Research News

Neuroscience in the post-genome era: an overview

Christopher A. Walsh

Although sequencing of the human genome has been taking place in rapidly accelerating fashion for years, the presentation of the entire sequence (to a first approximation, anyway) has now permitted a global view of its structure. The completion of the sequence has therefore invited all manner of aesthetic, philosophical, societal – as well as biological – discussions of its implications. This global view of the genome will undoubtedly forever change the face of biology. But how? Here are a few perspectives from the point of view of neuroscience.

Many of the first impressions gleaned from the review of the human genome sequence tend to be comparative or evolutionary. This is to be expected, because the most obvious thing to do when you are presented with the sequence of a new genome is to compare one sequence with previously sequenced genomes and see how it differs. Nonetheless, this first look has generated some interesting findings, and is likely to have a huge, immediate impact on our approach to understanding the evolution of humans and other species.

'…no more than 32 000 in the human genome. This is only a little more than twice the estimate for *Drosophila* (13 600), less than twice the estimate for *C. elegans* (19 000)…'

The obvious way to look at the human genome is in terms of the parts of it that encode proteins, and to see what those proteins are. Comparison of the numbers of genes has been a real surprise to those people (myself included) who expected that our 'higher' evolutionary status would be reflected in a greater number of distinct genes. It appears that there are no more than 32 000 in the human genome. This is only a little more than twice the estimate for *Drosophila* (13 600), less than twice the estimate for *C. elegans* (19 000), and only slightly greater than the estimate for the weed, *Arabidopsis thaliana* (25 500). In fact, some gene families are held fairly constant: for example, there are similar numbers of serine/threonine and tyrosine kinases in humans and the non-human, metazoan genomes[1,2]. By contrast, there are fewer members of the P450 family of detoxifying enzymes in humans than in worms, flies, or especially *Arabidopsis*[2]. The differences in the P450 genes presumably reflect the idea, recently proposed by James Watson, that organisms with large brains are capable of stepping out of the way of many dangers such as toxins, rather than having to deal with them after ingestion. In this way, the evolution of a brain, although it requires many new neural-related genes, actually produces potential savings in some non-neural genes. These differences show how supple evolutionary forces can be, and how nonrandom is the amplification of gene families. An original model held that vertebrates obtained four copies of many genes (Hox clusters, for example) present in single copy in flies or worms secondary to wholesale genome duplication and re-duplication, followed by paring away of extra copies of nonessential genes. Present data do not favor this model – most of all because many gene families are not tetraploid in humans. Instead the increase in gene number has been much more piecemeal, suggesting that evolution has been more selective in the gene families that are amplified and retained.

Where are the big increases in gene families from nonvertebrates to humans? One area that has clearly been preferentially expanded involves genes with neurobiological functions. For example, in the case of the semaphorins, which regulate axon and dendritic outgrowth and remodeling, there are 22 semaphorin genes reported in the human genome, versus six in the fly and two in the worm[1]. The human genome also shows substantial expansion of many other gene families that are crucial to brain development, for example, cadherin-family adhesion molecules, nerve growth factor (NGF) family members, transforming growth factor-β (TGF-β) family members, EPH ligands and ephrin receptors, and extracellular matrix proteins. Other neurobiologically related gene families that are greatly expanded are used in the adult function of the nervous system, such as ion channel families and genes related to myelin structure and function. Many of these expanded gene families probably reflect the increasing encephalization of vertebrates. Another surprise has come, however, in the origin of some of these new genes. For example, there is good evidence that the monoamine oxidase gene, a crucial enzyme in catecholamine metabolism and target of action of many psychotropic drugs, was acquired from a bacterium during the course of evolution. Dozens of other active human genes have been acquired from ancient retroviruses and other transposons[2].

'Changes in these transposable elements…might be crucial for evolutionary changes in brain size and shape by altering expression levels or expression patterns of nearby genes.'

The second focus of interest in the initial study of the genome sequence has been in the noncoding DNA, which might elucidate our evolution. The human genome is littered with countless potentially transposable elements, but it is reassuring to know that most of these elements are quiescent, and seemed to have become burned out many generations ago, rather than continuing to flip around. However, when examined globally, the locations of retained retrotransposons relative to the genes turn out to be highly nonrandom, suggesting that certain retrotransposons are retained for a good reason. Changes in these transposable elements, perhaps even more so than changes in the potions of genes that code for proteins, might be crucial for evolutionary changes in brain size and shape by altering expression levels or expression patterns of nearby genes. For example, the *LIS1* and *Reelin* genes encode proteins that are required for cerebral

 0166-2236/01/$ – see front matter PII: S0166-2236(00)01866-X

cortical development, and which are almost identical in sequence among all mammals and even among nonmammalian vertebrates. However, at least in the case of *Reelin*, its pattern of expression in the developing forebrain differs greatly between mammals who show a layered cerebral cortex and nonmammalian vertebrates in which the forebrain is not similarly laminated[3]. Perhaps, evolutionarily, it is changes in expression pattern and expression level that alter the formation of major brain structures, and perhaps this is, in turn, regulated by the insertion of so-called 'junk' DNA (Ref. 2).

Neuroscience in the post-genomic age

How will the availability of whole genome sequences change the practice of neuroscience? It is hard to predict for sure, just as it was hard to predict how understanding the double-helical structure of DNA would go on to change biology 50 years ago. However, a few trends are noticeable.

Humans as a convenient genetic model system for the mouse

Humans are likely to give mice a run for their money as a system for identifying the genes that are required for a working nervous system. The system of choice has evolved from flies, worms, zebrafish, to a recent interest in mice. Obviously, for some studies a mammalian system is needed. However, the availability of the human genome, and the millions of single nucleotide polymorphisms and mapping information that has come along with it[4], has and will continue to simplify identification of disease genes in humans. Suddenly, the human has advanced beyond the zebrafish and mouse to become one of those few organisms that are unusually well-characterized genetically. Of course, it will probably only be a few months until mice and zebrafish catch up. However, the presence of the entire human genome will forever make it easier to study questions of human brain evolution directly in the DNA. But increasingly, 'working with humans' – be it human disease genes in an animal model or even directly with human samples – will not just be good grantsmanship, it will also be important to stay current.

Computers replace plastic bags, filters and messy radioactive solutions for library screening and cloning

The human genome sequence has been reflected for many years in the improved reagents available. This has, in turn, led not only to the faster identification of human disease genes, but also to greatly increased ease of cloning or obtaining clones of cDNAs, and the ability to perform 'in silico' library screening, to find the gene of your choice already in some internet database somewhere. A good example of this approach was the identification of bitter taste receptors, in which the known location of a phenotype in mice led to a successful search for candidate genes in the syntenic region of the human genome sequence[5,6]. That search yielded the sequences of several genes that turned out to encode bitter taste receptors, as demonstrated physiologically.

All the genes, all the time

What the entire sequence of the human genome really symbolizes however is a new approach – systematic study of all genes, not just a few. How influential will this be? One camp feels that anyone not doing this kind of 'whole genome' analysis will be left behind on the trash-heap of neuroscience. This camp feels that the days of doing science from point-to-point, or from single molecule to single molecule, are numbered. Yet there is another, Luddite camp that emphasizes that there is as yet no evidence that the scientific method, which proceeds by controlling all variables except one – has been retired or invalidated. Only time will tell whether genome sequencing creates a startling new biological method, but we already know that it provides us with a marvelous new biological map.

References

1 Venter, J.C. *et al.* (2001) The Sequence of the Human Genome. *Science* 291, 1304–1351
2 Lander, E.S. *et al.* (2001) Initial sequencing and analysis of the human genome. International Human Genome Sequencing Consortium. *Nature* 409, 860–921
3 Bernier, B. *et al.* (1999) Reelin mRNA expression during embryonic brain development in the turtle Emys orbicularis. *J. Comp. Neurol.* 413, 463–479
4 Altshuler, D. *et al.* (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 407, 513–516
5 Adler, E. *et al.* (2000) A novel family of mammalian taste receptors. *Cell* 100, 693–702
6 Matsunami, H. *et al.* (2000) A family of candidate taste receptors in human and mouse. *Nature* 404, 601–604

Christopher A. Walsh
Dept of Neurology, Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, MA 02115
e-mail: cwalsh@caregroup.harvard.edu

Editorial policy

Most articles published in *TINS* are commissioned by the Editor. However, authors who wish to contribute to any section of the journal should contact the Editor with the names of all authors and a point-by-point outline of the proposed article including 10–12 key references. Outlines may also be sent to members of the Advisory Editorial Board. Completed manuscripts submitted without liaison with the *TINS* Editorial Office cannot be considered. *TINS* readers who wish to suggest that a particular topic be covered are also invited to send their suggestions to the Editor.

Correspondence and comment on any aspect of *TINS* or the field of neuroscience are welcomed, and should be sent directly to the Editor.