Date: **31 Oct 2011**
Version: **8**

The Brain Activity Map: **Technological Foundations**

Overview

This document is intended to provide an overview of the technical underpinnings that must be assembled to map the "functional connectome" of the brain. We start by listing the overarching technical goals of this national program, delineating the underlying assumptions made in defining these goals. We then formulate a list of requisite technological thrusts, with specific milestones, that we see as necessary to achieve these program goals. We conclude by drawing parallels of the project's scope with that of the Human Genome Project.

A) Technical Goals

We identify four principal goals:

- Ultra-large-scale, extremely high-density **measurement** (recording) of neural activity across widely distributed brain areas.
- Ultra-large-scale, distributed **stimulation** of neural circuits with arbitrarily complex spatiotemporal patterns and extremely high spatial resolution;
- Realization of large-scale **system integration and production *en masse*** to enable creation of powerful and robust "tools" permitting this work to be undertaken
- **Validation** of the technology, its wide-scale **deployment** to the academic and corporate neuroscience community, and **user training** will be required.

These are each outlined below.

Measurement (Imaging/Recording):

To elucidate the functional connectome, the type of measurements we envisage will involve the following:

a) Measurement with single-neuron resolution of both the **extracellular potential** (spiking) and **intracellular potential** (that is, spiking and, ultimately, subthreshold behavior) are of interest -- at the characteristic magnitudes and time scales on which they occur. (For both intracellular and extracellular potentials, sensitivity at least at the ~20uV level with 10KHz bandwidth is necessary.)

b) **Multiparametric physical/chemical measurements**, that is, recording the temporal dynamics of neurophysiological parameters beyond potentiometric measurement alone, are of interest. Such measurements may include: local chemical concentrations (*i.e.* the chemical analog of a local field potential), dynamical local forces involved in neural association, *etc*.

c) **Local interactions** with a "probe" or "reporter" located <100um from neuronal bodies will be required for measurements with single-neuron resolution and <1nm from the plasma membrane for optical voltage measurements. The methods to be developed will always be carried out through some form of remote readout (see below), however we anticipate that a **local reporter** will be often involved. This will be placed no further than the length scale (exponential

decrement) characterizing the spatial decay of the “fields” (electrical, chemical, force) of interest. Such probes/reporters may be arrays of narrow neural probes (shanks) inserted deep into the brain tissue, functional nanoparticles placed in the membrane either by direct intervention or developmental transport, or, ultimately, complex functional nanosystems capable of both local recording and local storing of the time records of the “fields” they measure.

d) **Deep interactions** (measurements) with neural tissue are required. As an example, for the rat, requisite interaction depths will typically be 0-10 mm. For primates, even deeper interaction lengths will be required.

e) **Non-local readout** of the neuronal fields[1] is essential to transport the acquired data to external measurement and analysis systems. In first phase efforts this will be achieved through direct connection involving electrical leads, optical fibers/waveguides, etc. Later generations may involve non-local interrogation fields, such as radio-frequency waves, of types that are capable of deep penetration of neural tissue. For the case of autonomous (untethered) probes, capable of local sensing and local data storage, an entirely different class of readout could involve tissue dissection and recovery of the individual probes -- in which case their internal data would be subsequently downloaded.

f) Highly **biocompatible interactions** with neural tissue are required, to prevent gliosis and attendant loss of sensitivity -- and, thereby, to enable chronic recording.

f) **Dense distributed recording with single-neuron resolution** to decipher a complete picture of the functional connectome.

g) Definitive identification of the **spatial location** within the brain of *each* of the local probes/reporters is essential, in order to correlate the measured functional connectome with the morphological connectome. For non-isotropic neuronal fields, **spatial orientation** of direction-sensitive probes/reporters may also be required.

Stimulation:

To elucidate the functional connectome, we anticipate that direct **stimulation** of specific neurons (followed by *measurement* of the response to such stimuli) will be essential. We envisage such stimulation will involve the following:

a) **Electrical stimulation** at the single-neuron and, possibly, at the single-synapse level. This could be carried out by direct electrical stimulation from the extracellular environment via electrophysiological probes, or, indirectly, through optical stimulation of membrane-bound, genetically inserted light-sensitive ion channels, or through photo-uncaging of active neurotransmitters (such as glutamate, GABA, Ach or glycine).

[1] Here we use “neuronal field” to mean the time-varying fields (electrical, chemical, mechanical, …) driven by the behavior of a *single* neuron. This is to avoid confusion with “local field”, used in neurophysiology to convey an average field generated by multiple neurons within some characteristic volume.

b) **Multiparametric physical/chemical neurostimulation**, that is, stimulation by alteration of neurophysiological parameters beyond direct neuronal depolarization. Such stimulating interactions may include: local delivery of neurochemical stimuli through microfluidic devices, mechanostimulation by imposition of dynamical local forces through, e.g. nanoelectromechanical actuators, *etc*.

c) **Stimulation of artificial actuators** through signals applied, presumably, <<100um from neuronal bodies, so as to achieve selective, single-neuron effects. Similar to the case for local recording, the stimulation methods to be developed will always be controlled through some form of remote input (see below), however we anticipate that **local trigger** might be involved. This local entity, although controlled remotely, would deliver the stimulus only to a restricted, specifically-targeted region. Such probes/triggers may be arrays of narrow neural probes (shanks) inserted deep into the brain tissue, or functional nanoparticles placed either by direct intervention or developmental transport.

d) **Deep interactions** (stimulation) with neural tissue are required. The previous examples for recording also hold here -- for the rat, requisite interaction depths are ~10mm. For primates, even deeper interaction lengths will be required (several centimeters).

e) **Non-local control** of the stimulating fields is essential to permit automated, spatially and temporally complex stimulation protocols under computer control. As in the case for recording, in first phase efforts this will be achieved through direct connection involving electrical leads, optical fibers/waveguides, etc. Later generations of technologies may involve non-local interrogation fields, such as radio-frequency waves, that permit deep penetration of neural tissue (unlike optical fields).

f) Highly **biocompatible interactions** between the probes/triggers and neural tissue are required, to prevent gliosis and attendant loss of effectiveness.

f) **Dense, distributed stimulation with single-neuron resolution, using natural sensory stimuli (e.g., visual, auditory, olfactory, tactile, and proprioceptive)** will ultimately be necessary, in order to decipher a complete picture of the functional connectome.

g) Deterministic programming of **spatial location** within the brain for *each* of the local probes/triggers is essential, in order to correlate the functional connectome with the morphological connectome. For non-isotropic stimulation fields, **spatial orientation** of direction-sensitive probes/triggers may also be required.

System Integration and Production

Academic research is, and will continue, driving specific innovations in nanodevices, nanoparticles, and synthetic & molecular biology to create new technology for this program. The complex technology we will assemble for this project requires the coherent concatenation of many such individual innovations. Systems-level engineering and **system integration** and of new powerful and robust “tools” (instruments) is required. Further, these tools need to be widely deployed to address the monumental task of deciphering how brains work; accordingly, **mass**

production of these new tools, with fidelity and uniformity, will be essential. Similarly, there will be a need for developing mathematical and computational strategies capable of amalgamating and analyzing the massively parallel time series data that will be generated from the measurements described above.

The integrative nature of this work and the capabilities required to achieve them places these technical goals well outside the scope of what can be accomplished by single-investigator-driven university research. Their pursuit must involve highly coordinated efforts between state-of-the-art microelectronic foundries, experts in instrument development and assembly, directed characterization and validation (discussed below), and centers devoted to computational innovation. This is unlikely to be jump-started in the corporate sector in the near term. No viable business model can be currently formulated to finance a program for prototype instrument development and manufacturing. Hence, we believe this essential work must be initiated through a sharply focused national initiative that would create a center, or several centers, that will amass the requisite technological competencies.

Validation, Technology Deployment, and User Training

In parallel with technological development, wide-scale **technology deployment** to the academic and corporate neuroscience community will be required, as well as the requisite **user training** in its use by a cadre of technological experts from the principal technology center(s). Early stage adopters of the technology will provide the absolutely essential role of **validation** of the technology by, for example, pilot neurophysiological studies at progressively increasing scales. These beta-testing projects will pave the way to wide-scale technology deployment of robust, well-characterized tools to the academic and corporate neuroscience community. Examples of locations that could serve as teaching sites for these new technologies are the neuroscience techniques summer courses held at the Marine Biological Laboratory in Woods Hole and in Cold Spring Harbor Laboratories.

B) Technical Milestones

To achieve the goals listed above, we anticipate that pursuit of the following technological thrusts will be necessary:

1. Large-Scale Measurement/Recording

- Our first goal is to measure activity from very large numbers of neurons in a model nervous system (for example: *C.elegans*, Drosophila and mouse neocortical brain slices. This will begin with a parallel efforts using optical, electrical and genomic techniques i) optical imaging (e.g., by two-photon methods) of neural circuits loaded with either calcium indicators or voltage-sensitive dyes, and ii) electrophysiological interrogation of neural tissue via neural probe arrays enabling highly multiplexed electrical recording and

(iii) indirect reading of neuronal activity using genomic signatures of it, such as DNA sequence errors that correlate with the spiking patterns generated by calcium-dependent polymerases

- Innovation to develop new classes of nanoparticle reporters of voltage appear as a crucial aspect of this work. For example, the creation of highly-sensitive potentiometric reporters at small enough scale to embed within the neuronal cell membrane would permit a new era of large-scale recording of the intracellular potential, simultaneously, of large numbers of neurons.
- Critical to the success of large-scale efforts to permit deep tissue imaging/recording will be the packaging developed to interface the multichannel technology with the subject (e.g. animal) under test. This should permit minimally invasive, free movement of the awake animal with minimal-to-no discomfort.
- Subsequent technological innovation will permit deep electrophysiological or optical measurements using dense arrays of electrical or optical probes based on integrated photonics coupled with next-generation nanoparticle or genetically-introduced voltage-sensitive reporters. Local introduction of calcium or voltage-sensitive reporters may require probe technology comprising co-integrated microfluidics to permit controlled local delivery of fluorophores or viral vectors.
- Among possible long-range technological goals might be the development of complex, untethered, nanoscale probes – which could be based on semiconductor nanotechnology or synthetic biology – that permit local sensing, local data storage, and, potentially, remote interrogation/activation by externally-imposed electromagnetic fields (e.g. at radio frequencies).

2. Large-Scale Stimulation

- Our second goal it to be able to stimulate individual neurons in the nervous system independently at first, then combinatorially and also by means of natural sensory stimuli. This will first begin with: i) optical stimulation (e.g. by two-photon uncaging or photoactivation) of neural tissue perfused with optogenetic constructs or caged compounds, and ii) electrophysiological interrogation of neural tissue via neural probe arrays enabling highly multiplexed electrical stimulation. Subsequent technological innovation will permit deep electrophysiological measurements using dense arrays of probes based on integrated photonics coupled with next-generation nanoparticle or genetically-introduced optogenetic reporters.
- For optogenetic stimulation, new classes of genetically introducible, light-sensitive ion channels providing higher sensitivity and more varied spectral coverage will need to be developed, particularly ones that can be excited with two-photon light, to permit single-cell resolution in vivo.

3. Local Reporters, Passive (Genomic and Nanoparticle Reporters and Stimulators)

- Our third goal is to enable patterned stimulation, in principle in any arbitrary spatio-temporal pattern. In a way, this will resemble “playing the piano” with the neural circuit.

- To achieve any arbitrary stimulation patterns, optical efforts will harness holographic methods based on phase-only spatial light modulators (SLMs), using two-photon excitation.
- For electrical stimulation, dense arrays of neural probes can be used for highly-multiplexed, complex patterns of neural stimulation. Work to optimize the stimulation electrode/tissue interface to permit stable, long-term stimulation is critical; use of nanotechnology-based approaches to deterministically engineer optimal electrode interfaces is a unexplored area that will be tackled.
- Zero-, one-, and two-dimensional artificial nanostructures and nanoparticles will play critical roles in this project.
    - Zero-dimensional nanostructures can be manipulated to produce a new generation of local optical reporters for neuroscience. These reporters will need to be capable of being embedded into neural membranes (thickness ~5nm) and of being sensitive to local electric fields as well as local chemical environments. The design of these nanostructures will draw from the newly established ability to control plasmonic behavior in metallic nanoparticles, quantum size effects in semiconductor heterostructures with designed asymmetries, and nanoparticles with embedded dopants possessing sharp emission spectra. These inorganic nanoparticle optical probes can be tuned for to match the photon energy requirements of the various excitation and detection systems. Further, compared to organic optical probes, they will be photochemically robust during extended interrogation. They will need to be combined with organic nanostructures, that is, biofunctionalized, to direct/embed them within neural membranes or synapses. They may be combined with selective molecular binding moieties to confer sensitivity to changes in local neurotransmitter concentrations.
    - One-dimensional structures such as nanotubes and nanowires may be used for highly local electrical measurements, for the delivery of photons to specific locations, and for the local release or collection of chemicals.
    - Two-dimensional nanostructures such as graphene may be engineered into artificial membrane patches, providing new interfaces of our electrical systems to biological membranes.
- Among possible long-range technological goals might be the development of complex, untethered, nanoscale probes – which could be based on semiconductor nanotechnology or synthetic biology – that permit local sensing, local data storage, and, potentially, remote interrogation/activation by externally-imposed electromagnetic fields (e.g. at radio frequencies).

4. Long-Term Brain/Probe Interaction

- The goal of the project is to develop long-term approaches that can observe and manipulate the response of whole brains to complex stimuli. This requires non-perturbative interactions between the technological interfaces that are introduced into brain tissue.

- The response of neural tissue to foreign entities, that is, to each individual probe/reporter/stimulator introduced for this work must be understood and controlled. This will require coordinated work between the technology developers and researchers in **anatomical neuropathology**.
- The **biocompatibility** of materials used in the development of the nanodevices and nanoparticles must be assessed. This will involve coordinated, iterative research in close collaborations between technology developers and neurophysiological researchers.

5. Database Assembly and Computational Analysis

- This project will generate immense volumes of data in the form of multichannel time records of neuronal activity. These channels will be indexed by position within the brain, and time-correlated with various complex stimuli (which also could be complex, multichannel time records of directly-applied neural stimuli).
- Data mining of this immense body of digital information will require new paradigms in multidimensional correlative analysis of massively parallel time series data.
- New models for brain processing must aid and inspire this numerical analysis. These models are likely to originate by new collaborations involving engineers with expertise in the theory of networks, time series analysis, and physicists with expertise in synchronization phenomena and the nonequilibrium thermodynamics of pattern formation.
- Centralized supercomputer facilities to enable the handling of the massive data sets and their numerical analysis and physical modeling will be essential for this work.

6. Large-Scale Systems Integration

- Large-scale integration of the micro- and nano-scale devices used as the “front ends” of these systems must be fabricated en masse in state-of-the-art semiconductor foundries. Sufficient resources to permit research and development to arrive at stabilized processes for the wafer-scale production of device arrays will be essential.
- Large-scale production techniques for nanoparticle probes will be required. Once the prototype reporters or stimulation particles are perfected, their mass production will likely involve engaging a commercial factory that adheres to “best practice” standards.
- Creation of systems-level instrumentation will require assembly of a cadre of engineering experts. Small-scale production of such complex instrumentation is feasible within a center with a dedicated professional staff. (An analogy here could be made to the building of space probes, which is undertaken in similar fashion at, for example, Caltech’s Jet Propulsion Laboratory. Such a paradigm can provide prototype instrument systems to a community of beta-testers selected and engaged to help drive the program forward.
- Mass production will ultimately be required to deliver the technology in large scale to the neurophysiology community who will actually map the functional connectome. This is probably best done by commercializing the beta-tested prototypes; this will permit infusion of commercial capital to optimally engineer the systems for production.

7. Local Reporters, Active (Smart Nanoprobes)

- Long term possibilities may include the creation of "smart nanoprobes" achieved through either synthetic biology or advanced nanotechnology.
- The synthetic biology approach would involve developing voltage sensitive reporters that would record time records of neural response embedded through protein or nucleic acid synthesis into artificial constructs (for example using an ion-sensitive processive polymerase). Post-experiment dissection and analysis would permit subsequent readout of these local time records of neural response (for example by in situ DNA sequencing).
- The nanotechnological approach would involve engineering local potentiometric sensors within, say, ten-micron-scale capsules. These capsules would also include sufficient memory to store time records of local neuronal signaling and/or a means of real-time readout of the data. Such entities could be powered by immersion of the brain into radio-frequency fields or red light. Post-experiment data acquisition could be carried out either through dissection and downloading (as discussed for the case of the synthetic biology approach) or by remote interrogation.
- Remote interrogation of implanted smart-nanoparticle probes will require both assigning a position to each nanoparticle probe, and sequential download of their information in a coherent manner.

C) The Brain Activity Map: **Overarching Perspective**

We briefly address the scope and complexity of the undertaking we envisage, and compare it to known, present benchmarks.

In 1990, the possibility of sequencing a viral genome of 100,000 base pairs (at an error rate of 0.001) was considered feasible, but sequencing one human genome of 3×10^9 bp was considered unrealistic. Today, one group with 20 machines is able to sequence 10,000 human genomes per year; in other words a single group can sequence **6×10^{13} bp** = $10^4 \times (6\times10^9$ bp) at an error rate of 1×10^{-5}. This generates about 3×10^{15} bytes (3PB) of data. It is noteworthy that the fast pace of continuing technological advances are, at present, multiplying this already-impressive capacity by roughly 5-fold per year.

At the outset of the Human Genome Project, even the most rudimentary task of genomic combinatorics – that is, comparing each short region to each other region – was predicted by some to require an unapproachable number, $>1\times10^{19} =(3\times10^9)^2$, of computer operations. Today, many clever linear algorithms (such as BLAST) have displaced the naïve N-squared algorithms (like Needleman-Wunsch) invoked at the outset of genomics analysis. These advances have allowed combinatorics analysis using 3×10^9, and not $(3\times10^9)^2$, separate calculations. Also, it has since been discovered that genomic data are much more

compressible than some originally thought; in general, 5×10^6 bytes (5MB) is sufficient per each 6×10^9 bp genome. This compressibility has translated into reduced requirements, from those originally projected, for genomic data.

We now draw analogies to the structural complexity of the brain. The human cortex comprises 1.6×10^{10} neurons (80% glia), the whole human nervous system 1×10^{11}; the mouse cortex 4×10^6 neurons (65% glia), the whole mouse nervous system 4×10^7.

To estimate **data storage capacities** required for a brain activity map we consider the anatomic connectome. Bock *et al*. (Nature 2011) have covered 1500 cell bodies with 1×10^{13} raw pixels. By analogy we can estimate that 7×10^6 mouse cortical cells would require something of order 5×10^{16} bytes. We note that this is less data than the current genome image data worldwide. Further, astronomy and astrophysics are already awash with data; currently 1 PB (1 petabyte = 10^{15} bytes) of public data is electronically accessible, and this volume is growing at 0.5 PB per year. The availability of this data has transformed astronomy research. Projections indicate that by 2020, more than 60 PB of archived data should be accessible to astronomers.

We turn to estimates of **data bandwidth** required for real-time imaging of brain activity. For an activity map of the brain, a bandwidth of 10^4 bytes per second per neuron may be required. For 3×10^9 actively firing neurons in the brain this would correspond to a raw data acquisition rate of 3×10^{13} bytes (30TB) per second. As a benchmark, in 2012 the Australian Square-Kilometer Array Pathfinder (ASKAP) radio telescope is on track to handle a data stream of 10TB/second from the telescopes to its digital correlators. This bandwidth will be aggressively pushed for the next generation radio telescopes, estimated to require 100X the 2012 bandwidth.

In more close consideration of the **complexity** of the grand challenge of mapping brain activity, we acknowledge that direct analogies to genomic bioinformatics and data handling are, most likely, somewhat limited. For example, the informatics associated with brain activity mappings are of much higher dimensionality than are linear genomics sequences. Brains are dynamical systems with operations on a very wide range of time scales. Their component neurons are complex dynamical systems in their own right, and the synapses between them are plastic over a vast hierarchy of time scales (from milliseconds to, presumably, years). The spectrum of behavior of even the simplest neural circuits (for example, two reciprocally connected inhibitory neurons) has many solutions – comprising a few stable and many unstable ones. These solutions are all dependent on dynamic parameters describing the neurons and the connections between them.

Hence, in addition to the sheer numbers, brain activity maps will differ from genomics in other ways. Prominent among these are: (i) resulting combinatorics, (ii) the state dependence of interactions between neurons (from short-term facilitation to more complex nonlinear interactions) and (iii) neuronal biophysics, which are extremely varied, adapted and complex. Further, to make headway, it is most likely that some foreknowledge will be required of the

function of neural circuits. To date, it has not been possible to predict function (as in: to compute invariance to size or contrast) from connectivity.

Given these considerations, we anticipate that connectivity studies will likely need to be embedded in functional studies within a "traditional" framework (e.g. hippocampal function and spatial coding, or memory formation and retrieval, etc.). In other words, brain activity mapping will likely have to focus first on near term studies of brain subsystems: networks, circuits, and areas. Acquired data will become truly valuable when both connectivity and functional studies are carried out in the same brain. For this reason, we will initially espouse focusing on technologies that enable ultra-dense neuronal sampling of small brains, or in restricted areas in behaving animals, to studies that focus on being exhaustive over the entire brain, at the expense of compatibility with simultaneous functional studies.

Authors:

Michael Roukes, Kavli Nanoscience Institute, Caltech
Rafael Yuste, Kavli Institute for Brain Science, Columbia University
Ralph Greenspan, Kavli Institute for Brain and Mind, UCSD
George Church, Genetics Department, Harvard Medical School
Paul Alivisatos, Lawrence Berkeley Laboratory, UCB

This working group, with members whose research spans neuroscience, nanoscience, genomics, and systems biology, formed at a recent workshop exploring the interface of neuroscience and nanoscience. The workshop was organized by The Kavli Foundation, The Gatsby Charitable Foundation, and the Allen Institute for Brain Science and was held in September 2011, at Chicheley Hall, home of the Kavli Royal Society International Centre, UK.