Three Controversial Hypotheses Regarding Primate Brains

Thomas Dean
Google

Stanford University. CS379C: *Computational Models of the Neocortex*. April, 2012.

Tech Report CS-12-01

Three Controversial Hypotheses Concerning Computation in the Primate Cortex

Thomas Dean, Greg S. Corrado and Jonathon Shlens

January 2012

Abstract:

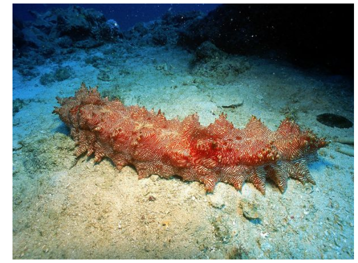
We consider three hypotheses concerning the primate neocortex which have influenced computational neuroscience in recent years. Is the mind modular in terms of its being profitably described as a collection of relatively independent functional units? Does the regular structure of the cortex imply a single algorithm at work, operating on many different inputs in parallel? Can the cognitive differences between humans and our closest primate relatives be explained in terms of a scalable cortical architecture? We bring to bear diverse sources of evidence to argue that the answers to each of these questions - with some judicious qualifications - are in the affirmative. In particular, we argue that while our higher cognitive functions may interact in a complicated fashion, many of the component functions operate through well-defined interfaces and, perhaps more important, are built on a neural substrate that scales easily under the control of a modular genetic architecture. Processing in the primary sensory cortices seem amenable to similar algorithmic principles, and, even for those cases where alternative principles are at play, the regular structure of cortex allows the same or greater advantages as the architecture scales. Similar genetic machinery to that used by nature to scale body plans has apparently been applied to scale cortical computations. The resulting replicated computing units can be used to build larger working memory and support deeper recursions needed to qualitatively improve our abilities to handle language, abstraction and social interaction.

(complete text in pdf)

www.cs.brown.edu/research/pubs/techreports/reports/CS-12-01.html

[1] To appear in the Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI-12), Toronto, Ontario (2012).

What *Holothuria glaberrima* (Wolpert's sea cucumber), *Caenorhabditis elegans* (Brenner's marine worm), *Aplysia californica* (Eric Kandel's sea slug), *Rana pipiens* (Lettvin and Maturana's frog), and *Rattus norvegicus* (Nicolelis' rats) can tell us about *Homo sapiens*?



Controversial Hypotheses

- Modular Minds — encapsulate specific function, evolved specific competences
- Single Algorithm — cortical components all implement the same basic algorithm[1]
- Quantity Suffices — human cognitive capabilities are different from those of other primates due to cortical quantity[2]

[1] This is a claim about *algorithms* — abstract computational recipes — and not about particular biological implementations.

[2] Cortical size is misleading since the real issue concerns whether a systematic increase in the size of the human neocortex yields more powerful computational capabilities that can account for the cognitive and cultural advances of *homo sapiens*.

- Modular Minds — encapsulate specific function, evolved specific competences

Originally proposed by Jerry Fodor in “Modularity of Mind” (1983). Here we refer to a Darwinian variant of Fodor’s idea in which modules are adaptations, the product of natural selection, evolved to underlie a specific cognitive competence.

- Single Algorithm — cortical components all implement the same basic algorithm

Consensus of the three speakers — James DiCarlo, Geoffrey Hinton, and Michael Lewicki — at the symposium entitled "Learning and high-level vision: from single neurons to computational theory” held in March 2010 at Stanford University.

- Quantity Suffices — human cognitive capabilities are different from those of other primates due to cortical quantity

“The difference in mind between man and the higher animals, great as it is, is one of degree not of kind.” — Charles Darwin in *The Descent of Man* which was first published in 1871.

“Take a chimp brain fetally, and let it go for two or three more rounds of cell division, and you get a human brain instead, and out pops symphonies and ideologies and hopscotch and everything else. What that tells is you is that with enough quantity you invent quality." — Stanford Professor Robert Sapolsky in answering a question following his talk entitled "Are Humans Just Another Primate?" at the California Academy of Sciences in February 2011.

- Modular Minds — encapsulate specific function, evolved specific competences

The cells of the eye implement a distinct function? Why wouldn't there be similar functional units in the body and in particular in the brain, say, for language, planning, object recognition, or even social dominance?

- Single Algorithm — cortical components all implement the same basic algorithm

The cortex evolved over a relatively short period of time and exhibits remarkable homogeneity in local structure. It seems plausible that this neural substrate might be exercised by the same basic algorithm running in parallel.

- Quantity Suffices — human cognitive capabilities are different from those of other primates due to cortical quantity

Human cortex has about three times the number of neurons as our closest non-human primate cousins with whom we share 98.8% of our 3B base pairs. There are only a few thousand that could plausibly influence the differences between our species and these seem inadequate to the job of undertaking a significant reprogramming of the brain.

Conclusions: A Preview

- Morphological modularity at the genomic level could enable computational scaling
- Algorithmic parsimony in the neocortex is probably valid at some granularity
- Additional steps of prenatal neurogenesis could increase combinatorial circuit depth

Hypothetically Speaking

- Karl Popper (1902-1994) — “In science you can let your hypotheses die in your stead.”

- George E. P. Box — “Essentially, all theories [hypotheses] are wrong, but some are [actually] *useful*.”

Testable Scientific Theories

- Camillo Golgi (1843-1926) — *reticular theory* — processes of contiguous cells are fused to create a massive network

- Santiago Ramón y Cajal (1852-1934) — *neuron theory* — the nervous system is made up of discrete individual cells

This was later generalized to *cell theory* which posits that all biological systems are made up of discrete individual cells.

Evidence in Neuroscience

- Histology — dissect, slice, stain, inspect, render — Cajal and Golgi to the Human Connectome Project
- Neurology — Harlow's 1848 observations of Phineas Gage to theories of consciousness and emotion
- Neurosurgery — Penfield and Sperry to non-invasive tumor treatment using stereotactic radiotherapy
- Neurophysiology — expose, probe, stimulate, record, analyze — Hubel and Wiesel to optogenetics
- Neuroimaging — Caton & Berger (EEG) to functional Nuclear Magnetic Resonance Imaging (fMRI)
- Evolutionary Neuroscience — Darwin[1] (1809-1882) to modern paleobiology and molecular phylogenetics

[1] Darwin, Charles. *The expression of the emotions in man and animals*. London: John Murray. 1872.

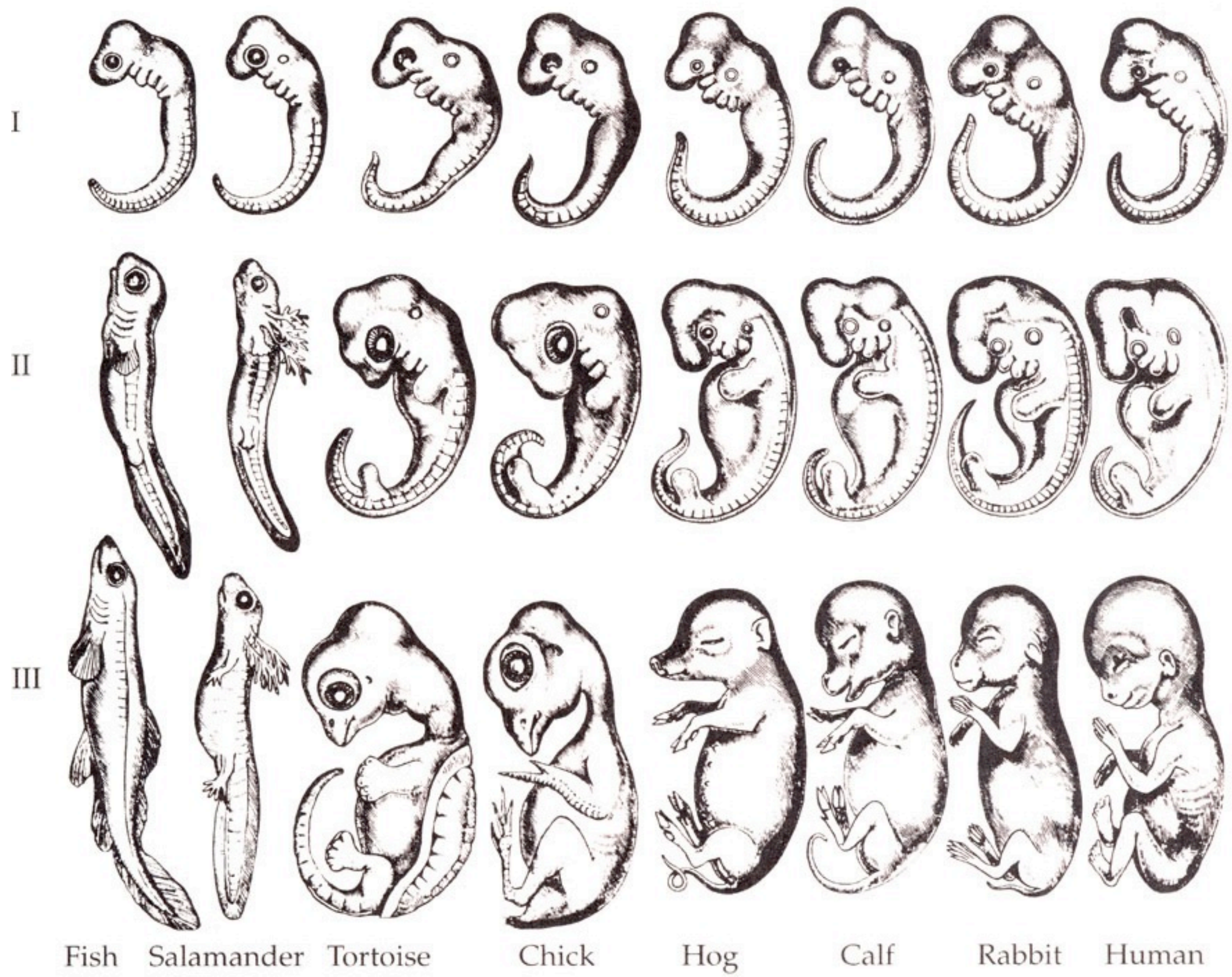
Discredited *Obvious* Theories

- The Triune Brain Theory — posits that the primate brain consists of the reptilian complex, the paleomammalian complex (limbic system), and the neomammalian complex (neocortex).
- Recapitulation Theory — congruence in form between same embryonic stages of different species is evidence that the embryos are repeating the evolutionary stages of their ancestral history.

Developed by Paul D. Maclean (1913-2007) and popularized in Carl Sagan's 1977 book *The Dragons of Eden.*

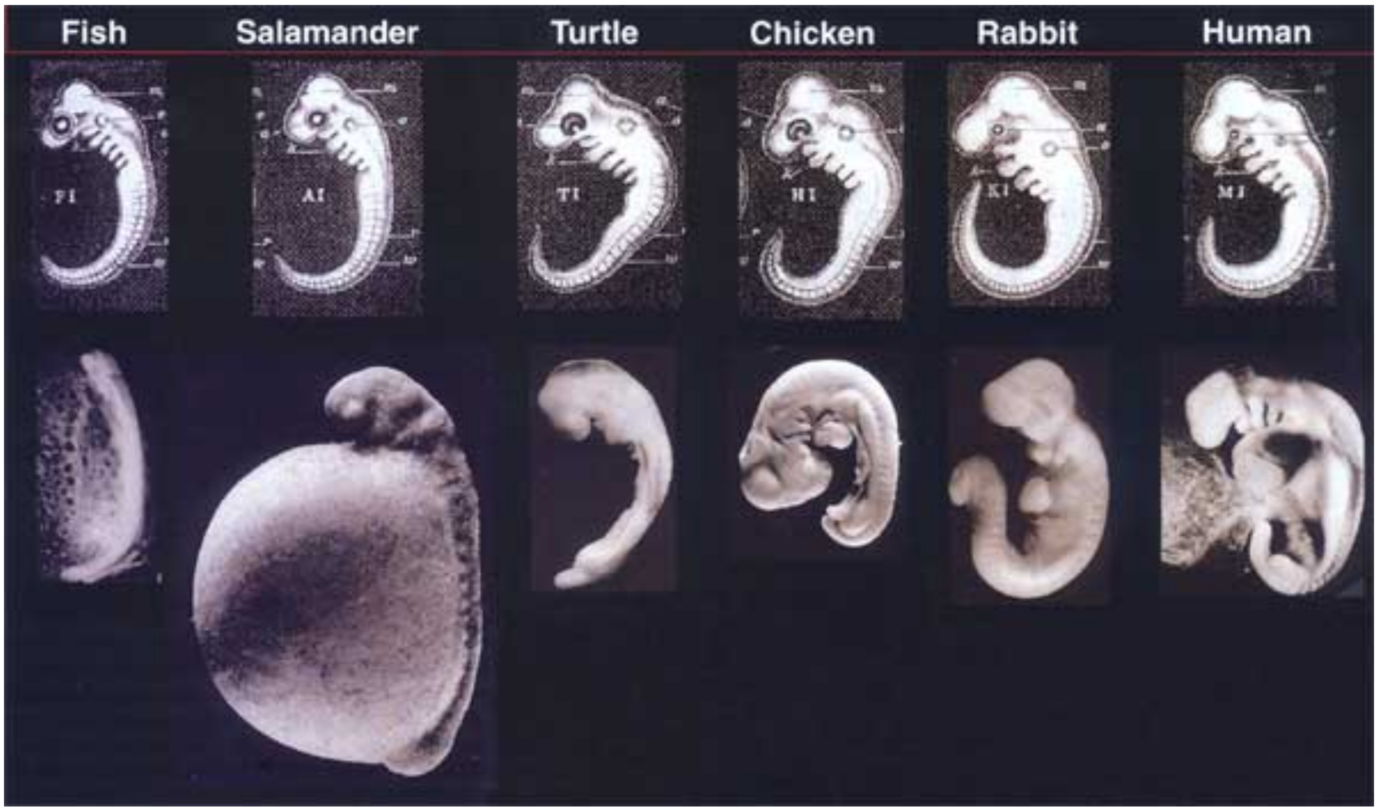
Developed by Ernst Haeckel (1834-1919) and summarized by the pithy phrase "ontogeny recapitulates phylogeny."

False Theories



Ernst Haeckel, *The History of Creation* (1868), translated by E. Ray Lankester, Kegan Paul, Trench & Co., London, 1883.

Falsified (Cooked) Evidence

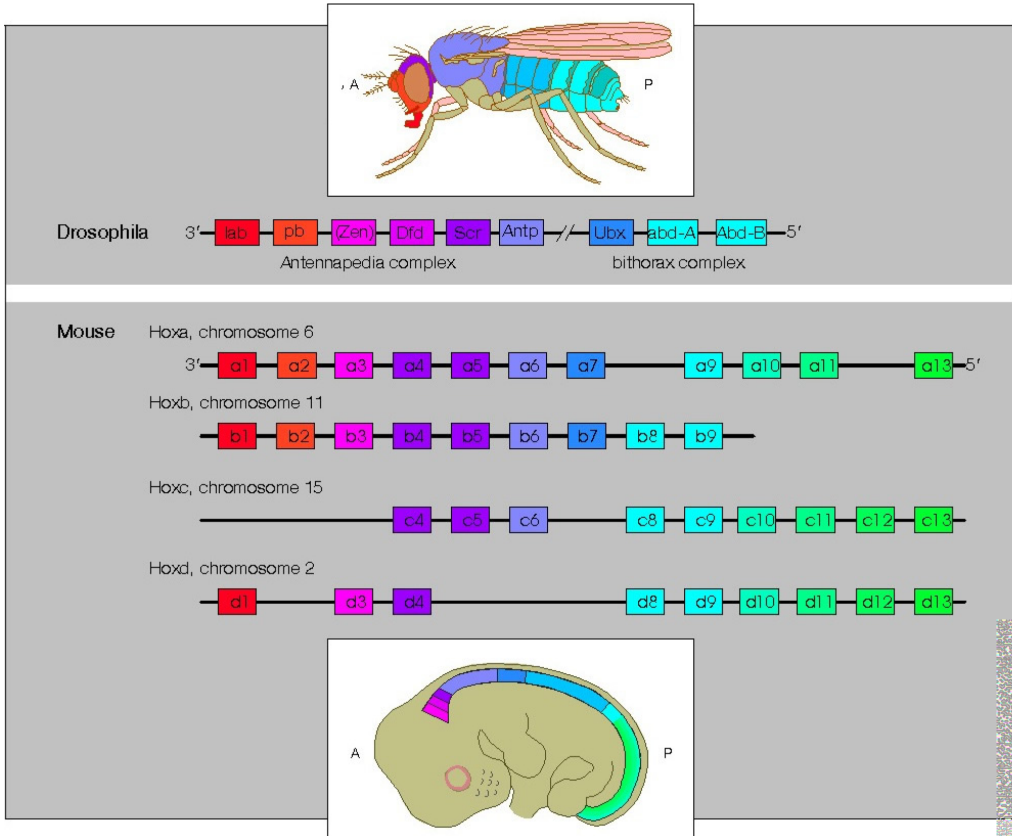


Robert J. Richards. Ernst Haeckel and the Struggles over Evolution and Religion. *Annals of the History and Philosophy of Biology*, Volume 10: 89-115. 2005.

Michael K. Richardson. Haeckel's Embryos Continued. *Science* 281 (5381): 1285-1289. 1998.

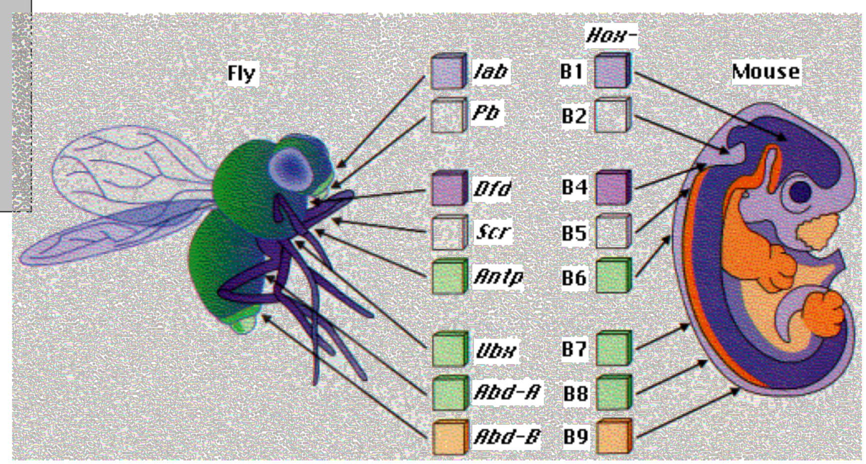
Michael K. Richardson and Gerhard Keuck. Haeckel's ABC of Evolution and Development. *Biological Reviews of the Cambridge Philosophical Society*, 77: 495-528. 2002.

Stranger Than Fiction

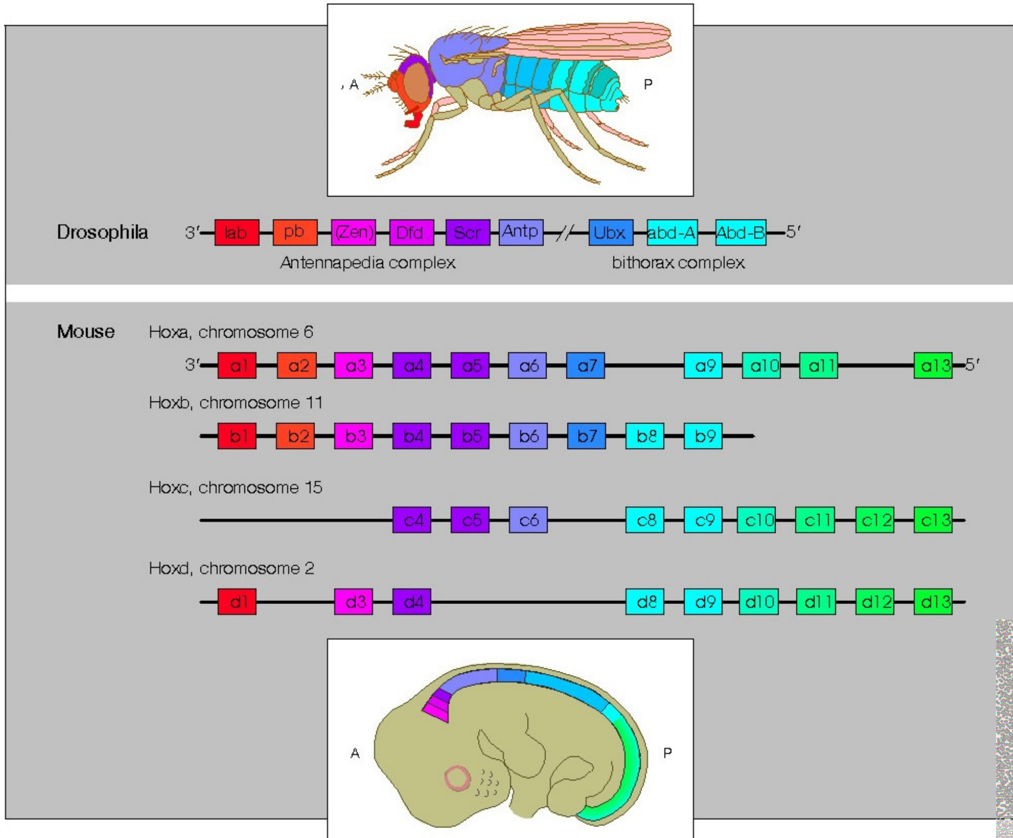


"What is true for *E. coli* is also true for the elephant."
— Jacques Monod (1910-1976)

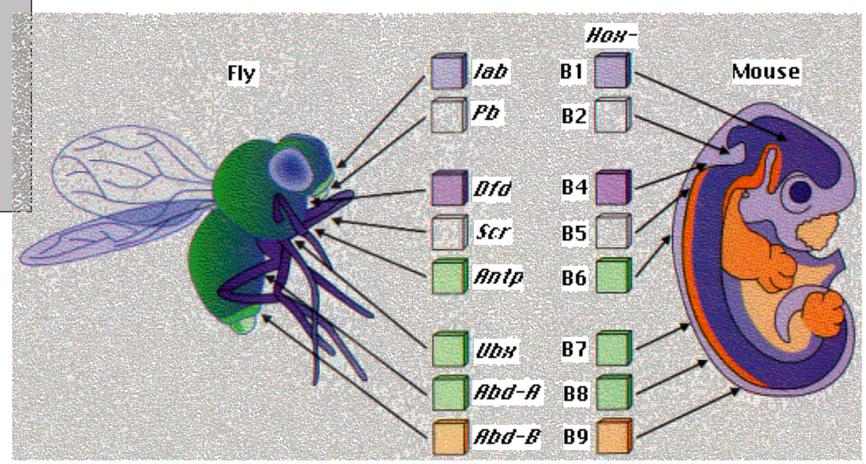
W. McGinnis, R. L. Garber, J. Wirz, A. Kuroiwa, and W. J. Gehring. A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. Cell, 37(2):403–408, 1984.



Morphological Modularity



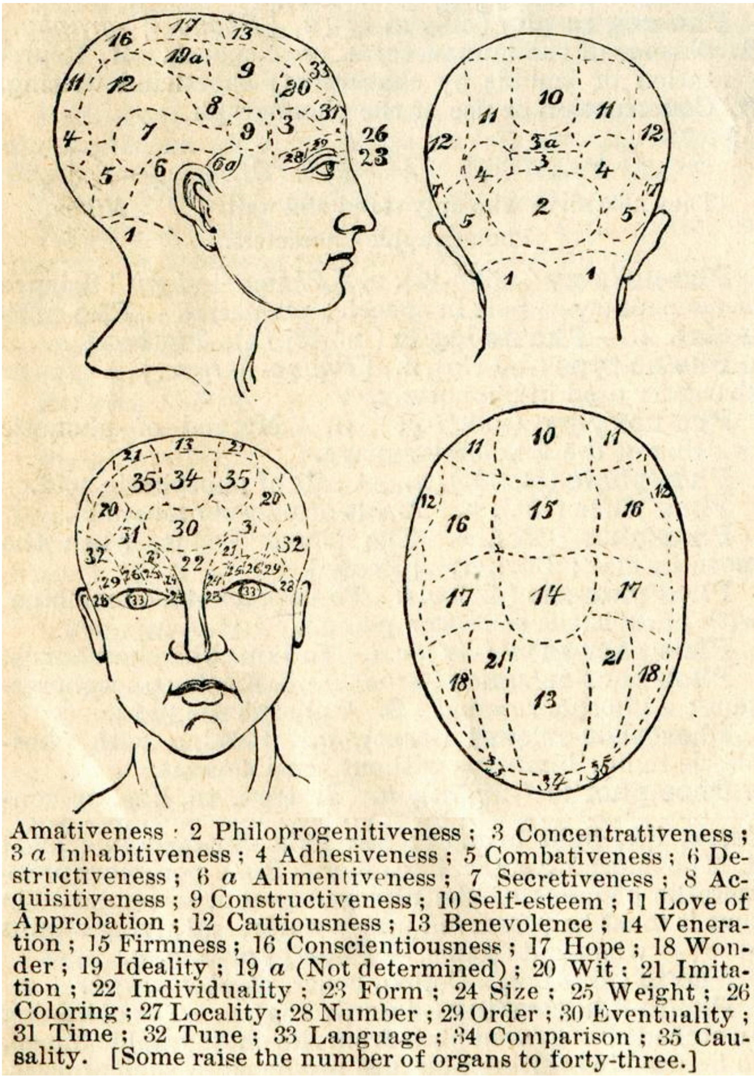
Animal bodies are modular in the engineering sense: they are divided into parts that can be developed and operated independently.

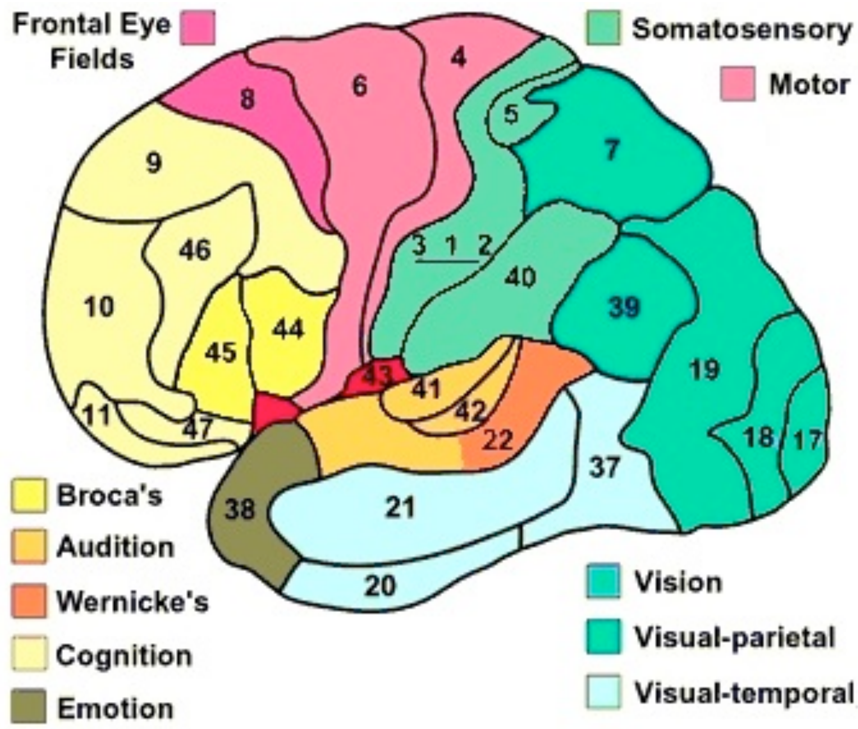


W. McGinnis, R. L. Garber, J. Wirz, A. Kuroiwa, and W. J. Gehring. A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. Cell, 37(2):403–408, 1984.

V.A. Casagrande and G. Purushothaman. Modularity. Encyclopedia of Perception, Volume 1, 561–566. 2009.

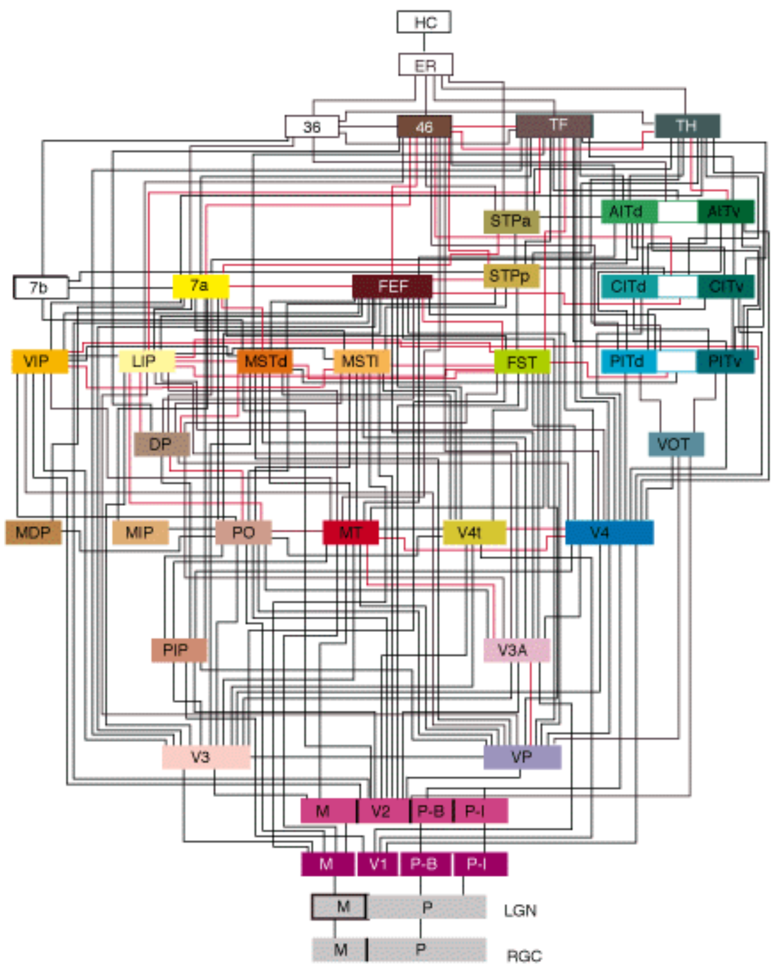
Franz Gall to Korbinian Brodmann

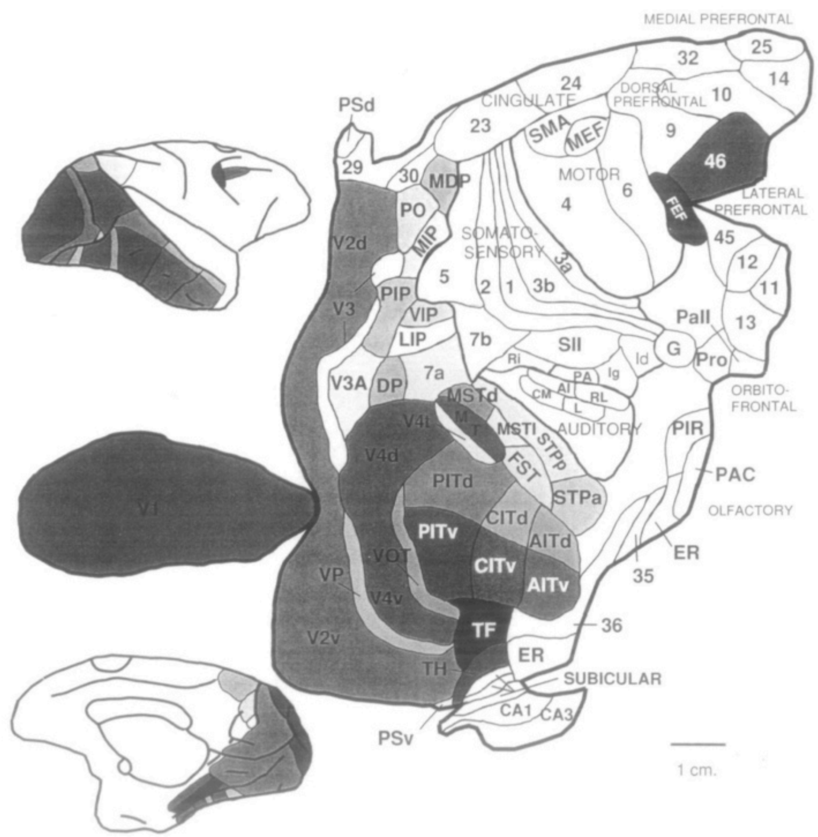




Korbinian Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*, Johann Ambrosius Barth Verlag, Leipzig, 1909.

Primate Cerebral Cortex





Felleman, D. J. and Van Essen, D. C. *Distributed hierarchical processing in primate cerebral cortex*. Cerebral Cortex, 1:1-47, 1991.

Evidence for Modularity

- Paul Broca (1824-1880) — Speech Production (Broca's Area) — in the posterior inferior frontal gyrus of the dominant hemisphere (left in 90% of humans).
- Carl Wernicke (1848-1905) — Language Comprehension (Wernicke's Area) — in the superior temporal gyrus of the dominant hemisphere.
- Patients with damaged Wernicke's area but intact Broca's area produce elaborate, syntactically-correct sentences devoid of meaning.
- In patients with a damaged Broca's area but intact Wernicke's area, meaning is preserved, but their sentences exhibit no syntactic deep structure, *e.g.*, patients exhibit little or no recursive embedding.

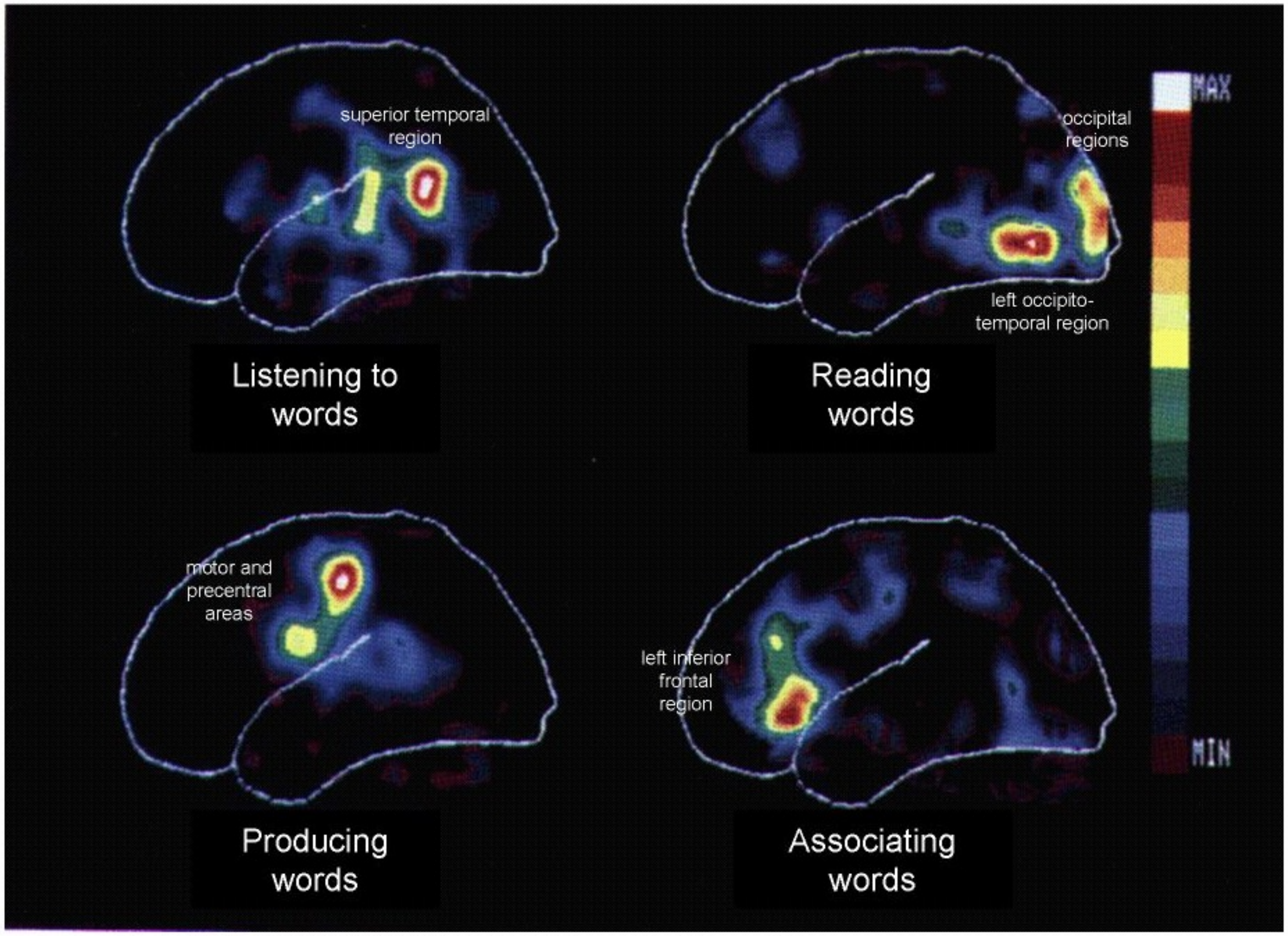


Figure 2.3. A historical image: the brain areas for language, as first revealed by PET scanning (data from Petersen et al., 1989; image courtesy of Marcus Raichle). Relative to the fixation of a small dot, silent reading (top right) activates processes of visual word recognition located in the rear part of the left hemisphere. Depending on the task, information is then transmitted to regions coding for speech sounds (top left), speech production (bottom left), or the manipulation of word meanings (bottom right).

Stanislas Dehaene. *Reading in the Brain: The Science and Evolution of a Human Invention*. Viking Press, 2009.

Evidence Against Modularity

- Language is late evolutionarily speaking but what about older functions such as those involved in emotion, primitive self awareness — what Antonio Damasio calls the *protoself*, basic object and face recognition, reading, planning, social functions such as determining status or whether someone is being truthful.
- Language isn’t really a safe bet either.

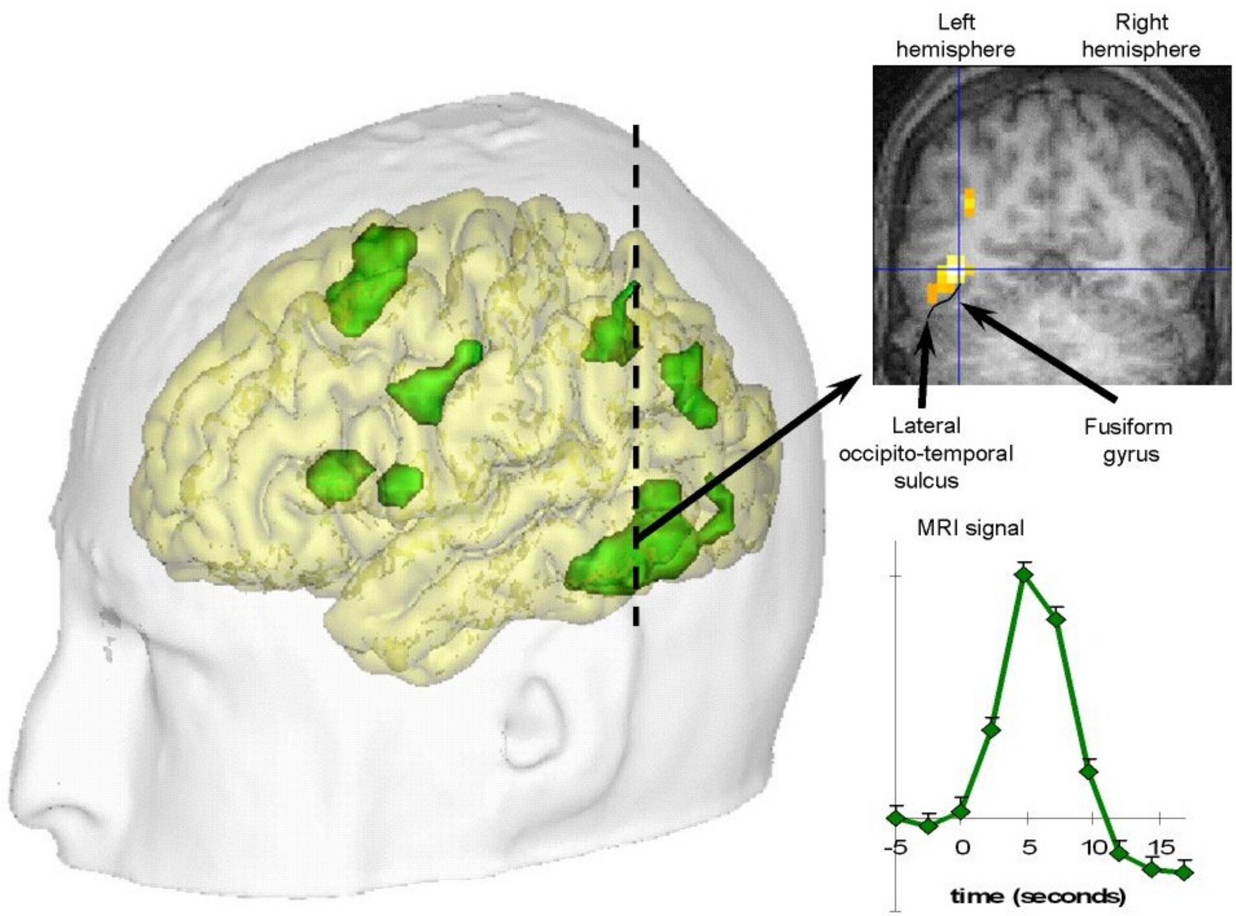
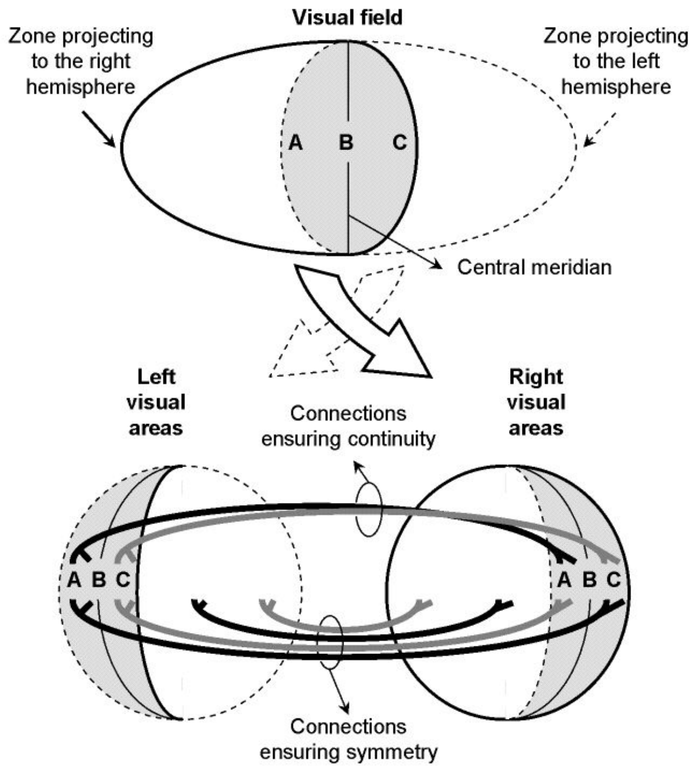
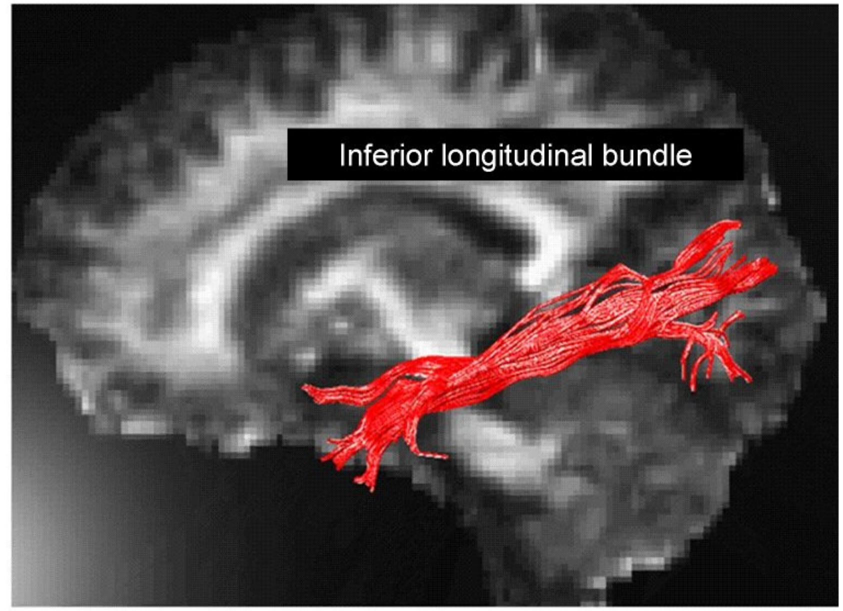


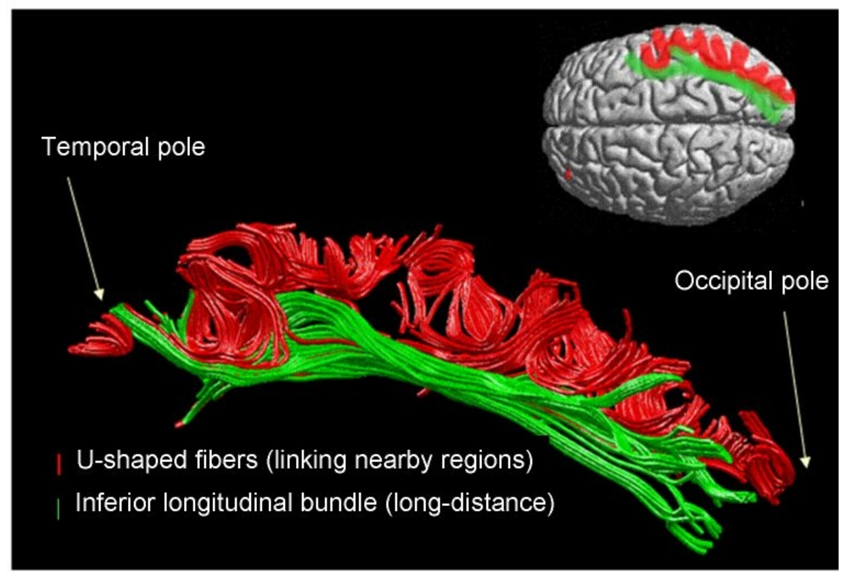
Figure 2.4. Functional magnetic resonance imaging (fMRI) can locate the brain areas involved in reading in just a few minutes. Participants read words presented at random intervals. After each word, reading areas show a characteristic increase in MRI signal which reaches a peak about five seconds later. The active network varies depending on the exact task and the nature of the control state. However, it always includes the visual word form area, the "brain's letterbox". This region is systematically located deep in the left lateral occipito-temporal sulcus, next to the fusiform gyrus.

Stanislas Dehaene. *Reading in the Brain: The Science and Evolution of a Human Invention*. Viking Press, 2009.





Diffusion Tensor Magnetic Resonance Imaging



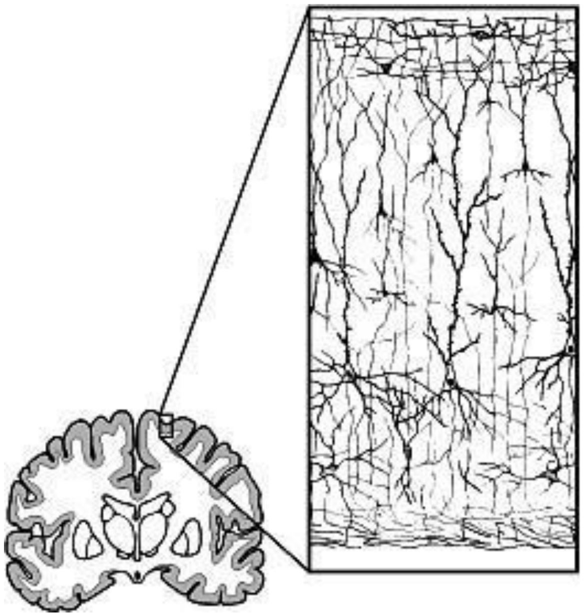
Stanislas Dehaene. *Reading in the Brain: The Science and Evolution of a Human Invention*. Viking Press, 2009.

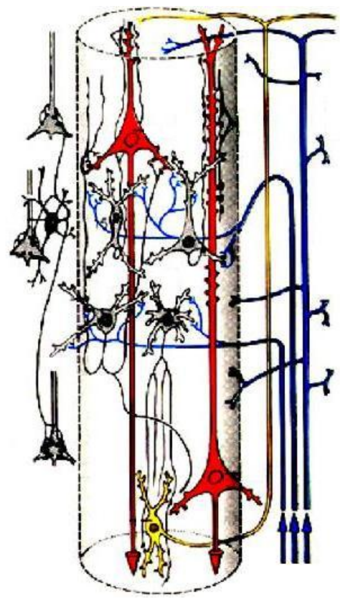
Single Algorithm Hypothesis

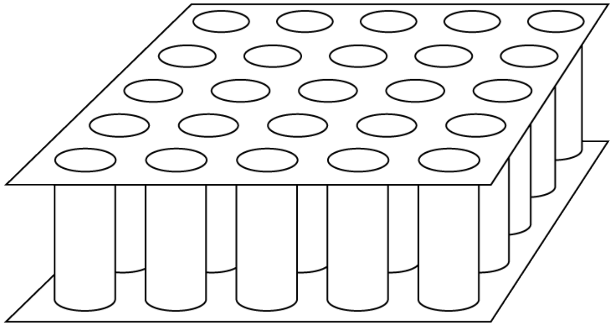
- Evidence against a uniform substrate
- Evidence for homogeneous substrate
- What do we mean by an *algorithm*?
- Does it involve more than a logic gate?
- Is it *just* a variant of Hebbian learning?
- Does it include *early* development?

Donald O. Hebb. *The organization of behavior*. New York: Wiley & Sons, 1949.

Terminology: Cortical Columns

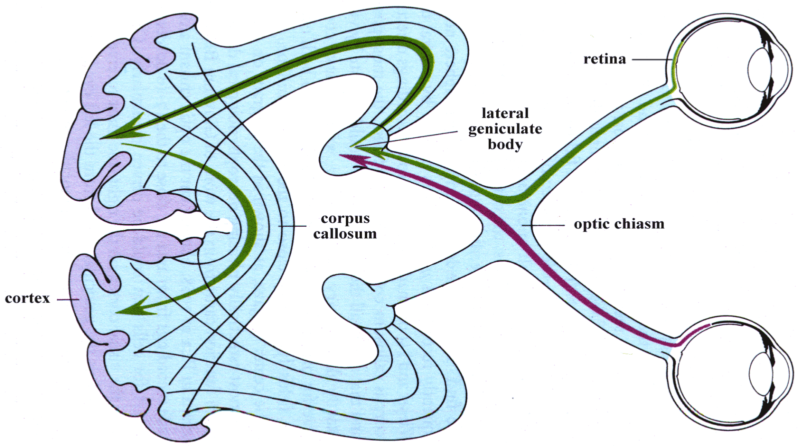


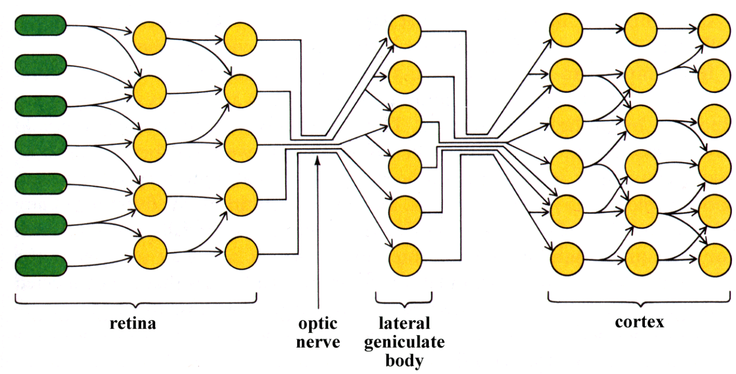


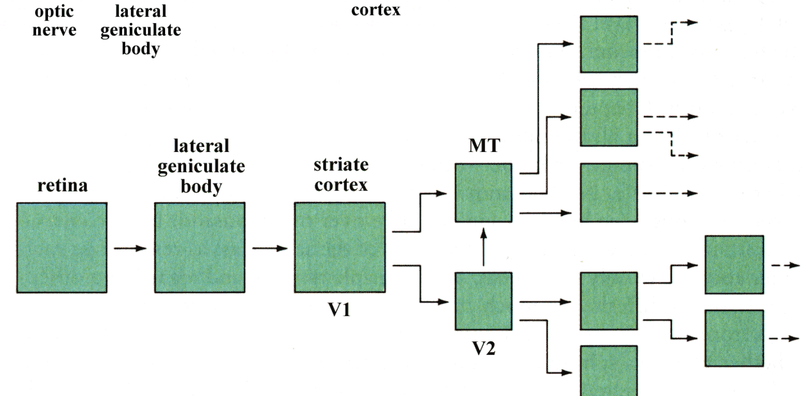


Vernon B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120(4):701–722, 1997.

Terminology: Receptive Fields

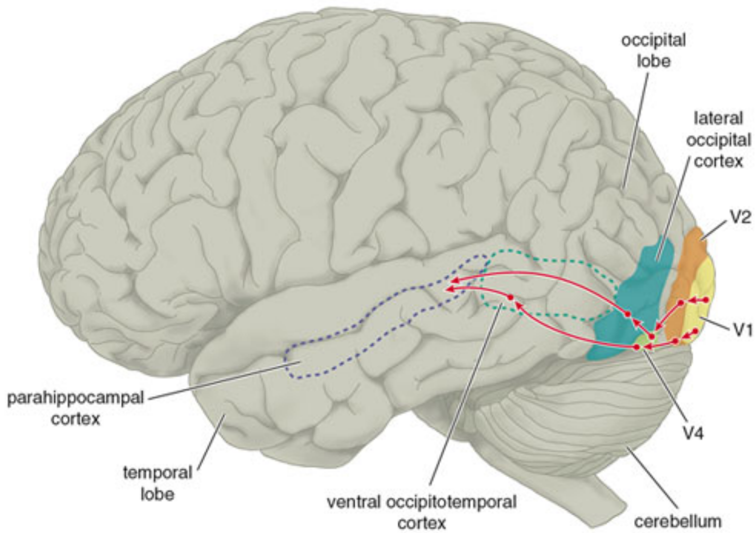




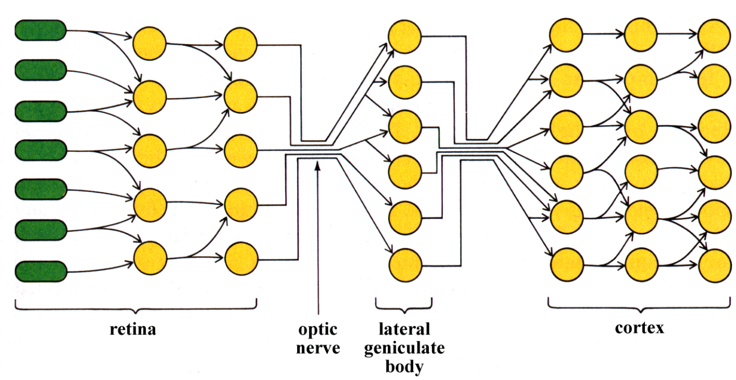


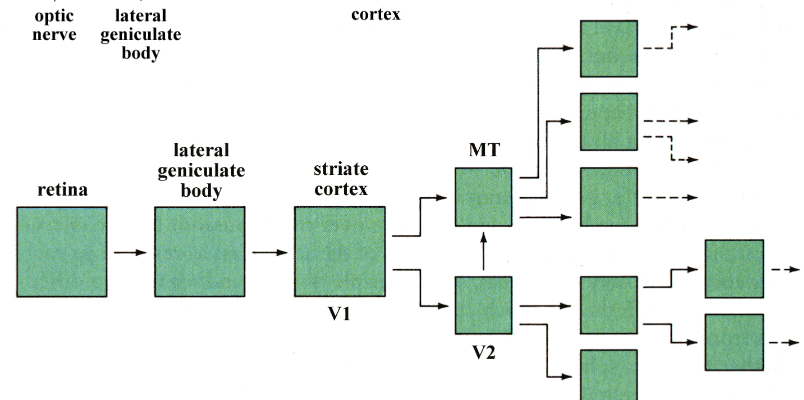
D. H. Hubel and T. N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106–154, 1962.

Terminology: Retinotopic Maps



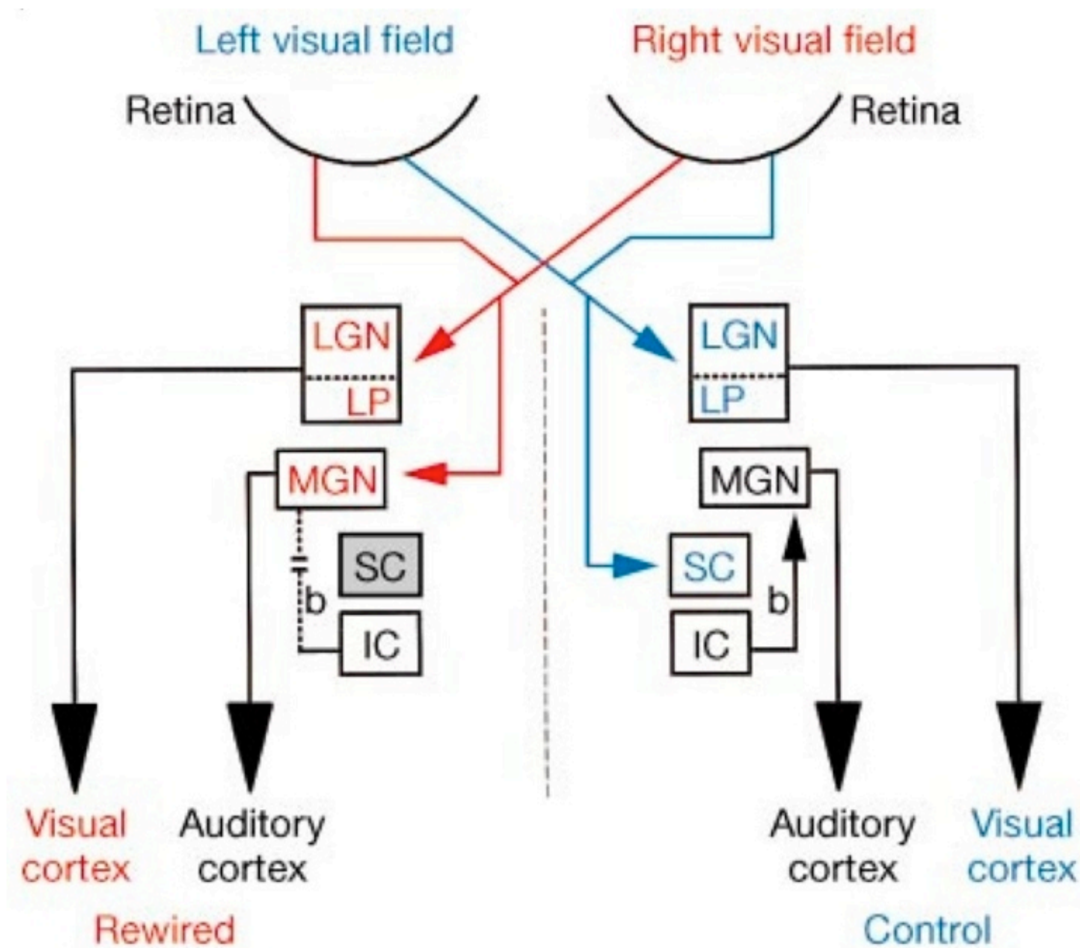
ventral visual pathway





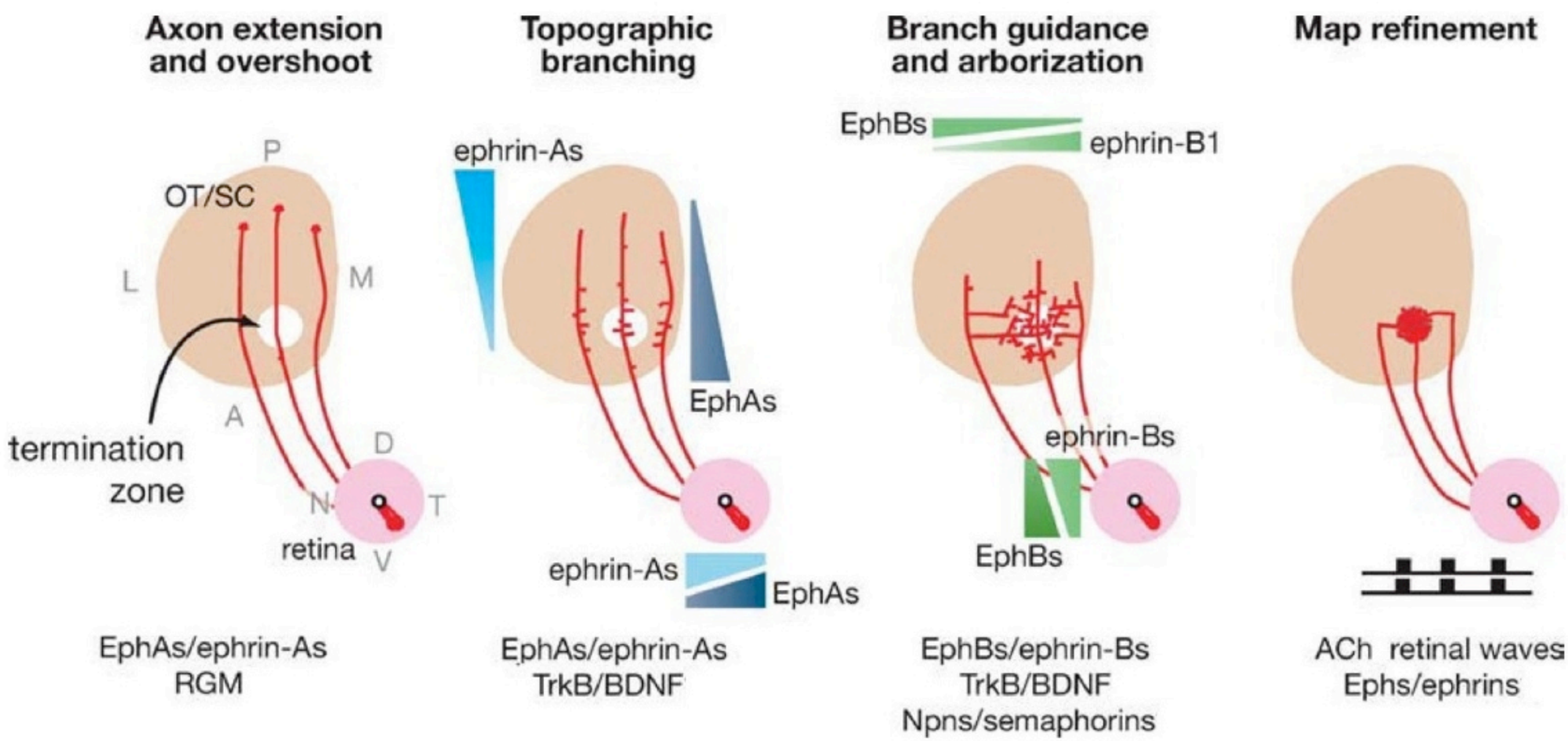
D. H. Hubel and T. N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106–154, 1962.

Rewiring Ferret Sensory Cortex



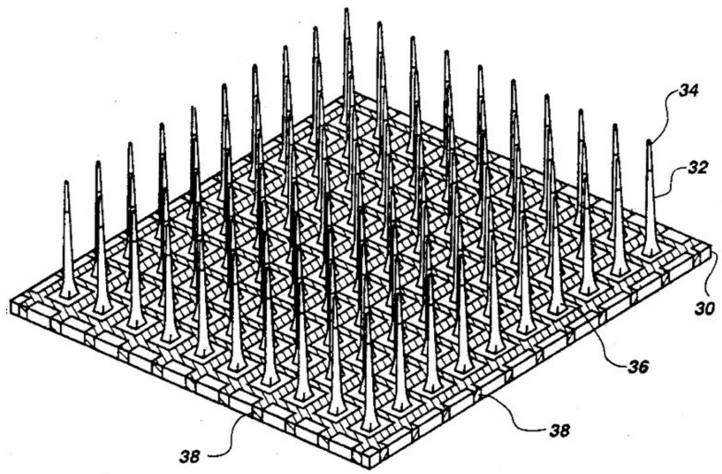
Laurie von Melchner, Sarah L. Pallas, and Mriganka Sur. Visual behavior mediated by retinal projections directed to the auditory pathway. *Nature*, 404(6780):871–876, 2000.

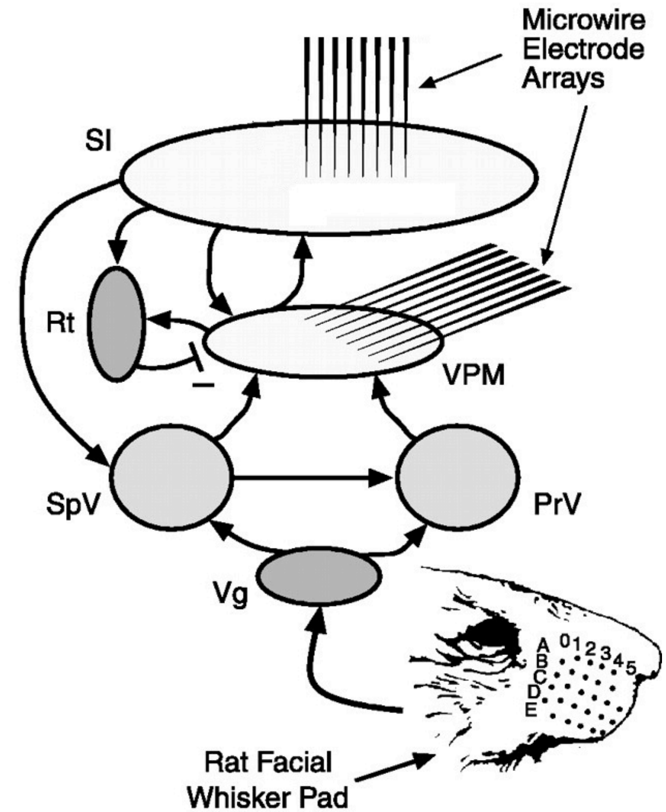
Retinotopic Organization Algorithm



Todd Mclaughlin and Dennis D. M. O'Leary. Molecular gradients and development of retinotopic maps. *Annual review of neuroscience*, 28(1):327–355, 2005.

Dynamic Somatosensory Cortex





[1] Rui M. Costa, Dana Cohen, , and Miguel A.L. Nicolelis. Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Current Biology, 14(13):1124–1134, 2004.

[2] David J. Krupa, Asif A. Ghazanfar, and Miguel A. L. Nicolelis. Immediate thalamic sensory plasticity depends on corticothalamic feedback. Proceedings of the National Academy of Sciences, 96(14):8200–8205, 1999.

[3] M.A. Nicolelis and E.E. Fanselow. Dynamic shifting in thalamocortical processing during different behavioural states. Philosophical Transactions of the Royal Society London B Biological Science, 357(1428):1753–1758, 2002.

Single Algorithm Hypothesis

- Local to Distant Function — inference matures integrating functional areas [1]
- Adolescent Development — lengthy aborization and myelination process [2]
- Extracortical Involvement — thalamic and cerebellar cortex involvement [3,4]

[1] J. Sepulcre, H. Liu, T. Talukdar, I. Martincorena, and B.T.T. Yeo. The organization of local and distant functional connectivity in the human brain. *PLoS Computational Biology*, 6(6), 2010.

[2] J. N. Giedd, J. Blumenthal, N. O. Jeffries, F. X. Castellanos, H. Liu, A. Zijdenbos, T. Paus, A. C. Evans, and J. L. Rapoport. Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, 2(10):861–863, 1999.

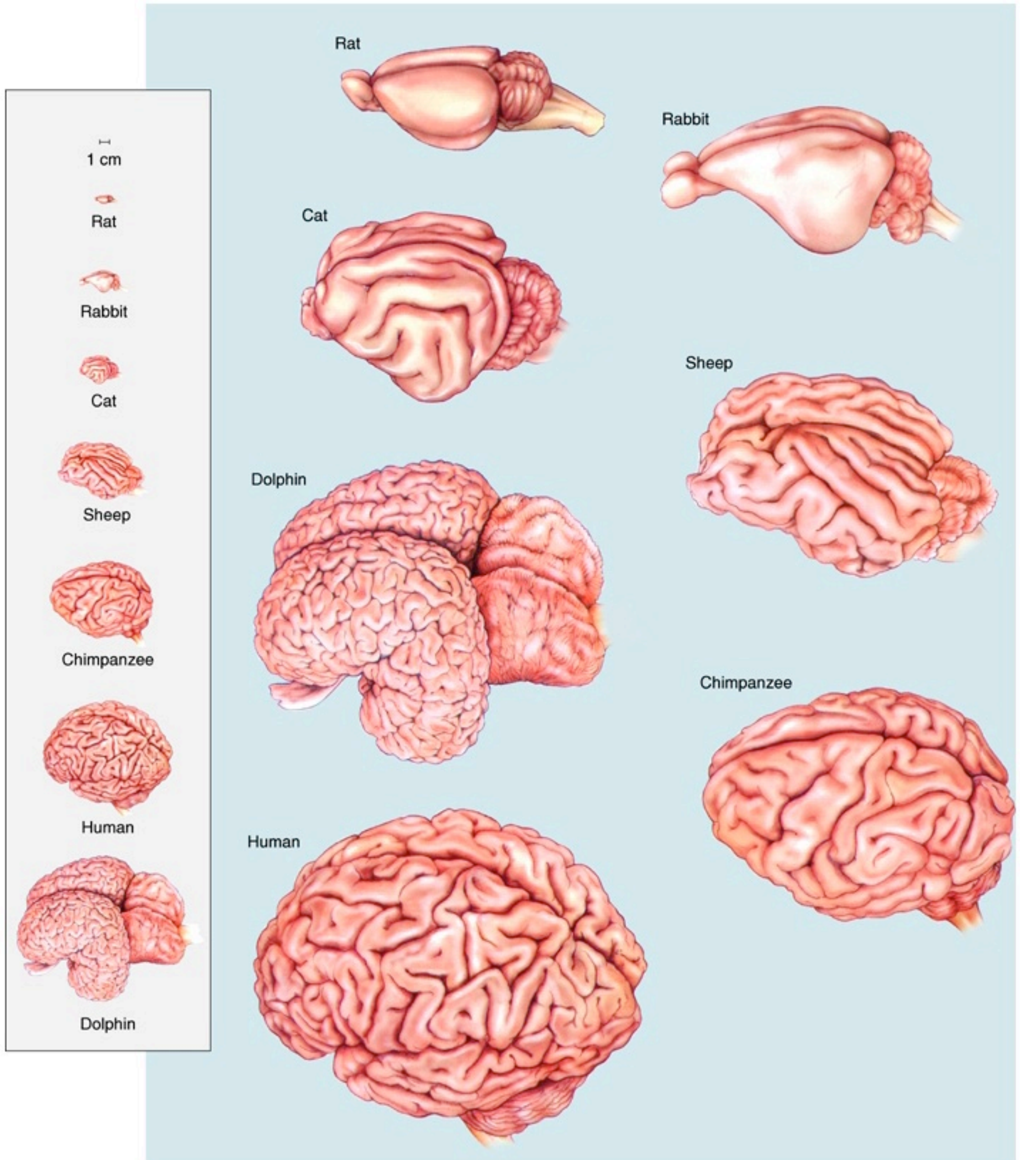
[3] P. Lieberman. On the nature and evolution of the neural bases of human language. *American Journal of Physical Anthropology*, 119(35):36–62, December 2002.

[4] H. Tiemeier, R. K. Lenroot, D. K. Greenstein, L. Tran, R. Pierson, and J. N. Giedd. Cerebellum development during childhood and adolescence: A longitudinal morphometric MRI study. *Neuroimage,* 49(1):63–70, 2010.

Quantity Suffices Hypothesis

- Why the term “quantity” is misleading?
- How does “neurogenesis” play a role?
- How do we build a brain that “scales”?
- What is the analog of more “cores”?
- Why is longer “development” critical?
- What can you do with a faster “bus”?

Mammalian Brain Size



Name	Neurons in the cerebral cortex	Image
Mouse	4,000,000	
Rat	21,000,000	
Dog	160,000,000	
Cat	300,000,000	
Chimpanzee	6,200,000,000	
Human	11,000,000,000	

According to Gordon Shepherd in *The Synaptic Organization of the Brain* [Oxford University Press, 1998, Page 6] we have a total of 10 billion neurons in the cerebral cortex. However, Christof Koch lists the total number of neurons in cerebral cortex at 20 billion in *Biophysics of Computation: Information Processing in Single Neurons*, [Oxford University Press, 1999, Page 87].

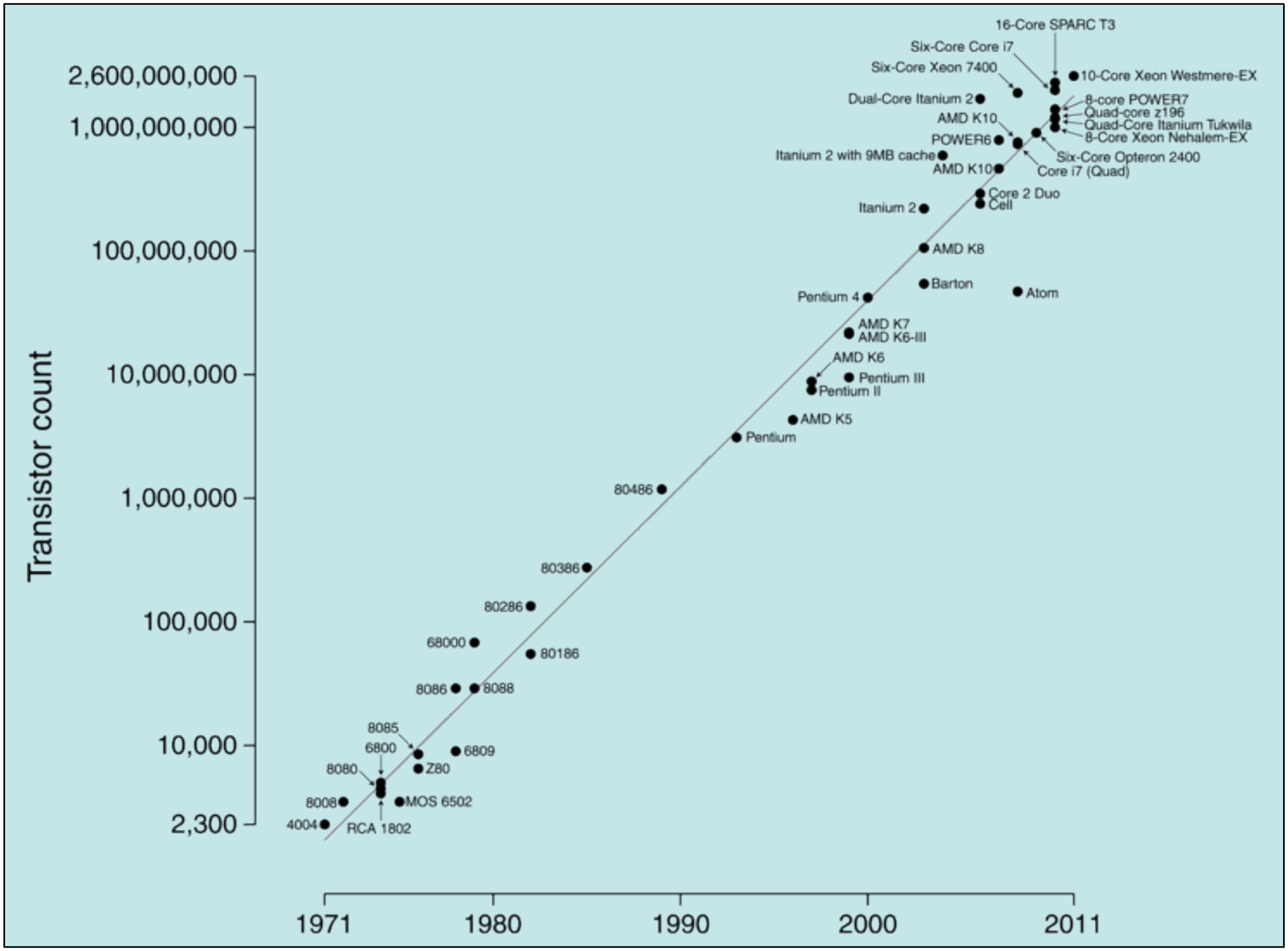
Human Cortex Differences

- Same mammalian types of neurons [1]
- Same basic primate functional areas [2]
- Some areas have more neurons [2]
- Some layers are more densely packed
- Same number connections per neuron
- More intra-regional connections
- Faster inter-regional connections
- Corpus callosum is somewhat thinner

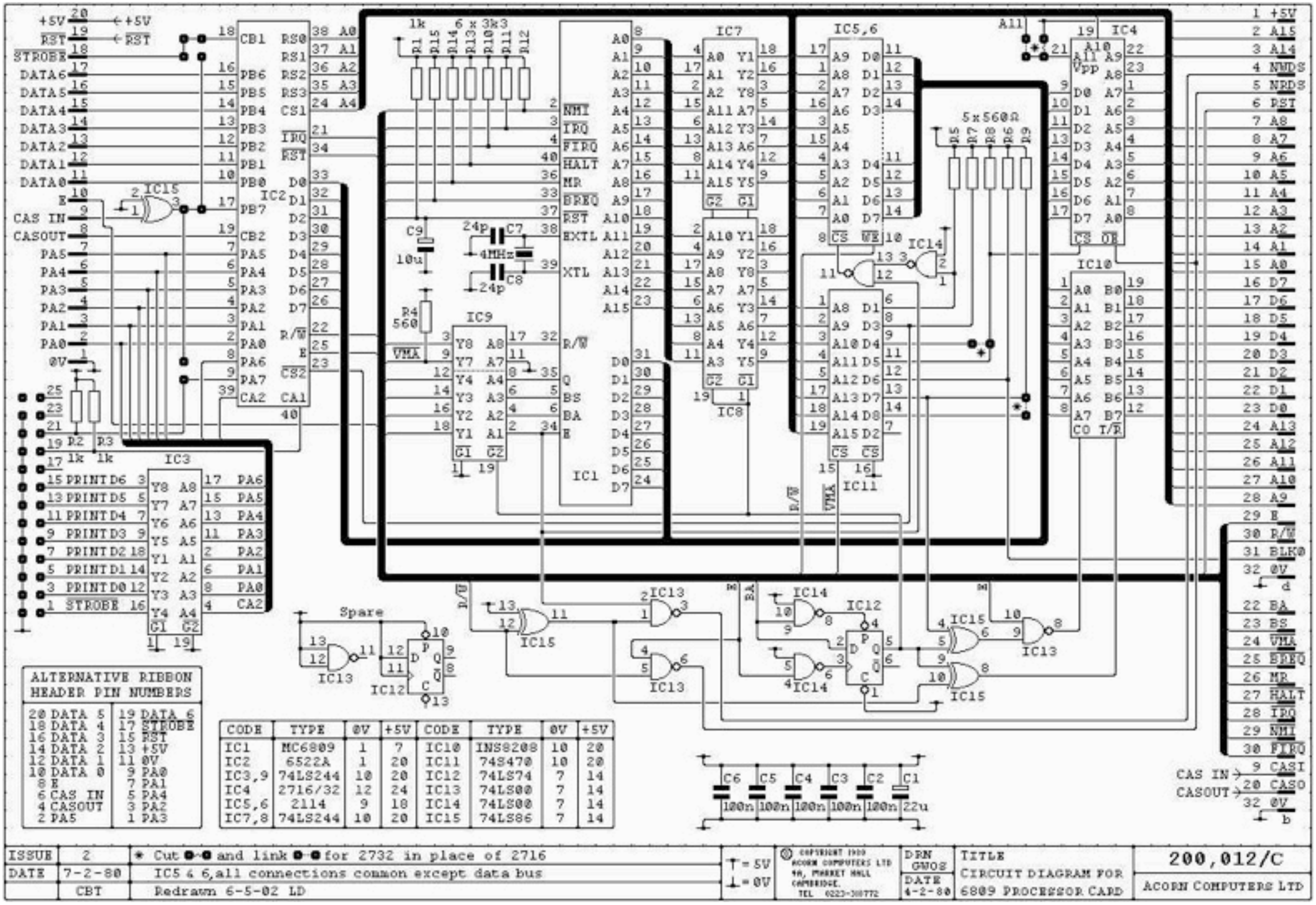
[1] Nimchinsky, E. A., Gilissen, E., Allman, J.M., Perl D.P., Erwin, J.M., and Hof, P.R. A neuronal morphologic type unique to humans and great apes. *Proceedings of the National Academy of Sciences, 96(9):5268–5273, 1999.*

[2] Semendefer, K., Armstrong, E., Schleicher, A., Zilles, K., and Van Hoesen, G.W. Prefrontal cortex in humans and apes: A comparative study of area 10. *American Journal of Physical Anthropology, 114(3):224–241, 2001.*

Processor Chip Transistor Count



Central Processing Unit



The Motorola 6809 is an 8-bit microprocessor CPU — with some 16-bit features — that was introduced in 1978.

Hardware, Software, Wetware

- Reusing versus refactoring software.
- Wetware cannot be made reentrant.
- Adding registers, cores, SIMD lanes.
- Recursion using structural replication.
- Deeper stack for sequential decisions.
- Faster, better-shielded networking.
- External memory, open-source library.

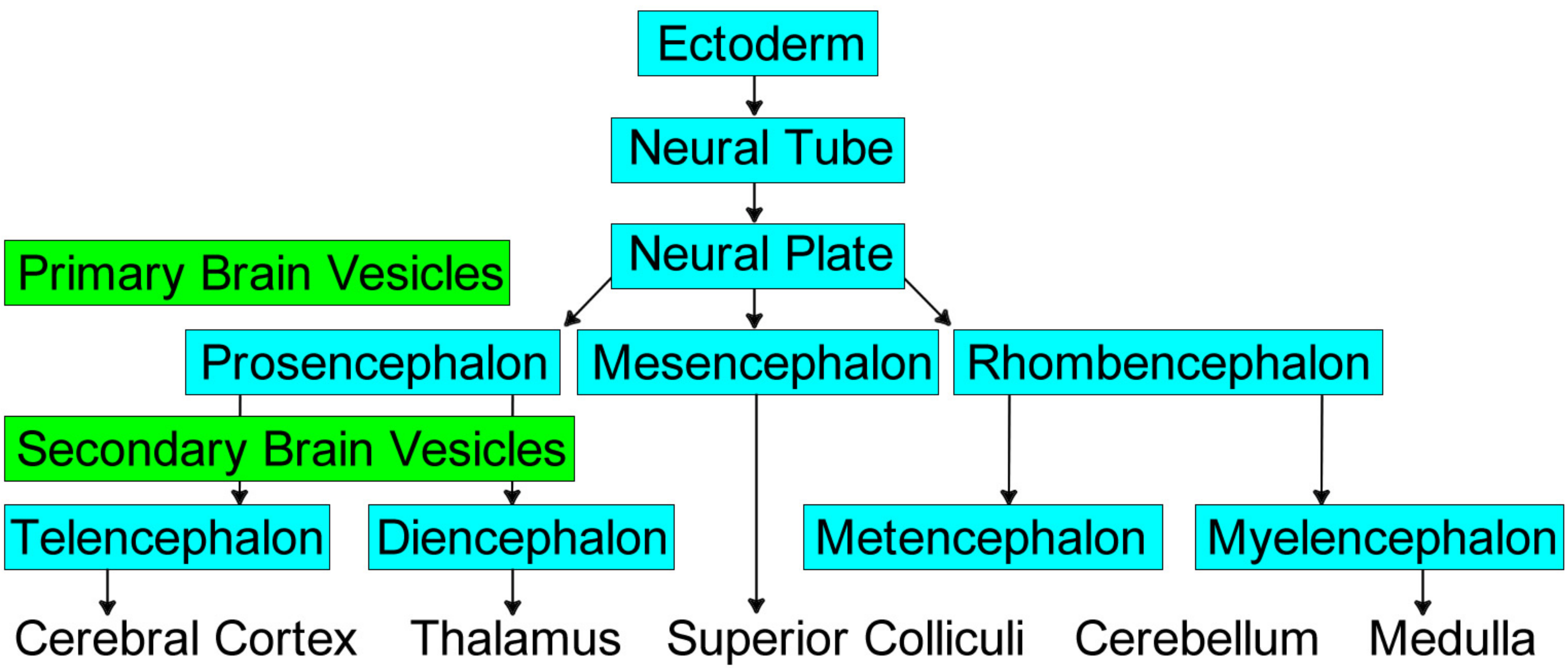
Code refactoring is a “disciplined technique for restructuring an existing body of code, altering its internal structure without changing its external behavior” [1].

A subroutine is *reentrant* if it can be interrupted in the middle of its execution and then safely called again — re-entered.

[1] Fowler, Martin (1999). *Refactoring. Improving the Design of Existing Code*. Addison-Wesley.

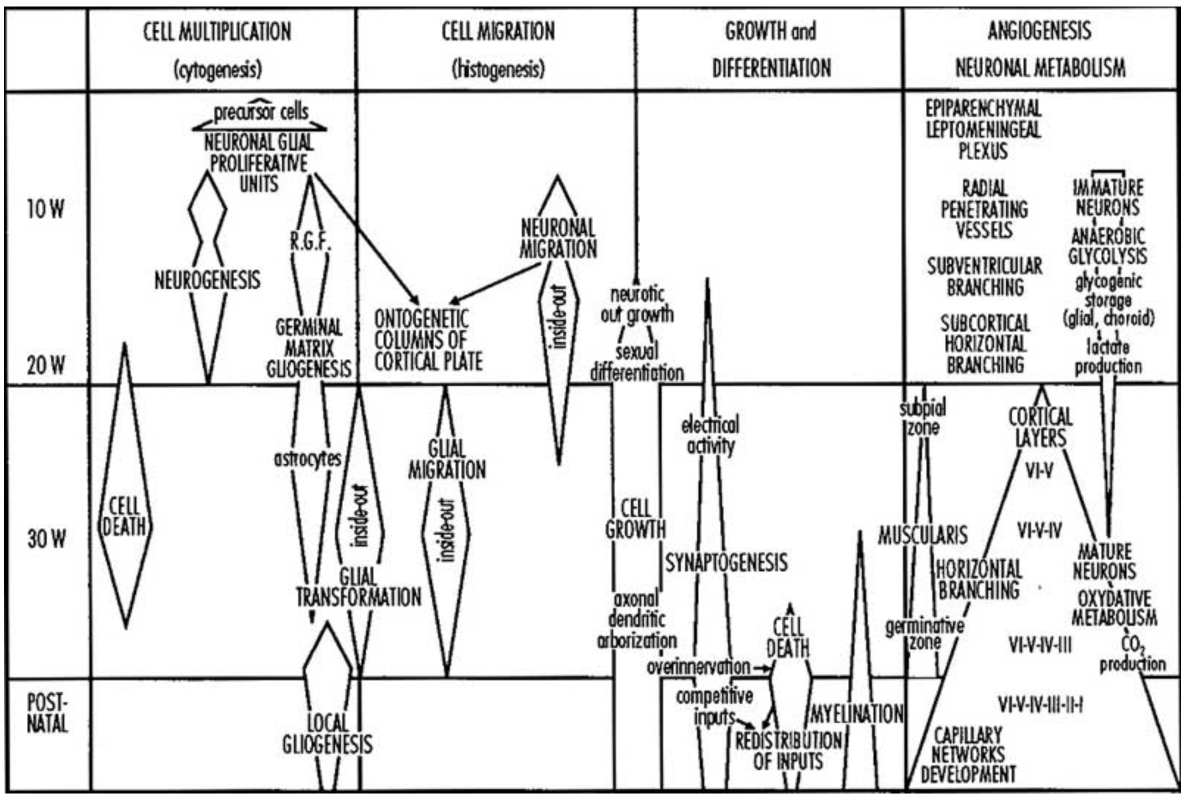
Primate Cortical Development

- Gastrulation forms ectoderm, mesoderm, endoderm
- Neurulation forms neural tube from embryo ectoderm
- Neuronal migrations from origin to their final position
- Cell differentiation into intermediate and mature form

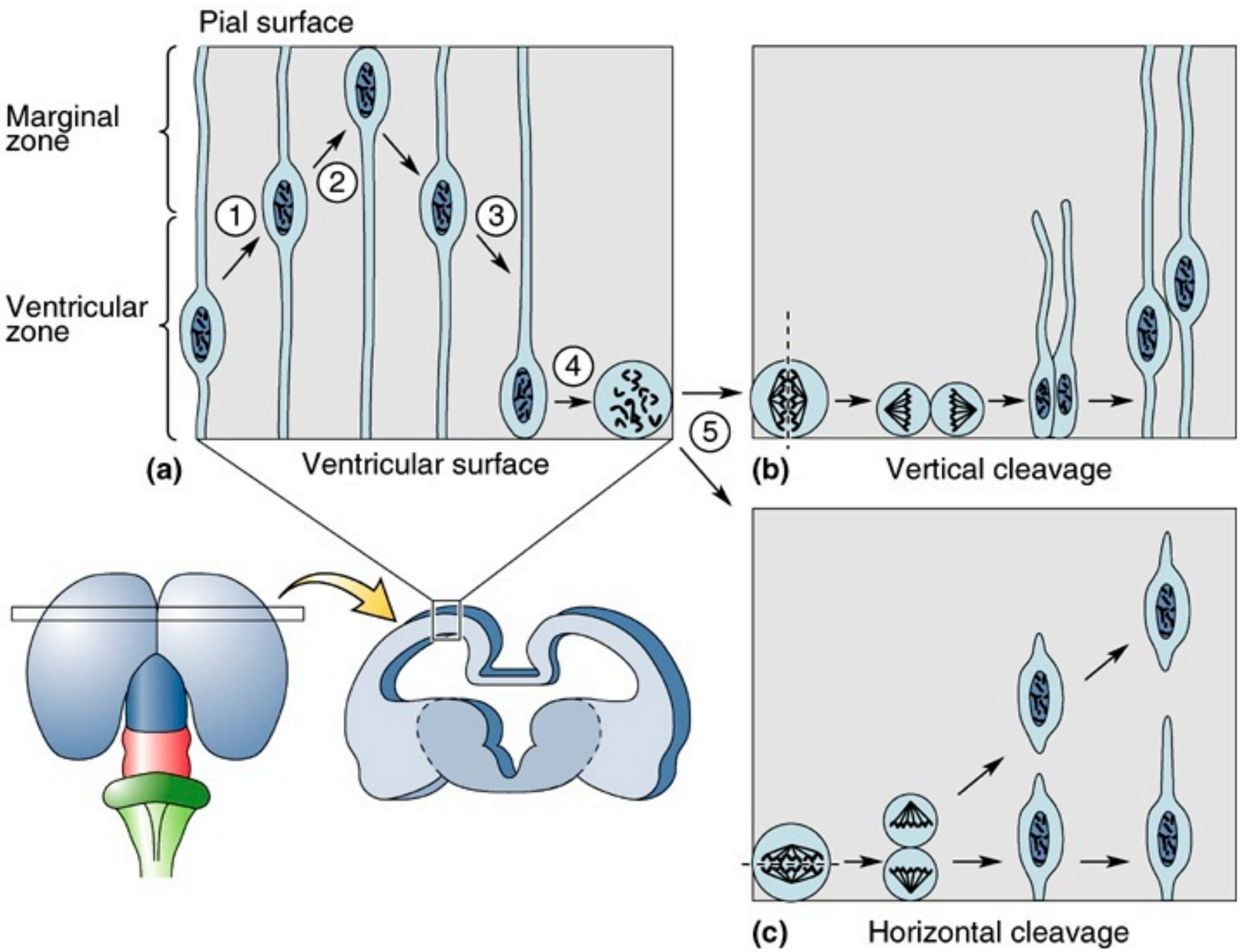


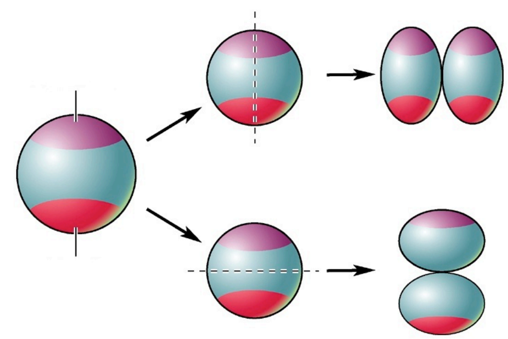
Primate Cortical Development

- Gastrulation forms ectoderm, mesoderm, endoderm
- Neurulation forms neural tube from embryo ectoderm
- Neuronal migrations from origin to their final position
- Cell differentiation into intermediate and mature form



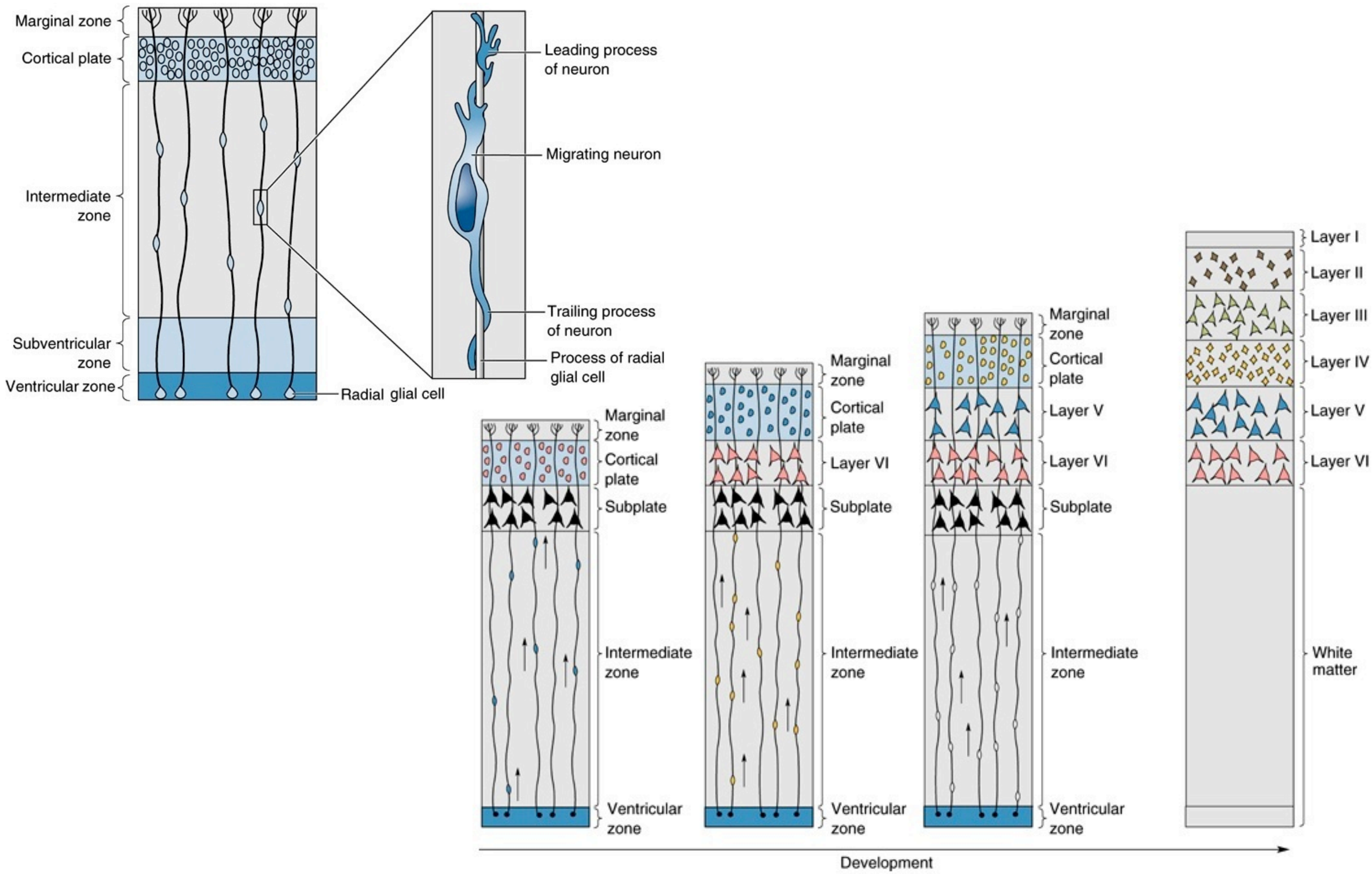
Genomic Regulatory Programs



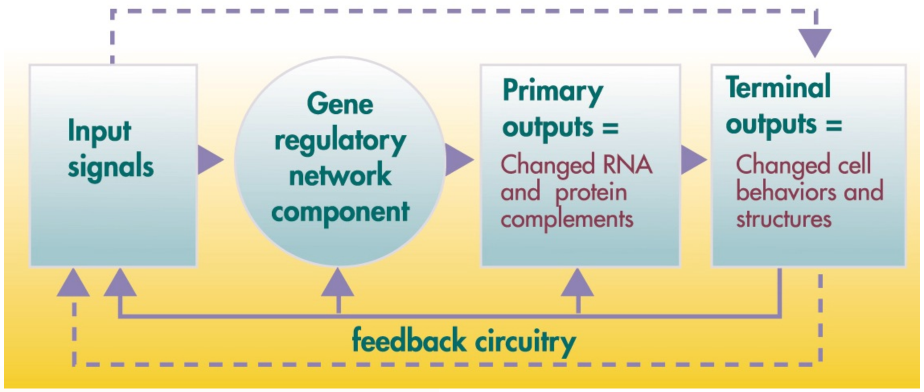


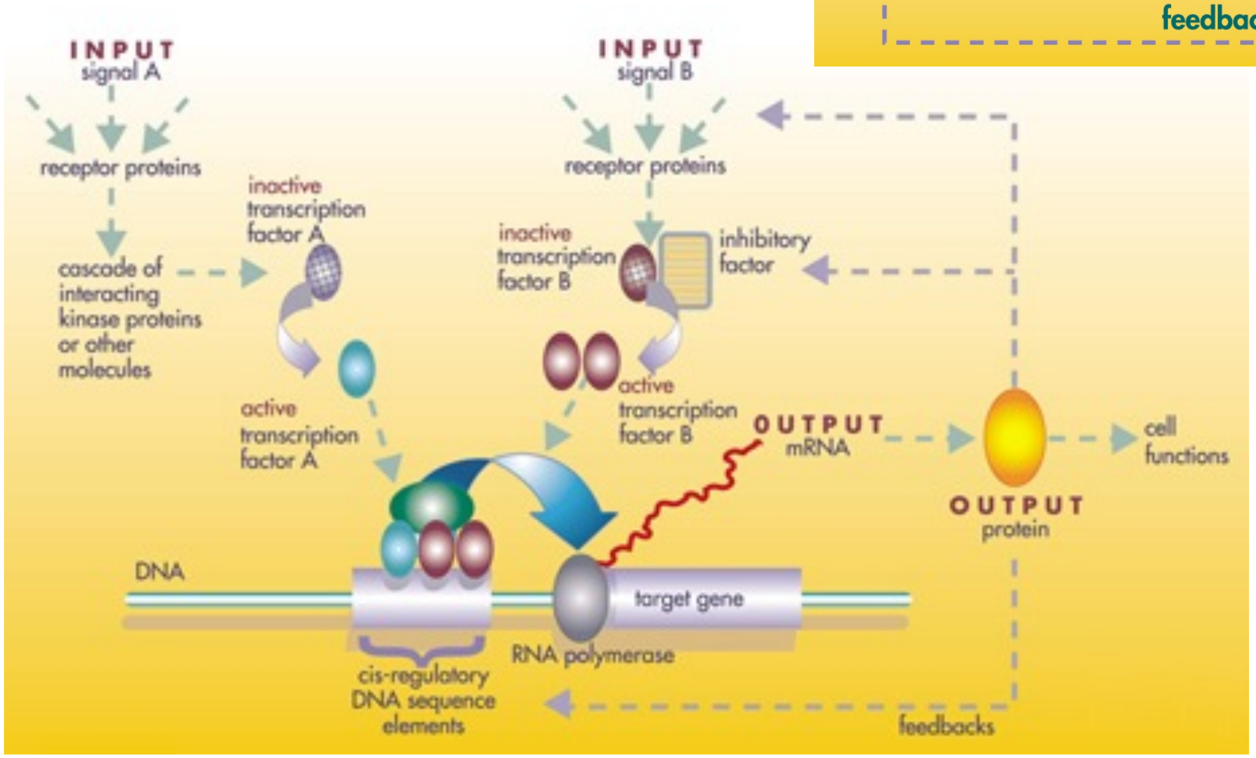
Mark F. Bear, Barry Connors, and Michael Paradiso. *Neuroscience: Exploring the Brain.* Lippincott Williams & Wilkins. 2006.

Genomic Regulatory Programs

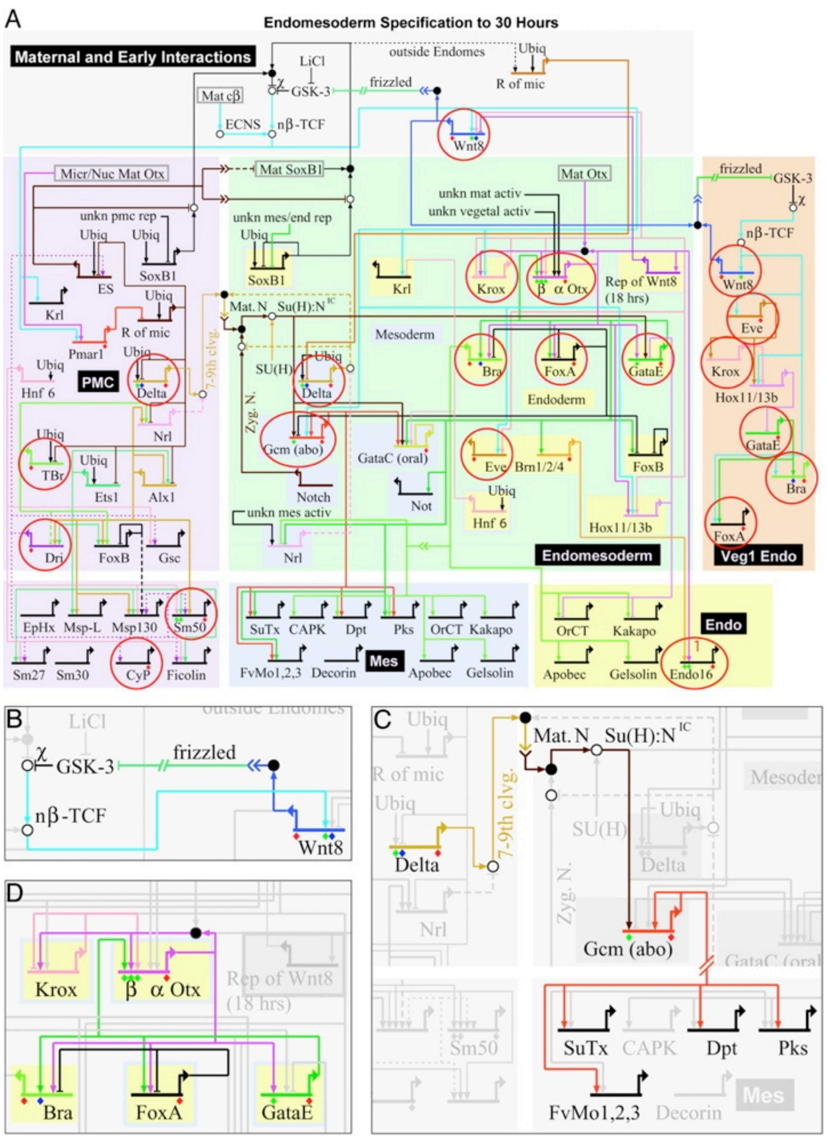


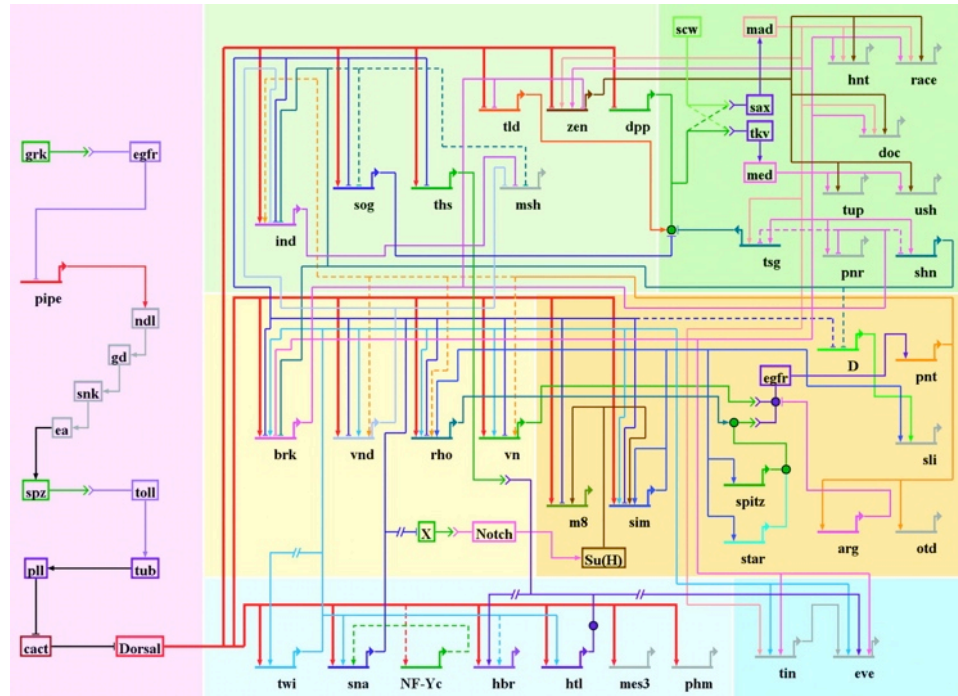
Genomic Regulatory Programs



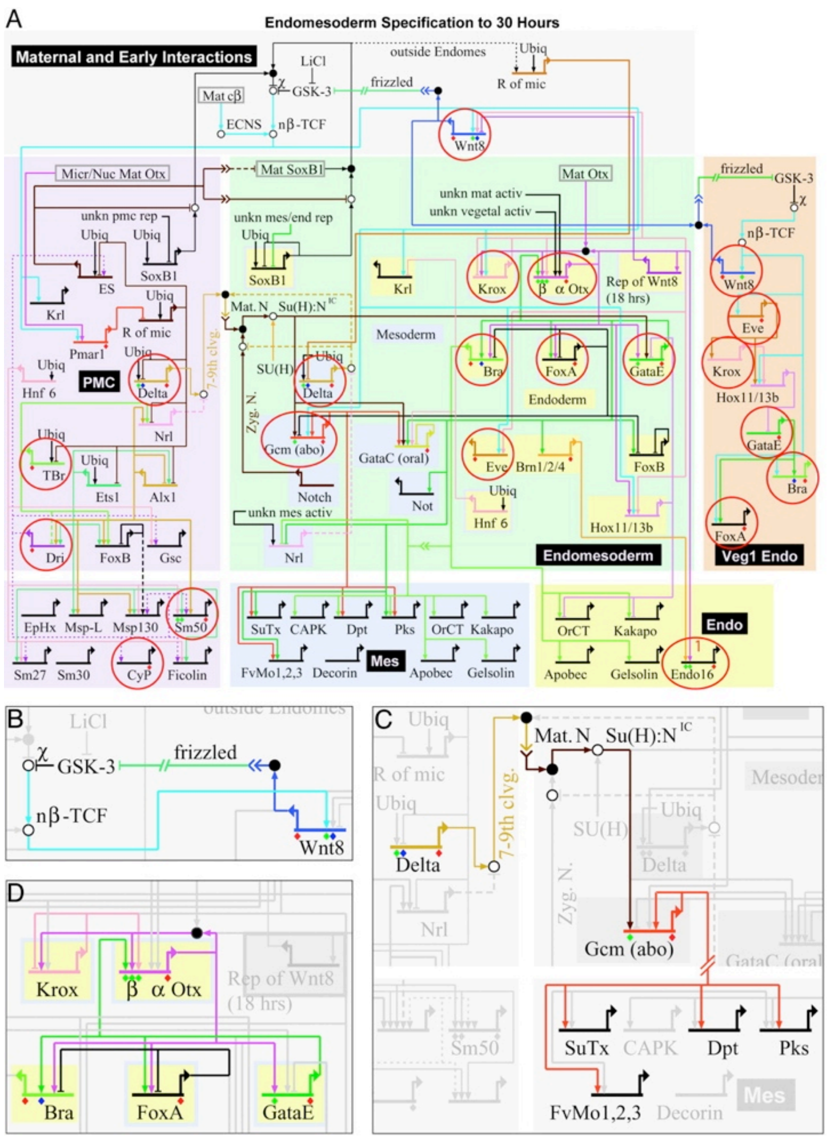


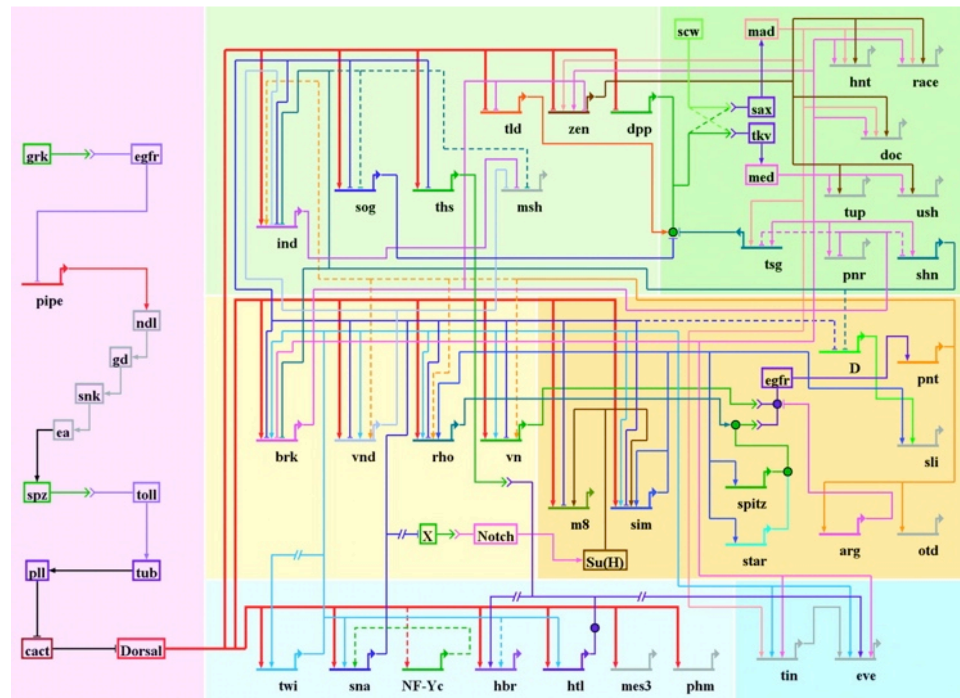
Genomic Regulatory Programs





Genomic Regulatory Programs





... and then there are all the post-transcriptional factors that control gene expression of RNA prior to translation

Modular Minds Hypothesis

- The idea of a circumscribed cognitive faculty is at odds with what we know about connectivity patterns of functional areas in primate cortex.
- In any case, this notion of modularity has little to do with related principles of design[1] which if biology[2] adhered to them would simplify efforts to reverse engineer the primate brain.

[1] In software engineering, modularity is a logical partitioning of the design that allows complex software to be manageable for the purpose of implementation and maintenance. — source: http://en.wikipedia.org/wiki/Modularity

[2] In terms of morphology — as opposed to cognitive function, natural selection seems adept at designing robust modular systems. Hox genes governing the basic body plans in most animals illustrate this modular design. The PAX-6 gene has the capability that if expressed in a fruit fly it builds a fruit-fly eye and if expressed in a mouse it builds a mouse eye [1].

[1] Callaerts P, Halder G, Gehring W. PAX-6 in development and evolution.. *Annual Review of Neuroscience.* 20 (1): 483–532.

Single Algorithm Hypothesis

- The experiments re-wiring ferret brains and ubiquity of topological map formation among primary sensory and motor cortex provide a compelling case for some core algorithmic principles at play, at least in these evolutionarily more recent cortical functional areas.
- Connections from cortical areas involved in decision making, language, speech, movement execution and planning to the cerebellar cortex and thalamic nuclei would require a more complicated set of algorithmic principles than Hinton, DiCarlo, Lewicki had in mind.

Quantity Suffices Hypothesis

- The basic cytoarchitecture of the primate neocortex — cortical columns — and developmental machinery for scaling the neural substrate along all of the key dimensions of computational power suffice to build a significantly more powerful computational engine.
- This scalable computational architecture[1] and the bootstrapping via cultural affordances from language to facilitate thinking and communication and external memory to simulate a universal Turing machine seem adequate to explain differences between humans and their closest primate cousins.

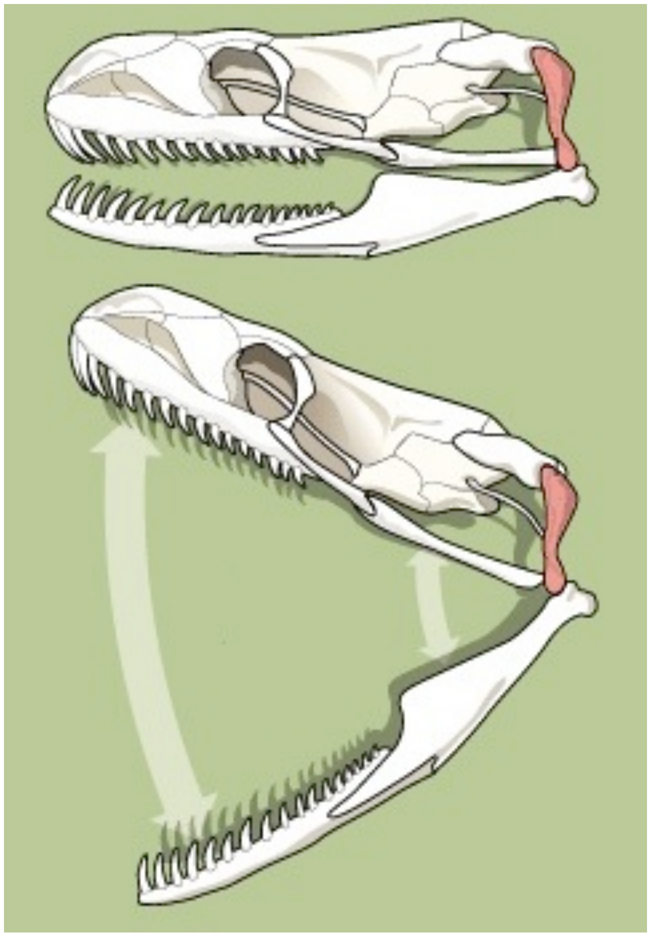
[1] With the most important consequence being an increase in the depth of combinatorial circuits that can be constructed from the neural substrate. See Vitaly Feldman and Leslie Valiant. Experience-induced neural circuits that achieve high capacity. *Neural Computation,* 21(12):2715–2754, 2009.

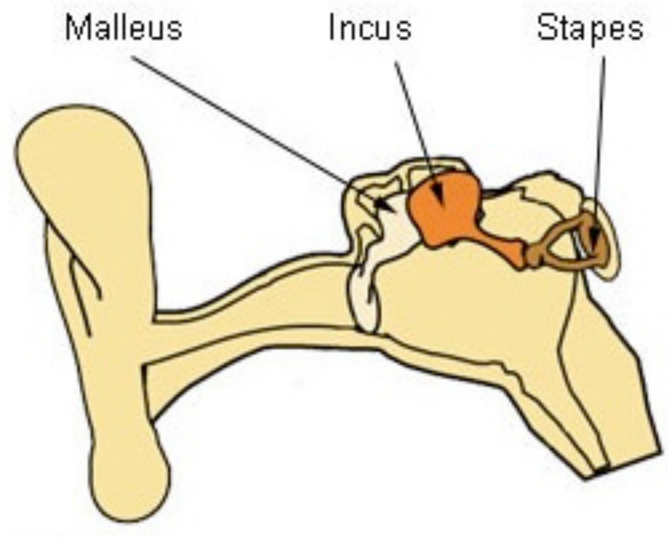
Deeper Combinatorial Circuits

- Recursive embedding to produce more complicated linguistic structures;
- Deeper, recursive epistemological and emotional theories of other minds;
- Managing with larger cliques and more complicated social arrangements;
- More sophisticated physical theories, mechanisms, and causal models;
- Deeper layers of abstraction, richer compositional models, motor neurons;

Supplementary Materials

Thrifty Natural Selection

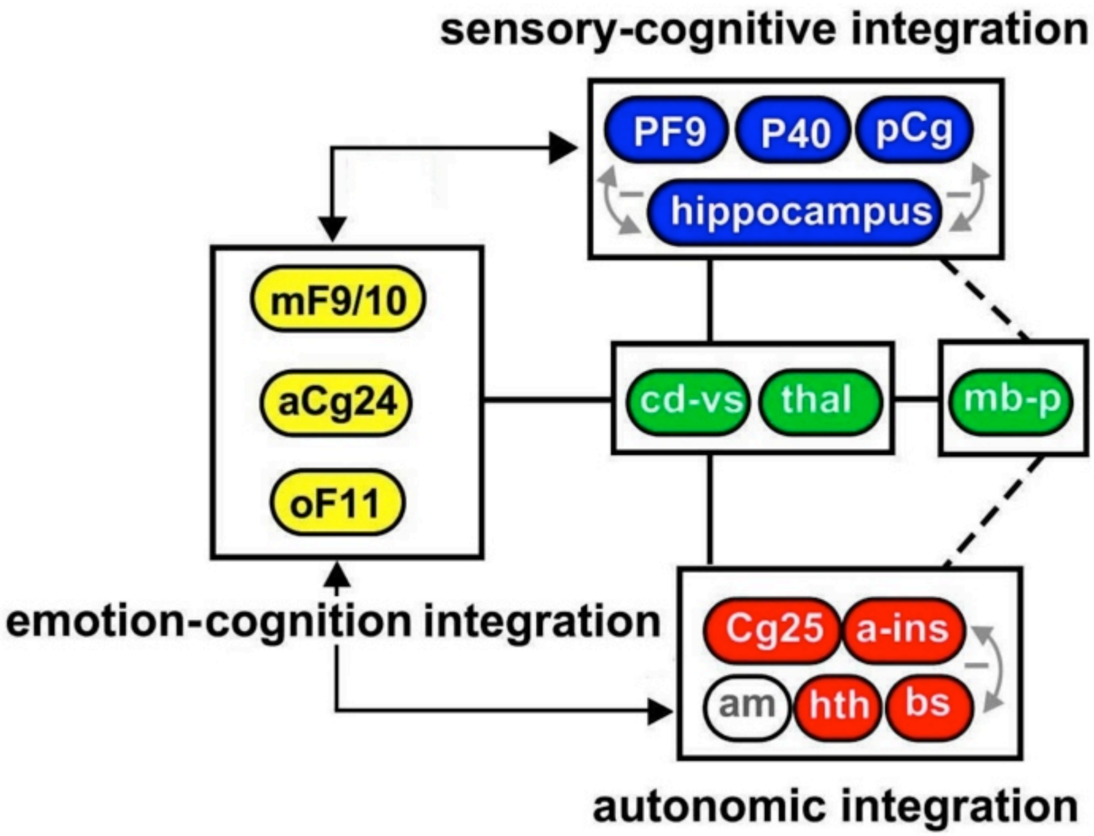


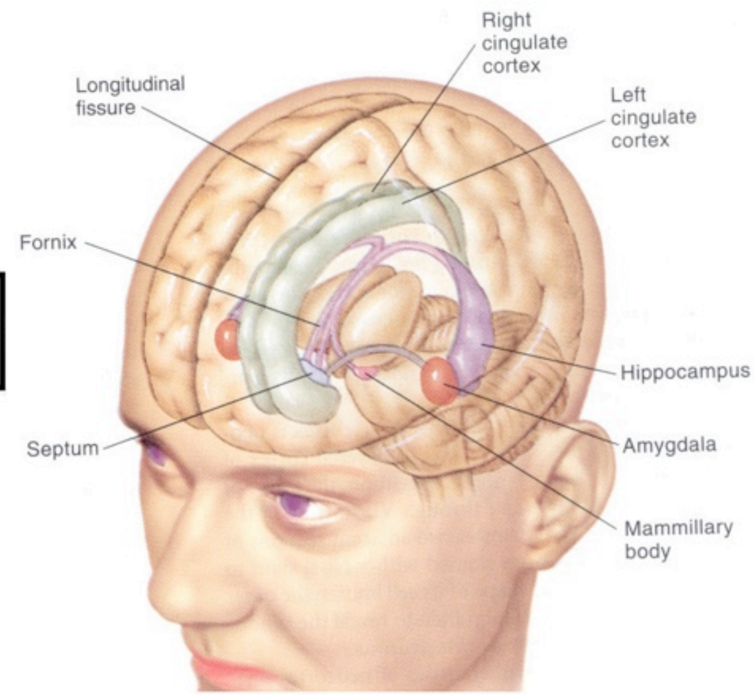


This is an example of *exaptation* in which a trait that has evolved to serve one function subsequently adapts to serve another.

Gould, Stephen Jay, and Elizabeth S. Vrba. Exaptation — a missing term in the science of form, *Paleobiology,* 8(1): 4-15 1982.

Mysterious Area 25



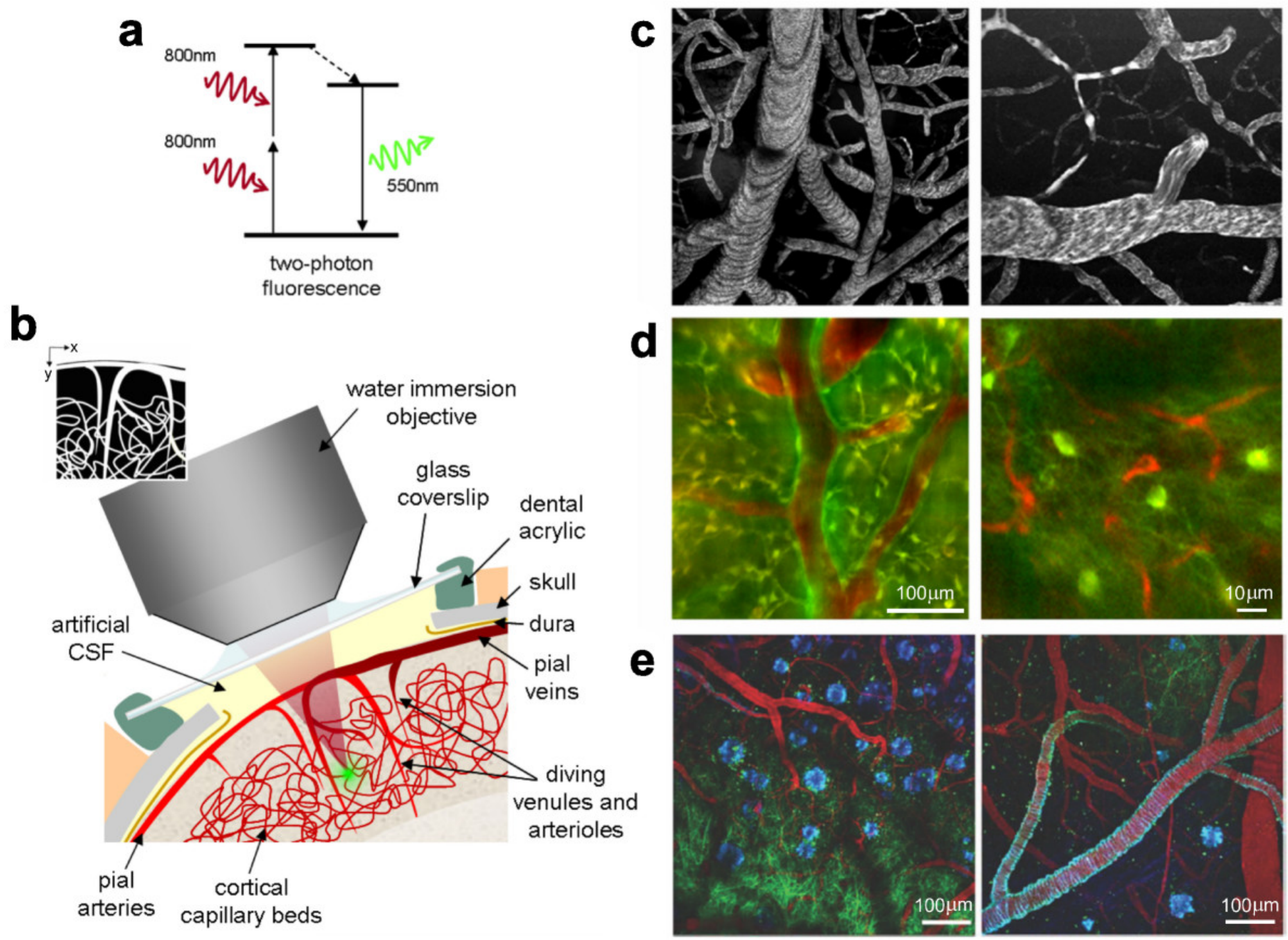


H. S. Mayberg, M. Liotti, S. K. Brannan, S. McGinnis, R. K. Mahurin, P. A. Jerabek, J. A. Silva, J. L. Tekell, C. C. Martin, and J. L. Lancaster. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. The American Journal of Psychiatry, 156(5):675–82, 1999.

D.A. Seminowicz, H.S. Mayberg, A.R. McIntosh, K. Goldapple, S. Kennedy, Z. Segal, and S. Rafi-Tari. Limbic-frontal circuitry in major depression: a path modeling metanalysis.NeuroImage, 22(1):409–418, 2004.

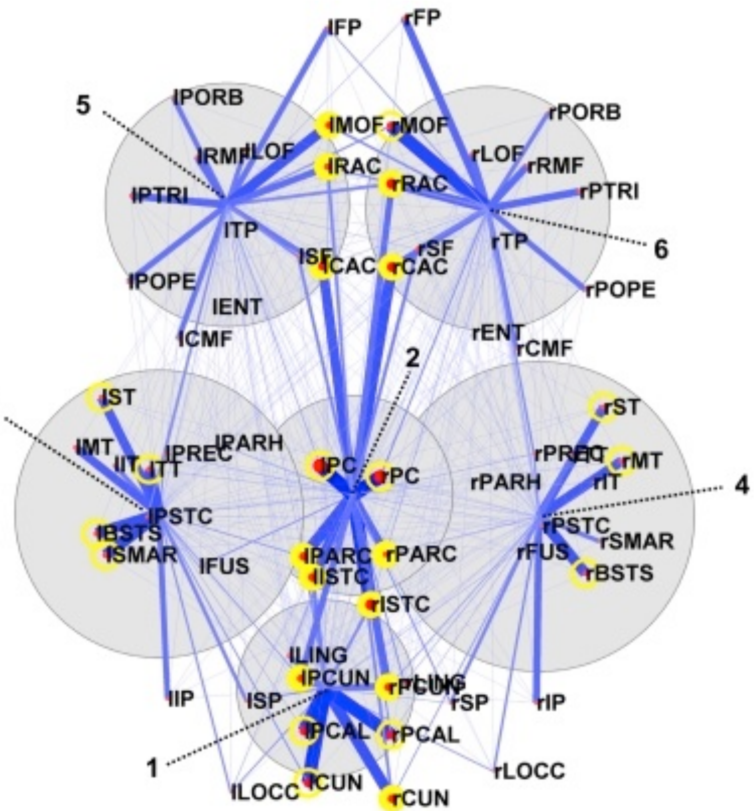
Brodmann’s Area 25 is located in subgenual cingulate cortex (Cg25) and has no spatial relationship to Roswell, New Mexico.

In Vivo Two-Photon Microscopy



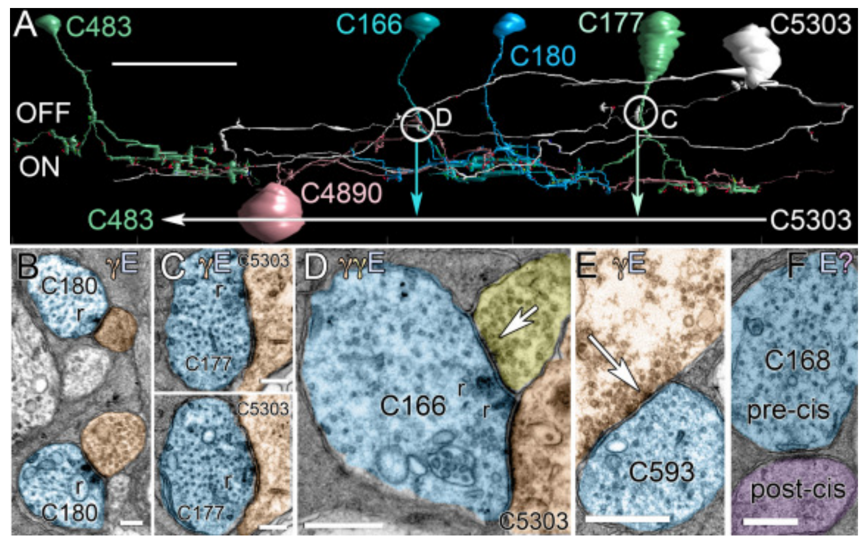
Human Connectome Project

Macroscale: Tensor Diffusion MRI



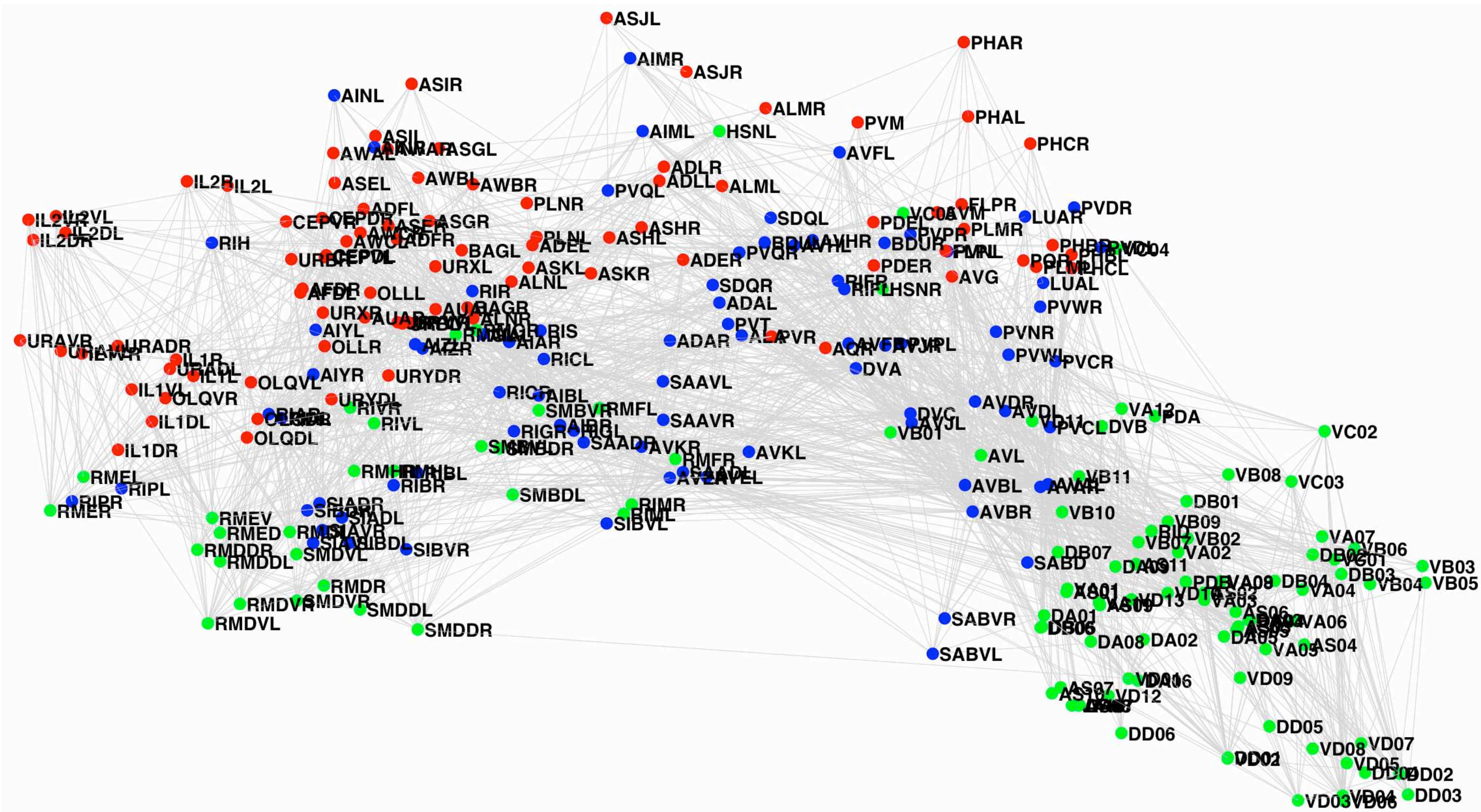
Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J. Honey, Van J. Wedeen, and Olaf Sporns. *Mapping the structural core of human cerebral cortex*. PLoS Biology, 6(7):159, 2008.

Microscale: Electron Microscopy



Anderson JR, Jones BW, Watt CB, Shaw MV, Yang JH, Demill D, Lauritzen JS, Lin Y, Rapp KD, Mastronarde D, Koshevoy P, Grimm B, Tasdizen T, Whitaker R, Marc RE. *Exploring the retinal connectome*. Molecular Vision, 17:355-365, 2011.

C. Elegans Connectome



Transcription & Translation