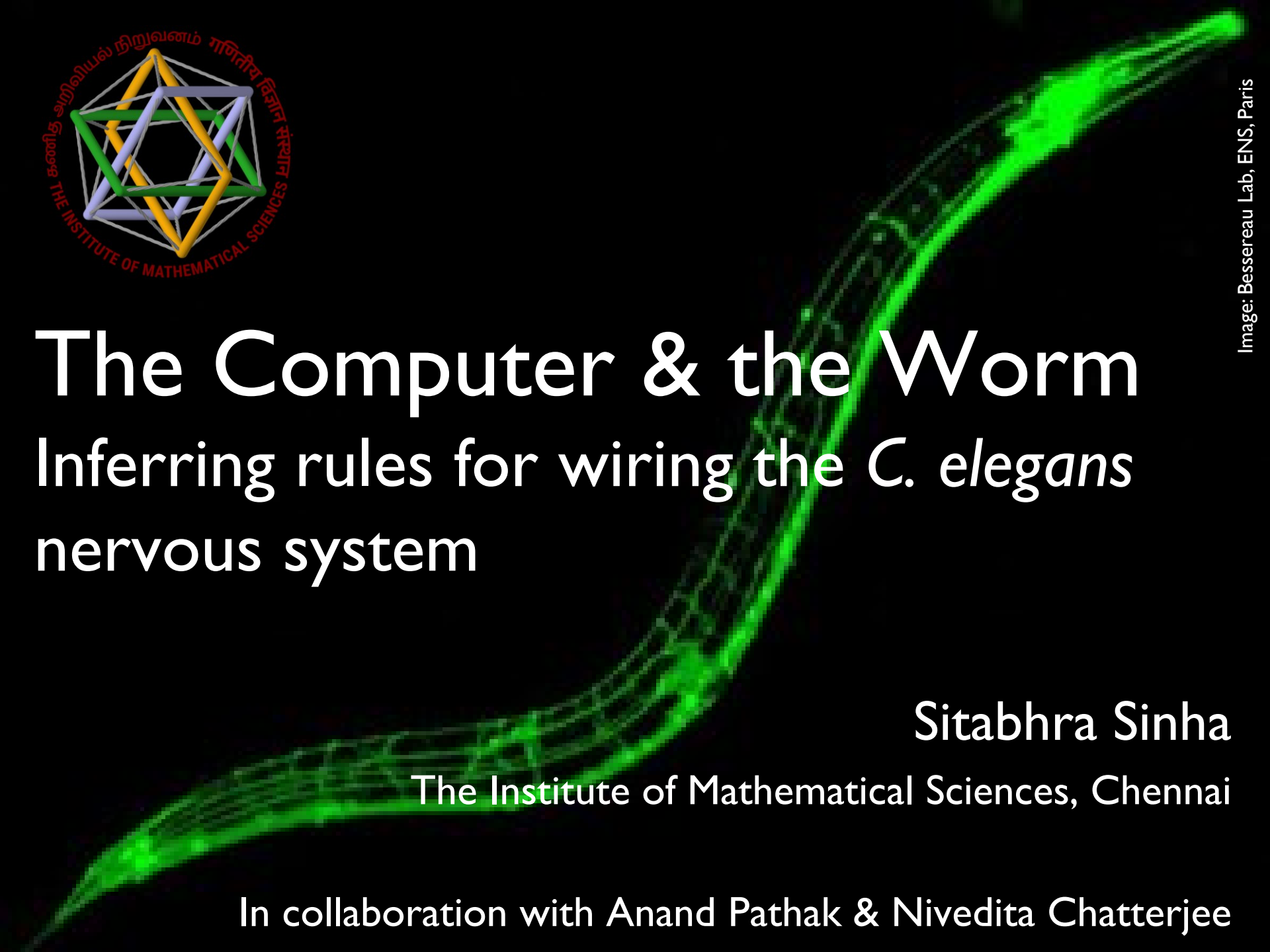


The Computer & the Worm

Inferring rules for wiring the *C. elegans* nervous system

Image: Bessereau Lab, ENS, Paris

Sitabhra Sinha

The Institute of Mathematical Sciences, Chennai

In collaboration with Anand Pathak & Nivedita Chatterjee

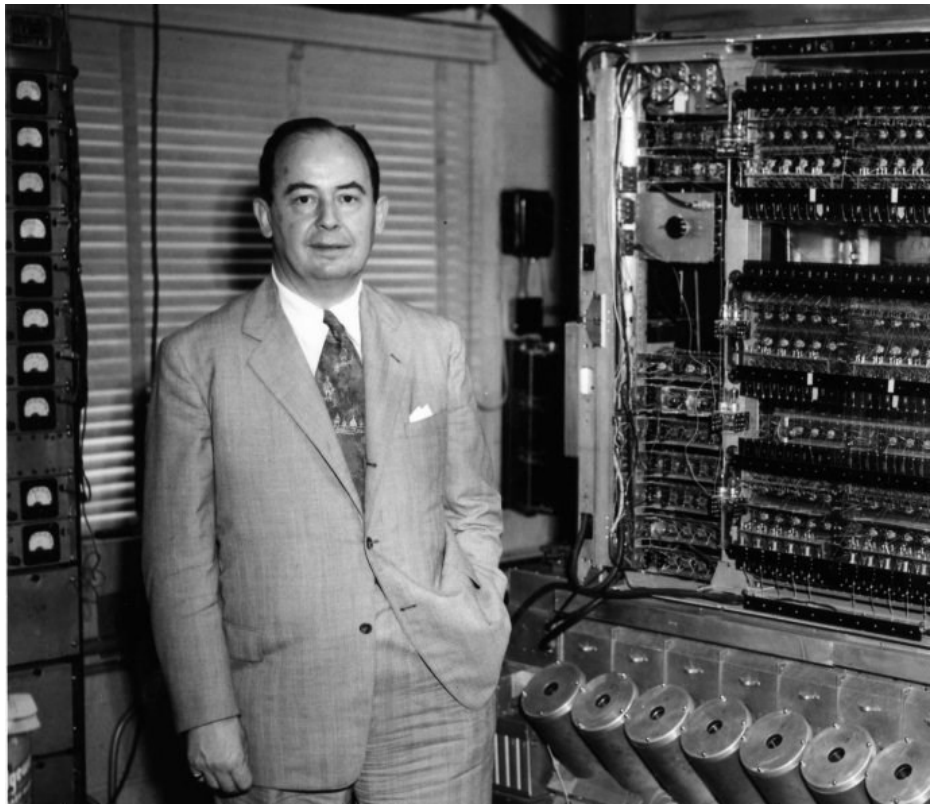


Image: Alan Richards

"A code… according to Turing's schema is supposed to make one machine behave as if it were another specific machine … must do the following things. It must contain (…) instructions that will cause the machine to examine every order it gets and determine whether this order has the structure appropriate to an order of the second machine. It must then contain… sufficient orders to make the machine (take) the actions… that the second machine would have taken (given) the order in question."

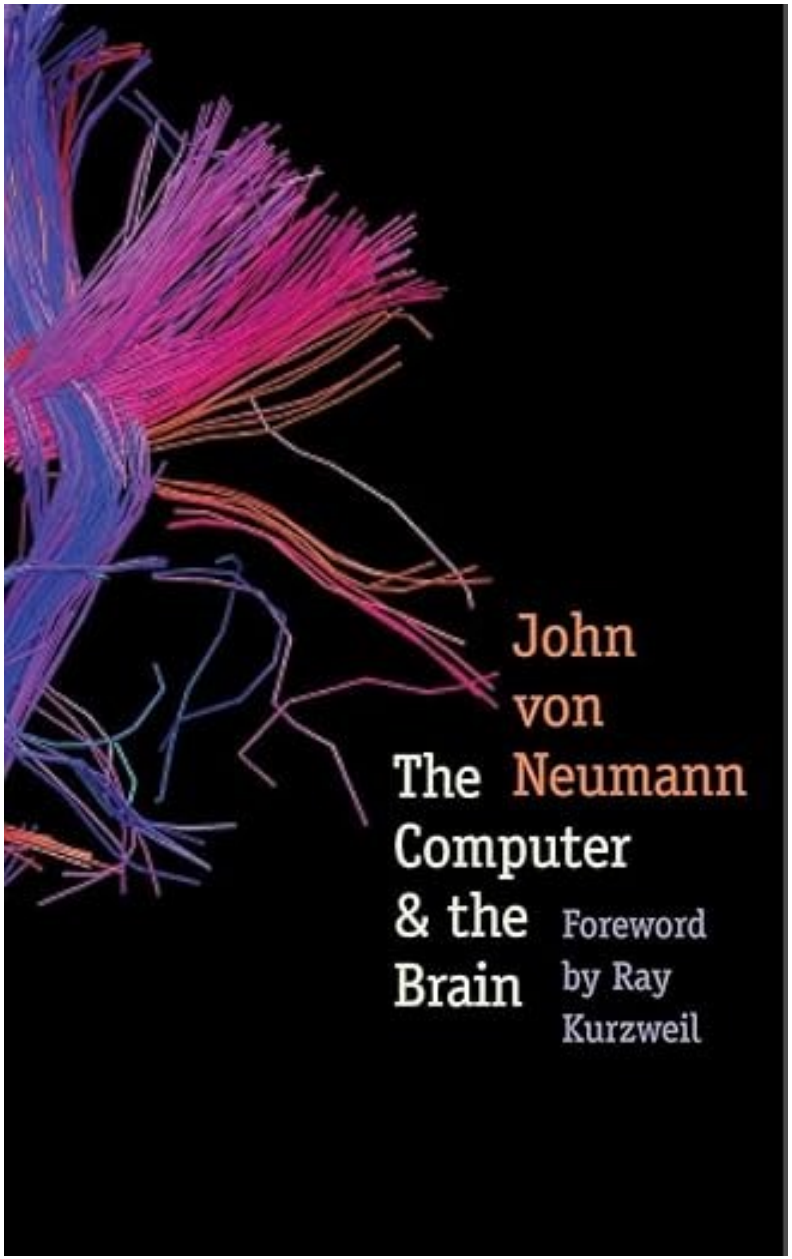
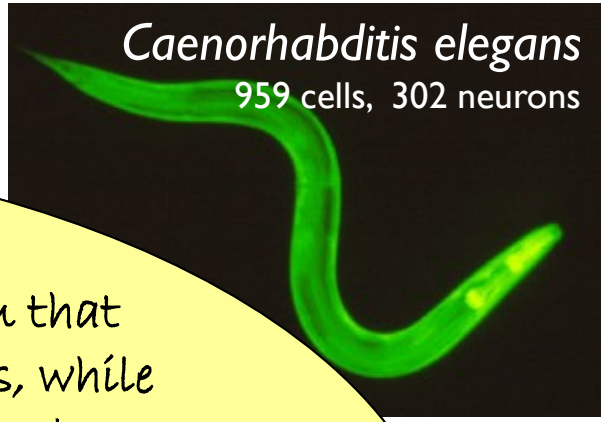


Image: www.amazon.in

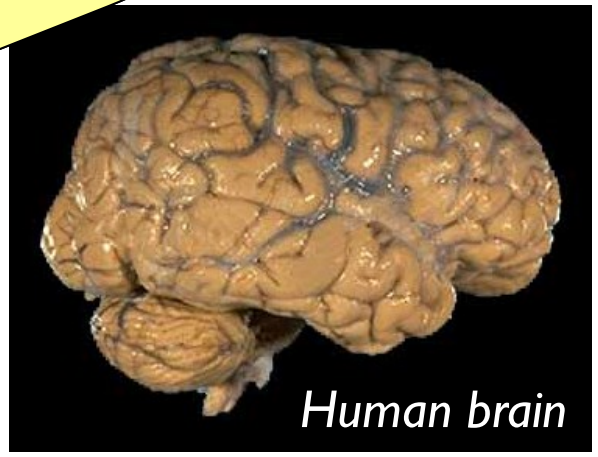
1958

Why the worm ?



Just imagine, here we have a natural organism that survives in the wild with just 300 odd neurons, while we struggle to mimic a single aspect of the mind, like memory, with neural networks having upwards of 10,000 model neurons... if we don't understand how the worm nervous system does its work, there's little hope we'll ever understand how the much more complicated human brain works, or make a reasonable computer model of it.

~ *Bikas K Chakrabarti (circa1998)*



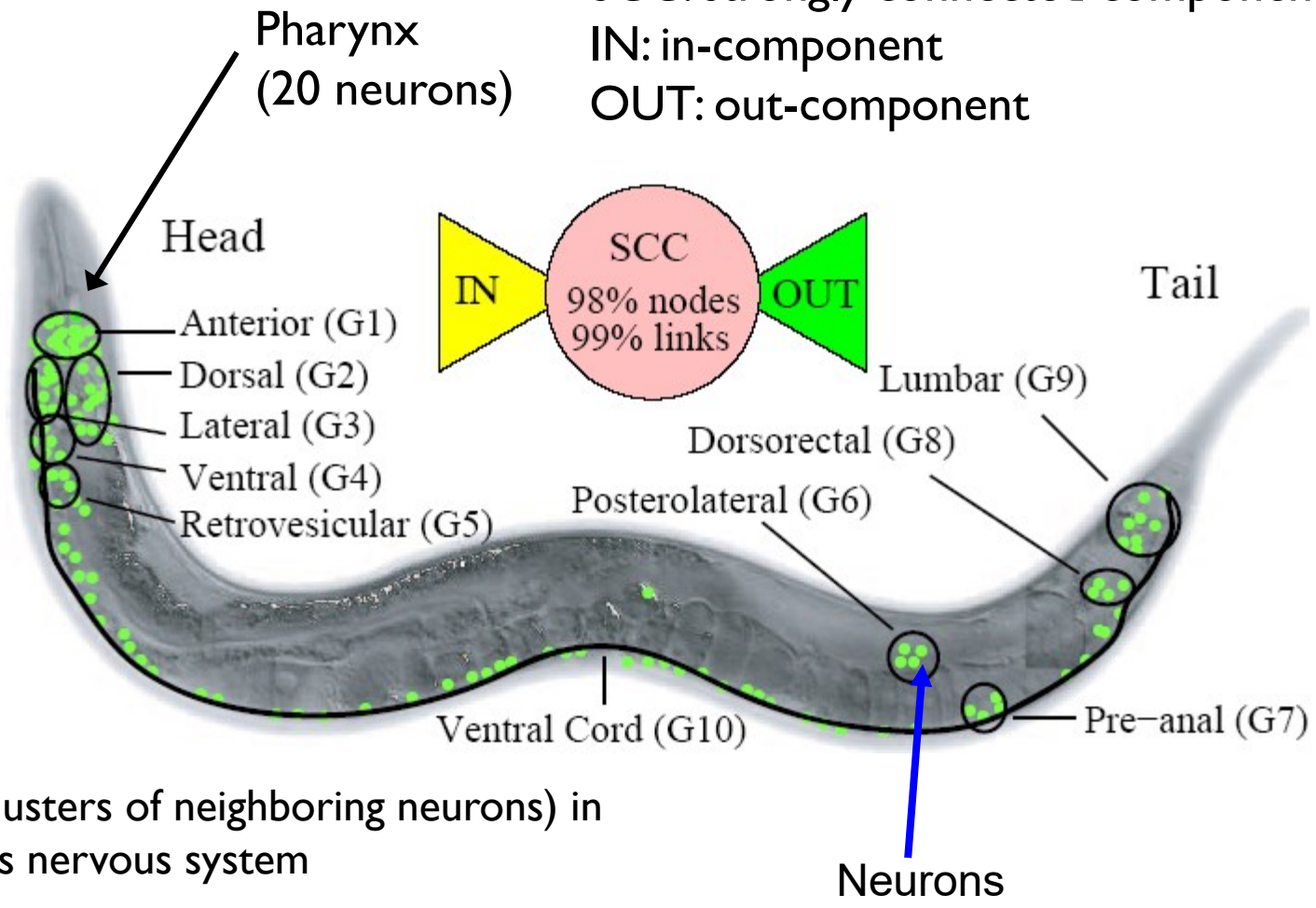
10^{11} cells, 10^{15} synapses

Structure of the nervous system

Focus on the 279 non-pharyngeal neurons forming a connected network in the hermaphrodite worm

Nivedita

SCC: strongly connected component
IN: in-component
OUT: out-component

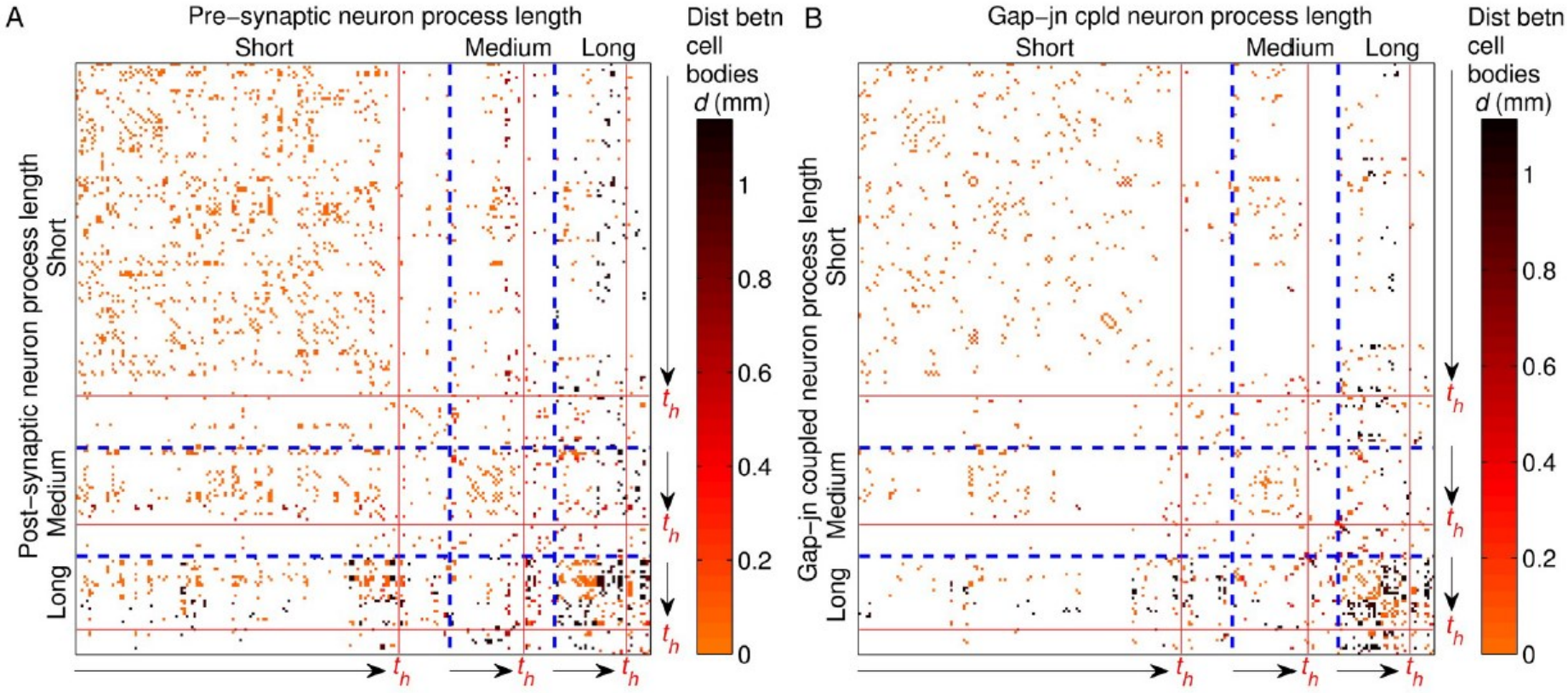


G: Ganglia (clusters of neighboring neurons) in the C. elegans nervous system

Connectivity of the somatic nervous system

Synaptic

Gap-junctional



Question:

Is the network modular ? How do you determine the modules if the connections are not localized within corresponding ganglia ?

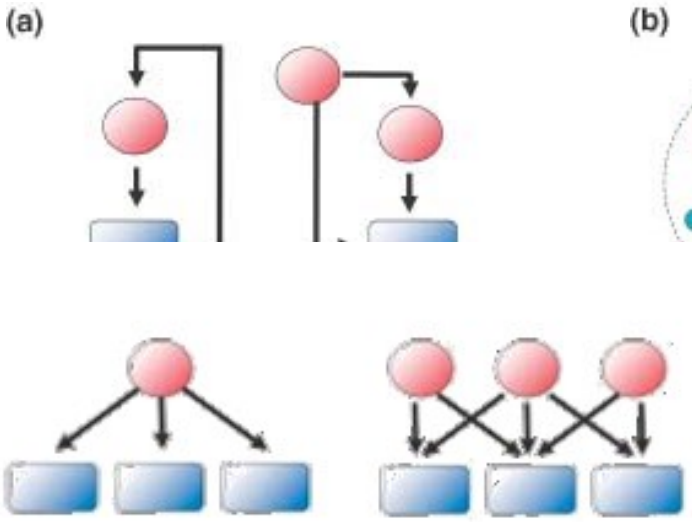
Modular Networks: dense connections *within* clusters (modules) & relatively few connections *between* modules

Raj

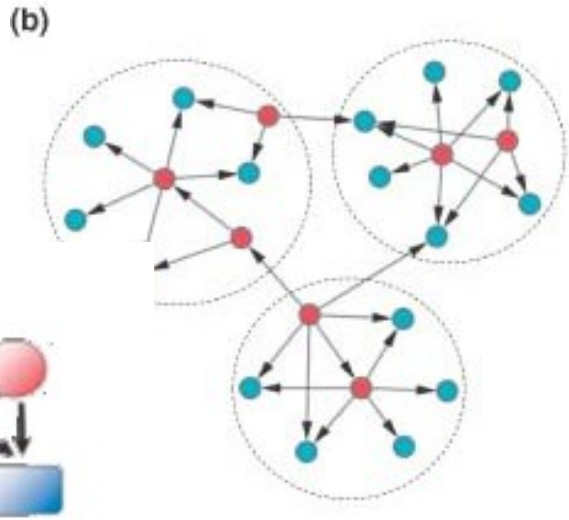
Modules: A *mesoscopic* organizational principle of networks

Going beyond ***motifs*** but more detailed than ***global*** description (*L*, *C* etc.)

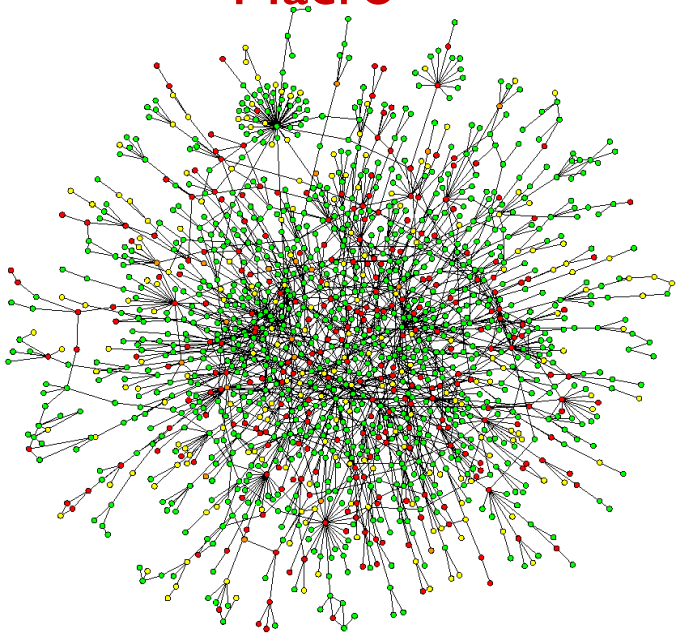
Micro



Meso



Macro

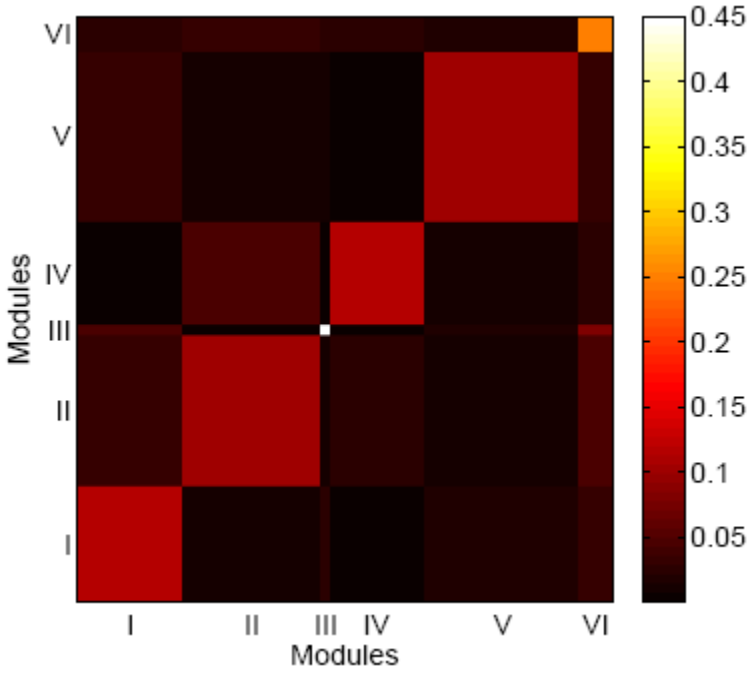


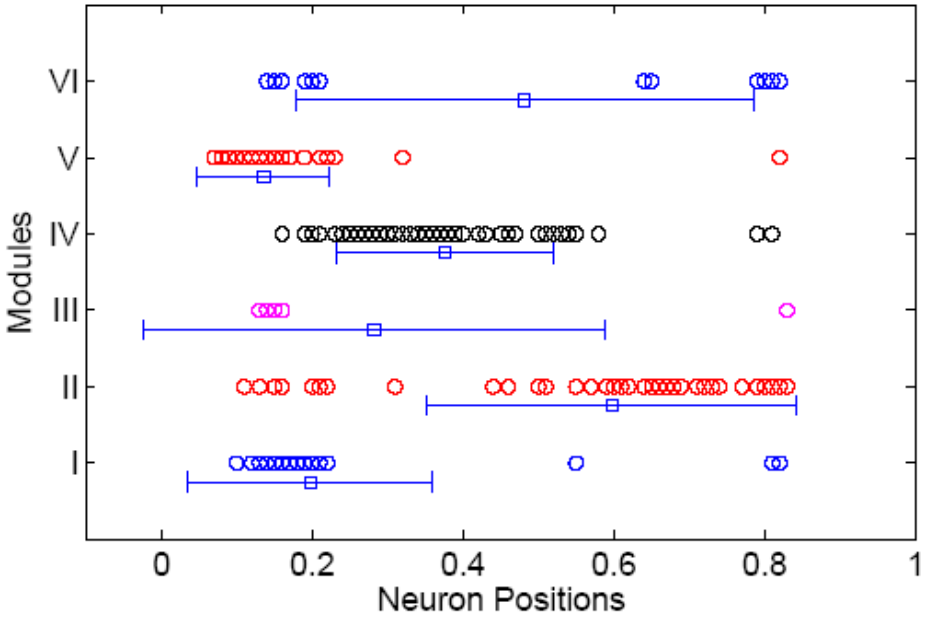
Kim & Park, WIREs Syst Biol & Med, 2010

What ? The Modular Structure of the Network

Decomposition of the somatic nervous system into 6 modules

Modules determined through a generalization of the spectral method (Leicht & Newman, 2008)





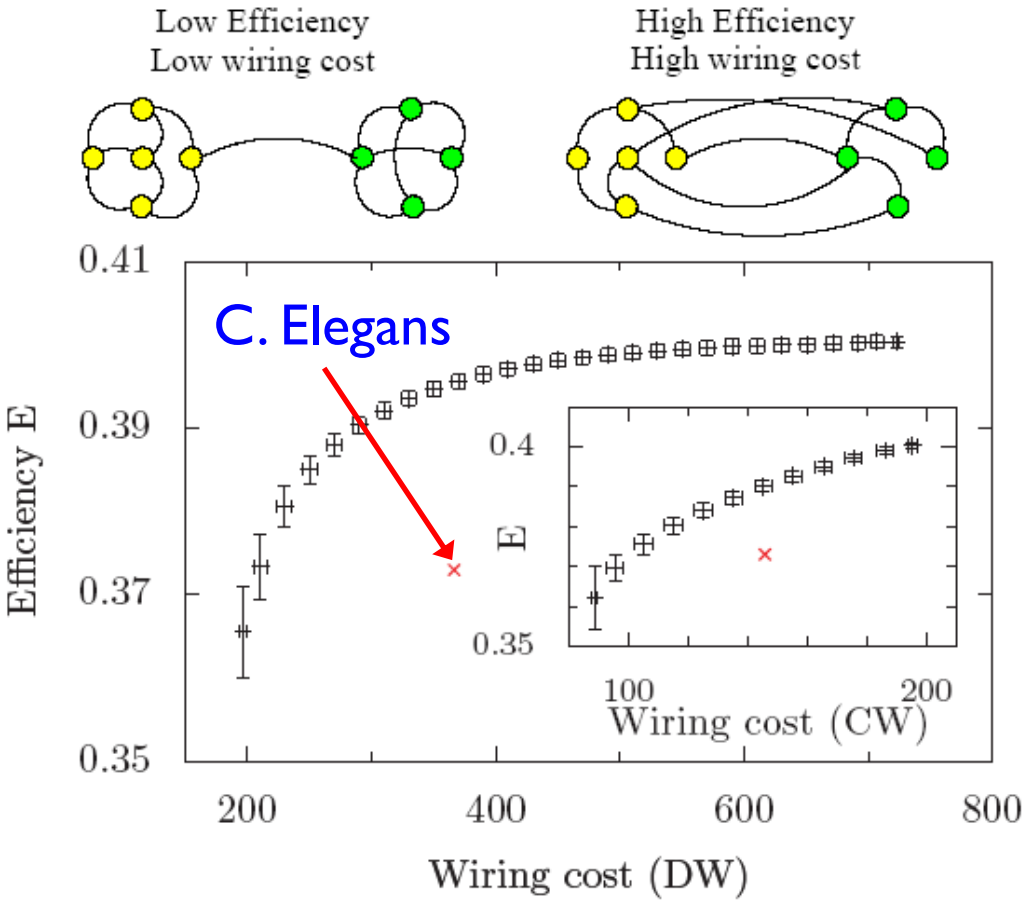
- Dense interconnectivity within neurons in a module, relative to connections between neurons in different modules
- Existence of the modules is not a trivial outcome of spatial location of constituent neurons

Optimizing for wiring cost and communication efficiency

Communication efficiency: $E = 1 / \text{avg path length}, \ell = 2 / N(N-1) \sum_{i>j} d_{ij}$

Wiring cost: $DW = \sum_{i>j} d_{ij}$ for all connected neurons

("dedicated wire" model)



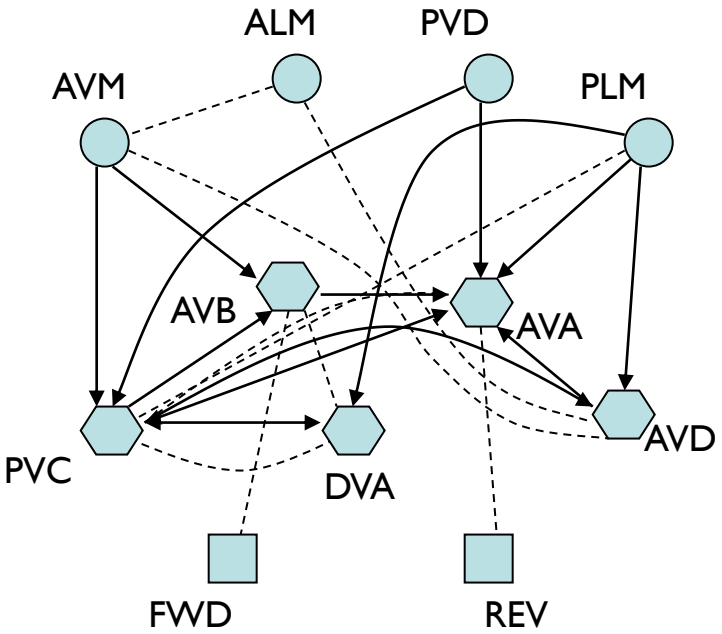
Trade-off between increasing communication efficiency and decreasing wiring cost

The network is sub-optimal !
⇒ presence of other constraints (possibly related to function) governing network organization

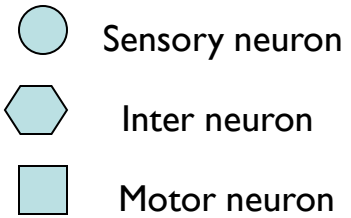
Linking Structure and Function

Functional circuits of C Elegans

- Touch sensitivity
- Egg laying
- Thermotaxis
- Chemosensory
- Defecation
- Locomotion
  when
  - Satiated: Feeding
  - Hungry: Exploration
  - Escape behavior: Tap withdrawal



Tap Withdrawal Circuit



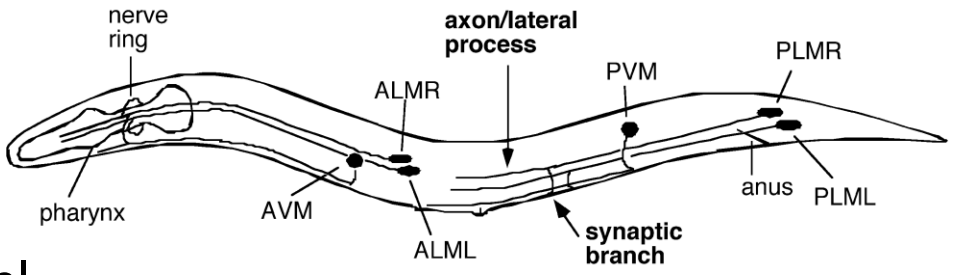
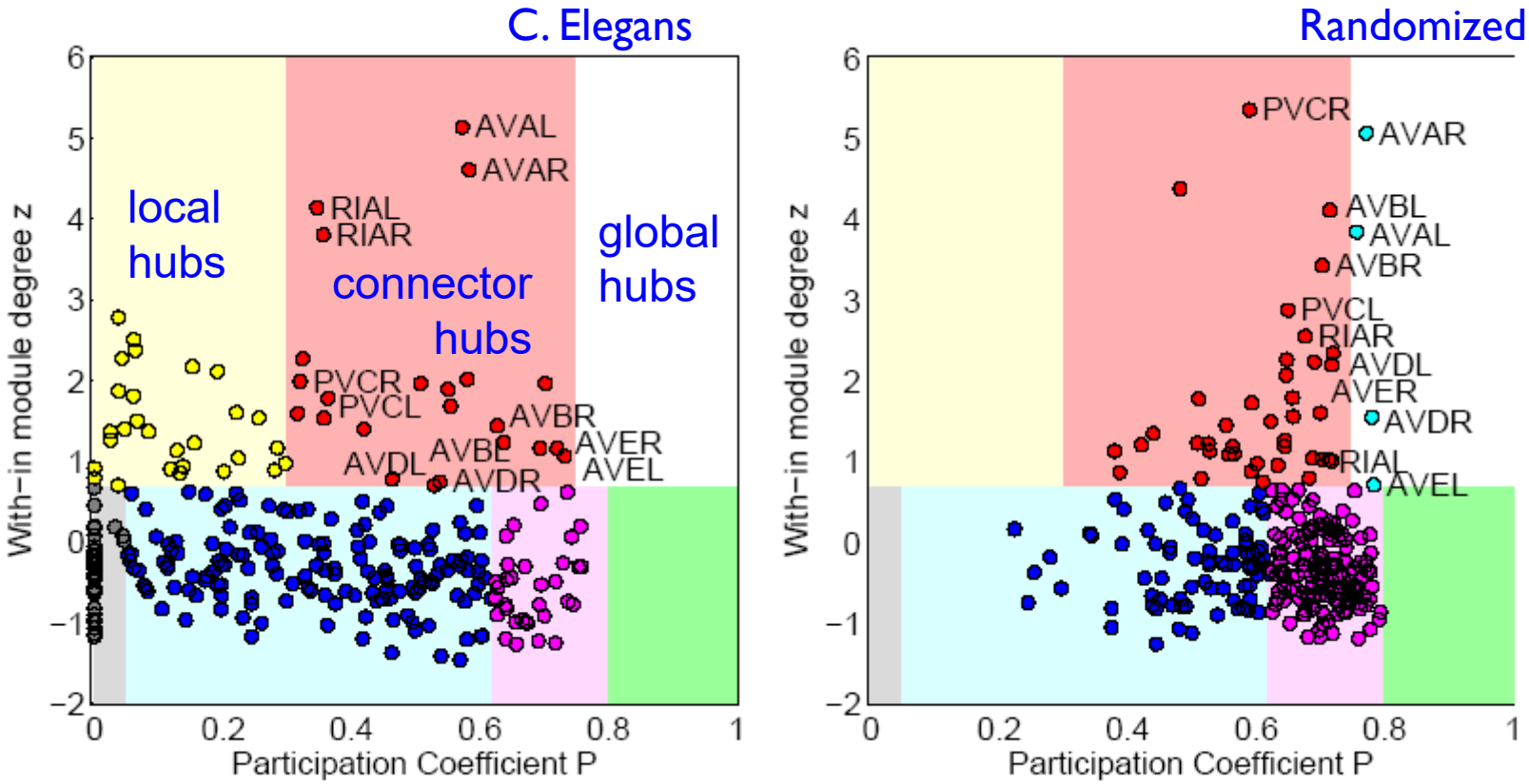


Image: sites.wustl.edu/nonetlab/c-elegans/

Why ? Mesoscopic network structure can alert us to critical functional role of neurons

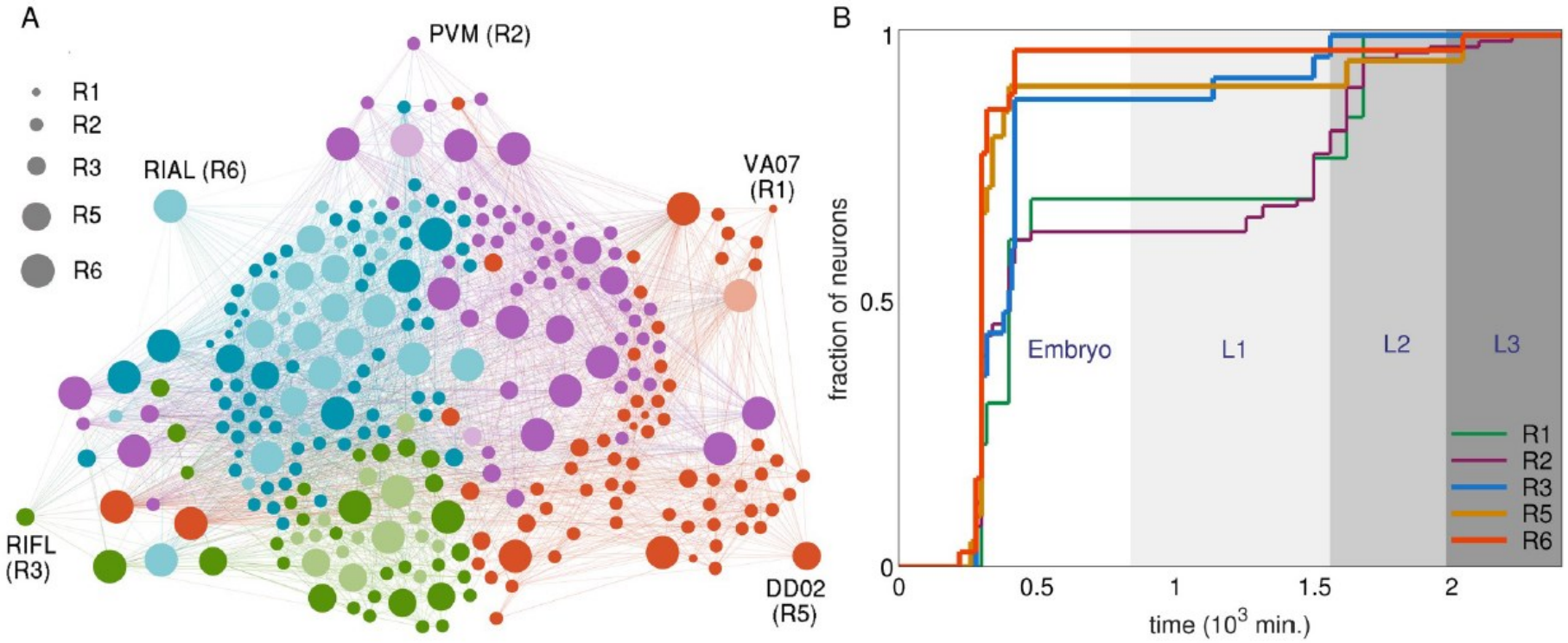


Importance of connector hubs: possibly integrating local activity for coherent response, 21 out of 23 already implicated in critical functions

AVKL and SMBVL are likely important for some as yet undetermined function

Neurons functioning as connectors between different modules lead in development

Anand

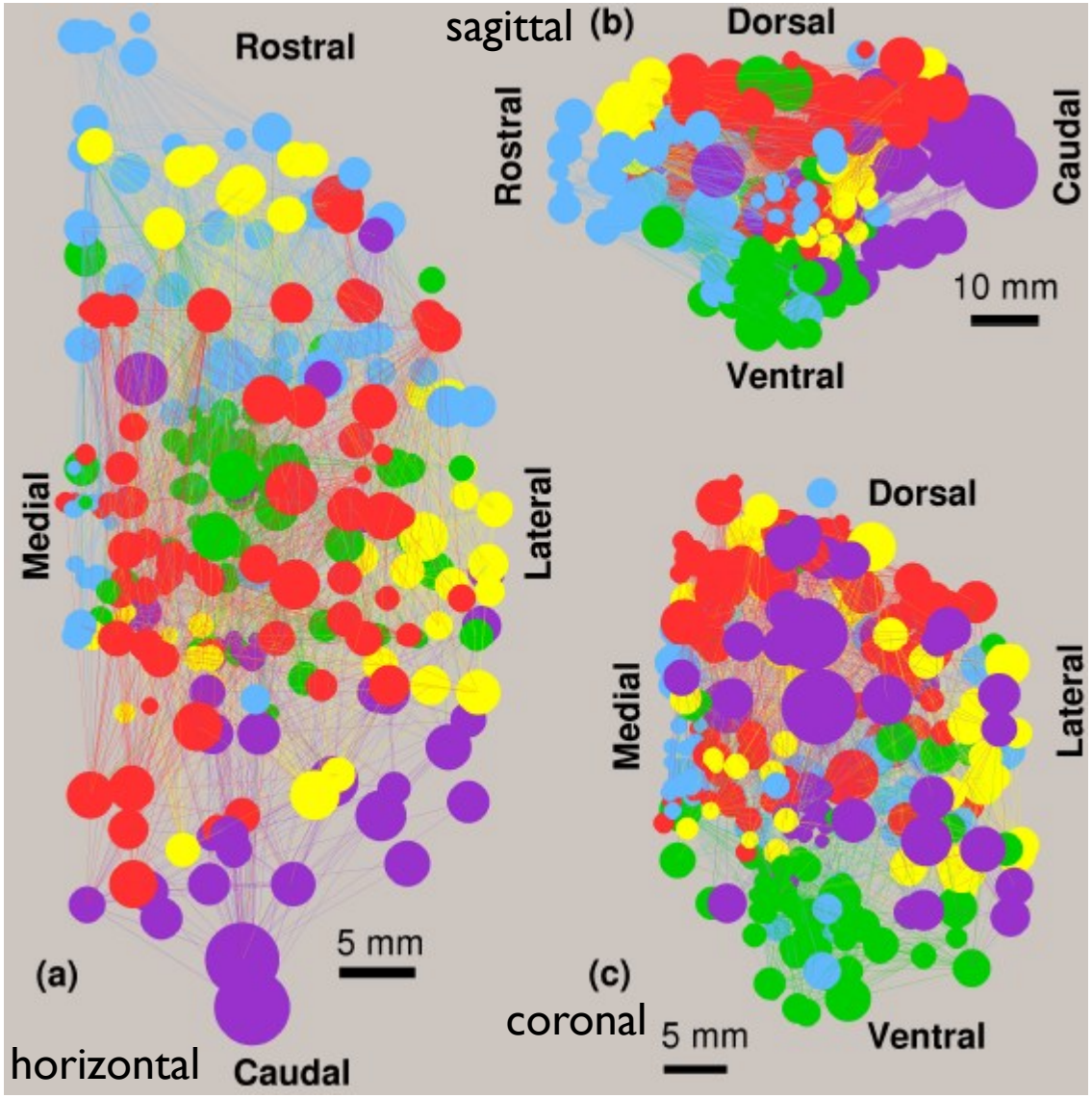


more than 90% of satellite connectors, provincial hubs and connector hubs appear before hatching

For the peripheral categories (R1 and R2), 70% or less differentiate by that time

Similar mesoscopic organization in the network of brain regions in the Macaque

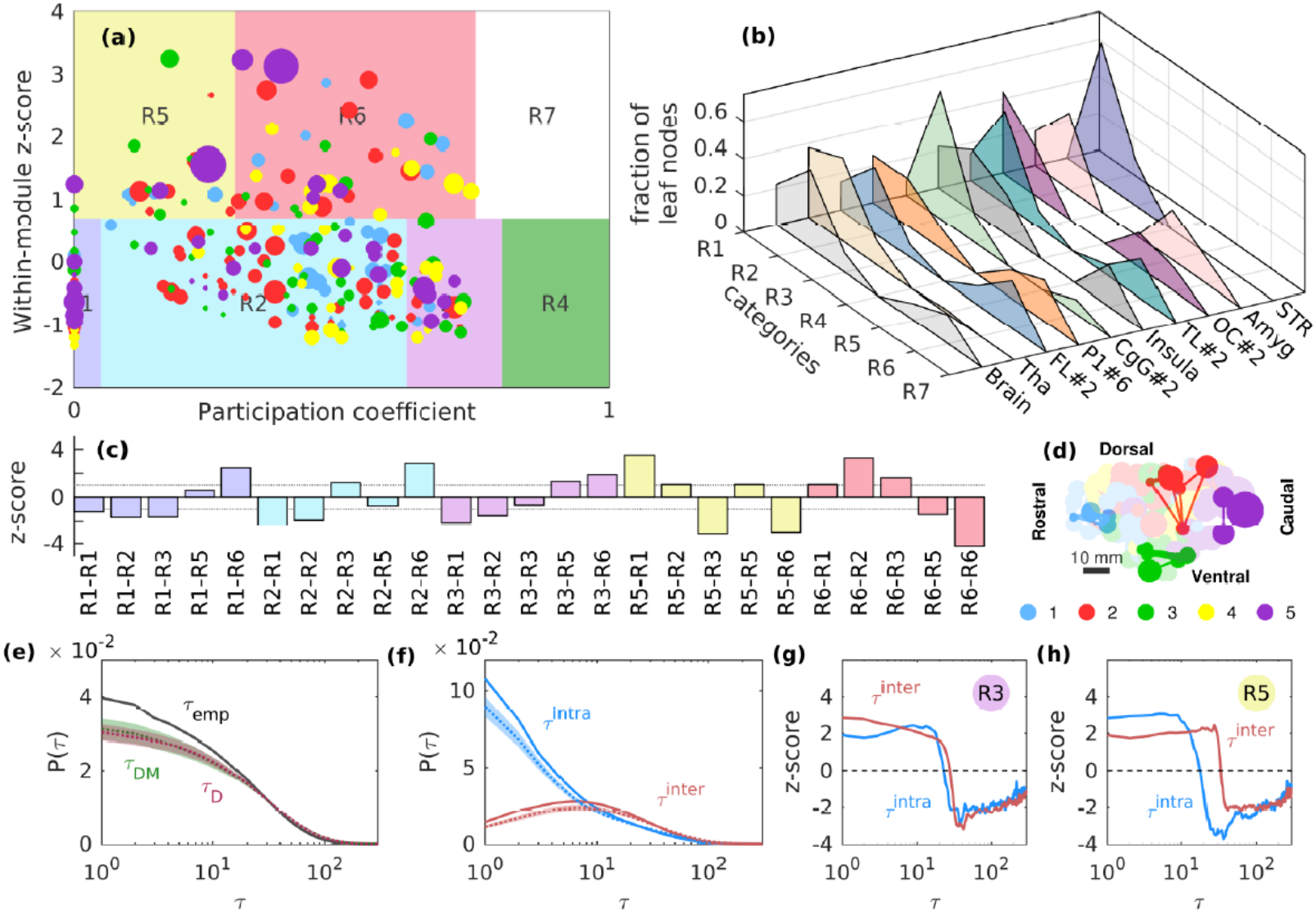
Anand



Circles: spatial positions of brain areas
Circle size: relative volumes
Links: fibre tracts connecting the areas

The network consists of 5 distinct densely connected communities (different colored nodes) that appear to be localized in space with some exceptions.

Information spreading within the Macaque brain enhanced by the specific pattern of intra- & inter-modular links

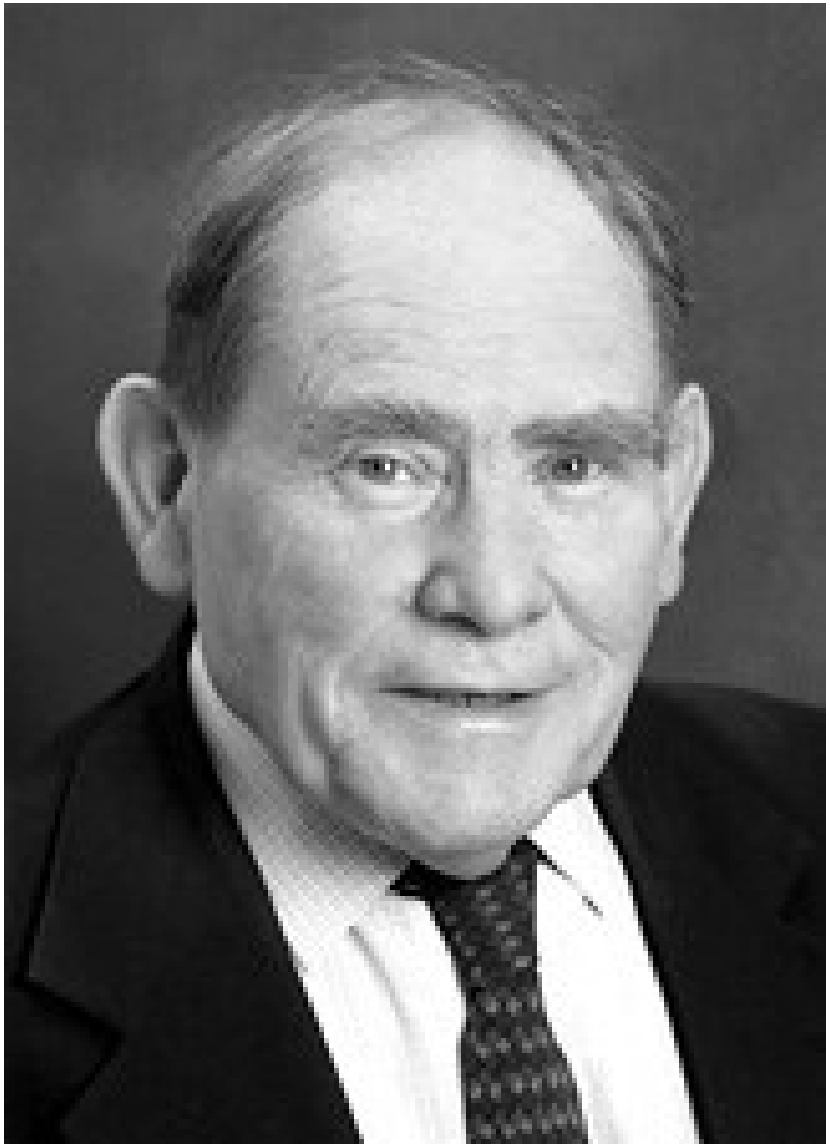


How ?

From a descriptive to a developmental perspective…

Key questions about nervous system development

- how are the neurons spatially localized in their specific positions ?
- how they connect to each other through synapses and gap junctions forming a network with a precisely delineated connection topology?
- what governs the temporal sequence in which different neurons appear over the course of development ?



Sydney Brenner

decomposed the problem of

how do genes affect behavior ?

Into understanding

(a) Understanding how genes specify the nervous system

how is it built?

and

(b) Understanding how behavior is produced by the activity of the nervous system

how does it work?

In a similar spirit, to resolve

The Wiring Problem

we view the system at a level intermediate between

•the detailed molecular machinery involving diffusible factors, contact mediated interactions, growth cone guidance, etc.,

complicated

and

•Spatial and network topological description of the nervous system of the mature worm.

complicated

We ask

Where, When and Who ?

Spatial: why is the neuron where it is relative to other neurons?

Temporal: why is it that certain neurons are born much earlier than others?

Topological: why does a neuron have the links it does?

Homophily based on multiple cellular properties governs neuronal connections

The questions about development of the worm nervous system are related to general principles expressed in terms of different types of

homophily

the tendency of entities sharing a certain feature to preferentially connect to each other

We identify **four** different types of homophily, related to:

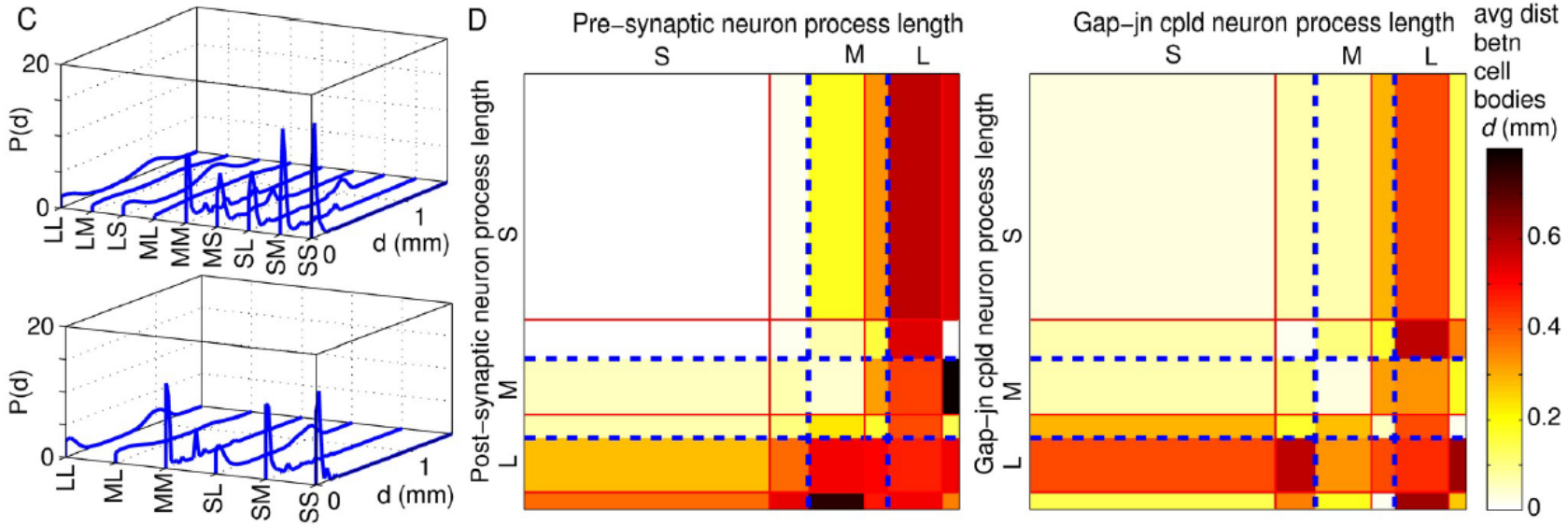
- ❑ process or neurite length of neurons,
- ❑ the time of their appearance,
- ❑ their lineage history, and
- ❑ bilateral symmetry

Anand

Process length homophily

An explicit preference for neurons to connect to other neurons whose neurites extend to similar distances as them (categorized into *long*, *medium* and *short* processes)

Demonstrated by comparing the empirical connectivity with that expected from randomized surrogates obtained by permuting process length category of each neuron



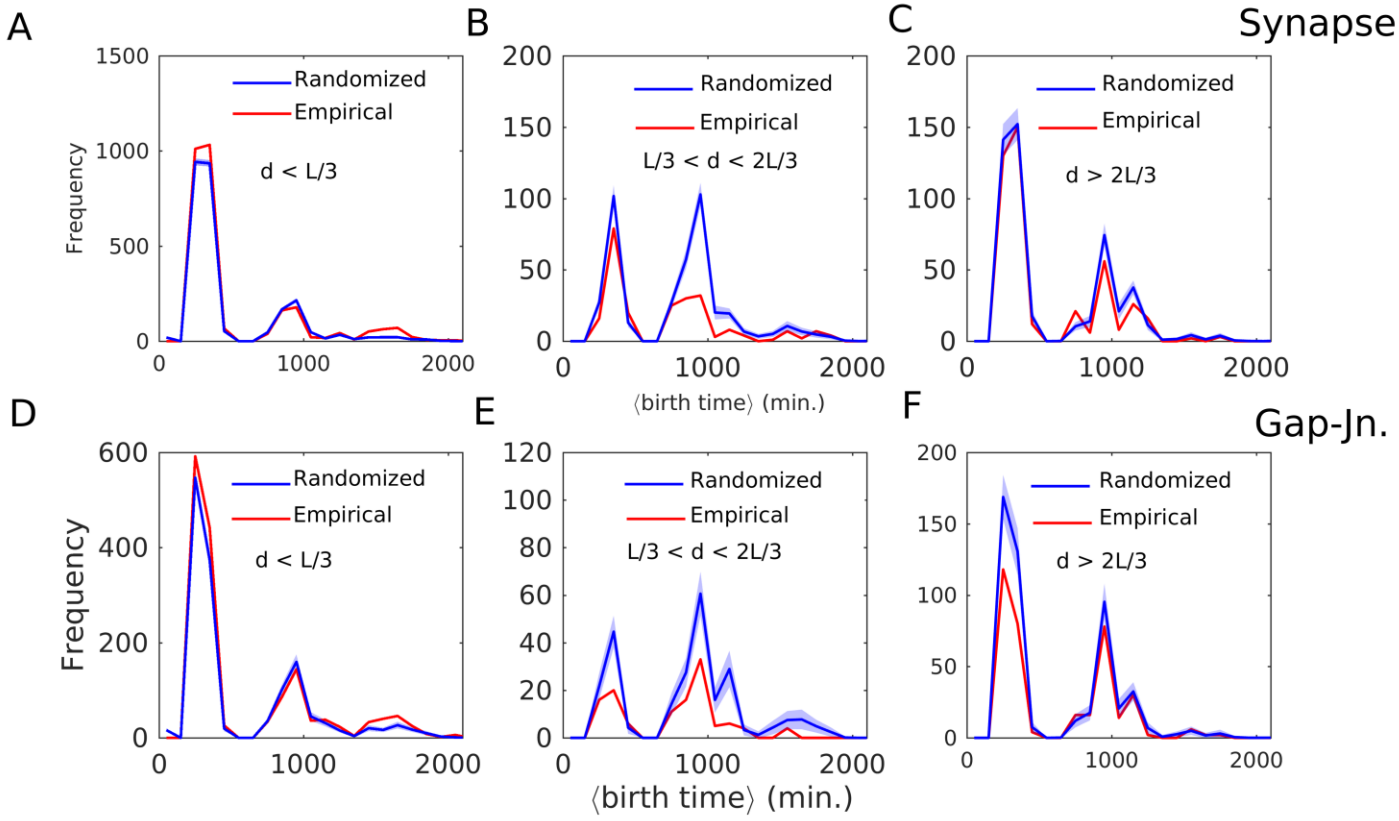
Process lengths affect the spatial arrangement of neurons

the distance *d* between cell bodies of connected pairs of neurons are distributed differently according to their respective process lengths, e.g., bimodal when at least one of the neurons have long or medium length process

Birth cohort homophily

The time of birth of cells determine their inter-connectivity with neurons preferring to connect to other members of their birth cohort (viz., early or later-born)

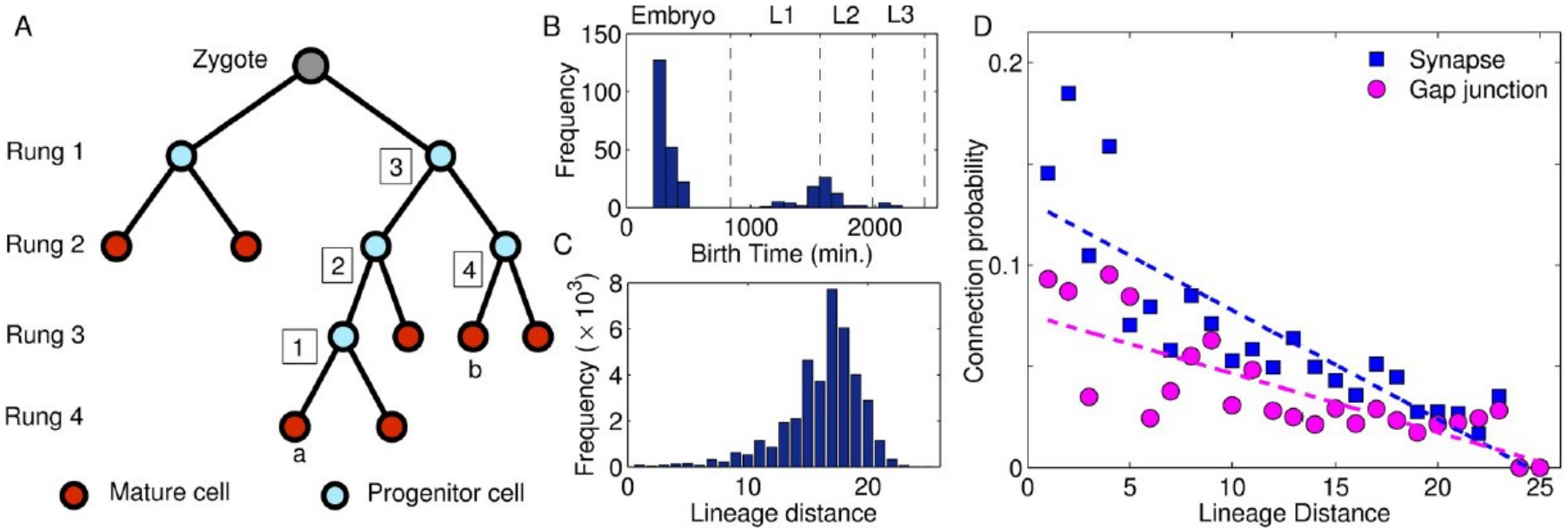
Restricted to neurons whose cell bodies are located in close physical proximity



Homophily ⇒ peaks of the empirical distribution have significantly higher values than the randomized distribution from a null model where connections occur independent of birth time

Lineage homophily

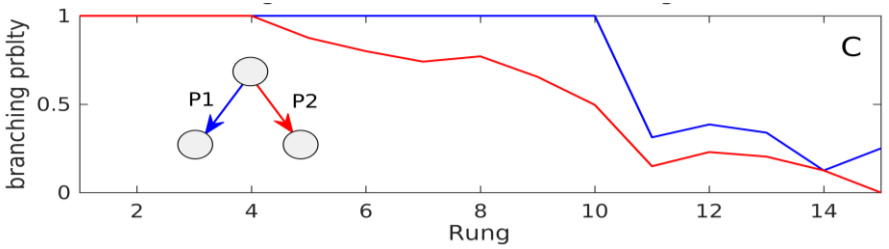
The lineage distance between a pair of mature cells is measured as the total number of cell divisions leading to each from their common progenitor



The probability of a pair of neurons to be connected through a synapse decreases with increasing lineage distance between them ($r = -0.87, p < 10^{-7}$)
For gap junctional connections the correlation is marginally weaker ($r = -0.79, p < 10^{-5}$).

Changes in cell body locations brought about by the appearance of cells born later through subsequent cell-divisions result in a weak correlation between connection probability and physical distance separating the cell bodies

Lineage relations from asymmetric stochastic branching process



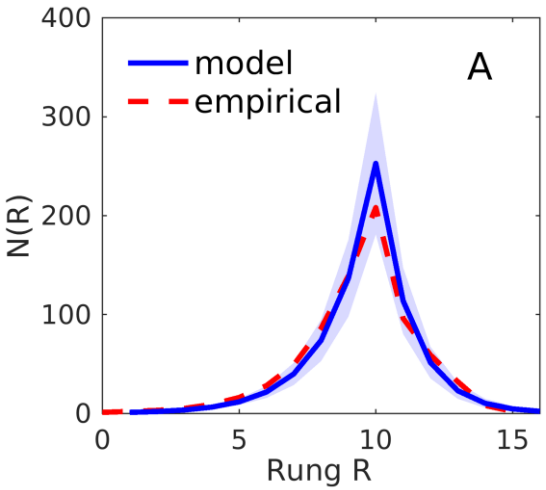
The Model:

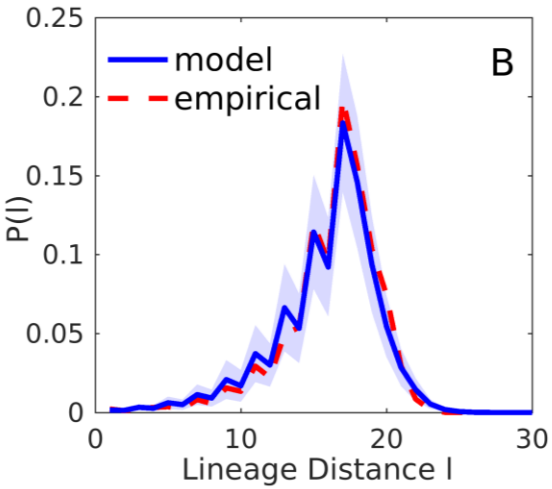
starting from single cell zygote, each cell division leads to at most two daughter cells, with independent probabilities P_1 and P_2 ($P_1 \geq P_2$)

Based on the probabilities P_1 and P_2, at each step of the generative process any one of the following three events can happen:

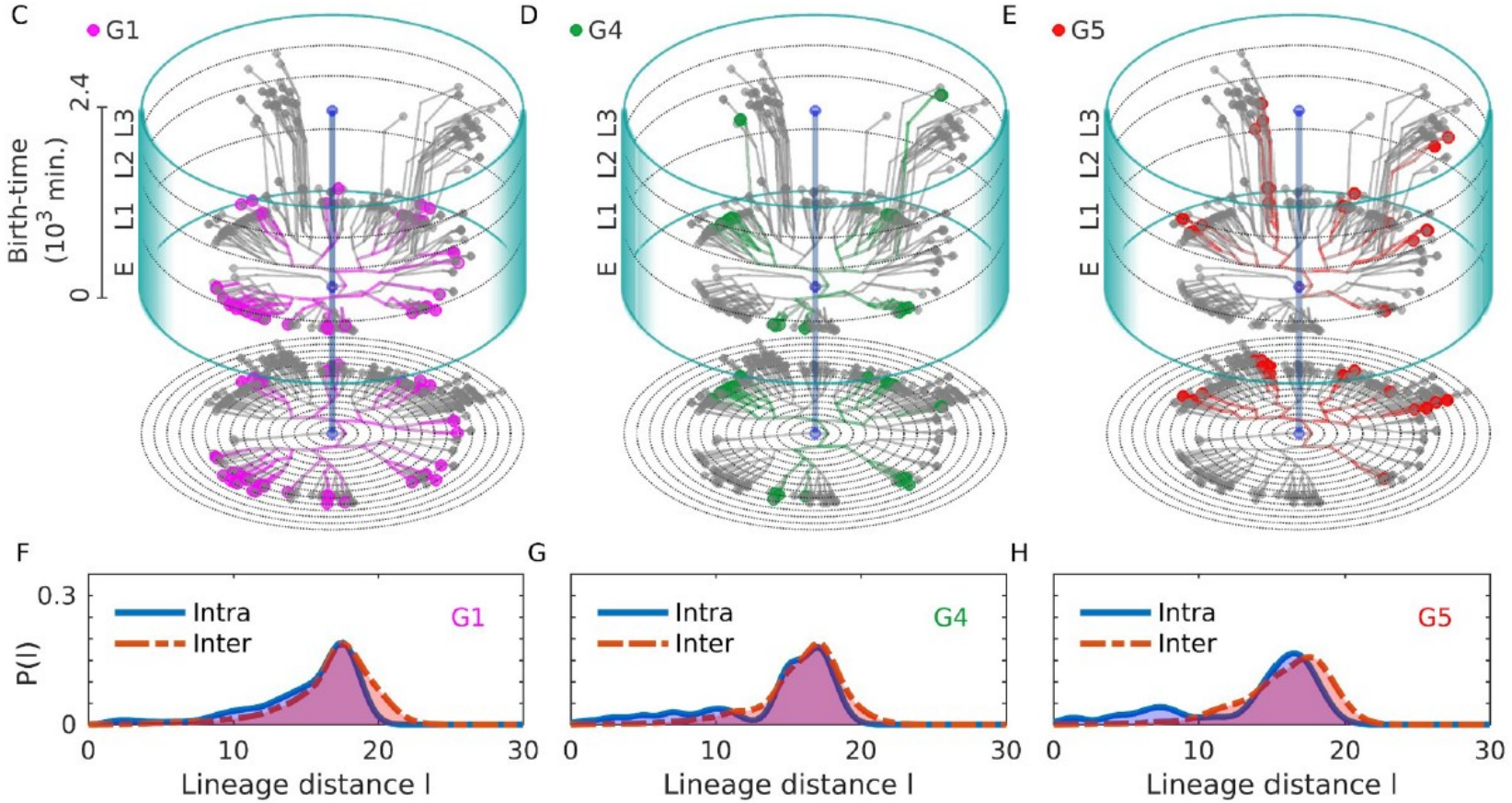
(i) proliferation occurs along both branches,
(ii) only one branch (the other branch leading to either apoptosis or a nonneural cell fate), and,
(iii) no branching → a terminal node (i.e., the cell differentiates into a neuron).

Estimation of P_1 and P_2 from the empirical lineage tree shows proliferation markedly reduces after rung 10.
Incorporated in the model by having $P_1 = 1$, $P_2 = 0.85$ until rung 9, and $P_1 = 0.25$, $P_2 = 0.2$, afterwards.



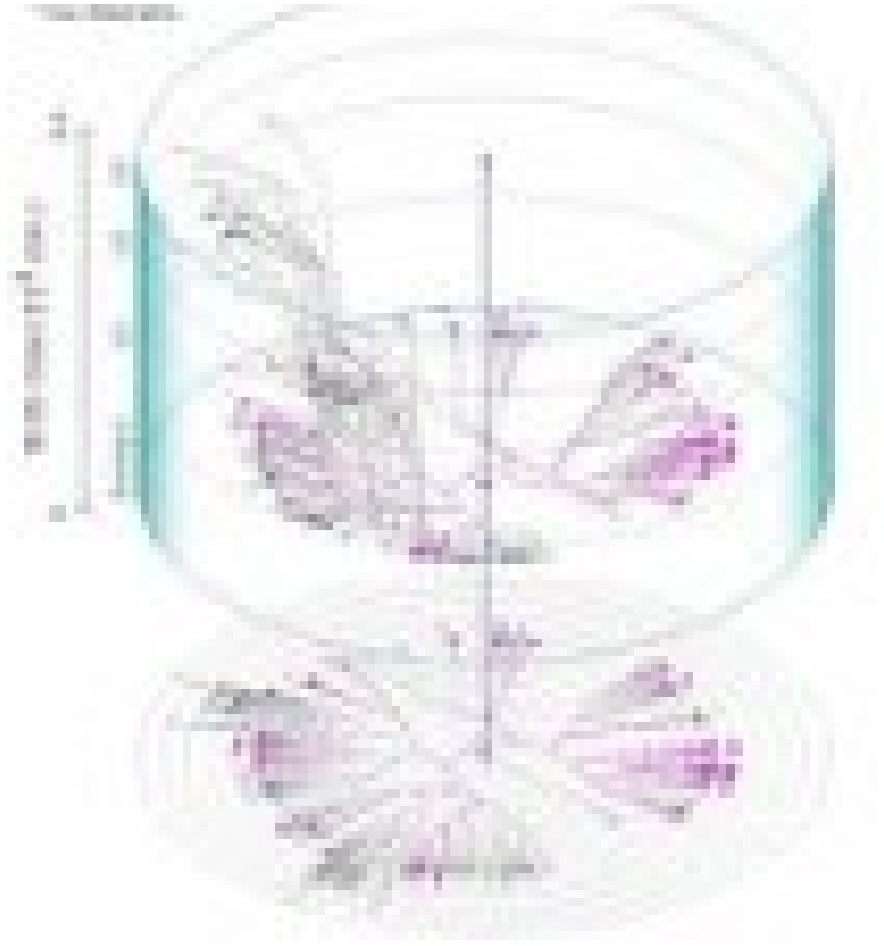


The different ganglia comprise clusters of closely related neurons



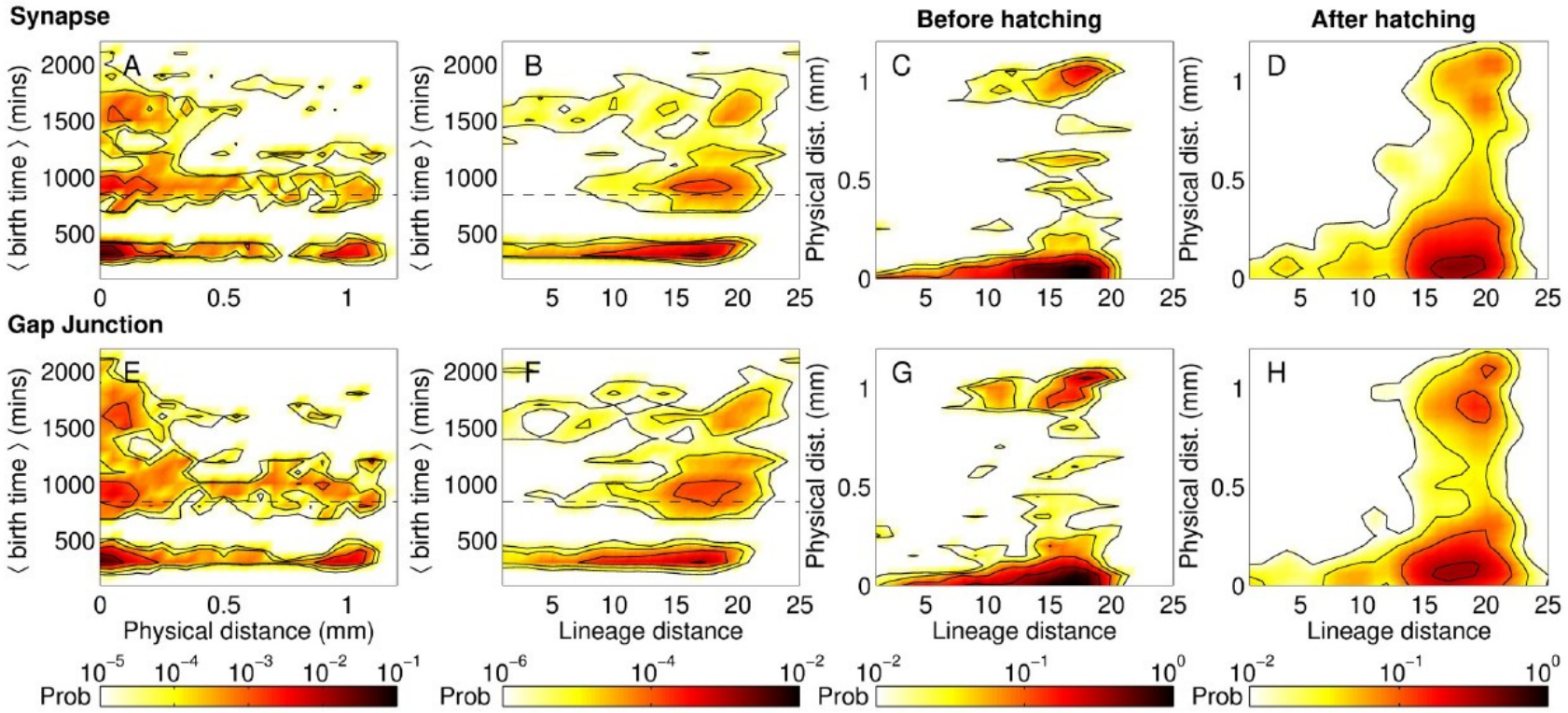
Each ganglion comprises several "families" of neurons emanating from different branches of lineage tree, each family composed of closely related cells sharing a last common ancestor separated from them by only a few cell divisions ⇒ bimodal distribution of l

Chronodendrograms for C elegans ganglia

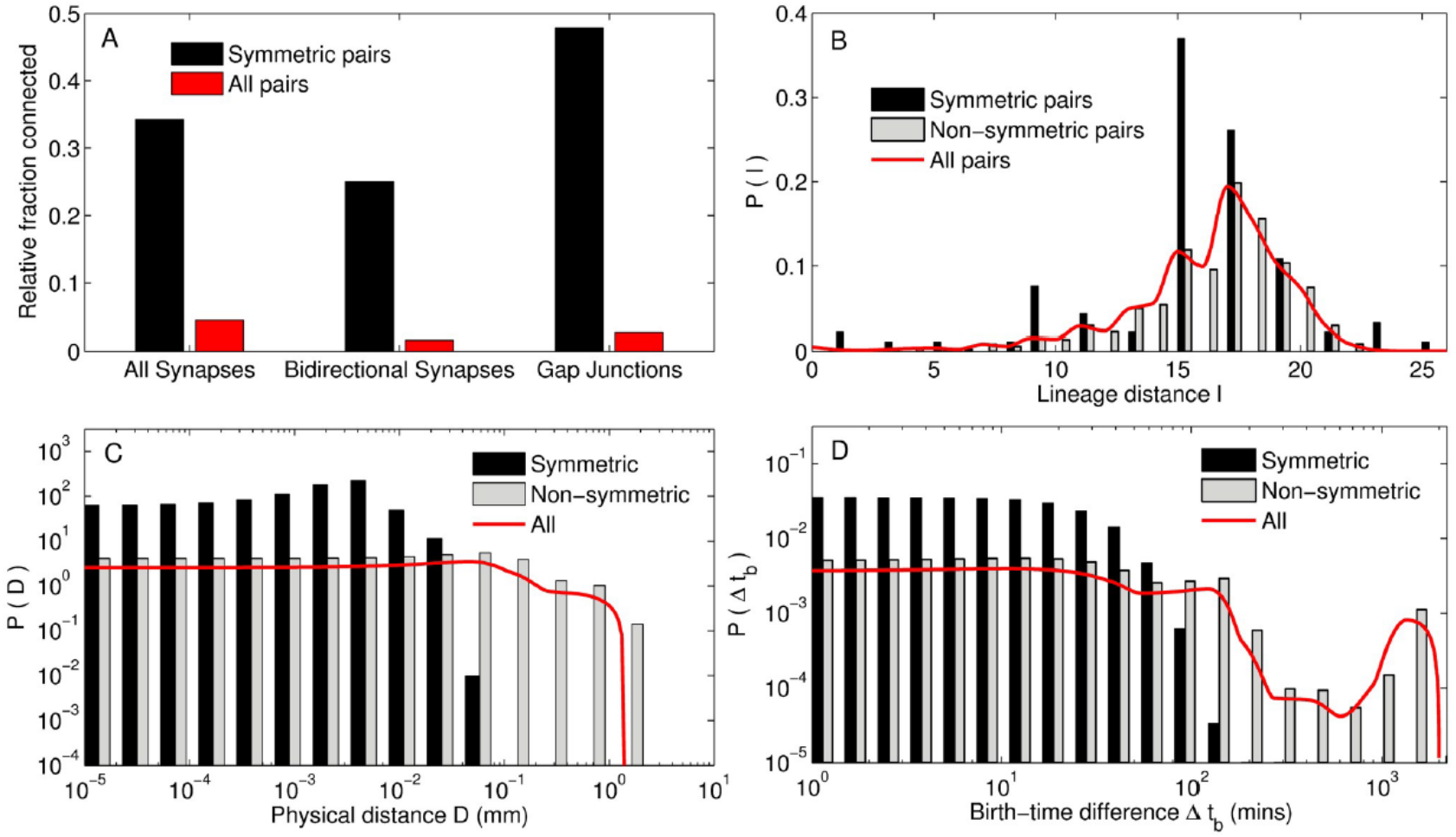


https://www.youtube.com/playlist?list=PLoko8zX-1aMZ5hF9nbIgLqVFm2TT8nxVj

Birth time and lineage relation together constrain the physical distance between cell bodies of connected neurons



Bilateral symmetric pairing homophily



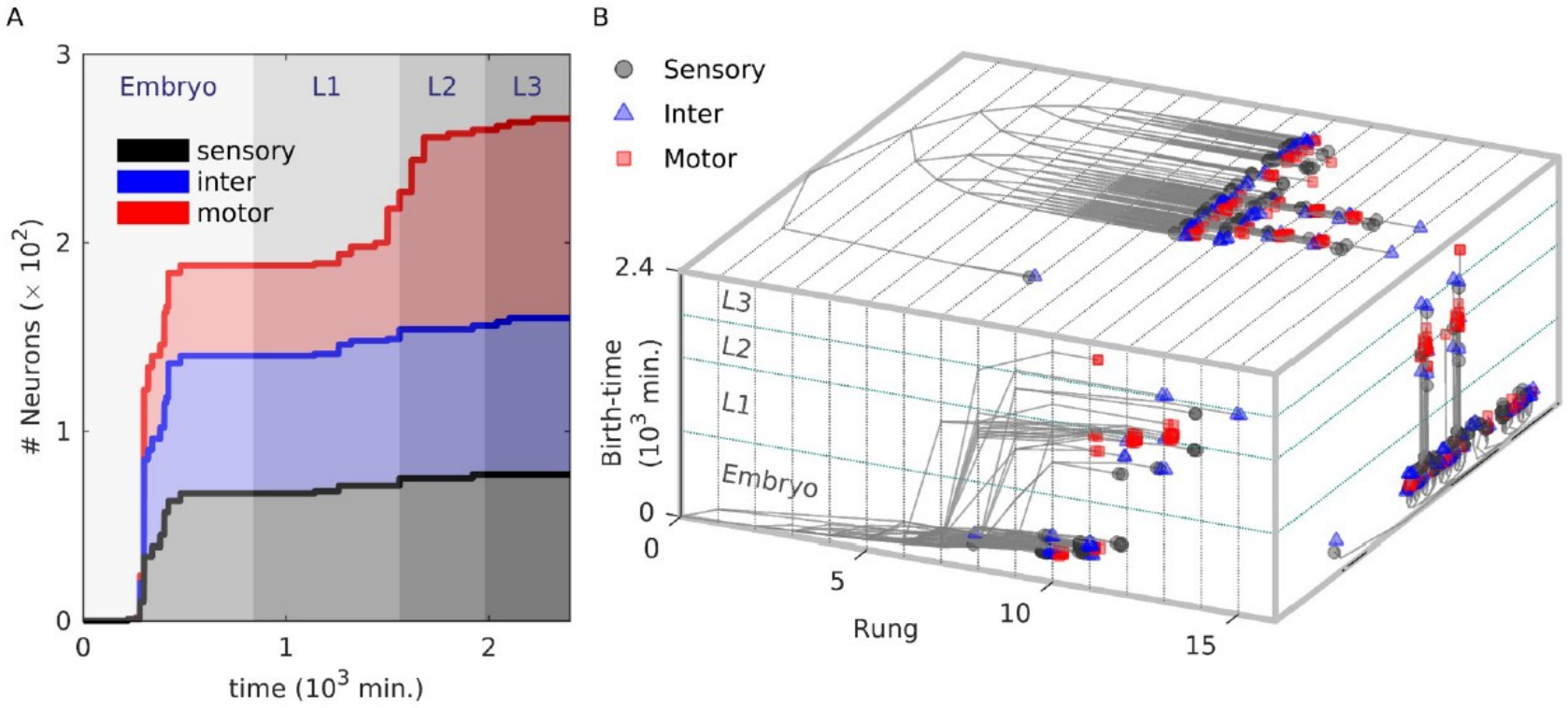
The major fraction (66%) of neurons in the somatic system occur in pairs

Located along the left and right sides of the body in a bilaterally symmetric fashion

Most bilaterally symmetric paired neurons also exhibit strong associations in their physical locations and birth times

Developmental histories of neurons show a bifurcation into early and late branches

Predominance of motor neurons in the late branch



Relative importance of the different types of homophily in determining the network connectivity

Estimated by using logistic regression analysis

Connection probability P between a pair of neurons: F (X_p, X_b, X_l, X_s)
X_p, X_b, X_l, X_s : independent predictor variables corresponding to the four attributes (process length, birth cohort, lineage and symmetric pairing) that show homophily

Synapses
symmetric pairing (β_s^{syn} = 1.78) > birth cohort (β_b^{syn} = 0.71) > process length (β_p^{syn} = 0.35) > lineage (β_l^{syn} = –0.06)
Gap junctions
symmetric pairing (β_s^{gap} = 3.22) > process length (β_p^{gap} = 0.22) > birth cohort (β_b^{gap} = 0.16) > lineage (β_l^{gap} = –0.08)

Magnitude of regression coefficients $\Rightarrow$ extent to which P is changed by altering the value of the corresponding X by a single unit (keeping other predictors unchanged)

E.g., increasing lineage distance between a pair of neurons by 6 units have approximately same effect on P via synapse as the difference in P between neurons belonging to same process length category and different process length categories

Outlook

We view the "wiring problem" for the *C. elegans* nervous system at a level intermediate between

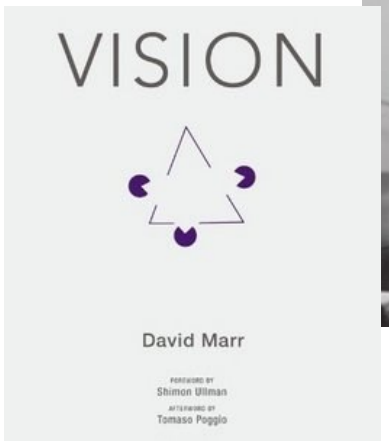
(i) the molecular mechanism-level details of developmental processes (involving diffusible factors, contact mediated interactions, growth cone guidance, etc.) and

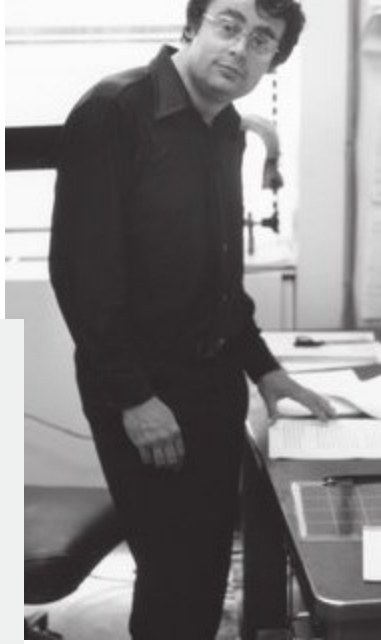
(ii) the resulting structural organization of the entire somatic nervous system

Aim: to uncover a set of guiding principles that govern the wiring and spatial localization of cell bodies, and which are implemented by (genetically encoded) molecular mechanisms

Marr's Three Levels of Analysis

David Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (1982)





David Marr
(1945-1980)

A complex system such as the brain can be understood at multiple levels of abstraction/generality

- **Goals & strategies level:**
  what the system does and why ?

- **Algorithmic level:**
  what are the steps by which the goals are achieved ?

- **Implementation or Hardware level:**
  how are the representations/computations physically realized ?

Outlook

We view the "wiring problem" for the *C. elegans* nervous system at a level intermediate between

(i) the molecular mechanism-level details of developmental processes (involving diffusible factors, contact mediated interactions, growth cone guidance, etc.) and

(ii) the resulting structural organization of the entire somatic nervous system

Aim: to uncover a set of guiding principles that govern the wiring and spatial localization of cell bodies, and which are implemented by (genetically encoded) molecular mechanisms

The principles are *strategies* for achieving specific network designs realized over the course of development

Future challenge: Delineating exactly how these governing principles (the various types of homophily) are implemented by molecular mechanisms

Thanks

Work done in collaboration with

Anand Pathak

(IMSc PhD, 2020)

Dartmouth College, NH

Nivedita Chatterjee

Vision Research Foundation, Shankar Nethralay, Chennai

Raj Kumar Pan

(IMSc PhD, 2010)

InlineMarket, Helsinki

A Pathak, N Chatterjee and SS (2020) Developmental trajectory of *Caenorhabditis elegans* nervous system governs its structural organization. *PLoS Comput Biol* **16**(1): e1007602.

R K Pan, N Chatterjee and SS (2010) Mesoscopic organization reveals the constraints governing *Caenorhabditis elegans* nervous system. *PLoS ONE* **5**(2): e9240