

Grant Agreement:	604102	Project Title:	Human Brain Project
Document Title:	Strategic Mouse Data for the HBP Mouse Brain Atlas and for Modelling: Package Two (data)		
Document Filename:	SP1 D1.4.4 FINAL.docx		
Deliverable Number:	D1.4.4		
Deliverable Type:	Data		
Work Package(s):	WPs 1.1, 1.2, 1.3, 1.4 (WPs involved in writing this document)		
Dissemination Level:	PU		
Planned Delivery Date:	M30/31 Mar 2016		
Actual Delivery Date:	Draft M29/29 Feb 2016; Revised draft M30/18 Mar 2016		
Authors:	Javier DEFELIPE, Pilar F. ROMERO, UPM (P59), Elaine MARSHALL , UEDIN, (P71) T1.4.1		
Compiling Editors:	Pilar F. ROMERO, UPM (P59), Elaine MARSHALL , UEDIN, (P71) T1.4.1		
Contributors:	<p>Chris PONTING, Grant BELGARD UOXF (P8) T1.1.2</p> <p>Seth GRANT; Marcia ROY; Zhen QUI; Mélissa KIERON, UEDIN (P71) T1.1.3</p> <p>Rafael LUJAN, UCLM (P100); Ryuichi SHIGEMOTO, IST (P88) T1.1.4</p> <p>Simon BERNECHE, SIB (P95), T1.1.5</p> <p>Torsten SCHWEDE, Ioannis XENARIOS, UNIBAS (P101), T1.1.5</p> <p>Antonino CATTANEO, SNS (P92), T1.1.6</p> <p>Enrico QUERUBINI, EBRI (P85), T1.1.6</p> <p>Michele MIGLIORE, CNR (P83), T1.1.6</p> <p>Bruno WEBER, Matthias SCHNEIDER, UZH (P63) T1.2.1</p> <p>Francesco PAVONE, Ludovico SILVESTRI, LENS (P34), T1.2.2</p> <p>Javier DEFELIPE, Lidia ALONSO, UPM (P59), T1.2.3</p> <p>Tamás FREUND, Szabolcs KALI, IEM HAS (P28), T1.2.4</p> <p>Yun WANG, WMC (P79), T1.2.4</p> <p>Francisco CLASCA, UAM (P57), T1.2.5</p> <p>Ángel MERCHÁN, UPM (P59), T1.2.6</p> <p>Douglas RAMSTRONG, UEDIN, P(71), T1.3.1, T1.3.2</p>		
Coordinator Review:	<p>EPFL (P1): Jeff MULLER, Martin TELEFONT</p> <p>UHEI (P45): Sabine SCHNEIDER, Martina SCHMALHOLZ</p>		
Editorial Review:	EPFL (P1): Guy WILLIS, Lauren ORWIN, Colin McKINNON		
Abstract:	The objective of SP1 is to generate neuroscientific concepts, knowledge, experimental datasets and tools, which will be used to build models for the		

simulation of the brain. These models will be integrated, for example, into neuromorphic systems (SP9) or neurorobotics controllers (SP10) in order to create cost-effective, energy-efficient, high-performance systems. An important role for SP1 is to provide data and knowledge to support activities undertaken by other SPs. This report provides a detailed account of the strategic mouse data for the HBP Mouse Brain Atlas and for modelling. Within this report we outline the datasets which have been deposited in the HBP Mouse Brain Atlas and those used for modelling. The package includes data on the distribution of major proteins; quantitative description of synaptic connections on neurons; numbers, distributions and relative densities of cells in selected regions and, where possible, across the whole brain; statistical parameters characterising spatial arrangements between neurons, glia and blood vessels; a high-resolution quantitative synaptic map of exemplar brain regions; EM blocks scans and volume analysis of exemplar brain regions with quantification of the neuropil organization. A short report herein describes each level of data and provides quantitative indicators of the completeness of the data compared to the targeted dataset and a projected full dataset to be generated by the research community. The report also includes a description of the way the data has been used in models of the mouse brain.

Keywords:

Single cell transcriptomics, synapse proteome, axonal projections, synaptic maps, ultrastructural data, brain vasculature, cell distributions, cell morphologies

Available at:

<https://www.humanbrainproject.eu/ec-deliverables>

Table of Contents

1. Introduction	5
2. Summary Table of Tools and Methodologies (Including Validation and Limitations)	6
3. Strategic Mouse Data for the HBP Mouse Brain Atlas and for Modelling.....	11
3.1 First Draft Transcriptomes of Major Hippocampal and Cerebellar Cell Types	11
3.2 First Synapse Proteome for a Specific Cell Type.....	13
3.3 Anatomical Reconstructions of Hippocampal Cells	16
3.4 Anatomical Reconstructions of Neocortical Cells	17
3.5 Single-cell and Bulk Connectomics of Thalamocortical and Corticocortical Projection Neurons	20
3.6 Distribution of Major Synapse Proteins	23
3.7 Numbers, Distributions and Relative Densities of Cells in Selected Regions	31
3.8 Synaptic Maps of Identified Neurons.....	36
3.9 High Resolution Optical Synaptic Maps and Electron Microscope Blocks Scans and Volume Analysis of Exemplar Brain Regions.....	41
3.10 High-Resolution 3D Maps of the Brain Vasculature	48
3.11 Potassium Channels Kinetics	51
3.12 Trans-Synaptic Signalling and Receptor Kinetics	55
3.13 Data Aggregation, Analysis and Dissemination	59
4. How Can Platform Developers Provide You With Feedback on The Data?.....	67
Annex A: References	71
Annex B: Dataset Information Cards.....	73

List of Figures and Tables

Table 1: SP1 tools and methodologies	6
Figure 1: Heat map showing list of genes and protein levels detected in complex with PSD95 in the CA1, CA3, CL6, and ST conditional lines.....	15
Figure 2: Synaptic proteome expression levels across the seven integral regions of the mouse brain	28
Figure 3: Heat-maps of synapse densities and synapse types coded by different colours across the 5 mouse brain sections	28
Figure 4: 3D reconstruction of dendritic shaft of CA1 pyramidal cells with immunogold labelling for SK2	29
Figure 5: 2D mapping of glutamate receptors in CA1 pyramidal cells by SDS-digested freeze-fracture replica labelling.....	30
Figure 6: Fine tuning of the segmentation algorithm for cell detection	34
Figure 7: PSD95 labelling of a layer III pyramidal neuron in the somatosensory cortex of a PSD95 EGFP mouse.	37
Figure 8: Synaptic maps of Identified Neurons	38
Figure 9: Pyramidal Cell Explorer	39
Figure 10: 3D Quantification Methodology	42
Figure 11: Imaging Processing.....	42
Figure 12: Synaptic Map of the Hippocampus.....	43
Figure 13: Example of a correlative optical and FIB/SEM synaptic map	44
Figure 14: Ontology	63
Figure 15: Simple PDS pepke Kappa nmda	64
Table 2: SP1 Numerical Indicators (D1.4.1).....	65

1. Introduction

The objective of SP1 is to generate neuroscientific concepts, knowledge, experimental datasets and tools, which will be used to build models for the simulation of the brain. These models will be integrated, for example, into neuromorphic systems (SP9) or neurorobotics controllers (SP10) in order to create cost-effective, energy-efficient, high-performance systems. An important role for SP1 is to provide data and knowledge to support activities undertaken by other SPs.

The aim of this deliverable is to provide a summary of the level of data generated within SP1 during M18-M30. This report also includes a description of the way the data has been used in models of the mouse brain. This report provides a detailed account of the strategic mouse data for the HBP Mouse Brain Atlas and for modelling. Within this report we outline the datasets which have been deposited in the HBP Mouse Brain Atlas and those used for modelling. The package includes data on the distribution of major proteins; quantitative description of synaptic connections on neurons; numbers, distributions and relative densities of cells in selected regions and, where possible, across the whole brain; statistical parameters characterising spatial arrangements between neurons, glia and blood vessels; a high-resolution quantitative synaptic map of exemplar brain regions; EM blocks scans and volume analysis of exemplar brain regions with quantification of the neuropil organization. A short report herein describes each level of data and provides quantitative indicators of the completeness of the data compared to the targeted dataset and a projected full dataset to be generated by the research community. The report also includes a description of the way the data has been used in models of the mouse brain.

The SP1 Molecular section has generated the first draft of the transcriptome and proteomes for specific cell types. During the Ramp-Up Phase, SP1 partners within WP1. Tasks 1.1.1-1.1.3 collaborated on developing and optimising techniques and skills to produce the first draft transcriptome and synapse proteome datasets. Tasks 1.1.2 and 1.1.3 have exceeded their target level dataset. The transcriptome dataset generated within HBP has been complemented and enhanced by the release of data from Sten LINNARSSON, (Science, 2015). The HBP transcriptome and synapse proteome datasets represents a novel resource to the wider community. All datasets have been submitted to HBP collaboration and have been used to develop cross-cutting projects in development of theoretical and computational models.

The SP1 Cellular section has generated the data scheduled for M1-M30. Furthermore, some tasks even have exceeded their target level dataset such as T1.2.3 and T1.2.6. The data generated include anatomical reconstructions of hippocampal and neocortical cells, single-cell and bulk connectomics of thalamocortical and corticocortical projection neurons, numbers, distributions and relative densities of cells in selected regions, synaptic maps of identified neurons, high resolution optical synaptic maps and electron microscope blocks scans and volume analysis of exemplar brain regions and high-resolution 3D maps of the brain vasculature.

2. Summary Table of Tools and Methodologies (Including Validation and Limitations)

Table 1: SP1 tools and methodologies

Tool/Methodology	Application	Validation	Limitations
Single cell isolation	Staining for known cell type markers to estimate proportions of each type dissociated	Imaging capture of single cells on Fluidigm C1 platform. 3. Staining for live cells.	Often lose processes in cell isolation; biased representation of cells captured
Single cell transcriptomics analysis	T1.1.2	1. Cross referencing with classical cell type markers as well as several genome-wide datasets on cell type-specific expression and spatial patterning 2. Working with collaborators to validate novel subtypes by in situ hybridisation and RNAscope	Assumes discrete differences, or at least extremes; harder to see a continuum with this approach; technical noise
Optimisation of Synaptosome and PSD complex isolation	T1.1.3	Micro-scaled version of synaptosomes and PSD complex isolation from <50mg of brain tissue (or 1 mouse hippocampus), compatible with quantitative LC-MS/MS analysis enabling identification of 1,906 synaptic proteins and quantitation of 1,215 synaptic proteins across 7 integral regions of the mouse brain.	Requires minimum 50 mg of starting material.
Analysis of human brain tissue samples from the MRC Edinburgh Human Brain Bank	T1.1.3	Proteome integrity developed based on the level of NR2B degradation/HUSPIR ratio (Bayes et al., 2014).	Can only be accurately performed using anti-NR2B monoclonal antibody from BD Biosciences.

Tool/Methodology	Application	Validation	Limitations
Western blotting, imaging and quantitation	T1.1.3	42 samples from n=4 undiseased/'control' human brains screened by western blotting using anti-NR2B antibody. HUSPIR ratio performed. Quantitative confocal microscopy methods developed for mapping of excitatory synapses (anti-PSD95 immunostaining), quantitative bioinformatics software package developed, quantitative western blotting for NR1, NR2B, PSD95, PSD93, Sap102 and Sap97 performed for all 12 neocortical regions for n=4 cases.	Requires identification and validation of antibody specificity.
Mass Spectrometry analysis of PSD proteins samples	T1.1.3	LC-MS/MS analysis methods developed for label-free quantitative mass spectrometry of 3ug of PSD proteins isolated from 300mg human brain tissue. Quantitation 1,902 PSD proteins across 12 neocortical regions of the human brain developed. Synaptic protein abundance atlas developed for the human neocortex.	Requires a minimum of 100 mg of starting material and samples having a HUSPIR ratio of equal to or greater than 1.2.
SDS-FRL technique	Subcellular localisation at the EM level - 2D mapping.	Considered the immunoEM with the highest sensitivity and now used by several laboratories.	Difficulty to follow or identify different compartments of the same neuron.
ImmunoFIB/SEM	Subcellular localisation at the EM level - 3D mapping	Considered the EM technique that allows analysis of long series of EM images.	Not usable for receptors or ion channels located at synaptic sites; only along the extrasynaptic plasma membrane.

Tool/Methodology	Application	Validation	Limitations
Software GPDQ (<i>Gold Particle Detection and Quantification</i>)	Quantification of immunogold particles for the SDS-FRL technique	Under validation.	Only usable for 2D mapping; not usable for 3D mapping of receptors or ion channels.
Metal particles of different elements	High resolution visualisation of two molecules in SDS-FRL	Under validation	Custom conjugation is necessary for antibodies, and multiple sequence-probe combinations are necessary for tagged proteins
Automated 3D segmentation with Confocal microscope	Image of stacks of neurons stained with anti-NeuN antibody; Sections stained with DAPI	Method validated in different cortical areas	The automatic segmentation algorithm must be supervised due to heterogeneity in the labelling, especially in those regions with very high cell density
Automatic segmentation of cells (upgrade)	Image of stacks of neurons stained with anti-NeuN antibody; Sections stained with DAPI	Method validated in different cortical areas Method to improve the accuracy for counting neuronal and non-neuronal cells	Automatic segmentation of different cell types must be tuned specifically for every cell population and brain region, due to heterogeneity of the labelling
3D Reconstruction with Neurolucida (Hippocampus)	Pyramidal cells and different types of Inhibitory neurons	Reconstructions of the soma, dendritic arbour and axon terminals of the neurons were checked by an expert anatomist. The reconstruction process itself is highly reproducible for the signal-to-noise ratio characteristic of this method. Tissue shrinkage due to processing was measured.	Detailed reconstruction of the exact branching pattern of the axonal arbour is very labour-intensive; therefore, only the 3D locations of the axonal boutons are routinely recorded (a full reconstruction of the axonal arbour can be done on demand). Precise measurement of dendritic diameters requires advanced post-processing and validation.
3D Reconstruction with Neurolucida (Neocortex)	Golgi cells; Excitatory and inhibitory neocortex neurons; synaptic connections; Cortical pyramidal cells and their projections	*	*

Tool/Methodology	Application	Validation	Limitations
<i>In utero</i> electroporation technique	Cortical pyramidal cells	Validated in specific cortical areas (Layer 5)	NA
Axonal projection analysis	<i>In vivo</i> single-neuron transfection, immunohistochemical visualisation and 3D digital reconstruction-measurement of complete thalamocortical projection axons and terminal varicosities	Method validated in 2 thalamic nuclei and several neocortical areas	Small varicosities may be underestimated. Analysis must thus include parallel stereological estimations (Aransay et al. 2015 Front. Neuroanat. May 19;9:59. doi: 10.3389/fnana.2015.00059)
Method of automated 3D counting of puncta	PSD95 puncta in hippocampus and somatosensory cortex	Validated to generate synaptic maps of identified individual neurons and mouse brain	The development of the segmentation algorithm in certain brain regions due to significant variation on the shape of the PSD
3D FIB-SEM Technology	Quantification of synaptic junctions in hippocampus and somatosensory cortex	Validated to generate synaptic maps of the hippocampus and somatosensory cortex, combined with confocal microscopy of PSD95 puncta	The resolution of confocal microscopy is limited, so the identification of PSD95 puncta may underestimate the number of synapses. Electron microscopy unambiguously identifies synapses, but it can only be performed on much smaller fields than confocal microscopy. For this reason we think that the correlative use of FIB/SEM and confocal microscopy will provide more reliable data than both techniques used independently
3D Reconstruction with X-ray tomographic microscopy	Vasculature system at sub-micrometre resolution	Tools for stitching, segmentation, and reconstruction tested on small datasets	NA
3D Reconstruction with serial two-photon microscopy	Large murine cortical samples	Tools for vascular reconstruction tested on small datasets	NA
3D Cell localisation with light sheet microscope	Fluorescently Labelled cells	Validated for cell localisation in the whole brain	Needs to be scaled up on parallel architectures

Tool/Methodology	Application	Validation	Limitations
CLARITY	Clearing and staining (optional) of entire mouse brain or of portion of human brain	Validated in both mouse and human brain tissues	Current protocol is time consuming. Clearing is not always optimal to enable light sheet microscopy
3D Reconstruction with serial two-photon microscopy	Sub-cellular reconstructions in brain subregions; Single thalamocortical axons	Validated to reconstruct large portion of mouse brain (e.g. hippocampus) [Costantini et al., Sci. Rep. 2015]	Imaging is a bit slow, and the apparatus must require less human interaction in order to apply the method to a large number of samples
Establishment of informatics analysis and data management pipeline for linking SP1 molecular and cellular data	Data mining molecular datasets at the level of molecular and cellular entity	The approach was validated in collaboration with SP5 against six published molecular studies	Rapidly moving scientific domain. Underpinning ontology appropriate to state-of-the-art at end 2015
Development and maintenance of informatics tools allowing integration of molecular data from multiple data sources	Bridging from raw molecular datasets to abstracted views required for modelling & simulation	Analysis method validated against community reference datasets. Simulation method validated against pre-existing simulator	Simulator at PoC stage, needs code optimisation

3. Strategic Mouse Data for the HBP Mouse Brain Atlas and for Modelling

3.1 First Draft Transcriptomes of Major Hippocampal and Cerebellar Cell Types

3.1.1 Data Description

3.1.1.1 Task / group responsible for generating data

T1.1.2, Chris PONTING, UOXF

3.1.1.2 Data, tools and methodologies storage location(s) (and links?)

The first draft transcriptomes of major hippocampal and cerebellar cell types derive from Task 1.1.2 and is led by Chris PONTING (UOXF). To generate these data, Smart-seq2 was implemented on the Fluidigm C1 platform, permitting for the first time the generation of full-length transcriptomes from individual cells in these regions.

Dataset Information Card (DIC):

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/7/

3.1.1.3 Description of data

Cells were dissociated from male C57BL/6J mice at P40 and run over a Percoll gradient before selecting the fractions that we determined were particularly enriched in neurons. We then prepared libraries from the wells that, on visual inspection in the C1, appeared to contain intact single cells, without obvious doublets, debris, or other contaminants. We discarded libraries that failed the QC process, ultimately procuring sequence from 1519 libraries from hippocampus and 573 libraries from cerebellum. In hippocampus we discarded sequenced libraries as outliers expressing a random mix of cell type markers, leaving 1322 single cell libraries, which we were able to cluster into high-level types in the following proportions:

- 697 neurons, of which 184 are unambiguously pyramidal neurons (61 subclusters) and 143 unambiguously interneurons (44 subclusters)
- 372 oligodendrocytes (46 subclusters)
- 74 astrocytes (25 subclusters)
- 60 microglia (21 subclusters)
- 23 endothelia (7 subclusters)
- 96 mural (27 subclusters)

We further hierarchically clustered each of these types into several subclusters each, as enumerated above. We are currently working out the appropriate level at which to separate these subclusters to represent true cell subtypes as opposed to transient cell states (we anticipate that only a subset of the subclusters above will represent bona fide cell types).

In cerebellum we clustered cells into at least five high-level cell types and 136 subclusters in total (only some of which we expect to represent bona fide cell types). An update on these cell types and subclusters will be posted by the end of March, prior to the end of the reporting period.

We therefore have, for each of these high level clusters, cells assigned to one of many subclusters of subtypes, and the levels at which genes are expressed in each of these clusters.

3.1.1.4 Completeness of data

These datasets exceed in scope and size the datasets we had anticipated producing, in that we captured more cell types and subtypes using the shotgun capture approach than we could have obtained with a more targeted approach. In addition to the mouse CA1 transcriptomes reported in Zeisel *et al.* (see below), there are a couple of ongoing efforts to produce single-cell transcriptomes in hippocampus. As far as we know, this is the only dataset reporting full-length whole-cell single-cell transcriptomes in mouse hippocampus. We are also not aware of efforts to produce such data in cerebellum, making this particularly interesting. Using this approach we detect more genes expressed in a cell of any given type than Zeisel *et al.*

There are two especially relevant datasets generated by the community over the past 30 months. The first, from HBP collaborator Sten LINNARSSON, has been discussed in detail in previous reports and is published in Zeisel *et al.* (Science, 2015). Briefly, this dataset represents shotgun captured cells with sequenced tags to count transcripts from juvenile mouse somatosensory cortex and CA1. The second, from the Allen Institute, has been published in Tasic *et al.* (2016) and describes single cells (especially neurons) targeted in young adult mouse primary visual cortex. These are excellent cortical complements to the data generated here in whole hippocampus and cerebellum.

3.1.1.5 Data quality and value

We had several layers of quality control at every stage, which ultimately resulted in high-quality full-length transcriptomes from 1322 hippocampal and 573 cerebellar cells, which clustered well according to known markers.

In addition to comprehensively covering all of hippocampus and cerebellum, the method we used permits examination of alternative splicing and provides a more comprehensive enumeration of the genes expressed in any given cell.

3.1.1.6 Data usage to date

We have delivered our data not only to the Neuroinformatics team (Martin TELEFONT and Catherine ZWAHLEN) but also directly to modellers (Daniel KELLER in Task 6.5.3 - molecular-level models of NGV coupling, to incorporate models of glia metabolism in cortical microcircuit simulation - and Paul TIESINGA and Rembrandt BAKKER in Task 5.4.1 - predictive neuroinformatics - neuronal addressing system).

3.1.1.7 Are the data considered final?

These data are not considered final, nor were they meant to be. Our timeline concerned producing a first draft of these data, and we continue to analyse and refine them.

3.1.2 Provenance

SP1, WP1.1, Task 1.1.2: Chris PONTING (UOXF)

3.2 First Synapse Proteome for a Specific Cell Type

3.2.1 Data Description

3.2.1.1 Task / group responsible for generating data

T1.1.3, Seth Grant, UEDIN

3.2.1.2 Data, tools and methodologies storage location(s) (and links?)

Cell-type specific proteomes from the brain regions of interest were isolated from knockin mice with Tandem Affinity Purification (TAP) tags inserted into the endogenous locus of post synaptic density-95 (PSD-95). One-step FLAG affinity purifications were used to isolate synaptic proteomes from conditional PSD-95 TAP tag knockin model from the following cell types: parvalbumin interneurons, CA1 pyramidal neurons, CA3 pyramidal neurons, and cortical pyramidal neurons.

Mass spectrometry (Orbitrap Velos Pro hybrid mass spectrometer, Thermo Scientific) was used to determine protein abundances, based on peptide intensity measured, in specific cell types. All raw data was processed using MaxQuant version 1.2.7.4. Peptide matches were made using the integrated Andromeda search engine, which accesses the mouse UniProtKB database. Data processing was done using Perseus. Statistical and cluster analysis was performed in Perseus.

Dataset Information Card (DIC):

Mouse Synaptic Proteome

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/9/

3.2.1.3 Description of data

Transgenic mice (male, 8 weeks, n=6-12 in each group) were used in this study. Conditional PSD95 TAP/+ mice were crossed with four lines of mice expressing region specific cell-specific expression of Cre recombinase:

- 1) CaMKII T29-Cre (T29) expresses Cre recombinase in the CA1 pyramidal cells of the hippocampus
- 2) Grik4-Cre expresses Cre recombinase in the CA3 pyramidal cells of the hippocampus
- 3) Adora2a-Cre expresses Cre recombinase in striatopallidal neurones within the striatum
- 4) Ntsr1-Cre expresses Cre recombinase in layer 4 cortex pyramidal neurons.

Using these crosses it was possible to purify PSD95 complexes from specific cell types in a manner that excludes PSD95 complexes in neighbouring cell types. For the mass spectrometry analyses, 25mg of protein was used per replicate (n=3-4) on the constitutively tagged PSD95 TAP line. A total of 789 proteins were detected across all sample sets. Of these,, 630 detected proteins were discarded due to either missing values from multiple replicates across samples or they were designated as background due to significant presence in the negative controls. Permutation based False Discovery Rate (FDR) was used for truncation and was set to a threshold of 0.05. A constant S0 value of 0 was used throughout. The remaining 159 proteins were significantly more abundant in the experimental samples compared to the negative controls, and were considered to be true components of the PSD95 complexes in specific cells (Figure 1).

3.2.1.4 Completeness of data

Four cell-types specific synaptic proteomes have been isolated and analysed, which exceeds the anticipated dataset of one.

These datasets are the first known cell type specific synapse proteomes.

There have been no other reports of cell type specific synapse proteomes, primarily because the state-of-the-art transgenic technology is the only known way to obtain this data from intact mouse brain. Various indirect approaches involving ribosome tagging have been developed but these do not measure synaptic proteins directly. The technology is now poised to be exploited for comprehensive studies across an expanded number of cell types.

3.2.1.5 Data quality and value

LS/MS data from key candidates was validated by comparison of data with the publicly available Allen Brain database containing genome wide spatial gene expression data from the mouse brain. No PSD95 was seen in any of the wild type controls, and the light and heavy immunoglobulin chains of the FLAG antibody are the only major visible bands.

Permutation based False Discovery Rate (FDR) was used for truncation and was set to a threshold of 0.05. A constant S0 value of 0 was used throughout. A normal distribution is used, and the mean and standard deviation is adjusted to simulate the signal of low abundance proteins. The data was also normalised to housekeeping genes (ribosomal) to ensure that all replicates within different sample sets contained the same amount of PSD95.

These datasets were designed to be integrated with mouse region synapse proteome data and synaptome mapping (see below) and with electrophysiological and behavioural data linked to the specific molecules. These are unique integrated datasets that permit new modelling and simulation linking genes, proteins, cells, circuits and behaviour.

3.2.1.6 Data usage to date

Hierarchical clustering by T1.1.3, (GRANT, UEDIN) has been used to produce a heatmap comparing the mean detected protein levels across brain regions samples. Differences in composition of PSD95 protein complexes were observed across brain regions and cell types. As outlined in section 3.3.1, this dataset will be processed in collaboration with Prof Douglas ARMSTRONG, UEDIN, WP1.3. This data will be used by groups involved in modelling biological data on synapses in computational models. As outlined in 3.5.1 we anticipate partners in SP5, SP6 and SP9 to interact with this data.

3.2.1.7 Are the data considered final?

These data are considered to be a robust draft of specific cell type synapse proteomes.

3.2.1.8 Publications connected to the gathered data

A manuscript is in preparation.

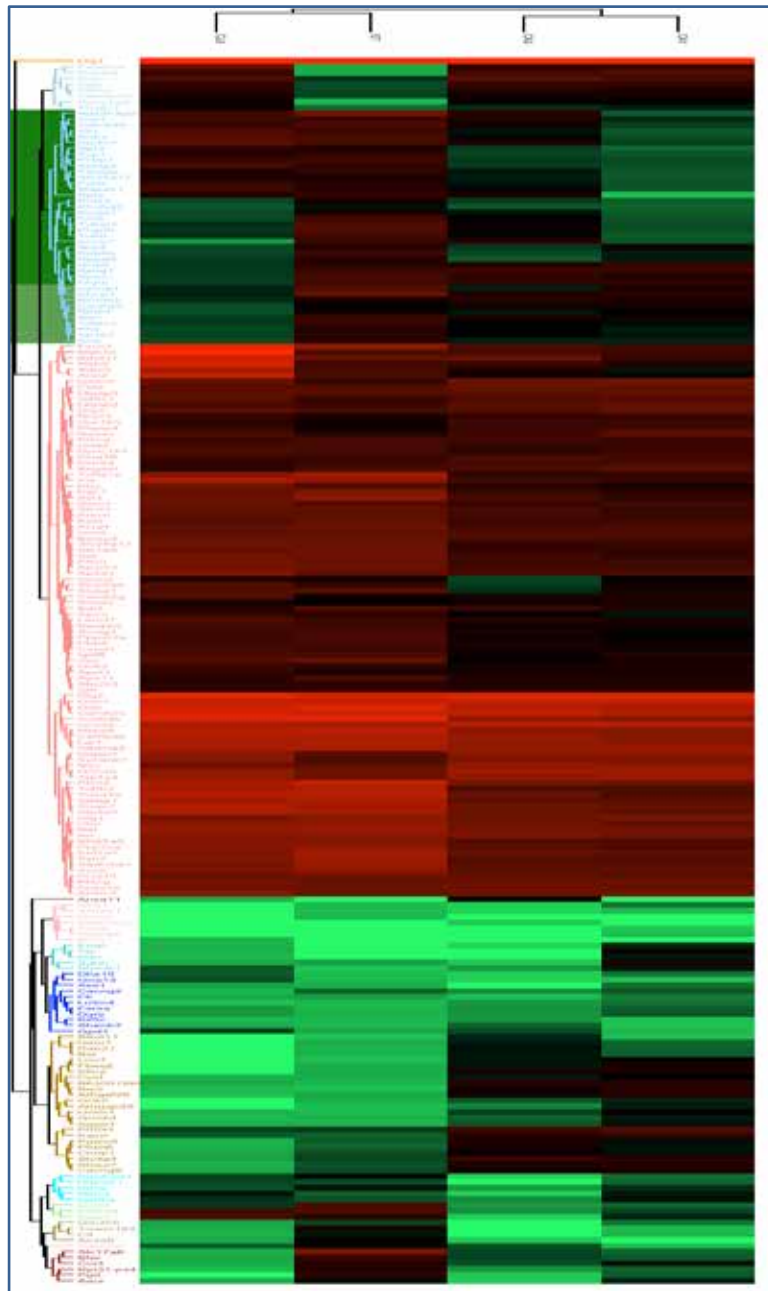


Figure 1: Heat map showing list of genes and protein levels detected in complex with PSD95 in the CA1, CA3, CL6, and ST conditional lines

CA1: CaMKII T29-Cre (T29) expresses Cre recombinase in the CA1 pyramidal cells of the hippocampus. CA3: Grik4-Cre expresses Cre recombinase in the CA3 pyramidal cells of the hippocampus: Adora2a-Cre expresses Cre recombinase in striatopallidal neurons within the striatum: Ntsr1-Cre expresses Cre recombinase in layer 4 or cortex.

3.2.2 Provenance

Data generated by SP1, T1.1.3, have been captured at UEDIN.

3.3 Anatomical Reconstructions of Hippocampal Cells

3.3.1 Data Description

3.3.1.1 Task / group responsible for generating data

T1.2.4/IEM HAS

3.3.1.2 Data, tools and methodologies storage location

Morphological reconstructions of neurons from all main areas of the mouse hippocampus were obtained by Prof. Tamás FREUND's laboratory at IEM HAS. The dataset also includes somatic whole-cell recordings using a standardised current injection protocol from the majority of the filled neurons. The data are currently stored at IEM HAS, and will be available through the HBP Neuroinformatics Portal, along with associated metadata and detailed information about the tools and methods applied.

Dataset Information Card (DIC):

Morphological reconstructions of mouse hippocampal neurons

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/85/

3.3.1.3 Description of data

All cells were filled in acute hippocampal slices from 8-week-old (P56-64) C57BL/6 mice (both males and females), and include both principal cells and a variety of different types of interneuron from areas CA3, CA1, and the dentate gyrus. As of February 18, 2016, 74 cells have been fully reconstructed (including dendrites, somata, and axonal bouton cloud); another 23 cells have been visualised and scanned but not yet reconstructed, and a further 27 neurons have been recorded and filled. The expected number of reconstructions by the end of M30 is close to the original target of 110.

This dataset collected in the Ramp-Up Phase should provide a good starting point for realistic hippocampal modelling in SGA1 (Task 6.2.4). An initial validation of the usability of the morphologies for model construction was carried out in the ramp-up phase (Task 6.4.4); more extensive modelling using these morphologies in the ramp-up phase was prevented by the late availability of an appropriate post-processing algorithm for the correct determination of dendritic diameters. At the same time, as this initial set of neurons was collected by random sampling from hippocampal slices, the dataset may not contain a sufficient sample from all the cell types which are important for modelling, and the reconstructions of cells with extended axonal arbours will be necessarily incomplete. Therefore, the planned continuation of this task in SGA1 (Task 1.2.5) will, on one hand, employ genetically labelled cell lines to specifically target critical neuronal populations, and, on the other hand, complement *in vitro* experiments with *in vivo* recordings to obtain more complete reconstructions.

As far as we know, this dataset is quite unique in that the number of available morphological reconstructions from the adult mouse hippocampus is quite small in existing databases (such as neuromorpho.org), and high-quality reconstructions of mouse hippocampal interneurons are especially sparse. In addition, very few morphological reconstructions also have associated physiological data from a standardised protocol, which makes it possible to systematically study structure-function relations in single neurons. Similar datasets are currently available from only a very limited set of brain regions (such as the somatosensory cortex of the rat).

3.3.1.4 Data quality and value

The quality of both morphological and physiological data was checked systematically by experts in the laboratory. Further analysis and validation will take place in the context of

model construction. One critical issue was the exact measurement of dendritic diameter, especially for thin higher-order processes whose size may be below the resolution of conventional light microscopy. Deconvolution-based techniques have been applied to overcome this issue, and electron microscopy can be used as a calibration and/or validation tool.

3.4 Anatomical Reconstructions of Neocortical Cells

3.4.1.1 Task / group responsible for generating data

T1.2.4/WMC

3.4.1.2 Data, tools and methodologies storage location

In addition to the task work using mouse brain tissue as planned for the first 30-month stage of HBP, additional reconstruction work has been done using rat and human tissues, in collaboration with SP5 and SP6. The reconstruction work of glia cells has not been performed because the technique of labelling glia cells is not yet developed well enough for the reconstruction of glia cells.

- 3D reconstructions, performed by WMC, of single neurons at a whole-brain-wide (n = 46): The dataset is hosted by HUST in Wuhan of China, waiting for an agreement between HUST and HBP to deliver the data to HBP. This is a collaboration between SP1, SP5 and SP6.
- 3D reconstructions of cortical neurons from the image stack of Golgi stained cortex of whole rat brain: BBP team.

Dataset Information Card (DIC):

3D reconstructions of cortical neurons at a whole-brain-wide,

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/169

- 3D reconstructions, performed by WMC, of cortical neurons from cortical tissues of human brains obtained from brain surgery in hospitals: Data are hosted by Dr. Yousheng Shu's Lab in Beijing of China. An agreement will be required for delivering the data to HBP. This is a collaboration between SP1, SP5 and SP6.

Dataset Information Card (DIC):

3D reconstructions of cortical neurons from cortical tissues of human brain obtained from brain surgery in hospitals

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/173/

- 3D reconstructions of cortical neurons from brain slices of wistar rats: BBP team.

Dataset Information Card (DIC):

3D reconstructions of cortical neurons from brain slices

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/172/

- 3D reconstructions of cortical neurons in primary visual and motor cortices from brain slices of mice: WMC

Dataset Information Card (DIC):3D reconstructions of cortical neurons in primary visual and motor cortices from brain slices

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/174/

3.4.1.3 Description of data

All reconstruction work was done using Neurolucida system. Using a Neurolucida offline system (NL360), single neurons in the somatosensory cortex were reconstructed at whole-brain-wide from a whole-brain image stack ($0.3 \times 0.3 \times 1 \mu\text{m}^3$ voxel size) made with an optical imaging method based on GFP fluorescence manipulation of the thy1 transgenic mouse (P60). Except for whole-brain-wide reconstruction of GFP fluorescence labelled and Golgi labelled neurons, the reconstruction of neurons were performed using brain slices.

- 3D Reconstructions carried out in T1.2.4 (neocortex):
 - 3D reconstructions of single neurons at a whole-brain-wide (n = 46): Using a high resolution fluorescent image stack of a GFP mouse brain (thy1-EGFP mouse brain, P60, both male and female), 36 cortical neurons and 10 striatum cells have been reconstructed into 3D model neurons at a whole-brain-wide. Cortical locations of somata include somatosensory cortex, motor cortex and striatum.
 - 3D reconstructions of cortical neurons from the image stack of Golgi stained cortex of whole rat brain (Sprague-Dawley (SD); one animal, male, age: 10-weeks): 41 of the total 46 reconstructed neurons have been sent to BBP team.
 - 3D reconstructions of cortical neurons from cortical tissues of human brains obtained from brain surgery in hospitals (n = 75): 3D reconstructions of cortical neurons from cortical tissues of human at a large range of ages from 8 to 80 years old. The brain tissues of patients were collected while a brain surgery was going in hospitals in China. Most neurons were from temporal cortical regions. The diseases of the patients include epilepsy and tumor, etc.
 - 3D reconstructions of cortical neurons from brain slices of Wistar rats (P14-P18; n = 75, both male and female): The neurons were initially recorded and filled in Dr. Henry MARKRAM's Lab. After histochemical staining, filled neurons were reconstructed into digital model neurons with a Neurolucida system. All 75 neurons of rat SSC have been sent to BBP team.
 - 3D reconstructions of cortical neurons in primary visual and motor cortices from brain slices of mice (n = 91, both male and female): 56 neurons from the primary visual cortex of C57BL/6 mice (V1B brain region, P14 - P21); 35 neurons from the motor cortex of ICR mice (P26 - P36). The reconstructed cells are all pyramidal cells.

All reconstruction work was done using Neurolucida system. Using a Neurolucida offline system (NL360), single neurons in the somatosensory cortex were reconstructed at whole-brain-wide from a whole-brain image stack ($0.3 \times 0.3 \times 1 \mu\text{m}^3$ voxel size) made with an optical imaging method based on GFP fluorescence manipulation of the thy1 transgenic mouse (P60). Except for whole-brain-wide reconstruction of GFP fluorescence labelled and Golgi labelled neurons, the reconstruction of neurons were performed using brain slices.

Four NL360 offline reconstruction systems with powerful computer work stations have been setup and are functioning, which enable 3D reconstruction of single neurons using high resolution image stacks of mouse and rat brains that are labelled with various labelling techniques.

Since the labelling technique of glia cells is still under investigation by MARKRAM's laboratory, the reconstruction of glia cells will be performed in the next stage of the HBP. The involvement of this activity in the SGA1 should be determined (either in any task of the core project led by any SGA1 partner or as a task in the context of a partnering project or as a task carried out by an external collaboration). The participation of China will follow the H2020 procedures in the SGA1 and thereafter.

3.4.1.4 Completeness of data

Dataset generated in T1.2.4 (neocortex) could be considered as the final dataset.

3.4.1.5 Data quality and value

The quality of the 3D morphological reconstructions was checked systematically by experts in the laboratory.

3.4.1.6 Data usage to date

The datasets generated have been shared with SP5 and SP6.

3.4.1.7 Publications connected to the gathered data

- Zhang, Y.F., Liu, L.X., Cao, H.T., Ou, L., Qu, J., Wang, Y., and Chen, J.G. *Otx1 promotes basal dendritic growth and regulates intrinsic electrophysiological and synaptic properties of layer V pyramidal neurons in mouse motor cortex*. Neuroscience, 2015. **285**:139-54. doi: 10.1016/j.neuroscience.2014.11.019. Epub 2014 Nov 20.
- Gao, Y., Liu, L., Li, Q., and Wang, Y. *Differential alterations in the morphology and electrophysiology of layer II pyramidal cells in the primary visual cortex of a mouse model prenatally exposed to LPS*. Neuroscience Letters, 2015. **591**:138-43. doi: 10.1016/j.neulet.2015.02.043. Epub 2015 Feb 19.

3.4.2 Provenance

3.4.2.1 Hippocampal Cells

All of the data were collected in Task 1.2.4 of the Ramp-Up Phase in the laboratory of Tamás FREUND at IEM HAS.

3.4.2.2 Neocortical Cells

Data generated have been captured on site at MARKRAM's (the neuron filling and staining of rat brain slices; this could be considered as an 'in kind' contribution) and WANG's laboratories.

3.5 Single-cell and Bulk Connectomics of Thalamocortical and Corticocortical Projection Neurons

3.5.1 Data Description

3.5.1.1 Task / group responsible for generating data

T1.2.5: The task group responsible for generating data is the Laboratory of Neural Circuits, Dept. of Anatomy and Neuroscience, Autonoma University (UAM), Madrid.

3.5.1.2 Data, tools and methodologies storage location

The data generated are .DAT file 3D reconstructions of the complete dendro-somatic and axonal domains of individual thalamocortical projection neurons. Importantly, each reconstruction is precisely registered to tissue contours of thalamic nuclei, striatal regions and cortical layers. Data also include both Neurolucida and stereological estimations of total axonal and dendrite lengths as well as axonal varicosity number, the later as proxy for terminal synapse distribution.

We use state-of-the-art single-cell transfection techniques with RNA or DNA constructs containing viral promoters able to drive in adult neurons massive levels of expression of fusion proteins such as GFP tagged for transport to the axonal membrane or synaptic domains. These techniques are applied *in vivo* by means of stereotactically-guided brain microinjection-electroporation. Labelling in isolated projection neurons (usually one cell per brain) are subsequently made stable, visible under transmitted light up to their more distant axon terminal branches using metal-enhanced immunohistochemical amplification methods. Detailed cytoarchitecture for precise atlas registration is obtained by applying light Thionin counterstain to the sections. Cells are finally and digitally reconstructed from 20-100 serial histological sections using a Neurolucida system through 20X-40X microscope objectives, (MFB Bioscience Inc., Vermont USA). Quantitative measurements are directly obtained from the Neurolucida files. Besides, bouton number in terminal axonal arborisations is estimated using the optical disector stereological method on a NewCAST system using Visiopharm software (Horshølm, Denmark).

In addition to the above single-cell studies, our task objectives also included a bulk-micropopulation mapping of thalamocortical connections in mice. We have applied these techniques to the study of multi-area projections of the posterior and lateral-posterior (pulvinar) nuclei.

Data are currently stored in our laboratory. Selected sample .DAT files have been already shared with SP5 groups to test the usability in the Neuroinformatics platform.

Dataset Information Card (DIC):

Thalamocortical cells full 3D reconstructions

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/94/

3.5.1.3 Description of data

The neurons analysed were labelled in adult (60-70 day old) male C57BL/6 mice (one cell per animal) with no apparent external lesion and full complement of vibrissae that are bred in the University Animal Facilities. A total of 103 mice were used. Animals are given water and food *ad libitum* in a 12/12hr light regime, housed at low density in large cages, and screened before operation for any visible skin damage or eye lesions.

Our datasets are morphological characterisations of thalamic projection neurons in first-order and higher-order nuclei of the visual and somatosensory thalamus.

As the labelled cells are long-range projection thalamus neurons, the dataset pertain to the cellular, --and even subcellular-- scales (dendrosomatic domain, axon varicosities) as well as the brain region/full brain scales (connectivity among nuclei/areas, axonal path and branching pattern).

The dataset is morphological/structural, because at present there are no available methods for consistently achieving both complete axonal labelling and electrophysiological recording of thalamic projection neurons with multi-branched axons. However, given the known homogeneity of neuron populations within selected thalamic domains in rodents, our structural data may be readily correlated with single cell recordings of neurons in the same nuclei.

3.5.1.4 Completeness of data

Although not specifically stated as one or “data production” goals, it must be pointed out a large part of the effort by our group during the Ramp-Up Phase has been invested to testing and developing methods and equipment for efficient and reliable labelling and 3D reconstruction of individual long-range projection neurons across large tissue volumes, potentially the entire brain. Developing these methodologies and workflows is crucial for the future phases of the project.

In this regard, we have succeeded in establishing two new protocols for single-cell labelling with high spatial precision based on the electroporation of RNA Sidbis or DNA AAV vectors. A paper reporting the RNA protocol is currently under review in *Frontiers in Neuroanatomy*.

In addition, we have labelled a total of 20 (out of 30 planned for M30) cells, of which 11 were found to be of sufficient quality and appropriate location for full 3D reconstructions.

During the M1-M30 period a paper has been published (Nakamura et al., 2015) by the laboratory of Takeshi KANEKO (Kyoto University) reporting the single cell architecture of thalamocortical projection cells of the lateral posterior complex of rats also using the Sindbis vector. We collaborate with Prof. KANEKO's group, and originally obtained from them the vector. Although, because of the species and size difference, the quantitative data from Sprague-Dawley rats cannot be used for the HBP atlas and modelling, the study findings validate our approach and provide a framework for reference and interspecies comparison.

3.5.1.5 Data quality and value

The neurons that we have 3D digitally reconstructed were selected among our best-labelled specimens. Prior to Neurolucida reconstruction the neurons were 2D drawn with at 200-400X under a microscope equipped camera lucida to detect any discontinuity because of tissue damage or incomplete labelling.

The measurements of axonal length produced by Neurolucida must incorporate a correction factor due to volume shrinkage that we have estimated at 10%.

It should be noted that axonal varicosities included in the DAT. provide a global 3D distribution of the larger ones. The total number is considerably underestimated because the smaller varicosities were beyond the resolution of the magnification used (200-400X) used for Neurolucida reconstructions. As a correction factor, we provide as metadata the varicosity number estimates obtained with the optical fractionator stereology method.

3.5.1.6 Data usage to date

To date, our pilot data (DIC: Thalamocortical cells full 3D reconstructions, see the link above) have been shared with WP5.1 (Tools for Brain Atlases) and WP5.5 (The Mouse Brain Atlas, Jaan BJALIE).

In addition, thalamocortical neurons labelled with the sindbis vector method produced in our laboratory were used during the Ramp-Up Phase to explore the suitability of visualising them with clearing + large volume 3D imaging in collaboration with LENS (Francesco PAVONE and Ludovico SILVESTRI, SP1) or Serial Two Photon scanning (Tissue Vision Inc., Cambridge, MA, USA) as well as the combination of the technique with serial reconstruction and measurement of labelled boutons at the electron microscope level (Joachim LÜBKE Forschungszentrum Jülich, Angel MERCHAN UPM, SP1, and Laszlo ACSÁDY, Hungarian Academy of Sciences Budapest).

Our data are final; however, additional metadata (virtual slice images of actual tissues sections containing the labelled axon arbour and somatodendritic domain will be incorporated once we have the high-resolution camera requested as part of SGA1.

3.5.1.7 Publications connected to the gathered data

- Porrero, C., Rodríguez-Moreno, J., Quetglas, J.I., Smerdou, C., Furuta, T., and Clascá, F. *A Simple and Efficient In Vivo non-viral RNA Transfection Method For Labeling the Whole Axonal Tree of Individual Adult Long-Range Projection Neurons*. (Frontiers in Neuroanatomy, 2016. <http://dx.doi.org/10.3389/fnana.2016.00027>).

(Description of the method used to generate data).

- Clascá, F., Porrero, C., Galazo, M.J., Rubio-Garrido, P., and Evangelio, M. *Anatomy and Development of Multispecific Thalamocortical Axons: Implications for Cortical Dynamics and Evolution*. In: Axons and Brain Architecture (Rockland KS., Ed.) London: Academic Press, 2016 pp. 69-92 ISBN 978-0128013939.

(Novel systematisation of thalamus neuronal morphotypes that will be used for description and ontology in the metadata).

- Aransay, A., Rodríguez-López, C., García-Amado, M., Clascá, F., and Prensa, L. (2015) *Long-range projection neurons of the mouse ventral tegmental area: a single-cell axon tracing analysis*. Frontiers in Neuroanatomy, 2015. 9:59. doi: 10.3389/fnana.2015.00059. eCollection 2015. PMID: 26042000.

(Description of the application of stereological methods to quantify the morphology of axonal arbourisations from individually labelled long-range projection neurons)

3.5.2 Provenance

Data generated have been captured on site at CLASCA's laboratory

3.6 Distribution of Major Synapse Proteins

3.6.1 Data Description: UEDIN

3.6.1.1 Task / group responsible for generating data

T1.1.3. UEDIN (Seth Grant)

3.6.1.2 Data, tools and methodologies storage location(s) (and links?)

Distribution of major proteins has been determined using two principle methodologies:

- 1) Mass spectrometry (MS) of synaptosome preparations. Protein abundances were determined based on peptide intensity measured and quantitated using Progenesis software (Linear Dynamics). All statistical analyses and visualisation scripts were previously published (Benjamini-Hochberg) and written in an R environment. Quantitative data, including MS run file name for all samples as well as the progenesis output for each sample, and brain region are contained within one single Microsoft Excel file uploaded to the HBP collaboration portal.
- 2) Single synapse synaptic proteome mapping (synaptome mapping) using a bespoke image analysis system. In order to map and define the distribution of individual synapses across the mouse brain, we combined state-of-the-art methods: i) synaptic proteins were fluorescently tagged using genetic labelling; ii) high resolution / high throughput imaging was performed using a confocal spinning disk microscope (SDM); iii) detection of individual synaptic puncta was carried out using an in-house developed method called Ensemble; and iv) unsupervised clustering was used to define different synapse types in order to analyse their spatial distribution in the mouse brain. A catalogue of synapse types at single-synapse scale is built using cutting-edge deep learning techniques in a purely unsupervised manner to explore of synapse proteomic diversity. Coarse-to-fine mapping of synapse types is done by voxel-wise max pooling strategy, and gives rise to re-delineation of new mouse brain synaptic zones.

Dataset Information Card (DIC):

Mouse Synaptic Proteome

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/9/

3.6.1.3 Description of data

For the preparation of the synaptic proteomes, seven integral regions of the (n=6, 8 weeks old) male adult C57bl6 mouse brain were dissected and synaptic proteins were isolated from: hippocampus, hypothalamus, striatum, cerebellum, frontal cortex, medial cortex, and caudal cortex. Synaptosomes were isolated from all seven regions. Synaptic proteins were identified and quantified by LC-MS/MS using a label-free quantitative strategy. MS analysis, quantitation and trends in protein abundances were calculated and statistically validated. A total of 1,900 synaptic proteins identified and 1,173 synaptic proteins quantified across all brain regions. Differential abundances of synaptic proteins between brain regions has identified and quantified the molecular heterogeneity of the synaptic proteome across the seven integral regions of the mouse brain (Figure 2).

Spatial proteome abundance maps have been generated from genetically labelled mice (in C57 background) carrying labelled PSD95/Dlg4 and SAP102/Dlg3. The spatial distribution of synapse proteome components at the single synapse level is being mapped using a high-throughput and automated imaging analysis pipeline. This characterises and maps diversities of individual synapses from across the mouse brain from the synapse level to the anatomical structural level, cell and synapse types (Figure 3).

3.6.1.4 Completeness of data

The brain region synapse proteome data collection is completed for the seven regions and a manuscript is in preparation. The synaptome mapping has produced the first whole-brain scale synaptome maps in the mouse and these maps are completed. A manuscript is in preparation.

There have been no mouse synapse proteome maps produced in the community to our knowledge.

3.6.1.5 Data quality and value

The synaptosome data has been analysed two different ways, one-way ANOVA and protein abundance trends. The false discovery rate (FDR) has been performed on all trends and differential abundances as described by Benjamini-HOCHBERG. This dataset is a valuable resource for scientists and modellers wanting to assess the heterogeneity of the synaptic proteome across brain regions. This dataset contains both pre- and post-synaptic proteins in different regions and therefore allows for modelling of trans-synaptic signaling to be performed on the seven Integral regions of the adult brain. Future value will be added once synaptosomes from disease models are analysed and compared to this dataset.

The synaptome mapping is a unique high quality resources with comprehensive data on ~0.5 billion synapses in >200 brain regions across the entire brain. The quality of the data has been validated with multiple forms of microscopy, automated and manual analysis. The maps are finally verified by correlating with ABI gene expression atlas, mesoscale Connectome atlas, and established synaptic proteome maps

3.6.1.6 Data usage to date

To date the data generated by T1.1.3, (Grant, UEDIN) has been used in for the following:

During the Ramp-Up Phase, visualisation and bioinformatics analysis has been performed to modelled disease-associated perturbations completed by GRANT (T1.1.3) and ARMSTRONG (WP1.3), SP1. The data is currently uploaded to the HBP Collaboratory for all HBP members to access and use in their modelling.

DIC: Mouse Synaptic Proteome (see the link above).

A manuscript is currently in preparation based on this dataset and modelling Parkinson's disease-associated proteomic changes across brain regions (Roy, M., Sorokina, O., Le Bihan, T., Tapia Gonzalez, S., de Felipe, J., Armstrong, D., Grant, S. (2016). Differential Quantitative Proteomics Reveals Brain Region-specific Abundances and Pathway Enrichment at Synapses.)

Data available at:

<https://collab.humanbrainproject.eu/#/collab/452/nav/4029?mode=run>

<https://collab.humanbrainproject.eu/#/collab/452/nav/4025?mode=run>

<https://collab.humanbrainproject.eu/#/collab/452/nav/4026?mode=run>

<https://collab.humanbrainproject.eu/#/collab/452/nav/4026?mode=run>

DIC: Mouse Synaptic Proteome (see the link above)

For SGA1 this data has been used to develop cross-cutting plans for moving from biological data on synapses to theoretical and computational models. As part of this, a meeting was held in Stockholm with Jeanette HELLGREN KOTALESKI, Sten GRILLNER and Anders LANSNER. Through this meeting a project plan was drawn up to detailing how a combination of neuron and network models could be used to investigate the effects of distinct synaptic functions associated within and across neurons. To evaluate whether neuromorphic computing platforms could be utilised in such models meetings were held

with the SpiNNaker team. SP1 members (WP1.1 and WP1.3) collaborated with SP4/SP9 members at the EITN to advocate that project needs were considered in the next iteration of Neuromorphic Platform development. Following discussions at the EITN meeting, we began developing theory based project proposals with Wolfgang MAASS, aimed at bringing issues of interest to molecular neurobiologists into the remit of theoreticians. GRANT (T1.1.3) and ARMSTRONG (WP1.3) developed a proposal for how SP1 could make the biological details for synaptic function more tractable for theoretical analysis during SGA1. The collaborative ideas detailed above will be developed further as part of CDP5 (see section Data Aggregation, Analysis and Disseminations for further information of WP1.3).

3.6.2 Data Description: UCLM & IST

3.6.2.1 Task / group responsible for generating data:

Task T1.1.4 (Neural Channelomics and Receptomics). PIs: Ryuichi Shigemoto (IST) & Rafael Luján (UCLM).

3.6.2.2 Data, tools and methodologies storage location(s) (and links?)

We are mapping the 2D and 3D distribution of neurotransmitter receptors (AMPA, NMDA and GABAB receptors) and ion channels (SK, GIRK and P/Q-type Ca²⁺ channels) in the hippocampus using newly developed immunoelectron microscopy techniques (the SDS-FRL and immunogold FIB/SEM techniques).

Our data storage is in our laboratories and found at Dataset Identification Card (DIC).

Dataset Identification Card (DIC):

Neural channelomics and receptomics

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/111/

3.6.2.3 Description of data

We provide an average of density of immunoparticles/ μm^2 (2D) and immunoparticles/ μm^3 (3D), size of clusters, composition of clusters, and spatial relation of clusters for different gold particles, obtained from more than 100 neurons of 28-day-old male C57BL/6 mice. We are studying those parameters, which represent the precise subcellular localisation of receptors and ion channels in the different compartments of pyramidal cells of the hippocampus.

3.6.2.4 Completeness of data

The dataset anticipated at the beginning of the project falls a bit short with the dataset delivered by M30 in regards to the use of the immunogold FIB/SEM technique. We could only generate around 80% of the data anticipated using such technique due to the lack of specific software that allows generation of fast 3D reconstruction of long series of serial sections and particularly the quantitative analysis in three dimensions. Regarding the SDS-FRL technique the anticipated dataset fits exactly with the dataset delivered by M30. However, we also generated 2D mapping of GABAB receptors, GIRK, SK and Cav2.1 in Purkinje cells of the cerebellum, not included in the original grant.

3.6.2.5 Data quality and value

Our data provides the spatial resolution of receptors and ion channels at the nanometre scale, with a level of detail never previously attained by the research community. Only a few laboratories can now use the SDS-FRL technique, but to our knowledge nobody is still using the immunogold FIB/SEM technique. Regarding the data generated by the community using the SDS-FRL technique over the past 30 months, the laboratory of Zoltan NUSSER provided the 2D mapping of Kv1.1 and Kv2.1 in pyramidal cells of the hippocampus (Kirizs

et al., 2014). In addition, the laboratory of Francesco FERRAGUTI, in collaboration with one of us (Ryuichi SHIGEMOTO), provided the precise subcellular localisation of the group I metabotropic glutamate receptor subtype mGluR1 in the cerebellar cortex (Mansouri et al., 2015), and the laboratory of Akos KULIK described the common nano-architecture of Cav2.1 channels in inhibitory and excitatory axon terminals in the hippocampus (Althof et al., 2015), but did not provide the detailed and systematic mapping information we are providing. Nevertheless, these articles and other articles published by both of the PIs demonstrate the high sensitivity of the SDS-FRL technique for the subcellular localisation of receptors and ion channels, and all of them provide valuable information for modelling.

3.6.2.6 Data usage to date

Our data have not been used so far by other SPs. We should discuss with modellers, Dan KELLER and Eilif MULLER, how our data in 3D reconstructed images or Excel format can be used for numerical simulations. Are the data considered final?

The data regarding the density of immunoparticles/ μm^2 (2D) and immunoparticles/ μm^3 (3D) of individual receptors and ion channels in a given neuron type, as well as the co-clustering and spatial relationship of two different proteins in a given subcellular compartment can be considered as final.

3.6.2.7 Publications connected to the gathered data

- Javdani, F., Holló, K., Hegedűs, K., Kis, G., Hegyi, Z., Dócs, K., Kasugai, Y., Fukazawa, Y., Shigemoto, R., and Antal, M. *Differential expression patterns of K(+) /Cl(-) cotransporter 2 in neurons within the superficial spinal dorsal horn of rats*. Journal of Comparative Neurology, 2015. **523**:1967-83.
- Mansouri, M., Kasugai, Y., Fukazawa, Y., Bertaso, F., Raynaud, F., Perroy, J., Fagni, L., Kaufmann, W.A., Watanabe, M., Shigemoto, R., and Ferraguti, F. *Distinct subsynaptic localization of type 1 metabotropic glutamate receptors at glutamatergic and GABAergic synapses in the rodent cerebellar cortex*. European Journal of Neuroscience, 2015. **41**:157-67.
- Luján, R., and Ciruela, F. (Eds). *Receptor and Ion Channel Detection in the Brain - Methods and Protocols*. Springer Science+Business Media New York, 2016. ISBN 978-1-4939-3063-0. <http://www.springer.com/gp/book/9781493930630>.
- Luján, R. *Pre-embedding methods for the localization of receptors and ion channels*. In: Rafael Luján & Francisco Ciruela (Eds). *Receptor and Ion Channel Detection in the Brain - Methods and Protocols*. Springer Science+Business Media New York, 2015 (in press). ISBN 978-1-4939-3063-0. <http://www.springer.com/gp/book/9781493930630>.
- Luján, R., and Watanabe, M. *Post-embedding immunohistochemistry in the localization of receptors and ion channels*. In: Rafael Luján & Francisco Ciruela (Eds). *Receptor and Ion Channel Detection in the Brain - Methods and Protocols*. Springer Science+Business Media New York, 2015 (in press). ISBN 978-1-4939-3063-0. <http://www.springer.com/gp/book/9781493930630>.
- Luján, R., and Aguado, C. *Localization and targeting of GIRK channels in mammalian central neurons*. In: Paul Slesinger (Ed.) *International Review of Neurobiology - A volume on GIRK potassium channels*. Elsevier Inc. Oxford, UK, 2015. ISBN 978-1-4939-3063-0. <http://www.springer.com/gp/book/9781493930630>.
- Harada, H., and Shigemoto, R. *High-Resolution localization of membrane proteins by SDS-digested freeze-fracture replica labelling (SDS-FRL)*. In: Rafael Luján & Francisco Ciruela (Eds). *Receptor and Ion Channel Detection in the Brain - Methods and Protocols*. Springer Science+Business Media New York, 2016. ISBN 978-1-4939-3063-0. <http://www.springer.com/gp/book/9781493930630>.

- Harada, H., and Shigemoto, R. *Immunogold protein localization on grid-glued freeze-fracture replicas*. In: Steven D. Schwartzbach & Thomas Schikorski (Eds). *High Resolution Imaging of Proteins in Tissues and Cells: Light and Electron Microscopy Methods and Protocols - Methods in Molecular Biology*. Springer Science+Business Media New York, 2016 (in press).

3.6.3 Provenance

3.6.3.1 UEDIN

All data generated within SP1, T1.1.3 (GRANT, UEDIN) has been prepared for data integration using standard vocabulary and ontologies. The molecular datasets generated by MS have been prepared with detailed protocols outlining information on strains and genotypes, genome assemblies, and the specific nomenclature rules used. Much of the data acquisition and processing workflow exploit HBP Neuroinformatics tools and resources. All Excel spreadsheet data files shared with the Neuroinformatics Platform include tabs and column sub-headings under both raw abundance and normalized abundance in which the individual MS datafile is listed. Each sample analysed has a unique MS run file number generated and associated with it. Therefore, all data generated by UEDIN can be identified by its listed unique MS run file documented and stored both in UEDIN GRANT lab as well as by the KPF in Edinburgh. This data has been deposited in the Neuroinformatics Platform. Full methods are described in manuscript currently in preparation. Image and mapping data have been captured and stored at UEDIN.

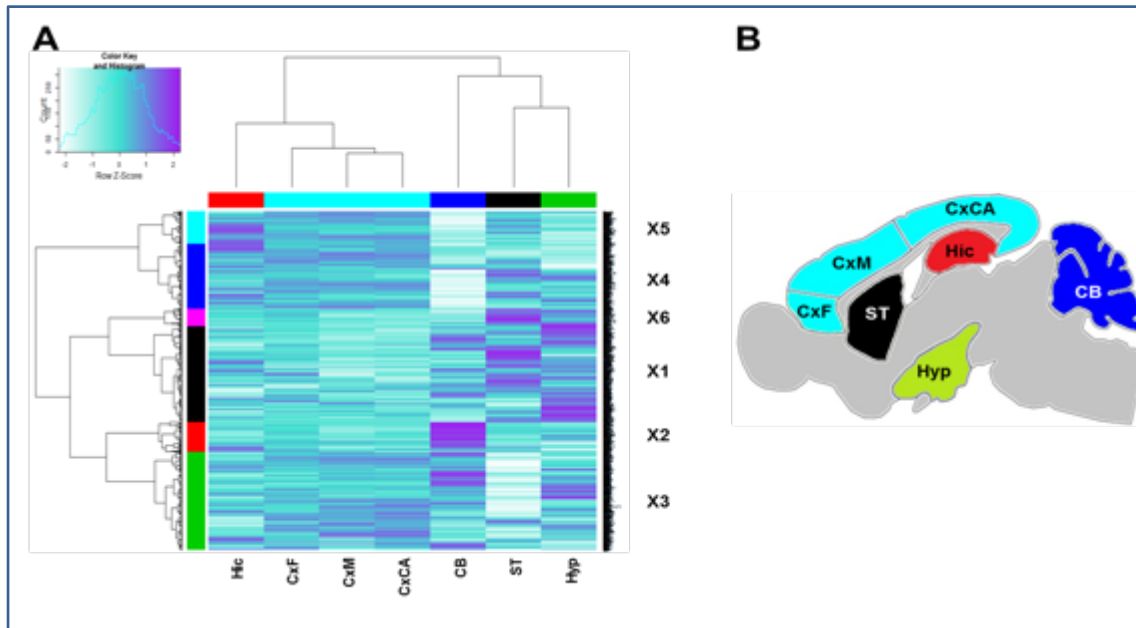


Figure 2: Synaptic proteome expression levels across the seven integral regions of the mouse brain

The mean abundance of 1,173 synaptic proteins isolated from n=6 two month old wildtype male mice identified and quantified by LC-MS/MS. (A) Hierarchical clustering both by brain region (x-axis) and by protein abundance (y-axis). The synaptic proteome was sub-clustered into six modules of proteins (turquoise, blue, magenta, black, red and green) that displayed similar trends in their levels of abundance across brain regions. (B) Diagram of seven integral mouse brain regions coloured by similarity of synaptic proteins identified and abundances.

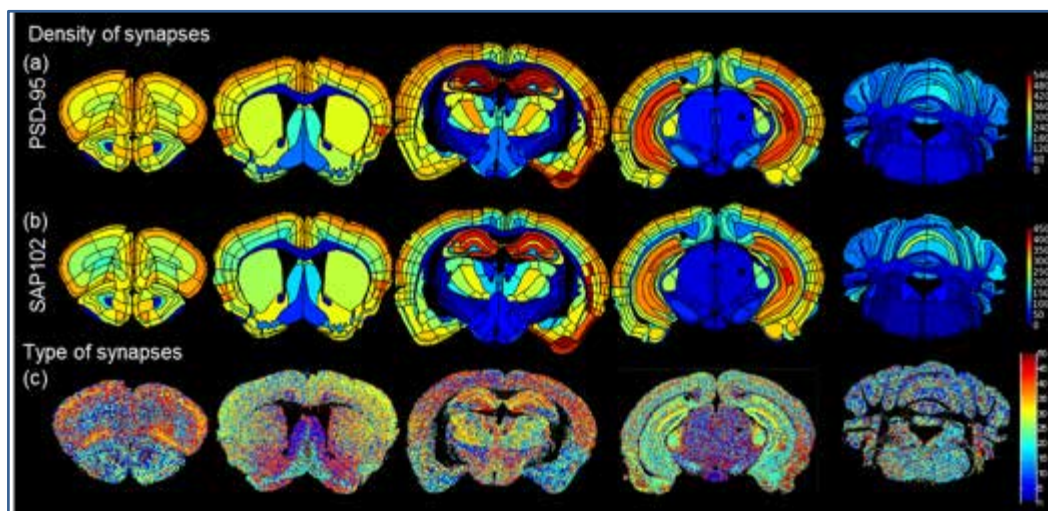


Figure 3: Heat-maps of synapse densities and synapse types coded by different colours across the 5 mouse brain sections

These reconstructed density maps (Fig. 3a,b) obtained from ~220 delineated anatomical regions show patterning of synapse densities with striking differences; regions expressing barely any PSD-95 and SAP102 synapses (hindbrain, midbrain, hypothalamus, white matter tracts) and regions with significantly increased synaptic density of both markers (hippocampus and cortical layers). We have also used methods that do not require manual delineation of regions (Fig. 3c).

3.6.3.2 UCLM & IST

Data generated within T1.1.4 have been captured at LUJAN's and SHIGEMOTO's laboratories.

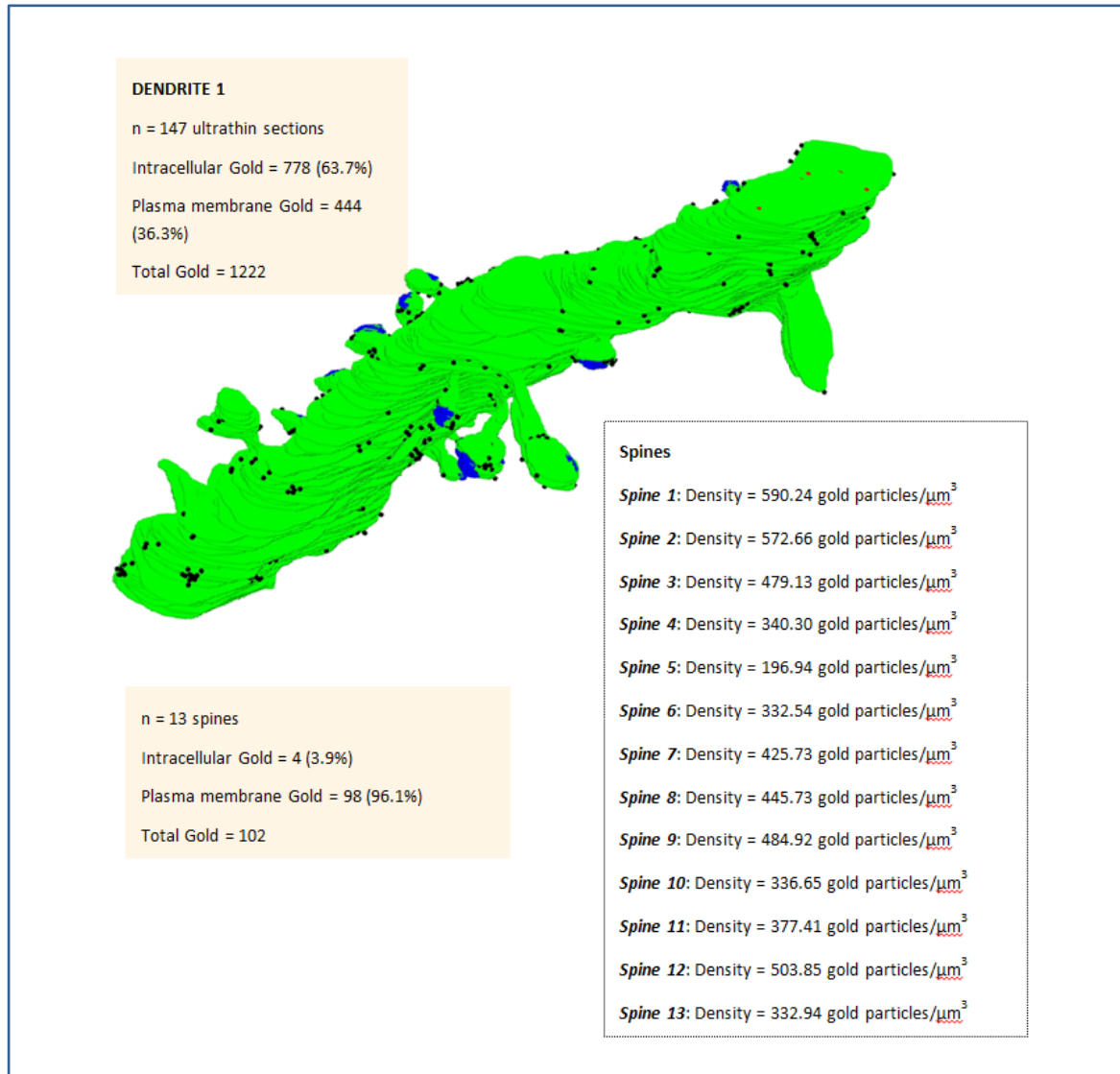


Figure 4: 3D reconstruction of dendritic shaft of CA1 pyramidal cells with immunogold labelling for SK2

Three-dimensional reconstruction of a dendrite of CA1 pyramidal cells immunolabelled for the potassium channel subunit SK2 in the *stratum radiatum*. Tissue was processed using the immunogold method using the pre-embedding immunogold method followed by the FIB/SEM technique. 3D reconstructions from 147 ultrathin sections were performed using the software Reconstruct. Black dots represent the location of gold particles for SK2 along the plasma membrane. Blue colour shows the PSD of excitatory synapses in dendritic spines. Immunoparticles were quantified along the plasma membrane and at intracellular sites. Among other things, the data demonstrate a high density and enrichment of SK2 in dendritic spines of pyramidal cells.

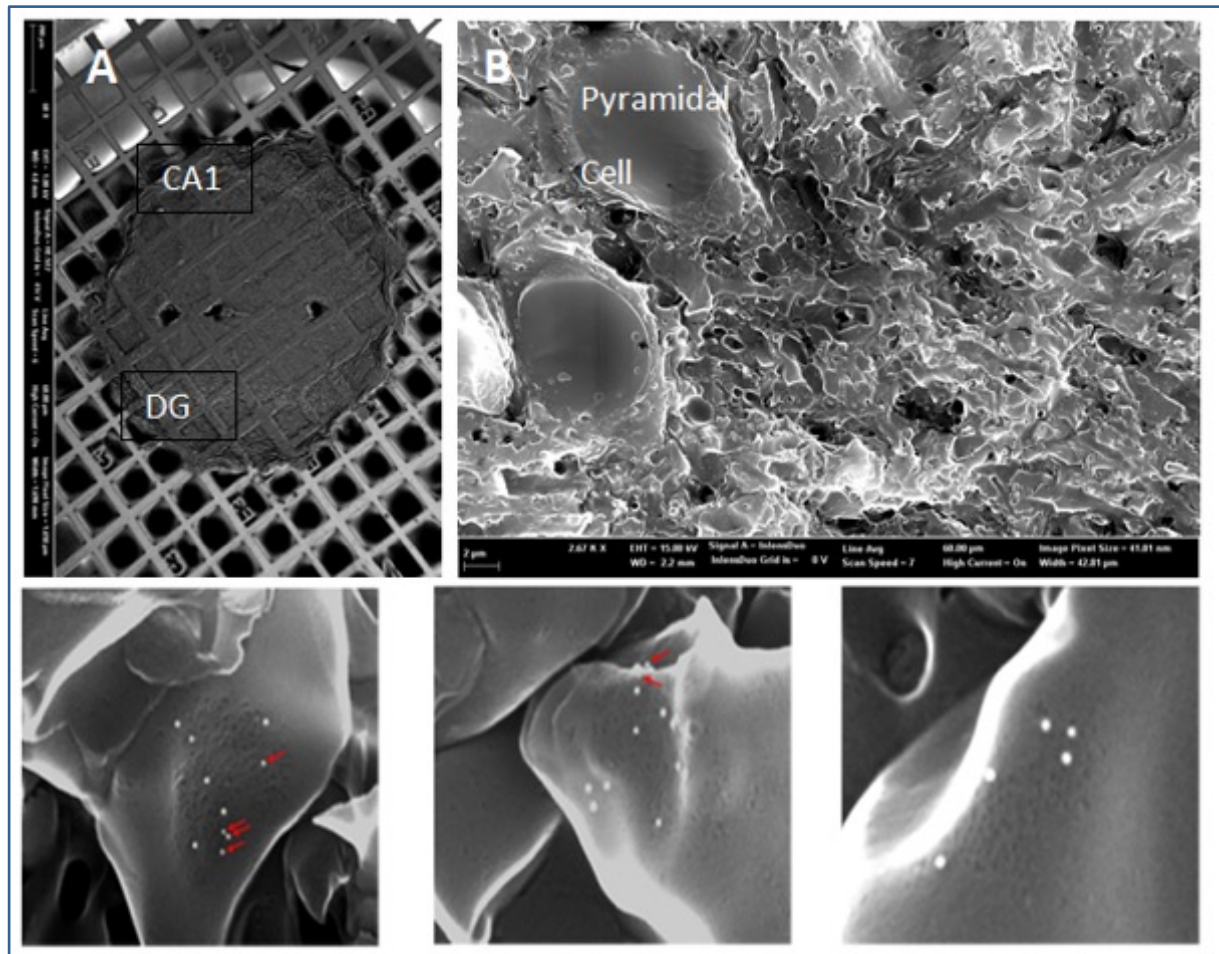


Figure 5: 2D mapping of glutamate receptors in CA1 pyramidal cells by SDS-digested freeze-fracture replica labelling

In grid-mapped replica showing whole layers of the CA1 area of the hippocampus (A), pyramidal cells and their dendrites with distance from cell bodies can be easily identified (B). Excitatory synapses can be identified as clusters of intramembrane particles and gold particles for AMPA receptors (C-E) are measured on 2D surface of synaptic and extrasynaptic membranes. Red arrows indicate 5 nm particles for NMDAR1. 10 nm particles are for AMPA receptors.

3.7 Numbers, Distributions and Relative Densities of Cells in Selected Regions

3.7.1 Data Description: LENS

3.7.1.1 Task(s)/group(s) responsible for generating data

T1.2.2, LENS

3.7.1.2 Data, tools and methodologies storage location(s) (and links?): LENS, CINECA

Dataset Information Cards (DICs):

Hippocampus reconstruction

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/16/?changelist_filters=o%3D4

Parvalbumin interneurons reconstructions

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/15/?changelist_filters=o%3D4

Purkinje cell distribution

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/14/?changelist_filters=o%3D4

Arc-dVenus

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/268/?changelist_filters=p%3D1%26o%3D3

3.7.1.3 Description of data

- Species, sex, age, number of specimen/subjects.: L7-GFP line, male, 10 pnd (post-natal days), 1 specimen; Arc-dVenus line, male, 56 pnd, 1 specimen; PV-Cre-tdTomato line, male, 56 pnd, 1 specimen.
- Scale (brain, brain region, cells, molecules), features (morphology/physiology/expression, etc.), locations, and description of entities, e.g. morphological characterisation of basket cells of the hippocampus; 3D localisation and counting of Purkinje cells of the mouse cerebellum; 3D expression of Arc-cells activated with visual stimulation in half mouse brain; 3D localisation of parvalbumin positive neurons in whole mouse brain.

3.7.1.4 Completeness of data

LENS delivered three 3D cellular map datasets as planned.

To our knowledge, currently there are no public projects aiming at producing whole-brain 3D distributions of cells.

Among the research community, the Allen brain institute worked on whole brain RNA mapping. The datasets produced by LENS are based on proteins expression, the information proposed are complementary to those produced by the Allen brain institute.

3.7.1.5 Data quality and value

A ground truth was performed in a small area of the dataset; a manual analysis was used to verify the quality of the data obtained using the counting methods algorithm.

In future, the localisation datasets obtained could be used to create models of whole mouse brain cells type distribution. They could be analysed to create models that correlate cells distribution with molecular expression. Moreover, they could be used to observe the variations of anatomical organization between healthy and ill animals.

3.7.1.6 Data usage to date

No Ramp-Up Phase Task used them.

3.7.1.7 Are the data considered final?

Yes.

3.7.1.8 Publications connected to the gathered data

- Frasconi, P., Silvestri, L., Soda, P., Cortini, P., Pavone, F.S., and Iannello, G. *Large-scale automated identification of mouse brain cells in confocal light sheet microscopy images*. Bioinformatics, 2014. **30**:587-593.
- Silvestri, L., Paciscopi, M., Soda, P., Biamonte, F., Iannello, G., Frasconi, P., and Pavone, F.S. *Quantitative neuroanatomy of all Purkinje cells with light sheet microscopy and high-throughput image analysis*. Frontiers in Neuroanatomy, 2015. **9**:68.

Both articles above describe the method used to generate and analyse the cerebellum dataset.

- Costantini, I., Ghobril, J.P., Di Giovanna, A.P., Allegra Mascara, A.L., Silvestri, L., Müllenbroich, M.C., Onofri L, Conti, V., Vanzi, F., Sacconi, F., Guerrini, R., Markram, H., Iannello, G., and Pavone, F.S. *A versatile clearing agent for multi-modal brain imaging*. Scientific Reports, 2015. **5**:9808.

This article describes the clearing method used in combination with the light sheet microscopy to image large mouse brain volumes.

3.7.2 Data Description: UPM

3.7.2.1 Task(s)/group(s) responsible for generating data

T1.2.3, UPM.

3.7.2.2 Data, tools and methodologies storage location(s) (and links?)

Data about density and distribution of different types of cells (excitatory, inhibitory neurons and glia) in the somatosensory cortex in GAD67 mice have been obtained using immunostaining, confocal microscopy and customised segmentation tools. Development of the segmentation tools and algorithms has been performed. Data are available as tiff image stacks and XLS files. Data has been aligned with requirements of Neuroinformatics Platform (SP5).

Dataset Information Card (DIC):

Cell number counting

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/33/?changelist_filters=o%3D4

3.7.2.3 Description of data

Species, sex, age, number of specimen/subjects.: *Mus musculus*, GAD67GPF strain, male, 8 weeks, n= 6 Scale (brain, brain region, cells, molecules), features (morphology/physiology/expression, etc.), locations, and description of entities, e.g. morphological characterisation of basket cells of the hippocampus.

We have analysed the somatosensory cortex at the cellular scale different populations of cells (excitatory and inhibitory neurons and glial cells) in all cortical layers.

3.7.2.4 Completeness of data

The algorithm for automatic segmentation is in a very advanced stage of development. Results have been already obtained and validated and most of the problems identified have been corrected and validated. Six datasets have been generated (including data from every cortical layer in each case), detailing the cell density and distribution of GABAergic interneurons in the somatosensory cortex. Quantitative data on the number and 3D distribution of GABAergic cells have been extracted from the segmentation.

The two-steps binarisation part of the algorithm has been completed and the algorithm can combine information of different channels (Figure 6).

3D data have been used by WP5.4 /T5.4.2 (Neuronal Structural Design and Predictions) by spatial statistical techniques to characterise the density and principle patterns of spatial distribution of GABAergic interneurons (currently ongoing). These data are essential to build the models of the cortical column.

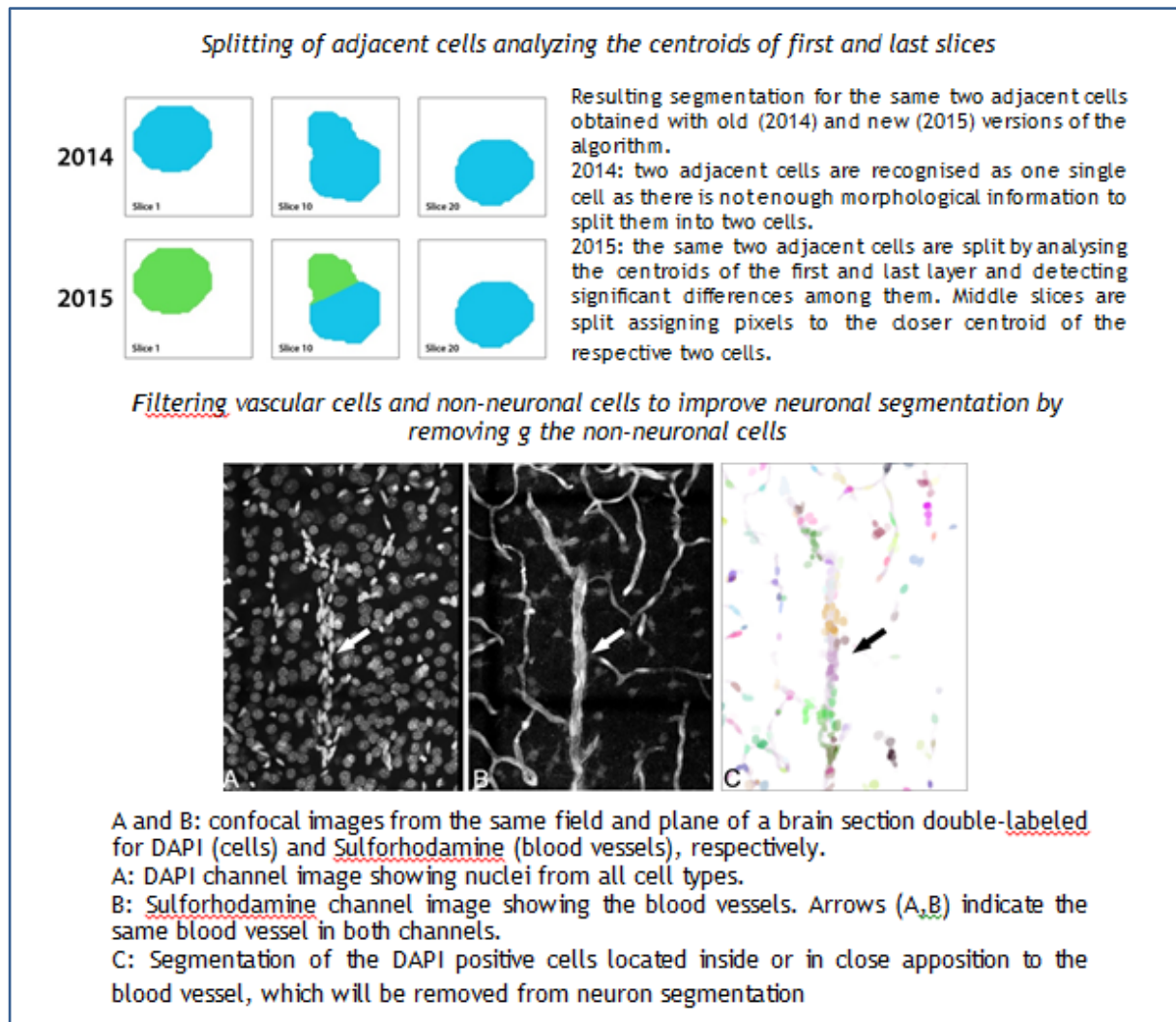


Figure 6: Fine tuning of the segmentation algorithm for cell detection

Projected data comprised numbers and distribution of excitatory and inhibitory neurons as well as glial cells, but only part of this cell subpopulation have been fully analysed

Several papers on neuron distribution and density have been published, in particular, a paper entitled “Reconstruction and simulation of neocortical microcircuitry” presenting the first digital reconstruction of a cortical minicolumn points out the necessity of 3-dimensional data of the different cells present in the cortical column to build more complete models.

3.7.2.5 Data quality and value

Data are still pending to be revised statistically to be considered as final.

Subjective analysis of these datasets indicates that numbers we obtained with our algorithm are within the range of the previously published data. Therefore, we are planning to use this technology during the SGA1.

3.7.2.6 Data usage to date

Data on number and distribution of cells are considered as strategic data and they will feed the models designed by the Neuroinformatics Platform (SP5) in task 5.1.1 and 5.1.2 defined in the SGA1.

3.7.2.7 Are the data considered final?

Data can be considered to be a robust draft. We currently completed with more cases and including more cell types.

3.7.2.8 Publications connected to the gathered data

Publications are still pending to be released to the scientific community.

- LaTorre, A., Alonso-Nanclares, L., Muelas,, S., Peña, J.M., and DeFelipe, J. 2013. *3D Segmentations of Neuronal Nuclei from confocal Microscope Images Stacks*. Frontiers in Neuroanatomy, 2013. 7:49. doi: 10.3389/fnana.2013.00049.

(In this paper, we present an algorithm to create 3D segmentations of neuronal cells from stacks of previously segmented 2D images. The algorithm not only reuses the information obtained in the 2D segmentation, but also attempts to correct some typical mistakes made by the 2D segmentation algorithms (for example, under segmentation of tightly-coupled clusters of cells). This method will be used to create an automatic tool to segment and calculate the number and spatial distribution of cells in different brain regions during the SGA1).

3.7.3 Provenance

3.7.3.1 LENS

WP1.2: T 1.2.2

3.7.3.2 UPM

The data generated have been captured on site at DEFELIPE's laboratory (Task 1.2.3).

3.8 Synaptic Maps of Identified Neurons

3.8.1 Data Description

3.8.1.1 Task(s)/group(s) responsible for generating data:

T1.2.6 (UPM)

3.8.1.2 Data, tools and methodologies storage location(s) (and links?)

Data have been delivered as tiff image stacks and XLS files. The format of these data has been aligned with the Neuroinformatics Platform (SP5) requirements. Transgenic PSD95GFP mice were used and their brains were fixed with paraformaldehyde and processed with standard histological techniques, immunocytochemistry, intracellular injections and new informatics tools developed in our laboratory in collaboration with neuroinformatic teams (see below).

Dataset Information Card (DIC):

Synapse maps

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/31/?changelist_filters=o%3D4

3.8.1.3 Description of data

We have used 14 PSD95GFP mice (males, P60) generated in the laboratory of Seth GRANT.

We have used a novel kind of analysis that combines the use of new genetic molecular tools (in collaboration with GRANT's laboratory) and the powerful neuroanatomical technique of intracellular injections in order to locate the main postsynaptic proteins within the different compartments of identified neurons. This approach enables large-scale synaptic maps of individual neurons to be generated. The results obtained will open new horizons in the field since the possibility of generating such extensive synaptic maps will open the way to analyse cortical circuitry at a level not previously feasible. Specifically, the main objective during the Ramp-Up Phase was to locate the postsynaptic protein PSD95 (which is one of the main components of the postsynaptic densities), in the different dendritic compartments of pyramidal neurons. For this purpose, cells were intracellularly injected in the hippocampus and somatosensory cortex of transgenic mice expressing the enhanced green fluorescent protein (EGFP) for PSD95 (EGFP-PSD95). This combined methodology reveals the cell architecture and the position of this synaptic protein in the different parts of the neuron, enabling large-scale molecular synaptic maps to be created.

This method involves first staining the section with the fluorescent dye DAPI (Sigma) to identify the cell bodies, and then the pyramidal cells can be individually injected with a fluorescent marker with the aid of an electrode. We have used Alexa fluor 594 for microinjection (Molecular Probes) to visualise dendrites in cytoarchitectonically identified pyramidal cell layer of CA1 of the hippocampus and layers II, III, V and VI of the somatosensory cortex. After injections, sections are analysed directly with the aid of a Zeiss LSM 710 Confocal microscope. Fluorescently labelled profiles are examined through separate channels, using excitation peaks of 585 and 491 nm to visualise Alexa fluor 594 and GFP, respectively. Acquisition of the stacks of high magnification microscopy images enables the reconstructions to be studied in 3D to quantify the distribution, size and shape of the PSDs in the injected pyramidal neurons. Figure 7 illustrates an example of an intracellularly injected cell and the subsequent processing of the images to define the location of the synaptic protein PSD95 in the dendrites of the neuron. Figure 7 represents the potential of the method to analyse the spine volumes in detail, and the PSD volumes,

and to reconstruct these structures along the dendrites of a neuron in order to create large-scale synaptic maps of identified neurons.

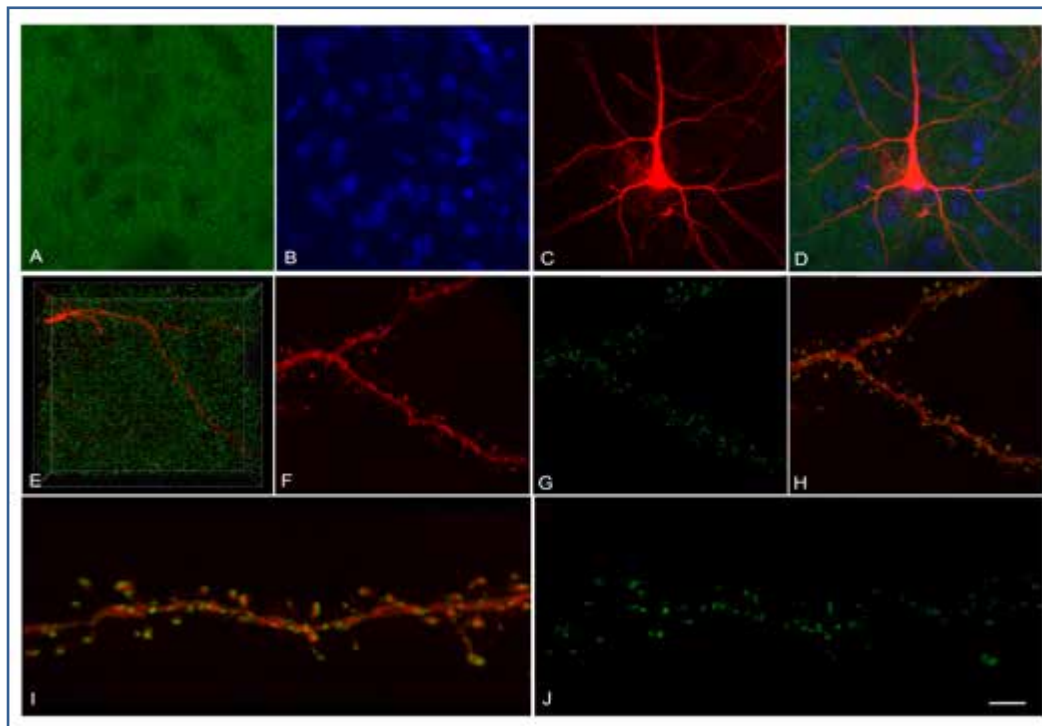


Figure 7: PSD95 labelling of a layer III pyramidal neuron in the somatosensory cortex of a PSD95 EGFP mouse.

A) Illustrates the punctuate PSD95 labelling. B) DAPI staining in the same field. C) An intracellularly injected pyramidal cell in the same microscopy field as in A and B. D) Shows the combination of the images in A-C. E-F) After injections we captured high magnification stacks of dendritic segments by confocal microscopy and we then used Imaris software (Bitplane, Switzerland) to create an artificial surface that fitted (by thresholding) the surface of the real dendrite. G) We then applied a mask to the green channel that set the voxels outside the surface to 0. Thus, we obtained a new green channel with the information contained within the dendritic surface created. H) By combining the new green channel and the real dendrite, the relative location of PSD95 (green) could be established with respect to the pyramidal cell (red). I,J) Show another example of dendritic spines (red) and PSD 95 labelling (green) at higher magnification. Scale Bar = 13 μm in AD, 8.3 in E, 5 μm in F-H and 2,5 μm in I, J.

Furthermore, remarkable differences exist in the microanatomy of pyramidal cells in different regions of the dendritic tree, and in distinct cortical layers and areas. Additionally, very little is known about the proportion of dendritic spines that contain PSD (synaptic versus non-synaptic dendritic spines), or the differences in the expression of PSD95 in these structures. Hence, we were studying the distribution of PSD molecules in shafts and spines along the injected pyramidal neurons in layers II, III, V and VI of the somatosensory cortex. Finally, the volume of dendritic spines and PSDs has been quantified in the different compartments of the neuron. Figure 8 represents the potential of the method to analyse the spine volumes in detail, and the PSD volumes, and to reconstruct these structures along the dendrites of a neuron in order to create large-scale synaptic maps of identified neurons.

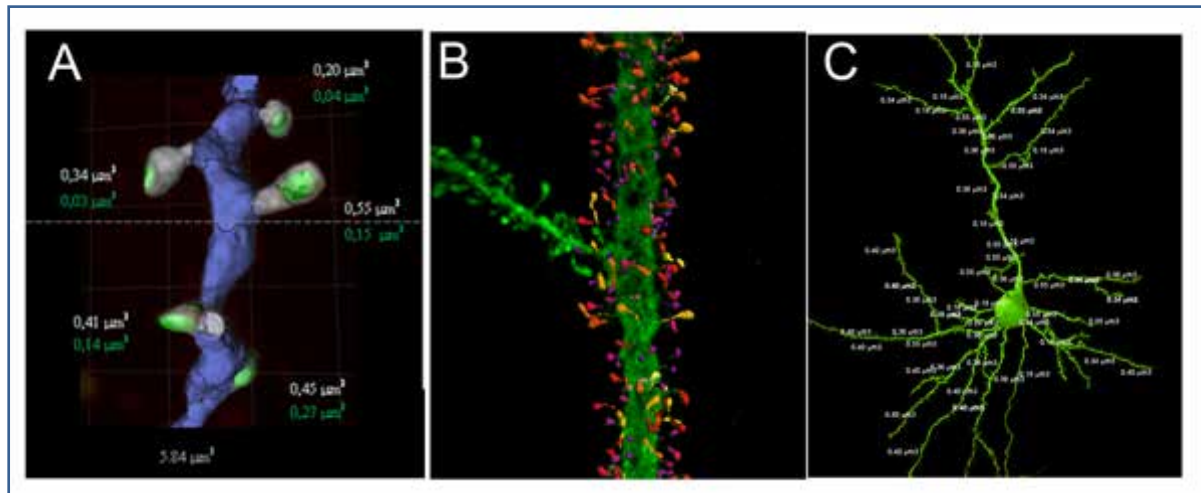


Figure 8: Synaptic maps of Identified Neurons

A) Estimation of the spine (gray) and PSD volumes (green) in a given dendritic segment where numerical values of the estimated dendritic spine and PSD volumes are also shown. **B,C)** Represent the method to reconstruct these structures along the dendrites of a neuron in order to create large-scale synaptic maps of identified neurons. In particular, **B** represents a segment of an intracellularly injected pyramidal neuron with dendritic spines of different volumes (different colours indicate the individual spine volumes). By applying this method to the PSD95 EGFP mice, it is possible to obtain a molecular synaptic map of the whole dendritic arbour of an individual neuron.

3.8.1.4 Completeness of data

We have obtained 3D confocal digital reconstructions of 60 dendrites, including main apical, colateral and basal dendrites from 20 individual identified pyramidal neurons in the hippocampus (n=10) and somatosensory cortex (n=10) together with the distribution of PSD95 puncta in the same stacks of images. We are currently generating the synaptic maps of the 60 dendrites by co-registration of both channels (Alexa and GFP).

Unfortunately, we have detected a problem with the data generated with the confocal microscope. The images that were obtained during the last 12 months show a distortion in the z-axis. Image processing to reconstruct dendritic spines is very time consuming because there are not automatic software tools to obtain an accurate 3D reconstruction of the dendritic spine. Thus, the reconstruction of these structures is mainly based on manual tracing using informatic tools. Therefore, it will take longer time than expected to generate the final maps because we have to acquire again the 3D confocal digital reconstructions of 46 dendrites and re-analyse the previous data.

In addition to the goal of the generation of synaptic maps on identified neurons, we have developed a new tool called PyramidalExplorer (Toharia et al., 2016) to interactively explore and reveal the detailed organisation of the microanatomy of pyramidal neurons with functionally related models (Figure 9). This tool consists of a set of functionalities that allow possible regional differences in the pyramidal cell architecture to be interactively discovered by combining quantitative morphological information about the structure of the cell with implemented functional models. We are planning to use this tool to analyse the synaptic maps to discover new aspects of the morpho-functional organization of the pyramidal neuron.

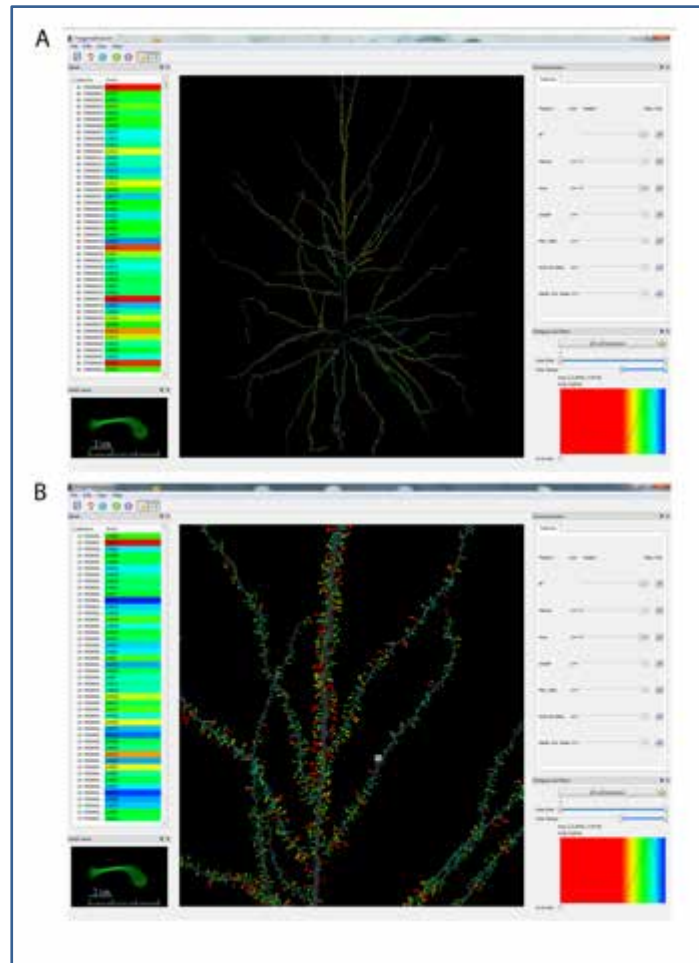


Figure 9: Pyramidal Cell Explorer

Graphical user interface (GUI) of PyramidalExplorer showing the cell comparison query result concerning Spine Area, from a pyramidal neuron in a global view (A) and in a zoom in view (B). Values are represented by colour code. Red colours represent highest values whereas blue colours represent the lowest values. White frame in the main window highlights the spine visualized in the Detail viewer widget. Taken from Toharia et al. (Front Neuroanat. 9:159, 2016)

As far as we know, this is the only dataset reporting large-scale synaptic maps of individual neurons from intact mouse brain.

The technology is now available to be exploited for comprehensive studies on the synaptic maps of pyramidal neurons. The dataset collected in the ramp-up phase is critical for the generation of realistic modelling of pyramidal cells in collaboration mostly with SP2 (T2.2.2), SP5 (Task 5.3.6) and SP6 (T6.2.1; T6.2.2) in the SGA1.

3.8.1.5 Data quality and value

The quality of the 3D morphological reconstructions was checked systematically by experts in the laboratory. In addition, further analysis and validation will be performed in the context of model construction.

These datasets were primarily designed to be integrated in computational models based on the morphological parameters found in real pyramidal cells and realistic synaptic weights.

3.8.1.6 Data usage to date

Parts of the experiments of synaptic maps we are generating are being used to validate single synaptic proteome mapping performed by T1.1.3. UEDIN (Seth GRANT) (synaptome mapping). Furthermore, the data will be used for the modelling and simulation efforts of SP4 and SP6 (SEGEV and GIUGLIANO) in the SGA1.

3.8.1.7 Are the data considered final?

These data are considered to be a robust draft of the synaptic maps of identified neurons we are generating. However, the tool PyramidalExplorer is available tool to analyse the morpho-functional organization of the pyramidal neurons.

3.8.1.8 Publications connected to the gathered data

- Toharia, P., Robles, O.D., Fernaud-Espinosa, I., Makarova, J., Galindo, S.E., Rodriguez, A., Pastor, L., Herreras, O., DeFelipe, J., and Benavides-Piccione, R. PyramidalExplorer: A new interactive tool to explore morpho-functional relations of human pyramidal neurons. *Frontiers in Neuroanatomy*, 2016. **9**:159.

(Using this tool will facilitate the study of the correlation of the pyramidal synaptic maps with the morpho-functional organization of the dendritic arbor of these neurons and to create a general model of pyramidal neuron).

3.8.2 Provenance

SP1: WP1.1, Task 1.1.3 (Seth GRANT) and WP1.2, Task 1.2.6 (Javier DEFELIPE)

3.9 High Resolution Optical Synaptic Maps and Electron Microscope Blocks Scans and Volume Analysis of Exemplar Brain Regions

3.9.1 Data Description

3.9.1.1 Task(s)/group(s) responsible for generating data

T1.2.6 (UPM)

3.9.1.2 Data, tools and methodologies storage location(s) (and links?)

The data we are generating are tiff image stacks and excel files. Data have been delivered as tiff image stacks and XLS files. Data has been aligned with the Neuroinformatics Platform (SP5) requirements. Transgenic PSD95GFP mice were used and their brains were fixed with paraformaldehyde and processed with standard histological techniques, confocal microscopy and electron microscopy (EM). The number and distribution of PSD95 puncta has been analysed using a novel 3D quantification methodology developed in our laboratory. For EM we have applied a recently developed method for three-dimensional reconstruction that involves the combination of focused ion beam milling and scanning electron microscopy (FIB/SEM). To analyse the data we have used new informatics tools developed in our laboratory in collaboration with neuroinformatics teams (see below). Data have been delivered as tiff image stacks and excel files. The format of these data has been aligned with the Neuroinformatics Platform (SP5) requirements.

Dataset Information Card (DIC):

Synapse maps

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/31/?changelist_filters=o%3D4

3.9.1.3 Description of data

We have used 5 PSD95GFP mice (male, P56-P60) generated in the laboratory of Javier DEFELIPE. This tissue has been used generate the optical synaptic maps and for the correlation studies with EM.

High resolution optical synaptic maps: We have studied the density and distribution of PSD95 puncta revealing the number and position of synapses in the hippocampus and somatosensory cortex in fixed brains, in collaboration with GRANT's laboratory. We have developed a novel 3D quantification methodology (Figure 10) to analyse the number and distribution of the PSD95 puncta. Specifically, this methodology was applied to tissue sections from the cerebral cortex and includes 3D filter algorithms and Fourier Transforms, followed by a segmentation process using 3D watershed and spot snacked segmentation. Thereafter, an automated counting process was performed using 3D connected components (Figure 11).

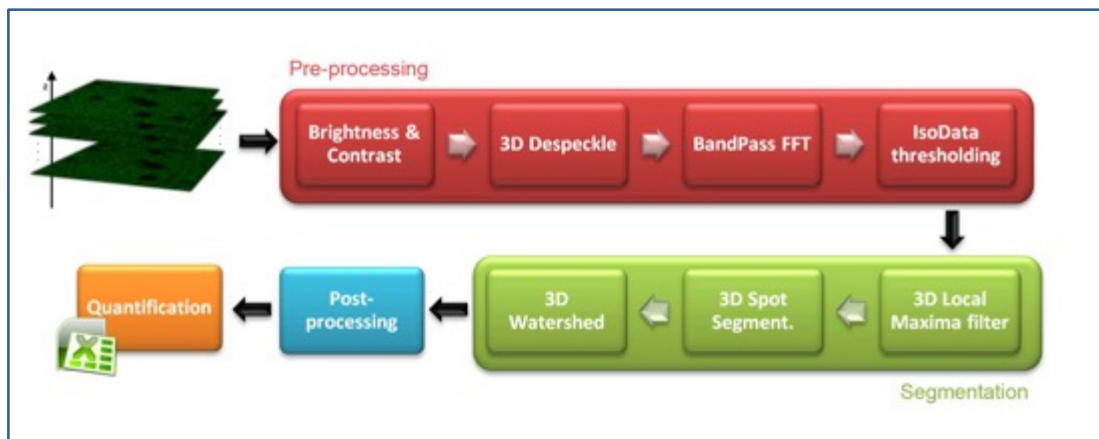


Figure 10: 3D Quantification Methodology

Diagram showing the main steps of the pre-processing task (red), segmentation (green), post-processing and quantification, performed in each stack of images in order to automatically estimate in 3D the number of PSD95 puncta.

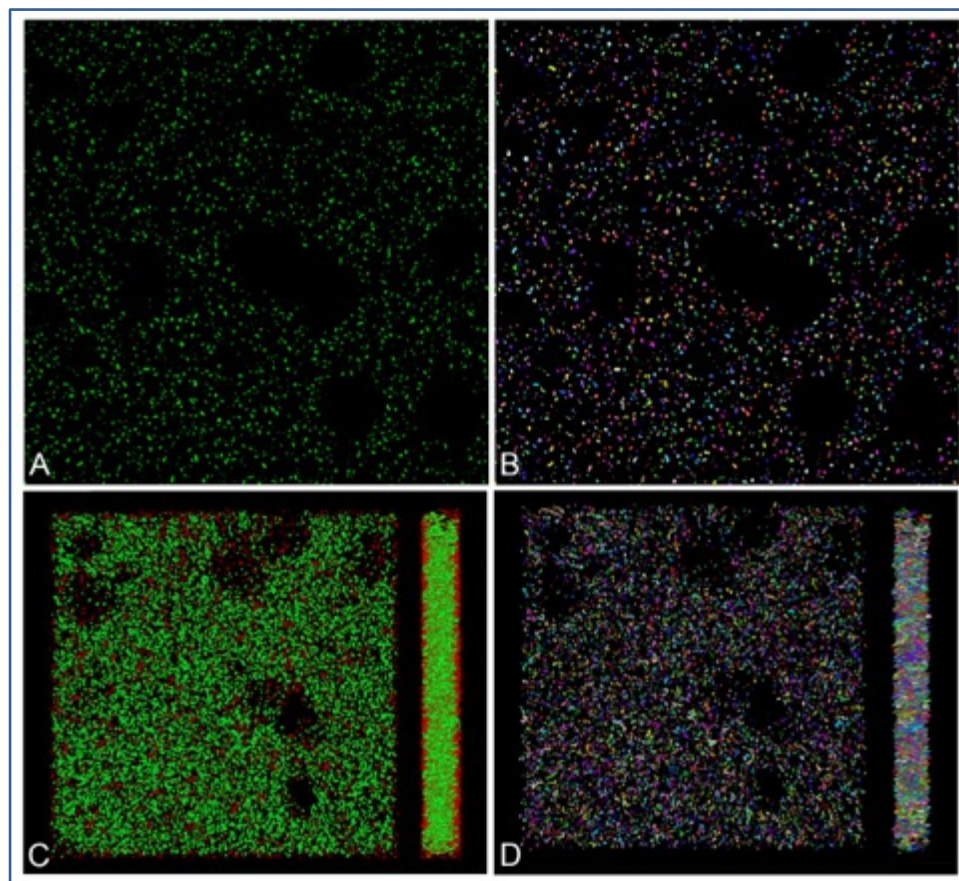


Figure 11: Imaging Processing

A) Original z projection stack of images (n=55) from layer III of somatosensory cortex of EGFP-PSD95 mouse after brightness/contrast correction algorithms and Despeckle denoise filtering processes. **B)** Shows the result after segmentation processing. Each color was randomly assigned to visualise individually detected PSD puncta. **C)** Illustrates the identified PSDs in 3D in frontal (left) and lateral (right) view. Puncta situated on edges of the stack are shown in red and were finally discarded during the post-processing step. **D)** Shows the result after segmentation processing of individually

coloured PSDs, in 3D in frontal (left) and lateral (right) view. Note that objects situated on the edges of the stacks were removed.

Furthermore, we manually reconstructed in 3D over 300 individual PSD95 puncta to generate a realistic model of PSD95 puncta distribution in the cerebral cortex to validate the accuracy of the counting. We demonstrate that with this methodology the accuracy of PSD95 counting is close to 90%.

In order to generate the optical synaptic maps, we obtained high resolution 3D confocal tiling stacks, including the whole hippocampus and somatosensory cortex. We applied the above-described counting method to estimate the number and distribution of PSD95 puncta in these regions. In addition, we have developed a new method to select particular areas within these brain regions (Figure 12). In this way, it is possible to obtain more accurate information about the density of PSD95 puncta in the neuropil since the tool allows the possibility to exclude areas that are lacking of puncta (e.g., cell bodies or blood vessels)

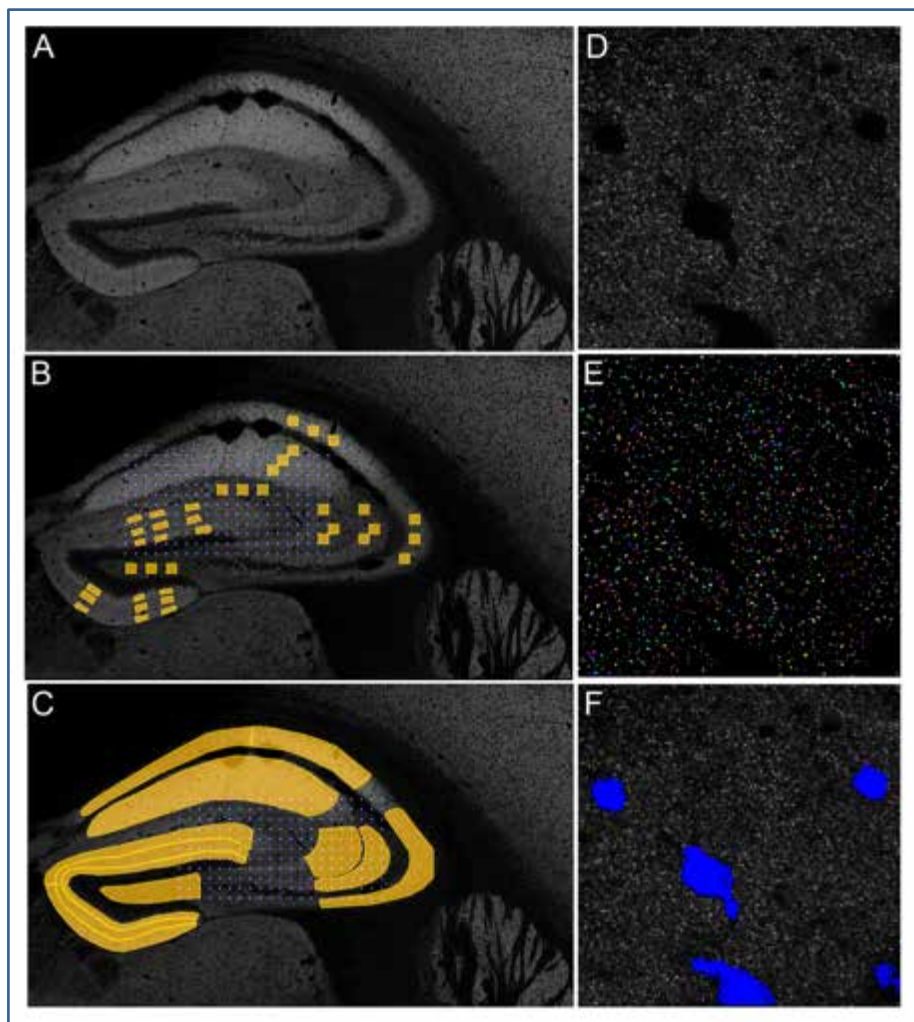


Figure 12: Synaptic Map of the Hippocampus

A) Tile scanning confocal image to illustrate the distribution of PSD95 labelling in the hippocampal formation of a PSD95 EGFP mouse. **B,C)** Same tile scanning confocal image shown in **A)** displaying a selection of particular stacks **B)** and regions of interest **C)** in the different areas of the hippocampal formation. **D)** Example of a stack from CA1 Stratum Oriens. **E)** Segmented PSD95 puncta from image shown in **D)**. **F)** Elimination of areas that lack PSD95 puncta (blue).

Correlative optical FIB/SEM synaptic maps: We have obtained 3D FIB/SEM stacks from the hippocampus and somatosensory cortex of three different PSD95GFP animals. The corresponding confocal scans showing PSD95 positive puncta have also been acquired and matched with EM data (Figure 13).

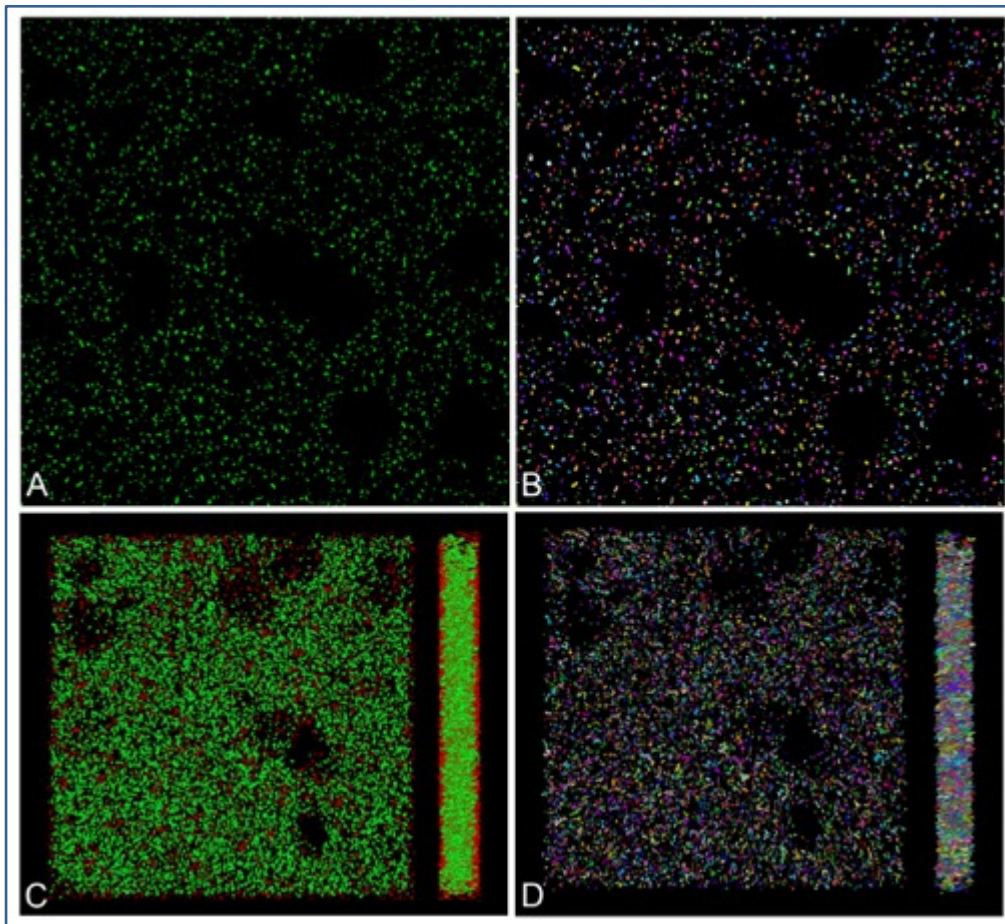


Figure 13: Example of a correlative optical and FIB/SEM synaptic map

First row: Confocal scans showing PSD95 positive puncta in different layers (SLM, SR and SO) of the hippocampus that were used for calculate the number and distribution of the PSD95 puncta. Second row: 3D FIB/SEM stacks obtained in the adjacent sections illustrated in the first row. This optical/EM correlation not only serves to validate the methodology for counting PSD95 puncta and generate optical synaptic maps, but also to generate EM synaptic maps of the neuropil in any region of the brain.

3.9.1.4 Completeness of data

3.9.1.4.1 High resolution optical synaptic maps

We have obtained 10 optical synaptic maps, including all hippocampal regions and all layers of the somatosensory.

3.9.1.4.2 Correlative optical and FIB/SEM synaptic maps: EM blocks scans and volume analysis

We have generated six synaptic maps comprising FIB/SEM and confocal images (three of the hippocampus, three of the somatosensory cortex); that is three times more than was planned to perform by M30.

This dataset is unique in that the volume of tissue that have been reconstructed at the EM level from the adult mouse hippocampus and somatosensory cortex is quite small in existing databases (see for example, Harris et al., Sci Data; 2:150046. doi: 10.1038/sdata.2015.46. eCollection 2015). Furthermore, high-quality synaptic optical maps of synaptic proteins are sparse.

The technology is now available to be exploited for comprehensive studies and generation of high resolution optical synaptic protein maps and 3D reconstruction using automatic EM techniques. However, data on optical synaptic protein maps and reconstructions at the EM synaptic resolution are sparse (see for example Kasthuri et al. *Cell* **162**, 648-661, 2015).

3.9.1.5 Data quality and value

The quality of the data was checked systematically by experts in the laboratory. In addition, further analysis and validation will be performed in the context of model construction.

The quantitative data on the number and distribution of synapses are essential for the building of models of the hippocampus and neocortex.

3.9.1.6 Data usage to date

The data being generated by this task feed the Neuroinformatics Platform (SP5) (SP5; Task 5.3.6; Task 5.4.3; Task 5.4.5; and, the Brain Simulation Platform (SP6, T6.2.2).

3.9.1.7 Are the data considered final?

These data still need to be revised and statistically analyzed to be considered final.

3.9.1.8 Publications connected to the gathered data

- Anton-Sanchez, L., Bielza, C., Merchan-Perez, A., Rodríguez, J.-R., DeFelipe, J., and Larrañaga, P. *Three-dimensional distribution of cortical synapses: a replicated point pattern-based analysis*. *Frontiers in Neuroanatomy*, 2014. **8**:85. doi:10.3389/fnana.2014.00085.

(In this article we were interested in exploring the three-dimensional spatial distribution of synapses in the cerebral cortex. The methodology used in this article will be applied for the mathematical analysis of the synaptic maps at both optical and the EM levels during the SGA1).

- Bosch, C., Martínez, A., Masachs, N., Teixeira, C. M., Feraud, I., Ulloa, F., et al. (2015). *FIB/SEM technology and high-throughput 3D reconstruction of dendritic spines and synapses in GFP-labeled adult-generated neurons*. *Frontiers in Neuroanatomy*, 2015. **9**:60. doi:10.3389/fnana.2015.00060.

(Here we show that FIB/SEM allows efficient, complete, and automatic 3D reconstruction of identified dendrites, including their spines and synapses, from GFP/DAB-labeled neurons. We will apply this technology to analyse the synaptic maps of labelled pyramidal neurons in both the rodent and human cerebral cortex during the SGA1).

- Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., et al. *Reconstruction and Simulation of Neocortical Microcircuitry*. *Cell*, 2015. **163**:456-492. doi:10.1016/j.cell.2015.09.029.

(This is a first-draft digital reconstruction of the microcircuitry of somatosensory cortex of juvenile rat. The reconstruction uses cellular and synaptic organising principles to algorithmically reconstruct detailed anatomy and physiology from sparse experimental data. This methodology provides an excellent roadmap for data-driven modelling and simulation of adult rodent and human brain circuitries that will be developed during the SGA1).

- Montes, J., Peña, J. M., DeFelipe, J., Herreras, O., and Merchan-Perez, A. *The Influence of Synaptic Size on AMPA Receptor Activation: A Monte Carlo Model*. PLOS ONE, 2015. 10:e0130924. doi:10.1371/journal.pone.0130924.

(Chemical synaptic transmission involves the release of a neurotransmitter that diffuses in the extracellular space and interacts with specific receptors located on the postsynaptic membrane. Computer simulation approaches provide fundamental tools for exploring various aspects of the synaptic transmission under different conditions. In particular, Monte Carlo methods can track the stochastic movements of neurotransmitter molecules and their interactions with other discrete molecules, the receptors. However, these methods are computationally expensive, even when used with simplified models, preventing their use in large-scale and multi-scale simulations of complex neuronal systems that may involve large numbers of synaptic connections. We have developed a machine-learning based method that can accurately predict relevant aspects of the behaviour of synapses, such as the percentage of open synaptic receptors as a function of time since the release of the neurotransmitter, with considerably lower computational cost compared with the conventional Monte Carlo alternative. The method is designed to learn patterns and general principles from a corpus of previously generated Monte Carlo simulations of synapses covering a wide range of structural and functional characteristics. These patterns are later used as a predictive model of the behaviour of synapses under different conditions without the need for additional computationally expensive Monte Carlo simulations. The resulting procedure is accurate, automatic, and it is general enough to predict synapse behaviour under experimental conditions that are different to the ones it has been trained on. Since our method efficiently reproduces the results that can be obtained with Monte Carlo simulations at a considerably lower computational cost, it is suitable for the simulation of high numbers of synapses and it is therefore an excellent tool for multi-scale simulations that will be performed during the SGA1).

- Márquez Neila, P., Baumela, L., González-Soriano, J., Rodríguez, J.-R., DeFelipe, J., and Merchán-Pérez, Á. *A Fast Method for the Segmentation of Synaptic Junctions and Mitochondria in Serial Electron Microscopic Images of the Brain*. Neuroinformatics, 2016. doi:10.1007/s12021-015-9288-z.

(Imaging techniques permit the automatic acquisition of a large number of serial sections from brain samples. Manual segmentation of these images is tedious, time-consuming and requires a high degree of user expertise. Therefore, there is considerable interest in developing automatic segmentation methods. We present a method that works with anisotropic voxels and that is computationally efficient allowing the segmentation of large image stacks. Our approach involves anisotropy-aware regularisation via conditional random field inference and surface smoothing techniques to improve the segmentation and visualisation. We have focused on the segmentation of mitochondria and synaptic junctions in EM stacks from the cerebral cortex, and have compared the results to those obtained by other methods. Our method is faster than other methods with similar segmentation results. Our image regularisation procedure introduces high-level knowledge about the structure of labels. We have also reduced memory requirements with the introduction of energy optimisation in overlapping partitions, which permits the regularisation of very large image stacks. This technology will be used during the SGA1 to accelerate the process of analysis of the EM images stacks).

- Luengo-Sanchez, S., Bielza, C., Benavides-Piccione, R., Fernaud-Espinosa, I., DeFelipe, J., and Larrañaga, P. *A univocal definition of the neuronal soma morphology using Gaussian mixture models*. Frontiers in Neuroanatomy, 2016. 9:137.

(High-throughput imaging methods have expanded quickly over the last few years, and the manual tracing of individual cells is a time-consuming task. Thus, it is necessary to

develop automatic techniques to acquire morphometric data on labelled neurons. The definition of the soma is fuzzy, as there is no clear line demarcating the soma of the labelled neurons and the origin of the dendrites and axon. Thus, the morphometric analysis of the neuronal soma is highly subjective. In this paper, we provide a mathematical definition and an automatic segmentation method to delimit the neuronal soma. We applied this method to the characterisation of pyramidal cells. This methodology will be used to analyse rodent and human pyramidal cells during the SGA1).

Furthermore, a manuscript is in preparation to describe the 3D quantification methodology to analyse the number and distribution PSD95 puncta.

3.9.2 Provenance

SP1: WP1.1, Task 1.1.3 (Seth GRANT) and WP1.2, Task 1.2.6 (Javier DEFELIPE)

3.10 High-Resolution 3D Maps of the Brain Vasculature

3.10.1 Data Description

3.10.1.1 Task(s)/group(s) responsible for generating data

T1.2.1, UZH

3.10.1.2 Data, tools and methodologies storage location(s) (and links?)

Detailed maps of the cerebrovascular system in the mouse brain are reconstructed in T.1.2.1 from different high-resolution imaging modalities including (A) absorption-based and phase contrast tomography (Prof. Marco STAMPANONI, Swiss Light Source, Villigen, Switzerland) and (B) confocal light sheet microscopy (Prof. Francesco PAVONE, European Laboratory for Non Linear Spectroscopy, Florence, Italy).

The acquired 3-D image stacks (raw image data) are analysed using a customised software framework for vessel segmentation and vascular reconstruction to obtain an explicit graph-based model of the complete cerebrovascular system.

Dataset Information Cards (DICs):

SRXTM Rat Cortex

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/138/

Cortex Vasculature Rat

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/139/

3.10.1.3 Description of data

3.10.1.3.1 Absorption-Based X-Ray Tomography

We have acquired four 3-D high-resolution angiographic datasets obtained from synchrotron radiation X-ray tomographic microscopy (SRXTM) of cylindrical samples of the somatosensory cortex (volume: 4.5mm^3 , optical magnification 20x, isotropic voxel spacing 700nm, 16 bit grayscale) of four rats (Sprague-Dawley, 10-12 months, male). An explicit graph-based model of cerebrovascular structures has been reconstructed from the image data with vessel diameters ranging from $60\mu\text{m}$ (pial vessels at the cortex surface) to $3\mu\text{m}$ at the capillary level. These datasets have been acquired to develop, implement, and validate the algorithms and protocols for image acquisition, vessel segmentation and reconstruction. Clearly, experiments on larger samples will use the targeted mouse models.

3.10.1.3.2 Phase-Contrast CT

Absorption-based imaging techniques are limited in the size of the sample (total attenuation). Phase-Contrast CT is viable alternative to image larger samples. Our current experimental setup yields an isotropic voxel size of $0.65\mu\text{m}^3$. The reconstruction of the phase-contrast sinograms involves additional postprocessing, most notably Paganin filtering, limiting the effective spatial resolution to about $1\mu\text{m}^3$. Data quality is most critically limited by imperfect sample preparation (incomplete filling of vessels due to occlusions or air-bubbles) and potential structural/positional changes of the sample during the acquisition lasting multiple hours (dislocation, shrinkage, sample-beam interaction). We have conducted multiple pilot studies and were able to significantly optimise the preparation and acquisition protocol including the choice of proper contrast agents. Besides the pilot studies, we have scanned an entire mouse brain of 1cm^3 volume perfused with 50% Indian ink and embedded in glue to ensure stability (effective resolution $1\mu\text{m}^3$,

16bit grayscale, 792 local tomographies, 30h scanning time, 7TB raw data). The local tomographies have been reconstructed and stitched together to form a consistent 3-D volume in a common coordinate frame.

3.10.1.3.3 *Confocal light sheet microscopy*

In contrast to X-ray-based imaging, multi-channel confocal light sheet microscopy allows for simultaneous acquisition of vascular features and further fluorescent labelled structures of the same sample in a single scan with perfect alignment (no need for further registration). Our current acquisition protocol uses an isotropic resolution of $0.52 \times 0.52 \times 2 \mu\text{m}^3$. The most limiting factor is PSF anisotropy resulting in critical axial resolution. Further improvement of the axial PSF, optimisation of the beam profile, and implementation of a virtual slit aperture is required to achieve sufficient spatial resolution for vascular reconstruction.

Even though many classical optical imaging methods provide high spatial accuracy, they are restricted to the investigation of (brain) tissue within a depth of a few hundreds of microns because of the limited penetration ability of the illuminating radiation. Therefore, imaging larger samples requires sectioning to obtain specimens of manageable thickness, as in all-optical histology. Sectioning, however, can severely distort or even damage the biological tissue, which requires non-trivial reconstruction.

3.10.1.4 **Completeness of data**

We are aware of only one whole-brain dataset (Xue, S et al., 2014) that has been acquired using micro-optical sectioning tomography (MOST) - a modification of knife-edge scanning microscopy (KESM) (Li. A. et al., 2010). We believe that the non-destructive nature of phase-contrast CT and confocal two-photon laser microscopy are clearly preferable over MOST/KESM.

3.10.1.5 **Data quality and value**

Regardless of the imaging modality, the acquired image data inevitably has certain shortcomings leading to artifacts in the reconstructed vasculature. Most notably, imperfect fillings lead to incomplete data resulting in gaps between vessel segments or even entirely missing subnetworks. Advanced postprocessing is hence critical for the reconstruction of complete vascular networks. While we have established a basic framework for automatic topology correction, there are still open problems particularly at the microvascular scale. The current dataset allows for the extraction of basic vascular statistics such as vascular volume fraction and intravascular distance. More involved statistics including topological and morphological information (bifurcation distance, tortuosity, bifurcation angles) have to be interpreted carefully in view of the incomplete vascular topology. Along the same lines, functional considerations of the vascular network as a whole, e.g., using computational fluid dynamics to simulate intravascular blood flow or tissue oxygenation, critically rely on a complete network topology.

3.10.1.6 **Publications connected to the gathered data**

- Schneider, M., Hirsch, S., Weber, B., Székely, G., and Menze, B.H. *Joint 3-D Vessel Segmentation and Centerline Extraction Using Oblique Hough Forests with Steerable Filters*. Medical Image Analysis, 2015. **19**:220-249.
(*Methodology for vessel segmentation and vascular reconstruction, absorption-based synchrotron dataset*).
- Rempfler, M., Schneider, M., Ielacqua, G.D., Xiao, X., Stock, S.R., Klohs, J., Székely, G., Andres, B., and Menze, B.H. *Reconstructing cerebrovascular networks under local physiological constraints by integer programming*. Medical Image Analysis, 2015. **25**:86-94.

(Methodology for topology correction).

- Schneider, M., Hirsch, S., Weber, B., Székely, G., and Menze, B.H. *TGIF: Topological Gap In-Fill for Vascular Networks - A Generative Physiological Modeling Approach*. In Polina Golland, Nobuhiko Hata, Christian Barillot, Joachim Hornegger, and Robert Howe, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2014*, volume 8674 of LNCS, pages 89-96. Springer Berlin/Heidelberg, 2014.

(Methodology for topology correction).

- Schneider, M., Reichold, J., Weber, B., Székely, G., and Sven Hirsch. *Tissue Metabolism Driven Arterial Tree Generation*. *Medical Image Analysis*, 2012. **16**:1397-1414.

(Methodology for topology correction).

- Patera, A., Astolfo, A., Mader, K.S., Schneider, M., Weber, B., and Stampanoni, M. *Towards the reconstruction of the mouse brain vascular networks with high-resolution synchrotron radiation X-ray tomographic microscopy*. *Medical Applications of Synchrotron Radiation (MASR2015)*, 2015. Poster session.

(Phase-contrast CT data, whole brain).

- Patera, A., Astolfo, A., Mader, K.S., Schneider, M., Weber, B., and Stampanoni, M. *Towards the reconstruction of the mouse brain vascular networks with high resolution synchrotron radiation X-ray tomographic microscopy*. *2nd International Conference on Tomography of Materials and Structures (ICTMS)*, 2015.

(Phase-contrast CT data, methodology for reconstruction and stitching).

- Xue, S., Gong, H., Jiang, T., Luo, W., Meng, Y., Liu, Q., Chen, S., and Li, A. *Indian-ink perfusion based method for reconstructing continuous vascular networks in whole mouse brain*. *PLoS One*, 2014. **9**: e88067.
- Li, A., Gong, H., Zhang, B., Wang, Q., Yan, C., Wu, J., Liu, Q., Zeng, S., and Luo, Q. *Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain*. *Science*, 2010. **330**:1404-1408.

3.10.2 Provenance

Absorption-based and phase contrast tomography:

Prof. Marco STAMPANONI, Swiss Light Source, Villigen, Switzerland

Reconstructed vascular models:

Prof. Bruno WEBER, SP1, WP12

Confocal light sheet laser microscopy:

Prof. Francesco PAVONE, SP1, WP12

3.11 Potassium Channels Kinetics

3.11.1 Data Description: Homology Models of K Channels

T1.1.5/SIB, UNIBAS

3.11.1.1 Task / group responsible for generating data

T1.1.5 - Group of Torsten Schwede (UNIBAS)

3.11.1.2 Data, tools and methodologies storage location(s) (and links?).

The data consists of three-dimensional atomistic models of mouse and human K-channels generated by template based (aka homology) modelling techniques. As models are generated using an automated homology modelling pipeline, it is preferable to refer to the original resources since models are updated on a regular basis as new crystal structures become available. Most improvements that we have made to the pipeline in the framework of this project (proper modelling of the oligomeric state of potassium channels, better assessment of the quality of the models and better modelling of loops) will be included into the next release of the productive version of SWISS-MODEL (<http://swissmodel.expasy.org>). Because of the widespread use of this resource (around 1000 models get generated on the SWISS-MODEL server every day), changes in the pipeline are thoroughly tested over long periods of time before they get included into the server. Meanwhile the 3D models that were produced will be deposited at ModelArchive.

- SwissModel server (<http://swissmodel.expasy.org>)
- SwissModel repository (<http://swissmodel.expasy.org/repository/>)
- Model Archive: (<http://www.modelarchive.org/>)

Dataset Identification Card (DIC):

Homology models of K channels

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/151/

3.11.1.3 Description of data

The following channels were modelled using the sequences peculiar to mouse and human:

Mouse:

Kv1.1,Kv1.2,Kv1.3,Kv1.4,Kv1.5,Kv1.6,Kv1.7,Kv1.8,Kv2.1,Kv2.2,Kv3.1,Kv3.2,Kv3.3,Kv3.4,Kv4.1,Kv4.2,Kv4.3,Kv5.1,Kv6.1,Kv6.2,Kv6.3,Kv6.4,Kv7.1,Kv7.2,Kv7.3,Kv7.4,Kv7.5,Kv8.1,Kv8.2,Kv9.1,Kv9.2,Kv9.3,Kv10.1,Kv10.2,Kv11.1,Kv11.2,Kv11.3,Kv12.1,Kv12.2,Kv12.3,HCN1,HCN2,HCN3,HCN4,K2p1.1,K2p2.1,K2p3.1,K2p4.1,K2p5.1,K2p6.1,K2p7.1,K2p9.1,K2p10.1,K2p12.1,K2p13.1,K2p15.1,K2p16.1,K2p17.1,K2p18.1,KCa1.1,KCa2.1,KCa2.2,KCa2.3,KCa3.1,KCa4.1,KCa4.2,KCa5.1,Kir1.1,Kir2.1,Kir2.2,Kir2.3,Kir2.4,Kir2.6,Kir3.1,Kir3.2,Kir3.3,Kir3.4,Kir4.1,Kir4.2,Kir5.1,Kir6.1,Kir6.2,Kir7.1

Human:

Kv1.1,Kv1.2,Kv1.3,Kv1.4,Kv1.5,Kv1.6,Kv1.7,Kv1.8,Kv2.1,Kv2.2,Kv3.1,Kv3.2,Kv3.3,Kv3.4,Kv4.1,Kv4.2,Kv4.3,Kv5.1,Kv6.1,Kv6.2,Kv6.3,Kv6.4,Kv7.1,Kv7.2,Kv7.3,Kv7.4,Kv7.5,Kv8.1,Kv8.2,Kv9.1,Kv9.2,Kv9.3,Kv10.1,Kv10.2,Kv11.1,Kv11.2,Kv11.3,Kv12.1,Kv12.2,Kv12.3,HCN1,HCN2,HCN3,HCN4,K2p1.1,K2p2.1,K2p3.1,K2p4.1,K2p5.1,K2p6.1,K2p7.1,K2p9.1,K2p10.1,K2p12.1,K2p13.1,K2p15.1,K2p16.1,K2p17.1,K2p18.1,KCa1.1,KCa2.1,KCa2.2,KCa2.3,KCa3.1,KCa4.1,KCa4.2,KCa5.1,Kir1.1,Kir2.1,Kir2.2,Kir2.3,Kir2.4,Kir2.6,Kir3.1,Kir3.2,Kir3.3,Kir3.4,Kir4.1,Kir4.2,Kir5.1,Kir6.1,Kir6.2,Kir7.1

3.11.1.4 Completeness of data

The dataset is in line with the anticipated one.

The dataset is complete; its quality is expected to increase with time as new structures become available.

3.11.1.5 Data quality and value

The quality of the models was assessed using: 1) GMQE (global model quality estimation), which predicts the quality of the model from the target-template sequence alignment, and 2) the newly developed QMEANbrane, which uses statistical potentials specific to membrane proteins, to assess the local and global quality of the modelled parts of the channels.

Model quality varies for different channels, but overall we find that reliable models were obtained for Kv, Kir and K2p channels. On the other hand, models of KCa and HCN channels are much less reliable because there is no high-resolution structure available for any member of these two families.

3.11.1.6 Data usage to date

The homology models will first be used by us within SGA1 Task 1.1.5 (KChannelKinetics: Modulation of action potential propagation by K⁺ channels) in which the functional properties of the different channels will be compared. The structural specificity of the different channels will be used in complement to electrophysiological data to define groups of channels and associated kinetic models.

3.11.1.7 Are the data considered final?

As stated above, models will improve as new crystal structures of potassium channels are released. The currently available models nevertheless fulfil the expectations.

3.11.1.8 Publications connected to the gathered data

- G. Studer, S. Bienert, N. Johner and T. Schwede, "PROMOD3 - A versatile homology modeling toolbox", Proceedings of 3DSIG 2015, Dublin, 114.
(Description of the new modelling engine for SwissModel).
- N. Johner, S. Bernèche, T. Schwede, "Structural constraints on evolution and their use for contact prediction", Submitted to Molecular Biology and Evolution (2016).
(Development of a new method to predict pairs of residues forming contacts in a protein from a multiple sequence alignment. These predictions are used to determine the relative orientation of the domains when modelling multi-domain proteins, which most potassium channels are.)

3.11.2 Data Description: Functional Annotation of K channels

3.11.2.1 Task / group responsible for generating data

T1.1.5 - Group of Ioannis XENARIOS (SIB)

3.11.2.2 Data, tools and methodologies storage location(s) (and links?)

Professional curators of the UniProt team read, analyse, compare, summarise and curate the data, conclusions and models found in scientific publications discussing the function and properties of a given protein.

The database is found at: <http://www.uniprot.org>

Dataset Identification Card (DIC):*K Channel Functional Annotation*

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/150/

3.11.2.3 Description of data

Updated entries by manual curation for several representative potassium channels (KCNA2, KCNB1, KCNC2, KCND2 and KCNK1).

- Completeness of data:

The delivered dataset in line with the anticipated one.

The manual curation protocol of UniProt ensures that the data is as complete as our current knowledge of a given protein is.

Many K channels are still not (fully) annotated in the UniProt database. The UniProt team will address this according to its protein priority list.

3.11.2.4 Data quality and value

The quality of the data found in the literature is controlled by an expert curator. A curator occasionally contacts authors of published work to clarify ambiguous points.

UniProt is used every month by ~500,000 users, which highlights the value of the data. Members of the HBP consortium are highly likely to be among these users.

3.11.2.5 Data usage to date

The specific K channel entries have been used within task T1.1.5 K Channel Kinetics (Simon BERNÈCHE) to verify that the channel's properties observed in the electrophysiology measurements used for the kinetic modelling are in line with the generally known properties of the channel.

3.11.2.6 Are the data considered final?

UniProt entries are updated on a monthly basis.

3.11.3 Data Description: K Channel Kinetics**3.11.3.1 Task / group responsible for generating data**

T1.1.5 - Group of Simon BERNÈCHE (SIB)

3.11.3.2 Data, tools and methodologies storage location(s) (and links?)

Our approach consisted in developing a single generic kinetic model that we hypothesised is common to all K channels. The rates are specific to each channel and are determined on the basis of available electrophysiology data. The generic model is based on our understanding of the molecular function of potassium channels, integrating structural and functional data as well as insights emerging from molecular mechanics simulations. The model notably includes the position of ions in the selectivity filter of the channel and presents the advantage that state transitions directly relate to structural changes in the protein. We have also developed two simpler kinetic models that serve to assess the usefulness of the added complexity in reproducing the experimental data.

The kinetic models and the associated rates will first be deposited in Channelpedia (<http://channelpedia.epfl.ch>).

Dataset Identification Card (DIC):

K Channel Kinetics

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/101/

3.11.3.3 Description of data

Three generic kinetic models are defined with different levels of complexity.

For each K channel, a set of rates is provided for each of the three kinetic models. The quality of the fit to the functional data is indicated for each model.

3.11.3.4 Completeness of data

The delivered dataset meets and exceeds the anticipated one. Instead of a single generic kinetic model, we are proposing three generic models of different complexity. Also, instead of only relying on scarce channel electrophysiological data found in the literature, we have profited from the exhaustive patch-clamp measurements performed at the Brain Mind Institute of the EPFL on all relevant K channels. We can thus provide reliable kinetic models for more channels than initially anticipated.

The dataset is complete, but the kinetic models are not fixed in stone and are expected to evolve.

Over the last decade, the mechanisms underlying ion channel function have been studied through structural approaches and, unfortunately, kinetic modelling has been mostly abandoned. Our kinetic models are the first to aim at bridging the structural and functional data of K channels.

3.11.3.5 Data quality and value

Quality of the model is assessed by its capacity to reliably fit the experimental data. Moreover, we check whether the obtained rates are identifiable (not degenerate in the sense that a change in the rate would lead to a worse fit of the data).

This exhaustive library of K channel kinetic models is a first. It aims to be a one-stop shop where theoretical neuroscientists will find reliable models taking into account the structural and biophysical properties of the channels. By design, these K channel models are independent of any neuronal data and thus should be appropriate for modelling any type of neurons.

3.11.3.6 Data usage to date

The kinetic models will first be used by us within SGA1 Task 1.1.5 (KChannelKinetics: Modulation of action potential propagation by K⁺ channels) to evaluate the signalling properties of each channel and to determine which of them should specifically be treated in neuron models. The results from Task 1.1.5 will be communicated on a regular basis to SP6.

3.11.3.7 Are the data considered final?

The data is not considered final as the development of the kinetic models should be an iterative process. As we get more channel electrophysiology data and we start working on neuron modelling, the K channel kinetic models will get further challenged and might have to be modified.

3.11.4 Provenance

SP1, WP1.1, Task 1.1.5: Torsten SCHWEDE (UNIBAS), Ioannis XENARIOS (SIB), Simon BERNÈCHE (SIB)

3.12 Trans-Synaptic Signalling and Receptor Kinetics

3.12.1 Data Description

3.12.1.1 Task / group responsible for generating data

T1.1.6/SNS, EBRI, CNR

3.12.1.2 Data, tools and methodologies storage location(s)

Dataset Information Cards have been created on the Neuroinformatics Platform and data uploaded:

<https://collab.humanbrainproject.eu/#/collab/348/nav/3158?state=uuid%3Db40f63dc-97a8-4d83-9919-97450d49d27d>

<https://collab.humanbrainproject.eu/#/collab/348/nav/3158?state=uuid%3D68a3e02f-5d7e-4052-9ed7-315d8f09d99e>

<https://collab.humanbrainproject.eu/#/collab/348/nav/3158?state=uuid%3Dad085507-cd77-4a8b-8336-d544dc95d2a4>

<https://collab.humanbrainproject.eu/#/collab/348/nav/3158?state=uuid%3D65d2cbda-bfc8-4cf2-92ee-c0174338dc9e>

<https://collab.humanbrainproject.eu/#/collab/348/nav/3158?state=uuid%3D1262aceb-5baf-4c45-82f9-4e5d25f3d3b8>

A modelling tool has been created on the Collaboratory (fitting task, <https://collab.humanbrainproject.eu/#/collab/348/nav/4922>).

Dataset Identification Cards (DICs):

Action potential dependent sIPSCs - from 3-4 months old Tg2576 female mice (A)

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/78/

Action potential dependent sIPSCs - from juvenile (P21-30) C57Bl6/J male mice - from CA1 pyramidal neurons receiving input from PV+ and CCK+ interneurons (B)

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/77/

Action potential dependent sIPSCs - from juvenile (P21-30) C57Bl6/J male mice - from CA1 pyramidal neurons receiving input from PV+ interneurons (C)

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/76/

Action potential dependent sIPSCs - from 3-4 months old wt (B6/SJL) female mice (D)

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/75/

Action potential dependent sIPSCs - from juvenile (P21-30) C57Bl6/J male mice (E)

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/74/?changelist_filters=o%3D4

Kinetic scheme of inhibitory events including gephyrin

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/275/

3.12.1.3 Description of data

- Experiment A:
 - Species: mice
 - Strain: Tg2576
 - Gender: female
 - Age: 3-4 months old
 - Number of specimen: 10 cells/3 animals
- Experiment B:
 - Species: mice
 - Strain: C57Bl6/J
 - Gender: male
 - Age: 3 weeks old
 - Number of specimen: 3 cells/3 animals
- Experiment C:
 - Species: mice
 - Strain: C57Bl6/J
 - Gender: male
 - Age: 3 weeks old
 - Number of specimen: 3 cells/3 animals
- Experiment D:
 - Species: mice
 - Strain: B6/SJL
 - Gender: female
 - Age: 3-4 months old
 - Number of specimen: 7 cells/3 animals
- Experiment E:
 - Species: mice
 - Strain: C57Bl6/J
 - Gender: female
 - Age: 3 weeks old
 - Number of specimen: 3 cells/3 animals
- Experiment F-G:
 - Species: mice
 - Strain: NL3 KO and NL3 R451C KI
 - Gender: male and female
 - Age: 3-8 days old

- Number of specimen: 6 cells/4 animals (for each strain)
- Experiment H:
 - Species: mice
 - Strain: NL3 KO and NL3 R451C KI
 - Gender: male and female
 - Age: 3-8 days old
 - Number of specimen: 6 cells/4 animals (for each strain)
- Experiment I:
 - Species: mice
 - Strain: C57Bl6/J and NL3 KO
 - Gender: male
 - Age: 4-5 weeks old
 - Number of specimen: 20 cells/6 animals
- Experiment L:
 - Species: mice
 - Strain: NL3 KO and NL3 R451C KI
 - Gender: male and female
 - Age: 6 days old (10 DIV)
 - Number of specimen: 3 cells/3 animals

Brain Area: CA1 Hippocampus, Target: CA1 principal cells, Patch-clamp recordings of inhibitory post-synaptic currents in acute slices

3.12.1.4 Completeness of data

Antibody characterisation: Each intrabody selected, namely anti -NLG1 clone 18, anti-NLG2 clone 1 and clone 645, anti-NLG3 clone 45, showed specificity only for the original bait screened with no cross reaction against the other NLGs baits in yeast. Anti-NLG2 clone 1 and clone 645 were tested *in vitro* for specific interaction with NLG2 in co-immunoprecipitation assay.

Electrophysiology data: Only control data (WT, NLGL3KO/KI from acute and organotypic hippocampal slices) been delivered up to now.

Functional data obtained by selectively silencing gephyrin and NLGs with intrabodies will constitute the basis for studies on transynaptic signalling.

In addition, the intrabody tool will be used to study the functional properties of other proteins for which traditional tools are not available

We have been part of the HBP for 23 months. During this period, we have generated intrabodies against gephyrin, and NLG1, NLG2 and NLG3. We are now in the process of testing the function of the intrabodies by expressing them via viral injection in organotypic hippocampal slices. This will allow elucidating how changes in gephyrin and/or neuroligin expression alter the functional organisation of excitatory and inhibitory synapses and their plasticity processes.

So far, there is no evidence in literature on intrabodies against NGLs family since these proteins are only being studied using KO and KI mice.

These data will advance the knowledge obtained from NL3 R451C knock-in and NL3 knock-out mice.

3.12.1.5 Data quality and value

Data generated so far from Controls, NL3 R451C knock-in and NL3 knock-out mice at different postnatal developmental stages have been compared with those present in the literature from the groups of A. TRILLER, T. SUDHOF, P. SHEIFFELE, N. BROSE

We believe our data will be very valuable for the users because they will provide new insights into the function of synapses in physiological and pathological conditions.

3.12.1.6 Data usage to date

Data was used to test the fitting procedure (fitting task, Collaboratory, <https://collab.humanbrainproject.eu/#/collab/348/nav/4922>); the results from the fitting task will be used in the Brain Simulation Platform (T6.2.4)

3.12.1.7 Are the data considered final?

Electrophysiological data are not final since they need to be compared with those obtained after silencing gephyrin and neuroligins with their respective intrabodies.

The fitting task will be public on March 30th. It will be constantly improved and upgraded according to user requests.

3.12.1.8 Publications connected to the gathered data

- Lupascu et al., A general procedure to study subcellular models of transsynaptic signaling at inhibitory synapses, 2016, submitted.

(Implementation and use of a procedure to fit of individual inhibitory spontaneous synaptic events; gathered data is used to test the procedure, and for a preliminary analysis using an initial kinetic model of gephyrin-dependent subcellular pathways).

3.12.2 Provenance

SP1/ WP1/T1.1.6

3.13 Data Aggregation, Analysis and Dissemination

3.13.1 Data Description

3.13.1.1 Task / group responsible for generating data

WP1.3/UEDIN

3.13.1.2 Data, tools and methodologies storage location(s) (and links?)

A: List of synaptic target proteins

This list contains names of synaptic proteins for which synapse modellers would like to have quantitative information, e.g. localisation, abundance, binding partners. It should therefore be helpful to experimentalists in determining which proteins to quantify in the highest level of detail.

Location:

https://collaboration.humanbrainproject.eu/documents/portlet_file_entry/10727/target_proteins_HBP2.txt/2efd1525-5c57-4578-bcc0-c91ae4fba366?status=0

Dataset Identification Card (DIC):

List of synaptic target proteins

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/105/

Synaptic target proteins:

file:///C:/Users/Pilar/AppData/Local/Temp/target_proteins_HBP2_static.html

B: Curated data from neuroproteomics publications

Location: Awaiting SP5 adding data to Neuroinformatics Platform; in the meantime direct from Oksana SOROKINA <oksana.sorokina@ed.ac.uk>.

We have identified ~50 publications containing neuroproteomic data. We have extracted and put in to a standard format data tables from a subset of these papers. We have attached metadata describing the experimental protocols used to generate the data in each table. In order to describe the protocols in a standard, query-able fashion we have developed, in association with SP5, ontology to describe the methods. Data from six papers were extracted as a validation and sent to SP5, who will incorporate it in the Neuroinformatics Platform. Once the data has been entered in the Neuroinformatics Platform we will restart work on extracting data.

Dataset Identification Card (DIC):

Curated data from neuroproteomics publications

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/274/?changelist_filters=o%3D4

C: KappaNEURON software package

Location: <https://github.com/davidcsterratt/KappaNEURON>

This software package incorporates the rule-based SpatialKappa simulator into the NEURON simulator, thus allowing complex molecular simulations, e.g. of the postsynaptic density, to exist in the context of electrical activity. Rule-based modelling defines interactions between binding sites rather than between complexes and complexes are created dynamically throughout the simulation.

Dataset Identification Card (DIC):*KappaNEURON software package*

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/97/

D: Community Detection Modularity Suite

Location: <http://sourceforge.net/projects/cdmsuite/>

CDM is a software package featuring: C++ and R implementations of Newman & Girvan Modularity-based community detection algorithms; R implementations of the edge Betweenness Random Walk algorithm; Boot-strapping facilities to test cluster robustness; OpenMP implementation of Newman & Girvan Geodesic and Random Walk edge; Betweenness algorithms; Overlapping community detection model, including the partition density and fuzzy modularity metrics for community detection.

Dataset Identification Card (DIC):*Community Detection Modularity Suite*

https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/130/

3.13.1.3 Description of data

- Species, sex, age, number of specimen/subjects:
 - A: Mouse, mapped to human
 - B: Varies according to paper
 - C: N/A
 - D: N/A
- Scale (brain, brain region, cells, molecules):
 - A: Molecules/brain.
 - B: Varies according to paper.
 - C: Integrates simulations at the molecular level with neuron-level simulations of electrical activity.
 - D: Molecules

3.13.1.4 Completeness of data

A: This piece of work was not anticipated, but it falls into the remit of T1.3.1. From the cross-SP meeting between SP1, SP2, SP5 and SP6 in the HBP Summit 2014 it was evident that several modelling groups at the cellular (e.g. WP6.4.5), intracellular (e.g. WP6.4.1 and WP6.4.2) and molecular levels (e.g. WP6.3) wanted to know more about the molecular composition of the biological objects being modelled. The quantities most in demand were the subtype identity of proteins involved, their specific localisation and companions inside the cell and the total protein amounts. This is a huge task for data-producing groups at HBP, or indeed for any other group, so efforts should be focused on just a few targets. We therefore agreed that the modellers would work on producing a list of target proteins.

B: This satisfies the requirements of MS13, “Formats and ontologies for molecular experiments”.

C: As per Task 1.3.2, KappaNEURON is an informatics tool that allows integration of molecular data from multiple data sources. It satisfies MS14 Informatics tools v1.

D: As per Task 1.3.2, Community Detection Suite is an informatics tool that allows integration of molecular data from multiple data sources. It satisfies MS15 Informatics tools v2.

A: As far as we are aware, no other group is generating this data.

B: As far as we are aware, no other group is generating this data.

C: The KappaNEURON simulator is unique: there is no other simulator that integrates rule-based molecular and electrical simulations.

D: The spectral modularity clustering algorithm produces more meaningful clusters than other algorithms (see figure).

3.13.1.5 Data quality and value

A: The data is not definitive, but rather a wish-list of proteins that modellers would like experimentalists to quantify. We have mapped each protein/gene to: the ID of the protein in the Mouse Genome Informatics Database; ENTREZ ID of mouse gene; ENTREZ ID of human gene; HUGO name of human gene; function (those not yet mapped have been marked NA); reason the protein was included in the list; description/name of the gene/protein; and name of gene, as given in Mouse Genome Informatics Database.

B: The generation of the ontology to describe methods was piloted on two papers, and involved in-depth discussions with members of SP5 to make it appropriate and useful. We are confident that by going through this process we have a fairly robust ontology, but refinements will be needed as we curate more papers.

C: The KappaNEURON simulator has been verified by comparing reference simulations performed using KappaNEURON and standard NEURON (Sterratt et al 2015).

D: The Community Detection suite has been validated against common clustering methods and this evaluation is published (McLean et al 2016).

A: The value of the list of target proteins is to experimentalists, looking to produce the datasets likely to be most use to modellers.

B: The ontology and data curation will be of use to modellers searching for parameter values for models.

C: The KappaNEURON simulator allows for models of synaptic proteomes with combinatorially large numbers of molecular complexes to be simulated in the context of electrical activity. It is therefore ideally suited for models of synaptic plasticity.

D: The Community Detection suite provides, to our knowledge, the only clustering algorithm that performs well without the requirement of iterative fine tuning of parameters on molecular interaction networks of the size, scale and sparseness commonly published.

3.13.1.6 Data usage to date

A: The list of target proteins has been used to facilitate discussions between experimentalists and modellers.

B: As this data is not in the platform yet, no-one has used it.

C: We have used KappaNEURON for a demonstration simulation of synaptic plasticity in Sterratt et al (2015).

D: CDM Suite has been used in analyses by ourselves and collaborators for the analysis of brain proteomics data. Specifically, we have been working on using it to help analyse the

some of the data produced by WP1.1; we are currently exploring how to add it to genenetwork.org. An early release version has been used by an external group in Turin for proteomic analysis (in preparation)

3.13.1.7 Are the data considered final?

A: No - we may need to extend the list in the light of modellers' views and experimental advances. Discussions in CDP2 and CDP5 may also lead to changes in the list.

B: No - there are more papers to curate. When the data start being entered in the Neuroinformatics Platform, we can resume our side of the work.

C: No - the code is still under development and could benefit from optimisations.

D: Yes.

3.13.1.8 Publications connected to the gathered data

A: None

B: None

C: Sterratt et al (2015) (Description of the algorithm and implementation of KappaNEURON)

D: McLean et al (2016) Description of the algorithm, validation and evaluation against existing algorithms against a variety of public datasets. Also included a CytoScape plugin to promote user uptake.

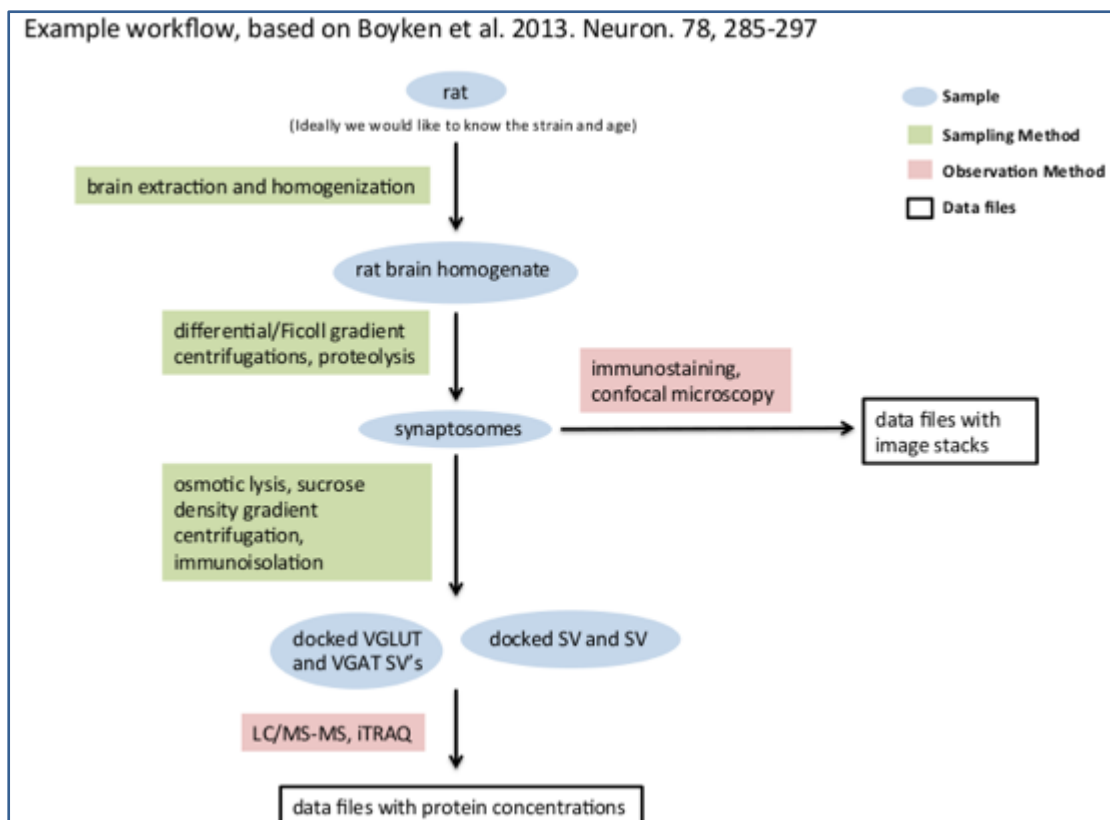


Figure 14: Ontology

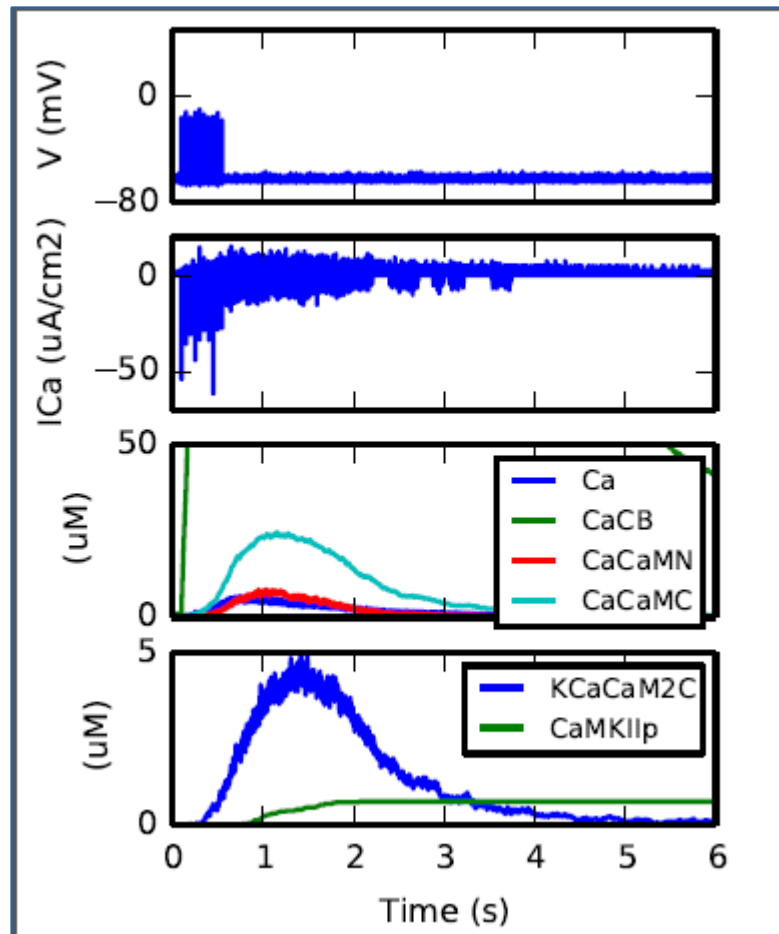


Figure 15: Simple PDS pepke Kappa nmda

3.13.2 Provenance

A: As explained earlier, the list of target proteins arose out of discussions between SP1, SP2, SP5 and SP6 in the HBP Summit 2014. We also included molecules that have appeared in previous computational models. The provenance of each item in the list is given in the data table.

B: The provenance of each table, including the name of the annotator, has been carefully marked up.

C: The coding work was carried out by David STERRATT.

D: The coding work was carried out by Colin MCLEAN.

Table 2: SP1 Numerical Indicators (D1.4.1)

Task	Indicator	Target (Month)									
		6		12		18		24		30	
		Planned	Achieved	Planned	Achieved	Planned	Achieved	Planned	Achieved	Planned	Achieved
T1.1.2	No. of neocortical neuron types with first draft SCTs	0	0	0	0	6	22	6	22	6	22
T1.1.2	No. of hippocampal neuron types with first draft SCTs	0	0	0	0	0	0	4	~20 (est.)	4	~20 (est.)
T1.1.2	No. of cerebellum neuron types with first draft SCTs	0	0	0	0	0	0	0	0	2	~5 (est.)
T1.1.3	No. of brain region synapse proteomes	0	0	0	0	15	A total of 42 brain region synapse proteome profiles generated by LC-MS/MS. (7 integral regions of the mouse brain from n=6 mice)	20	42	25	42
T1.2.1	No. of maps of mouse vasculature	0	0	4	4	6	6	6	6	0	0
T1.2.2	No. of cell distribution maps of the whole mouse brain	0	0	0	0	1	1	2	1	3	3
T1.2.6	No. of synaptic maps of individual identified neurons	0	0	0	0	12	12	24	16* see section 3.7, page 33)	60	16* (see section 3.7, page 33)
T1.2.4	No. of mouse cell morphologies reconstructed (hippocampus)	0	0	20	13 cells fully reconstructed & 13 visualised and scanned	50	23 cells fully reconstructed & 19 visualised and scanned	80	36 cells fully reconstructed & 24 visualised and scanned	110	74 cells fully reconstructed & 23 visualised and scanned (as of February 18, 2016)

Task	Indicator	Target (Month)									
		6		12		18		24		30	
		Planned	Achieved	Planned	Achieved	Planned	Achieved	Planned	Achieved	Planned	Achieved
T1.2.4	No. of mouse cell morphologies reconstructed (neocortex)	8	8 (mouse)	20-40	20-40 (mouse) & 30 single PCs and interneurons (rat) & 19 neurons (rat)	30-50	50 neurons (41 Golgi cells with incomplete structures and 9 PCs)	40-60	>60	50-70	> 70
T1.2.5	No. of axonal projections traced	1	1	4	6 Thalamocortical SS & 1 Thalamocortical visual	12	11 axons	20	13	30	11 4 Visual TC (3 complete 3D) 10 somatosensory (2 complete 3D)
T1.2.5	No. of terminal axon branches and monosynaptic targets identified	4	4	10	10	20	20 targets	30	26	30	TBC
T1.2.6	No. of synaptic maps of mouse brain	0	0	0	0	1	1	2	3	2	6

4. How Can Platform Developers Provide You With Feedback on The Data?

First Draft Transcriptomes of Major Hippocampal and Cerebellar Cell Types

- Contact: Chris PONTING, UOXF (Email chris.ponting@dpag.ox.ac.uk)

First Synapse Proteome for a Specific Cell Type

- Contact: Seth GRANT (seth.grant@ed.ac.uk)

Anatomical Reconstructions of Neocortical and Hippocampal Cells

Hippocampal Cells

- Contact: Attila GULYÁS (gulyas@koki.hu) or Szabolcs KÁLI (kali@koki.hu) in the Laboratory of Cerebral Cortex Research at the Institute of Experimental Medicine, Hungarian Academy of Sciences (IEM HAS)
- We solicit feedback on the morphological and physiological data from both a neuroinformatics and a modelling point of view, e.g., regarding the appropriateness of the data format, the completeness of the metadata, and the usefulness of the data for the construction of detailed neuronal models. Discussions of the anatomical and physiological protocols and comparisons with methods used by other laboratories would also be useful.
- Collaborations are sought to increase the efficiency and precision of morphological reconstruction, which is currently the most time-consuming (and rate-limiting) part of the workflow. Improved methods for the automatic, precise measurement of dendritic diameters based on confocal image stacks would be especially helpful.

Neocortical Cells

- Contact: Yun WANG (wangy62@gmail.com)

Single-cell and Bulk Connectomics of Thalamocortical and Corticocortical Projection Neurons

- Contact: Francisco CLASCÁ (UAM) (francisco.clasca@uam.es)
- We have a regular liaison established with developers in three SP5 groups:
 - Rembrandt BAKKER and Paul TIESINGA at Donders University (3D neuronal reconstructions)
 - Martin TELEFONT, William HILLCOCK and Elisabetta LAVARONE at EPFL (Formats for neuronal reconstruction and metadata requirements)
 - Jan BJAALIE at Oslo University (Long-range projection neuron registration to the HBP Digital Mouse Atlas)
- Over the Ramp-Up phase, we have held several focal meetings with these Developers, both presential or videoconference (Madrid March 2014; Firenze, July 2014, Heidelberg September 2014, March 2015 (videoconference); Madrid September 2015; Madrid February 2016. These extended interactions have been extremely helpful for defining

the specific requirements of the platform developers about the data to be produced and for breaking scientific language barriers, which is by no means a non-trivial issue given our widely different backgrounds: biomedical and computational. In turn, we think that we have helped developers perceive the critical factors involved in producing relevant usable data about long-range projection neurons. From these interactions we have defined a realistic and proven set of standards and procedures that we look forwards to implement during the following phases of HBP. Elisabetta LAVARONE will act as an “embedded”, permanent liaison with the EPFL group. We also look forwards to advancing the indexation atlasing effort with the Oslo group.

Distribution of Major Proteins

UEDIN

- Contact: Seth GRANT (seth.grant@ed.ac.uk)

UCLM & IST

- Contact: Rafael LUJÁN, UCLM, Email: Rafael.Lujan@uclm.es
- UCLM & IST: Example of data of density of ion channels in a form useful for numerical simulations.
- UCLM & IST: Software development/modification for automatic image analysis to extract gold particles of different sizes in EM images, to measure the number, density, and clustering of these different types of particles. Generation of antibodies good enough for EM applications to visualise receptors and ion channels on neuronal membranes.

Numbers, Distributions and Relative Densities of Cells in Selected Regions

LENS

- Contact: Ludovico SILVESTRI - silvestri@lens.unifi.it - +390554572504
- Define standards - at high detail level - for both data (file format, spatial coordinates, etc.) and metadata to maximize usability of experimental data in the platforms. Define a list of neuronal populations to look at in the SGA1. We are planning to map inhibitory neurons using markers such as parvalbumin, vaso-intestinal peptide (VIP), somatostatin. But we can add/change markers if platform developers believe that other neuronal populations would be more critical to map.
- We are looking for some help in registration of whole mouse brain images onto standard atlases (INCF or Allen).

UPM

- Contact: Javier DEFELIPE, UPM, Cajal Cortical Circuits Laboratory. Email: defelipe@cajal.csic.es
- The most helpful feedback would be from the developers of the Brain Atlas to find out the precise technical requirements of the Platforms to integrate the strategic data that are being generated. Discussions of the anatomical methodologies and comparisons with techniques used by other laboratories would also be useful.
- We would like to establish close collaborations with the Neuroinformatics Platform to improve the efficiency and quality of processing of images generated both with the

confocal microscope and with the electron microscope. Also, it is important to develop sophisticated analytical tools in the Platforms for the interpretation of the data generated.

Synaptic maps of identified neurons

- Contact: Javier DEFELIPE, UPM, Cajal Cortical Circuits Laboratory. Email: defelipe@cajal.csic.es
- The most helpful feedback from the Platforms would be to know how our data are being used by the Neuroinformatics Platform and modellers and how useful are the data for developers and users. Discussions of the anatomical methodologies and comparisons with techniques used by other laboratories would also be useful.
- We would like to establish close collaborations with the Neuroinformatics Platform to improve the efficiency and quality of automatic morphological reconstructions. This is the most time-consuming part of the workflow. The improvement of available methods would help to accelerate the generation of data.

High resolution optical synaptic maps and electron microscope blocks scans and volume analysis of exemplar brain regions

- Contact: Javier DEFELIPE, UPM, Cajal Cortical Circuits Laboratory. Email: defelipe@cajal.csic.es
- The most helpful feedback would be from the developers of the Brain Atlas to find out the precise technical requirements of the Platforms to integrate the strategic data that are being generated. Discussions of the anatomical methodologies and comparisons with techniques used by other laboratories would also be useful.
- We would like to establish close collaborations with the Neuroinformatics Platform to improve the efficiency and quality of processing of images generated both with the confocal microscope and with the electron microscope. Also, it is important to develop sophisticated analytical tools in the Platforms for the interpretation of the data generated.

High-resolution 3D maps of the brain vasculature

- Contact: Prof. Bruno WEBER, University of Zurich, bweber@pharma.uzh.ch
- Desired feedback: completeness of data, what kind of vascular information/labels are used/required.
- Collaborations: HPC infrastructure to upscale vascular reconstruction to whole brain (CINECA).

Potassium channels kinetics

- Who to contact: Niklaus JOHNER (niklaus.johner@unibas.ch)
- It would be useful to know about the restrictions that models implemented within the Brain Simulation Platform might impose on our kinetic models. What mathematical form or algorithm could ultimately be applied for simulating the activity of ion channels?

- We would be interested in the development of subcellular models highlighting the role of K channels and of reduced neuron models taking into account K channel properties.

Trans-synaptic signalling and receptor kinetics

- Contact: For modelling Michele MIGLIORE, michele.migliore@cnr.it; For electrophysiology: Enrico CHERUBINI: e.cherubini@ebri.it, Silvia MARINELLI: s.marinelli@ebri.it, Cristina MARCHETTI: c.marchetti@ebri.it; For intrabodies: Simonetta LISI: simonetta.lisi@sns.it, Martina GORACCI: martina.goracci@sns.it
- Information on most suitable data format, most useful experimental protocols for model implementation.
- A collaboration already exists with the group of A. BARBERIS at the Italian Institute of Technology in Genova who is interested in exploring, using quantum dots, how selective intrabodies against gephyrin and NLGs affect GABA receptor trafficking at inhibitory synapses. Further collaboration with scientists working on scaffold and/or synaptic adhesion molecules to better understand their function in synapses formation and stabilization is envisaged.

Data aggregation, analysis and dissemination

- Contacts:
 - A: David STERRATT: david.c.sterratt@ed.ac.uk>
 - B: Oksana SOROKINA: oksana.sorokina@ed.ac.uk
 - C: David STERRATT: david.c.sterratt@ed.ac.uk>
 - D: Colin MCLEAN: cmclean5@staffmail.ed.ac.uk
- What types of feedback are most helpful/constructive?
 - A: Discussion of what other proteins should be in the list.
 - B: Interaction with the users of A to refine how we exploit the larger protein dataset in simulations. Note that the larger protein set contains proteins that are less relevant for simulation but cover the majority of human neurological disease genes.
 - C: Testing the code. Comments on the algorithm. Suggestions for improving the code. Patches and pull requests.
 - D: The application of clustering analysis is more specific to the molecular groups in SP1 and potentially SP2
- We are looking to develop models of synaptic plasticity that are at a detailed molecular level.

Annex A: References

- Althof, D., Baehrens, D., Watanabe, M., Suzuki, N., Fakler, B., and Kulik, Á. *Inhibitory and excitatory axon terminals share a common nano-architecture of their Cav2.1 (P/Q-type) Ca(2+) channels*. *Frontiers in Cellular Neuroscience*, 2015. **11**:315. doi: 10.3389/fncel.2015.00315. eCollection 2015.
- Harris, K.M., Spacek, J., Bell, M.E., Parker, P.H., Lindsey L.F., Baden, A.D., Vogelstein, J.T., and Burns, R. A resource from 3D electron microscopy of hippocampal neuropil for user training and tool development. *Scientific Data*, 2015. **2**:150046. doi: 10.1038/sdata.2015.46. eCollection 2015.
- Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Vázquez-Reina, A., Kaynig, V., Jones, T.R., Roberts, M., Morgan, J.L., Tapia, J.C., Seung, H.S., Roncal, W.G., Vogelstein, J.T., Burns, R., Sussman, D.L., Priebe, C.E, Pfister, H., and Lichtman, J.W. Saturated reconstruction volume of a neocortex. *Cell*, 2015. **162**:648-661.
- Kirizis, T., Kerti-Szigeti, K., Lorincz, A., and Nusser, Z. *Distinct axo-somato-dendritic distributions of three potassium channels in CA1 hippocampal pyramidal cells*. *European Journal of Neuroscience*, 2014. **39**:1771-83. doi: 10.1111/ejn.12526. Epub 2014 Mar 7.
- Li, A., Gong, H., Zhang, B., Wang, Q., Yan, C., Wu, J., Liu, Q., Zeng, S., and Luo, Q. *Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain*. *Science*, 2010. **330**:1404-1408.
- Linnarsson, S. *NEUROSCIENCE. A tree of the human brain*. *Science*, 2015. **350**:37. doi: 10.1126/science.aad2792.
- Mansouri, M., Kasugai, Y., Fukazawa, Y., Bertaso, F., Raynaud, F., Perroy, J., Fagni, L., Kaufmann, W.A., Watanabe, M., Shigemoto, R., and Ferraguti, F. *Distinct subsynaptic localization of type 1 metabotropic glutamate receptors at glutamatergic and GABAergic synapses in the rodent cerebellar cortex*. *European Journal of Neuroscience*, 2015. **41**:157-67. doi: 10.1111/ejn.12779. Epub 2014 Nov 6
- McLean, C., He, X., Simpson, I.T., and Armstrong, J.D. *Improved Functional Enrichment Analysis of Biological Networks using Scalable Modularity Based Clustering*. *Journal of Proteomics & Bioinformatics*, 2016. **9**:009-018, doi:10.4172/jpb.1000383.
- Nakamura, H., Hioki, H., Furuta, T., and Kaneko, T. *Different cortical projections from three subdivisions of the rat lateral posterior thalamic nucleus: a single-neuron tracing study with viral vectors*. *European Journal of Neuroscience*, 2015. **41**:1294-310. doi: 10.1111/ejn.12882.
- Sterratt, D.C., Sorokina, O. and Armstrong, J.D. Hybrid Systems Biology: Second International Workshop, HSB 2013, Taormina, Italy, September 2, 2013 and Third International Workshop, HSB 2014, Vienna, Austria, July 23-24, 2014, Revised Selected Papers, vol. 7699 of Lecture Notes in Bioinformatics, chap. Integration of Rule-Based Models and Compartmental Models of Neurons, pp. 143-158. Springer International Publishing, Cham, 2015. Preprint at <http://adsabs.harvard.edu/abs/2014arXiv1411.4980S>
- Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., Bertagnolli, D., Goldy, J., Shapovalova, N., Parry, S., Lee, C., Smith, K., Bernard, A., Madisen, L., Sunkin, S.M., Hawrylycz, M., Koch, C., and Zeng, H. *Adult mouse cortical cell taxonomy revealed by single cell transcriptomics*. *Nature Neuroscience*, 2016. **19**:335-46. doi: 10.1038/nn.4216. Epub 2016 Jan 4.



Xue, S., Gong, H., Jiang, T., Luo, W., Meng, Y., Liu, Q., Chen, S., and Li, A. *Indian-ink perfusion based method for reconstructing continuous vascular networks in whole mouse brain*. PLoS One, 2014. **9**: e88067.

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., Rolny, C., Castelo-Branco, G., Hjerling-Leffler, J., and Linnarsson, S. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA seq. Science, 2015. **347**:1138-1142.



Annex B: Dataset Information Cards

Task	Partner	Data / model name	DIC name	DIC registered	Link
1.1.2	UOXF	First Draft Transcriptomes of Major Neocortical, Hippocampal and Cerebellar Cell Types	Single cell transcriptomes from mouse hippocampus, cerebellum, and cortex	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/7/
1.1.3	UEDIN	Distribution of major synaptic proteins- 7 regions, synapse proteome maps	Mouse Synaptic Proteome	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/9/
1.1.3	UEDIN	First Synapse Proteome for 4 Specific Cell Types	Mouse Synaptic Proteome	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/9/
1.1.4	UCLM	FIB/SEM immunogold technique data sets, pyramidal cells of hippocampus	Neural channelomics and receptomics	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/111/
1.1.4	UCLM	SDS-FRL technique data sets, pyramidal cells of hippocampus and cerebellum (extra data set)			
1.1.5	SIB	Potassium channel kinetics	Homology models of K channels	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/151/
			K Channel Functional Annotation	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/150
			K Channels Kinetics	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/101
1.1.6	CNR		Kinetic scheme of inhibitory events	yes	https://dataset-information-



			including gephyrin		manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/275
			Action potential dependent sIPSCs - from 3-4 months old Tg2576 female mice	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/78/
			Action potential dependent sIPSCs - from juvenile (P21-30) C57Bl6/J male mice - from CA1 pyramidal neurons receiving input from PV+ and CCK+ interneurons	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/77/
			Action potential dependent sIPSCs - from juvenile (P21-30) C57Bl6/J male mice - from CA1 pyramidal neurons receiving input from PV+ interneurons	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/76/
			Action potential dependent sIPSCs - from 3-4 months old wt (B6/SJL) female mice	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/75/
			Action potential dependent sIPSCs - from juvenile (P21-30) C57Bl6/J male mice (E)	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/74/?_changelist_filters=o%3D4
1.2.1	UZH	High-resolution 3D maps of the brain vasculature	Cortex Vasculature Rat	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/139/
1.2.1	UZH	3-D high-resolution angiographic datasets obtained from synchrotron radiation X-ray tomographic microscopy (SRXTM) of cylindrical samples of the rat somatosensory cortex	SRXTM Rat Cortex	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/138
1.2.1	UZH	Whole mouse brain			
1.2.2	LENS	Reconstruction of the hippocampus of a thy1-GFP-M transgenic mouse, where a subset of	Hippocampus reconstruction	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog



		pyramidal neurons and of interneurons are labeled			og/datasetinformationcard/16/?_changelist_filters=o%3D4
		3D localization of parvalbumin positive neurons in whole mouse brain	Parvalbumin interneurons distribution	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/15/?_changelist_filters=o%3D4
		3D localization and counting of Purkinje cells of the mouse cerebellum	Purkinje cells distribution	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/14/?_changelist_filters=o%3D4
			Arc-dVenus	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/268/?_changelist_filters=p%3D1%26o%3D3
1.2.3	UPM	Six data sets have been generated (including data from every cortical layer), density and distribution of different types of cells in the somatosensory cortex in fixed brains	Cell Number Counting	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/33/?_changelist_filters=o%3D4
1.2.3	UPM	3D distribution of GABAergic cells	Not available yet (extra data)		
1.2.4	IEM HAS	Anatomical Reconstructions of Hippocampal Cells	Morphological reconstructions of mouse hippocampal neurons	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/85/
1.2.4	WMC	Anatomical Reconstructions of Neocortical Cells	3D reconstructions of cortical neurons in primary visual and motor cortices from brain slices, mouse	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/174/
			3D reconstructions of cortical neurons from cortic tissues of human brain obtained from brain surgery in hospitals	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/173
			3D reconstructions of cortical neurons from brain slices, rat	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/172



			3D reconstructions of cortical neurons at a whole-brain-wide, mouse	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/169
1.2.5	UAM	Single-cell Thalamocortical and Corticocortical Projection Neurons	Thalamocortical cells full 3D reconstructions	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/94/
1.2.5	UAM	Bulk-micropopulation mapping of thalamocortical connections in mice	Thalamocortical cells full 3D reconstructions	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/94/
1.2.6	UPM	Synaptic maps of identified neurons	Synapse maps	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/31/?_changelist_filters=o%3D4
1.2.6	UPM	High resolution optical synaptic maps and electron microscope blocks scans and volume analysis of exemplar brain regions			
1.3.1	UEDIN	Data sources and tools for molecular and cellular informatics	List of Synaptic Target Proteins	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/105
1.3.1	UEDIN	Data sources and tools for molecular and cellular informatics	Curated data from neuroproteomics publications	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/274/?_changelist_filters=o%3D4
1.3.2	UEDIN	Data sources and tools for molecular and cellular informatics	Community Detection Modularity Suite	yes	https://dataset-information-manager.herokuapp.com/admin/dataset_information_catalog/datasetinformationcard/130