The case for sequencing the genome of the electric eel


*Department of Biology 42451, University of Louisiana at Lafayette, Lafayette, LA 70504, U.S.A., †Section of Neurobiology, School of Biological Sciences, University of Texas at Austin, P. O. Box 7608, Austin, TX 78713-7608, U.S.A., ‡Department of Biological Sciences, Florida International University, Miami, FL 33199, U.S.A., §Department of Biology, Foster Hall - 3AF, New Mexico State University, Las Cruces, NM 88003, U.S.A., ‖Department of Arts and Sciences, University of Phoenix, Lafayette, LA 70504, U.S.A. and \|Biotechnology Center, 425 Henry Mall, University of Wisconsin, Madison, WI 53706, U.S.A.

(Received 6 February 2007, Accepted 4 July 2007)

A substantial international community of biologists have proposed the electric eel _Electrophorus electricus_ (Teleostei: Gymnotiformes) as an important candidate for genome sequencing. In this study, the authors outline the unique advantages that a genome sequencing project of this species would offer society for developing new ways of producing and storing electricity. Over tens of millions of years, electric fish have evolved an exceptional capacity to generate a weak (millivolt) electric field in the water near their body from specialized muscle-derived electric organs, and simultaneously, to sense changes in this field that occur when it interacts with foreign objects. This electric sense is used both to navigate and orient in murky tropical waters and to communicate with other members of the same species. Some species, such as the electric eel, have also evolved a strong voltage organ as a means of stunning prey. This organism, and a handful of others scattered worldwide, convert chemical energy from food directly into workable electric energy and could provide important clues on how this process could be manipulated for human benefit. Electric fishes have been used as models for the study of basic biological and behavioural mechanisms for more than 40 years by a large and growing research community. These fishes represent a rich source of experimental material in the areas of excitable membranes, neurochemistry, cellular differentiation, spinal cord regeneration, animal behaviour and the evolution of novel sensory and motor organs. Studies on electric fishes also have tremendous potential as a model for the study of developmental or disease processes, such as muscular dystrophy and spinal cord regeneration. Access to the genome sequence of _E. electricus_ will provide society with a whole new set of molecular tools for understanding the biophysical control of electromotive molecules, excitable membranes and the cellular production of weak and strong electric fields. Understanding the regulation of ion channel genes will be central for efforts to induce the differentiation of electrogenic cells in other tissues and organisms and to control the intrinsic electric behaviours of these cells. Dense genomic sequence information of _E. electricus_ will also help elucidate the genetic basis for the origin and adaptive diversification of a novel vertebrate tissue. The value of existing resources within the community of electric fish research will be greatly enhanced across a broad range of physiological and environmental sciences by having a draft genome sequence of the electric eel.
INTRODUCTION

In this paper, the authors outline the advantages to society of dense (seven- to 11-fold) shotgun sequencing of the genome of the electric eel *Electrophorus electricus* (Linnaeus 1766) (Teleostei, Gymnotiformes). The electric eel is one of several hundred vertebrate species capable of producing external electric fields in the water around their body. These electric fields are generated day and night throughout the life of the fish, range in amplitude from weak (millivolts) to strong (10–600 V) and may be used in a variety of natural behaviours, including communication, navigation, predation and defence (Albert & Crampton, 2006). Bioelectrogenesis is known only among vertebrates, and because the flow of electrical current requires an aquatic medium, all species capable of bioelectrogenesis are fishes.

The study of bioelectrogenesis is emerging as an important model system in the neural and behavioural sciences, including studies aimed at the molecular, cellular, organismal and ecosystem levels of organization. Electric fishes are the subjects of investigation in topics as varied as the physical structure and evolution of sodium channels (Noda *et al.*, 1984; Zakon *et al.*, 2006), the molecular basis of neuronal and synaptic plasticity (Bell *et al.*, 2005; Gomez *et al.*, 2005), neuronal regeneration and adult neurogenesis (Zupanc, 2001), changes in electric signals mediated by steroid and peptide hormones (Zakon, 2005), modelling electroreceptive images (Nelson *et al.*, 2002; Pereira *et al.*, 2005), assembly rules for neural networks (Bastian *et al.*, 2004) and ecological physiology (Julian *et al.*, 2003). The literature on electric fish physiology, behaviour and phylogeny is large and expanding rapidly; a search (15 June 2007) on the ISI Web of Science for original papers published 1973–2007 with the terms ‘electroreception’, ‘electric fish’ or ‘electric organ’ returned 2998 references, with an average of about 88 per year. A similar search on the terms ‘electric eel’ and ‘*Electrophorus*’ returned 739 papers.

Electric fishes produce electric discharges by means of specialized electric organs. Each of these electric organs is composed of hundreds (or in the case of *E. electricus*, thousands) of electrically charged cells called electrocytes (Fig. 1). The summed activation of electrocyte depolarizations is referred to as the electric organ discharge (EOD). In most electric fishes, the electrocytes are myogenically derived; i.e. electrocytes arise from the skeletal muscle precursors called myoblasts during development and regeneration. However, in one family of South American electric fish, Apteronotidae, the electrocytes of adults are derived from spinal electromotor neurons (Bennett, 1971). Muscle determination and differentiation in vertebrates centres on a core regulatory network, which is composed of two families of transcription factors; the basic helix-loop-helix muscle regulatory factors belonging to the MyoD family and the myocyte enhancer factor 2 group of MADS-box regulators (Atchley *et al.*, 1994; Yun & Wold, 1996; reviewed in Puri & Sartorelli, 2000; Wu *et al.*, 2000; Puri *et al.*, 2002; Palacios & Puri, 2006). Members of this gene network interact to regulate transcription of downstream tissue-specific differentiation genes (Ochi & Westerfield,
Immunological studies in a Neotropical electric fish [*Sternopygus macrurus* (Bloch & Schneider 1801)] have shown that electrocytes derive from the fusion of differentiated skeletal muscle fibres that primarily express type II myosin heavy chain (MHC) (Unguez & Zakon, 1998a, b).

Mature electrocytes are larger than muscle fibres, do not contain sarcomeres, or express MHC or tropomyosin, but do express keratin, a protein not expressed in muscle. Electrocytes are not contractile, although they do retain some muscle proteins, including cytoskeletal (Mermelstein *et al.*, 1997, 2000; Cartaud *et al.*, 2000), sarcomeric (Unguez & Zakon, 2002; Unguez *et al.*, 2002), neuromuscular junction proteins (Asher *et al.*, 1994) and even myogenic transcription factor proteins (Kim *et al.*, 2004; Neville & Schmidt, 1992; Ellisman & Levinson, 1982). Electrocytes are also known to differ from myocytes in an
increased expression of fast-gated Na\(^+\) channels, Na\(^+\)/K\(^+\) ATPase pumps and acetylcholine receptors (AChR) (Asher et al., 1994; Gotter et al., 1998).

Electric fishes are masters of ion channel regulation. By means of exquisite spatio-temporal control over the depolarization of excitable membranes, individuals of a given species are able to precisely mould distinct, structurally complex electric fields in the water around their body (Caputi, 1999). Two aspects of the EOD, in particular, the waveform and duration of individual pulses are based largely on the properties of the ion channels embedded within the electrocyte membranes (Ferrari et al., 1995). The EOD waveform refers to the number of phases of alternating positive and negative voltage in a head–tail recording, and the pulse duration is measured in milliseconds (Fig. 2). Many electric fish are also masters of electric discharge modulation. Some species can modulate the pulse duration and amplitude of the EOD, others modulate the amplitude only, and yet others do not appear to modulate the EOD at all (Markham & Stoddard, 2005). The repetition rate of the electric discharge train may also be modulated, and this is regulated by pacemaker neurons of brain-stem medulla. Sex steroids have been shown to modify all three of these EOD parameters in different species (Stoddard et al., 2006).

![Figure 2. Electric organ discharge (EOD). The discharge of the electric organ creates an electric field in the water around the fish. (a) Waveforms of species with a pulse- (Gymnotus varzea above) and wave- (Eigenmannia cf. virescens below) type EOD, as measured from head to tail. Tracings from W. Crampton. (b) Sketch of an instant of the electric field at a time when the head of the fish is at its peak positive value. Anterior to left. Contour lines are of constant potential (isopotentials in microvolts). Redrawn by M. Nelson (Knudsen, 1975). (c) Sketch of the same instant as in (b) showing lines of current flow perpendicular to isopotential lines. Ion currents flow down the voltage gradient (arrows). In the trunk region of the fish, the lines of current flow intersect the body at approximately right angles. When an object is brought close to the fish, it alters the pattern of current flow. Redrawn by M. Nelson from (Heiligenberg, 1977).](image-url)
To summarize, electric organs represent an important and largely unexplored resource for insights into the molecular basis of bioelectrical generation and ionic regulation. Diversity in the membrane excitability of electrocytes contrasts strongly with the relatively simple excitability of vertebrate skeletal muscle from which electrocytes evolved (Zakon et al., 2006). Ion channel genes, both structural (i.e. protein coding) and regulatory (e.g. transcription factors), are therefore important targets for understanding the biophysical control of electromotive molecules, excitable membranes and the cellular production of weakly electric fields (Zakon, 2002). Understanding the regulation of ion channel genes in electric fish will be central for efforts to induce the differentiation of electrogenic cells in other tissues and organisms and to control the intrinsic electric behaviours of these cells.

BACKGROUND TO BIOELECTROGENESIS AND ELECTRIC COMMUNICATION

The physiological basis of electrocyte function is well understood. For example, in the electric eel, Na\(^+\)/K\(^+\)-ATPase pumps embedded within the electrocyte membrane act continually to generate a weak voltage gradient (155 mV) across the membrane (K\(^+\) = −90 mV and E Na\(^+\) = +65 mV). Current flow is generated by the opening of voltage-gated ion channels in the electrocyte membrane, which may be depolarized by action potentials from innervating electromotoneurons or the depolarization of nearby electrocytes. When stimulated, activated AChR generate endplate potentials, triggering Na\(^+\) channel-mediated action potentials peaking at 65 mV on the innervated membrane. The non-innervated membrane contains no voltage-gated Na\(^+\) channels and maintains the −90 mV resting potential. The result is a transcellular potential difference of approximately 155 mV. Since each cell is stimulated simultaneously, electrocyte transcellular potentials summate. The potentials of three electrocytes, for example, culminate to produce 465 mV. Currents generated by stimulated electrocytes flow down electrocyte columns in the posterior to anterior direction. The circuit is closed by current flowing out the head, through the water, and back into the tail region.

Current flow generated by the electric organs produces an external electrical field that extends out to about one body length from the body surface (Knudsen, 1975). Any object within this field that has a conductivity different from that of water distorts the lines of current flow and therefore casts a sort of ‘electrical shadow’ onto the body surface. Electroreceptor cells embedded within the skin of the fish are sensitive to changes in the direction of ion fluxes across their membranes and transduce this local information into the language of action potentials. This information is then transmitted by nerves to the brain for integration with data from other sensory modalities. The whole circuit – from electric discharge to electroreception – forms the basis of active electrolocation, the capacity to construct cognitive images of the external electrical environment from changes in the self-generated EOD. Electric fish use electrolocation to orient themselves within the local environment and to navigate at night in murky tropical waters.

In addition to electrolocation, electric fish use their EOD in social interactions, such as aggressive and courtship behaviours. Multiple lines of evidence suggest that the EOD carries important cues used in recognizing members of
the same species. Phylogenetic comparisons of EOD differences among closely related species show that EOD pulse duration is the most rapidly evolving aspect of the electric signal among closely related pulse (Gymnotus) and wave (Sternopygus) types of species (Crampton & Albert, 2005). Ecological comparisons among members of multispecies sympatric assemblages have shown strict segregation of species with regards to a combination of pulse duration and repetition rate (Crampton & Albert, 2005). Behavioural playback studies in the laboratory have demonstrated that pulse duration and repetition rate are the two most important cues used by females in mate preference trials.

Diversity in characteristics of the electric organ and electrocytes forms the basis for production of species-specific signals (Hopkins, 1995). In Neotropical (gymnotiform) electric fishes, such as the electric eel, the electric organs are complex and physiologically heterogeneous (Caputi & Trujillo-Cenoz, 1994). Depending on the species, electric organs may exhibit regional specializations in patterns of nervous innervation, the size and shape of individual electrocytes and the repertoire of ion channel proteins. There are in addition several post-effector mechanisms that influence patterns of ion current flows through the body and therefore help shape the external EOD (Caputi & Budeli, 2006; Pereira et al., 2007). These include especially body shape, electrocyte tube configuration and patterns of electric organ insulation. All of these specializations result in the production of a distinct EOD.

Signal complexity is determined largely by three features: the anatomical organization of the cells of the electric organ, the biophysical properties of the ion channels in these cells and their surrounding connective tissues and skin and the body shape of the fish (Caputi, 1999). There is considerable information available on how the anatomical features of the electric organ result in the generation of complex waveforms. It is also well known how the structure of the electric organs have changed within groups to give rise to signal diversity (Zakon, 1996). In contrast, from a molecular perspective, little is known of the genes or proteins involved in determining the size and shape of electrocytes or the repertoire of receptors in the electrocyte membranes. However, because the membrane biophysics of electrocytes depend on ion channels and these molecules are known and their properties so well studied, the authors are in a good position to look at how expression and evolution of ion channel genes shape communication signals (Smith, 1999).

**A PROPOSAL**

The authors propose that among electric fish, the electric eel *E. electricus* be accorded the highest priority for genomic sequencing due to a unique suite of physiological and phylogenetic features. The electric eel is alone among freshwater fishes in possessing both strong (up to 600 V) and weak (millivolt) electric discharges. The electric eel grows to a relatively large size (2 m) and therefore provides large amounts (tens of grams) of electric organ tissue required for proteomic and metabolomic studies. Being a freshwater fish, the electric eel is more amenable as a laboratory model than is a marine species, and several closely related gymnotiform species are regularly bred in captivity (Kirschbaum & Schugardt, 2002). The diploid genome size of *E. electricus* is
c. 1 pg/cell or c. 1.0 gigabase of DNA distributed over 52 chromosomes. This is almost exactly the mean for other non-polyploid teleost fishes, c. 56% that of the zebrafish genome and c. 30% of the human genome. Karyotype analysis shows that *E. electricus* is not polyploid (Fonteles-Santos et al., 2002). Analysis of 42 cDNA clones with a total length of c. 47 kb shows an average GC composition of 41.9 + 0.05% (unpubl. data). The authors anticipate that c. seven-fold passes on genomic sequence will provide sufficient overlap among the shotgun-sequenced products to reconstruct the gene order and chromosomal arrangements of the genome (Lindblad-Toh et al., 2005). A seven-fold sequence coverage would permit a high-quality assembly covering >95% of the electric eel genome. The authors also anticipate that an 11-pass shotgun sequence would provide sufficient overlap among sequences to reconstruct a complete genomic scaffold. Additional information from the genomes of other genomic species would be used to establish synteny. The authors also propose to immediately initiate low-cost cDNA sequencing using new pyro-sequencing techniques, to provide a facile means of DNA chip-based gene expression studies and information to help annotate the genome sequence.

From a phylogenetic perspective, the electric eel is a member of the same superorder (Ostariophysi) as the zebrafish *Danio rerio* (Hamilton 1822) (Cypriniformes: Cyprinidae) from which it is separated by c. 80 million years (Alves-Gomes, 2001). This is approximately the same divergence time as that between humans and rats. Genomic studies on gymnnotiform electric fishes therefore gain most from the use of the numerous molecular tools developed for zebrafish. The results of research into the genomics of the electric eel will complement those of other teleost fish genomic models for which genomic sequencing is complete or underway (http://genamics.com; Table I). The recent or forthcoming publications of genomic sequences for these taxa provide an excellent foundation for sequencing weakly electric fishes and at substantial savings in time and other resources.

The objectives of an electric eel genome project closely match the mission of the U.S. Department of Energy Joint Genome Institute Community Sequencing Programme (DOE/JGI CSP). The intent of this programme is to conduct direct sequencing and to generate informatics capacity towards issues of scientific and societal importance in organisms other than those related directly to human disease or traditional model organisms. Dense genome sequences are currently available for c. 1400 species. Because of their medical and environmental importance and small genome sizes most of the species selected for sequencing are prokaryotic Eubacteria and Archaea (c. 800 spp.). As of this time there are 602 eukaryotic genome sequencing projects either completed or underway, of which 252 are for metazoans (animals) and nine are fishes (Table I).

The design of this proposal was crafted at the ‘International Electric Fish Genome Workshop’ held at the University of Wisconsin-Madison Biotechnology Center on 22 May 2005, with participants from universities all over North America. The workshop resulted in a proposal to the DOE/JGI CSP to generate raw sequencing data for electric eel genomic DNA, to assemble these into a draft genome sequence and to house the data at JGI, allowing community access according to the JGI data-sharing policy. The proposal received favourable reviews but was not accepted on the first round. However, declining costs of
high throughput DNA sequencing and an excess sequencing capacity at genome centres will eventually make possible the sequencing of an electric fish.

SCIENTIFIC QUESTIONS TO BE ADDRESSED

The availability of dense genomic data in the electric eel will permit for the first time genome-wide understanding of the factors that promote electrocyte differentiation. These data will allow experiments to be designed to address such questions as: Are new genes needed for the evolution of novel cell and tissue types? How are gene networks modified in the origins of new neuronal circuits and behaviours? If known genes are utilized, how are their expression patterns altered? Have genes been co-opted for the formation of novel structures? What genes used in the production of electric signal design differ among closely related species, within geographically widespread species, during development and during sexual differentiation?

For example, experiments could compare gene expression among different tissues (e.g. muscle v. electric organ), developmental stages, sexes and geographically disparate populations to address the Neo-Darwinian hypothesis: Do phylogenetic (interspecific) differences in electric discharges derive from within species variation in the expression, concentration and post-translational modifications of neural and electrocyte membrane proteins?

Complete genome sequences enable a more thorough understanding of biology, particularly of complex traits. Data from an electric eel genome project will contribute directly to scientific questions in four emerging areas of

Table I. Teleost fish species for which genomic sequencing is complete or underway. Stat., sequencing status; I, incomplete; C, complete

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common name</th>
<th>Order</th>
<th>Stat.</th>
<th>Type</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncorhynchus mykiss</td>
<td>Rainbow trout</td>
<td>Salmoniform</td>
<td>I</td>
<td>EST</td>
<td>TIGR</td>
</tr>
<tr>
<td>Salmo salar</td>
<td>Brown trout</td>
<td>Salmoniform</td>
<td>I</td>
<td>Genome</td>
<td>Genome British Columbia</td>
</tr>
<tr>
<td>Danio rerio</td>
<td>Zebrafish</td>
<td>Cypriniform</td>
<td>I</td>
<td>Genome</td>
<td>Washington University, Sanger Institute</td>
</tr>
<tr>
<td>Ictalurus punctatus</td>
<td>Channel catfish</td>
<td>Siluriform</td>
<td>I</td>
<td>EST</td>
<td>TIGR, Auburn University</td>
</tr>
<tr>
<td>Oreochromis niloticus</td>
<td>Nile perch</td>
<td>Perciform</td>
<td>I</td>
<td>Genome</td>
<td>University of New Hampshire</td>
</tr>
<tr>
<td>Oryzias latipes</td>
<td>Medaka</td>
<td>Beloniform</td>
<td>I</td>
<td>Genome</td>
<td>National Institute of Genetics Japan</td>
</tr>
<tr>
<td>O. latipes</td>
<td>Medaka</td>
<td>Beloniform</td>
<td>I</td>
<td>EST</td>
<td>TIGR, National Institute of Genetics Japan</td>
</tr>
<tr>
<td>Gasterosteus aculeatus</td>
<td>Threespine stickleback</td>
<td>Gasterosteiform</td>
<td>I</td>
<td>Genome</td>
<td>Broad Institute</td>
</tr>
<tr>
<td>Tetraodon nigroviridis</td>
<td>Spotted grebe pufferfish</td>
<td>Tetraodontiform</td>
<td>I</td>
<td>Genome</td>
<td>Genoscope, Broad Institute</td>
</tr>
<tr>
<td>Takifugu rubripes</td>
<td>Fugu</td>
<td>Tetraodontiform</td>
<td>C</td>
<td>Genome</td>
<td>JGI</td>
</tr>
</tbody>
</table>
research: (1) the origin of electric cells and tissues, (2) comparative studies of regeneration, (3) the neurophysiology of electrorception, a ‘sixth’ vertebrate sense and (4) biodiversity of megadiverse tropical aquatic ecosystems. Because the generation of the EOD is physiologically simple and directly related to the anatomical and molecular aspects of the electric organ, it is possible to take a reductionistic approach to these questions (Zakon et al., 1999; Stoddard, 2002a, b; Stoddard et al., 2006).

ORIGIN OF ELECTROCYTES, A NOVEL VERTEBRATE TISSUE

What differences in gene expression regulate differentiation of muscle cell precursors into mature myocytes and electrocytes?

A key event in evolution of electric fish was the origin of the electric organ from muscle (Zakon, 2002). This involved changes in the development of non-contractile electrocytes from skeletal muscle tissues (Unguez & Zakon, 1996; 1998a, b; Unguez et al., 2001, 2002; Unguez & Zakon, 2002). The origin of the electric organ from muscle remains poorly understood, but such understanding is a reachable goal in the near future (Zakon, 2005). One of the principle motivations for an electric eel genome sequencing project is to advance our understanding of the signals that control the activation of stem cells or dedifferentiation of adult cells. Defining normal molecular pathways of regeneration would serve as the basis of comparison to neoplastic processes (Fig. 3). Results of an electric eel genome sequencing project would be used to: (1) identify the molecular basis of progenitor cells during growth and regeneration, (2) characterize the transcriptional programme that diverges from the myogenic programme to give rise to the non-contractile current-producing cells of the electric organ and (3) identify molecular correlates that can distinguish among components of distinct electrical motor circuits (e.g. electromotoneurons and somatomotoneurons).

COMPARATIVE STUDY OF REGENERATION

What are the genes and gene products in Neotropical electric fish that facilitate complete regeneration of functional circuits in mature spinal cord after amputation?

Neotropical (gymnotiform) electric fish exhibit the extraordinary ability to replace spinal cord, skin, skeleton, muscle and the muscle-derived electric organ following recurring amputations (Patterson & Zakon, 1993, 1996, 1997). Indeed, among vertebrates only gymnotiforms and urodele amphibians (salamanders) possess the capacity to fully regenerate injured or excised portions of the spinal cord as adults. However, in terms of complexity and structural organization, the spinal cord of gymnotiforms much more closely resembles that of mammals than does the spinal cord of amphibians (Fig. 4). Therefore, information on the complete genome sequence of E. electricus will greatly advance research into the genetic basis of spontaneous spinal cord regeneration in other gymnotiform fishes.

Mechanisms to repair or replace damaged cells and body structures are universal among multicellular animals (Metazoa), yet the ability to fully regenerate
complex tissues is substantially reduced in many derived vertebrate groups, such as mammals and birds (Chernoff et al., 2003). Humans and other mammals have a very limited regenerative capacity, as evident by the complete loss of tissues and organs as a result of degenerative diseases and injuries (McClellan, 1998). It is well known that resident, tissue-specific stem cells serve to respond to injury or disease by repairing or replacing damaged tissue among vertebrates (Zupanc & Clint, 2003). It is also becoming clear that in some vertebrates, mature cells can re-enter the cell cycle, i.e. dedifferentiate, and contribute to the regeneration process (Chernoff, 1996; Zhang et al., 2000; Carlson et al., 2001; Echeverri et al., 2001; Echeverri & Tanaka, 2002; Beck et al., 2003; Maden & Hind, 2003). An understanding of the molecular and cellular mechanisms underlying dedifferentiation, redifferentiation and transdifferentiation of tissue that occurs during regeneration is critical to understanding the biology of cell phenotype determination and maintenance and also has the potential to reveal ways to enhance tissue restoration in humans.

**NEUROPHYSIOLOGY OF THE ELECTROSENSORY SYSTEM, A VERTEBRATE ‘SIXTH SENSE’**

What are the genomic, developmental and behavioural circumstances involved in the origin of vertebrate sensory systems?
The availability of genomic sequences for an electric fish will revolutionize studies on the expression, regulation and distribution of ion channels and neurotransmitter systems in the electrosensory system (Northcutt et al., 2000; Freitas et al., 2006). Active research in these areas include investigations into the neural physiology and molecular biology of electroreception (Maler & Monaghan, 1991; Maler et al., 1991; Maler & Mugnaini, 1994; Maler, 1996; 1999a, b; Maler & Hincke, 1999) ion channels in the electrosensory system (Turner et al., 1994, 1995, 1996, 2002; Turner & Borg, 1995; Turner & Moroz, 1995; Turner & Maler, 1999), neuronal physiology and mathematical modelling of electrosensory signals (Chacron et al., 2000, 2001a, b, 2003a, b, 2005a, b), comparative neuromorphology (Albert et al., 1998), neural physiology and biochemistry in the electrosensory system (Lewis & Maler, 2001; 2002a, b, 2004) and the physiology of electrosensory signal processing (Krahe et al., 2000, 2002; Krahe & Gabbiani, 2004).

Fig. 4. Organization of the adult spinal cord in selected vertebrate species. (a) An amphibian, the Northern leopard frog *Rana pipiens*. (b) A mammal, the house mouse *Mus musculus*. (c) A gymnotiform, the Rat-tail knifefish *Sternopygus macrurus*. (d) The electric eel (*Electrophorus electricus*). All sections from abdominal body regions. (a) and (c) stained with haematoxylin for cell bodies; (b) and (d) stained with luxol blue (dark) for myelinated axons and Nissl (light) for cell nuclei. Note that the mammalian spinal cord (b) more closely resembles knifefishes (c and d) than it does amphibians (a) in being organized into discrete white and grey matter zones with dorsal and ventral horns and motoneuron nuclei. Adult amphibian spinal cords are secondarily simplified, resembling juveniles of other vertebrates, with fewer cell types and a large central canal (Roth et al., 1997). Ependymal layer is visible as dark cells in (a) and (c) lining the central canal.

© 2008 The Authors
These multifold investigations into the physiology and molecular biology of the electric sense are providing unique insights into the mechanisms that neurons use to decode sensory signals in the vertebrate central nervous system (Zakon, 2003). The electric sense of Neotropical electric fishes such as the electric eel is a relatively simple and tractable system with broad applications for discovering principles in other vertebrate sensory systems including humans. One of the most striking findings is that neuronal responses to electrosensory signals depend on the precise control of the expression of membrane channels and receptors (Fernandez et al., 2005; Novak et al., 2006; Stoddard et al., 2006). This control is achieved through many mechanisms, but most critical is the cell specific control of gene expression, mostly at the transcriptional level (Maler & Mugnaini, 1994; Zakon, 2002). Obviously, the availability of genomic sequences will be essential for real progress in this area.

Recent studies using cloning of key proteins involved in mammalian, including human, memory (N-methyl-D-aspartic acid receptors and their downstream effectors such as Ras-GRF1,2) show remarkable homology with the electric fish (Zakon et al., 2002; Dunlap & Larkins-Ford, 2003) and mouse genes, suggesting that research on electric fish may have direct benefits for human brain function and pathology. The availability of a dense electric eel genome on these areas of research would greatly accelerate efforts to harness the remarkable ion channels that form the basis of both electrosensation and EODs.

EVOLUTION OF SPECIES-RICH TROPICAL AQUATIC ECOSYSTEMS

What are the differences in gene expression among closely related electric fish species that underlie the generation of diverse electric signals?

Neotropical freshwaters constitute the largest epicontinental aquatic fauna on earth, with more than 6000 species, accounting for perhaps 10% of all vertebrate species (Vari & Malabarba, 1998). The evolutionary and ecological mechanisms underlying the origin and maintenance of this enormous diversity remain poorly understood, and the genetic basis of adaptive diversification in Neotropical freshwaters is almost entirely unexplored.

The species-specific electric signals of Neotropical electric fishes represent unparalleled opportunities for studies into the physiological and genomic basis of diversification (Zakon, 2003). These fishes possess all the major patterns of species diversity observed among Neotropical fishes as whole (Lundberg, 1998a; Lundberg et al., 1998b). They are: (1) ancient, with origins in the Late Cretaceous period (c. 80 Ma), (2) diverse, with approximately 200 known species (152 described) and many others anticipated from future collections and (3) geographically widespread, with species inhabiting all Neotropical river basins from southern Mexico to northern Argentina. Even at the species level, gymnotiform lineages are ancient, in many cases dating to the Middle Miocene (c. 12 Ma), and widespread, with single species commonly ranging over 10^6–10^7 km² (Albert et al., 2005). Gymnotiformes are also like other Neotropical fishes in being distributed in polyphyletic regional species assemblages. In this regard, they are quite unlike many of the celebrated cases of adaptive radiation on islands and lakes. Insular adaptive radiations have proven useful for illustrating
some of the genetic and ecological mechanisms underlying speciation and adaptation, but they are rare: *i.e.* most local and regional species assemblages on Earth are polyphyletic; *i.e.* to say they have multiple origins in space and time.

Because the evolutionary origins and ecological maintenance of species is closely linked to mechanisms of reproductive isolation and because electric signals are transmitted as action potentials in the same ‘currency’ as that of the nervous system, electric fishes provide a unique window for studies on the genomic and physiological basis of adaptive diversification in the tropics (Albert & Crampton, 2005). Four properties of the EOD waveform make it an attractive target for studies into the genomic basis of adaptive diversification. The EOD waveform is: (1) generated day and night throughout the entire life of an individual and therefore continually accessible to monitoring, (2) stereotyped, relatively consistent from pulse to pulse and possessing aspects that are characteristic of each species, (3) physiologically simple, arising from the summed activation kinetics about eight distinct electrocyte membrane proteins and (4) adaptive, functionally associated with water local quality parameters (*e.g.* flow rate, dissolved O$_2$, temperature, conductivity (Crampton, 1998). For example, species that live in highly conductive water possess multiple rows of electrocytes that are capable of generating more current (Hopkins, 1999).

In summary, the fortunate combination in gymnotiform electric fishes of an exceptional electrogenic-electrosensory system and otherwise typical patterns of biodiversity and biogeography permits studies into the behavioural and ecological mechanisms underlying species differences simply unavailable in other fishes. Furthermore, because the physiology of the electric discharge is both simple and stereotyped, these species-level differences can be readily studied at the cellular and genomic level. Complete genome sequence data on the electric eel will dramatically accelerate understanding of the genetic mechanisms underlying the evolution of electric signals and the adaptive diversification of gymnotiform species in Neotropical freshwaters.

**RATIONALE FOR COMPLETE GENOME SEQUENCES**

A complete genome sequence of the electric eel will dramatically advance research in a broad range of disciplines. These emerging fields have implications beyond the area of electric fish or even teleost research, such as clinical and biotechnological applications. Access to a completely annotated genome of the electric eel will open avenues to rapidly identify the actions of regulatory elements and other genes that are transiently expressed or in low copy numbers. Complete genome data will rapidly illuminate general features of electric organ and electrocyte development. These generalities can be utilized to guide biotechnology research into methods for transferring and adapting biochemical pathways from electric teleost fish to cell culture systems ranging from yeast to mammals.

Genomic DNA is valuable because the key genes of interest that regulate cell and tissue formation may be rarely expressed (*e.g.* transcription factors) or not expressed at all (*e.g.* cis-regulatory elements). In their report of the recently sequenced dog genome the authors note: ‘Fifty per cent of the most highly conserved non-coding sequence in the genome show striking clustering in 200
gene-poor regions, most of which contain genes with key roles in establishing or maintaining cellular identity, such as transcription factors or axon guidance receptors (Lindblad-Toh et al., 2005: 804).

Whole genome comparisons of both coding and non-coding regions are required to: (1) understand the abundances and distributions of members of gene families and repetitive elements, (2) enable whole genome comparisons to other teleost fish and more distant vertebrate species, (3) provide target material for the isolation and/or amplification of specific sequences for RNA and protein levels studies and (4) examine the roles of genome-wide sequence structures as loci of recombination. The sequencing of targeted coding regions or the piecemeal examination of bacterial artificial chromosome (BAC), fosmid or plasmid clones simply does have sufficient power to reveal the whole genome changes involved in the specialization of gene structure and function that occurred during evolutionary diversification.

Whole genome sequences also provide high-quality genomic information of immediate use in constructing high-density DNA microarrays. Expression microarrays will be important for examining gene expression patterns associated with alternate phenotypic or disease states in electric fish models. Microarrays based on whole genome sequence (e.g. tiling arrays) have the advantage of including all sequences in the genome, including ‘cryptic’ genes not predicted using computer algorithms. In addition, such expression arrays, as well as ‘re-sequencing’ or SNP arrays can be used in novel fast-track approaches to the positional cloning of QTLs that contribute to normal and abnormal physiological and developmental variations and are detected and localized by QTL linkage mapping approaches.

Using an instrument known as the maskless array synthesizer (MAS) developed at The University of Wisconsin (in a collaboration of engineers, geneticists and biochemists in the laboratories of Sussman, Cerrina and Blattner), high-density DNA microarrays may be created at a relatively low cost within only a few hours after a sequence is obtained (cf. Singh-Gasson et al., 1999). This technology has been commercialized by NimbleGen Systems, Inc. (Madison, WI, U.S.A.) and is now widely used for studies of transcription, epigenetics and evolution, including the ENCODE human genome project (E.P.C., 2004). Whole genome sequence arrays, such as tiling arrays, can give important insights into elements involved in co-regulation, development and neuronal development. Tissues of interest to the research community vary widely from muscle tissue, electric organ tissue, neuronal tissue and embryonic tissue.

The authors propose the following strategy for the *E. electricus* genome based on experience and insights gained, while mapping and sequencing recent genomes, such as the mouse and opossum. A c. seven-fold whole genome shotgun component consisting of both small (plasmids) and large insert clones (BAC, fosmid) ordered and oriented on the chromosomes, will provide the community of users with rapid access to most of the *E. electricus* genome sequence. A BAC clone-based physical map, along with paired end sequences from fosmids and the mapped BAC clones, will provide a framework by which the genome sequence can be accurately assembled and ordered and oriented on the chromosomes. In addition, anchoring to the genome will be confirmed by performing fluorescence in situ hybridization with fosmids from major
supercontigs. The authors understand from previous successful proposals for complete genomic sequencing in species with large genomes, such as the domestic dog *Canis familiaris* (Ostrander *et al.*, 2002; Lindblad-Toh *et al.*, 2005) and the short-tailed opossum *Monodelphis domestica*, that utilizing different library sizes provides a framework for accurate assembly. As with the mouse genome sequence, the proposed approach will result in anchored ‘supercontigs’ (sequence contigs connected by at least two read-pair links) of >10 Mb in average length.

Prior to the initiation of genomic sequencing at DOE/JGI CSP, the authors propose to begin establishing some of the minimal resources needed to help assemble and annotate these data. This work would take advantage of recently developed rapid and inexpensive methods of obtaining short reads from many hundreds of thousands of cDNA clones. For example, one instrument, commercialized by 454 Inc., has been used to obtain nearly 200 000 cDNA sequences of c. 92 nucleotides each from cDNA isolated from the legume model plant *Medicago truncatula* (alfalfa) (Cheung *et al.*, 2006). This sequence only requires c. $2000 investment and would provide enough information to produce gene expression chips using the MAS to allow immediate studies of relative gene expression among the organs in *E. electricus*. This information in turn would provide important information to help annotate the genome when a complete genome sequence is obtained.

**LIMITATIONS OF *E. ELECTRICUS* AS A MODEL GENOMIC ORGANISM**

Despite the many advantages that genomic sequences of *E. electricus* would offer, in some regards this species is not an ideal model organism for genetic studies. Individual electric eels grow to a large adult size (2 m) and are difficult to breed in captivity. In addition, *E. electricus* is unique among Gymnotiformes in not possessing the otherwise exceptional ability of this group to regenerate the spinal cord after injury or amputation of the caudal appendage. Therefore, many scientific questions that the authors would like to address using genetic manipulation (e.g. origin of electric cells and tissues, gene regulatory networks related to new neuronal circuits) would not be accessible in this species. There is in general a dearth of genomic resources available for Gymnotiformes as a whole, and *E. electricus* in particular. There are currently little or no data about the repeat nature of this genome or information to construct a physical map of the genome, although a BAC library is available for the chocolate-ghost *Apteronotus leptorhynchus* (Ellis 1912) from R. Dunn and colleagues.

In this regard, there are two other gymnotiform electric fishes for which genomic sequences would be valuable. The feather-tail knifefish *Brachyhypopomus pinnicaudatus* (Hopkins 1991) (Hypopomidae) is readily bred and raised in captivity, with a 5 month cycle from egg to egg (pers. obs.). This species does not possess a strong electric discharge, but the electrocytes of its weakly electric organ are physiologically and developmentally very similar to those of the electric eel. *Brachyhypopomus pinnicaudatus* would make an exceptional model organism for targeted genomic studies using genome-wide and knock-down approaches, including surveys of upregulation, or RNA and/or protein analyses...
under different environmental/experimental conditions. Another gymnotiform species, the black-ghost *A. leptorhynchus* (Apteronotidae) cannot be easily bred in captivity but exhibits a spectacular phenotypic trait in the possession of neurogenic (neurally derived) electric organs as adults. Comparison of neurogenic and myogenic organs will provide keys to development of electric tissue from two different types of stem cells.

The genomes of these three gymnotiform species (*E. electricus*, *B. pinnicaudatus* and *Apteronotus albifrons* (Linnaeus 1766)) are around 1 gigabase each, so getting complete genomic sequence data is not a trivial matter. Nevertheless, the authors believe that complete genomic DNA of any one of these species would be more useful to the user community than would partial genomes of all three species, mainly because the transcription factors regulating electrocyte differentiation are expected to be highly conserved among gymnotiform species. In other words, dense (7–11× pass) on genomic sequence from any of these species would be more valuable than complete cDNA sequences from all three species.

**PROSPECTUS**

The availability of dense genomic sequences for electric fish will be of tremendous use for the large community of scientists currently working directly with electric fish. Such information will accelerate the role of electric fish as a model organism by uniting the technological power of emerging tools in genomics and bioinformatics with the depth of physiological, ecological and evolutionary background knowledge now available in these species.

Understanding the genomic basis of bioelectrogenesis will also benefit a much larger scientific community. Because electrogenesis is achieved by harnessing and modifying biophysical mechanisms common to other excitable membranes (*e.g.* neural, muscular and cardiac cells), information gained on the genomics of the electric eel will provide new insights into the biophysical function of excitable cells generally. From a practical standpoint, harnessing the full potential of bioelectrogenesis will require a complete understanding of mechanisms underlying the development, regulation and modulation of electrogenic cells. Results of these comparative studies into the fundamental genetic and epigenetic control of muscle and neural development will help accelerate the functional annotation of the human genome and understanding of the roles of human genes in health and disease.

Results of an electric eel genome project would facilitate the discovery of genes such as transcription factors that are expressed only rarely or transiently but which are fundamental in development. These data would greatly advance research into the development and regeneration of electrocytes, a novel vertebrate tissue type, and the electrosensory system, a novel vertebrate sensory system. The availability of genomic sequences in electric fish will enable real progress on the development of new technologies to use electricity in medicine, agriculture and other aspect of applied research.

Understanding the genetic basis for the origin and maintenance of electric cells and tissues will have a profound influence on many emerging areas of biotechnology research in the biomedical and agricultural sciences. The authors briefly describe several possible applications, which, although by no means
exhaustive, provide a glimpse of the power gained from understanding the genomic basis of electrocyte differentiation.

**BIOBATTERIES IN VITRO**

Single-celled photosynthetic eukaryotes could be genetically engineered to convert sunlight directly into usable electric currents. These cells could serve as biobatteries that would store and discharge electricity under highly specified circumstances. For example, certain aquatic algae or yeasts could be designed to maintain strong voltage potentials in solution by means of enhanced expression of Na\(^+\) pumps. Voltage-gated protein channels and gap junctions could be added or modified to co-ordinate membrane depolarization under the influence of common external cues, such as chemical or light gradients or ambient electrical fields. By altering genes for cell adhesion molecules, mechanisms of mitosis and motility could be altered so that cells populations would spontaneously organize into polarized columns (filaments) and surfaces (biofilms). The organization of a population of electrogenic cells in series or in parallel would enhance electromotive force (volts) and current (amps), respectively, of the cell array. Unlike mechanical or chemical batteries, genetically engineered biobatteries made of aquatic algae or yeasts would spontaneously grow, repair and replace themselves with minimal operator input.

**BIOBATTERIES IN VIVO**

In the context of multicellular tissues (in vivo), the conversion of normal cells into electrogenic cells would be a unique resource in supplying power to implanted electric devices (e.g. nanots) or the excitation of conventional biological targets, such as muscles, nerves and glands. In a medical context, electrogenic cells could be used in the control of many disorders involving the dysfunction of electrically excitable membranes (e.g. Parkinson’s disease, epilepsy, muscular dystrophy and cardiac arrhythmia, etc.). For example, Duchenne muscular dystrophy results from mutations in the dystrophin gene, which impairs attachment of the cytoskeleton to the cell membrane (Webster & Blau, 1990; Lovering et al., 2005; Goncalves et al., 2006). The application of weak electric currents from electrocytes induced in surrounding tissues could serve as an external cue to help polarize tubulin attachment activity (Kirschner & Mitchison, 1986; Gerhart & Kirschner, 1997). Being derived from native tissues, transformed electrogenic cells would automatically deliver desired electrical currents directly to target cells without risk of challenging the immune system. Because ion currents are precisely the currency of excitable membranes, electrogenic cells would deliver biologically meaningful stimuli at appropriate amplitudes and frequencies. Electrogenic cells could also be used to control the localization and direction of weak electrical currents or fields on biological or inanimate surfaces. Electrogenic tissues could generate weak currents to constrict localized portions of the dermis, the walls of blood vessels or other epithelial surfaces (like botox) and could be used in the control of blood pressure and other cardiovascular diseases. In a clinical setting, all such gene therapy
approaches would be most effective in association with a carefully designed physical therapy programme (Lovering et al., 2005).

**BIOREPORTERS**

Electrogenic cells could also be used as bioreporters to ‘electrically label’ cells, in the same way as green fluorescent protein and other photolabelled cells report their location and abundance by emitting light. An advantage of emitting weak electric fields *in vivo* is that the host animals come naturally equipped with a complex nervous system that is specifically designed to detect and interpret weak electric signals. Electrically labelled bioreporter cells could also be used *ex vivo* for ecotoxicological studies in remote aquatic environments, where the emission of weak electric fields could be readily monitored by available technologies (Cardon & Gage, 2006; Harms et al., 2006; Mancuso, 2006). For example, electrically emitting bioreporter cells could document spatio-temporal patterning in resource availability. Electrogenic bioreporter cells could serve as indicators for variety of natural or synthetic chemical or biological indicators, including soil and sediment types, pollutants and other ecosystem stressors.

In summary, the community of electric fish biologists is well positioned to use a dense genome sequence of the electric eel in several emerging areas of research related to the production, control and evolution of bioelectrogenesis. Advances in these areas are likely to have profound impacts in both basic and applied research into the regulation of excitable membranes. A complete genome sequence for the electric eel will bring the study of bioelectrogenesis into the genomic era of biology.

Many of the ideas in this paper were developed at the International Electric Fish Genome Workshop held at the University of Wisconsin-Madison Biotechnology Center on 22 May 2005. Special thanks to all the workshop participants. We also thank W. Crampton for Fig. 2(a), M. Nelson for Fig. 2 (b), (c) and N. Maclean for the invitation to prepare this review.

**References**


**Electronic Reference**