are performed (scholar.google.com and previously scirus.com). If these searches fail to return usable molecular or electrophysiological information for a given neuron type, the corresponding property is labeled as ‘currently unknown’ (gray boxes in Figure 2 and empty entries in Table 1).

For inclusion in the knowledge base, molecular and electrophysiological data must be linked to a morphologically defined neuron type. As with the establishment of morphological types, the linking process is blind to neuron names used by authors. Rather, links require either the co-presentation of axonal–dendritic information or the citation, for that specific evidence, of a source that has a morphological description of the neuron. Only in two cases can linking be achieved without complete knowledge of the morphology. The first applies to principal cells within their layers (i.e., granule cells
in DG and pyramidal cells in CA1 SP). These neuron types are readily identified as positive or negative for a biomarker because the layers are >90% homogeneous for the principal cells (Cepek et al., 2013) and the somata are densely packed. In the second case, certain groups of neurons can be identified by their axonal tracts, for example, granule cells by their mossy fibers and medial EC LII stellate cells by their perforant path projection. In all cases, the linking information is explicitly included with the evidence reported at Hippocampome.org.

**Molecular biomarker expression**

The 20 most studied biomarkers in hippocampal research (Figure 2; extended listing: hippocampome.org/markers) were targeted for literature searches across all 122 neuron types. These include calcium-binding proteins (PV, CB, CR), receptors and transporters (CB1, sub P rec, muscarinic receptor 2, serotonin receptor 3, vesicular glutamate transporter 3, metabotropic glutamate receptor 1α, GABA<sub>α</sub>, α1 subunit), neuropeptides (CCK, SOM, enkephalin, NPY, VIP, neurogranin), and a miscellaneous group of cytoskeletal and extracellular matrix proteins (α-actinin 2, reelin), transcription factors (COUP-TFI), and enzymes (nNOS).

When information is available, a particular neuron type is characterized as positive, negative, or mixed positive-negative for a biomarker. In the case of mixed expression, the data are evaluated to determine whether the mixed information might be attributed to differences in species (e.g., rats vs mice), techniques (e.g., protein detection vs mRNA detection), or subcellular localization (soma vs neurite). When a single population of neurons is shown to be divisible into clearly negative and clearly positive subpopulations, this is taken as an indication of biomarker subtypes. When we are unable to determine whether mixed biomarker data is attributable to species/technique/subcellular localization differences or subtypes, the data are annotated as unresolved.

**Electrophysiological properties**

Electrophysiological property values are extracted from the literature and compiled, when available, for all Hippocampome.org neuron types. The knowledge base includes passive (R<sub>mem</sub>, τ<sub>mem</sub>, resting membrane potential or V<sub>rest</sub>), spike (AP<sub>amp</sub>, AP<sub>width</sub>, V<sub>peak</sub>), (AHP), and other response parameters (maxFR, sAHP, hyperpolarization sag ratio). Values are extracted from published reports either from text (or tables) or by digitizing figures (plotdigitizer.sourceforge.net) and reported as mean, range, standard deviation, and number of measurements. Although these properties can be defined and measured in multiple ways, Hippocampome.org standardizes the data according to a single definition (hippocampome.org/eophys-defs). This standardization, along with full hand-curation, distinguishes our approach from the semi-automated mining by NeuroElectro (Tripathy et al., 2014).

**Maintenance and further development**

The growth, evolution, and accuracy of Hippocampome.org content rely on two main mechanisms: continuous literature mining (hippocampome.org/ongoing-mining) and community feedback (hippocampome.org/feedback). We update the bibliographic listing at quarterly intervals (based on citation alerts for many of the core review articles as well as perusals of new issues of most mainline neuroscience publications), adding new relevant references and linking to the knowledge base the articles that have been annotated. Moreover, we welcome suggestions for improvements, corrections, and additions. Addressing of this feedback will also be incorporated into the Frequently Asked Questions (FAQs) listing (hippocampome.org/FAQ) for future reference by all users. In order to ensure reliable resource citation, we adopt a numbered versioning system to release additions of new neuron types and specification of additional properties for existing neuron types. Publication of this article marks the v.1.0 release of Hippocampome.org.

**Web portal and database**

The web portal and associated database infrastructures facilitate access to and utilization of morphological, molecular, electrophysiological, and connectivity information. The implementation leverages the model-view-control software design. The model component, which defines the database interface, is provided solely by server-side code. The view component, which renders the web pages, and the control code, which implements decision logic, are both served up by the server,
but are run in the user’s browser. The underlying relational database ensures flexibility in establishing relations between data records.

Hippocampome.org is deployed on a CentOS 5.11 server running Apache 2.2.22 and runs on current versions of Firefox, Chrome, Safari, and Explorer. Knowledge base content is served up to the PHP 5.3.27 website from a MySQL 5.1.73 database. Django 1.7.1 and Python 3.4.2 provide database ingest capability of comma separated value annotation files derived from human-interpreted peer-reviewed literature. Hippocampome.org code is available open source at github.com/Hippocampome.org.

Pairwise correlation analysis
We explored pairwise correlations between 205 properties of Hippocampome neuron types, including neurotransmitter; axonal, dendritic, and somatic locations in the 26 partitions and 6 areas; the projecting (inter-areas) or local (intra-area) nature of axons and dendrites; axon and dendrite co-presence within any partition; axonal and dendritic presence in a single layer only or in ≥3 layers; clear positive or negative expression of any biomarkers; and high (top third) or low (bottom third) values for seven electrophysiological properties (excluding highly stimulus-dependent sag ratio, saHNP, and maxFR). To evaluate the correlations between these categorical properties, we use 2 x 2 contingency matrices with Barnard’s exact test, which provides the greatest statistical power when row and column totals are free to vary (Lydersen et al., 2009).

Acknowledgements
We thank Maurizio Bergamino, Faramarz Faghihi, Ben Holmes, Sean Macksey, Keivan Moradi, Mona Sullman, Siva Venkadesh, and many student interns (hippocampome.org/thx) for invaluable help. This project is supported by grants R01NS39600 (NIH), MURI N00014-10-1-0198 (ONR), Nakfi (Keck), CENTEC (AFOSR), Robust Intelligence (NSF), and Northrop Grumman.

Additional information

<table>
<thead>
<tr>
<th>Funding</th>
<th>Grant reference</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Institutes of Health (NIH)</td>
<td>R01NS39600</td>
<td>Giorgio A Ascoli</td>
</tr>
<tr>
<td>Office of Naval Research (ONR)</td>
<td>MURI N00014-10-1-0198</td>
<td>Giorgio A Ascoli</td>
</tr>
<tr>
<td>Air Force Office of Scientific Research (AFOSR)</td>
<td>CENTEC</td>
<td>Giorgio A Ascoli</td>
</tr>
<tr>
<td>National Science Foundation (NSF)</td>
<td>Robust Intelligence</td>
<td>Giorgio A Ascoli</td>
</tr>
<tr>
<td>Northrop Grumman</td>
<td></td>
<td>Giorgio A Ascoli</td>
</tr>
<tr>
<td>Keck</td>
<td>Nakfi</td>
<td>Giorgio A Ascoli</td>
</tr>
</tbody>
</table>

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions
DWW, CMW, CLR, AOK, GAA, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; DJH, Conception and design, Analysis and interpretation of data, Drafting or revising the article

Author ORCIDs
Dielk W Wheeler, http://orcid.org/0000-0001-8635-0033
David J Hamilton, http://orcid.org/0000-0001-6265-9917

References


Name-calling in the hippocampus (and beyond): coming to terms with neuron types and properties

D. J. Hamilton · D. W. Wheeler · C. M. White · C. L. Rees · A. O. Komendatov · M. Bergamino · G. A. Ascoli

Received: 11 March 2016 / Accepted: 24 May 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Widely spread naming inconsistencies in neuroscience pose a vexing obstacle to effective communication within and across areas of expertise. This problem is particularly acute when identifying neuron types and their properties. Hippoccampome.org is a web-accessible neuroinformatics resource that organizes existing data about essential properties of all known neuron types in the rodent hippocampal formation. Hippoccampome.org links evidence supporting the assignment of a property to a type with direct pointers to quotes and figures. Mining this knowledge from peer-reviewed reports reveals the troubling extent of terminological ambiguity and undefined terms. Examples span simple cases of using multiple synonyms and acronyms for the same molecular biomarkers (or other property) to more complex cases of neuronal naming. New publications often use different terms without mapping them to previous terms. As a result, neurons of the same type are assigned disparate names, while neurons of different types are bestowed the same name. Furthermore, non-unique properties are frequently used as names, and several neuron types are not named at all. In order to alleviate this nomenclature confusion regarding hippocampal neuron types and properties, we introduce a new functionality of Hippoccampone.org: a fully searchable, curated catalog of human and machine-readable definitions, each linked to the corresponding neuron and property terms. Furthermore, we extend our robust approach to providing each neuron type with an informative name and unique identifier by mapping all encountered synonyms and homonyms.

Keywords Hippocampus · Neuron · Type · Property · Nomenclature

1 Introduction

From its beginning, neuroscience has been tied to ad hoc neuron naming, which is subject to the whims of researchers with diverse interests. It has always been the inclination of neuroscientists to name neurons based on certain observed properties. Already in the 1800s, researchers leveraged ongoing progress in optical microscopy and newly discovered staining techniques to identify neuron types and their morphological features. Historical examples include Betz’ naming of “giant pyramids” [1] and Cajal’s description of “psychic cells” (nowadays known as pyramidal neurons) as characterized by “…a dendritic shaft and tuft directed toward the cerebral surface [and] the existence of collateral spines on the dendritic processes…” [2]. Thousands of reports describing neurons and their characteristics have been published since, and several dozens of distinct types of neurons had been already recognized before the turn of the millennium in each of several prominent neural systems, such as among the “GABAergic non-principal cells” of the hippocampus [3].

The often subjective and arbitrary naming of neurons led to a cluttered literature landscape in which breakdowns in
communication can hinder the understanding of the structure and function of the brain. A comprehensive solution would require establishing a broadly applicable and widely accepted classification scheme defining neuron types based on their properties. However, despite early efforts focused on identifying key neuronal properties with precise terminology [4], to this date there is a high level of disorganization when it comes to reporting neuronal property information. Although community efforts exist for the expert curation of neuroanatomical terms pertaining to brain regions [5] and grass root scholarly collation of neuroscience terminology [6], the continuously increasing pace of data acquisition is paradoxically yielding an ever more fractured lexicon, creating serious impediment to progress.

We have previously proposed an ontological approach to defining neurons based on necessary and sufficient part-relation-value triple-store techniques [7]. In the absence of comprehensive data and unbiased sampling, however, it may be impossible to select a priori the appropriate defining properties [8]. Using too few or too many constraints results in under-defining or over-defining a neuron type. The former case (“over-lumping”) leads to a few large groups of neurons that share very few properties; the latter (“over-splitting”) leads to myriad types of doubtful interpretation. To complicate this matter further, the continuous gradation of key properties may require a shift to fuzzy classification approaches [9].

A recent empirical assessment of inter-investigator agreement on morphological classes of neocortical interneurons demonstrated a variable level of consensus across neuron types and properties [10]. One of the most reliable identifiers of neuron types is the presence or absence of axons and dendrites within well-defined neuroanatomical boundaries. Spotlighting this, Hippocampome.org [11] recently established unambiguous definitions of neuron types primarily based on axonal and dendritic distributions across all the main subregions and layers of the hippocampal formation. This classification approach yielded an initial catalog of 122 neuron types identified from the scientific literature. It is important to stress that the classification criteria employed by Hippocampome.org operate independently of previously used names.

In this framework, a neuron type is initially identified by its (putative) neurotransmitter and the presence of axons and dendrites in the distinct layers of dentate gyrus, CA3, CA2, CA1, subiculum, and entorhinal cortex. Each type is further characterized by available information on biomarker expression and electrophysiological features. This relatively simple characterization allows dense curation of the published literature through text mining and annotation. The resulting information is instantiated as a machine-readable electronic relational knowledge base that is publicly and freely available, facilitating web accessibility and computational analytics. With critical properties compiled in an easily accessible portal, Hippocampome.org provides a unique opportunity to establish a consistent set of definitions and a naming protocol that could be expanded to other cortical areas, aiding research and scientific communication.

The remaining of this report is organized as the following. The next section provides illustrative examples of the terminological confusion regarding neuron types and properties from the hippocampal literature. The following section outlines the three steps toward a solution: first, we describe the design of a database to define, store, browse, search, and retrieve human-interpretable but machine-readable definitions of neuron types based on their properties, as recently implemented at Hippocampome.org. Second, we introduce a newly deployed functionality that maps all relevant property terms to corresponding concepts, linking their occurrence in the published evidence to community-accepted definitions. Third, we offer a formal definition of the resulting neuron types and detail the process to assign each of them with a unique common name. The last section closes the paper with concluding remarks.

2 A neuronal “Tower of Babel”

The nomenclatures of neuron types and of their features are both vexed with ambiguities, resulting in a “many-to-many” mapping between neurons and names as well as inconsistent definitions of properties. We illustrate below representative examples of the most common scenarios from the hippocampal literature.

When neurons are described in a publication, they are typically named in isolation, out of context with respect to the rest of the brain circuit and the literature. Sometimes neuron types or individual neurons are indicated solely by a non-descriptive label (e.g., “Type I” cells [12] or “cell #7” [13], and occasionally they are not named at all. When proper terms are used, it may still be difficult to discern whether a word is meant to be a name or merely a description, as when referring to “multipolar cells” [12, 14]. The result is often a baffling web of associations between names and neuron types.

Consider for instance the term “CA1 Bistratified cell originally chosen over 20 years ago to name a group of hippocampal neurons with axons and dendrites prominently invading the oriens and radiatum layers without crossing into lacunosum-molecular [15]. Different authors later used the exact same noun referring to the morphological pattern of a different neuron type with axons
distributed in the CA1 oriens and radiatum layers (though also extending into the subiculum), but dendrites limited to oriens [16]. Unfortunately, neurons with these distinct characteristics had already been bestowed the different name of “CA1 trilaminar cells” in an earlier article [17]. Nevertheless, the label “CA1 trilaminar cell” was also used to describe yet another neuron type that had a similar axonal distribution, but dendrites invading lacunosum-moleculare [16]. But the confusion does not end here, as other labs independently referred to this latter morphology as either “CA1 Schaffer-associated” [18] or “CA1 apical dendrite innervating” [19].

We should note that these are not exceptional instances, but absolutely frequent occurrences, as depicted by several additional examples in Fig. 1 [20–29]. There are also multiple cases of the same referencing article calling a morphologically defined type by different synonyms, such as “perforant path-associated” and “CA1 R-LM” referring to neurons with axons and dendrites in CA1 stratum lacunosum-moleculare and dendrites in radiatum [18] (Fig. 2a). At the same time, these are not sterile spelling quibbles, because the specific laminar pattern of dendrites and axons defines the potential connectivity of the circuit and therefore the computational functions of neurons.

The confusion is not limited to neuron types but also affects the nomenclature of neuronal features, including morphological, electrophysiological, and molecular terminology. Qualitative phraseology is especially common in reporting morphological properties. An examination of the evidence collated in Hippocampus.org pertaining to the relative abundance of axons in an anatomical location of interest reveals ample use of terms such as “most,” “majority,” and “usually.” Furthermore, categorical terms are often employed to indicate continuous spatial distributions, as in “superficial/deep layer X,” “proximal/distal area Y,” and “septal/temporal region Z.” A clear consensus of how such terms should be adopted and interpreted, and what terms are to be avoided, reduces ambiguity. Hippocampus.org proposes a set of protocols for the description of neurites and their locations (hippocampus.org/full-interp).

The electrophysiological lexicon suffers not only from ambiguous descriptors but also from inconsistent definitions of the parameters themselves. For example, some investigators measure action potential amplitude from the resting membrane potential to the peak of the spike [30]. A complementary subset of studies, however, calculates action potential amplitude relative to the spike threshold.

<table>
<thead>
<tr>
<th>Cited Name</th>
<th>Hippocampus Neuron Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA1 bistratified</td>
<td>CA1 Bistratified</td>
</tr>
<tr>
<td>CA1 basket</td>
<td>CA1 Bistratified</td>
</tr>
<tr>
<td>CA1 trilaminar</td>
<td>CA1 Radial Trilaminar</td>
</tr>
<tr>
<td>CA1 Schaffer-associated</td>
<td>CA1 Trilaminar</td>
</tr>
<tr>
<td>CA1 apical dendrite innervating</td>
<td>Schaffer Collateral-Associated</td>
</tr>
<tr>
<td>CA1 oriens-bistratified</td>
<td>CA1 Oriens-Bistratified</td>
</tr>
<tr>
<td>CA1 O-R</td>
<td>CA1 O-LM</td>
</tr>
<tr>
<td>CA1 O-LM</td>
<td>CA1 O-LM</td>
</tr>
<tr>
<td>CA1 horizontal</td>
<td>CA1 O-LM</td>
</tr>
<tr>
<td>CA1 horizontal O-A</td>
<td>CA1 Recurrent O-LM</td>
</tr>
<tr>
<td>CA1 O-A</td>
<td>CA1 Recurrent O-LM</td>
</tr>
<tr>
<td>CA1 perforant path-associated</td>
<td>CA1 Perforant Path-Associated</td>
</tr>
<tr>
<td>CA1 R-LM</td>
<td>CA1 Perforant Path-Associated QuadD</td>
</tr>
<tr>
<td>CA3 O-LM</td>
<td>CA3 Perforant Path-Associated QuadD</td>
</tr>
<tr>
<td>CA3 LMR-Targeting</td>
<td>CA3 Perforant Path-Associated QuadD</td>
</tr>
<tr>
<td>CA3 radiatum</td>
<td>CA3 Perforant Path-Associated QuadD</td>
</tr>
</tbody>
</table>

Fig. 1 Relationships between cited names [3, 15–29] and neuron types. This bipartite graph highlights the naming confusion that is typical within the neuroscience community today
potential [31]. The relationship between the minimum and the steady-state membrane potentials resulting from a hyperpolarizing current is similarly ambiguous. The sag ratio quantifies the relative difference between the peak hyperpolarization and steady-state hyperpolarization [32]. Alternatively, the sag percentage reports the fractional change in membrane potential from peak to steady state relative to the steady state [33]. Figure 2B schematically shows the differences between these parameter definitions. Plainly, the use of identical or similar names for terms with different electrophysiological meanings can lead to the propagation of confusion and, worse, incorrect interpretations of data that are incorporated into the literature moving forward.

Molecular biomarkers bear an overabundance of synonyms, homonyms, hyponyms, hypernyms, and abbreviations. There is movement toward standardizing the naming of proteins, but it is debatable whether the efforts are alleviating or augmenting confusion. For instance, the entire family of mammalian neuronal transporters has been given the official name of "solute carrier family [X] member [Y]." The new names confer that the proteins are transporters, but provide little information beyond that. As an example, some authors now refer to vesicular glutamate transporter 2 (Gene ID: 84487, ncbi.nlm.nih.gov/protein/NP_445879.1) by the abbreviation Slc17a6, short for the official full name "solute carrier family 17, member 6," while others keep the familiar vGlut2. If these two alternatives were not enough, the marker is also known by the symbols Dnpi and Vgl [34–37] (Fig. 2c).

One of the worst cases of molecular biomarker terminology confusion in neuroscience involves glutamate receptors. Metabotropic glutamate receptors (mGluRs) are not to be confused with three classes of ionotropic receptors (GluRs): AMPA, kainate, and NMDA, sometimes referred to as AMPARs, kAs, and the NRs [38]. In the promising new naming schema for glutamate receptors, metabotropic receptors retain use of mGluR, while AMPA receptors use GluA, kainate GluK, and NMDA GluN [39]. It is yet to be seen how widely used either of these schemata will be. Alas, even if the entire research community compactly embraced them today, the problem of linking new information with previous publications would remain.

3 Resolving the neuron-type crossword puzzle

The solution to both the naming dilemma and property-based neuronal classification lies in establishing and consistently applying an unambiguous, clearly defined, unique nomenclature with links to antecedent synonyms. With property terms, scholarly resources can serve as broadly accepted references and dictionaries, such as the Medical Subject Headings (MeSH) by the US National Library of
Medicine [40] and NeuroLex by the National Institutes of Health-contracted Neuroscience Information Framework [41]. However, using such services requires turning attention away from the material with the confusing or unknown term, navigating external site(s), finding and processing the definition(s), then refocusing attention to the original material. A terms portal integrated into the original material would greatly simplify the process.

3.1 Data schema for property-based classification of hippocampal neurons

To solve the neuronal naming problem, the neuroscience community would ideally adopt a robust approach to classification. Using the distributions of axons and dendrites across identifiable anatomical areas is advantageous for a number of reasons. Axonal and dendritic patterning is fundamental to all neurons, yet sufficiently information-rich to allow grouping at a useful level of abstraction on the spectrum from considering all neurons the same (as would be the case if spike integrator were the chosen property) and each individual neuron unique (as would result if using exact matches of the neurite arbors). In addition, neurite patterns are more stable and less dependent upon experimental conditions than molecular markers and electrophysiology, respectively. Lastly, as demonstrated below, this approach naturally provides the means of creating unique, concise, informative names of neuron types.

We designed an open-source online system enabling machine-readable information accessibility. Knowledge about each Hippocampome.org neuron type, including the names, synonyms, properties, and evidence, is stored in a relational database sourcing a user-friendly web-accessible interface. Figure 3 depicts the conceptual organization of the database based on three general categories: neuron types, neuron properties, and published evidence. Links between data and relations are captured in separate relation tables to both increase flexibility and reduce complexity, thereby facilitating continuous development and long-term maintainability.

Converting information published for human consumption into machine-readable form dictates system level decisions to minimize the energy cost of processing. We chose a three-step workflow. The first step is for researchers (doctoral students, postdocs, and faculty) to identify and study relevant articles, gleaning salient information and encoding it into spreadsheets. The second step involves python code to ingest these spreadsheets into data tables, populating along the way relation tables. The third step consists of rendering the resulting structured data in web pages dynamically leveraging the database. Performing the most time consuming tasks up front (steps one and two) allows for fast web-based lookup access by the end-user community. The data/relation table design adds a layer of complexity to the database, but simplifies the resultant query implementation complexity, considerably speeding up real-time interactive retrieval.

3.2 Neuron term machine-readable definition identifier

In order to facilitate the collation of machine-readable definitions of relevant terms, we designed and implemented a novel functionality of Hippocampome.org for online assistance in disambiguating neuron property nomenclature (Fig. 4). This new resource (Hippocampome.org) integrates key neuron term descriptors into a curated catalog of web-accessible human- and machine-readable definitions. Users can browse, search, and filter terms from drop-down menus augmented with autocomplete-as-you-type function. After selecting one or more terms, the portal returns the mapped concept with mouse/cursor-layover display of all available synonyms and the context in which they appear, along with a list of available definitions and direct hyperlinks to the corresponding source providers. Users can also search for specific keywords of interest within the definitions. Furthermore, when browsing Hippocampome.org and all cited evidence within, terms with available definitions are now highlighted: users can display a definition...
Fig. 4. Neuron term machine-readable definition identifier: an online portal for conceptual mapping of neuronal properties fully integrated in Hippocampome.org.

The first challenge in deploying this novel functionality was to identify the set of terms requiring machine-readable definitions. This research leveraged two primary sources of terms: Petilla [4] and the article excerpts cited as evidence in Hippocampome.org [11]. The Neuron Registry [7] constituted a third minor source of terms. The Petilla terminological consistency consists of a finite list of (∼232) published terms. Hippocampome.org, in contrast, contributes a less neatly bounded set of terms exceeding 10 K discrete tokens (as estimated by the wordle.net utility, Fig. 5). To parse these tokens into a manageable set, we filtered the Hippocampome.org terms at each extreme of the occurrence count spectrum. This preprocessing step eliminated words with very large (>1000) occurrence counts, including uninformative strings such as “a,” “the,” and “of,” as well as words with very small (<100) occurrence counts, representing rare and typically uninteresting terms like “outside-out” and “sheetlike.” Lastly we hand-curated the remaining set of approximately 700 terms to remove nonscientifically relevant words yielding a final corpus of 490 evidence-derived terms. An additional 782 terms corresponded to neuron names, anatomical regions, biomarkers, and electrophysiological parameters stored in Hippocampome.org. In all, due to minor overlaps among the above lists, this collation accounted for 1478 distinct terms.

To find machine-readable definitions we devised a preferred portal/repository approach. For general neurobiological terms, we first searched Neurolex.org, MeSH browser (nlm.nih.gov/mesh), the Bioportal services from the National Center for Biomedical Ontology [42], and the US Public Health Service CRISP database [43]. The terms from Hippocampome.org evidence primarily refer to the rodent hippocampus, thus it is essential that the extracted definition be relevant to these target domains. Since the same word can have different meanings, most definitions retrieved by the initial automated search were largely out of context, requiring a slow step of manual curation. We preferentially assigned evidence terms from Hippocampome.org definitions and links most relevant to the rodent hippocampal formation. Similarly, we linked the Petilla terms to definitions in the context of GABAergic interneurons of the cerebral cortex.

For protein definitions, we harnessed the Ontology Look-up Service [44] of the Gene Ontology Consortium [45] as the sole reference given the depth and breadth of coverage for this type of molecular data. Because the molecular terms are generally regular and systematically databased, we successfully automated API-based pulling from established sources (e.g., the National Center for Biotechnology Information). For term not found in these primary resources, we reverted to Google searches, prioritizing definitions from scholarly or institutional sources such as the Allen Brain Atlas [46], Scholarpedia.org, and the US National Institute of Standards and Technology (nist.gov). For residual blanks, we resorted to dictionaries like Merriam-Webster or Wikipedia.

The last step of manual curation involved concept mapping to group together distinct terms linking to textually different but logically analogous definitions. For
example, “action potential” and “spike” are synonyms for which multiple machine-readable definitions exist. This mapping yielded 810 distinct concepts from the 1478 unique terms, with a total of 924 unique definitions from 1378 distinct resource links. Table 1 summarizes the neuron term counts, including number per category (i.e., morphological, molecular, and electrophysiological) and unique instances. Table 2 organizes this information by resources providing the machine-readable external links to the term definitions.

3.3 Neuron type naming

The classification schema introduced by Hippocampome.org [11] defines neuron types based on their properties, starting from morphological patterns and with the added specification of molecular and electrophysiological features. For example, Hippocampome.org defines dentate gyrus granule cells as excitatory neurons with axons in the hilus, CA3 lucidum/pyramidal, and CA2 pyramidal, dendrites in the inner and outer molecular layer, and soma in the granular layer. These definitions are now available as an explicit list (hippocampome.org/neuron-types) and linked from the term definition portal described above.

It is difficult to quantify how many unique neuron types have been defined to date in the hippocampal formation due to ambiguity and overlap of descriptors across research labs. We constrain the number of Hippocampome.org [11] neuron types (e.g., 122 in the initial release) by limiting the primary characterization properties to axonal/dendritic patterns and excitatory/inhibitory neurotransmitters.

Furthermore, Hippocampome.org neuron types are assigned both a formal name and a unique number identifier (e.g., DG (e) 2201p-CA3_00110 Granule; type 1000). The formal name contains several components (hippocampome.org/formal-name): (a) the abbreviation of the subregion where the soma is located, (b) a symbol specifying the putative major neurotransmitter (i.e., “e” for glutamatergic, excitatory neurons or “i” for GABAergic, inhibitory neurons), and (c) a numeric encoding for the presence or absence of neurites within the subregion of soma location. In neuron types whose axons extend outside of their home subregion, the numerical encoding continues with a “p” (for projecting) followed by codes analogous to (a) and (c) to specify the subregions receiving the projection. Finally, the formal name ends with a unique, human-friendly label that attempts to maximize usability and understanding of neuron types within the research community. Figure 6 illustrates the selection process for determining this “common name.”

In the most clear-cut cases, a single name dominates the literature as universally recognized and understood. In such “canonical” cases, we adopt these standard names, as in Granule, Mossy, CA3 basket, and CA1 pyramidal cells. In other situations, a neuron type may not be as broadly known, but is only cited in a single way. In these cases, we straightforwardly adopt the single cited name, such as in Semilunar Granule, CA3 Giant, and CA3 Granule cells. The remaining cases represent the confusing scenarios in which the literature describes the same neuron types with multiple names and different neuron types with the same name.

If one name or acronym is clearly dominant, with more frequent citations than all other names, we adopt it as the common name, as in the cases of HIPP, MOPP, HICAP, and MOLAX interneurons. Other neuron types, however, have multiple, approximately equally cited names, especially in the less-studied entorhinal cortex. In these cases, to avoid playing favorites, we hybridize the cited names, as is LI-II Multipolar-Pyramidal, LI-II Pyramidal-Fan, and MEC LI-III Pyramidal-Multiiform. Lastly, there are neuron types for which all cited names entail potential confusion with similar or identical names already assigned to other neuron types based upon the rules above. In these scenarios, we are forced to either modify a cited name in order to differentiate it (e.g., Mossy MOLDEN, DG Basket CCK+, and CA3c Pyramidal) or to create a new name altogether (e.g., AIPRIM, HIPROM, MOCAP, CA3 SO–SO). We try to use this final clause sparingly (only 4 names out of 122 in Hippocampome.org are entirely new), but minor modifications of pre-existing names are often unavoidable (46 out of 122).

4 Discussion

The basis of communication is language. Unfortunately, the language of neuroscience is lacking a common terminology with respect to neuron types and their associated discriminating properties. Paraphrasing Shakespeare: “What’s in a name? That which we call a [neuron] by any other name would [fire] as [frequently].” By first establishing neuron types based on their necessary and sufficient common characteristics, and then methodically applying a naming protocol, it is possible to establish a basis for systematic neuron naming. This work differs from prior efforts in the level of comprehensiveness. There have not been any all-inclusive compilations of neuron types within the entire rodent hippocampal formation based on peer-reviewed published literature for the past two decades [3]. Scientific laboratories most often work independently, and researchers performing experiments typically name neurons for their convenience. Hippocampome.org dynamically integrates these data across all known experimental evidence.
<table>
<thead>
<tr>
<th>Category</th>
<th>Term source</th>
<th>Summary</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuron type</td>
<td>Hippocampone.org</td>
<td>641</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>122</td>
<td><a href="http://hippocampone.org/formal-typedefs">http://hippocampone.org/formal-typedefs</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA1 Ivy interneuron</td>
<td>CA1 Ivy</td>
</tr>
<tr>
<td>Morphology</td>
<td>Petilla</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Soma - Shape - Triangular</td>
<td>Pyramidal</td>
</tr>
<tr>
<td>Hippocampone.org</td>
<td>34</td>
<td>34</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>33</td>
<td>CA1:SP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA1:SP</td>
<td><a href="http://atlas.brain-map.org/#atlas=1&amp;plate=10096023&amp;structure=399&amp;x=4594.285714285715&amp;y=3.95.428641183386&amp;zoom=3&amp;resolution=11.97&amp;z=5">http://atlas.brain-map.org/#atlas=1&amp;plate=10096023&amp;structure=399&amp;x=4594.285714285715&amp;y=3.95.428641183386&amp;zoom=3&amp;resolution=11.97&amp;z=5</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Field CA1, pyramidal layer</td>
</tr>
<tr>
<td>Molecular marker</td>
<td>Petilla</td>
<td>175</td>
<td>176</td>
</tr>
<tr>
<td></td>
<td></td>
<td>752</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transcription factor - Nkx2.1</td>
<td>Transcription factor - Nkx2.1</td>
</tr>
<tr>
<td>Hippocampone.org</td>
<td>97</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95</td>
<td>PV</td>
</tr>
<tr>
<td>Electrophysiology</td>
<td>Petilla</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>36</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spike amplitude</td>
<td>APampl</td>
</tr>
<tr>
<td>Hippocampone.org</td>
<td>10</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>APampl</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://biportal.bioontology.org/ontologies/4295?terms&amp;conceptIds=X77KG">http://biportal.bioontology.org/ontologies/4295?terms&amp;conceptIds=X77KG</a></td>
<td></td>
</tr>
<tr>
<td>Evidence &amp; notes</td>
<td>Hippocampone.org</td>
<td>490</td>
<td>429</td>
</tr>
<tr>
<td></td>
<td></td>
<td>434</td>
<td>436</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gap junction</td>
<td>Electrical synapse</td>
</tr>
</tbody>
</table>

Sum: 1504 1049 1434 9097
Distinct: 1578 810 1178 924
Name-calling in the hippocampus (and beyond): coming to terms with neuron types and properties

We have striven to find human-friendly names that are recognizable to, at least, those who are familiar with hippocampal neurons. In many cases, however, these names have minimal informational content to those unfamiliar with the type. The part of the formal name that is most informative is the numeric encoding of the neurite pattern (detailed description: hippocampome.org/find-term). Knowledge of the pattern of dendrites and axons conveys information about potential connectivity of the neuron type within the circuit. Therefore, incorporation of this pattern into the name allows instantaneous envisioning of the location of the neurites and by extension the connectivity of the type. In addition, this numeric encoding is unique for most neuron types with only subtypes discriminated by their primary neurotransmitter, post-synaptic target specificity, or molecular marker and/or electrophysiology profiles having the same pattern. In these cases, the human-friendly part of the name provides uniqueness (e.g., "CA1 2232 Basket" and "CA1 2232 Basket CCK+^{-}). This method of naming neurons results in extremely informative, concise names without necessitating the memorization of many acronyms. Furthermore, it is applicable to any brain region that is divisible into parcels.

Going beyond Hippocampome.org, the same approach to defining neuron types can be extended outside the hippocampal formation. For example, CA1 neurons that project to other brain regions such as the lateral septum, medial septum, and/or hypothalamus can be characterized by extending the axonal/dendritic patterns to encompass those regions.

Nomenclature confusion could be mitigated with increased awareness of the neurons, molecules, and properties and how they fit in the historical context. This is a lot to ask of researchers, but resources like Hippocampome.org provide significant assistance. Hippocampome.org demonstrates that the necessary and sufficient discriminating property of neurite patterning is a workable and advantageous foundation upon which to build a neuron type library. Enhancing such a library with a terms definition portal further reduces terminology confusion. Coupled, these resources begin clarifying the muddled state of the literature and re-illuminating the path to neuroscience progress.
Fig. 6 Decision logic for assigning common names to Hippocampome.org neuron types

Acknowledgments This work was supported by National Institutes of Health (NIH) National Institute of Neurological Disorders and Stroke (NINDS) Grant R01 NS086001, National Science Foundation (NSF) Robust Intelligence Grant IIS-1302256, Air Force Office of Scientific Research (AFOSR) Grant FA9550-10-1-0385, Keck Nordic, Office of Naval Research (ONR) Grant MURI N00014-10-1-019801.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


153


D. J. Hamilton MS current research interests include neuroinformatics, hippocampal formation, neuron classification.

D. W. Wheeler PhD current research interests include network dynamics, hippocampal formation, neuron classification.

C. M. White PhD current research interests include molecular biomarkers, hippocampal formation, neuron classification.

C. L. Rees MS current research interests include graph-theoretic analysis, hippocampal formation, neuron classification.

A. O. Komendantov PhD current research interests include electrophysiological models, hippocampal formation, neuron classification.

M. Bergamino PhD current research interests include biomedical engineering, hippocampal formation, neuron classification.

G. A. Ascoli PhD current research interests include computational neuroanatomy, hippocampal formation, neuron classification.
REFERENCES


experience. *International Journal of Clinical and Experimental Pathology, 6*(12), 2955–2961.


Eastwood, S. L., & Harrison, P. J. (2003). Interstitial white matter neurons express less reelin and are abnormally distributed in schizophrenia: towards an integration of molecular and morphologic aspects of the neurodevelopmental hypothesis. *Molecular Psychiatry, 8*(9), 821–831.


effects of maternal deprivation on NPY and Agouti-related protein mRNA. *Endocrinology, 142*(11), 4771–4776.


recommendations for the nomenclature of neuropeptide Y, peptide YY, and pancreatic polypeptide receptors. *Pharmacological Reviews, 50*(1), 143–150.


Müller, C., Pongratz, S., Pidlich, J., Penner, E., Kaider, A., Schemper, M., ... Ferenci, P. (1998). Treatment of pruritus in chronic liver disease with the 5-hydroxytryptamine receptor type 3 antagonist ondansetron: a randomized,


Tatemoto, K. (2004). Neuropeptide Y: History and Overview. In P. M. C. Michel (Ed.), *Neuropeptide Y and Related Peptides* (pp. 1–21). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-18764-3_1


Christopher Loren Rees was born in Blacksburg, Virginia. Even as a toddler, he showed an intrinsic bent for organization and analysis (i.e. informatics) when he arranged Matchbox cars by color and function in his highchair. His father, a professor at Virginia Tech, and his mother, a homemaker and community volunteer, always made Christopher and his brother feel incredibly loved and important. Together, they spent time talking at meals (including at the corner booth at Hardee’s, where he learned to read), did crossword puzzles, traveled, and did projects. They built an amazing treehouse, which the front page of the newspaper called a “Tree Palace,” as well as an 8-foot-deep secret underground tunnel (with electricity) into the shed that was the envy of every kid in the neighborhood, thus rendering moot its status as “secret.” Someday, they will finish the suspension bridge from the treehouse to the nearby pine tree.

As a child and teenager, Christopher played piano for 10 years, tuba in marching and symphonic bands for 5, and baseball and soccer. It was as an 8-year-old member of this soccer team where Christopher first met his future wife, Rachel. Sadly, at that time, he was frequently too shy to talk to her. She was very pretty. Plus, she was better at soccer than he was.

Christopher’s interest in the brain began in high school when his grandfather was diagnosed with, and later passed away from, Alzheimer’s disease. In college at Virginia Tech, while majoring in computer and electrical engineering, Christopher had his first experience with what might be called brain research. In an independent study, he built a robot whose “brain” and ability to walk were based on central pattern generators. Upon graduating, he pursued medicine. He spent time living with a medical missionary doctor in Togo, West Africa, where he helped to diagnose malaria, treat injuries ranging from gunshot wounds to snake bites, and was allowed to scrub in on and assist with surgical procedures from leg amputations to typhoid perforations. Though he enjoyed the human side of medicine, he missed research and learning with computers that he had been exposed to at Virginia Tech.

He went back to school and graduated with a Master’s degree in bioinformatics from George Washington University in 2009, studying gene dysregulation in Autism and Schizophrenia. During this time, two life-changing and providential events occurred. First, a professor named Dr. Vaisman from George Mason University came to speak at a weekly seminar. In conversations afterwards, Christopher became intrigued with his research and with the bioinformatics program at GMU, to which he decided to apply. Secondly, Christopher asked Rachel to marry him – the smartest thing he has ever done.
As a then-young Ph.D. student about to begin at GMU, on the basis of a webpage professing interests in neuroscience and soccer, Christopher contacted Dr. Giorgio Ascoli of the Krasnow Institute about the possibility of doing a one-semester lab rotation. Dr. Ascoli asked when Christopher might be available to talk about suitable projects and “working out a useful experience for all involved.” Christopher replied that any time would be fine. In what might have been the first test of many, Dr. Ascoli set an interview for 9pm the following night. A perfect project was found and a lab rotation became two rotations, which turned into a Ph.D. dissertation, which yielded several published articles with more to come. From Christopher’s point of view, the experience could not have been more “useful,” challenging, or fun.