



**Laboratory of Functional
and Structural Genomics**

<http://nucleus3d.cent.uw.edu.pl>

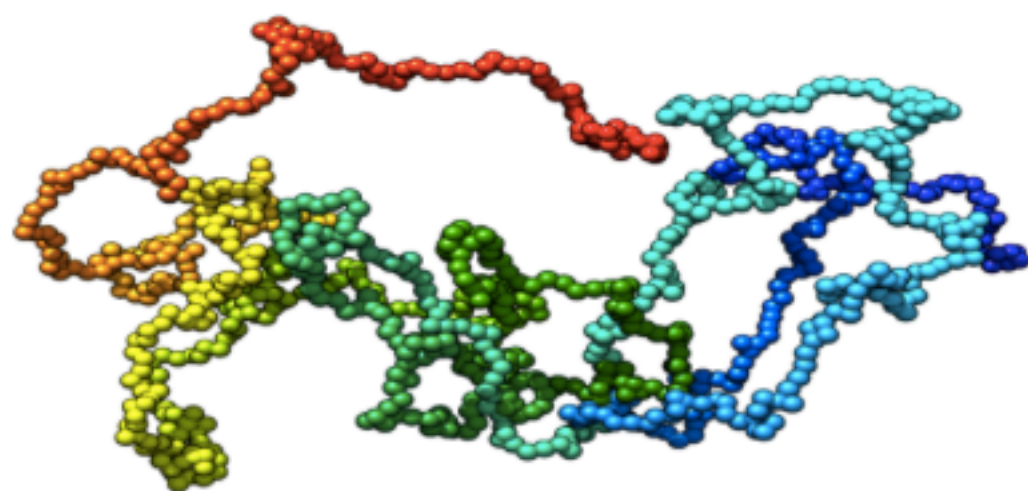
kontakt:

Profesor Dariusz Plewczyński d.plewczynski(at)cent.uw.edu.pl

Mgr Paulina Urban p.urban(at)cent.uw.edu.pl

Mgr Michał Denkiewicz m.denkiewicz(at)cent.uw.edu.pl

Ania Bugaj a.bugaj(at)cent.uw.edu.pl



STRUCTURAL GENOMICS

Modeling of the 3D structure of chromatin inside the cell nucleus based on the ChIA-PET and Hi-C data.



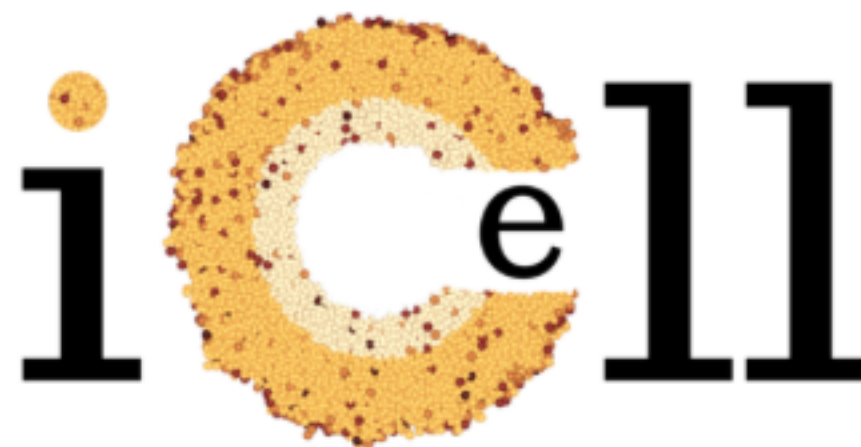
FUNCTIONAL GENOMICS

Studying how different genomic regions interact and cooperate together in order to gain better understanding of the genotype-phenotype relationship.



BIOSTATISTICS AND COGNITIVE COMPUTING

Field of expertise: prediction, machine learning, statistics, computational methods, data preparation.



BIOLOGICAL SYSTEMS MODELING

Within the framework of iCell project we develop the novel computational modelling framework for multiscale analysis of tumor growth.

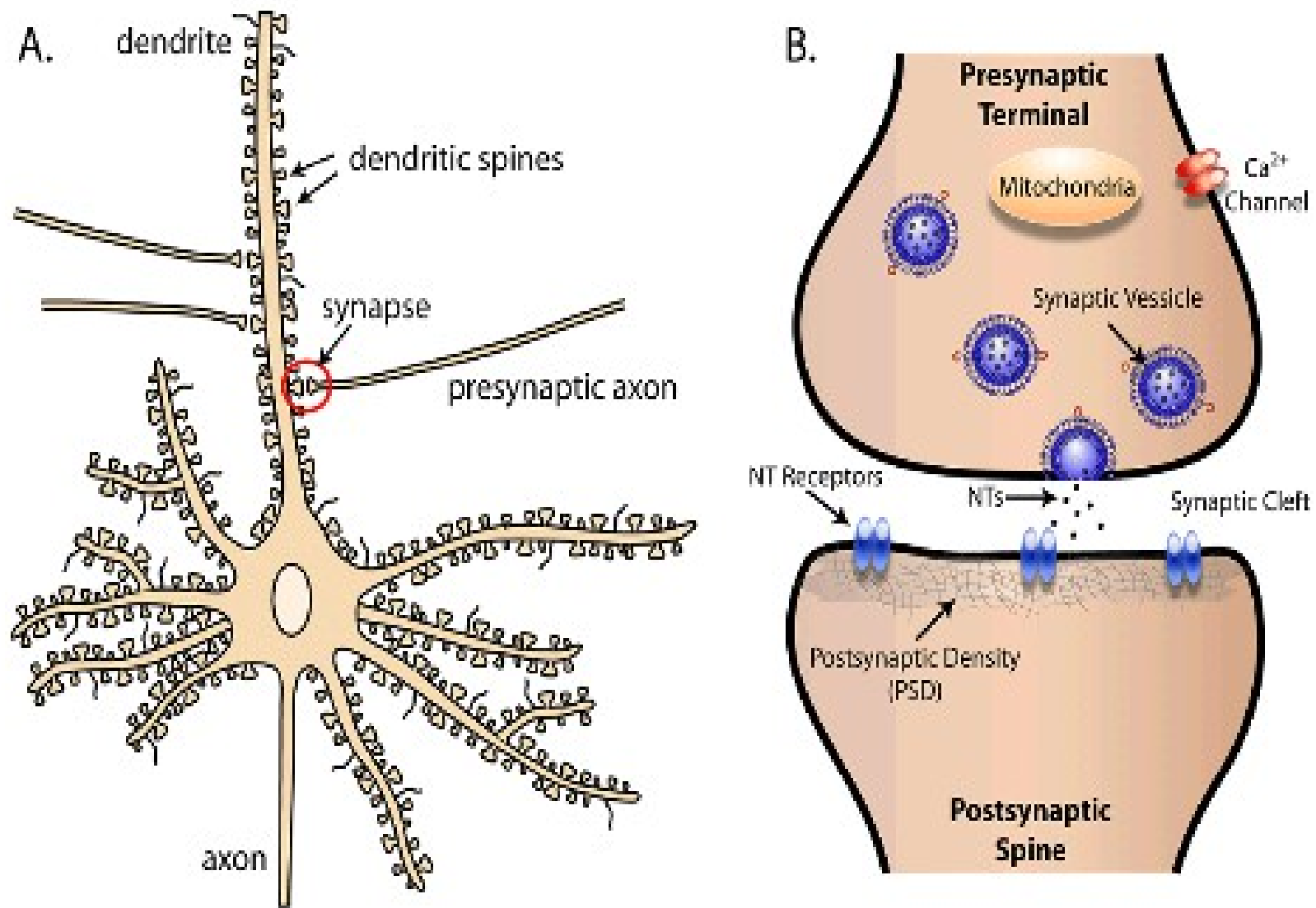
KOLCE DENDRYTYCZNE

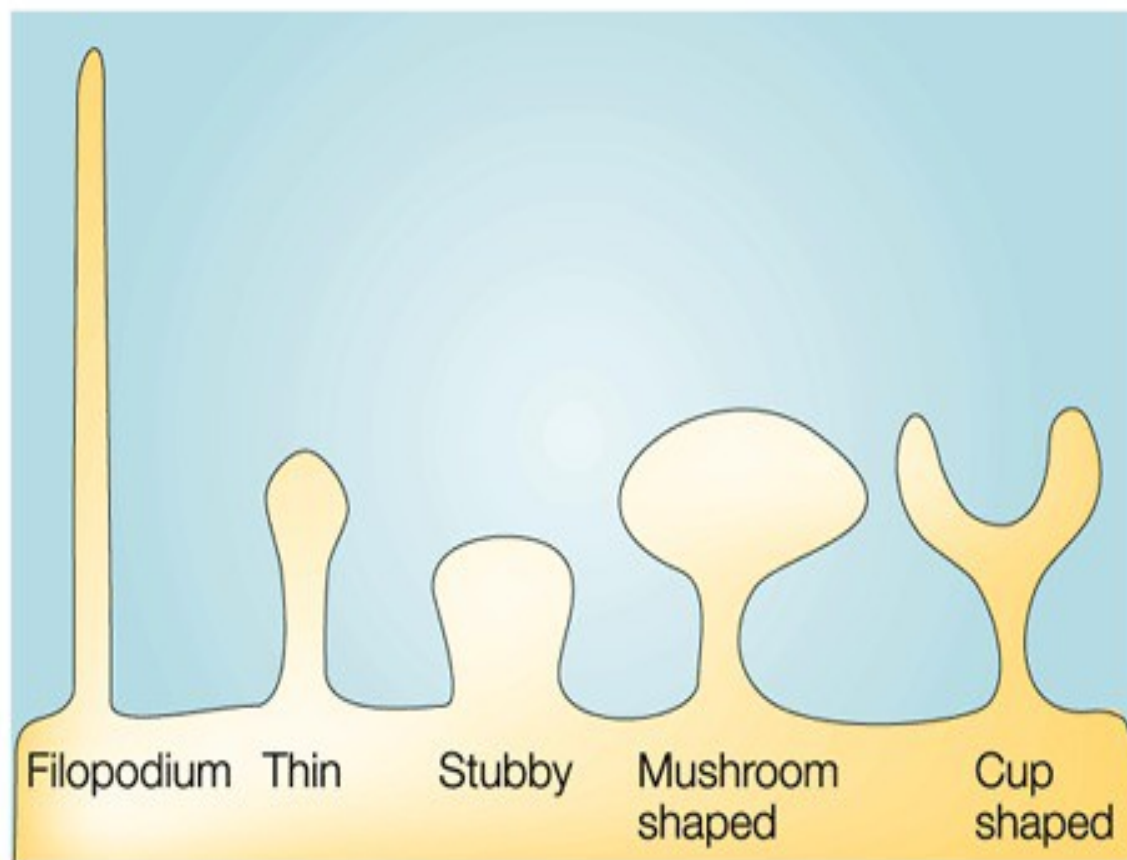
Analiza danych

Paulina Urban

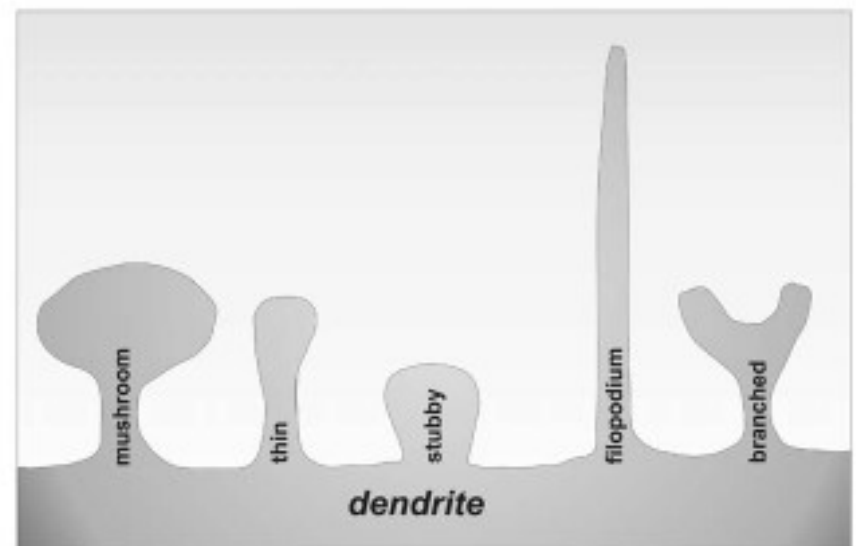
email:p.urban@cent.uw.edu.pl

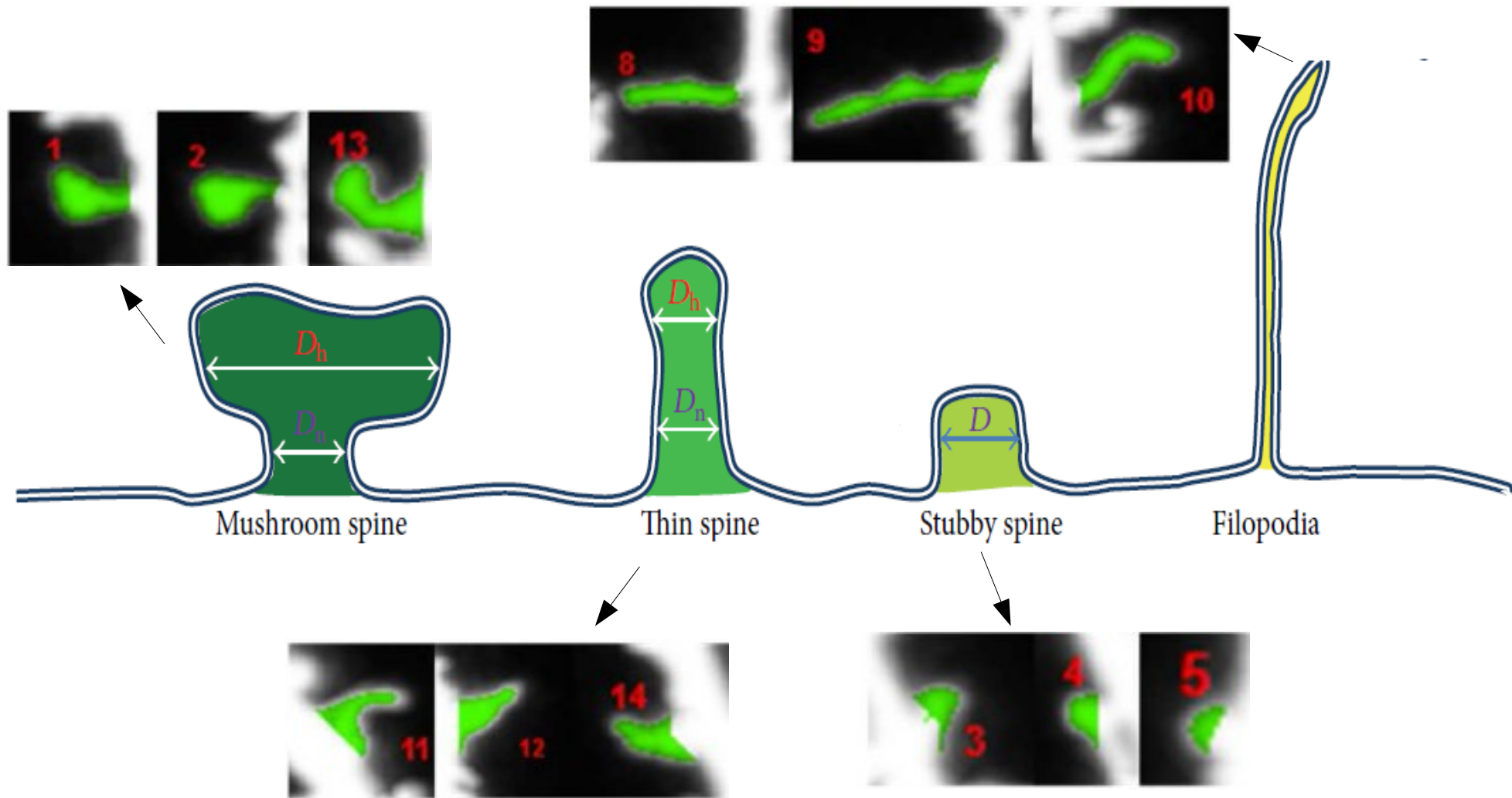
Background: Dendrites and dendritic spines

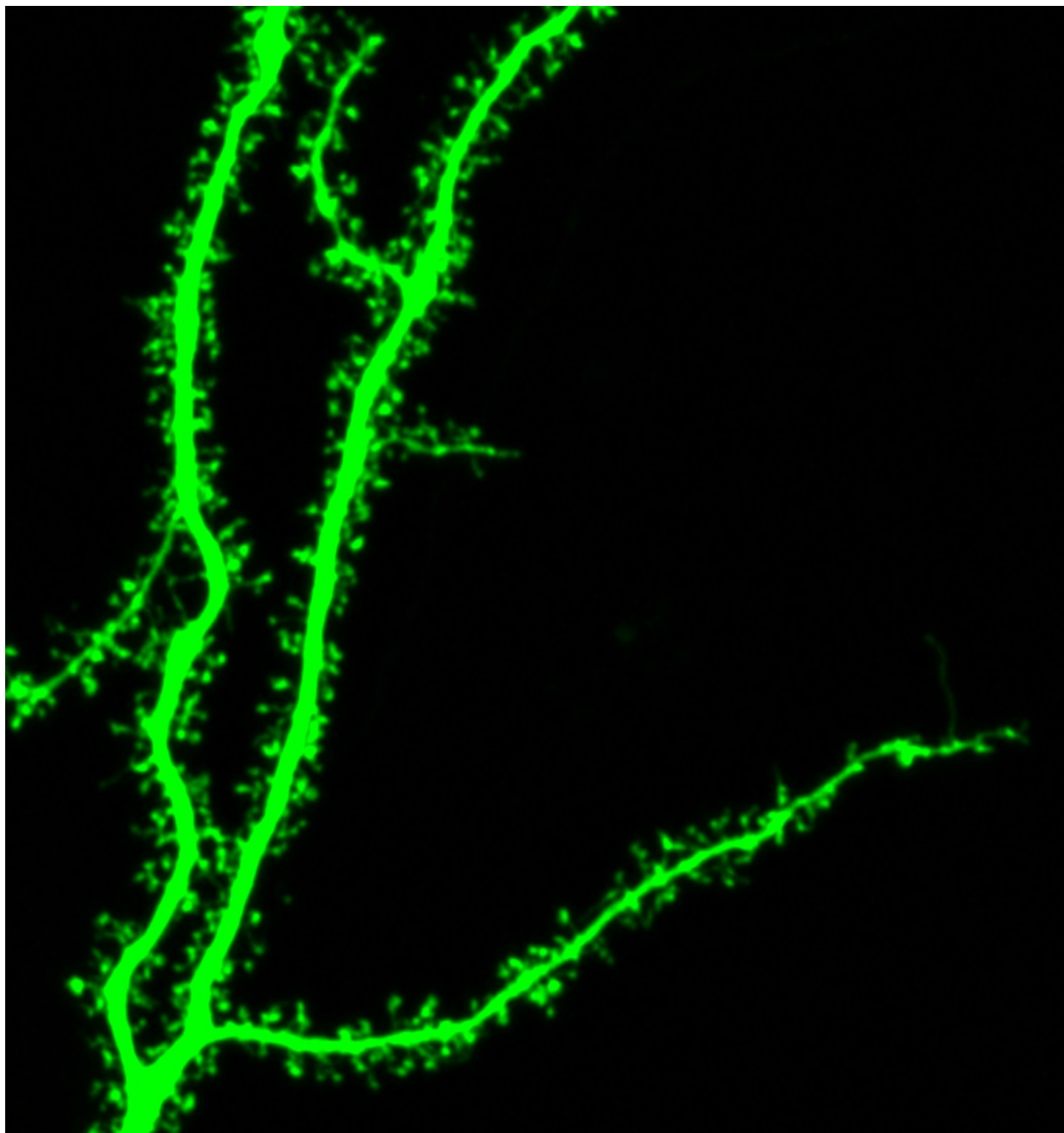


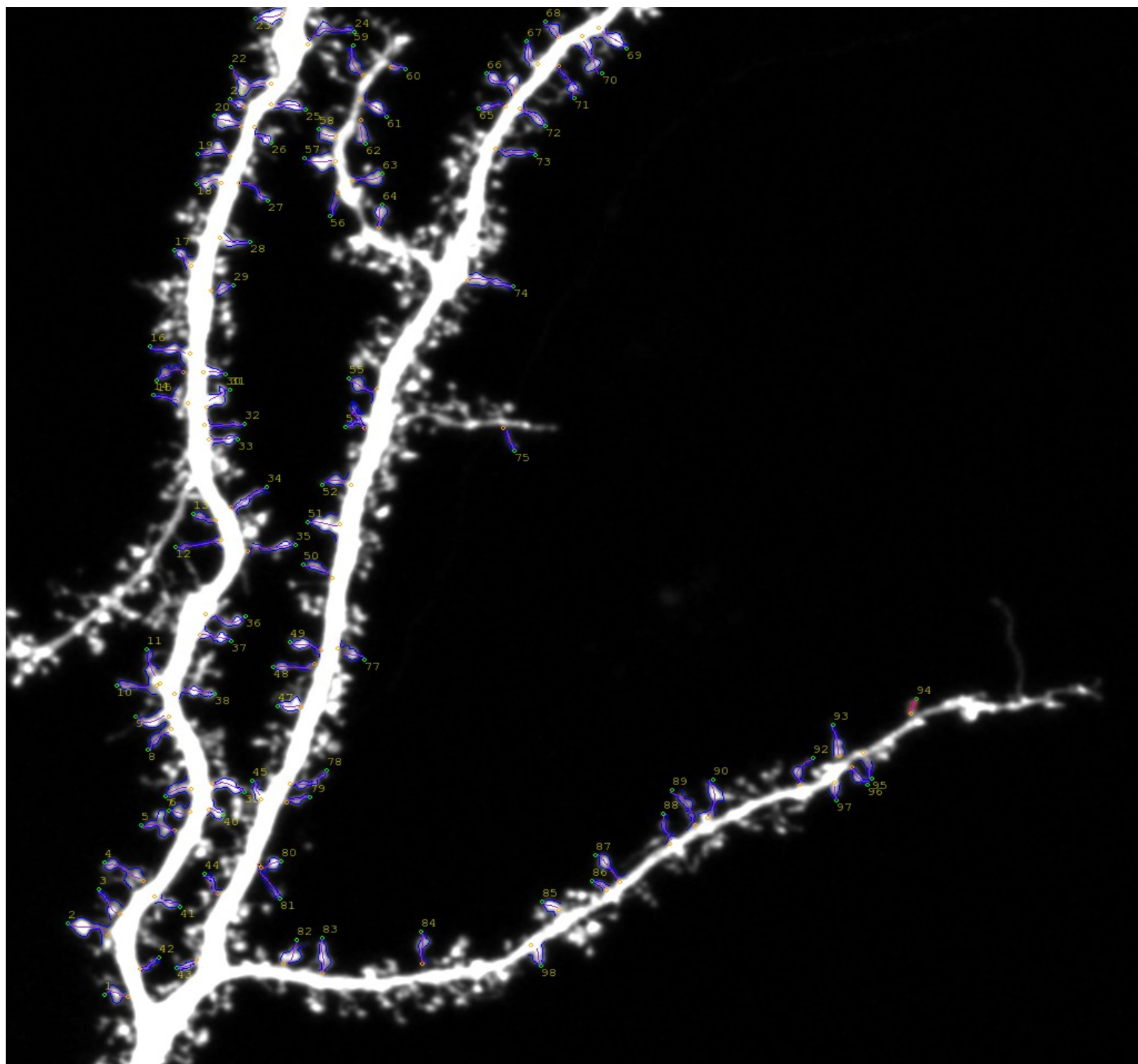


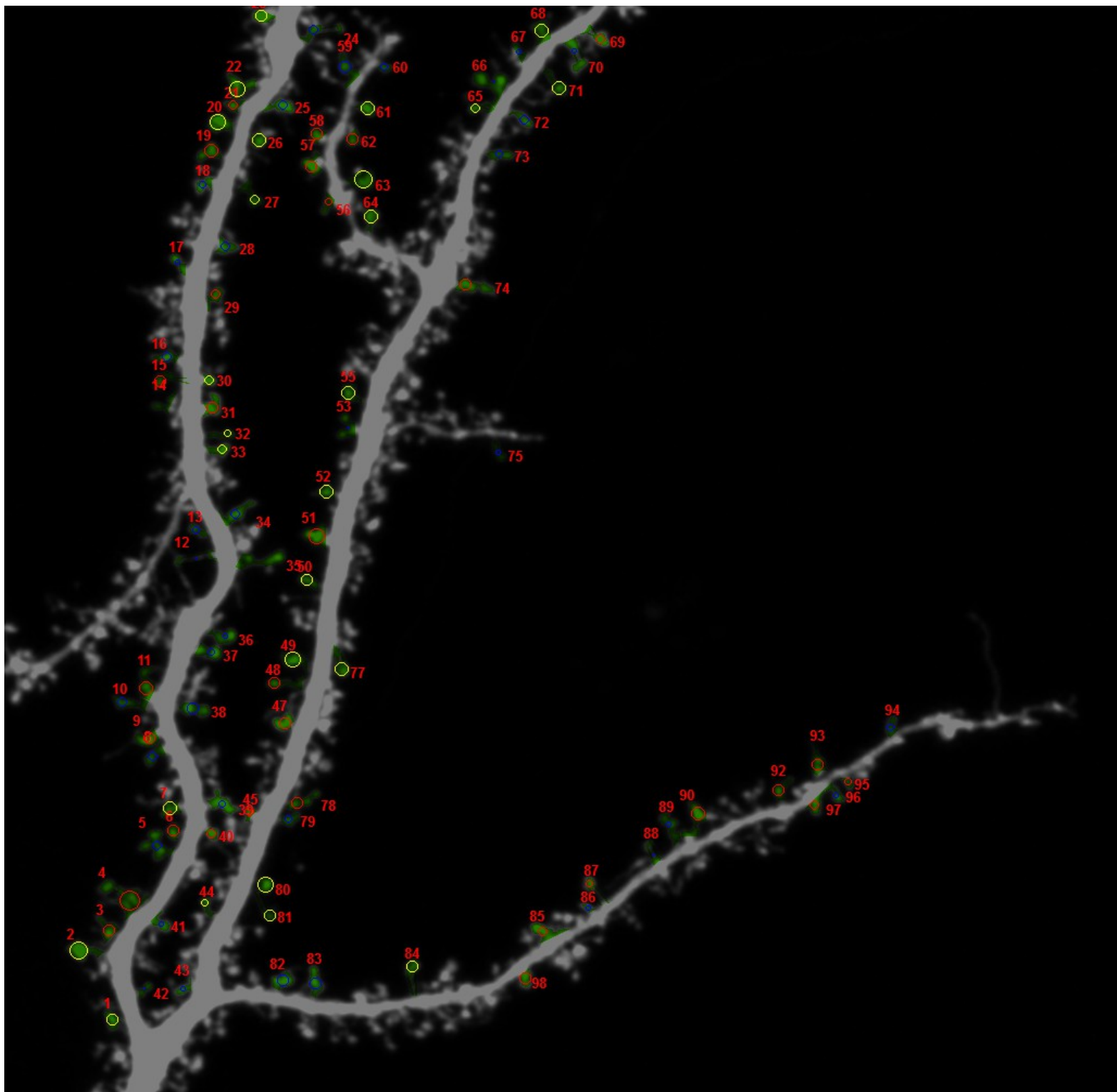
Nature Reviews | **Neurosci**



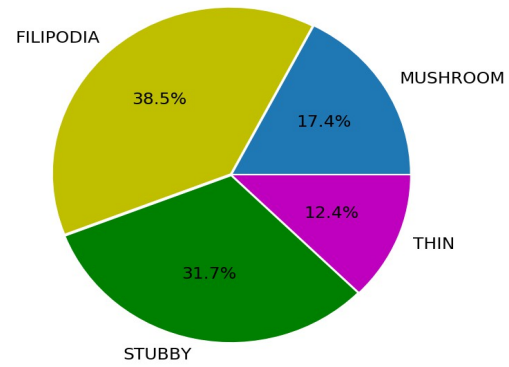
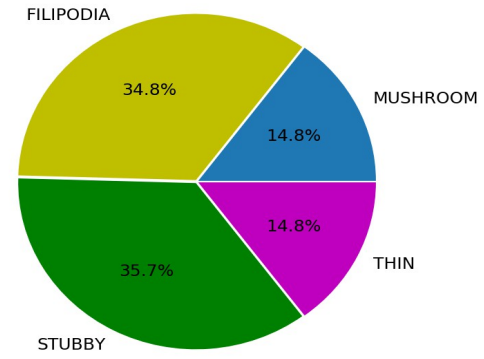
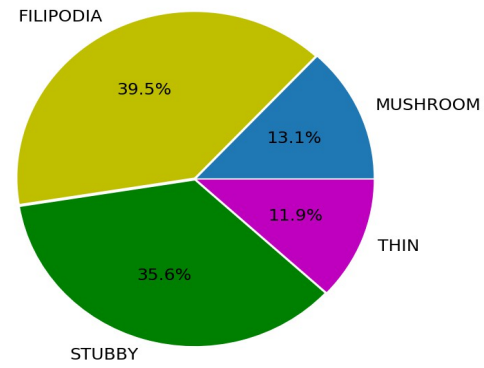
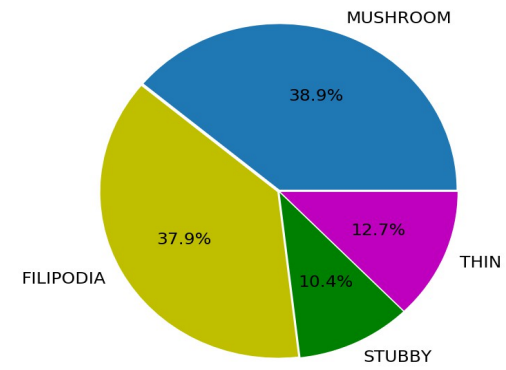
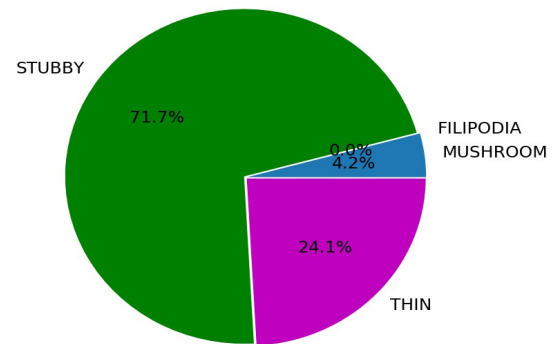


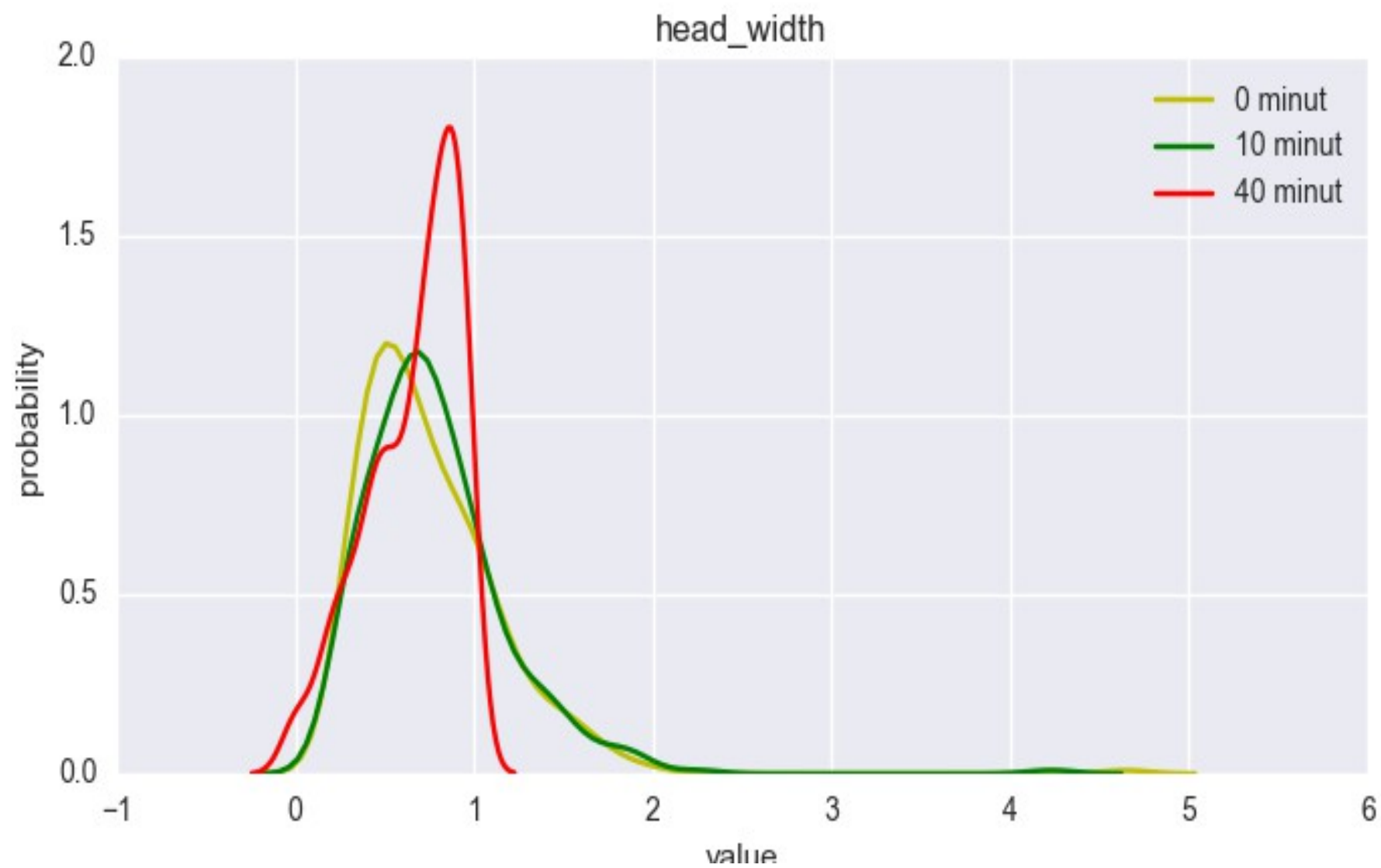


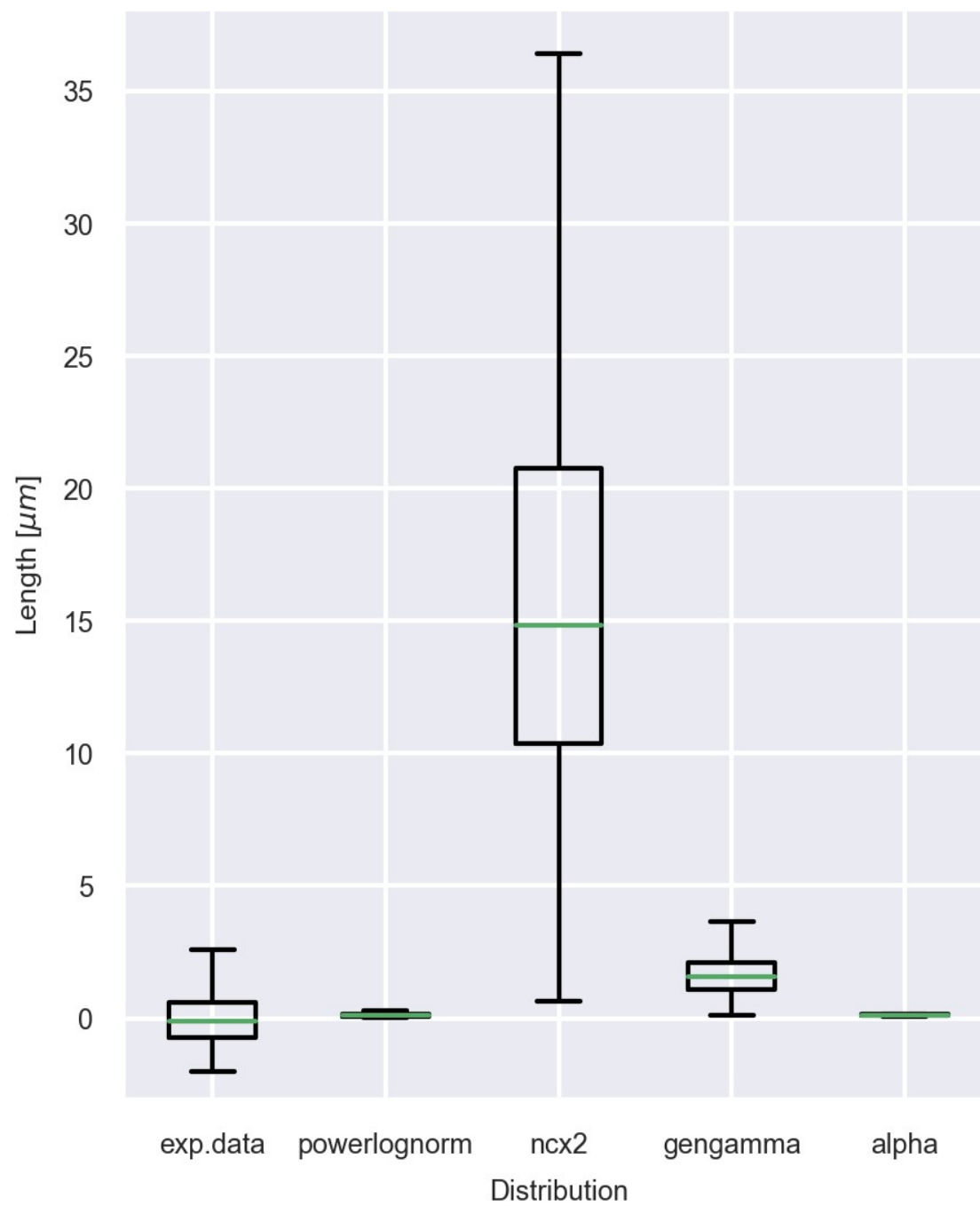




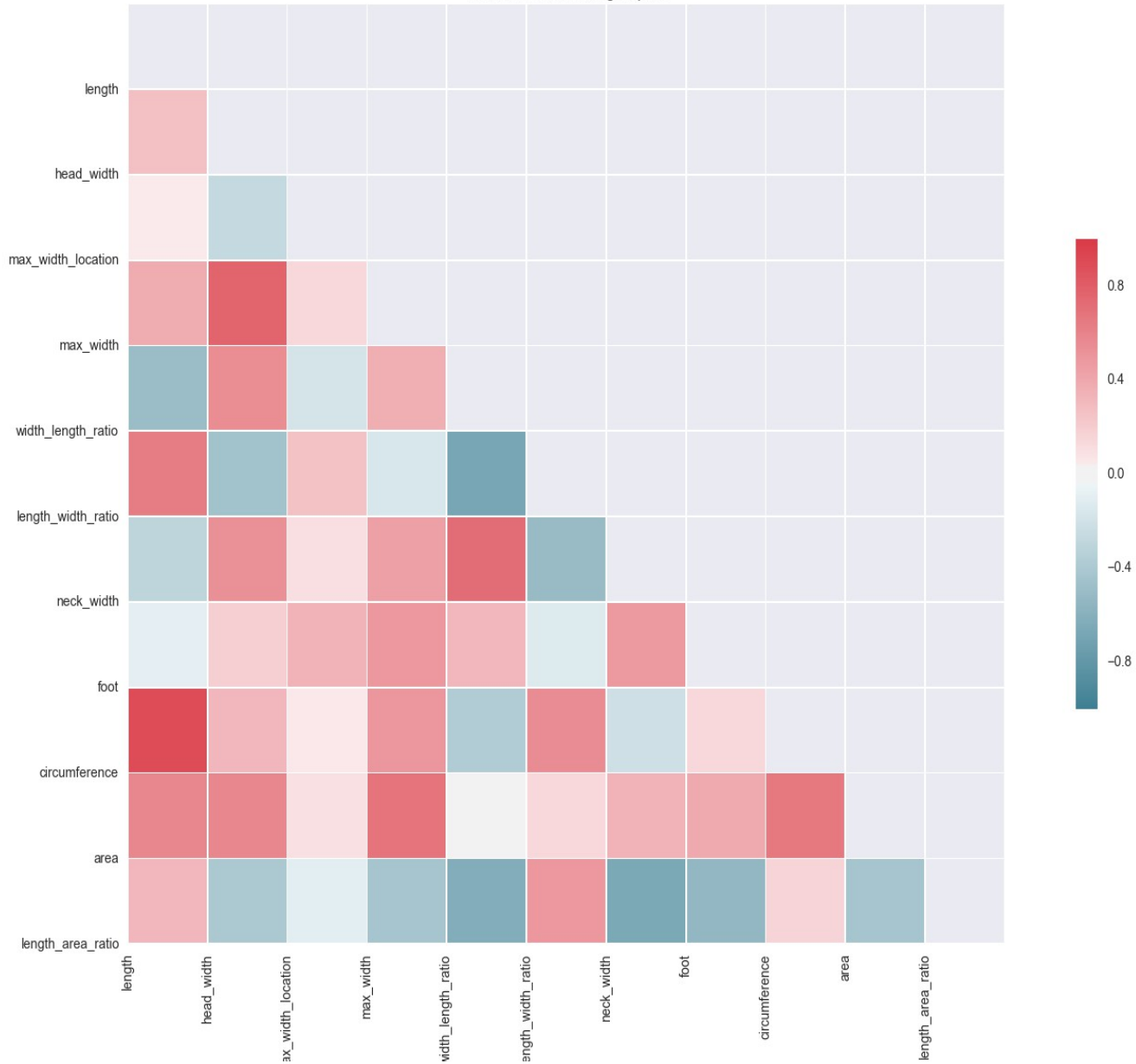
	A	B
1		0
2	0	(1, 4.56262989365, 1.90125992293, 0.189655172414, 1.90125992293, 0.416702640197, 2.39979281035, 0.376961536499, 0.703566617817, 11.5455338146, 2.57423593714, 1.77242102319, 'pSuper_1')
3	1	(1, 0.880828194368, 0.673361870308, 0.583333333333, 0.886484599141, 0.764464483101, 1.30810524505, 0.673361870308, 0.91, 3.67399870513, 0.576221359934, 1.52862815511, 'pSuper_2')
4	2	(1, 1.35865487786, 1.16666713201, 0.529411764706, 1.16666713201, 0.858692778442, 1.16456085937, 0.989949493661, 1.04062481231, 5.38553560008, 1.14201332792, 1.18970141998, 'pSuper_3')
5	3	(1, 0.983958952031, 0.576109308887, 0.928571428571, 0.667757440992, 0.585501364359, 1.70793795006, 0.576109308887, 0.800690226907, 3.46225486424, 0.501791749489, 1.96089105298, 'pSuper_4')
6	4	(1, 2.57463418484, 0.870337398674, 0.147058823529, 0.870337398674, 0.338043130088, 2.95820240375, 0.252388589282, 0.744110206891, 7.85243484035, 1.41668775528, 1.81736178296, 'pSuper_5')
7	5	(1, 2.61999299647, 0.727088369949, 0.942857142857, 0.728951307798, 0.277515386846, 3.60340380173, 0.14, 0.798122797569, 8.73267093383, 1.54109833704, 1.70008164534, 'pSuper_6')
8	6	(1, 2.14364474394, 0.975178936112, 0.241379310345, 0.975178936112, 0.454916300319, 2.19820656965, 0.0989949493661, 0.527103493884, 6.04072056108, 1.06933066933, 2.00466030333, 'pSuper_7')
9	7	(1, 1.41525397035, 0.89695911496, 0.3, 0.89695911496, 0.633779613944, 1.57783554093, 0.395979797464, 1.33386143261, 6.75759899045, 1.02443516705, 1.38149686371, 'pSuper_8')
10	8	(1, 1.49503858339, 0.551667414232, 0.590909090909, 0.551667414232, 0.368998780607, 2.71003605582, 0.221359436212, 0.31304951685, 4.82473370912, 0.652778931521, 2.29026782453, 'pSuper_9')
11	9	(1, 1.45883778697, 1.09621087698, 0.277777777778, 1.09621087698, 0.751427531403, 1.330800321, 0.541385194188, 0.490692259483, 5.01848390584, 0.974461684084, 1.49707044494, 'pSuper_10')
12	10	(1, 1.39326043897, 0.845793065403, 0.736842105263, 0.99006959973, 0.607060275125, 1.64728288273, 0.513466580662, 0.92153507028, 5.82253870729, 1.12608941596, 1.23725560264, 'pSuper_11')
13	11	(1, 1.33582412945, 1.10602483066, 0.444444444444, 1.10602483066, 0.827971891119, 1.2077704699, 0.389026553285, 0.442718872424, 4.84289920907, 0.962918348528, 1.38726625315, 'pSuper_FAKCy3_psd95_647_1')
14	12	(1, 1.99647690473, 1.06127984247, 0.444444444444, 1.06127984247, 0.531576318243, 1.88119742299, 0.417820276486, 0.376961536499, 6.28548340845, 1.27213033754, 1.56939650429, 'pSuper_FAKCy3_psd95_647_2')
15	13	(1, 1.58628044747, 0.821880160774, 0.428571428571, 0.821880160774, 0.518117815852, 1.93006294978, 0.156524758425, 0.221359436212, 4.83790412539, 0.738634096025, 2.14758627582, 'pSuper_FAKCy3_psd95_647_3')
16	14	(1, 1.94140391406, 1.06951830381, 0.28, 1.06951830381, 0.550899426989, 1.81521336021, 0.42, 0.42, 4.70988949884, 1.28622534314, 1.5093808596, 'pSuper_FAKCy3_psd95_647_4')
17	15	(1, 1.61136478137, 0.753936259198, 0.434782608696, 0.753936259198, 0.467886767736, 2.13726924751, 0.14, 0.168471028356, 4.25825819323, 0.801992954551, 2.00920066969, 'pSuper_FAKCy3_psd95_647_5')
18	16	(1, 1.97104018368, 0.646920790769, 0.192307692308, 0.646920790769, 0.328212887858, 3.04680296537, 0.174912769949, 0.425793377121, 5.63724677055, 0.79152021224, 2.49019564276, 'pSuper_FAKCy3_psd95_647_6')
19	17	(1, 2.37522767851, 0.66426936779, 0.323529411765, 0.66426936779, 0.279665555348, 3.57569954854, 0.235475552175, 0.577234787586, 6.99620906845, 1.07050068143, 2.21880071607, 'pSuper_FAKCy3_psd95_647_7')
20	18	(1, 2.12865464776, 0.855014343035, 0.275862068966, 0.855014343035, 0.401668886935, 2.48961279434, 0.14, 0.14, 5.0156211292, 0.883679872274, 2.4088527017, 'pSuper_FAKCy3_psd95_647_8')
21	19	(1, 0.900705668256, 1.02698559571, 0.307692307692, 1.02698559571, 1.14020110221, 0.877038268128, 0.74838010564, 0.442718872424, 3.27078550532, 0.659833505531, 1.36504991139, 'pSuper_FAKCy3_psd95_647_9')
22	20	(1, 1.82010818046, 0.974987179403, 0.653846153846, 1.1038408167, 0.535675400984, 1.86680216819, 0.221359436212, 0.31304951685, 6.31129455492, 1.17707427422, 1.54629849647, 'pSuper_FAKCy3_psd95_647_10')
23	21	(1, 2.26978939725, 0.698392118163, 0.666666666667, 0.909105328697, 0.307690272502, 3.25002149684, 0.324143400116, 0.691197782114, 6.65401715552, 1.35794210448, 1.67149202441, 'pSuper_Glur1_Cy3_NR1_647_1')
24	22	(1, 1.27835676729, 0.658722279252, 0.352941176471, 0.658722279252, 0.515288295181, 1.94066119753, 0.21, 0.28, 3.47907908184, 0.497171677982, 2.5712582271, 'pSuper_Glur1_Cy3_NR1_647_2')
25	23	(1, 1.73344020533, 0.938569207431, 0.333333333333, 0.938569207431, 0.541448850987, 1.8468965225, 0.14, 0.31304951685, 5.61589259404, 0.845283764066, 2.05071986358, 'pSuper_Glur1_Cy3_NR1_647_3')
26	24	(1, 0.911565618171, 0.411222200007, 0.307692307692, 0.411222200007, 0.451116399972, 2.21672277945, 0.31304951685, 0.564358042381, 3.40604573976, 0.340784132763, 2.67490628387, 'pSuper_Glur1_Cy3_NR1_647_4')
27	25	(1, 1.86611864951, 0.77, 0.230769230769, 0.77, 0.412621137568, 2.42353071365, 0.19125363393, 0.650615999894, 5.84303523007, 0.860468590689, 2.16872372764, 'pSuper_Glur1_Cy3_NR1_647_5')
28	26	(1, 1.48935695017, 0.869302595139, 0.380952380952, 0.869302595139, 0.583676461873, 1.71327792934, 0.356931365951, 0.356931365951, 4.96511308048, 0.94190846273, 1.58121198513, 'pSuper_Glur1_Cy3_NR1_647_6')
29	27	(1, 1.25765966337, 0.63, 0.833333333333, 0.7, 0.500930433206, 1.99628517996, 0.63, 0.63, 4.37586129853, 0.716708909921, 1.75477051557, 'pSuper_Glur1_Cy3_NR1_647_7')
30	28	(1, 3.01598229131, 1.05212390936, 0.325581395349, 1.05212390936, 0.348849498351, 2.86656568155, 0.14, 0.156524758425, 7.7896972212, 1.52239884201, 1.98107237609, 'pSuper_Glur1_Cy3_NR1_647_8')
31	29	(1, 2.40722419682, 1.26297391746, 0.548387096774, 1.26297391746, 0.524659863059, 1.90599676173, 0.312723254693, 0.63, 8.53600515765, 1.66599435089, 1.44491738255, 'pSuper_Glur1_Cy3_NR1_647_9')
32	30	(2, 1.4632663924, 0.948480416798, 0.25, 0.948480416798, 0.648193945905, 1.5427481332, 0.14, 0.28, 3.93049974311, 0.72727309139, 2.01199028222, 'pSuper_2')
33	31	(2, 0.615873034813, 0.576504643862, 0.555555555556, 0.628904835649, 0.936077099133, 1.06828807256, 0.504777178565, 0.564358042381, 2.70487759667, 0.336949950924, 1.82778787509, 'pSuper_3')

A**B****C****D****E**



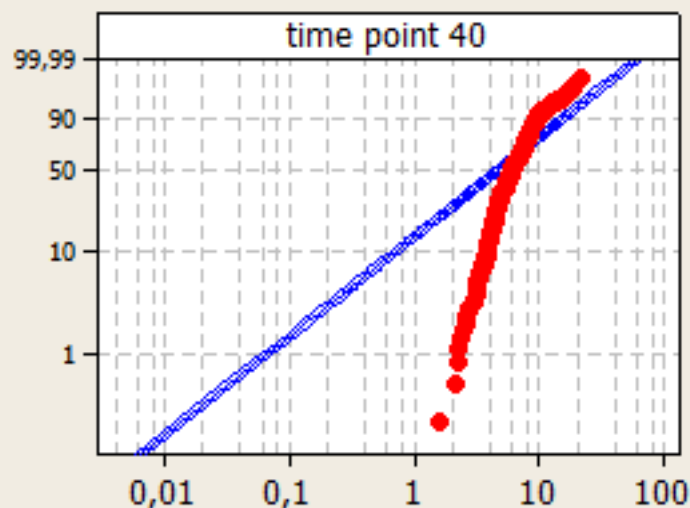
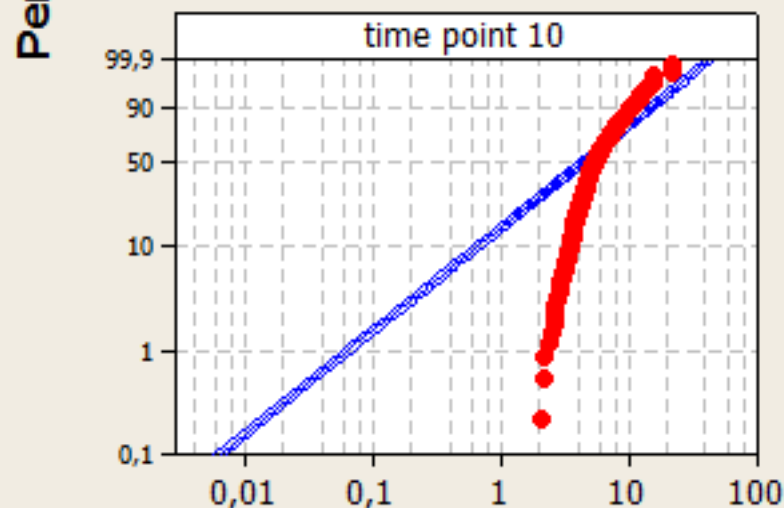
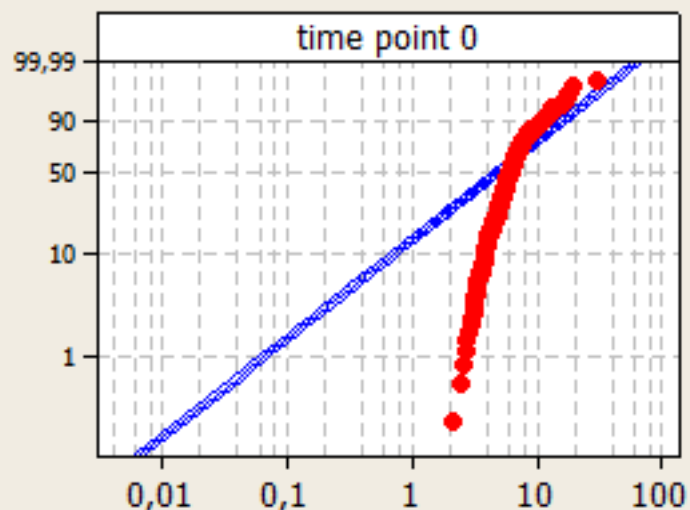
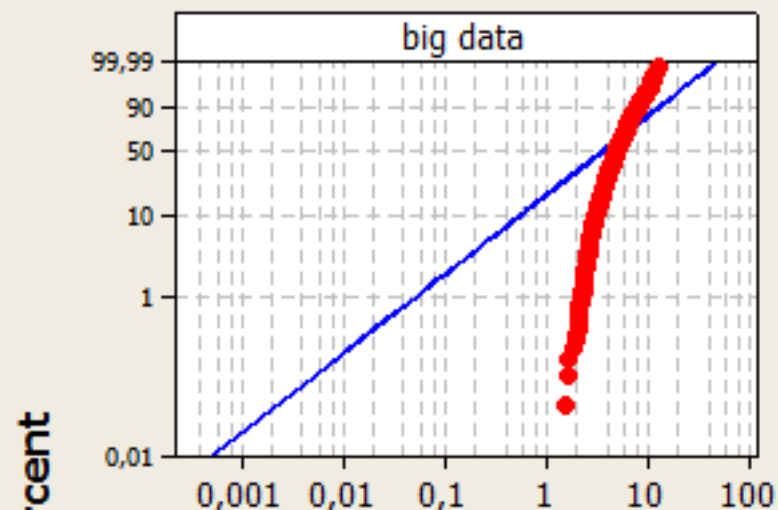


Time 0 data from magic spine

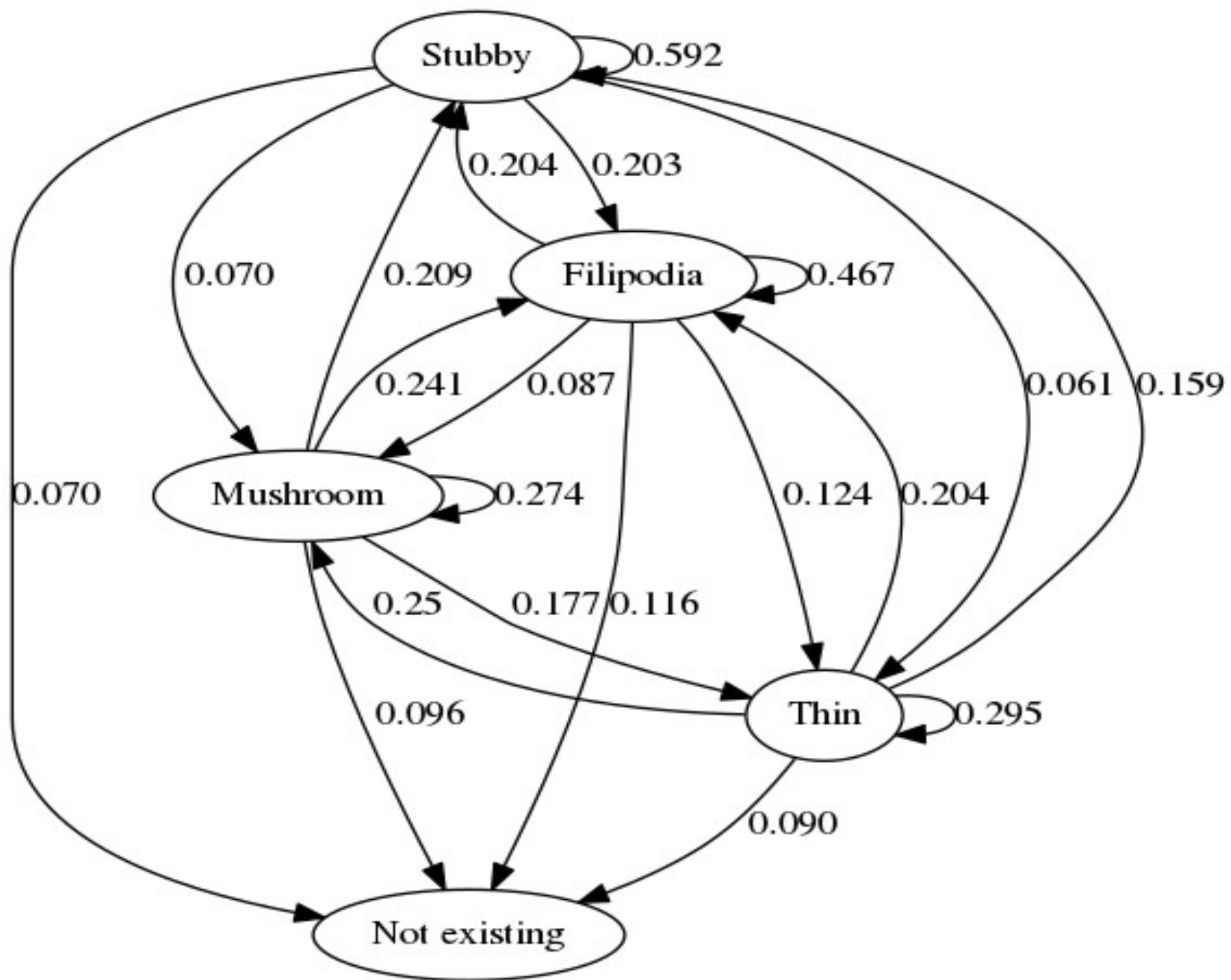


Circumference_Magic spine

Exponential - 95% CI



big data	
Mean	5,344
N	1751
AD	355,864
P-Value	<0,003
time point 0	
Mean	6,712
N	334
AD	58,868
P-Value	<0,003
time point 10	
Mean	6,385
N	314
AD	50,353
P-Value	<0,003
time point 40	
Mean	6,416
N	326
AD	59,230
P-Value	<0,003



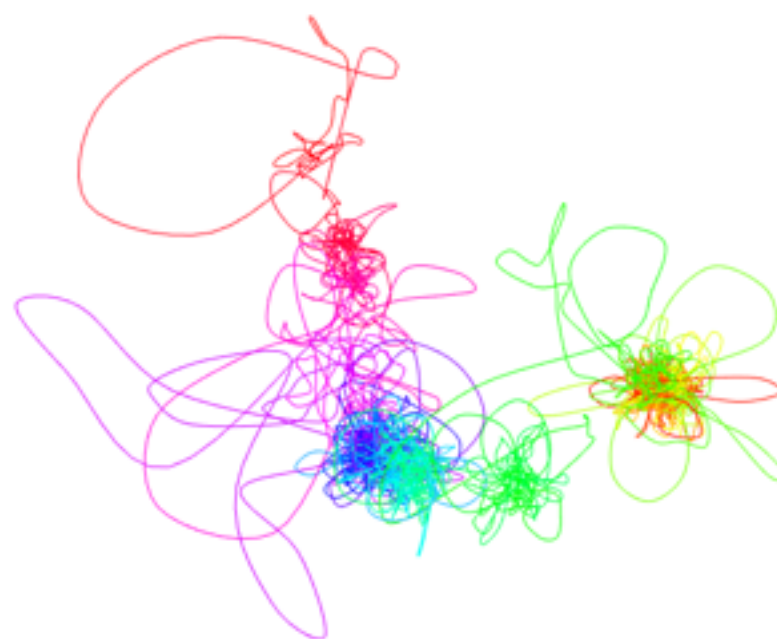
Dla kogo

- Podstawy neurobiologii
- Programowanie Python/Matlab
- Podstawy analizy danych i statystyki
-

Ania Bugaj

kontakt: a.bugaj@cent.uw.edu.pl

Pracownia Badawcza



Depression - Proteomics

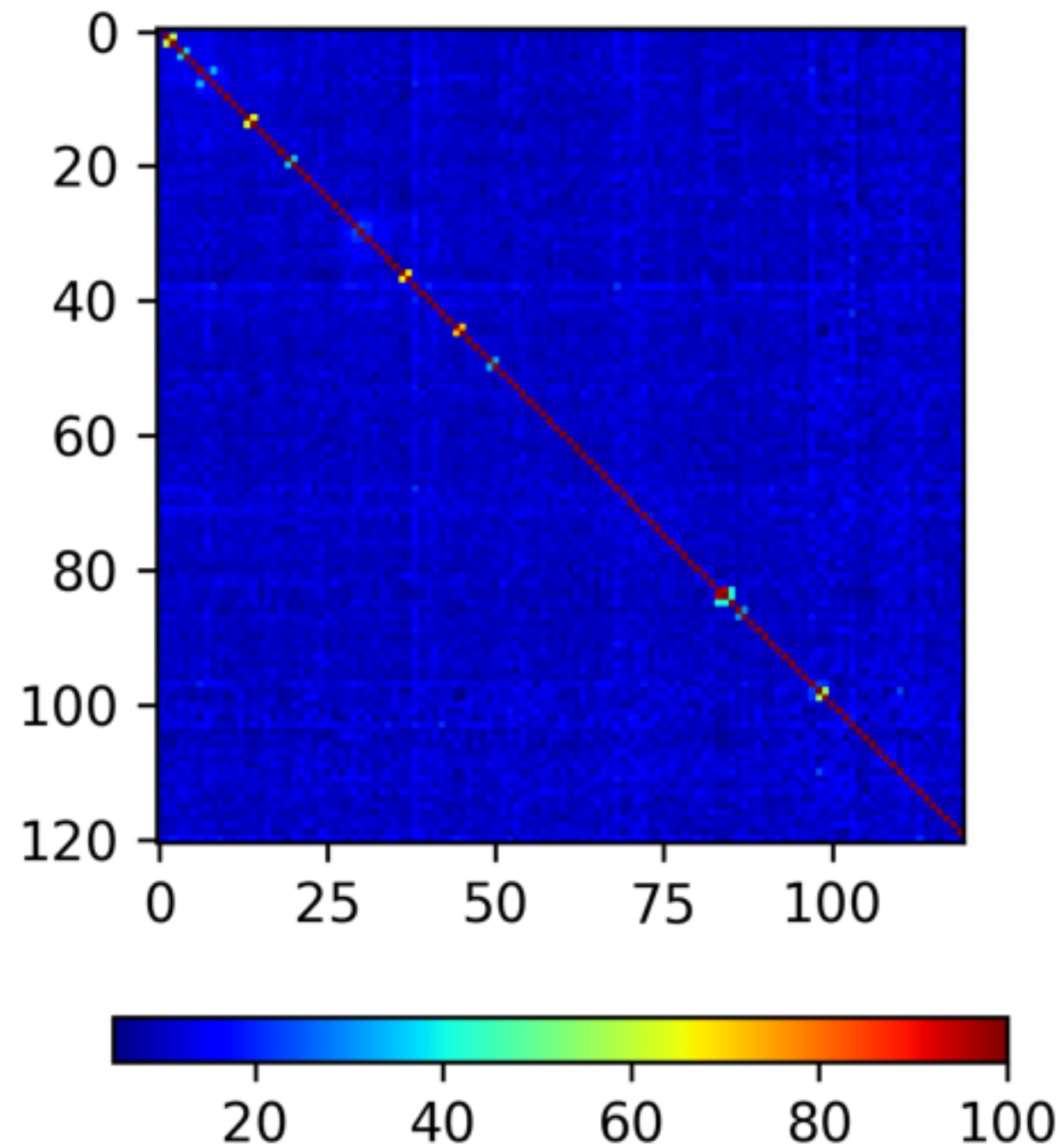
Multiple Sequence Analysis of experimentally derived proteins (IBD PAN)

- (a) FFT-NS-1, FFT-NS-2 — Progressive methods
- (b) FFT-NS-i, NW-NS-i — Iterative refinement method
- (c) L-INS-i, E-INS-i, G-INS-i — Iterative refinement methods using WSP and consistency scores

Analysis - MSA

- Uniprot IDs —> FASTA
- MSA using MAFFT (**M**ultiple **A**lignment using **F**ast **F**ourier **T**ransform)

Results



Analysis - DOMAINS

- CDD - Conserved Domains Database, NCBI
- CDSearch

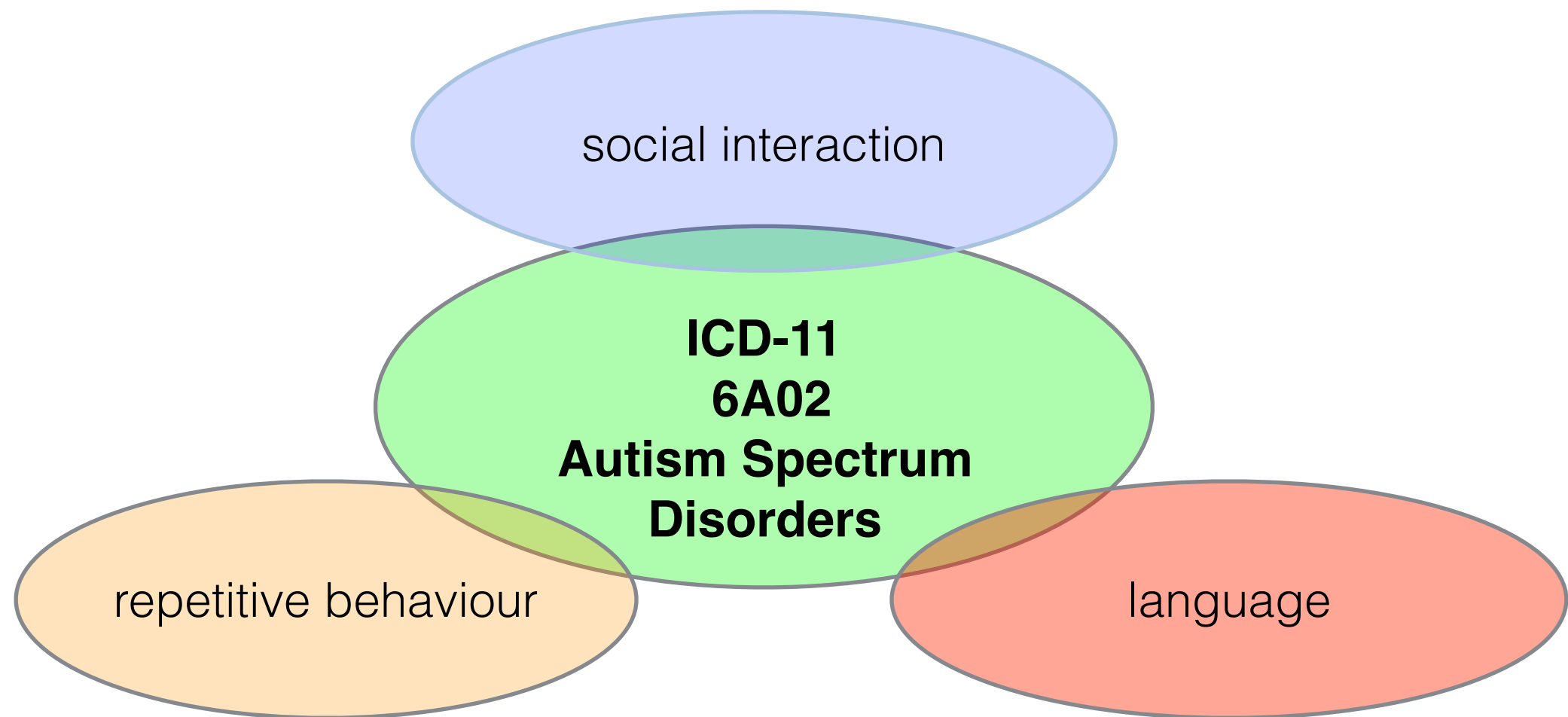
Autism Spectrum Disorders



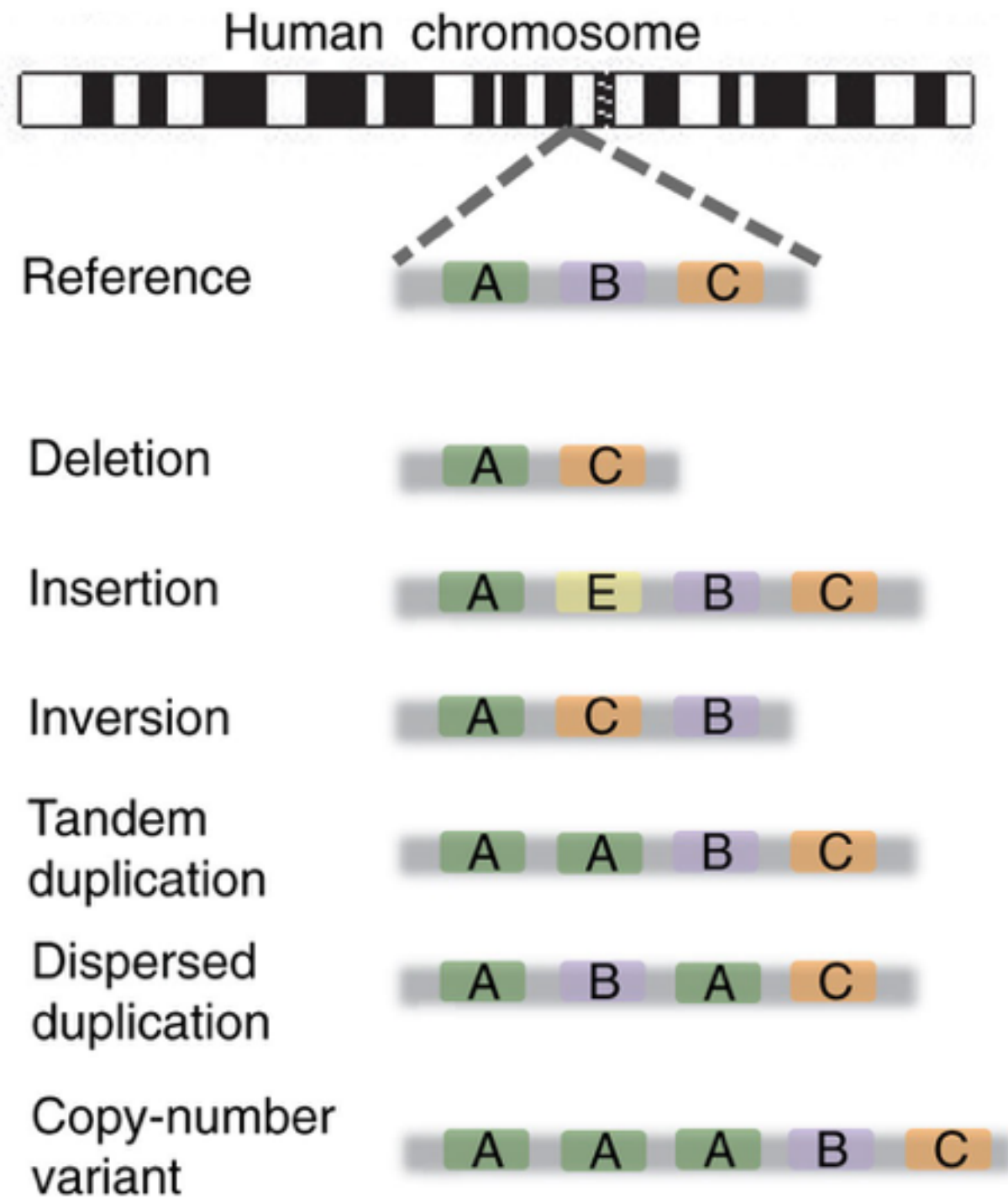
**Laboratory of Functional
and Structural Genomics**

CeNT CENTRUM
NOWYCH
TECHNOLOGII

Autism Spectrum Disorder



Structural Variants



Data

NCBI Resources How To [annamariabugaj@gmail.com](#) [My NCBI](#) [Sign Out](#)

dbVar dbVar Autism Search Help

Advanced



dbVar

dbVar is NCBI's database of genomic structural variation – insertions, deletions, duplications, inversions, mobile element insertions, translocations, and complex chromosomal rearrangements

NCBI Resources How To [annamariabugaj@gmail.com](#) [My NCBI](#) [Sign Out](#)

ClinVar ClinVar Search ClinVar for gene symbols, HGVS expressions, conditions, and more Search Help

Advanced

Home About Access Help Submit Statistics FTP

ACTGATGGTATGGGGCCAAGAGATATATCT
CAGGTACGGCTGTCATCACTTAGACCTCAC
CAGGGCTGGGCATAAAAGTCAGGGCAGAGC
CCATGGTGCATCTGACTCCTGAGGAGAAGT
GCAGGTTGGTATCAAGGTTACAAGACAGGT
GGCACTGACTCTCTCTGCCTATTGGTCTAT

ClinVar

ClinVar aggregates information about genomic variation and its relationship to human health.

Human Gene

The Human Gene module provides researchers around the world with instant access to the most up-to-date information on all known human genes associated with autism spectrum disorders (ASD).



SFARI GENE

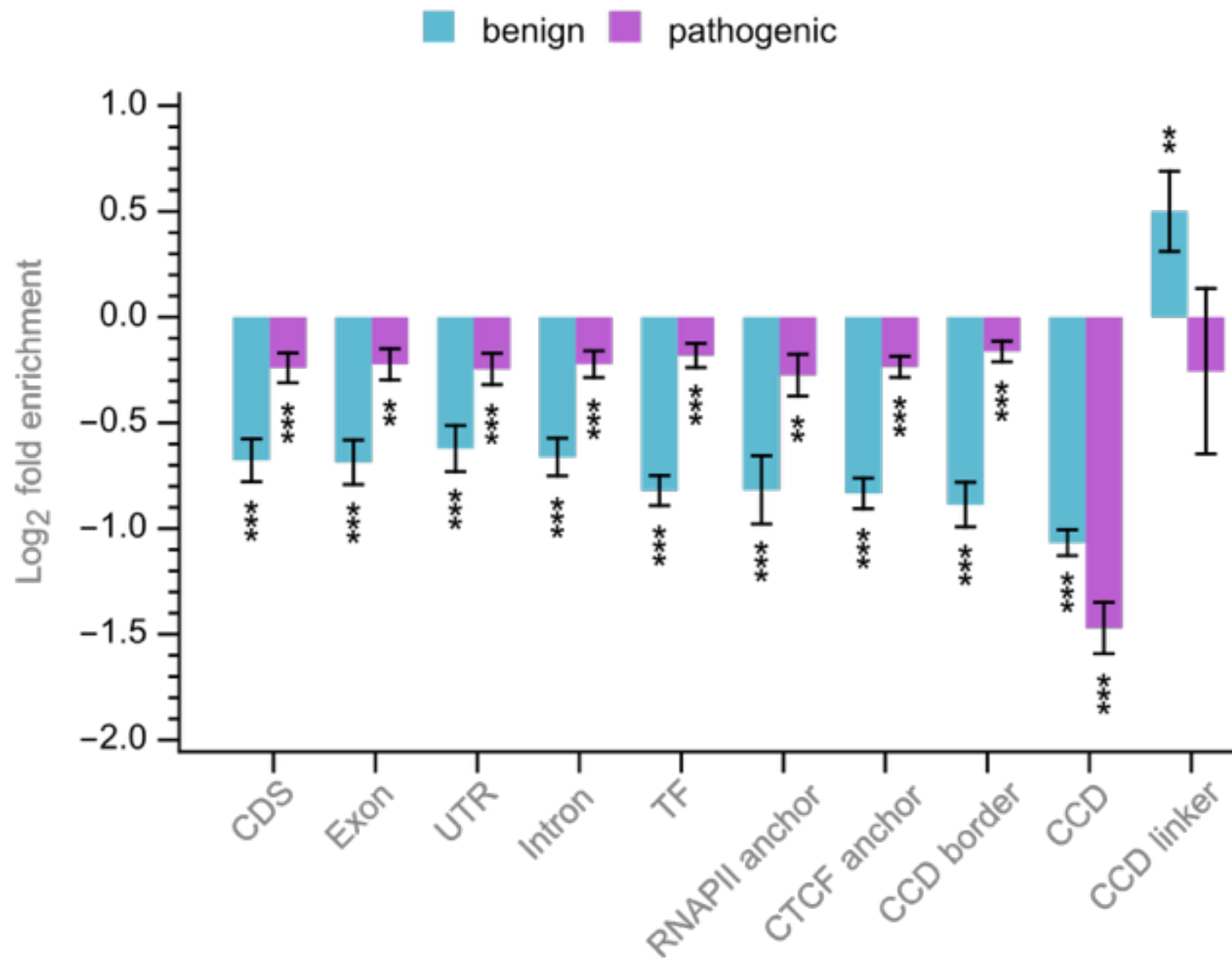
Explore the Database

Methods and Tools



- Lift Over to 19hg
- Mapping on the reference genome
- BedTools
- Enrichment Analysis
- Enrichment Plot

Results

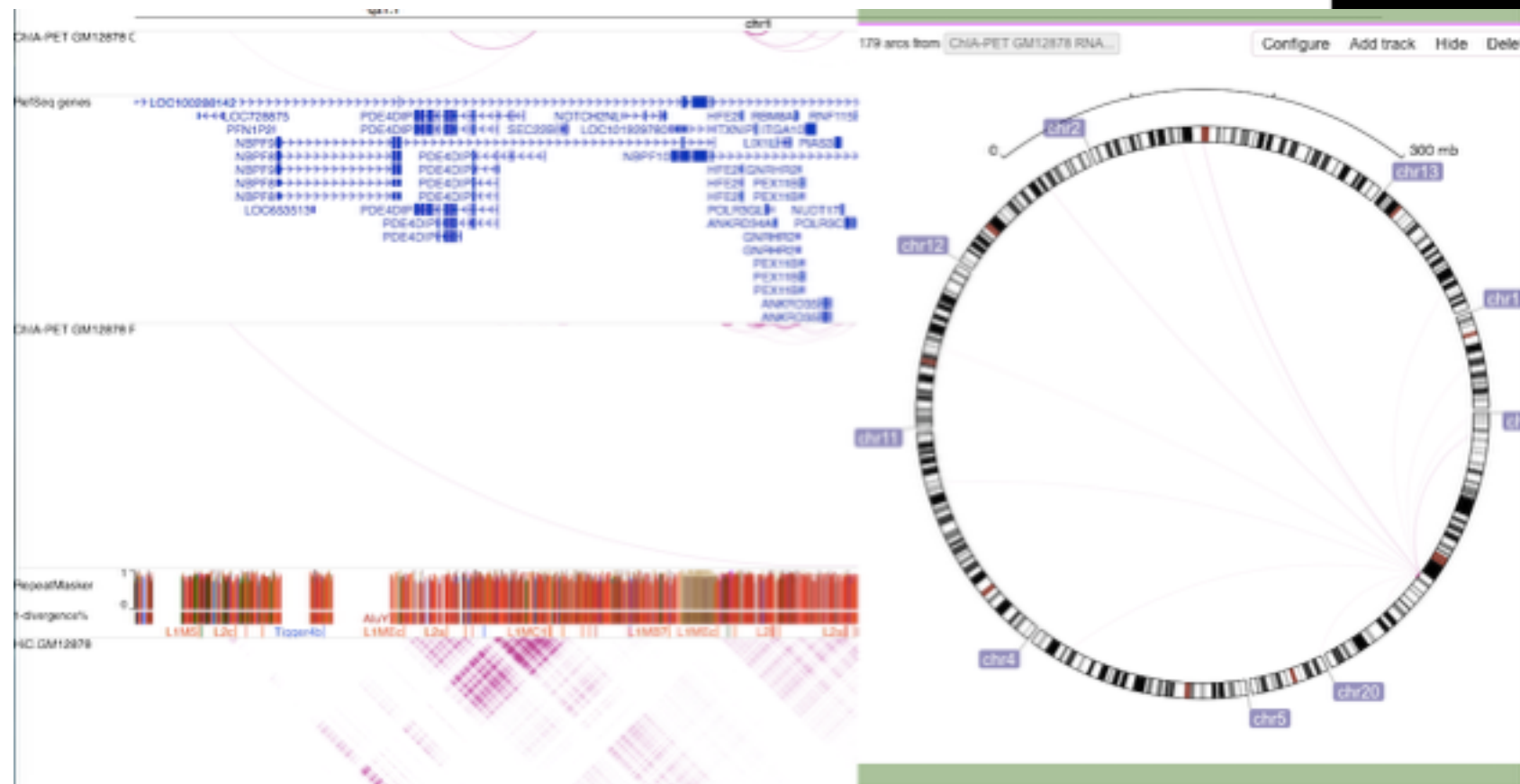
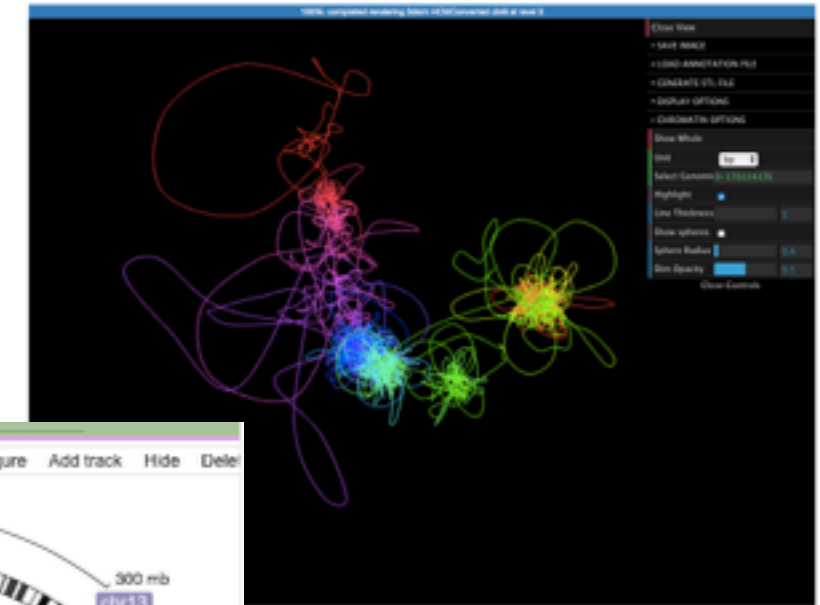


Next Steps

- Creating a data base of confirmed SVs and SNPs

	Standard	Standard	Standard	Standard	Standard
1	chr	start	end	description	study
2	chr16	29500000	30010000	microdeletion	http://www.nejm.org/doi/
3	chr1	146467203	147858208	duplication	https://doi.org/10.1016/
4	chr2	50145643	51259674	del/dup	https://doi.org/10.1016/
5	chr3	195747398	196191434	deletion	https://doi.org/10.1016/
6	chr7	72773570	74144177	duplication	https://doi.org/10.1016/
7	chr7	73978801	74144177	duplication	https://doi.org/10.1016/
8	chr15	23683783	28471141	duplication	https://doi.org/10.1016/
9	chr15	26971834	27548829	duplication	https://doi.org/10.1016/

- Analyzing data



I would like to thank

Dariusz Plewczyński, DSc, PhD

Paulina Urban, MSc

Michał Denkiewicz, MSc

Michał Sadowski, BSc

Przemysław Szałaj, MSc

Zofia Parteka, BSc

Agnieszka Kraft, BSc

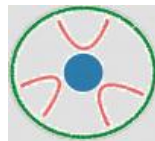
from LGFiS CeNT, UW

Laboratorium Genomiki Funkcjonalnej i Strukturalnej – projekt "Analiza sieciowa powiązań w strukturze chromatyny i interakcji białko-białko"

Michał Denkwicz

`m.denkwicz@cent.uw.edu.pl`

14.11.2017

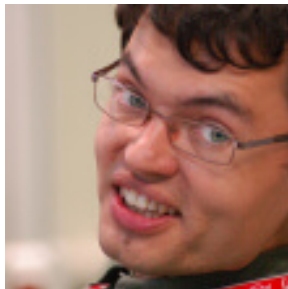


Laboratorium Genomiki Funkcjonalnej i Strukturalnej.
Centrum Nowych Technologii UW.

<http://nucleus3d.cent.uw.edu.pl/>

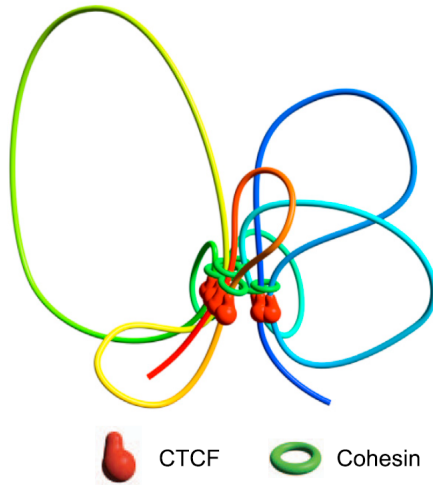
<http://www.cent.uw.edu.pl/pl/badania/lab/lgfs>

iCell: przetwarzania informacji w organizmach żywych. Rola struktury przestrzennej i hierarchii skal w sterowaniu procesami biofizycznymi w komórce.



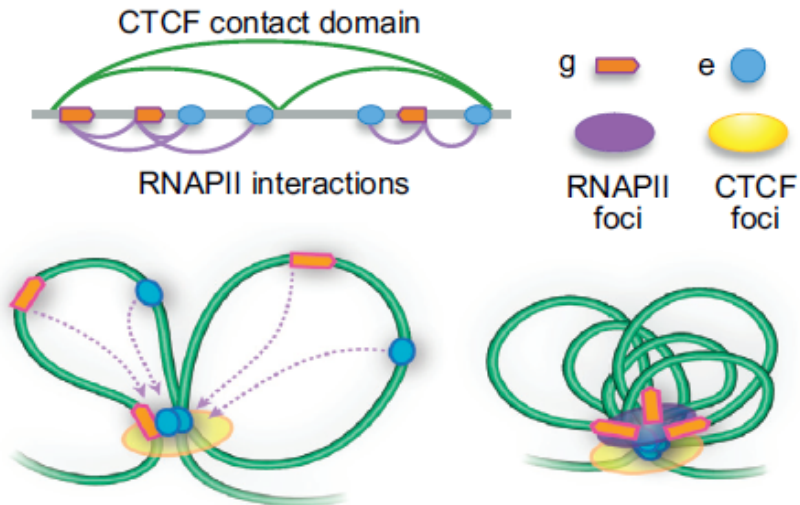
Ryc. 1: dr hab. Dariusz Plewczynski

Struktura 3d chromatyny



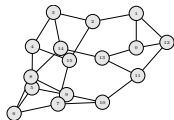
Ryc. 2: Modeling chromatin structure [Tang et al., 2015]

Znaczenie struktury

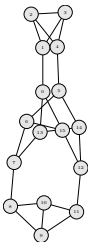


Ryc. 3: CTCF domains [Tang et al., 2015]

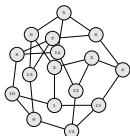
Topologia sieci



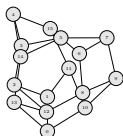
Max max closeness



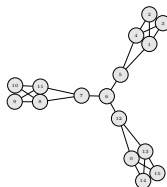
Max var constraint



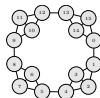
Min avg betweenness



Min avg clustering



Max max betweenness



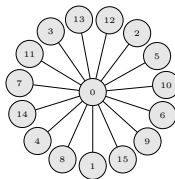
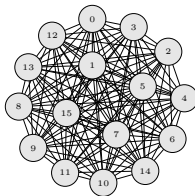
Max avg clustering



Min max closeness



Max avg betweenness



Miary centralności

Closeness of node u : $\sum_{s \neq u} \frac{1}{d(u,s)}$

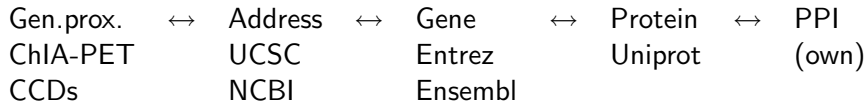
Betweenness of node u : $\sum_{s \neq u, t \neq u} \frac{\text{\#shortest paths from } s \text{ to } t \text{ containig } u}{\text{\#shortest paths from } s \text{ to } t}$

These can be aggregated (*min*, *max*, *avg*) across the network, to obtain a whole-network measure.

Project: Genome proximity vs PPI proximity

- ▶ Idea: proximal genes should code interacting proteins.
- ▶ PPI networks: based on previous project [Saha et al., 2014].
 - ▶ old, needs updating/expanding?
- ▶ Gene proximity:
 1. Same or other domain (0-1)
 2. Based on ChIA-PET contacts [Tang et al., 2015].

Data sources



Projekt

Możliwość współpracy z LGFiS:

1. Analiza sieciowa:
 - 1.1 Sieci białko-biało (PPI).
 - 1.2 Struktura chromatyny.
2. Integracja informacji w różnych bazach danych biologicznych.

Wymagania:

1. Programowanie (python), zwł. przetwarzanie danych.
2. Przeszukiwanie literatury w języku angielskim.
3. Chęć przyswojenia nowej wiedzy.
4. Motywacja.

Literatura I

Indrajit Saha, Julian Zubek, Tomas Klingström, Simon Forsberg, Johan Wikander, Marcin Kierczak, Ujjwal Maulik, and Dariusz Plewczynski. Ensemble learning prediction of protein–protein interactions using proteins functional annotations. *Molecular BioSystems*, 10(4):820–830, 2014.

Zhonghui Tang, Oscar Junhong Luo, Xingwang Li, Meizhen Zheng, Jacqueline Jufen Zhu, Przemyslaw Szalaj, Pawel Trzaskoma, Adriana Magalska, Jakub Włodarczyk, Blazej Ruszczycki, et al. Ctfc-mediated human 3d genome architecture reveals chromatin topology for transcription. *Cell*, 163(7):1611–1627, 2015.