NeuroMorpho.Org Implementation of Digital Neuroscience: Dense Coverage and Integration with the NIF

Maryam Halavi · Sridevi Polavaram ·
Duncan E. Donohue · Gail Hamilton · Jeffrey Hoyt ·
Kenneth P. Smith · Giorgio A. Ascoli

Abstract Neuronal morphology affects network connectivity, plasticity, and information processing. Uncovering the design principles and functional consequences of dendritic and axonal shape necessitates quantitative analysis and computational modeling of detailed experimental data. Digital reconstructions provide the required neuromorphological descriptions in a parsimonious, comprehensive, and reliable numerical format. NeuroMorpho.Org is the largest web-accessible repository service for digitally reconstructed neurons and one of the integrated resources in the Neuroscience Information Framework (NIF). Here we describe the NeuroMorpho.Org approach as an exemplary experience in designing, creating, populating, and curating a neuroscience digital resource. The simple three-tier architecture of NeuroMorpho.Org (web client, web server, and relational database) encompasses all necessary elements to support a large-scale, integrate-able repository. The data content, while heterogeneous in scientific scope and experimental origin, is unified in format and presentation by an in house standardization protocol. The server application (MRALD) is secure, customizable, and developer-friendly. Centralized processing and expert annotation yields a comprehensive set of metadata that enriches and complements the raw data. The thoroughly tested interface design allows for optimal and effective data search and retrieval. Availability of data in both original and standardized formats ensures compatibility with existing resources and fosters further tool development. Other key functions enable extensive exploration and discovery, including 3D and interactive visualization of branching, frequently measured morphometrics, and reciprocal links to the original PubMed publications. The integration of NeuroMorpho.Org with version-1 of the NIF (NIFv1) provides the opportunity to access morphological data in the context of other relevant resources and diverse subdomains of neuroscience, opening exciting new possibilities in data mining and knowledge discovery. The outcome of such coordination is the rapid and powerful advancement of neuroscience research at both the conceptual and technological level.

Keywords Neuronal morphology · Digital reconstruction · Data sharing · Neuroscience information framework · Axonal arbors · Dendritic trees

M. Halavi · S. Polavaram · D. E. Donohue · G. A. Ascoli (✉)
Center for Neural Informatics,
Structure, & Plasticity, and Molecular Neuroscience Department,
Krasnow Institute for Advanced Study, George Mason University,
Fairfax, VA, USA
e-mail: ascoli@gmu.edu

G. Hamilton · J. Hoyt · K. P. Smith
MITRE Corporation,
McLean, VA, USA

Introduction

The ability of the brain to process information, control behavior, and drive consciousness depends substantially on the formation and preservation of proper connections between axons and dendrites from different regions and neuronal classes. Etiological investigations in neurological and psychiatric disorders provide ample evidence for the crucial role of appropriate development and maintenance of neural circuits in healthy brain function. Each neuron must have proper synaptic partners in order to function effectively and accurately. The rich and unique morphological properties that are characteristic of each cell type in the nervous system play an essential role in targeting and invading different regions and layers, and in establishing the proper circuitry. The anatomy of the dendritic tree is a major determinant of synaptic integration (Rall and Rinzel

1973; Segev and London 2000; Gulledge et al. 2005; London and Hausser 2005) as well as cell excitability and the neural firing behavior (Mainen and Sejnowski 1996; Cuntz et al. 2007). Morphological diversity also relates to the intrinsic functional differences between neuron classes (Vetter et al. 2001; Stiefel and Sejnowski 2007). As a result, dendritic morphology influences both neural computation (Eilers and Konnerth 1997; Duch and Levine 2000; Tossit and Stocker 2000; Schierwagen and Claus 2002) and network function (Chklovskii 2004; Chen et al. 2006). Understanding the relationship between structure and function in the nervous system implies investigation of the complexity of and dynamic interactions among numerous neuron types (Bota and Swanson 2007a, b). These studies often combine quantitative analyses and computational models based on synaptic biophysics and realistic neuronal morphology as provided in three-dimensional (3D) digital reconstructions.

NeuroMorpho.Org is to date the largest repository service for neuromorphological reconstructions (Fig. 1). Digitally reconstructed neurons can be used and re-used in various research projects with different scientific aims (Ascoli 2007). Accessibility is one of the main obstacles that prevent researchers from re-using data. While a large amount of morphological data is collected in neuroscience, only a limited fraction is practically shared (Liu and Ascoli 2007). Only few laboratories explicitly refuse to share their data; many others agree in principle, but are unwilling or unable to invest the time and technical efforts required for data sharing (Ascoli 2006a). One of the main contributions of Neuro-Morpho.Org is to provide technical support and computational infrastructure to facilitate the task of researchers eager to share their morphological data. A major resulting benefit for the users is the option to quickly and efficiently identify the data of their particular interest and gather related information and metadata. A significant portion of such data may otherwise not become available or require additional time to be discovered. Moreover, an important level of scientific and numerical quality assurance is attained, because data are collected from peer-reviewed publications and files are pre-processed for inclusion in NeuroMorpho.Org.

NeuroMorphpo.Org also extends data usability and visibility by making morphological reconstructions accessible through the Neuroscience Information Framework (NIF). The NIF is a pioneering initiative that fosters the seamless integration of neuroscience resources from different domains of expertise into a single search engine (Gardner et al. 2008). The close link between Neuro-Morpho.Org and NIF was naturally inspired by the wide relevance and applicability of neuronal morphology in many other neuroscience resources. Currently in its third year of existence, NeuroMorpho.Org continues to improve its functionality and expand its data content. The main goal of this paper is to describe the NeuroMorpho.Org approach to digital neuroscience, its integration with other NIF resources, and some of the most instructive challenges and solutions in this process. A more coincise and less technical description for the general neuroscience readership, aimed at users rather than developers, was recently published (Ascoli et al. 2007). Interested readers are also invited to browse and download the data and a wealth of additional information and explanation online at Neuro-Morpho.Org, and to provide feedback.

Fig. 1 Screen shot of the main page of NeuroMorpho.Org, the starting point for navigating through the site with a complete list of options on the left

Design of NeuroMorpho.Org

To maximize its value and durability, NeuroMorpho.Org adopted available bioinformatics standards and guidelines, while making necessary adjustments to accommodate the specific needs of its morphological content. Neuronal reconstructions are heterogeneous in terms of both their numerical specification (file format, size, resolution, etc.) and the scope of the originating studies (e.g. anatomical, behavioral, electrophysiological, pharmacological or developmental). The NeuroMorpho.Org infrastructure had to develop a unified set of design principles and choice of metadata to achieve several important goals:

- to store the various data types according to a predefined structured schema;
- to support different reconstruction formats, while providing a standardized output;
- to facilitate complex queries and data retrieval;
- to enable rapid development and deployment while remaining user-friendly;
- to be interoperable with other neuroscience resources;
- to optimize reliability, security, robustness, and scalability.

There have been several previous efforts to create repositories for neuronal reconstructions (reviewed in Ascoli 2006b). An extended list of these existing archives is available at NeuroMorpho.Org under the "Tools and Links" menu. In general, all data available through these databases are mirrored in NeuroMorpho.Org. There are however substantial differences that make NeuroMorpho.Org unique among such resources. The foremost is that data in NeuroMorpho.Org are contributed by a large (and continuously increasing) number of different researchers rather than an individual laboratory. A primary goal of NeuroMorpho.Org is to achieve and maintain dense coverage of all the publicly available digital reconstructions rather than provide a static venue to distribute a particular subset of neuronal morphologies.

Metadata Development

Reconstructions in NeuroMorpho.Org are organized according to cell types, animal species, brain regions, technical protocols, tracing methods, numerous morphometrics measures, and several other dimensions. In general, metadata greatly adds to the significance and scientific value of data, enables effective user searches, as well as integration with related resources. Three general categories of metadata in NeuroMorpho.Org relate to (1) the data source, such as the laboratory and researcher providing the reconstructions, the reference articles, and the archive internet address (if any); (2) the subject of the study, such as animal species and strain, area of the brain, and neuron

type; and (3) the experimental methodology, such as histological protocols, reconstruction hardware and software, and format of the original data. The absence of a widely accepted neuroscience ontology was a cause of considerable hardships in the process, which stresses the need for a standard terminology in the field. We attempted to follow the classifications used in the original studies describing the reconstructions. In particular, the detailed information retrieved from peer-reviewed journal publications was initially evaluated, and relevant data was extracted. As more studies and data were processed, metadata was optimized accordingly. Several examples of metadata and their descriptions are provided in Table 1. The database schema is continuously updated and publicly posted on NeuroMorpho.Org under the "Tools and Links" tab.

Web Architecture

The basic NeuroMorpho.Org framework consists of a standard three tier architecture (web client, web server, and relational database). This organization is scalable, robust, and flexible, while at the same time allowing for easy management and network deployment (Fig. 2). All the original data provided by researchers are stored in a back-end relational database, including raw and processed files, images, and metadata (Fig. 2a). NeuroMorpho.Org utilizes MySQL V5.0 as the database management system, with several add-on custom applications (written in Java, C + +, and Matlab). The web application server, Apache Tomcat 5.5, runs on a 2.0 GHz Intel dual quad core processor machine under the Linux Fedora 8 operating system.

Web-accessible, secure, and user-friendly interaction with the images and metadata in the back-end database is enabled by MRALD (Blake et al. 2002). MRALD is a platform and database independent application, is easily customized to new domains, and has been deployed in mission critical systems (e.g., aviation) continuously since 2001. MRALD's form builder enables system designers to rapidly generate intuitive hyper-text markup language (HTML)-based data retrieval forms; forms can be associated with specific users for access control. Data interaction is also possible via custom java server pages (JSPs) and keyword search. Hidden from the user, MRALD translates requests into structured query language (SQL) and interrogates the underlying database via Java database connectivity (JDBC). It can return results in multiple formats, including HTML, extensible markup language (XML), comma or configurable separated values (CSV), tab-delimited text, Excel spreadsheet files, and in other, user-defined, formats. MRALD's web-based administration features include form update, insertion, and deletion; a schema visualization tool; user account management; and the ability to assign data and users to collaborative

Table 1 Examples of metadata extracted from peer-review publications and their descriptions

Sample metadata	Description
Archive_name	The name of the primary laboratory that has performed the study.
Neuron_name	The name of the morphological data file as assigned by the providing lab.
Maximum_age and Minimum_age	The age of the animals. This is typically reported as a range. Therefore, a maximum and/or a minimum age are included.
Age_scale	The scale used for reporting animal age. Usually, age is reported either in days, months, or years as indicated by a preceding letter (D, M or Y).
Development	Qualitative classification of age (Young, Adult or Old). In some publications this classification is provided without reporting the precise animal age.
Sex	Gender of the animal. It is not always specified. Therefore, it is classified explicitly as: male (M), female (F), either sex (M/F), or not reported (NR).
Brain_region	The different areas of the brain that neurons belong to. We organized the brain regions in three different classification levels, starting from main anatomical regions toward the finer distinction of the layers in the cortex.
Cell_type	The morphological class assigned to the neuron. Each neuron is classified based on three cell type levels. At the first level, reconstructions are distinguished as: Principal cell, Interneuron or Axonal terminal.
Original_format	Information regarding the original format of each data file.
Protocol	Classification of the experimental method based on whether the dye injection occurred in the live animal, in a slice, or in neuron culture.
Slicing_thickness	The thickness of the slices cut during tissue processing. It may serve as an important determinant of the reconstruction completeness.
Slicing_direction	The orientation of the sampled tissue cut.
Stain	The type of dye used for staining the neurons.
Objective_Magnification	The magnification used to trace and reconstruct the neurons.
Objective_type	The type of the objective used in the reconstruction process.
Reconstruction_method	The software used for tracing the neurons.
URL reference	The web address of the data file, if it is available on-line.
Notes	Any additional, important information, not stored in other fields.

communities (Smith et al. 2004). MRALD's internal workflow processing for translating HTML into SQL is customizable, and can be extended by a developer to insert new steps (e.g., filters) into the normal processing pipeline (Fig. 2b). MRALD is freely available to the academic community at neuroinformatics.mitre.org.

On the front-end web page, NeuroMorpho.Org offers three complementary search interfaces, namely by metadata, by morphometric measure, and by keywords. In the first interface, search terms are available to the user as dropdown menus grouped in four general sections (Animal, Experiment, Anatomy, or Source). Within each section, menus reflect the underlying metadata types (e.g. sex and age; histological protocol and reconstruction method; brain region and cell type; deposition date and original format). Sub-menus appear dynamically as appropriate (Fig. 3). For example, upon selection of the term "Rat" under "Species", a "Strain" sub-menu is offered including relevant options ("Sprague Dawley", "Long-Evans", etc.). In the second search interface, value ranges can be assigned to morphometric features such as soma surface, number of branches, arbor length or volume, height, width, etc. Features can be specified alone or in combination to construct simple or more complex queries. The third interface allows users to
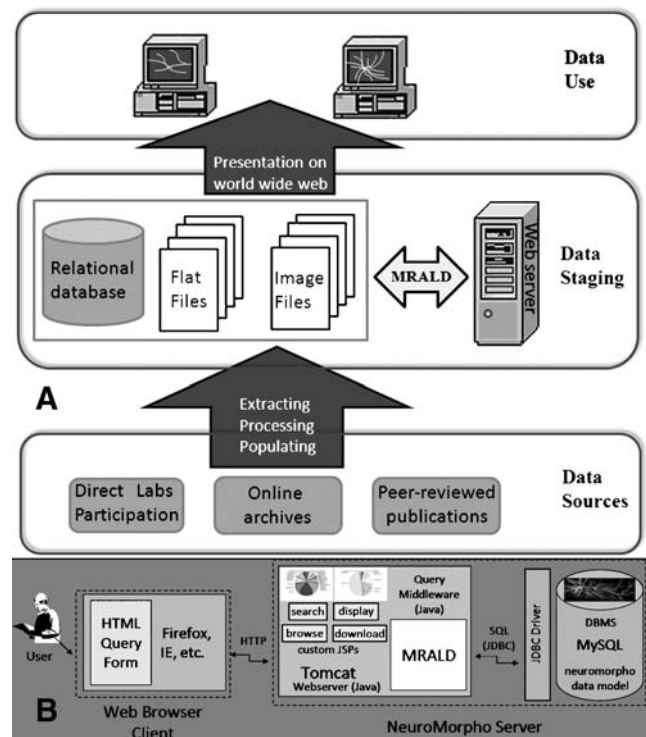

Fig. 2 Structure and information flow among the components of NeuroMorpho.Org. (a) Organization of the processing pipeline. (b) Role of MRALD in the three-tier data retrieval architecture

Fig. 3 Representative search through the Metadata page. Filter criteria can be combined and fine tuned through cascades of drop-down menus to increasing levels of detail. In this example, Monkey is selected as species, Macaque as strain, Lucifer Yellow as staining method, and Local projecting Pyramidal cell as the cell type and subtype of the principal cell class. These selections returned 20 neurons as the results set

retrieve data by typing keywords in a simple search bar (Ascoli 2007b).

When activated by any of these three interfaces, MRALD processes the queries, generates the results dynamically, and sends them to the web client. The number of cells matching a given set of search criteria can also be requested before visualizing the results. The simplest option to display the search results is in a Summary format, in which each neuron is represented with a thumbnail image and an abridged set of its metadata. Clicking on one of these neuron entries calls the individual page of that reconstruction, with links to all raw and processed data, metadata, and related files (described below).

Alternatively, users can browse through search results (or the entire content of NeuroMorpho.Org) after sorting them by any of four criteria: brain region, animal species, cell type, and laboratory name. An overall view is rendered as a mouse-sensitive pie chart (http://cewolf.sourceforge.net), and results are visualized as they are loaded to minimize wait time even for massive data sets (Fig. 4). The result of this organization is that a user can navigate from a conceptual query to the raw data in three clicks.

The data presentations by brain regions, cell types, and animal species have clear biological meaning, to be organized according to the NIF standard ontology (Bug et al. 2008). The view by laboratory name can be useful for data contributors to demonstrate to funding agencies and promotion committees that they have followed through with an effective data sharing plan.

Data Presentation

NeuroMorpho.Org does not require any user registration or login to search and download data. For each neuron in the database, both graphic representations and flat files are made available through direct links for visualization and download. Flat files include the original reconstruction file as provided by the laboratory of origin, the version converted into a standardized format, the log detailing all modifications, and a document listing any remaining notes or irregularities (see *Standardization process* section below). Users may choose to download one or all of these four files for any number of neurons as a single compressed archive. Each neuron is illustrated with a static two-dimensional image as well as a 3D animation of the
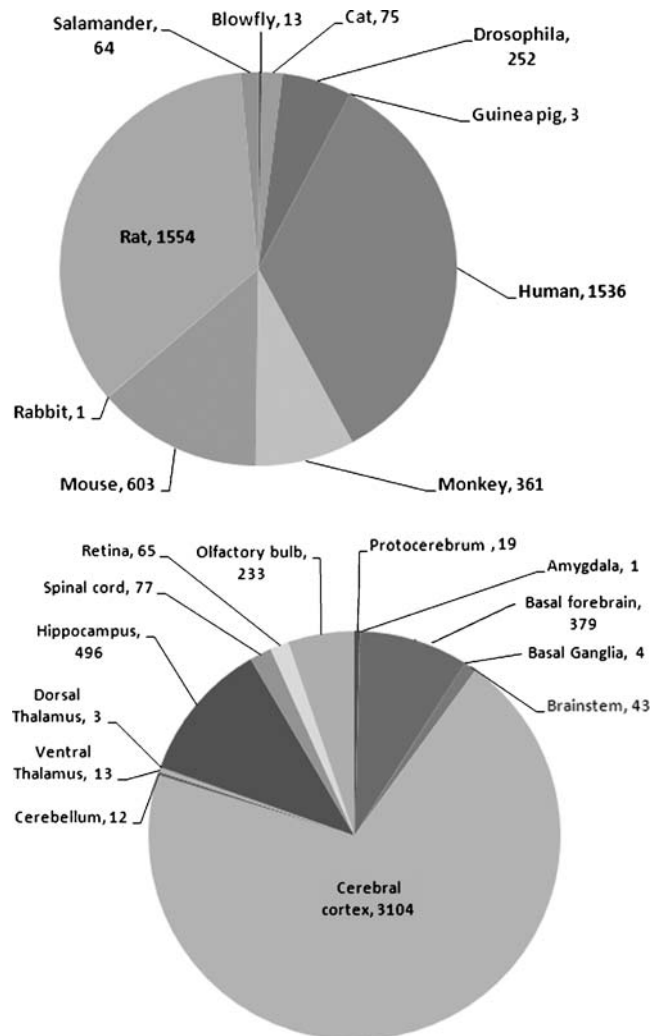

Fig. 4 Data content diversity in NeuroMorpho.Org: pie chart representations of different species (*top*) and brain regions (*bottom*)

extending arborization while it rotates around the cell central axis. Moreover, to allow for interactive 3D manipulation of neurons, the Cell Viewer Application Cvapp (Cannon et al. 1998) was custom modified, streamlining functionality and enabling automatic online deployment through the Java Network Launching Protocol (JNLP).

The database also stores the PubMed identity (PMID) for all referenced papers (see *Data model and data management* section below), and a corresponding XML file, created through Java Server Pages (JSP), is accessed by the NIF Broker that mediates the Entrez LinkOut functionality service provided by PubMed (Marenco et al. 2008). This architecture design allows direct reciprocal access between the peer-reviewed reference and the raw data. In particular, a link from the individual neuron page to the PubMed abstract of the publication(s) describing the experiment provides the users with a broader perspective on the reconstructions. To access the reconstructions from PubMed, users can follow the LinkOut option on the top right corner Links menu (see Fig. 2 in Marenco et al. 2008), which leads to NeuroMorpho. Org through the Neuroscience Database Gateway, a precursor of the NIF (see Fig. 3 in Marenco et al. 2008).

Integration of NeuroMorpho.Org with the NIF

Neuroscience tools, data, information, and knowledge can often be related to the structure of neurons. The long recognized centrality of neuronal morphology in the neuroscience community led to the early identification of Neuro-Morpho.Org as a foundational resource in the development of the NIF. At the same time, integration of NeuroMorpho. Org with the NIF facilitates queries in a contextually rich environment, which would not be otherwise available within the restricted domain of digital neuromorphological reconstructions. Moreover, the integration process itself serves as a useful step for benchmarking the technological standards within the field. NIF has faced and successfully overcome numerous challenges to achieve interoperability among resources with respect to hardware, software, communication protocols, user application, and data compatibility.

Resource Registration, Concept Mapping, and Query Mediator Protocols

As with other resources, NeuroMorpho.Org is catalogued as an external public database in the NIF registry. NIF registration requires basic information about the resource's setup, such as name, URL, administrative and technical contacts, content type, and data availability. In addition to this "superficial" registration, functional interactions of NeuroMorpho.Org with NIF also required a deeper registration at the level of the database schema design. In particular,

an essential element of such deep registration is the consistent mapping of corresponding concepts among resources. This step implied sharing access to the standard Java Database Connectivity (JDBC) driver and a complete documentation of the database tables and attribute definitions.

Most attribute names of the tables in the NeuroMorpho. Org database were matched to existing concept identifiers from NIFSTD, the NIF standardized terminology pool (Bug et al. 2008). For example, the terms "Purkinje cell" and "Granule cell" in NeuroMorpho.Org are respectively mapped to "birnlex 867-Purkinje" and "nifext_153-Dentate gyrus granule cell" in NIFSTD. NIFSTD also incorporates cell type relations and terms from other related efforts, such as the Open Biomedical Ontologies (OBO: Bard et al. 2005) and the sub-cellular anatomy ontology (SAO: Larson et al. 2007). The partially hierarchical organization of NeuroMorpho.Org metadata (e.g. with the primary distinction of cell types among "principal cells", "interneurons", and "axonal terminals") facilitated local term mapping onto the NIFSTD and other relation ontologies (Fig. 5). As it continues to evolve, the NIFSTD vocabulary will be used as a controlled terminology for NeuroMorpho.Org, including the adoption of terms beyond cell type (e.g. species and brain regions). The ongoing growth of NeuroMorpho.Org in return will feed back to the NIF vocabularies with new terms necessary to describe neuronal reconstructions added at each successive release.

Deep registration and concept mapping acquire practical utility from the user's perspective if the corresponding resources have programmatically cross-accessible query interfaces. In particular, this deeper registration allows users to "drill down" into the content of dynamic resources such as NeuroMorpho.Org, as explained in the next paragraph. Typical application generated requests are exemplified in the next section as well. NeuroMorpho.Org is also technically interoperable in that it allows direct external access to the data via URL embedded queries that accept name-value pairs specifying data source, SQL query, and output format.

Advanced Functionality Enabled by Integration

The NIFSTD terminology mapping allows concept based search on federated resources. The NIF search engine retrieves data by interrogating other resources in parallel based on their respective semantics as opposed to the original query string. This potential "drill down" usage of NIF with respect to the research domain of NeuroMorpho. Org can be illustrated with a few examples. At the simplest level, users gain access to data and information from multiple resources at once. For instance, a NIF keyword search for "Purkinje cell" retrieves results from the Cell Centered Database (CCDB: Martone et al. 2002), SenseLab (Shepherd et al. 1998), and NeuroMorpho.Org. Thus, the
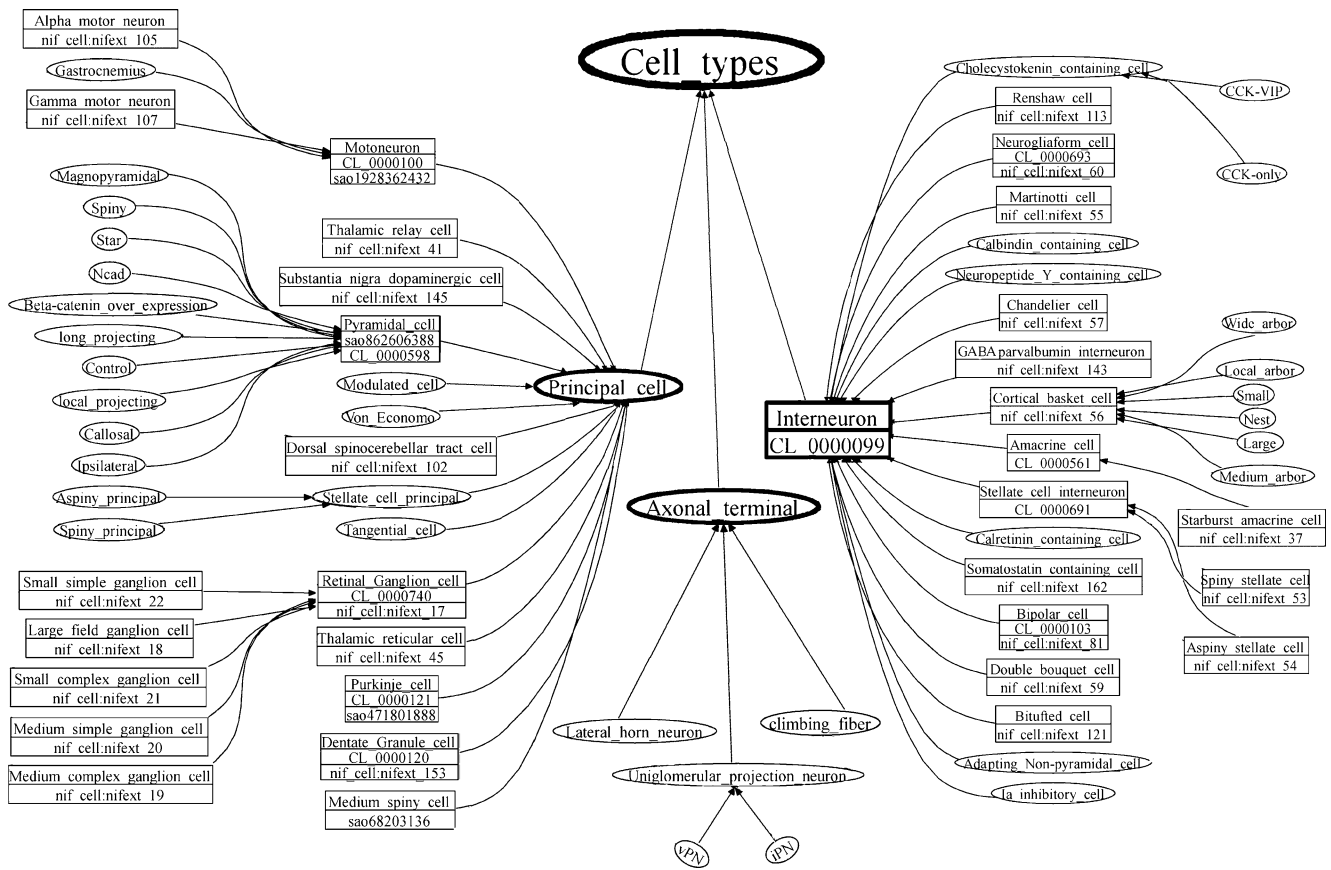
Fig. 5 A schema representation of cell type mapping between NeuroMorpho.Org and three relevant ontologies: NIF Standard (NIFSTD), Open Biomedical Ontologies (OBO), and Subcellular Anatomy Ontology (SAO). Each NeuroMorpho.Org cell type is represented as a tree node in the hierarchy. Rectangular nodes are mapped onto equivalent terms in NIFSTD (prefix "nif_cell"), OBO (prefix "CL"), or SAO (prefix "sao"). The nodes in oval shape are not mapped. Arrows represent "is-a" relationships (e.g. "Chandelier cells are interneurons")

user is provided at once with subcellular microscopy images, physiological properties, and morphological reconstructions for this cell type.

A more complex situation could be envisioned with a NIF search for "Nucleus accumbens". NeuroMorpho.Org returns 379 basal forebrain reconstructions: 146 large aspiny cells and 232 medium spiny neurons from the rat, and 1 medium spiny neuron from the mouse. One of the two corresponding reference article reports that nicotine may cause enduring changes in both classes of these cells. The CCDB-retrieved high resolution tomographic images show the distribution of spines on a single dendritic branch. Combination of spine density from CCDB data with the arbor metrics from NeuroMorpho.Org suggests a hypothetical effect of nicotine on subcellular volumes. Conveniently, NIF also returns results from the National Institutes of Health grant database CRISP (Computer Retrieval of Information on Scientific Projects), broadening the context of this search to a related list of ongoing funded research.

In another scenario, a user is designing a research project to understand the kinetics of K^+ channels in dentate granule cells of Sprague Dawley rats. A NIF search seamlessly yields 44 digital dendritic reconstructions from NeuroMorpho.Org, 21 compartmental models from SenseLab, available monoclonal antibody for Kv subunits from NeuroMab.org, a set of interacting protein along with the gene kcnip3 from Entrez Protein (ncbi.nlm.nih.gov/Database), and gene expression patterns for KcnK1 from Gensat.org. Combining all of this information gives the user a wider perspective on the topic of interest and the available data. Depending on the goal of the user, such direct capability to mine heterogeneous databases can positively impact various stages of research. As the usage statistics of NIF grows, so does the motivation for sharing and organizing data, which in turn will also require expanding the scope of interoperability of this resource.

Data Model and Management

There are four criteria for inclusion of anatomical data in NeuroMorpho.Org. (1) The content of each entry must

Table 2 Historic review of version releases to date and their number of reconstructions

Version	Release date	Data content
3.0	07/15/2008	4462
2.2	02/29/2008	2185
2.1	12/20/2007	1906
2.0	10/15/2007	1826
1.2	08/29/2007	1412
1.1	05/04/2007	1160
1.0	12/22/2006	1029
Beta	09/20/2006	1000
Alpha	08/01/2006	932

include branching structures (axons and/or dendrites) of individual neurons with a unique root (typically the soma), as opposed to e.g. fiber tracts or disconnected neuropil. (2) The data must explicitly represent the arbor connectivity in vector format (e.g. midline positions and diameter), rather than membrane surface contours or volumetric intensities (image stacks or voxel sets). (3) The data must be freely available except for requirements of proper credit assignment. (4) The data must be described or used in a peer-reviewed publication. The "dense coverage" goal of NeuroMorpho.Org is to include essentially all data that

meet these four criteria. The initial data were gathered by mining available on-line archives and through direct peer-to-peer requests to individual labs known to the database curators. This first round of collection yielded just shy of 1,000 digital reconstructions, and resulted in the first NeuroMorpho.Org release in August 2006 with 932 neurons. Several version updates have since considerably expanded the data content (Table 2) and improved the site functionality.

Metadata Extraction

Upon receipt of data meeting the above four criteria, each neuron is assigned a unique identification number, which is used as the primary organizational key throughout the application. Data content in NeuroMorpho.Org combines flat files, images, and textual metadata, all in several formats. Flat files consist of raw and processed digital morphologies (see *Standardization process* below). Images include both static (300×240 pixel) and animated illustrations of each cell (see *Data presentation* above), created with Cvapp and MatLab, respectively, with the aid of in-house scripts to semi-automate the process. Detailed information in text-based format is retrieved from the
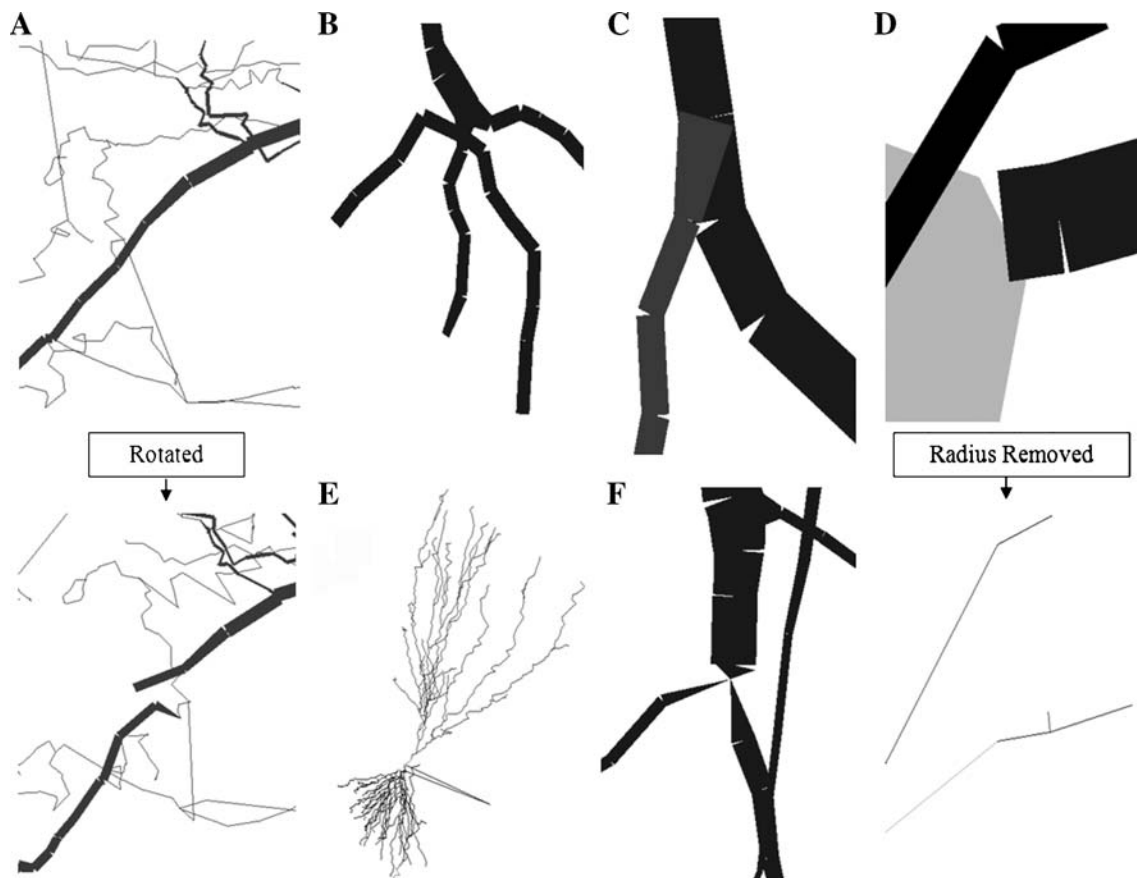
Fig. 6 Examples of morphological irregularities in the original data before processing for inclusion in NeuroMorpho.Org

corresponding journal article publication. A considerable amount of effort is expended reviewing publications to gather and annotate the metadata for insertion in the repository. When the necessary information is missing in the original scientific reports and related online sources, we contact the authors directly. Additional metadata is extracted in the form of morphological measurements with L-Measure, which is freely available (Scorcioni et al. 2008).

The main structure of the underlying relational database is organized around 27 metadata tables. The related images and flat files are loaded into the system and indexed in the database for retrieval in the detailed page dynamically generated for each neuron upon query. Similarly, the abstract of and PubMed link to each paper are inserted in the database for reference. The morphological measurements are also stored for use in the morphometry search. The core table links the primary identification key to all metadata fields as well as to these expanded tables in support of complex data retrieval. This flexible data model allows the seamless addition of new metadata fields and related information as needed or desired.

Standardization Process

When neurons are digitally reconstructed in various laboratories and by different operators, the resulting morphologies come in diverse formats and often include several types of peculiarities, idiosyncrasies, and irregularities (some of which are graphically represented in Fig. 6). In order to ensure a base level of homogeneity and compatibility with available analysis, visualization, and modeling tools, all reconstructions in NeuroMorpho.Org are converted into the common SWC format using L-Measure (Scorcioni et al. 2008). The resulting files undergo a process of standardization using a custom Java software program which reads SWC files, fixes some irregularities automatically, and flags other anomalous lines as potential irregularities to be corrected by hand.

Table 3 Abridged examples of cell types included in Neuro-Morpho.Org from different brain regions

Region	Cell class	Number of neurons
Basal forebrain	Medium spiny cell	233
Basal Ganglia	Dopamine cell	4
Cerebellum	Climbing fiber	2
	Purkinje cell	10
Cerebral cortex	Adapting non-pyramidal cell	1
	Basket cell	64
	Bipolar cell	6
	Bitufted cell	47
	Chandelier cell	4
	Double bouquet cell	9
	Martinotti cell	36
	Neurogliaform cell	30
	Neuropeptide Y (NPY) containing cell	11
	Parvalbumin (PV) containing cell	8
	Pyramidal cell	2509
	Somatostatin (SOM) containing cell	33
	Stellate cell	32
	Von Economo	29
Dorsal Thalamus	Relay cell	5
Hippocampus	Calbindin (CB) containing cell	18
	Calretinin (CR) containing cell	29
	Cholecystokinin (CCK) containing cell	14
	Granule cell	82
Olfactory bulb	Uniglomerular projection neuron	233
Protocerebrum	Lateral horn neuron	19
Retina	Amacrine cell	1
	Ganglion cell	64
Spinal cord	Dorsal Spinocerebellar tract (DSCT) cell	8
	Ia inhibitory cell	8
	Motoneuron	100
	Renshaw cell	4
Ventral Thalamus	Modulated cell	12
	Reticular cell	1
Visual Lobe	Tangential cell	13

All fixed and flagged lines are documented in a standardization log file with an alphanumeric code (defined in the help), a text description of the anomaly, and any action taken to remedy it. The numerical code denotes the type of error (e.g. "4.1" indicates a point with zero diameter). Automatically fixed lines are designated as "type A"; flagged lines are designated as "type B1" unless they are corrected by hand, in which case they are changed to "type B2". Any additional irregularities noted by the operator, but not detected by the program, are designated as "type C". To facilitate processing of reconstructions with complicated or numerous corrections, the visualization and editing program Neuromantic (also freely available at www.rdg.ac.uk/neuromantic) has been modified to automatically track changes and append corresponding descriptions to the comment section of the SWC file. At the end of this process, the standardization program is run once again on the resulting SWC files to create a log of potential "remaining issues".

For further quality assurance, all files, images, and data are first uploaded into a temporary website for inspection and approval by the original providers of the raw data. This step also ensures that proper credit is assigned, and minimizes the risk of inaccurate representations. After implementing the eventual changes requested by data owners, the new release is uploaded onto the main site.

Access Statistics

Currently, NeuroMorpho.Org contains more than 4,400 neurons from ten different species, fourteen distinct brain regions, and more than 35 cell types (Table 3 and Fig. 4). Additional neurons have already been collected and are being processed for inclusion in the next release. Furthermore, another 4,000 neurons have been identified in the literature and will be requested from the authors. As the number of data sets in this repository expands, so will the opportunities to mine for new scientific insights, such as developmental rules and structural-functional relations. Recent citations already demonstrate usage of data from NeuroMorpho.Org in the published literature (Crook et al. 2007; Bar-Yehuda and Korngreen 2008; Fleidervish and Libman 2008; Wen and Chklovskii 2008). Several encouraging projects are also active regarding related tool development (e.g. www.neuronland.org).

The application keeps track of all access to the web site (Fig. 7). A complete history is maintained to aid site improvement and future planning. Moreover, these reports show how often each and every data set is downloaded and which queries are most used. The usage log file is evaluated monthly to harness the desired statistics and track the number of downloads. At the time of this writing, NeuroMorpho.Org has had more than 10,000 visitors from 55

Fig. 7 Usage of NeuroMorpho.Org. (**a**) A chronological illustration of the number of site access and file downloads on a logarithmic scale. (**b**) Distribution of session profiles based on the relative proportion of downloads per hit
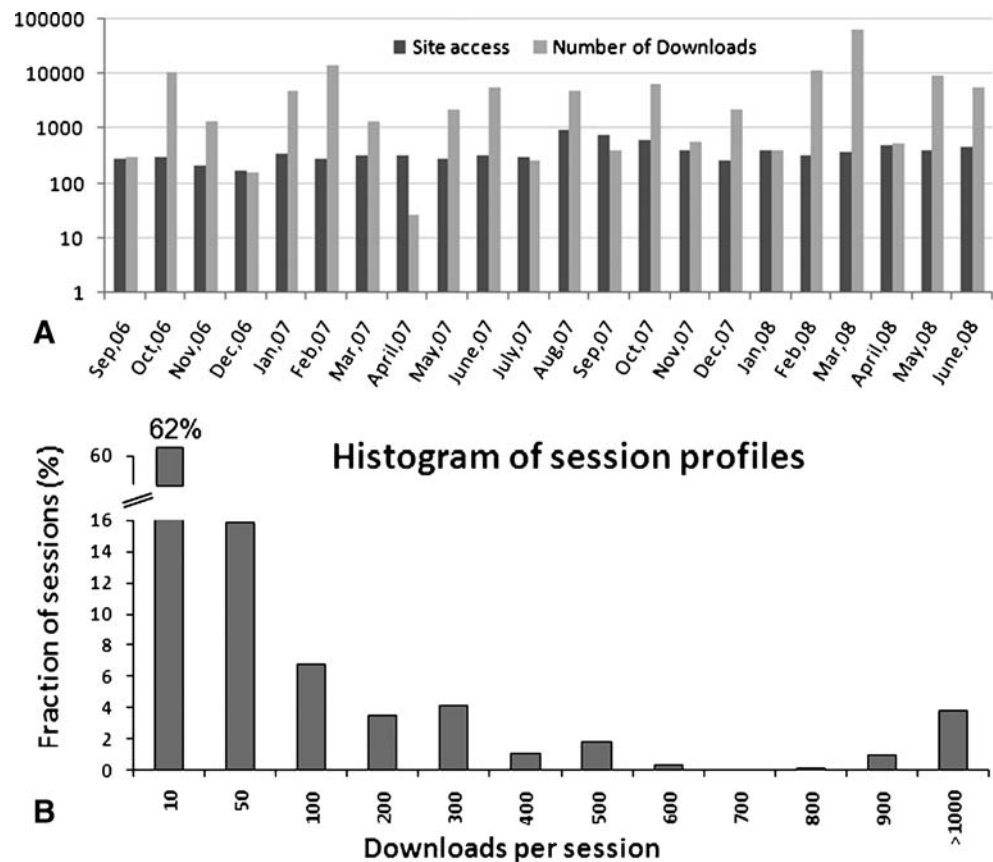
countries worldwide, with a steady flow of at least 250 to 300 visits per month (Fig. 7a). A total of more than 250,000 files have already been downloaded (up to 70 downloads per neuron). Although some users are downloading massive amounts of data at a time, the majority of searches and downloads are targeted at specific subsets of reconstructions (Fig. 7b). These results suggest that scientists are using NeuroMorpho.Org to access neuronal reconstructions, indicating that this resource has become a successful reference to browse neuroanatomical information, evaluate morphometric measurements, and download the data of interest.

An even broader awareness of NeuroMorpho.Org is expected to result from the NIF integration, as preliminarily evidenced from the usage statistics after the recent NIF release. Increasing usage of NeuroMorpho.Org and all other NIF-related resources will ultimately prove that integration benefits the whole neuroscience community, adding value to the data, and providing greater opportunity for new discoveries. At the same time, it is important to remember that a critical determinant of the success of NeuroMorpho.Org is the intent to cover densely all the digital reconstructions available for public sharing. In light of the NIF aim and scope, the importance of earning the user's confidence that, if a data file can be obtained, it will be found in such a database, is a key lesson for those seeking to create similar digital resources.

Information sharing statement All data, metadata, information and computational infrastructure of NeuroMorpho.Org are publicly available.

References

Ascoli, G. A. (2006a). The ups and downs of neuroscience shares. *Neuroinformatics*, 4(3), 213–216. doi:10.1385/NI:4:3:213.

Ascoli, G. A. (2006b). Mobilizing the base of neuroscience data: the case of neuronal morphologies. *Nature Reviews. Neuroscience*, 7 (4), 318–324. doi:10.1038/nrn1885.

Ascoli, G. A. (2007). Successes and rewards in sharing digital reconstructions of neuronal morphology. *Neuroinformatics*, 5 (3), 154–160. doi:10.1007/s12021-007-0010-7.

Ascoli, G. A., Donohue, D. E., & Halavi, M. (2007). NeuroMorpho. Org — A central resource for neuronal morphologies. *The Journal of Neuroscience*, 27, 9247–9251. doi:10.1523/JNEUROSCI.2055-07.2007.

Bard, J., Rhee, S., & Ashburner, M. (2005). An ontology for cell types. *Genome Biology*, 6, R21. doi:10.1186/gb-2005-6-2-r21.

Bar-Yehuda, D., & Korngreen, A. (2008). Space-Clamp problems when voltage clamping neurons expressing voltage-gated conductances. *Journal of Neurophysiology*, 99, 1127–1136. doi:10.1152/jn.01232.2007.

Blake, B., Hamilton, G., & Hoyt, J. (2002). Using component-based development and web technologies to support a distributed data management system. *Annals of Software Engineering*, 13, 13–34. doi:10.1023/A:1016537224754.

Bota, M., & Swanson, L. W. (2007a). Online workbenches for neural network connections. *The Journal of Comparative Neurology*, 500, 807–814. doi:10.1002/cne.21209.

Bota, M., & Swanson, L. W. (2007b). The neuron classification problem. *Brain Research. Brain Research Reviews*, 56, 79–88. doi:10.1016/j.brainresrev.2007.05.005.

Bug, W., Ascoli, G. A., Gardner, D., Grethe, J. S., Gupta, A., Fennema-Notestine, C., et al. (2008). The NIFSTD vocabulary: Rationale, design considerations and scope. *Neuroinformatics*, this issue.

Cannon, R. C., Turner, D. A., Pyapali, G. K., & Wheal, H. V. (1998). An on-line archive of reconstructed hippocampal neurons. *Journal of Neuroscience Methods*, 84(1–2), 49–54. doi:10.1016/S0165-0270(98)00091-0.

Chen, B. L., Hall, D. H., & Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12), 4723–4728. doi:10.1073/pnas.0506806103.

Chklovskii, D. B. (2004). Synaptic connectivity and neuronal morphology: two sides of the same coin. *Neuron*, 43, 609–617.

Crook, S., Gleeson, P., Howell, F., Svitak, J., & Silver, R. A. (2007). MorphML: level 1 of the NeuroML standards for neuronal morphology data and model specification. *Neuroinformatics*, 5 (2), 96–104. doi:10.1007/s12021-007-0003-6.

Cuntz, H., Borst, A., & Segev, I. (2007). Optimization principles of dendritic structure. *Theoretical Biology & Medical Modelling*, 4, 21. doi:10.1186/1742-4682-4-21.

Duch, C., & Levine, R. B. (2000). Remodeling of membrane properties and dendritic architecture accompanies the postembryonic conversion of a slow into a fast motoneuron. *The Journal of Neuroscience*, 20(18), 6950–6961.

Eilers, J., & Konnerth, A. (1997). Dendritic signal integration. *Current Opinion in Neurobiology*, 7, 385–390. doi:10.1016/S0959-4388(97)80067-0.

Fleidervish, I. A., & Libman, L. (2008). How cesium dialysis affects the passive properties of pyramidal neurons: implications for voltage clamp studies of persistent sodium current. *New Journal of Physics*, ▪▪▪, 10.

Gardner, D., Akil, H., Ascoli, G. A., Bowden, D. M., Bug, W., Donohue, D. E., et al. (2008). The Neuroscience Information Framework: a data and knowledge environment for neuroscience. *Neuroinformatics*, this issue.

Gulledge, A. T., Kampa, B. M., & Stuart, G. J. (2005). Synaptic integration in dendritic trees. *Journal of Neurobiology*, 64, 75–90. doi:10.1002/neu.20144.

Larson, S. D., Fong, L. L., Gupta, A., Condit, C., Bug, W. J., & Martone, M. E. (2007). A formal ontology of subcellular neuroanatomy. *Front. Neuroinform*, 1(3).

Liu, Y., & Ascoli, G. A. (2007). Value added by data sharing: long-term potentiation of neuroscience research. A commentary on the 2007 SfN satellite symposium on data sharing. *Neuroinformatics*, 5(3), 143–145. doi:10.1007/s12021-007-0009-0.

London, M., & Hausser, M. (2005). Dendritic computation. *Annual Review of Neuroscience*, *28*, 503–532. doi:10.1146/annurev.neuro.28.061604.135703.

Mainen, Z. F., & Sejnowski, T. J. (1996). Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature*, *382*, 363–366. doi:10.1038/382363a0.

Marenco, L., Ascoli, G. A., Martone, M. E., Shepherd, G. M., & Miller, P. L. (2008). The pilot NIF LinkOut Broker: A web resource to facilitate federated data integration using NCBI identifiers. *Neuroinformatics*, this issue.

Martone, M., Gupta, A., Wong, M., Qian, X., Sosinsky, G., Ludäscher, B., et al. (2002). A cell-centered database for electron tomographic data. *Journal of Structural Biology*, *138*, 145–155. doi:10.1016/S1047-8477(02)00006-0.

Rall, W., & Rinzel, J. (1973). Branch input resistance and steady attenuation for input to one branch of a dendritic neuron model. *Biophysical Journal*, *13*, 648–687.

Schierwagen, A., & Claus, C. (2002). Dendritic morphology and signal delay in superior colliculus neurons. *Neurocomputing*, *38–40*, 343–350. doi:10.1016/S0925-2312(01)00417-9.

Scorcioni, R., Polavaram, S., & Ascoli, G. A. (2008). L-Measure: a web-accessible tool for the analysis, comparison and search of digital reconstructions of neuronal morphologies. *Nature Protocols*, *3*(5), 866–876. doi:10.1038/nprot.2008.51.

Segev, I., & London, M. (2000). Untangling dendrites with quantitative models. *Science*, *290*, 744–750. doi:10.1126/science.290.5492.744.

Shepherd, G. M., Mirsky, J. S., Healy, M. D., Singer, M. S., Skoufos, E., Hines, M. S., et al. (1998). The Human Brain Project: neuroinformatics tools for integrating, searching and modeling multidisciplinary neuroscience data. *Trends in Neurosciences*, *21*, 460–468. doi:10.1016/S0166-2236(98)01300-9.

Smith, K., Jajodia, S., Swarup, V., Hoyt, J., Hamilton, G., Cornett, T., et al. (2004). Enabling the sharing of neuroimaging data through well-defined intermediate levels of visibility. *NeuroImage*, *22*, 1646–1656. doi:10.1016/j.neuroimage.2004.03.048.

Stiefel, K. M., & Sejnowski, T. J. (2007). Mapping function onto neuronal morphology. *Journal of Neurophysiology*, *98*(1), 513–526. doi:10.1152/jn.00865.2006.

Tossit, M., & Stocker, R. F. (2000). Metomorphosis in drosophila and other insects: the fate of neurons throughout the stages. *Progress in Neurobiology*, *62*, 89–111. doi:10.1016/S0301-0082(99)00069-6.

Vetter, P., Roth, A., & Häusser, M. (2001). Propagation of action potentials in dendrites depends on dendritic morphology. *Journal of Neurophysiology*, *85*, 926–937.

Wen, Q., & Chklovskii, D. B. (2008). A cost-benefit analysis of neuronal morphology. *Journal of Neurophysiology*, *99*(5), 2320–2328. doi:10.1152/jn.00280.2007.