Neurotree: An academic genealogy of neuroscience

Stephen V. David[1], Benjamin Y. Hayden[2]
[1]Institute for Systems Research, University of Maryland, College Park
[2]Department of Neurobiology, Duke University

1. Abstract

Neurotree (http://neurotree.org) is an actively growing online genealogy of neuroscience. The website currently includes mentor-trainee relationships between over 4500 neuroscientists. This number continues to grow as interested members of the neuroscience community make contributions. Training relationships are presented in a tree format typical of genealogical diagrams, allowing users to navigate the "family tree" intuitively.

Neurotree is publicly accessible. Any user can add information to the web site, and any user who requests editing privileges can edit existing information.

Neurotree serves several purposes. We have traced several mentorship lines back over more than two centuries, making the web site a valuable source of historical information. Neurotree also allows active researchers to link to their own web sites so that interested visitors can learn more about current research. Finally, Neurotree provides data on how mentor relationships span different branches of neuroscience. Here we describe the data set and present some graph-theoretical analyses, focusing on how the data can be clustered and on individual researchers who provide links between clusters.

2. Tree architecture

Neurotree has been developed entirely using open-source software. Data about researchers and the mentor relationships between them are stored in a relational database (MySQL), and the web interface is generated dynamically by a scripting system that operates on the web server (apache/PHP).

People	Connections
pid (unique #) first name middle name last name institution research area photo URL home page ... + technical details (added by, date added, etc)	cid (unique #) pid1 pid2 connection type (grad student, post-doc, etc.) institution start year stop year ... + technical details (added by, date added, etc)

The database consists of two basic tables, **people** and **connections**. Each mentor relationship stored as a row in the connections table. Pointers (pid1 and pid2) index the two people involved in the relationship, and the connection type variable specifies the nature of the relationship (graduate student, postdoctoral researcher, etc.).

Auxiliary tables are used to link neuroscientists based on research areas and institutions for indexing and searching.

Fun neurogenealogy fact: Famous cousins

Did you know that the following historical figures were academic cousins?
Ivan Pavlov and Sigmund Freud (through Johannes Muller)
Donald Hebb and Stephen Kuffler (through Charles Sherrington)
Max Wertheimer and Alois Alzheimer (through Wilhelm Wundt)
Franz Nissl and Karl Lashley (2nd cousins through Wundt)

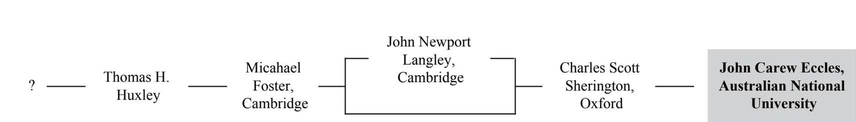
3. Example Neurotree entry

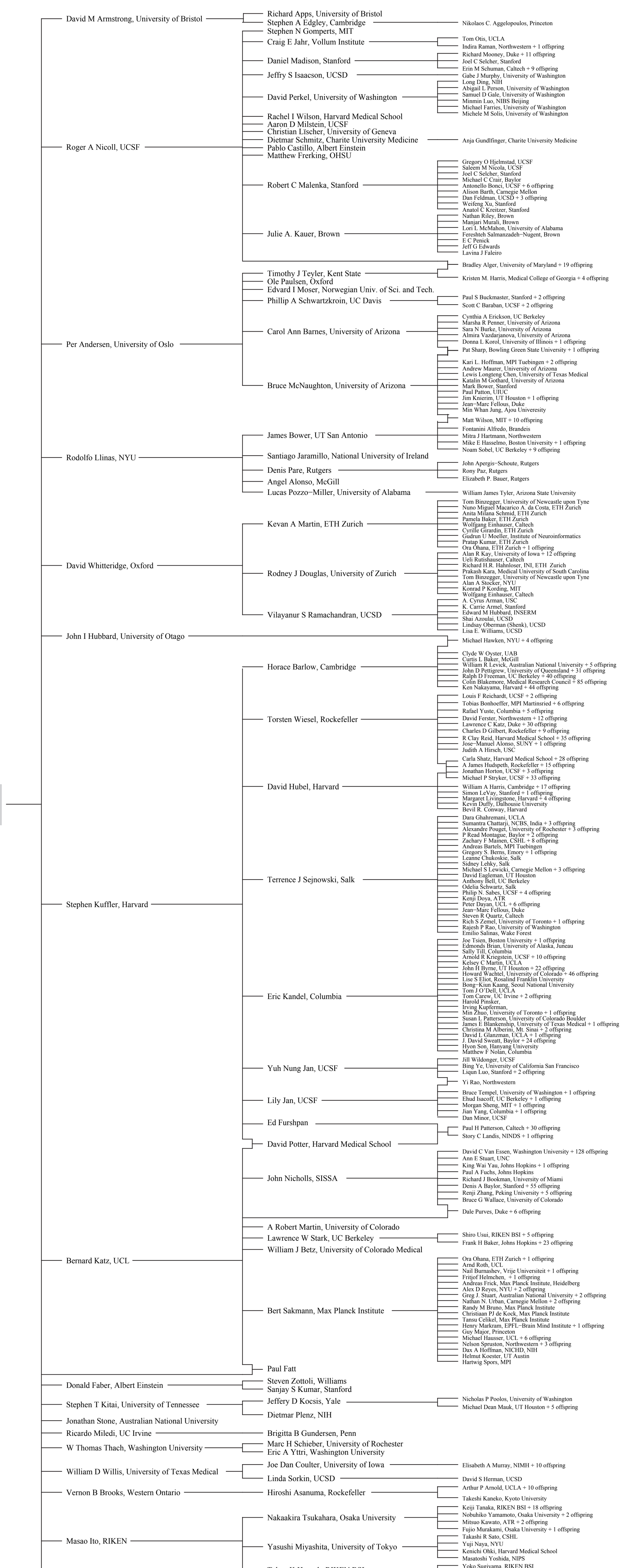
The influential physiologist John Eccles trained a large number of neuroscientists who themselves went on to train many others. We use Eccles as an example to illustrate the interface and information available in Neurotree.

Screenshot: search page

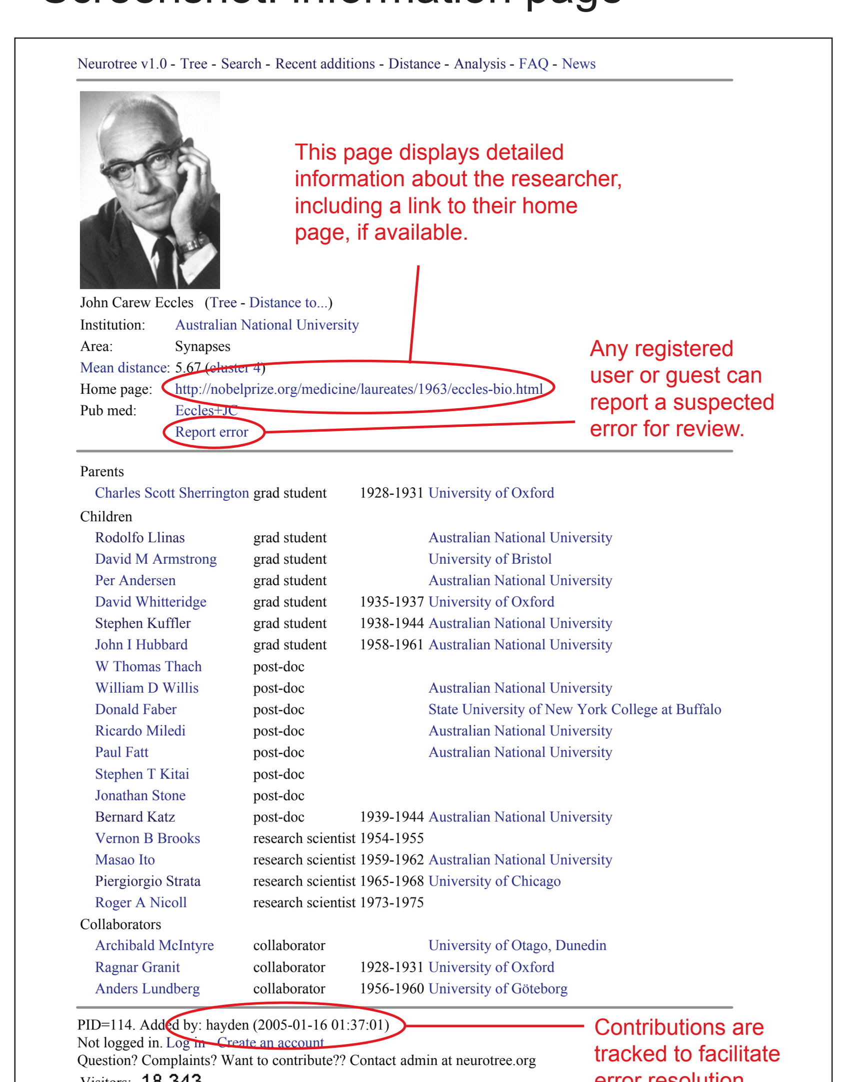
Family tree

The ancestors and descendants of any Neurotree entry can be displayed in a tree diagram. The tree shown here has been modified from the on-line version to take advantage of the denser presentation capabilities of print.





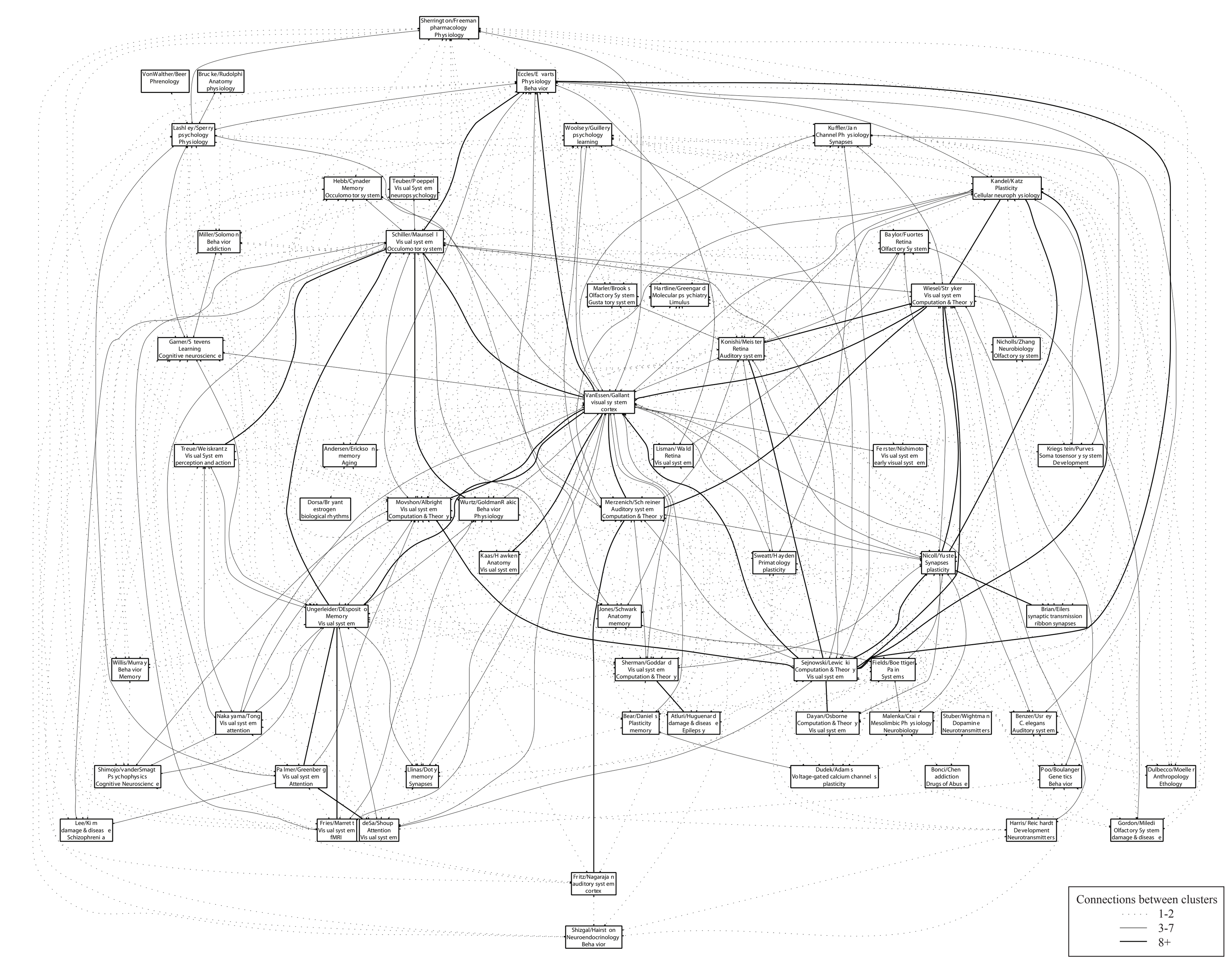
Screenshot: information page



4. Cluster analysis

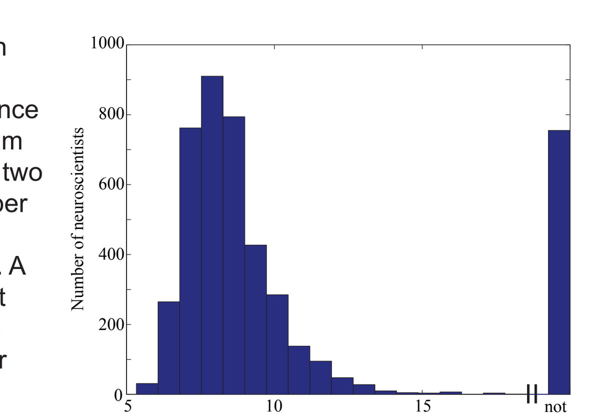
In order to identify major groups in Neurotree, we treated the tree as a graph with nodes defined as people and edges defined as mentor relationships (each with equal weight). Clusters were generated by a partitioning algorithm that uses spectral factorization (C. Lim, S. Bohacek, J. Hespanha, K. Obraczka. Hierarchical max-flow routing. In *Proc of the IEEE GLOBECOM*, Nov. 2005.) to minimize distortion as individual nodes are collapsed into aggregate clusters.

This cluster map was drawn using Graphviz (graphviz.org). Each cluster is a box labeled with the two member nodes with the strongest average connection strength across the whole tree and by the two most common research areas in that cluster. Clusters are sorted from top to bottom roughly from oldest to youngest. Line weight indicates the number of mentor connections between clusters.



5. Central nodes

Each Neurotree node (i.e., each neuroscientist) can be characterized by its mean distance from every other node (histogram at right). The distance between two nodes corresponds to the number of mentorship steps required to travel between them in the tree. A large number of nodes have not yet been connected to the main tree. These are shown at the far right of the histogram.

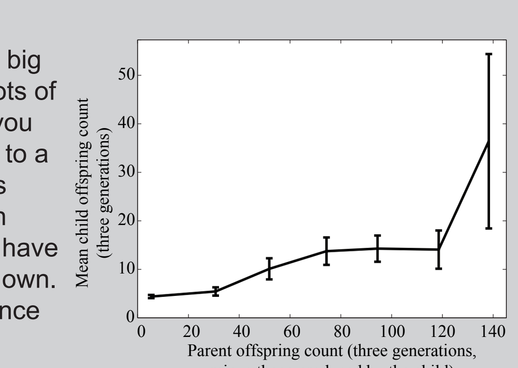


Below are the 50 nodes with shortest mean distance. The list contains a strong bias toward systems and, in particular, the visual system. This suggests either that visual neuroscientists are highly promiscuous or that the population of the tree is biased by having been started in a vision lab. This question will only be answered with more data!

1 (5.33)	Peter Schiller - Visual system - MIT	26 (5.96)	R Clay Reid - Visual system - Harvard Medical School
2 (5.39)	Terrence J Sejnowski - Computation & theory - Salk	27 (5.98)	Wolf Singer - Visual system - MPI Frankfurt
3 (5.45)	Stephen Kuffler - Visual system - Harvard	28 (6.00)	Vernon Mountcastle - Somatosensory system - Hopkins
4 (5.50)	Torsten Wiesel - Visual system - Rockefeller	29 (6.04)	Nikos K Logothetis - Visual system - MPI Tuebingen
5 (5.54)	Michael P Stryker - Development, visual system - UCSF	30 (6.06)	Leanne Chukoskie - Oculomotor system - Salk
6 (5.59)	John HR Maunsell - Visual system - Harvard Medical School	31 (6.06)	Yang Dan - Visual Cortex - UC Berkeley
7 (5.64)	David C Van Essen - Visual system - Washington University	32 (6.08)	Robert Desimone - Systems - MIT
8 (5.67)	John Carew Eccles - Synapses - Australian National University	33 (6.08)	Ken Nakayama - Visual system - Harvard
9 (5.70)	Mark Konishi - Auditory system - Caltech	34 (6.09)	Carla Shatz - Development - Harvard Medical School
10 (5.74)	Leslie Ungerleider - Visual system, cognition - NIMH	35 (6.09)	Charles Gross - Visual system - Princeton
11 (5.79)	Jack L Gallant - Visual system - UC Berkeley	36 (6.10)	Eric I Knudsen - Auditory brainstem - Stanford
12 (5.80)	Michael M Merzenich - Auditory system, plasticity - UCSF	37 (6.11)	Andreas Bartels - Visual system - MPI Tuebingen
13 (5.80)	James A Mazer - Visual system - Yale	38 (6.11)	Ralph D Freeman - Visual system - UC Berkeley
14 (5.80)	David Hubel - Visual system - Harvard	39 (6.12)	Mriganka Sur - Development, visual system - MIT
15 (5.84)	Stephen G Lisberger - Oculomotor system - UCSF	40 (6.13)	John Nicholls - Regeneration - SISSA
16 (5.86)	Horace Barlow - Computation & theory - Cambridge	41 (6.14)	Edward M Callaway - Visual cortex - Salk
17 (5.87)	Sidney Lehky - Visual system - Salk	42 (6.14)	J Anthony Movshon - Visual system - NYU
18 (5.87)	William T Newsome - Visual system, decision - Stanford	43 (6.15)	Mark D'Esposito - Executive control - UC Berkeley
19 (5.89)	Hans-Lukas Teuber - Neuropsychology - MIT	44 (6.16)	Roger A Nicoll - Neurobiology - UCSF
20 (5.89)	Lawrence C Katz - Sensory systems - Duke	45 (6.16)	Thomas D Albright - Visual system - Salk
21 (5.90)	Mortimer Mishkin - Systems - NIMH	46 (6.17)	Philip N. Sabes - Sensorimotor control - UCSF
22 (5.91)	Eric Kandel - Learning and memory - Columbia	47 (6.17)	Zachary F Mainen - Olfaction - CSHL
23 (5.92)	Richard A Andersen - Systems - Caltech	48 (6.17)	Michael N Shadlen - Decision & reward - Washington
24 (5.93)	Robert H Wurtz - Oculomotor system - NEI	49 (6.18)	Bernard Katz - Synaptic physiology - UCL
25 (5.96)	Otto D Creutzfeldt - Visual system - Kraepelin Institute	50 (6.18)	Sue Corkin - Memory - MIT

Fun neurogenealogy fact: Big families stay big

Is it better to study in a big lab or a small one? If lots of grandchildren is what you want, you should head to a big lab. Neuroscientists trained by mentors with many offspring tend to have many offspring of their own. But note the high variance for very large labs!



Acknowledgments

We thank Michael Wu for his graphic design work and, of course, the many contributors (especially hanks, JLand52, kathleen, cab, brian, lneedleman, kwoonwong, jgolearly) who continue to build Neurotree into a resource of increasing value.

NeuroTree