

Allen Institute Large-scale Datasets and Modeling Tools

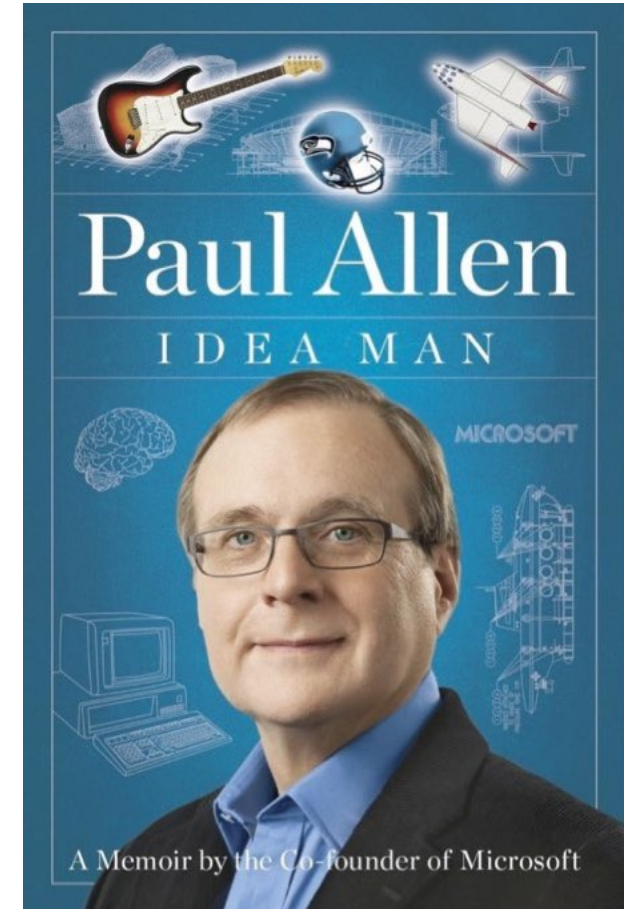
Anton Arkhipov
Associate Investigator



ALLEN INSTITUTE *for*
BRAIN SCIENCE

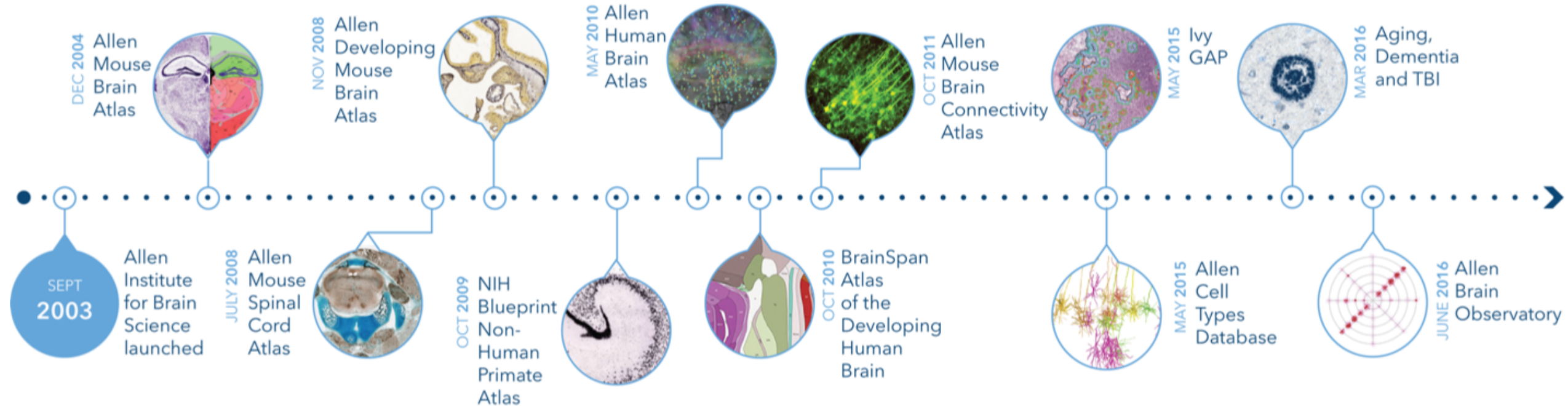
A Brief History

- An independent, non-profit medical research organization, founded in 2003, supporting basic research in the brain sciences
- Culture between a university and a biotech startup, focused on a handful of large projects that can be done at scale and that require tight interactions across disciplines
- Ten-year program initiated in 2012 for building cellular-level observatories for mice and human cortex



Allen Institute - Online Public Resources

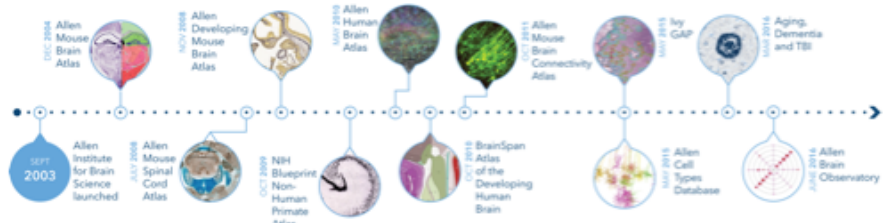
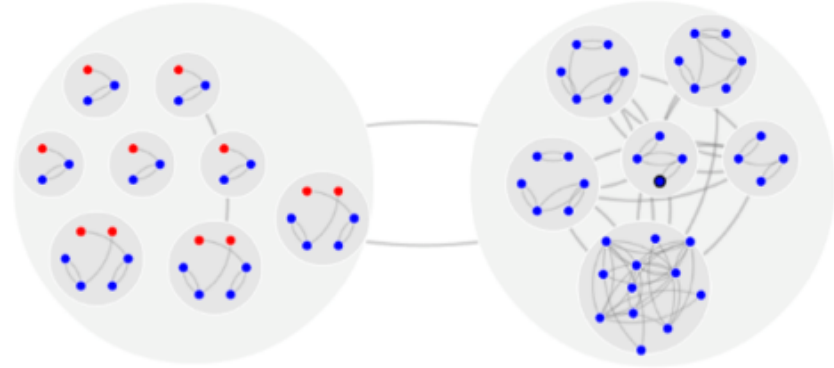
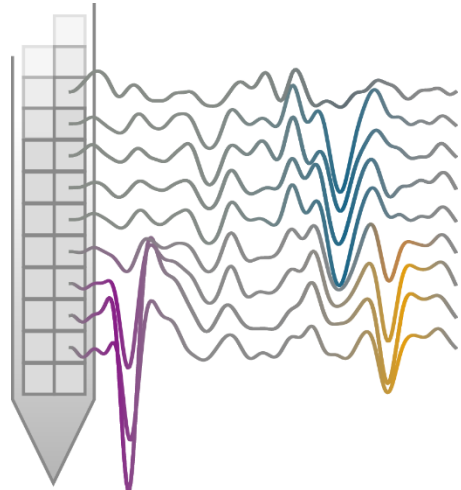
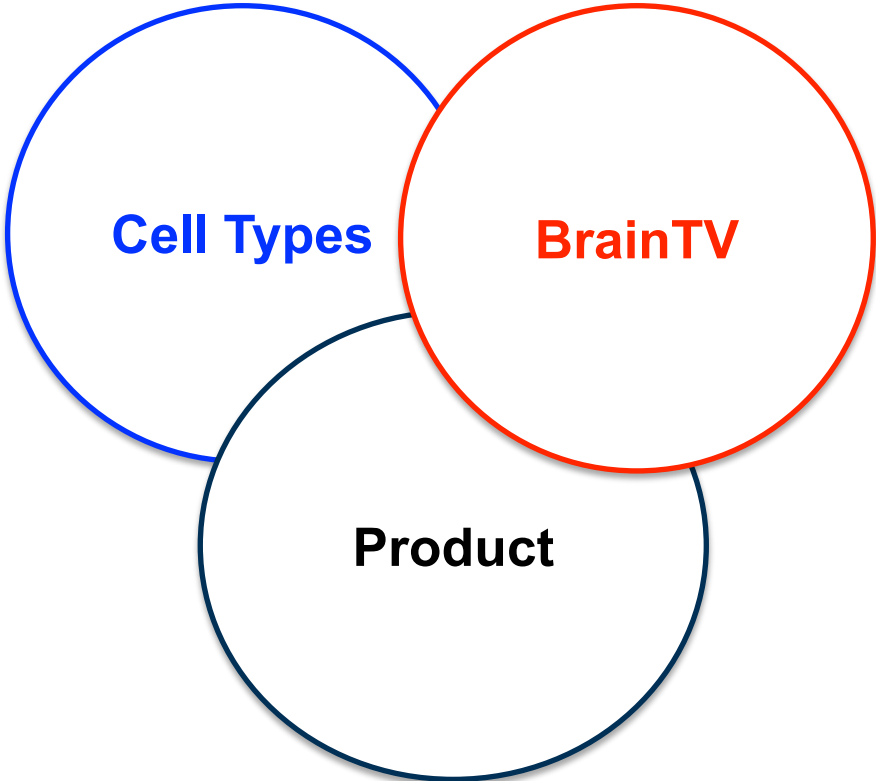
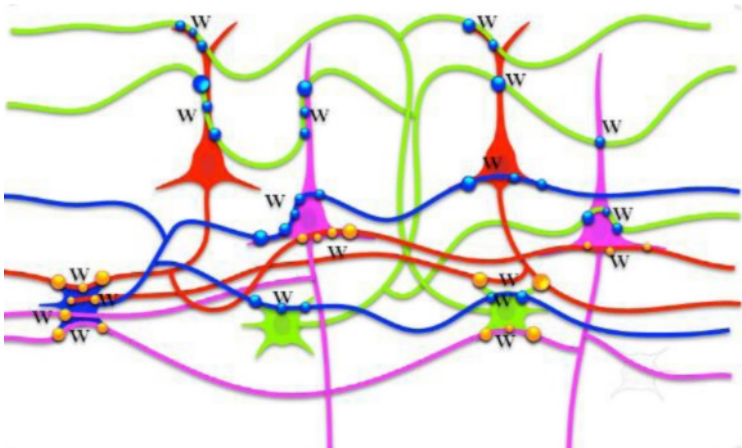
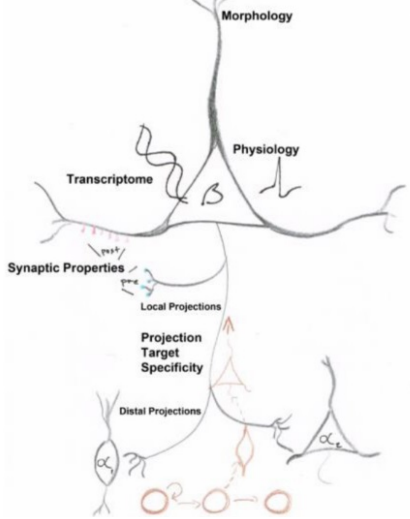
www.brain-map.org



All data are

- publicly accessible via API as soon as they pass QC
- freely available without any commercial restrictions
- accessible 1-2 years prior to any publications

Institute Programs



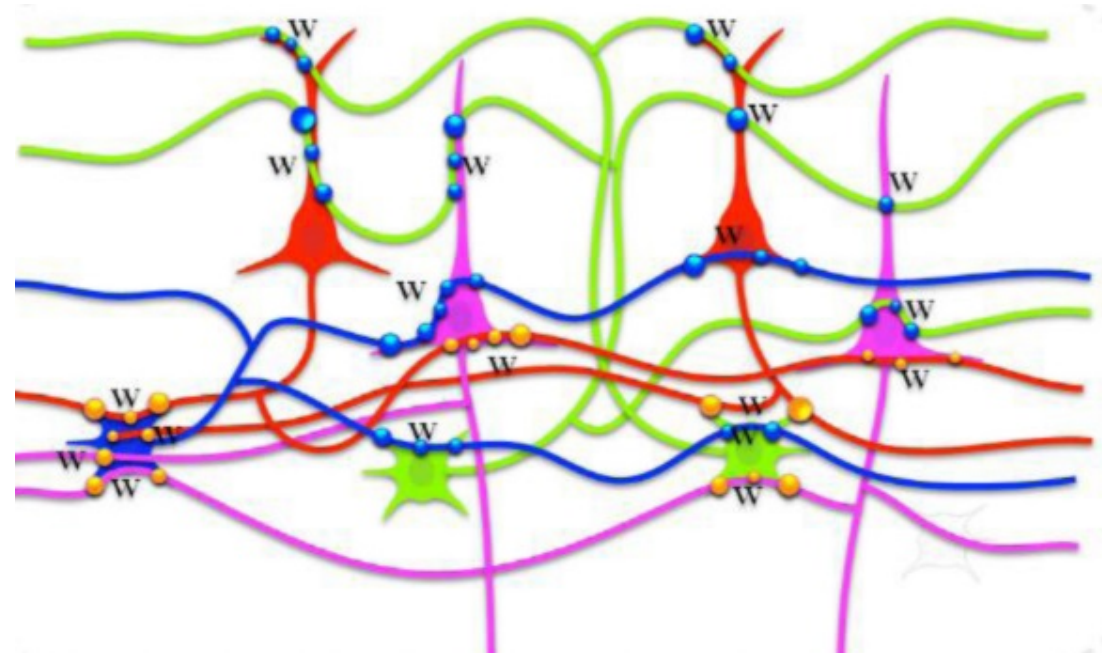
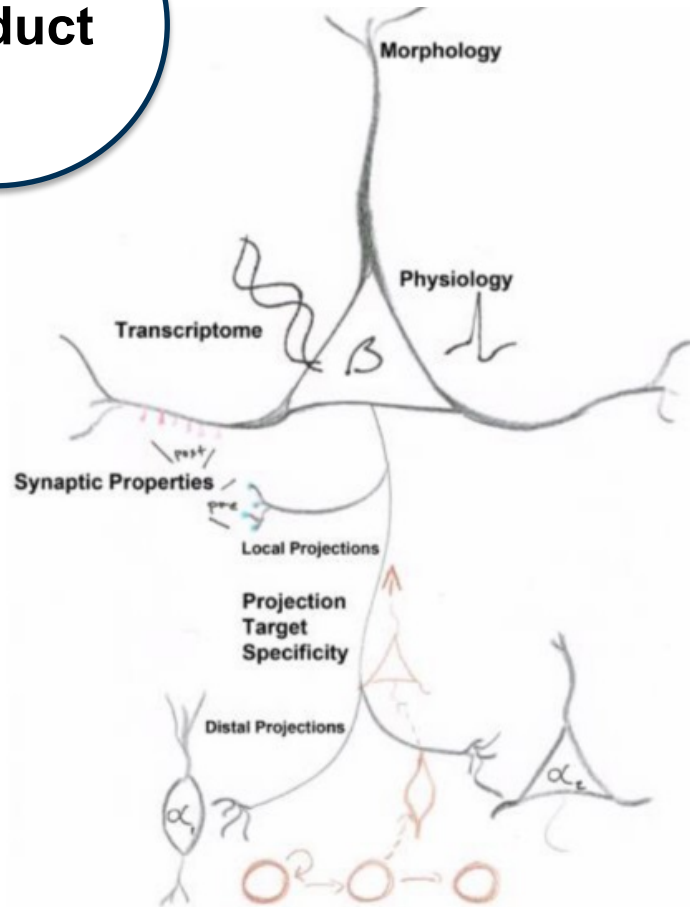
Cell Types Program

Cell
Types

BrainTV

Product

Characteristics of cells and connections

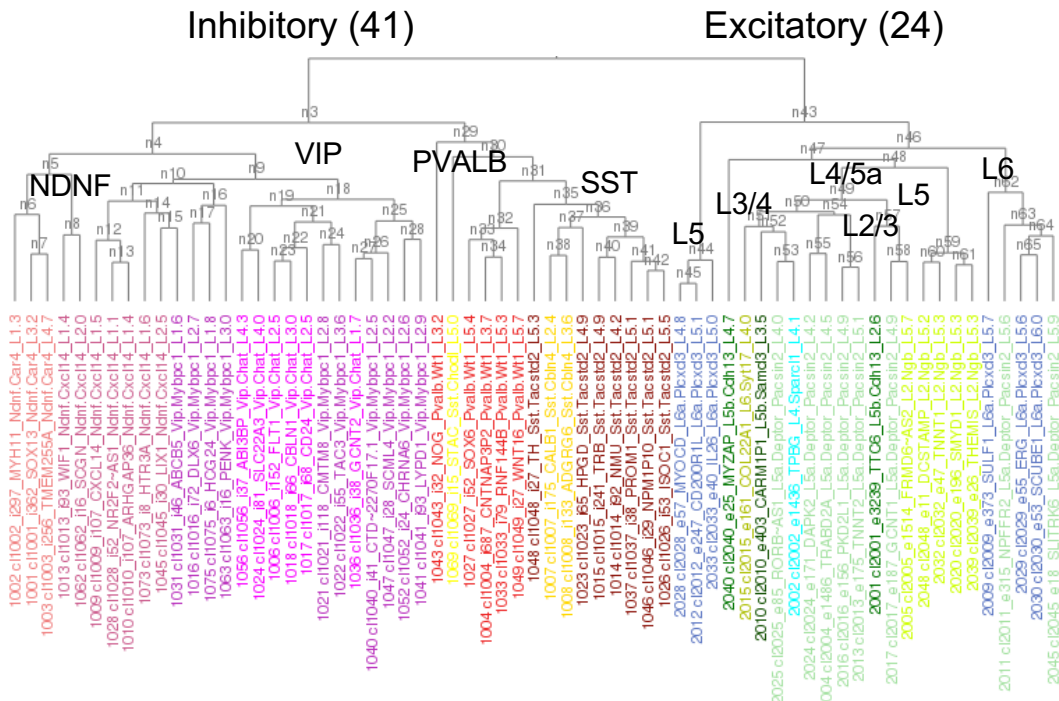


Transcriptomics (Mouse and Human)

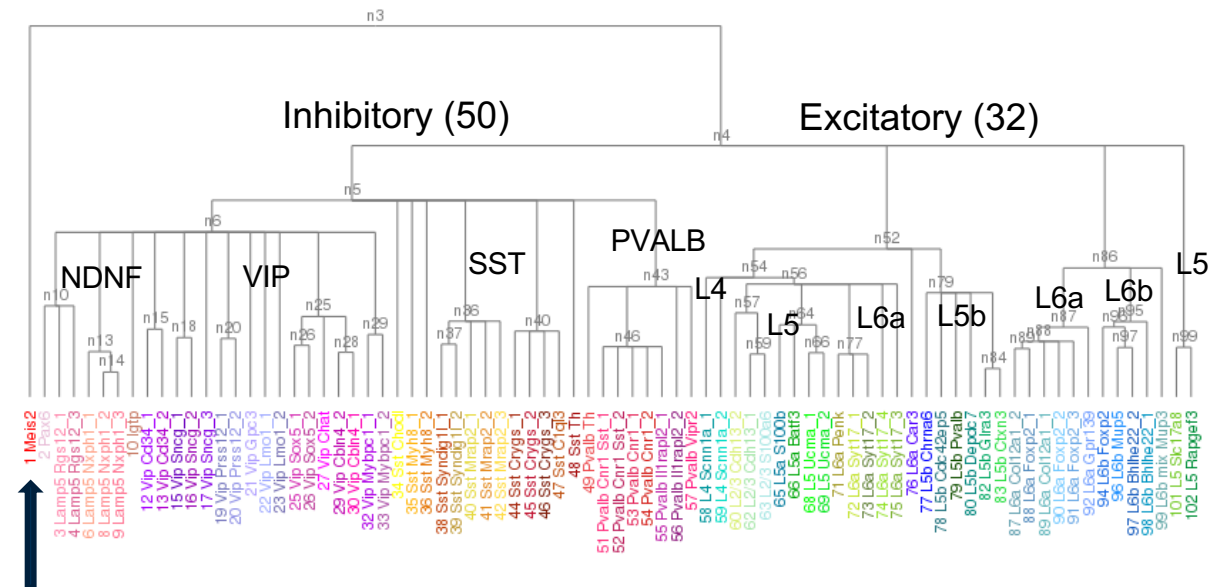
Deliverables: Single-cell transcriptomic data for mouse and human cortical cells, transcriptomic cell type taxonomy.

Scope: Tens of thousands of cells.

Human temporal cortex



Mouse visual cortex



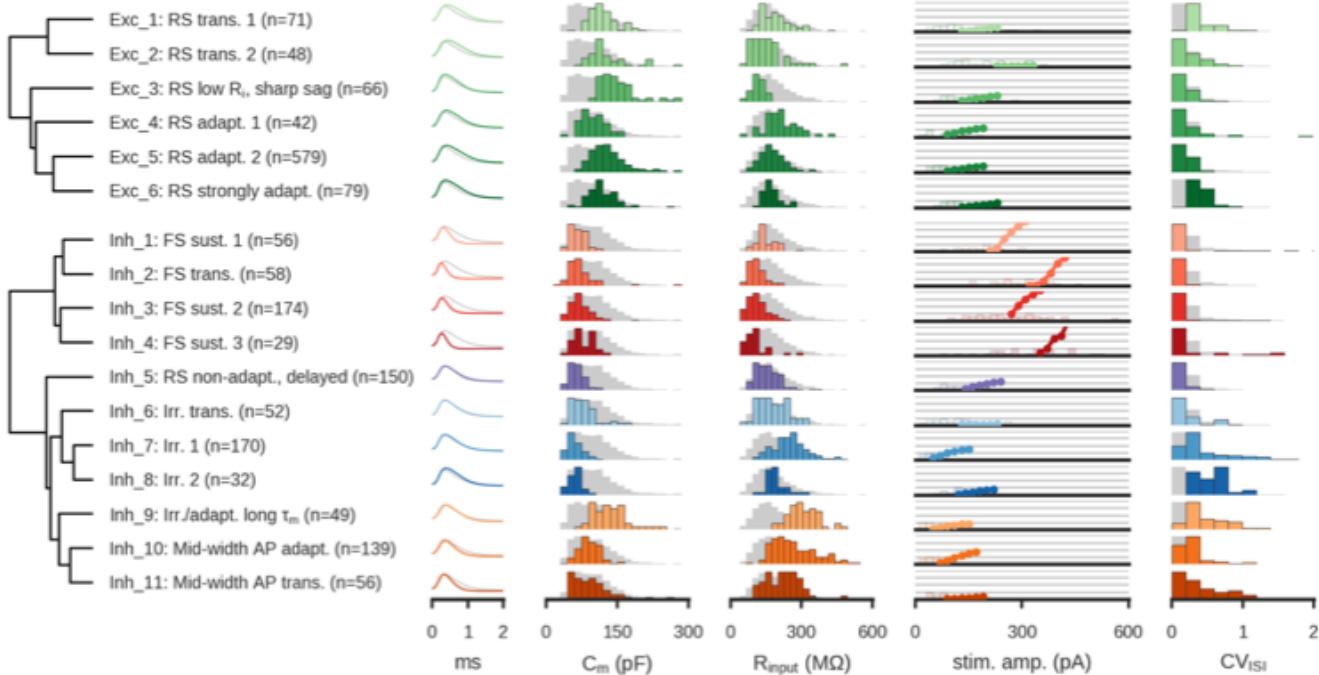
Rare type in deep layer 6 has not yet been sampled in human

Electrophysiology and Morphology of Neurons *in Vitro* (Mouse and Human)

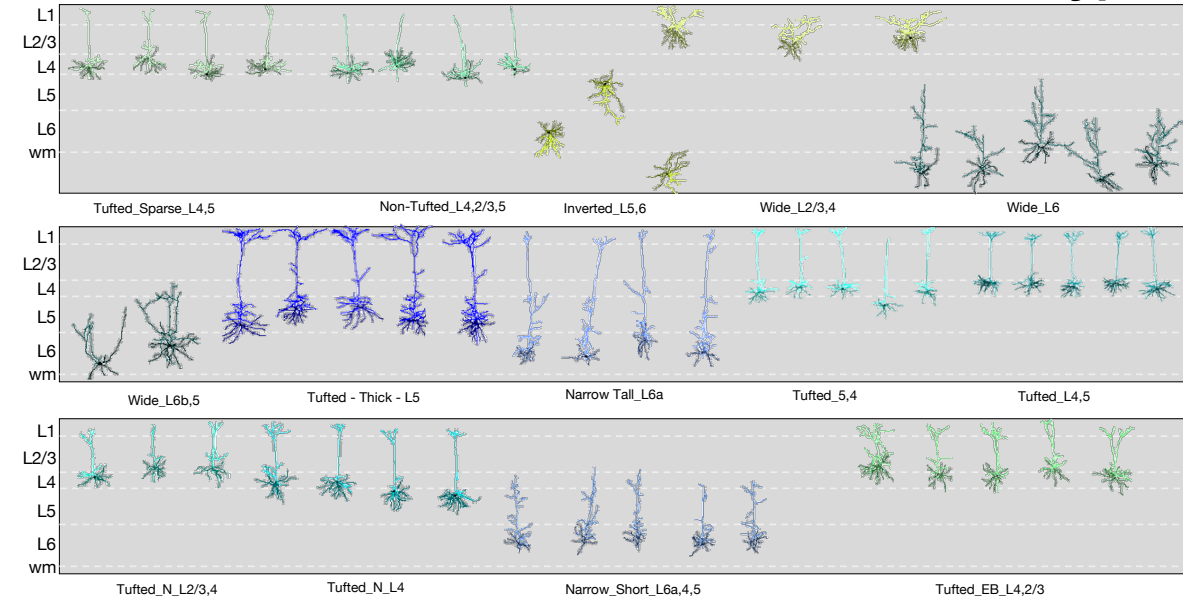
Deliverables: Ephys and morphology data, as well as models, for cortical cells; morphological and physiological classification of cell types.

Scope: Thousands of cells in the mouse visual cortex and human middle temporal gyrus (MTG).

e-types



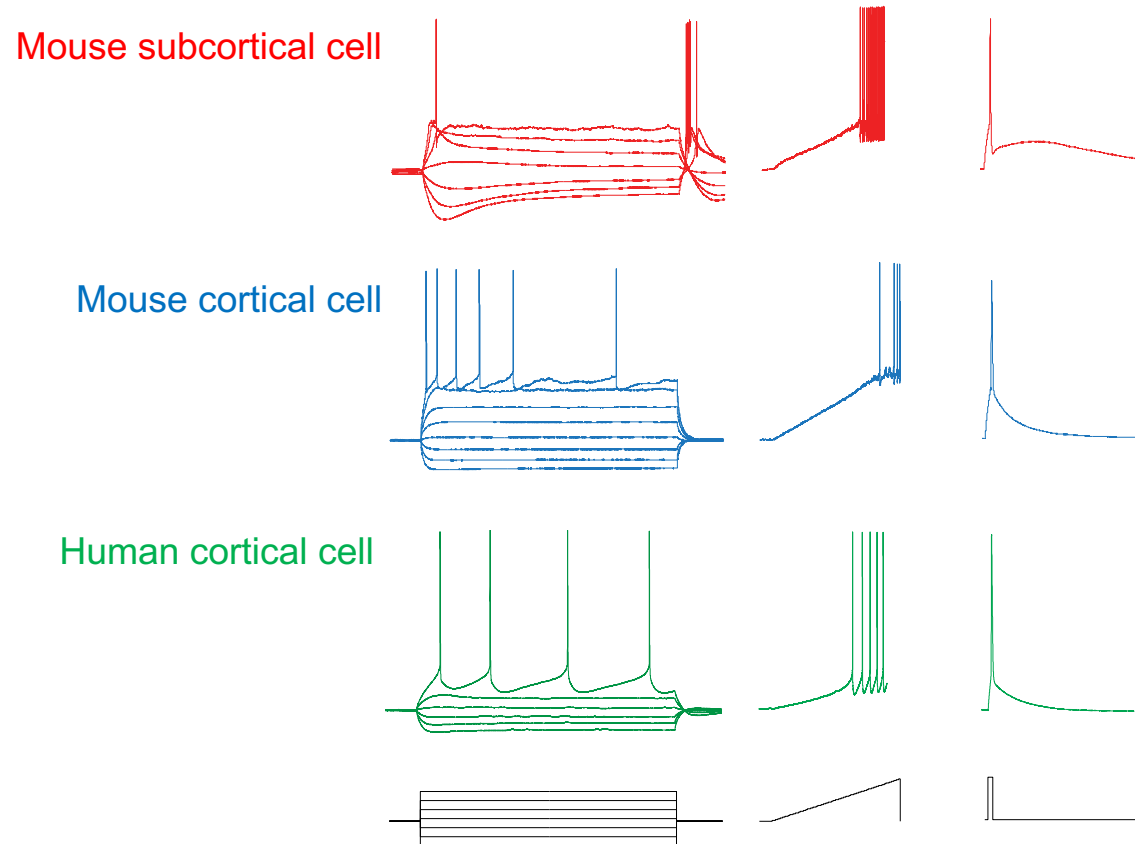
m-types



Mouse and Human Patch-Seq

Deliverables: Transcriptomic, ephys, and morphology data, as well as models, for mouse and human cortical cells.

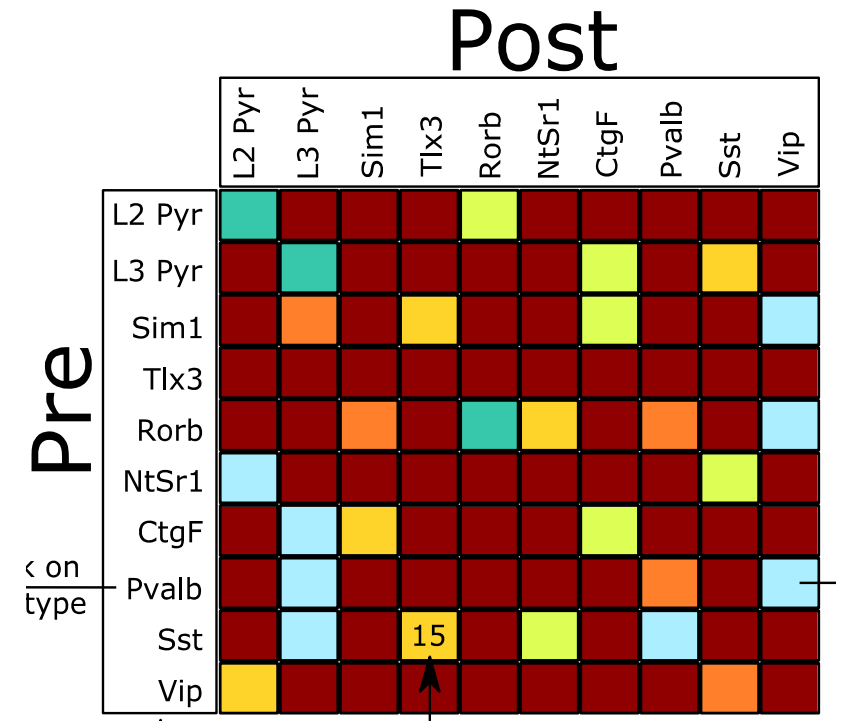
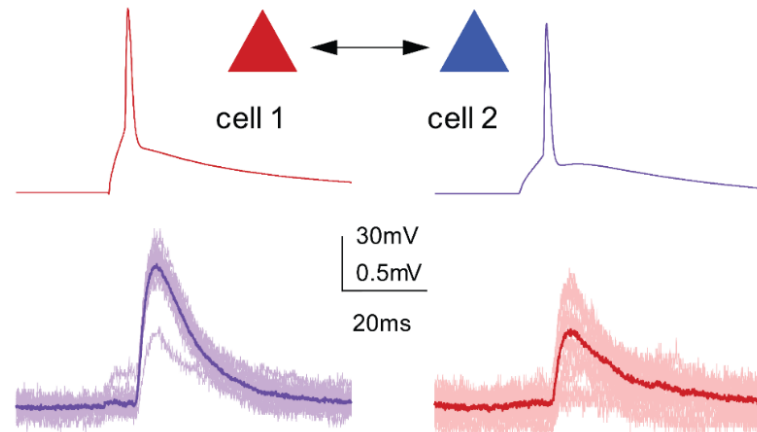
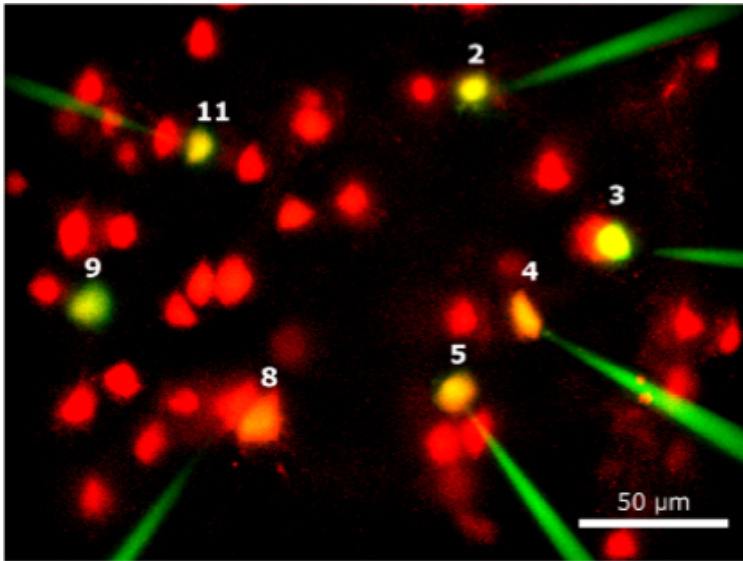
Scope: Thousands of neurons in mouse and human.



Mouse and Human Synaptic Physiology

Deliverables: Connection probabilities, synaptic strength, and short-term plasticity properties for connections between cortical cell types in mouse and human.

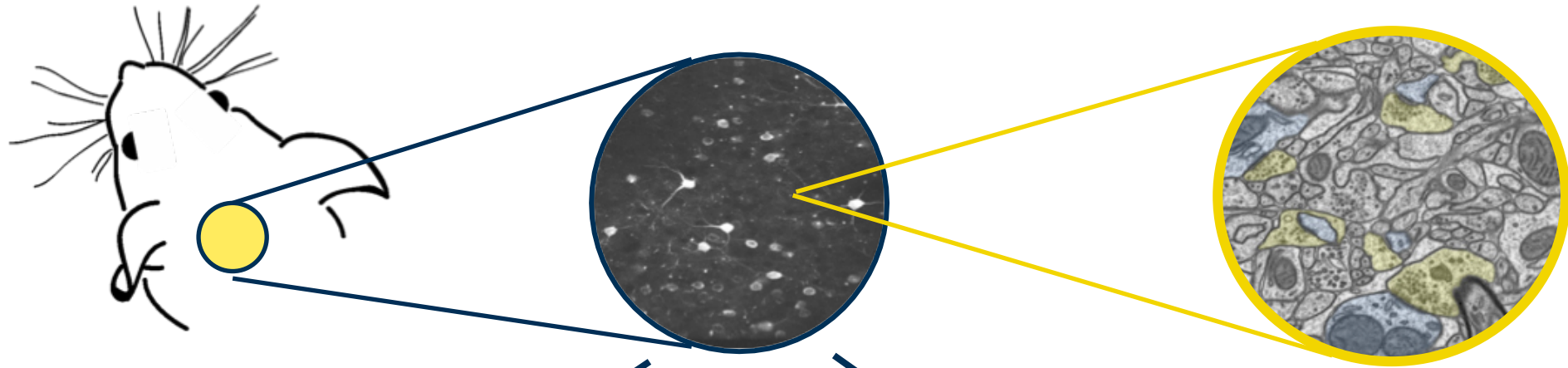
Scope: Initially 50-100 connections per matrix element; approximately 10x10 matrix (i.e., ~10 “cell types”) in mouse and as close to this plan as possible in human. Expand to study specific cell types and correspondence of transcriptome to connectivity in later years.



Mouse and Human EM Connectomics

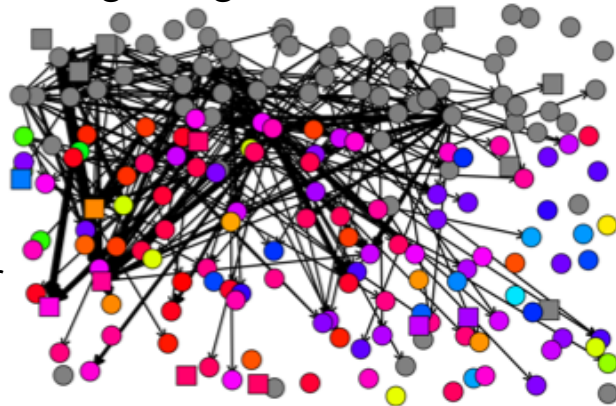
Deliverables: Reconstruction of neuronal morphologies and the connectivity matrix in the local cortical circuit – 1 mm³ in mouse V1 and 2 mm³ in human MTG.

Scope: ~100,000 neurons (in the mouse, most of the neurons would be functionally characterized *in vivo*).



Cell Types

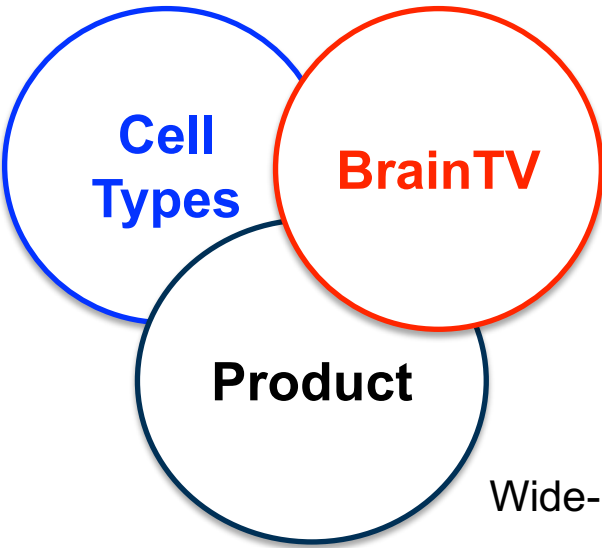
Wiring Diagram



Automated reconstructions are by Sebastian Seung Lab, Princeton University

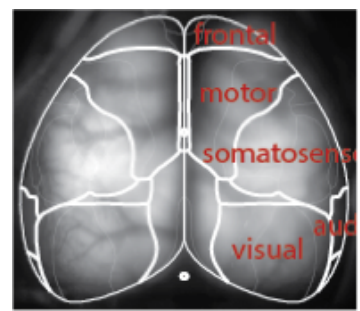
Ca²⁺ imaging is by
Andreas Tolias lab, Baylor
and by the Allen Institute

BrainTV Program

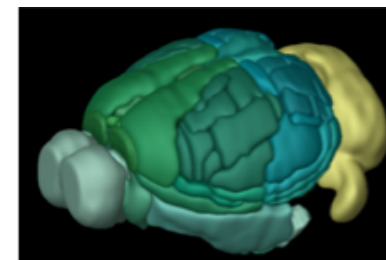


Activity and Computations in Cortical Circuits *in Vivo*

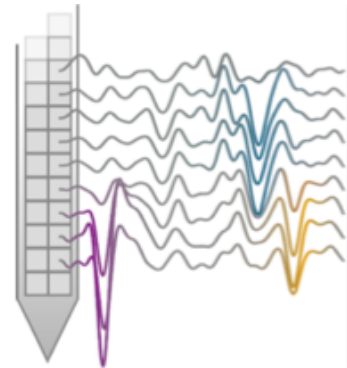
Wide-Field Imaging



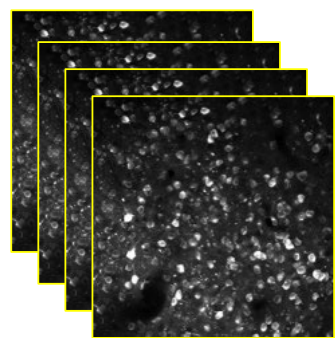
Population Statistics & Machine Intelligence Networks



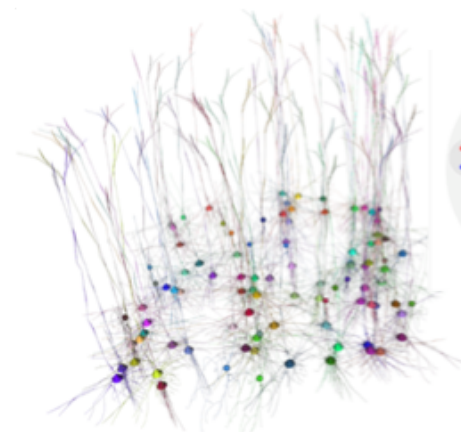
Ephys – Neuropixels Probes



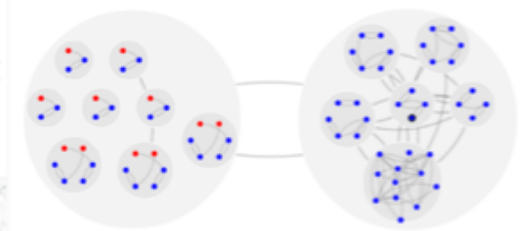
2p cellular imaging



Biophysical Modeling



Point Neuron & Filter Networks



Visual Coding

HOME

GET STARTED

BRAIN OBSERVATORY

TOOLS

OVERVIEW

EXPERIMENTS

CELLS

SDK

TRANSGENIC CHARACTERIZATION

DOCUMENTATION

HELP

Deliverables:

Ophys and ephys data for neuronal responses in the mouse cortex *in vivo*.

Scope:

Tens of thousands of cells for ophys; tens of thousands of units for ephys; ~10 cortical areas.

Allen Brain Observatory

<http://observatory.brain-map.org/visualcoding>

The screenshot displays the Allen Brain Observatory interface. At the top, navigation tabs include HOME, GET STARTED, BRAIN OBSERVATORY (highlighted), and TOOLS. Below these are secondary tabs: OVERVIEW, EXPERIMENTS, CELLS, SDK, TRANSGENIC CHARACTERIZATION, DOCUMENTATION, and HELP.

Transgenic Mouse Lines

- Cux2-CreERT2;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Rorb-IRES2-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Scnn1a-Tg3-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Nr5a1-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Rbp4-Cre_KL100;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Emx1-IRES-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)

Cortical Imaging Locations

A diagram of a mouse brain section shows various cortical areas: AUDd, AUDp, TEa, VISi, VISal, VISr, VISa, VISam, VISpm, VISp, VISpl, and VISpo. A 1mm scale bar is provided. Below the brain diagram, a depth scale is shown with horizontal bars representing imaging locations at 175, 275, and 375 micrometers.

Allen Brain Observatory

About the Allen Brain Observatory

The Allen Brain Observatory provides standardized *in vivo* surveys of the mouse visual cortex, featuring visually evoked calcium responses from neurons in selected areas and Cre lines.

Key features of the data

- Searchable data for calcium imaging signals across visual areas and depths.
- A variety of data views capturing visual cortex activity and cell population statistics.
- Standardized spatial responses to five stimuli surveyed from transgenic mouse lines.
- Raw data and analysis tools available for download via the Allen Brain Observatory Development Kit (ABO-SDK).

Search the data

- Click on the EXPERIMENTS tab to view experimental details.
- Click on a Visual Stimulus to view more.

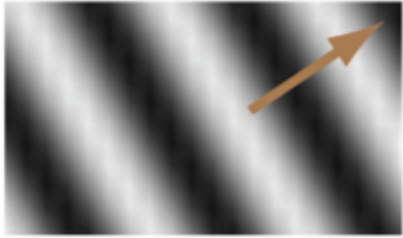
Visual Stimuli and Cell Responses

Learn about visual stimuli

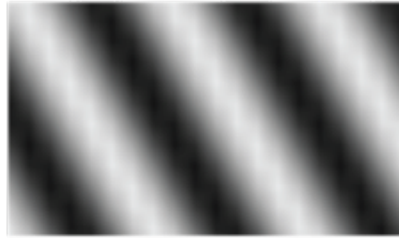
- Drifting Gratings
- Static Gratings
- Natural Scenes
- Natural Movies
- Locally Sparse Noise

Allen Brain Observatory – in Vivo Ca^{2+} Imaging

Drifting
Gratings



Static
Gratings



Locally sparse
noise



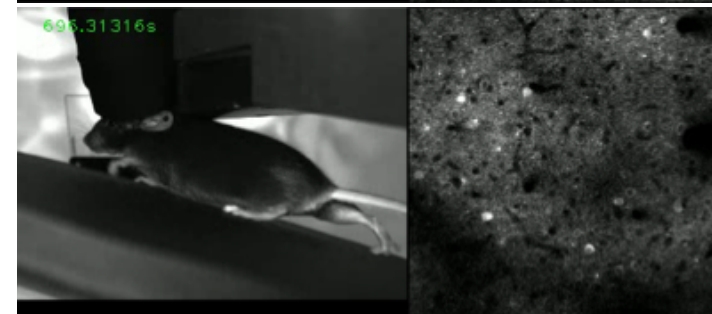
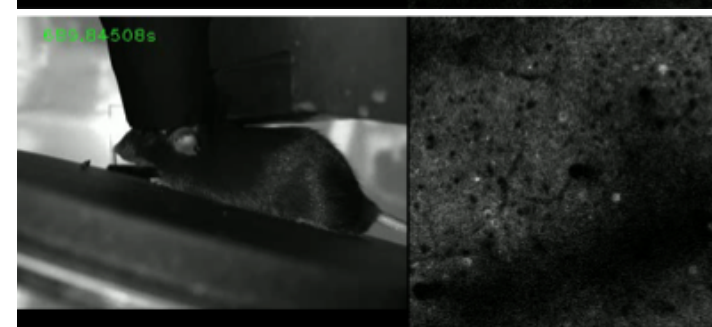
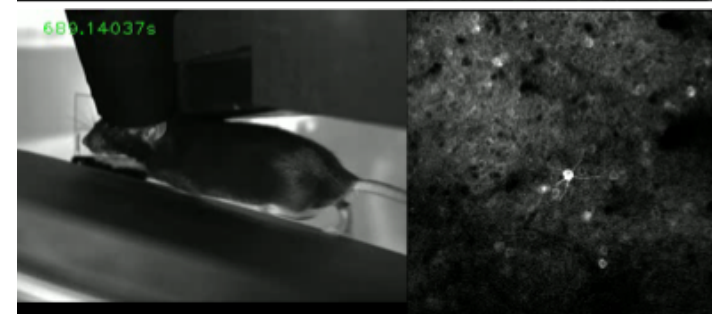
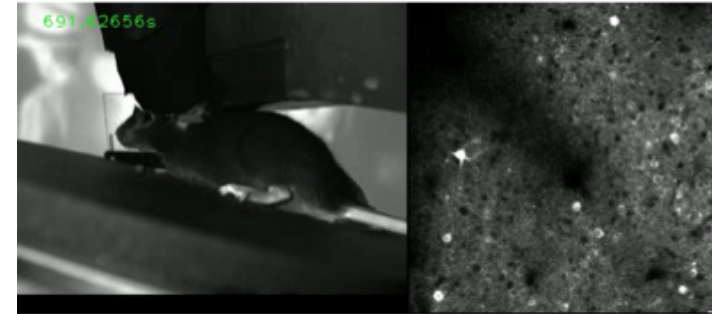
Natural
Scenes



Natural
Movies



Spontaneous
Activity



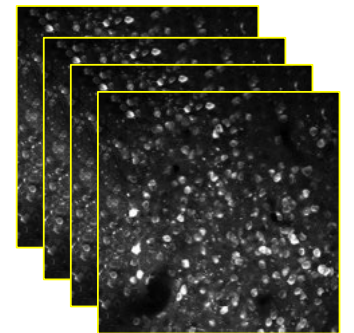
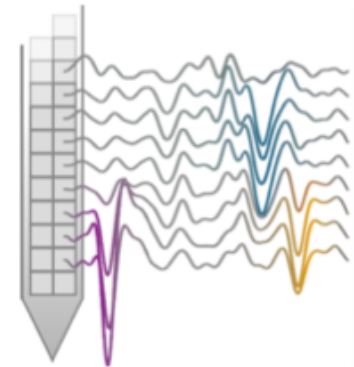
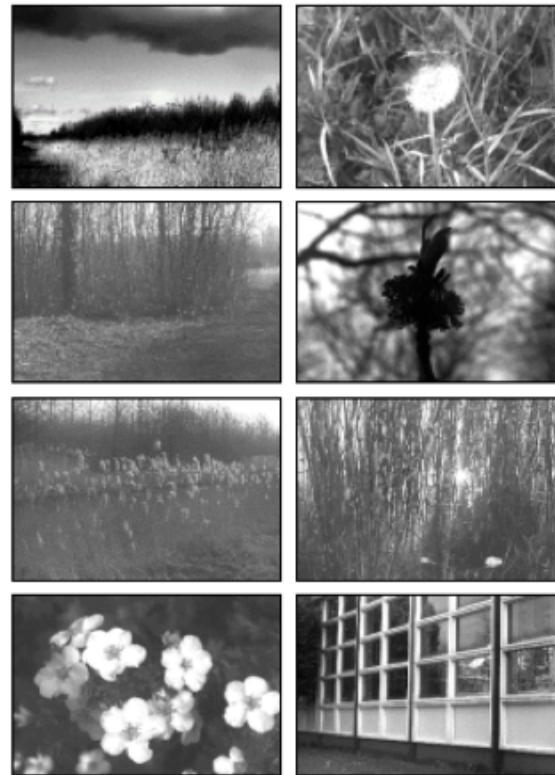
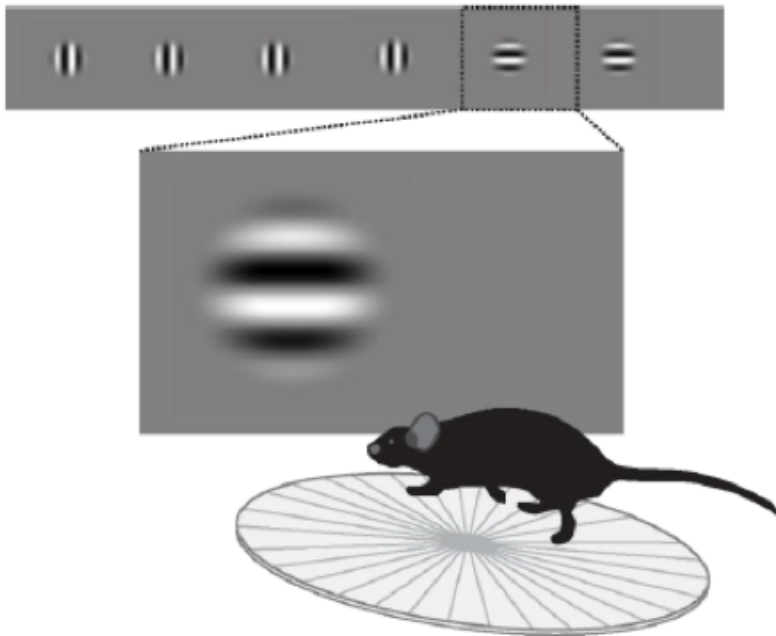
Visual Behavior

Deliverables:

Ophys and ephys data for single-cell responses in the mouse cortex in vivo, in the context of a change detection behavior.

Scope:

Tens of thousands of cells for ophys; tens of thousands of units for ephys.

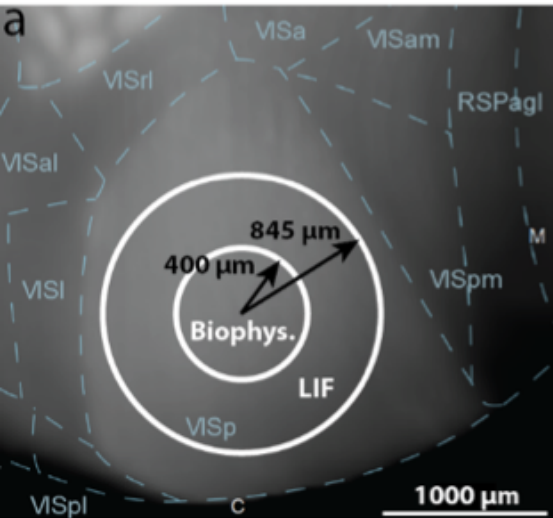


Models of V1 and of Multiple Cortical Areas

V1 Models

Deliverables: A series of models of the V1 cortical column, employing biophysical and point-neuron approaches.

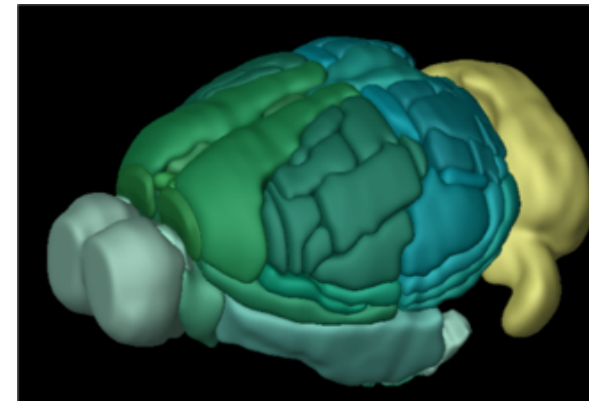
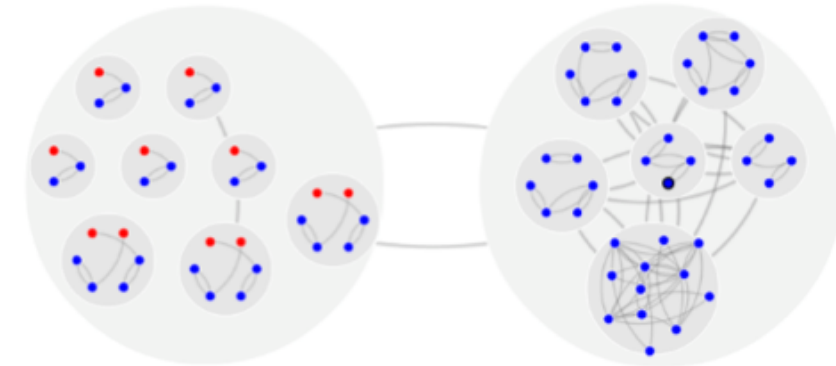
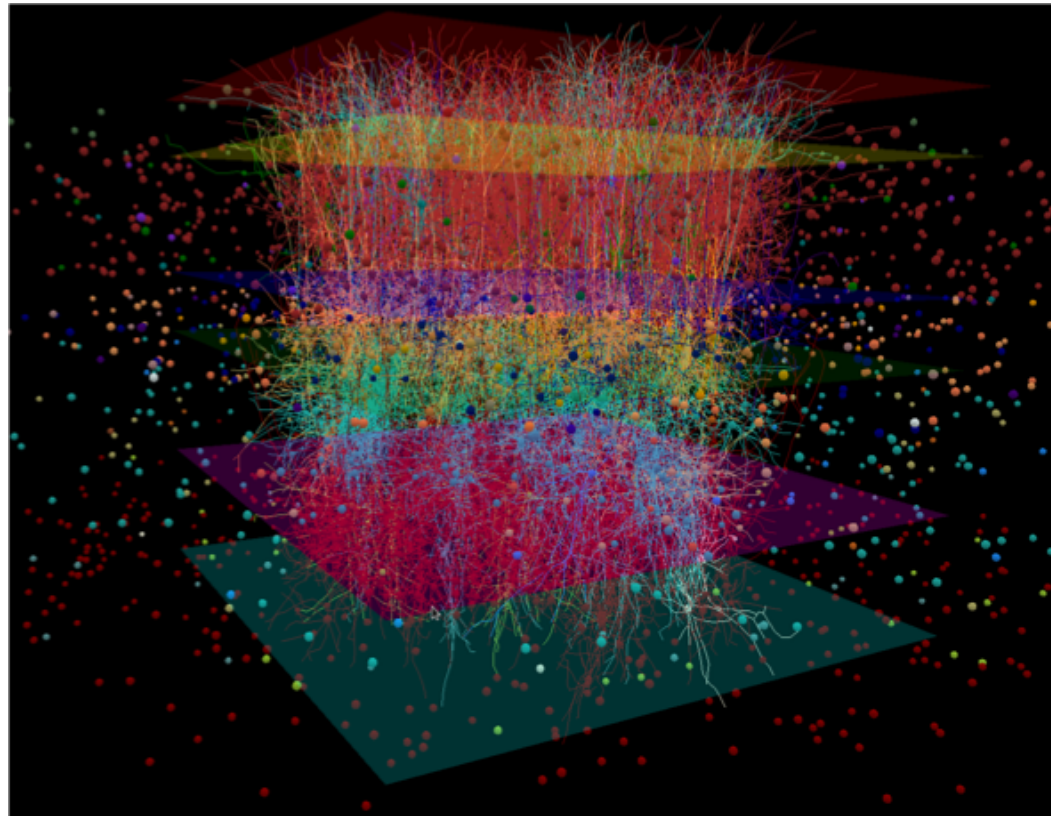
Scope: Representation of all cortical layers in V1, ~200,000 cells.



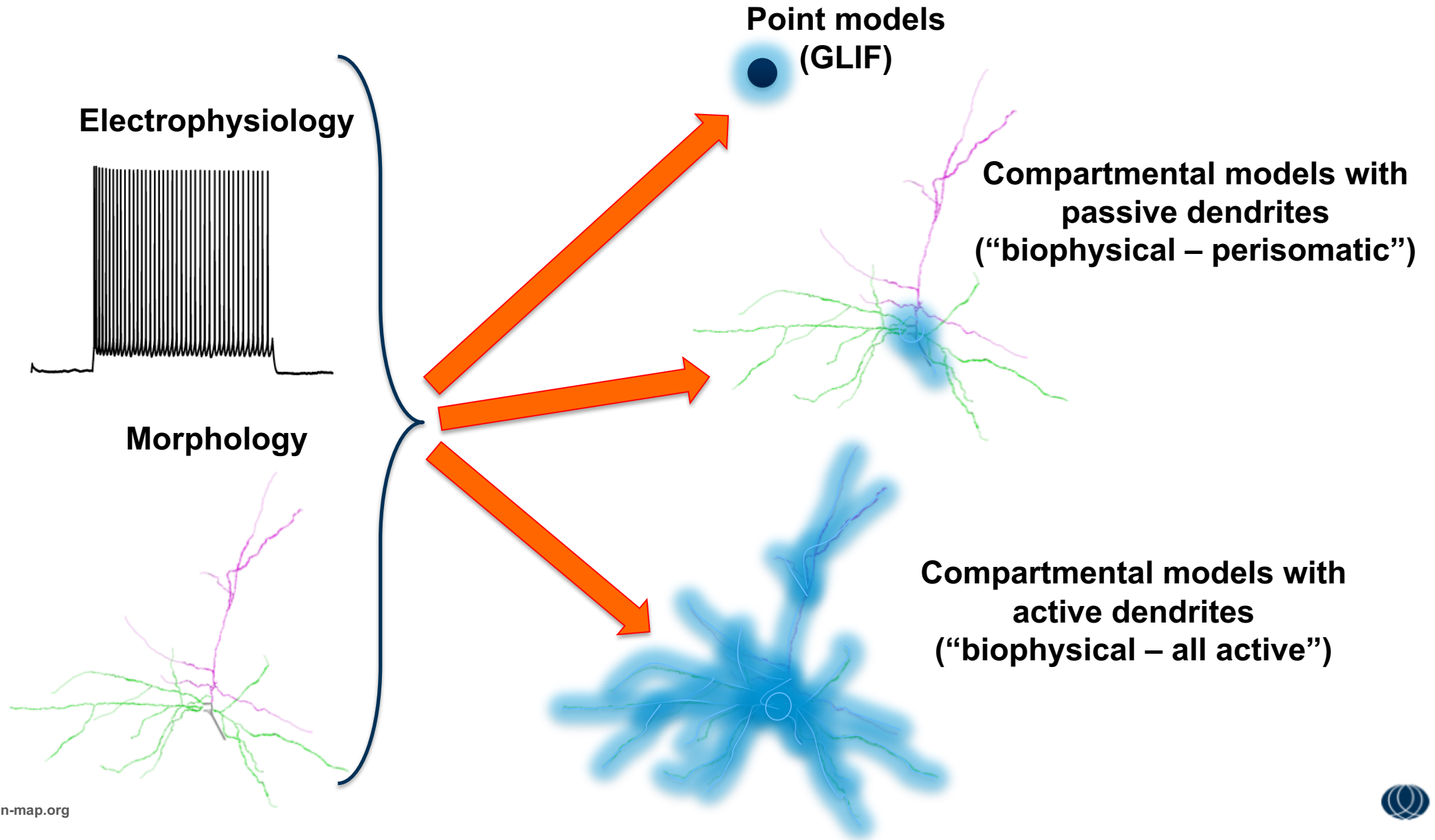
Cortex Models

Deliverables: A series of models of multiple cortical areas in the context of behavior, employing point-neuron and machine intelligence approaches.

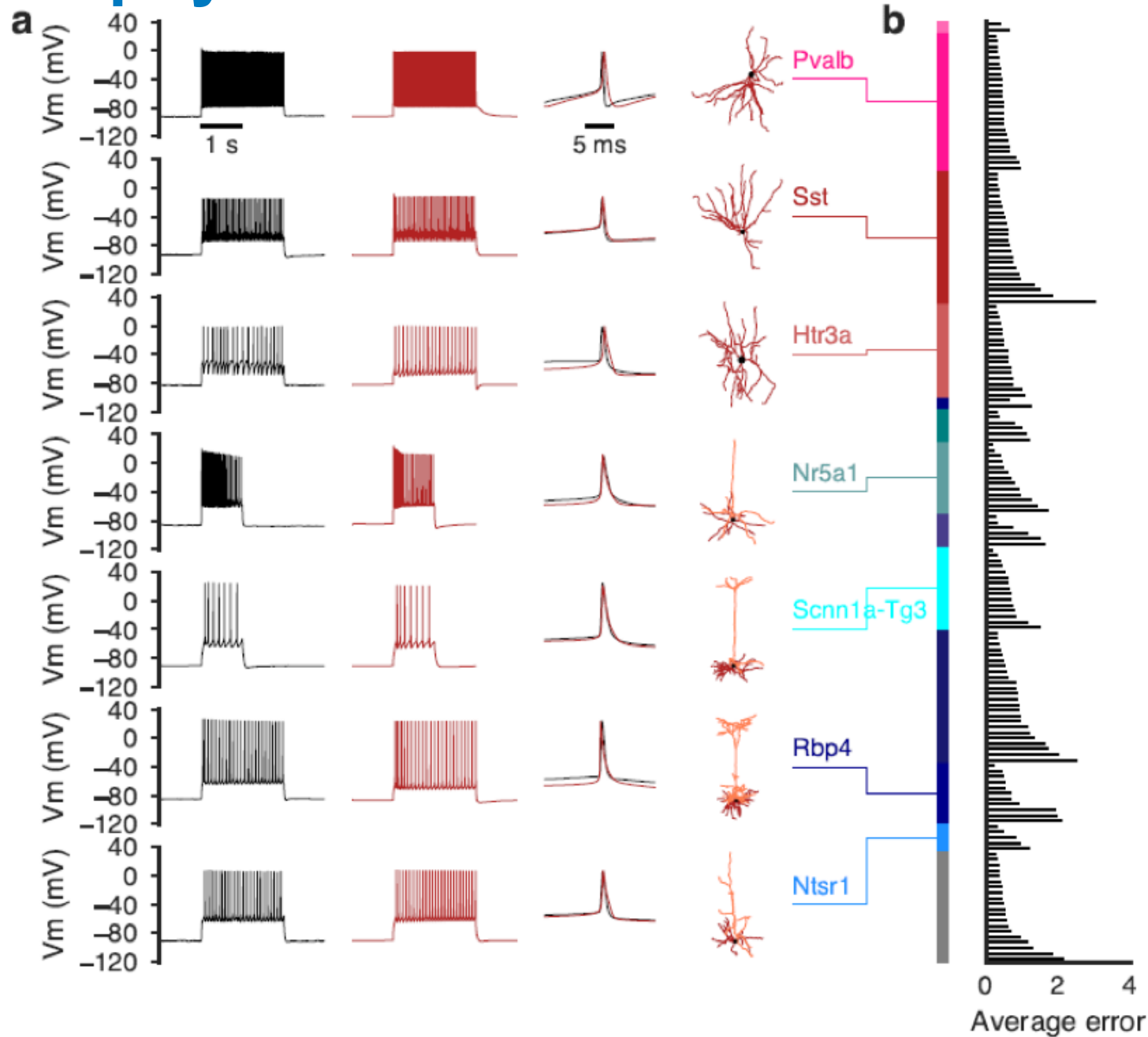
Scope: Several visual and, potentially, a number of non-visual areas.



From Slice Data to Neuronal Models



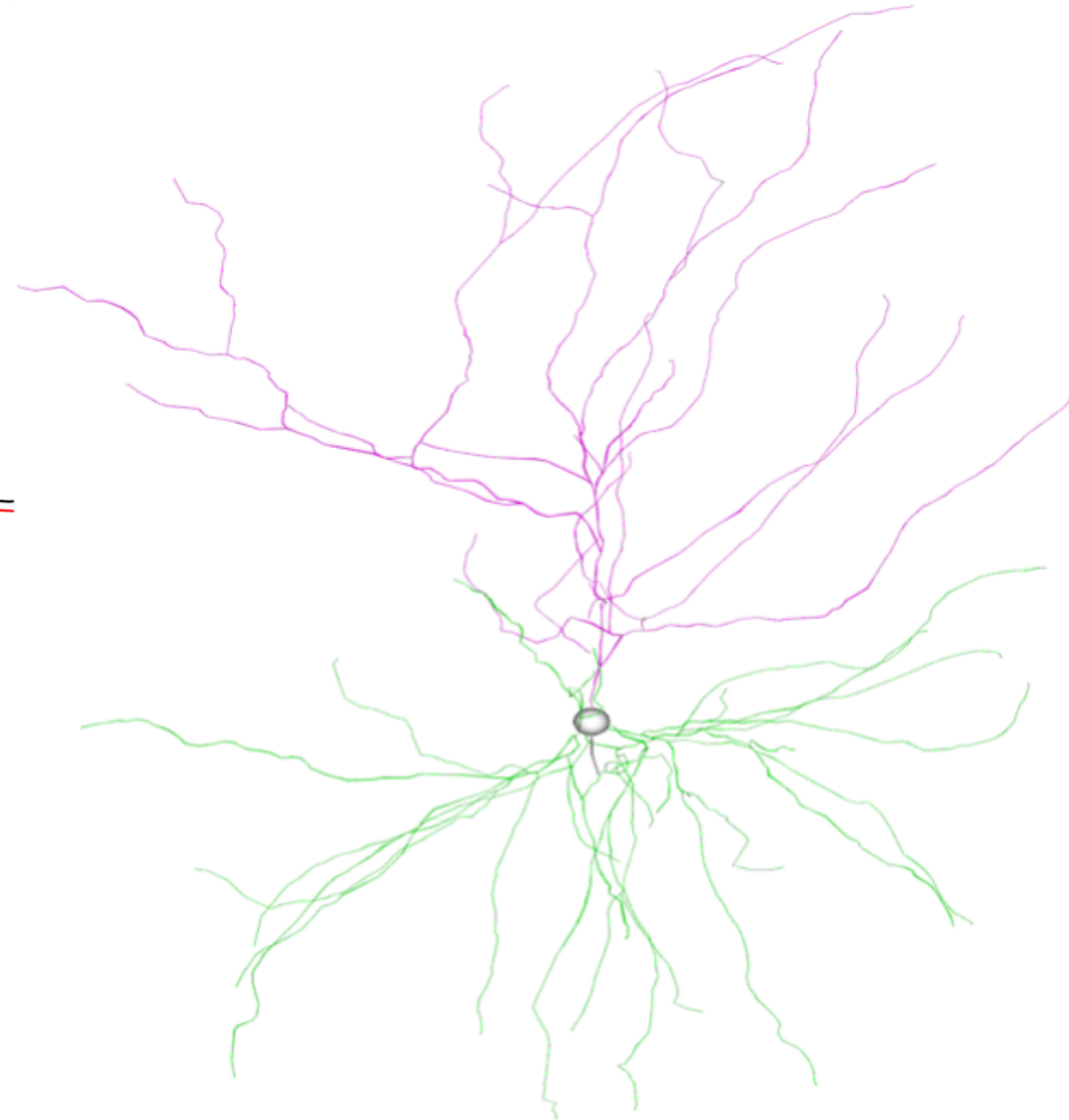
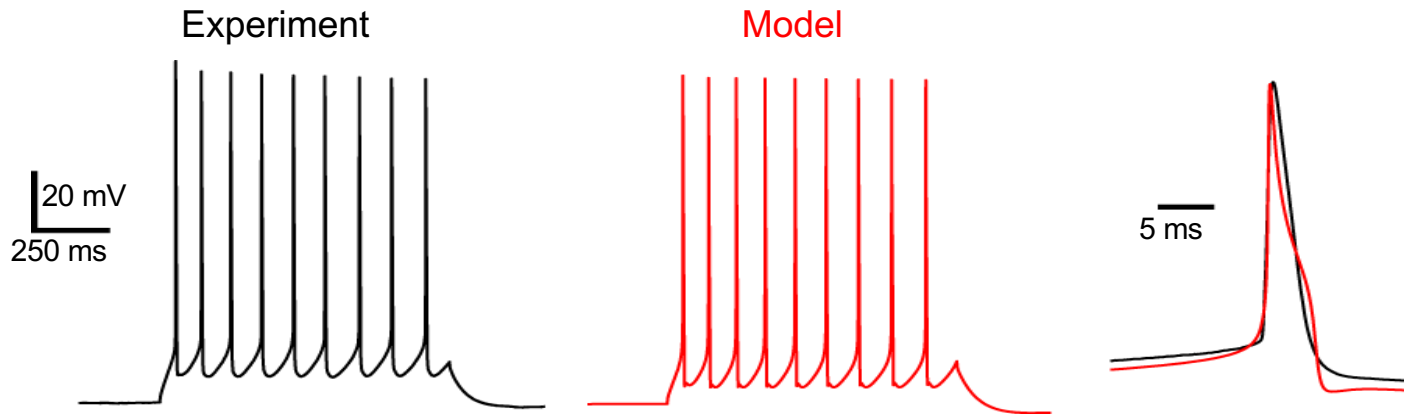
Biophysical Models of Individual Neurons



Gouwens et al.,
Nature Communications, in press.

Teeter et al.,
Nature Communications, in press.

Biophysical Models of Human Cells

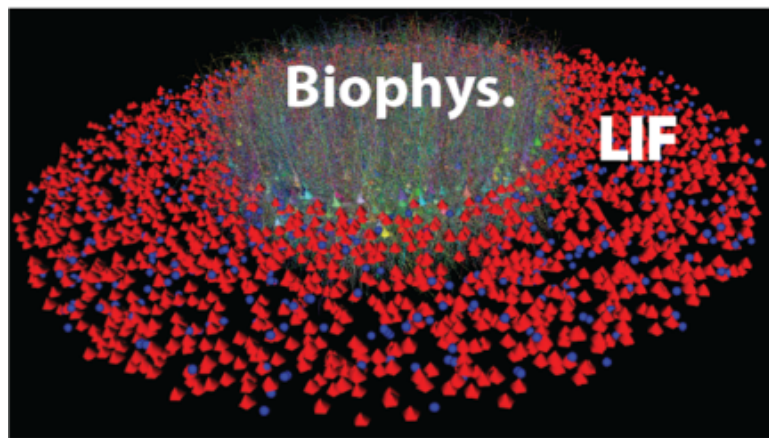
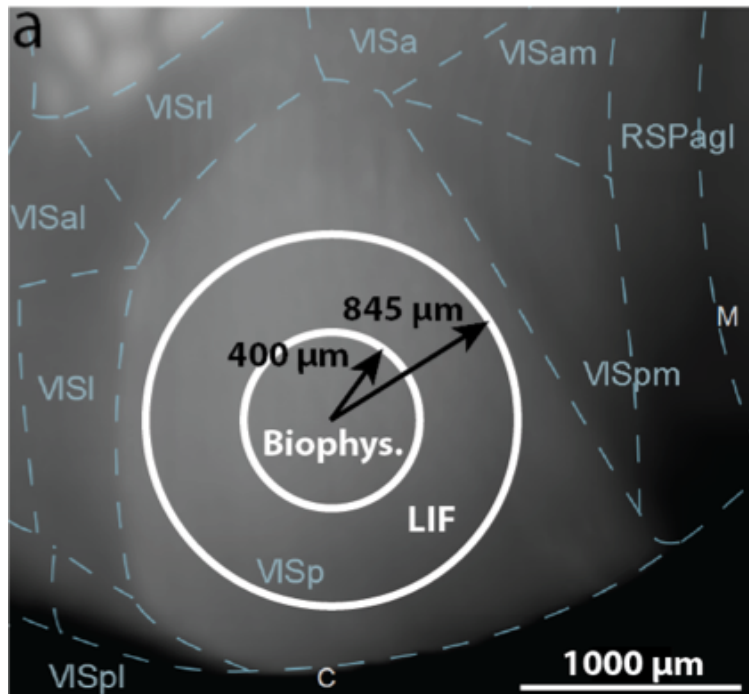
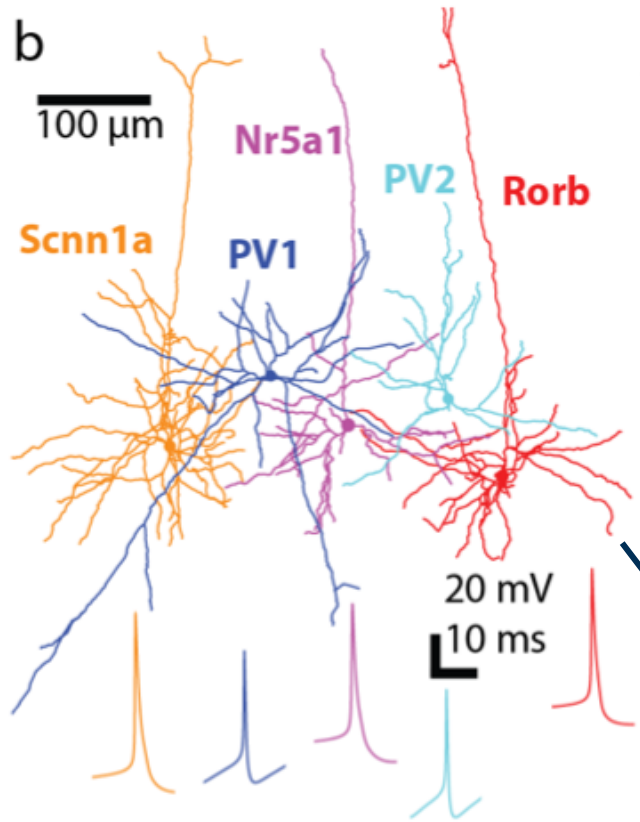


- The same optimization procedure as for mouse cells has worked successfully for the human neurons
- 43 biophysical models of human neurons are currently available in the Allen Cell Types Database



Biophysical Model of the Layer 4 of Mouse V1

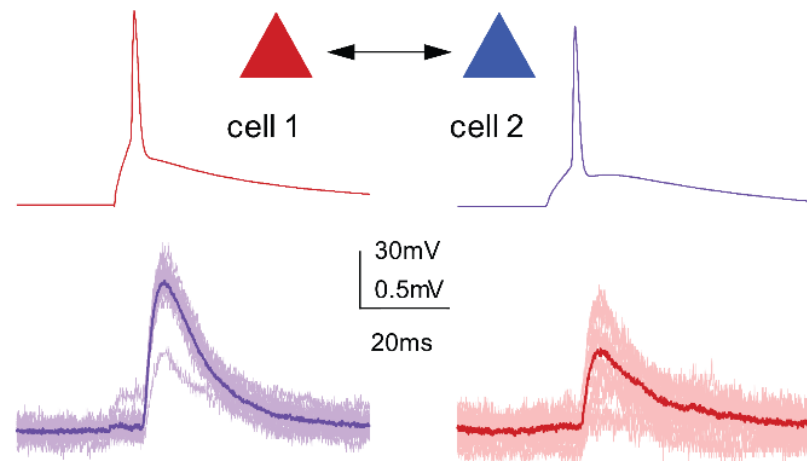
Models of Individual Cells
celltypes.brain-map.org



10,000 biophysically detailed neurons

35,000 LIF point neurons

Connectivity and Synaptic Properties



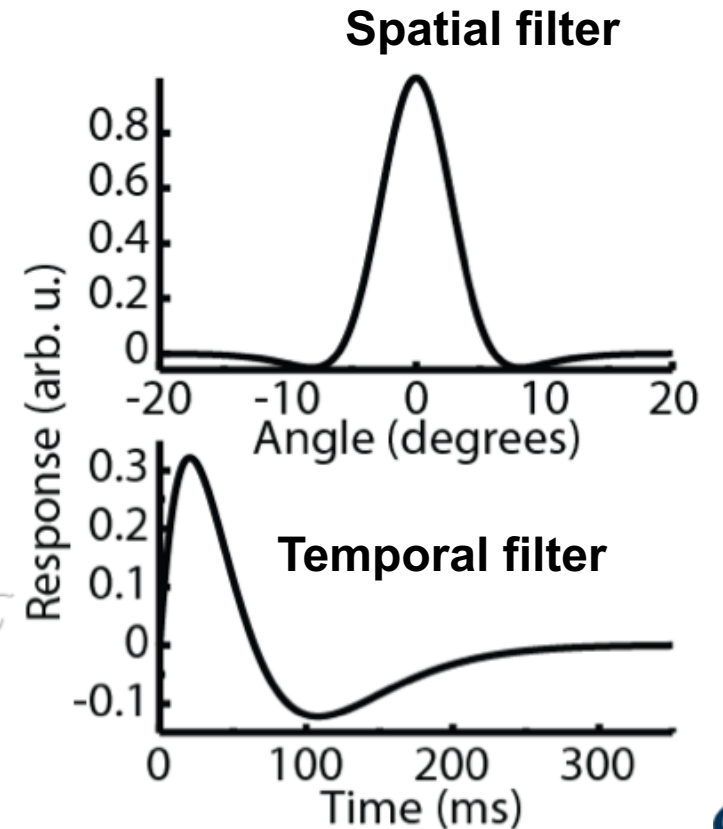
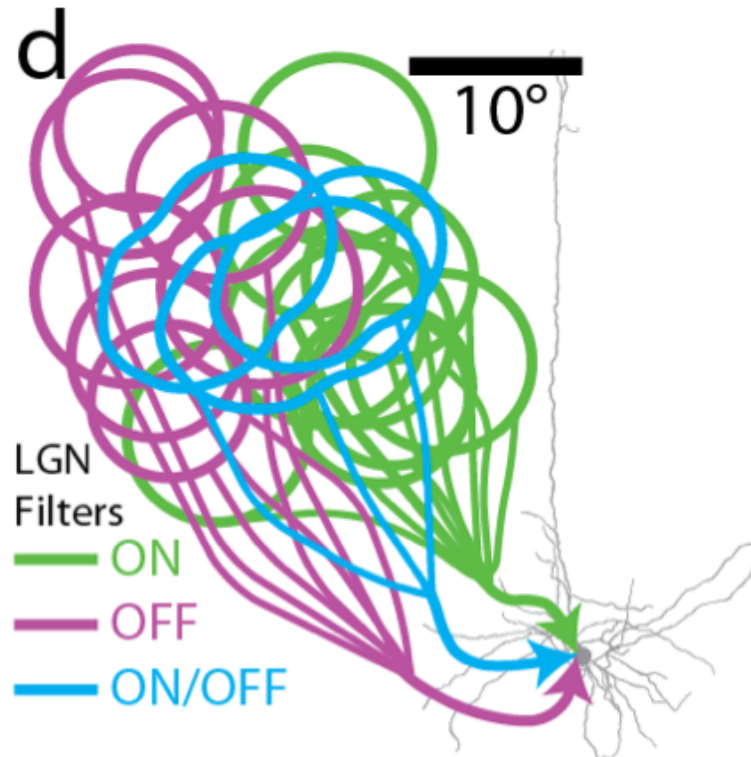
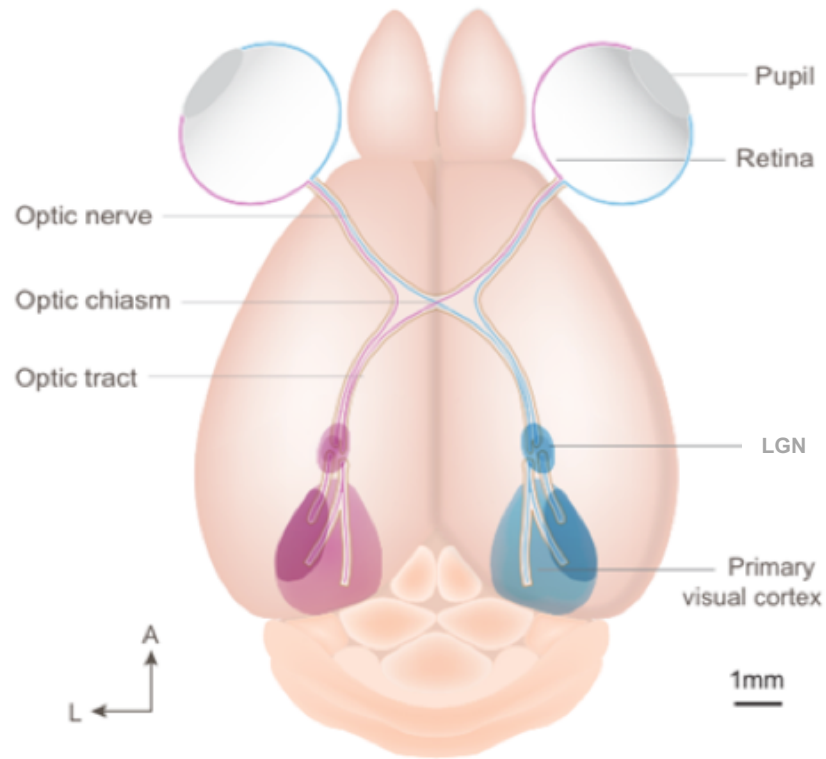
Wu et al., Mol. Brain (2009)
Perin et al., PNAS (2011)
Ko et al., Nature (2011)
Cossell et al., Nature (2015)
Lee et al., Nature (2016)

...



Model of Thalamic Inputs

- 9,000 LGN filters
- 3 'types': **ON**, **OFF**, **ON/OFF**
- Retinotopy based connectivity rules from LGN to L4.

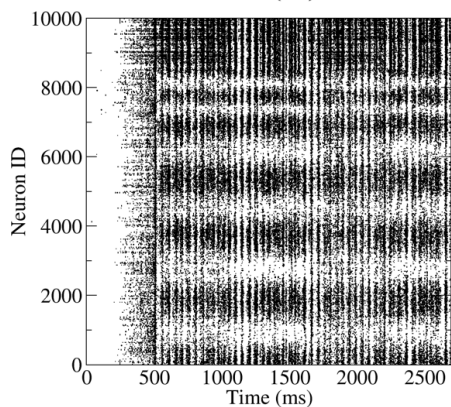
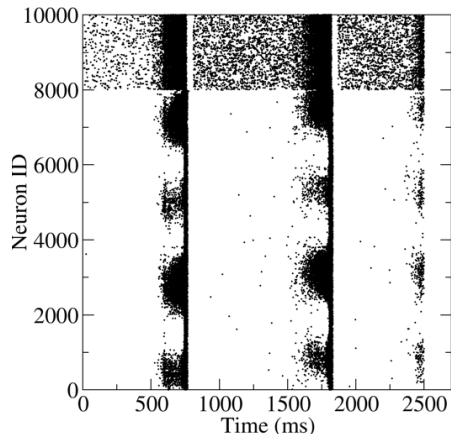
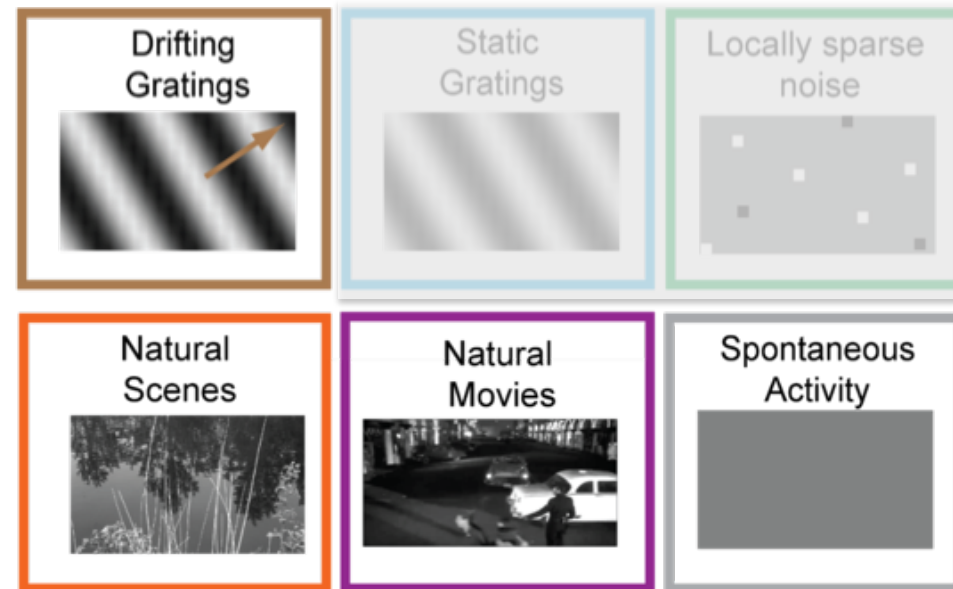


Model Tuning and Production Runs

- 3 model instantiations (“3 mice”: M1, M2, M3)
- For each, sample alternative connectivity rules

Test data:

- Spontaneous activity (20 trials)
- Gratings (8-32 conditions, 10 trials each)
- Natural Images (10 images, 100 trials each)
- Natural Movies (3 movies, 10 trials each)
- Full-field flashes (2 conditions, 10 trials each)
- Moving Bars (4 conditions, 10 trials each)



Training data:

- Spont. rate from 1 trial
- Peak rates for 1 grating trial (a single SF/TF/orientation)
- Peak response to a full-field flash from 1 trial

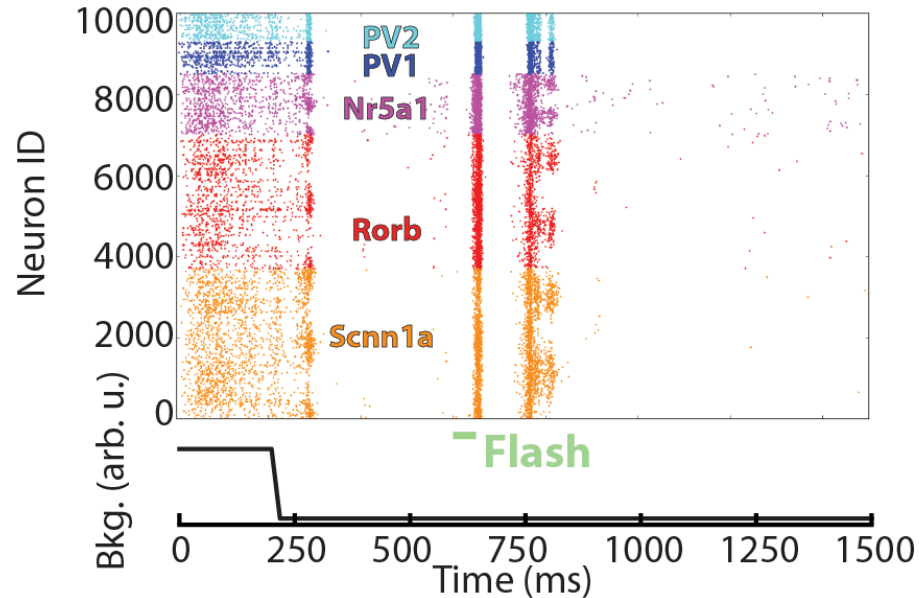
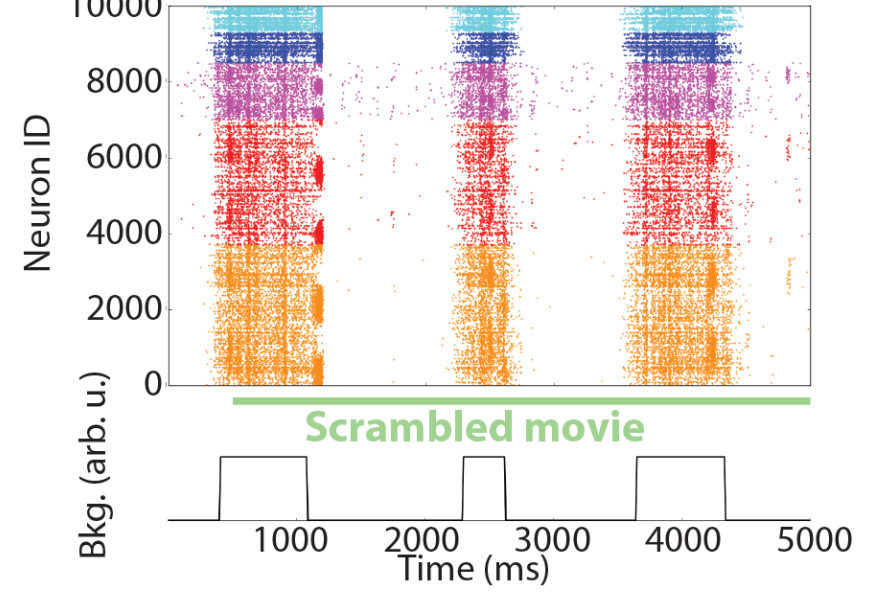
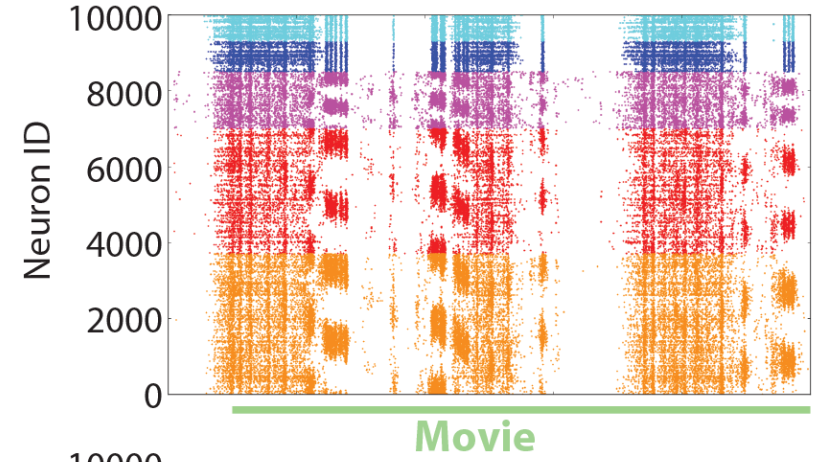
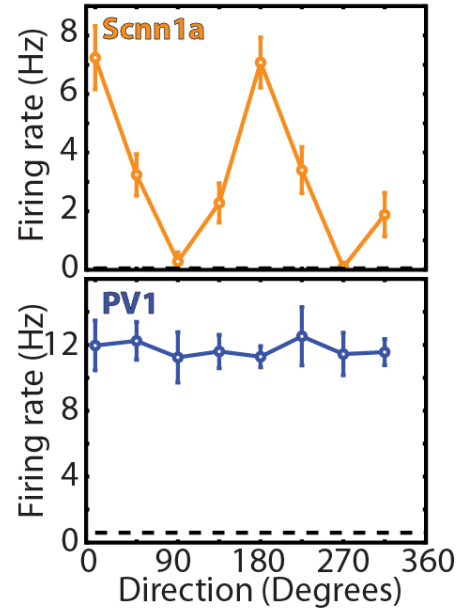
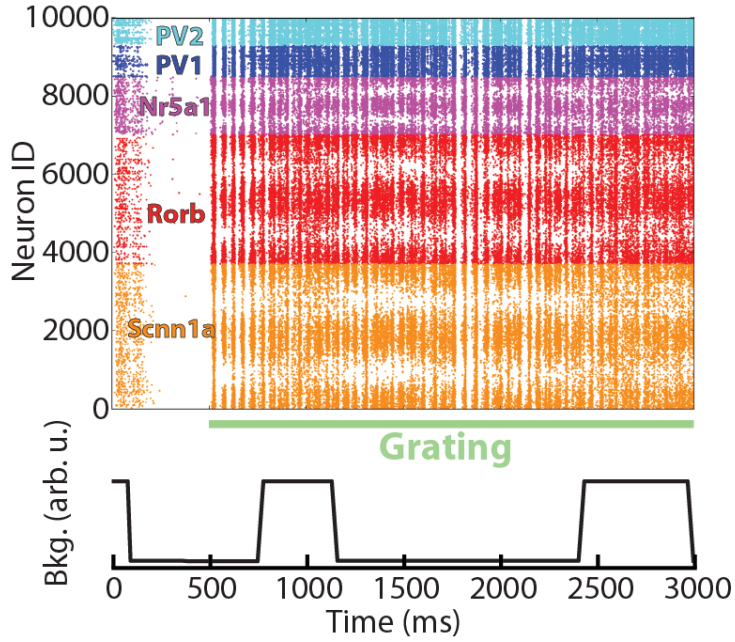
Total number of simulations: ~3,600

Total simulated time: ~3 hours

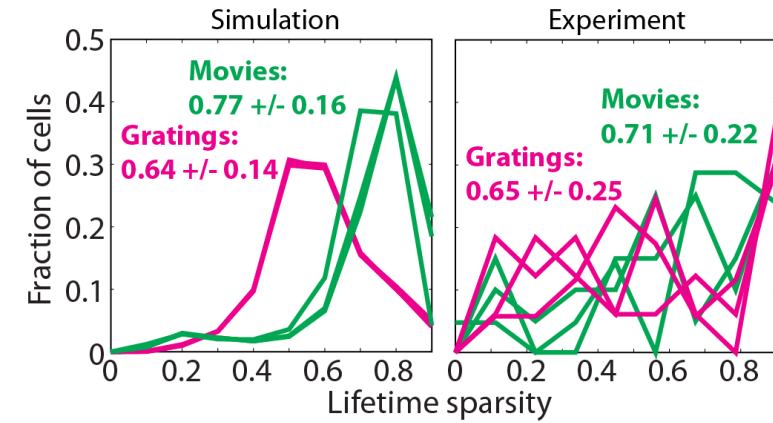
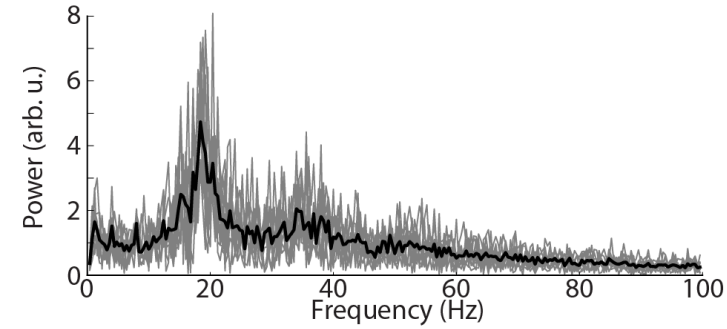
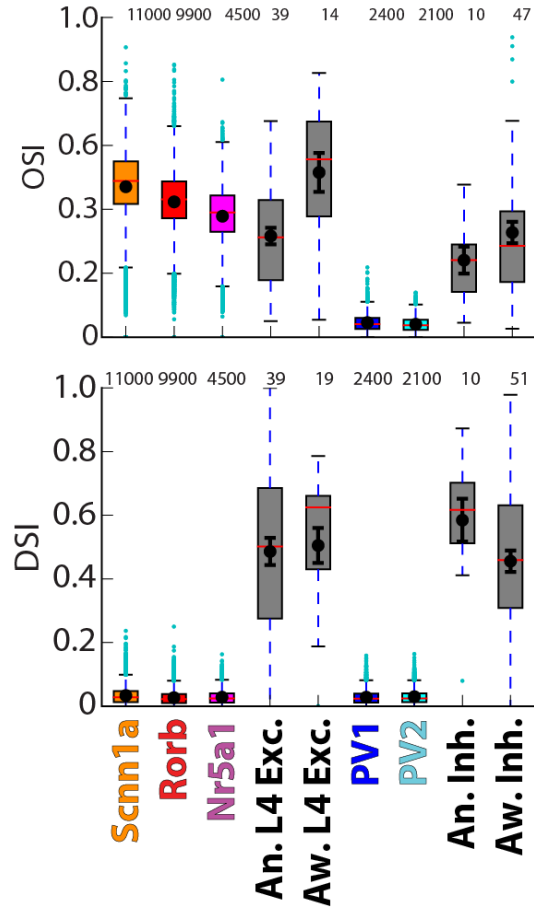
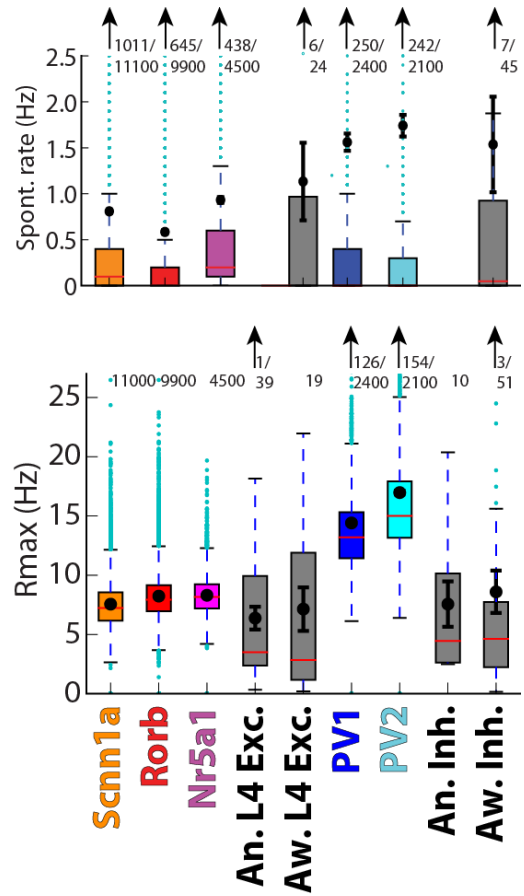
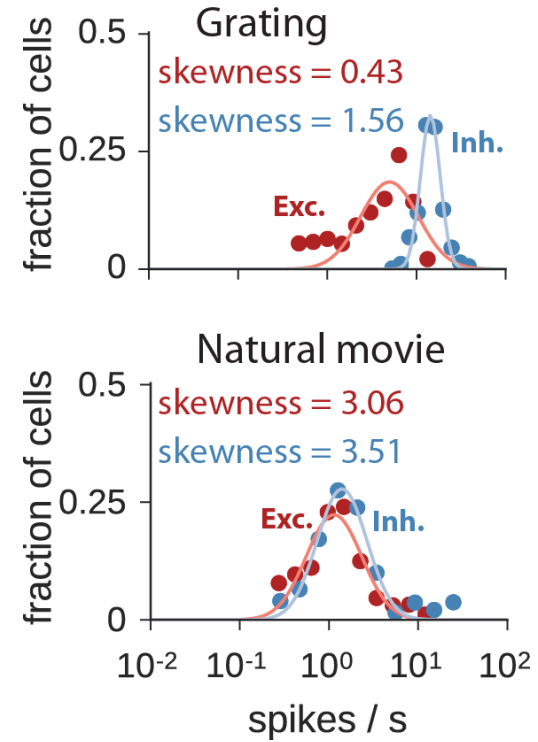
Computational costs: ~6 months on 1,500 processors



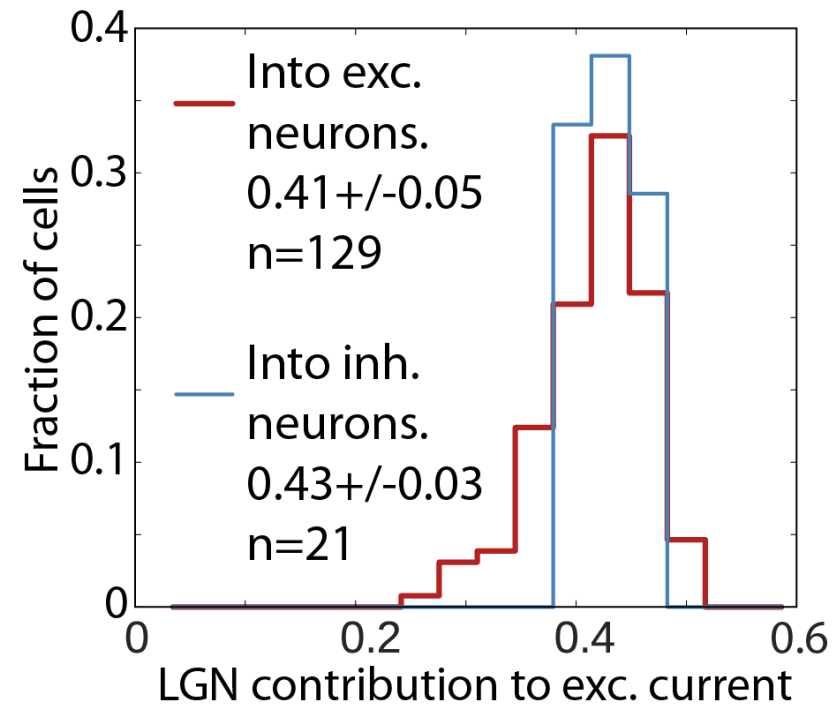
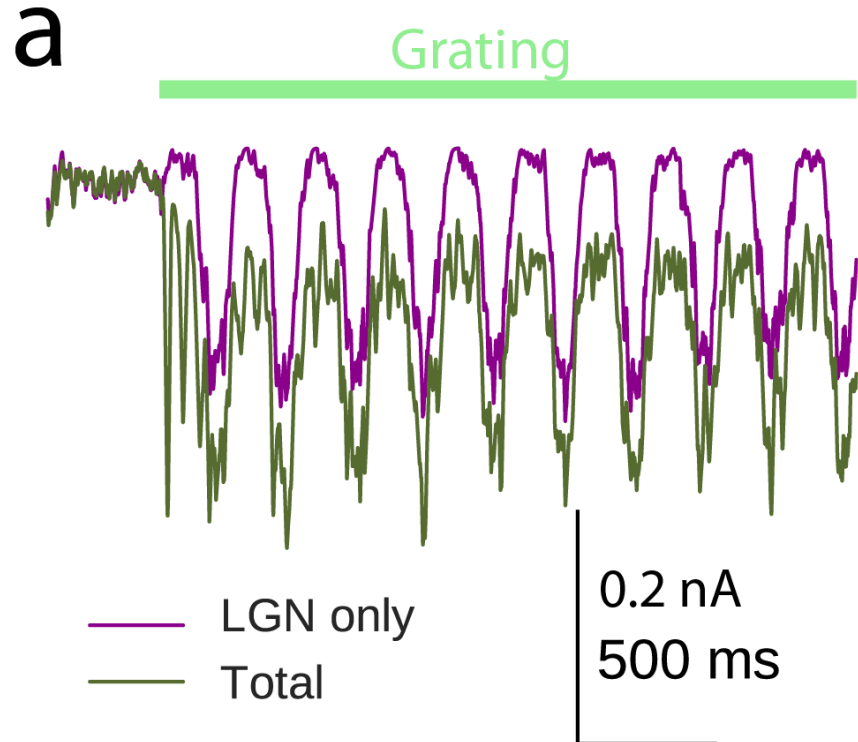
Model Responses



Benchmarking the Model against Experiments



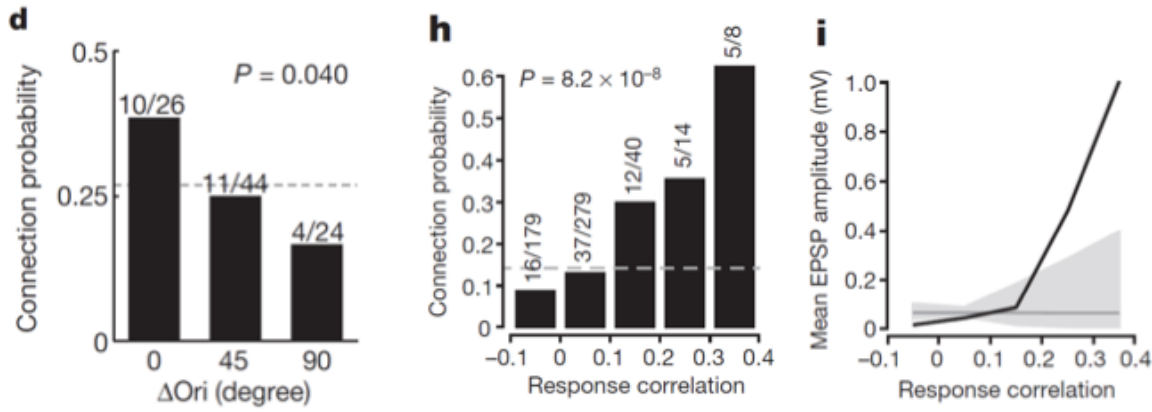
Cortical Amplification



Experiment (2 Hz):
0.36 \pm 0.02
for exc. neurons

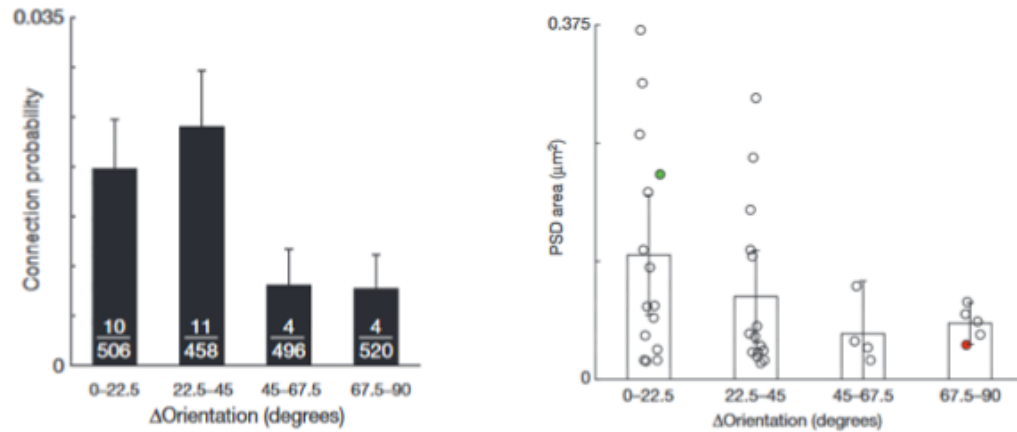
(Lien and Scanziani,
Nature Neurosci., 2013)

Like-to-Like Connectivity



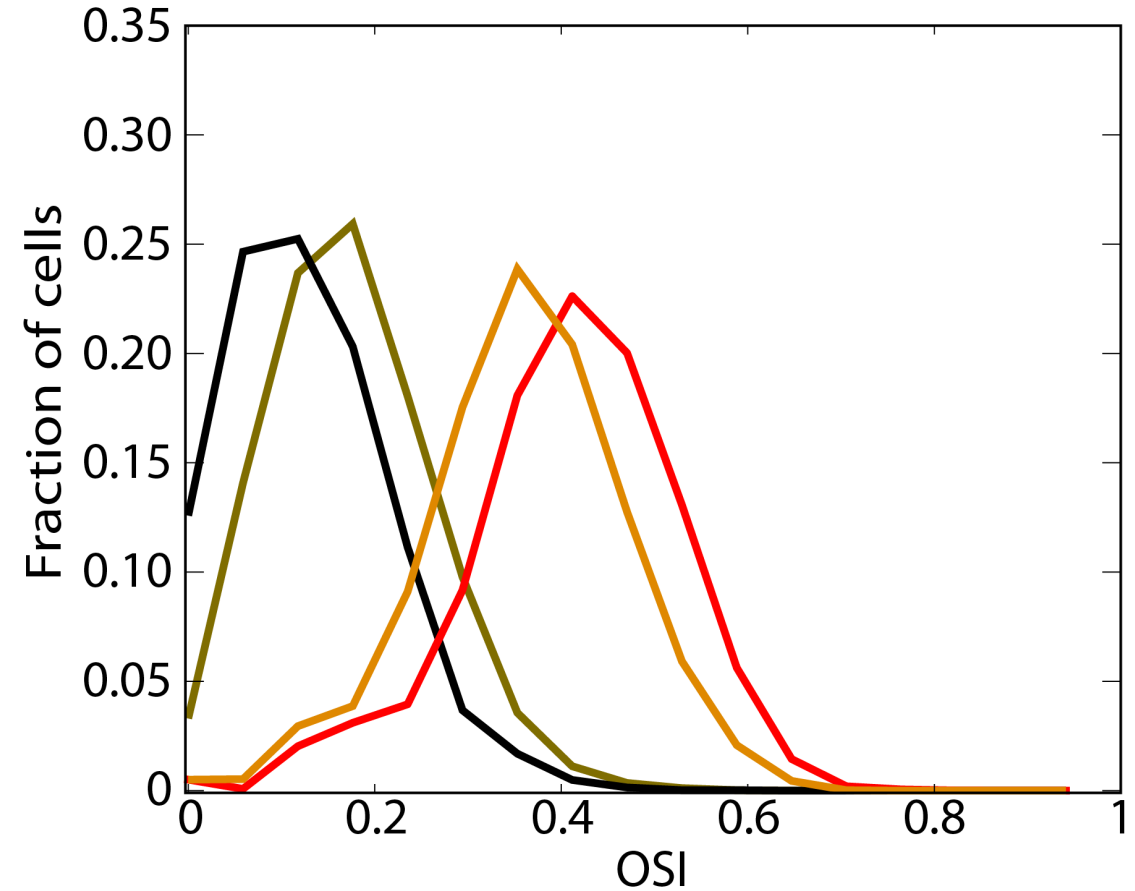
Ko et al., Nature (2011)

Cossell et al., Nature (2015)

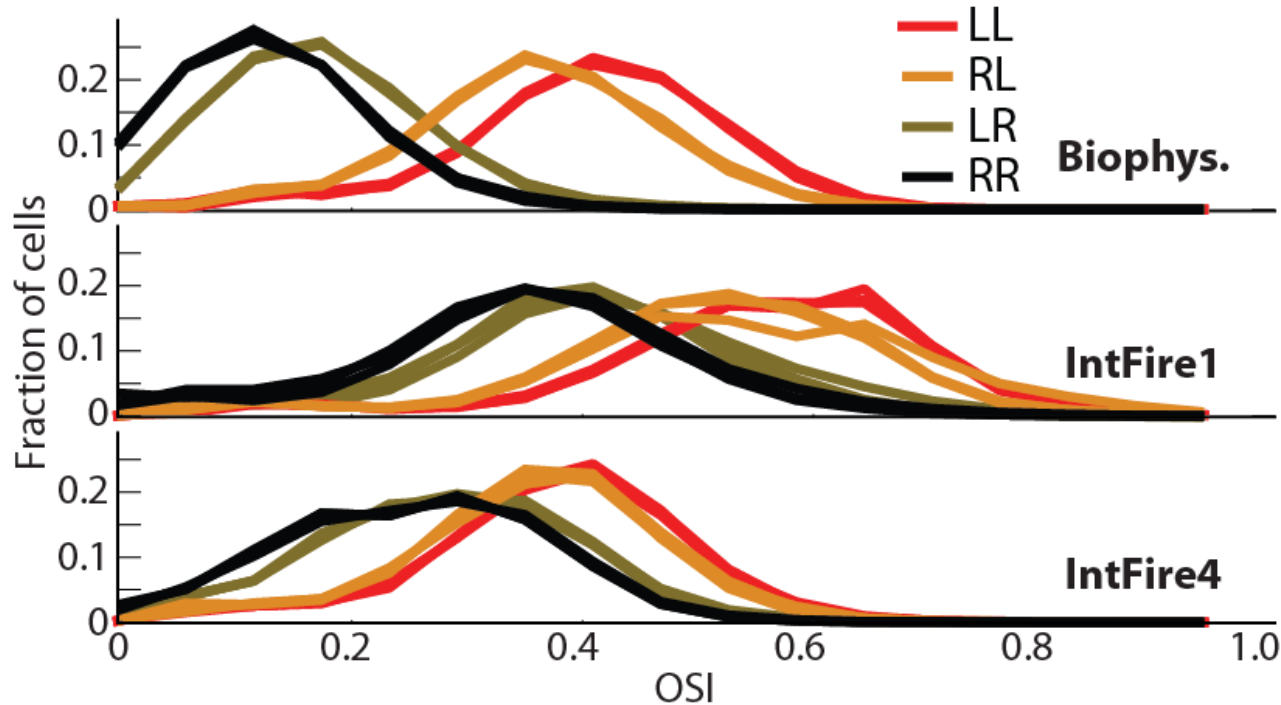
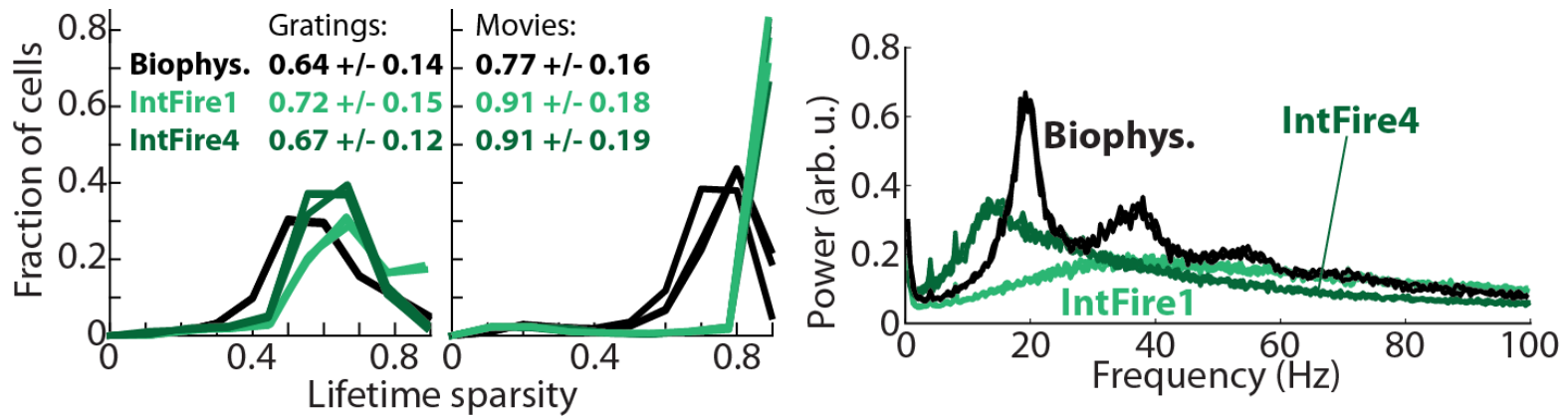
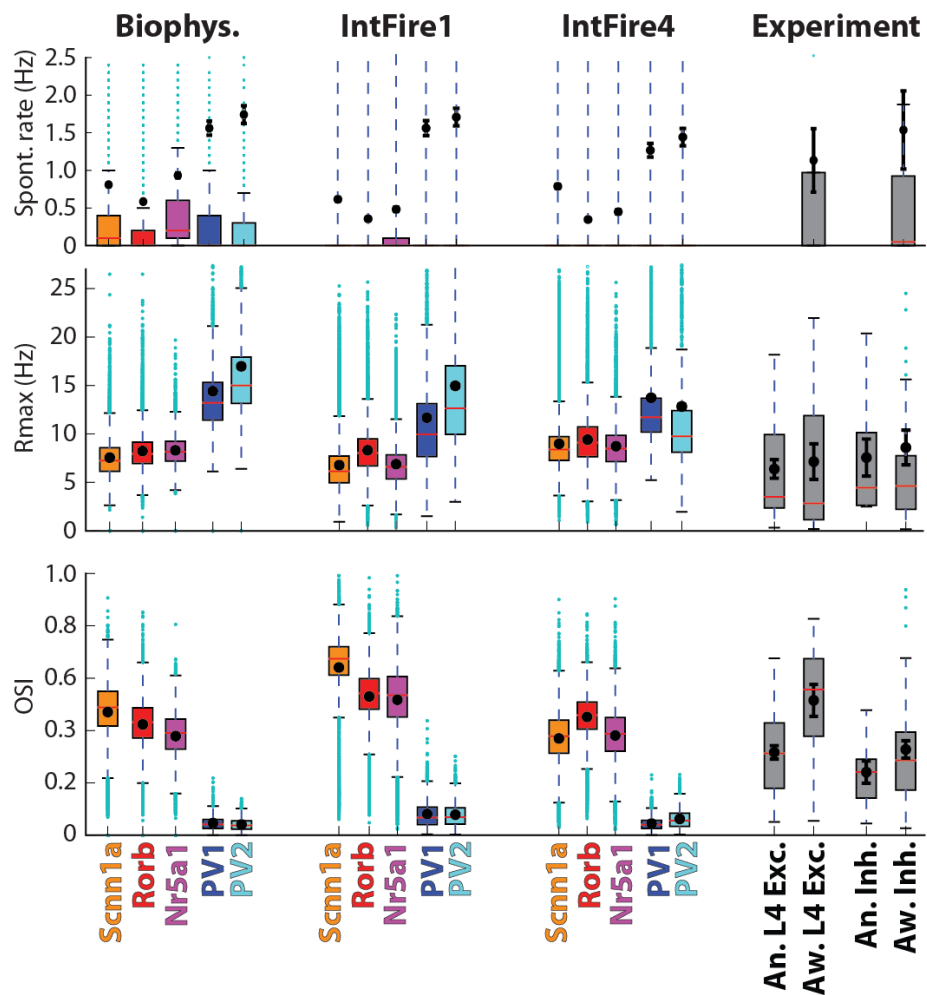


Lee et al., Nature (2016)

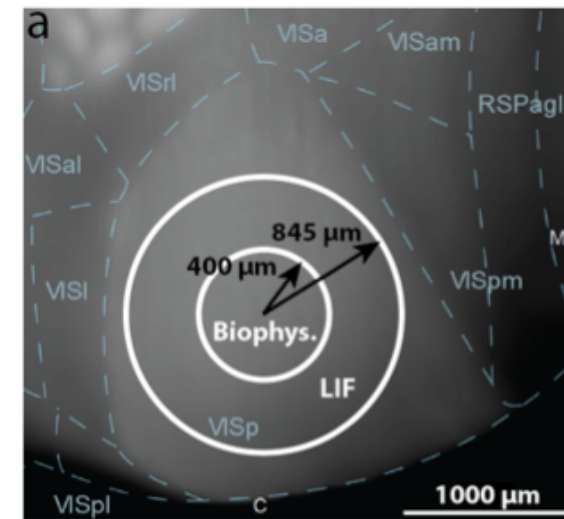
- Like-to-like connectivity and weights
- Like-to-like connectivity, random weights
- Random connectivity, like-to-like weights
- Random connectivity and weights



Replacing Biophysical Cells by Point-Neuron Models



Biophysically Detailed Model of Mouse V1



Cell Type	L1 Htr3a
N of cells	978
N of models	2

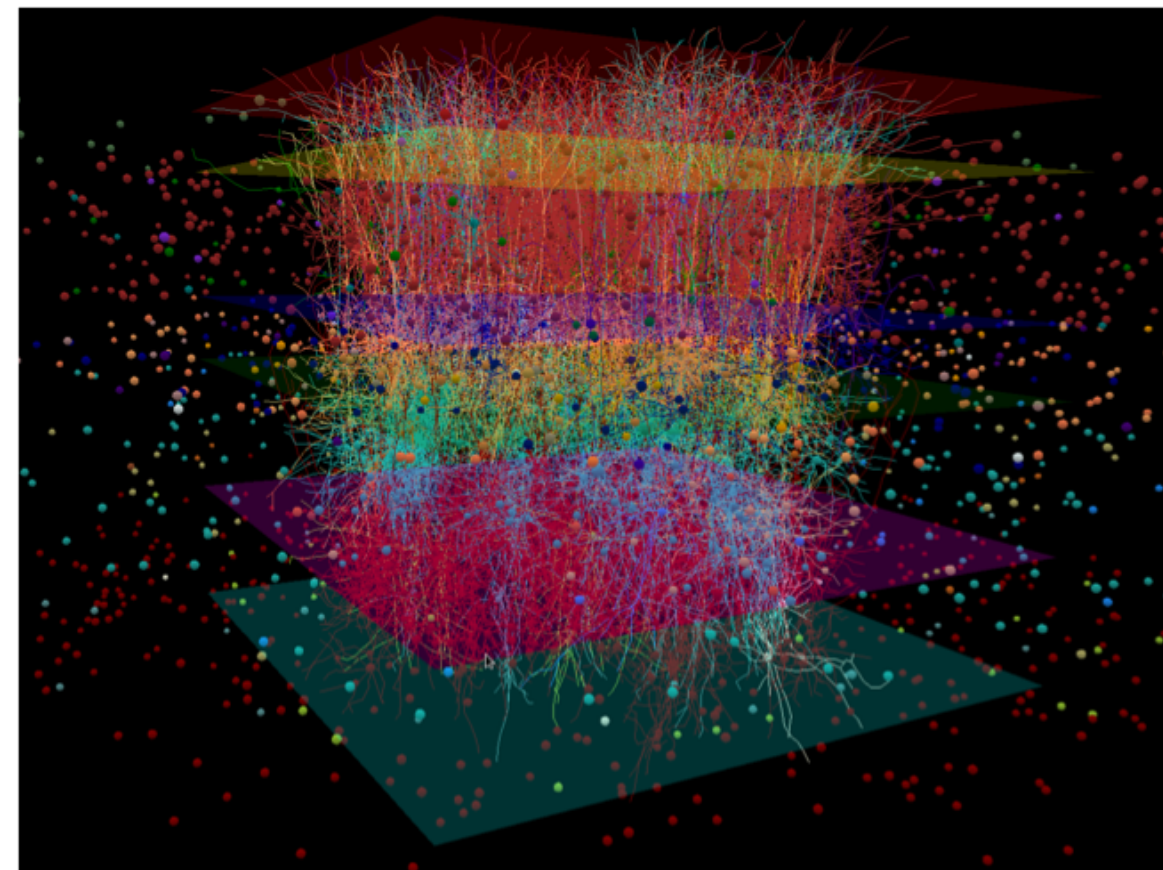
Cell Type	L2/3 Cux2	L2/3 PV	L2/3 SST	L2/3 Htr3a
N of cells	12689	640	459	1107
N of cell models	1	3	4	8

Cell Type	L4 Scnn1a	L4 Rorb	L4 Nr5a1	L4 other E	L4 PV	L4 SST	L4 Htr3a
N of cells	3072	2620	1248	3194	1025	553	270
N cell models	6	11	9	5	2	2	1

Cell Type	L5 Rbp4	L5 other E	L5 PV	L5 SST	L5 Htr3a
N of cells	6010	1448	673	594	98
N of cell models	9	18	10	5	4

Cell Type	L6 Ntsr1	L6 PV	L6 SST	L6 Htr3a
N of cells	12893	1052	1056	192
N of cell models	5	4	3	2

The model contains **51,978 biophysical cells (114 unique models); point-neuron models for the rest; 230,000 cells in total.**



Allen Institute's Databases and Allen Software Development Kit (Allen SDK)

<http://brain-map.org/>

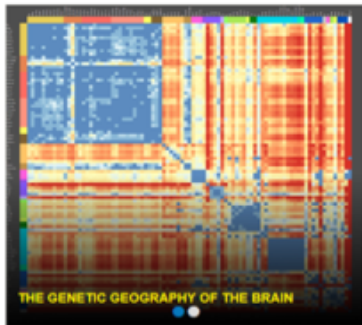
<http://alleninstitute.github.io/AllenSDK/>

ALLEN BRAIN ATLAS
DATA PORTAL

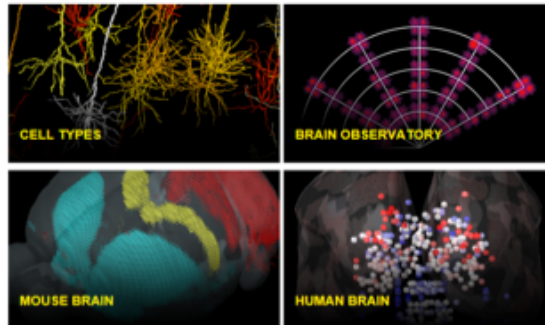
HOME GET STARTED DATA TOOLS

Search...

SCIENCE VIGNETTES



DATA & TOOLS



WHAT'S NEW

- Latest Data Release October 19, 2017
- Events & Training

ALLEN INSTITUTE PUBLICATIONS

View a full list of publications from the Allen Institute for Brain Science.

TUTORIALS

DEVELOPING MOUSE	+
NON-HUMAN PRIMATE	+
MOUSE SPINAL CORD	+
GLIOBLASTOMA	+
AGING, DEMENTIA AND TBI	+
REFERENCE ATLASES	+

CONTENTS

Install Guide

Data Resources

- Brain Observatory
- Cell Types
- Mouse Connectivity
- Reference Space
- API Access

Models

- Generalized LIF
- Biophysical

Examples

Authors

Source Documentation

- allensdk.api package
- allensdk.brain_observatory package
- allensdk.config package
- allensdk.core package
- allensdk.ephys package
- allensdk.model package
- allensdk.morphology package
- allensdk.test_utilities package

Github Profile

QUESTIONS

Send any questions using the [Send Us a Message](#) link below, or submit your question to [StackOverflow](#) using with the 'allen-sdk' tag.

If you encounter any problems using the AllenSDK, please create an issue on [Github's issue tracker](#).

WELCOME TO THE ALLEN SDK

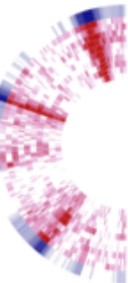
The Allen Software Development Kit houses source code for reading and processing Allen Brain Atlas data. The Allen SDK focuses on the Allen Brain Observatory, Cell Types Database, and Mouse Brain Connectivity Atlas.

ALLEN BRAIN OBSERVATORY

The [Allen Brain Observatory](#) is a data resource for understanding sensory processing in the mouse visual cortex. This study systematically measures visual responses in multiple cortical areas and layers using two-photon calcium imaging of GCaMP6-labeled neurons targeted using Cre driver lines. Response characterizations include orientation tuning, spatial and temporal frequency tuning, temporal dynamics, and spatial receptive field structure.

The mean fluorescence traces for all segmented cells are available in the Neurodata Without Borders file format ([NWB files](#)). These files contain standardized descriptions of visual stimuli to support stimulus-specific tuning analysis. The Allen SDK provides code to:

- download and organize experiment data according to cortical area, imaging depth, and Cre line
- remove the contribution of neuropil signal from fluorescence traces
- access (or compute) dF/F traces based on the neuropil-corrected traces
- perform stimulus-specific tuning analysis (e.g. drifting grating direction tuning)



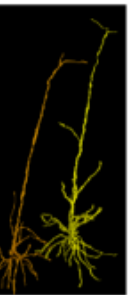
ALLEN CELL TYPES DATABASE

The [Allen Cell Types Database](#) contains electrophysiological and morphological characterizations of individual neurons in the mouse primary visual cortex. The Allen SDK provides Python code for accessing electrophysiology measurements ([NWB files](#)) for all neurons and morphological reconstructions ([SWC files](#)) for a subset of neurons.

The Database also contains two classes of models fit to this data set: biophysical models produced using the NEURON simulator and generalized leaky integrate and fire models (GLIFs) produced using custom Python code provided with this toolkit.

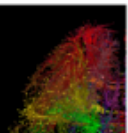
The Allen SDK provides sample code demonstrating how to download neuronal model parameters from the Allen Brain Atlas API and run your own simulations using stimuli from the Allen Cell Types Database or custom current injections:

- Biophysical Models
- Generalized LIF Models



ALLEN MOUSE BRAIN CONNECTIVITY ATLAS

The [Allen Mouse Brain Connectivity Atlas](#) is a high-resolution map of neural connections in the mouse brain. Built on an array of transgenic mice genetically engineered to target specific cell types, the Atlas comprises a unique compendium of projections from selected neuronal populations throughout the brain. The primary data of the Atlas consists of high-resolution images of axonal projections targeting different anatomic regions or various cell types using Cre-dependent



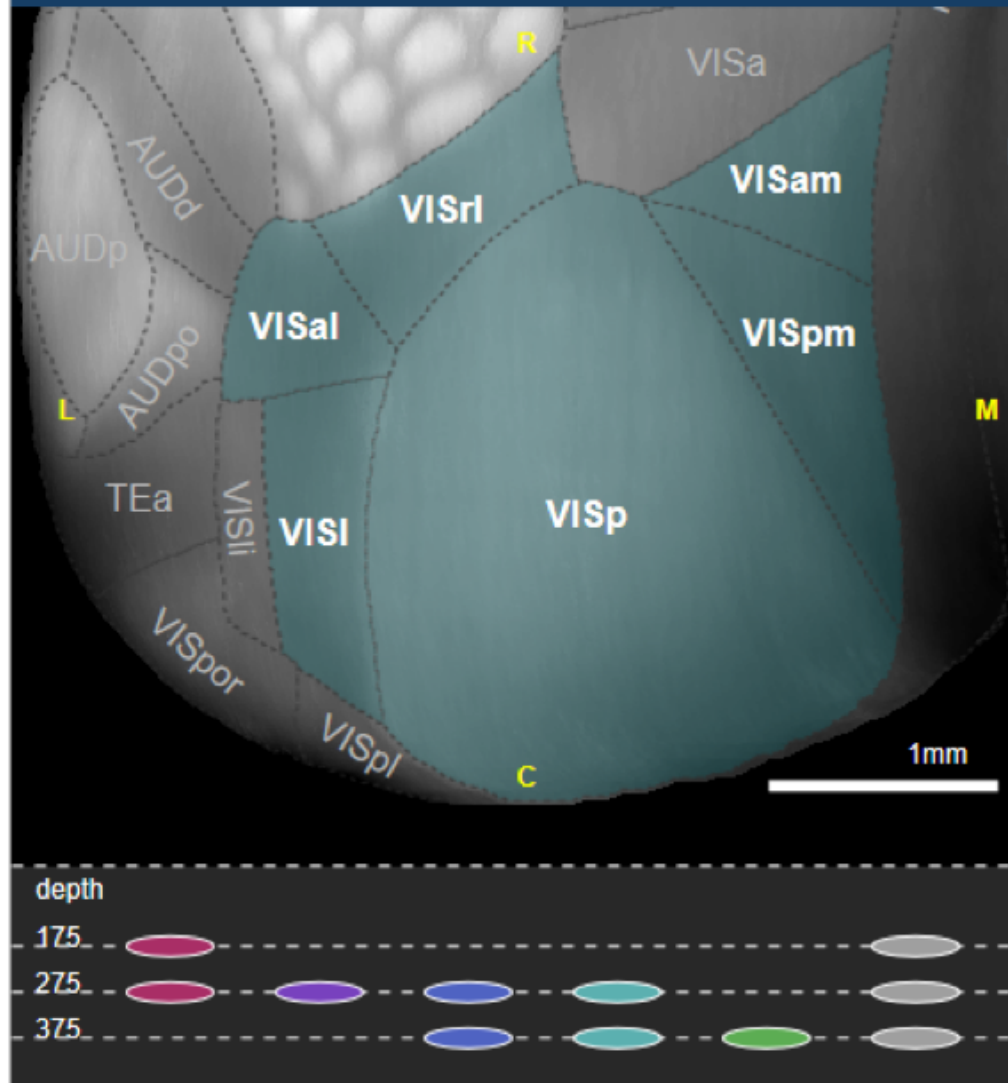
Allen Brain Observatory

<http://observatory.brain-map.org/visualcoding>

Transgenic Mouse Lines

- Cux2-CreERT2;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Rorb-IRES2-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Scnn1a-Tg3-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Nr5a1-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Rbp4-Cre_KL100;Camk2a-tTA;Ai93(TITL-GCaMP6f)
- Emx1-IRES-Cre;Camk2a-tTA;Ai93(TITL-GCaMP6f)

Cortical Imaging Locations



- Mouse Visual Cortex
- 6 Cre lines
- 6 areas
- 181 datasets
- ~45,000 neurons

Allen Cell Types Database

<http://celltypes.brain-map.org>

- Mouse and Human Cortical Neurons
- 1058 mouse and 279 human cells with ephys recordings
- 295 mouse and 107 human morphology reconstructions
- 712 mouse and 157 human cells with GLIF models
- 205 mouse and 43 human cells with biophysical models

ALLEN BRAIN ATLAS
DATA PORTAL

HOME GET STARTED CELL TYPES TOOLS

OVERVIEW CELL FEATURE SEARCH DOWNLOAD DOCUMENTATION HELP

Overview

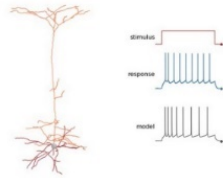
This brain cell database is a survey of biological features derived from single cell data, from human and mouse.

The database contains **electrophysiological**, **morphological**, and **transcriptomic** properties gathered from individual cells, and models simulating cell activity. At this early stage of data generation, survey coverage focused on select areas of cerebral cortex, and thalamic neurons.

Browse electrophysiological response data and reconstructed neuronal morphologies using the Cell Feature Search tool. Transcriptomic data can be accessed through the Download page.

Use the Allen Software Development Kit (SDK) to programmatically access and analyze raw data, and to run models.

Data can be downloaded by selecting individual experiments in the Cell Feature Search tool, by accessing transcriptomic RNA-Seq files via the Download page, or through the Allen SDK or API.



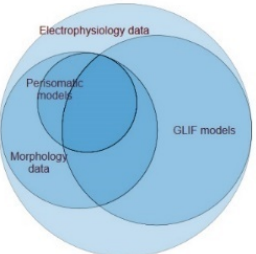
Human Data

Cells are acquired from donated brain tissue in the temporal or frontal lobes based on structural annotations described in The Allen Human Brain Reference Atlas. For electrophysiological and morphological analyses in the cortex, cells are selected based on soma shape and laminar location.

For transcriptomic analysis, individual layers of cortex are dissected, and neuronal nuclei are isolated. Laminar sampling is guided by the relative number of neurons present in each layer.

Inferior temporal gyrus (ITG)

Donor Profiles



This interactive Venn diagram shows how many cells are available for each data modality (electrophysiology, morphology, transcriptomics) and models. Select a category to view the subset of cells.

There are **279** human cells for which we have Electrophysiology data.

Include Transcriptomic Data



ALLEN BRAIN ATLAS
DATA PORTAL

HOME GET STARTED CELL TYPES TOOLS

OVERVIEW CELL FEATURE SEARCH DOWNLOAD DOCUMENTATION HELP

Morphology Summary

Donor	H17.03.008	Max Euclidean Distance	367.8
Species	Human	Number of Stems	4
ID	55570553	Number of Bifurcations	17
Hemisphere	right	Average Contraction	0.887
Area	middle temporal gyrus	ParentDaughter	0.962
Layer	4		
Sex	Female		
Age	60 yrs		
Dendrite Type	sparsely spiny		
Apical Dendrite	NA		
Clinical Observation	epilepsy		



Browse Morphology Data

Projected top view

Projected side view

3D neuron reconstruction

Truncated

Not Reconstructed

Soma

Axon

Dendrite

Apical Dendrite

View image stack

Contrast

Download Morphology

Download Morphological Measurements

Reset

145 145

SEND US A MESSAGE >

Twitter Facebook YouTube LinkedIn

About the Allen Institute
Allen Institute Publications for Brain Science

Copyright ©2017, Allen Institute for Brain Science. All Rights Reserved.

The Dynamic Brain Summer Workshop at Friday Harbor Lab



https://github.com/AllenInstitute/SWDB_2017

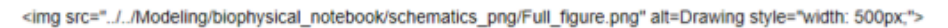
```
In [1]: import os
import sys
import pandas as pd
import numpy as np
import random

from modelingsdk.builder.networks import SynNetwork

from matplotlib import pyplot as plt
import matplotlib.image as mpimg
%matplotlib inline
```

Modeling biophysically detailed cells with NEURON

In this tutorial, we will create a network of 8 cells. Four of these cells will be biophysically detailed and four will be Leaky-Integrate-and-Fire (LIF) models.

Drawing style="width: 500px;">

Build recurrent network

For building the 8 node network, we will save the nodes (neurons) and edges (connections) separately.

```
In [2]: # Output file names
directory_name = 'network/recurrent_network/'

# Create directory if it doesn't exist
if not os.path.exists(directory_name):
    os.makedirs(directory_name)
```

We will explore the directory structure of how files are being saved later. Just know for now that we will be creating files and saving them in a directory called "network".

Building Nodes

In this section we will be creating the nodes only. They are not receiving input nor connected to each other yet. The schematic below illustrates where we will be.



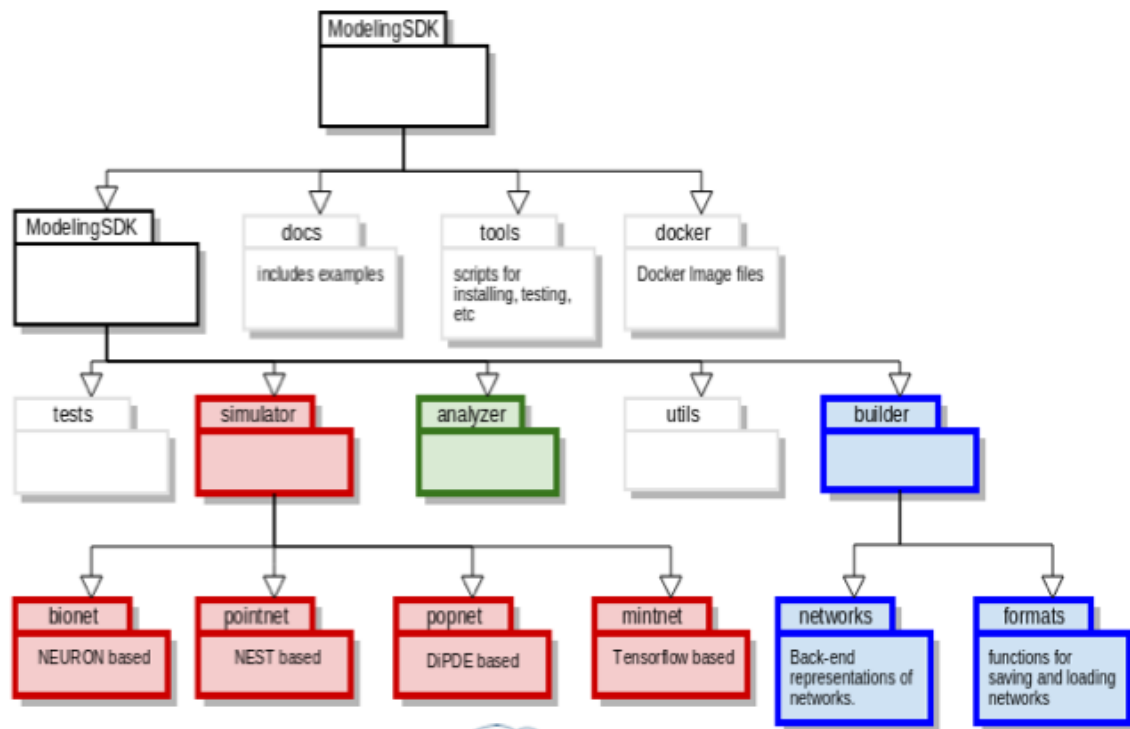
Brain Modeling ToolKit (BMTK)

Deliverables:

A software suite for building, simulating and analyzing models at different levels of resolution.

Scope:

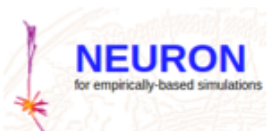
Software for 4 levels of resolution (biophysical, point neuron, population statistics, machine intelligence).



builder - An API for building large networks

simulator - A collection of interfaces for quickly running network simulations on a variety of simulators.

analyzer - Functions for analyzing and visualizing network and simulation data.



DiPDE



<https://alleninstitute.github.io/bmtk/>



THANK YOU

We wish to thank the Allen Institute for Brain Science founders, Paul G. Allen and Jody Allen, for their vision, encouragement and support.

ALLENINSTITUTE.ORG
BRAIN-MAP.ORG



ALLEN INSTITUTE *for*
BRAIN SCIENCE